

UNSUPERVISED CLASSIFICATION OF DYNAMIC FROTHS.

G. Forbes* and G. de Jager**

* Centre for Minerals Research, Dept. of Chemical Engineering, University of Cape Town, Private Bag, Rondebosch, 7701, South Africa

** Digital Image Processing Group, Dept. of Electrical Engineering, University of Cape Town, Private Bag, Rondebosch, 7701, South Africa

Abstract: Machine vision systems typically classify images of a flotation froth surface into one of a distinct set of classes. This process typically involves having an experienced operator identify a set of froth classes. After this, a machine vision system is trained to identify these froth classes. Identifying these froth classes is particularly challenging for froths which have “dynamic” bubble size distributions. Using unsupervised clustering algorithms, it is possible to automatically learn these froth classes without user input. Validation of this technique is done by showing that the identified froth classes have statistically different relationships between the froth velocity and concentrate grade.

Key words: machine vision, froth flotation, unsupervised classification, bubble size distribution.

1. INTRODUCTION

1.1 Flotation

Flotation is a separation process used in many mining operations to upgrade the desired mineral concentration before further downstream processing. The operation of the flotation process is a complex one which is not entirely understood. Each flotation cell has numerous input parameters (reagent dosage, froth depth, air flow rate) and is also affected by numerous disturbance variables (ore type, mill performance). Typically, plant operators inspect the state of the froth visually, taking into account such parameters as velocity, bubble size, texture, colour and stability. Based on the state of the froth, the operator might make changes to one or more of the input parameters in order to achieve optimal performance.

As a result of this, numerous machine vision systems have been developed to analyse the state of the froth in a manner similar to that of an experienced plant operator. The advantage of such an instrument is the availability of precise, unbiased measurements 24 hours a day.

1.2 Froth Classification

Machine vision systems that monitor froth flotation cells typically classify the state of the froth into a number of discreet classes [1, 2, 3]. These froth classes are usually dependent on the bubble size distribution (texture) of the froth. The reason for using both bubble size and texture measurements is that it is not always possible to accurately segment individual bubbles in a froth. This means that it is not always possible to determine an accurate bubble size distribution (BSD) for a froth. Under such circumstances, texture measures can be used which

allow for the discrimination of froths with different bubble size distributions. However, the texture measures do not provide the user with an accurate bubble size distribution.

The froth classes are usually determined by studying the fluctuations in a flotation cell over a long period of time (typically a number of days). Visually dissimilar froth classes are then identified by experienced operators. The machine vision system is then trained to be able to identify whether or not the cell being monitored is in one of these predetermined froth classes. One of the disadvantages of such a method is that there is no guarantee that all possible froth classes will be identified during the training process. This means that the system will not be able to provide useful information when an unknown froth class is identified.

1.3 “Dynamic” Bubble Size Distributions

When the bubble size distribution of the froth being monitored does not change rapidly over a short period of time, identification of froth classes by an experienced operator is a relatively simple process. This is not always the case; froths exist which have “dynamic” bubble size distributions. An example of such a froth is shown in Figure 1, where two frames of video footage which have been taken within one second of each other are shown.

This paper will be dealing with froths which have these “dynamic” bubble size distributions. One of the biggest difficulties with these froths is identifying different froth classes. This is because two different froths will look similar at times when viewed side by side. This dynamic nature of the froth makes it very difficult to classify froths into the appropriate classes. It is also a very time consuming task, that is likely to have multiple operators

classifying the same data set into different resultant subsets. Having a solution to finding these froth classes with minimal operator intervention is of utmost importance as it results in the availability of consistent results within a reasonable time frame.

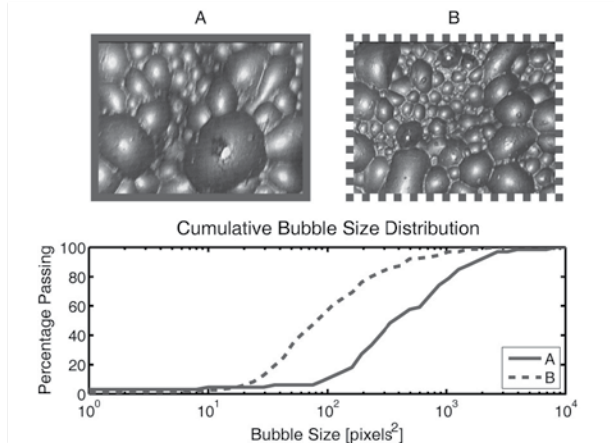


Figure 1: An example of a froth with a “dynamic” BSD. The BSDs are taken from two frames of video which are within 1 second of each other.

