Abstract

We are in an age often referred to as the information age. In this information age, because we believe that information leads to power and success, from the technologies such as computers, satellites, etc., we have been collecting tremendous amounts of information. Our ability to analyze and understand massive datasets lags far behind our ability to gather and
store data. Image and video data contains abundant, rich information for data miners to explore. On one hand, the rich literature on image and video data analysis will naturally provide many advanced methods that may help mining other kinds of data. On the other hand, recent research on data mining will also provide some new, interesting methods that may benefit image and video data retrieval and analysis. Today, a lot of data is available everywhere but the ability to understand and make use of that data is very less. Whether the context is business, medicine, science or government, the datasets themselves are of little value. What is of value is the knowledge that can be inferred from the data and put to use. We need systems which would analyze the data for us. This paper basically aims to find out important pixels of an image using one of the classification technique named as decision tree (ID-3). Our aim is to separate the important and unimportant pixels of an image using simple rules. Further one of the compression techniques named as Huffman algorithm is applied for image compression. Finally, resultant image is stored with lesser space complexity.

Reference

- Jiawei Han. 2008. Data Mining for Image/Video Processing: A promising research frontier. Department of Computer Science, University of Illinois at Urbana-Champaign.
- J.Han and M. Kamber. 2006. Data Mining: Concepts and Techniques (2nd Ed.). Morgan Kaufmann.
- Jiawei Han. 2009. Research challenges for data mining in science and engineering. Department of Computer Science and Engineering, University of Illinois at Urbana-Champaign.
Commun. ACM 38, Issue11, pp55-64.

**Index Terms**

Computer Science

Information Science

**Key words**

KDD

Data Mining

Image Processing

Compression Ratio

Information Gain