Tina Memo No. 2010-012
Internal Report rejected from BMVC 2011

Towards a Quantitative Segmentation of Martian Terrains.

P. Tar, N.A. Thacker.

Last updated
13 / 3 / 2011

ISBE, Medical School,
University of Manchester,
Stopford Building, Oxford Road,
Manchester, M13 9PT, UK.

Abstract

There is demand within the planetary science community for automated image analysis tools. Researchers in these areas commonly require estimated quantities of geological features appearing within satellite images for comparison against theoretical models. Such comparisons can only be meaningful if quantities are accompanied by appropriate estimates of uncertainty (e.g. error bars). Here, we show it is possible to use Bayes Theorem, in conjunction with a lower variance bound and error propagation, to make unbiased probabilistic measurements of quantities of features within theoretically computed confidence intervals. We also investigate methods of representing local image patch information and show a binary encoding, similar to BRIEF, may be an optimal solution for modelling local structures. Our approach is applied to the segmentation of Martian terrain types where it is shown to be reliable and generalisable.

1 Introduction

Imaging of planetary surfaces is a high profile element of missions to explore the solar system. Over the last 10-15 years a series of orbiting missions have returned images of the Martian surface with unprecedented resolution (down to 0.25 m per pixel [1]), while the current missions Messenger [2] and Dawn [3] will provide high resolution coverage of the surfaces of Mercury and Vesta respectively. Such missions present an unrivalled opportunity to understand the history of geological processes that have shaped the surfaces of the terrestrial planets and ice moons.

Planetary surfaces respond to and record processes driven from both above and below the surface. The nature and relative importance of these processes varies from place to place and, in any one place, through time; the current morphology of the landscape is thus a product of (and hence record of) ancient and recent variations in processes, in some cases driven by climate change. For example, variation in the rate of uplift of a surface alters the slope of the land, affecting the direction and pattern of drainage networks, and degree of dissection of the surface. In geomorphological studies on Earth, the scale and pattern metrics of landscape organization are qualitatively and quantitatively analysed to determine the principal active processes responsible for the evolution of the terrain [4]. Recent developments in planetary imaging allow such techniques to be adapted to understand the evolution of other planets.

Automated image analysis has the potential to be a key component in this enterprise, especially in the case of Mars. The quantity of available imagery has surpassed the capacity of individuals to make first pass assessments seeking the presence or absence of textures diagnostic of particular processes. In addition to selecting promising candidates for further study, possible applications of an automated 'sifting' system include the identification of images containing evidence of drainage networks or dessication, whose distribution across a planetary surface is diagnostic of past or present climate patterns [5] [6]. Another potential application relates to the determination of relative surface ages from size frequency distributions (SFDs) of impact craters [7]. The ability to place geomorphic features in a chronological sequence, and hence to establish relative timing in landscape, is vital to developing and testing, quantitatively, models of landscape evolution.

For any data to be scientifically useful, outputs must be quantitatively valid, which for the purposes of this paper will be defined as 'reflecting underlying ground truths within clearly stated confidence intervals'. For example, summary outputs may involve plots showing the quantities of features present accompanied by error bars allowing results to be meaningfully compared to theory by scientific researchers. Automatically generated SFDs must be capable of generating such error bars to fulfil scientific researchers' needs. Achieving this quantitative validity for a variety of outputs, including SFDs, is the focus of our current research. On some level each of these applications involves the recognition and counting of pixels belonging to different classes of geological/geomorphological feature.

As a proof of concept this paper will focus on a simpler problem of pixel-wise segmentation of Martian terrains at a fixed scale and orientation. The aim is to produce quantitative estimates of the amounts, in pixels, of different terrain types within a set of test images. This is achieved using a flexible local image descriptor (FLID) to encode local image patch information, a classifier based upon Bayes Theorem, and the application of a lower variance bound along with error propagation for the construction of error bars.

Existing work on the analysis of planetary surfaces has concenrated on the detection of impact craters using Hough transforms [9] [11] [12] [14] [15] [16], ellipse or circle fitting [13] [15], various types of template matching [9] [14], binary image processing on shadows [10], the Haar transform [17] [18], grey level co-occurrences and appearance modelling [19]. To improve accuracy some methods pre-process images to determine regions of interest using grey level thresholds [10] and Grey Level Co-occurrence Matrices (GLCM) [13]. Some methods use post-processing such

as morphological operations to join edges or remove isolated edges [8]. Despite such filtering many false positives are reported which have been further filtered using size thresholds [10], neural networks and models in eigenspace [13], or cross-referencing against alternative detection methods [14].

