LATEST DEVELOPMENTS AND RESULTS IN AUTOMATIC SCF COUNTING PART I: DEVELOPMENT OF ALGORITHMS AND PRELIMINARY RESULTS

Michel Giner
Unité de Recherche Biométrie et Informatique, Centre de Coopération Internationale en Recherche Agronomique pour le Développement, département des Cultures Annuelles (CIRAD CA), Montpellier, France
Jean-Paul Gourlot
Laboratoire de Technologie Cotonnière, CIRAD CA, Montpellier, France
Bruno Bachelier
Institut de Recherche Agronomique, Maroua, Cameroon
Ehoud Ahronowitz, Marc Hugon and Guillaume Damiand Laboratoire d'Informatique, Robotique et de Microélectronique (LIRMM), Université des Sciences et Techniques du Languedoc, Montpellier, France

Abstract

Different image-processing methods were used to count seed coat fragments (SCF) in cotton fiber webs. The results obtained showed that, similar to SCF detection on yarn, it is possible to use this technique to predict the "SCF potential" of lines during varietal improvement programs. The last method was tested on 440 images derived from carded cottons of various geographic origins. Counts obtained automatically by the analysis of images acquired by a camera were similar to SCF counts made visually by an "expert"; the relationship between the two counts being:
square root of the automated count = 0.867 * square root of the visual count + 0.393 (r = 0.957).

Introduction

Seed Coat Fragments (SCF) are torn off the seeds when the cotton fiber is separated from the seeds by ginning. Cleaning steps (lint-cleaner, card) do not entirely remove these fragments as some are still attached to fibers. Various studies have shown that the size of the SCF tends to decrease during cleaning operations and that their total mass also decreases [Baldwin *et al*. 1995, Mangialardi, 1988, 1992].

Ring spinning does not involve any cleaning steps after carding except for the cleaning that occurs during the combing phase. The SCF thus remain in the fiber up to the spinning phase and here cause breakages or result in the production of yarn with many imperfections. In open end spinning some of the SCF are removed by the opener when the fibers are separated. The SCF that remain cause yarn breakages or at least yarn imperfections that may or may not be removed by the yarn cleaning operations [Frydrych, 1989; Schenek, 1992; Jones, 1996; Aubry and Renner, 1996].

Productivity is reduced in both cases and the number of yarn imperfections increases. Literature data shows that SCF make up a large proportion of foreign particles in the fiber, i.e. about 27 % of yarn defects (Anthony *et al*., 1988; Hebert and Thibodeaux, 1989) and 17 % of those in fabric (Mangialardi, 1992).

SCF, corresponding to fiber impurities that cause problems during high-speed spinning, are responsible for many complaints. The different organizations involved in the spinning industry are now convinced that productivity can only be improved if SCF are reduced or eliminated (Curran, 1992).

Several SCF detection methods have been developed, e.g. visual counting of the number of SCF in a mass of brightly illuminated fiber waste separated on a Shirley Analyzer (ASTM D2812-76), and SCF count under high magnification (ASTM D2496). Using these methods, Mangialardi (1988) demonstrated that the count was greatly influenced by cotton variety and that the mechanical processing techniques applied to the fiber during ginning also had a profound effect on the number of SCF.

At CIRAD, studies performed in the context of multisite varietal trials used a Zellwegger Uster regularimeter evaluating Ring Spinning 20 tex yarn to provide a detailed analysis of imperfections (Frydrych 1989, 1992). Using these analyses, Frydrych showed that it is possible to demonstrate varietal and geographic differences by the SCF content of yarns analyzed. Subsequently, Gourlot *et al* (1995) showed that the heritability of this SCF count is high (about 76 %). Using such information, this method is therefore able to characterize the SCF potential of production lines and thus provides useful information for breeders during varietal improvement programs.

However, these programs require a large number of analyses to evaluate the genetic variability of the populations from which the low-impurity lines are selected. Manufacturing yarn and checking the nature of each imperfection detected by the regularimeter is effective but economically unviable. This led us to develop the Trashcam procedure.

