

# Cratermatic Finder Routine Description

Michael Mendenhall

August 2006

## 1 Notation

### 1.1 The Topographic Landscape

The dataset that *Cratermatic* deals with is a rectangular  $n \times m$  array of elevation values, which we will denote by  $[\mathbf{Z}]$ , the  $(i, j)$  element of which is a real number denoted by  $Z_{i,j}$ . It is often conceptually easier to think of the topography as a continuous, differentiable function  $z : S \rightarrow \mathbb{R}$  over some domain  $S \subset \mathbb{R}^2$ .

### 1.2 Derivatives

For actual computations on the discrete dataset, *Cratermatic* replaces the continuous derivative  $\frac{\partial z}{\partial x}$  by the discrete derivative  $Dx([\mathbf{Z}])$ ,

$$\begin{bmatrix} -\frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix} * [\mathbf{Z}] : Z_{i,j} \mapsto -\frac{1}{2}Z_{i-1,j} + \frac{1}{2}Z_{i+1,j}$$

(where  $*$  denotes discrete convolution). Similarly,  $\frac{\partial z}{\partial y}$  is replaced by  $Dy([\mathbf{Z}])$ ,

$$\begin{bmatrix} -\frac{1}{2} \\ 0 \\ \frac{1}{2} \end{bmatrix} * [\mathbf{Z}] : Z_{i,j} \mapsto -\frac{1}{2}Z_{i,j-1} + \frac{1}{2}Z_{i,j+1}$$

For an analytic continuous function of two variables, these two partial derivatives are sufficient to specify the slope in any direction at any point. For the discrete dataset, however, it is useful to define

additional derivatives in other directions, specifically the diagonal derivatives  $Dxy([\mathbf{Z}])$

$$\begin{bmatrix} -\frac{1}{2\sqrt{2}} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2\sqrt{2}} \end{bmatrix} * [\mathbf{Z}] : Z_{i,j} \mapsto -\frac{1}{2\sqrt{2}}Z_{i-1,j-1} + \frac{1}{2\sqrt{2}}Z_{i+1,j+1}$$

and  $Dxy'([\mathbf{Z}])$ ,

$$\begin{bmatrix} 0 & 0 & \frac{1}{2\sqrt{2}} \\ 0 & 0 & 0 \\ -\frac{1}{2\sqrt{2}} & 0 & 0 \end{bmatrix} * [\mathbf{Z}] : Z_{i,j} \mapsto -\frac{1}{2\sqrt{2}}Z_{i-1,j+1} + \frac{1}{2\sqrt{2}}Z_{i+1,j-1}$$

### 1.3 The C-Transform

The craterfinding transform (C-Transform) is defined for continuous topography as

$$C \circ z : \vec{\mathbf{x}} \mapsto \iint_S e^{-\frac{|\vec{\mathbf{x}}-\vec{\mathbf{x}}'|^2}{2r^2}} \vec{\nabla}_z(\vec{\mathbf{x}}') \cdot \frac{\vec{\mathbf{x}}' - \vec{\mathbf{x}}}{|\vec{\mathbf{x}}' - \vec{\mathbf{x}}|} d\vec{\mathbf{x}}'$$

To remove interference from unwanted areas of the terrain (for example, craters that have already been found), the region of integration may exclude those areas (this is implemented by setting the gradient to 0 in the excluded areas).

### 1.4 The Watershed Transform

For an  $m \times n$  topography dataset  $[\mathbf{Z}]$ , let the watershed transform of  $[\mathbf{Z}]$  be denoted by  $\mathcal{W}$ , a set of disjoint subsets of the  $m \times n$  lattice whose union covers the  $m \times n$  lattice.

## 2 The "findcraters" routine

Starting the procedure with target radius  $r = 5$ ,

1. Create a blurred copy of the topography  $[\mathbf{G}]$  by applying a Gaussian blur of radius  $r/5$  to  $[\mathbf{Z}]$
2. Calculate the watershed transform  $\mathcal{W}$  of  $[\mathbf{G}]$
3. Calculate the gradient  $[\nabla\mathbf{G}]$  of  $[\mathbf{G}]$
4. Calculate the radius  $r$  C-Transform  $[\mathbf{C}]$  of  $[\mathbf{Z}]$
5. Find crater candidate regions in  $[\mathbf{C}]$ :

- (a) Calculate the discrete second derivatives of  $[C]$  in 4 directions:  $D_x(D_x([C]))$ ,  $D_y(D_y([C]))$ ,  $D_{xy}(D_{xy}([C]))$ , and  $D_{xy'}(D_{xy'}([C]))$
- (b) Create a binary image  $[B]$  where  $B_{i,j} = 1$  if all four second derivatives are positive at  $(i, j)$ , and  $B_{i,j} = 0$  otherwise
- (c) Find the path-connected regions of 1's in  $[B]$ , each of which is an upward-convex region of  $[C]$  which may be the interior of a crater. Denote this set of regions by  $\mathcal{B}$ .