The segmentation of Martian terrains can be seen as a texture analysis problem. Recent BMVC work on texture analysis includes the use of Texture Fragmentation and Reconstruction (TFR) [23] and Markov Random Fields (MRF) [24] where textures are described in graph form. Textures have also been modelled using extended Fields-of-Experts [25] with performance demonstrated on synthesis tasks. Classification of textures has been performed using Random Forests (RF) by randomly sampled pixels [26], and Naive Bayes Classifiers (NBC) compared with Naive Credal Classifiers (NCC) using Local Binary Pattern descriptors [27]. In an attempt to mitigate against erroneous classifications the latter argues that returning a set of potential classifications from an NCC is better than forcing a single class decision.

A quantitative error assessment of craters detected using Support Vector Machines (SVM) has been made by comparing results against confidence intervals set by human performance [8]. However, this work did not explicitly generate confidence intervals for the SVM results.

Methods have also been used to extract other features from Digital Elevation Maps (DEM). These include river channels using paths of steepest decent [20] [21] and fault scarps using wavelets [22].

From a quantitative perspective, we are not aware of any reported methods which attempt to estimate quantitation error. Performance is reported only in absolute terms based upon discrete counts of successes and failures. Research has been done into the quantitative use of Gaussian Process Latent Variable Models [28], for which results have been reported which show interpolated values within theoretically computed confidence intervals. Our work, whilst very different in method, has similar aims in trying to systematically produce meaningful error bars on summary outputs.

2 Methodology

At the core of our system is Bayes Theorem:

$$P(k|X) = \frac{P(X|k)P(k)}{\sum_{l \in K} P(X|l)P(l)} \tag{1}$$

where k is a classification of Martian terrain from the set of all terrains, K; X is a vector of evidence observed at a location within an image; $P(k|X)$ is the probability that class k is the cause of observation X; $P(X|k)$ is the probability density of X for class k; and $P(k)$ is the probability of classification k existing within the image under analysis. From this our investigations focus on two main issues:

- the selection of an appropriate representation vector, X, to encode local image patch information;
- and the quantitative use of the resulting probabilities, $P(k|X)$, to estimate quantities of terrain types present within test images.

Inspired by representations which make use of relative intensity differences, such as the Haar transform used by Viola-Jones [29], and BRIEF [30], we have developed a flexible local image descriptor (FLID) for comparing sets of pixel pairs. The FLID has the following parameters:

- A set of N pixel pairs $\{(\alpha_1, \beta_1), (\alpha_2, \beta_2), \ldots, (\alpha_N, \beta_N)\}$;
- an operating radius, r, within which pairs are randomly selected;
- a level of quantisation, q, which determines the level of detail at which pixel intensities will be compared;
- and threshold, θ, which determines the significance of the quantisation levels.

For an image location a vector, $X = (x_1, x_2, \ldots, x_N)$, can be constructed for pixel pairs sampled around the location within radius r. Each element, x_i, is computing by taking the difference in intensity between pixel α_i and β_i. This difference is quantised to integer values between 0 and q if asymmetrical encoding is used, or between $-q/2$ to $q/2$ if symmetrical encoding is used. The interval, with respect to the original grey levels, between each quantisation level is determined by θ. Any grey level differences above or below the limits of quantisation are clipped.

If threshold, θ, is set to 1 standard deviation of image noise the determination of vector elements amounts to a series of hypothesis tests. In the case of asymmetrical encoding these tests return the number of standard deviations by which pixel α is brighter than β, returning 0 if α is equal to or darker than β. In the case of symmetrical encoding these tests also return negative numbers corresponding to cases where β is brighter than α.

To keep the representation simple, probability density estimates, $P(X|k)$, are sampled into N-dimensional histograms. Storage requirements are made tractable by encoding histograms using a tree structure. Elements of vector X dictate a path through the trees to the leaf nodes where frequencies of observed patterns are recorded. Tree nodes are dynamically generated during training such that only observed patterns occupy memory.

Estimating the quantity of pixels belonging to classes of known Martian terrains could be approached by accumulating a discrete count of classification labels. Such labels could be determined from SVM, RF, Boosting or maximum class conditional probabilities from Bayes Theorem for example. Error bars for such discrete counts could be empirically determined via Monte-Carlo simulations, which need careful tuning to match data samples. We suggest an alternative approach using a continuous probabilistic count of pixels followed by the application of a lower variance bound and error propagation to compute error bars from incoming data. For comparison we also include a discrete count based upon labels giving maximum class conditional probabilities. The discrete count baseline is given by:

$$C(k|R) = \sum_{d \in R} \delta(P(k|X_d, R) > max(P(l|X_d, R), \ldots)) \tag{2}$$

where $C(k|R)$ is the count of class labels k within region R; d is a point within the region; $P(k|X_d, R)$ is the probability observation X_d belongs to class k; and function δ returns 1 if and only if the condition within is true. Due to the accumulation of misclassified observations we believe C, or any similar count, will be biased and have a comparatively large variance. Our proposed probabilistic count can provide an unbiased estimate by compensating for misclassifications:

$$Q(k|R) = \sum_{d \in R} P(k|X_d, R) \tag{3}$$

where $Q(k|R)$ is the estimated quantity of pixels within region R belonging to class k. The prior terms within Bayes Theorem can be substituted with Q and Expectation Maximisation (EM) can be employed to converge upon appropriate estimates.

