

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/311911775>

Recent development of feature extraction and classification multispectral/hyperspectral images: a systematic lite....

Article in *Journal of Physics Conference Series* · December 2016

DOI: 10.1088/1742-6596/801/1/012045

CITATIONS

0

READS

64

3 authors:



Andie Setiyoko

Indonesian National Institute of Aeronautics ...

7 PUBLICATIONS 3 CITATIONS

SEE PROFILE



I Gede Wahyu Surya Dharma

University of Indonesia

2 PUBLICATIONS 0 CITATIONS

SEE PROFILE



Toto Haryanto

University of Indonesia

17 PUBLICATIONS 4 CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:



Protein Secondary Structure Prediction Using Support Vector Machine [View project](#)



Research Methodology [View project](#)

Recent development of feature extraction and classification multispectral/hyperspectral images: a systematic literature review

This content has been downloaded from IOPscience. Please scroll down to see the full text.

2017 J. Phys.: Conf. Ser. 801 012045

(<http://iopscience.iop.org/1742-6596/801/1/012045>)

View [the table of contents for this issue](#), or go to the [journal homepage](#) for more

Download details:

IP Address: 120.188.65.44

This content was downloaded on 26/04/2017 at 03:35

Please note that [terms and conditions apply](#).

You may also be interested in:

[A multiple classifier system based on Ant-Colony Optimization for Hyperspectral image classification](#)

Ke Tang, Li Xie and Guangyao Li

[One method for HJ-1-A HSI and CCD data fusion](#)

Wencheng Xiong, Yun Shao, Wenming Shen et al.

[Maximum relevance, minimum redundancy band selection based on neighborhood rough set for hyperspectral data classification](#)

Yao Liu, Yuehua Chen, Kezhu Tan et al.

[Parallel computation for blood cell classification in medical hyperspectral imagery](#)

Wei Li, Lucheng Wu, Xianbo Qiu et al.

[Multispectral digital lensless holographic microscopy: from femtosecond laser to white light LED](#)

J Garcia-Sucerquia

[Multispectral Electrical Impedance Tomography using Optimization over Manifolds](#)

A Fouchard, S. Bonnet and O. David

[The Effectiveness of Hydrothermal Alteration Mapping based on Hyperspectral Data in Tropical Region](#)

R R D Muhammad and A Saepuloh

[Comparison of two Classification methods \(MLC and SVM\) to extract land use and land cover in Johor Malaysia](#)

B Rokni Deilmai, B Bin Ahmad and H Zabihi

Recent development of feature extraction and classification multispectral/hyperspectral images: a systematic literature review

A Setiyoko¹, I G W S Dharma², and T Haryanto³

^{1,2,3}Faculty of Computer Science, University of Indonesia Depok, Indonesia, andie.setiyoko@ui.ac.id, i.gede614@ui.ac.id, toto.haryanto@ui.ac.id

¹Remote Sensing Technology and Data Center-LAPAN, Jl Lapan 70, Pekayon, Pasar Rebo, Jakarta Timur, Indonesia 13710, andie.setiyoko@lapan.go.id

Abstract. Multispectral data and hyperspectral data acquired from satellite sensor have the ability in detecting various objects on the earth ranging from low scale to high scale modeling. These data are increasingly being used to produce geospatial information for rapid analysis by running feature extraction or classification process. Applying the most suited model for this data mining is still challenging because there are issues regarding accuracy and computational cost. This research aim is to develop a better understanding regarding object feature extraction and classification applied for satellite image by systematically reviewing related recent research projects. A method used in this research is based on PRISMA statement. After deriving important points from trusted sources, pixel based and texture-based feature extraction techniques are promising technique to be analyzed more in recent development of feature extraction and classification.