1.4 Objectives

The specific objectives of this paper are to show that unsupervised classification algorithms can be used to automatically detect a user specified number of froth classes. It will also be shown that the froth classes identified are not random groupings of froth classes, but are in fact real froth classes that are statistically significantly different in terms of the metallurgical performance of the cell at the time of operation when the froth class was observed.

2. UNSUPERVISED CLASSIFICATION OF FROTHS

2.1 Froth Data Set

The data set used in this work consists of 105 video segments. Each of these video segments is one minute in duration (1500 frames) and was captured from the first cell of the copper rougher circuit at Kennecott Utah Copper Concentrator in January 2006. At the same time that the video footage was captured, metallurgical samples of the feed, concentrate and tailings of the cell being monitored were taken. These samples were later analysed to determine their elemental composition.

Some example images of the froth video segments collected are shown in Figure 2. It is important to note that the still images do not capture the dynamic nature of the froths.

2.2 Bubble Size Distribution Measurements

The bubble size distribution for each of the frames of

video from all 105 videos was calculated using the improved watershed segmentation technique [4]. As has been shown previously, further reduction of the data in the bubble size distribution to a mean, median or p80 value is not appropriate for dynamic froths such as the one being examined here [5].

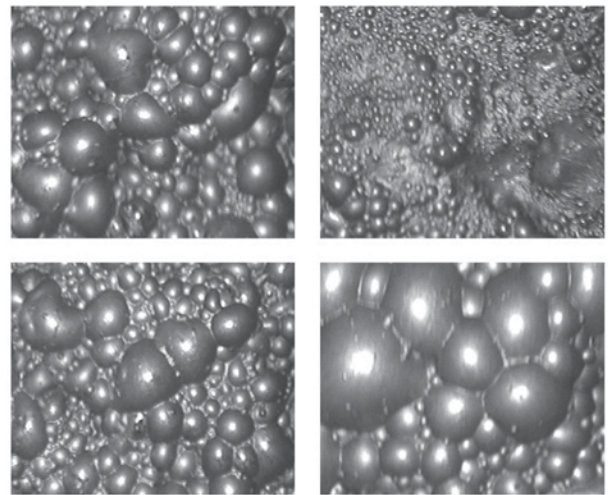


Figure 2: Example images of the froth of the first rougher cell at Kennecott. Note the different bubble sizes. Single still images do not provide an accurate description of the “dynamic” nature of the froth.

2.3 Frequently Occurring Bubble Size Distributions

The frequently occurring BSD algorithm is inspired by the work of Varma and Zisserman [6], but instead of finding image *textons* that occur frequently over an image, *frequently occurring BSDs* are determined that are found from the bubble size data.

A random sample of frames is taken from the entire data set of froth videos, such that the sample is representative of all the different bubble size distributions that can be found in the data set. The cumulative bubble size distributions are then calculated for each of the frames in the sample. Once this has been done, an unsupervised furthest-neighbour clustering algorithm is used to split the sample into a user specified number of classes. This is achieved by firstly creating an intra-distance matrix for the entire set of cumulative BSD samples. The Kolmogorov-Smirnov distance measure [7] is used to calculate the distance between two cumulative BSDs. The Matlab statistics toolbox is used to perform the unsupervised furthest-neighbour clustering. The intra-distance matrix is passed to the linkage function, the output of which is passed to the cluster function. This generates the user specified number of classes. For this work, eight clusters are typically used. This value is chosen so that the number of classes is small enough to ensure that there is still a visual difference between the images from which the identified BSDs are generated. The mean cumulative bubble size distribution can then be calculated for each of

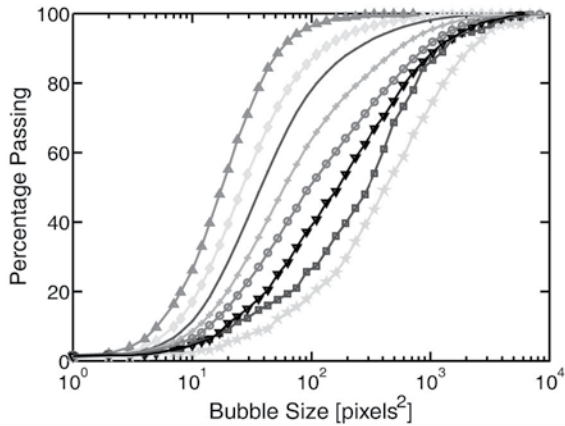


Figure 3: Frequently occurring BSDs learnt from 500 samples.

