AUTOMATION OF AREA MEASUREMENT OF SUNSPOTS

LAJOS GYŐRI
Heliophysical Observatory of the Hungarian Academy of Sciences,
Gyula Observing Station, 5701 Gyula, P.O. Box 93, Hungary

(Received 21 January 1997; accepted 11 December 1997)

Abstract. When drawing up a database for sunspots from a large collection of white-light films, a need for the automation of the process arises. The concepts used at the automation of the area measurements of sunspots are described. As an example, sunspot groups NOAA 5521 and 5528 are processed and the areas obtained are compared to the measurements published in the literature. Similar values are obtained, except umbral areas published by Steinegger *et al.* (1996) which are significantly larger than ours. We find that the differences may be attributed to the fact that the definition proposed by Steinegger *et al.* (1996) for the penumbra–umbra border of a sunspot is not equivalent to those used for the measurements of others of the umbral area.

1. Introduction

We have about 100 000 white-light photoheliograms at Debrecen Observatory and at its Gyula Observing Station. To process this large number of heliograms, for example to compile a heliographic coordinate and area database for sunspot groups, it is expedient to minimize the human involvement in the process. The minimization of the human interaction in the process of the compilation of the database provides three advantages: the first one is the economic benefit of reducing cost, the second one arises from the fact that the homogeneity of the database is more relevant using instrumental decision instead of human judgment, and the third one is the speeding up the processing time.

In order to achieve a significant automation, a program package was developed. This software, using scanned or CCD image of a sunspot group determines automatically the contours characterizing the umbra-penumbra and the photosphere-penumbra boundaries and from this the umbral and penumbral areas and the heliographic positions of the spots are determined.

The matter was divided into three parts. This paper is the first part which provides the concepts used in the determination of the area of the sunspots. The second paper will include the determination of the heliographical coordinates of the sunspots, the corrections used at the reduction and the automatic identification of the spots of two heliograms taken nearly at the same time. Generally, we take three heliograms nearly at the same times (within ten minutes). The automatic identification provides the possibility of making averages over these heliograms, resulting in better accuracy. In the third paper, the observational material, the instrumentation and the assessment of the accuracy will be addressed. Additionally, the effect of different factors on accuracy, as was revealed in the testing runs of the programs, will be discussed.

Solar Physics **180:** 109–130, 1998.

To digitize a photograph usually two devices are used: an image scanner which is suitable for digitizing large images but usually distorts the image in an unpredictable way (depending on the quality of its step motor) or a CCD camera.

We use an 8-bit CCD camera (the image size is 768 × 576 pixels) to digitize a sunspot group on our photoheliogram. The diameter of the full-disc solar image on our photoheliogram is about 100 mm. The scale of the digitized image is about $0.25''$ pixel^{-1}.

If we had a suitable scanner then whole photographic image (not only a sunspot group) of the Sun should be digitized and, in this case, the digitized image would be suitable for obtaining the global heliographic coordinates of the sunspots without using any other instrument. But if we have only the digitized image of a sunspot group (as in the case of a CCD camera) then, for positioning this image on the Sun, it is necessary to determine the center and the radius of the photoheliogram and the position of some spots (at least three) with a coordinate measuring instrument.

If we want to determine the area of a sunspot then first we must answer the question: what is a sunspot? Here, we are not interested in the physical properties of a sunspot but rather in the properties of its image. If we have a look at an image containing a sunspot group then we can observe two essential features of the spots: they are darker then the surrounding photosphere and they have borders (that is the change of the intensity from its photospheric value to the spot value takes place within a relatively short distance).

On the basis of the above-mentioned properties of the sunspot, the method for area determination of the sunspots can be divided into two main categories: the method using only the lower intensity of the sunspots (in this paper this method will be addressed as a threshold method) and the method making use both of their lower intensity and the gradient of the intensity of the sunspot image (this one will be referred to as a border method).

Threshold methods: these methods use a prescribed threshold intensity value (cutoff value) to decide whether an image point belongs to a spot or not (in fact two threshold values are prescribed, one for the penumbra and one for the umbra). On the basis of how the cutoff values are determined, these methods can be divided into two groups. The first group uses cutoff values inferred from statistical studies of the spot-to-photosphere intensity ratios (Chapman and Groisman, 1984; Chapman *et al.*, 1989; Brandt *et al.*, 1990). Chapman and Groisman (1984) adjusted a cutoff value for intensity of penumbra so that a general agreement between areas published in SGD and their own area measurements should hold. Chapman *et al.* (1989) tried to put the threshold method on a firm basis, incorporating the essence (the large intensity gradient) of the border method into it. They found that the intensity of the steepest descent at the outer edge of the penumbra is at 8.5% below the photospheric value. But this goal can be achieved only on a statistical basis because according to our experience (see Section 5 of this paper) the intensities of the borders of different spots (especially in case of the umbrae) are different.

The second group of the threshold method tries to determine the cutoff values for penumbra and umbra, utilizing some properties of the histogram and the cumulative frequency diagram of the image of the sunspot group Steinegger *et al.* (1996). We shall return to this method in Section 5 of this paper.

Border methods: here, at first the borders of the spots are determined, making use of the sensitivity of the eye to the abrupt changes in the intensity in an image. The following methods can be put here. In Debrecen Observatory the video image of a sunspot group is used to measure its area. An isodensity line is fitted to the edge of the spot. This is achieved by changing the intensity of the isodensity line until the best fit (according to the judgment of the measurer) is obtained (Dezső, Gerlei, and Kovács, 1987). The spot areas in the *Greenwich Photoheliograph Results* (De La Rue, 1869) were measured using a glass plate with accurately ruled squares in contact with photographic film. The numbers of the squares contained in a sunspot were counted and transformed into area on the solar disc. The methods using sunspot drawings can also be placed among the border methods.

The border methods presented above for area determination of the sunspots are time consuming, tiresome and demand a lot of attention. These may be the reasons why the observational materials for sunspots piled up in observatories all over the world are, for the most part, unprocessed. The reasons which have given impetus to the measurements of the heliographic coordinate and area of the sunspots at the beginning of these measurements are still valid. In fact, the comparison of the sunspot areas with the high accuracy measurements of the solar constant might yield new results in the field of solar-terrestrial relations.

In the light of the above-mentioned facts it would be useful to work out a method which, making use of the image processing techniques, automates the border method of the areal measurement of the sunspots.

2. Basic Concepts

The first idea that comes to mind when solving the task of looking for boundaries in an image is the use of the gradient image. High values for the gradient indicate the presence of a border, an abrupt change in the gray-level value. But the trouble with the gradient image in our case is that the gradients along the boundaries (photosphere–penumbra, umbra–penumbra) do not represent a contiguous, clearly cut closed line mainly due to filamentary structure being present in a sunspot group. Some manipulation of the gradient image, such as global and local thresholding (see Section 4.4 for the definition of these terms), may help but does not solve the problem. This situation can be easily seen in Figures 1–3. Although, it cannot be directly interpreted as photosphere–penumbra and umbra–penumbra boundaries, the gradient image contains a lot of information about these boundaries, as can be seen by comparing the gradient image (Figures 1–3) with the original image (Figures 5 and 6). This information can be exploited as follows: let us assign a

