Visibility Classification of Rocks in Piles

Tobias Andersson, Matthew J. Thurley
Luleå University of Technology
Systems and Interactions
Department of Computer Science and Electrical Engineering
Luleå, Sweden
{tobias.andersson, matthew.thurley}@ltu.se

Abstract

Size measurement of rocks is usually performed by manual sampling and sieving techniques. Automatic on-line analysis of rock size based on image analysis techniques would allow non-invasive, frequent and consistent measurement. In practical measurement systems based on image analysis techniques, the surface of rock piles will be sampled and therefore contain overlapping rock fragments. It is critical to identify partially visible rock fragments for accurate size measurements. In this research, statistical classification methods are used to discriminate rocks on the surface of a pile between entirely visible and partially visible rocks. The feature visibility ratio is combined with commonly used 2D shape features to evaluate whether 2D shape features can improve classification accuracies to minimize overlapped particle error.

1 Introduction

In the mining and aggregate industry, measurements of the size of rocks are important. Measurements of the size of rocks can be used for quality control before delivery to customers. Also, accurate measurement of fragments size can be used for process control to allow efficient handling of the material.

Manual sampling followed by sieving with a square mesh is often used for quality control. The manual sampling is performed infrequently and is time-consuming. For efficient process control fast feedback of a measurement technique is required.

Image analysis techniques promise a quick, inexpensive and non-contact solution to determining the size distribution of a rock pile. Such techniques capture information of the surface of the rock pile which is then used to infer the pile size distribution. Work has been published on size and shape analysis of rock fragments. Comparison of manual sampling and estimates of rock fragment's size using 2D imaging analysis has been published [1, 2, 9, 19, 20]. 3D surface measurement of rocks has been applied to segmentation where rocks had little or no overlap [11], or to shape measurements of individual rock fragments [12]. However, in practice, rocks are not loaded and transported individually. Rocks are often transported on conveyor belt systems where the rocks lie in a pile. So, overlapping rock fragments will be present in a practical measurement system. Therefore, the system must minimize overlapping particle error. Overlapping particle error describes the fact that many particles are only partially visible and a large bias to the smaller size classes results if they are treated as small entirely visible particles and sized using only their visible profile. To minimize overlapped particle error, Thurley and Ng have presented work where a 3D feature called visibility ratio have been used to classify the visibility of rocks [18]. In earlier work we have classified pellets on a surface of a pile into entirely and partially visible pellets to overcome overlapped particle error [4].

In the presented research, the 3D visibility ratio is combined with commonly used 2D shape features to evaluate whether 2D shape features can improve classification performance. As it is critical to identify partially visible rocks and exclude them from any size estimates, not only minimum error rate decision rule is used. Minimum risk decision rule is used in this research to further increase the identification rate of partially visible pellets.

2 3D Surface Data

A sample of crushed basalt rock fragments was manually sieved into the three size classes 11.2-16.0, 16.0-22.4 and 22.4 - 31.5 mm. The rocks where color coded to allow manual identification of rock sizes in mixed rock piles. 3D surface data of 10 rock piles was collected using the
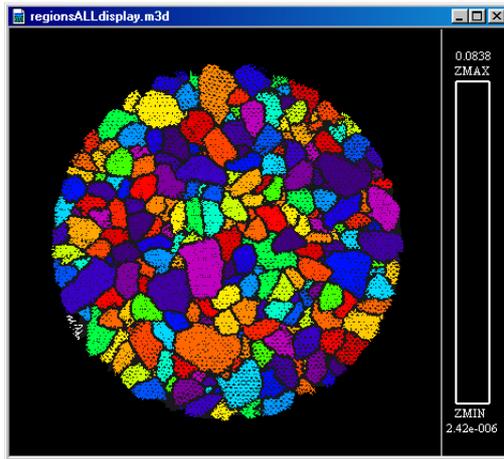
Figure 1. 3D surface model viewed from above. The sample is color coded to display the manual segmentation.

Monash Shape Measurement system developed by Alexander and Ng [3]. Thurley and Ng [17] details the collection of the rock piles used in this research. The 10 rock piles was divided into two separate sets. The first set, called *training set*, will be used to develop algorithms for visibility classification. The second set, called *hold out set*, will be held out during development of visibility classifiers. Thus, the second set will only be used to validate the classifiers performance.

