ASTRONOMY & ASTROPHYSICS
SUPPLEMENT SERIES

Astron. Astrophys. Suppl. Ser. **127,** 587-595 (1998)

Discrimination of point-like objects in astronomical images using surface curvature

A. Llebaria, P. Lamy, and P. Malburet

Laboratoire d'Astronomie Spatiale du CNRS, BP. 8, F-13376 Marseille Cedex 12, France

Received April 14; accepted June 4, 1997

Abstract. We present a method for the discrimination of point-like objects in astronomical images using surface curvature. The principal curvatures k_1 and k_2 are calculated using a local quadric approximation to the log of the intensity of the original image. The analysis of the (k_1, k_2) diagram allows to easily separate various shapes and to establish a criterion for discrimination. A mask is then generated and the invalid pixels are replaced by a local, pyramidal interpolation using a non-linear, multiresolution method. We present two applications, the discrimination and removal of stars from the dust tail of comet P/Halley observed while it was crossing the Milky Way, and the removal of cosmic ray impacts from images of the solar corona obtained with a space-coronagraph.

Key words: techniques: image processing — comets: general; P/Halley — sun: corona

1. Introduction

The accurate detection and subsequent removal of point-like objects in astronomical images is a frequent problem. Let us mention, for example, the case of stars pervading an extended object or cosmic ray impacts on CCD images. The problem is often complicated by various distortions affecting the point-like objects, e.g. saturations, variable Point Spread Function in the field-of-view, irregular images, which preclude the use of fitting methods. Indeed the classical method of correlation with or without filtering (implemented for instance in the "inventory" package of MIDAS) does not work well on saturated or irregular images. We propose a new approach to the problem based on the properties of surface curvature which is related to peaks, pits and flats in images and therefore contain the key information to discriminate point-like objects.

The intensity values of an image can be viewed as a two-dimensional surface in a three-dimensional space with the two spatial coordinates forming two of the three dimensions and the intensity axis forming the third dimension. The properties of two-dimensional surfaces embedded in a three-dimensional space are classically studied in differential geometry (Do Carmo 1976; Doneddu 1978; Faux & Pratt 1987). Surface curvature describes how much this surface is bending in a given direction at a particular point; it involves the first and second partial derivatives with respect to the spatial coordinates and thus involves the spatial neighborhood of the point. Surface curvature as an intrinsic tool to analyze images is an old topic in the field of image processing. It has been introduced to detect and characterize peaks, pits, flats (Paton 1974; Peuker & Douglas 1975; Toriwaki & Fukumura 1978; Halarick et al. 1983; Laffey et al. 1985; Nackman 1984), edges (Halarick 1984), and corners (Denger & Billie 1984). Texture analysis and computer vision use it extensively. Curiously, astronomical data processing has ignored this method (as far as we know). It is due, perhaps, to odd side effects induced into the classification process by the frequent very large photometric range of astronomical images. We shall show how to alleviate this difficulty by taking the logarithm of the intensity.

In the sections below, we first recall the definition of the minimum and maximum local curvatures k_1 and k_2 in a continuous image and then we present the algorithm to calculate these curvatures in a digital image. We then extensively discuss the (k_1, k_2) diagram in terms of the morphology of objects. Next we explain how to define the neighborhood of point-like objects to be removed and how to restore the background. The method is illustrated on:

i) a wide-field CCD image of the tail of comet P/Halley taken in 1986 when the comet was in the southern Milky Way,
ii) a coronographic image of the solar corona obtained with the SOHO/LASCO space experiment.

Send offprint requests to: A. Llebaria

2. Method

2.1. The logarithmic transformation

Practical experiments with the surface curvature of astronomical images have soon revealed to us the superiority of dealing with the logarithm of the intensity, $\log(B)$, rather than with the intensity itself. The justifications of this transformation are easily understood:

i) The range of curvatures will be narrower on the $\log(B)$ images than on the intensity images insuring a better detection of faint objects;
ii) Point-like objects such as stars may be viewed, at least in a first approximation, as having a Gaussian shape. The logarithmic transformation will generate similar paraboloids (i.e. having the same parameters) for which the maximum of curvature takes place at their tops. Practically, the logarithmic transformation of the star images will exhibit a coarse paraboloidal shape which will be easily localized from the maximum of curvature.

2.2. Surface curvature

Let x and y be the two spatial coordinates and $z = \log(B)$, the log transform of the intensity value. We now closely follow the formalism introduced by Peet & Sahota (1985) and specify any point on the surface by its position vector:

$$\mathbf{R}(x, y) = x\boldsymbol{i} + y\boldsymbol{j} + z(x, y)\boldsymbol{k}.$$

The first fundamental form of the surface is the expression for the element of arc length of curves on the surface which pass through the point under consideration.
It is given by:

$$I = \mathrm{d}s^2 = \mathrm{d}\mathbf{R}\,.\,\mathrm{d}\mathbf{R} = E\mathrm{d}x^2 + 2F\mathrm{d}x\mathrm{d}y + G\mathrm{d}y^2$$

where

$$E = 1 + \left(\frac{\partial z}{\partial x}\right)^2 \quad F = \frac{\partial z}{\partial x}\frac{\partial z}{\partial y} \quad G = 1 + \left(\frac{\partial z}{\partial y}\right)^2.$$

