Advances in Intelligent Systems and Computing 704

Bidyut B. Chaudhuri
Mohan S. Kankanhalli
Balasubramanian Raman *Editors*

Proceedings of 2nd International Conference on Computer Vision & Image Processing

CVIP 2017, Volume 2

Springer

Advances in Intelligent Systems and Computing

Volume 704

Series editor

Janusz Kacprzyk, Polish Academy of Sciences, Warsaw, Poland
e-mail: kacprzyk@ibspan.waw.pl

The series "Advances in Intelligent Systems and Computing" contains publications on theory, applications, and design methods of Intelligent Systems and Intelligent Computing. Virtually all disciplines such as engineering, natural sciences, computer and information science, ICT, economics, business, e-commerce, environment, healthcare, life science are covered. The list of topics spans all the areas of modern intelligent systems and computing such as: computational intelligence, soft computing including neural networks, fuzzy systems, evolutionary computing and the fusion of these paradigms, social intelligence, ambient intelligence, computational neuroscience, artificial life, virtual worlds and society, cognitive science and systems, Perception and Vision, DNA and immune based systems, self-organizing and adaptive systems, e-Learning and teaching, human-centered and human-centric computing, recommender systems, intelligent control, robotics and mechatronics including human-machine teaming, knowledge-based paradigms, learning paradigms, machine ethics, intelligent data analysis, knowledge management, intelligent agents, intelligent decision making and support, intelligent network security, trust management, interactive entertainment, Web intelligence and multimedia.

The publications within "Advances in Intelligent Systems and Computing" are primarily proceedings of important conferences, symposia and congresses. They cover significant recent developments in the field, both of a foundational and applicable character. An important characteristic feature of the series is the short publication time and world-wide distribution. This permits a rapid and broad dissemination of research results.

More information about this series at http://www.springer.com/series/11156

Bidyut B. Chaudhuri · Mohan S. Kankanhalli
Balasubramanian Raman
Editors

Proceedings of 2nd International Conference on Computer Vision & Image Processing

CVIP 2017, Volume 2

Springer

Editors
Bidyut B. Chaudhuri
Computer Vision and Pattern Recognition Unit
Indian Statistical Institute
Kolkata
India

Balasubramanian Raman
Department of Computer Science and Engineering
Indian Institute of Technology Roorkee
Roorkee, Uttarakhand
India

Mohan S. Kankanhalli
School of Computing
National University of Singapore
Singapore
Singapore

ISSN 2194-5357 ISSN 2194-5365 (electronic)
Advances in Intelligent Systems and Computing
ISBN 978-981-10-7897-2 ISBN 978-981-10-7898-9 (eBook)
https://doi.org/10.1007/978-981-10-7898-9

Library of Congress Control Number: 2017963008

Printed on acid-free paper

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
part of Springer Nature
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

Preface

The Second International Conference on Computer Vision and Image Processing (CVIP 2017) was organized at Indian Institute of Technology Roorkee (IITR), Greater Noida Extension Center, during September 09–12, 2017. The conference was endorsed by International Association of Pattern Recognition (IAPR) and was primarily sponsored by MathWorks. CVIP 2017 brought together delegates from around the globe in the focused area of computer vision and image processing, facilitating exchange of ideas and initiation of collaborations. Among 175 paper submissions, 64 (37%) were accepted based on multiple high-quality reviews provided by the members of our technical program committee from ten different countries. We, the organizers of the conference, were ably guided by its advisory committee comprising distinguished researchers in the field of computer vision and image processing from seven different countries. A rich and diverse technical program was designed for CVIP 2017 comprising five plenary talks and paper presentations in seven oral and two poster sessions. Emphasis was given to the latest advances in Cybernetic Health, Perception of Visual Sentiment, Reshaping of Human Figures in Images and Videos using 3D Morphable Models, Vision and Language, and Challenges in Biometric System Development. The papers for the technical sessions were divided based on their theme relating to Computer Vision Applications, Document Image Analysis, Machine Learning and Uncertainty Handling, Surveillance and Security, Summarization, Retrieval and Recognition, and Low-level Computer Vision. This edited volume contains the papers presented in the technical sessions of the conference, organized session-wise. Organizing CVIP 2017, which culminates with the compilation of the volume of proceedings, has been a gratifying and enjoyable experience for us. The success of the conference was due to synergistic contributions of various individuals and groups including the international advisory committee members with their invaluable suggestions, the technical program committee members with their timely high-quality reviews, the keynote speakers with informative lectures, the local organizing committee members with their unconditional help, and our sponsors and endorsers with their timely support. Finally, we would like to thank Springer for agreeing to publish the proceedings in their prestigious Advances in Intelligent Systems and Computing (AISC) series. We hope the technical contributions made by the authors in these volumes presenting the proceedings of CVIP 2017 will be appreciated by one and all.

Kolkata, India — Bidyut B. Chaudhuri
Singapore, Singapore — Mohan S. Kankanhalli
Roorkee, India — Balasubramanian Raman

Committees

General Chairs

Bidyut Baran Chaudhuri, ISI Kolkata, India
Mohan Kankanhalli, NUS, Singapore

Organizing Chairs

Balasubramanian Raman, IIT Roorkee, India
Sanjeev Kumar, IIT Roorkee, India
Partha Pratim Roy, IIT Roorkee, India
Vinod Pankajakshan, IIT Roorkee, India

Program Chairs

Debashis Sen, IIT Kharagpur, India
Sudipta Mukhopadhyay, IIT Kharagpur, India
Dilip Prasad, NTU, Singapore
Mukesh Saini, IIT Ropar, India

Workshop Chairs

Brajesh Kaushik, IIT Roorkee, India
N. Sukavanam, IIT Roorkee, India

Plenary Chairs

Dharmendra Singh, IIT Roorkee, India
R. D. Garg, IIT Roorkee, India

International Advisory Committee

A. G. Ramakrishnan, Indian Institute of Science, Bangalore, India
Arpan Pal, Tata Consultancy Services, Kolkata, India
B. Krishna Mohan, IIT Bombay, India