6. Restrict each element  $X \in \mathcal{B}$  to those segments in  $\mathcal{W}$  whose minimum point is contained in  $X$ :

- (a) For each  $W \in \mathcal{W}$ , find some minimum point  $\vec{p}_W \in W$  such that  $[G](\vec{p}_W) \leq [G](\vec{q}) \forall \vec{q} \in W$
- (b) For each  $W \in \mathcal{W}$ , if  $\vec{p}_W \notin X$ , set  $X = X \setminus W$

As a result of this operation, no  $W \in \mathcal{W}$  can contain points from more than one  $X \in \mathcal{B}$ . The future expansion of the regions  $X \in \mathcal{B}$  will be limited to the union of regions  $W \in \mathcal{W}$  which intersect  $X$ : for each  $X \in \mathcal{B}$ , let  $X'$  denote  $\bigcup\{W \in \mathcal{W} : W \cap X \neq \emptyset\}$ .

7. Remove each element of  $\mathcal{B}$  for which 70% or more of the points have been identified as a crater in a previous step (this prevents double-counting of craters)

8. Expand each region  $X \in \mathcal{B}$  to cover (we hope) the whole crater. This is done using a "cellular automata"-like approach, in which a set of "active points" expands to a new set of "active points" to cover more of the crater. Each "active point"  $a$  is associated with a set of characteristics: a position  $a_p$ , a starting point  $a_{p_0}$ , an owner  $a_X \in \mathcal{B}$ , a highest gradient value  $a_g \in \mathbb{R}^+$ , a distance  $a_d \in \mathbb{R}^+$ , and a "strength"  $a_s \in \mathbb{R}^+$  which is used to determine when the point has travelled too far over too low slopes. For each region  $X \in \mathcal{B}$ , we also track the highest gradient yet seen in the region,  $X_g \in \mathbb{R}^+$ , which starts out with the value 0.

(a) Create the set  $A$  of initial active points

- i. The initial active points are each  $\vec{p}_W$  contained in some  $X \in \mathcal{B}$ .
- ii. The owner of each active point is the  $X \in \mathcal{B}$  which contains it
- iii. The points start with  $a_g = 0$ ,  $a_d = 0$ ,  $a_s = 1$

(b) Create a new set of active points  $A'$  by expanding out from the current set of active points  $A$ . For each  $a \in A$  with owner  $X \in \mathcal{B}$  for which  $a_s \geq 10^{-r/5}$ , we consider the possibility of creating a new point  $a' \in A'$  for each of the 8 pixels surrounding  $a$  which are also in  $X'$ . Let  $\vec{d}\mathbf{v} \in (\{-1, 0, 1\} \times \{-1, 0, 1\}) \setminus \{(0, 0)\}$  denote the direction of movement from  $a$  to  $a'$ , and let  $dz = ([G](a'_p) - [G](a_p)) / |\vec{d}\mathbf{v}|$  denote the slope of the move. Expansion to  $a'$  is subject to the following restrictions:

- i. If  $a'$  has been expanded onto in a previous step, do not expand (no need to repeatedly visit the same area)
- ii. If  $a_p \in a_X$  (i.e. we are still inside the owner region  $X$ ), skip the rest of these restrictions
- iii. If  $dz < 0$ , do not expand (only travel uphill)
- iv. If  $\frac{(a'_p - a_{p0}) \cdot [\nabla \mathbf{G}](a'_p)}{|a'_p - a_{p0}| \cdot |[\nabla \mathbf{G}](a'_p)|} > -0.7$ , do not expand (only expand radially outward from starting point)

If several  $a \in A$  could potentially expand onto  $a'$ , the one with the minimum  $a_d + |dv|$  is given priority. On expansion, the newly created  $a'$  inherits the characteristics of  $a$  with the following modifications:

- $a'_d = a_d + |\vec{d}\mathbf{v}|$
- $a'_g = \max(a_g, dz)$ . We also update  $X_g = \max(X_g, dz)$  for owner region  $a_X$ .
- If  $a'_p \notin a_X$  but  $a_p \in a_X$  (i.e. we have just moved outside the owner region), set  $a'_d = 1000$  and  $a'_{p0} = a_p$ . Thus, for points outside the original region, distance and direction of travel is measured relative to the boundary of the region.
- When  $a'_p \notin a_X$  (outside of the original owner region), the strength  $a'_s$  may be reduced by low slope  $dz$  relative to previously recorded high values:
  - if  $dz < 0.6 \cdot a'_g$  then  $a'_s = a'_s \cdot dz / a'_g$
  - For owner region  $a'_X$  of  $a'$ , if  $dz < 0.2 \cdot X_g$  then  $a'_s = a'_s \cdot 3dz / X_g$

(c) So long as  $A' \neq \emptyset$ , repeat the preceding step with  $A = A'$ .

