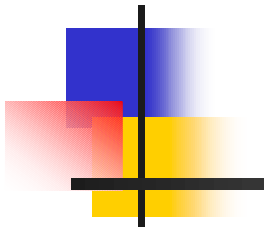


Visual and Textual Content Based Indexing and Retrieval



International Journal Digital Library, 2: 269-287, 2000

Introduction



- To access the content of multimedia information efficiently, there are different descriptions of content textual and visual descriptions that refer to textual and visual documents.
- Content textual retrieval system uses non-structured (keyword, free-text) as a basis for indexing and retrieval.
 - E.g. best-match information retrieval techniques are used to search a ranked list of relevant documents based.
 - Content description is semantically powerful.
 - Free text can describe the content naturally and semantically.
 - But the queries based are not allowed directly on the visual properties of the images.

Introduction (*continue...*)



- Visual retrieval system uses visual image features in indexing and queries.
 - The features are extracted automatically from image.
 - Users formulate their queries either by supplying sample images or by specifying a combination of visual features.

Paper Contribution



- This paper presented a content-based visual retrieval system that enables users to formulate their queries.
- User queries can be expressed in a form of example images or specifying a combination of textual and visual features.
 - Textual attributes described the semantic of an image.
 - Visual features described the image content such as color, shape that were extracted automatically or semi-automatically.

Paper Contribution (*continue...*)



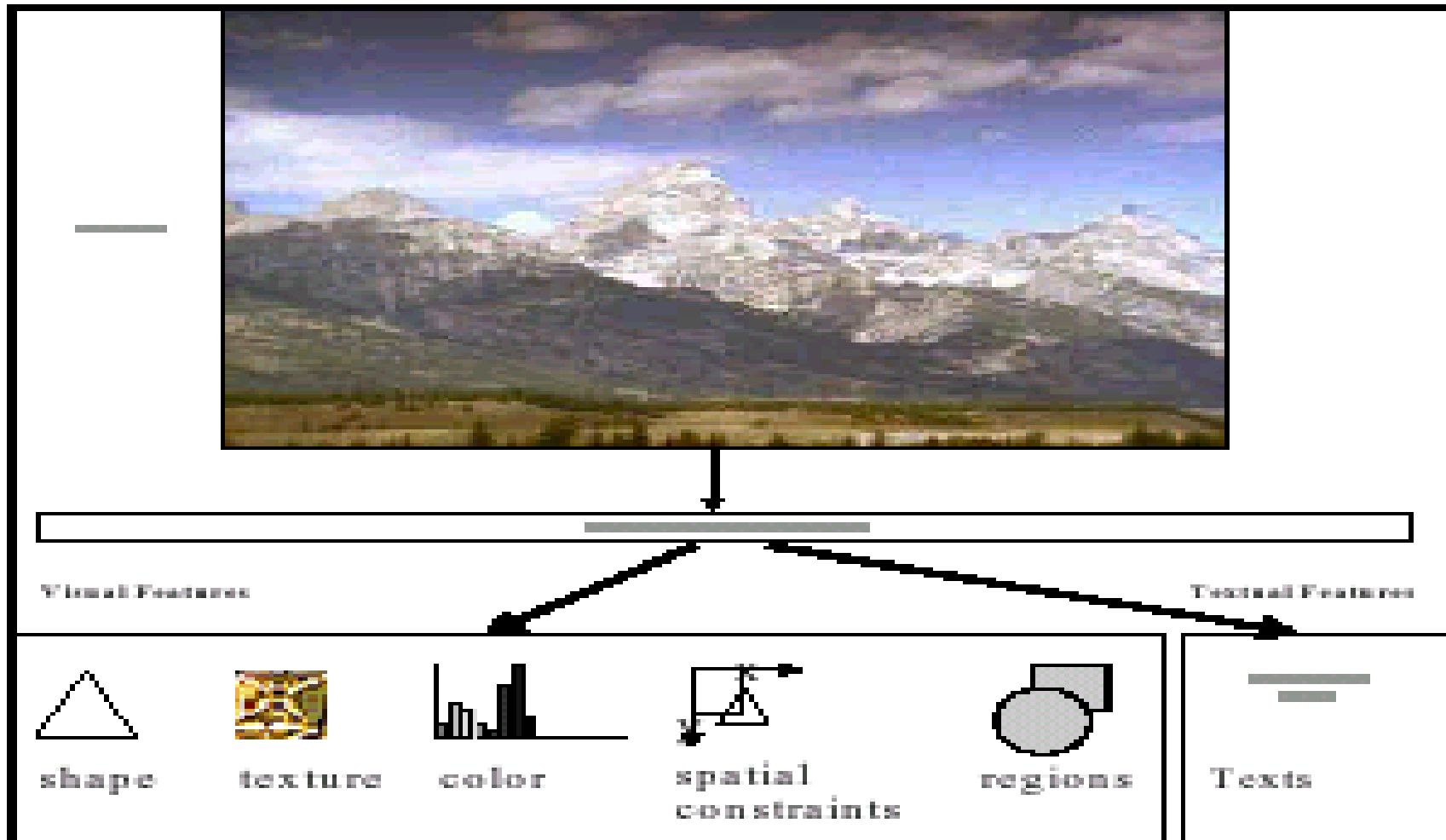
- The prototype is independent of a specific application domain.
 - E.g. the user may define own class of images that he/she looked for and store these classes in the database and reused them whenever he/she wanted.
- The approach supported different concepts compositions to obtain more complex concepts.
 - E.g. Sea concept + plane concept = plane on the sea concept (the plane on the sea).