As the probabilistic count is in a continuous form it becomes possible to compute analytically an estimate of its expected error. We believe the main sources of uncertainty are the values of Q converged upon during EM and the uncertainties within the finite density estimates, $P(X|k)$. A lower bound on Q can be determined using the Cramer Rao Bound and error propagation can be used to track the effects of the Poisson errors found within the histogram based densities. Combined, these gives an approximate error of

$$var(Q(k|R)) \approx \frac{Q(k|R)^2}{\sum_{d \in R} P(k|X_d, R)^2} + Q(k|R) - \sum_{d \in R} P(k|X_d, R)^2 \tag{4}$$

As $\sum_{d \in R} P(k|X_d, R)^2 \leq Q$, the first term, (the CRB of the Likelihood) can be shown to be greater than the last two. These latter terms are for a worst case density estimate (based upon a histogram as here), and would be smaller for (valid) paramteric forms. Systematic errors, due to density variations across regional samples, are expected to be negligible for small samples but become significant and ultimately dominate for large Q. Below we show how this statistical component of the expected error compares to empirically observed data for both discrete, C, and probabilistic, Q, counts.

3 Experiments

Density estimates were constructed for 10 Martian terrains using training images each 900x760 pixels in size. Corresponding 300x760 pixel images were tiled together in 6 random groups of 3 forming datasets A-F shown in Figure 1. The parameter space of the FLID was systematically scanned covering a range of N, r, q and θ values. Both asymmetrical and symmetrical quantisations were tested.

For each configuration of parameters all datasets were analysed. During analysis the datasets were processed using sub-regions covering approximately 4500 pixels. During each iteration discrete and probabilistic counts were

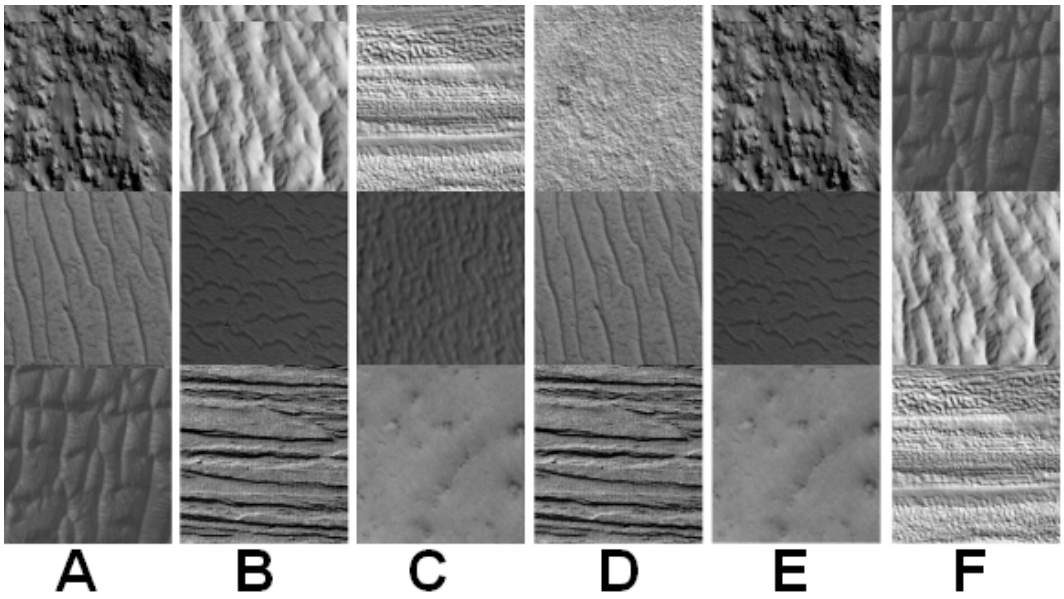

Figure 1: Reduced versions of the 6 datasets, A-F, analysed during testing. There are 10 textures in total randomly combined into triplets.

accumulated for each terrain type and theoretical errors were computed. Residuals between counts and known ground truths were also recorded for each iteration so real error distribution could be plotted and compared to the theory.