The 3D surface data of the training and hold out set was manually interrogated to segment the data. A sample of one rock pile is shown in figure 1 where the manual segmentation is shown. Each rock was also labelled with its visibility class based on how much of the rock that was visible. The training set has a total number of 1069 rocks on the surface of the piles. 568 rocks are labelled as entirely visible and the remaining 501 rocks are partially visible. The hold out set has a total number of 1010 rocks on the surface of the piles. 515 rocks are labelled as entirely visible and the remaining 495 rocks are partially visible.

3 Classification

Estimating the size of entirely visible rocks have been successful using the best-fit rectangle [18]. However, best-fit rectangle do not discriminate effectively between size classes of partially visible rocks. It is critical to identify partially visible rocks and exclude them from any size estimates.

In this research we extract features to describe each rock, select a set of feature using binary logistic regression, use classifiers based on minimum error rate and minimum risk decision rules and validate the classifiers performance using the hold out method.

3.1 Feature Extraction

A 3D feature called visibility ratio have been used by Thurley and Ng to classify the visibility of rocks in piles [18]. In image analysis, shape analysis is a common approach to describe and classify regions, that comprise specific objects, in an image. 2D shape features have been used to detect broad-leaved weeds in cereal crops [15], to investigate the suitability of an imaging system to measure shape of particles [6] and for detection and classification of rocks [16].

In earlier work [4], where we presented methods for visibility classification of pellets in piles, features with scale were used. The features visibility ratio, equivalent area diameter, minor axis and major axis were effective to discriminate between entirely and partially visible pellets. The intended application, an industrial prototype, used the same measurement setup as in the laboratory. So, any classifier that worked would be suitable for the industrial prototype.

This rock data do not consist of the same type of rocks, the same kind of range of sizes, or the same kind of 3D data that might be used in an industrial system. Thus, it is very important to intelligently choose classifiers that will be sufficiently generic. For example, as the rocks in the sample range from sieve size classes of 11.2 mm to 31.5 mm and the resolution of the laboratory imaging system makes it possible to detect smaller sizes than that, it is important to not use features with scales. That is, measures like the area or length of the minor axis of the region would bias a classifier so that small regions are more likely to be partially visible. Therefore, in this research we only use scale less features and the 11 features that are extracted to describe each rock are described below:

- **Visibility ratio** [18] is a boundary following algorithm that accommodates *sparse, irregularly spaced* 3D coordinate data to allow the determination of entirely visible and partially visible rocks.

- **Aspect ratio** [6] is calculated by equation 1 and is the ratio between the minor and major axis length of the region. The minor and major axis is the non-perpendicular shortest and longest axis across the region.

$$AR = \frac{Major}{Minor} \qquad (1)$$

- **Circularity** of a region is commonly calculated by equation 2. This technique to calculate the circularity of regions is popular but is shown to give incorrect

results [5] due to fact that the perimeter is not a quantity which can be approximated well in digital images. However, we include this measure in the analysis.

$$C = \frac{Perimeter^2}{4\pi Area} \qquad (2)$$

- **Ratio 1** [15] is a measure of the region's "compactness" and is calculated by equation 3. That is, the measure is the major axis squared divided by the area.

$$Ratio_1 = \frac{Major^2}{Area} \qquad (3)$$

- **Ratio 2** [15] is also a measure of the region's "compactness" and is calculated by equation 4.

$$Ratio_2 = \frac{Minor^2}{Area} \qquad (4)$$

- **Ratio 3** [15, 14] (or called perimeter by Niekum) is yet another measure of the region's "compactness" and is calculated by equation 5. Note that the quantity perimeter is in the equation and therefore the feature should be treated carefully.

$$Ratio_3 = \frac{Perimeter^2}{Area} \qquad (5)$$

- **Eccentricity** [13] is the ratio between the foci of an ellipse and its major axis length. The ellipse has the same second-moments as the region. The value range between 0 for a circular region and 1 for a line segment.

- **Solidity** [13] is the ratio between the region area and the area of the region's convex hull. The convex hull is the smallest convex polygon that contain the region.

- **Extent** [13] is the ratio between the region area and the area of the region's Feret-box. A Feret-box is calculated by finding the maximum and minimum value of a region's i- and j-coordinates.

