IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661,p-ISSN: 2278-8727, Volume 18, Issue 3, Ver. I (May-Jun. 2016), PP 14-19
www.iosrjournals.org

Clustering Of Images Based On Image Properties

Nithyananda C R[1], Ramachandra A C[2]
[1]*(Department Of Computer Science and Engineering, East Point College Of Engg. and Technology, India)*
[2]*(Department Of Electronics and Communication Engineering, Alpha College Of Engineering, India)*

Abstract: *Image Properties are used for the analysis of quality of the given image. The Intensity images, Contrast images, Weibull images and Fractal images are extracted from the input images. Eight basic image properties are calculated for these images. Analysis is made for the different Properties. Clustering is made on different types of images based on discriminative properties. Images are classified by using K-means method. Then analysis is made on the different clusters.*
Keywords: *Clustering, Image Property, K-means method, Normalization.*

I. INTRODUCTION

Process of extraction of various features from an image is used for analyzing the given image. The set of features are combined together to form feature vectors. Clustering is a method of grouping of similar objects or entities into groups. Grouping is based on the input values. The data representing the properties of images which are stored in the form of array or matrix. These values are considered as input for clustering. Clustering helps to identify or distinguish the images which can be used for Pattern Recognition and texture analysis.

II. LITERATURE SURVEY

Graytone Spatial Dependencies Textural Features for image classification is proposed by Robert M Haralick et. al.,[1]. Three different types of images : Areal images, photomicrographs and Earth Resources Technology Satellite images are used for classification. Piecewise Linear Decision Rule and Min-Max Decision Rules are used for classification. Authors concluded that Textural Features are general applicability for image Classification. Anne H Schistad Solberg et. al., [2] investigated the four Texture Features namely Features based on local image statistics, Gray level Cooccurrence matrices, Fractal features and Lognormal field model features for ERS-1 SAR images to landuse classification. The classification accuracy is increased by combining texture features. Entropy, Contrast, Angular Second moment, Inverse Difference moment, Inertia and Cluster shade texture features are computed from the Coccurrence matrix. The features : Kurtosis, Skewness, homogeneity and Power to mean ratio are computed by using local statistics.

Valery V Starovoitov et. al., [3] proposed a method to prepare a Cooccurrence Matrix for binary images which are sufficient for structure detection of an image. By using Joint feature with Cooccurrence matrix, computation time of analysis and memory is decreased. Jianguo Zhang et. al., [4] observed that Texture Features such as Central Mean, Global Mean and Variance are invariant to Rotation and Translation. Area, Compactness and Perimeter Texel properties are invariant to Rotation and Translation. For Translation invariance, Fourier Spectrum is used. Rotation invariant Features such as Ring and Wedge Filter, Gabor Model and Log-polar representation are obtained by applying polar coordinates on Frequency domain. Cheung Ming Lai et. al.,[5] proposed Fast Fractal image coding based on a single kick-out condition. Zero contrast prediction which reduces the computational time and reduces the additional memory is also implemented to speed up the encoding process.

Zho Wang et. al., [6] proposed the alternative quality measures for traditional quality assessments called Structural Similarity (*SSIM*). Highly structured natural image signals are independent of average contrast and luminance. Symmetry, Boundedness and Unique maximum conditions are satisfied for the similarity measure. Fumitaka Hosotani et. al., [7] introduced a method to denoise an image using 2D nonhormonic analysis. The proposed method resulted for good PSSNR but fails w.r.t. Structure Similarity Index Matrix. Unsupervised, Cluster based learning, image retrieval using similarity information is introduced by Yixin Chen et. al.,[8]. Images are represented in the form of nodes in a graph, each edge is with proper weight. NCut partitioning method is used to retrieve the clustered images. Optimizing the Support vector Machine parameters and Multiclass Support Vector Machine (*SVM*) parameters for image classification by Structural risk minimization and cross validation is introduced by Amit David et. al., [9]. Missclassification is avoided by thresholding maximum absolute *SVM* value. Error rate is tabulated for Linear Classifier, Gaussian *SVM*, 7-nearest Neighbor, Multilayer perceptron Neural Networks, Bayesian Neural Networks etc.,.

