Tina Memo No. 2010-011
Internal.

The Application of Appearance Models to Martian Impact Craters

Paul Tar

Last updated
20 / 6 / 2010

WWW.TINA-VISION.NET

Imaging Science and Biomedical Engineering Division,
Medical School, University of Manchester,
Stopford Building, Oxford Road,
Manchester, M13 9PT.

The Application of Appearance Models to Martian Impact Craters

Paul Tar

Abstract

This document describes preliminary work regarding a PhD project which seeks to construct automated image analysis systems for use in planetary sciences. Simple linear Active Appearance Models of grey level image patches (AAMs) are used here in an attempt to automatically identify craters in satellite images of the Martian surface. We describe a series of experiments which investigate AAM of various orders, and find that they do not provide a suitable solution to this task, due to two features of these approaches. The first is that the similarity function used to optimise model location (based on Likelihood) is incapable of distinguishing between evidence for an object and entirely featureless regions of an image. Secondly, Martian craters seem to conform to a class of appearance behaviour for which high order terms of a linear model describe nothing more than a stochastic (essentially noise) process.

1 Introduction

Appearance modelling is a method of locating and tracking known objects within images. They can be considered a general approach to the representation of local image structure, as such they have been applied as the method of choice for a vast number of image analysis tasks, especially within facial recognition system. An appearance model typically contains a linear combination of components which can be fitted to images using parameters of maximum likelihood. Due to the technique's popularity in locating and tracking objects it seem sensible to investigate the possibility of applying such models to the detection of geological structures. The following pages investigate and evaluate the appropriateness of applying appearance models to the detection of Martian impact craters. The software employed in the investigation is directly borrowed from the Tina Vision libraries. This software was previously demonstrated in the tracking of lips to augment speech recognition [4] and also aorta tracking in medical image sequences. The method is briefly reviewed including details of how it has been applied in the new context. Experiments are described which assess the appearance model's ability to locate individual impact craters. Results are discussed and their outcomes, along with wider theoretical issues, are evaluated against the requirements of a quantitative crater counting system.

2 Method

The appearance model software found within the Tina Vision libraries utilises a Point Distribution Model (PDM) to capture the shape of an object's boundary. It allows for both an inner and outer contour to be defined. Here, only the outer contour is applied to model the annular ridges around impact crater structures as viewed perpendicular to the Martian surface. 12 points are evenly spaced about crater rims.

To obtain a mean crater rim shape and main modes of shape variation a training set is first marked-up by hand using consistent user enforced criteria. The ith shape in the training of N samples is described by a vector v_i:

$$v_i = [x_{i1}, y_{i1}, x_{i2}, y_{i2}, ..., x_{iN}, y_{iN}]^T$$

where each (x_{ij}, y_{ij}) pair are coordinates of the jth point of the ith training sample shape. The shapes are normalised by scaling to a unit width, zero translation and rotation using

$$x_i = M(\frac{1}{s}, -\theta)[v_i - t_c]$$

$$t_c = [t_{x1}, t_{y1}, ..., t_{xN}, t_{yN}]^T$$

where x_i is a normalised shape by the function M which rotates by θ, scales by s and translates by t_c. This normalisation permits a more representative model to be produced from numerous, slightly different samples. Given the set of normalised shapes, the mean shape, $\bar{x}$, and the covariance matrix, S is computed. Principal

Component Analysis [5] is then used to determine the eigenvectors and eigenvalues of S exposing the main modes of shape variation. A model crater rim shape can then be synthesised using

$$x = \bar{x} + P_s b_s$$

where $P_s = [p_1, p_2, ..., p_t]$ is a matrix of the most significant column eigenvectors (i.e. most significant forms of shape model variation) and b_s a vector of weights indicating the quantity of each eigenvector to use in the shape model.

The grey level appearance around a rim is captured by 1 pixel wide profile cross sections, g_{ij}, 10 pixels long, centred upon the 12 points of the crater PDM. Each jth profile is taken perpendicular to the jth point in training sample, i. Rather than calculating individual mean profiles and covariance matrices for each model point, the profiles are concatenated to form a shape-wider profile vector, h_j, for each training image:

$$h_i = [g_{i1}, g_{i2}, ..., g_{iN}]^T$$

PCA again is used to calculate the eigenvectors and eigenvalues of the covariance matrix S_g thereby exposing the main modes of grey level profile variation. Model profiles can then be reconstructed using:

$$h = \bar{h} + P_g b_g \tag{1}$$

Where h is a reconstructed vector of grey levels, $\bar{h}$ is the mean grey level profile, P_g is a matrix of eigenvectors corresponding to the largest modes of grey level profile variation and b_g are the profile model weights. This process is also described in [4].

