Lecture Notes in Computational Vision and Biomechanics 25

Jude Hemanth
Valentina Emilia Balas *Editors*

Biologically Rationalized Computing Techniques For Image Processing Applications

Springer

Lecture Notes in Computational Vision and Biomechanics

Volume 25

The research related to the analysis of living structures (Biomechanics) has been a source of recent research in several distinct areas of science, for example, Mathematics, Mechanical Engineering, Physics, Informatics, Medicine and Sport. However, for its successful achievement, numerous research topics should be considered, such as image processing and analysis, geometric and numerical modelling, biomechanics, experimental analysis, mechanobiology and enhanced visualization, and their application to real cases must be developed and more investigation is needed. Additionally, enhanced hardware solutions and less invasive devices are demanded.

On the other hand, Image Analysis (Computational Vision) is used for the extraction of high level information from static images or dynamic image sequences. Examples of applications involving image analysis can be the study of motion of structures from image sequences, shape reconstruction from images, and medical diagnosis. As a multidisciplinary area, Computational Vision considers techniques and methods from other disciplines, such as Artificial Intelligence, Signal Processing, Mathematics, Physics and Informatics. Despite the many research projects in this area, more robust and efficient methods of Computational Imaging are still demanded in many application domains in Medicine, and their validation in real scenarios is matter of urgency.

These two important and predominant branches of Science are increasingly considered to be strongly connected and related. Hence, the main goal of the LNCV&B book series consists of the provision of a comprehensive forum for discussion on the current state-of-the-art in these fields by emphasizing their connection. The book series covers (but is not limited to):

- Applications of Computational Vision and Biomechanics
- Biometrics and Biomedical Pattern Analysis
- Cellular Imaging and Cellular Mechanics
- Clinical Biomechanics
- Computational Bioimaging and Visualization
- Computational Biology in Biomedical Imaging
- Development of Biomechanical Devices
- Device and Technique Development for Biomedical Imaging
- Digital Geometry Algorithms for Computational Vision and Visualization
- Experimental Biomechanics
- Gait & Posture Mechanics
- Multiscale Analysis in Biomechanics
- Neuromuscular Biomechanics
- Numerical Methods for Living Tissues
- Numerical Simulation
- Software Development on Computational Vision and Biomechanics
- Grid and High Performance Computing for Computational Vision and Biomechanics
- Image-based Geometric Modeling and Mesh Generation
- Image Processing and Analysis
- Image Processing and Visualization in Biofluids
- Image Understanding
- Material Models
- Mechanobiology
- Medical Image Analysis
- Molecular Mechanics
- Multi-Modal Image Systems
- Multiscale Biosensors in Biomedical Imaging
- Multiscale Devices and Biomems for Biomedical Imaging
- Musculoskeletal Biomechanics
- Sport Biomechanics
- Virtual Reality in Biomechanics
- Vision Systems

More information about this series at http://www.springer.com/series/8910

Jude Hemanth · Valentina Emilia Balas
Editors

Biologically Rationalized Computing Techniques For Image Processing Applications

Springer

Editors
Jude Hemanth
Karunya University
Coimbatore
India

Valentina Emilia Balas
Faculty of Engineering
Aurel Vlaicu University of Arad
Arad
Romania

ISSN 2212-9391 ISSN 2212-9413 (electronic)
Lecture Notes in Computational Vision and Biomechanics
ISBN 978-3-319-61315-4 ISBN 978-3-319-61316-1 (eBook)
DOI 10.1007/978-3-319-61316-1

Library of Congress Control Number: 2017945695

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Contents

Artificial Bee Colony Algorithm for Classification of Semi-urban LU/LC Features Using High-Resolution Satellite Data

J. Jayanth, T. Ashok Kumar, Shivaprakash Koliwad and V.S. Shalini

Abstract Attempts to classify high-resolution satellite data with conventional classifier show limited success since the traditional-per-pixel classifiers examine only the spectral variance ignoring the spatial distribution of the pixels corresponding to the land use/land cover classes. The work is carried out in two stages on panchromatic sharpened IRS P-6 LISS-IV (2.5 m) multispectral (MS) imagery of the year 2014 of Mangalore coastal zone along the west coast of Karnataka state of India. In the first stage, in order to overcome the limitations experienced in the parametric and nonparametric classifications, the swarm intelligence optimisation technique based on Artificial Bee Colony (ABC) algorithm has been studied for twelve land cover classes that are mapped. In the second stage, to bring out a greater separability between the spectrally overlapping classes, a texture-based image classification approach has been introduced and a methodology is developed to determine the optimal window size, interpixel distance and the best combinations of texture bands in multispectral data. The five texture measures, viz. entropy (ENT), angular second moment (ASM), contrast (CON), MEAN and homogeneity (Hmg) derived from the grey-level co-occurrence matrix (GLCM), are investigated

J. Jayanth (✉)
Department of Electronics and Communication Engineering,
GSSS Institute of Engineering & Technology for Women,
Mysore, Karnataka 570 016, India
e-mail: jayanthnov8@gmail.com

T. Ashok Kumar
PES Institute of Technology and Management, Shivamogga,
Karnataka 577 204, India
e-mail: ashokkumar0868@gmail.com

S. Koliwad
Department of Electronics & Communication Engineering,
Malnad College of Engineering, Hassan, Karnataka 573 202, India
e-mail: spksagar2006@gmail.com

V.S. Shalini
Department of Electronics and Communication Engineering,
ATME College of Engineering, Mysore, Karnataka 570 028, India
e-mail: shalinivs29@gmail.com

J. Hemanth and V.E. Balas (eds.), *Biologically Rationalized Computing Techniques For Image Processing Applications*, Lecture Notes in Computational Vision and Biomechanics 25, DOI 10.1007/978-3-319-61316-1_1

in the study. The major observations and contributions of this work are as follows: in the first stage, the image classifier employing the ABC algorithm exhibits higher classification accuracy when compared with maximum likelihood classifier. In the second stage, the results show that combining textural features and spectral bands in classification approach has proven very useful in delineating the spectrally overlapping classes, particularly at higher class hierarchy level.

Keywords Artificial Bee Colony · MLC · GLCM · Texture

1 Introduction

In recent years, remote sensing data classification became attractive due to its technical, economic and environmental benefits. Basically, this process is difficult because regions of a landscape are frequently not properly segregated which may lead to overlap in the specified region of interest. Precisely, for each pixel, grey-level values are allocated depending on the mediocre reflectance of different land cover classes on the given study area. Supervised, unsupervised and semi-supervised are the three popular learning techniques for land cover classification [1–3]. Classification of Remotely Sensed (RS) data has been carried out by several authors through computational artificial intelligence techniques such as fuzzy logic [4], neural network [5], support vector machine [6], K-means [7]. Metaheuristics have been also widely used for remote sensing data classification, for example particle swarm optimisation [8], ant colony optimisation [9], bee colony [10], artificial immune system [11] and genetic algorithm [12].

RS data classification for specified boundaries of classes with overlapping regions has been a challenging task for the research communities [13, 14]. This provides a reason to develop a classification model, which takes concern of this issue. Literature survey also indicates that texture information [5, 6, 8, 9, 15], surface temperature [6], digital elevation model (DEM) [12] have been studied as ancillary data in the RS data classification which is captured by different sensors. Among the several types of ancillary data, the grey-level co-occurrence matrix (GLCM)-based texture statistics derived from the image data are reported to be performing satisfactorily in RS classification [11]. Ultimate goal in all classification techniques is to best utilise the spectral, spatial and temporal resolutions and polarisation signature of the data and other inherent characteristics associated with it, and devise classification techniques which show improvements in accuracy, stability and speed. The motivation behind this work is to provide a faultless classifier, which can furnish accurate information about various land use/land cover (LU/LC) which is a prerequisite and indispensable issue in the efficient monitoring and management of natural resources and quick delivery of the end products. But it seems that the present traditional approach of classification fails to make the best use of the rich data obtained through the modern sensor and cripples in taking advantage of the capabilities of today's high-speed computational machines. In the above context, the current research work is primarily

aimed at developing an efficient and reliable classification strategy by integrating advanced image processing techniques through swarm intelligence to facilitate improving classification accuracy of high spatial resolution of RS data over semi-urban LU/LC features. In this regard, ABC algorithm is used to classify data and image texture analysis is used for characterising the spatial variations for extracting the information from the data.

