

QUANTIFICATION OF CRYSTAL CONGLOMERATES USING IMAGE ANALYSIS

D J LOVE¹, S D PEACOCK² and G T SCHUMANN³

¹*David Love Process Engineering cc, Durban, South Africa*

²*Tongaat-Hulett Sugar Limited, Durban, South Africa*

³*Formerly of Tongaat-Hulett Sugar Limited, Durban, South Africa (Retired)*

E-mail: Steve.Peacock@Huletts.co.za

Abstract

A regular crystal form, in particular the absence of conglomerates, is an important aspect of refined sugar quality because irregularities increase the quantity of mother liquor which remains on the surface of the crystals or is contained within occlusions in the crystals themselves. This aspect of sugar quality is of greater importance when processing lower purity liquors due to the increased impurity load of the mother liquor retained by the crystals in these circumstances.

The current technique for the quantification of conglomerates in sugar samples, namely the *crystal regularity index*, is based on the subjective classification of photographed crystals in terms of three or four groupings. Each of the groupings is given an arbitrary weighting in terms of crystal regularity and the total score for the sample is expressed as a percentage of the maximum score which could be obtained by a sample of crystals of perfectly regular shape. This technique is both slow and dependent on subjective classification by one or more observers.

An attempt has been made to develop an improved technique for conglomerate quantification, automating the procedure as far as possible and replacing the subjective element of the original analytical method by using computerised image analysis procedures. After the extensive investigation of a number of possible classification criteria, some preliminary methods have been obtained which are based on the analysis of crystal shape and the characteristics of the crystal edges.

Keywords: conglomerates, image analysis, sugar quality, refined sugar, raw sugar, crystal regularity index

Introduction

The production of crystals of regular form is an important aspect of refined sugar quality. The presence of conglomerates is particularly detrimental because such irregularities increase the quantity of mother liquor which remains on the surface of the crystals or is contained within occlusions in the crystals themselves. This aspect of sugar quality is even more important when processing lower purity liquors due to the increased impurity load of the mother liquor which is retained by the crystals under these circumstances. A high proportion of conglomerates in refined sugar also lowers its bulk density, leading to possible difficulties in packaging. While conglomerates form more readily at high liquor purities, they have been encountered in raw sugars, leading to products having high colour, high moisture content and poor filterability.

The current method used to quantify the extent of conglomeration in sugar samples is slow and dependent on a subjective classification carried out by one or more observers. An attempt was therefore made to develop an improved technique, using computerised image analysis, which would automate the procedure as far as possible and eliminate the subjective element of the original method.

Crystal regularity index

The concept of the *crystal regularity index* (CRI) as a means of quantifying the regularity of a sample of sugar crystals was originally developed by Hill (1965). The CRI, which may vary between 100% for a set of perfect crystals and 0% for a set of conglomerates, is obtained by classifying each of the individual crystals of a sample as belonging to one of three categories of crystal regularity (*good*, *moderate* or *conglomerated*). Each *good* crystal is awarded two points, each *moderate* crystal one point and *conglomerated* crystals are awarded no points. The CRI is calculated by expressing the points total obtained from the analysis of 200 crystals as a percentage of the maximum obtainable score. The CRI assessment process is described in more detail in Appendix 1.

The major disadvantages of the CRI technique are the time involved in the carrying out the tests and the subjective nature of the crystal classification. Even Hill (1965) recorded appreciable differences between the levels of assessment of four different observers when analysing the same photographs of a sample of refined sugar (up to 30% CRI). While the technique may be useful for ordering a set of sugar samples according to their degree of crystal regularity (provided the tests are carried out by the same observer), it provides a poor quantitative measure of the regularity of any particular sample.

A slightly modified version of the CRI measure was described by Plews (1970) in which four categories of crystal regularity are employed, carrying quality ratings of three, two, one and zero points. As for the original method, a set of photographic standards was provided to assist in the classification of individual crystals into the four categories of regularity (namely *good*, *moderate*, *irregular* and *conglomerated*). This modification to the original method obviously changes the maximum quality score which may be achieved by any sugar sample. However, it is doubtful whether any improvement in the technique could be obtained simply by changing the arbitrary scale which is employed in the assessment of the crystals.