Reconstructed models are fitted to test images using the following least mean squared error (MSE) cost function, minimised via a simplex algorithm [1] to determine best fitting model parameters, i.e. shape and grey scale modes of variation weights:

$$E_p = (h - \bar{h})^T (h - \bar{h}) - b_g^T b_g \tag{2}$$

Where h here is the image profile and b_g are the weights for the modes.

The major aim is to find a single crater model sufficiently expressive to locate craters in various stages of degradation. The shape model should account for various perturbations in the circular appearance of craters when viewed perpendicular to a surface. The main modes of shape variation may possibly provide information regarding the direction of impacts, as shallow angle impacts produce elliptical craters in the direction of the impacting body's trajectory. The grey level model should account for variation in crater rim illumination, potentially providing information about the level of crater erosion as shadow and illuminated highlights become less pronounces as craters are degraded.

3 Experiments

Three appearance models were created during the experiments. The first model was trained upon 50 well defined, relatively pristine craters which will be referred to as the Distinctive model. The second model was trained upon 50 poorly defined craters in states of relatively advanced degradation, which will be referred to as the Degraded model. Finally, a comprehensive model based upon 100 training examples (50 distinctive, 50 degraded) was created which will be referred to as the Wider model. All sample craters were prepared such that each was 32 pixels in diameter centred within a 64 pixel squared patch of Martian plane with roughly the same orientation. Craters were consistently marked-up by forming a central axis dividing the shadowed and illuminated sides then the 12 points of the PDM were allocated clockwise, starting with the shadowed side first. Each point was placed as close as possible onto the crest of the crater rims. This mark-up approach is illustrated in Figure 1.

Various experiments were conducted to determine the performance of each model when tasked with localising craters within 50 test images. These experiments are outlined below:

- The Distinctive model was tested on 50 pristine craters of similar size and appearance as to those which constituted the model's training set. The number of successfully localised craters were recorded with the search space constrained to include only a specified number of modes of shape and profile variation. The mean only up to the first 6 modes were tested. Each mode was permitted to drift to within 3 standard deviations of the mean.

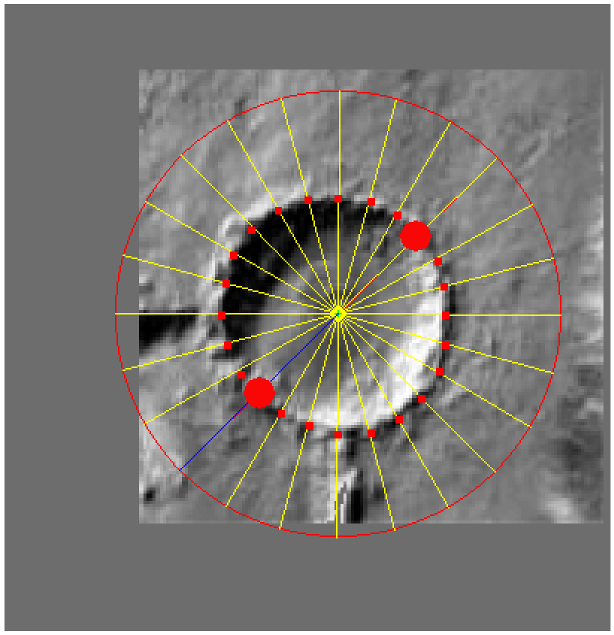


Figure 1: Crater mark-up method. Large red dots indicate ends of a central axis dividing shadow and illuminated sides. Points are defined clockwise about the rim.

- The Degraded model was tested on 50 degraded craters of similar size and appearance as to those which constituted the model's training set. The number of successfully localised craters were recorded with the search space constrained to include only a specified number of modes of shape and profile variation. The mean only up to the first 6 modes were tested. Each mode was permitted to drift to within 3 standard deviations of the mean.
- The Wider model was tested on 50 craters of similar size and appearance as to those which constituted the model's training set. Half of these craters were pristine and half eroded. Again, the number of successfully localised craters were recorded with the search space constrained to include the mean up to the first 6 modes of variation. Each mode was permitted to drift to within 3 standard deviations of the mean.
- The performance of the Distinctive model was tested for a second time upon 50 distinctive craters, this time using 6 modes of both shape and profile variation. Here, the search space was constrained by limiting drift to within different distances from the mean between 0 to 3 standard deviations.
- The performance of the Distinctive model was tested again using the best performing parameters, this time with the mean model shifted by various offsets to assess the sensitivity of the search when craters are closer or further away from the expected location.

