Detecting craters by training random forest based on existing crater map and spatial structural information

Yan-Wen Wang, Cheng-Zhi Qin*, Wei-Ming Cheng
State Key Laboratory of Resources and Environmental Information System, Institute of Geographic Sciences and Natural Resources Research, CAS
Beijing, China
* Corresponding author: qincz@lreis.ac.cn

A-Xing Zhu
Jiangsu Center for Collaborative Innovation in Geographical Information Resource Development and Application and School of Geography
Nanjing Normal University
Nanjing, China

Department of Geography
University of Wisconsin-Madison
Madison, Wisconsin, USA

Abstract—**A new approach to detecting craters based on DEM is proposed. Main difference between the proposed approach and existing approaches is the use of geomorphometric information with existing crater map through machine learning, while existing approaches mainly consider shape information of craters and ignore spatial structural information of crater interior. The proposed approach includes two stages: 1) to use existing crater map to train a random forest classifier based on multi-scale landform element information, which is then applied to extracting crater candidates; and 2) to use radial topographic profiles of craters on the map to train the other random forest classifier, which is then used to identify how likely each candidate extracted in the first stage is crater. Experimental result of the case study was quantitatively evaluated, compared with that from a representative of existing approaches.**

I. Background

Detecting craters is important for not only scientific aims (such as inferences about the ages and history of planets' surfaces), but also engineering applications (such as spacecraft landing and working). Because manual delineation of craters is too inefficient to fit the actual applications, automatic (or semi-automatic) approaches of detecting craters are needed.

Existing automatic (or semi-automatic) crater detection approaches (CDAs) can be classified as two main types according to the type of main data source used, i.e. image-analysis-based CDAs and terrain-analysis-based CDAs. Image-analysis-based CDAs detect craters based on their optical features or highlight-shadow patterns recorded on remote sensing images, especially on grayscale photographs [1-3]. These features and patterns are mainly originated from the change of light and shade due to craters' rims. However the terrain information, which is key to identifying geomorphic objects, cannot be directly contained in images and thus is considerately ignored in these CDAs.

Terrain-analysis-based CDAs, which are based on digital elevation model (DEM), could use all kinds of terrain information to detect craters, thus have advantages over the image-analysis-based CDAs [4]. Current terrain-analysis-based CDAs often conduct a depression-filling process on DEM to extract round and symmetric boundaries (i.e., possible rims) of crater candidates [4-7]. Then craters are identified among these candidates by morphometric analysis or a machine learning classifier (e.g., C4.5) which is trained based on experts' labeled craters depicted by morphologic attributes (such as diameter, roundness, and depth) [4]. These CDAs mainly consider the morphometric information of craters (sometimes only shapes of crater boundaries) and ignore the spatial structure inside crater. This situation means the limited ability to detect complex craters which might be overlapped and degraded.

This study proposes a new terrain-analysis-based CDA which use geomorphometric information with existing crater map through machine learning, so to effectively consider spatial structural information of crater interior during crater detection.

II. Method

Existing crater map (for part area or adjacent area of the working region of crater detection) implicitly contains domain knowledge on identifying crater which could be applied to the working region. This map can provide plenty of available samples to train machine learning classifier for CDA, which is more efficient and automatic than the conventional way of manual sample collection or manual rule assignment for classifier.

Furthermore, geomorphometric information derived from DEM can depict crater from not only morphologic (as the situation in existing CDAs) but also spatial structural perspectives, which could be used as features for machine learning. Above basic idea directs the design of the proposed approach in this study.

Similar to existing CDAs, the proposed approach also includes two stages (Fig. 1) to fulfill two tasks respectively, i.e. extracting crater candidates, and identifying craters among candidates.

