UNIT-5

Information Retrieval

- **Information retrieval (IR)** is the science of searching information in documents, in textual data, searching for documents themselves, searching for metadata which describe documents, or searching within databases,

- There is a common confusion, however between data retrieval, document retrieval, information retrieval, and text retrieval, and each of these has its own bodies of literature, theory and technologies.

- IR is interdisciplinary, based on computer science, mathematics, library science, information science, information architecture, linguistics, statistics and physics.
- Automated IR systems are used to reduce information overload.

- Many universities and public libraries use IR systems to provide access to books, journals, and other documents.
- Web search engines such as Google, Yahoo search and Live Search (formerly MSN Search) are the most visible IR applications

Retrieval Process

- An information retrieval process begins when a user enters a query into the system.
- Queries are formal statements of information needs, for example search strings in web search engines.
- In information retrieval a query does not uniquely identify a single object in the collection. Instead, several objects may match the query, perhaps with different degrees of relevancy.

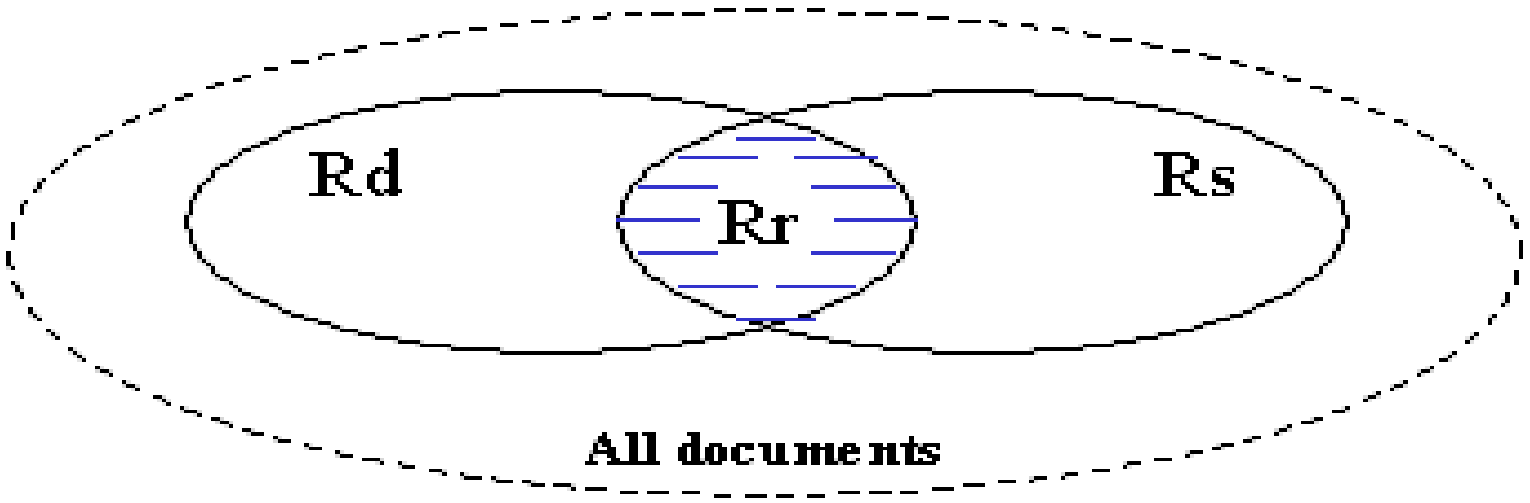
- An object is an entity which keeps or stores information in a database.
- User queries are matched to objects stored in the database.
- Depending on the application the data objects may be, for example, text documents, images or videos.

- Most IR systems compute a numeric score on how well each object in the database match the query, and rank the objects according to this value.
- The top ranking objects are then shown to the user.
- The process may then be iterated if the user wishes to refine the query.

Performance measures

- Many different measures for evaluating the performance of information retrieval systems have been proposed.
- The measures require a collection of documents and a query.
- All common measures described here assume a ground truth notion of relevancy: every document is known to be either relevant or non-relevant to a particular query.

Performance measures



Rd = relevant documents
Rr = retrieved relevant documents
Rs = response set

Recall = Rr/Rd
Precision = Rr/Rs

- The collection contains 2, preferably overlapping sub-sets:
- The set ***Rd*** of all documents in the collection that are relevant to a particular query, and
- The set ***Rs*** of all documents retrieved by the system in response to a query.
- The overlap, labeled ***Rr***, represents the relevant documents that were retrieved.
  Obviously, the goal is to have the highest possible overlap, i.e. that ***Rr*** is as large as possible.

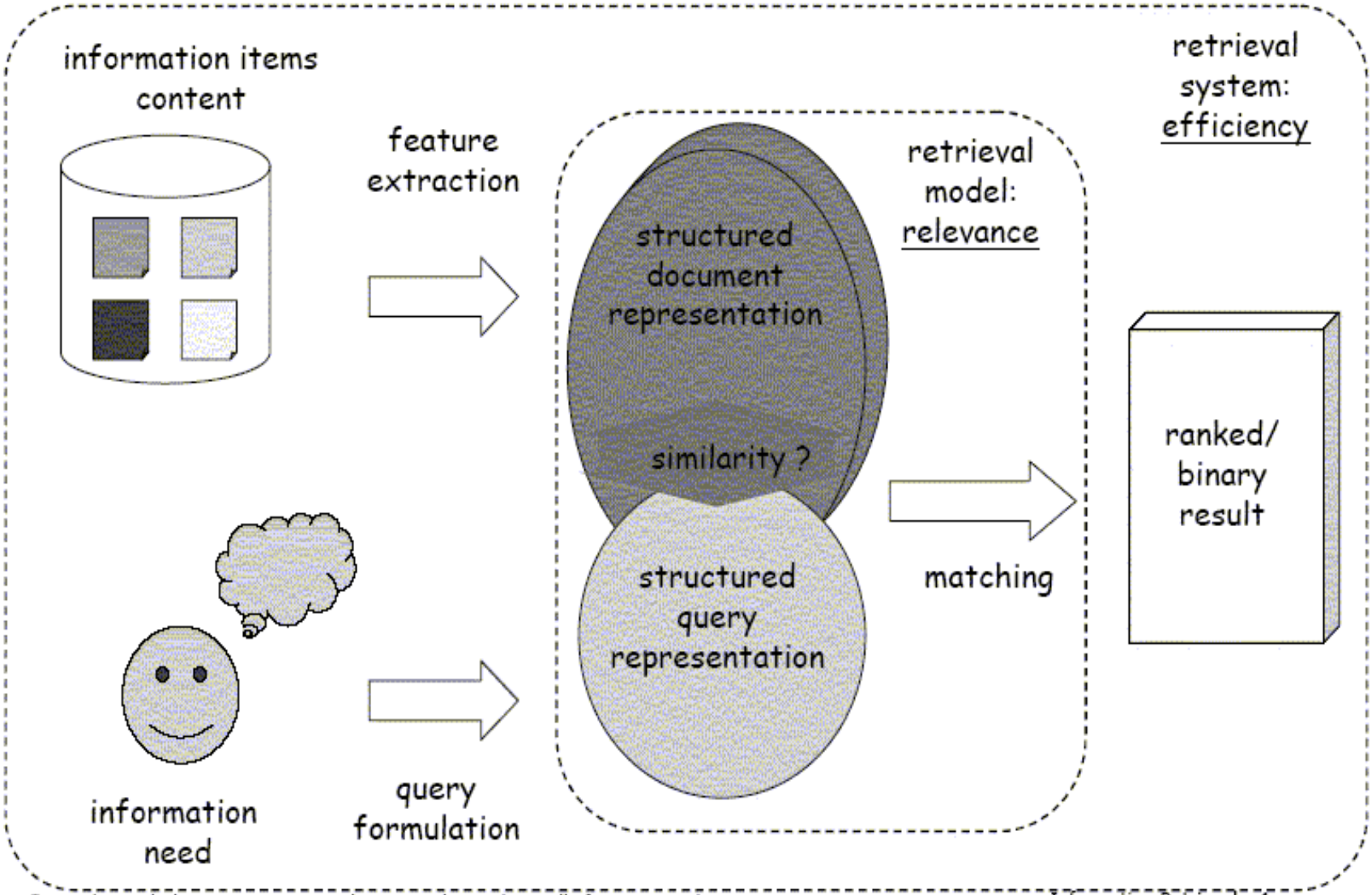
- The measures most commonly used in evaluation of information retrieval system performance are:

1. ***Recall***, ***R***, defined as the fraction of relevant documents retrieved (Rr) of all relevant documents (Rd). i.e.:

   R = Rr / Rd

2. ***Precision***, ***P***, defined as the fraction of relevant documents (Rr) in the response set (Rs). I.e.:

   P = Rr / Rs


information items
content
feature
extraction
retrieval
system:
efficiency
retrieval
model:
relevance
structured
document
representation
similarity ?
structured
query
representation
matching
ranked/
binary
result
information
need
query
formulation

F-measure

- The weighted harmonic mean of precision and recall, the traditional F-measure or balanced F-score is:
- F = 2(precision . Recall)/(precision+recall)
- This is also known as the F_1 measure, because recall and precision are evenly weighted.

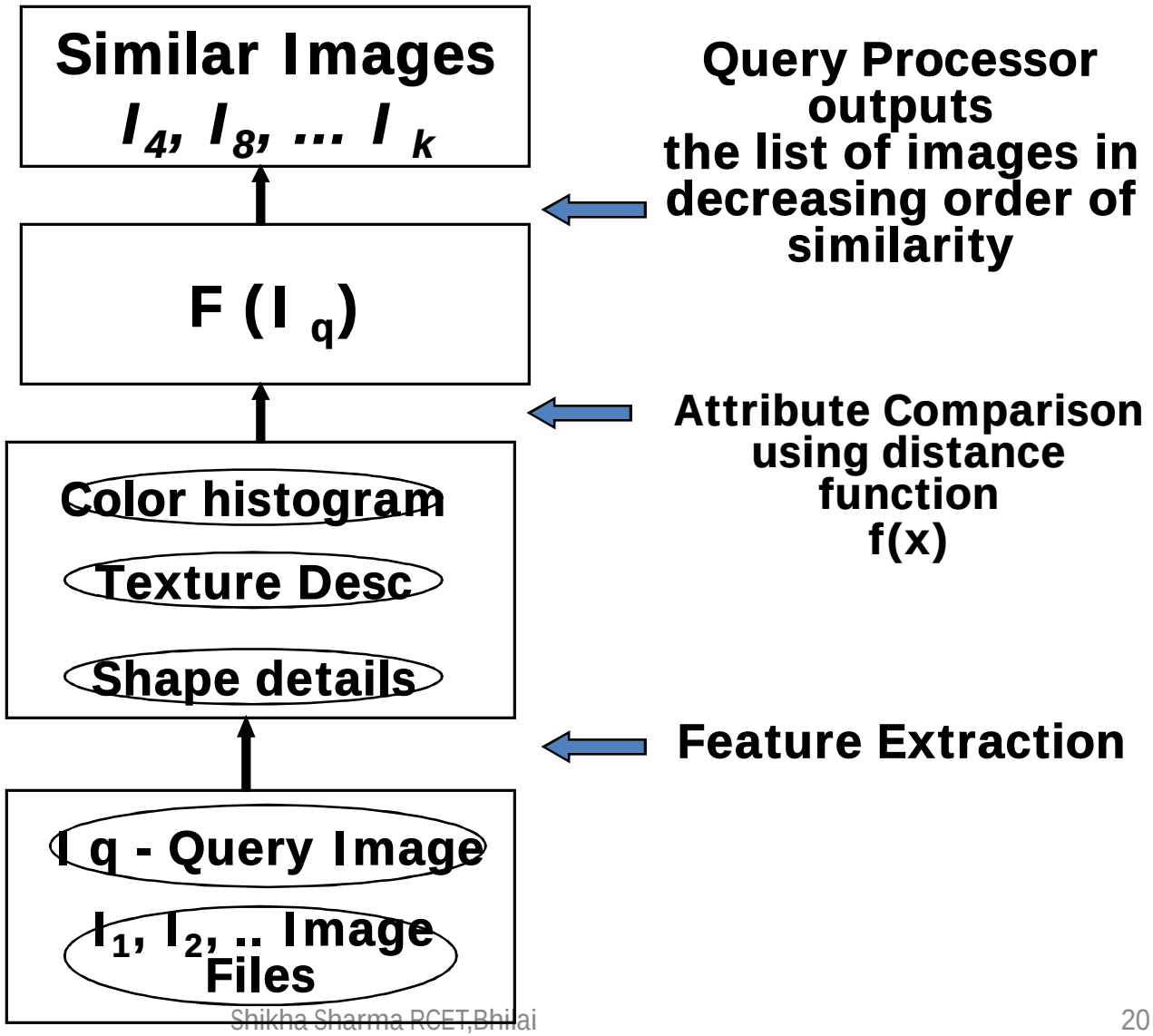
Image Retrieval

- An **image retrieval** system is a computer system for browsing, searching and retrieving images from a large database of digital images.
- Most traditional and common methods of image retrieval utilize some method of adding metadata such as captioning, keywords, or descriptions to the images so that retrieval can be performed over the annotation words.
- Manual image annotation is time-consuming, laborious and expensive; to address this, there has been a large amount of research done on automatic image annotation.

- The first microcomputer-based image database retrieval system was developed at MIT, in the February 1987 issue of IEEE Transactions on Industrial Electronics .
- Another method of image retrieval is content-based image retrieval (CBIR), which aims at avoiding the use of textual descriptions and instead retrieves images based on their visual similarity to a user-supplied query image or user-specified image features.

- Image retrieval phase
  - Query by example
  - Query by color
  - Query by shape
  - Query by Texture Similarity
  - Query by other category bit.