1. Introduction

A multispectral band is a representation of the spectral resolution and a panchromatic band is representativeness of spatial resolution or pixel size of a particular image. Spectral resolution with regard to how detailed the division of the wavelength on the satellite sensor. For example Landsat 8 with OLI (Operational Land Imager) multispectral bands that divided into 8 bands (1-7,9) or WorldView 3 that contained 8 bands. Development of spectral resolution leads to hyperspectral data, which divide wavelength in very narrow range. Hyperspectral data contained a large number of bands, AVIRIS with 224 bands and HYDICE with 210 bands [1]. Multispectral and hyperspectral data have the ability to do classification and feature extraction process because each band has a character to display certain objects. For example, Landsat 8 band 1 representing interactions at wavelengths can detect water quality [2]. Multispectral and hyperspectral data has been used for various applications such as land cover classification for agriculture, for example, there was a research project conducted crop type mapping in a highly fragmented and heterogeneous agricultural landscape by using multitemporal Landsat 8 [3]. There was also project research to do an analysis of land cover change of greenness trends using 3 generations of Landsat images [4]. Some examples of cases also used on specific interests such as identifying burned area by using Landsat TM/ETM+ [5], drought monitoring using Landsat 8 OLI imagery [6].

Multispectral images with various band included in it, have information that can be captured. Feature extraction has an important role to capture meaningful information from the data so can give more precision on supervised or unsupervised technique.

Principal Component Analysis (PCA) besides as a feature extraction, it is a technique for reducing high dimensional of data. PCA has been applied successfully on Landsat-7 ETM imaginary for a high dimensionality reduction method and also image enhancement[7]. Combination PCA with classification such Support Vector Machine (SVM) has been succeeded applied for handling image fusion with filtering



called image fusion and recursive filtering (IFRF)[8]. The salt field from Fused Landsat TM data have been mapped based on reflectance-spectra. The inverse PCA in this research applied to transform back to the seven band [9].

Since texture analysis proposed by Haralick as feature extraction on image processing [10], this technique has been applied on Remote Sensing research field. One of the texture analysis using Grey Level Co-occurrence Matrix (GLCM). GLCM has been conducted for detection ice detection on Bohai sea in Tiongkok[11]. The result has been validated using co-temporal HJ1B-CCD on 30 m imaginary and confusion matrix for accuracy. GLCM also has been implemented to analyze texture from Synthetic Aperture Radar (SAR) data for mapping global urban area at 30 m resolution. Another implementation of GLCM is for analyzing multiband texture on multispectral images [12].

A hyperspectral image is containing a lot of bands that have a specific purpose to be classified. Those regions pixel can be assigned as a unique label that able to specify belonging classes [13]. However a hyperspectral image relatively containing low spatial resolution and unavoidable noise which obtained from transmission and reception limit [14]. To classify the region from hyperspectral image, spectral-spatial classification network proposed by Pan, with input from PCANet that label region pixels already determined [15], while Li proposed Framework of Spectral-Mixture-Analysis-Basis HIS SR as a framework to reconstructed low resolution hyperspectral. Before coming into a framework, PCA is applied to a hyperspectral image to de-correlate the atom in the spectral dictionary [14]. In contrast, Lu proposed two main steps in set-to-set distance-based spectral-spatial classification. The first step, they generate both of training and testing sets and the second step is set-to-set distance-based classification. In step two, set-to-set distance-based classification is divided into two subtasks, representation of sets and label assignment. The Affine Hull (AH) model is used to adaptively represent the characteristic of each set [13].

The objective of this systematic review is to develop a better understanding regarding object feature extraction and classification based on an image, particularly satellite image. The research will compare a number of techniques to extract feature information and to classify the specific object from satellite image in the recent time.

2. Method

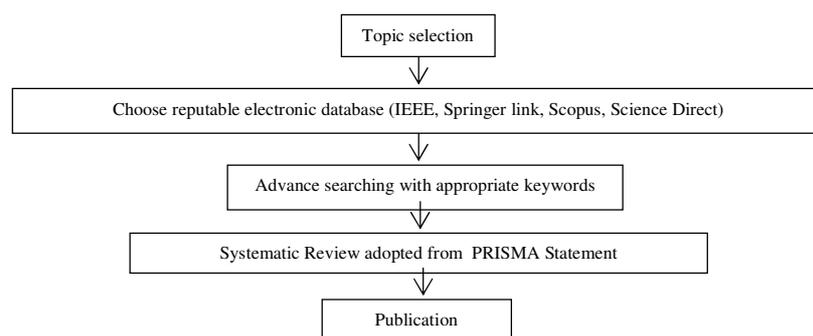


Figure 1. Flowchart of the systematic literature review

A method used in this paper is based on PRISMA Statement developed by Moher et al [16]. Literature searching were involved comprehensive and scooping steps has been done by Catenanu et al [17]. Scientific papers regarding feature extraction and classification on Satellite data are selected from reputable four electronic included IEEE, Springer link, Scopus, Science Direct to give comprehensive result and discussion as representative of updated technique in feature extraction and classification.

