Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 167 (2020) 382–391

Procedia Computer Science

www.elsevier.com/locate/procedia

International Conference on Computational Intelligence and Data Science (ICCIDS 2019)

Prediction of Surface Roughness by Machine Vision using Principal Components based Regression Analysis

Ketaki Joshi[a]*, Bhushan Patil[b]

[a]*Research Scholar (Mechanical Engineering), Fr. Conceicao Rodrigues College of Engineering, Bandra, Mumbai, India*
[b]*Professor and Dean (Research and Development), Fr. Conceicao Rodrigues College of Engineering, Bandra, Mumbai, India*

Abstract

Machine vision provides imaging based solution for inspection of quality. For components to perform their intended functions, surface quality is equally important as their dimensional quality. Surface Roughness (R_a) is a widely accepted measure for evaluating surface quality. Traditional methods for measuring surface roughness may not be feasible in the industries insisting on 100% inspection due to the efforts and time involved in measurements. Machine vision can provide an automated, economic, fast and reliable solution.

This paper presents surface texture characterization of free hand grinding surfaces using machine vision approach and evaluation of their surface roughness using regression model based on machine vision parameters. Standard slip gauges sets of free hand grinding surfaces are selected for developing the regression model. Surface images are acquired in MATLAB software and processed for texture characterization using grey level co-occurrence (GLCM) matrix features. Then principal component analysis is carried out in SPSS software to define directions of unique variances in GLCM features. Furthermore, relationship between surface roughness value and GLCM features-based principal components is modeled using multiple regression analysis. Regression model developed can be used to predict unknown roughness values for free hand ground specimens. The approach demonstrated provides an automated, non-contact type method for measurement of surface roughness.

*Keywords:*Machine Vision; Regression Analysis; Surface Roughness; Grey Level Co-occurrence Matrix

* Corresponding author. Tel.: +91-22-67114104; fax: +91-22-67114100.
*E-mail address:*ketaki.joshi@fragnel.edu.in

10.1016/j.procs.2020.03.242

1. Introduction

Machine vision technology uses image based data to inspect quality of components. Surface quality of industrial components is a critical quality characteristic from functional, ergonomic and aesthetics aspects. Machine vision approach can be used characterize surface texture using the concept that an image is represented as a two-dimensional image intensity function characterized by two parameters: amount of light incident on the surface and amount of light reflected from the surface. Amount of light incident on the surface depends on illumination whereas that reflected from object is a function of surface irregularities and texture [1]. Machine learning can be carried out to model the relationship between vision-based texture parameters and surface roughness (R_a) value.

The paper present a unique machine vision approach for surface roughness evaluation using principal components based regression analysis for free hand grinding surfaces. It works on the principle of texture characterization of machined surface image using texture analysis techniques, followed by principal component analysis to represent original texture data in reduced orthogonal dimensional space featuring unique variances in the data and finally modeling the relationship between R_a and these principal components using multiple regression analysis. Proposed approach is capable of predicting surface roughness of free hand grinding surfaces from image data using regression model developed. The method can prove to be advantageous over traditional approach on account of time and efforts involved in inspection.

1.1. Grey Level Co-Occurrence Matrix based Parameters for Surface Texture Characterization

One of the commonly used statistical techniques to characterize surface texture using image data is second-order statistic grey level co-occurrence matrix (GLCM). It estimates joint probability $P_{d\theta}(i,j)$ of two pixels in the image separated by a distance d along the direction θ to have intensity values i and j [2]. GLCM is square matrix having dimensions equal to number of intensity levels present in the image and is effective in case of sufficient number of grey levels being present in the matrix [3]. According to Haralick (1979), distance d should be 1 and direction θ can be 0^0, 45^0, 90^0 or 135^0[4].Contrast, absolute value, angular second moment (energy), correlation, entropy, inverse difference, maximum probability are some of the common features extracted from GLCM [5].

Lu et al. (2006) used co-occurrence matrix for speckle pattern of grinding surface textures to estimate surface roughness. Contrast, correlation, homogeneity and energy were extracted along different directions and features curves were studied. It was found that parameter of normalized exponential trend curve for energy feature has strong correlation with surface roughness [6].

Shome et al. (2009) developed a machine vision based non-contact system for estimating surface roughness using GLCM parameters. Contrast, energy, entropy, homogeneity, range and standard deviation were extracted from images of turned components and surface roughness was modelled using neural networks approach [7].