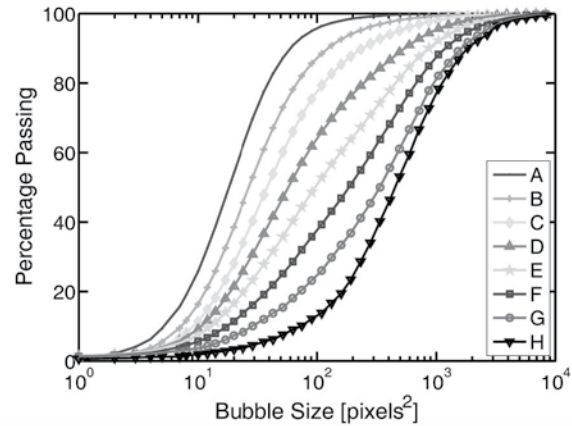


Figure 4: Frequently occurring BSDs learnt from 9000 samples.

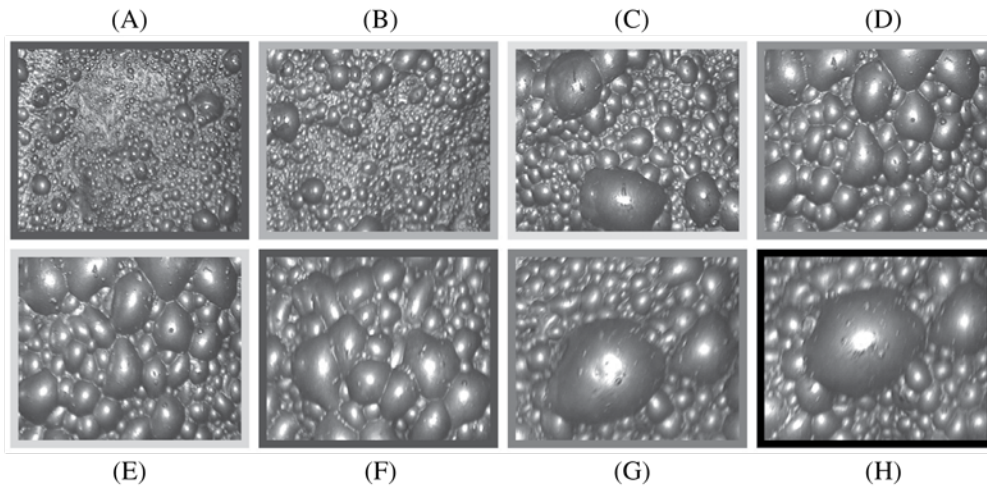


Figure 5: Example images of the froths represented by the cumulative BSDs in Figure 4.

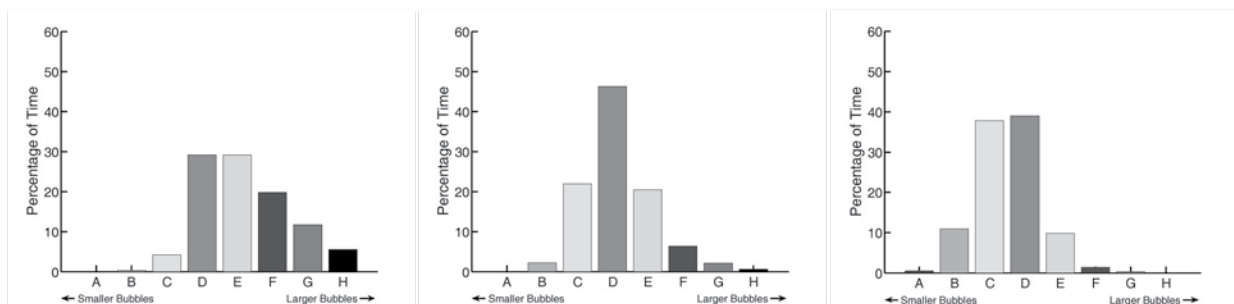


Figure 6: The results of using unsupervised classification to determine three froth classes. The labels of the histograms correspond to the labels of the BSDs in Figure 4.

segments. This resulting set of cumulative bubble size distributions is known as the frequently occurring BSDs [8].