A typical additive Gaussian image noise model is assumed, and to make the threshold values, θ, statistically meaningful each of the training and testing images' grey levels were scaled to have unit noise. To keep the experiments simple each dataset was given equal quantities of the 3 assigned terrains, and correspondingly, uniform priors were used for each classification.

4 Results

Summarised below is a cross-section of results from the FLID parameter space showing trends and characteristics which are representative of our full experimental outputs.

The percentage of correctly classified pixels, i.e. accuracy of the discrete count, is indicative of the representation's ability to create densities providing good class separability. By this measure the best performance was achieved using a sampling radius, r, of 16 pixels for each dataset using both symmetrical and asymmetrical encodings. For symmetrical encoding the optimal number of pixel pairs, N, was 10 for each dataset, whereas it ranged from 18 to 24 for asymmetrical encoding. The optimal quantisation, q, was 3 for each dataset for symmetrical encoding and 2 for asymmetrical. The best threshold, θ, was around 5 to 7 for most cases. The total storage required for density estimation in all cases was in the order of 100,000 populated histogram bins. Discrete count performance was typically between 70% and 90% successful in correctly classifying pixels. Probabilistic counts in all cases very closely matched actual ground truths with their predicted errors being within the same order of magnitude as the observed errors.

Successful classification is shown in Figure 2 against total storage required for density estimates (populated histogram bins) for different levels of quantisation. As can be seen best results are achieved when the quantisation is kept to a mimimum.

The impact of changes in threshold, θ, upon successful classification and total storage requirements is shown in Figure 3 for asymmetrical quantisation. Within this plot other FLID parameters were kept at their optimal values. Similar results were observed for symmetrical quantisation. Segmentation success curves for each dataset are shown exhibiting data dependant optimal thresholds. However, in all cases the optimal threshold is closely approached between 5 and 7 standard deviations of image noise (images were preprocessed to give unit noise). Storage requirements drop rapidly as the threshold increases as fewer pixel comparisons pass the hypothesis tests.

The impact of adjusting the number of pixel pairs, N, upon successful classification is shown in Figure 4 for asymmetrical quantisation. Separate plots are shown for each dataset. Overlaid upon each plot are curves for different levels of quantisation, q, between 2 and 9. All other FLID parameters were kept at their optimal values.

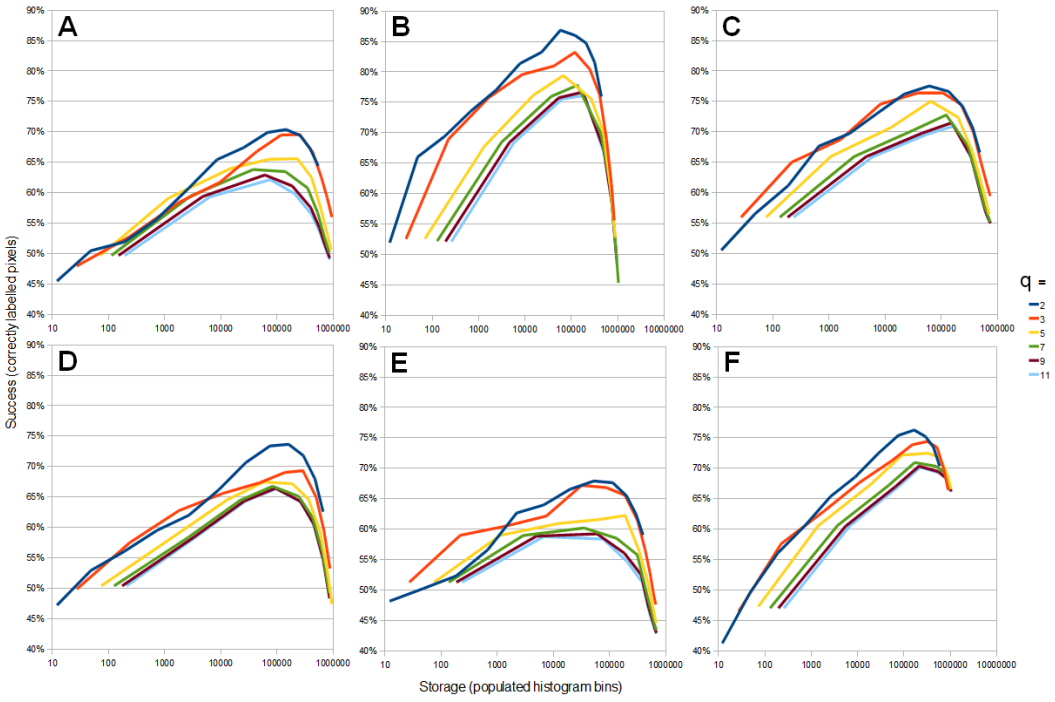

Figure 2: Classification success rates against total storage required for histogram based density estimates. Each plot corresponds to a different dataset. Each curve represents a different level of quantisation, q.