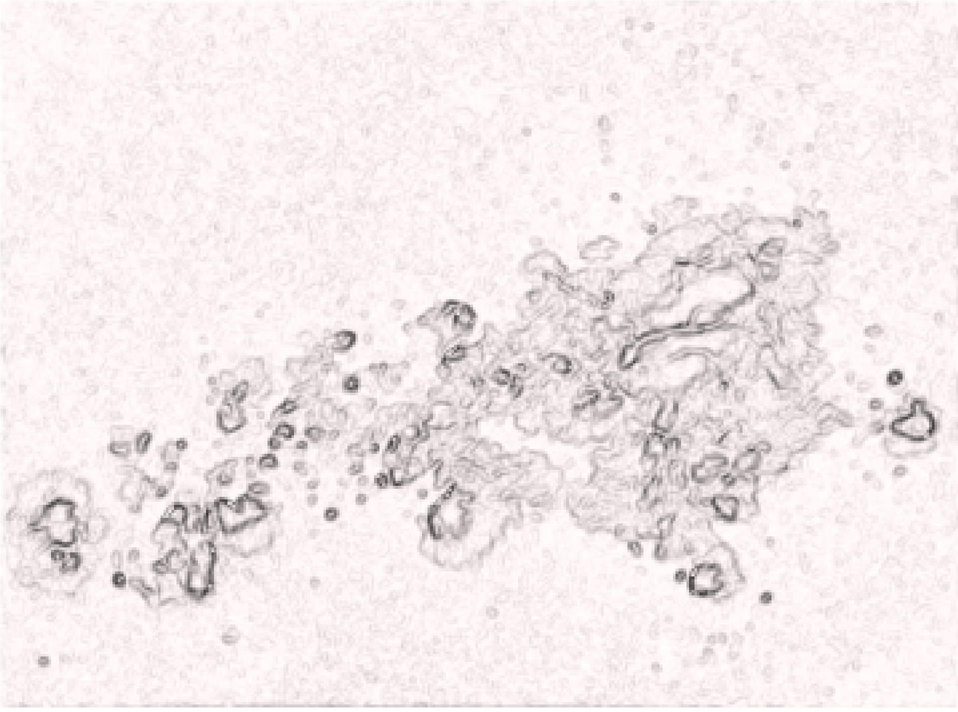

Figure 1. Gradient image of sunspot group in NOAA region 5528 (15 June 1989, 07:43:03 UT).

gradient value to every intensity contour of the image by summing up the gradient values along the points of the contour (the term contour means here an isodensity line) and dividing the sum by the numbers of the contour points. Let AGAC (average gradient along contour) denote this value.

It is reasonable to choose the contour having maximum AGAC among the contours of a spot as the umbra–penumbra boundary of the spot. This will be the guide line for finding the umbra–penumbra boundary, although, in some cases it must be modified. Similarly the contour having the first local maximum (counted from the photosphere) in AGAC among the contours of a spot can be taken as the photosphere–penumbra boundary of the spot, although, this definition must also be modified in some cases. As an example Figure 4 plots the AGAC values along the intensity contours of spot S of sunspot group in NOAA region 5528 (see Figures 5 and 6 for spot S).

3. Overview of the Algorithm

The outline of the processing scheme is as follows:

– Checking the CCD image. By taking too large a transillumination, it is possible to burn in (to overilluminate) the umbra and so to distort the spot boundaries. Here,

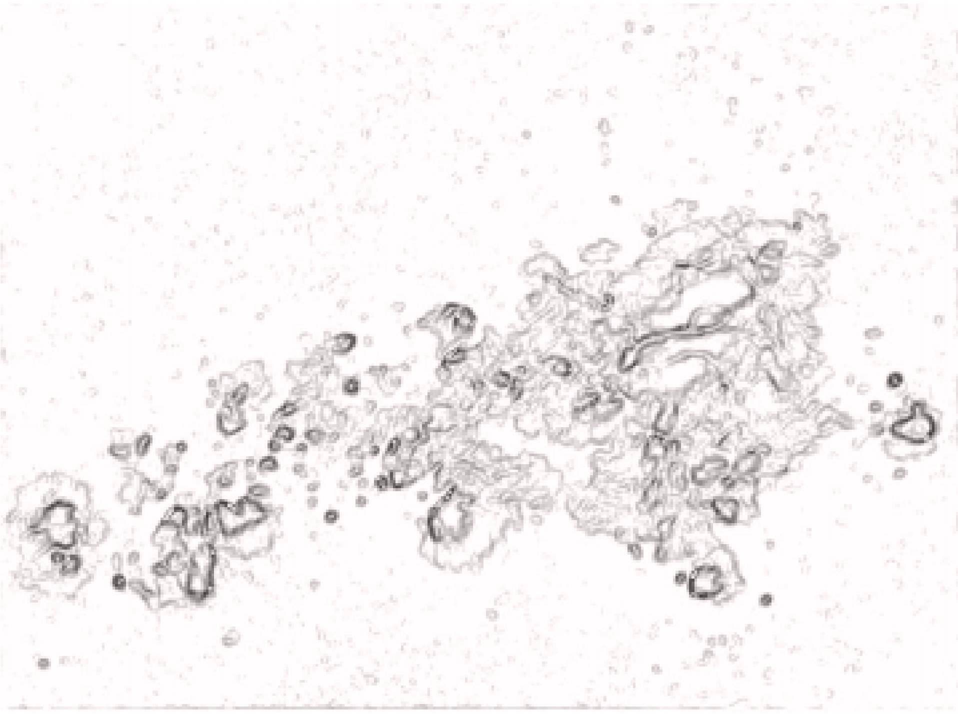

Figure 2. Same as Figure 1, but with global thresholding. The threshold value is the average photospheric gradient of the image.

using the histogram of the image, this overillumination is checked. Additionally, as we shall see later, if the spot is near the solar limb, the CCD image must be properly oriented relative to the limb to make allowance for limb darkening correction. This proper orientation is checked as well.

– Filtering out cross-threads. For orientation purpose, two spider lines are placed in the primary focus of the heliograph. It must be checked that these lines are in the image and if so, they must be filtered out.

– Background correction. This is used to correct for the unevenness of the transillumination used in taking the CCD image of a sunspot group and for the sensitivity distribution of the CCD array.

– Limb-darkening correction. The limb darkening of the solar disc distorts the contours belonging to a given intensity (gray level), so it must be corrected for.

– Fourier filtering the image. To increase the sharpness (especially in bad seeing conditions) of the boundaries of the spots it is useful to do some spatial frequency manipulation of the image.

– Determining the *Beginning Intensity Level*. Our CCD images have 256 intensity levels. To determine the intensity contours of the image it is reasonable to begin with the level close to the level of the penumbra–photosphere boundaries but slightly below it. This is the *Beginning Intensity Level* of the processing.

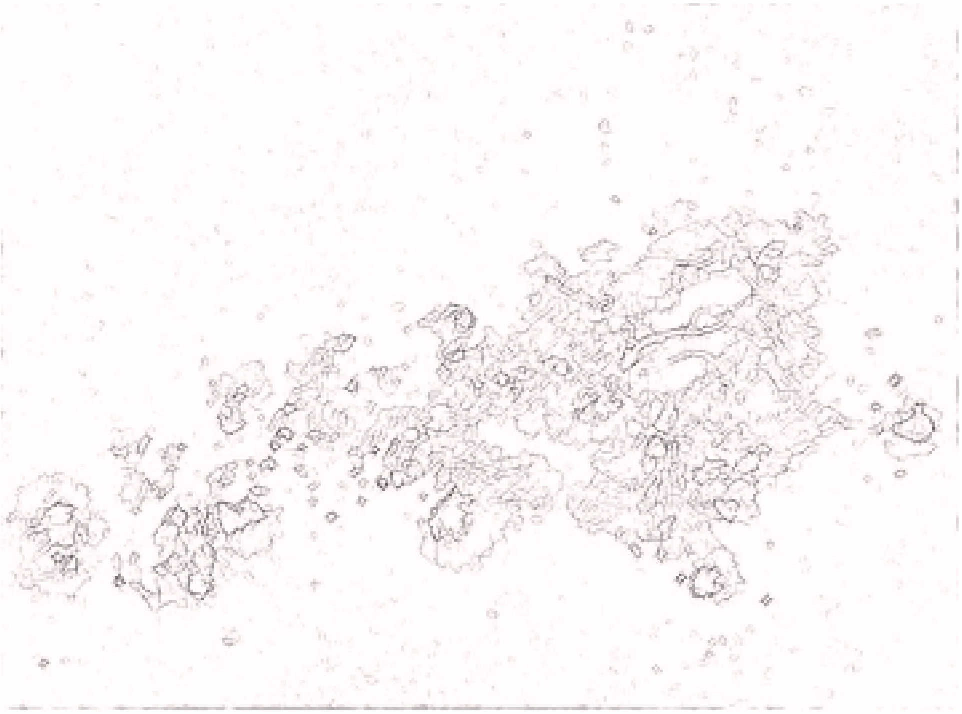

Figure 3. Same as Figure 2, but applying a local thresholding with a 3 × 3 environment.