Typical CBIR System



Applications

The need to retrieve images from a collection is there in the following disciplines.

- Satellite Image databases, Geographic Information System (GIS)
- Architectural Design
- Crime prevention (Finger prints and face recognition)
- Journalism and TV production
- Art galleries and Museum Management
- Advertising and Publishing
- Fashion and Fabric design
- Medicine

Unfortunately not many from these disciplines know about CBIR.

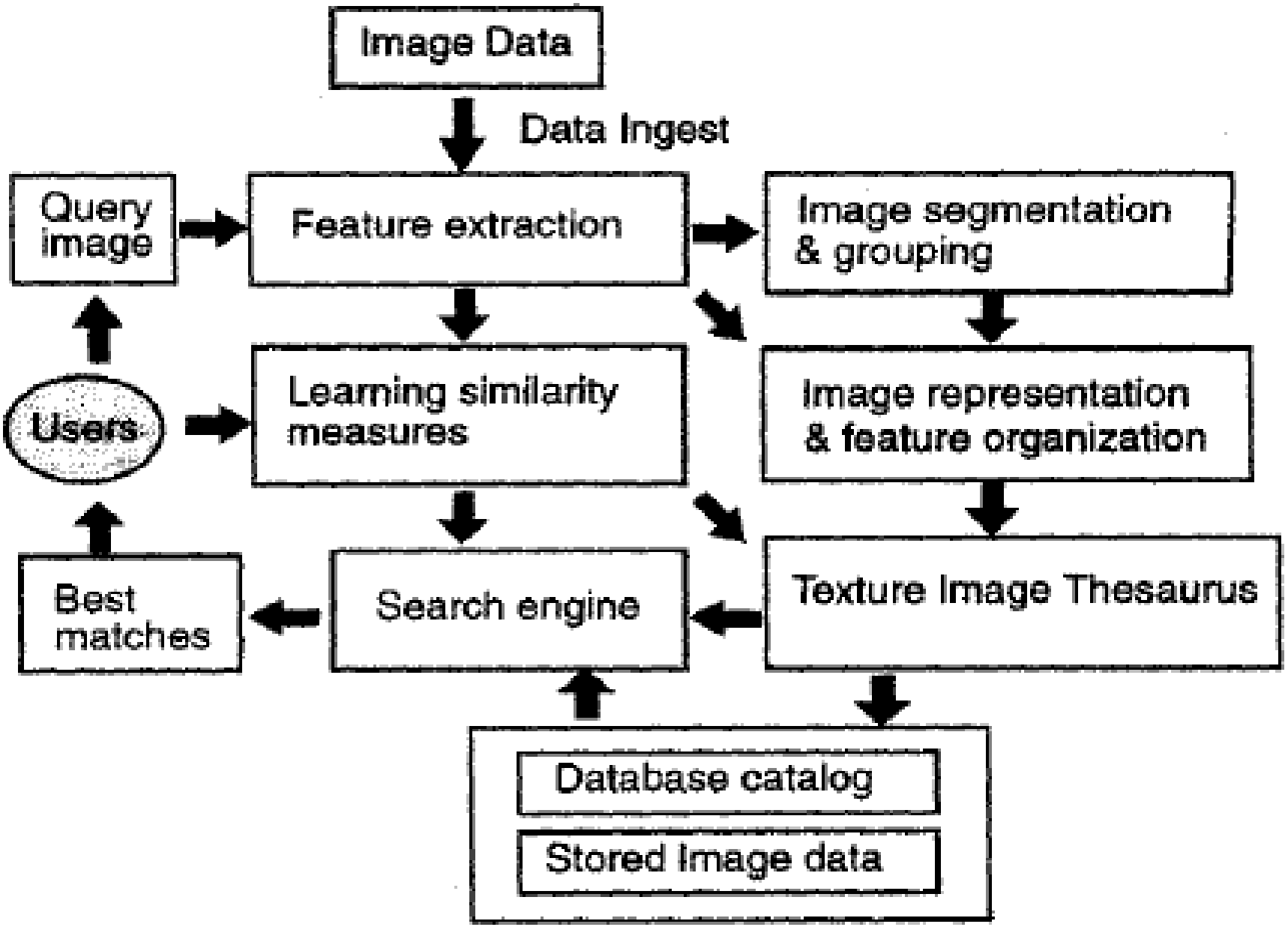
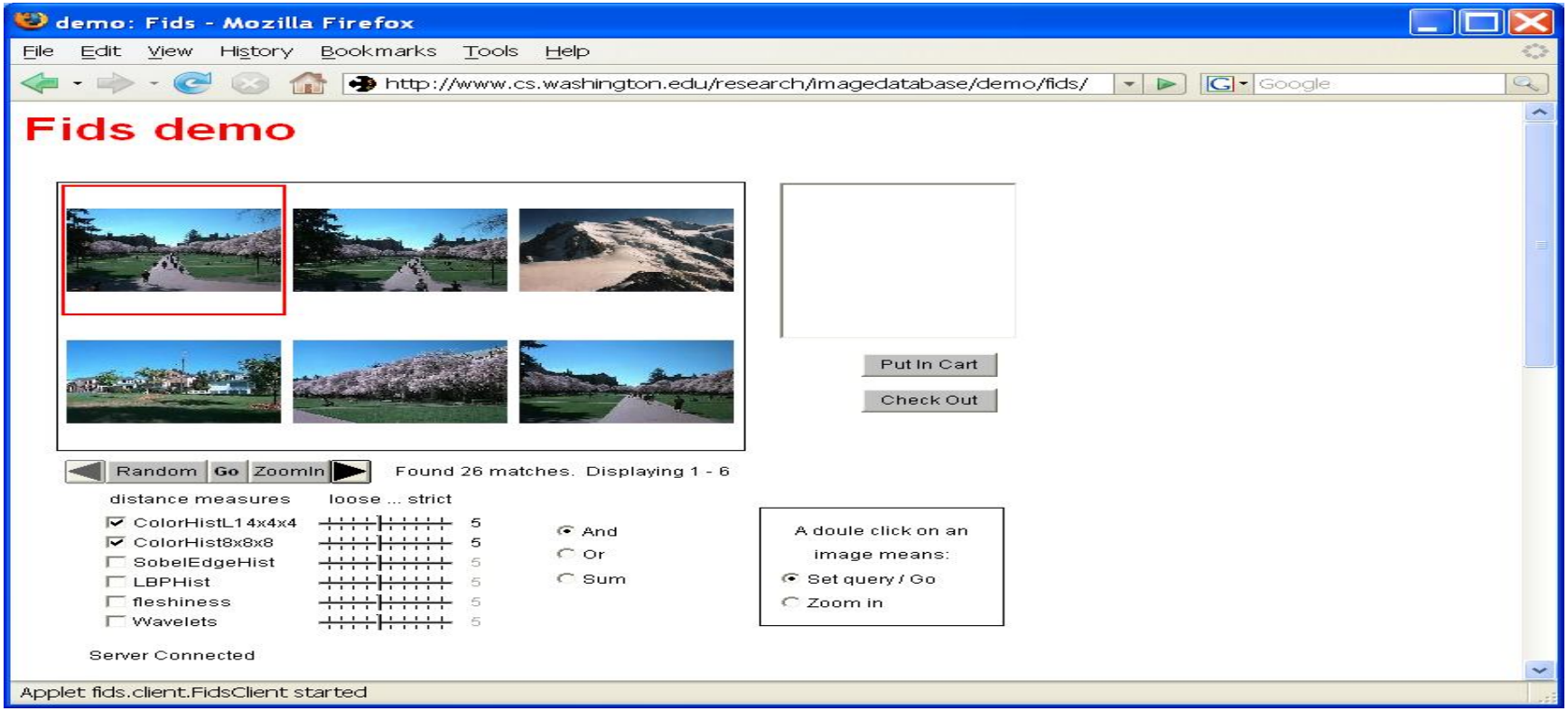


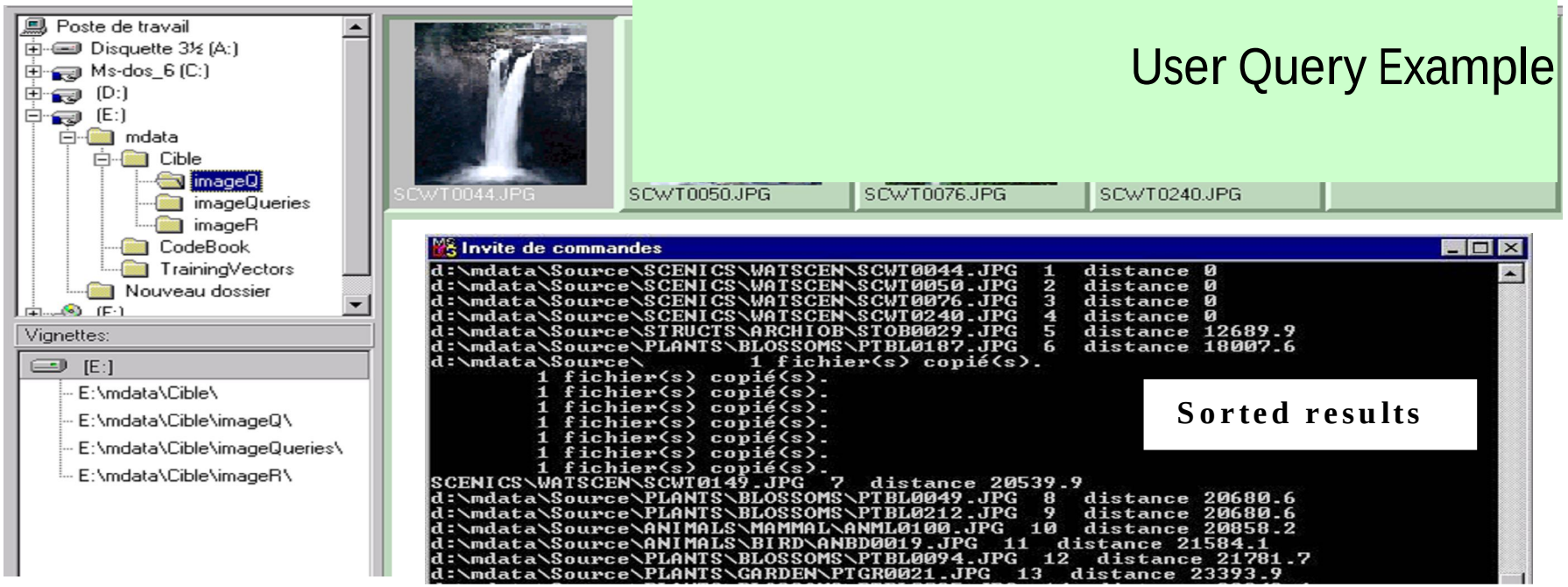
FIGURE 1. Content-based image retrieval system for browsing large-scale aerial photographs.

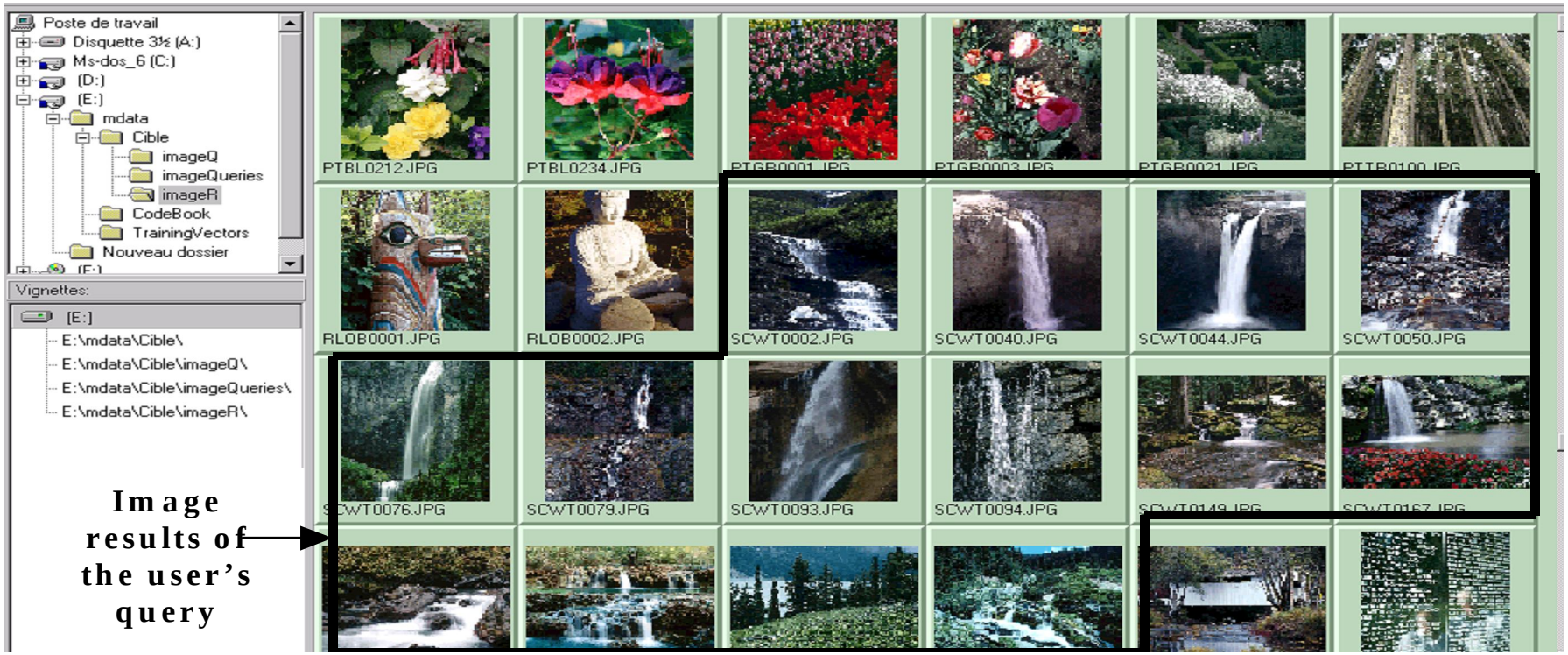
Image Retrieval Phase

- Query by example (QBE)
  - Allow to select sample image to search.