The second fundamental form arises from the curvature of these curves at the point of interest and in the given direction:

$$II = e\mathrm{d}x^2 + 2f\mathrm{d}x\mathrm{d}y + g\mathrm{d}y^2,$$

where

$$e = \frac{\partial^2 z}{\partial x^2}\Delta \quad f = \frac{\partial^2 z}{\partial x \partial y}\Delta \quad g = \frac{\partial^2 z}{\partial y^2}\Delta,$$

and

$$\Delta = (EG - F^2)^{-1/2}.$$

Casting the above expression into matrix form with:

$$V = \begin{pmatrix} \mathrm{d}x \\ \mathrm{d}y \end{pmatrix} \quad A = \begin{pmatrix} E & F \\ F & G \end{pmatrix} \quad B = \begin{pmatrix} e & f \\ f & g \end{pmatrix}$$

the two fundamental forms become:

$$I = V^{\mathrm{t}}AV \qquad II = V^{\mathrm{t}}BV.$$

Then the curvature of the surface in the direction defined by V is given by:

$$k = \frac{V^{\mathrm{t}}BV}{V^{\mathrm{t}}AV}.$$

Extreme values of k are given by the solution to the eigenvalue problem:

$$(B - kA)V = 0$$

which gives k_1 and k_2, the minimum and maximum curvatures, respectively. k_1 and k_2 are invariant under rigid motions of the surface. Two other curvatures are often defined, the Gaussian or total curvature $K = k_1k_2$ and the mean curvature $M = (k_1 + k_2)/2$ but it turns out that the principal curvatures k_1 and k_2 are best suited to the detailed characterization required by our problem.
As curvature is a geometric property of surfaces, we must introduce a continuous representation of the discrete 2-dimensional array of data in order to calculate local derivatives and then, local curvatures. We retain the so-called facet model (Halarick 1984) which considers a square window of $L \times L$ data points (L is an odd integer) centered on the pixel of interest and calculates the 2-dimensional polynomial fit of degree N ($N \geq 2$), satisfying the criterion of least square errors (other functional representations may be chosen, see Laffey et al. 1985). This polynomial represents the local continuous surface with continuous derivatives up to order N. The window size L can, for instance, be adapted to the spatial extent of the Point Spread Function (PSF). It is very often the case in astronomical images that the PSFs are undersampled; then the lowest size corresponding to $L = 3$ is retained so as to minimize the error introduced by the facet model. This, in turn, maximizes the chance of detecting point-like objects but may result in a large number of false alarms due to local noise (e.g. noise spikes). When the ultimate goal is to restore the background, a trade-off should be reached which insures a minimal percentage of valid pixels as discussed in Sect. 4. When the PSFs are oversampled, then a larger size is recommended to limit the false alarms. For the simple facet model of second order, i.e. a 3×3 window implemented in our examples, the surface $z(x, y) = \log B$ is approximated by a quadric

$$z(x, y) = a_{00} + a_{10}x + a_{01}y + a_{20}x^2 + a_{02}y^2 + a_{11}xy$$

and the practical calculation of k_1 and k_2 is extremely simple.

It is quite convenient to discuss the surface curvature in the (k_1, k_2) space as illustrated in Fig. 1. With the above convention $k_2 \geq k_1$ and only the corresponding half plane

is therefore allowed. Also, according to the above definitions, if a curve on the surface is concave up, the surface curvature in that direction is positive and vice-versa. The allowed half-plane of Fig. 1 is divided in three regions by the k_1 and k_2 axis. The two axis themselves correspond to parabolic points, ridges on the k_1 axis ($k_1 < k_2 = 0$) and valleys on the k_2 axis ($k_2 > 0 = k_1$). The second quadrant $k_2 > 0 > k_1$ is the region of the saddle points. The elliptic points are located in the remaining two half-quadrants, distinguished by their concavity : pits for $k_2 > k_1 > 0$ and peaks for $k_1 < k_2 < 0$. Obviously, a flat is at the origin $k_1 = k_2 = 0$.

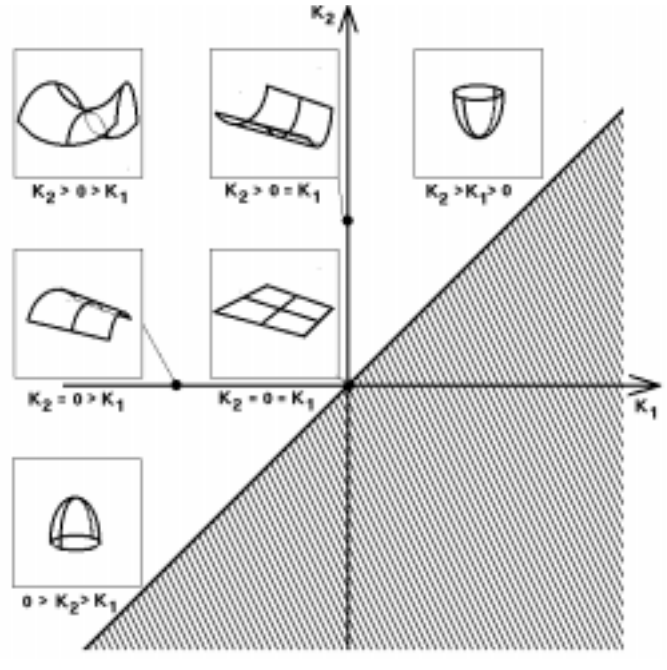

Fig. 1. Surface curvature visualized in the plan of the two principal curvatures k_1 and k_2

