ESTIMATION OF FALSE DETECTIONS FOR EVALUATION OF CRATER DETECTION ALGORITHMS. G. Salamunićcar[1] and S. Lončarić[2], [1]AVL-AST d.o.o., Av. Dubrovnik 10/II, HR-10020 Zagreb-Novi Zagreb, Croatia, gsc@ieee.org, [2]Faculty of Electrical Engineering and Computing, University of Zagreb, Unska 3, HR-10000 Zagreb, Croatia, sven.loncaric@fer.hr.

Introduction: A method for estimation of false detections for crater detection algorithms (CDAs) is proposed. In combination with known ground truth (GT) and other available analyses, the proposed method can improve evaluation of CDAs.

Evaluation of CDAs: CDA applications include dating of a planetary surface [1], autonomous landing to planets [2] and asteroids [3], and advanced statistical analysis [4]. The most common evaluation of CDAs is visual comparison of detected craters with topography. More advanced evaluations include comparison with craters' catalogue(s) [5-7] and cumulative size-frequency distribution(s) [1] provided by planetary scientists, as well as the use of Receiver Operating Characteristics (ROC) [8-12]. For those evaluations, GT (locations and sizes of known craters as labeled by geologists) is required.

Limitation of Known GT: For estimation of GT, a catalogue with 17582 craters was assembled, wherein each crater is aligned with MOLA topography and confirmed by three independent sources: (1) catalogue from N. G. Barlow et al. [13]; (2) catalogue from J. F. Rodionova et al. [14]; and (3) revised version of catalogue used in previous work [15]. The new GT catalogue can confirm with very high accuracy that detection of some crater by a CDA is correct - True Positive (TP). For a crater that is not in the GT catalogue, the best that can be done is to consider it as incorrect detection - False Positive (FP). However, as shown in Fig. 1, where we have three registered craters from the above catalogue, six Quasi-Circular-Depressions (QSDs) marked by Frey et al. [16] are not included in GT. Accordingly, correct detection of such QSD by a CDA will be evaluated as FPs, which may not be correct.

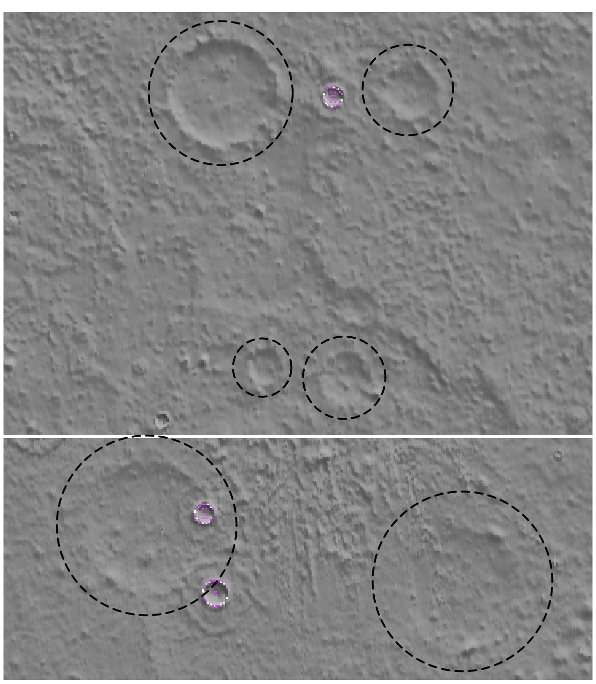

Figure 1: Three registered craters that are part of GT and six QSDs marked by Frey et al. that are not.

Further Classification of FPs: While classification on TPs and FPs is required for e.g. FROC analysis of CDAs, GT can only confirm for TPs with high accuracy that they really are correct detections, as shown above. Accordingly, contribution to future evaluations of CDAs is expected from further classifications of FPs (detected craters that are not part of GT) on FPs that most probably are incorrect detections and FPs that should be considered in further extension of the current GT catalogue.

Similar Topography Without Craters: Artificial topography as much as possible similar to the real one but without any crater can be very useful. For such a topography, ideal CDAs would not detect any crater. However, what should be expected is that CDAs will detect some craters, for which we know with 100% accuracy that they are incorrect detections. It is to expect that percentage of those incorrect detections will be smaller than percentage of FPs from real topography. Difference are those craters that are probably correctly detected but that are not included in GT catalogue. Main limitation of the approach are possibilities in creating artificial topography, and so the methodology should not be used as replacement of e.g. FROC analysis, but as a complimentary evaluation of CDAs.

Topography Inversion: One way to crate artificial topography without any craters is to invert existing topography, or, in other words, to multiply with -1 value of altitude associated to each pixel from e.g. MOLA data in 1/64° resolution. Craters will become hills, while hills with quasi-circular shape will become false craters which should not be detected by CDAs. While such an artificial topography is very similar in a number of characteristics to the real one, it should also be noted that there are number of disadvantages. Volcanoes in inverted topography will become similar to craters, and will cause underestimation of results. On the other hand, in inverted topography we will not have valleys and canyons like in the real one, which will cause overestimation of results. However, it is known that in inverted topography CDAs should not detect craters because there are none, and so such an evaluation is helpful.