Gian Luca Foresti, University of Udine, Italy
Jonathan Wu, University of Windsor, Canada
Josep Lladós, Universitat Autònoma de Barcelona, Catalonia, Spain
Michael Blumenstein, Griffith University, Australia
Phalguni Gupta, IIT Kanpur, India
Pradeep Atrey, State University of New York, Albany, USA
Prem K. Kalra, IIT Delhi, India
Santanu Choudhury, IIT Delhi, India
Subhasis Chaudhuri, IIT Bombay, India
Umapada Pal, Indian Statistical Institute, Kolkata, India

Publication Chairs

Debi Prosad Dogra, IIT Bhubaneshwar, India
Rajarshi Pal, IDBRT Hyderabad, India
Biplab Banerjee, IIT Roorkee, India

Area Chairs

Ananda S. Chowdhury, Jadavpur University, India
Arnav Bhavsar, IIT Mandi, India
Christian Micheloni, University of Udine, Italy
Gaurav Bhatnagar, IIT Jodhpur, India
Ibrahim Venkat, Universiti Sains Malaysia, Malaysia
Kidiyo Kpalma, Institut National des Sciences Appliquées de Rennes, France
Maheshkumar H. Kolekar, IIT Patna, India
Pritee Khanna, IIIT Jabalpur, India
Rajiv Ratn Shah, IIIT Delhi, India
Shanmuganathan Raman, IIT Gandhinagar, India
Subrahmanyam Murala, IIT Ropar, India
Vijayan K. Asari, University of Dayton, USA

Technical Program Committee

A. V. Subramanyam, IIIT Delhi, India
Abhishek Midya, NIT Silchar, India
Ajoy Mondal, Indian Statistical Institute, India
Alireza Alaei, Université de Tours, France
Amanjot Kaur, IIT Ropar, India
Amit Kumar Verma, NIT Meghalaya, India
Ananda Chowdhury, Jadavpur University, India
Anil Gonde, Shri Guru Gobind Singhji Institute of Engineering and Technology, India
Anindya Halder, North-Eastern Hill University, India
Ankush Mittal, Graphic Era University, India
Ashis Dhara, IIT Kharagpur, India
Aveek Shankar Brahmachari, Stryker Global Technology Center, India

Badrinarayan Subudhi, NIT Goa, India
Bijaylaxmi Das, IIT Kharagpur, India
Dinabandhu Bhandari, Heritage Institute of Technology, India
Dwarikanath Mohapatra, IBM Research, Australia
Enmei, Nanyang Technological University, Singapore
Gan Tian, School of Computer Science and Technology, Shandong University, China
Gao Tao, North China Electric Power University, China
Gaurav Gupta, The NorthCap University, Gurgaon, India
Grace Y. Wang, Auckland University of Technology, New Zealand
Guoqiang, Ocean University of China, China
Harish Katti, Indian Institute of Science, Bangalore, India
Hemanth Korrapati, National Robotics Engineering Center, Carnegie Mellon University, USA
Himanshu Agarwal, Maharaja Agrasen Institute of Technology, New Delhi, India
Jatindra Dash, IIT Kharagpur, India
Jayasree Chakraborty, Research Fellow, Memorial Sloan Kettering Cancer Center, USA
K. C. Santosh, Department of Computer Science, The University of South Dakota, USA
Kaushik Roy, West Bengal State University, India
Krishna Agarwal, University of Tromso, Norway
Ma He, Northeastern University, Shenyang, China
Mandar Kale, IIT Kharagpur, India
Manish, University of Nantes, France
Manish Chowdhury, KTH, Sweden
Manish Narwaria, Dhirubhai Ambani Institute of Information and Communication Technology, India
Manoj Kumar, Babasaheb Bhimrao Ambedkar Central University, India
Manoj Thakur, IIT Mandi, India
Meghshyam G. Prasad, Kolhapur Institute of Technology, India
Minakshi Banerjee, RCC Institute of Information Technology, India
Naveen Kumar, NIT Kurukshetra, India
Nidhi Taneja, Indira Gandhi Delhi Technical University for Women, India
P. Shivakumara, University of Malaya, Malaysia
Padmanabha Venkatagiri, National University of Singapore, Singapore
Partha Pratim Kundu, Indian Statistical Institute, India
Prabhu Natarajan, National University of Singapore, Singapore
Puneet Goyal, IIT Ropar, India
S. K. Gupta, IIT Roorkee, India
Sankaraiah Sreeramula, Fusionex International, Malaysia
Santosh Vipparthi, MNIT Jaipur, India
Sarif Kumar Naik, Philips, India
Shrikant Mehre, IIT Kharagpur, India
Sobhan Dhara, NIT Rourkela, India

Subramanyam, IIIT Delhi, India
Suchi Jain, IIT Ropar, India
Sudhish N. George, NIT Calicut, India
Suman Mitra, Dhirubhai Ambani Institute of Information and Communication Technology, India
Tanmay Basu, Ramakrishna Mission Vivekananda University, India
Tu Enmei, Rolls-Royce, NTU, Singapore
Vijay Kumar B. G., Australian Centre for Robotic Vision, Australia
Vikrant Karale, IIT Kharagpur, India
Xiangyu Wang, Agency for Science, Technology and Research, Singapore
Ying Zhang, Institute for Infocomm Research, Singapore
Zhong Guoqiang, Ocean University of China, China

Publicity Chairs

Navneet Kumar Gupta, IIT Roorkee, India
Asha Rani, IIT Roorkee, India
Priyanka Singh, State University of New York, Albany, USA
Suresh Merugu, IIT Roorkee, India