User Query Example

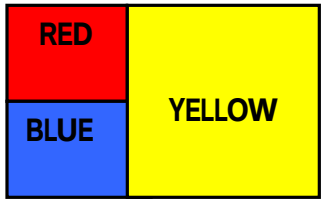


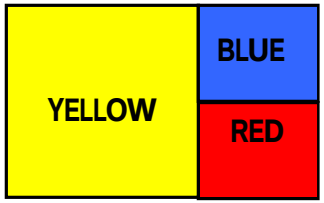


Retrieval by Similarities - Color Similarity

Color Similarity:

Color distribution similarity has been one of the first choices because if one chooses a proper representation and measure it can be partially reliable even in presence of changes in lighting, view angle, and scale.





Results (continued)

Retrieval result examples:

- Example 1: query mode, using color features

(a) (b) (c) (d) (e)

Query mode result sample: (a) The query image.

(b) -(e) The similar image returned by the retrieval system. They are ranked by the color similarity measurement, with (b) having the highest ranking.

CONTENT BASED IMAGE RETRIEVAL by COLOR

File Cluster Tools Help

Query Image

Retrieved Images



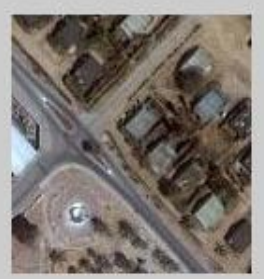







Image Retrieval Phase (cont.)

- Query by color anglogram
  - Histogram intersection measures is a fairly standard metric to analyze histogram base on features.
  - Image is divided into 5 sub-images, upper right, upper left, lower right, lower left, and the center image.

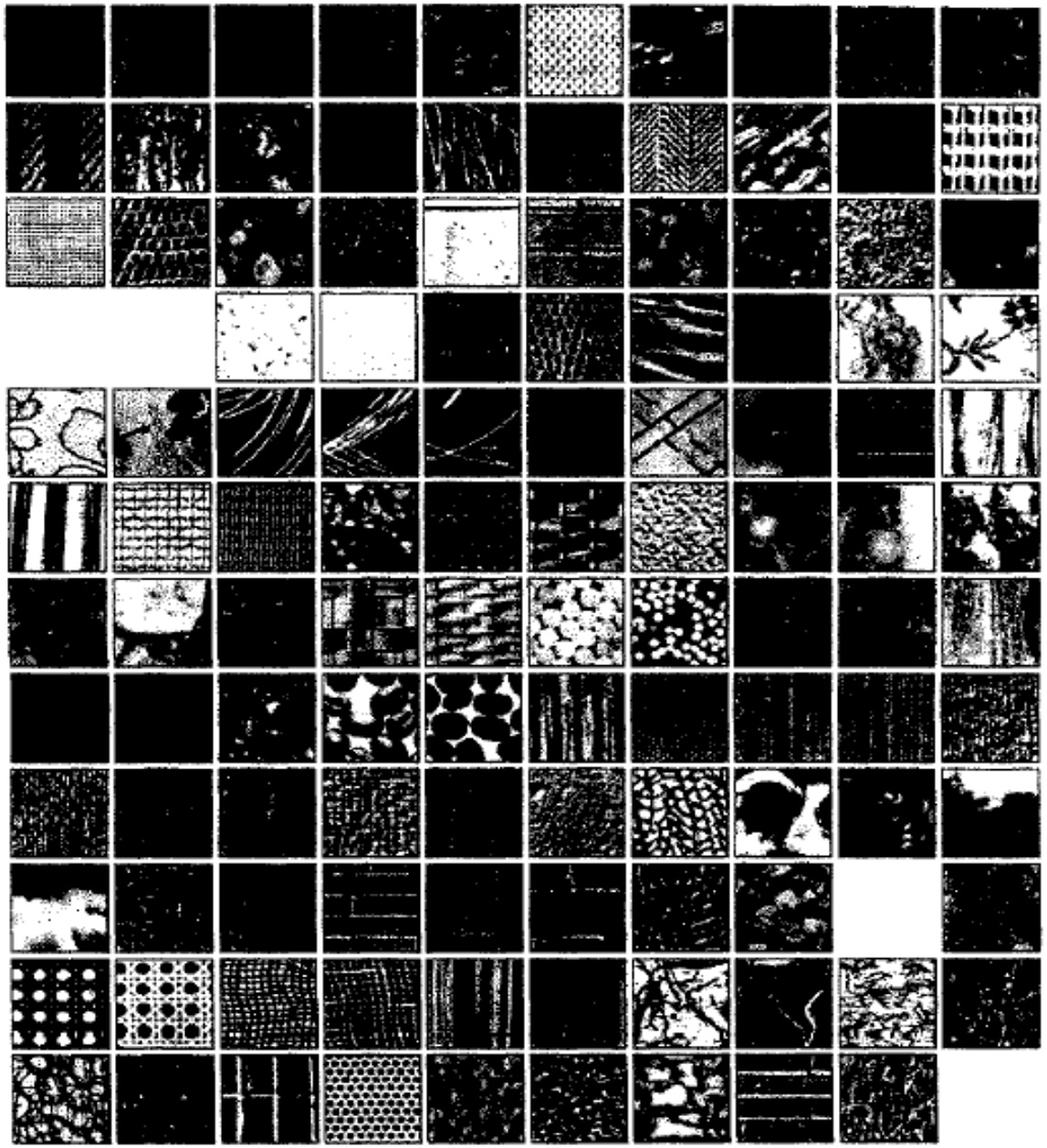
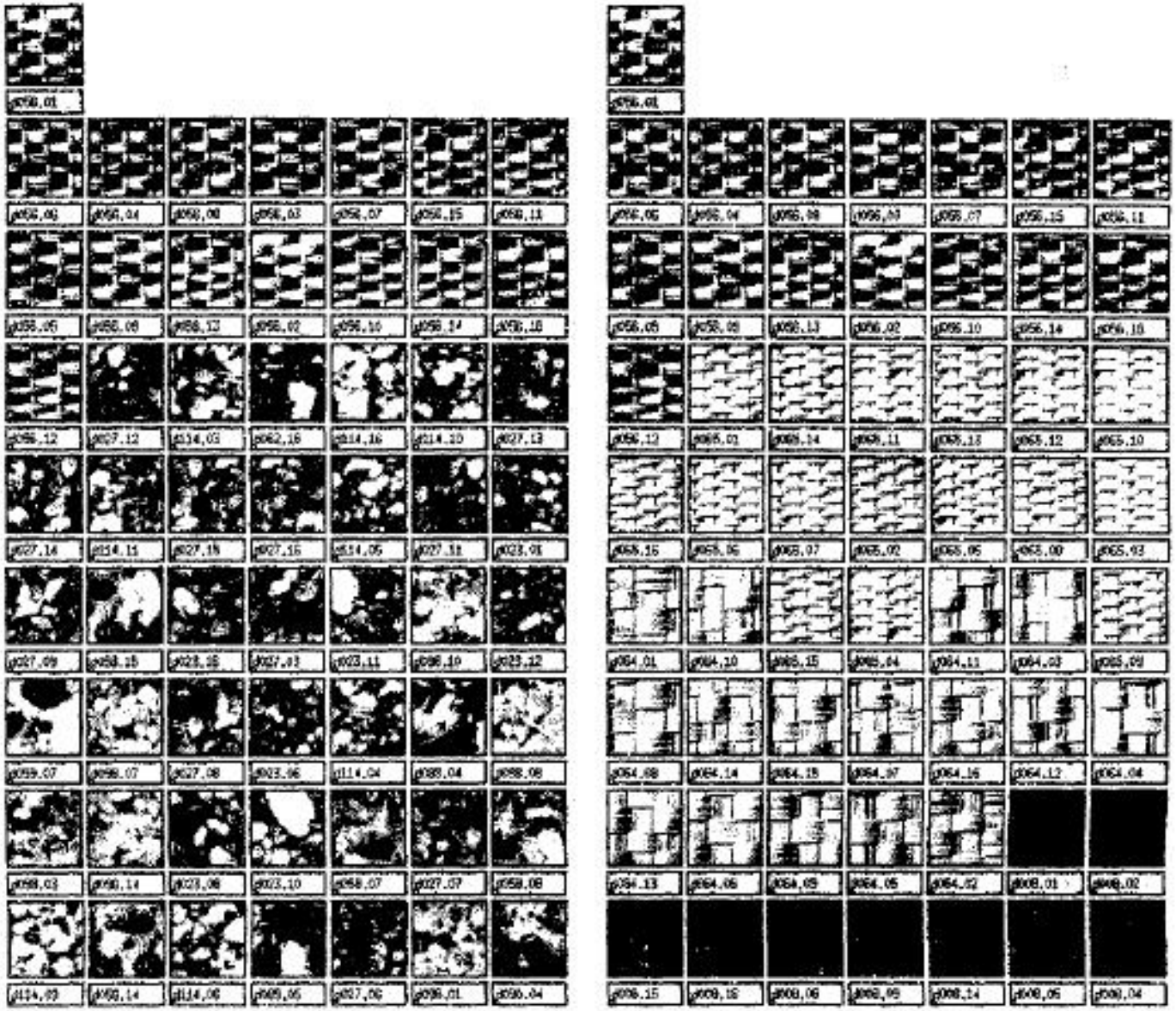


FIGURE 5. Example images of 116 different textures. From right to left and top to bottom are the D1-D112 textures from the Brodatz album. The last 7 textures are O1-O7 from the USC database. Missing D* textures are left blank (white space).



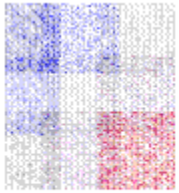
(a) before learning similarity measures (b) after learning similarity measures

Figure 3: Pattern retrieval with and without learning. Each query pattern has 15 other similar patterns in the database. The input query (d056.01) is shown at the top of the column in each case. With or without learning, the Gabor features provide a very good representation in retrieving all the other 15 images from the same texture class. However, note the degradation in visual similarity after that for the case without learning. The images are ordered according to decreasing similarity from left to right and top to bottom. For the case with learning similarity, the performance continues without any marked degradation in perceptual similarity, even after 50 patterns are retrieved.

Retrieval by Similarities - Texture Similarity

Texture Similarity:

- Texture reflects the texture of entire image.
- Texture is most useful for full images of textures, such as catalogs of wood grains, marble, sand, or stones.
- Texture images are generally hard to categorize using keywords alone because our vocabulary for textures is limited



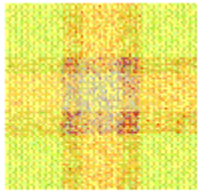


Image Retrieval Phase (cont.)

- Retrieve by shape
  - Each image is divided into 256 block.
  - Each block is approximated with hue and saturated value.
  - Corresponding feature points are mapped perceptually base on the saturated value.
  - Feature histogram is obtained by measure the largest angle of the nearest feature points.

Retrieval by Similarities - Shape Similarity

Shape Similarity:

- Shape represents the shapes that appear in the image.
- Shapes are determined by identifying regions of uniform color.
- Shape is useful to capture objects.
- Shape is very useful for querying on simple shapes.

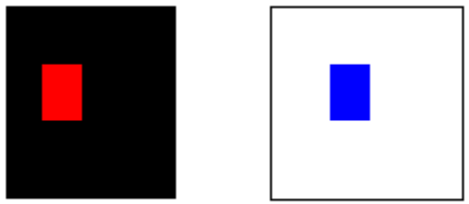


Image Retrieval Phase (cont.)