- **Elongation** [19] is the ratio between the minor and major axis of the Best-Fit rectangle.

- **Rectangularity** [19] is the ratio between the region area and the Best-Fit rectangle area.

3.2 Classification methods

The distribution of feature values in a data set is important to investigate in order to choose the right classification method. Inspection of our data set shows that the feature values are not multivariate normally distributed. The type of the response value also need to be considered when classification method is chosen. The response variable's type is binary for the visibility classification as the visibility class of a rock is either entirely visible or partially visible.

Johnson [10] suggests to use logistic regression as a classification method when the features values distribution are not multivariate normal. A rigorous description of logistic regression can be found in An Introduction to Generalized Linear Models [7].

Logistic regression can be used when the response variable are binary, ordinal or nominal. In the case when the response variable can only take two values, the method is called binary logistic regression. The form of the logistic regression model is shown in equation 6 where y is the response variable, $\mathbf{x}$ is a feature vector, β_0 is a constant and β_1 is a vector of parameters. $P(y = 1|\mathbf{x})$ denotes the probability that $y = 1$ given the observed feature vector $\mathbf{x}$. The model, or more specifically, β_0 and β_1 is fit to the known data via the use of maximum likelihood estimation.

$$P(y = 1|\mathbf{x}) = \frac{e^{\beta_0 + \beta_1' \mathbf{x}}}{1 + e^{\beta_0 + \beta_1' \mathbf{x}}} \qquad (6)$$

When the model is fitted to the known data, the conditional probabilities $P(y = 0|\mathbf{x})$ and $P(y = 1|\mathbf{x})$ can be calculated for new data and a decision be made by the classifier to what class the new data most probable belong. Following the notation in the book Pattern Classification [8], a decision rule that minimizes the average probability of error is the *minimum error rate* decision rule. The *minimum error rate* decision rule is to decide $y = 0$ when

$$P(y = 0|\mathbf{x}) > P(y = 1|\mathbf{x})$$

As it is critical to identify partially visible rocks and exclude them from any further analysis, a minimization of the error rate may result in too many partially visible rocks being classified as entirely visible. An alternative to the above classification rule is the *minimum-risk* decision rule. One way to define the *minimum-risk* decision rule is to decide $y = 0$ when

$$R(y = 0|\mathbf{x}) < R(y = 1|\mathbf{x})$$

The conditional risks $R(y = 0|\mathbf{x})$ and $R(y = 1|\mathbf{x})$ is defined by equation 7 and 8, where λ_{ij} is the loss incurred for deciding i when the true class is j. This will be used in our experiment to penalize misclassification of partially visible rocks as entirely visible rocks.

$$R(y = 0|\mathbf{x}) = \lambda_{00} P(y = 0|\mathbf{x}) + \lambda_{01} P(y = 1|\mathbf{x}) \qquad (7)$$

$$R(y = 1|\mathbf{x}) = \lambda_{10} P(y = 0|\mathbf{x}) + \lambda_{11} P(y = 1|\mathbf{x}) \qquad (8)$$

3.3 Feature Selection

We extract 11 features to describe each visible rock on the surface of the pile. An important step when designing a classifier is to select a set of features that can be used to discriminate between classes efficiently. Different procedures for selection of features have been proposed to guide researchers to select an efficient set of features. In this research, we use forward selection of features. Forward selection is an iterative technique that begin with an empty model. If some features are calculated to be statistically significant to include in the model, the feature that is most significant is added to the model. It is important that only one feature is added to the model at each step in the forward selection procedure. In the next iteration all remaining features are tested for inclusion to the model. The procedure stops when no new features meet the criteria of inclusion in the model.

To test whether a feature is statistical significant we use the likelihood ratio test. The likelihood ratio test is used to compare the goodness-of-fit between nested models. That is, to test a simpler model against a more complex model. The likelihood ratio statistic is calculated by equation 9, where L_0 and L_1 is the maximized value of the log-likelihood function for the full model and the reduced model.

$$LR = -2(L_0 - L_1) \tag{9}$$

The sampling distribution of the likelihood ratio statistic is approximately a chi-square distribution. The degrees of freedom of the distribution are equal to the difference in the number of parameters between the two models. In a forward selection procedure for binary logistic regression the degrees of freedom of the chi-square distribution is equal to 1 when LR is considered between consecutive steps. The chi-square distribution with one degree of freedom at a significance level of 0.01 equal to 6.63. Thus, the LR between two models have to be larger than 6.63 for a feature to be determined statistically significant and included in the model.