To ensure that the random sample of frames taken from the entire data set was representative, a test was

performed by looking at the resultant frequently occurring BSDs that were found for different numbers of samples. The results from these tests are shown in Figures 3 and 4. These figures show the resulting frequently occurring BSDs when the number of samples drawn are 500 and 9000 respectively. From these figures

Figures 3 and 4. These figures show the resulting frequently occurring BSDs when the number of samples drawn are 500 and 9000 respectively. From these figures it is evident that a vast increase in the size of sample used in the unsupervised clustering does not have a significant impact on the resultant frequently occurring BSDs. The only difference is that the smoothness of the BSDs is increased when larger sample sizes are used. The difference in labelling of the frequently occurring BSDs is a result of the order in which the outputs from the clustering algorithm are generated, so the differences in labelling can be ignored.

Figure 5 shows example images of froths corresponding to the frequently occurring BSDs shown in Figure 4.

2.4 Characterisation of Dynamic Froths

Using the frequently occurring cumulative bubble size distributions, it is possible to characterise each video segment as a histogram. The histogram has the same number of bins as the number of frequently occurring BSDs that were identified in Section 2.3. The histogram shows the percentage of time that the froth has a bubble size distribution similar to the frequently occurring bubble size distributions.

The chi-squared distance measure [9] can be used to provide a measure of dissimilarity between the characteristic histograms of different froths. It is possible to create a dissimilarity matrix for the entire data set of characteristic histograms of froth video segments. This can be used in an unsupervised clustering algorithm (furthest-neighbour) to classify the data set into classes with similar characteristic histograms. Once again, the Matlab statistical toolbox functions: linkage and cluster are used to do the clustering.

The results from using these unsupervised clustering algorithms are shown in Figure 6. Note that the labels of the bars in Figure 6 corresponds to the bubble size distributions with the same labels in Figures 4 and 5. The characteristic histograms are clustered into three froth classes. This number is chosen for two reasons: firstly, experience tells us that flotation cells typically have between three to five different froth classes under normal operating conditions and secondly, to maximise the amount of concentrate data per froth class, which is important to ensure that statistically meaningful results are obtained.

3. VALIDATION

Froth velocity is an important performance indicator which is typically used for mass pull and concentrate grade prediction. In this section, the link between froth velocity and concentrate grade is used to validate the froth classes identified by the clustering algorithm described in Section 2.

3.1 Metallurgical Responses of Froth Classes

For each of the identified froth classes, the relationship between the froth velocity and metallurgical content of the concentrate can be modelled by linear regression. Different froth classes will have different trends, and so it is possible to use this information to determine if the froth classes identified are correct or just random collections of froths.

If the froth classes identified by the unsupervised classification algorithm are no more than a random selection of froths, then the relationships between the froth velocity and concentrate metallurgy for each of the froth classes will not be statistically significantly different from each other.

An example of such a set of regression lines is shown in Figure 7 which corresponds to a set of three froth classes which have been generated by randomly selecting their membership. The values in Table I show the results from an analysis to determine whether or not the regression lines from each of the randomly created froth classes are statistically different. The values are all less than ninety-five percent. This indicates that one cannot say with confidence that the lines are statistically different, and must therefore accept the null hypothesis which is that there is no difference between these froth classes. This is exactly what is to be expected from randomly allocated froth classes.

3.2 Statistical Calculations

This section gives a brief overview of the statistical tests used for the comparison of regression lines from different froth classes. For more detail, the reader is referred to [9,10]. The following series of F-Tests are performed in order to determine whether or not the regression lines are statistically different:

1. F-Test for the comparison of data sets' variance
2. F-Test for the comparison of the slopes of the regression lines

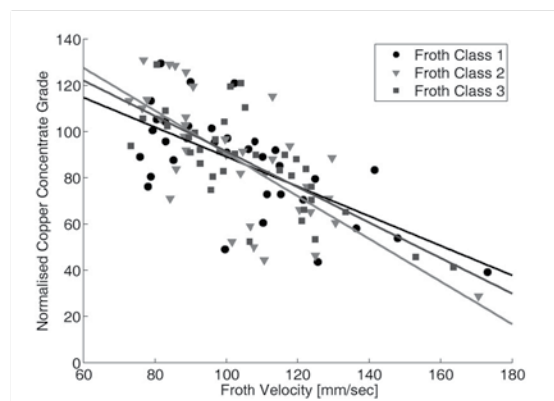


Figure 7: Linear trends relating froth velocity to concentrate grade for three randomly allocated froth classes.