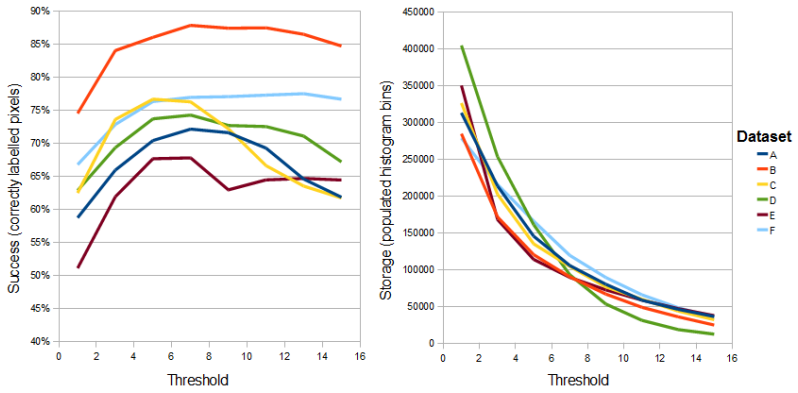

Figure 3: Effects of changes in threshold, θ, upon segmentation success (left) and total storage requirements (right). Results for different datasets are presented by curves A-F.

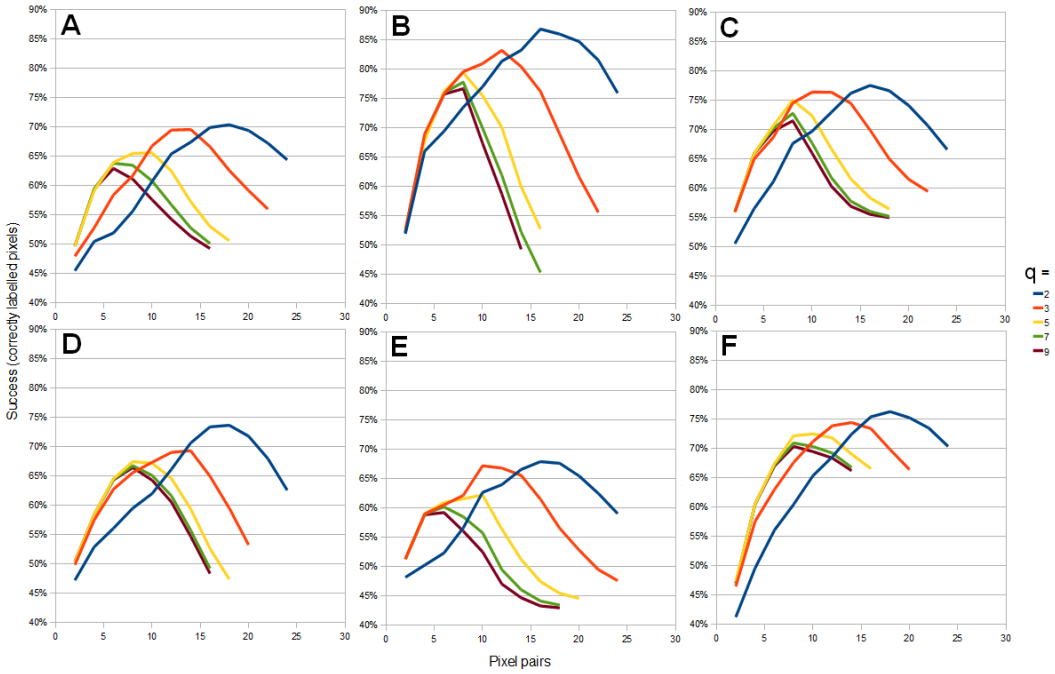

Figure 4: Segmentation performance against numbers of pixel pairs used in descriptor. The different curves indicate performance over quantisations, q, ranging from 2 to 9. Each plot relates to a different dataset.