The results of the forward selection can be seen in table 1. In the first step, it is interesting to note that all features are statistically significant. With a likelihood ratio of 674.1, the visibility ratio stands out to be the most effective feature for discrimination between entirely and partially visible rocks. This was expected as the visibility ratio is designed to measure the visibility of fragments in a pile based on the 3D data. Solidity is the second most effective feature with a likelihood ratio of 436.2. This is also expected. Our belief before we did this experiment was that if any feature would improve the classification of visibility of rocks, solidity would contribute with most information about the visibility of rocks. This is because when a rock is partially

occluded by other rocks, the rock will probably have concave parts that are occluded and the solidity measure is a measure of this effect. The other features are statistically significant but do not improve the model as much as the features visibility ratio and solidity. And in the second step solidity is selected with a likelihood ratio of 136.6. In steps 3 to 5 Rectangularity, Circularity and ratio 1 are selected with likelihood ratios 15.9, 9.3 and 10.4 respectively. In the sixth step of the procedure no further features are selected as any improvement of additional features are statistically insignificant ($LR < 6.63$).

3.4 Estimating the classifiers performance

The estimated classification accuracy for the model at each step is shown in a confusion matrix, where the number of entirely or partially visible rocks classified as entirely or partially visible is shown. Note that this estimated performance is generated with the resubstitution method. That is, the same data that was used to train the classifier is used to estimate the how the classifier works on the same data. It has to be emphasized that the use of resubstitution to estimate a classifier's performance will generate overoptimistic results.

A confusion matrix is shown in table 2 for the estimated classification accuracies that use the *minimum error rate* decision rule. Note that with only the visibility ratio, an overall classification accuracy of 83.8 % is achieved. The overall classification accuracy increase as features are added to the model. At the fifth and final step the overall classification accuracy reaches 88.4 %. Interesting to note is that the classification accuracy for partially visible rocks are consistently worse than for entirely visible rocks. The consequence being that further analysis of the classified entirely visible rocks will contain a high degree of partially visible rocks.

4 Validation of Visibility Classification

As the resubstitution method give overoptimistic results, the classifiers performance are validated using the holdout method. The holdout method is a technique where a classifier is developed on a specific training set. A hold out set, separate from the training set, is used to validate how well the classifiers perform on new data. This method gives an unbiased estimate of classifiers performance. As our data is divided into two separate sets of rocks we use the first set as the training set and the second set as the test set.

First, the new data in the hold out set is used to validate the performance of the 5 different classifiers using the *minimum error rate* decision rule. A confusion matrix of the classification accuracies is shown in table 3. Note that the

Feature	Step 1		Step 2		Step 3		Step 4		Step 5		Step 6	
	LR	p *	LR	p *	LR	p *	LR	p *	LR	p *	LR	p *
Visibility ratio	**674.1**	**0**										
Aspect ratio	79.5	0	25.3	0	0.8	0.36	1.0	0.32	3.0	0.08	5.5	0.02
Circularity	195.2	0	60.2	0	11.2	0.00	**9.3**	**0**				
Ratio 1	196.6	0	64.4	0	2.9	0.09	4.8	0.03	**10.4**	**0**		
Ratio 2	193.7	0	68.2	0	3.4	0.07	2.5	0.11	4.1	0.04	0	0.84
Ratio 3	195.2	0	60.2	0	11.2	0.00	9.3	0	0.0	0.83	0	0.88
Eccentricity	74.7	0	23.6	0	1.8	0.18	3.6	0.06	5.9	0.01	0	0.88
Solidity	436.2	0	**136.6**	**0**								
Extent	297.1	0	91.6	0	2.5	0.11	5.8	0.02	6.6	0.01	2.4	0.12
Elongation	53.8	0	19.0	0	0.3	0.56	0.5	0.48	1.9	0.17	2.9	0.09
Rectangularity	185.7	0	44.0	0	**15.9**	**0.00**						

Table 1. Forward selection of features where a significance level of 0.01 is used for inclusion of new features to the model. The likelihood ratio, LR, and its corresponding p-value is shown for each feature.

