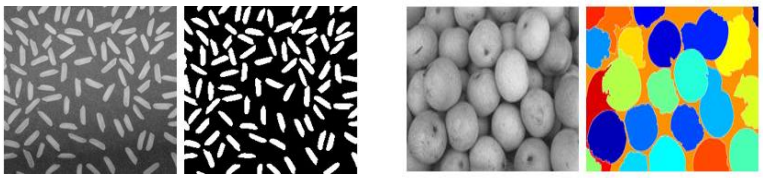
Topic : Image Segmentation.

Q. What is Image segmentation?

Image segmentation is the process of partitioning an image into multiple segments. The goal of segmentation is to simplify and/or change the representation of an image into something that is more meaningful and easier to analyze.

This is typically used to identify objects or other relevant information in images. There are many different ways to perform image segmentation, including:

I) Thresholding methods
II) Color-based Segmentation
III) Edge detection methods
IV) Texture methods



Q. What are the two basic properties on which image segmentation algorithm are based?

Image segmentation algorithms are based on one of the two basic properties of pixel intensity values, **discontinuity** and **similarity**.

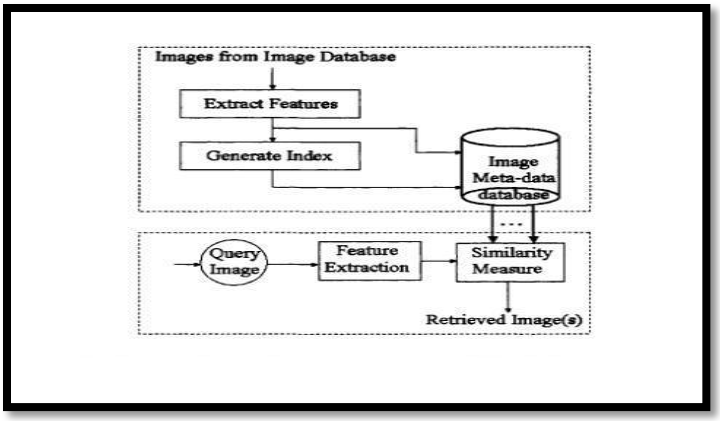
Discontinuity approach is to partition an image based on abrupt changes in pixel intensity values, these detects edges in an image. Edge detection reduces the size of the information while preserving the necessary structural information.

Second approach stated that partitioning image into regions that are **similar** according to a set of predefine criteria. Thresholding is the example of this method.

Topic: Similarity based retrieval.

Retrieval from Multimedia database: it is complex process to retrieval the data from multimedia database as it requires a large amount of browsing, filtering and searching techniques. We can search and retrieve images by manually (by manual annotation method) or automatic. If we search large database like multimedia information (audio, video etc) then manual searching or text based search concept is too time consuming. If we focus on CBIR technique that deals with the content of the Image or multimedia information will retrieve better result.

- ✖ **What is CBIR (content based image retrieval)?**
- ❑ "Content based" means that the search will analyze the content of the image rather than the meta-data, such as keywords, tags etc.
- ❑ The term 'Content' might refer to the color, shape, textures or any other information derived from the image.

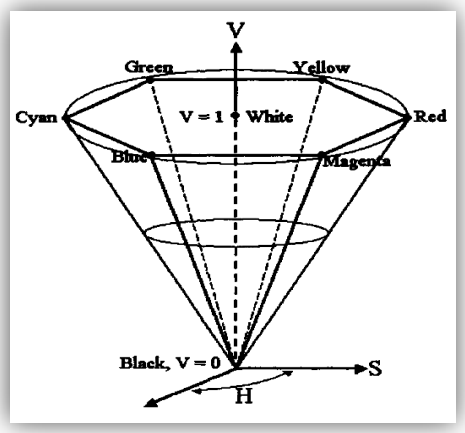


Topic: Image retrieval by colors:-

- A color pixel in a digital image is represented by three color channels (usually Red, Green and Blue).
- The HSV color space is widely used in the CBIR community.
- Hue of a color represents the relative color appearance; that is, 'redness', 'greenness,' and so on,
- Value indicates the darkness of the color (or the perceived luminance), and
- Saturation represents the strength of the color.