Paper Contribution (*continue...*)



- The approach implemented a powerful shape representation by using the Fourier mathematical descriptor to reconstruct the image shape.
- Knowledge was structurally represented at different levels of semantics by concept and can be automatically generated from user queries. The knowledge representation associated with visual and textual features.

Visual Features



Visual Features (*continue...*)



- **Shape** => characterised by boundaries, length, width area.
 - The accuracy of the shape description is adapted by Fourier descriptors.
- **Color** => is characterised by color distribution.
 - Image color distribution is represented by color histogram $H(M)$.
 - $H(M)$ is a vector $(h_{c_1}, h_{c_2}, h_{c_3}, \dots, h_{c_n})$ where each h_{c_j} represents the number of pixels of the color c_j in image M .
 - A color histogram computed on :
 - a whole image when browsing through the database.
 - an image area when the considered image is divided into many identical areas.
 - An image region

Visual Features *(continue...)*



- **Color Extraction** => based on image format (e.g. jpeg)
 - the region and image colors are extracted and represented by RGB (Red Green Blue) model.
 - The colors are transformed and represented in HSV (Hue Saturation Value).
 - RGB histogram is structured by 3 attributes (red,green,blue).
 - Each attribute takes as values an object structured by 2 attributes (the variation and color mean).
 - HSV histogram contains colors histograms, distance set, methods that compute the attributes of colors histogram.
 - Comparing the two colors regions \Leftrightarrow computing the distance between the histogram of the target region and the histogram of source.

Visual Features (*continue...*)



- **Region Texture** => is represented to be 2 classes.
 - The first class is structured by a set of attributes that model the 4 statistical moments (computed by 4 suitable methods -coarseness, variance, contrast, directionality) and the number of pixels of the region.
 - The second class is represented by the histogram of the gray differences.