2 Maximum Likelihood Classifiers

The maximum likelihood classifier (MLC) is one of the most commonly used algorithms for supervised classifications in RS data analysis [5–8]. Classification error is been reduced in MLC algorithm, assuming that the spectral data in the content has been distributed normally. Due to this impression, analyst chooses the most homogenous training sites by picking up many smaller training sites rather than few larger training sites. Due to this assumption, distribution of classes is been categorised depending on the pattern which is entirely outlined by the *MEAN vector* and the *covariance matrix* [6]. Hence, MLC needs a training data set that contains a relationship between the multivariate object properties and known classes. Jayanth J et al. [8] mentioned that MLC algorithm provides less accurate results for various land cover classes. MLC algorithm provides an accurate result when compared with parallelepiped classifier, but computational time is more [9]. However, one of the major drawbacks in MLC is the complexity involved due to multimodal distribution in the classes which may lead to poor classified results [7]. Hence, MLC method itself cannot solve the problem of spectral confusion in RS classification.

3 Artificial Bee Colony

In 2005, Dervis Karaboga defined Artificial Bee Colony (ABC) algorithm, which has been inspired by the intelligent behaviour of honey bees. Bees choose their food source in search space by flying in different aspects through waggle dance or randomly without using experience. When the bees are updated with their new nectar value, it is compared with old nectar content which is stored in the memory and updates with the highest nectar value. To manage the balance exploration and exploitation process, local search and global search methods are combined to validate and test the instance.

- Pixels (DN values) in the image are represented by bee.
- LC classes, such as pool, grass dry, are the food sources.

The four essential elements in the proposed algorithm are: initialising a set of features, fitness function, local search strategy and prediction strategy.

Classification procedure of ABC algorithm is shown in Fig. 1.

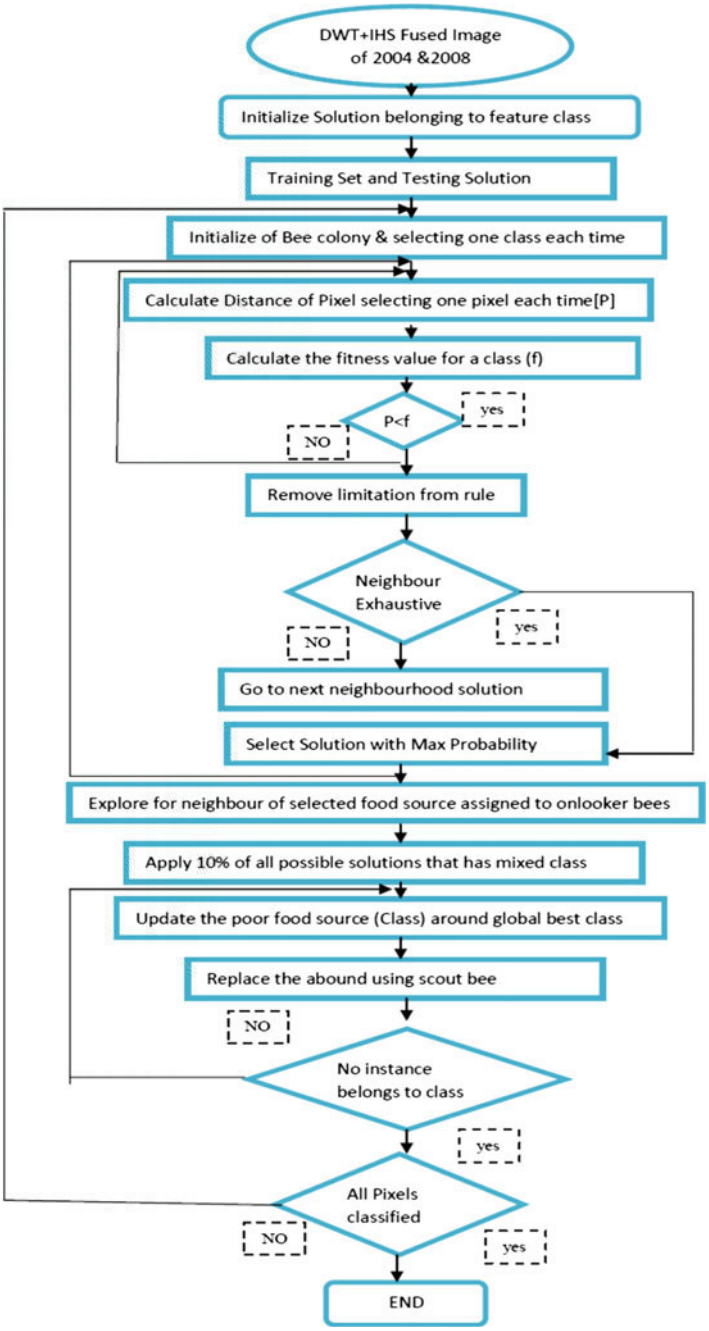

Fig. 1 Flow chart for Artificial Bee Colony algorithm classifier

3.1 Advantages of ABC

The principal advantage of ABC classifier over the traditional classifiers is as follows: bees being very optimal well-defined workers, they distribute the workload among themselves and their dancing behaviour helps in optimal design.

- ABC obeys proximity principle which responds to quality factor of the class.
- ABC can handle high-dimensional data and represents the acquired knowledge from waggle dance which is easy to intuit and assimilate.
- Simple and able to handle noisy data by creating a lower bound and upper bound for all the attributes.

3.2 Factors Affecting the Performance of the Artificial Bee Colony

Only the most commonly affecting issues are dealt with the following:

- Since bees do not depend on prior assumption about the data, it interrupts the performance of training data at the end of the list while assigning task to the agent, and may affect the performance of the classifier.
- ABC uses sequential covering algorithm, and hence, a small variation in data can make classification unstable and cause overfitting inside the data.
- Since each task has been selected by agents, the updating of agents from new task and its resource allocated to its agents can consume classification time.

4 Extraction of Textural Features

Texture represents the pixel relationship between spatial values and their grey-level properties which carries useful information for feature discrimination purpose in images. There is no single, unambiguous, widely accepted definition of texture in the literature. Analysis of texture in an image is referred to a mathematical model which shows differences in spatial characteristics of the image as a means of information extraction as shown in Table 1.

Since the first principal component (PC-1) contains the information which is common to all bands and accounts for maximum variance by removing the redundancy of information [16], the PC-1 image was extracted from the Linear Image Sensing Sensor (LISS)-IV data of 2.5 m spatial resolution to obtain GLCM and derive textural features thereof. All the textural features were extracted in

Table 1 Definitions of the texture measures based on GLCM

$\text{Entropy} = \sum_{i=0}^{G-1}\sum_{j=0}^{G-1} p(i,j)\log_2[p(i,j)]$	$\text{Mean}: \sum_{i=0}^{G-1}\sum_{j=0}^{G-1} ip(i,j)$
$\text{Energy (ASM)} = \sum_{i=0}^{G-1}\sum_{j=0}^{G-1}[p(i,j)]^2$	$\text{Homogenity(IDM)}: \sum_{i=0}^{G-1}\sum_{j=0}^{G-1}\frac{[P(i,j)]^2}{1+(i-j)^2}$
$\text{Contrast (Inertia)}: \sum_{i=0}^{G-1}\sum_{j=0}^{G-1}(i-j)^2 p(i,j)$	

ENT Entropy; *ASM* Angular second moment; *CON* Contrast; MEAN; *Hmg* Homogeneity

Geomatica V 10.0 at grey level (quantization levels) 16. We observed that the higher the grey level, the darker the image is and no marked difference is seen [7]. In the first phase of the work, five texture features (ENT, ASM, CON, MEAN and Hmg) [8] were computed from the GLCM at window sizes (W): 5 × 5, 7 × 7, 11 × 11, 17 × 17, 25 × 25, 35 × 35 and 51 × 51 at $\theta = 0°$ (assuming texture to be isotropic) keeping interpixel distance (D) unity. The above five features were also extracted at D: 3 and 7 keeping W: 25 × 25 for ENT, ASM, CON and Hmg, and 35 × 35 for MEAN.

Despite the fact that textural features exhibit non-Gaussian distribution of data, and Jeffries–Matusita distance performs better as a feature selection criterion for multivariate normal (Gaussian) classes [12], we employed the transformed divergence (TD) and Jeffries–Matusita (JM) distance measure as feature selection criterion for determining the best combination of multispectral bands and textural features. The JM distance varies between 0 and 1414, where 0 signifies no separability and the upper bound 1414 indicates high separability. Both the TD and JM distance measures are based on the measure of the statistical separation between pairs of class signatures [2], and the same has been adopted by [1–4].