- Query by shape



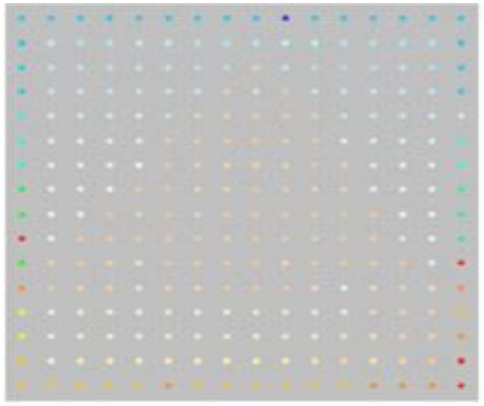
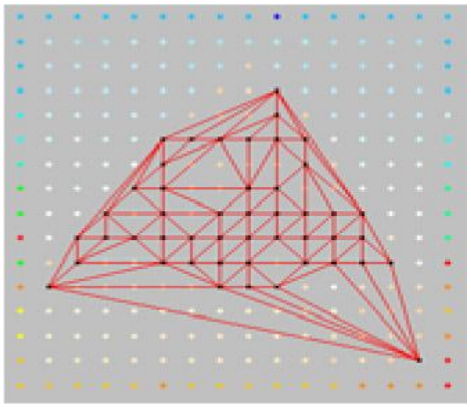
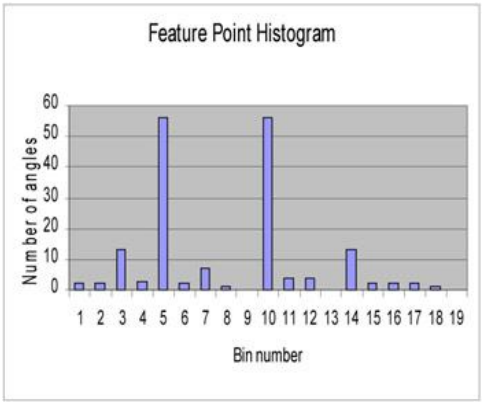
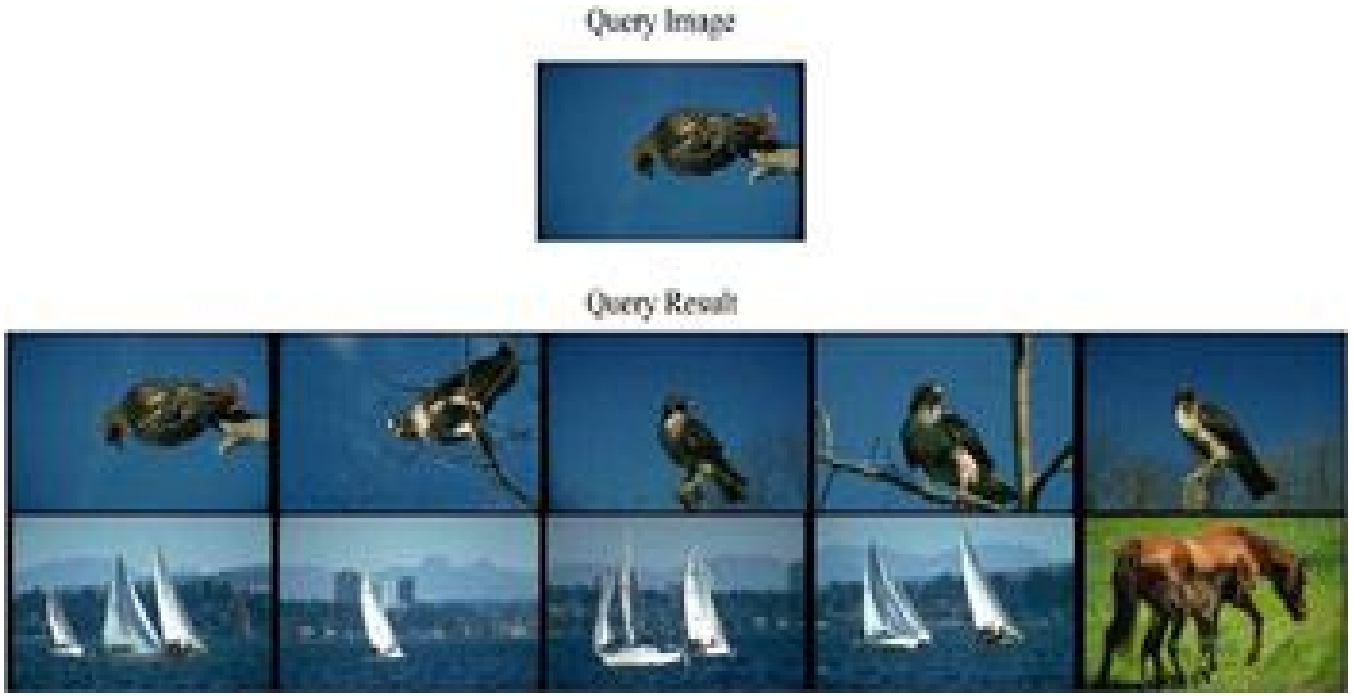





Image Retrieval Phase (cont.)

- Query by shape anglogram sample output:



Retrieval by Similarities - Spatial Similarity (1)

Spatial Similarity:

– Symbolic Image

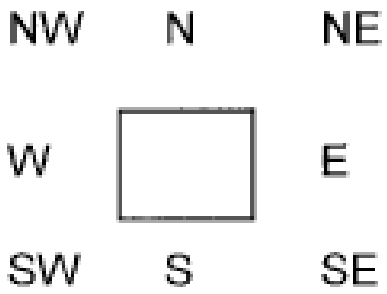
Spatial similarity assumes that images have been segmented into meaningful objects, each object being associated with is centroid and a symbolic name. This representation is called a symbolic image.

- Similarity Function

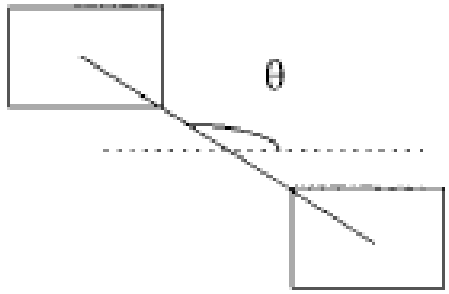
It is relatively easy to define similarity functions for such image modulo transformations such as rotation, scaling and translation.

Retrieval by Similarities - Spatial Similarity (2)

Directional Relations



(a) strict and mixed directional relations



(b) slope directional relations

Retrieval by Similarities - Spatial Similarity (3)

Topological Relationship

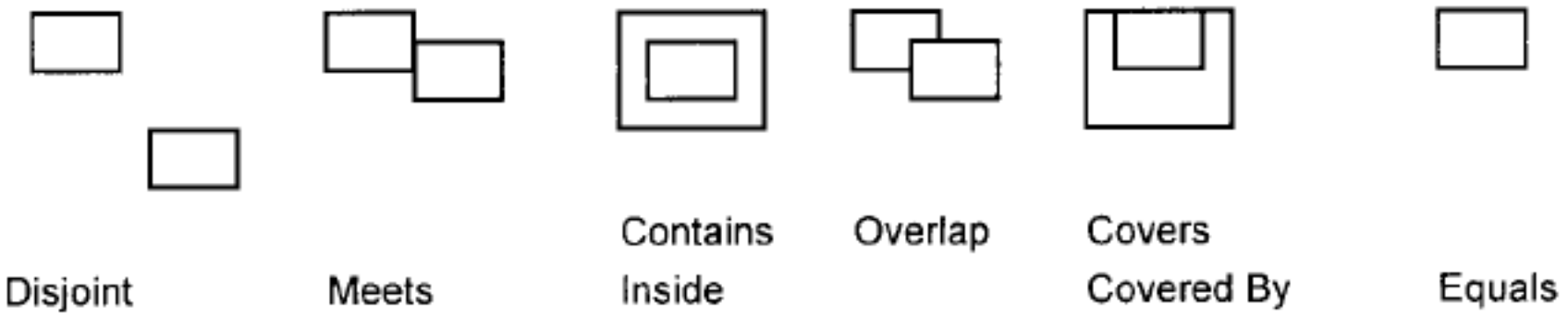


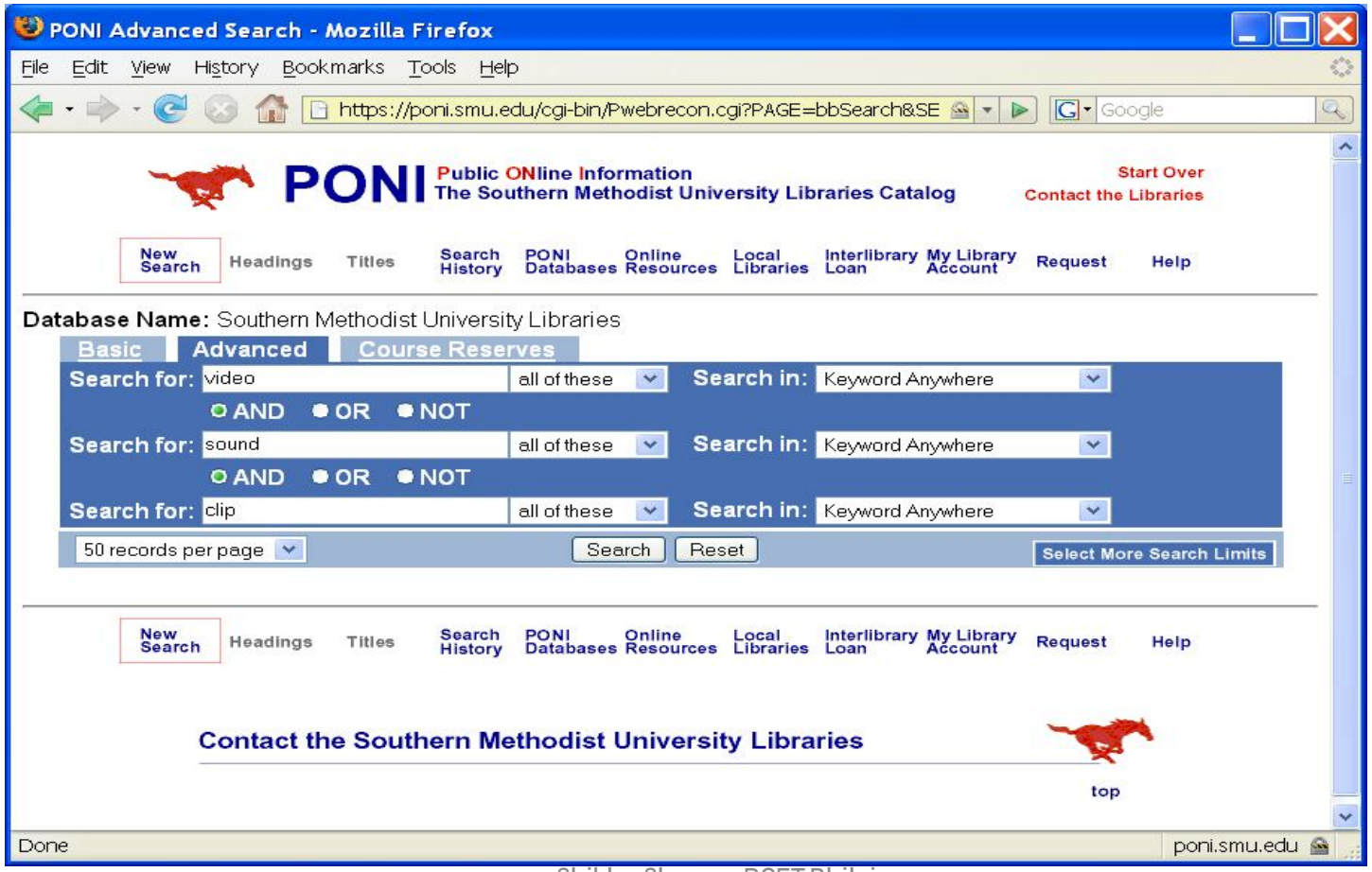
Image Retrieval Phase (cont.)

- Query by color and other category selection combination.
  - Use training dataset: sky, sun, land, water, boat, grass, horse, bird, human, pyramid, column, tower, and snow.
  - Sun(5%), grass (15%), Sky(20%) combine with the LSI matrix to return better results.

Content-Based Video Retrieval System

Introduction

- Traditional Library search method

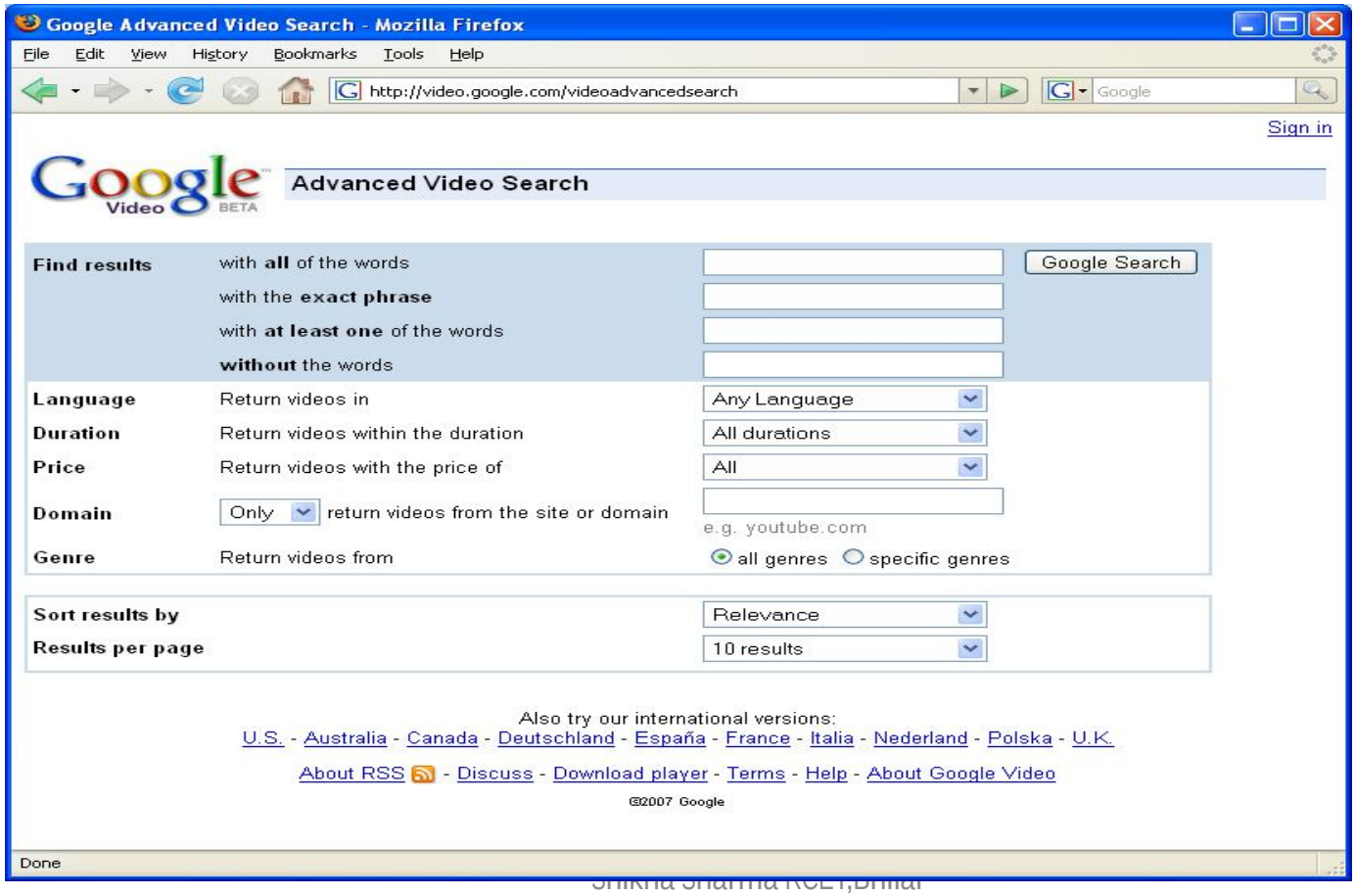


Introduction (cont.)

