GE-International Journal of Engineering Research
Vol. 4, Issue 6, June 2016 IF- 4.721 ISSN: (2321-1717)

Website: www.aarf.asia Email : editor@aarf.asia , editoraarf@gmail.com

# A NEW PROPOSED TECHNIQUE FOR FEATURE EXTRACTION FROM SATELLITE IMAGES

**Ms. Alifiya Jahagirdar[1], Dr. Manisha Patil[2], Dr. Vrushsen Pawar[3], Dr. Vilas Kharat[4]**

[1] Research Student, Department of Computer Science, Savitribai Phule Pune University, India.

[2] Research Scholar, Department of R&D, Zensar Technologies, Pune, India.

[3] Professor and Head, Department of Science, Water and Land Management Institute, Aurangabad, India.

[4] Head Department of Computer Science, Department of C.S, Savitribai Phule Pune University, Pune, India.

## ABSTRACT

*During the last few decades large amount of satellite images have been captured to study different aspects related to agriculture, atmosphere, land use and land cover etc. It increases the demand of analyzing and generating accurate data from the images and provides the results as needed. A popular method is used for calculating the co-occurrence probabilities and determining features within remotely sensed digital imagery. Typically, the co-occurrence features are calculated by using a grey level co-occurrence matrix(GLCM) to store the co-occurring probabilities. In our research paper we have studied existing feature extraction techniques and results and also designed and developed a new innovative feature extraction technique which will reduce the system processing time and space complexity and give better recognition rate as compared to other existing techniques.*

**Keywords:** Feature extraction, Hash table, Linked list, Co-occurrence probabilities, Satellite images

## I. INTRODUCTION

Problems of pattern recognition and image processing have gained interest in past twenty years. Due to this a lot of researchers are developing methods those are experimental and also software and hardware to be used in the design of pattern recognition and image processing systems. Pattern recognition field has become more popular sine 1960's.Pattern recognition includes many methods used for different application in various fields.Application areas of pattern recognition are like bioinformatics, image analysis, data mining, remote sensing, speech recognition, GIS and many more. To get the desired results features have to be extracted and then analyzed for recognition and classification process.

## II. DEFINITIONS OF PATTERN RECOGNITION

Authors Duda and Hart in 1973 defined the pattern recognition is a field concerned with machine recognition of meaning regularities in noisy of complex environments. [1]Pavlidis defined pattern recognition in his book: "the word pattern is derived from the same root as the word patron and, in his original use, means something which is set up as a perfect example to be imitated. Thus pattern recognition means the identification of the ideal which a given object was madeafter." [2].In 1978Gonzalez and Thomas defined pattern recognition as a classification of input data via extraction important features from a lot of noisy data. [3]Researcher Fukunaga in 1990 defined pattern recognition as" A problem of estimating density functions in a high-dimensional space and dividing the space into the regions of categories of classes."[5].1992(Schalkoff) defined Pattern Recognition as"The science that concerns the description or classification (recognition) of measurements"[6].1993(Srihari,Govindaraju) defined pattern recognition as a discipline which learn some theories and methods to design machines that can recognize patterns in noisy data or complexenvironment. [7]1996(Ripley) outlined pattern recognition in his book: "Given some examples of complex signals and the correct decisions for them, make decisions automatically for a stream of future examples" [8].2002 (Robert P.W. Duin) described the nature of patternrecognition is engineering; the final aim of Pattern recognition is to design machines to solve the gap between application andtheory. [9].2003(Sergios Theodoridis,) Pattern recognition is a scientific discipline whose aim is the classification of the objects into a lotof categories or classes. Pattern recognition is also a integral partin most machine intelligence system built for decision making.**[**10]

**2.1 Pattern Recognition**: The process is divided into different steps as given below

1. Data acquisition and sensing: Measurements of physical variables, Important issues: bandwidth, resolution, sensitivity, distortion, SNR, latency, etc.
2. Pre-processing: Removal of noise in data, Isolation of patterns of interest from the background.
3. Feature extraction: Finding a new representation in terms of features.
4. Model learning and estimation: Learning a mapping between features and pattern groups and categories.
5. Classification: Using features and learned models to assign a pattern to a category.
6. Post-processing: Evaluation of confidence in decisions, Exploitation of context to improve performance, Combination of experts.