Wang et al. (2009) used GLCM parameters including contrast, mean, correlation, homogeneity, angular second moment, dissimilarity, entropy and variance for evaluating surface roughness [8].

Gadelmawla (2011) investigated the relationship between surface roughness and GLCM features. Total 24 GLCM parameters were selected for the study like angular second moment, contrast, correlation, entropy, sum of entropy, variance and sum of variance etc. Correlation coefficients between each of the GLCM feature and R_a value were calculated and estimation equations were built up. The values predicted were in good agreement with the actual values [9].

Datta et al. (2012) studied application of machine vision system to evaluate surface roughness of turned components using GLCM features. They studied the relationship between roughness and GLCM contrast, homogeneity and concluded that these parameters are suitable for tool condition monitoring using vision data [10].

Nathan et al. (2014) used GLCM parameters to find surface roughness of end milled AA6061. Images of milled surfaces at different values of speed, feed and depth-of-cut according to L27 orthogonal array were processed and GLCM features such as contrast, correlation, energy and homogeneity were determined. Then the texture features and stylus based surface roughness values were compared to establish their relationship [11].

Vishwanatha and Srinivasa (2018) used GLCM features including contrast, energy, correlation and homogeneity extracted from fused images of dual-tree wavelet transform. Then the parameter values were used to model surface roughness using neural networks approach [12].

1.2. Principal Component Analysis

Exploratory factor analysis is the most suitable technique for analysing patterns of complex multidimensional, relationships present in huge data involving large number of variables and observations. Researchers use this technique to examine underlying patterns or relationships for determining the possibility of condensing this information and representing it by a smaller set of factors or components. Primary purpose of factor analysis is to define the underlying structure or relationship among variables involved in analysis. It was invented by Karl Pearson in 1901 and has two variants 'principal component analysis' and 'common factor analysis' [13]. Principal component analysis (PCA) derives factors having unique variance to reduce dimensionality of the data whereas common factor analysis finds common variance to identify latent structure in the data.

Principal components (PC) are linear combinations of original variables representing the information contained by them in the new coordinate system. This system has its axes along the directions of maximum unique variability in data. In order to reduce dimensionality of the data, principal components with Eigen values greater than 1 are used for data representation [14]. General form of principal components is as given in the equation (1).

$$\mathbf{Z_i = c_{i1}x_1 + c_{i2}x_2 + \ldots + c_{ip}x_p} \tag{1}$$

The constants $\mathbf{c_{ij}}$'s are determined using Eigenvectors and associated Eigen values as shown in equation (2).

$$\mathbf{C'\Sigma C = \Lambda} \tag{2}$$

In equation (2), **C** represents the matrix having its columns representing eigenvectors of **p** quality parameters, $\mathbf{\Sigma}$ represents covariance matrix and $\mathbf{\Lambda}$ represents a $\mathbf{p \times p}$ diagonal matrix of Eigenvalues.

1.3. Multiple Regression Analysis

Multiple regression analysis establishes the relationship between a single dependent metric variable and a number of independent metric variables. Multiple regression equation is determined by the method of least squares which minimizes sum of squared residuals. Independent Variables are added in regression model based on their additional predictive power. The variables are added using stepwise estimation where the model starts by selecting the best predictor of dependent variable and additional independent variables are then added based on their greatest incremental explanatory power. The variables are added as long as their partial correlation coefficients are statistically significant [14].

General form of multiple regression equation is as given in equation (3) where Y represents output, X_i represent input parameters, b_0 is the intercept, b_i are regression coefficients and ε is the residual.

$$Y = b_0 + b_1X_1 + b_2X_2 + \ldots\ldots + b_nX_n + \varepsilon \tag{3}$$

In current study, dependent variable is surface roughness (R_a) and independent variables are principal components extracted from GLCM features.

2. Problem Definition

The present paper aims at developing a machine vision based system for predicting surface roughness using principal component based regression model, with the case example of standard slip gauges of free hand grinding surfaces. Images of slip gauge templates selected for the study with their surface roughness values are as shown in fig.1.