HSV Color space:

- ✖ The 'Value' (V) varies along the vertical axis of the cone.
- ✖ The 'Hue' (H) varies along the periphery of the circle of the cone and is represented as an angle about the vertical axis,
- ✖ The 'Saturation' (S) varies along the radial distance



Multimedia Database

Q. What is Multimedia Data?

There are number of data types that can be characterized as multimedia data types. These are typically the elements for the building blocks of ore generalized multimedia environments, platforms, or integrating tools. The basic types can be described as follows :

• **Text:** The form in which the text can be stored can vary greatly. In addition to ASCII based files, text is typically stored in processor files, spreadsheets, databases and annotations on more general multimedia objects.
• **Images**: There is great variance in the quality and size of storage for still images. Digitalized images are sequence of pixels that represents a region in the user's graphical display. The space overhead for still images varies on the basis of resolution, size, complexity, and compression scheme used to store image. The popular image formats are jpg, png, bmp, tiff.
• **Audio**: An increasingly popular data type being integrated in most of applications is Audio. Its quite space intensive. One minute of sound can take up to 2-3 Mbs of space. Several techniques are used to compress it in suitable format.
• **Video**: One on the most space consuming multimedia data type is digitalized video. The digitalized videos are stored as sequence of frames. Depending upon its resolution and size a single frame can consume upto 1 MB. Also to have realistic video playback, the transmission, compression, and decompression of digitalized require continuous transfer rate.
• **Graphic Objects**: These consists of special data structures used to define 2D and 3D shapes through which we can define multimedia objects. These includes various formats used by image, video editing applications. Examples are CAD / CAM objects

Q. How is Multimedia Data Different?

Conceptually it should be possible to treat multimedia data in the same way as data based on the data types (e.g. numbers, dates and characters).However, there are few challenges that arises from multimedia data.

- ❑ The content of multimedia data is often captured with different "capture" techniques (e.g., image processing). Multimedia processing techniques need to be able to handle different ways of content capture including automated ways and/or manual methods.
- ❑ Queries posed by the user in multimedia databases often cannot come back with a textual answer. Rather, the answer to a query may be a complex multimedia presentation that the user can browse at his/her relaxation.
- ❑ Multimedia data is large and affects the storage , retrieval and transmission of multimedia data.
- ❑ In case of video and audio databases time to retrieve information may be critical ex(Video on demand).

- Automatic feature extraction and Indexing: In conventional databases user explicitly submits the attribute values of objects inserted into the database. In contrast, advanced tools such as image processing and pattern recognition tools for images, to extract the various features and content of multimedia objects. As size of data is very large we need special data structures for storing and indexing.

Q. Basic Approaches for Data Retrieval?

- **Conventional database system :** This is the widely-used approach to manage and search for structured data. All data in a database system must conform to some predefined structures and constraints (i.e., schema's). To formulate a database query the user must specify which data objects are to be retrieved, the database tables from which they are to be extracted and predicate on which the retrieval is based. A query language for the database will generally be of the artificial kind, one with restricted syntax and vocabulary, such as SQL.

- **Information retrieval (IR) system**

  IR system is mainly used to search large text collections, in which the content of the (text) data is described by an indexer using keywords or a textual abstract, and keywords or natural language is used to express query demands. For example for an image or video we have to describe it in words or in a way need to store lot of metadata (textual form).

- **Content based retrieval (CBR) system**

This approach is used to retrieve desired multimedia objects from a large collection on the basis of features (such as color, texture and shape, etc.) that can be automatically extracted from the objects themselves. Although keyword can be treated as a "feature" for text data, traditional information retrieval has much more higher performance than content-based retrieval because keyword has the proven ability to represent semantics, while no features have shown convincing semantic describing ability. But major drawback of this method is that it lacks precision.

- **Graph or tree pattern matching**:

  This approach aims to retrieve object sub-graphs from an object graph according to some denoted patterns.

Advanced Data Structure to represent Multimedia data :

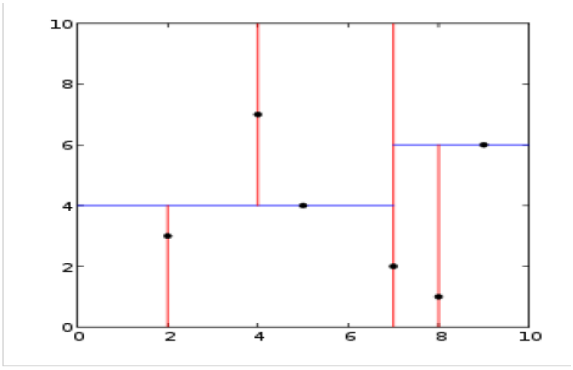
Many modern database applications deal with large amounts of multidimensional data. Examples include multimedia content-based retrieval (high dimensional multimedia feature data), time-series similarity retrieval, data mining and spatial/spatio-temporal applications.

Why are access methods important?