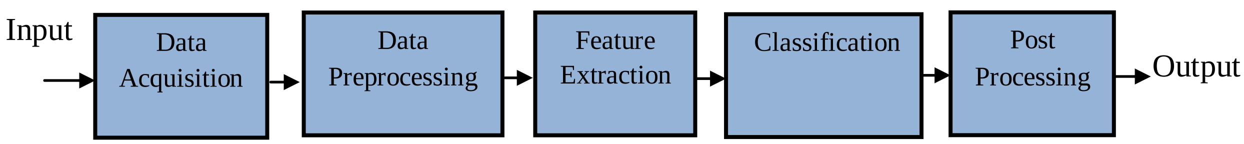


**Figure 1: Pattern Recognition System[12]**

## III. FEATURE EXTRACTION

Feature extraction is a key function in various image processing applications. A feature is an image characteristic that can capture certain visual property of the image. Mapping the image pixels into the feature space is known as feature extraction [13]. For automatic identification of the objects from satellite image data, they are to be associated with certain attributes which characterize them and differentiate them with each other. The similarity between images can be determined through features which are represented as vector [13]. Feature Extraction extracts various attributes of an object and associates it with feature vector and thencharacterizes it. Feature Extraction is an important steps to classify an image and identify the objects. There are various contents of an image such as color, texture, shape etc. It is used to represent and index an image or an object. Important feature is texture for many image types, which is the pattern of information or arrangement of the structure found in a picture. The feature is used in different applications such as image processing, remote sensing and content-based image retrieval. These features can be extracted in several ways. There are three methods used for texture feature extraction

1. Gray Level Co-Occurrence Matrix(GLCM)
2. Gray Level Co-occurrence Matrix Using Linked List (GLCMLL)
3. Gray Level Co-occurrence Matrix Hybrid Structure (GLCMHS)

## IV. STUDY OF EXISTING FEATURE EXTRACTION TECHNIQUES

The most common way is using a Gray Level Co-occurrence Matrix (GLCM). GLCM contains the second-order statistical information of neighboring pixels of an image. Textural properties can be calculated from GLCM to understand the details about the image content. However, the calculation of GLCM is very computationally intensive.[15]

### A. Gray level co-occurrence matrix (GLCM)

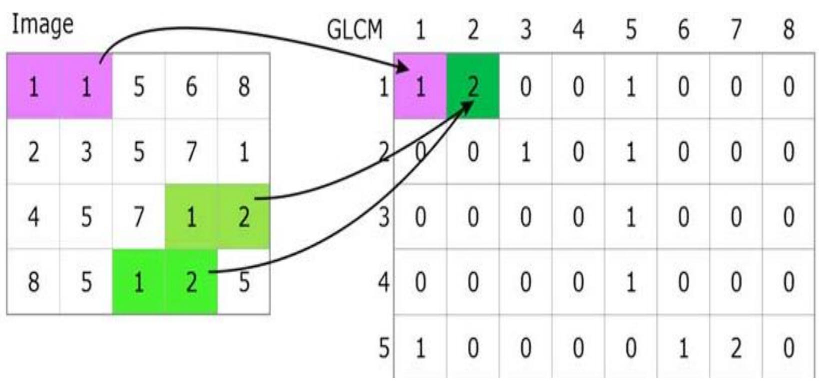


**Figure 2: Illustration of image (left) is transformed into the Gray level co-occurrence matrix (GLCM) [16]**

The overall computational complexity time is computation time of calculation the GLCM, normalization of the GLCM, and calculation of texture features. Most of the time is spent for calculation of GLCM. There are different methods for decreasing the GLCM time consumption. In one method the image is represented by four or five bits instead of eight bits that makes to reduce the size of GLCM but it makes to remove some information about the image. Another method is that to reduce the size of GLCM by storing just non-zero values. Clausi and Jernigant [17] describe the gray Level Co-occurrence Linked List (GLCLL).