	Known	Predicted		
		Entirely visible	Partially visible	% Correct
Step 1	Ent. visible	**482**	86	84.9
	Part. visible	87	**414**	82.6
	Overall %			83.8
Step 2	Ent. visible	**498**	70	87.7
	Part. visible	82	**419**	83.6
	Overall %			85.8
Step 3	Ent. visible	**500**	68	88.0
	Part. visible	79	**422**	84.2
	Overall %			86.2
Step 4	Ent. visible	**506**	62	89.1
	Part. visible	73	**428**	85.4
	Overall %			87.4
Step 5	Ent. visible	**512**	56	90.1
	Part. visible	68	**433**	86.4
	Overall %			88.4

Table 2. Classification accuracy estimated on the training set for each step of the feature selection procedure.

	Known	Predicted			
		Entirely visible		Partially visible	
		Number	Ratio *	Number	% Correct
Step 1	Ent. visible	**418**	**4.86**	97	81.2
	Part. visible	86	1	**409**	82.6
	Overall				81.9
Step 2	Ent. visible	**444**	**4.11**	71	86.2
	Part. visible	105	1	**390**	78.8
	Overall				82.6
Step 3	Ent. visible	**438**	**4.38**	77	85.0
	Part. visible	100	1	**395**	79.8
	Overall				82.5
Step 4	Ent. visible	**443**	**4.92**	72	86.0
	Part. visible	90	1	**405**	81.8
	Overall				84.0
Step 5	Ent. visible	**445**	**4.78**	70	86.4
	Part. visible	93	1	**402**	81.2
	Overall				83.9

* The ratio of accurately classified entirely visible rocks to misclassified partially visible rocks.

Table 3. Classification accuracy validated on the hold out set for each step of the feature selection procedure.

classifier with only the visibility ratio have an overall performance of 81.9 %. The overall performance is best for the classifier with 4 features with a classification accuracy of 84.0 % and the classifier with 5 features is close to that with 83.9 %. Most interestingly is that the classifier with only the visibility ratio misclassifies partially visible rocks best with 82.6% classified correctly. When more features are added to the classifier, the classifier do perform better in classifying the entirely visible rocks correct but not the important to identify partially visible rocks. Our belief is that the effect of occlusion affect the rocks visible shape randomly and that 2D shape features do not measure these effects efficiently.

Classification accuracy is not the only important factor in this application. The ratio of accurately classified entirely visible rocks to misclassified partially visible rocks is an important criteria in the statistical significance of the result. For the simple classifier that only consists of the vis-

ibility ratio, a ratio of 4.86 to 1 is achieved. That is, the accurately classified entirely visible rocks occur at a rate of 4.86 the misclassified partially visible rocks. From table 3 we observe that there is no clear advantage to adding more features to the model as the one feature ratio is 4.86 and the multiple feature ratios range between 4.22 and 4.92.

To maximize the ratio of accurately classified entirely visible rocks to misclassified partially visible the *minimum-risk* decision rule is used. The loss function is set to make misclassification of partially visible rocks 3 times more costly than misclassification of entirely visible rocks. The classification accuracies is shown in table 4. As expected the classification accuracy of partially visible rocks increase for the 5 classifiers. Again, we observe that the simple classifier that only consists of the visibility ratio is effective with a ratio 8.37 to 1 and that no clear improvement of the ratio

	Known	Predicted			% Correct
		Entirely visible		Partially visible	
		Number	Ratio *	Number	
Step 1	Ent. visible	**343**	**8.37**	172	66.6
	Part. visible	41	1	**454**	91.7
	Overall				78.9
Step 2	Ent. visible	**385**	**7.40**	130	74.8
	Part. visible	52	1	**443**	89.5
	Overall				83.2
Step 3	Ent. visible	**392**	**8.34**	123	76.1
	Part. visible	47	1	**448**	90.5
	Overall				83.2
Step 4	Ent. visible	**399**	**8.31**	116	77.5
	Part. visible	48	1	**447**	90.3
	Overall				84.0
Step 5	Ent. visible	**401**	**8.53**	114	77.9
	Part. visible	47	1	**448**	90.5
	Overall				83.9

∗ The ratio of accurately classified entirely visible rocks to misclassified partially visible rocks.