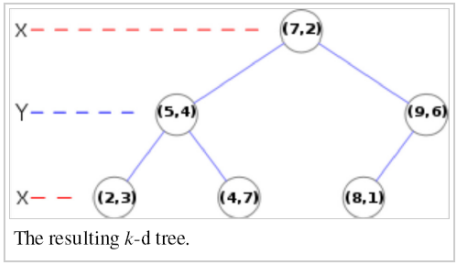
The main purpose is to is to support efficient spatial selection, for example range queries or nearest neighbor queries of spatial objects. Without a spatial index, every object in database need to be checked whether it meets selection criterion, i.e. complete linear scan of relational database. Clustering is needed to group those objects which are often requested together. Otherwise, many different disk pages will have to be fetched, resulting in slow response. For spatial selecting the clustering implies storing objects which are close together in reality also close together in the computer memory (instead of scattered over the whole memory).

k-d Trees

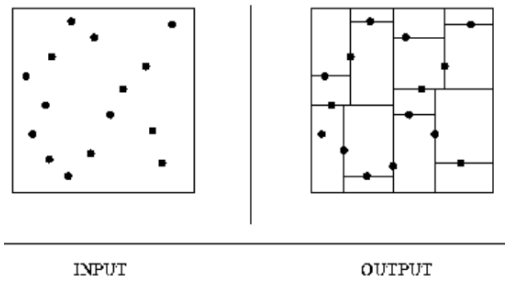
Early structure used for indexing in multiple dimensions. The k-d tree is used to store k-dimensional points data. For example Image may have many attributes such as spatial position, texture, color. So k-d trees can be used to represent such data.

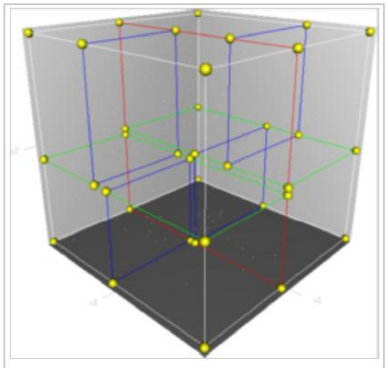


k-d tree decomposition for the point set (2,3), (5,4), (9,6), (4,7), (8,1), (7,2).



The resulting *k*-d tree.





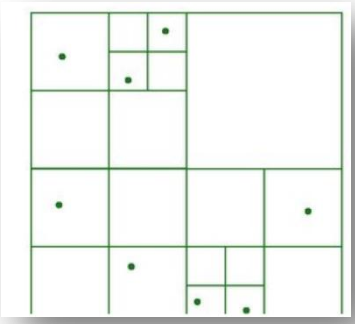
The *k*-d tree is a binary tree in which every node is a k-dimensional point. Every non-leaf node can be thought of as implicitly generating a splitting hyper plane that divides the space into two parts, known as half-spaces. Points to the left of this hyperplane are represented by the left subtree and right is called right subtree.

Division of Space by Quad-trees:

Each node of a quad-tree is associated with a rectangular region of space; the top node is associated with the entire target space. Each non-leaf nodes divides its region into four equal sized quadrants correspondingly each such node has four child nodes corresponding to the four quadrants and so on. Leaf nodes have between zero and some fixed maximum number of points

Common uses of Quad-trees are:

1. Image Representation
2. Spatial Indexing
3. Efficient collision detection in two dimensions.
4. Storing sparse data, such as a formatting information for a spreadsheet or for some matrix calculations .



Representing Image Using Quad-tree:

Let's say we divide the picture area into 4 sections. Those 4 sections are then further divided into 4 subsections. We continue this process, repeatedly dividing a square region by 4. We must impose a limit to the levels of division otherwise we could go on dividing the picture forever. Generally, this limit is imposed due to storage considerations or to limit processing time or due to the resolution of the output device. A pixel is the smallest subsection of the quad tree.

To summarize, a square or quadrant in the picture is either :

1. entirely one color
2. composed of 4 smaller sub-squares

To represent a picture using a quad tree, each leaf must represent a uniform area of the picture. If the picture is black and white, we only need one bit to represent the colour in each leaf; for example, 0 could mean black and 1 could mean white. Now consider the following image : The definition of a picture is a two-dimensional array, where the elements of the array are colored points

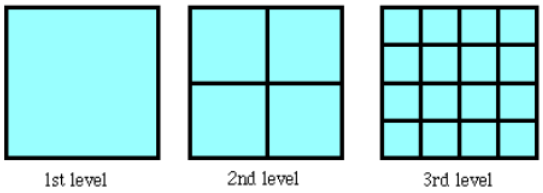


Figure 2.3: First three levels of quad tree

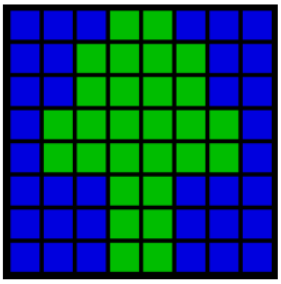


Figure 2.4: Given Image

Advantages of quad trees include:

1. Quad trees can be manipulated and accessed much quicker than other models.
2. Erasing a picture takes only one step. All that is required is to set the root node to neutral.
3. Zooming to a particular quadrant in the tree is a one step operation.
4. To reduce the complexity of the image, it suffices to remove the final level of nodes.
5. Accessing particular regions of the image is a very fast operation. This is useful for updating certain regions of an image, perhaps for an environment with multiple windows. The Main disadvantage is that it takes up lots of space.