5 Materials

5.1 Data Products Used

Indian Remote Sensor (IRS)-P6 LISS-IV satellite data captured on 16 April 2004 and 26 December 2007 consisting of three multispectral (MS, 5.8 m spatial resolution) bands recorded at green (0.52–0.59 μm), red (0.62–0.68 μm) and infrared (0.77–0.86 μm) wavelengths, and a panchromatic imagery of CARTOSAT-1 captured on 7 January 2008 (2.5 m spatial resolution); Census Data from Mangalore Urban Development Authority have been used in this study. All the satellite data are geo-referenced and projected with reference to global positioning system (GPS) readings.

Table 2 LU/LC classification hierarchy levels and details of the training and validation sites

Level-I	Level-II	Training sites		Validation sites
		MLC	Strata	
Built-up	RCC and asbestos sheet-roofing (RCC_sheet)	1032	11	239
	Clay tiled roofing (tiled_roof)	441	11	42
Roads	National highway (NH)	355	4	68
	Interior tar roads (int_rd)	490	5	147
Wastelands	Open ground/playgrounds (open_gnd)	1080	4	34
	Beach sand (sand)	270	2	21
Water	Pool water (pool)	93	1	18
	Sea water (sea)	884	2	19
Veg	Veg_mix[a] (veg_mix)	1848	6	178
	Acacia tree clusters (acacia)	1661	5	43
Grass dry	Grass dry (dry_grass)	424	3	66
	Total:	**8578**	**54**	**875**

[a]Thick leaf/fruit yielding trees such as mango, jack fruit, coconut, banyan

5.2 Study Area

The area under investigation is the Mangalore coastal region on the coastal belt of Karnataka state, India, between 12° 51′ 32″–12° 57′ 44″ N latitudes and 74° 51′ 30″–74° 48′01″ E longitudes. The image dimension of the study area is 1664 × 2065 pixels in MS data.

6 Results and Discussion

The objective of the behavioural study is to investigate the response of the Artificial Bee Colony (ABC) algorithm for classification on hierarchy level I and II over a semi-urban LU/LC area (coastal region of mangalore district) is shown in Table 2. The secondary objective of this analysis is to investigate the effectiveness of integrating GLCM-derived texture measures, viz. ENT, ASM, CON, MEAN and Hmg in classifying the panchromatic sharpened data of 2.5 m resolution at class hierarchy level-I and level-II, and to determine the optimal window size (W) and interpixel distance (D) for the texture measures under investigation.

6.1 Performance of ABC at Class Hierarchy Level-I and Level-II

The study was extended to compare the performance of both the classifiers at class hierarchy level-I and level-II. Hence, the training data sets, correspondingly the

Table 3 Classwise comparison of the accuracy of MLC and ABC at class hierarchy level-I

Class name	Producer's accuracy (%)		User's accuracy (%)	
	Classifier type			
	MLC	ABC	MLC	ABC
Grass_dry	72.18	73.68	93.20	91.59
Wasteland	91.11	88.89	96.47	95.24
Road	73.28	88.55	57.14	65.91
Built-up	70.05	82.35	77.51	77.39
Veg	94.58	81.28	87.67	92.70
Water	100	100	100	100
OCA (%)	81.42	83.44		

Table 4 Classwise comparison of the accuracy of MLC and ABC at hierarchy level-II

Class name	Producer's accuracy (%)		User's accuracy (%)	
	Classifier type			
	MLC	ABC	MLC	ABC
Grass_dry	75.94	69.92	90.99	92.08
Open_gnd	83.02	86.79	89.80	93.88
Rd_interior	83.70	83.70	60.63	62.10
Roof_RCC	85.00	83.75	79.07	83.75
Roof_Sheet	81.08	89.19	76.92	70.21
Roof_tiled	82.86	71.43	50.00	60.98
Sand	89.19	89.19	100	97.06
Veg_acacia	52.00	74.40	89.04	79.49
Veg-Mix	76.92	80.77	84.51	87.50
Water_Pool	100	100	93.75	100
Water_Sea	100	100	100	100
OCA (%)	77.88	80.03		

validation sets, were merged into 6 classes from 11. The classification results are tabulated in Tables 3 and 4.

At CHL-I, ABC shows an improvement of 3% OCA when compared with MLC algorithm (ABC 83.44% and MLC 81.42%). At class hierarchy level-I (CHL-I), ABC shows an improvement of 15.2% in producer's accuracy (PA) (88.55%) and 8.77% in user accuracy (UA) (65.91%) over MLC (PA 73.28% and UA 57.14%) for urban class *road*. Further, for class *built-up*, the ABC shows an improvement of 12.3% in PA over MLC. For the class wasteland, ABC shows 2.2 and 1.2% lesser PA and UA when compared with MLC. The class *veg* (vegetation) has experienced a trade-off in its UA and PA between both the classifiers. Hence, in both the class hierarchy levels, the ABC is able to maintain approximately 2% higher OCA (2.02 and 2.15% CHL-I and level-II, respectively) over MLC. It would also a

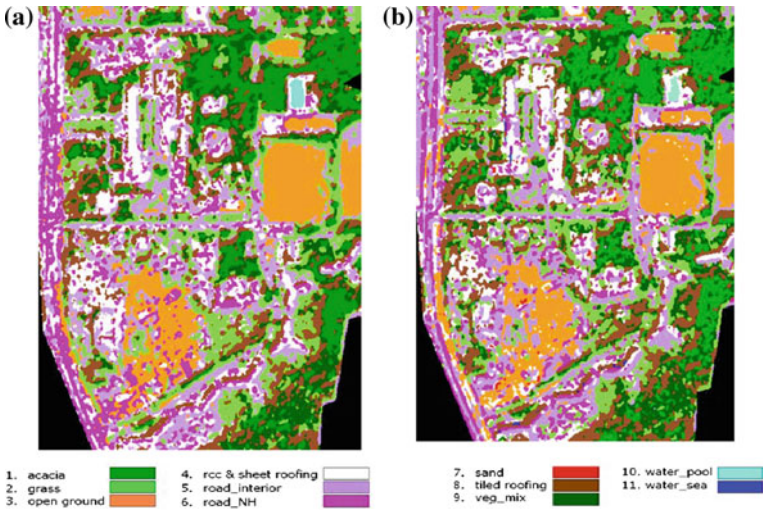

Fig. 2 Land cover classification in the area of Mangalore using **a** MLC and **b** ABC

noteworthy observation that as the class hierarchy level is increased from 6 classes to 11 classes, the OCAs of both the classifiers are found to be decreasing by an amount of 3–3.5% (Fig. 2).

Based on the classification results, ABC algorithm has been selected as the best candidate for classification and evaluation of textural features at class hierarchy level-I and level-II.

6.2 *Texture: Selection of Optimal Window Size and Interpixel Distance*

The aim of this work was to determine the optimal window size (W) and interpixel distance (D) for all the five textural measures. The extracted textural bands with five features (ENT, ASM, CON, MEAN and Hmg) at 9 different window sizes keeping $\theta = 0°$ were combined (stacked) separately with the three bands of the multispectral (MS) data. The supervised classification based on ABC algorithm, trained with 8578 locations, was carried out on the above data sets and one MS data set at CHL-II. Thereafter, a 3×3 majority filter was passed on all the classified images as a post-classification smoothing requirement. Finally, the accuracy assessment was carried out on the classified images using a validation data set comprising of 875 test points (details are in Table 2). The overall classification accuracy (OCA) and the kappa coefficients were computed and tabulated in Tables 5 and 6.