Evaluation of Methodology: For testing purposes, Radon/Hough transform (RH) [17] based CDA was implemented. Input parameters are global topography of Mars (MOLA data) in 1/64° resolution, and defined slope, required in order to mark some pixel as 1 or 0 before RH is performed. Detections where center is on lower altitude than rim are marked in Table 1 as Craters (*C*), while others are marked as Hills (*H*) and actually represent "craters" detected in inverted topography, which should not be detected by CDA. The next column is *GT*, percentage of *C* found in GT catalogue with 17582 craters. Percentage of estimated false detections using inverted topography approach *H/C* is shown in next column, in form *1 - H/C* to be comparable to *GT*. More precise computation of *1 - H/C* is shown in last column as *P*, where results for craters/hills are evaluated separately for different number of votes, after which the sub-results are combined into final proportion of correctly detected craters. With increase of required slope, percentage of TPs first increases and then decreases. Explanation is that for small required slope, density of pixels marked as 1 is so high, that rims of some craters can not be detected. For large required slope, rims of craters with too small slope can not be detected.

Conclusion: Both methodologies confirmed similar pattern regarding the RH CDA method. They both show that best compromise for RH is inside some particular slope range as shown in Table 1. However, desired behavior for CDA is not to depend on slope. While 14134 craters from GT were correctly detected (80.39%) for slope 4%, the number of false detections is much higher as confirmed by both methodologies. As shown in Table 1, there are differences between estimations given by *GT* and *P*, which is expected because all craters that are not in GT are not necessarily false detections. For more advanced CDAs, lower *H* and higher *P* is expected. In order to evaluate that, estimation of false detections should be used in combination with known GT and FROC analysis.

References: [1] Sawabe Y. et al. (2005) *Adv. Space Res.*, in press. [2] Cheng Y. et al. (2005) *IEEE Conference on Computer Vision and Pattern Recognition*. [3] Leroy B. et al. (2001) *Image and Vision Computing*, *19*, 787-792. [4] Salamunićcar G. (2004) *Adv. Space Res.*, *33*, 2281-2287. [5] Michael G. G. (2002) *Microsymposium 36*, MS068. [6] Michael G. G. (2003) *Planetary and Space Science*, *51*, 563-568. [7] Michael G. G. (2003) *Microsymposium 38*, MS067. [8] Burl M. C. et al. (2001) *6th Int. Symposium on Artificial Intelligence, Robotics, and Automation in Space*, AM118. [9] Honda R. et al. (2002) *Progress of Discovery Science*, 395-407. [10] Vinogradova T. et al. (2002) *IEEE Aerospace Conference*, 7-3201 - 7-3210. [11] Vinogradova T. and E. Mjolsness (2001) *AGU Fall Meeting*, Abstract #P42A-0563. [12] Wetzler P. G. et al. (2005) *7th IEEE Workshop on Applications of Computer Vision*, 178-184. [13] Barlow N. G. et al., *http://webgis.wr.usgs.gov/mars.htm*. [14] Rodionova J. F. et al., *http://selena.sai.msu.ru/Home/Mars_Cat/Mars_Cat.htm*. [15] Salamunićcar G. (2003) *LPS XXXIV*, Abstract #1403. [16] Frey H. V. et al. (2002) *GRL*, *29*, 10.1029/2001GL013832. [17] Deans S. R. (1981) *IEEE PAMI*, *3*, 185-188.

Table 1: Results for RH based CDA for specified slope and required number of votes. The number of detected craters/hills, % of C in GT, and two estimations of false detection rate (1-FDT) are shown.

slope & % of votes/votes^2	craters (C)	hills (H)	GT %	1-H/C %	P %
2, 0/0	**96754**	**101496**	**14.59**	**0**	**5.93**
2, 70.71/50	70899	71176	19.73	0	8.10
2, 100/100	35232	31018	31.53	11.96	11.96
4, 0/0	**75869**	**71310**	**18.63**	**6.01**	**23.00**
4, 70.71/50	49641	40187	27.24	19.04	28.22
4, 100/100	20245	13240	40.30	34.60	34.60
6, 0/0	**63744**	**55314**	**20.54**	**13.22**	**31.85**
6, 70.71/50	39053	27742	30.41	28.96	36.92
6, 100/100	12038	6526	46.68	45.79	45.79
8, 0/0	**55743**	**44769**	**20.91**	**19.69**	**37.93**
8, 70.71/50	30789	19793	32.02	35.71	42.03
8, 100/100	6880	3305	54.13	51.96	51.96
10, 0/0	**49498**	**37417**	**20.24**	**24.41**	**41.74**
10, 70.71/50	23479	14298	33.12	39.10	44.61
10, 100/100	3920	1754	60.33	55.26	55.26
12, 0/0	**44049**	**31346**	**18.86**	**28.84**	**45.11**
12, 70.71/50	17235	9965	34.18	42.18	45.99
12, 100/100	2202	1126	57.99	48.86	48.86
14, 0/0	**38738**	**26562**	**17.32**	**31.43**	**46.72**
14, 70.71/50	12094	6832	36.89	43.51	45.53
14, 100/100	1230	765	45.45	37.80	37.80
16, 0/0	**33310**	**22556**	**16.19**	**32.28**	**47.06**
16, 70.71/50	8243	4750	40.24	42.38	42.64
16, 100/100	690	617	27.10	10.58	10.58
18, 0/0	**27963**	**19334**	**15.71**	**30.86**	**45.88**
18, 70.71/50	5480	3293	42.94	39.91	39.91
18, 100/100	439	479	11.39	0	0
20, 0/0	**23122**	**16316**	**14.90**	**29.44**	**43.67**
20, 70.71/50	3738	2391	40.34	36.04	36.04
20, 100/100	305	391	3.93	0	0