ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 4, April 2016

Computer Classification of Jharkhand Coal Based on Feature analysis and Repeatability using Image Processing

Deepak Kumar[1], S. S. Tripathy[2]

PG Student [I&C], Dept. of ECE, BIT MESRA, Ranchi, Jharkhand, India

Assistant Professor, Dept. of ECE, BIT MESRA, Ranchi, Jharkhand, India

ABSTRACT: Multiple ways are used in this paper for analysis and classification of coal images which is based on computer application method. In this paper, 2 basic method Feature analysis and Repeatability are used for classification of coal images. This paper deals with the analysis of coal which is taken from Jharkhand coal mines. The desirable property of a feature detector is repeatability which is based on the mean and standard deviation of intensity values and feature extraction of coal which gives the basic data set value based on which coal is classified. Haralick feature of the image gives the basic information for classification of the image among the others. GLCM (Gray Level Co-occurrence Matrix) is used to calculate the pixel occurrence to calculate the features. Before applying this method on images, pre-processing is done on the image for removing the noises from the image using thresholding and morphological operation with dilation followed by erosion to make the background white for betterment of the result. In this topic, both Haralick and Repeatability data set values are used for classification of coal images. Calculated values of Haralick and Repeatability of all query images are compared with their dataset value so that number of images classifies easily based on that values. These values also related with the coal chemical properties like carbon, hydrogen and nitrogen to reduce the chemical analysis of coal. By combining both feature values, the classification is easily done and give 90% efficiency.

KEYWORDS: Haralick feature; Image analysis; Thresholding; Morphology; Repeatability, Reproducibility

I. INTRODUCTION

Coal is an organic compound which is hard but brittle in nature. Coal forms from the plants. Plants and animals died and decomposed into the ground for thousands of years and due to the presence of heat and pressure inside the earth they converted into the natural resources. Coal is basically divided into 4 categories i.e. Anthracite, Bituminous, Lignite, and Peat. Peat is generally not lying in the properties of coal.so basically coal is divided into 3 categories. Anthracite is the oldest type of coal and Lignite is the classified in the laboratory according to chemical and physical properties. Physical properties [4] such as calorific value, moisture, volatile matter and ash and Chemical properties such as carbon, hydrogen, sulfur, nitrogen, and oxygen. Feature [1] extraction gives the information about the properties of the image which varies an image among others. Texture can generally define either by texture intensity or spectral properties. Haralick feature [5] is basically divided into 6 types of features i.e. Homogeneity, Contrast, Energy, Entropy, and Dissimilarity. Repeatability [2] defines as the determination of the abundance of a phase as that value of different Between 2 single determination and converted into Reproducibility [8] to determine the volume% of the component for classification of the image among others. Snapshot of coal can be taken using the camera with the resolution of 13 megapixels. RGB (Red Green Blue) image should be converted into the gray image because GLCM [9] of the image is based on the gray level intensity values. GLCM is used to determine the feature extraction of the image. GLCM is used to pixel occurrence to calculate the features. For better comparison, values of this feature should be in the range of 0 to 1. If calculated value is not in this range then that particular value should be normalized. Preprocessing applies on the image before the calculation of features of the image. The properties of the image are increased after applying preprocessing [12] on the image. If an image is not having good quality or any noise is present in the image then remove that noise using the filter according to the choice of image that we want which can increase the quality of the image and by using local histogram equalization we can also increase the quality of the image.

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 4, April 2016

Background subtraction [7] also be done before calculation of feature of the image because it affects the featured result of the image. Background of coal image should be fully white for the betterment of result. This paper presents worked on the white background image which is done by preprocessing. The objective of this paper is to focus on the both Tamura and Repeatability values.