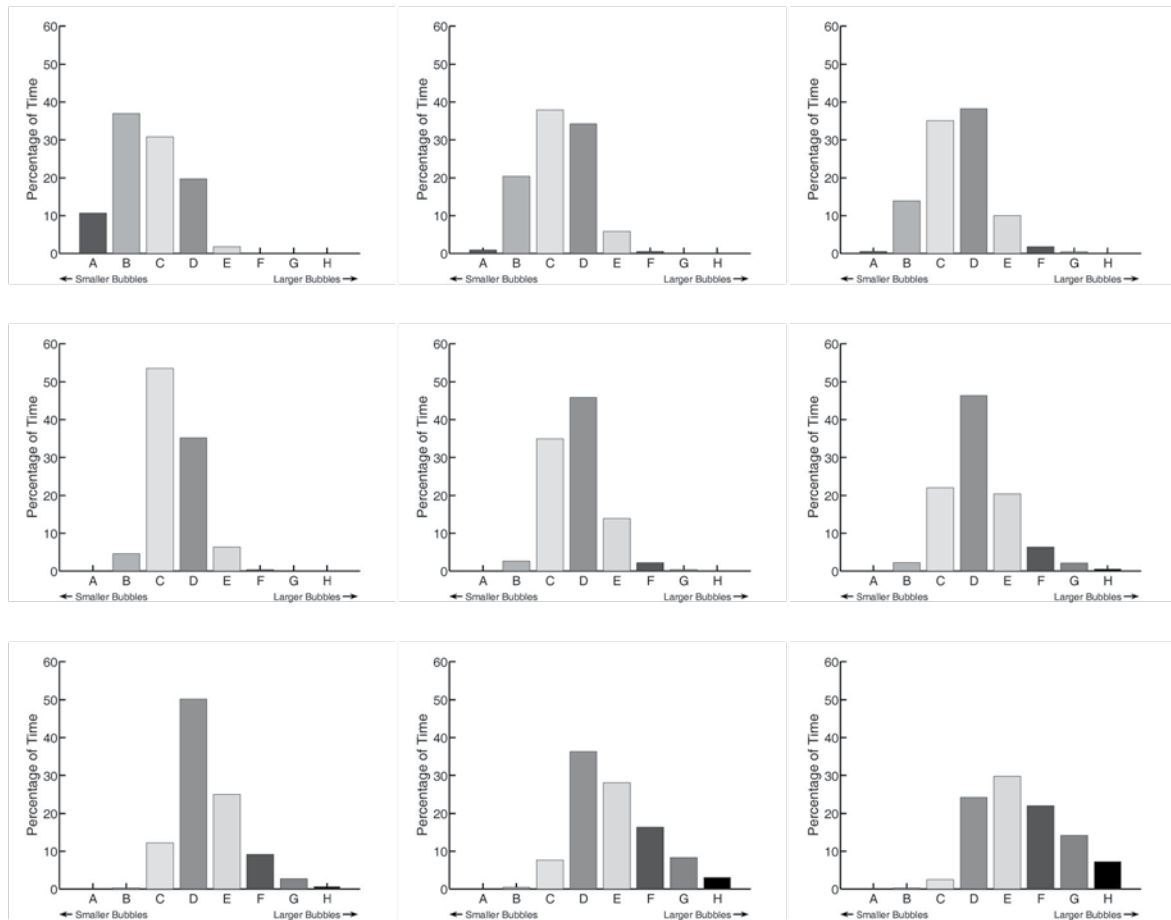


Figure 8: The results of using unsupervised classification to determine nine froth classes. The labels of the histograms correspond to the labels of the BSDs in Figure 4.

Table I: Summary of comparative statistics for three randomly allocated froth classes.