– Determining the gradient image.

– Determining each contour of every intensity level below the *Beginning Intensity Level*. Here, the AGAC value belonging to the contour is also determined.

– Determining the contours belonging to a *Local Minimum* or a *Local Maximum*. Every *Local Minimum* and *Local Maximum* is determined, and a subset of the contour set determined above is assigned to a *Local Maximum* or a *Local Minimum* of the image. Two subsets belonging to two *Local Minima*, of course, may contain common contours if the two *Local Minima* have merged at the intensity level of these contours.

– Filtering out *Dark Penumbral Filaments*, *Granular Local Minima* and *Bright Regions*. Not every local minimum in the image of a sunspot group is a spot. It can be a *Dark Penumbral Filament* – an unusually dark region between bright penumbral filaments – or a *Granular Local Minimum*, i.e., a real dark granule or one caused by the seeing. *Bright Regions* in a sunspot penumbra show up as *Local Maxima*. They must be handled separately and their areas (if their intensities are near to that of the photosphere) must be subtracted from the penumbral area.

– Determining the penumbral boundaries of the spots.

– Determining the umbral boundaries of the spots.

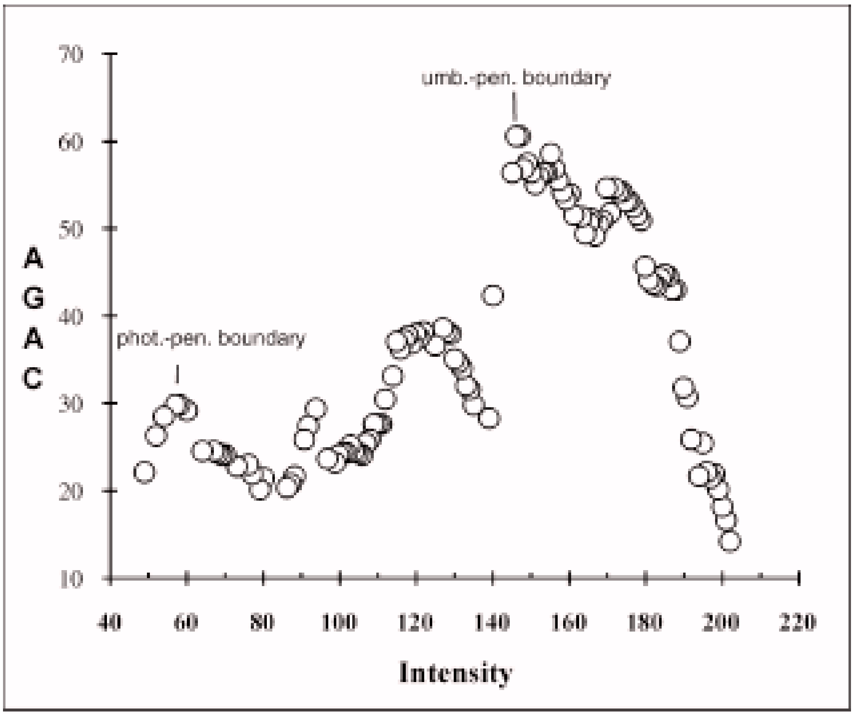

Figure 4. AGAC (average gradient along contour) versus intensity of the contour for spot S shown in Figures 5 and 6. The intensity contour corresponding to the first local maximum is the photosphere-penumbra boundary, and the intensity contour corresponding to the global maximum is the penumbra –umbra boundary of spot S. The other local maxima in the figure reflect inner structures in the penumbra and umbra. Note: As the inverse image is used to determine the boundary contours of the spots, thus, the photosphere is at the lower intensities and the umbra is at the higher ones.

– Merging *Local Minima* having the same *Umbra–Penumbra Boundary* into one umbra.

– Up to now, we considered the boundaries of the spots as iso-density lines, but sometimes different parts of a boundary could belong to different iso-density lines. To take into account this fact, a correction is applied to the boundaries of the spot.

– Determining the umbral and penumbral areas and centers of gravity of the spots. This is done by determining the pixels inside the boundary contour. Here, the area of the *Bright Regions* having the intensity of the photosphere inside a penumbra are subtracted from the total area of the spot.

In the next sections a little more detailed description of some steps of the algorithm are addressed.

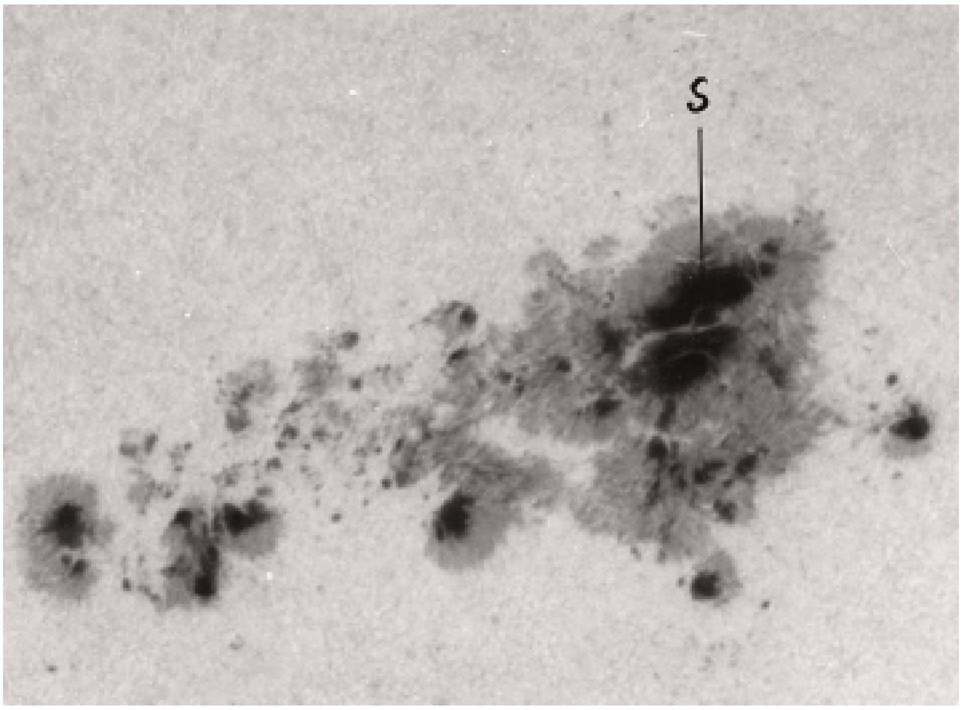

Figure 5. NOAA region 5528 (15 June 1989, 07:43:03 UT).

4. Discussion of the Steps of the Algorithm

4.1. CORRECTION FOR LIMB DARKENING

To correct a sunspot group for limb darkening, two methods are used. If the sunspot group is farther than 100 arc sec from the limb, then the limb darkening over a spot group can be approximated by a plane. From the CCD image of the spot group the position of this plane can be determined, and after that the correction for limb darkening can be done.

When the sunspot group is near the limb (within 100 arc sec) the CCD image is taken in such a way that the columns of the image are parallel to the tangent of the solar limb at the spot group (an accuracy of about five degrees, of the parallelism is enough). Having made corrections for faculae, the average of the profile of the first and the last rows of the image (of course care is taken not to have spots on the edges of the image) represents fairly the limb darkening along the other rows of the image as well, thus they can be used to correct it.