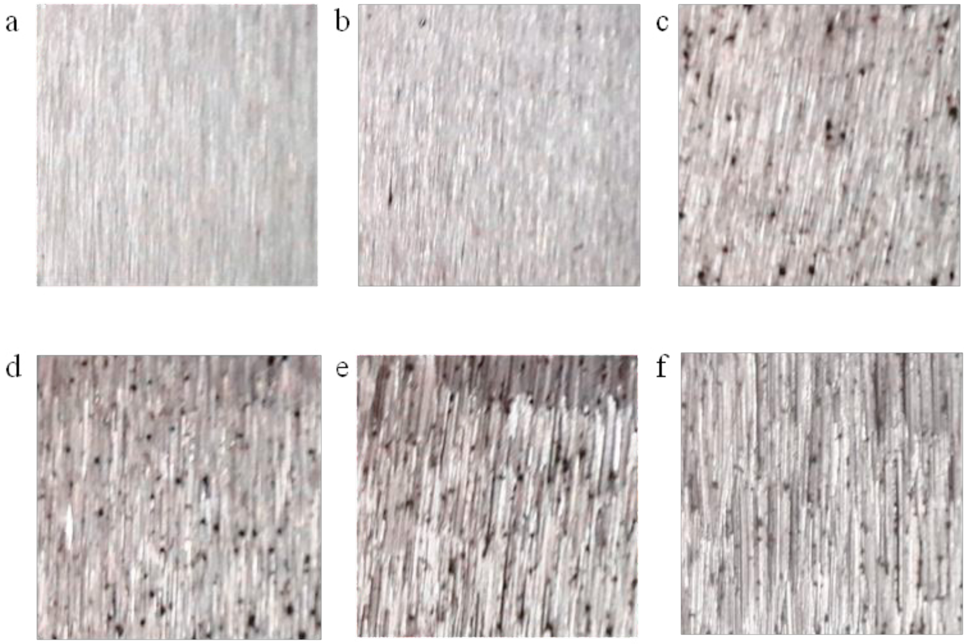

Fig. 1. Images of Free Hand Grinding Surface Templates used for Analysis with R_a Values in µm
(a) 0.8; (b) 1.6; (c) 3.2; (d) 6.3; (e) 12.5; (f) 25

3. Methodology

The methodology employed for developing machine vision system to evaluate surface roughness of surfaces selected for the study includes three major steps: surface texture characterization using machine vision, representation of data in principal component space and development of regression model of relationship between these principal components and R_a value.

For surface texture characterization using machine vision, images of slip gauges for free hand grinding surfaces are captured using digital camera having 12 MP resolution and acquired in MATLAB software for image processing. The images are preprocessed to convert color image into grayscale image. Each image of slip gauge template having original size of 800×800 pixels is then cropped to get 16 images of 200×200 size as shown in fig.2.

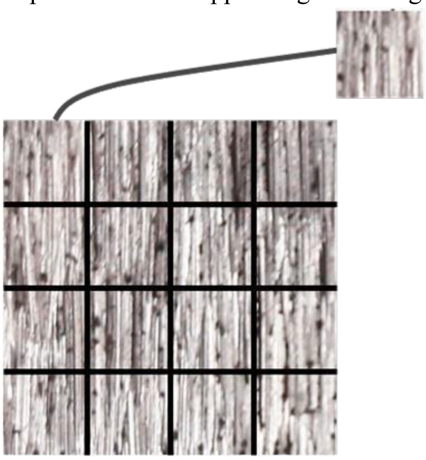

Fig. 2.Image of Grinding Template Cropped for Extraction of GLCM Features

Surface texture is then characterized using GLCM parameters as given in the equations (4) to (7) below. The parameters are extracted by considering correlation in image intensity values along different directions such as 0^0, 45^0, 90^0 and 135^0 using offset [0,1], [-1,1],[-1,0] and [-1,-1]. By default, global GLCM parameters are calculated along 0^0.GLCM features are extracted for all the 96 (6 slip gauges * 16 cropped images) images.

$$\text{Contrast} = \sum_i \sum_j (i-j)^2 \mathbf{P_d}(i,j) \tag{4}$$

$$\text{Correlation} = \frac{\sum_i \sum_j (i-\mu_x)(j-\mu_y)\mathbf{P_d}(i,j)}{\sigma_x \sigma_y} \tag{5}$$

$$\text{Homogeneity} = \sum_i \sum_j \frac{\mathbf{P_d}(i,j)}{1+|i-j|} \tag{6}$$

$$\text{Energy} = \sum_i \sum_j {\mathbf{P_d}}^2(i,j) \tag{7}$$

In second step, GLCM data for the samples is imported in SPSS software to carry out principal component analysis for determining directions of unique variances in the data. GLCM data is represented in reduced orthogonal space using principal components having Eigen value greater than 1.