3. Discrimination of point-like objects on the (k_1, k_2) diagram

Let us now consider how the above theoretical considerations apply to real astronomical images. Figure 2 is a sketch of the (k_1, k_2) diagram based on a CCD image which exhibits an extended source, the tail of comet P/Halley, a rich star field and various defects and artefacts. The procedure applied to this image will be detailed in Sect. 5 as a practical example. The features of this (k_1, k_2) diagram are sufficiently general to localize the various domains of interest and to establish the appropriate criteria for the discrimination.

The main feature of the diagram is a "butterfly" pattern composed of two wings A and B and a central core C. Both wings pertain to stars, wing A to their top and

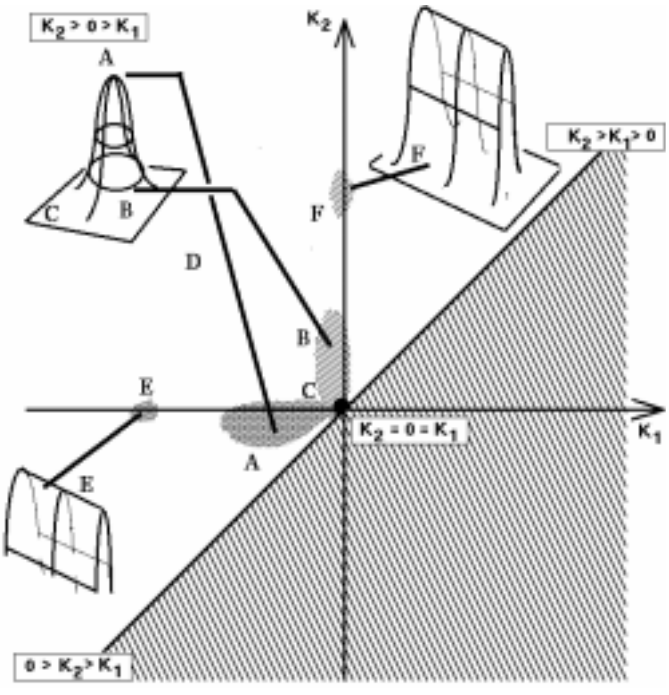

Fig. 2. A practical k_1, k_2 diagram for an astronomical image. The regions of interest are related to specific parts or curvatures of real objects

wing B to their bottom parts. These later parts are fairly asymmetric saddle points leaning to valleys and explaining the closeness of the wing B to the k_2 axis. In fact, the orientation of the wing, as given for instance by the slope k_2/k_1 of their symmetry axis, is a very good estimate of the anisotropy of the surface curvature. For instance, elongated (e.g. trailed) star images will induce a rotation of wing A. The central core C obviously corresponds to the flat parts of the image, essentially the extended object and the background. Finally, two additional features are conspicuous in the diagram, the blobs E and F respectively localized on the k_1 and k_2 axis. They correspond to elongated structures in the original image which involve both ridges (top parts) and valleys (bottom parts). These structures are typically bad columns in the CCD and saturated star images which spill off along the column direction.

It is now a very simple matter to establish criteria for discriminating objects in the original image. These criteria are written in terms on conditions on k_1 and k_2 or on a combination of the two. A mask or several masks are generated and then applied to the original image. Two simple cases may be given as practical examples. Let ϵ_i denote some small positive values of the curvatures. Two criteria may be used to detect the top of the stars

i) $k_1 < -\epsilon_1$
ii) $(k_1^2 + k_2^2)^{1/2} > \epsilon_2$.

Conversely the possible criteria for discriminating the extended source and background are

i) $\mid k_1 \mid < \epsilon_3$ and $\mid k_2 \mid < \epsilon_3$
ii) $(k_1^2 + k_2^2)^{1/2} < \epsilon_4$

respectively delimiting a square and a circular region in the (k_1, k_2) diagram. Our experience is that the two criteria lead to nearly the same result. For practical matters, the modulus of the curvature $k = (k_1^2 + k_2^2)^{1/2}$ has the advantage of combining the principal curvatures so that a single parameter has to be dealt with. Inspection of the histogram of the k values generally allows a straightforward determination of the border between point-like objects and extended objects.

4. Restoration of the background

Many applications of detecting point-like objects in astronomical images require, in a subsequent phase, to restore the background. Here we mean the "flat" parts which will usually consist in an extended object and a true sky background. Applying the criteria discussed in the above section allows to generate a mask or several masks which are combined into a single, final one. The ideal case of point-like objects with Gaussian shapes would lead to a mask composed of perfect circular areas at the location of each "star". Defects or saturations in the real images introduce distortions of those areas. They may even present a hole if the image is severely saturated and will appear as rings which may further be partly open. These problems can be alleviated by a simple morphological operation which consists in a slight geometrical extension, a "dilatation", of the above areas. Usually, an extension by one or two pixels is adequate and allows to generate the final mask. It is therefore a binary image where the valid pixels of the background are set to 1 and the invalid pixels of the point-like objects to be removed are set to 0. The purpose of the next and final phase is to replace patches of invalid pixels in the original intensity image by a local interpolation, so-called pyramidal interpolation, using a non-linear multiresolution method which reminds of the general lines of the Haar wavelet analysis and synthesis.