II. PREPROCESSING

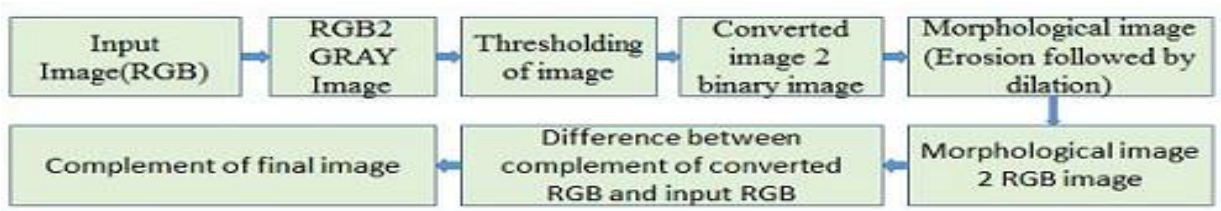


Figure 1 preprocessing of query image

The block diagram methods followed in the proposed method for the betterment of result. RGB image is converted into gray image and thresholding and leveling is done on the grayscale image. Thresholding is used to minimize the Interclass variance of the black and white pixels. For morphological operation converted image again converted into a binary image. Morphological operation based on erosion followed by dilation. In dilation process value of pixels present out of image, boundary is assigned the minimum value i.e. all pixels value beyond the image border is set to be 0 and in dilation process value of pixels present out of image border are assigned to maximum value i.e. all pixels present out of image boundary are assigned to be 1. After the morphological operation, computed image is changed into RGB format because there is a requirement of comparison between the real image and the computed image that can be done only in the same type of image. The difference between the computed RGB image and the real image gives the desired result into complement form so after difference takes the complement of the final image that gives the real image with only a white background.

III. REPEATABILITY

Repeatability gives the closeness between the dataset result and query image, It gives the result under the same measuring instrument used under the same condition over the same location. Since features are used as the starting point and main primitives for subsequent algorithms, the overall algorithm will often only be as good as its feature detector. Consequently, the desirable property of a feature detector is repeatability. Reproducibility which enhances the result and gives the result of measurement of the same measure under the different condition. Reproducibility is also calculated by using the repeatability which is based on the mean and standard deviation of image intensity value.

A. Mean

The mean is statistician's jargon for the average value of intensity values of the image.

$$\mu = \frac{1}{N_g}\sum_{i=1}^{N_g} x_i * p(g)$$

P(g)=GLCM feature of image
Ng=size of the image either in x or y-axis.

B. Standard deviation

The standard deviation is a measure of how far the image intensity fluctuates the mean.

$$\sigma^2 = \frac{1}{N_g - 1}\sum_{i=0}^{N_g-1} (x_i - \mu)$$

$$\text{Repeatability}(R_r) = {\sigma_i}^2 + {\sigma_j}^2 - \sigma_i * \sigma_j$$

C. *Reproducibility*

It enhances property for a feature detector calculates using repeatability value.

Reproducibility=2√2*Repeatability

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 4, April 2016

$$\text{Reproducibility}(R_s)=2\sqrt{2} * R_r$$

IV. HARALICK FEATURE

Haralick texture feature is a texture representation for the human perception which gives a great help for finding the properties of any object using its image based on 14 Haralick features but in this paper worked on only 5 Haralick feature include Homogeneity, Contrast, Dissimilarity, Energy, and Entropy. Texture intensity of Haralick is calculated using the GLCM which is also known as gray tone spatial dependence matrix. GLCM is based on the pixels value, direction length between two pixels either in horizontal, vertical or in diagonal direction and distance between pixel pair.

A. Homogeneity

It is used to define similarity between the pixels and neighbor pixels.

$$f_{Homogeneity}=\sum_{i=0}^{m}\sum_{j=0}^{n}\frac{p(i,j)}{1+(i-j)^2}$$

p(I,j)=GLCM matrix

m x n is the size of the image

B. Contrast

It is used to measure the amount of local variation in the image.

$$f_{contrast}=\sum_{i=0}^{m}|i-j|^2\sum_{i=0}^{m}\sum_{j=0}^{n}p(i,j)$$

C. Dissimilarity

It is used to measure the variation in intensity value of pixels in a pair.