While still employing the three categories of crystal regularity originally proposed by Hill (1965), Hibbert *et al.* (1975) modified the basic CRI method by using two independent observers to carry out the analysis and taking an average of the resulting scores. When increased levels of accuracy are required, the authors recommend calculation of the CRI based on the analysis on four sieved size fractions instead of the original two. While both of these modifications to the technique may lead to increased precision in terms of a quantitative measure of crystal regularity, they substantially increase the length of time required to carry out the measurements.

Image analysis

The current study aimed to eliminate the major disadvantages of the current CRI techniques by automating the procedure as far as possible, so as to reduce the time required for analysis, and by using computerised image analysis to remove the subjective nature of the existing methods.

There are two interacting aspects of computerised image analysis which must be considered when carrying out any project of this nature. The first involves the preparation of an acceptable image for further processing, while the second relates to the nature of the image analysis techniques used. Improving the quality of the image generated means that less sophisticated analysis techniques can be employed. Similarly, by making use of more advanced image analysis tools the requirements for image quality may be relaxed. The nature of the particular problem being investigated will determine whether it is more cost effective to develop improvements in image quality or to develop more advanced image analysis tools.

During the course of the current work, some slight modifications were made to the image generation technique used for conglomerate analysis. However, the main focus was on the development of analysis techniques, as it was felt that there was more scope for advances to be made in this area. In particular, the use of a wide range of public domain image analysis tools which have recently become available was investigated.

Image preparation

The technique used for image generation is broadly similar to that conventionally employed for conglomerate analysis, as described in Appendix 1. The major advance over the original method developed by Hill (1965) is the use of a digital scanner for image capture at a high resolution, as opposed to photographic enlargement.

In order to facilitate the process of discerning the crystals from their background during the computerised analysis of the image, the use of an alcohol-soluble dye for changing the colour of the sugar crystals was investigated. This procedure did marginally improve the contrast of the images obtained and it was used for the generation of all the images in the current study. However, it is doubtful whether the additional time required for the dyeing of the crystals could be justified on this basis when carrying out a routine crystal conglomerate analysis procedure.

Initial investigations

The analysis of sugar crystals for the presence of conglomerates was based on the use of public domain image analysis software called *Image J* (<http://rsb.info.nih.gov/ij/index.html>).

This powerful program processes a number of image formats and performs a number of basic functions, including (among others):

- carrying out standard image processing techniques (scaling, rotation, flipping, contrast enhancement, sharpening, smoothing, edge detection, filtering, etc.),
- calculating area, perimeter and pixel value statistics, and
- measuring distances and angles.

In addition, a number of public domain *Java* plug-ins for *Image J* are available to carry out image analysis techniques involving granulometry, texture, waviness and roughness, facet orientation, particle shape characteristics, fractal dimension, wavelet transformations and many others.

The initial investigations into conglomerate analysis were carried out by a student employed on a temporary basis by Tongaat-Hulett Sugar¹. A number of images, containing both crystals of regular form and conglomerates, were analysed using *Image J*. All of the image analysis tools available within the software were utilised in an attempt to develop a quantitative

¹ Samantha Hendricks (2003).

measure which could be used to identify conglomerates. Unfortunately, none of these techniques was found to be successful.

Of particular interest during this initial investigation period were the *shape factors* which can be calculated using *Image J*, such as the form factor:

$$\frac{4 \cdot \pi \cdot A}{P^2}$$

and the roundness:

$$\frac{4 \cdot A}{\pi \cdot D^2}$$

where A is the area of the crystal of interest, P is its perimeter length and D is the length of its major axis.

It may be expected that a highly conglomerated crystal would have a relatively large perimeter length in comparison to its surface area and the form factor was therefore considered to be a promising candidate to identify conglomeration. Similarly, it could be expected that a conglomerated crystal would be more elongated (i.e. have a relatively long major axis) than a regular sucrose crystal in relation to its surface area. However, in practice neither of these measures was found to be a reliable identifier of conglomerates. This is due to the fact that sugar crystals do not exist as ‘good’ and ‘bad’ grains, but cover a wide spectrum of regularity. In fact, as the regularity of a crystal decreases, it typically becomes more rounded and therefore actually ‘improves’ in terms of the form factor and roundness measures, before becoming worse again. These two measures are thus not good indicators of conglomeration.