Table 4. Classification accuracy validated on the hold out set for each step of the feature selection procedure. The minimum-risk decision rule is used were misclassification of partially visible rocks is 3 times more costly.

	Known	Predicted			% Correct
		Entirely visible		Partially visible	
		Number	Ratio *	Number	
Step 1	Ent. visible	**306**	**4.78**	209	59.4
	Part. visible	64	1	**431**	87.1
	Overall				73.0
Step 2	Ent. visible	**320**	**6.03**	195	62.1
	Part. visible	53	1	**442**	89.3
	Overall				83.2
Step 3	Ent. visible	**335**	**6.20**	180	65.0
	Part. visible	54	1	**441**	89.1
	Overall				76.8
Step 4	Ent. visible	**347**	**6.67**	168	67.4
	Part. visible	52	1	**443**	89.5
	Overall				78.2

∗ The ratio of accurately classified entirely visible rocks to misclassified partially visible rocks.

Table 5. Classification accuracy validated on the hold out set for each step of the feature selection procedure when 2D shape features are used. The minimum-risk decision rule is used as in table 4.

could be seen by adding 2D shape features.

For comparison, a classifier is composed using the same procedure as before but leaving out the 3D visibility ratio. This will give an example of how effective the visibility ratio is and what is possible with only 2D shape features. The features selected are solidity, ratio 3, ratio 1 and rectangularity in that order. The confusion matrix when the classifiers performance are validated with the hold out data using the *minimum-risk* decision rule is shown in table 5. The best classifier, when only 2D features is used, is achieved when all four selected features are used. The ratio is 6.67 to 1 for the four feature classifier.

5 Comments on the results

Finally, we comment on the somewhat contradicting result that features are said to significantly improve the model but validation shows no improvement of classification performance. This is discussed by Johnson [10], where he emphasized that feature selection give guidance to effectively select a few possible subsets to consider to be used for classification. For example, in our experiment we have a total number of 11 features. If we would consider all possible subsets, a total number of subsets would be $2^{11} - 1 = 2047$. In our experiment we only need to consider the five subsets that where selected in table 1. Even as features are tested to significantly improve the model, all features may not actually be required to efficiently discriminate between different classes. This is exemplified by Johnson in figure 2, where

the ellipsoids of concentration for distinct bivariate normal populations. In figure 2 it is clear that both variables X_1 and X_2 are statistically significant and either or both may be included in a feature selection procedure. However, when it comes to validating a classifiers performance it is clear that only feature X_1 is actually needed to accurately discriminate between the populations.

6 Conclusion

Visibility classification of rocks in piles have been presented. The 3D visibility ratio, with the capacity of discriminating particles between visibility classes have been used in combination with other commonly used 2D shape features. The classifiers are based on binary logistic regression and discriminate between rocks that are entirely visible and partially visible.

Using forward selection of features with a significance level of 99 %, four of the additional 2D shape features are selected and can be said to improve the classifier. The estimated classification accuracies on the training data also indicate that the performance of the classifier with the additional features are higher than for a classifier that only use the 3D visibility ratio. But when validation is performed using the hold out set, there is no clear improvement in classification accuracy when additional features are used. Interestingly the misclassification of partially visible rocks increase as additional features are used.

As it is critical to identify the partially visible rocks, a minimum-risk decision rule is used to increase the ratio of accurately classified entirely visible rocks to misclassified
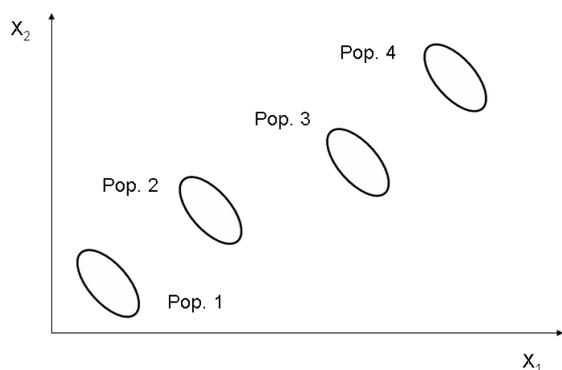
Figure 2. An example where both X_1 and X_2 are statistically significant. However, for classification between the populations only the variable X_1 is needed.

partially visible rocks. Using the simple classifier that only consist of the visibility ratio, a ratio of 8.37 accurately classified visible rocks to 1 misclassified partially visible rock is achieved. And no clear advantage is seen by adding additional features.