$$f_{Dissimilarity}=\sum_{i=0}^{m}\sum_{j=0}^{n}|i-j|*p(i,j)$$

D. Entropy

It is used to measure the complexity of image that depends on the pixel intensity value.

$$f_{Entropy}=-\sum_{i=0}^{m}\sum_{j=0}^{n}p(i,j)*\log(p(i,j))$$

E. Energy

It is used to measure of local homogeneity of the image.

$$f_{ASM}=\sum_{i=0}^{m}\sum_{j=0}^{n}\{p(i,j)\}\text{^}2$$

$$f_{Energy}=\sqrt{f_{ASM}}$$

V. EXPERIMENTAL PROCEDURE

Coal images present in the database are of size 450 X 450. Snapshot of the object taken using a 13-megapixel camera with the distance of approximately 20cm in the sunlight. All the values of feature Note that the equation is centered using a center tab stop. Extraction component are set after calculation using the algorithm of feature extraction and according to that values of query image is compared with the dataset values of others image and similarity between the query image and dataset image is calculated. The database images are tested in the laboratory and their chemical properties are calculated using carbon%, hydrogen%, nitrogen% values so we can know that what type of coal is this and from where it belongs. By calculating the rank, it is concluded that what type of coal is this and according to the area it is concluded that from which area this coal is belonging. Using this method exact coal is found according to our requirement with less cost and high efficiency.

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 4, April 2016

TABLE 1: LAB Experimental data of Jharkhand coal

COAL AREA	CARBON%	NITROGEN%	HYDROGEN%	TYPE OF COAL
DHANBAD	86.16	1.581	5.500	Low-volatile bituminous
WEST BOKARO	74.05	0.726	4.421	Medium volatile bituminous
CHATRA	60.12	1.20	5.000	High volatile bituminous
RAMGARH	70.45	1.10	5.000	Medium volatile bituminous

TABLE 2: Haralick feature Data set values of Jharkhand coal

AREA OF COAL	Homogeneity	Contrast	Energy	Entropy	Dissimilarity
DHANBAD	0.0029-0.0030	0.6500-0.9000	0.8500-1.0000	0.0370-0.0389	0.0017-0.0025
RAMGARH	0.0027-0.0029	0.4000-0.6000	0.0000-0.2000	0.0340-0.0350	0.0030-0.0037
CHATRA	0.0027-0.0029	0.0000-0.4500	0.2000-0.4000	0.0350-0.0369	0.0022-0.0030
BOKARO	0.0031-0.0032	0.9250-1.0000	0.4500-0.7500	0.0390-0.0400	0.0012-0.0016

Experimental data of Haralick feature is calculated by analysis of 30 sample from each area

TABLE 3: Repeatability& Reproducibility data set value of Jharkhand coal

COAL AREA	Repeatability(Rr)	Reproducibility(Rs)
DHANBAD	0.5000-0.6800	0.6500-0.8500
WEST BOKARO	0.6800-1.000	0.8500-1.0000
CHATRA	0.0000-0.2000	0.0000-0.2500
RAMGARH	0.2000-0.4500	0.2500-0.5000

Set values for all coal images is calculated by analysis of 30 samples from each area which are maximum repeated value in that given range. Another unknown sample feature value is compared with a set value and calculates the efficiency of feature and Repeatability values by comparing both method. All values of both algorithm must be in the range of 0 to 1. If values are not in this range make them "Between" 0 to 1 by Normalization method.

$$X_{normalized} = \frac{X - X_{min}}{X_{max} - X_{min}}$$

where X=current value

X_{min} is minimum value among all values

X_{max} is maximum value among all values

VI. RESULT AND DISCUSSION

TABLE 4: Haralick features result of Bokaro Coal mine

BOKARO COAL	Homogeneity	Contrast	Energy	Entropy	Dissimilarity
Sample no 1	0.0032	0.9870	0.4242	0.0399	0.0013
Sample no 2	0.0031	0.9299	0.6085	0.0394	0.0013