- Other search engines still using description search method.
- Current image search method: by description.

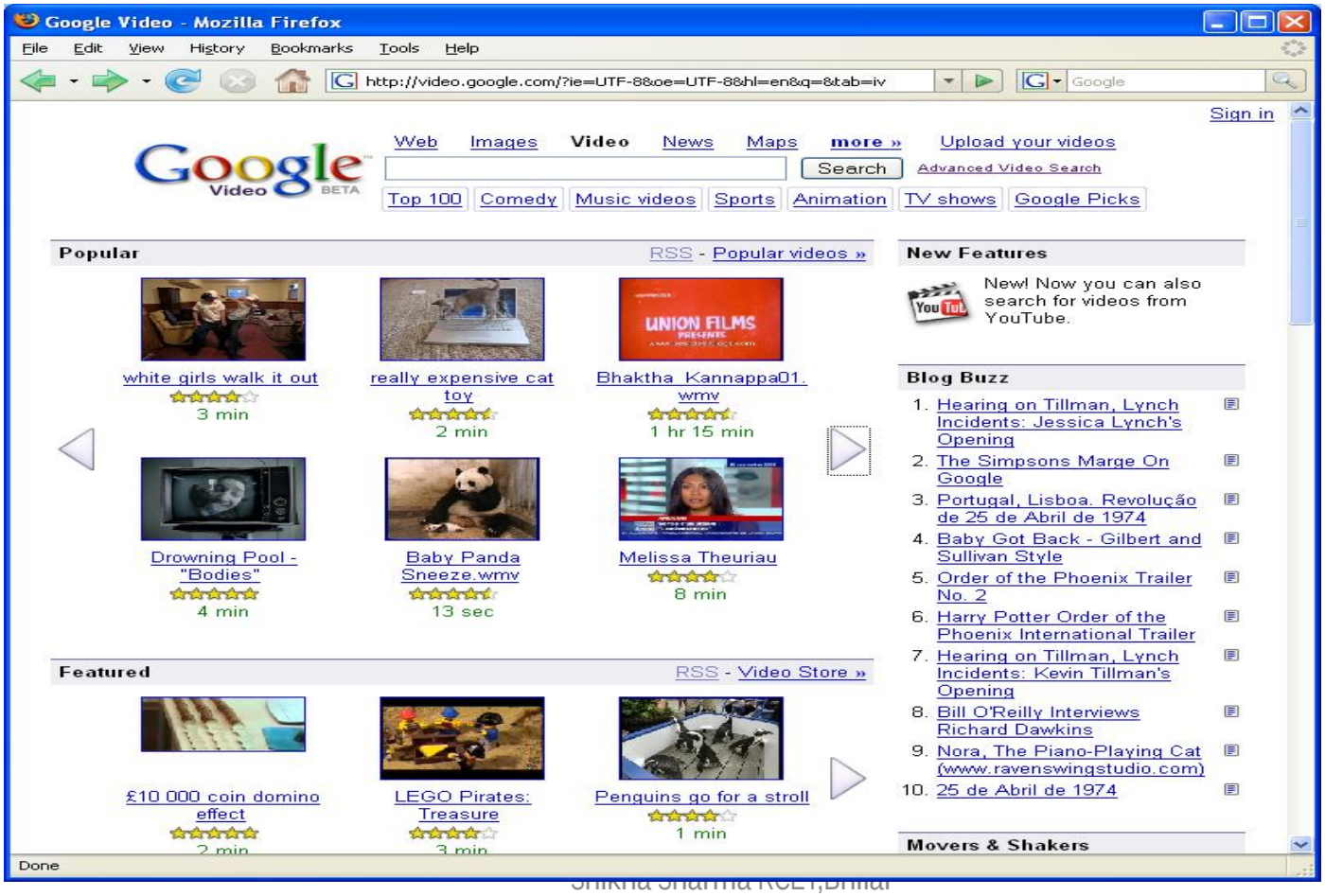
Introduction (cont.)

- Sample of Google Video Search:



Introduction (cont.)

- Google Video Archive selections:

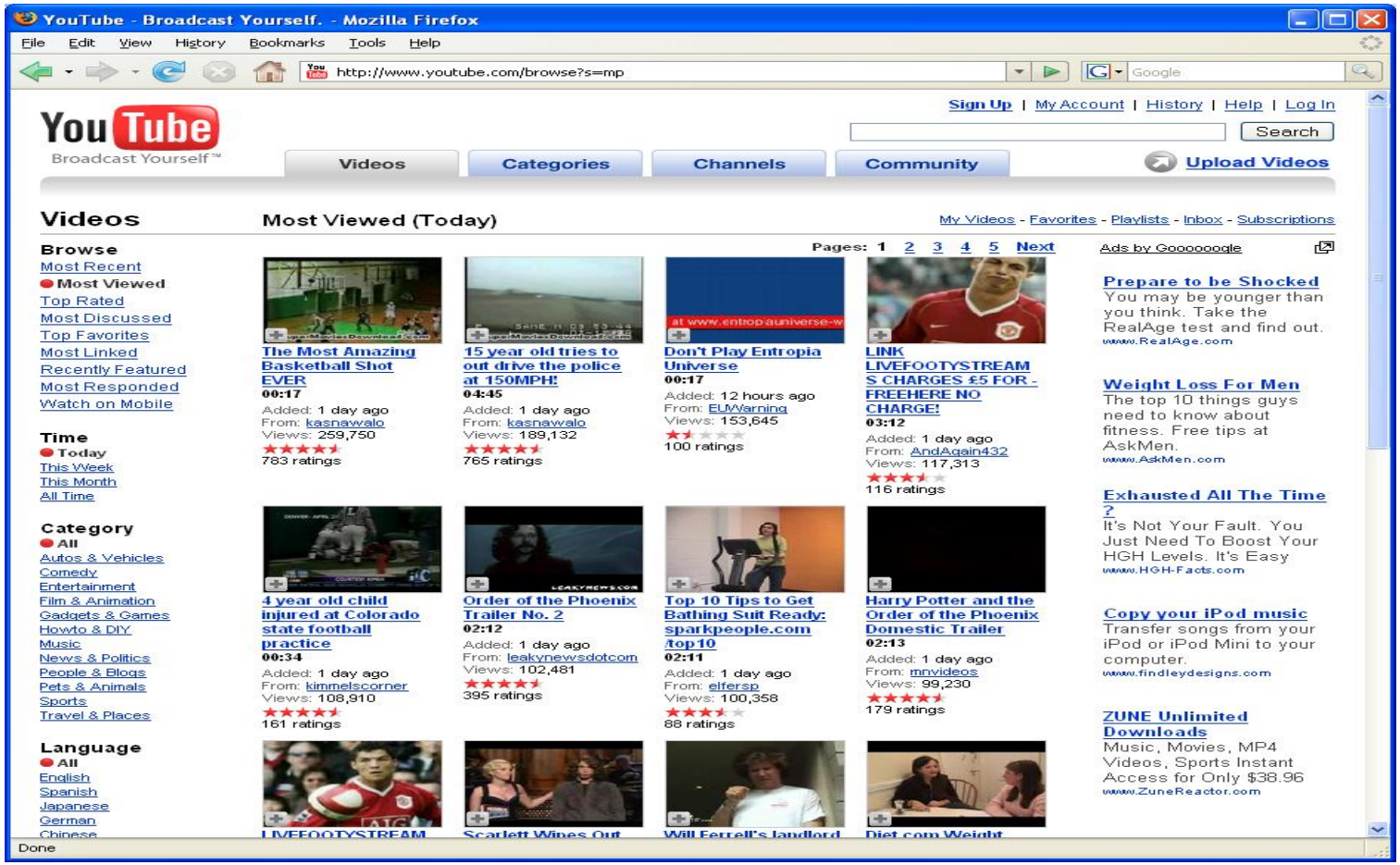


Introduction (cont.)

- Picture is worth a thousand words.
- More than words can express.
- Growing number video clips on MySpace and YouTube, there is a need for a video search engine.

Introduction (cont.)

- Sample YouTube Video page:



Introduction (cont.)

- Therefore, we need a better search technique – Content-Based Video Retrieval System (CBVR).

Introduction (cont.)

- What good is video retrieval?
  - Historical Achieve
  - Forensic documents
  - Fingerprint & DNA matching
  - Security usage

Overview (cont.)

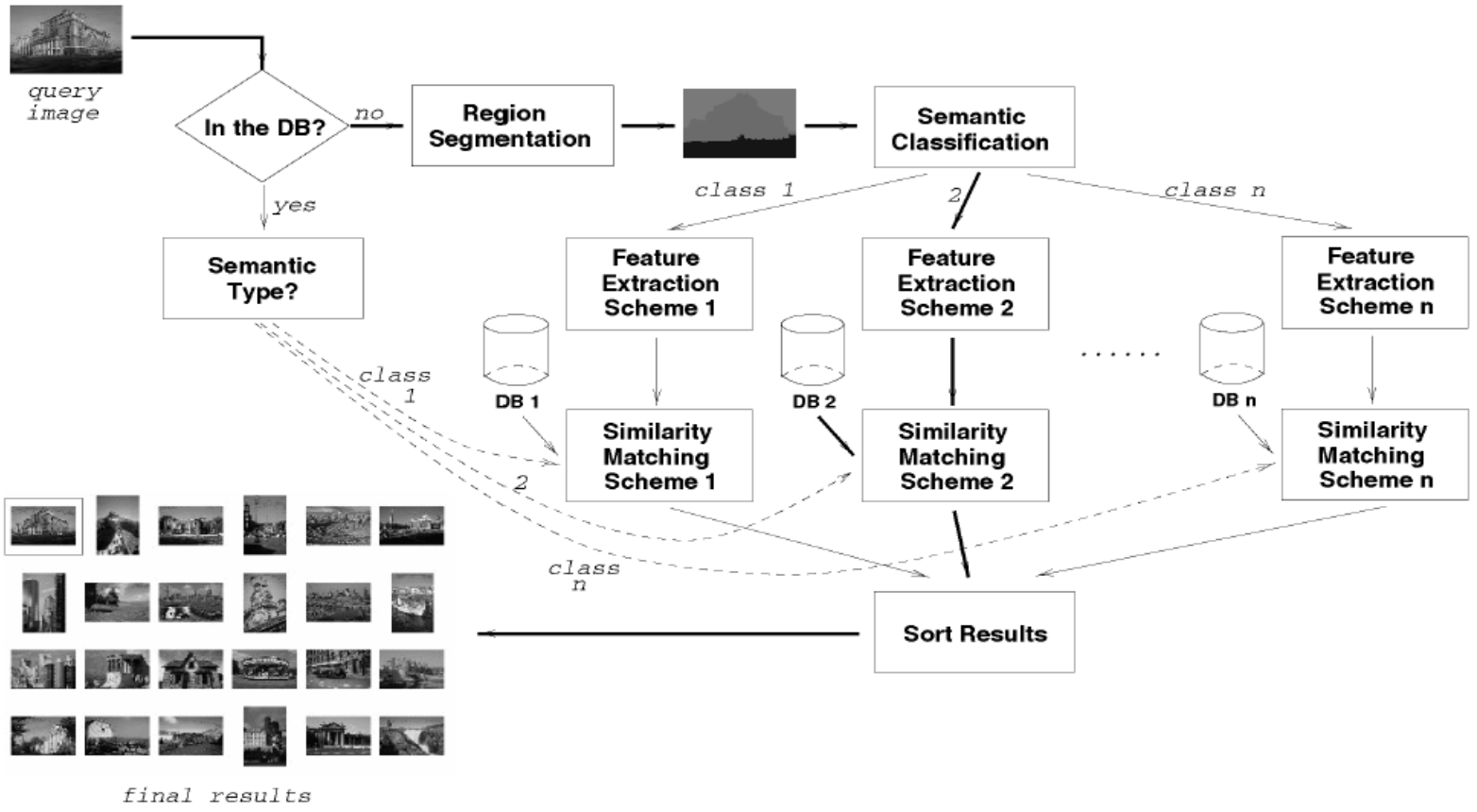
- CBVR has two Approaches:
  - Attribute based
  - Object based
- CBVR can be done by:
  - Color
  - Texture
  - Shape
  - Spatial relationship
  - Semantic primitives
  - Browsing
  - Objective Attribute
  - Subjective Attribute
  - Motion
  - Text & domain concepts

Overview (cont.)