9. Fill in any holes in each region  $X$  so that each region is simply connected
10. Eliminate any regions which cover an area of less than  $\frac{\pi r^2}{4}$  or more than  $8\pi(r+5)^2$
11. Check the characteristics of each region  $X$  to see whether it is indeed a crater. The crater identification process is based on 7 somewhat arbitrary tuning parameters  $k_1, \dots, k_7$  (described below), for which the optimal values are a matter for future research.

- find the center-of-mass of the region  $\vec{\mathbf{c}} = \frac{\iint_X \vec{\mathbf{x}} d\vec{\mathbf{x}}}{\iint_X d\vec{\mathbf{x}}}$
- Express the shape of  $X$  in terms of the radius around  $\vec{\mathbf{c}}$  as a fourier series:

$$r(\theta) = r_0 \cdot \left( 1 + \sum_{n=1}^{\infty} a_n \sin n\theta + b_n \cos n\theta \right)$$

$$\pi r_0^2 \equiv \iint_X d\vec{\mathbf{x}}; \quad a_n \equiv \frac{\iint_X \sin n\theta d\vec{\mathbf{x}}}{\iint_X d\vec{\mathbf{x}}}; \quad b_n \equiv \frac{\iint_X \cos n\theta d\vec{\mathbf{x}}}{\iint_X d\vec{\mathbf{x}}}$$

- Let  $r_i(\theta)$  denote the  $i^{\text{th}}$  partial sum of the series,  $r_i(\theta) = r_0 \cdot (1 + \sum_{n=1}^i a_n \sin n\theta + b_n \cos n\theta)$ .
- Define  $m_i \equiv \sqrt{a_i^2 + b_i^2} \forall i > 0$ , the magnitude of the contribution from the  $i^{\text{th}}$  component. For an ideal (perfectly round) crater,  $m_i = 0 \forall i > 0$ .
- Eliminate as non-craters any regions for which  $m_2 > k_1 = 0.25$  (region is too narrow; ideally 0)
- Eliminate as non-craters any regions for which  $m_3 > k_2 = 0.10$  (region is too lumpy; ideally 0)
- To assess how well the first  $n$  terms of the series reflect the shape of the region, we can measure how far inside or outside of the  $n$ -term sum boundary the actual region lies. With  $X$  denoting, as usual, the actual crater region, let  $X_n$  denote the region defined by the first  $n$  terms of the radial fourier series. The "error region" is thus  $E = (X \setminus X_n) \cup (X_n \setminus X)$ . Define

$$d_n \equiv \frac{1}{\pi r_0^2} \sqrt{\iint_E |\vec{\mathbf{x}} - r_n(\theta) \hat{\mathbf{x}}|^2 d\vec{\mathbf{x}}}$$

which provides a (scale invariant) measure of deviation from the  $n$ -term sum (with perfect agreement indicated by  $d_n = 0$ ).

- Eliminate as non-craters any regions for which  $d_2 > k_3 = 0.06$  (ideally 0)
- Eliminate as non-craters any regions for which  $d_3 > k_4 = 0.02$  (ideally 0)
- Express the the average gradient  $\vec{\mathbf{g}}$  as a function of angle  $\theta$  around the center point  $\vec{\mathbf{c}}$  in terms of a Fourier series

$$\vec{\mathbf{g}}(\theta) = \sum_{n=0}^{\infty} (a_n \sin n\theta + b_n \cos n\theta) \hat{\mathbf{x}} + \sum_{n=0}^{\infty} (c_n \sin n\theta + d_n \cos n\theta) \hat{\mathbf{y}}$$

$$a_n \equiv \frac{\iint_X \frac{\partial}{\partial x} z(\vec{\mathbf{x}}) \cdot \sin n\theta d\vec{\mathbf{x}}}{\iint_X d\vec{\mathbf{x}}}, \text{ etc.}$$

For an ideal crater,  $b_1 \approx c_1$  are the only nonzero components.  $b_0$  and  $d_0$  indicate the level of average background slope, and can mostly be ignored.

- Eliminate as non-craters any regions for which  $\min(\frac{|b_1|}{|c_1|}, \frac{|c_1|}{|b_1|}) < k_5 = 0.5$  (balance between  $x$  and  $y$  direction slopes, ideally 1)
- define  $s_1 = \sqrt{b_1^2 + c_1^2}$ ,  $s'_1 = \sqrt{a_1^2 + d_1^2}$ ,  $s_2 = \sqrt{a_2^2 + b_2^2 + c_2^2 + d_2^2}$
- Eliminate as non-craters any regions for which  $\frac{s'_1}{s_1} > k_6 = 0.33$  (ideally 0)
- Eliminate as non-craters any regions for which  $\frac{s_2}{s_1} > k_7 = 0.33$  (ideally 0)

12. If  $r < 20$ , restart the craterfinding process with new target radius  $2r$ . Otherwise, so long as both the image width and height of  $[Z]$  are  $> 40$  pixels, scale the image  $[Z]$  down by a factor of 2 in each direction to produce reduced image  $[Z']$ , and apply the same craterfinding routine to  $[Z']$  with target radius  $r = 20$ .