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 4, April 2016

Sample no 3	0.0031	0.9667	0.6985	0.0389	0.0012
Sample no 4	0.0032	0.9891	0.7923	0.0400	0.0011
Sample no 5	0.0031	1.0000	0.7334	0.0398	0.0015

TABLE 5: Haralick features result of ChatraCoal mine

CHATRA COAL	Homogeneity	Contrast	Energy	Entropy	Dissimilarity
Sample no 1	0.0029	0.2747	0.2196	0.0369	0.0020
Sample no 2	0.0029	0.4726	0.3174	0.0367	0.0022
Sample no 3	0.0028	0.4901	0.3762	0.0352	0.0026
Sample no 4	0.0028	0.4246	0.4449	0.0355	0.0022
Sample no 5	0.0028	0.0000	0.4739	0.0354	0.0024

TABLE 6: Haralick features result of Ramgarh Coal mine:

RAMGARH COAL	Homogeneity	Contrast	Energy	Entropy	Dissimilarity
Sample no 1	0.0028	0.4701	0.2060	0.0351	0.0028
Sample no 2	0.0027	0.6122	0.0000	0.0340	0.0037
Sample no 3	0.0027	0.5718	0.1537	0.0344	0.0037
Sample no 4	0.0028	0.6255	0.0968	0.0352	0.0037
Sample no 5	0.0028	0.4824	0.0684	0.0347	0.0036

TABLE 7: Haralick features result of Dhanbad Coal mine:

DHANBAD COAL	Homogeneity	Contrast	Energy	Entropy	Dissimilarity
Sample no 1	0.0029	0.6305	0.9811	0.0386	0.0020
Sample no 2	0.0030	0.7517	0.8890	0.0377	0.0026
Sample no 3	0.0030	0.7806	0.8755	0.0380	0.0025
Sample no 4	0.0030	0.8693	1.0000	0.0383	0.0023
Sample no 5	0.0030	0.8948	0.9046	0.0385	0.0022

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 4, April 2016

TABLE 8: Repeatability and Reproducibility result of Dhanbad coal mine:

DHANABD COAL	Sample no1	Sample no 2	Sample no 3	Sample no 4
Repeatability	0.5100	0.5320	0.6544	0.6965
Reproducibility	0.6745	0.7075	0.8015	0.8405

TABLE 9: Repeatability and Reproducibility result of Ramgarh coal mine:

RAMGRAH COAL	Sample no1	Sample no 2	Sample no 3	Sample no 4
Repeatability	0.2155	0.2689	0.3000	0.4256
Reproducibility	0.2706	0.3100	0.3569	0.4659

TABLE 9: Repeatability and Reproducibility result of Bokaro coal mine:

BOKARO COAL	Sample no1	Sample no 2	Sample no 3	Sample no 4
Repeatability	0.7562	0.7752	0.8615	0.9425
Reproducibility	0.8615	0.8845	0.9265	0.9664

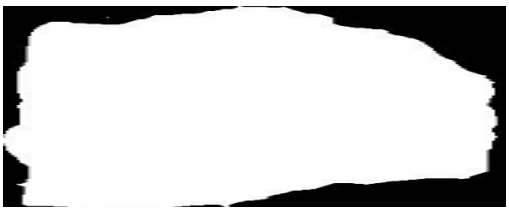
TABLE 10: Repeatability and Reproducibility result of Chatra coal mine:

CHATRA COAL	Sample no1	Sample no 2	Sample no 3	Sample no 4
Repeatability	0.0012	0.1269	0.1456	0.1965
Reproducibility	0.0125	0.1505	0.1989	0.2345

VII. PREPROCESSING RESULT

Real image with noise background Morphological binary image

(a) (b)