Rectangle fitting

Following the failure of the built-in *Image J* techniques to reliably identify conglomerates, it was decided to investigate a number of customised image analysis methods, making use of the image processing capabilities of the MATLAB programming language. In order to carry out this task, the original scanned pictures were processed using *Image J* to yield good quality black and white images of the sugar crystals which were exported in a TIF file format suitable for importing into MATLAB. As part of the image generation process, *Image J* also produced a text file containing useful information about the crystals which could be used in the image analysis process (for example, the surface areas and perimeters of each crystal, the x- and y-co-ordinates of the crystal centroids and the x- and y-co-ordinates of the top left-hand corner of each crystal, etc).

The first customised image analysis technique which was investigated was the fitting of a rectangular ‘box’ around each of the crystals. While single well-formed sugar crystals are reasonably rectangular, conglomerates are typically quite irregular in shape. It was hoped that this difference could be exploited to quantify the presence of conglomerates in a sugar sample.

In carrying out the rectangle fitting process, the image file and the data file produced by *Image J* were read into MATLAB and converted to forms suitable for further processing. Each of the crystals in the image was then separately analysed. As an initial guess for the best-fit rectangle, a square of the same surface area as the crystal itself was positioned on the centroid of the crystal.

The MATLAB simplex optimisation algorithm was then used to find the rectangle which best fitted the crystal shape by manipulating the following five variables related to the positioning, size and rotation of the rectangle in relation to its crystal:

- the rectangle length,
- the rectangle width,
- the x-co-ordinate offset of the centre of the rectangle from the crystal centroid,
- the y-co-ordinate offset of the centre of the rectangle from the crystal centroid, and
- the angle of rotation of the rectangle from the vertical.

Simple linear geometry was employed to determine the position of the resulting rectangle on the image. The objective function to be minimised in the best-fit optimisation process was calculated by adding together the number of pixels within the rectangle which were blank (i.e. not occupied by the crystal) and the number of pixels outside the rectangle which belonged to the crystal. This objective function was standardised by dividing the resulting number by the total surface area (in pixels) of the crystal, so as not to prejudice larger crystals at the expense of smaller crystals.

The results of the rectangle fitting process are shown in Figure 1, which contains two sets of 30 crystals of refined sugar each. The first of these sets contains crystals of regular shape, while the second set contains conglomerates. The figure shows the picture after completion of the analysis process, with the best-fit rectangles overlaid in red.