The appearance model was considered to have successfully localised a crater if minimising the cost function achieved a match of 9 or more points within plus or minus 2 pixels of the crater rim. This match criteria was consistently applied in all experiments.

4 Results

Results showed the most effective model was the Distinctive model when applied to similarly distinctive craters. When allowed to search within 3 standard deviations of the mean the best result recorded 46 out of 50 correctly localised craters. This result was achieved when at least 3 modes of shape variation were permitted and only 1 mode of grey level profile. Performance degraded as more modes of grey level profile were included in the search. Performance of the Degraded and Wider models were very poor, with best performances of 10 out of 50 and 24 out of 50 respectively. As with the Distinctive model, best performance was achieved when at least 3 or more modes of shape variation were permitted, but performance degraded quickly as more than 1 or 2 modes of grey level profile were included in the search.

Results restricting searches to within various standard deviations of the mean shape and profile yielded similar results, as best performance was achieved when grey profile modes were permitted to drift by only a fraction of σ. Performance reduced quickly when profiles drifted away from the mean grey levels by more than a fraction.

Sensitivity to offset images became significant after craters moved further than 4 pixels away from the mean model (12.5% of crater diameter), reducing to only 15 out of 50 successfully localised craters when moved 16 pixels off centre (50% of diameter).

Full results are presented in Appendix B.

5 Discussion

A human observer can easily identify and locate every sample crater within both the pristine and degraded sets, which raises the question as to what exactly prevents the appearance model from working in all cases. This question is especially pertinent in the case of the degraded craters where successful location rates were as low as zero with some parameters, whilst only at best achieving 20% with optimal parameters. Before attempting to shed light on this question it is worth returning to the original aims of the experiments: to assess the ability to create a unified model for all craters and subsequently utilise the modes of variation to assess impact types and levels of erosion. It is clear from the results that even locating degraded craters is challenging and creating a single appearance model (Wider model) to account for both distinctive and degraded craters only serves to reduce performance overall. The experiments would indicate a simple, single model for all impact craters is not viable using this method. Further, using modes of variation to understand levels of erosion is impractical, as eroded craters can barely be located let alone measured using this method. With this in mind, what exactly is going wrong?

The high localisation rates for pristine craters should not be surprising, as these craters were (subjectively at least) selected for their distinctiveness in terms of completeness, sharpness of shadows and illumination highlights. It should also be no surprise that allowing more degrees of freedom in the shape model (i.e. modes of shape variation) assisted in localising the exact contours of crater rims, as the Simplex algorithm minimising the search cost function could better place PDM markers. However, increasing the number of grey level profile modes consistently reduced performance after more than the first major mode of variation was permitted. The Distinctive, Degraded and Wider models all exhibited this degradation in performance as greater degrees of freedom in grey levels were included. This degradation was also evident when 6 modes of grey level variation were permitted, but only within a fraction of the standard deviation, where more freedom (up to 3σ) only diminished localisation ability. An obvious criticism could be the training sample size. Perhaps with a larger database of sample craters a better grey level profile model could be created. However, there are other more likely explanations, especially in the case of degraded craters, but equally apply to many distinctive cases too. The rest of this discussion covers various possible problems, and where possible suggests improvements which could be made to overcome them.

The mean squared error cost function assumes residuals between grey level models and observations are normally distributed. This is an assumption which has a strong chance of being violated with degraded craters where parts of rims may be missing or occluded by erosive processes. Such craters may exhibit many outliers making the cost function inappropriate from a statistically valid perspective. It may be advantageous to measure the distribution of residuals from the training data and then adjust the cost function accordingly using a flexible non-parametric residual distribution thereby more closely tailoring the cost function to the specific appearance model.

The search attempts to minimise the cost function by adjusting any possible parameter permitted with no penalty for how extreme parameters may be stretched. All searches were limited to at most 3σ from the mean model. It was shown that the further from the mean modes the search was allowed to stray, the worse performance became. The search appears to find lower cost function minima at points in the distributions which are increasingly unlikely, especially in image areas containing less information. Consider a mode of grey level variation which inversely linked a crater shadow with its opposing illumination highlight. When the model moves one direction along this mode the contrast between shadow and illumination increases. Conversely, when moved along the other direction the contrast reduces. If the search is permitted to stray far away from such a mean in a contrast reducing direction, eventually a uniform plane containing little information presents a preferable cost function minima than an actual crater. This possibility may be alleviated by probabilistically weighting the output of the cost function by the probability of the observation. To do this in a quantitatively valid way knowledge of the true distribution of the latent variables (i.e. modes of variation) would be essential.