Method used

As a first step in the procedure we mix about 10 grams of cotton on our laboratory opener. The fiber web is rolled around a cylinder on leaving the opener. This web is cut into four parts that are placed one on top of the other on entering our laboratory card (Shirley Platt). Four layers of

Reprinted from the *Proceedings of the Beltwide Cotton Conference*
Volume 2:1633-1637 (1997)
National Cotton Council, Memphis TN

fiber web are rolled around the reception cylinder (circumference about 0.70 m, width about 0.20 m) to form the test web (about 18 g/m5). The hypothesis was put forward that only the SCF remain on the fiber web after carding. This web is then pressed by a plate against a sheet of glass through which a camera is pointed to acquire an image. Because the web is thin, and the contrast between the SCF (dark color) and the background (fibers) is optimal, a white screen is placed between the pressure plate and the fiber web. Two or three images are taken for each web then analyzed. During development we routinely took 3 pictures of each web and prepared 4 webs for each sample.

The computerized representation of the image acquired by the camera takes the form of a succession of coloured pixels arranged in lines and columns (scanning combined with digitalization at regular intervals). Thus, the image obtained is made up of a set of pixels, each with a set of coordinates (x: position on the line, y: line number) and a color level. This color is generally specified using the RGB (Red Green Blue) system or using a gray scale that ranges from white (value 255) to black (value 0) in 256 grey levels. The image is analyzed by evaluating the color at each set of pixel coordinates. Using this concept, the differences between the coordinates and the pixel counts gives a sizing and surface area evaluation, i.e. morphological information.

Development process

Global threshold approach

A Trashmeter installed on an HVI measuring line (Motion Control Incorporated) was used to acquire the images. Two lamps, placed on either side of the camera's field of vision (images composed of 16 gray levels), illuminated the cotton pressed onto the glass plate.

The SCF showed a darker gray level than the background. A simple threshold was therefore sufficient to discriminate the objects by isolating the darker spots.

The images analyzed were of too poor quality to permit the detection of SCF. Thus the threshold selected isolated dark zones that instead of corresponding to SCF were in fact patches caused by lighting irregularities.

This showed that if the illumination of the surface acquired by the camera is irregular, the gray levels change depending on the position of the lamp. The threshold used to digitalize the images, and including the darkest areas (produced by the SCF or lighting defects) was calculated in the following manner:

- all the pixels with gray levels above the threshold (whiter) were considered as background, with a value of 255,
- all the others were attributed a value of 0 and were considered as SCF.

This biased threshold was unable to detect the SCF over the entire surface of the image as the SCF count and sizing information were incorrect in comparison with the results given by visual counts of the fiber web (Figure 1).

Images acquired in gray levels by a surveillance camera (Hymatom) fitted with 12 mm lens (Cosmicar) were used to count SCF by 3 other different approaches with images composed of 256 gray levels.

Pre-treatment followed by global threshold

To remove the bias introduced by sample lighting, one image on a white background was acquired every x images and from this were subtracted the images of the fiber web, pixel by pixel (difference between the two matrices). This method reduced the variations observed in background intensity and rendered the SCF count comparable regardless of the pixel's position in the image (Figure 2).

The threshold value was selected after studying diagrams depicting the cumulated frequency of pixel gray levels in a large number of images. It was noted that the percentage of pixels with a dark color (possibly corresponding to SCF) was very low, at about 0.036 % (Figure 3).

The gray level corresponding to this percentage was selected as the threshold. The images were now digitalized to detect the SCF. The number of SCF were then counted and morphological information (position, size, surface area, etc) concerning the SCF was stored.

This method relies on a stable background for the different images but may systematically count interfering objects. The threshold was set at 0.036 % of dark pixels on the gray level frequency diagram. However, the percentage of dark pixels in the image is proportional to the total surface area covered by the SCF. Consequently, this threshold does not detect SCF in an equivalent manner in fiber webs containing different SCF contents.

Localized threshold approach (V2)

The threshold methods described above were seen to be of limited use as they were excessively rigid. They consider, wrongly, that contrasts between gray levels are comparable at all points in the image. A different approach was therefore employed to take account of local variations in the gray levels of the images.

A gray level curve, corresponding to each y line in the image, may be produced by describing the position (x) of the pixel on that ligne : Level = $f_y(x)$. The derivative of this function at a given point evaluates the rate at which the gray level varies at this point. As it is unable to describe the basic function, the procedure used compares the gray levels of 2 opposite pixels on a scale of 5 pixels to analyze changes in the difference between the gray levels (Figure 4). If this difference exceeds a certain reference value (relative threshold), this triggers the process used to create a new object.

SCF are detected by analyzing the image line by line. Generally, each SCF generally covers several lines on the image. If the algorithm detects part of a SCF on a particular line, it joins this to the part detected on previous lines on condition that a topological relationship is established. If this is not the case, a new object is created. A local threshold is therefore used (as opposed to a global threshold for the entire image).