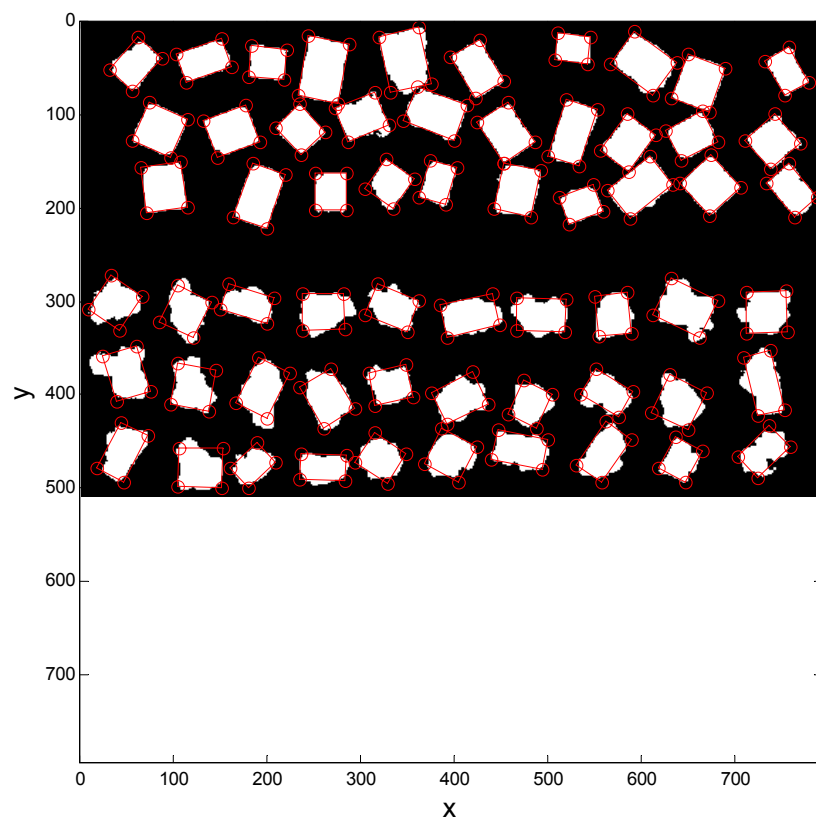


Figure 1. Fitted rectangles obtained for 60 refined sugar crystals.

Figure 2 contains a bar graph of the standardised objective function calculated for the 60 crystals. The difference in objective function values between the regular crystals and the conglomerates is readily apparent.

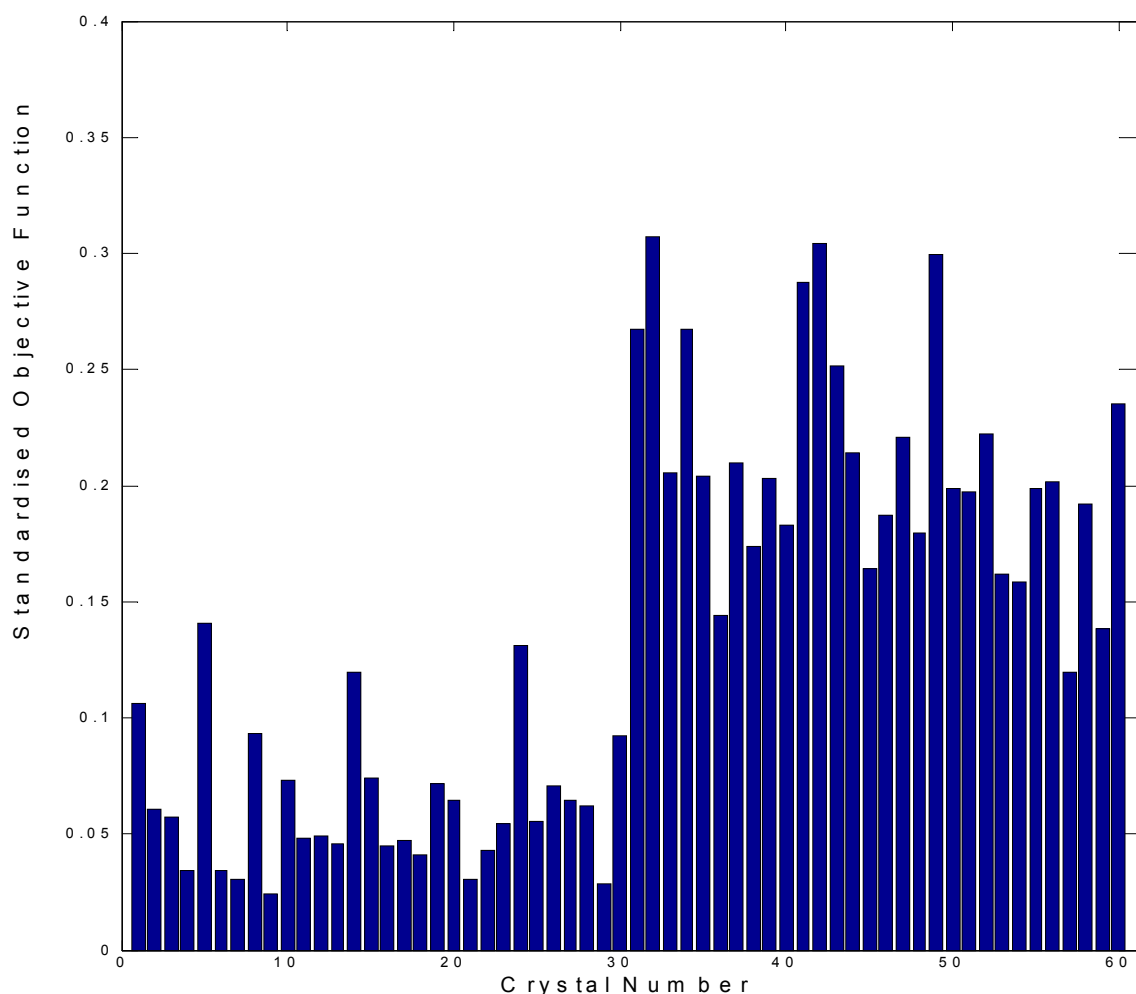


Figure 2. Results of the rectangle fitting process.