This method is preferred to a global surface fitting procedure because it preserves the original intensity values of the background where it is valid, allowing both better local and global adjustments due to the multiresolution process. Incidentally, a strictly local interpolation is inapplicable because the remaining valid points may appear as a set of unconnected domains.

The procedure starts by shrinking the original $N \times M$ pixels image to a $N/2 \times M/2$ image. Each pixel in the latter corresponds to an area of 2×2 pixels in the former. Their value is a mean of valid pixels in the 2×2 related area. If 3 or 4 pixels in this area are invalid, the mean is classified as invalid. After this first phase, we obtain an image of means and a corresponding binary image or mask of valid or invalid means. Repeating this procedure builds another couple of images (means and mask) of lower resolution. The image of means is called the analysis image.

The procedure is repeated l times until either the extinction of patches of invalid pixels is obtained, or N_{l+1} or M_{l+1} are too small (that is equal to or less than 4 pixels by side). If some final pixels are invalid, they are replaced by values estimated with a global adjustment of a quadric surface over the remaining valid pixels. After this, all pixels are valid at the lowest resolution level and the synthesis process can start.

The synthesis is done step by step as follows:

1. expand the synthesis image built in the previous $l - 1$ step by 2×2,
2. substitute invalid pixels in the analysis image at this step l with the corresponding pixels from the above expanded image,
3. take this image (with substituted pixels) as the new synthesis image at this step l.

The procedure stops when the initial resolution is recovered. The method restores, in each patch of invalid pixels, the low resolution estimate from neighbouring valid pixels.

This procedure needs a minimal percentage of valid pixels ($\sim 50\%$) to obtain meaningful results. When the number of valid pixels is very low, the threshold to obtain a significant mean from a 2×2 area in an analysis step can be reduced to one valid pixel.

Turning to practical considerations, the iterative nature of this algorithm allows an easy implementation, in spite of its complex appearance. Due to the pyramidal nature, it uses less dynamic memory (i.e. for an initial $M \times N$ pixels image, we need a temporary storage equal to $2(M \times N + (M \times N)/4 + (M \times N)/16 + ...)$, and its complexity is $O(M \times N)$ operations, far better than spline 2-D interpolation methods. And last but not least, it is very stable even with ill-conditionned data (for instance, very irregular spatial distributions of valid data).

5. Two practical applications

We now illustrate the practical use of the method on two specific, very different examples.

5.1. The case of Comet P/Halley

Figure 3 displays a wide-field ($9^\circ \times 5^\circ$) CCD image of comet P/Halley obtained on 17 April 1986 at the European Southern Observatory in Chile (Lamy et al. 1987). The camera consisted of a cooled, large-format CCD (630×1024) pixels behing a Canon f/2.8, 100 mm lens. This particular image was obtained with a B filter and with an exposure time of 15 minutes. In April 1986, the comet was crossing the southern Milky Way resulting in a very rich star-field surperimposed onto the comet. The fast f-number led to a non-negligible distortion of the star images which are further slightly trailed due to an imperfection of the guiding. Several star images are saturated, the most severe ones presenting the characteristic linear