We compared the number of SCF detected by Trashcam using the localized threshold technique with the distribution of the different types of neps identified on the regularimeter for yarns of different sizes produced by Carded Ring Spinning (Table 1).

The results obtained showed that the Trashcam method was sufficiently accurate for use in improvement programs as it was possible, like in the method using counts produced by the regularimeter, to separate low SCF potential lines from the others.

These promising results encouraged Bachelier to use a Trashcam in Cameroon during a varietal improvement program. Bachelier demonstrated that the number of SCF in the progeny of a given cross is very variable. The good correlation, established on 5 crosses, between the average counts obtained on each F2 generation and on each F3 generation (R^2 = 94.7 %) means that efficient selection may be performed on the basis of this criterion.

The correlations revealed, though already very good (Table 1), can be improved. Illumination of the fiber web causes considerable disparities in the contrast between SCF and the fiber web. The SCF detected generally show gray levels very different from those of the fiber. Comparisons with visual analyses of the web showed that certain SCF were covered with fibers or oriented with their lightest side outward. In these cases, the difference between the gray levels was insufficient for this type of image analysis to detect the SCF.

An approach based on the pixel in its environment (V3)
The preliminary work conducted by Charnier (1995) then by Kamel (1995) in the *Laboratoire d'Informatique, de Robotique et de Microélectronique* in Montpellier, showed that their algorithmic approach could be applied to the detection of SCF in a cotton web. Their purely algorithmic approach provides a wealth of information concerning the image by describing the characteristics of each pixel and the area in which it is situated. However, as the information collected is excessively voluminous, we continued the work underway in the context of our collaboration with M. Hugon and G. Damiand, in parallel with the work conducted by Fiorio (1995). This work resulted in the production of a third algorithm, called Trashcam V3. This more sensitive software was able to take account of the gray level shifts noted around the SCF and could therefore be used to provide a better description of the frontier between background and SCF.

The process therefore consisted of analyzing the image line by line and pixel by pixel. Neighboring pixels were then compared as regards their gray levels. When this difference exceeded a set value (the sensitivity), it can be assumed that the border of a SCF has been detected.

This analysis was not always able to completely close the borders of a SCF because the gradual shift in the different gray levels was sometimes below the sensitivity value (e.g. in the case of SCF covered with fibers) (Figure 5).

To complete the description and close the borders of the objects, a concept of "chance" was developed specially to manage these gradual shifts in gray levels. The hypothesis was made that the bigger the object, the higher the probability that gradual shifts will be found, and the greater the risk of impurities being left undetected. This risk was compensated by the "chances" of not prematurely losing impurities that show pronounced gradual shift. The number of "chances" increases in proportion to the size of the object. These "chances" are used to describe the final outline of the object by attributing pixels that would have been discarded if sensitivity thresholds alone were employed for the analysis.

This pixel analysis mean that powerful tests, unusable in previous development work, could now be employed. Thus, the objects are only considered to be SCF if they correspond to the criteria set by a series of about 30 parameters.

The detection of new defects, not revealed by the previous version, is thus possible by improving the evaluation of the surface area of all the SCF.

Results

These parameters were set from the results obtained using about 20 images. The algorithm described in this way was validated for 440 digitalized images of fiber webs acquired by the camera. The counts obtained using the "localized threshold approach" and the "pixel in its environment" approach were compared with visual counts made of images on the screen. Square root transformation was used to meet linear regression conditions (normal, independent variables and homogeneous variances).

First comparison: Relationship between 2 visual counts
By visualizing each of the 440 images twice on a computer screen, an "expert" counted the black points reputed to correspond to SCF. Figure 6 illustrates the relationship between these two counts. All confidence intervals in this document were calculated with a type I risk of $\alpha = 5\%$, and it may be concluded that:

- an excellent relationship was obtained: square root of count2 = 0.94 (±0.028) * square root of count1 + 0.37 (±0.143), giving a correlation of 0.953. This relationship is therefore difficult to improve as the images were not sufficiently contrasted to systematically reveal impurities in the same manner:

~ over time: the operator may suffer of eye fatigue,
~ the probability to leave impurities increases with the number of impurities to be counted,
~ cotton showing few impurities tended the Aexpert@ to over-evaluate local contrasts which will be counted as SCF.

– for the same reasons, the correlation between automated counting and visual counting is unlikely to exceed that shown by two visual counts.

- the 2 visual counts were similar, and their mean was used as the reference when comparing the different versions of the algorithm.