In last step, PC scores are used as independent variables and $R_a$values of slip gauge templates as dependent variables to carry out regression analysis. The output provides a regression equation modeling the relationship between GLCM features based PCs and surface roughness R_a value. Flowchart for the methodology is as shown in fig.3.

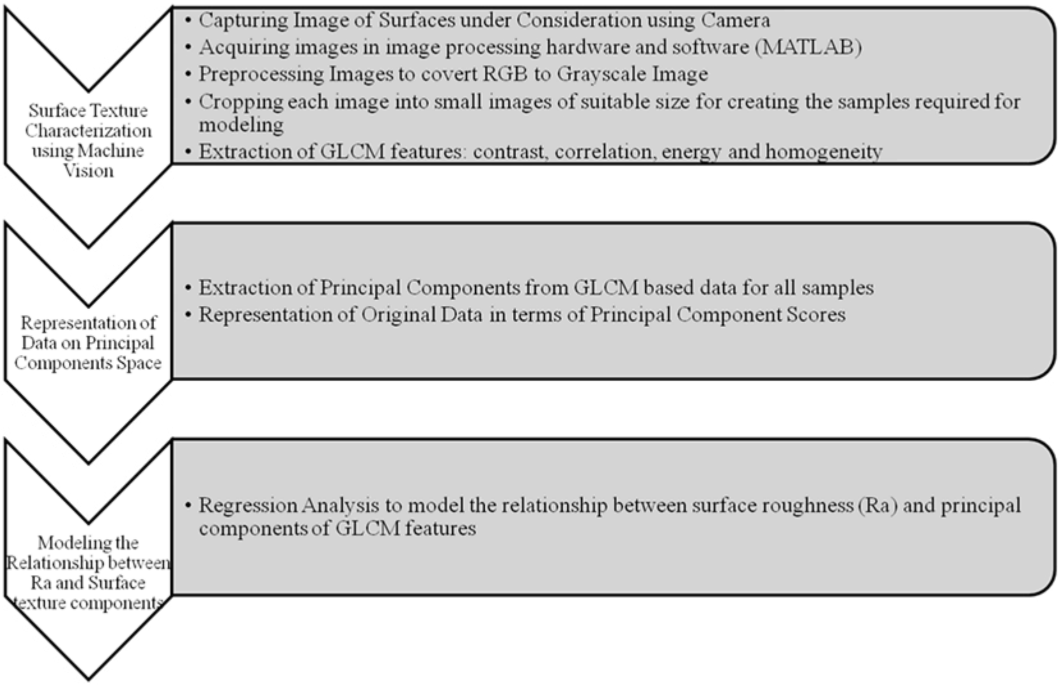

Fig. 3.Methodology for Developing Principal Components based Regression Model for Predicting Surface Roughness using Machine Vision

4. Results and Discussions

As explained in the methodology, GLCM features are extracted for images under consideration using MATLAB. Then the data is analyzed using PCA and multiple regression analysis in SPSS. The results of PCA and multiple regression analysis are as discussed below.

4.1. Results of Principal Component Analysis

The test for Kaiser-Meyer-Olkin (KMO) sampling adequacy and Bartlett's test of Sphericity check appropriateness of applying PCA. KMO measure checks interrelations among the data and Bartlett's test provides statistical significance of correlations among some of the variables in correlation matrix[14].Table 1 indicates the results of KMO and Bartlett's test.Value of KMO measure (.864) is greater than 0.5 and Bartlett's test indicates significant correlations among the data, both indicating appropriateness of applying PCA.

Table 1.Results of KMO and Bartlett's Test

KMO and Bartlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.864
Bartlett's Test of Sphericity	Approx. Chi-Square	8798.909
	Df	120
	Sig.	.000

Table 2 provides communalities and extraction referring to common variance in the data structure. All the variables have shared variance exceeding 0.8.