Web site

Himanshu Buckchash, Webmaster, IEEE UP Section

Contents

About the Editors

Prof. Bidyut B. Chaudhuri is INAE Distinguished Professor and J. C. Bose Fellow of Computer Vision and Pattern Recognition Unit at Indian Statistical Institute, Kolkata. He received his B.Sc (Hons.), B.Tech, and M.Tech from Calcutta University, India, in 1969, 1972, and 1974, respectively, and Ph.D. from IIT Kanpur in 1980. He did his Postdoc work during 1981–82 from Queen's University, UK. He also worked as a Visiting Faculty at Tech University, Hannover, during 1986–87 as well as at GSF institute of Radiation Protection (now Leibnitz Institute), Munich, in 1990 and 1992. His research interests include digital document processing, optical character recognition; natural language processing including lexicon generation, ambiguity analysis, syntactic and semantic analysis in Bangla and other Indian languages; statistical and fuzzy pattern recognition including data clustering and density estimation; computer vision and image processing; application-oriented research and externally funded project execution; and cognitive science. He is a Life Fellow of IEEE, International Association for Pattern Recognition (IAPR), Third World Academy of Science (TWAS), Indian National Sciences Academy (INSA), National Academy of Sciences (NASc), Indian National Academy of Engineering (INAE), Institute of Electronics and Telecommunication Engineering (IETE), West Bengal Academy of Science and Technology, Optical Society of India, and Society of Machine Aids for Translation. He has published over 400 papers in journals and conference proceedings of national and international repute.

Dr. Mohan S. Kankanhalli is a Dean of School of Computing and Provost's Chair Professor of Computer Science at National University of Singapore. Before that, he was the Vice-Provost (Graduate Education) for NUS during 2014–2016 and Associate Provost (Graduate Education) during 2011–2013. He was earlier the School of Computing Vice-Dean for Research during 2001–2007 and Vice-Dean for Graduate Studies during 2007–2010. He obtained his B.Tech (Electrical Engineering) from IIT Kharagpur in 1986 and M.S. and Ph.D. (Computer and Systems Engineering) from the Rensselaer Polytechnic Institute in 1988 and 1990, respectively. He subsequently joined the Institute of Systems Science. His research interests include multimedia computing, information security, image/video processing, and social media analysis. He is a Fellow of IEEE. He has published over 250 papers in journals and conference proceedings of international repute.

Dr. Balasubramanian Raman is currently an Associate Professor in the Department of Computer Science and Engineering at IIT Roorkee, India. He completed his Ph.D. (Mathematics and Computer Science) from IIT Madras, Chennai, in 2001. His areas of interest include computer vision—optical flow problems, fractional transform theory, wavelet analysis, image and video processing, multimedia security: digital image watermarking and encryption, biometrics, content-based image and video retrieval, hyperspectral and microwave imaging and visualization, and volume graphics. He has published over 100 papers in refereed journals and contributed seven chapters in books.

A Novel Method for Logo Detection Based on Curvelet Transform Using GLCM Features

G. V. S. S. K. R. Naganjaneyulu, Ch Sai Krishna and A. V. Narasimhadhan

Abstract Automatic logo detection is a key tool for document retrieval, document recognition, document classification, and authentication. It helps in office automation as it enables the effective identification of source of a document. In this paper, a novel approach for logo detection using curvelet transform is proposed. The curvelet transform is employed for logo detection because of its ability to represent curved singularities efficiently when compared to wavelet and ridgelet transforms. The proposed algorithm consists of five steps, namely segmentation, noncandidate elimination, computation of curvelet coefficients, gray level co-occurrence matrix (GLCM) features extraction, followed by classification using a pretrained support vector machine classifier. The proposed algorithm is tested on a standard dataset, and the performance is compared with the state-of-the-art methods. The results show good improvement in the accuracy when compared with the competitors.

Keywords Logo detection · Curvelet transform · GLCM features

1 Introduction

The detection of logo can be considered as an important clue for document image analysis and retrieval. Logos are special visual objects that are commonly used in business and government documents as a declaration of document source and ownership. Logos are generally used to aid and promote instant public recognition for an organization. The document retrieval is easier using logo recognition than a keyword search. Methods for document retrieval using logo detection are presented in [12, 18, 29, 37]. A method for categorization of documents based on logo detection is described in [26].

G. V. S. S. K. R. Naganjaneyulu (✉) · A. V. Narasimhadhan
National Institute of Technology Karnataka, Surathkal, India
e-mail: snd.786@gmail.com

C. Sai Krishna
Rajiv Gandhi University of Knowledge Technologies, Krishna, Nuzvid, India

B. B. Chaudhuri et al. (eds.), *Proceedings of 2nd International Conference on Computer Vision & Image Processing*, Advances in Intelligent Systems and Computing 704, https://doi.org/10.1007/978-981-10-7898-9_1

Many algorithms have been developed by several researchers for logo detection and recognition. A brief literature review is presented in Table 1. These methods are broadly categorized into three types, namely heuristic methods [16, 21, 27, 31, 32], spatial density-based methods [2, 25, 28], and multi-resolution methods [36].

1.1 Heuristic Methods

Seiden et al. [27] used the top-down X-Y cut segmentation algorithm to analyze a binary document to extract 16 features from each connected component, and a rule-based classifier for classification of the logos. Wang et al. [32] extracted all feature rectangles from the entire document to detect the logos. A feature rectangle is a minimum virtual rectangle which fully embraces at least one foreground pixel and their edges have all background pixels. Once all the feature rectangles are extracted from whole document image, classification is performed using a decision tree classifier.

1.2 Spatial Density-Based Methods

Pham [25] and Ahmed [2] proposed logo detection methods that are based on spatial density of foreground pixels within a given window. The fundamental assumption in these methods is that the spatial density related to logo regions is greater than that of non-logo regions. The preprocessing steps are followed by the estimation of spatial density using mountain function [33], and classification is performed using a decision tree classifier.