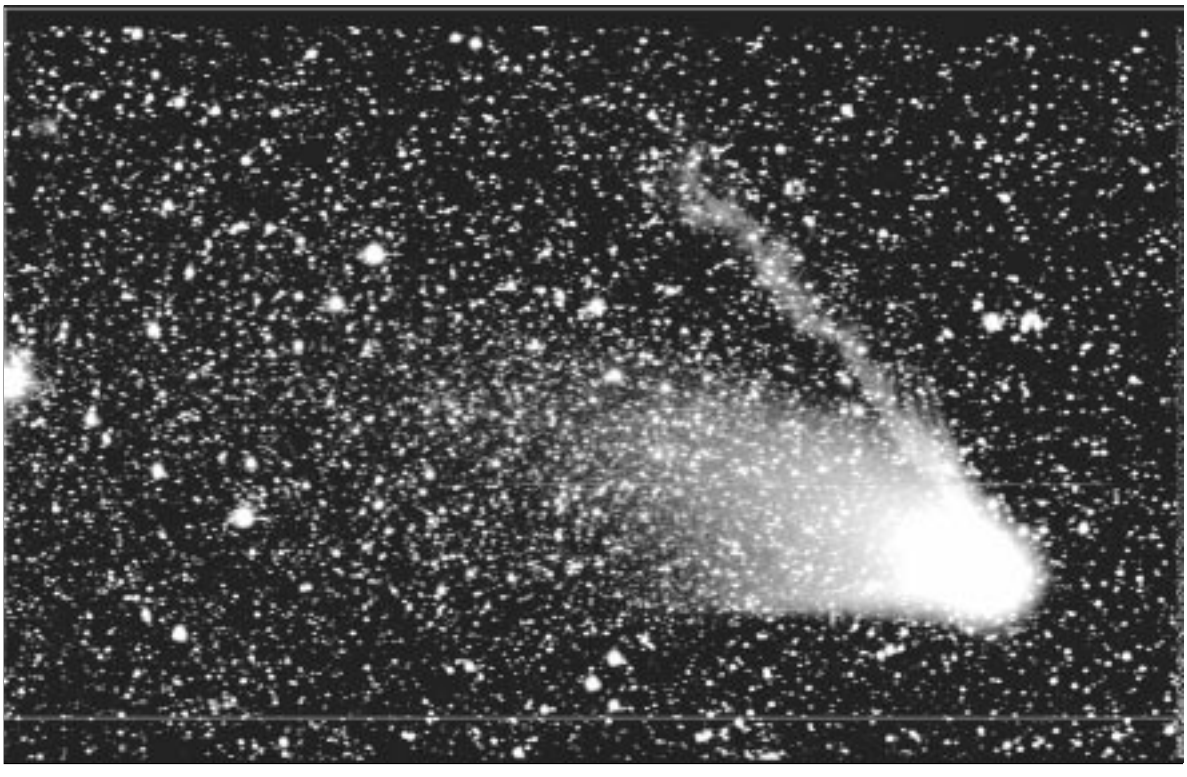

Fig. 3. The logarithm of the intensity image of comet P/Halley after specific corrections described in the text

Fig. 4. The bidimentional histogram of the principal curvatures (k_1, k_2 diagram)

extension along a column. The quantitative analysis of the dust tail to retrieve for instance the physical properties of the dust and its production rate, requires the removal of all star images to perform its accurate photometry. Following the classical processing of CCD images (bias removal, flat-fielding ...) and specific corrections for vignetting and atmospheric transmission, we started applying our method by taking the log of the intensity.

Figure 4 shows the (k_1, k_2) diagram of principal curvatures. The characteristic features discussed in Sect. 3 are present: the butterfly pattern and the two blobs on the k_1 and k_2 axis. Note the conspicuous angular offset of the A wing from the bissector ($k_1 = k_2$), to be compared to the closeness of the B wing to the k_2 axis, offset produced by the trailed star images.

Figure 5 represents the histogram of the modulus of the curvature $k = (k_1^2 + k_2^2)^{1/2}$.

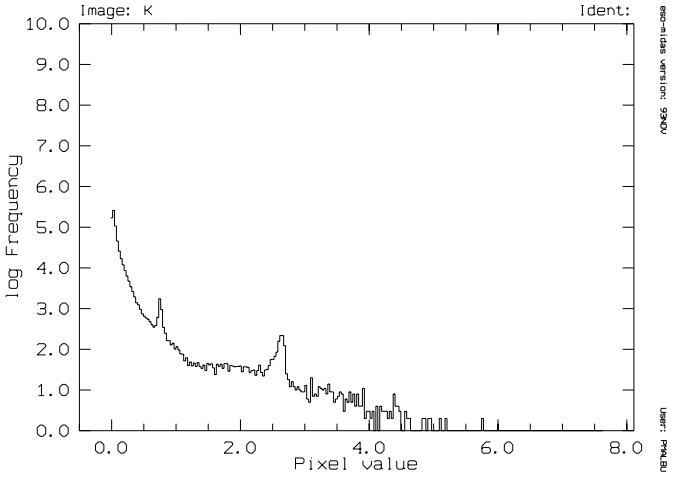

Fig. 5. The histogram of the modulus of the curvature and the selected threshold value $k = 0.09$

Figure 6a displays the first mask generated with a discrimination level of $k > 0.09$. The few, heavily saturated stars produced flat tops which were not properly detected. A second mask was generated by applying a threshold to the intensity to cope with this problem.

Figure 6b display the final mask resulting from previous masks (OR operation) and the morphological operation described in Sect. 4. The percentage of invalid pixels amounts to 32% which ensured a good restitution of the background.

Figure 7a shows the intensity of the final, "restored" image of comet P/Halley with an intensity range similar to that of the original image (Fig. 3).

Figure 7b is similar to Fig. 8 except for a streched intensity range to visualize the full extent of the dust tail. The effenciency of the method is excellent: only traces of the few most intense (saturated) stars are persisting although at very faint levels.

This method was successfully applied to a set of 32 wide-field images of comet P/Halley.

5.2. The case of the coronographic image

Figure 8 displays a CCD image of the solar corona obtained on 4 May 1996 with the LASCO/C3 externally-occulted coronagraph aboard the SOHO spacecraft located at the Sun-Earth L1 Lagrangian point. This instrument has an effective focal length of 77.6 mm and an effective f-number of 9.3. The 512×512 pixels image has been extracted from an original 1024×1024 pixels image centered on the Sun and obtained with a blue filter ($450-520$ nm bandpass) with an exposure time of 10 minutes.

The $8^\circ \times 8^\circ$ field-of-view includes the outer K and F coronae (the Sun and inner corona are blocked by the occulter), comet Hyakutake, a few stars and numerous impacts of cosmic rays.