The use of PCA to extract latent variables controlling the shape and grey level profiles assumes the components can be uncorrelated linearly. Complex interactions between variables may help account for the degradation in performance when greater number of profile modes are permitted in searches. There is a strong possibility the higher modes of variation constitute nothing but noise and inappropriately correlated information lost within them.

A non-linear method of dimensionality reduction perhaps could assist in the extraction of more appropriate latent variables, however problems with such methods have been highlighted [2].

If problems with residual distributions, latent variable distributions and non-linear correlations could be addressed then it might become more plausible to create a unified crater model with quantitatively meaningful modes of variation. At the moment the modes of variation appear to be at best random and at worst plainly inappropriate and detrimental.

6 Application to Crater Counting

The appearance model only takes into account the appearance of the object to which it is trained. As such it is only capable of interpreting a cost function minima as a crater, whether it is a crater or not. Without an appreciation of what other objects could exist and how their cost function values compare against actual craters, there is no way of providing a quantitative confidence level of the correctness of fits. Significant modifications would have to be made to the existing software to introduce this essential functionality, without which corrections could not be made to crater counting estimates.

At present the software finds a single cost function minima in an image meaning it can not actually count craters at all. For crater counting every minima, or indeed every point in an image must be interpretable probabilistically as being a crater or otherwise. The offset experiment showed sensitivity to missing a genuine crater cost function minima increased significantly after moving around 12.5% away from the crater centre. With this knowledge it would be sensible to repeat searches on a lattice of comparable distances, recording each closest minima as a potential crater.

Finally, a working system must be capable of working at multiple wide ranging scales, as craters counting relies upon profiling the frequency of craters of different sizes. Significant additions would have to be made to introduce this additional functionality.

7 Conclusions

The application of appearance modelling has been demonstrated to be successful in locating the majority of well-defined Martian impact craters when only a single crater is presented within an image close to an expected location. However, the limitations of the method have prevented the localisation of the majority of less well-defined craters, especially those which are heavily eroded thereby preventing their use in general cases. It has also been demonstrated creating a unified model for both distinctive and degraded craters only serves to reduce performance in comparison to a model of well-defined craters only. The lack of an associated confidence level also restricts usage in any scientific context, as for any given localisation there is no indication of the probability of the correctness of the model fit.

A major finding is the lack of utility of higher modes of grey level profile variation. Contrary to expectation, results show there is no benefit in searching for model fits outside the largest eigenvector of the grey level profile covariance matrix. In fact, providing more degrees of freedom within the search consistently reduces performance. Increasing freedom within the modes of shape variation does provide some benefit, however the inherently ridged nature of geological structures such as craters makes any such benefits negligible beyond the first mode. Reducing the appearance model to a mean grey level with little shape variation effectively turns the method into an elaborate template matching system. The matching of ridged templates has already been demonstrated elsewhere [3] and shown to be comparable in performance with this more complex method making the additional complications difficult to justify. It should be noted that even for applications where the use of appearance models are common (such as face recognition) a separate system is often required to generate initial hypotheses (ie: a face detection system is required to initialise use of the face recognition system).

It can be argued a larger training set could help overcome some of the issues noted. To illustrate the effect training set size has upon the model Appendix B includes results from an earlier attempt to model craters using 30 samples. The old results used craters similar to those used in the Distinctive model. Despite including more training data, the new model's performance is not a significant improvement upon the original. The prospect of marking-up many hundreds of examples is not a desirably one, especially as ridged template matching can already achieve comparable performance.

In summary, an appearance modelling approach along the lines of the system described does not provide a convincing starting point for the creation of a quantitative crater counting system for the following two key reasons:

1. The maximum likelihood fit of parameters operates irrespective of the presence or absence of signal from the desired target. The method does not differentiate between noise, clutter and actual objects of interest making it impossible to trust any fit unless it is known a priori that the objects of interest are present. This lack of quantification of signal strength precludes the method's use in detection and limits it to the parametrisation of signals already known to exist.

2. It is easy to forget that the application of a linear model pre-supposes correlated behaviours which can be usefully identified in data. We may tend to believe that we can continue to add new degrees of freedom to a model so long as we have enough data to estimate the parameters. However, we find here that the higher modes of variation within linear models, especially grey level models, are dominated by essentially stochastic processes. This is to perhaps be expected, as the intrinsic structure of a crater is quite simple, and the processes which introduce differences between craters are due to random processes which do not have large scale correlations. There is no useful information contained within more sophisticated models, and extra freedom serves only to reduce specificity during image search.

Although both of these issues might be expected to be a general feature of many object classes, these problems seem to us to be under-reported in the general literature.