1.3 Multi-resolution Analysis-Based Methods

Shirdhonkar et al. [28] used discrete wavelet transform (DWT) [24] to calculate spatial density. The document image is divided into different nonoverlapping blocks of fixed size. The DWT coefficients of these blocks of the document image are computed. Using the DWT coefficients, two features namely energy and standard deviation of each window from all the sub-bands are computed to differentiate logo and non-logo blocks in the document image. In [28], author made an assumption that the complete logo is present in a single block. A segmentation step can be used to avoid such an assumption. The DWT coefficients can represent point singularities effectively, but fails to represent curve singularities. Ridgelet transform [6, 10] can represent line singularities; however, it also fails in representing curve singularities. Curvelet transform can represent the curved singularities in a better way when compared to wavelets and ridgelets [7, 22, 23, 30]. The logos generally contain curve singularities, and therefore curvelet transform is a good choice for detection of logos.

Table 1 Literature review of logo detection methods

Papers	Key features of algorithm
Hassanzadeh et al. [16]	Morphological operations, decision tree classifier, and merging separated parts in logo
Shirdhonkar et al. [28]	Discrete wavelet transform, spatial density using two features—mean and variance
Wang [31]	RAG data structure and Bayesian modeling
Li et al. [21]	Axial lines, shape descriptors
Wang et al. [32]	Feature rectangles
Ahmed et al. [2]	Colorimetric uniformity and spatial compactness using mountain function
Zhu et al. [36]	Multi-scale approach
Pham [25]	Mountain function
Seiden et al. [27]	X-Y cut segmentation, rule-based classifier
Doermann et al. [11]	Geometric invariants

On the other hand, in literature, an approach with two features and a decision tree-based classifier for logo detection is typically used. Usage of more number of features and a sophisticated classifier like support vector machine classifier (SVM) [8] can produce a better accuracy. In this work, logo detection is performed by means of segmentation, noncandidate elimination, computation of GLCM features from each sub-band of curvelet coefficients, followed by classification using SVM classifier.

The organization of the paper is as follows. In Sect. 2- theory related to the methods employed is presented. In Sect. 3, a brief description of the different blocks is provided. In Sect. 4, the performance analysis of the proposed method and comparison with the state of the art are analyzed, followed by conclusion in Sect. 5.

2 Theory

2.1 Fast Discrete Curvelet Transform

Curvelet transform [5] is a nonadaptive approach of representing images at multiple scales and multiple angles. It is a higher dimensional generalization of the wavelet transform and overcomes the limitation of wavelets in terms of representing curved singularities.

In the proposed method, fast discrete curvelet transform (FDCT) [4] is employed. The FDCT of an image is defined as,

$$c(j, \mathbf{k}, l) = \Sigma_{n_1, n_2 \epsilon P_j} \hat{f}[n_1, n_2 - n_1 \tan\theta_l] \tilde{U}_j[n_1, n_2] e^{i2\pi(\frac{k_1 n_1}{L_{1,j}} + \frac{k_2 n_2}{L_{2,j}})}. \tag{1}$$

where $c(j, \mathbf{k}, l)$ is representing discrete curvelet coefficient at scale 2^j, location $\mathbf{k} = [k_1 \quad k_2]$, angle θ_l and U_j (with length $L_{1,j}$ and width $L_{2,j}$) is a Cartesian (rectangular) generalization of the polar (circular) window that isolates the digital frequencies $[n_1, n_2]$ at the wedges, $\theta_l = 2\pi 2^{\lfloor \frac{-j}{2} \rfloor} l$, $l = 0, 1, 2, ...$ and $P_j = \{(n_1, n_2) : n_{1,0} \leq n_1 \leq n_{1,0} + L_{1,j}, n_{2,0} \leq n_2 \leq n_{2,0} + L_{2,j}\}$. The curvelet transform is used in texture analysis in [3].

2.2 GLCM Features

The GLCM is a tool to measure texture as it represents the distribution of the co-occurring values in an image. The texture features generated and used are Haralick features [15]. GLCM is a matrix in which the number of rows and columns is equal to the number of gray levels G in the image. The GLCM matrix element $P(i, j|\Delta x, \Delta y)$ is the relative frequency with which two pixels, separated by a pixel distance $(\Delta x, \Delta y)$, occur within a given neighborhood. Given an image I of size $N \times N$, the element of in ith, jth column of normalized GLCM matrix P can be defined as,

$$P(i, j) = \sum_{x=1}^{N} \sum_{y=1}^{N} \delta(I(x, y) - i)\delta(I(x + \Delta x, y + \Delta y) - j) \tag{2}$$

where $\delta(.)$ is impulse function. Let us assume that μ is the mean value of P, and μ_x, μ_y, σ_x, σ_y are the means and standard deviations of the marginal densities P_x, P_y. The marginal densities P_x and P_y are related to $P(i, j)$ as described below.

$$P_x(i) = \sum_{j=1}^{G} P(i, j), \ P_y(j) = \sum_{i=1}^{G} P(i, j) \tag{3}$$

The expressions for $P_{x+y}(k)$ and $P_{x-y}(k)$ are given below.

$$P_{x+y}(k) = \sum_{i=1}^{G} \sum_{j=1}^{G} P(i, j)_{i+j=k}; \ k = 2, 3, \ldots 2N \tag{4}$$

$$P_{x-y}(k) = \sum_{i=1}^{G} \sum_{j=1}^{G} P(i, j)_{|i-j|=k}; \ k = 0, 1, \ldots N - 1 \tag{5}$$

The 12 GLCM features used in this work are defined in Table 2. Angular second moment gives the homogeneity in the image. Contrast provides an insight into the local gray level variation in image, while inverse difference moment gives the similarity among nonzero entries in the image. The correlation represents gray value