Second comparison: "localized threshold (V2)" approach / mean visual count

Figure 7 shows a close relationship between the two variables (r=0.924), giving the following relation : square root of V2 count = 0.76 (±0.029) * square root of mean visual count - 0.298 (±0.151). However, the slope of the regression was significantly different from 1. This indicates that the number of SCF missed by the automated counting system increases with the total number of SCF in the image. The pixel by pixel analysis of spots considered to be SCF shows that the algorithm malfunctions in cases where there is little contrast between the SCF and the background.

Third comparison: "pixel in its environment (V3)" approach / mean visual count

Figure 8 shows a close correlation between the two variables (r=0.957), giving the following relation: square root of V3 count = 0.867 (±0.024) * square root of mean visual count + 0.393 (±0.125).

There was a statistically significant difference between this slope and that observed between V2 and the visual count. This slope showed that better account had been taken of the gradual shifts though without reaching equality. The expert considered a spot that was darker than the background to be a SCF, even if the border could not be clearly seen (cases where fibers were covering the SCF). To detect this spot the algorithm had to work pixel by pixel, building up the edges from sufficiently marked contrasts in the gray levels. If the spot detected by the algorithm did not show a closed border, it was considered as background, and was not counted as a SCF.

Interface

Trashcam V3 runs in the Windows 3.1 environment. It includes the management of an image acquisition system, analysis of the data with algorithmic processing, and expresses the results in the form of SCF surface area distribution histograms. The 3 files produced (global information, detailed information by SCF and data concerning the size distribution histograms) take the form of text files that can be imported into other software packages. We supplemented the interface with various function that render the Trashcam image analysis application easy and rapid to use. For instance, one particular option automatically labels the samples analyzed, regardless of the number of individual measurements chosen.

The image obtained may be saved in BMP format to be processed later. It is also possible to batch analyze several images with a single command by selecting "Open several ..." in the menu. Here, Trashcam can be used without the image acquisition system.

Conclusion

The most recent version of Trashcam includes a user-friendly interface that offers functions designed to facilitate its use. A new algorithm gives a more precise evaluation (size and number) of the SCF in the images taken of card webs acquired by a camera.

Despite a certain degree of visual counting inaccuracy due to insufficiently contrasted images and eye fatigue experienced by the "expert", ..., we noted that - in accordance with our objectives - the count provided by the V3 system was closer than the V2 system to the visual count. In fact, the improvements made in the management of gradual shifts and in sensitivity allow the V3 algorithm to provide a better definition of impurity borders and thus furnish a more accurate evaluation of SCF surface areas.

Thus, the Trashcam system can evaluate the SCF potential of a sample and avoid the necessity of manufacturing yarn. This leads to 4 advantages:

– investment in equipment is considerably reduced,
- the fiber mass required for the test can be reduced to 12 grams,
- the time required to perform the test (including sample preparation) is reduced to 15 min per sample, compared with 4 hours,
- the SCF potential of new lines can be evaluated at the preliminary stages of varietal selection and in a large number of samples.

However, we consider that the quality of the image should be improved to enhance the contrast between SCF and the background. In this way the algorithm used to manage the gradual shifts in the image will be used less often to define impurity borders, and less SCF will be classified as background.

Bibliography

Anthony S., Meredith W.R., Williford J.R., Mangialardi G.J., 1988, "Seed coat fragments in ginned lint : the effect of varieties, harvesting and ginning practices", Textile Research Journal 58 (2), pp. 111-116.

American Society for Testing and Materials, 1977, "Non-lint content of cotton", Part 33, D2812-76, pp. 576-583.

American Society for Testing and Materials, 1977, "Seed coat fragments and *finiculi* in cotton fiber samples", Part 33, D2496-75, pp. 491-496.

Aubry E. et Renner M., 1996, "Le point sur la filature B rotor", LéIndustrie Textile, n° 1280, Octobre 1996, pp. 28-34.

Bachelier B., 1996, "Variabilité de la teneur en SCF", présentation lors des 7° journées Coton, CIRAD, Montpellier, France.

Baldwin J. C., Qaud M. and Schleth A., 1995, "AFIS Seed Coat Nep measurement", Proceedings of the Beltwide Cotton Conference, Memphis, pp. 1250-1253.

Kamel H., 1995, "Analyse d'images à la volée", Mémoire de Diplôme d'Etude Approfondie en Informatique, soutenance le 27 juin 1995, Université Sciences et Techniques du Languedoc, Montpellier 2.