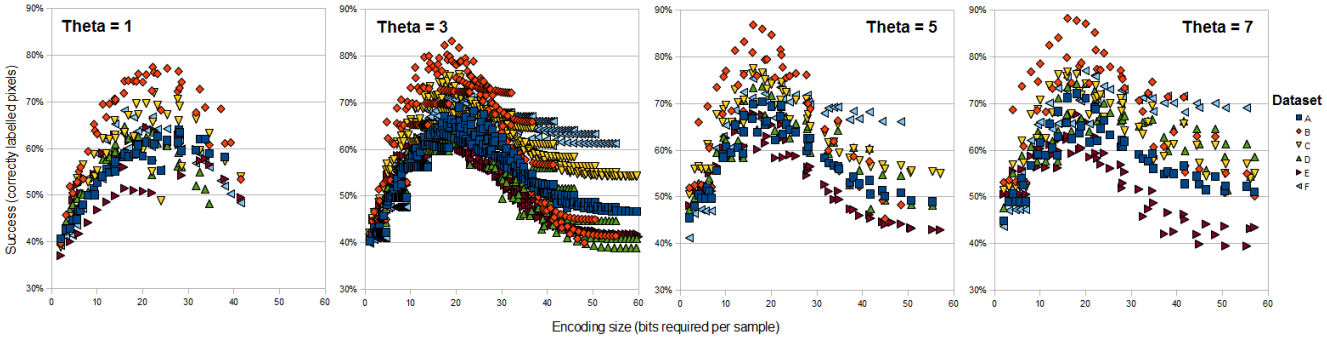

Figure 5: Number of bits required to encode vector X, given by $N \log_2 q$, plotted against segmentation performance for various thresholds. Data points represent performance across a range of quantisations, q, and number of pixel pairs, N. Results for all datasets are overlaid in each plot.

Similar results were observed for symmetrical quantisations with lower values of q peaking with greater numbers of pairs. The representation appears to favour a wider context, give by an increased number of pixel pairs, rather than greater pixel intensity information, given by increased quantisation. For a finite quantity of training data the best absolute peak performance is achieved when the quantisation is set to a minimum with a wider context.

Successful classification plotted against the size of encoding for vector X is shown in Figure 5. The 4 plots show results for thresholds at 1, 3, 5 and 7 respectively with broad curves for each dataset overlaid. Points correspond to a range of quantisations, q, and pixel pairs, N. As encoding size is a function of both q and N points near the peaks correspond to both low N with higher q, and low q with higher N, corroborating results shown in Figure 4.

Actual error distributions determined by multiple iterations and recordings of residuals are shown in Figure 6 for a selection of Martian terrains. The top row of plots shows empirically observed errors on probabilistic counts in blue, corresponding comfortably with theoretical errors which are overlaid in red. The bottom row shows the discrete counterparts exhibiting irregular and biased error distributions.

5 Discussion

The Martian terrains used in our experiments are not exceptional and so could be replaced with arbitrary textures, such as Brodatz [31], without changing our results. The successful classification of any texture depends upon

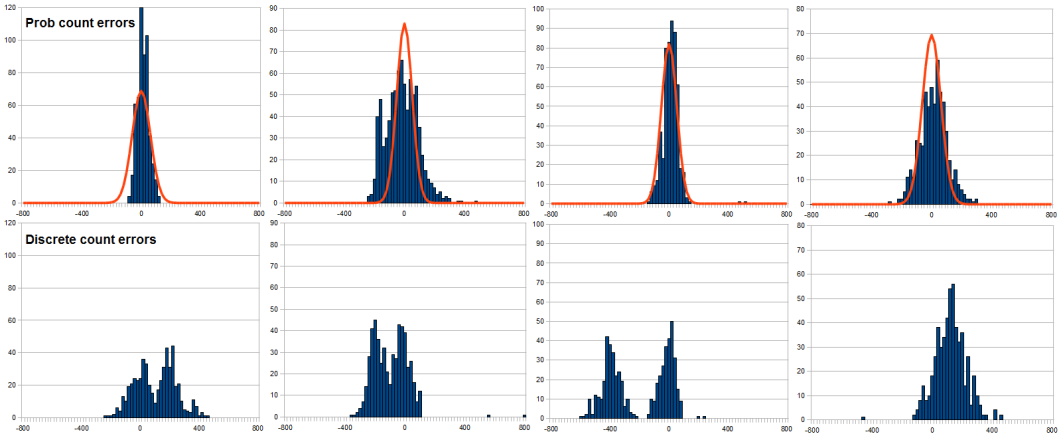

Figure 6: Typical error characteristics of the discrete, C, (bottom) and probabilistic, Q, (top) counting methods for a selection of datasets and classes. Theoretical errors are given by the red curves.