There are possible lines of investigation which may improve the method, such as taking better account of the statistical distribution of the data expected to be observed and using a non-linear method of extracting the modes of variation. However, for now the technique should not be considered for quantitative crater counting.

8 References

1. Press W et al. Numerical Recipes in C, SE, Cambridge University Press, ISBN 81-85618-16-X, 1992
2. Tar P. Automated Planetary Image Analysis and Associated Literature, Images Science and Biomedical Engineering Division, University of Manchester, Tina Memo No 201000, 2010
3. Magee M et al. Automated Identification of Martian Craters Using Image Processing, South West Institute, San Antonio, Texas, Lunar and Planetary Science Conf. 2003
4. Luettin J, Thacker NA, Speech Reading Using Probabilistic Models, Images Science and Biomedical Engineering Division, University of Manchester, Tina Memo No 1997005, 1997
5. Jolliffe I T, Principle Component Analysis, New York Springer, University of Geneva, ISBN: 0-387-96269-7, 1986

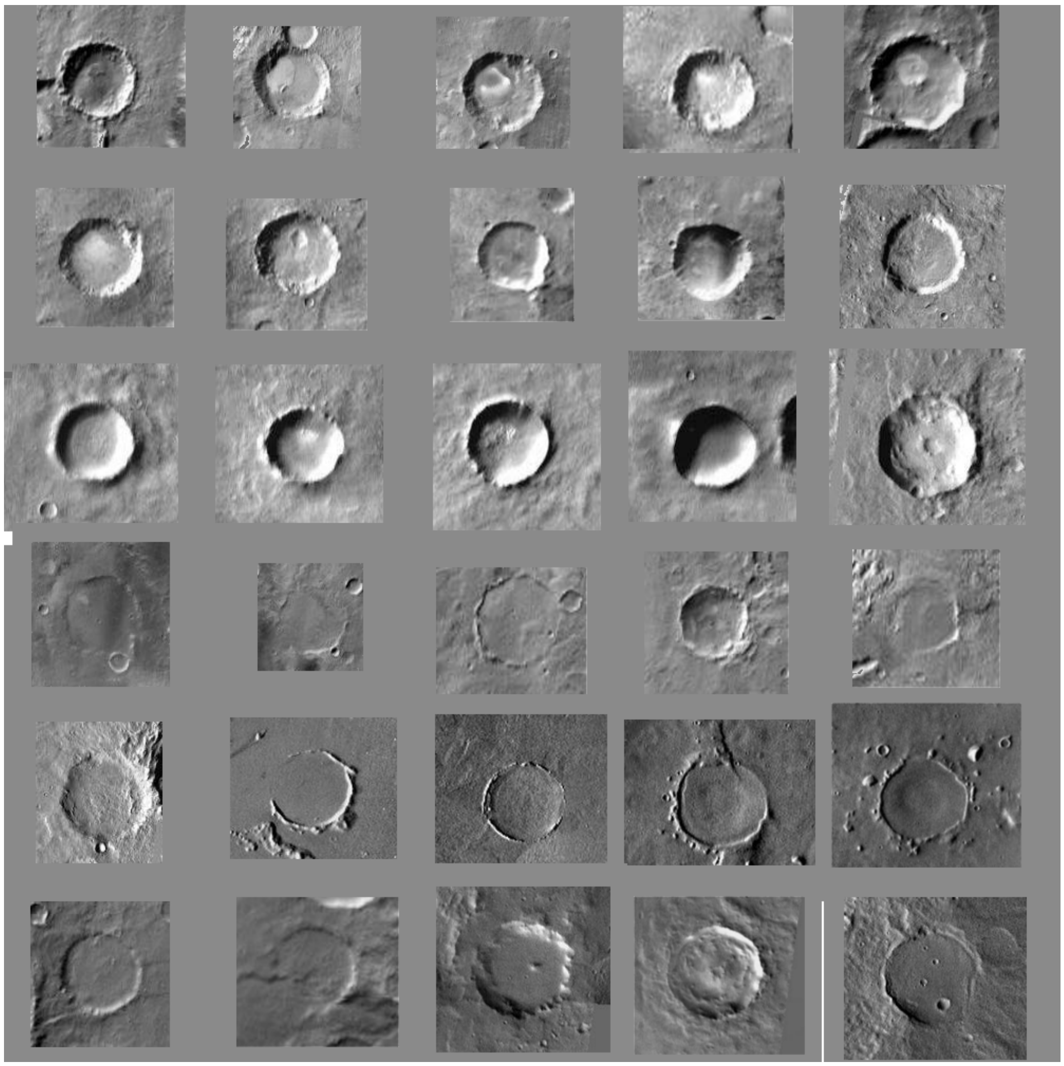
Notes from NAT

The inability of Likelihood to identify relevant structures when used in this context might suggest that a technique which maximises apparent signal would fare better than an ASM. Template matching strategies seem to me to do precicely this. However, any suggestion that the most appropriate image content might be identified by looking for high linear model parameters seems too restrictive. After all, for some applications we may well be interested in locating uniform regions. Alternatively, a Bayesian approach might be considered, which takes appropriate account of how background and signal image patches are distributed in the parameter space defined by the linear model. However, as use of prior distributions during search (MAP) would lead to biased parameter estimates this might be one of those circumstances when the approach suggested in Tina memo 2005-008 might be in order. Unfortunately this would require all candidate local Likelihood optima to be identified before selecting the best using Bayes Theorem, thus losing most of the benefits of forming the problem as an optimisation and making downhill search inappropriate.

On the issue of model complexity, it is unfortunate that conventional approaches choose the number of dimensions simply on the basis of the proportion of explained variance, rather than making the stochastic components in the data an explicit part of the modelling process. This identifies the break between genuine statistics (where the stochastic part of the statistical model is used in a self-consistent manner, such as in factor analysis) and 'statistical' models of appearance (see Tina memo 2010-009).