- CBVR has two phases:
  - Database Population phase
    - Video shot boundary detection
    - Key Frames selection
    - Feature extraction
  - Video Retrieval phase
    - Similarity measure

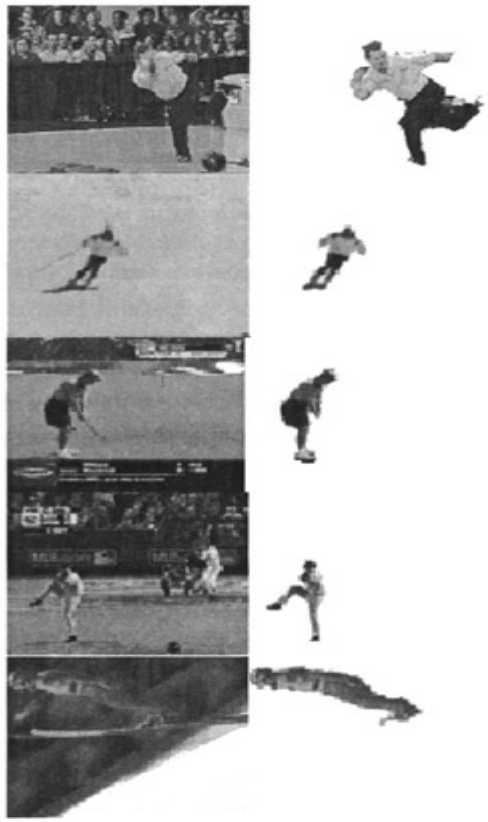
Overview (cont.)

- How CBVR works:



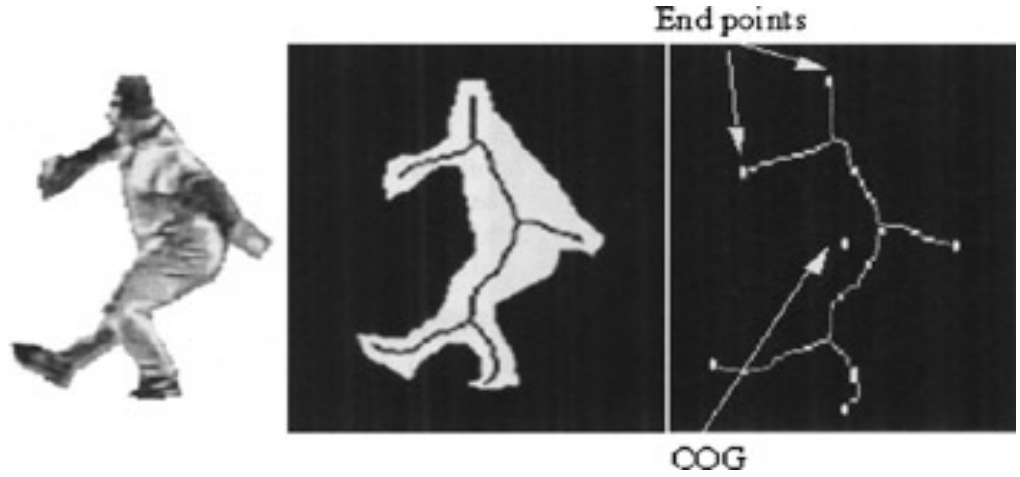
Database Population Phase

- Here are the three major procedures:
  - Shot boundary detection – partition, segments



Database Population Phase (cont)

– Key frames selection – select characteristics
– Extracting low-level spatial features like color, texture, shape, etc.



Database Population Phase (cont.)

- Video is complex data type – audio & video
- Audio can be handled by query by humming.
- Voice recognition system using Patricia-like tree to construct all possible substrings of a sentence.
- Audio is categorized by: speech, music, and sound.
- Audio retrieval methods: Hidden Markov Model, Boolean Search with multi-query using Fuzzy Logic.

Database Population Phase (cont)

- Most simple database storage: description of video as index along with the video.
- Human effort is involved in this case.
- We are searching for automatic video indexing and digital image storage method – Latent Semantic Indexing (LSI)

Database Population Phase (cont.)

- LSI is using vector space model – low rank approximation of vector space represent image document collection.
- Original matrix is replaced by an as close as possible matrix, where its column space is only the subspace of the original matrix column space.
- By reducing the rank of the matrix, noises (duplicate frames) are reduce to improve storage and retrieval performance.
- Term indexing is referred to the process of assigning terms to the content of the video.

Database Population Phase (cont.)

- Closest terms in the database is returned based on the similarity measure between the query images and the resulting ones.
- Cosine similarity measure is used in the vector space model.
- Cosine similarity measure on Term-by-video matrix:

$$\cos(t_1, v) = \frac{\sum_{h=1}^{k} t_1 v}{\sqrt{\sum_{h=1}^{k} t_1^{\,2} \sum_{h=1}^{k} v^2}}$$

Database Population Phase (cont.)

- Enterprise database like Oracle introduces new object type: ORDImage, which contains four different visual attributes: global color, local color, texture and shape.
- ORDImageIndex provides multidimensional index structure to speed up stored feature vectors.

Database Population Phase (cont.)

- Oracle example of joining two images of Picture1 and Picture2:

```
CREATE TABLE Picture1(
          author VARCHAR2(30),
          description VARCHAR2(200),
          photo1 ORDSYS.ORDImage,
          photo1_sig ORDSYS.ORDImageSignature
);
CREATE TABLE Picture2(
          mydescription VARCHAR2(200),
          photo2 ORDSYS.ORDImage,
          photo2_sig ORDSYS.ORDImageSignature
);
SELECT p1.description, p2.mydescription
FROM Picture p1, Picture p2,
WHERE
   ORDSYS.IMGSimilar(p1.photo1_sig, p2.photo2_sig,
   'color="0,6" texture="0,2" shape="0,1"
   location="0,1"', 20)=1;
```

Note: Weighted sum of the distance of the visual attributes is less than or equal to the threshold, the image is matched.

Summary

- Content-Based Video Retrieval system contains two phases:
  - Database population phase
    - Shot boundary detection
    - Key frames selection
    - Extract low-level features

Conclusion

- Content-based Video Retrieval system is not a sound system.
- Video stream will become the main stream in the years to come.
- Better off if we had a efficient CBVR search engine ready.
- Still many area needs to be improved.

IEEE 1394 / FireWire

Introduction

- The **IEEE 1394 interface** is a serial bus interface standard for high-speed communications
- real-time data transfer, frequently used by personal computers.
- Used in digital audio, digital video, automotive, and aeronautics applications.
- Originally Created by Apple and standardized as IEEE1394 in 1995

Introduction

- Apple computer corporation version is called Fire Wire while Sony and other corporation's version is called i.Link
- when released into the IEEE working group developed into an interconnect that could integrate both entertainment and consumer electronics.

Features

- Fast transfer of data
- Lots of devices on the bus
- Ease of use
- Hot pluggable
- Provide power through the cable
- Plug-and-play
- Low cabling cost
- Low implementation cost

Standards

- Currently there are three main standards and hundreds of IEEE recommendations.
- Three standards:
  - IEEE 1394
    - Original Standard, release in 1995
    - Supports transfer rates of 100,200,400Mbps
  - IEEE 1394a
    - Fixed interoperability problems that were left out of the spec for IEEE 1394
    - Added performance enhancements into the spec over the original
  - IEEE 1394b
    - Enhanced transfer rates for 800 , 1600, 3200Mbps
    - Separated the Standard from the Protocol

History

- Apple invented Fire Wire in the early 1990s
- adopted by IEEE Trade Association
  - became known as 1394
- name Fire Wire still used by Apple
- Others have adopted the name i.Link
  - trademarked by the Sony Corporation
- 1394 comes from the specification that defines the technical traits of the interface

Benefits

- speed
- ideal for an interface between high end digital electronics to exchange data
- Hot Swap and Daisy Chain
- peer-to-peer capabilities
- provides a power source in the media

Shortcomings

- not robust enough to operate as a true network interface
- Many devices don’t require such high throughput yet
  - USB still sufficient
- Many computer companies haven’t started including the interface on their personal computers

Types of Transfers

- Isochronous –
  - Broadcast in a one-one one-many fashion
  - No error correction or re-transmission
  - Uses up-to 80% of the available bandwidth
- Asynchronous –
  - Targets a specific node at a specific address
  - Error Checking and Re-Transmission
  - No Guarantee on the amount of bandwidth

4 Level Protocol Stack

- Physical Layer
  - Used for configuration of the bus
  - Electrical Signaling
  - Mechanical Connectors / Cabling
  - Arbitration Mechanism
  - Serial Encoding / Decoding
- Link Layer
  - Checking and Appending CRCs
  - Determines the type of transaction
  - Handles Isochronous transactions
- Transaction Layer
  - Handles Asynchronous transactions (request/response)
- Serial Bus Management Layer
  - Removed from the 1394b standard

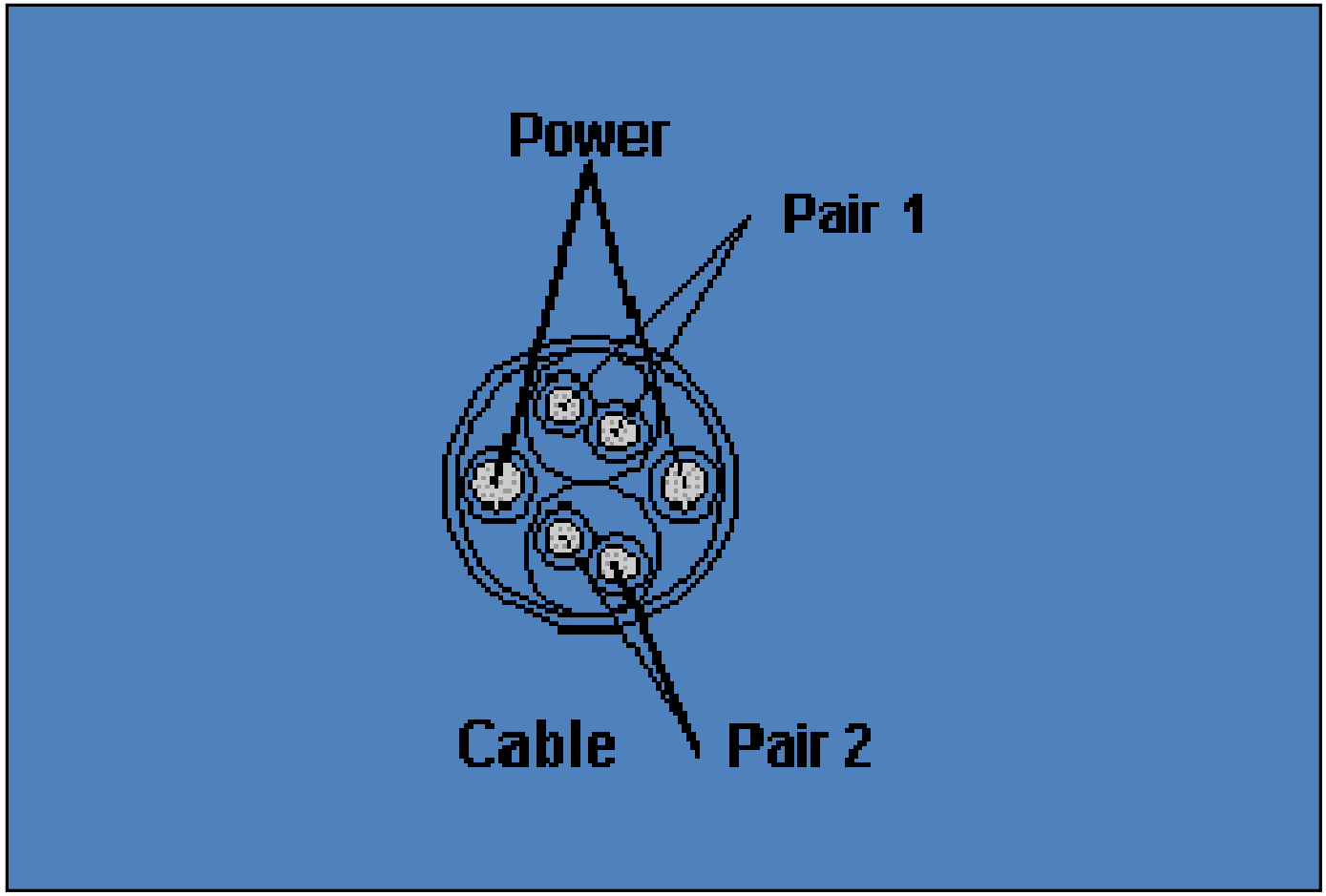
Topology

- IEEE 1394 is a peer to peer protocol, meaning a specific host is not needed for communication
- Configuration of the bus occurs when ever a new device is plugged in