Based on the analysis of a number of refined sugar samples, a target value for the standardised objective function of 0.13 was selected. Typically, any crystals receiving an objective function value above this level were found to be conglomerated. Using this target level as a criterion, the number of conglomerates present in any sugar sample can be determined and expressed as a fraction of the total number of crystals present (a value of 52% was received for the sample shown in Figure 1).

Rectangle fitting was found to be a reasonably reliable measure of conglomeration for refined sugar samples. However, the method was found to be unsuitable for raw sugars due to the vagaries of the more elongated shape of the crystals. When creating a distributed set of crystals using a piece of sieve mesh and adhesive tape, the raw sugar crystals were found to have a tendency to enter the apertures vertically. As a result, the generated images show the raw crystals ‘end-on’, as shown in Figure 3. Because of the particular shape of raw sucrose crystals, even the well-formed raw grains have a non-rectangular shape in this orientation, making any differentiation between regular crystals and the conglomerates impossible.

Of all those investigated in the current study, the rectangle fitting technique was found to be the best method for identifying conglomerates in refined sugar samples. In order to make use of the technique for the analysis of raw sugar samples, however, great care would need to be taken to ensure that all of the crystals in the captured image lie in the correct orientation for the image analysis process. Further work with regard to image generation is required in order to ensure that this can be reliably achieved during routine operation.

In the hope of avoiding these crystal orientation problems with samples of raw sugar, a number of other image analysis techniques were investigated.

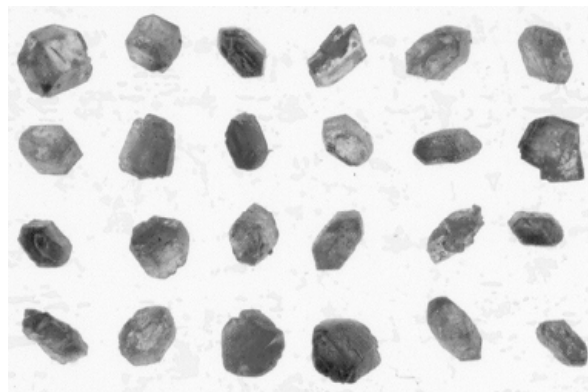


Figure 3. Distributed raw sugar crystals obtained using the sieve mesh technique.

Pixel intensity

By studying a large number of crystal images, it was found that crystals of regular form were usually quite uniform in colour, while conglomerates were more patchy and irregular in colouration due to factors such as the presence of mother liquor inclusions and variations in thickness. Using grayscale images, an attempt was thus made to study the variation of pixel intensity (i.e. the degree of 'whiteness' of the image pixels) within individual crystals as a means of identifying conglomerates.

The calculation of the standard deviation of the pixel intensity for all the pixels within a single crystal is a reasonably simple task. However, this statistic is not particularly useful for conglomerate identification because it only measures deviation from the mean intensity. Thus both of the hypothetical crystals in Figure 4 below would yield the same standard deviation because half of the pixels are white and half are black. However, a crystal with the colour characteristics of the hypothetical crystal on the right hand side of the image is far more likely to be conglomerated than one similar to the hypothetical crystal in the left of the image.

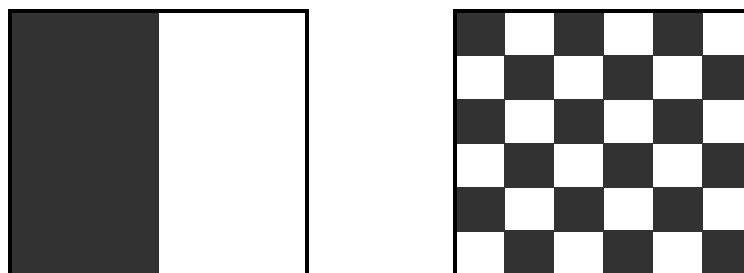


Figure 4. Hypothetical crystals with equal standard deviations.