Visual Features *(continue...)*



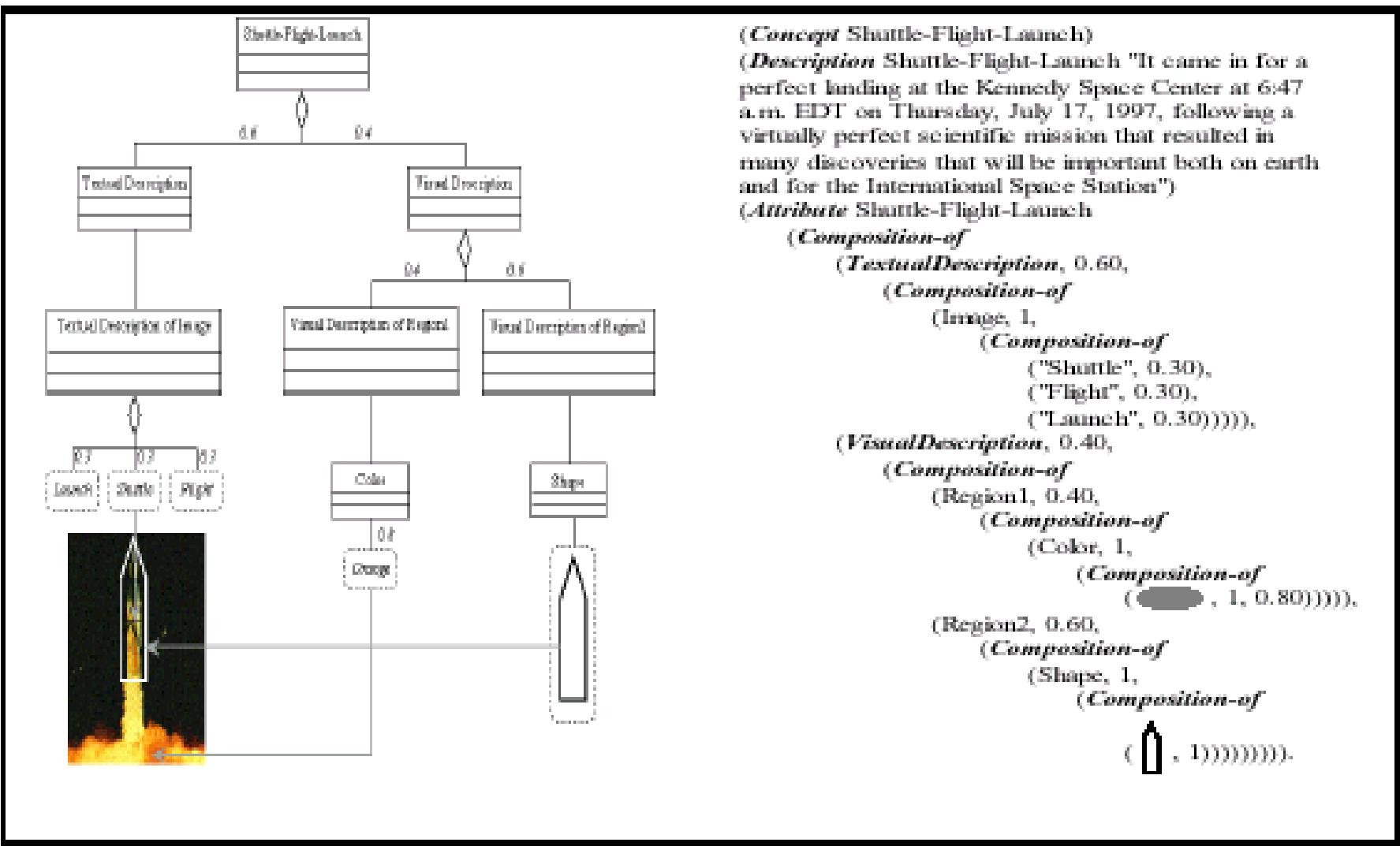
- **Spatial Localization** => locates the relations between image regions and computes the distance between two points in those regions.
 - Each region is considered as a minimal rectangle.
 - The image is subdivided into 9 (arbitrary choose) areas of identical dimensions.
 - The position of a region is deduced from the position of the gravity centre in one of 9 areas.
 - The localization is stored in the database.

Concept-Based Query



- The representation contains information that can be used to identify the concept in the image in terms of other sub-concepts and image content attributes.
 - E.g. If images of canary birds are looked for, shapes, colors, textures that characterize a canary.
 - E.g. “Shuttle Flight Launch in a daylight” \leq has two sub-concepts ie. Shuttle-Flight-Launch and Daylight.
- **Description**
 - Composition-of-relationship is used to describe the composition of content attribute in terms of a weight.
 - E.g. A weight of 40% for region 1, a weight of 60% for region2.

Concept-Based Query (continue...)



(*Concept Shuttle-Flight-Launch*)
 (*Description Shuttle-Flight-Launch* "It came in for a perfect landing at the Kennedy Space Center at 6:47 a.m. EDT on Thursday, July 17, 1997, following a virtually perfect scientific mission that resulted in many discoveries that will be important both on earth and for the International Space Station")
 (*Attribute Shuttle-Flight-Launch*

(*Composition-of*
 (*TextualDescription*, 0.60,
 (*Composition-of*
 (Image, 1,
 (*Composition-of*
 ("Shuttle", 0.30),
 ("Flight", 0.30),
 ("Launch", 0.30))))),
 (*VisualDescription*, 0.40,
 (*Composition-of*
 (Region1, 0.40,
 (*Composition-of*
 (Color, 1,
 (*Composition-of*
 (Orange, 1, 0.80))))),
 (Region2, 0.60,
 (*Composition-of*
 (Shape, 1,
 (*Composition-of*
 (ShuttleShape, 1)))))))).