Table 2 Definition of 12 features extracted from GLCM matrix

Angular second moment	$\sum_{i=0}^{G-1}\sum_{j=0}^{G-1}\{P(i,j)\}^2$
Contrast	$\sum_{n=0}^{G-1} n^2\{\sum_{i=1}^{G}\sum_{j=1_{\lvert i-j\rvert=n}}^{G} P(i,j)\}$
Correlation	$\sum_{i=0}^{G-1}\sum_{j=0}^{G-1}\frac{(i\times j)\times P(i,j)-(\mu_x\times\mu_y)}{\sigma_x\times\sigma_y}$
Variance	$\sum_{i=0}^{G-1}\sum_{j=0}^{G-1}(i-\mu)^2 P(i,j)$
Inverse difference moment	$\sum_{i=0}^{G-1}\sum_{j=0}^{G-1}\frac{1}{1+(i-j)^2}P(i,j)$
Entropy	$-\sum_{i=0}^{G-1}\sum_{j=0}^{G-1}P(i,j)\times log(P(i,j))$
Average	$\sum_{G}^{2G-2} i P_{x+y}(i)$
Sum entropy	$-\sum_{i=0}^{2G-2}P_{x+y}(i)log(P_{x+y}(i))$
Difference entropy	$-\sum_{i=0}^{G-1}P_{x-y}(i)log(P_{x-y}(i))$
Inertia	$\sum_{i=0}^{G-1}\sum_{j=0}^{G-1}(i-j)^2\times P(i,j)$
Shade	$\sum_{i=0}^{G-1}\sum_{j=0}^{G-1}(i+j-\mu_x-\mu_y)^3\times P(i,j)$
Prominance	$\sum_{i=0}^{G-1}\sum_{j=0}^{G-1}(i+j-\mu_x-\mu_y)^4\times P(i,j)$

linear dependencies in the image. Entropies give the statistical disorderliness in the image. The average and variance represent the central tendency and the spread in GLCM matrix (not of the image). Inertia, cluster prominence, and shade are higher order statistic related to texture. The GLCM features are used in various applications involving texture classification in [17, 35]. GLCM features are also used along with multi-scale transforms in [13, 34].

3 Description of the Algorithm

The block diagram of the proposed method is shown in Fig. 1. The proposed algorithm initially performs segmentation, using morphological operations [14] and connected component analysis (CCA) [9]. The improbable logo candidates are removed based on heuristic features in noncandidate elimination phase. Curvelet coefficients of the remaining logo candidates are obtained using FDCT. Subsequently, 12 features are computed from GLCM matrix obtained from each wedge of curvelet coefficients that are extracted in eight different directions in order to obtain the texture clues in all the directions. A pretrained SVM classifier is used for removal of false logo blocks using the GLCM features. The detailed explanation of each stage is discussed in the following subsections.

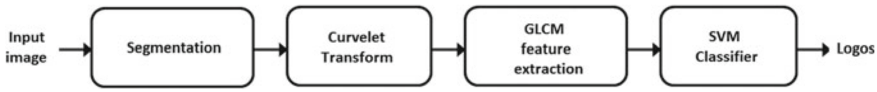


Fig. 1 Block diagram of the proposed method

3.1 Segmentation

Segmentation is the foremost step in logo detection. The inclusion of segmentation may increase the computational complexity; nevertheless, it gives an advantage of avoiding the assumption that the logo has to be in a single block under consideration, as mentioned in [28]. The proposed scheme employs a segmentation algorithm which is based on connected components analysis. In a document image, a segment is a portion which has all connected components near to each other. Hence, connecting the nearest connected components for clubbing them together to make a segment is achieved by applying morphological operations. A morphological dilation operation with a rectangular structuring element of size 15×90 is performed on the complement of the input binary image. This dilation operation is followed by a closing operation with the same structuring element in order to connect the similar patterns in the image. CCA is performed on the resulting image from the morphological operations to find the coordinates of all possible bounding boxes. A typical output of segmentation is shown in Fig. 2.

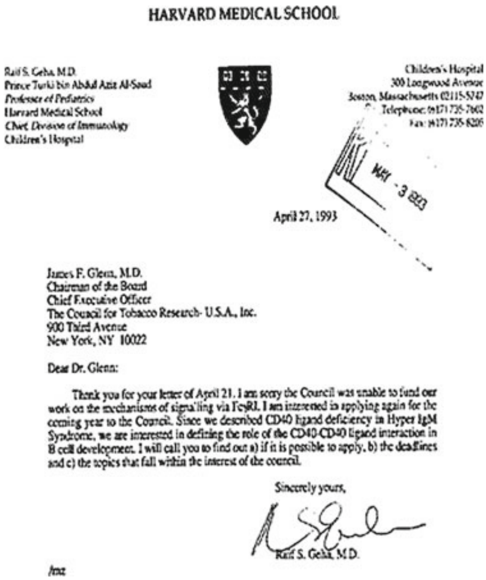
3.2 Noncandidate Elimination

The purpose of this step is to remove obvious non-logo candidates in order to reduce the load on classifier and computational complexity. Out of all the blocks obtained from CCA, the obvious false positives are removed using two heuristic clues, namely aspect ratio and foreground pixel density. The connected components with aspect ratio less than 3 and foreground pixel density greater than 70% are considered as logo candidates. These thresholds are obtained by computing the values of aspect ratio and spatial density of 100 logo objects. All probable logo candidates from segmentation step are first filtered with the aspect ratio criteria, and those do not satisfy the criteria are identified as non-logo candidates. The remaining probable logo candidates are further refined using the foreground pixel density criteria. Using the aspect ratio, clue large paragraphs containing text are removed, and using foreground density, criteria typically headings are eliminated as non-logo objects. The remaining possible logo candidates after the removal of non-logo candidates by enforcing the aforementioned constraints are considered for further processing.