Table 5 OCA for five textural measures at seven different window sizes (W)

Window size (W):	5 × 5	7 × 7	11 × 11	17 × 17	25 × 25	35 × 35	51 × 51
Entropy (ENT)	64.11	65.83	67.89	75.09	78.17	77.6	76.34
Energy (ASM)	64.23	63.43	65.71	70.17	74.63	72.34	69.26
Contrast (CON)	66.29	67.43	67.66	71.43	76	76	73.71
MEAN	62.63	67.2	68.69	67.77	68.34	69.37	68.34
Homogeneity (Hmg)	64.46	65.14	67.87	70.51	72.69	71.09	72.69

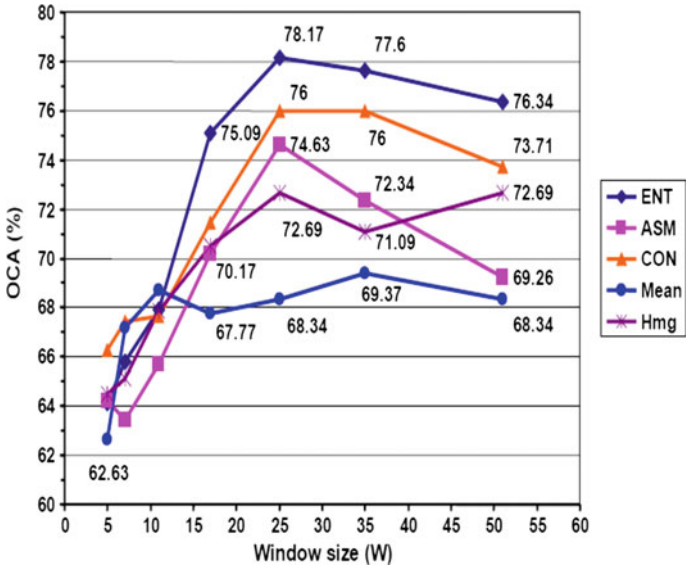

Fig. 3 Plot of OCA for five textural measures at seven different window sizes (W)

From Table 5; Fig. 3, the optimal window size for each of the five textural measures was determined from the band combination which provided the highest OCA, and the same is assumed to be the effect of the contribution of integrating the respective textural feature on MS band. Thereafter, the five textural features were extracted for D: 3 and 7 keeping W: 25 × 25 for features: ENT, ASM, CON and Hmg, and 35 × 35 for MEAN, and combined with MS bands creating an additional set of 10 data sets.

Further, Table 6; Fig. 4 show the classification accuracy obtained for the combinations of ENT (W: 25), ASM (W: 25), CON (W: 25), MEAN (W: 35) and Hmg (W: 25) at optimal window sizes (indicated within the brackets) at interpixel distance of 1, 3 and 7, keeping $\theta = 0°$. The difference in the magnitude of OCA is 1.71, 1.3, 0.46, 0.46 and 1.8% for ENT, ASM, CON, MEAN and Hmg, respectively, between D: 1 and 3. Only the texture measure Hmg shows some significant increase in OCA by an amount of 1.8% from D: 1 to 3. Also, Fig. 4 visualises that

Table 6 OCA for five textural measures at three different interpixel distances (D) at optimal window size (W)

Interpixel distance (D):	1	3	7
Entropy (ENT) (W: 25)	78.17	76.46	73.37
Energy (ASM) (W: 25)	74.63	73.26	70.97
Contrast (CON) (W: 25)	76	76.46	72.8
MEAN (W: 35)	69.37	69.83	69.03
Homogeneity (Hmg) (W: 25)	72.69	74.51	66.63

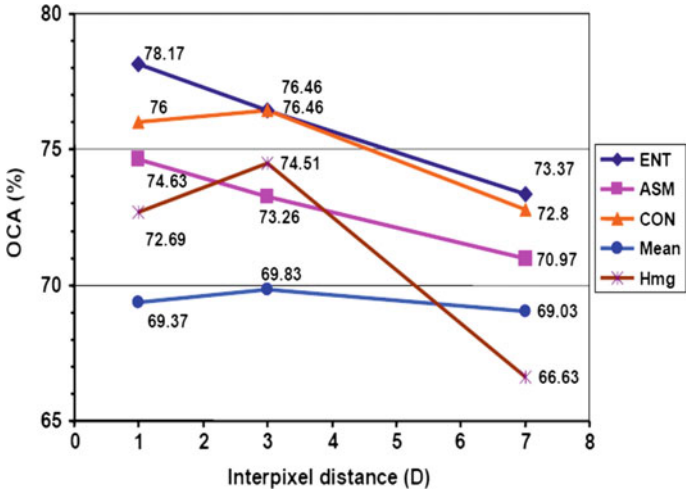

Fig. 4 Plot of OCA for five textural measures at three different interpixel distances (D) at optimal window size (W)

two of the features, viz. ENT, ASM, are moving downwards and the remaining three features, viz. CON, MEAN and Hmg, are moving upwards on OCA for change in D from 1 to 3. But among them, MEAN is not showing any significant difference in its OCA for variable D. Above D: 3, all the features exhibit a downfall in their performance. Therefore, an interpixel distance in the range of 1–3 exhibits satisfactory performance for all the texture features under investigation.

6.3 Effectiveness of Texture Feature Combinations

The objective of this experiment is to understand how a single and combinations of more than one GLCM texture measures perform with MS bands in MS data classification. The classification was carried out by employing MLC. The results of the classification using multiple texture features are provided in Table 7. A single

Table 7 Overall classification accuracy (OCA) and Kappa coefficient for data sets having combination of multispectral and texture bands at CHL-II

Sl. No.	Band ID No.	Band combinations	OCA (%)	Kappa statistics
1	1	MS	62.63	0.5721
2	2	1	78.17	0.7406
3	3	2	74.63	0.7007
7		1 + 2[a]		
8	7	1 + 3	76.57	0.7217
13		2 + 5[a]		
17		1 + 2 + 3[a]	75.09	
20	15	1 + 3 + 4	73.83	0.6867
21	16	1 + 3 + 5	76	0.7144
22	17	1 + 4 + 5	72.57	0.677
23	18	2 + 3 + 4	73.26	0.6802

MS Multispectral band, *1* Entropy, *2* ASM, *3* Contrast, *4* MEAN, *5* Homogeneity, [a]Band combinations excluded from further analysis as the combinations resulted in total misclassification of one or more classes, *ID No* Identification number provided only to the band combinations which resulted in valid classified images

texture at a time is able to provide a complete all-class classified image and hence treated as a valid classification. Still, the combinations of two textures together have shown a satisfactory performance by maintaining accuracy level approximately higher than 75% (e.g. ENT+CON, MEAN+Hmg, ASM+CON, CON+MEAN, ENT +Hmg and ENT+MEAN band combinations in Table 7). It is interesting to note that bands MEAN and Hmg alone have produced an OCA of 69.37 and 74.51%, respectively; however, when they are combined together, the OCA is increased to just 75.31%. Contrary to this observation, bands ENT and CON have produced an OCA of 78.17 and 76.46%, respectively, when used separately, but when combined together (MS+ENT+CON), the accuracy still stands at 76.57% leaving no hope of any reasonable improvement in accuracy for combination of two textures. Instead, the combination has degraded the single texture accuracy of the ENT from 78.17 to 76.57% with no appreciable increase in accuracy from 76.46 to 76.57% in respect of CON. Likewise, a simultaneous combination of three, four and five textures in MS has merged a large number of classes together, leaving as small as only two classes in the classified images as illustrated in Fig. 5. Thus, greater number of texture bands in MS has made the classification unreliable and produced invalid classified images. Therefore, it is difficult to generalise which band combination of two or more than two texture features together is most efficient in classification. It is also apparent from the above that particularly in the present classifier MLC, an increase in the number of texture features need not necessarily contribute to the improvement in accuracy; instead, such combinations degrade the classifier's performance.