To ensure that relevant papers obtained, appropriate keywords are selected. For instance, the keyword (“PCA”) and (“multispectral”), (“GLCM”) and (“Multispectral”), (“texture feature”) and (“hyperspectral”), (“hyperspectral”) and (“classification”) conducted for searching the journal. Briefly, there are five steps review the paper started from Topic selection finished with paper publication. Flowchart of systematic review shown on the figure 1.

3. Result

There are about 20 paper journals selected for this systematic review. Most of them focus on feature extraction and classification satellite images. PCA and texture analysis like GLCM widely applied on various images. We only include the journal with scope related to the topic. One of the specific journals for this topic is ISPRS Journal of Photogrammetry and Remote Sensing published by Elsevier (1,7,11,12,19,20,21). The others paper journal selected are from Journal Remote Sensing of Environment on Science Direct (2, 4 and 5). The review process considering parameters from PRISMA statement such as 1) study selection study characteristics, description study, data quality and discussion.

3.1 Study Selection

Overall, 17 potentially papers relevant to updated techniques of feature extraction and classification were identified and selected to be reviewed. All papers derived from trusted online journal.

3.2 Study Characteristic

There are various techniques nowadays to be used for feature extraction and classification data, particularly for multispectral data or hyperspectral data. 4 papers regarding Eigen value has been chosen to be reviewed in this project. Murugan *et al* have developed experiment research by conducting PCA (principal component analysis) to extract information feature from remote sensing image [18]. A similar technique has been used by Lin *et al* [19], by analyzing Eigen feature to classify LIDAR cloud point. There were also newly developed techniques as modified PCA to do feature extraction or classification. Zhong *et al* developed the new technique called blind spectral inmixing based on sparse component analysis (SCA) for hyperspectral data[1]. While Zalbaba *et al* also developed novel-folded PCA to improve feature extraction and data reduction applied to hyperspectral image and SAR (synthetic aperture radar) in remote sensing [20].

Furthermore, Lu *et al* [7], developed the concept of spectral-spatial classification method where each test and training sets are adaptively represented by AH model. This model utilizes both similarity and variance of pixels to adaptively characterize the sets. In addition, AH-model-based able to represent each set with a similarity between testing and training sets is transformed into a geometric distance. Finally, labeling test sets are coming from a minimum distance between test and training sets.

A proposed method classification accuracy improved greatly from combining both spatial and spectral information of the HSI in a set-by-set manner. In contrast to Lu *et al*, both Pan *et al* [9] and Li *et al* [8] are using PCA as method to reconstructed the hyperspectral images. Li *et al* used PCA as a method to reconstruct the hyperspectral images. Li *et al* used PCA as a method to decorrelate the atoms in the spectral dictionary then followed by eigenvectors due to their powerful expressive ability. Moreover, after spectral dictionary describing the spectral characteristics, singular value decomposition (SVD) is used in group-based sparse representation to obtain the adaptive dictionary for each group to effectively analyze spatial similarity and spectral correlation from two different dimensions of a group. Slightly similar to proposed method that Li *et al* proposed, Pan *et al* [9] proposed classification method by using modified PCA named PCANet and their classification method named spectral-spatial classification network.