The application of our method was quite straightforward in this case. The mask was only modified to retain the comet (whose head is slightly saturated) and the bright fringe surrounding the occulter.

Figure 9 displays the final result after correction, and Fig. 10, the difference with the original image. All artifacts and star images have been removed, including very faint ones which were hardly noticeable on the original image. This method has been implemented in the routine processing of the ~ 100 images produced daily by the LASCO experiment.

6. Conclusion

We have presented a new method for the discrimination of point-like objects using surface curvature. The characterization by the curvature offers a very general and flexible mean which can be adapted to any specific problem without any necessary a priori knowledge of the shape of the point-like objects. A few trials is basically what is needed to determine the threshold for discrimination. Two applications have illustrated this method. In the first one, stars were removed from an extended object, the dust tail of comet P/Halley. In the second one, cosmic ray impacts were removed from single CCD images of the solar corona. In both cases, the point-like objects are efficiently detected and corrected using a pyramidal interpolation.

References

Brueckner G.E., Howard R.A., Koomen M.J., et al., in: Solar Phys. 162, 357

Denger, Billie, 1984, in: Proc. 7th. Int. Conf. Pattern Recogn., Montreal, p. 863

Do Carmo M.P., 1976, in: Differential Geometry of Curves and Surfaces. Prentice-Hall, Englewood Cliffs

Doneddu A., 1978, in: Géométrie différentielle. Intégrales multiples. Vuibert, Paris

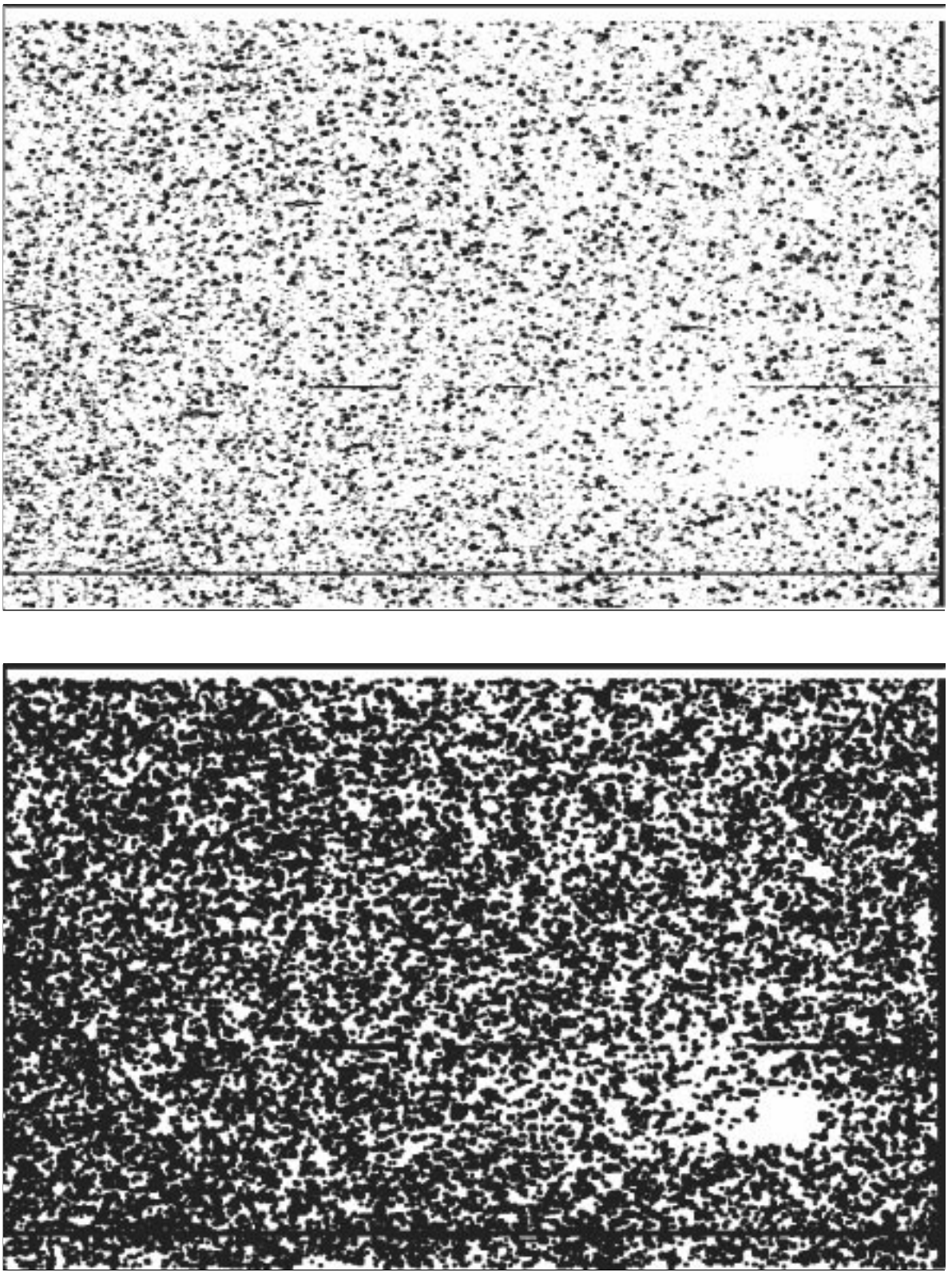

Fig. 6. a) The mask resulting from thresholding the modulus of the curvature at $k = 0.09$. **b)** The final mask as described in the text