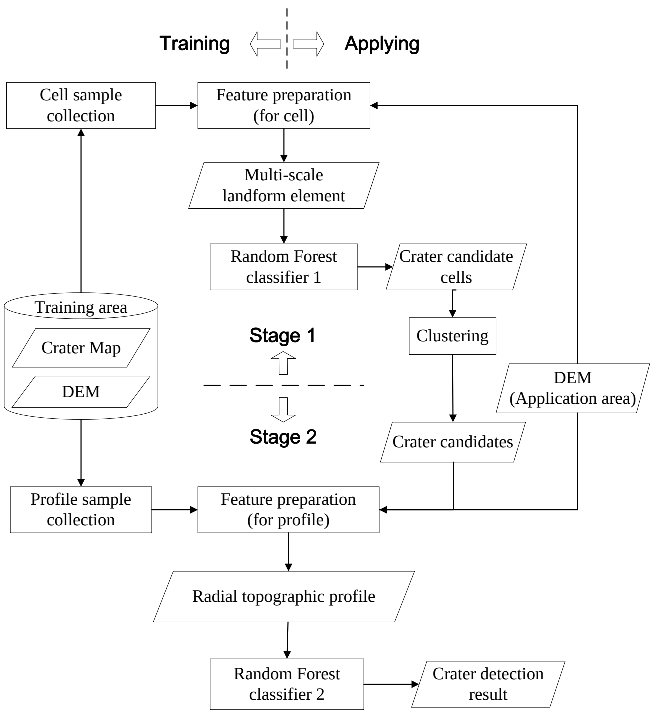

Figure 1. Workflow of the proposed approach

A. Stage 1: extracting crater candidates

Note that the depression-filling way often used for extracting crater candidates in existing CDAs ignores the spatial structural information of crater interiors and might not work for those overlapped and/or degraded craters. In this stage we use a machine learning method (i.e., the random forest, RF for short) to extract possible cells. Samples for training a RF classifier are directly collected from existing crater map. Positive samples are cells within craters on the map, while negative samples are the cells selected outside craters.

Feature selection is key to RF training. From the geomorphometric perspective, craters normally show a center-periphery structure with lower, wider flat interior (with uplifted center sometimes) surrounded by a higher, narrower rim with a larger slope. The terrain of a cell inside crater, together with its surrounding area, often shows characteristic convexity-concavity variation across different analysis scales (i.e., size of analysis window). Due to the local variation in DEM, the topographic attributes (e.g., slope gradient, curvature, and so on) or landform element type derived normally at a specific analysis scale cannot depict this characteristic well. In this stage the features for RF input are selected to be the multi-scale landform element [8] which is proposed to revise the Geomorphons method [9] across a series of analysis scale [8].

After trained with multi-scale landform element information on samples from existing crater map, the RF classifier is used to extract crater candidate cells in the application area. Then a clustering process is conducted on these crater candidate cells to form crater candidates (i.e., boundaries of individual craters).

B. Stage 2: identifying craters among candidates

Note that the center-periphery structure of craters can be depicted typically with radial topographic profiles. In this stage the other RF classifier is trained with radial topographic profile samples collected based on existing crater map. Positive samples for training this RF classifier are radial topographic profiles originated from the centers of circular craters on the map, while negative samples are those randomly selected outside the craters. For each radial topographic profile sample, the features as RF input are organized as a 1×10 vector with each feature recording a normalized elevation value in order along the profile.

After trained with radial topographic profile samples based on existing crater map, this RF classifier can be used to identify how likely each candidate extracted in the first stage is crater. A crater candidate will be identified to be crater when this crater candidate has a large enough count of radial topographic profiles classified to be crater's profiles by the trained RF.

III. EXPERIMENT

A. Study areas and data

LOLA (Lunar Orbiter Laser Altimeter) [10] compiled by experts was adopted as the crater map for training the proposed approach, as well as evaluation data. Lunar DEM with a resolution of 500 m from the Chang'E-1 satellite [11] was used in this experiment.

Both training area and application area are in middle and lower latitude on the lunar farside (Fig. 2). The training area (160.0 °W~149.5 °W, 28.8 °N~36.9 °N) is about 7.8×10^4 km^2, and the application area (107.4 °E~133.8 °E, 13.6 °N~33.3 °N) is about 4.8×10^5 km^2 (Fig. 2).