In order to combat this deficiency, the calculation of the pixel intensity deviation for each crystal was based on a 'nearest-neighbour analysis'. The intensity of each of the interior pixels within an individual crystal was compared against the intensities of its eight nearest neighbours. The sum of the total squared deviations for all pixels in the crystal was then divided by the surface area of the crystal (to avoid prejudicing the larger crystals) and the square root of this number was then taken, to yield an average intensity deviation per pixel. So as not to disadvantage crystals which were lighter in shade over those which were darker, the average intensity deviation was divided by the mean pixel intensity for the crystal of interest, thereby yielding a measure similar to a coefficient of variation for the pixel intensity.

Unfortunately, however, tests carried out on a wide range of sugar samples showed the pixel intensity method to be unable to reliably discriminate between conglomerates and regular crystals.

Edge tracing

Typically, crystals of a regular form are characterised by a few straight sides of reasonable length, while conglomerates generally have a larger number of short edges. In addition, when tracing around the edge of a regular crystal, all of the corner angles usually rotate in the same direction (i.e. either clockwise or anti-clockwise). However, when tracing around the edge of conglomerate crystals, some of the angles rotate in the 'incorrect' direction (*re-entrant angles*). The characteristics of the crystal edges were thus investigated as a possible means of identifying the presence of conglomerates in a sugar sample.

Using the MATLAB programming language, an edge tracing algorithm was developed to follow the outer edge of each individual crystal one pixel at a time. At each step, the change in angle of the crystal edge was recorded. For example, for a rectangular crystal form, the edge tracing algorithm would report most pixels as having no change in angle (i.e. the four straight edges), while four of the edge pixels would yield a 90° change in angle (i.e. the four corners)². An example of the output of the edge tracing algorithm for a refined sugar crystal of regular form is shown in Figure 5. The graph shows the cumulative change in angle as the entire crystal edge is traced.

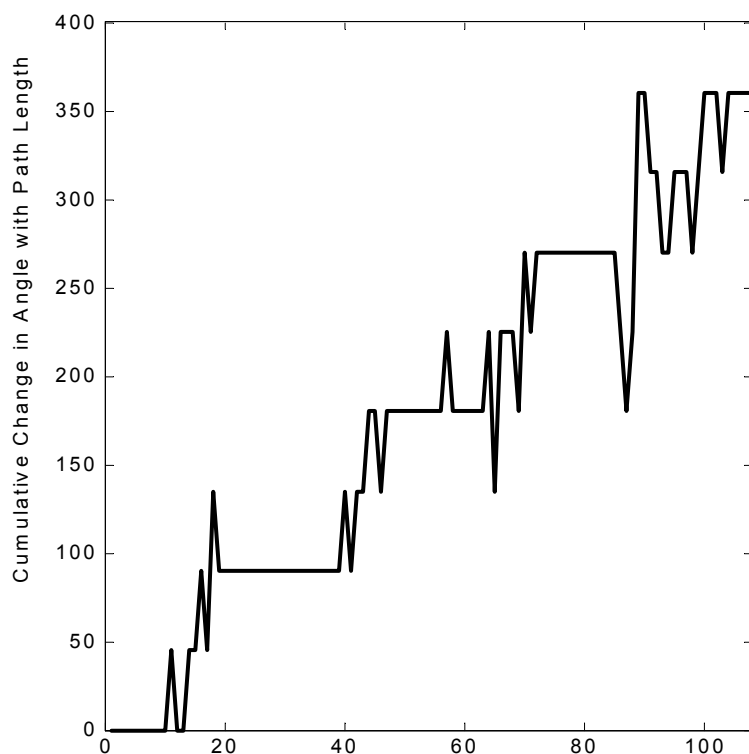


Figure 5. An example of the results of the edge tracing algorithm.

² Any analysis of this sort is obviously limited by the nature of the pixels in the image. As each pixel only has eight nearest neighbours, the change in angle from pixel to pixel can only take on eight discrete values. However, the edge tracing algorithm still proved useful, notwithstanding this limitation in terms of angle resolution.