### B. Gray Level Co-occurrence Matrix Using Linked List (GLCMLL)

In GLCMLL just non-zero values of GLCM are stored in a linked list, and each linked list node contains the two co-occurring gray-values, the co-occurrence probability of these two pair gray-values, and a link to the next node. When a new pair (i, j) comes, first there is a search for finding i, if it is found then there is a search for j. If there is a (i, j) in the list, their probability is increased, else the new node is added to the list. The GLCLL increases the calculation time.

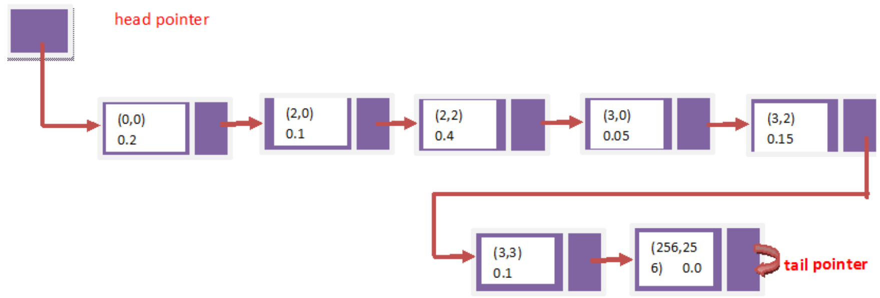


**Figure 3:GLCLL structure for determining image texture features.**

### C. Gray Level Co-occurrence Matrix Hybrid Structure (GLCMHS)

In 2001, Clausi and ZhaoThe [18] Gray Level Co-occurrence Hybrid Structure (GLCHS) that is based on an integrated hash table and linked list. Each node of the linked list includes two integer elements to store the gray-value pairs and two pointers to the previous and the next node. In the hash table, one element stores the probability of the GLCM and another stores the linked list pointer. Access to the hash table is provided by using (i, j). Each entry in the hash table has a pointer. A null pointer indicates that a particular co-occurring pair (i, j) does not have a representative node. Each new node inserted at the end of the linked list and its gray level values would be set. If the pointer is not null, then the probability of the existing node on the linked list is increased. The hash table allows rapid access to an (i, j). GLCHS is faster than GLCLL, and is useful for large image but it results increased a complexity of implementation due to a two dimensional hash table with longer linked list [18, 17].