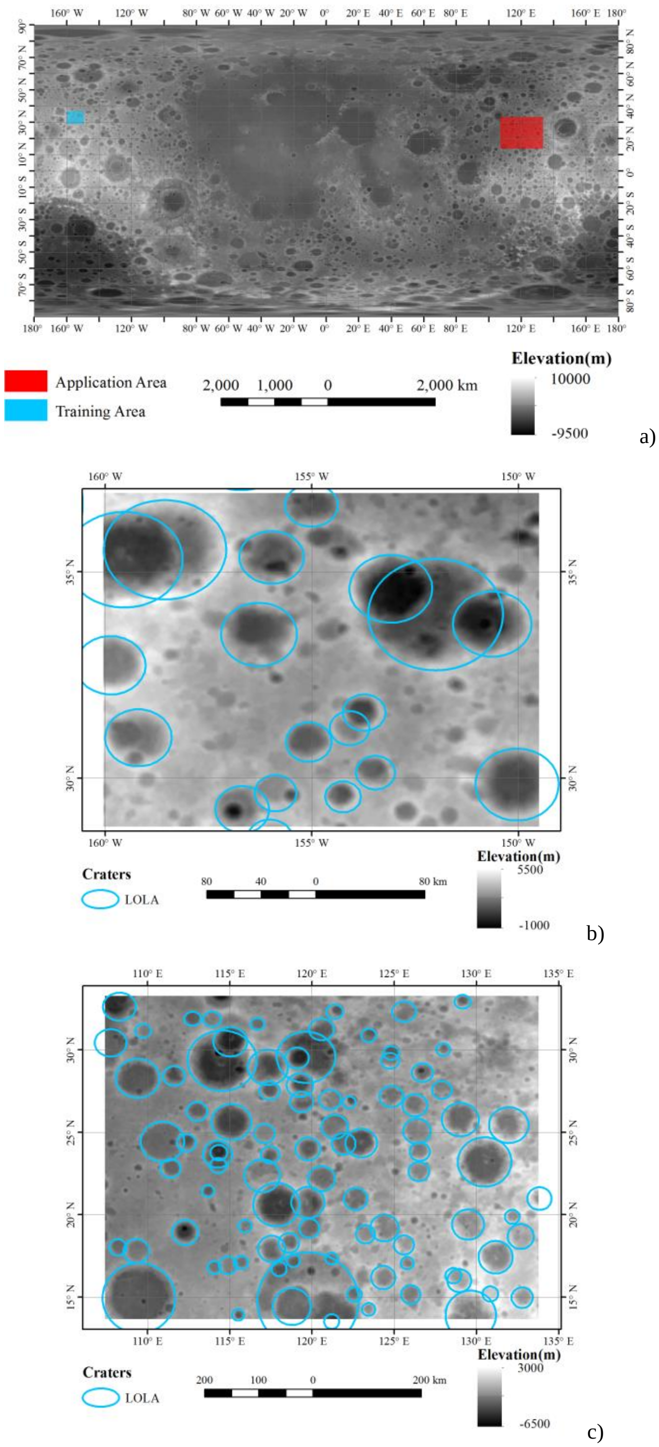

Figure 2. Study area maps: a) location map; b) training area; c) application area.

B. *Application of the proposed approach*

For the first stage of the proposed approach, the analysis scales (i.e., the radii of circular analysis window) for deriving multi-scale landform element information were set to be from 3 km to 60 km with a step of 1 km. Thus the features for each cell being input RF are organized as a 1×58 vector with each feature recording a landform element type identified at the corresponding analysis scale. During training the RF for each stage of the proposed approach, the maximum of iterations was set to be 200. For creating crater candidates from crater candidate cells, DBSCAN algorithm was performed with a neighborhood searching radius of 2.5 km (i.e., 5 cells) and a minimum of 10 neighboring crater candidate cells for valid clusters.

In the second stage of the proposed approach, 12 radial topographic profiles (starting from due north with a step of 30°) created from each crater and crater candidate were input the RF during training and applying, respectively. The proposed approach identified a crater candidate to be a crater when more than half of the 12 radial topographic profiles of the crater candidate were classified as radial topographic profiles of a crater by the trained RF.

C. *Evaluation method*

Compared with the AutoCrat approach [4], the proposed approach was evaluated based on the LOLA crater map in the application area. Whether a crater '*P*' identified by the proposed approach matches a crater '*T*' on LOLA map was judged based on whether the ratio of the intersection area to the union area between *P* and *T* is larger than a preset *C_threshold* value (0.3~0.7 tested in this experiment).

D. *Experimental results and discussion*

In the application area which has 92 craters according to LOLA, the proposed approach and AutoCrat extracted 99 and 83 craters, respectively (Fig. 3; Table 1). When *C_threshold*=0.5, a total of 63 craters identified by the proposed approach matched craters in LOLA, while this number for AutoCrat was 49. When 40 craters in LOLA were consistently identified by both approaches, 23 and 9 craters in LOLA were identified only by the proposed approach and AutoCrat, respectively. This shows that the proposed approach performed better than AutoCrat. For other *C_threshold* values tested, there are similar situation which the proposed approach performed well (Table 1). The count of identified craters which were totally nonoverlapping with the craters in LOLA was 17 for the proposed approach and 22 for AutoCrat. Some of them might be craters missed in LOLA, which requires further analysis.

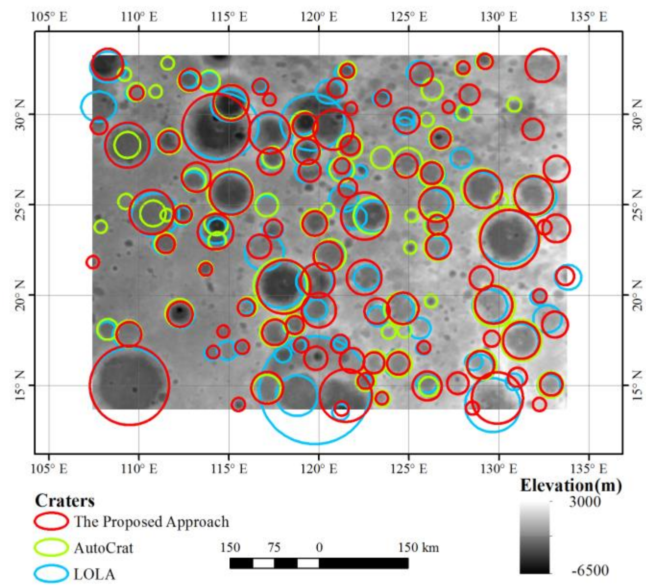

Figure 3. Craters identified by the proposed approach in the application area.