As a means of identifying conglomerate crystals, the results of the edge tracing algorithm for each individual crystal were compared to a set of 'standards', namely:

- a circle (characterised by a regular, linear change in angle from 0-360° as the algorithm follows the edge of the circle around),
- a rectangle (characterised by four straight sides and four corners of 90°),
- an irregular hexagon (characterised by six straight sides and six corners), and
- an irregular octagon (characterised by eight straight sides and eight corners).

Each of the standard forms was fitted to the actual data found by the edge tracing algorithm using a simplex optimisation routine. The deviation of the actual edge tracing data from these best-fit standard forms was then investigated as a possible means of identifying conglomerate crystals (for example, it was considered that a regular crystal form would be more likely to conform more closely to the shape of an irregular hexagon than would a conglomerated crystal). Unfortunately, however, none of these techniques was found to reliably identify conglomerates.

The final technique investigated for identifying conglomerates was a crude check for autocorrelation in the angle data measured for each individual crystal. Single pixels yielding re-entrant angles were found to be quite common in the data produced by the edge tracing algorithm (there are a number of single-pixel negative changes in the cumulative angle which are readily apparent in Figure 5). This is a result of the poor resolution in the angle changes which results from each pixel only having eight nearest neighbours and does not necessarily indicate the actual existence of a re-entrant angle 'corner' on the crystal itself. However, it was assumed that two successive re-entrant (negative) changes in edge angle would indicate the existence of a substantial re-entrant angle on the crystal edge.

A simple algorithm was thus implemented in MATLAB to trace the edges of the individual crystals in an image and to search for occasions of two successive negative angle changes. It may be expected that this would happen reasonably infrequently for crystals of regular shape, but would be quite widespread in conglomerates.

The results of an analysis for the sixty crystals presented in Figures 1 and 2 are shown in Figure 6. By counting the number of crystals which exhibit more than two re-entrant angles and expressing the resulting figure as a percentage of the total number of crystals present, it was hoped that the presence of conglomerates in the sample could be quantified (a figure of 51.7% was obtained for the analysis shown in Figure 6). While some differences between the regular crystals and the conglomerates are apparent, the technique is not sufficiently powerful to be used as a routine method for quantifying conglomeration in raw and refined sugar samples.

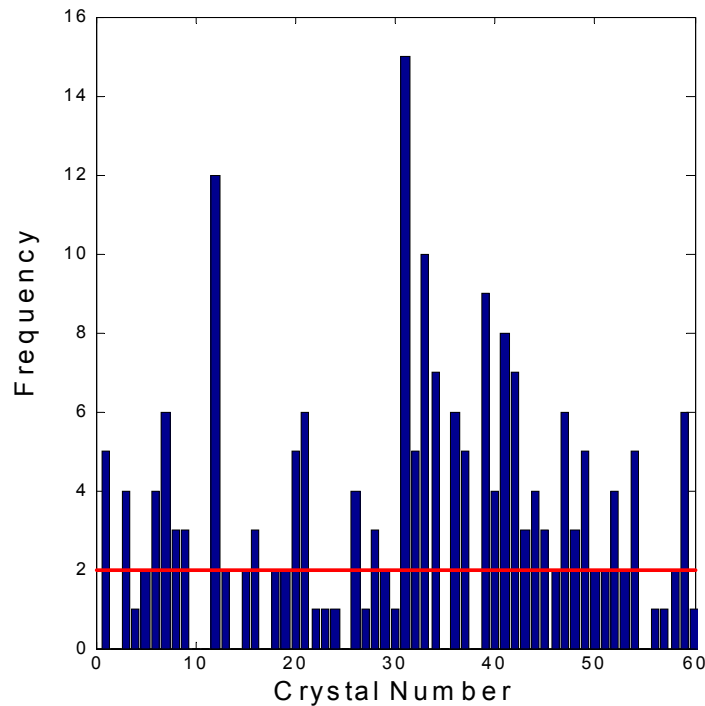


Figure 6. Results of the re-entrant angle analysis for sixty refined sugar crystals.

Conclusions

The current method used to quantify the extent of conglomeration in sugar samples is slow and dependent on a subjective classification carried out by one or more observers. Some preliminary techniques have been investigated, using computerised image analysis, which make progress towards automating the procedure and eliminating the subjective element of the original method.