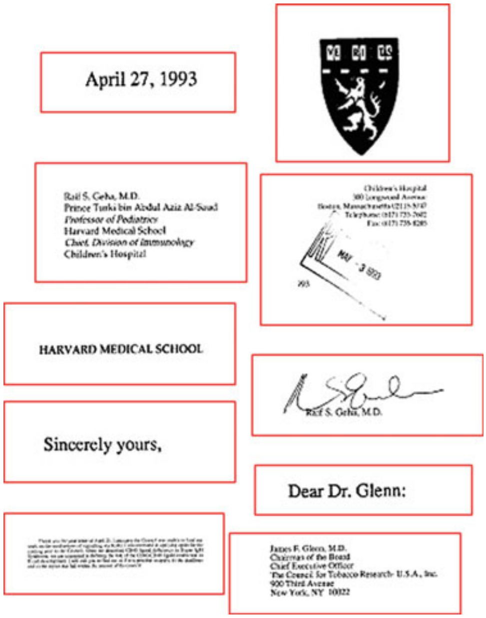
3.3 FDCT

For all the possible logo candidates, FDCT is calculated with number of scales equal to 4. The number of scales is chosen as 4 by trial and error procedure. The computation of FDCT is performed using the non-equispaced FFT method proposed in [4]. Total number of wedges resulting from the computation of FDCT in this work is 130.

Fig. 2 Example image and its segmented blocks



(a) Input image



(b) Corresponding output blocks of segmentation

3.4 GLCM Features

Texture in logos is an important clue as logos generally contain patterns inside them along with the text. The GLCM features along with the curvelet coefficients are used to extract the textural clues in the logos. In the proposed scheme, 12 features mentioned in Table 2 are extracted from each wedge of FDCT obtained from the previous step. A total of 130 wedges, 8 directions, and 12 features results in 12, 480 features per each block are computed as a part of feature extraction and used for training a SVM classifier.

3.5 SVM Classifier

A pretrained SVM classifier [8] is used to separate logo candidates from all the possible logo candidates. SVM classifier is a supervised binary classifier based on linear discriminants. It considers the extreme data samples near the boundary called support vectors and maximizes the margin between the classes. In nonlinear boundary case where it is not a linearly separable problem, SVM treats the nonlinear function as a linear function by projecting the data into a higher dimensional case. SVM provides an optimal decision for binary classification problem and is chosen for the binary classification problem of logo objects. In this work, 20 logos and 200 non-logos are used for training the SVM classifier with linear kernel which outputs the true logo objects.

4 Experiments and Results

The proposed method is tested on Tobacco-800 dataset[1] [1, 19, 20] which is composed of 1151 binary images. Tobacco 800 is a public dataset which is a part of IIT CDIP dataset that contains about 42 million document images collected from UCSF. This dataset is widely used by many researchers [2, 16, 25, 28, 36] for testing the logo detection methods. A total of 20 logo candidates and 200 non-logo candidates are used for training and the remaining for testing. Performance of the proposed work is evaluated using below performance metrics.

$$Accuracy = \frac{TP}{TP + TN} \tag{6}$$

$$Precision = \frac{TP + FN}{TP + FP + FN + TN} \tag{7}$$

[1]The database is available at http://www.umiacs.umd.edu/~zhugy/tobacco800.html.

Fig. 3 Examples of logos detected (true positives)

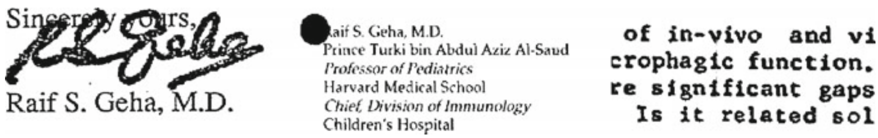


Fig. 4 Examples of non-logos detected as logos (false positives)

Table 3 Results of algorithm

Category	Logos	Non-logos
Training	20	200
Correctly classified	66	1747
Miss classified	6	29
Total	92	1976

where
TP: True positives (Total number of correctly detected logos)
TN: True negatives (Number of non-detected logos)
FP: False positives (Not a logo but algorithm detected it as logo)
FN: False negatives (Not a logo and the algorithm detected as non-logo).
Some of the correctly detected logos are shown in Fig. 3, and false positives are shown in Fig. 4.

The proposed algorithm gives an accuracy of 91.47% and precision of 98.1%. The results obtained using the proposed method are presented in Table 3. From Fig. 3, it can be seen that if there is a text block associated with the logo, the proposed method is able to obtain the logo completely along with the text part, which is an advantage. Figure 4 shows typical false positives, and most of the times they are cluttered text regions which almost look like a logo. The accuracy obtained using the proposed method is compared with the state-of-the-art algorithms and is given in Fig. 5. It can be observed that the proposed algorithm outperforms the other methods in terms of accuracy of detection of logos. The significant improvement in the accuracy can be attributed to the ability of curvelet transform in representing the curve discontinuities. The inclusion of segmentation in proposed method removes the assumption that entire logo needs to be in a single block which is fed to the classifier. However, the computational complexity increases because of inclusion of segmentation involving morphological operations.

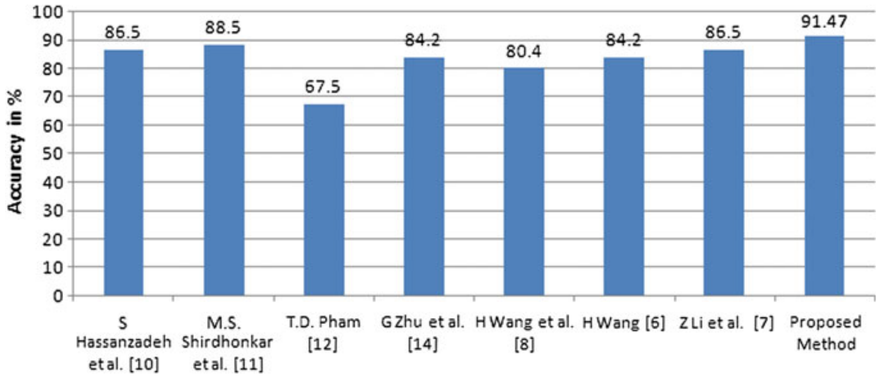


Fig. 5 Comparison of accuracies of different algorithms for logo detection

5 Conclusion

In this work, a novel approach for logo detection using curvelet transform has been proposed. The proposed algorithm gives an accuracy of 91.47% with a precision of 98.1%. It has been observed that the accuracy of the proposed algorithm is better than the other competitors with a competitive precision. The inclusion of segmentation automated the process of selection of the probable logo candidates. Consequently, the proposed method has removed the assumption that the logo must be in a single block that is sent to the classifier which is a drawback in other methods. A better precision has been obtained because of the noncandidate elimination. The algorithm is able to include the text part near to the logo, which is an advantage.