Assay	Froth Class A	Froth Class B	Confidence of Difference in Slope	Confidence of Difference in Intercept	Confidence of Difference in Mean	Confidence of Difference in (Overall)
Copper [Cu]	1	2	84.08	10.37	55.23	84.08
	1	3	53.04	34.01	66.95	66.95
	2	3	56.61	29.64	64.78	66.78

Table II: Summary of comparative statistics for three froth classes.

Assay	Froth Class A	Froth Class B	Confidence of Difference in Slope	Confidence of Difference in Intercept	Confidence of Difference in Mean	Confidence of Difference in (Overall)
Copper [Cu]	1	2	7.89	99.91	99.95	99.95
	1	3	5.11	99.99	100.00	100.00
	2	3	61.40	99.91	99.96	99.96

Table III: Summary of comparative statistics three froth classes generated using the alternative approach.

Assay	Froth Class A	Froth Class B	Confidence of Difference in Slope	Confidence of Difference in Intercept	Confidence of Difference in Mean	Confidence of Difference in (Overall)
Copper [Cu]	1	2	63.07	99.99	100.00	100.00
	1	3	95.30	100.00	100.00	100.00
	2	3	94.58	100.00	100.00	100.00

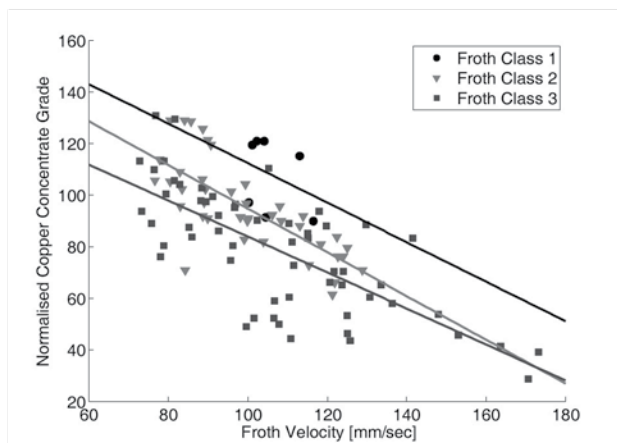


Figure 9: Linear trends relating froth velocity to concentrate grade for three froth classes.

3. F-Test for the comparison of the intercepts of the regression lines
4. F-Test for the comparison of the mean of the regression lines.

3.3 Verification of Unsupervised Clustering

Figure 9 shows the resulting trends from using the fitting of a linear regression model relating froth velocity to normalised concentrate grade for each of the three froth classes determined by the unsupervised clustering algorithm (the copper concentrate grade has been normalised for confidentiality reasons. The normalisation process does not however affect the trends observed.) It is clear from the figure that the trends have different mean values, unlike the random allocation of froths in Figure 7.

Statistical analysis of the differences between these regression lines is shown in Table II. All three of the regression lines are different from each other with at least 99.95% confidence.

These results show that the techniques used here to automatically determine the froth classes present in a set of videos of dynamic froths give meaningful results and not just a random selection of froth classes.

4. AN ALTERNATIVE APPROACH

4.1 User Intervention

An alternative approach to the previously mentioned classification method is now discussed. Unlike the previous technique, which only relies on the user's input for the number of froth classes to be determined, this method makes use of user intervention in deciding how to merge a larger set of froth classes into a smaller, more manageable set of froth classes.

Once again, the unsupervised froth classification techniques described earlier were used to determine nine froth classes. These nine froth classes are shown in

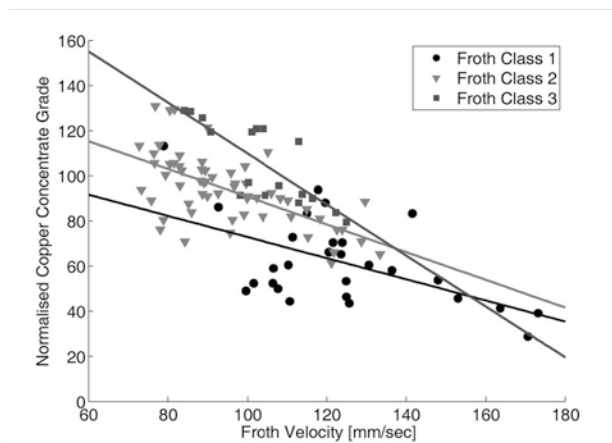


Figure 10: Linear trends relating froth velocity to concentrate grade for three froth classes using the alternative approach.