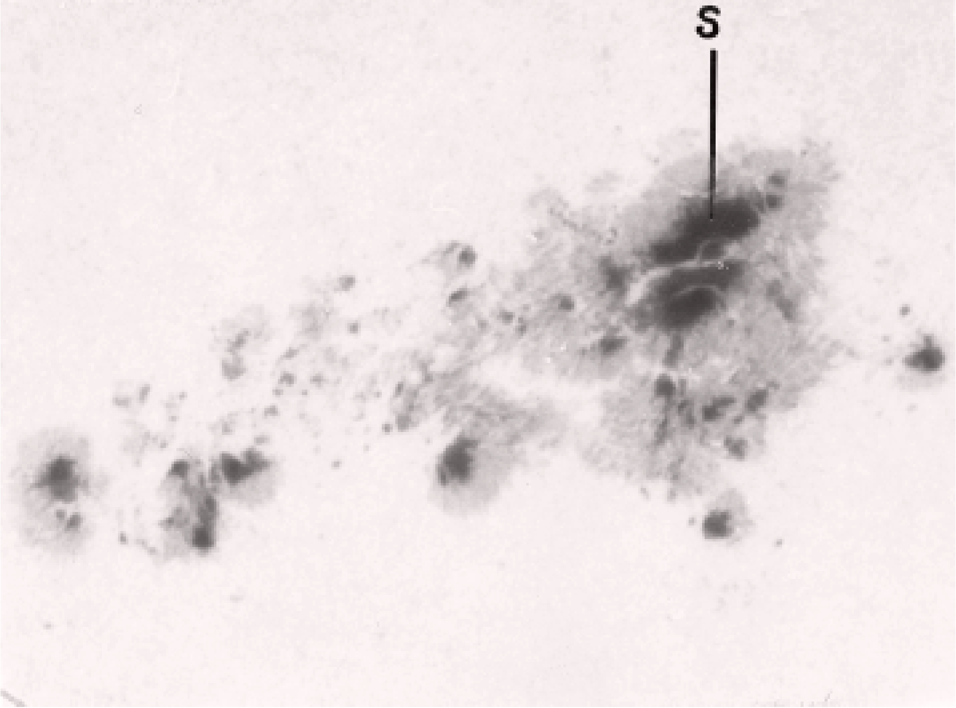

Figure 6. Same as Figure 5, but under-exposed, when taking the photographic copy of the original photoheliogram, to make the umbra structure more visible.

4.2. FOURIER FILTERING

The purpose of the Fourier filtering is to enhance the borders (photosphere–penumbra, penumbra–umbra) of the sunspots. The very short spatial wavelengths are attenuated to decrease noise in the image. It is known (Frieden, 1979) that the simple expedient of decreasing a finite band of the image spectral value at frequencies centered about the origin will result in an enhanced output. For this reason, long wavelengths and the zero spatial frequency are also attenuated. Attenuation of the long wavelengths has also the effect of decreasing the unevenness of the illumination remaining after background and limb darkening corrections. The short wavelengths are amplified to emphasize boundaries.

Figure 7 shows the multiplication factor, $g(\lambda)$, as a function of the spatial wavelength, λ, with which the corresponding Fourier component of the image is multiplied. This function is obtained by prescribing the values of the function at five wavelengths and connecting them with straight lines. The five prescribed points (best suited to our observations) are: (0.25,0); (1,1); (2,1.8); (3,1); (194,0.8). The first number in the data pairs is the spatial wavelength given in arcsecs and the second one is the multiplication factor for this wavelength. We have chosen the maximum value for the gain factor at wave length $2''$ because the diffraction limit

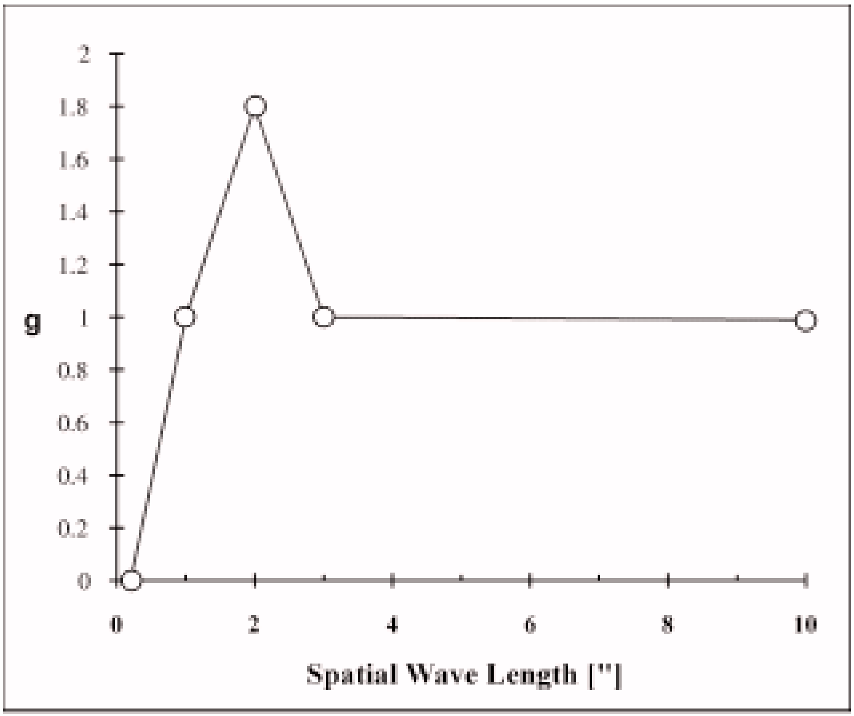

Figure 7. The factor $g(\lambda)$ with that the corresponding Fourier component of the image is multiplied versus the spatial wavelength, λ.

for our photoheliograph is about $1''$. The zero spatial frequency of the image is attenuated by the factor 0.5.

4.3. BEGINNING INTENSITY LEVEL

Using the edge columns and rows of the image, the intensity level and the standard deviation of the photosphere is determined (provided that there are no spots on the edges of the image). Subtracting twice the standard deviation from the photospheric level, a preliminary beginning level of intensity for processing is obtained. If the numbers of the contours belonging to this intensity (i.e., the numbers of the preliminary penumbrae) are larger than a prescribed value then the beginning intensity value is decreased until the number of the contours obtained are smaller than the prescribed value. So, we achieve the result that the beginning intensity level of the processing is near enough to the intensity of the penumbra–photosphere boundary.

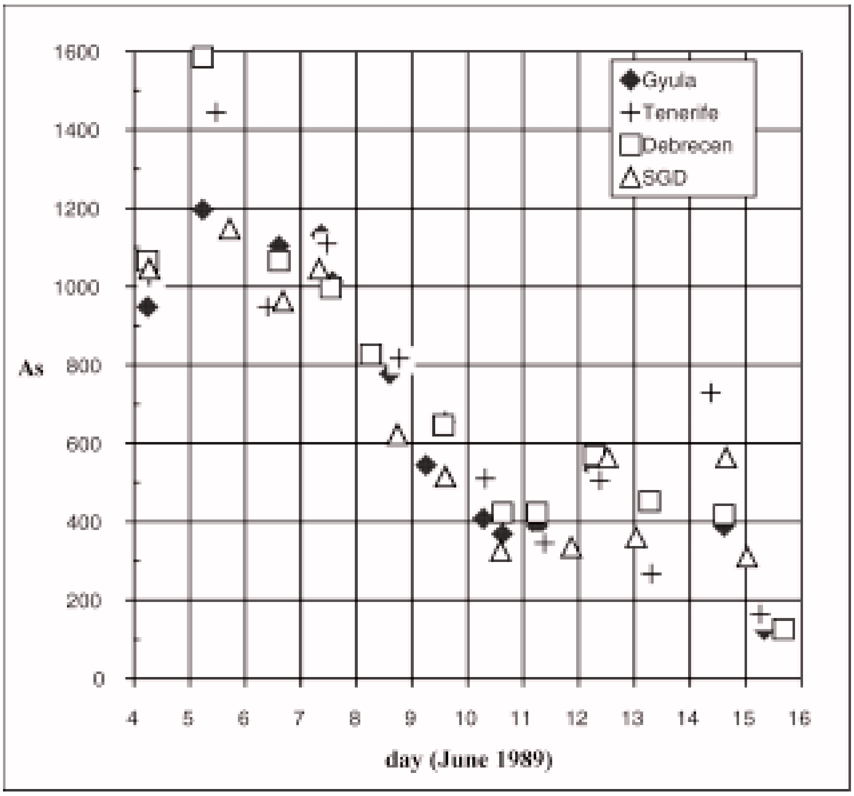

Figure 8. Comparison of the measurements for spot area As (penumbra + umbra) measured in units of 10^{-6} solar hemispheres of the sunspot group in NOAA region 5521 from different observatories.