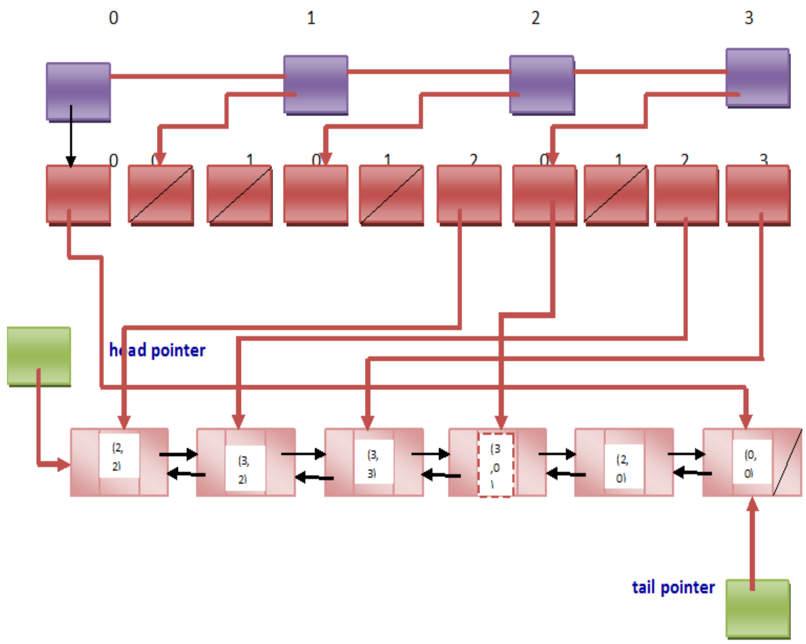


**Figure 4: GLCHS Structure For Determining Image Texture Features**

## V. LIMITATIONS WITH EXISTING GLCM TECHNIQUE

As we discussed above about GLCM , computation time of calculation the GLCM, normalization of the GLCM and calculation of texture features of GLCM is quite complex and hence Clausi and Jernigant [17] proposed the gray Level Co-occurrence Linked List (GLCLL). They tried to decrease complexity and system processing time using linked list but as we know that linked list internally uses pointers and it always leads to memory leaks issues. And memory leak issues will always leads to low performance of the system and more run time object creation. Hence to overcome system thoroghput issue, Clausi and ZhaoThe [18] proposed Gray Level Co-occurrence Hybrid Structure (GLCHS) that is based on an integrated hash table and linked list. But as hash table is thread safe in nature and hence unable to provide good throughput in concurrent environment. It will always leads to system crash if the system is getting used by huge number of users at the same time. Hence to overcome all the above issues, for the present research work we have designed, developed new enhanced GLCM feature extraction technique which is called as eGLCM

## VI. PROPOSED FEATURE EXTRACTION TECHNIQUE

**Enhanced Gray Level Co-occurrence Matrix – Using Integrated approach of Hashmap and LinkedList (eGLCM)**

Enhanced Gray Level Co-occurrence (eGLCM) that is based on an integrated hash map and linked list. Each node of the linked list includes two integer elements to store the gray-value pairs and two pointers to the previous and the next node.

In the hash map, one element stores the probability of the GLCM and another stores the linked list pointer. Access to the hash map is provided by using key element i.e. the probability of GLCM. Each key in the hash map has a value i.e. linked list pointer. A null key indicates that a particular co-occurring pair of key value does not have a representative node or value. Each new node inserted at the end of the linked list and its gray level values would be set. If the key is not null, then the probability of the existing node on the linked list is increased. The hash map allows rapid access to get the value of key pair. As we know that Hashmap is non-synchronized, it does not take more time as compared to hashtable processing time. Hash map also gives better result in concurrent environment in term memory and CPU utilization.

Hence for the present research work hash map has been selected to integrate with linked list for GLCM and named as Enhanced Gray Level Co-occurrence Matrix (eGLCM). To verify the accurate results of eGLCM, proposed eGLCM technique has been implementing using following JAVA API.

```
HashMap<String,LinkedList<String>>testMap=newHashMap<int,LinkedList<String>>();
```

## VII. RESEARCH METHODOLOGY

To develop enhanced GLCM feature extraction technique, following architecture has been designed, developed and used for the feature extraction process of pattern recognition system

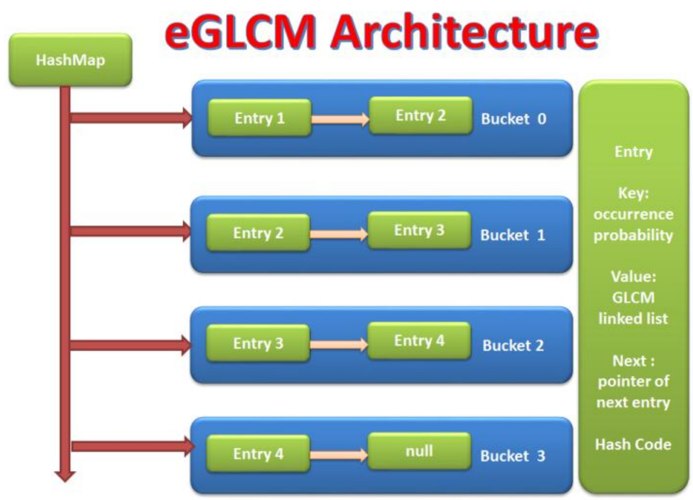


**Figure 5: eGLCM Structure for Determining Image Texture Features**

eGLCM architecture is mainly based on integrated functionality of GLCM, linked list and Hash map. As we know that Hash map contains key and value pair and hence in the HashMap Entry, four members are defined. Each instance or entry of the HashMap represents a node on the single linked list. The two integer members (x1,x2) store the grey level pairs. Two self-referential pointers are defined to access previous (*prev) and next (*next) ListNode nodes. Linked list nodes are defined to represent the first (head) and the last (tail) nodes. In the hash map entry, one float member (key) stores the grey level co-occurrence probability and the other stores the linked list pointer (*list ptr). The list ptr points to the corresponding node on the linked list associated by the grey level pair