Figure 8. The metallurgical (froth velocity vs. concentrate grade) responses of these classes were analysed, and froths which were grouped together. It is necessary to do this, as having nine froth classes results in the metallurgical data effectively being divided by nine, thus reducing the statistical confidence that can be put on the observed trends. The result of this grouping of froth classes resulted in the middle column of froth classes in Figure 8 being chosen as the final froth classes.

The final metallurgical responses of these froth classes is shown in Figure 10. It is clear that different linear regression models exist for the separate froth classes. This is confirmed in Table III which shows the results for testing the statistical significance of the regression lines being the same. All of the trends relating froth velocity to concentrate grade for these froth classes have a 99.9% confidence that they are statistically different. It is also interesting to note that for this set of froth classes, the trends have an almost 95% confidence that the slopes are statistically different. This information is particularly useful from an operational point of view, as it provides the operator with additional information which can be used for the improvement of operation of the cell being monitored.

5. CONCLUSION

It has been shown that unsupervised clustering techniques can be used to separate a set of dynamic froths into visually similar froth classes. The results from the clustering have been validated by showing that each of the froth classes identified have statistically different regression lines relating froth velocity to concentrate grade. This would not be the case if the froth video segments were divided into random classes.

An alternative approach which relies more on user intervention results in very similar froth classes being identified. This is a further indication that the technique is giving appropriate clusterings of data. The alternative

approach also provides additional information for the operators as the trends modelled have statistically different slopes relating froth velocity to concentrate grade.

The major advantages of having a system which is able to make use of unsupervised clustering are the consistency of the results, the speed at which they can be obtained and the fact that this method will enable new froth classes to be identified in an online setting. This is unlike manually identifying froth classes which is a difficult, time consuming and operator dependent process, which invariably results in a poor set of froth classes to work with.

ACKNOWLEDGEMENTS

The authors would like to thank the following for their financial support: the NRF, the Department of Labour, Rio Tinto, Anglo Platinum and UCT Department of Chemical Engineering.

6. REFERENCES

- [1] G. Bartolacci, P. Pelletier, J. Tessier, C. Duchesne, P.A. Boss, and J. Fournier, "Application of numerical image analysis to process diagnosis and physical parameter measurement in mineral processes part i: Flotation control based on froth textural characteristics," *Centenary of Flotation Symposium*, Brisbane, Queensland, pp. 73 – 84, 2005.
- [2] A. Cipriano, M. Guarini, R. Vidal, A. Soto, C. Seplveda, D. Mery, and H. Briseo, "A real time visual sensor for supervision of flotation cells," *Minerals Engineering*, vol. 11, pp. 489 – 499, 1998.
- [3] M. Lu, W. Liu, F. Wang, and Y. Wang, "Using image analysis method to evaluate the performance of coal flotation for industrial column," *19th International Pittsburgh Coal Conference*, Pittsburgh, 2002.
- [4] G. Forbes and G. de Jager, "Texture measures for improved watershed segmentation of froth images," in *Fifteenth Annual Symposium of the Pattern Recognition Association of South Africa*, Grabouw, pp. 1–6, 2004.
- [5] G. Forbes and G. de Jager, "Bubble size distributions for froth classification," *Sixteenth Annual Symposium of the Pattern Recognition Association of South Africa*, Langebaan, South Africa, pp. 99–104, 2005.
- [6] M. Varma and A. Zisserman, "A statistical approach to texture classification from single images," *International Journal of Computer Vision*, vol. 62, no. 1–2, pp. 61–81, 2005.
- [7] Y. Rubner, J. Puzicha, C. Tomasi, and J. M. Buhmann, "Empirical evaluation of dissimilarity measures for color and texture," *Computer Vision and Image Understanding Journal*, vol. 84, no. 1, pp. 25–43, 2001.
- [8] G R Forbes, G de Jager, and D J Bradshaw, "Effective use of bubble size distribution measurements," *XXIII International Mineral Processing Congress*, Istanbul, Turkey, pp. 554–559, 2006.
- [9] T. J. Napier-Munn, "An introduction to comparative statistics and experimental design for minerals engineers," 2005.
- [10] T. J. Napier-Munn, "Analysing plant trials by comparing recovery-grade regression lines," *Minerals Engineering*, vol. 11, no. 10, pp. 949 – 958, 1998.