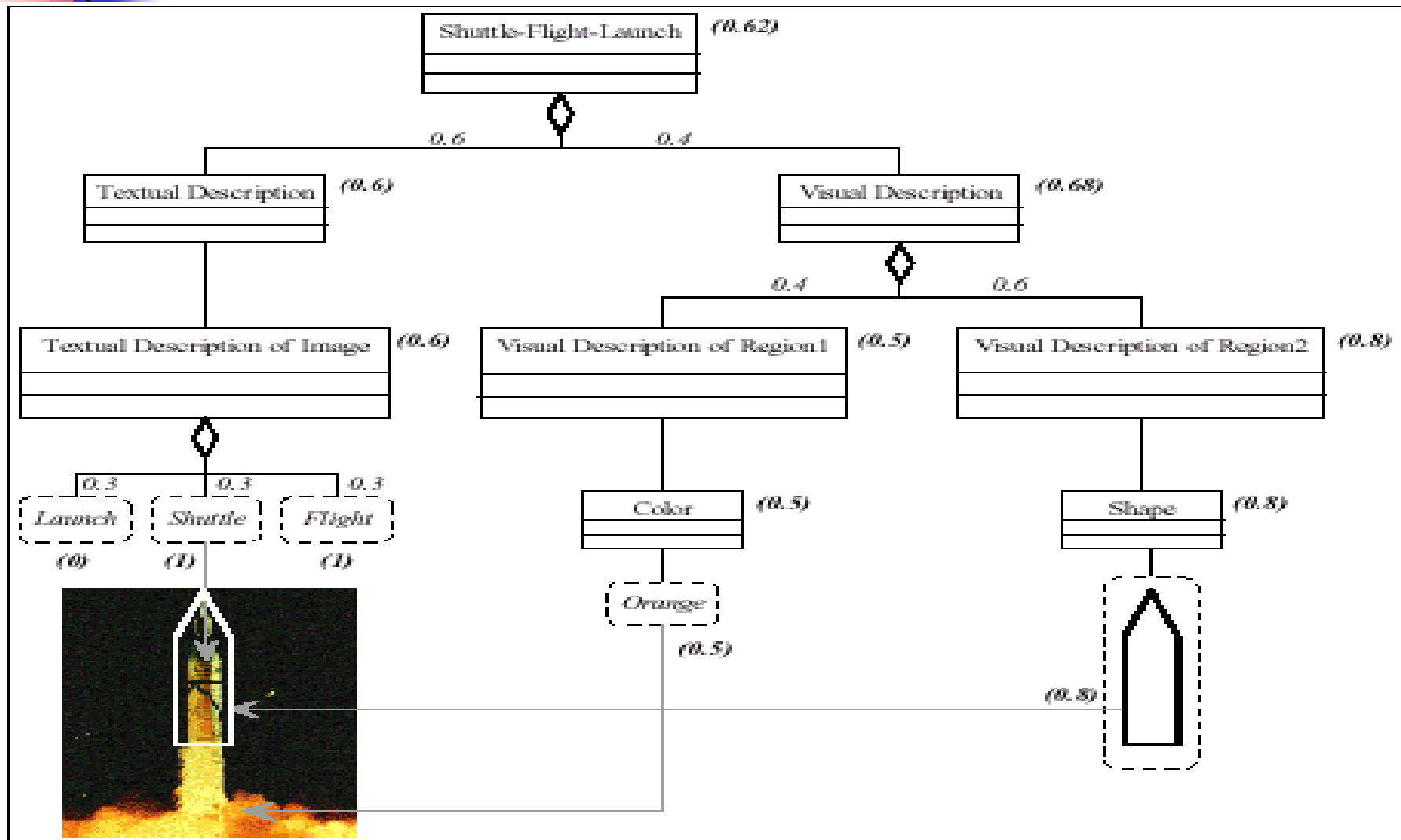
Concept-Based Query (*continue...*)

- **Query Processing** => the system constructs the query concept in order to match it with the content of images stored in the database.
- The matching => similarity distance between the concept specified in query and the database image content.
- For color, texture and shape => implementing a set of similarity distance.
 - Let H = histogram source of image (query image).
 - Let I = histogram of target of image (image in database).
 - $D(H,I) = ((H-I).A.(H-I)^T)^{1/2}$; A = similarity matrix(n*n)

Concept-Based Query (*continue...*)

- For spatial location
 - There is no similarity measure.
 - The system considers only the image that respect the localization specified in the query.
- For the text
 - The distance = 1 if the words matched are the same or synonymous, otherwise distance = 0.