## VIII. RESULTS ANALYSIS AND INTERPRETATION

| Feature Extraction methods | Computational Time (in miliseconds) | | | | | | |
|---|---|---|---|---|---|---|---|
| | Grey Level G | Baren | Water | Rly-Line | Built-up | Forest | Agriculture |
| eGLCM method | 256 | 26 | 118 | 280 | 400 | 700 | 986 |
| | 128 | 27 | 110 | 210 | 363 | 520 | 632 |
| | 64 | 25 | 80 | 138 | 211 | 250 | 290 |
| | 32 | 19 | 50 | 72 | 93 | 95 | 102 |
| | 16 | 12 | 22 | 29 | 33 | 35 | 40 |
| GLCHS method | 256 | 30 | 124 | 291 | 497 | 734 | 990 |
| | 128 | 29 | 113 | 239 | 383 | 536 | 662 |
| | 64 | 26 | 89 | 158 | 228 | 261 | 303 |
| | 32 | 21 | 54 | 77 | 99 | 99 | 109 |
| | 16 | 15 | 27 | 34 | 39 | 43 | 46 |
| GLCLL | 256 | 79 | 351 | 835 | 1547 | 2420 | 3509 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| method | 128 | 76 | 322 | 718 | 1242 | 1778 | 2354 |
| | 64 | 69 | 256 | 484 | 703 | 899 | 1049 |
| | 32 | 56 | 157 | 232 | 288 | 332 | 376 |
| | 16 | 39 | 75 | 95 | 110 | 123 | 135 |
| GLCM method | 256 | 564.5 | 214 | 217 | 997 | 298 | 22410.5 |
| | 128 | 5397 | 5480 | 5591 | 5735 | 5889 | 5979 |
| | 64 | 1374 | 1434 | 1510 | 1545 | 1589 | 1625 |
| | 32 | 362 | 394 | 417 | 427 | 439 | 451 |
| | 16 | 102 | 115 | 124 | 129 | 137 | 144 |

**Table 1: Average Computation Time Required Determining Texture Features for Each of GLCHS, GLCLL, GLCM GLCHS and eGLCM**

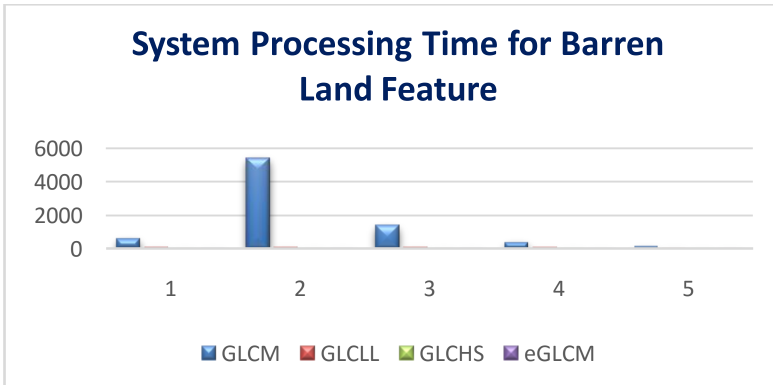


**Figure 6: Chart showing the System Processing Time for Barren Land Feature**

In our research we tried reducing the system processing time used for feature extraction using different GLCM techniques. In the above table it shows the time required for different features of satellite images like barren land, water, railway line, agriculture and built up areas. We have also shown the system time processing required for feature extraction by bar chart graph. It clearly shows that the blue bar which represents the GLCM method requires more time compared to GLCLL represented in red bar. When we tried GLCHS techniques represented in green color it gave better result than GLCLL techniques. But when we studied and implemented our innovative technique eGLCM it required lesser processing time for barren land feature as compared to other techniques used for feature extraction from satellite imagery.