Automatic Labeling which describes the face visual attributes such as age, nose, gender, jaw shape, children, wearing sun glass, hair etc., is implemented by Neeraj Kumar et. al.,[10]. New dataset of faces called Face Traces with labeled attribute and PubFig with identities are proposed. Unsupervised, Affinity matrix based spectral clustering dimension is proposed by Hsin Chien Huang et. al., [11]. One Dimensional Search and Eigen vectors are used for finding the distance matrix or similarity. Benqin Song et.al., [12] developed a method which integrate the Extended Multiattribute Profiles and Sparse representation to exploit spatial and spectral features. Any given low dimensional structures may be represented in sparse and classified based on gathered information.

Image Properties are computed using following equations:

$$Brightness = \frac{\sum_{m=1}^{M}\sum_{n=1}^{N} I(m.n)}{M*N} \quad (1)$$

where M and N represents the dimension of the image $I(m,n)$.

$$Std.Deviation = \frac{1}{MN}\sum_{i=0}^{N-1}\sum_{j=0}^{M-1}\left(I_{ij} - \bar{I}\right)^{\mathbf{2}} \quad (2)$$

where M,N represents the size of the intensity and the average intensity of an image are represented by I and $\bar{I}$ respectively.

$$Entropy = -\sum P_k \, log_2 \, P_k \quad for\ k = 1\ to\ l \quad (3)$$

where k is the random value and P_k is the Probability value for k.

$$Kurtosis(x) = E\left\{\frac{(x-\mu)^4}{\sigma^4}\right\} \quad (4)$$

where E is Expectation, σ is deviation, μ is Standard Deviation for random variable x.

In older mathematical Statistics textbooks, Skewness is calculated either by mean and mode or mean and median.

$$Skewness = \frac{m_3}{m_2^{\frac{3}{2}}} = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^3}{\left[\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2\right]^{\frac{3}{2}}} \quad (5)$$

where m_2 and m_3 are second and third Moment represents variation and symmetric estimator. $\bar{x}$ and x_i are sample mean and Median for i elements.

$$Visibility = \frac{I_{max} - I_{min}}{I_{max} + I_{min}} \quad (6)$$

where I_{max} and I_{min} are the maximum and minimum luminance.

$$Spatial\ Frequency = \sqrt{RF^2 + CF^2} \quad (7)$$

where Row Frequency (RF) and Column Frequency (CF) are calculated by Equation (8) and (9).

$$RF = \sqrt{\frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N}[I(m,n) - I(m,n-1)]^2} \quad (8)$$

$$CF = \sqrt{\frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N}[I(m,n) - I(m-1,n)]^2} \quad (9)$$

where I is the image with M and N as Row and column size.

III. APPROACH

The image depository contains different types of images such as good visible, moderate visible and blur images. These images are converted into Intensity image, Weibull image, Contrast image and Fractal images after suitable transformations. All the Intensity images are selected for grouping. Similarly the Contrast images, Weibull images and Fractal images are selected for grouping separately. From each image types, the various properties like Contrast, Brightness, Skewness, Entropy, Separability, Kurtosis. Visibility and Spatial Frequency are extracted. These represent the feature of the particular image. The image property values are computed for Intensity image, Contrast image, Weibull and Fractal images. For each image, 32 image properties are extracted. These values are stored in the matrix nx32, where n represents the number of images.

A. Normalizing the data

Normalization is required to maintain the uniformity of different set of values. From image property matrix, we have to normalize each column based on its Maximum and minimum values. Equation (10) can be used to find the Normalized values for each element.

$$Normalize(x) = \frac{x - Min}{Max - Min} \quad (10)$$

Where *x* is the element considered for normalization, Min is the Minimum value of the particular column elements and Max is the maximum value of the column elements. If *Max* equals to *Min*, then Normalized value can be set to 0.5. We normalized the property values in the range 0 to 1, i.e., Min = 0 and Max = 1.

B. Clustering by K-means Method

K-means Algorithm is Vector Quantization Method to partition n observations into K-Clusters for the nearest mean. Aim is to minimize the square of the distance (Within-Cluster Sum of Squares) to the clusters from the points.

K-means Algorithm can be defined as follows:

1. Input the number of Clusters k.
2. If input is not provided, then randomly choose k value.
3. For the nearest class, group each data point.
4. For each cluster compute the sample mean.
5. For each data point, reclassify based on nearest mean.
6. Stop if there is small change in the mean. Otherwise repeat step 4.