References

1. Agam, G., Argamon, S., Frieder, O., Grossman, D., Lewis, D.: The Complex Document Image Processing (CDIP) Test Collection Project. Illinois Institute of Technology (2006), http://ir.iit.edu/projects/CDIP.html
2. Ahmed, Z., Fella, H.: Logos extraction on picture documents using shape and color density. In: ISIE IEEE International Symposium on Industrial Electronics, pp. 2492–2496. IEEE (2008)
3. Arivazhagan, S., Ganesan, L., Kumar, S.T.: Texture classification using curvelet statistical and co-occurrence features. In: 18th International Conference on Pattern Recognition, ICPR, vol. 2, pp. 938–941. IEEE (2006)
4. Candes, E., Demanet, L., Donoho, D., Ying, L.: Fast discrete curvelet transforms. Multiscale Modeling & Simulation 5(3), 861–899 (2006)
5. Candes, E.J., Donoho, D.L.: Curvelets: A surprisingly effective nonadaptive representation for objects with edges. FPO AE 09499-0039 (1999)
6. Candès, E.J., Donoho, D.L.: Ridgelets: A key to higher-dimensional intermittency? Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences 357(1760), 2495–2509 (1999)

7. Candes, E.J., Donoho, D.L.: Curvelets: A surprisingly effective nonadaptive representation for objects with edges. Tech. rep., DTIC Document (2000)
8. Cortes, C., Vapnik, V.: Support-vector networks. Machine learning 20(3), 273–297 (1995)
9. Dillencourt, M.B., Samet, H., Tamminen, M.: A general approach to connected-component labeling for arbitrary image representations. Journal of the ACM (JACM) 39(2), 253–280 (1992)
10. Do, M.N., Vetterli, M.: The finite ridgelet transform for image representation. IEEE Transactions on Image Processing 12(1), 16–28 (2003)
11. Doermann, D.S., Rivlin, E., Weiss, I.: Logo recognition using geometric invariants. In: Proceedings of the Second International Conference on Document Analysis and Recognition, pp. 894–897. IEEE (1993)
12. Folkers, A., Samet, H.: Content-based image retrieval using fourier descriptors on a logo database. In: Proceedings, 16th International Conference on Pattern Recognition, vol. 3, pp. 521–524 (2002)
13. Francis, S.V., Sasikala, M., Saranya, S.: Detection of breast abnormality from thermograms using curvelet transform based feature extraction. Journal of medical systems 38(4), 1–9 (2014)
14. Gonzalez, R.C., et al.: Re woods, digital image processing. Addison–Wesely Publishing Company (1992)
15. Haralick, R.M.: Statistical and structural approaches to texture. Proceedings of the IEEE 67(5), 786–804 (1979)
16. Hassanzadeh, S., Pourghassem, H.: Fast logo detection based on morphological features in document images. In: IEEE 7th International Colloquium on Signal Processing and its Applications (CSPA), pp. 283–286. IEEE (2011)
17. Jafarpour, S., Sedghi, Z., Amirani, M.C.: A robust brain mri classification with glcm features. International Journal of Computer Applications 37(12), 1–5 (2012)
18. Jain, R., Doermann, D.: Logo retrieval in document images. In: Document Analysis Systems (2012)
19. Legacy: The Legacy Tobacco Document Library (LTDL). University of California, San Francisco (2007), http://legacy.library.ucsf.edu/
20. Lewis, D., Agam, G., Argamon, S., Frieder, O., Grossman, D., Heard, J.: Building a test collection for complex document information processing. In: Proc. 29th Annual Int. ACM SIGIR Conference, pp. 665–666 (2006)
21. Li, Z., Schulte-Austum, M., Neschen, M.: Fast logo detection and recognition in document images. In: 20th International Conference on Pattern Recognition (ICPR), pp. 2716–2719. IEEE (2010)
22. Ma, J., Plonka, G.: The curvelet transform. Signal Processing Magazine, IEEE 27(2), 118–133 (2010)
23. Ma, J., Plonka, G.: A review of curvelets and recent applications. IEEE Signal Processing Magazine 27(2), 118–133 (2010)
24. Mallat, S., Zhong, S.: Characterization of signals from multiscale edges. IEEE Transactions on Pattern Analysis & Machine Intelligence (7), 710–732 (1992)
25. Pham, T.D.: Unconstrained logo detection in document images. Pattern recognition 36(12), 3023–3025 (2003)
26. Rusinol, M., Llados, J.: Logo spotting by a bag-of-words approach for document categorization. In: 10th International Conference on Document Analysis and Recognition, ICDAR, pp. 111–115. IEEE (2009)
27. Seiden, S., Dillencourt, M., Irani, S., Borrey, R., Murphy, T.: Logo detection in document images. In: Proceedings of the International Conference on Imaging Science, Systems, and Technology, pp. 446–449 (1997)
28. Shirdhonkar, M., Kokare, M.: Automatic logo detection in document images. In: IEEE International Conference on Computational Intelligence and Computing Research (ICCIC), pp. 1–3. IEEE (2010)
29. Shirdhonkar, M., Kokare, M.B.: Automatic logo based document image retrieval. International Journal of Knowledge-based and Intelligent Engineering Systems 19(1), 27–33 (2015)