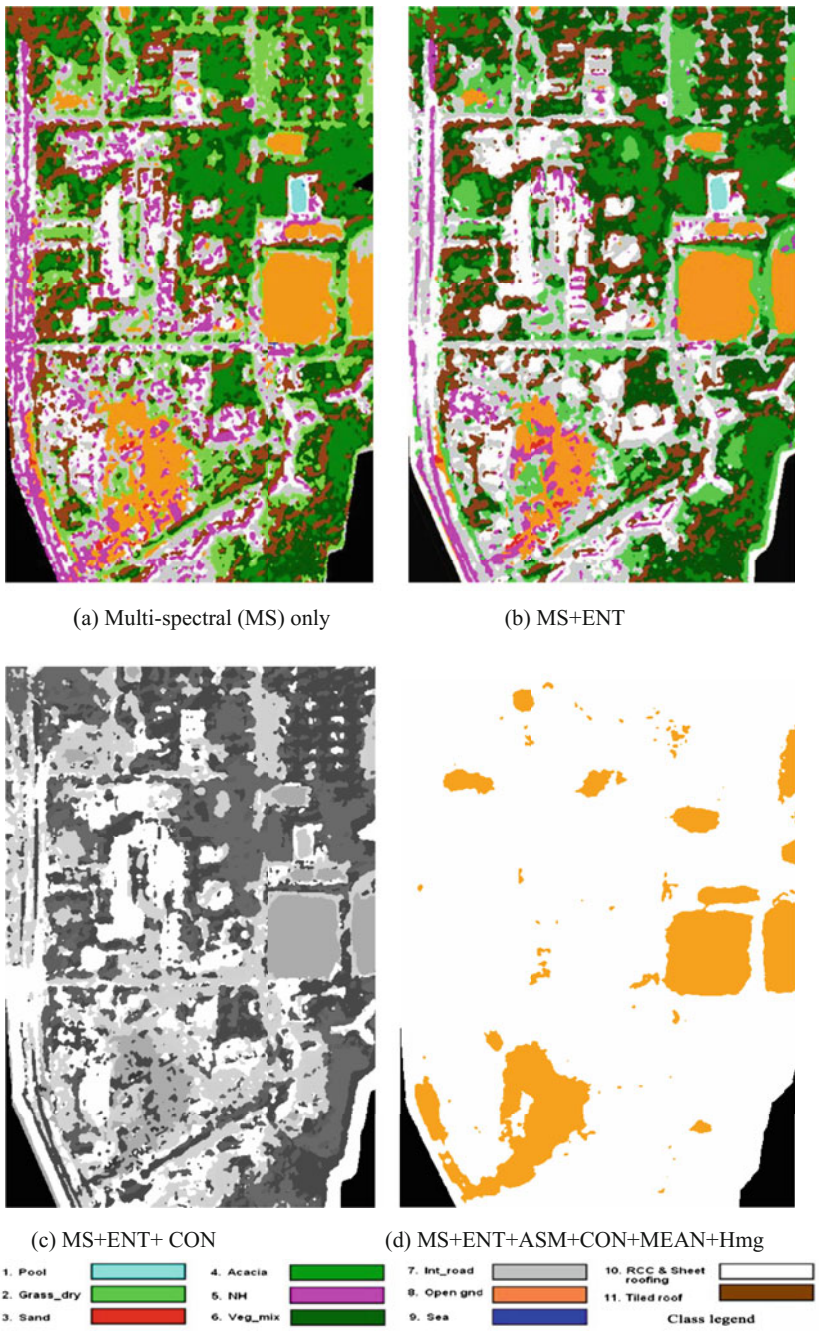

Fig. 5 Subset of classified images at CHL-II

Nevertheless, the results indicate that incorporation of texture measures into classification improves the overall classification accuracy. The discussion would be more complete if a quantitative study is made on the confusion matrices to understand the effect of the texture features on individual LU/LC classes. Hence, only two confusion matrices corresponding to MS band and MS+ENT bands are presented here in Tables 8 and 9, respectively, as the latter provides the highest accuracy in the whole study carried under CHL-II. The confusion matrices corresponding to Fig. 5c and d are not able to classify due to misclassification by the ABC classifier. So in the following discussion, the MS band and MS+ENT band combinations are referred to as case-1 and case-2, respectively.

The PA and UA for classes *pool water* and *sea water* are 100% in both the cases (Tables 8 and 9) since they are spectrally distinct and spatially homogeneous. The next two spatially homogeneous classes are *sand* and *open_gnd*. The PA for *sand* remains the same (71.43%) in both the cases, but the UA attained 100% in case-2 from the 88.24% of case-1. It is obvious because classes such as *pool water*, *sea water* and *sand* exhibit the highest TD of 2000 which indicates total separability when paired with other classes. However, even though the class *open_gnd* is spatially homogeneous, it is not spectrally distinct. The TD between the *open_gnd* and *rd_interior* is not satisfactory (1897.97). Hence, 10 validation points of class *rd_interior*, 4 each of *sand* and *RCC_sheet,* have been misclassified as *open_gnd* and reduced the UA of *open_gnd* to 64.71% in case-1, whereas these misclassifications are markedly lessened in case-2 and a UA of 88.57% is achieved for *open_gnd*. This also results in an improvement in Kappa value from 0.63 to 0.881 in comparison with case-1.

The class *roof_tiled* represents individual houses of approximately 15 m $\times$ 15 m in area located amidst of thick vegetation (*veg_mix*) along the sides of *interior road*. The major source of misclassification of *roof_tiled*, as seen in case-1, is from *veg mix* and *RCC_sheet*. In case-2, texture could make the UA and Kappa see a small improvement of 3.69% and 0.039, respectively, over MS band, but no difference in PA (92.86%). Hence, texture did not make a notable improvement on class *roof_tiled* which is distributed as relatively small individual entities over the study area. For class *RCC_sheet* (*RCC* and *asbestos sheet-roofing*), the most of the misclassification has come from the *interior road*, and a major portion of the validation sites has also been misclassified to *rd-NH, rd_interior road* and *roof_tiled* in MS data. But inclusion of feature ENT has significantly reduced the misclassification of *RCC_sheet* to *rd_NH*, *grass_dry* and *roof_tiled*, resulting in an improvement in PA from 53.14 to 76.99% and Kappa value from 0.64 to 0.724 when compared with MS band. This in turn has greatly benefited *rd_NH*, and hence, class *NH* has experienced an improvement in its UA from 48.33 to 82.35% and Kappa from 0.43 to 0.808 in MS+ENT band. The above analysis draws attention to the fact that texture has no effect on spatially homogeneous and spectrally distinct classes such as water bodies and sand. However, texture has shown an appreciable performance on spatially homogeneous but spectrally overlapping classes such as *open ground*, *RCC*, *NH* and *interior roads*.

Table 8 Confusion matrix and conditional kappa values of the classification results obtained for MS band alone for 11 classes

classes	1	2	3	4	5	6	7	8	9	10	11	Row total	UA%
1	**18**											18	100
2		**55**					12			7		74	74.32
3			**15**							2		17	88.24
4				**38**		101						139	27.34
5					**58**		9			53		120	48.33
6		5		4		**62**				1	2	74	83.78
7		4			3		**84**	1		32	1	125	67.20
8			4				10	**33**		4		51	64.71
9									**19**			19	100
10		2	2		7	1	32			**127**		171	74.27
11				1		14				13	**39**	67	58.21
Column total	18	66	21	43	68	178	147	34	19	239	42	**875**	
PA%	100	83.33	71.43	88.37	85.29	34.83	57.14	97.06	100	53.14	92.86		**OCA**
													62.63
kappa	1.0	0.72	0.87	0.23	0.43	0.79	0.60	0.63	1.0	0.64	0.56		

Class legend: *1* pool; *2* grass_dry; *3* sand; *4* acacia; *5* NH; *6* veg_mix; *7* inl_rd; *8* open_gnd; *9* sea; *10* rcc_sheet; *11* tiled_roof

Table 9 Confusion matrix and conditional kappa values of the classification results obtained for MS+ENT (W: 25 × 25, D = 1) bands for 11 classes

classes	1	2	3	4	5	6	7	8	9	10	11	Row total	UA%
1	**18**											18	100
2		**56**				2	3	1		2		64	87.50
3			**15**									15	100
4				**26**		28						54	48.15
5		1			**56**		1			10		68	82.35
6		3		17		**132**				2	2	156	84.62
7		5			6		**108**	1		32	1	153	70.59
8			4				10	**31**				35	88.57
9									**19**			19	100
10		1	2		6	1	35	1		**184**		230	80.00
11				1		15				9	**39**	63	61.90
Column total	18	66	21	43	68	178	147	34	19	239	42	**875**	
PA%	100	84.85	71.43	60.47	82.35	74.16	73.47	91.18	100	76.99	92.86		**OCA**
													78.17
kappa	1.0	0.864	1.0	0.454	0.808	0.806	0.646	0.881	1.0	0.724	0.599		

Class legend: *1* pool; *2* grass_dry; *3* sand; *4* acacia; *5* NH; *6* veg_mix; *7* int_rd; *8* open_gnd; *9* sea; *10* rcc_sheet; *11* tiled_roof

Another major spectrally overlapping classes are the mixed vegetation (*veg_mix*) and acacia forest (*veg_acacia*), which exhibit the lowest TD measure of 434.44. The confusion matrix corresponding to MS shows that out of the given 178 validation sites to *veg mix*, 101 have been misclassified as *veg_acacia*, 14 as *roof_tiled* and 1 site to *RCC_sheet*. However, there is no serious misclassification from other classes. Hence, even though it has produced the lowest PA of 34.83%, it could maintain a reasonably high UA of 83.78%. On the contrary, the same MS band has produced a relatively higher PA of 88.37% for *acacia*, but has made a fall in UA to a very low value of 27.34%, since 101 sites of *veg_mix* are misclassified as *veg_acacia*. A trade-off is seen between the PA and UA of these two classes. But the addition of texture band on MS could raise the PA of *veg_mix* from 34.83% to 74.16% with no appreciable improvement in the UA and Kappa. Contrary to this, texture has made a greater contribution to the UA (27.34–48.15%) of class *acacia* at the cost of a fall in PA from 88.37 to 60.47%, but without losing in Kappa. The Kappa, in turn, is improved from 0.23 to 0.454, approximately double that of the MS band. Finally, the class *grass_dry* is also benefited by texture as it could record an increase in Kappa from 0.72 to 0.864 in comparison with case-1. Hence, it is evident from above that addition of texture in MS band has the potential to improve the accuracy even for those classes which are spectrally overlapping and spatially non-homogeneous. For all the 11 classes, texture has maintained higher Kappa value over the MS band.