4.4. GRADIENT IMAGE

To determine the AGAC value belonging to an intensity contour, we need the gradient image. To approximate the gradient in the case of a discrete function, we used the method suggested by Sobel (Tennenbaum *et al.*, 1969). But before using the gradient image we try to remove the gradient values that do not belong to a sunspot border. To get rid of the photospheric noise, a global thresholding is applied using the average gradient value of the photosphere as a threshold value, i.e., if the gradient value in an image point is lower than the threshold value, then the gradient of this point is set to zero. Comparing Figure 1 with Figure 2 it can be seen that a part of the surplus gradients (not representing a spot border) is removed. If we apply a local thresholding (i.e., retaining the gradient value of a pixel only if it is a local maximum inside a prescribed environment of the pixel in question) then further success can be achieved in removing the surplus gradients from the gradient image as can be seen by comparing Figure 2 with Figure 3.

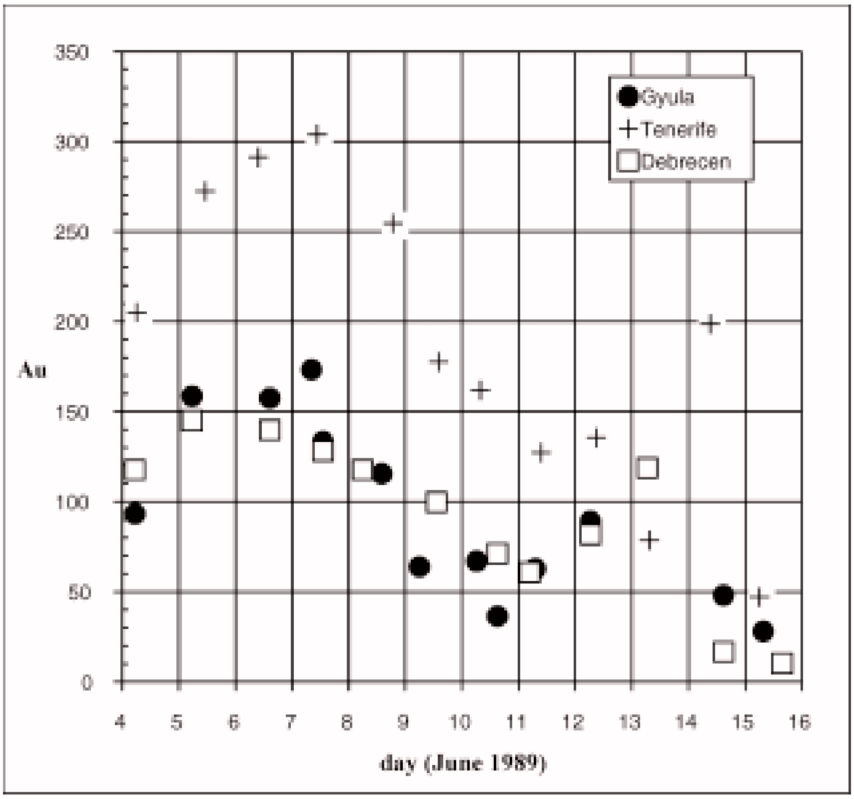

Figure 9. Comparison of the measurements for umbral area Au measured in units of 10^{-6} solar hemispheres of sunspot group in NOAA region 5521 from different observatories.

Not only the magnitude of the gradient is determined but also its orientation. For the determination of the average gradient along an intensity contour, the gradient is weighted by a factor determined from the angle between the direction of the gradient and the tangent to the intensity contour.

4.5. FILTERING OUT *Dark Penumbral Filaments, Granular Local Minima,* AND *Bright Regions*

Some dark lanes between the penumbral filaments in a sunspot group with a well-developed umbra can be so dark and so wide that it is picked up by the program as a *Local Minimum*. Since this lane is not as dark as an umbra and its boundary is not as sharp as the boundary of an umbra, using these properties, they may be filtered out. Note, that using only the criteria for the intensity of the umbra to filter out *Dark Penumbral Filaments*, the penumbral spots (spots above the intensity of the umbra) would also be filtered out.

The level of the intensity where the processing begins is somewhere between the intensity of the photosphere and the intensity of the penumbra-photosphere

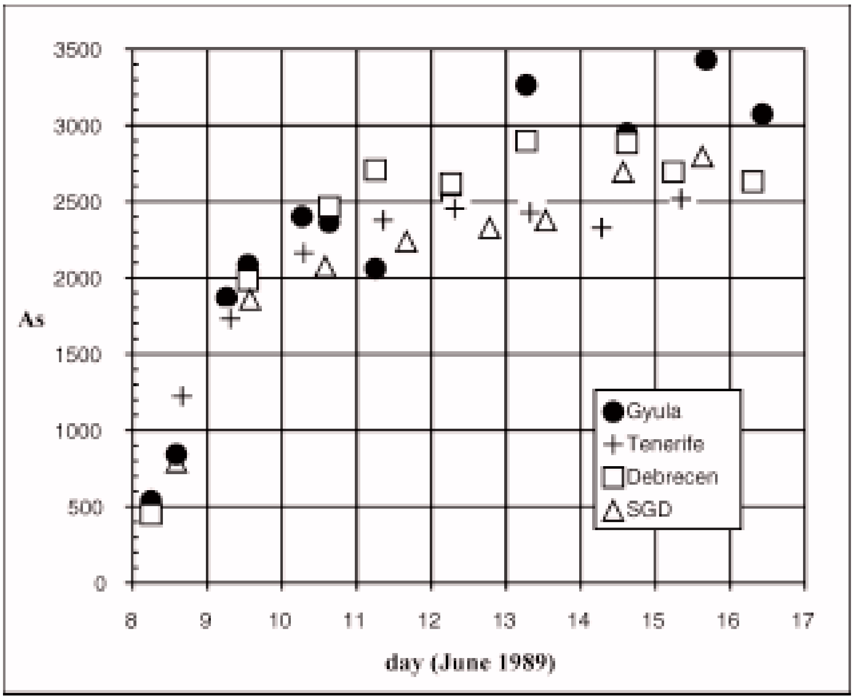

Figure 10. Comparison of the measurements for spot area As measured in units of 10^{-6} solar hemispheres (penumbra + umbra) of the sunspot group in NOAA region 5528 from different observatories.

boundary, so some of the darker granulae are picked up by the program as *Local Minima*. To filter out *Granular Local Minima* it is necessary to determine the difference between *Granular Local Minima* and pores. For this purpose, the following method is used. *Local Minima* having small area are singled out and their average intensity and the standard deviation of their intensity is determined. Subtracting twice the standard deviation from the average intensity, an intensity criterion for *Granular Local Minima* is obtained. Using the AGAC values of the largest spots in the spot group, a gradient criterion for *Granular Local Minima* can also be set up. This criterion says: if the maximum AGAC value of a *Local Minimum* does not reach a certain percentage of the average AGAC value belonging to the penumbra-photosphere boundary of some well developed spots then the *Local Minimum* is a *Granular Local Minimum*.

The contour of a *Bright Region* (photospheric material inside a penumbra) at a given intensity level is inside another contour belonging to the same intensity level. Using this feature of the *Bright Regions*, they are managed separately from the *Local Minima*. If the intensity of a *Bright Region* is below the intensity of the *Penumbra–Photosphere Boundary* of the spot containing the *Bright Region*, then the area of the *Bright Region* is subtracted from the total area of the spot.