30. Starck, J.L., Candes, E., Donoho, D.: The curvelet transform for image denoising. IEEE Transactions on Image Processing, 11(6), 670–684 (Jun 2002)
31. Wang, H.: Document logo detection and recognition using bayesian model. In: 20th International Conference on Pattern Recognition (ICPR), pp. 1961–1964. IEEE (2010)
32. Wang, H., Chen, Y.: Logo detection in document images based on boundary extension of feature rectangles. In: 10th International Conference on, Document Analysis and Recognition, ICDAR, pp. 1335–1339. IEEE (2009)
33. Yager, R.R., Filev, D.P.: Approximate clustering via the mountain method. IEEE Transactions on Systems, Man and Cybernetics 24(8), 1279–1284 (1994)
34. Youssef, S.M.: Ictedct-cbir: Integrating curvelet transform with enhanced dominant colors extraction and texture analysis for efficient content-based image retrieval. Computers & Electrical Engineering 38(5), 1358–1376 (2012)
35. Zhang, J., Tong, L., Wang, L., Li, N.: Texture analysis of multiple sclerosis: a comparative study. Magnetic resonance imaging 26(8), 1160–1166 (2008)
36. Zhu, G., Doermann, D.: Automatic document logo detection. In: Ninth International Conference on Document Analysis and Recognition, ICDAR, vol. 2, pp. 864–868. IEEE (2007)
37. Zhu, G., Doermann, D.: Logo matching for document image retrieval. In: 10th International Conference on Document Analysis and Recognition, ICDAR, pp. 606–610. IEEE (2009)

Bayesian Approach for Landslide Identification from High-Resolution Satellite Images

Pilli Madalasa, Gorthi R K Sai Subrahmanyam, Tapas Ranjan Martha, Rama Rao Nidamanuri and Deepak Mishra

Abstract Landslides are one of the severe natural catastrophes that affect thousands of lives and cause colossal damage to infrastructure from small to region scales. Detection of landslide is a prerequisite for damage assessment. We propose a novel method based on object-oriented image analysis using bi-temporal satellite images and DEM. The proposed methodology involves segmentation, followed by extraction of spatial and spectral features of landslides and classification based on supervised Bayesian classifier. The proposed framework is based on the change detection of spatial features which capture the spatial attributes of landslides. The proposed methodology has been applied for the detection and mapping of landslides of different sizes in selected study sites in Himachal Pradesh and Uttarakhand, India. For this, high-resolution multispectral images from the IRS, LISS-IV sensor and DEM from Cartosat-1 are used in this study. The resultant landslides are compared and validated with the inventory landslide maps. The results show that the proposed methodology can identify medium- and large-scale landslides efficiently.

P. Madalasa · R. R. Nidamanuri
Department of Earth and Space Sciences, Indian Institute of Space Science and Technology, Trivandrum, India

G. R. K. Sai Subrahmanyam (✉)
Department of Electrical Engineering, Indian Institute of Technology Tirupati, Tirupati, Andhra Pradesh, India
e-mail: rkg@iittp.ac.in

D. Mishra
Department of Avionics, Indian Institute of Space Science and Technology, Thiruvananthapuram, India

T. R. Martha
Geosciences Group, National Remote Sensing Centre, Hyderabad, India

B. B. Chaudhuri et al. (eds.), *Proceedings of 2nd International Conference on Computer Vision & Image Processing*, Advances in Intelligent Systems and Computing 704, https://doi.org/10.1007/978-981-10-7898-9_2

1 Introduction

Natural catastrophes such as landslides, avalanches, floods cause acute environmental problems, substantial loss of life and colossal damage to infrastructure in mountain regions. Landslides are ranked third in terms of death due to natural calamity. Landslide is defined as the movement of a mass of earth or debris or rock from a higher altitude due to gravitational force [1]. The most vulnerable regions of landslides in India are the Western Ghats and the Himalayas. Western Ghats are geologically stable and are on uplifted plateau. The Himalayan mountain region is one of the youngest mountains which is tectonically active. While Western Ghats experience intermittent landslides, Himalayan regions experience huge and massive landslides frequently with impact sites spreading across several north and north-eastern states of India. Every year, one person dies for every 100 sq kms in Himalayan regions due to landslides. Overall, $\sim$12.6% of total land area is affected by landslides in India. Spatial prediction and mapping of landslides are valuable for disaster preparedness and management. Even though landslides are highly unpredictable, there is a spatial pattern in their occurrences and landslide can occur within a neighbourhood multiple times. Therefore, the detection of sites of historical landslides helps to predict the occurrences of landslides and is vital for assessing the damage caused to landscape.

The conventional methods of landslide detection mainly rely on the visual interpretation of aerial and satellite images and aided by moderate field measurements. Traditionally, interpretation can be done on vertical or oblique stereoscopic aerial photographs. In [2] nearly 60,000 earthquake-induced landslide scarps were mapped by visual interpretation. Visual interpretation is subjective, labour-intensive and time-consuming. Digital methods for analysis of remote sensing satellite images for landslide mapping have been evolving [3]. From the satellite imagery landslide regions are identified based upon its spectral, spatial, contextual information and morphometric feature [4]. In [4] a study, object regions in multispectral images pertinent to landslides are identified based on spectral information, shape and morphometric characteristics. The resultant objects are recognised as landslides in further classification and are categorised as debris slides, debris flows and rock slides using adjacency and morphometric criteria.

Landslide regions can be identified based upon a change detection analysis. Landslides are also mapped by tracking the differences in orthorectified co-registered images and digital elevation models (DEMs) acquired over the same geographical position at different times [4]. Using a change detection-based Markov random field (CDMRF) technique, [5] developed a near-automatic method for landslide mapping from bi-temporal aerial photographs at pixel level. With the advent of computational power, especially machine learning techniques for remote sensing image classification have gradually shifted from the visual interpretation to computer-aided methods. In [6] developed a semi-automated landslide detection and mapping, which combines OBIA and SVM classifier. In this, the classifier was trained using the spectral, spatial and texture features.