Finally, it is quite evident from the two classified images shown in Fig. 5b, c that the data set having texture feature entropy (ENT) integrated with MS is fairly free from the *salt-and-pepper* noise, which is very common and annoying in classified images (Fig. 5a). This improvement is also seen in other classified images too which employ the texture and MS bands together. Besides, since linear structures such as roads and buildings are seen more distinct, the shape information is found to be greatly pronounced in a texture-embedded classified image.

6.4 Investigation of Texture at Class Hierarchy Level-I and Level-II

The purpose of this phase of investigation is to examine the effect of the texture features at classification hierarchy level-I, in comparison with level-II. Results of the work have been tabulated in Table 10. The corresponding plot is also available in Fig. 6. Here, 6 data sets having texture measures with optimal W and D in combination with MS bands, but one feature at a time, were investigated. At class hierarchy level-I, the OCA obtained for all data sets is remarkably higher than level-II. It is also obvious from Fig. 6 that the MS band has gained an improvement of 18.17% in OCA (from 62.63 to 80.80%) when CHL is decreased from level-II to level-I without any texture integrated in it. The incorporation of texture at CHL-I

Table 10 OCA (%) and Kappa for various combinations of spectral and texture measures one at a time at CHL-I and CHL-II (6 and 11 classes, respectively)

Band combinations:		MS	ENT	ASM	CON	MEAN	Hmg
Class hierarchy level-I	OCA %	80.80	8.23	81.37	87.54	84.69	85.49
	Kappa	0.749	0.791	0.753	0.835	0.798	0.808
Class hierarchy level-II	OCA %	62.63	78.17	74.63	76.46	69.83	74.51
	Kappa	0.572	0.740	0.700	0.720	0.650	0.699

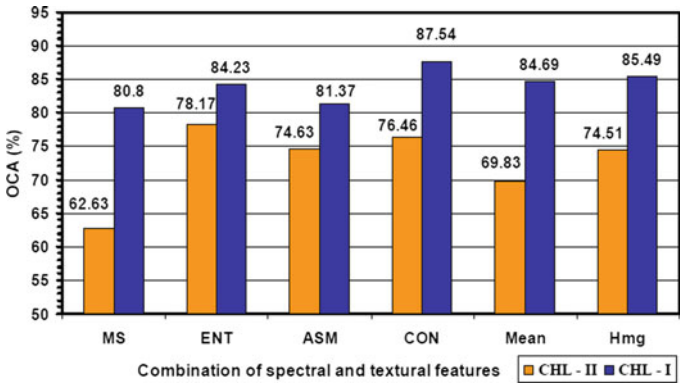

Fig. 6 Plot of OCA versus various combinations of spectral and texture measures one at a time at class hierarchy level-I and level-II (6 and 11 classes, respectively)

has improved the OCA in the range of 0.57% (with ASM: 81.37%) to 6.74% (with CON: 87.54%) in comparison with MS band of CHL-I (80.80%). Whereas, even though, the OCA (62.63%) is small for MS band at CHL-II, the difference in improvement seen upon adding texture ranges from 6.74% (with MEAN: 69.83%) to as high as 15.54% (with ENT: 78.17%). Hence, texture is found to be very effective in improving the classification accuracy at classification hierarchy level-II than level-I.

In summary, it is seen that texture feature ENT (W: 25 × 25, D: 1) not only produced the highest accuracy of 78.17% in class hierarchy level-II, but also maintained higher accuracy over its counterparts at all window sizes greater than W: 11 × 11 and at all interpixel distances. In contrast to the above, it is apparent from the plot in Fig. 6 that in CHL-I, the highest accuracy of 87.54% is obtained for the texture feature CON at W: 25 × 25, D: 1. The ENT and CON have also occupied the first and the second positions, respectively, in performance at CHL-II when investigated one texture at a time with MS. But the combination MS+ENT +CON, which is expected to perform better than ENT, stood second in accuracy (OCA: 76.57%) among all the band combinations with MS. Meanwhile, in CHL-II, the lowest OCA (69.37%) is recorded for texture band MEAN at W: 35 × 35, D: 1, but shows a steady performance for a wide range of W and D. Hence, it is difficult to ascertain which combination of the texture features of more than one band is

important in improving the overall classification accuracy. In the course of the analysis, one of the principal findings is that the urban or semi-urban LU/LC classes such as *RCC* and *sheet-roofing*, *roads* and *open grounds* have been substantially benefited by integrating texture features in the classification of multispectral bands.

7 Conclusion

The study demonstrated the effect of integration of spectral and textural features in classification of semi-urban Lu/LC features on LISS-IV data of 2.5 m spatial resolution at CHL-I and CHL-II. Textural measures, viz. entropy, energy, contrast, MEAN and homogeneity, were investigated with multispectral bands and various texture feature combinations. The following conclusions were drawn from the present study:

1. This work has presented a new method for classifying RS data using Artificial Bee Colony algorithm. As a result, ABC algorithm has provided high-quality result on the study area. OCA for ABC method shows an improvement of 3% in comparison with MLC.
2. Combining textural and spectral features in high-resolution satellite data has proven to be effective in improving the classification accuracy of spectrally overlapping but spatially homogenous urban land use/land cover features such as RCC roofing, asbestos sheet-roofing, interior road, NH and open grounds (playground). Texture combinations have not made marked difference over spectrally distinct and spatially homogeneous classes such as sand and water bodies in comparison with multispectral bands. Integration of texture greatly reduced the 'salt-and-pepper' noise in the classified images.
3. Literature survey indicates that there is no unique textural window size describing a land cover class as it depends on the spatial resolution of the data, the land cover spread which varies from one geographical area to another, and human cultural habitation; hence, it is solely problem dependent [16–18].

Acknowledgements The authors graciously thank Dr. Dwarakish, Associate Professor, NITK, Surathkal; DR. K.S. Shreedhara, UBDTCE, Davangere; and the staff at KSRSAC, Mysore, for their assistance and help rendered during field work.

References

1. Hala E, Mohamed NH (2012) Mapping potential landfill sites for North Sinai cities using spatial multicriteria evaluation. Egypt J Remote Sens Space Sci 15(2):143–150
2. Du P, Tan K, Xing X (2012) A novel binary tree support vector machine for hyperspectral remote sensing image classification. Optics Commun 285:3054–3060. doi:10.1016/j.optcom.2012.02.092