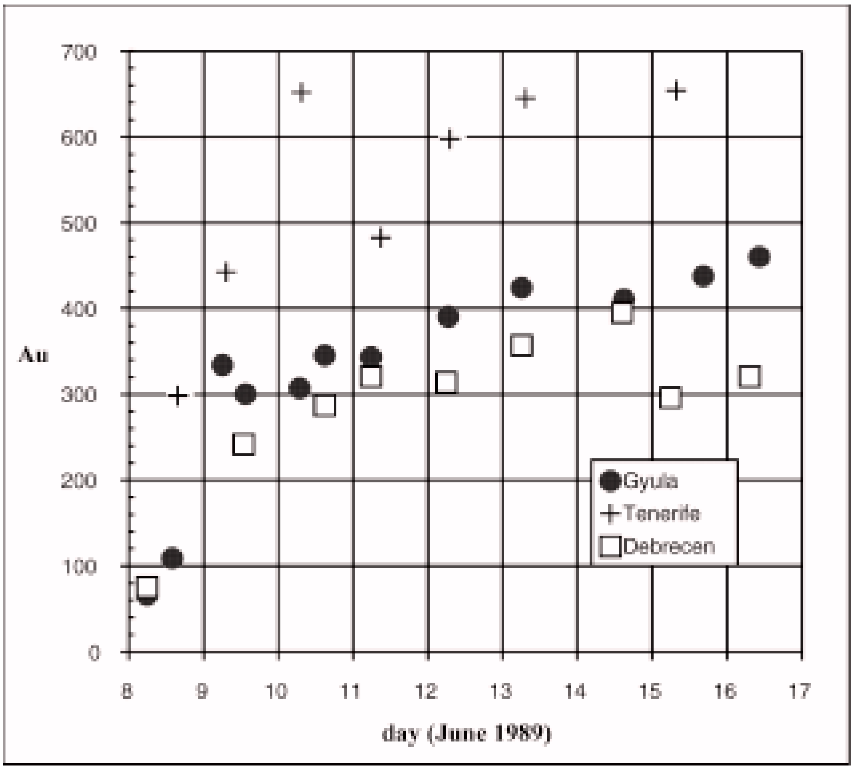

Figure 11. Comparison of the measurements for umbral area Au measured in units of 10^{-6} solar hemispheres of sunspot group in NOAA region 5528 from different observatories.

It may sometimes happen that a *Granular Local Minimum* or a *Dark Penumbral Filament* is not filtered out (we are not too severe in specifying the parameters directing the filtering processes to avoid filtering out pores as well). But this is no problem because we have three heliograms taken nearly at the same times (within ten minutes) and only those spots are retained which are contained in all three heliograms.

4.6. DETERMINATION OF THE *Penumbra–Photosphere Boundary* OF THE SPOTS

The determination of the *Penumbra–Photosphere Boundary* of the *Local Minima* comprises the following steps:

– Those *Local Minima* having the same last contour (numbering of the contours of an *Local Minimum* begins from the umbra) are united into a preliminary penumbra.

– Inside a preliminary penumbra, the *Local Minimum* with minimum intensity (*Minimum Intensity Spot*) is ascertained.

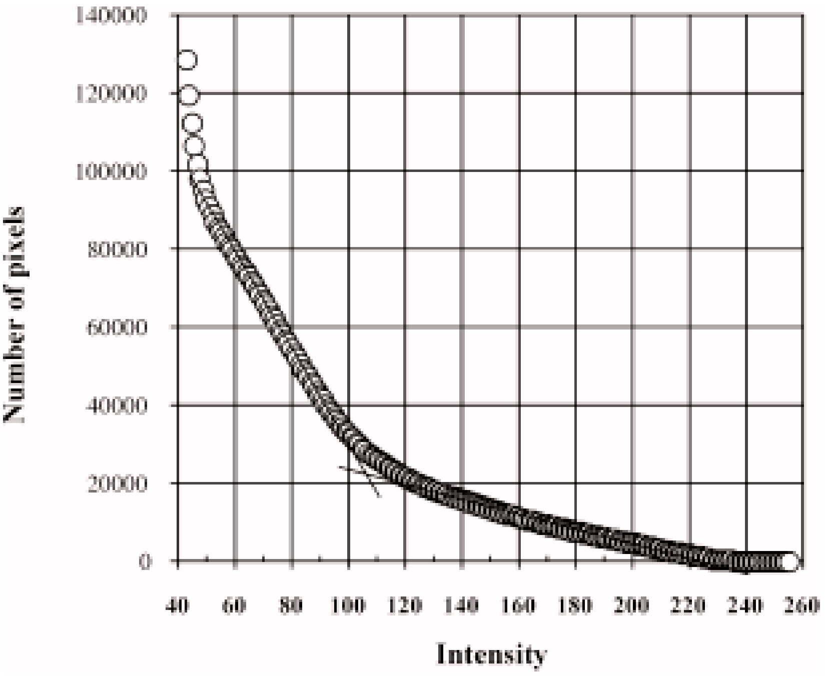

Figure 12. Cumulative frequency diagram for the sunspot group in NOAA region 5528 (15 June 1989, 07:43:03 UT). Straight lines were fitted to the penumbral and umbral parts of the diagram, and from their intersection a cutoff value for umbral intensity was derived. Note: to determine the cumulative frequency diagram we used the inverse image, so, the photosphere is at the lower intensities and the umbra is at higher ones.

– The contour representing the *Penumbra–Photosphere Boundary* of the *Minimum Intensity Spot* is determined, using the concepts mentioned before.

– If, in addition to the *Minimum Intensity Spot*, there is a *Local Minimum* (or there are *Local Minima*) in a preliminary penumbra, then, at the determination of the *Penumbra–Photosphere Boundary* of this *Local Minimum*, two cases are considered. If the contour representing the *Penumbra–Photosphere Boundary* of the *Minimum Intensity Spot* is among the contours of the *Local Minimum* in question, then this contour will also be the *Penumbra–Photosphere Boundary* of the *Local Minimum*, otherwise the relative separation contour with respect to the *Minimum Intensity Spot*, of the *Local Minimum* will be the *Penumbra–Photosphere Boundary* of the *Local Minimum*. In the latter case, of course, the preliminary penumbra is separated into two penumbrae. The explanation of the second case is that the *Penumbra–Photosphere Boundary* of a small spot very near to a large spot can have higher intensity than the *Penumbra–Photosphere Boundary* of the large one.

As defined before, the *Penumbra–Photosphere Boundary* of a *Minimum Intensity Spot* is the contour with the first local maximum (counted from the photosphere) in AGAC among the contours of the spot. But sometimes (especially in bad seeing

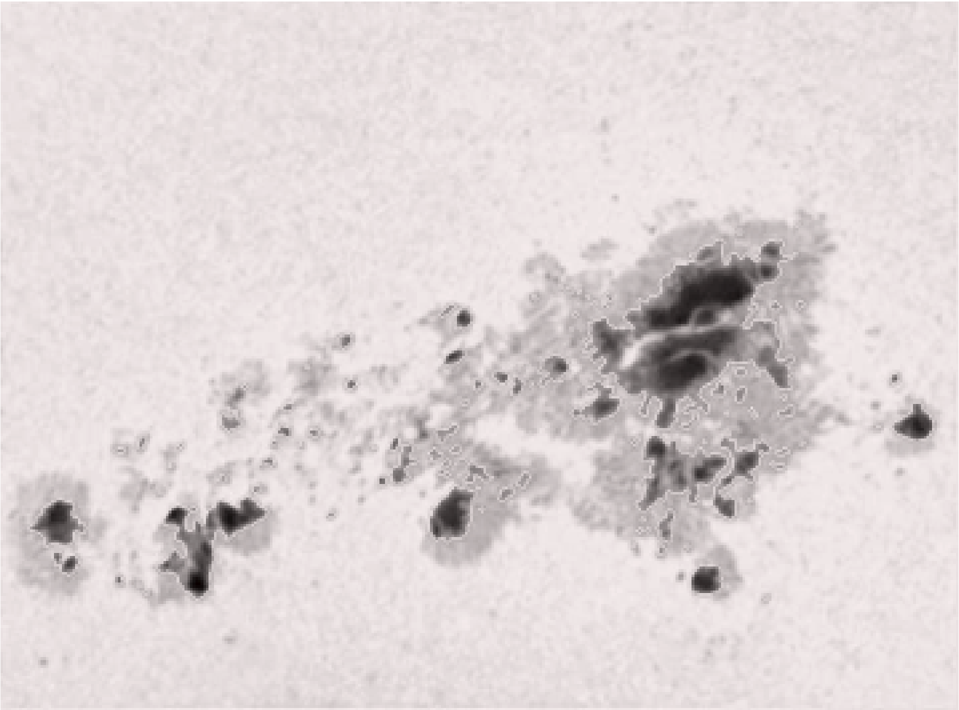

Figure 13. Contours corresponding to intensity level 107 are superimposed on the image of sunspot group (NOAA 5528, 15 June 1989, 07:43:03 UT).