Charnier P., 1995, "Outils algorithmiques pour le codage interpixel et ses applications", thèse de Sciences pour l'Ingénieur, spécialité informatique, Université des Sciences et Techniques du Languedoc, Montpellier, France, ref. 95MON20002, 165 pages.

Curran J.M., 1992, "What are the qualitative needs of the modern spinning industry?", Technical Seminar at the 51st Plenary Meeting of the International Cotton Advisory Committee, Cotton Marketing Systems and Quality Evaluation, Liverpool, September 1992, pp. 3-6.

Fiorio C., 1995, "Approche interpixel en analyse d'images, une topologie et des algorithmes de segmentation", thèse de Sciences pour l'Ingénieur, spécialité informatique, Université des Sciences et Techniques du Languedoc, Montpellier, France, ref. 95MON20179, 201pages.

Frydrych R., Gutknecht J. ,1989, "Identification et comptage des diverses imperfections rencontrées sur le fil de coton", Coton et Fibres Tropicales 44 (1) pp. 59-65.

Frydrych R., 1992, "Evaluation du nombre d'imperfections sur fil au régularimPtre Uster GGP et plaquette de fil, ainsi que sur voile de carde", Présentation lors des 3° journées Coton, CIRAD, Montpellier, France.

Gourlot J-P., Frydrych R., Hequet E., Thollard F., Constantin O., Bachelier B., 1995, "Seed coat fragment counting and sizing in card web", Proceedings of the Beltwide Cotton Conference, San Antonio, pp. 1245-1249.

Hebert J.J. and Thibodeaux D.P. , 1989, "A study of seed coat fragment propensity in cotton cultivar", Proceedings of the Beltwide Cotton Conference, Memphis, pp. 110.

Jones P. C., "The influence of Seed Coat Neps in yarn manufacturing", 1996, Proceedings of the Beltwide Cotton Conference, Nashville, pp. 1454-1460.

Mangialardi G. J., 1988, "A review of Seed Coat Fragment research", Proceedings of the Beltwide Cotton Conference, New Orleans, pp. 592-594.

Mangialardi G. J., 1992, "Lint cleaning effect on Seed Coat Fragment size distribution in cotton", Textile Research Journal 62 (6), 335-340.

Schenek A. and Schleth A., 1992, "Electro-optical trash particle counting and sizing", International Cotton Conference Bremen, pp. 199 - 208.

Table 1: Illustration of the number of SCF detected on a Zellweger Uster UT3 regularimeter by Trashcam counting of different carded yarns (Gourlot *et al*., 1995)

Yarn type	Prediction	R^2
20 tex	SCFyarn = 11.43 * TrV2 - 62.94	83.9
27 tex	SCFyarn = 9.66 * TrV2 - 61.30	83.2
37 tex	SCFyarn = 7.35 * TrV2 - 39.99	78.9

SCFyarn : the number of SCF detected in the yarn by the R. Frydrych method (1989).
TrV2 : the number of SCF counted in the fiber web images using the Trashcam V2 method described by Gourlot *et al* (1995).

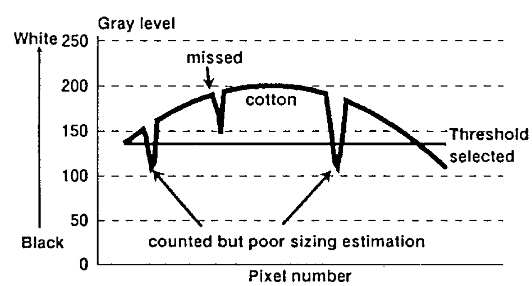

Figure 1: Typical gray level curve for a single line of the image.

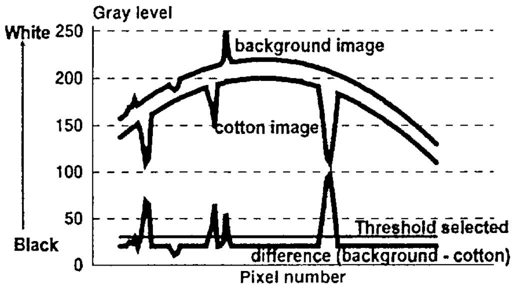

Figure 2: Typical gray level curve for a single line of the image and processing by pixel-pixel subtraction.