Table 2.Communalities (independent Variables)

	Initial	Extraction
GLOBCONT	1.000	.972
GLOBCORR	1.000	.867
GLOBENE	1.000	.956
GLOBHOM	1.000	.970
@45CONT	1.000	.971
@45CORR	1.000	.863
@45ENE	1.000	.957
@45HOM	1.000	.972
@90CONT	1.000	.969
@90CORR	1.000	.890
@90ENE	1.000	.959
@90HOM	1.000	.971
@135CONT	1.000	.970
@135CORR	1.000	.873
@135ENE	1.000	.957
@135HOM	1.000	.973

Extraction Method: Principal Component Analysis.

Eigen values of principal components extracted from the data are represented in Scree plot. Fig. 4 shows Scree plot indicating the presence of two principal components having Eigen value greater than 1. These principal components with Eigen values greater than 1 explain 94.308% of the overall variance in the data. PC1 explains 59.843% variance whereas PC2 explains 34.465% variance in the data. In comparison, PC1 captures relatively higher variance in GLCM feature values for different surfaces compared to PC2.

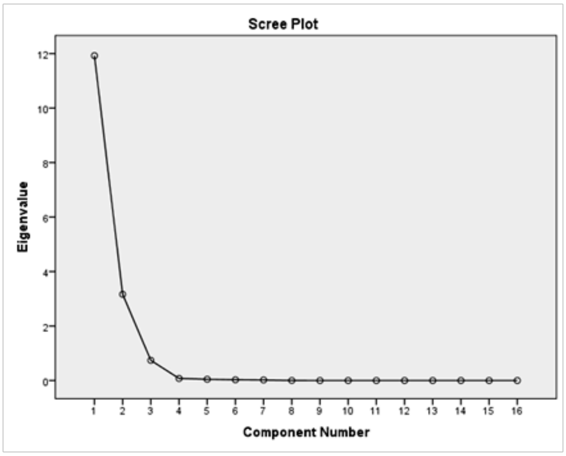

Fig. 4. Scree Plot

All above analysis indicates unidimensionality of the measures ensuring construct validity. That means all the items loaded on each factor are strongly associated.

Rotated component matrix is provided by PCA with 'Varimax' rotation for redistribution of variance in order to achieve simpler and meaningful patterns in the data. All values less than 0.4 are suppressed in the analysis and measures having cross loading or improper loadings are dropped. This provides the final rotated component matrix having unidimensional measures as shown in table 3.

Table 3.Rotated Component Matrix for Independent Variables

Rotated Component Matrix[a]		
	Component	
	1	2
GLOBCONT		.970
GLOBCORR	-.930	
GLOBENE	.871	
@45CONT		.967
@45CORR	-.929	
@45ENE	.872	
@90CONT		.974
@90CORR	-.939	
@135CONT		.966
@135CORR	-.934	
@135ENE	.872	

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.
a. Rotation converged in 3 iterations.

The results of PCA indicate variance explained by principal components being greater than 90% and loading greater than 80% with negligible cross-loading. This ensures discriminant and convergent validity. Factor loadings indicate that variables having similar texture characteristics are correlated into a principal component. This ensures nomological validity. Table 3 indicates that the factors loaded on first principal component are contrast and correlation along 0^0, correlation and energy along 45^0, correlation along 90^0 as well as correlation and energy along

135^0 whereas factors loaded on second principal component are contrasts along 0^0, 45^0, 90^0 and 135^0.

4.2. Results of Multiple Regression Analysis

Adjusted R^2 value in regression model summary table 4 indicates that 66.2% variation in surface roughness is explained by regression model. This adjusted R^2 value is much above the minimum R^2 value that can be statistically significant with a power of 0.8 at significance level of 0.05, for sample size of 96 and number of independent variables (principal components) as two.

Table 4. Model Summary for Regression Model

Model Summary[b]				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.818[a]	.669	.662	4.9320
a. Predictors: (Constant), PC2, PC1				
b. Dependent Variable: Surf_Rough				

ANOVA summary table of statistical significance test for overall model fit is given in table 5.

Table 5.ANOVA Statistical Test for Overall Model Fit

ANOVA[a]						
Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	4580.270	2	2290.135	94.149	.000[b]
	Residual	2262.183	93	24.325		
	Total	6842.453	95			
a. Dependent Variable: Surf_Rough						
b. Predictors: (Constant), PC2, PC1						

Total sum of squares is 6842.453 whereas reduction in the error by regression is 4580.270/6842.453*100 = 66.94%. The regression result is statistically significant with F-ratio of 94.149 and significance level of 0.000.