conditions), there is no local maximum of the AGAC on the contours of the penumbra. In this case, the contour having the largest rate of the change of the AGAC from one contour to the next is chosen as the *Penumbra–Photosphere Boundary* of the *Minimum Intensity Spot*.

4.7. DETERMINATION OF THE *Umbra–Penumbra Boundary* OF THE SPOTS

The *Umbra–Penumbra Boundary* of a *Local Minimum* was defined as the contour with maximum AGAC among the contours of the *Local Minimum*. But this definition can only be taken as a preliminary definition of the *Umbra–Penumbra Boundary* and needs some refinement. If the *Umbra–Penumbra Boundary* of a *Local Minimum* (denoted by A) contains the *Umbra–Penumbra Boundary* of another *Local Minimum* (denoted by B) then this means that (at the level of the intensity of the *Umbra–Penumbra Boundary* of *Local Minimum* A) *Local Minimum* B and *Local Minimum* A were the same spot, but after the separation *Local Minimum* B had sharper border than *Local Minimum* A. In this case the preliminary *Umbra–Penumbra Boundary* is taken as *Umbra–Penumbra Boundary* for *Local Minimum* B and the separation contour of *Local Minimum* A with respect to *Local Minimum* B is taken as the *Umbra–Penumbra Boundary* for *Local Minimum* A. If the preliminary *Umbra–Penumbra Boundary* of *Local Minimum* A contains the preliminary *Umbra–Penumbra Boundary* of several *Local Minima* then the innermost relative separation contour of *Local Minimum* A is taken as *Umbra–Penumbra Boundary* for *Local Minimum* A.

Another modification of the preliminary *Umbra–Penumbra Boundary* of a *Local Minimum* is needed when the intensity level of the *Local Minimum* preliminary *Umbra–Penumbra Boundary* is higher then the level prescribed for a *Local Minimum* to be an umbra. According to the paper of Beck and Chapman (1993) the umbral boundary intensity is about 65 percent of the intensity of the photosphere. To use this information about the intensity of the umbra we must try to transform our CCD intensity into real intensity on the Sun. The intensity of a point on the Sun's surface becomes the intensity on the CCD image due to the following processes:

First, the light coming from the Sun's surface causes silver deposition on the film. The connection between the intensity i of the light and the density ρ of the silver deposition (in the linear portion of the $D - \log E$ curve of the film) is (Andrews and Hunt, 1977)

$$k\rho = \gamma \ln it + D_0 \,, \tag{1}$$

where γ is the gamma of the film, t is time of the exposure, and k and D_0 are constant parameters.

Then, the Sun's film image is transilluminated. The connection between the incident I_i and the transmitted I_t intensity of the light is

$$I_t = I_i \exp(-k\rho) \,. \tag{2}$$

Last, the transmitted intensity is converted to a CCD intensity by the CCD camera. In the case of linear connection between the transmitted intensity and the CCD intensity x the transformation between them is

$$I_t = ax + b \,, \tag{3}$$

where a and b are constants specified for the CCD camera.

Using relations (1), (2), and (3), the connection between the relative intensities of the photosphere on the CCD image and on the Sun can be written

$$x/x_p = (1 - x_0/x_p)(i/i_p)^\gamma + x_0/x_p \,, \tag{4}$$

where x_p is photospheric intensity on the CCD image, i_p is the intensity of the photosphere on the Sun and $x_0 = -b/a$.

There are two problems with using this formula to determine the maximum intensity level acceptable for an umbra. The first one is practical: we do not know the exact value of gamma because it changes from photoheliogram to photoheliogram depending markedly on the conditions during development (Kitchin, 1984).

The second one is conceptual: according to our practice the *Umbra–Penumbra Boundaries* of the pixels in a sunspot group can have different intensity levels. But, fortunately, we only need the maximum intensity level for an umbra to decide whether a spot can be an umbra (the boundary of the umbra is not based on intensities but on gradients). Thus, we determine a preliminary maximum intensity level for umbrae using an average gamma and the relative intensity (65% of the photospheric intensity) of the umbra boundary on the Sun, and afterwards this value is increased by an amount obtained from experience to get the maximum acceptable umbral intensity level. Using this method of prescribing the maximum umbral level, sometimes, it can occur that a penumbral spot is regarded as an umbra.

If the intensity of the preliminary *Umbra–Penumbra Boundary* of a *Local Minimum* is higher than the maximum prescribed umbral intensity level, then we go back on the contours of the *Local Minimum* until the prescribed maximum intensity level for the umbra is reached and the contour obtained this way will be the *Umbra–Penumbra Boundary* of the *Local Minimum*. If the prescribed maximum intensity level for the umbra can not be reached, the *Local Minimum* is taken as a penumbral spot with zero umbral area.

4.8. MERGING *Local Minima* HAVING THE SAME *Umbra–Penumbra Boundary* INTO ONE UMBRA

During the life of a sunspot group an umbra can break up into several umbrae or sometimes nearby umbrae can merge into one umbra. How does the program manage this process? In our approaches the umbra–penumbra boundary is a line with maximum gradient value. So, if two (or more) *Local Minima* have the same contour as an *Umbra–Penumbra Boundary* this means that they have already merged or they have not yet separated. Following this line of argument, *Local Minima* having the same contour as the umbra–penumbra boundary are merged into one umbra. But the merging is only done from the areal viewpoint (the merged *Local Minima* have a common umbral area determined by the common boundary). From the viewpoint of the heliographic coordinate, the *Local Minima* comprising an umbra are handled separately; their centers of gravity, from which their heliographic coordinates are ascertained, are determined by using their separation contours.

In our interpretation, the term spot means a local minimum (a pixel representing the center of gravity of a sunspot) retained after different filtering mechanisms in a sunspot group. A spot can be a penumbral one if its intensity is not low enough or can be an umbral one if it is inside an umbra. A spot is specified by three data: its position, the umbra that contains it and the penumbra that contains it. The hierarchy of notations spot, umbra and penumbra is as follows: a penumbra contains umbrae and an umbra contains spots.

Table I

Area differences between other and Gyula measurements in percentage of the Gyula measurement and in units of 10^{-6} solar hemisphere (m.h.) for sunspot groups in NOAA regions 5521 and 5528

Observatory	Mean		Stand. dev.	
	%	m.h.	%	m.h.
Spot area (penumbra+umbra				
Tenerife	6	−93	27	358
Debrecen	6	26	13	274
SGD	4	−119	40	278
Umbral area				
Tenerife	101	141	68	70
Debrecen	−11	−27	28	43

4.9. IMPROVEMENT OF THE BOUNDARY OF THE UMBRA AND THE PENUMBRA

According to the definitions given earlier, the boundaries (*Penumbra–Photosphere Boundary* and *Umbra–Penumbra Boundary*) are iso-density lines. There are large sunspot groups with very complicated inner structure. It can occur with this kind of sunspot that different parts of the penumbral border of the spot have different intensities. The seeing can also give rise to changes in the intensity of the boundary of a spot. To make allowances for these, the boundary obtained above is divided into intervals and the average gradient along an interval is compared to the values of the average gradient along similar intervals of the neighboring intensity contours of the boundary and the interval of the boundary is replaced by the one having the maximum average gradient value.