two main factors: having sufficiently accurate density estimations; and having minimal overlaps between density estimates. The first relates to the complexity of representation and quantity of training data available; the second depends upon the distinctiveness of texture classes. The greater the complexity, i.e. larger N and q, the wider the pattern space of vector X. Greater complexity provides more unique and potentially discriminating density estimates. However, the larger the pattern space the larger, yet less well populated, density estimates become. Trends within our results using FLID indicate representational complexity is best increased by widening context via the inclusion of greater numbers of pixels rather than increasing grey-level quantisation. This can be seen in Figure 2 where, for all datasets, performance is inversely proportional to q. Also, the minimum $q = 2$ curve almost always remains above others for any given quantity of storage used in density estimation. Figure 4 shows, for lower q, more pixel pairs can be utilised, again giving highest peak performance at the minimum possible quantisation. Similar results were found using symmetrical encoding where the lowest quantisation possible involved 3 distinctive levels. It should be noted that symmetrical binning with $q = 3$ is equivalent to asymmetrical binning with $q = 2$ if pixel pairs in the latter are selected to complement each other in opposite directions. Given the evidence it is reasonable to suggest a 'this-pixel-is-brighter-than-that-pixel' binary vector is a generally optimum representation for encoding local image structure.

Pixel values can be changed by illumination, surface albedo, transparencies etc. making absolute intensity differences meaningless. This may leave brighter and darker comparisons as the only useful information regarding structure. Structure may thus be viewed as patterns of pixel rank orders. Work presented by Kisilev et al. [32] shows how sets of pairwise comparisons can be used to reconstruct, up to a rank order, the shape of an underlying function. The same method could be applied to a binary representation to reconstruct image patches up to a rank order. This observation provides an explanation for the efficiency of the encoding which utilise a subset of random pixels to model a wider structure of stable relative pixel intensities.

A binary vector corresponds closely to the Binary Robust Independent Elementary Features (BRIEF) representation [30] which has been demonstrated to outperform other encodings including SIFT [33] and SURF [34]. The BRIEF work focused on matching features between images utilising large numbers of pixel pairs within extended regions and the Hamming distance as a similarity measure. Here, probability densities were built using thousands of small local image patches with relatively small numbers of pixel pairs. Despite the differences, the underlying representation of pairwise binary comparisons remains the same. A drawback of BRIEF is the need to smooth images to suppress instabilities due to noise. Our innovation of an adjustable significance threshold, θ, removes this need for smoothing and the inevitable loss of spatial detail it entails.

In comparison to their discrete counterparts the probabilistic counts are unbiased and repeatable across all datasets. If hard decisions are avoided and probabilities of classes are used instead, mistakes in classification can be compensated for resulting in more representative estimates of ground truths. Probabilistic counts are still not exact, but they lend themselves to straight forward computation of error bars. Actual error distributions have been shown to closely match the theory, whereas actual errors on discrete counts are irregular and difficult to predict. Scientific applications demand results are representative of ground truths within stated confidence intervals. These demands are met by our probabilistic technique. Whilst it would be possible to create similar results from standard classifiers it would required Monte-Carlo simulations or large quantities of mark-up to analyse their error characteristics. Our methods remove the need for any such simulations or comparisons with mark-up.

Our method does however rely upon the sample estimates of probabilities encoded in the tree structure being representative of all equivalent textured regions. However, spatial variations are expected and data under analysis

will rarely have exactly the same density functions as training data. The effect of this is a systematic error which grows as the mismatch between training and actual data increases. Our experiments classified pixels in batches of approximately 4500. We have found larger batches produce errors significantly greater than theory would predict because of these systematic effects. As we suspect that real scientific studies will require the ability to deal with larger regions, a focus of future work will be to address this issue.

5.1 Conclusions

A systematic comparison of pixel sampling methods suggests a binary representation similar to BRIEF is an optimal method of encoding local image patch information. This encoding allowing for the creation of compact density estimates based upon multidimensional histograms. We have presented a quantitative method for segmenting Martian terrains using these estimates which could also be applied to general texture segmentation tasks. It must be emphasised that the focus of our work has not been to merely create a new method of texture segmentation; rather it has been to demonstrate the possibility of performing pattern recognition and classification in a scientifically meaningful way. We have demonstrated the ability to conduct a probabilistic count which provides unbiased estimates of the quantities of different classes of feature present within data. The continuous formulation of this count allows straightforward computation of error bars, avoiding the need for complex Monte-Carlos which would be required in a discrete counting scheme. As a baseline for comparison a discrete count, imitating typical outputs of standard classifiers, has shown to be capable of working well in absolute terms but is typically biased and results are highly dataset dependant. Our probabilistic alternative is both reliable and general.