The fitting of best-fit rectangular boxes around the individual crystals in an image was found to be the most reliable means of identifying conglomerates in sugar samples. This technique is suitable for routine use in the analysis of refined sugar samples. However, due to problems in ensuring the proper orientation of the crystals during the image generation process, the rectangle fitting technique was found to be a poor predictor of conglomeration in raw sugar samples.

Some of the other image analysis techniques investigated, such as the analysis of re-entrant angles using edge tracing, did show promise. However, further work will be required before they are suitable for regular use.

REFERENCES

Hibbert D, Phillipson RT, Woodwark W, Bonner BT and Mackay B (1975). Crystal regularity and its influence on white sugar quality. *Int Sug J* 77: 3-8 and 35-37.

Hill S (1965). A method of determining a 'crystal regularity index' for white sugar. *Int Sug J* 67: 201-204.

Plews RW (1970). *Analytical Methods Used in Sugar Refining*. Elsevier Publishing Company Limited, Amsterdam, The Netherlands. pp 85-87.

APPENDIX 1

A summary of the CRI method

The first operation in determining the CRI for a sample of refined sugar is to measure the mean aperture (MA) of the sugar sample using routine sieve analysis. Two narrow size fractions are then collected by further sieving of a sub-sample of the sugar, such that one fraction has a size range below the MA and one fraction has a size range above the MA. Each of the two size fractions is then separately processed.

A distributed set of crystals is obtained using a piece of sieve mesh with an appropriate aperture size. A piece of transparent self-adhesive tape is stuck to the underside of the mesh and a soft brush is used to distribute the crystals over the mesh. When the sieve mesh is inverted, a single crystal per screen aperture should adhere to the tape, while the unattached crystals fall away. The adhesive tape is then peeled away and attached to a glass plate or microscope slide and the result is suitable for the preparation of a photograph of the evenly distributed crystals. In the original method developed by Hill (1965), a photographic enlarger was used to create the required crystal photographs. With the benefits of modern technology, it is more cost effective to make use of a high resolution scanner which can produce enlarged digital images of good quality.

By comparison against the set of photographed standards shown in Figure 7 below, a total of 100 crystals from each of the two size fractions is divided into the three categories of crystal regularity. To avoid bias in the analysis, it is important that the 100 crystals selected for analysis from each photograph are chosen randomly (e.g. by analysing all crystals in a randomly chosen set of the crystal rows in the photograph). *Good* crystals are given two points, *moderate* crystals are given one point and *conglomerated* crystals are given no points. The total score obtained over the 200 crystals in the two size fractions is expressed as a percentage of the maximum achievable score (400 points) in order to obtain the CRI value.

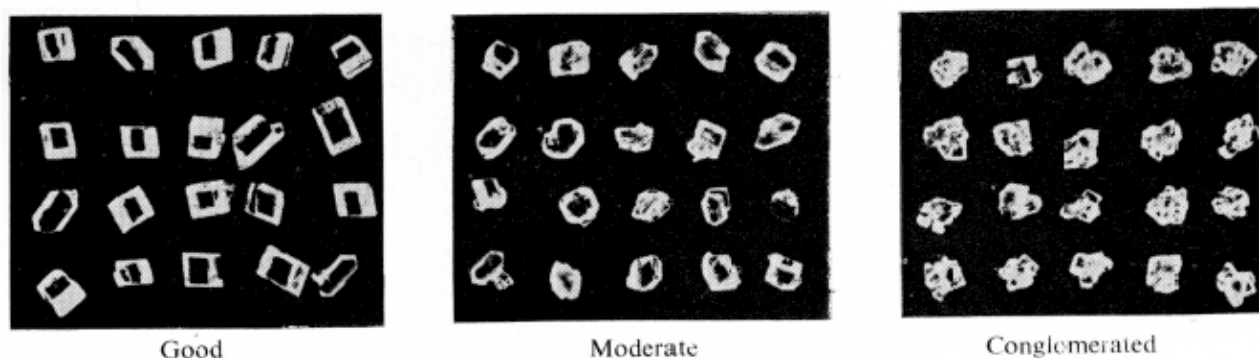


Figure 7. Crystal regularity standards (Hill, 1965).