5. Results

To assess the reliability of the method proposed here, areas for the sunspot groups in NOAA region 5521 and 5528 were determined and compared with the results from Tenerife (Steinegger *et al.*, 1996), from Debrecen (Kálmán and Gerlei, 1994) and from *Solar Geophysical Data* Bulletin (*SGD*). Following Steinegger *et al.* (1996) we used the largest SGD value from different observatories per day and sunspot groups. The details of this comparison can be seen in Figures 8–11. As the scatter in the area measurements for sunspots depends on the area itself, it is expedient to use the percentage difference in the area measurements for statistical purposes (i.e. the area differences between Gyula and other measurements in percentage of the Gyula

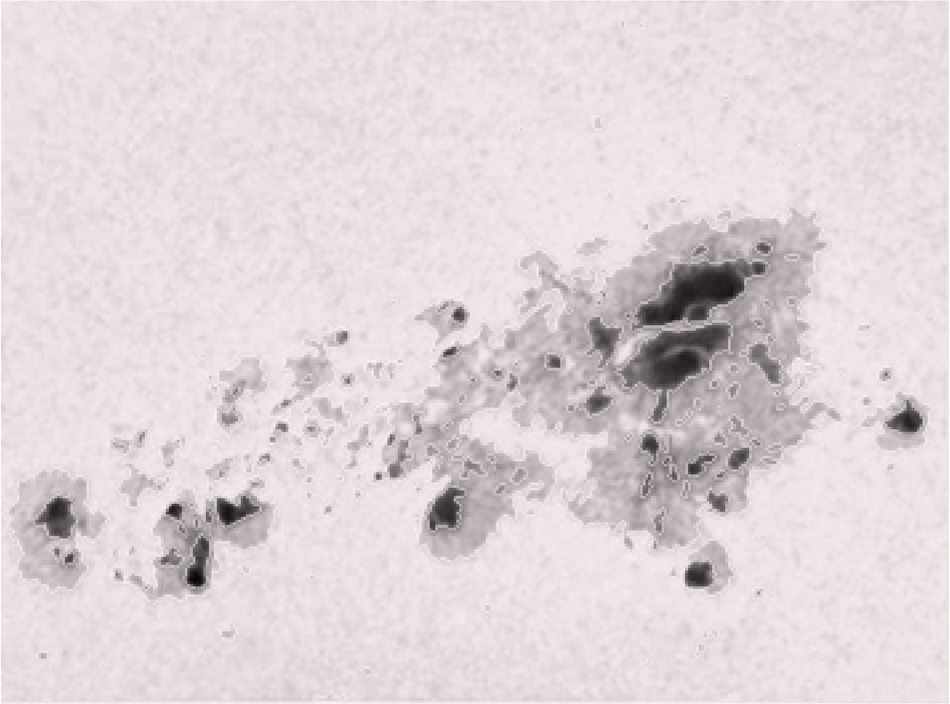

Figure 14. Border contours (umbral and penumbral) obtained by the method suggested in this paper are superimposed on the image of sunspot group (NOAA 5528, 15 June 1989, 07:43:03 UT).

measurement are used here). Table I provides the summarizing statistics which express the main aspect of the comparison of the different spot area measurements. No definite differences were found in the mean of the percentage differences of the spot areas but the scatter is significant especially in the case of the SGD.

The umbral areas obtained by Steinegger *et al.* (1996) are about 100% larger than those measured by us. In order to clarify this discrepancy we followed the method described in their paper. For this purpose we used the sunspot group in NOAA region 5528 (15 June 1989). First we determined the cumulative frequency diagram for the sunspot group (see Figure 12). The lines fitted to the umbral and the penumbral parts of the diagram intersect at intensity level 107. In Figure 13 the contours corresponding to intensity level 107 are superimposed on the sunspot group. Figure 14 shows the border contours (umbral and penumbral) obtained by the method proposed by us. Figures 13 and 14 are printed versions of the images on the computer screen and after printing the quality of the image has changed for the worse. For this reason and to it make easier to judge the goodness of the fit of the contours to the borders of the spots, two photographic copies (one of them is underexposed to make the umbra structure more visible) of the original photoheliograms are provided (Figures 5 and 6). Comparing Figures 13 and 14 we can see that the umbral contours obtained using the method suggested by Steinegger *et al.* (1996) embrace a significantly larger area than the contours obtained by the method proposed in this paper. The umbral area determined by the contours of Figure 13 is 48% larger then the one obtained by the contours of Figure 14. This value is very near to the 51% which is the percentage difference between the Gyula and Tenerife measurements for the umbral area for this day (see Figure 11). A similar result was obtained for sunspot group NOAA region 5528 on 13 June 1989. Using the method suggested by Steinegger *et al.* (1996) we obtained 62% larger area than with the method proposed by us here. The difference between Gyula and Tenerife measurements for umbral area for this day is 67% (see Figure 11).

Using the umbral contours in Figure 14, the following statistics can be established for their intensities: Their average intensity is 120, the standard deviation is 19 and their intensities change between 101 and 160. These statistics question whether a unique cutoff value can be adopted for every umbra. In their paper Steinegger *et al.* (1996) concluded: ‘Obviously, at some of the observatories certain parts of the umbra are identified as penumbra.’ In the light of the above-mentioned facts, we think that the opposite statement is valid, that is, in their measurements for umbral area, certain parts of the penumbra are identified as umbra.

6. Discussion

Using the concepts presented above, computer programs were written to single out spots inside a sunspot group and to determine their positions and areas. The programs can use CCD or scanned images of sunspot groups. The proposed method has the advantage over other methods (Dezső *et al.*, 1987; De La Rue, 1869) used for this purpose that it gets rid of the human involvement in the measuring process so it may result in better homogeneity and accuracy of the database compiled using this method. It is another advantage, especially near the limb, that before determining spot areas a limb correction is applied. It can also be mentioned as a benefit that taking into account the bright regions of the penumbra, a more accurate spot area can be determined.

Acknowledgements

This work was supported by the Hungarian Foundation for Scientific Research under grant No. OTKA 007422, and by the Hungarian–American Joint Fund for Science and Technology under contract No. 95a-524.

References

Andrews, H. C. and Hunt, B. R.: 1977, *Digital Image Restoration*, Prentice-Hall, Inc., Englewood Cliffs, New Jersey.

Beck, J. G. and Chapman, G. A.: 1993, *Solar Phys.* **146**, 49.
Brandt, P. N., Schmidt, W., and Steinegger, M.: 1990, *Solar Phys.* 129, 191.
Chapman, G. A. and Groisman, G.: 1984, *Solar Phys.* **91**, 45.
Chapman, G. A., Herzog, A. D., Laico, D. E., Lawrence, J. K., and Templer, M. S.: 1989, *Astrophys. J.* **343**, 547.
De La Rue, W.: 1869, *Phil. Trans. Roy. Soc.*, **159**.
Dezső L., Gerlei O., and Kovács Á.: 1987, Debrecen Photoheliographic Results for the Year 1977.
Frieden, B. R.: 1979, in T. S. Huang (ed.), *Picture Processing and Digital Filtering*, Springer-Verlag, Berlin.
Kálmán, B. and Gerlei, O.: 1994, private communication.
Kitchin, C. R.: 1984, *Astrophysical Techniques*, Adam Hilger Ltd, Bristol.
Steinegger, M., Vázquez, M., Bonet, J. A., and Brandt, P. N.: 1996, *Astrophys. J.* **461**, 478.
Tennenbaum, J. M., Kay, A. C., Binford, T., Falk, Feldman, J., Grape G., and Sobel, I.: 1969, in D. A. Walker and L. M. Norton (eds.), ‘The Stanford Hand-Eye Project’, *Proc. IJCAI*, p. 521.