TABLE I. MATCH BETWEEN CRATERS IDENTIFIED BY THE PROPOSED APPROACH AND CRATERS IN LOLA IN THE APPLICATION AREA, COMPARED WITH THE RESULTS FROM AUTOCRAT

C_threshold	The proposed approach		AutoCrat	
	Match	Mismatch	Match	Mismatch
0.7	43	56	37	46
0.6	57*	42	44	39
0.5	**63***	**36**	**49**	**34**
0.4	69*	30	49	34
0.3	73*	26	51	32

*: The proposed approach performed better than AutoCrat, with a very significant level (P<0.01).

IV. SUMMARY

Combining random forest with existing crater map and geomorphometric information, the proposed approach can effectively consider spatial structural information of crater interior during crater detection.

The performance of the proposed approach will be further investigated, so to explore potential improvement on both the design of the proposed approach and its parameter-settings.

The design of the proposed approach provides an example of the mining and use of the domain knowledge implicitly contained in an existing geomorphic type map through machine learning with geomorphometric information (not only morphologic information but also spatial structural information). This could be potentially useful for the design of approach to extracting other geomorphologic types (such as volcanos, river terraces, and alluvial fans), especially in regions where have been only partly mapped by experts.

ACKNOWLEDGMENTS

This study was supported by the National Natural Science Foundation of China (No. 41422109, 41431177).

REFERENCES

[1] Urbach, E. R., and T. F. Stepinski, 2009. “Automatic detection of sub-km craters in high resolution planetary images”. Planetary and Space Science, 57(7), 880-887.

[2] Salamunićcar, G., and S. Lončarić, 2010. “Method for crater detection from Martian digital topography data using gradient value/orientation, morphometry, vote analysis, slip tuning, and calibration”. IEEE Transactions on Geoscience and Remote Sensing, 48(5), 2317-2329.

[3] Ding, M., Y. Cao, and Q. Wu, 2013. “Novel approach of crater detection by crater candidate region selection and matrix-pattern-oriented least squares support vector machine”. Chinese Journal of Aeronautics, 26(2), 385–393.

[4] Stepinski, T. F., M. P. Mendenhall, and B. D. Bue, 2009. “Machine cataloging of impact craters on Mars”. Icarus, 203(1), 77-87.

[5] Bue, B. D., and T. F. Stepinski, 2007. “Machine detection of Martian impact craters from digital topography data”. IEEE Transactions on Geoscience and Remote Sensing, 45(1), 265-274.

[6] Wan, C., W.-M. Cheng, Z.-P. Zhou, S.-M. Zhao, and Y. Xia, 2012. “Automatic extraction of lunar impact craters from Chang’E-1 satellite photographs”. Science China – Physics, Mechanics & Astronomy, 55(1), 162-169.

[7] Liu, D., M. Chen, K. Qian, M. Lei, and Y. Zhou, 2017. “Boundary detection of dispersal impact craters based on morphological characteristics using lunar digital elevation model”. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 10(12), 5632-5646.

[8] Kang, X., Y. Wang, C. Qin, W. Cheng, S. Zhao, A. Zhu, and W. Zhang, 2016. “A new method of landform element classification based on multi-scale morphology”. Geographical Research, 35(9), 1637-1646. (in Chinese with English abstract)

[9] Jasiewicz, J., and T. F. Stepinski, 2013. “Geomorphons: A pattern recognition approach to classification and mapping of landforms”. Geomorphology, 182, 147-156.

[10] Kadish, S. J., C. I. Fassett, J. W. Head, D. E. Smith, M. T. Zuber, G. A. Neumann, and E. Mazarico, 2011. “A global catalog of large lunar crater (≥20 KM) from the Lunar Orbiter Laser Altimeter”. In 42nd Lunar and Planetary Science Conference, March 7-11, The Woodlands, Texas. LPI Contribution No. 1608, p.1006.

[11] Li C.-L., X. Ren, J.-J. Liu, X.-D. Zhou, L.-L. Mu, J.-Y. Wang, R. Shu, Y.-L. Zhou, H.-B. Zhang, C. Lǜ, J.-Z. Liu, W. Zuo, Y. Su, W.-B. Wei, W. Bian, M. Wang, C. Xu, D.-Q. Kong, X.-Q. Wang, F. Wang, L. Geng, Z.-B. Zhang, L. Zheng, X.-Y. Zhu, J.-D. Li, and Z.-Y. Ouyang, 2010. “Laser altimetry data of Chang’E-1 and the global lunar DEM model”. Science China - Earth Science, 53(11), 1582–1593.