IV. FIGURES AND TABLES

Different images of various dimensions are stored in image depository. For our method of implementation, 512x512 size images are taken as test inputs. Original images, converted images in the form of Intensity images, Contrast images, Weibull images and fractal images for Airplane, Baboon, Lena Good Visible image and Lena blur images are shown in Fig. 1 to Fig. 5.

Fig. 1. Airplane, Baboon, Lena(Good Visible) and Lena (Blur) original Images

Fig. 2. Airplane, Baboon, Lena(Good Visible) and Lena (Blur) Intensity Images

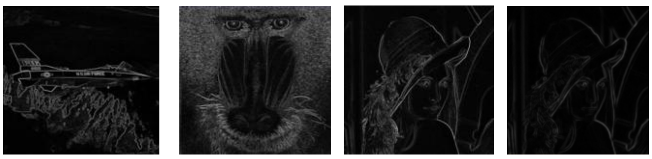

Fig. 3. Airplane, Baboon, Lena(Good Visible) and Lena (Blur) Contrast Images

Fig. 4. Airplane, Baboon, Lena(Good Visible) and Lena (Blur) Weibull Images

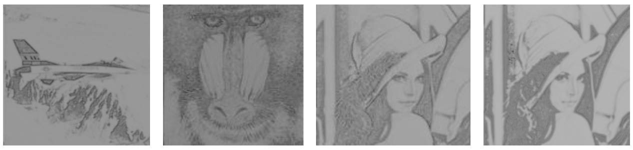

Fig. 5. Airplane, Baboon, Lena(Good Visible) and Lena (Blur) Fractal Images

The Normalized values for 9 different types of images for Intensity Properties, Contrast Properties, Weibull Properties and Fractal Properties are listed in Tables (Table 1 to 4). Abbreviations used in the tables are: Img-Image, Br-Brightness, Std.Dev-Standard Deviation, Ent-Entropy, Skw-Skewness, Krt-Kurtosis, Sep-Seprarability, Sp.Fr-Spatial Frequency and Vis-Visibility.

Table 1. Intensity Image Properties

Img.No.	**Br**	**Std. Dev**	**Ent**	**Skw**	**Krt**	**Sep**	**Sp.Fr**	**Vis**
1	0.58	0.53	0.93	0.65	0.38	0.31	0.65	0.4
2	0.25	0.92	0.96	0.53	0	0.82	1	1
3	0.42	1	1	0.64	0.01	1	0.74	0.91
4	0	0.7	0.91	1	0.5	0.57	0.66	0.83
5	0.88	0.02	0.12	0.38	0.43	0.62	0.01	0.01
6	0.13	0.69	0.9	0.88	0.32	0.68	0.63	0.75
7	0.98	0.11	0.46	0.11	0.15	0.39	0.06	0.06
8	1	0.07	0.35	0.65	0.17	0.59	0.01	0.03
9	0.14	0.81	0.97	0.8	0.18	0.88	0.39	0.89
10	0.49	0.04	0.26	0.64	0.4	0.34	0.01	0.03
11	0.24	0.33	0.79	0.52	1	0	0.59	0.3
12	0.39	0	0	0.5	0.3	0.29	0	0
13	0.85	0.57	0.96	0	0.38	0.22	0.96	0.35
14	0.98	0.01	0.15	0.41	0.44	0.2	0.02	0.01
15	0.71	0.01	0.15	0.41	0.44	0.2	0.02	0.01
16	0.84	0.02	0.17	0.06	0.2	0.39	0.02	0.01

Table 2. Contrast Image Properties

Img. No.	**Br**	**Std. Dev**	**Ent**	**Skw**	**Krt**	**Sep**	**Sp. Fr**	**Vis**
1	0.76	0.68	0.91	0.07	0.01	0.27	0.62	0.08
2	1	1	1	0	0	0.48	1	0.11
3	0.63	0.78	0.87	0.13	0.01	0.25	0.69	0.25
4	0.6	0.79	0.83	0.12	0.01	0.41	0.73	0.33
5	0.01	0.18	0.04	0.65	0.23	1	0.2	1
6	0.51	0.75	0.79	0.14	0.01	0.36	0.61	0.45
7	0.08	0.2	0.28	0.57	0.17	0.69	0.21	0.8
8	0.01	0.02	0.1	0.93	0.85	0.67	0.02	0.72
9	0.37	0.59	0.7	0.2	0.01	0.36	0.46	0.54
10	0.02	0.01	0.12	0.9	0.79	0.61	0.01	0.64
11	0.74	0.73	0.9	0.11	0.01	0.3	0.65	0.1
12	0	0	0	1	1	0.85	0	0.62
13	0.85	0.78	0.91	0.17	0.01	0	0.77	0
14	0.04	0.02	0.14	0.92	0.84	0.63	0.02	0.45
15	0.03	0.02	0.15	0.88	0.76	0.59	0.01	0.61
16	0.01	0.01	0.08	0.93	0.85	0.69	0.01	0.58