Table 1. General Description of Studies

Study	Objective	Method	Result	Further Research Proposed
Murugan <i>et. al</i> (2011), India	Reduction in the dimension of RSI by data compression technique using Eigen matrix	PCA	A prominent hill structure at the middle of the RSI is shown very clearly. The linearity of the structural hill is lucidly displayed sharply	Applying various algorithms such as filtering, edge detection and classification
Zhong <i>et. al</i> (2016), China	To consider the properties of sparsity for hyperspectral remote sensing imagery	Sparse component analysis (BSUSCA)	The results show that BSUSCA performs better than the other unmixing algorithms, in an overall view	SCA-based endmember number estimation algorithm will be added to the model
Zalbaba <i>et. al</i> (2014), UK	To address challenges regarding memory management	Novel Folded-PCA (F-PCA)	Overall computational cost reduced to roughly 10%, Folded-PCA requires less than 1% of memory as those in conventional PCA	To apply this approach in other analysis of large dimensional dataset is required
Lin <i>et. al</i> (2014), Taiwan	The introduction of a method to extract Eigen-features from a local point set by using the weighted PCA	SVM classification and Eigen feature analysis	The accuracy of our classification improves compared to the classification using standard Eigen-features	To investigate the application of the method on the point clouds captured from laser terrestrial mobile mapping systems
Kautsias <i>et. al</i> (2009), Greece	Enhancing signal of burnt surface	Forward/ backward PCA	Forward / backward PCA improved spectral discrimination	Not mention specifically
Kang <i>et. al</i> (2013), China	To propose Image Fusion Recursive Filter (IFRF) for feature extraction of hyperspectral image	IFRF Support Vector Machine (SVM)	IFRF efficiently can reduce the dimension of the hyperspectral image.	Adoption of EPFs method to process fused band images and investigation of the considering the correlation degree of adjacent band
Zhang <i>et. al</i> (2011), China	The objective of research is Mapping the salt filed from Fused Landsat TM data	PCA	Salt farm mapped 91.95 % correctly and rose up to 94.6 % for panchromatic band (15 m). The Highest accuracy with PCA-fused data is 98.8%.	Not mention specifically
Su, <i>et. al</i> (2013), China	To detect Sea Ice base on texture information	GLCM	GLCM texture analysis is the best method compared to other methods based on confusion matrix. The result are validated by co-temporal HJ1B-CCD 30m imagery by visual interpretation	Texture analysis will be carried out for different types of Sea Ice using remote sensing technique.
Ban <i>et al</i> (2014), Sweden	To evaluate spaceborne SAR data for improved global urban mapping	KTH-Pavia urban extractor, GLCM	KTH- Urban extractor is effective technique for extracting urban from SAR image data	KTH-Pavia can be applied for monitoring urbanization if we have historical images.
Lu <i>et. al</i> (2016), China	To classify hyperspectral images (HSI)	Distance Based Classification	Distance based classification have the highest accuracy and faster running time among others methods such as SVM, SVM -CK, LBP etc	Investigating about ability of set-based method for semi-supervised or semi-unsupervised framework
Li <i>et. al</i> (2016) , China	The Development the framework for utilizing spectral mixture analysis spatio spectral group sparsity	Sparse representation matrix	Sparse matrix representation not only maintain spectral consistency but also produce image more detail	Computation times within an acceptable range and the acceleration of the algorithm will be focused
Pan <i>et al</i> (2016), China	To construct simple nonlinear spectral-spatial network	Deep Learning Neural Network	Nonlinear Spectral-Spatial Network (NSSNet) is simplified Deep Learning based performs some deep learning method	The future work will be devoted to reducing training data and improving the accuracies

The literature has been reviewed were quantitative experimental research. It can be understood from research design or research methodology conducted and how the result validated.

Commonly the multispectral/hyperspectral data analysis used to be recognized whether supervised or unsupervised technique. From 17 paper selected, all paper regarding the supervised technique. Although [7][8][9][11][12] [21] focus on the feature extraction technique, a classification method is a way to recognize the images. Meanwhile, [13]–[15] specifically focus on the classification for multispectral/hyperspectral images.

3.3 Description of the Study

Research selected related to Eigen value came from multi countries such as India, United Kingdom, China, Taiwan, Greece, and Sweden. According to data used, Murugan et al were using multispectral data in their research [18]. Zhong et al and Zalbaba et al were using hyperspectral data [1] [20], but Zalbaba et al were also added SAR image in their experiment [20]. Lin et al applied their project to LIDAR cloud point [19]. Table 2 shows specification of data used by each research.