Multicollinearity introduces detrimental effect on the predictive power of regression model which is checked by variance inflation factor (VIF). It indicates the effect of other independent variables on the standard error of a regression coefficient with high value indicating multicollinearity.Maximum permissible VIF is 10 [14]. Table 6 shows that all VIF values are well below maximum value, indicating multicollinearity well within the control.

Table 6.Regression Coefficients

Coefficients[a]								
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	-9.574	3.571		-2.681	.009		
	PC1	12.933	1.593	.541	8.119	.000	.802	1.247
	PC2	-1.022	.162	-.419	-6.292	.000	.802	1.247
a. Dependent Variable: Surf_Rough								

The assumption of normality of error term is ensured from nature of histogram and normal probability plot given in fig. 5 and fig. 6. Normal probability plot provides graphical comparison of the nature of sample distribution to normal distribution. Its diagonal line represents normal distribution and actual distribution is plotted against this line [14].Fig. 5 indicates close resemblance of actual distribution with normal distribution line, thereby validating the assumption of normality of the error term.

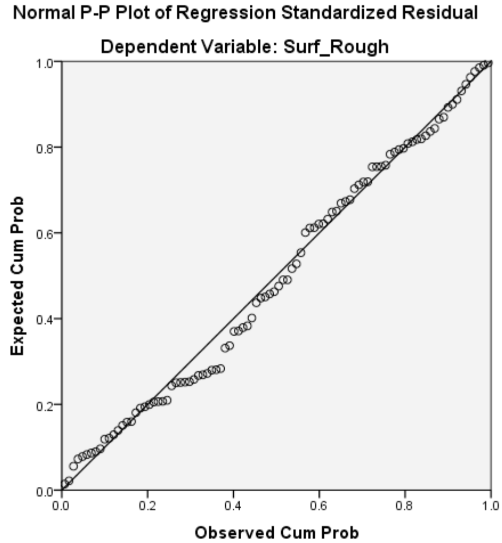

Fig. 5.Normal Probability Plot of Regression Standardized Residuals

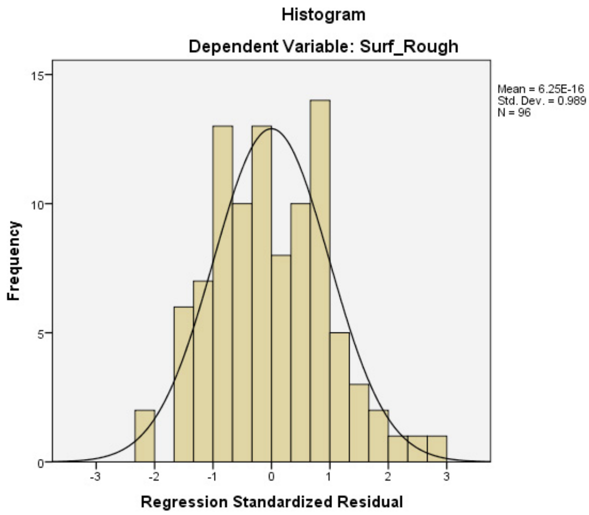

Fig.6. Histogram of Regression Standardized Residuals

Standardized regression coefficients calculated from standardized data are as given in table 6. Standardized beta coefficients help in comparing the coefficients in terms of their relative explanatory power of dependent variable [14]. As indicated in table 6, standardized beta coefficient for first PC is greater compared to that for second PC. It indicates that first principal component has relatively higher explanatory power of surface roughness.
Mathematically, the model regression equation for prediction of surface roughness (R_a) would be written as given in equation (8) below.

$$R_a = -9.574 + 12.933 * PC1 + (-1.022)*PC2 \quad (8)$$

5. Conclusion

The paper presented a machine vision approach for evaluating surface roughness of free hand ground specimens. Images of standard slip gauges for free hand grinding surfaces were captured using camera and acquired in MATLAB for processing. Images were cropped and processed to extract GLCM texture parameters: contrast, correlation, energy and homogeneity along different directions such as 0^0, 45^0, 90^0 and 135^0.Principal component analysis was carried out on GLCM features data for all the slip gauge surfaces, to identify directions of unique variances in the data. The results of PCA were validated for construct validity, discriminant validity, convergent validity and nomological validity. Results of PCA indicated the presence of two principal components based on GLCM features, explaining 94.308% of the total variance present in data. Factors loaded on first principal component were contrast and correlation along 0^0, correlation and energy along 45^0, correlation along 90^0 as well as correlation and energy along 135^0. Factors loaded on second principal component were contrasts along 0^0, 45^0, 90^0 and 135^0.