For comparison, we have also shown what is possible when the visibility ratio is not used and only 2D shape features are used. The ratio of accurately classified entirely visible rocks to misclassified partially visible rocks is 6.67 to 1 when four 2D shape features is used.

Finally, in practice a simple classifier is preferred and as no clear advantage have been seen by adding additional features we conclude that the visibility ratio is effective alone and should be used by itself in a practical measurement system.

References

[1] S. Al-Thyabat and N. Miles. An improved estimation of size distribution from particle profile measurements. *Powder Technolgy*, 166(3):152–160, Jun 2006.

[2] S. Al-Thyabat, N. Miles, and T. Koh. Estimation of the size distribution of particles moving on a conveyor belt. *Minerals Engineering*, 20:72–83, Jan 2007.

[3] B. Alexander and K. Ng. 3-d shape measurment by active triangulation using an array of coded light stripes. *SPIE Optics, Illumination and Image Sensing for Machine Vision II*, 850:199–209, 1987.

[4] T. Andersson, M. Thurley, and O. Marklund. Visibility classification of pellets in piles for sizing without overlapped particle error. *IEEE Proceedings of the 9th Biennial Conference of the Australian Pattern Recognition Society on Digital Image Computing Techniques and Applications*, pages 508 – 514, 3-5 Dec. 2007.

[5] M. J. Bottema. Circularity of objects in images. *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing*, 4(6):2247–2250, June 2000.

[6] R. Carter and Y. Yan. Measurement of particle shape using digital imaging techniques. *Journal of Physics: Conference Series*, 15:177–182, 2005.

[7] A. J. Dobson. *An introduction to generalized linear models.* Chapman & Hall/CRC, 2nd edition, 2002. ISBN: 1-58488-165-8.

[8] R. O. Duda, P. E. Hart, and D. G. Stork. *Pattern classification.* John Wiley & Sons, 2nd edition, 2001. ISBN: 0-471-05669-3.

[9] J. Fernlund. The effect of particle form on sieve analysis: a test by image analysis. *Engineering Geology*, 50:111–124, 1998.

[10] D. E. Johnson. *Applied Multivariate Methods for Data Analysts.* Duxbury Press, 1998. ISBN: 0-534-23796-7.

[11] H. Kim, A. Rauch, and C. Haas. Automated quality assessment of stone aggregates based on laser imaging and a neural network. *Journal of Computing in Civil Engineering*, pages 58–64, January 2004.

[12] J. Lee, M. Smith, L. Smith, and P. Midha. A mathematical morphology approach to image based 3d particle shape analysis. In *Machine Vision and Applications*, volume 16(5), pages 282–288. Springer–Verlag, 2005.

[13] Matlab®. Matlab image processing toolbox 5, users guide. page 560, 2007. Release 2007a.

[14] S. Niekum. Reliable rock detection and classification for autonomous science. Master's thesis, Carnegie Mellon Univeristy, December 2005.

[15] A. Perez, F. Lopez, J. Benlloch, and S. Christensen. Colour and shape analysis techniques for weed detection in cereal fields. In *Proceedings of 1st European Conference for Information Technology in Agriculture.* Elsevier, February 2000.

[16] D. Thompson, S. Niekum, T. Smith, and D. Wettergreen. Automatic detection and classification of features of geologic interest. *Proceedings of the IEEE Aerospace Conference.*

[17] M. J. Thurley and K. C. Ng. Identifying, visualizing, and comparing regions in irregularly spaced 3d surface data. *Computer Vision and Image Understanding*, 98(2):239–270, February 2005.

[18] M. J. Thurley and K. C. Ng. Identification and sizing of the entirely visible rocks from a 3d surface data segmentation of laboratory rock piles. *Computer Vision and Image Understanding*, 111(2):170–178, 2008. DOI: 10.1016/j.cviu.2007.09.009.

[19] W. Wang. Image analysis of particles by modified ferret method - best-fit rectangle. *Powder Technolgy*, 165(1):1–10, Jun 2006.

[20] W. Wang and O. Stephansson. Comparison between sieving and image analysis of aggregates. *Measurement of Blast Fragmentation - Proceedings of the FRAGBLAST 5 Workshop*, pages 141–148, 1996.