3. Zheng J, Cui Z, Liu A, Jia Y (2008) A K-means remote sensing image classification method based on AdaBoost. In: fourth international conference on natural computation ICNC '08, vol 4, pp 27–32. doi:10.1109/ICNC.2008.903
4. Yang H, Du Q, Chen G (2012) Particle swarm optimization-based hyperspectral dimensionality reduction for urban land cover classification. IEEE J Sel Top Appl Earth Observations Remote Sen 5(2):544–554. doi:10.1109/JSTARS.2012.2185822
5. Liu X, Li X, Liu L, He J, Ai B (2008) An innovative method to classify remote-sensing images using ant colony optimization. In: IEEE transactions on geoscience and remote Sensing 46(12). doi:10.1109/TGRS.2008.2001754
6. Jayanth J, Koliwad S, Ashok Kumar T (2015) Classification of remote sensed data using Artificial Bee Colony algorithm. Egypt J Remote Sens Space Sci 18(1):119–126. doi:10.1016/j.ejrs.2015.03.001
7. Jayanth J, Ashok Kumar T, Koliwad S, Krishnashastry S (2015) Artificial bee colony algorithm for classification of remote sensed data. In: international conference on industrial instrumentation and control (ICIC)
8. Jayanth J, Koliwad S, Ashok Kumar T (2015) Identification of land cover changes in the coastal area of Dakshina Kannada district, south India, during the year 2004–2008. Egypt J Remote Sens Space Sci 19(1):119–126. doi:10.1016/j.ejrs.2015.03.001
9. Jayanth J, Ashok Kumar T, Koliwad S (2012) Comparative analysis of Image fusion techniques for remote sensing. In: international conference on advanced machine learning technologies and applications (AMLTA 2012) Cairo, Egypt, December 8–10. Proceedings of communications in computer and information Science. In: Hassanien AE, Salem AM, Ramadan R, Kim T (eds) (2012) vol 322 Springer, Berlin, pp 111–117
10. Kumar U, Dasgupta A, Mukhopadhyay C (2012) Sequential maximum a posterior (SMAP) algorithm for classification of urban area using multi-resolution spatial data with derived geographical layers. In: Proceedings of the India conference on geo-spatial technologies and applications, department of computer science and engineering, Indian Institute of Technology Bombay (IITB), 12–13 April 2012
11. Chang Che-Wei, Ho Chien-Chang, Chen Jyh-Horng (2012) ADHD classification by a texture analysis of anatomical brain MRI data. Frontier Sys Neuro Sci 12:123–129
12. Choodarathnakara AL, Ashok Kumar T, Koliwad S, Patil CG (2012) Soft classification techniques for RS data. IJCSET 2(11):1468–1471
13. Srikrishna Shastri C, Ashok Kumar T, Koliwad S (2016) Advances in classification techniques for semi urban land features using high resolution satellite data. Int J Adv Remote Sen GIS 5(3):1639–1648
14. Ashok Kumar T, Koliwad S, Dwarakish GS (2008) Effectiveness of decision tree algorithm in land use/land cover classification of NITK, Surathkal campus using high resolution satellite data. Int J Earth Sci Engg 18(1):56–65
15. Ashok Kumar T, Koliwad S, Dwarakish GS (2008) Classification of high-resolution satellite imagery: a non-parametric approach. National Institute of Technology Karnataka Research Bulletin, vol 17 no 1, pp 21–33
16. Ashok Kumar T, Koliwad S, Dwarakish GS, Patil CG (2008) Non-parametric classifier: an integrated approach with texture for land-cover classification of pan-sharpened LISS-IV imagery. In: Souvenir cum abstracts, Indian national cartographic association (INCA)-XXVIII international congress on collaborative mapping & space technologies Ahmedabad, Gujarath, India, November 4–6
17. Tuceryan M, Anil Jain K (1998) Texture Analysis. The handbook of pattern recognition and computer vision. World Scientific Publishing Co, pp 207–248
18. Ashok kumar T (2013) Advanced Image processing Techniques for Land feature Classification. Lambert Academic Publishing, pp 25–86

Saliency-Based Image Compression Using Walsh–Hadamard Transform (WHT)

A. Diana Andrushia and R. Thangarjan

Abstract Owing to the development of multimedia technology, it is mandatory to perform image compression, while transferring an image from one end to another. The proposed method directly highlights the salient region in WHT domain, which results in the saliency map with lesser computation. The WHT-based saliency map is directly used to guide the image compression. Initially, the important and less important regions are identified using WHT-based visual saliency model. It significantly reduces the entropy and also reserves perceptual fidelity. The main aim of the proposed method is to produce the high-quality compressed images with lesser computational effort and thereby achieving high compression ratio. Due to the simplicity and high speed of WHT, the proposed visual saliency-based image compression method is producing reliable results, in terms of peak signal-to-noise ratio (PSNR), compression ratio, and structural similarity (SSIM), compared to the state-of-the-art methods.

Keywords Saliency detection · Image compression · Walsh–Hadamard transform · PSNR

1 Introduction

Salient region detection drastically attracts the attention toward many of the computer vision and pattern recognition tasks such as image compression, object recognition, content-based image retrieval, image collection browsing, image editing, visual tracking, and human–robot interaction. It aims to detect the salient

A. Diana Andrushia (✉)
Department of ECE, Karunya Universtiy, Coimbatore, Tamil Nadu, India
e-mail: andrushia@gmail.com

R. Thangarjan
Department of CSE, Kongu Engineering College, Erode, Tamil Nadu, India
e-mail: thangs_68@yahoo.com

J. Hemanth and V.E. Balas (eds.), *Biologically Rationalized Computing Techniques For Image Processing Applications*, Lecture Notes in Computational Vision and Biomechanics 25, DOI 10.1007/978-3-319-61316-1_2

region of an image under biological plausibility. In this paper, Walsh–Hadamard transform (WHT)-based visual saliency detection is used for the image compression application.

Demand on data compression is increasing rapidly as the modern technologies are growing high. High storage capacity is required for uncompressed images. Many images are having the common characteristics as their neighboring pixels are highly correlated with redundant information [1]. Image compression-based techniques are aiming to reduce the redundant information by eliminating the spectral and spatial redundancies. It results in the reduction of consumption of expensive resources in the form of transmission bandwidth and hard disk space. Image compression techniques are generally classified into two types. One is spatial coding and another one is transform coding. In the transform coding type, discrete cosine transform (DCT), discrete Fourier transform (DFT), Walsh–Hadamard transform (WHT), etc., are used to perform natural image compression. Each transform is having its own advantages and disadvantages in the compression domain. Transform type coding is used in the proposed method.

WHT-based visual saliency detection and WHT-based image compression based on saliency map are the two main phases of this proposed method. The proposed WHT-based visual saliency detection is transform domain approach. So the frequency domain approaches are only considered for the performance comparison. As the WHT-based compression also transform domain approach so frequency domain methods only considered for fair performance comparisons in the compression phase.

The proposed work is experimented through MIT dataset. It is one of the benchmark datasets for visual saliency detection which consists of the indoor, outdoor images. The performance metric of receiver operating characteristics (ROC), area under the curve (AUC), precision, recall, and F-measure is obtained to analyze the proposed visual saliency detection. Peak signal-to-noise ratio (PSNR), structural similarity (SSIM), and compression ratio are obtained for saliency-based image compression. The performance metrics are yielding significant results while comparing with state-of-the-art methods.

WHT is used in the saliency detection as well as image compression. WHT is chosen because the number of computations in this transform is significantly less compared to the other transforms, and WHT is the key transform to provide energy compactness. Finding the salient information in image/audio/video will reduce the number of computations and lesser hardware in the compression techniques. The proposed visual saliency-based compression method achieves 90% compression ratio. It is due to the lesser computations of WHT transform. This is the most added advantage of the proposed method compared to the state-of-the-art methods.

The reminder part of this chapter constructed as follows: Sect. 2 clearly explains about the backgrounds of visual saliency detection, saliency-based image compression. Section 3 elaborates the proposed methodology. Section 4 explains the experimental results with the performance metrics. Finally, the conclusion of the chapter is given in the last section.

2 Backgrounds

2.1 Saliency Detection

The world is full of visual data. Humans selectively perceive the visual information that is getting in through their eyes. Visual attention is the process of selecting particular information from the plenty of raw data perceived. For example, while sincerely watching a cricket match in the play ground, a sudden change in the action of the umpire with red shirt picks the attention of the spectators in the gallery despite the colossal load of visual inputs such as actions of the batsmen, bowlers, and fielders is there. The eyes of the spectators gaze the umpire momentarily before shifting to other events in the visual scene. The phenomenon of drawing the focus of attention to certain regions in a given scene or image is called visual attention [2]. In the jargon of computer vision, these regions are known as *salient* regions.

Detection of salient regions finds application in a wide spectrum of processes such as automatic object detection, image retrieval, remote sensing, automatic target detection, image and video segmentation, robotics, scene understanding, computer–human interaction, driver assistance, action recognition, background subtraction, image and video compression, video summarization, medical imaging, and automatic cropping. The cognitive process that directs human to select highly relevant data from a scene/image is named as visual attention.

Recently, a number of computational models have been developed to highlight salient regions. As far as computational models are concerned, there are two types of models in the literature, namely bottom-up and top-down approaches. Bottom-up approach works from the low-level visual features and moves up to construct a saliency map. The top-down approach is goal-driven, and it uses prior knowledge to achieve the goal such as target detection, scene classification, and object recognition [3]. The top-down approach starts from a particular visual processing task.

The computational models of visual attention either bottom–up or top–down can be constructed in spatial domain or in frequency domain, in order to highlight the salient regions as saliency map. Spatial domain methods require more computation time to obtain the features compared to the frequency domain methods [4]. The computational complexities of these models are very high, and these models are not performed with multi-scale spatial and frequency analysis. Many models which come under frequency domain approaches have used only local features to identify the salient regions.

The very first method of saliency detection is developed by Itti et al. [5]. Local contrast information is used to develop the method. The local features are only used by Ma et al. [6], Harel et al. [7], and Goferman et al. [8] to obtain the visual saliency detection. The global features are considered in [9–12] to construct the visual saliency model. However, the accurate identification of salient regions should also involve the global features. In recent years, many researchers have shown more interest to build computational visual models in the transform domain. Mainly

Fourier transform (FT) and wavelet transform (WT) have also been extensively used to highlight the salient regions in the transform domain. Every approach has its own pros and cons.