Furthermore, multiple regression analysis was carried out to model the relationship between surface roughness value and GLCM-based principle components. Results of regression analysis were validated for statistical significance, multicollinearity within control and normality of error term. Regression equation developed for predicting surface roughness of free hand ground specimens is $R_a = -9.574 + 12.933 * PC1 + (-1.022) * PC2$.

Machine vision system proposed here provides contact-less, automated means of measuring surface roughness that can replace traditional metrology instruments such as stylus based surface roughness testers, optical profile meters or interferometers. The system offers potential of providing industry-ready solution for 100% automated inspection of surface quality. The research can be furthered by developing regression models for other machining processes and using different predictive modelling tools.

References

[1] Palani, S., and U. Natarajan. (2011) "Prediction of surface roughness in CNC end milling by machine vision system using artificial neural network based on 2D Fourier transform." *The International Journal of Advanced Manufacturing Technology* 54, no. 9-12: 1033-1042.

[2] Al-Kindi, Ghassan A., and BijanShirinzadeh. (2007) "An evaluation of surface roughness parameters measurement using vision-based data." *International Journal of Machine Tools and Manufacture* 47, no. 3-4: 697-708.

[3] Materka, Andrzej, and Michal Strzelecki. (1998) "Texture analysis methods–a review." *Technical university of lodz, institute of electronics, COST B11 report, Brussels*: 9-11.

[4] Haralick, Robert M. (1979) "Statistical and structural approaches to texture." *Proceedings of the IEEE*67, no. 5: 786-804.

[5] Kumar, Rajneesh, P. Kulashekar, B. Dhanasekar, and Br Ramamoorthy. (2005) "Application of digital image magnification for surface roughness evaluation using machine vision." *International Journal of Machine Tools and Manufacture* 45, no. 2: 228-234.

[6] Lu, Rong-Sheng, Gui-Yun Tian, Duke Gledhill, and Steve Ward. (2006) "Grinding surface roughness measurement based on the co-occurrence matrix of speckle pattern texture." *Applied Optics* 45, no. 35: 8839-8847.

[7] Shome, D., P. K. Ray, and B. Mahanty. (2009) "Non-contact estimation of surface roughness in turning using computer vision and artificial neural networks." *International Journal of Industrial and Systems Engineering* 4, no. 4: 349-367.

[8] Wang, Xin, and Nicolas D. Georganas. (2009) "GLCM texture based fractal method for evaluating fabric surface roughness." In *2009 Canadian Conference on Electrical and Computer Engineering*, pp. 104-107. IEEE.

[9] Gadelmawla, E. S. (2011) "Estimation of surface roughness for turning operations using image texture features." *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture* 225, no. 8: 1281-1292.

[10] Datta, Anurup, SamikDutta, Surjya K. Pal, RanjanSen, and SudiptaMukhopadhyay. (2012) "Texture analysis of turned surface images using grey level co-occurrence technique." In *Advanced Materials Research*, vol. 365, pp. 38-43. Trans Tech Publications.

[11] Nathan, D., G. Thanigaiyarasu, and K. Vani. (2014) "Study on the relationship between surface roughness of AA6061 alloy end milling and image texture features of milled surface." *Procedia Engineering* 97: 150-157.

[12] Vishwanatha, J. S., and P. SrinivasaPai. (2018) "Modelling and prediction of surface roughness in Ti-6Al-4V turned surfaces: use of DTCWT image fusion and GLCM." In *IOP Conference Series: Materials Science and Engineering*, vol. 376, no. 1, p. 012133. IOP Publishing.

[13] Pearson, Karl. (1901) "LIII. On lines and planes of closest fit to systems of points in space." *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 2, no. 11: 559-572.

[14] Hair, J. F. (2010) "Multivariate data analysis: Pearson College Division.".