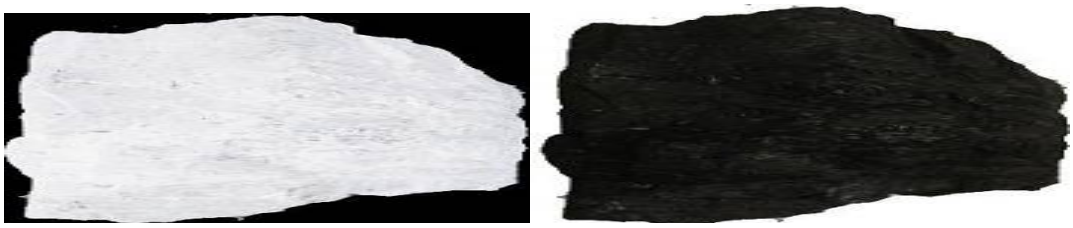
Morphological binary image (c)

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 4, April 2016



Extracted RGB image (d) Real extracted fully Noiseless (e)

VIII. CONCLUSION

In the pre-processing, image is extracted with a white background image using the concept of morphology. In dilation process value of pixels present out of image, boundary is assigned the minimum value i.e. all pixels value beyond the image border is set to be 0 means beyond the image area all is black. In dilation process value of pixels present out of the image, theborderis assigned to maximum value i.e. all pixels present out of image boundary are assigned to be 1 means total image area is black. Non-homogenous coal texture was proposed using the textural information to classify each type of coal.all the coal samples can make different with one another. Two algorithms based on the Haralick feature, Repeatability and Reproducibility values are compared with their dataset value for each and every image and number of images classifies easily based on that values. By combining both feature value, the classification is easily and give 90% efficiency.

REFERENCES

[1] D.Antonypandirranjan, K.Suganya, "Performance Analysis For Texture Classification" IEEE, International Conference on Emerging Trends in Science, Engineering and Technology, India,pp.251-254, march,2012
[2] Michael Cloke, Edward Lester, Martin Allen, "Repeatability of Maceral analysis using image analysis systems", Elsevier science Ltd, India,Vol 74 pp.654-658,march 1995.
[3] Dehmeshki, M. F. Daemi , N. J. Miles ,B. P. Atkins , R. E. Marston, "Classification of Coal Images by a Multi –Scale Segmentation Techniques", IEEE Conference on Mineral Resource Energy, Vol.95,pp.271-276,1995.
[4] Brian H. Bowen, Marty W. Irwin, "Coal Characteristics", The Energy Centre at Discovery Park Purdue University, Vol.8, Oct 2008.
[5] Nathan Loewke, "Haralick Texture Analysis for Stem Cell
[6] G. O'Brine, B.Jenkins, J. Esterle, H. Beath,"Coal characterization by automated coal petrography", Elsevier conference, at fuel 82,pp.1067-1073, Dec 2002.
[7] R. C. Gonzalez, and R. E. Woods, "Digital Image Processing," 2nd ed., Prentice Hall, Upper Saddle River, NJ, 2002.
[8] Michoke, Edward Lester, Martin Allen, Nick J. Miles,"Repeatability of MaceralAnalysis using Image Analysis Systems" Fuel, Vol.74 .pp.653-658,1995.
[9] M. A. Tahir, A. Bouridane, F. Kurugollu, and A. Amira, "Accelerating the computation of GLCM and Haralick texture features on reconfigurable hardware," ICIP '04. 2004 International Conference on image processing, Vol. 5, pp. 2857-2860,2004.
[10] Azhar M. Kadim, "Iris Texture Recognition Using Co-occurrence Matrix Features with K_means Algorithm ",Journal of Al-Nahrain University science,Vol.14 (4), pp.185-190,December 2011.
[11] Gleb Beliakov, Simon James, and Luigi Troiano, "Texture recognition by using GLCM and various aggregationfunctions",IEEE International Conference on Fuzzy logic,Vol6,pp265-270,Nov 2009.
[12] N. G. Rognin, M. Arditi, L. Mercier, "Parametric imaging for characterizing focal liver lesions in contrast-enhanced ultrasound" , IEEE International Conferene on Ultrson, Vol 57, pp 2503-2511,Nov 2010.