Our techniques can also be applied to get system processing time for other features like railway lines, built-up areas, agriculture and water etc.

## IX. CONCLUSION

In the present research work we have studied the existing methods used for feature extraction like GLCM, GLCLL and GLCHS. We implemented the methods and from the results interpretation given above in table 1 proved that existing GLCM techniques requires more time and space for system processing. Hence to overcome this problem, new and enhance GLCM feature extraction technique is designed and developed. This new feature extraction technique is called as Enhanced Gray Level Co-occurrence Matrix (eGLCM) which takes less system processing time as compared to other existing GLCM techniques.It also gives better recognition rate then other GLCM techniques in term of pattern recognition system.

## REFERENCES


1. Editorial, Advances in Pattern Recognition, Pattern Recognition Letters 26,395-398,2005
2. T. Pavlidis. Structural Pattern Recognition, Springer Verlag,Berlin Heidelberg New York, 1977.
3. Gonzalez,R.C.Thomas,M.G .Syntatic Pattern Recognition: An Introduction (Addison Wesley, Reading, MA, 1978).
4. Watanabe, Pattern Recognition: Human and Mechanical. Wiley, New York, 1985
5. K. Fukunaga. Introduction to statistical patternrecognition (2nd ed). Academic Press, Boston 1990
6. RJ Schalkoff. Pattern Recognition: Statistical, Structural and Neural Approaches. John Wiley & Sons,1992
7. Srihari, S.N., Covindaraju, Pattern recognition, Chapman & Hall, London, 1034-1041, 1993
8. B. Ripley, Pattern Recognition and Neural Networks,Cambridge University Press, Cambridge, 1996
9. Robert P.W. Duin, Structural, Syntactic, and Statistical andPattern Recogition : Joint Iapr International workshops Sspr2002 and Spr 2002, Windsor, Ontario, Canada, August 6-9,2002 Proceedings
10. SergiosTheodoridis, Konstantinos Koutroumbas , patternrecognition , Pattern Recognition ,Elsevier(USA)),1982
11. Jie Liu[1, 2],Jigui Sun[1, 2],Shengsheng Wang[1, 2], "Pattern Recognition: An overview", IJCSNS International Journal of Computer Science and Network Security, VOL.6 No.6, June 2006.

12. Jayashree R. Pansare, Shravan H. Gawande, Maya Ingle, “Real-Time Static Hand Gesture Recognition for American Sign Language (ASL) in Complex Background” Journal of Signal and Information Processing, Vol.3 No.3(2012), Article ID:22132,4 pages.
13. Bhuravarjula, Hari Hara Pavan Kumar, and VNS Vijaya Kumar. "A Novel Image Retrieval Method using Color Moments."International Journal Of Electronics And Computer Science Engineering (Ijecse, ISSN: 2277-1956) 1.04 (2012): 2432-2438.
14. S. Belongie, J. Malik, and J. Puzicha, “Shape matching and object recognition using shape contexts,” Pattern Anal. Mach. Intell. IEEE Trans. On, vol. 24, no. 4, pp. 509–522, 2002.
15.http://cepublications.et.tudelft.nl/publications/300_implementing_texture_feature_extraction_algorithms_on_fpga.pdf
16.https://www.researchgate.net/figure/295249792_fig3_Figure-3-Illustration-of-how-the-image-left-is-transformed-into-the-gray-level.
17. D. A. Clausi and M. E. Jernigan, A Fast Method to Determine Co-occurrence Texture Features Using a Linked List Implementation, Remote Sensing of Environment 36 (1996), no. 1, 506–509.
18. D. A. Clausi and Y. Zhao, Rapid Determination of Co-occurrence Texture Features, IEEE Int. Geoscience and Remote Sensing Symposium 4 (2001), 1880–1882.