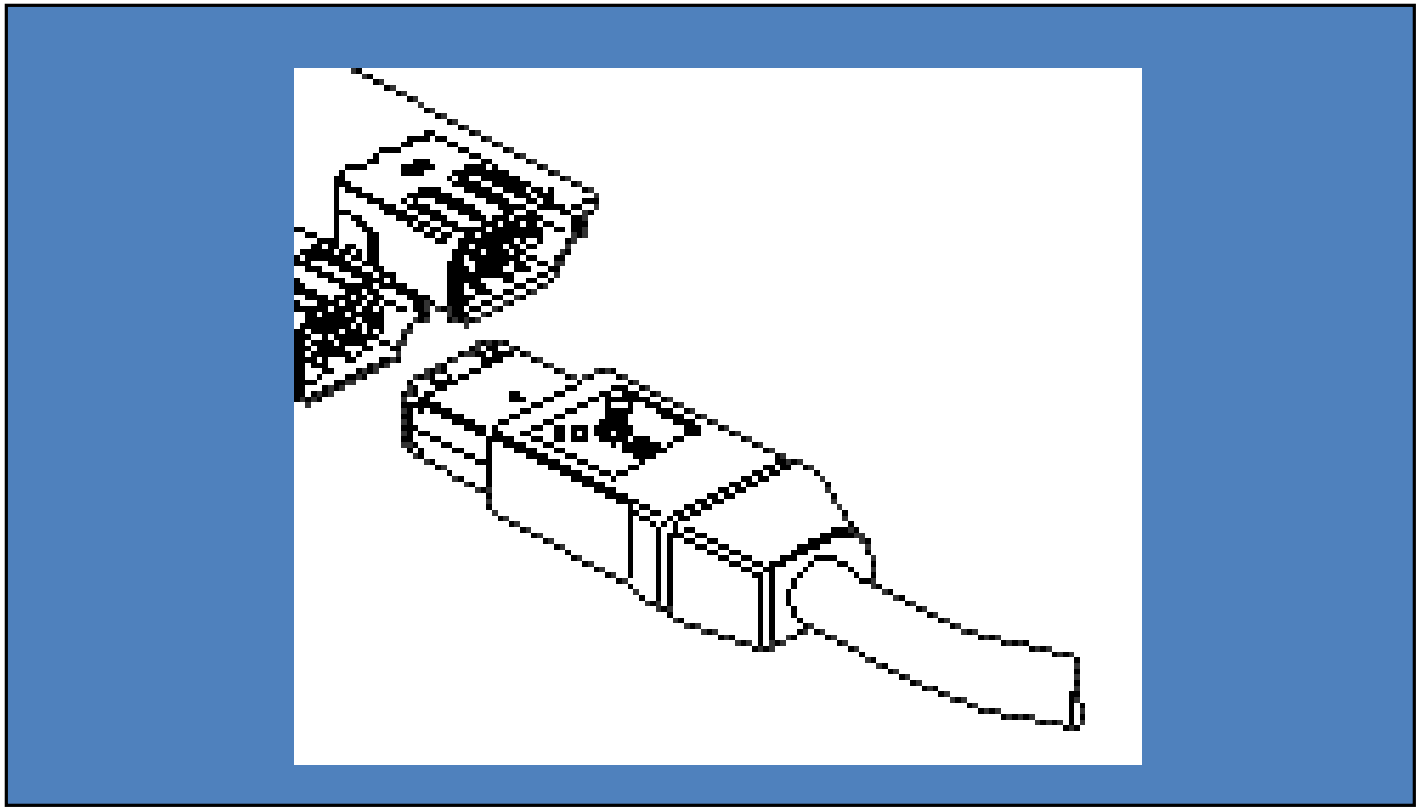
- Uses a 64bit address scheme based on the IEEE 1212 standard

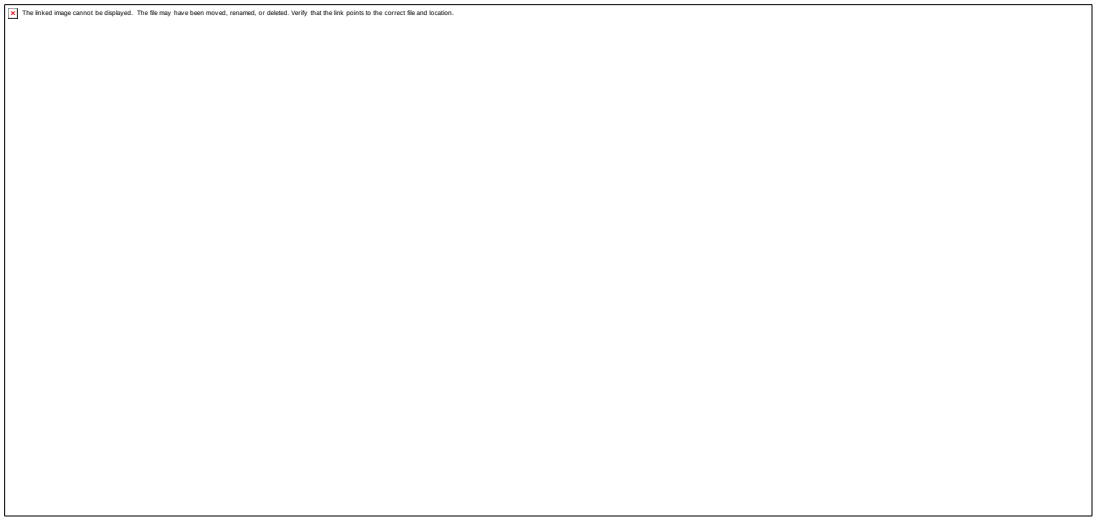
- 64 bits decomposed as follows:
  - 10 bits for the Bus ID – Allowing 1024 busses
  - 6 bits for the Device ID – Allowing 64 Devices
  - 48 bits for storage

Cable (1394a)



Connector (1394a)

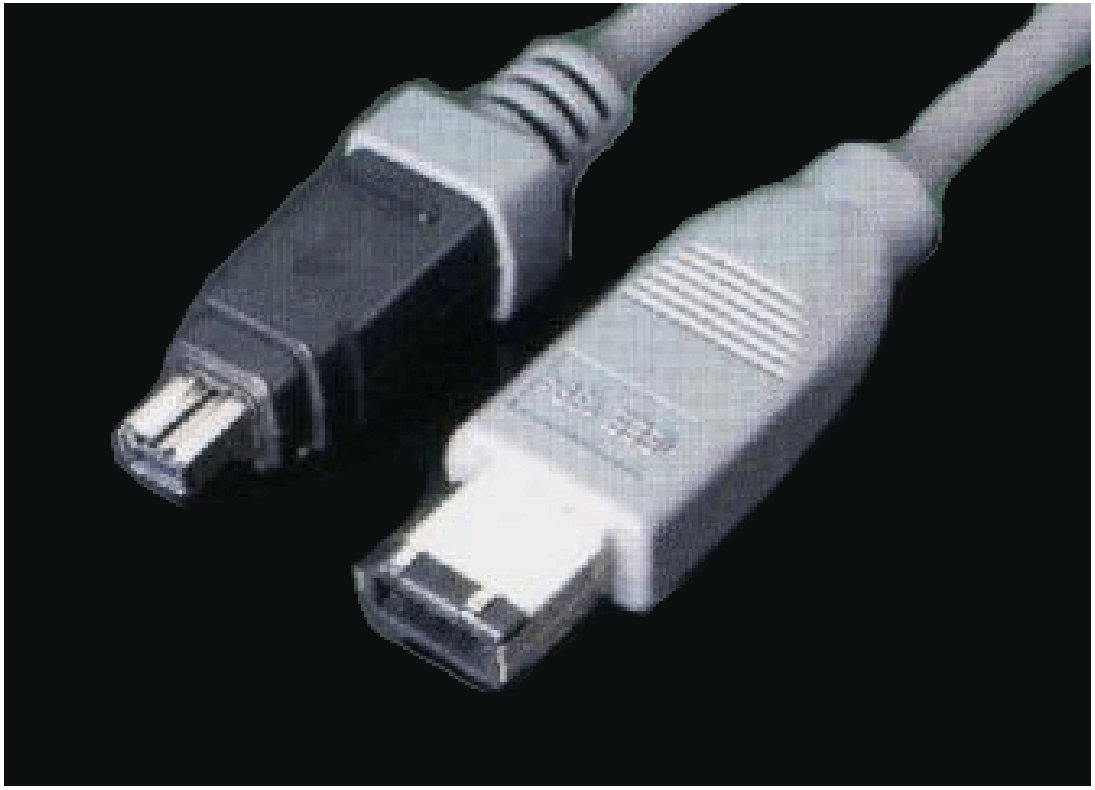



The linked image cannot be displayed. The file may have been moved, renamed, or deleted. Verify that the link points to the correct file and location.

Connecting media

- 6-conductor cable
  - twisted pairs of wire for data transmission
  - two power conductors
  - power wires (8 to 30 volts DC, 1.5 Amps) supply power
- 4 pin connectors are designed mostly for battery powered

- IEEE 1394 cables are limited to about 4.5 meters in length
  - signal distortion



IEEE 1394 Applications: Peripherals

- A way to transport data from digital cameras, scanners, printers, etc...
- Up to 40 times faster than USB
- Firewire cards can adapt to older devices

IEEE 1394 Applications: Networking

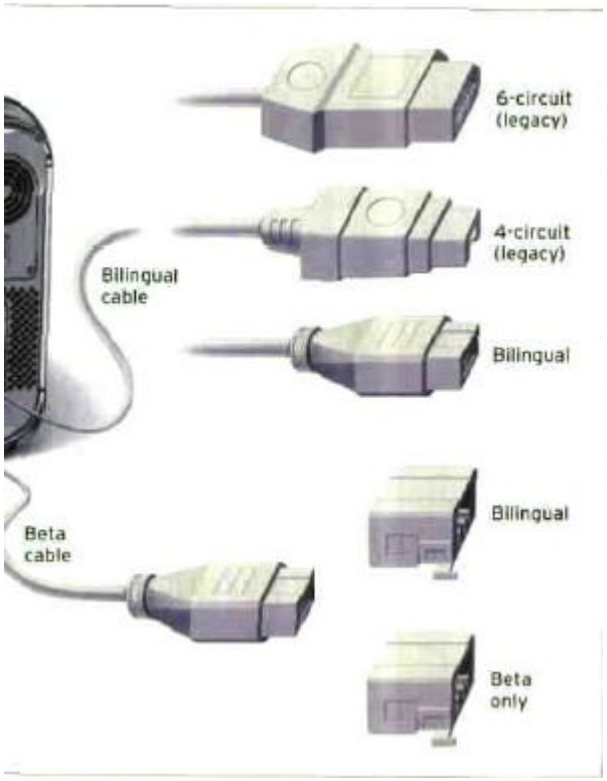
- Great for home networking
- Connect devices with differing transfer rates
- Simple plug in styles
- “Hot Plug” capability useful for troubleshooting
- Compatible with future devices

IEEE 1394 Applications: Industrial

- “Hot Plug” capability minimizes downtime and loss of productivity
- Industrial cameras can detect imperfections on fast mass production lines
- Cameras can be linked to other cameras or devices for a quick transfer of data over the network.
- Faster and simpler than traditional applications which makes work more efficient

The Future of IEEE 1394

- USB2
  - Max transfer rate of 480 Mbps
- IEEE 1394b
  - Sustained transfer rate of 800 Mbps
  - Max transfer rate of 3200 Mbps (High Implementation)
  - Unique features



The Future of IEEE 1394

- Computer Hardware
  - Hard Drives
  - CD-ROMS, DVD-ROM, DVD-RAM
  - Digital Video
  - Other Peripherals

The Future of IEEE 1394

- Consumer Electronics
  - DVI will most likely be replaced by firewire
  - Firewire can be daisy chained with other electronic devices
  - More options for copy protection
  - Ability to record in digital
  - Devices can be controlled remotely

The Future of IEEE 1394

- Networking
  - Link devices in a laboratory
    - Easy to transfer data to computers
    - Old devices can be linked with a bridge
    - Standardized interface makes data recording easier
  - Link consumer electronics and computers
    - Robust interface can share video, audio, IP traffic
    - Ability to operate without a PC

1394 Bus

- When connecting devices with 1394 cables you are connecting them to a 1394 bus
- share a linear electronic path
- Daisy chaining
- signal is received by all the devices

Peer-to-Peer

- communicate with one another without the need for a computer
  - could transfer data from a hard drive without the need for computer assistance
  - hook too cameras together and dub from one to the other without any computer
- share the same resource without any need for special support
- Also supports “Hot Swapping”

Speed

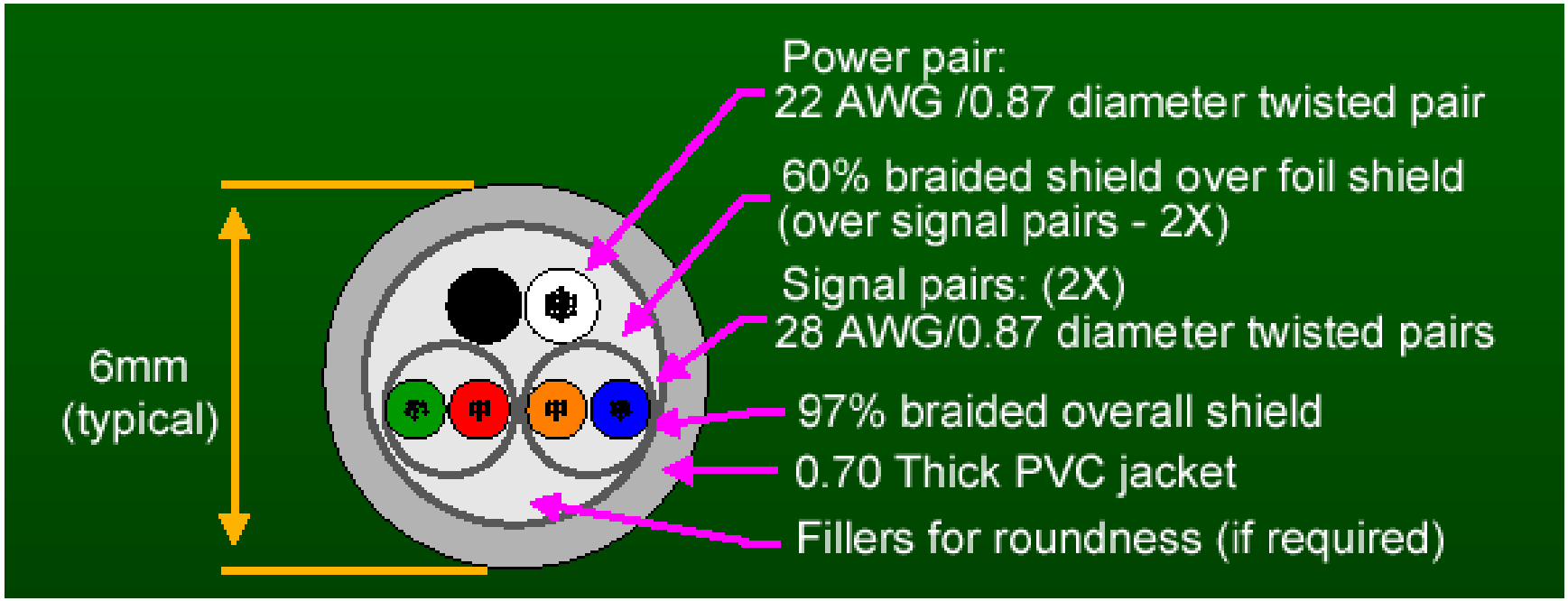
- Very fast
  - Up to 400 Mbps
- Much faster than competitors
  - USB only 11 Mbps