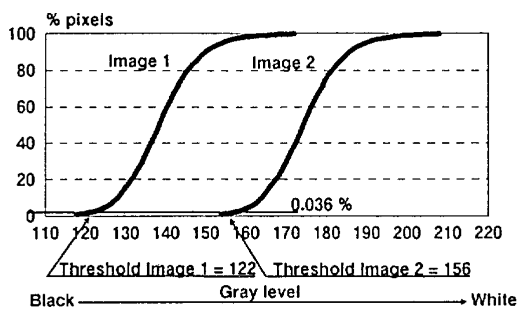

Figure 3: Threshold selection on cumulated frequency diagrams of the gray levels in 2 images.

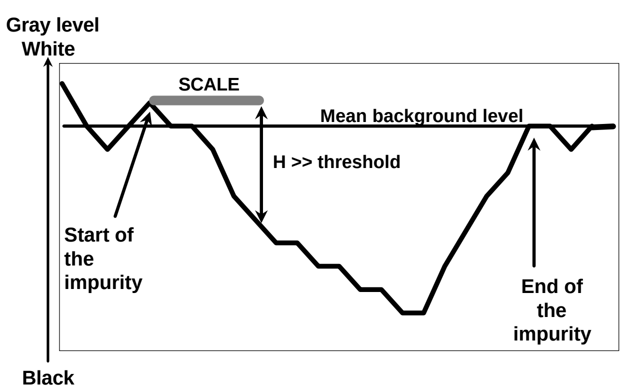

Figure 4: Detailled description of a line in the image showing a SCF, analyse with algorithm V2.

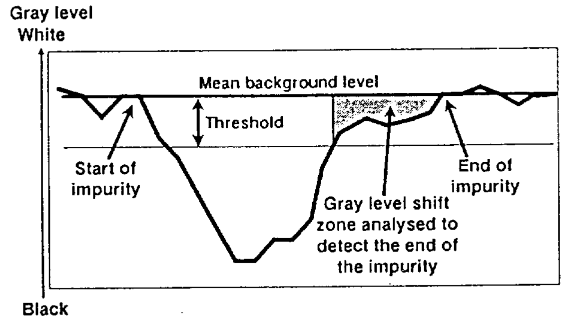

Figure 5: Detailed description of an image line: Processing of a gradual gray level shift.

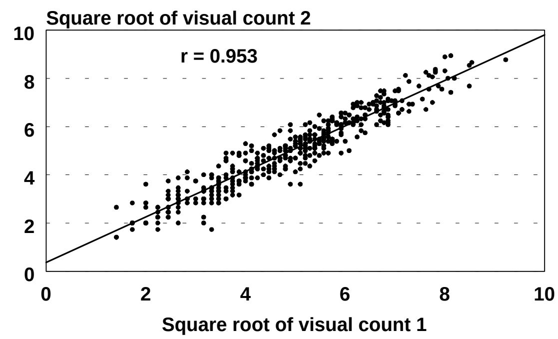

Figure 6: Correlation between square roots of 2 visual counts.

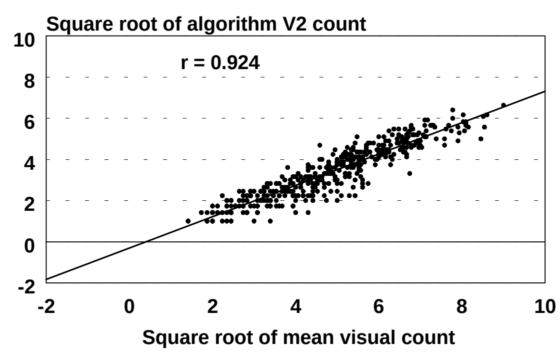

Figure 7 : Correlation between square root of the algorithm count using the "localized threshold" (V2) approach and square root of the mean values for the visual count.

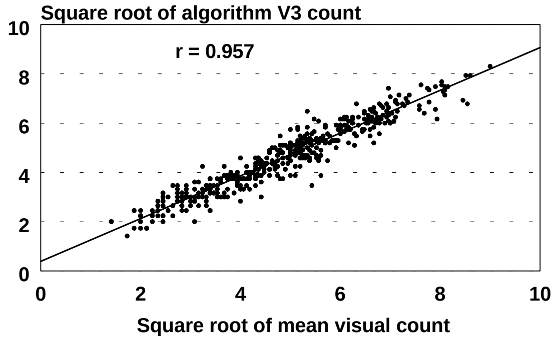

Figure 8: Comparison between square root of the algorithm count using the "pixel in its environment" (V3) approach and square root of the mean values for the visual count.