measures the change of slope angle in the direction of tangent to contour and is calculated analytically from a fourth-order polynomial fitted to a patch of surface composed of a $3 \times 3$ neighborhood centered on a focus pixel. Curvature is used for distinguishing between convex and concave crater walls and ridges. $\kappa > 0$ corresponds to convex topography, whereas $\kappa < 0$ corresponds to concave topography.

3) **Flood** — Flood, $f(x, y)$, is a binary variable such that pixels located inside topographic basins have $f(x, y) = 1$, and all other pixels have $f(x, y) = 0$. The flood variable is particularly useful to distinguish between two flat landforms, inter crater plateaus and crater floors, because crater floors are usually enclosed basins and plateaus are not.

A pixel-based feature vector, $v(x, y) = \{s, \kappa, f\}(x, y)$, describes the topography of the landscape at the level of an individual pixel. The raster of vectors $v$ offers a mathematical representation of a site's landscape. Pixel-based feature vectors are used in the segmentation stage of our proposed method.

## III. Segmentation

Segmentation is an intermediate data-preparation stage in our framework. The segmentation procedure subdivides the landscape into mutually exclusive and exhaustive segments containing pixels having approximately uniform feature vectors. These segments constitute topographic objects, which are classified and merged into coherent physical landforms by the supervised learning algorithms applied in the later stages of our framework. Raster segmentation has been the subject of intense study in the domain of image analysis. In the context of image analysis, a variety of techniques have been proposed [1], [4], [11], [14], [29], [37], [46], all of which could, in principle, be extended to landscape segmentation. However, requirements for an effective segmentation for the purpose of classification are different from those encountered in the field of computer vision. In particular, for the purpose of classification, it is desirable to have relatively small, approximately equal-sized segments (a requirement at odds with typical problems in image analysis). Having small segments eliminates the danger of a particularly large segment being misclassified, which would avoid producing a grossly incorrect map. Additionally, a misclassification of a small segment carries a small impact in the overall map accuracy. Moreover, having approximately equal-sized segments assures that statistics of terrain attributes are calculated from comparable ensembles of member pixels.

We investigate two different segmentation algorithms. The dividing algorithm splits the landscape on the basis of abrupt discontinuities in pixel-based feature vectors. The agglomerative algorithm initially treats each pixel as an individual segment. These initial segments are combined into larger segments as long as a user-defined criterion for the uniformity of constituent pixel-based feature vectors holds. Both algorithms use the same pixel-based feature vectors $v(x, y)$.

### A. *Watershed-Based Segmentation*

Our first approach to segmenting a landscape incorporates the watershed transform [5] applied to a gray-scale image that encapsulates gradients of pixel-based feature vectors. This image is calculated using a (computationally simple) homogeneity measure $H$ [18]. The homogeneity measure $H$ is calculated using a moving square window of width $2L$ + 1 (where L is user-defined) applied over the raster of feature vectors $v(x, y)$. Consider a focal pixel $(x_c, y_c)$ having a single feature (a component of a feature vector), for example, slope $s(x_c, y_c)$. For every pixel in the window we calculate a separation vector $\mathbf{d}_i = (x_i - x_c, y_i - y_c)$. From the separation vector we construct a gradient vector,

$$\mathbf{g}_i = (s(x_i, y_i) - s(x_c, y_c)) \times \frac{\mathbf{d}_i}{\|\mathbf{d}_i\|} \tag{1}$$

and we use gradient vectors calculated for all pixels in the window to calculate the homogeneity measure H,

$$H = \| \sum_{i=1}^{(2L+1)^2} \mathbf{g}_i \| \tag{2}$$

A pixel located in a region that is homogeneous with respect to $s$ has a small value of $H$. On the other hand, a pixel located near a boundary between two regions characterized by different values of $s$ has a large value of $H$. A raster constructed by calculating the values of $H$ for all pixels in the landscape can be interpreted as a gray-scale image and is referred to as the $H$-image. We denote the $H$-image by $H$. White areas on $H$ represent boundaries of homogeneous regions, whereas the dark areas represent the actual regions. The extension of the $H$-image concept to encapsulation of gradients in a field of multi-dimensional feature vectors is straightforward. In order to calculate $H$ for our three dimensional feature vectors $v(x, y)$, we calculate the three individual $H$ values separately for each pixel and combine them to obtain the overall value of $H$ at that pixel:

$$H = \sqrt{w_s H_s^2 + w_\kappa H_\kappa^2 + w_f H_f^2} \tag{3}$$

where $w_s$, $w_\kappa$, and $w_f$ are weights introduced to offset different numerical ranges of the attributes. An example of the $H$-image is given in Fig. 2A. The watershed transform of $H$ results in an over segmentation of the $H$-image (and thus the landscape), an undesirable feature in computer vision, but a desirable feature in the context of segmentation-based classification.

### B. *$K$-Means Based Segmentation*

Our second approach to segmenting a landscape is agglomerative and incorporates a contiguity-enhanced variant of the standard $K$-means clustering algorithm [43], which uses – in addition to terrain attributes – spatial coordinates of pixels as features. In this approach the pixel-based feature vector is given as $v = (s, \kappa, f, x, y)$. The additional spatial features ($x$ and $y$) control the size of the segments while providing the resultant segments with very desirable geometric properties. For example, in areas where terrain features are approximately uniform, the local gradient of $v$ is dominated by changes in $x$ and $y$ leading to the formation of round-shaped segments. On the other hand, in areas where change