FT gives promising results for applications involving only stationary signals. The amplitude and phase spectrum of Fourier transform is used in [13]. Guo et al. [12] used the phase spectrum of quaternion Fourier transform (PQFT) to highlight the saliency, and it is also applied for efficient video compression. Hou et al. [14] used Fourier transform and log spectrum to construct the spectral residual approach for the saliency detection.

WT has the capacity to provide multi-scale spatial and frequency analysis because it codes the signal at different bands and bandwidths. WT can represent singularities in a much better manner than FT can. And moreover, WT can be applied for nonstationary signals also [15]. WT is used in [16] to find the salient object. WT-based orientation feature maps are obtained in different scale. The order map is also found by using Fourier analysis. The local, global information are used in the WT-based salient point detection [17]. WT-based salient detection is used in multi-scale image retrieval problem.

Wavelets are very good to represent point singularities, but when it comes to directional features they fall short [18]. The main reason is that wavelets are generated by isotropic elements. The quantity of decomposition level should be very large when approximating a curve using WT. The disadvantages of WT have been overcome by using multi-directional and multi-scale transforms.

The higher directional wavelet transforms of ridgelet, curvelet, shearlet are also used for the visual saliency detection. The directional features are captured effectively, and the potential salient regions are identified. Bao et al. [19] proposed visual saliency detection based on shearlet transform, in which the local and global contrasts are used to obtain local and global saliency map. Initially, the potential salient regions are identified in order to update the feature maps in shearlet domain.

Even though the transform based saliency detection methods are producing reliable results, suffered highly from computation complexity. Larger computations are required for these methods.

In general, the transform domain visual saliency detection methods are using the following steps.

- Transform the input image into transform domain,
- Obtain the feature maps for various features,
- Combine the various feature maps,
- Use the top-down features if required for particular application,
- Apply inverse transforms to get the saliency map.

The saliency map is the topographical map which shows the visual saliency in the visual scene. So the visual saliency detection methods are showing the outputs as visual saliency map. Saliency-based image segmentation, image compression, and image retrieval are the popular areas of research.

The Walsh–Hadamard transform (WHT) has lesser computations and extremely fast transform. It is computed only by addition and subtraction. Lesser hardware is required for the practical implementation [20]. The highly correlated pixels are captured by the WHT in the visual space. Hence, in this chapter WHT is used to detect the salient regions in transform domain.

2.2 Visual Saliency-Based Image Compression

Usually, the important regions of an image may be small and highly degraded at low bitrates. The standards of compression such as JPEG/JPEG-2000, MPEG 4 are not handling the salient regions well. Guo et al. [21] and Hadi et al. [22] investigated saliency-based compression techniques. To adhere the saliency values, the transmitted coefficients are modified. These methods cannot handle salient regions well and also suffer from complicated computations. Barua et al. [23] developed wavelet-based image compression technique for images and videos. The algorithm is designed to obtain the saliency values in wavelet domain and then corresponding image/video coefficients are transmitted. It preserves the important region of an image/video.

Nabil Ouerhani et al. [24] proposed adaptive color image compression based on biologically inspired visual attention. The initial stage perceptual salient regions of interest are identified automatically. The adaptive coding scheme allocates higher number of bits for the salient regions. The results are compatible with the JPEG standards.

Li et al. [25] performed video compression based on computational models of visual attention. The salient regions are encoded with higher quality compared with non-salient regions. The salient regions are awarded with higher priority rather than others. But it may generate visible artifacts in the non-salient part where the quality of image is poorer. The artifacts also sometimes draw the end-user's attention. In several cases, the high level of artifacts becomes salient and captures the viewer's attention. But the notable artifacts are not to be salient.

Hadizadeh [26] dissertation reveals the visual saliency methods for video compression and transmission. Saliency-based video coding is investigated. The main concept is that high salient regions are having higher ability to percept than lesser salient regions. The quality of image/video is handled toward the user most attended regions. This method effectively performs video coding expect in two major cases. If any region is richly salient, then its saliency will be increased after the compression, provided the quality of the image/video remains high. The reason is that the users are noticing the high salient regions in the scene. If the region is lesser salient, then its saliency will be decreasing after the compression task, because the lower saliency regions are ended with lesser quality.

Ho-Phuoc et al. [27] proposed the visual saliency-based data compression for image sensors. The adaptive image compression is presented in each block. First the saliency value is obtained, and then the Haar wavelet transform is applied for the compression. This framework gives lesser memory and compact operators. The data stored in the image sensors are very much reduced, and image quality is not altered.

Zundy et al. [28] proposed the content-based image compression using visual saliency methods. Initially, the saliency map is obtained from the video, automatically or by user input. The salient regions are performed with nonlinear image scaling. Salient image regions are given higher pixel count, and non-salient regions are given lesser pixel count. Existing compression techniques are utilized to compress the nonlinearly down-scaled images, and in the receiver end it is up-scaled. This method supports for anti-aliasing effect which reduces the aliasing in highly scaled regions.

In order to reduce redundant information in the dynamic scenes, the visual saliency in videos is proposed by Tu et al. [29]. Based on the video visual saliency map, the redundant information is removed. In this paper, video visual saliency is the catalyst of video compression technique. DCT is utilized to perform the video compression, and this technique is adopted for MPEG-1, MPEG-4, and H.265/HEVC standards.

Yu et al. [30] used visual saliency to guide image compression. At individual scales, the saliency is measured through Laplacian pyramid. The proposed compression algorithm decreases the entropy of the image with respect to the saliency map in each scale. Dhavale et al. [31] proposed visual computational model-based image compression. It successfully locates the regions of interest of the human and thereby applies for image compression.

Duan et al. [32] proposed the image compression technique based on saliency detection and independent component analysis. The input image is transformed first using ICA. The transformed coefficients are numbered with set zero coefficient percentage. The sparse nature of independent component analysis is used in this method. It is compared with DCT-based compression method.

Many of the state-of-the-art methods are failed to show the energy compactness. In the field of image compression, energy compactness is the key point and these methods also suffer from higher computational complexity. In order to revoke the key points, the WHT is used in the proposed image compression method. The redundant pixels are captured by the WHT with lesser computations.

2.3 Walsh–Hadamard Transform (WHT)

Discrete Fourier transform (DFT), discrete cosine transform (DCT), and Walsh–Hadamard transform (WHT) are widely used in the image processing applications. These linear image transforms are chosen in the image processing application because of their flexibility, energy compaction, and robustness. These transforms effectively extract the edges and also provide energy compaction in the

state-of-the-art methods. Among all these transforms, WHT is very gorgeous one because of its simplicity and its computational efficiency. The major properties of WHT are same as that of other image transforms. The basis vector components of WHT are orthogonal, and it is having binary values (±1) only.

WHT is orthogonal, non-sinusoidal transform which is used in image filtering, speech processing, and medical signal analysis. To be more specific the lunar images/signals are well processed, coded, and filtered by WHT. It is known well because of its simplicity and fast computation. WHT is the substitute of Fourier transform. It is computationally simpler because it requires no multiplication or division operations. Every computation is performed by simple addition and subtraction operation. WHT is one of the very fast transforms which can be implemented in $O(N\log_2 N)$ additions and subtractions. So the hardware implementation of WHT-based applications is also so simpler [20].

So it is beneficial in terms of energy consumption and lesser computation. WHT is real, orthogonal, and symmetric $H = H^* = H^{\mathrm{T}} = H^{-1}$. Walsh–Hadamard transform is represented in terms of Walsh–Hadamard transform matrix (WHTM).

It consists of set of N rows denoted by H_{j} for $j = 0, 1, 2 \ldots N - 1$. The properties of WHT matrix are:

i. H_j takes values as +1 and −1,
ii. The size of WHT matrix is usually the power of 2,
iii. H_j [0] = 1 for all j.

The size of the WHTM is generally the power of two. The second-order Hadamard matrix is given by, $H = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$.

The WHTM with the order of 4 is $\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \end{pmatrix}$.

Each row in the matrix is called as basis vector of WHTM [33]. Generally, the basis vectors are orthonormal and orthogonal. Orthonormal means dot product of each basis vector themself is one. Orthogonal means dot product of any two basis vectors is zero.

The computation of WHT involves very simple step only, when the image is projected into basis images, each pixel is multiplied by ±1, whereas the FFT needed complex multiplication. So WHT is more efficient than FFT in terms of computation complexity.

Consider the image C of size $N \times N$ with the pixels of $c(x, y)$, the 2D WHT is defined as

$$H(u, v) = \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} c(x, y) g(x, y, u, v) \tag{1}$$