Table 2. Literature Data Used

Study	Type	Data	# of bands
Murugan <i>et al</i> (2011), India	Multispectral	IRS-1C LISS III image	4
Zhong <i>et al</i> (2016), China	Hyperspectral	AVIRIS, HYDICE	224, 210
Zalbaba <i>et al</i> (2014), UK	Hyperspectral and SAR image	AVIRIS, micro-Doppler	224, -
Lin <i>et al</i> (2014), Taiwan	LIDAR cloud point	LIDAR	1
Kautsias <i>et al</i> (2009), Greece	Multispectral	Landsat 7 ETM+	7
Kang <i>et al</i> (2013), China	Hyperspectral	Indiana Pine, Pavia Salinas	220
Zhang <i>et al</i> (2011), China	Multispectral	Landsat TM Plus (ETM+)	7
Su, <i>et al</i> (2013), China	Hyperspectral	MODIS	4 (r,g,b, nir)
Ban <i>et al</i> (2014), Sweden	SAR-CVV	ENVISAT ASAR	1
Lu <i>et al</i> (2016), China	Hyperspectral	AVIRIS, ROSIS-03	200, 204, 103
Li <i>et al</i> (2016), China	Hyperspectral & Multispectral	ROSIS-03, HIS, PHI Xiaqiao	103, 191, 80
Pan <i>et al</i> (2016), China	Hyperspectral	AVIRIS, ROSIS-03	200, 103

3.4 Discussion

Regarding PCA method and its modification used in many research papers mentioned previously, there are advantages could be taken from these methods. PCA derives Eigenvalue of a bundle of images whether it is multispectral or hyperspectral to get the best value of image in the form of Eigen matrix image. There were various improvements after applying the method. Murugan et al and Zhong et al were proved that objects in the image were more clearly seen visually [18] [1]. Murugan et al focused on prominent hill structure, linearity of hill structure, water bodies, river coarse that were clearly identified after PCA was applied [18]. Zhong *et al* were interested in interpreting many complicated mineral types, including alunite, kaolinite, chalcedony, muscovite, montmorillonite, andradite, buddingtonite [1]. The method developed, BSUSCA can clearly identified the similar but different mineral species. Zhong et al have stated also that road and roof can be classified sharply [1]. Either than that, Lin et al applied modified Eigen-feature technique to classify LIDAR cloud point [19]. The result of the process was 2 classes, building and no-building. This result is similar with 2 previous papers, that the method was strong to detect the edge in the image in building the contrast between two objects.

PCA and its variance still become an interesting technique for capturing satellite images. Some variance of PCA such as Independent Component Analysis (ICA), forward backward PCA, inverse PCA are adopted for solving the issues on multispectral analysis. Another issue regarding PCA and its modification method is efficiency. By applying novel Folded-PCA (F-PCA), Zalbaba et al have been proved that overall computational cost reduced to roughly 10% and Folded-PCA requires less than 1% of memory as those in conventional PCA [20]. This research can be applied to process big data, such as multispectral data or even hyperspectral data that contained a large number of bands. Texture based analysis also applied by researchers for multispectral images. GLCM technique can be conducted to enrich pixel by pixel information of the data. GLCM give more comprehensive analysis and can improve global classification rate up to 13.5% [21]. On the other hand, the development of machine learning technique has an important role in this research. Even [15] claimed that their method, (NSSNet) outperform from some deep learning technique.

4. Conclusion

The review has shown that use multispectral images with its behavior have an important role in any research field. The technique of feature extraction and classification are still to be developed by researchers for better accuracy. Pixel-based and texture-based feature combined with robust supervised algorithm ensure that the research is promising to be continued for the future work.