6 References

1. McEwen AS et al. (2007), Mars Reconnaissance Orbiter's High Resolution Imaging Science Experiment (HiRISE), Journal of Geophysical Research, vol 112, E5
2. Hawkins SE et al. (2007), The Mercury dual imaging system on the MESSENGER spacecraft, Space Science Reviews, vol 131, 1-4, 247-338
3. Russell C et al. (2007), Dawn Mission to Vesta and Ceres - Symbiosis Between Terrestrial Observations and Robotic Exploration, Earth Moon and Planets, vol 101, 65-91
4. Blum M D et al. (2000), Fluvial responses to climate and sea-level change, Sedimentology, vol 47, 2-48
5. Head J W et al. (2003), Recent ice ages on Mars, Nature, vol 426, 6968, 797-802
6. Wei Luo et al. (2009), Journal of Geophysical Research, vol 114, E11010
7. Neukum G (2001), Cratering records in the inner solar system in relation to the lunar reference system, Space Science Reviews, vol 96, 55
8. Plesko S C et al. (2006), A Statistical Analysis of Automated Crater Counts in MOC and HRSC Data, 37th Lunar and Planetary Science Conference
9. Svein O K et al. (2007), Automatic and semi-automatic detection of possible meteorite impact structures in the Fennoscandian shield using pattern recognition of spatial data, Proceedings ScanGIS, 227-235
10. Smirnov A (2002), Exploratory Study of Automated Crater Detection Algorithm, Dept. of Computer Science, University of Colorado, Internal Report, www.cs.colorado.edu / rossnd /fcdmf /CraterPaper.pdf
11. Stepinski T (2006), Machine Detection of Martian Craters from Digital Topography, 37th Lunar and Planetary Science Conference, LPI
12. Salamuniccar G et al. (2006), Estimation of False Detections for Evaluation of Crater Detection Algorithms, 37th Lunar and Planetary Science Conference
13. Kim J et al. (2004), Quantitative Assessment of Automated Crater Detection on Mars, 20th ISPRS Congress
14. Magee M et al. (2003), Automated Identification of Martian Craters Using Image Processing, 34th Lunar and Planetary Science Conference
15. Earl J et al. (2005), Automatic Recognition of Crater-like Structures in Terrestrial and Planetary Images, 36th Lunar and Planetary Science Conference
16. Honda R et al. (2000), Data Mining System for Planetary Images: Crater Detection and Categorization, Proceedings of the International Workshop on Machine Learning of Spatial Knowledge, ICML, 103-108
17. Martins R et al. (2008), Crater Detection by a Boosting Approach, Geoscience and Remote Sensing Letters, vol 6 i 1, 127-131
18. Ding w et al. (2011), Sub-Kilometer Crater Discovery with Boosting and Transfer Learning, ACM Transactions on Computational Logic

19. Tar P (2010), The Application of Appearance Models to Martian Impact Craters, The University of Manchester, Internal Report:Tina memo 2010-011, www.tina-vision.net /docs/memos/2010-011.pdf
20. Stepinski T (2004), Extraction of Martian Valley Networks from Digital Topography, Journal of Geophysical Research Vol 109, page 9
21. Stepinski T (2006), Properties of Martian Highlands Drainage From Themis Image and MOLA Topography, 37th Lunar and Planetary Science Conference
22. Vaz DA et al. (2006), Automatic Detection and Classification of Fault Scarps on MOLA data, 37th Lunar and Planetary Science Conference
23. Gaetano R et al. (2010), Graph-based Analysis of Textured Images for Hierarchical Segmentation, BMVC 2010
24. Jia W et al. (2010), Classifying Textile Designs using Region Graphs, BMVC 2010
25. Heess N et al. (2009), Learning Generative Texture Models with Extended Fields-of-Experts, BMVC 2009
26. Santner J et al. (2009), Interactive Texture Segmentation using Random Forests and Total Variation, BMVC 2009
27. Corani G et al. (2010), Robust Texture Recognition Using Credal Classifiers, BMVC 2010
28. Rasmussen C E, Williams K I (2006), Gaussian Processes for Machine Learning, MIT Press, ISBN 026218253X
29. Viola P, Micheal J (2001), Robust Real-time Object Detection, Second International Workshop on Statistical and Computational Theories of Vision
30. Calonder M et al. (2010), BRIEF: Binary Robust Independent Elementary Features, EPFL, ECCV conference paper
31. Brodatz P (1966), Textures: A Photographic Album for Artists and Designers, Dover Publications, ISBN 0486406997
32. Kisilev, Pavel and Freedman (2010), Parameter Tuning by Pairwise Preferences, Proceedings of the British Machine Vision Conference, BMVA Press
33. Lowe D (2004), Distinctive Image Features from Scale-Invariant Keypoints, Computer Vision and Image Understanding 20, 91-110
34. Bay H et al. (2008), SURF: Speeded Up Robust Features, Computer Vision and Image Understanding 10, 346-359