R-Tree:

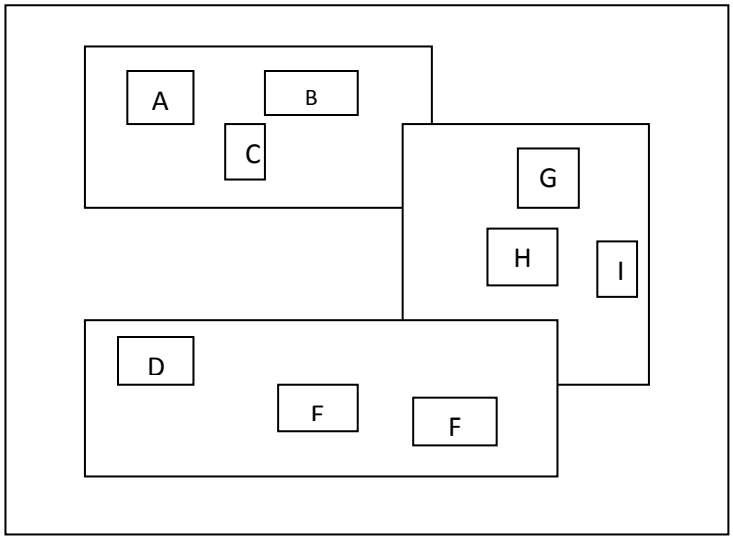
R-trees are tree data structures used for spatial access methods, i.e., for indexing **multi-dimensional information** such as **geographical coordinates,** rectangles or polygons.

The key idea of the data structure is to **group nearby objects** and represent them with their **minimum bounding rectangle** in the next higher level of the tree; the **"R" in R-tree is for rectangle**.

A **common real-world usage for an R-tree** might be to store spatial objects such as **restaurant locations or the polygons that typical maps are made of: streets, buildings, outlines of lakes, coastlines,** etc. and then find **answers quickly to queries** such as **"Find all museums within 2 km of my current location", "retrieve all road segments within 2 km of my location"** (to display them in a navigation system) or "find the nearest gas station" (although not taking roads into account).

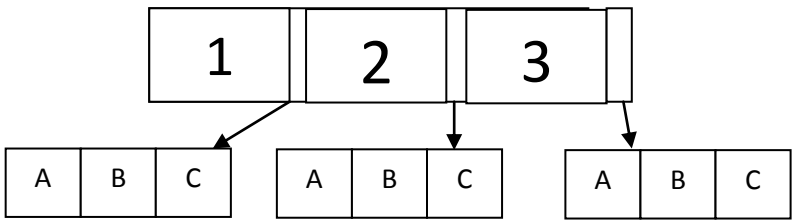
Since all objects lie within this bounding rectangle, a query that does not intersect the bounding rectangle also cannot intersect any of the contained objects. At the leaf level, each rectangle describes a single object; at higher levels the aggregation of an increasing number of objects.

The R-tree was proposed by Antonin Guttman in 1984and has found significant use in both theoretical and applied contexts.



BOUNDING BOX AND POLYGONS

R-TREE



Q. What is QBIC?

QBIC stands for Query by Image Content. It is Content Base Image Retrieval Techniques (CBIR). CBIR is also known as QBIC. CBIR or QBIC is the application of computer vision to image retrieval technique. This technique is applied to searching digital images in large database. Previously text based search is not so efficient that's why Content of the image based search is gaining more popularity. Different implementations of QBIC make use of different types of user queries:

i) Query by example
ii) Query by sketch.

Q. What is Virage?

Visual Information retrieval (VIR) Image engine analysis and automatically ingests/absorb and understands all rich media content and allows users to search extensive video asset with pinpoint accuracy. It offers a full spectrum of retrieval methods including search by audio, scene, key-frame, image etc.

Q. What is OLAP?

online analytical processing, or **OLAP,** it is an approach to quickly provide the answer to complex analytical quires. OLAP is part of the broader category of business intelligence, which also encompasses relational database, report writing and data mining.