Appendix A: Model Training Data

Examples of distinctive (top) and degraded (bottom) craters used during training and testing.



Appendix B: All Results

Results from all experiments follow.

	Shape modes						
Grey modes	0	1	2	3	4	5	6
0	40	41	42	43	43	43	43
1	37	46	46	46	46	46	45
2	15	20	26	27	30	25	26
3	22	21	29	26	29	26	28
4	18	19	25	24	24	26	25
5	20	20	29	32	32	32	26
6	17	26	30	32	30	31	28

Number of successfully located craters out of a set of 50 relatively pristine craters when the Distinctive model is applied. The number of modes of shape variation permitted increases from left to right. The number of modes of grey level profile variation increases from top to bottom. This is displayed graphically in Figure 2.

	Shape modes						
Grey modes	0	1	2	3	4	5	6
0	1	8	9	8	8	9	8
1	3	10	7	8	9	8	9
2	1	9	9	9	9	10	10
3	0	6	7	6	6	5	5
4	0	5	4	5	6	6	4
5	0	4	4	4	5	4	4
6	0	4	4	4	5	4	6

Number of successfully located craters out of a set of 50 relatively eroded craters when the Degraded model is applied. The number of modes of shape variation permitted increases from left to right. The number of modes of grey level profile variation increases from top to bottom. This is displayed graphically in Figure 3.

	Shape modes						
Grey modes	0	1	2	3	4	5	6
0	3	7	14	16	14	16	14
1	7	6	19	19	24	22	20
2	1	1	5	6	5	8	8
3	1	1	5	5	4	5	5
4	0	0	9	8	7	8	7
5	0	2	5	4	5	7	5
6	1	3	10	12	11	12	11

Number of successfully located craters out of a set of 50 mixed craters when the Wider model is applied. The number of modes of shape variation permitted increases from left to right. The number of modes of grey level profile variation increases from top to bottom. This is displayed graphically in Figure 4.

	Shape SD		
Grey SD	0.0	1.5	3.0
0.0	36	47	46
0.5	31	45	46
1.0	25	43	47
1.5	29	45	48
2.0	25	42	42
2.5	16	37	37
3.0	17	31	28

Number of successfully located craters out of a set of 50 craters when the Distinctive model is applied with 6 modes of shape and profile variation. The number of standard deviations from the mean permitted increases from left to right, top to bottom. This is displayed graphically in Figure 5.

Offset from Origin:	0	2	4	6	8	10	12	14	16
Successes:	46	46	45	36	26	17	14	19	15

Number of successfully located craters out of a set of 50 relatively pristine craters when the Distinctive set model is applied. The distance craters are offset from their expected location increases from left to right. This is displayed graphically in Figure 6.