Table 3. Weibull Image Properties

Img. No.	**Br**	**Std. Dev**	**Ent**	**Skw**	**Krt**	**Sep**	**Sp. Fr**	**Vis**
1	0.59	0.42	0.71	0.58	0.37	0.36	0.66	0.23
2	0.33	1	1	0.35	0	0.79	1	1
3	0.32	0.7	0.68	0.6	0.01	1	0.58	0.73
4	0.23	0.63	0.84	1	0.34	0.61	0.68	0.73
5	0	0	0	0.55	0.17	0.78	0.1	0.43
6	0.39	0.6	0.82	0.81	0.23	0.67	0.52	0.53
7	0.77	0.57	0.98	0	0.14	0.43	0.34	0.21
8	0.14	0.39	0.75	0.56	0.12	0.65	0	0.68
9	0.44	0.79	0.79	0.76	0.14	0.86	0.69	0.62
10	0.1	0.01	0.12	0.25	0.39	0.33	0.02	0.26
11	0.52	0.06	0.14	0.47	1	0	0.47	0.01
12	1	0.32	0.65	0.86	0.49	0.29	0.53	0
13	0.96	0.33	0.65	0.72	0.54	0.18	0.76	0.01
14	0.65	0.14	0.36	0.24	0.67	0.07	0.35	0.01
15	0.9	0.37	0.76	0.37	0.29	0.39	0.44	0.06
16	0.24	0.06	0.27	0.15	0.37	0.31	0.17	0.21

Table 4. Fractal Image Properties

Img. No.	Br	Std. Dev	Ent	Skw	Krt	Sep	Sp. Fr	Vis
1	0.73	0.67	0.87	0.37	0.1	0.09	0.59	0.09
2	0.93	1	0.88	0.58	0.21	0.28	1	0.03
3	0.73	0.71	0 .82	0.34	0.08	0.08	0.63	0.1
4	0.94	0.84	0.92	0.57	0.27	0.12	0.92	0.02
5	0.17	0	0	0.84	0.81	0.79	0.08	0.19
6	1	0.93	0.85	0.67	0.38	0.38	0.87	0.01
7	0.1	0.04	0.16	0.91	0.86	0.86	0.06	0.4
8	0.09	0.02	0.06	0.92	0.89	0.9	0.04	0.39
9	0.97	0.78	0.81	0.64	0.4	0.27	0.77	0
10	0.49	0.08	0.34	0	0.47	0	0.27	0.01
11	0.67	0.78	1	0.25	0	0.15	0.83	0.14
12	0.18	0.05	0.37	0.73	0.62	0.58	0.12	0.32
13	0.86	0.58	0.77	0.66	0.42	0.25	0.64	0
14	0	0.05	0.1	1	1	1	0	1
15	0.01	0.06	0.17	0.97	0.95	0.95	0.02	0.89
16	0.05	0.04	0.07	0.96	0.94	0.95	0.042	0.57

K-means Algorithm is used for clustering for the Fractal images with Entropy Properties. Clustering is implemented for two, three and four clusters(Fig. 6). The Centroids are calculated for each clusters.

Output for two clusters:
Entered objects :
0.87, 0.88, 0.82, 0.92, 0, 0.85, 0.16, 0.06, 0.81, 0.34, 1, 0.37, 0.77, 0.1, 0.17, 0.07 .