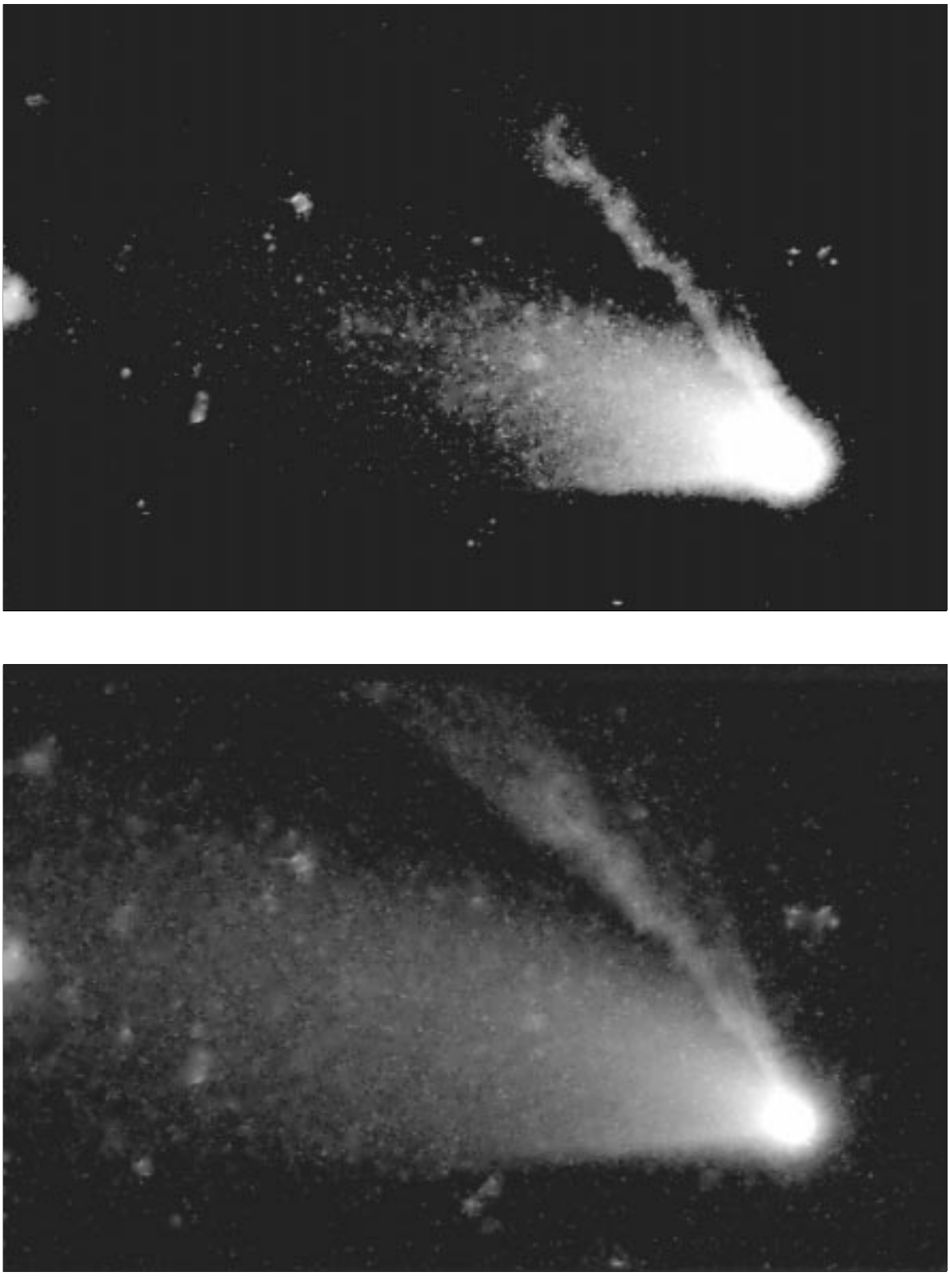

Fig. 7. The intensity of the final, "restored" image of comet P/Halley using the intensity ranges. **a)** $900 < I < 1900$, similar to Fig. 3. **b)** $50 < I > 1900$