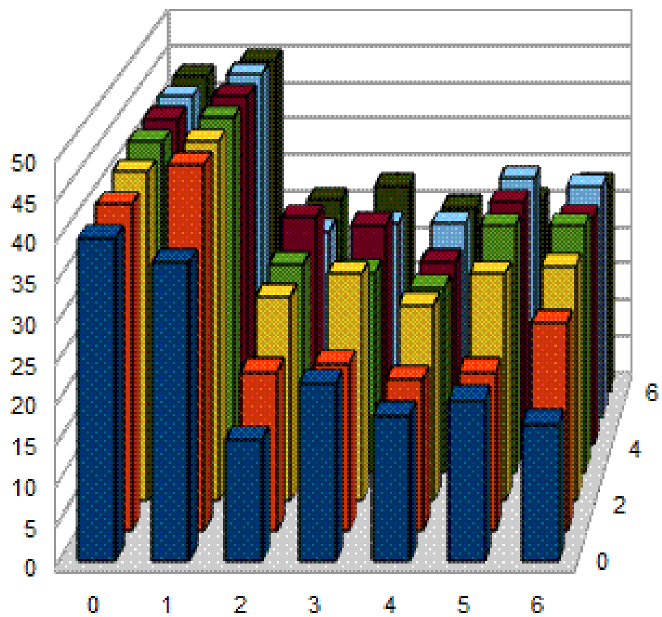


Figure 2: Number of successfully located craters out of a set of 50 relatively pristine craters when the Distinctive model is applied. The number of modes of shape variation permitted increases into the page. The number of modes of grey level profile variation increases left to right.

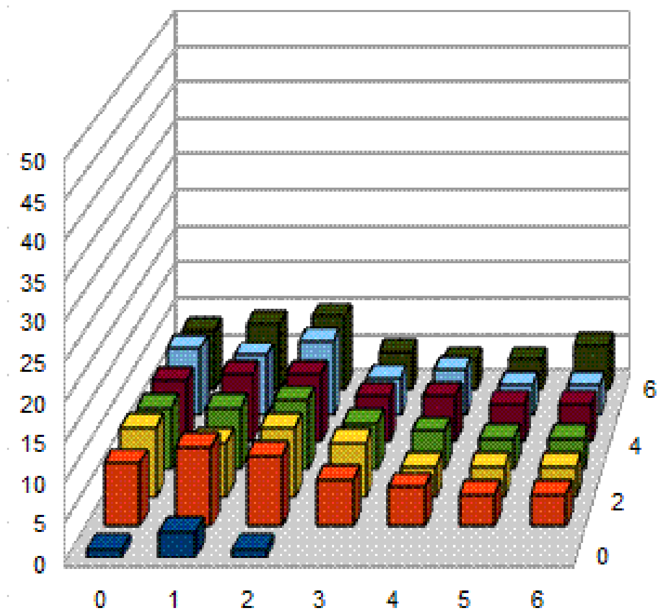


Figure 3: Number of successfully located craters out of a set of 50 relatively eroded craters when the Degraded model is applied. The number of modes of shape variation permitted increases into the page. The number of modes of grey level profile variation increases from left to right.

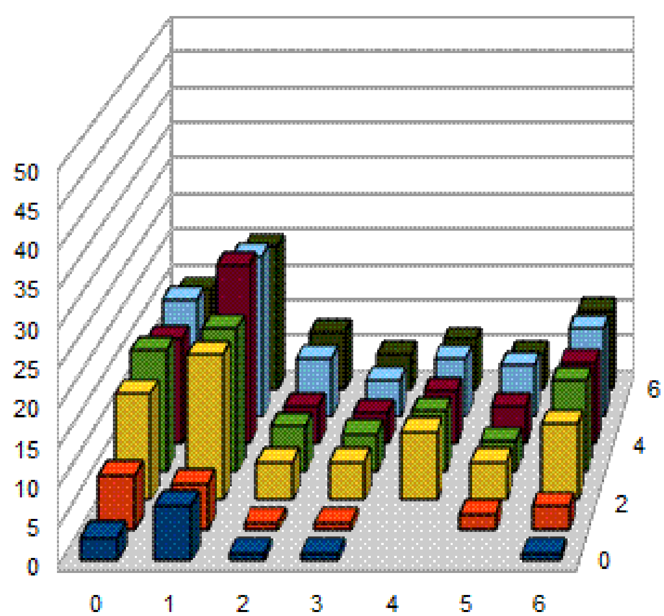


Figure 4: Number of successfully located craters out of a set of 50 mixed craters when the Wider model is applied. The number of modes of shape variation permitted increases into the page. The number of modes of grey level profile variation increases from left to right.