Centroids :
C1 = 0.1
C2 = 0.8
Clustering :
Group 1 : 0, 0.16, 0.06, 0.34, 0.37,0.1, 0.17. 0.07
Group 2 : 0.87, 0.88, 0.82, 0.92, 0.85, 0.81, 1, 0.77

Output for three clusters:
Entered objects :
0.87, 0.88, 0.82, 0.92, 0, 0.85,0.16, 0.06, 0.81, 0.34, 1, 0.37, 0.77, 0.1, 0.17, 0.07

Centroids :
C1 = 0.1
C2 = 0.8
C3 = 0
Clustering :
Group 1 : 0.16, 0.06, 0.34. 0.37, 0.1, 0.17, 0.07
Group 2 : 0.87, 0.88, 0.82, 0.92, 0.85, 0.81, 1, 0.77
Group 3 : 0

Output for four clusters:
Entered objects :
0.87, 0.88, 0.82, 0.92, 0, 0.85, 0.16, 0.06, 0.81, 0.34, 1, 0.37, 0.77, 0.1, 0.17, 0.07.

Centroids :
C1 = 0.1
C2 = 0.7
C3 = 0
C4 = 0.8
Clustering :
Group 1 : 0.16, 0.06, 0.34. 0.37, 0.1, 0.17, 0.07
Group 2 : 0.77
Group 3 : 0
Group 4 : 0.87, 0.88, 0.82, 0.92, 0.85, 0.81, 1

Fig. 6. Fractal Image Entropy Property Clustering output by using K-means Algorithm for different number of Clusters

V. CONCLUSIONS

From the image depository, the intensity images, Contrast images, Weibull images and Fractal images are extracted. Brightness, Standard Deviation Entropy, Skewness, Kurtosis, Separability, Spatial Frequency and Visibility Properties are computed. Total thirty two properties are calculated and tabulated as Feature Vectors for each image. The images are grouped into clusters by applying K-means Clustering method. The proposed method classify the images which are acceptable by the user.

REFERENCES

[1] Robert M Haralick, K Shanmugam and Its'hak Dinstein, Textural Features for Image Classification, IEEE Transactions on Systems, man and Cybernetics, 3, 1973, 610-621.
[2] Anne H. Schistad Solberg and Anil K. Jain, Texture Fusion and Feature Selection Applied to SAR Imagery, IEEE Transactions on GeoScience and Remote Sensing, 35, 1997, 475-479.
[3] Valery V Starovoitov, Sang-Yong Jeong and Rae-Hong Park, Texture Periodicity Detection: Features, Properties, and Comparisons, IEEE Transactions on Systems, man and Cybernetics, 28, 1998, 839-849.
[4] Jianguo Zhang and Tieniu Tan, Brief review of invariant texture analysis methods, International Journal Pattern Recognition, 35, 2002, 735–747.
[5] Cheung Ming Lai, Kin Man Lam and Wan-Chi Siu, A Fast Fractal Image Coding Based on Kick-Out and Zero Contrast Conditions, IEEE Transactions on Image Processing, 12, 2003, 1398-1403.
[6] Zhou Wang, Alan C Bovik, Hamid R Sheikh and Eero P Simoncelli, Image Quality Assessment: From Error Visibility to Structural Similarity, IEEE Transactions on Image Processing, 13, 2004, 1-14.
[7] Fumitaka Hosotani, Yuya Inuzuka, Masaya Hasegawa, Shigeki Hirobayashi, and Tadanobu Misawa, Image Denoising With Edge-Preserving and Segmentation Based on Mask NHA, IEEE Transactions on Image Processing, 24, 2015, 6025-6033.
[8] Yixin Chen, James Z Wang and Robert Krovetz, CLUE: Cluster-Based Retrieval of Images by Unsupervised Learning, IEEE Transactions on Image Processing, 14, 2005,1187-1201.
[9] Amit David and Boaz Lerner, Support vector machine-based image classification for genetic syndrome diagnosis, Elsevier Pattern Recognition Letters, 26, 2005, 1029–1038.
[10] Neeraj Kumar, Alexander C. Berg, Peter N Belhumeur and Shree K Nayar, Describable Visual Attributes for Face Verification and Image Search, IEEE Transactions on Pattern Analysis and Machine intelligence, 33, 2011, 1962-1977.
[11] Hsin Chien Huang, Yung-Yu Chuang and Chu-Song Chen, Affinity Aggregation for Spectral Clustering, IEEE Conference on Computer Vision and Pattern Recognition, 2012, 773-780.
[12] Benqin Song, Jun Li, Mauro Dalla Mura, Antonio Plaza and Jon Atli Benediktsson, Remotely Sensed Image Classification Using Sparse Representations of Morphological Attribute Profiles, IEEE Transactions on GeoScience and Remote Sensing, 52, 2014, 5122-5235.