5. Reference

- [1] Y. Zhong, X. Wang, L. Zhao, R. Feng, L. Zhang, and Y. Xu, "Blind spectral unmixing based on sparse component analysis for hyperspectral remote sensing imagery," *ISPRS J. Photogramm. Remote Sens.*, vol. 119, pp. 49–63, 2016.
- [2] J. A. Concha and J. R. Schott, "Retrieval of color producing agents in Case 2 waters using Landsat 8," *Remote Sens. Environ.*, vol. 185, pp. 95–107, 2015.
- [3] A. Asgarian, A. Soffianian, and S. Pourmanafi, "Crop type mapping in a highly fragmented and heterogeneous agricultural landscape : A case of central Iran using multi-temporal Landsat 8 imagery," *Comput. Electron. Agric.*, vol. 127, pp. 531–540, 2016.
- [4] Z. Zhu *et al.*, "Including land cover change in analysis of greenness trends using all available Landsat 5, 7, and 8 images: A case study from Guangzhou, China (2000-2014)," *Remote Sens. Environ.*, vol. 185, pp. 243–257, 2015.
- [5] A. Bastarrika, E. Chuvieco, and M. P. Martín, "Mapping burned areas from landsat TM/ETM+ data with a two-phase algorithm: Balancing omission and commission errors," *Remote Sens. Environ.*, vol. 115, no. 4, pp. 1003–1012, 2011.
- [6] E. Ozelkan, G. Chen, and B. B. Ustundag, "Multiscale object-based drought monitoring and comparison in rainfed and irrigated agriculture from Landsat 8 OLI imagery," *Int. J. Appl. Earth Obs. Geoinf.*, vol. 44, pp. 159–170, 2016.
- [7] N. Koutsias, G. Mallinis, and M. Karteris, "A forward/backward principal component analysis of Landsat-7 ETM+data to enhance the spectral signal of burnt surfaces," *ISPRS J. Photogramm. Remote Sens.*, vol. 64, no. 1, pp. 37–46, 2009.
- [8] X. Kang, S. Li, and J. A. Benediktsson, "Feature extraction of hyperspectral images with image fusion and recursive filtering," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 6, pp. 3742–3752, 2014.
- [9] Y. Zhang, P. X. Hu, and J. Gao, "A reflectance spectra-based approach to mapping salt fields using PCA-fused Landsat TM data," *Adv. Sp. Res.*, vol. 47, no. 9, pp. 1490–1496, 2011.
- [10] R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural Features for Image Classification," *Syst. Man Cybern. IEEE Trans.*, no. 6, pp. 610–621, 1973.
- [11] H. Su, Y. Wang, J. Xiao, and L. Li, "Improving MODIS sea ice detectability using gray level co-occurrence matrix texture analysis method: A case study in the Bohai Sea," *ISPRS J. Photogramm. Remote Sens.*, vol. 85, pp. 13–20, 2013.
- [12] Y. Ban, A. Jacob, and P. Gamba, "Spaceborne SAR data for global urban mapping at 30m resolution using a robust urban extractor," *ISPRS J. Photogramm. Remote Sens.*, vol. 103, pp. 28–37, 2014.
- [13] T. Lu *et al.*, "Set-to-Set Distance-Based Spectral – Spatial Classification of Hyperspectral Images," vol. 54, no. 12, pp. 7122–7134, 2016.
- [14] J. Li, Q. Yuan, H. Shen, and S. Member, "Hyperspectral Image Super-Resolution by Spectral Mixture Analysis and Spatial – Spectral Group Sparsity," vol. 13, no. 9, pp. 1250–1254, 2016.
- [15] B. Pan, Z. Shi, N. Zhang, and S. Xie, "Hyperspectral Image Classification Based on Nonlinear Spectral–Spatial Network," *IEEE GEOSCIENCE AND REMOTE SENSING LETTERS*, pp. 3286–3289, 2016.
- [16] D. Moher *et al.*, "Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement," *Syst. Rev.*, vol. 4, no. 1, p. 1, 2015.
- [17] A. Cetateanu and A. Jones, "How can GPS technology help us better understand exposure to the food environment? A systematic review," *SSM - Popul. Heal.*, vol. 2, pp. 196–205, 2016.
- [18] S. Murugan, "Significance of Eigen Matrix in Spectral Domain of Remote Sensing Images (RSI)," *International Journal of Computer Applications (0975 – 8887)*, vol. 35, no. 10, pp. 1–6, 2011.
- [19] C. H. Lin, J. Y. Chen, P. L. Su, and C. H. Chen, "Eigen-feature analysis of weighted covariance matrices for LiDAR point cloud classification," *ISPRS J. Photogramm. Remote Sens.*, vol. 94, pp. 70–79, 2014.
- [20] J. Zabalza *et al.*, "Novel Folded-PCA for improved feature extraction and data reduction with hyperspectral imaging and SAR in remote sensing," *ISPRS J. Photogramm. Remote Sens.*, vol. 93, pp. 112–122, 2014.
- [21] A. Safia and D. C. He, "Multiband compact texture unit descriptor for intra-band and inter-band texture analysis," *ISPRS J. Photogramm. Remote Sens.*, vol. 105, pp. 169–185, 2015.