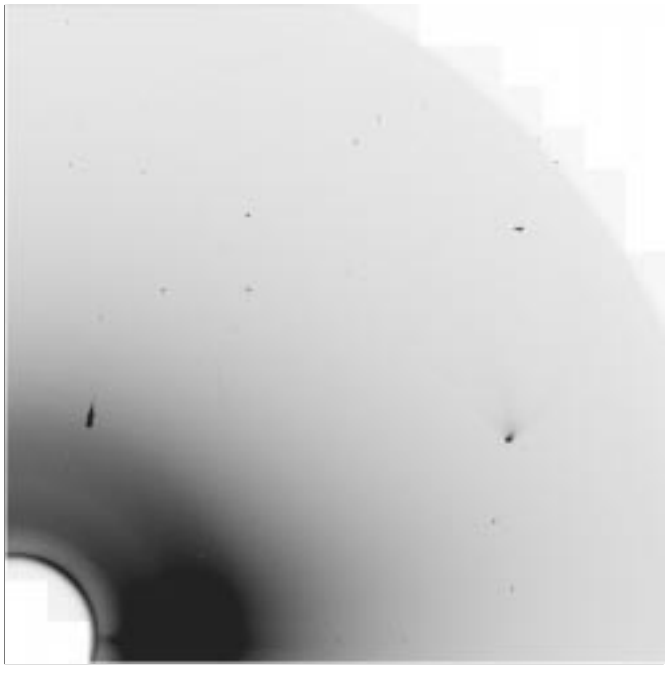

Fig. 8. Intensity image og the solar corona (north-west quadrant). The long exposure time caused saturation of the equatorial K corona and the head of comet Hyakutake. The Sun and inner corona are blocked by the occulter-bottom left corner)

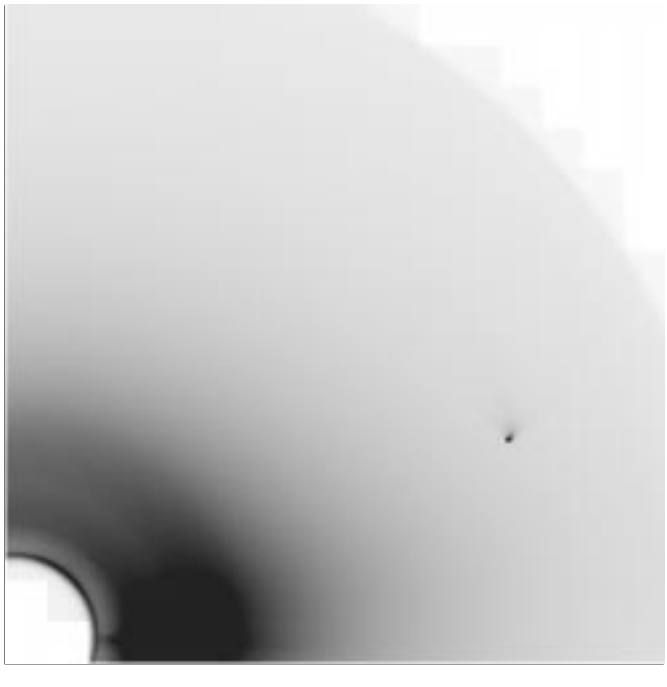

Fig. 9. Intensity of the final "restored" image of the solar corona including comet Hyakutake

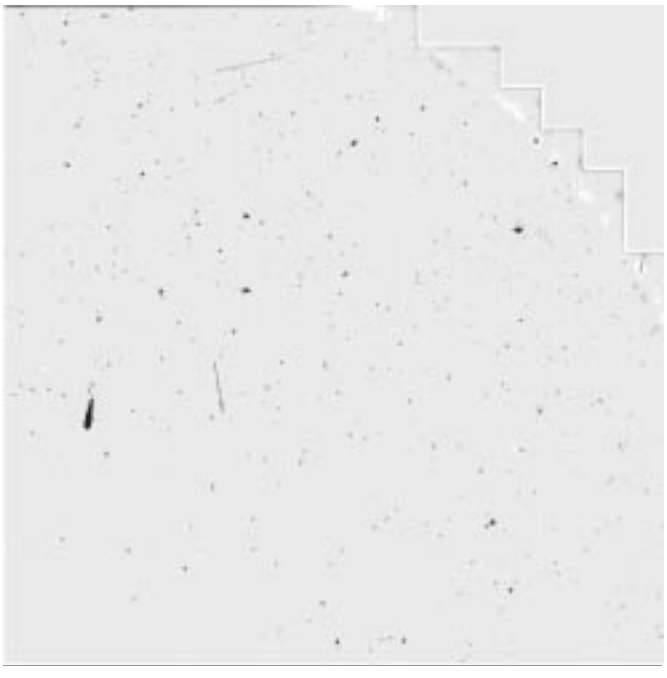

Fig. 10. Difference of the original and corrected images revealing mostly impacts of cosmic rays and a few stars

Faux I.D., Pratt M.J., 1987, in: Computational geometry for design and manufacture. Ellis Horwood, New York
Halarick, 1984, IEEE Transaction on Pattern Analysis and Machine Intelligence PAMI-6, 58
Halarick, Watson, Laffey, 1983, in: International Journal of Robotics Research 2, 50
Laffey, Watson, Halarick, 1985, in: Computer, Vision, Graphics, and Image Processing 29, 143
Lamy P.L., Pederson H., Vio R., 1987, A&A 187, 661
Nackman, 1984, IEEE Transaction on Pattern Analysis and Machine Intelligence PAMI-6, 442
Paton, 1974, in: Computer, Graphics, and Image Processing 4, 40
Peet F.G., Sahorta T.S., 1985, IEEE Transaction on Pattern Analysis and Machine Intelligence PAMI-7, 734
Peucker, Douglas, 1975, in: Computer Graphics, and Image Processing 4, 375
Toriwaki, Fukumura T., 1978, in: Computer Graphics, and Image Processing 7, 30