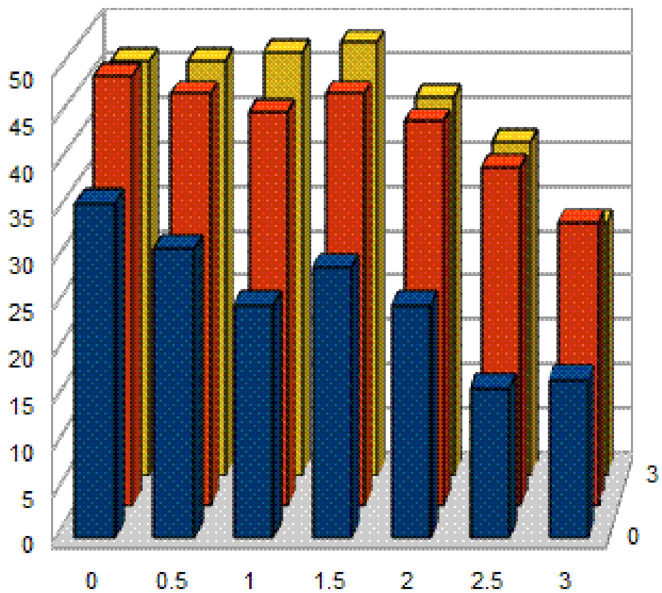


Figure 5: Number of successfully located craters out of a set of 50 craters when the Distinctive model is applied with 6 modes of shape and profile variation. The number of standard deviations from the mean permitted increases from left to right (profile), into the page (shape).

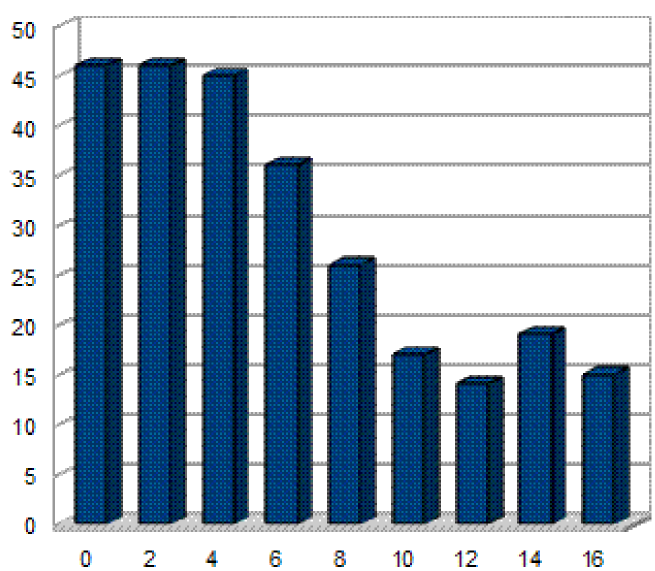


Figure 6: Number of successfully located craters out of a set of 50 relatively pristine craters when the Distinctive set model is applied. The distance craters are offset from their expected location increases from left to right.

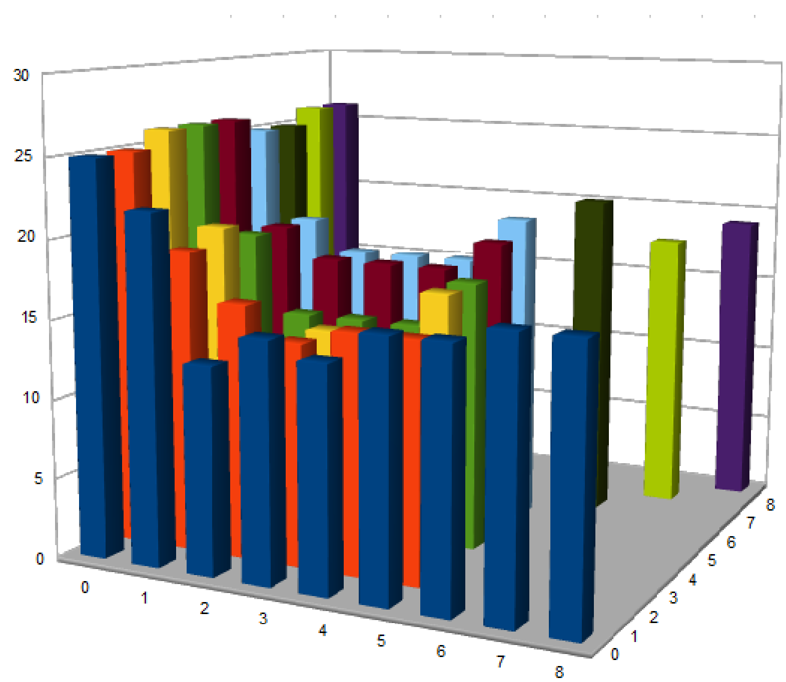


Figure 7: Results from a previous crater model constructed using 30 training samples and 30 test images. Modes of shape variation increase into the page. Modes of grey level profile variation increase left to right. Note the similarity between this chart and the one found in figure 2