IEEE Specifications

Bus Architecture

- Serial version of IEEE 1212
- Similar logical architecture to the memory bus
- Automatic address assignment

The Cable

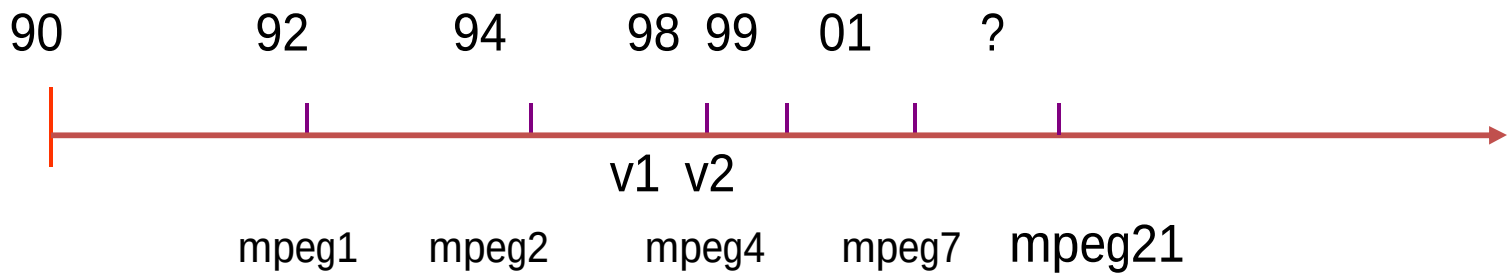

Power pair:
22 AWG /0.87 diameter twisted pair
60% braided shield over foil shield
(over signal pairs - 2X)
Signal pairs: (2X)
28 AWG/0.87 diameter twisted pairs
97% braided overall shield
0.70 Thick PVC jacket
Fillers for roundness (if required)
6mm
(typical)

Product Revisions

- IEEE 1394A
  - Made the bus more efficient by changing the way it behaved
- IEEE 1394B
  - Changed the medium of data conveyance to allow for greater speed and distances (fiber optics, UTP5)

Overview of MPEG-7

From MPEG-1 to MPEG-7



- MPEG-3, ever defined, but abandoned
- MPEG-5 and -6, not defined

MPEG Family

MPEG-1 – Coding of moving pictures and audio for digital storage media (CD-ROM, MP3), 11/92

MPEG-2 – Generic Coding of moving pictures and audio information (DVD, Digital TV), 11/94

MPEG-4 – Coding of Audiovisual Objects for MM appls Ver1 09/98, Ver2 11/99

MPEG-7 – Multimedia content description for AV material 08/01

MPEG-21 – Digital AV framework: Integration of multimedia technologies, 11/01

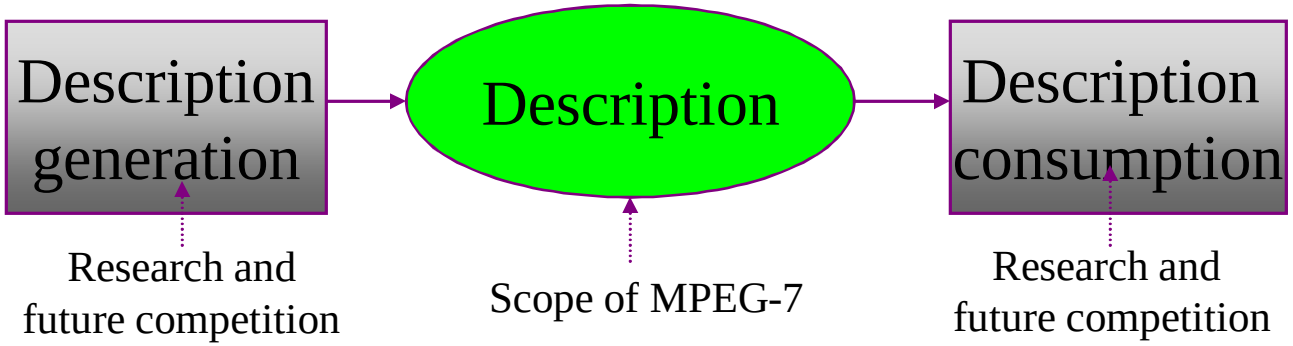
Why is MPEG-7 needed

- Digital audiovisual information increasing
  - more and more available contents
  - all kinds of sources of information
- Use of the digital audiovisual information
  - description of the contents
  - fast search of the contents

Objective of MPEG-7

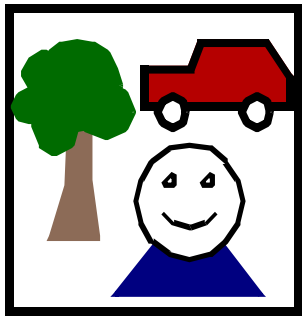
- Standardize content-based description for various types of audiovisual information
  - Enable fast and efficient content searching, filtering and identification
  - Describe several aspects of the content (low-level features, structure, semantic, models, collections, creation, etc.)
  - Address a large range of applications
- Types of audiovisual information:
  - Audio, speech
  - Moving video, still pictures, graphics, 3D models
  - Information on how objects are combined in scenes

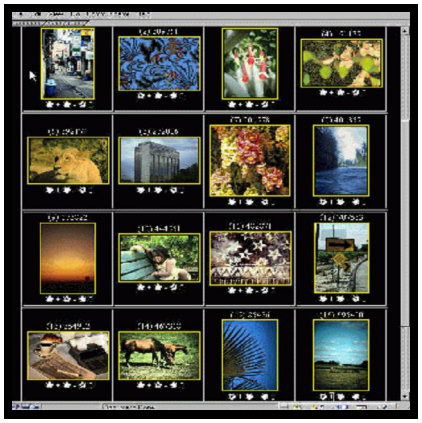
Scope of MPEG-7



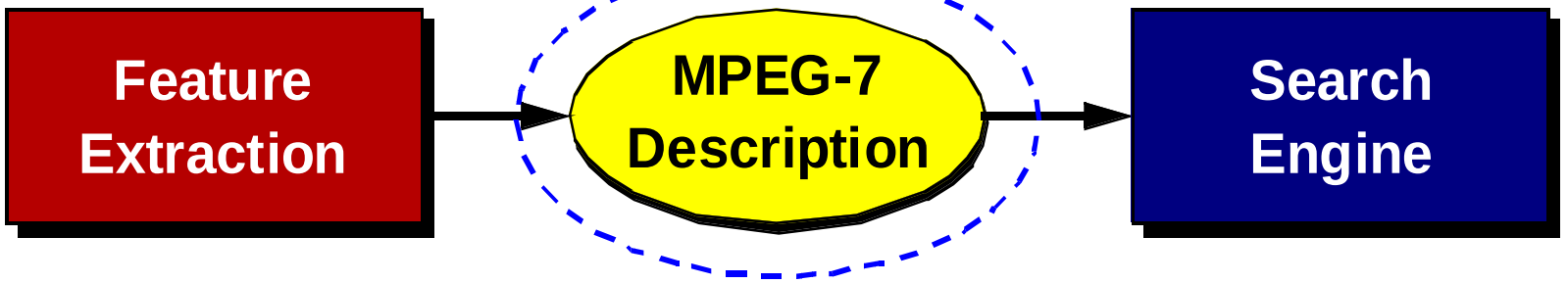
- The description generation (feature extraction, indexing process, annotation & authoring tools,...) and consumption (search engine, filtering tool, retrieval process, browsing device, ...) are non normative parts of MPEG-7.
- The goal is to define the minimum that enables interoperability.

Scope of MPEG-7





standardization



Feature Extraction:
Content analysis (D, DS)
Feature extraction (D, DS)
Annotation tools (DS)
Authoring (DS)

MPEG-7 Scope:
Description Schemes (DSs)
Descriptors (Ds)
Language (DDL)
Ref: MPEG-7 Concepts

Search Engine:
Searching & filtering
Classification
Manipulation
Summarization Indexing

Audio in MPEG-7

- Audio content description (yes)
- Sound retrieval and classifier (yes)
- Speech synthesis (no)
- Speech recognition (no)
- Probability Models (yes)

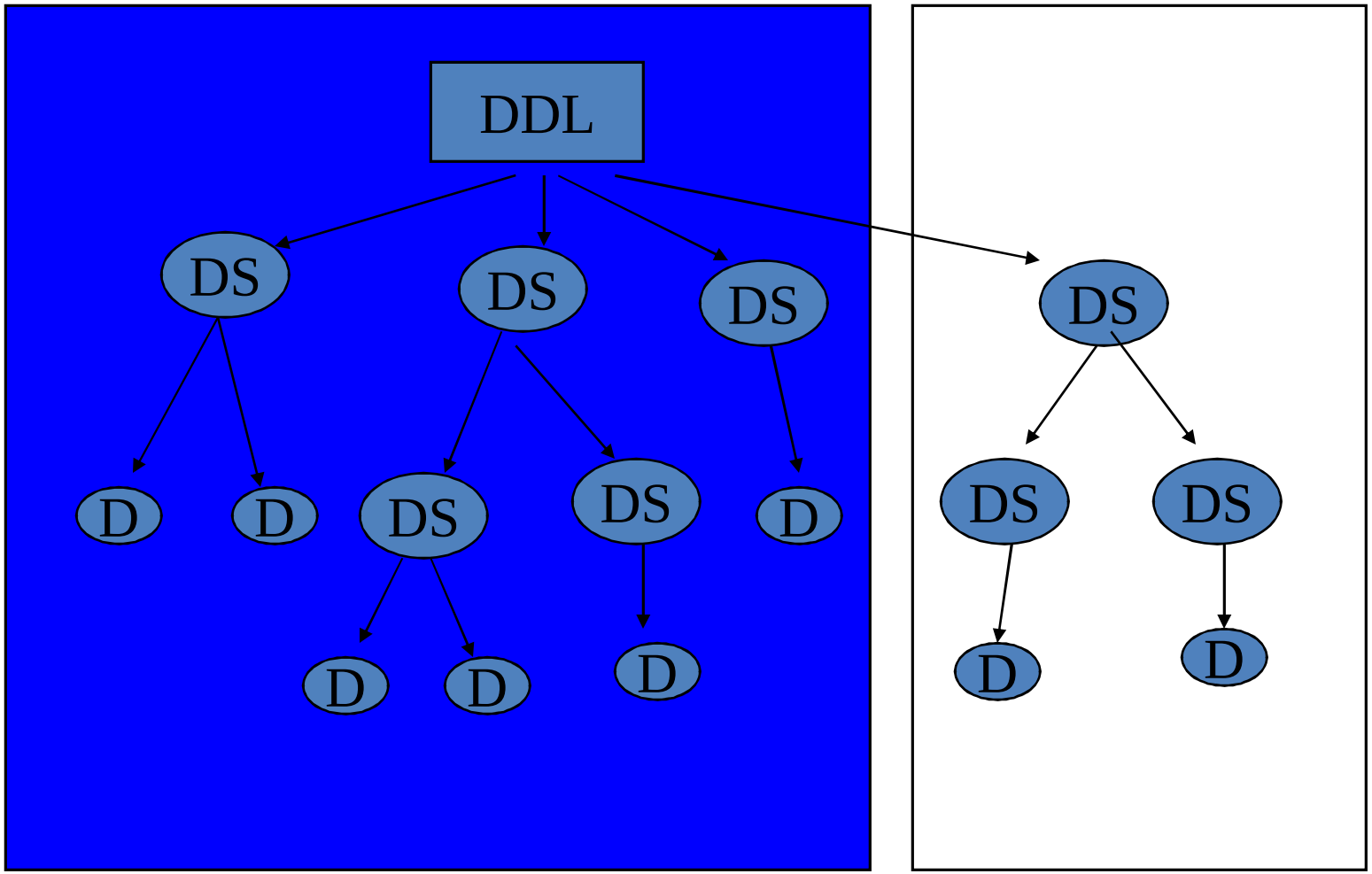
Parts of the MPEG-7 Standard

- **ISO / IEC 15938 - 1:** ***Systems***
- **ISO / IEC 15938 - 2:** ***Description Definition Language***
- **ISO / IEC 15938 - 3:** ***Visual***
- **ISO / IEC 15938 - 4:** ***Audio***
- **ISO / IEC 15938 - 5:** ***Multimedia Description Schemes***
- **ISO / IEC 15938 - 6:** ***Reference Software***

Main elements of MPEG-7

- **Descriptors (D):** representations of features, that define the syntax and the semantics of each feature representation (low-level).
- **Description Schemes (DS):** that specify the structure and semantics of the relationships between their components, which may be both Ds and DSs (high-level).
- **A Description Definition Language (DDL):** based on XML Schema, to allow the creation of new DSs and Ds, and to allow the extension and modification of existing DSs
- **System tools:** to support multiplexing of descriptions, synchronization issues, transmission mechanisms, coded representations, management and protection of intellectual property

Relations of main elements



Description Definition Language

- *Description Definition Language (DDL) is a language that define what description is valid, and allows the creation of new Description Schemes and Descriptors. It also allows the extension and modification of existing Description Schemes*
- *DDL is used to define a set of formal rules*
  - *ordering of the elements*
  - *occurrences of elements*

  ……..
- *XML + MPEG-7 extensions*

XML: Base for DDL

- **Why choose XML as the base for the DDL?**
  - The popularity of XML
  - The interoperability with other standards in the future

- **Why XML should be extended for MPEG-7?**
  - SGML > XML
  - Structural extensions
  - Datatype extensions

DDL parser

DDL parser is a software to check if a description is valid