Concept-Based Query (*continue...*)



Experimental Results



- **Evaluation method**

- Retrieval goodness is a capacity of retrieval system that retrieves effectively information relevant to user.
- A given query and a given number of images retrieved
 - $\text{recall} = \frac{|\# \text{ relevant images retrieved}|}{|\text{total relevant images}|}$
 - $\text{precision} = \frac{|\# \text{ relevant images retrieved}|}{|\# \text{ retrieved images}|}$

- **Data Sets**

- collection of 30,000 images - animals, flowers, people, scenery
- all images were catalogued and each image was associated description.
- 100 queries were selected from a test collection.

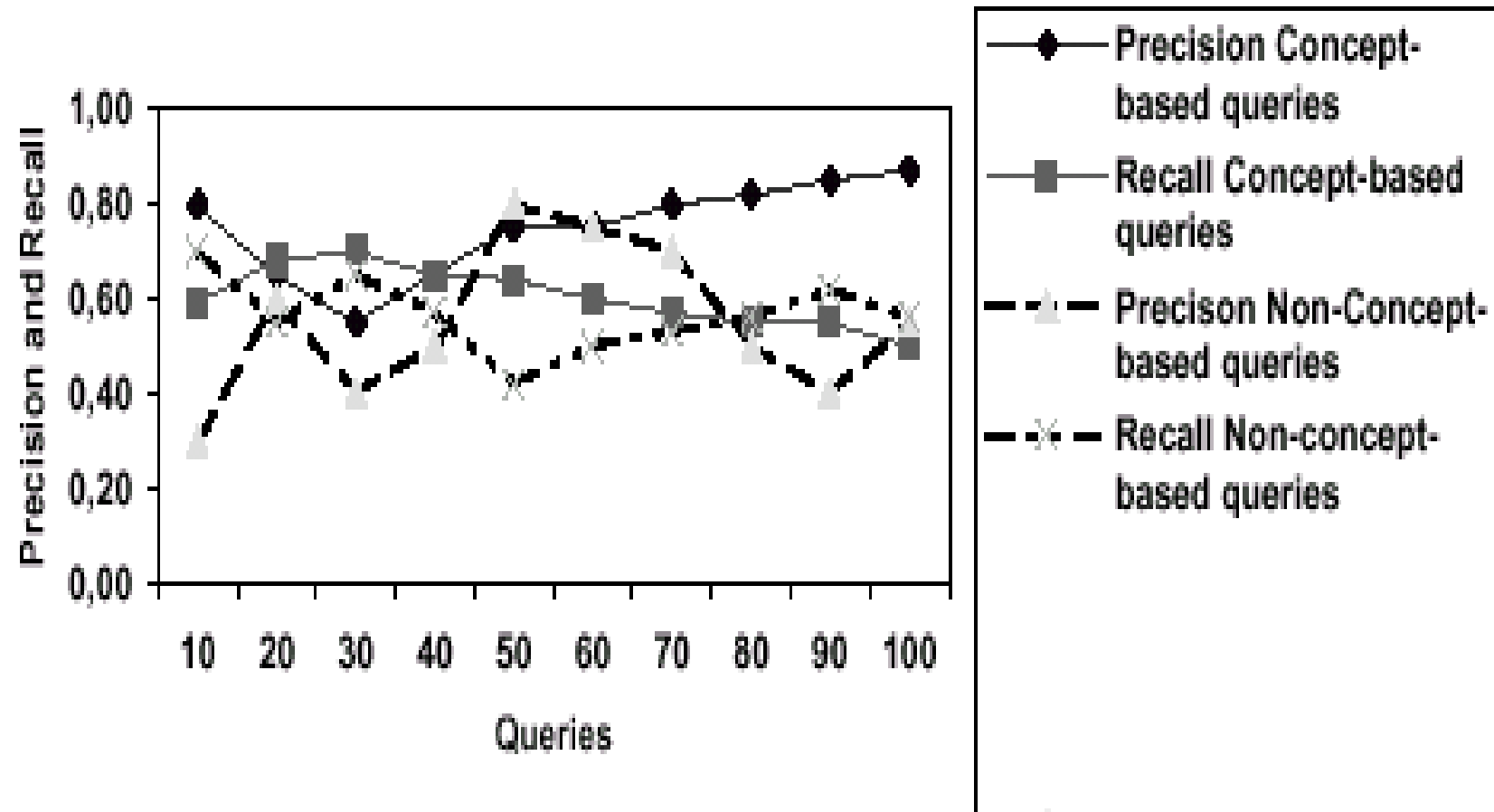
Experimental Results (*continue...*)

- **Results and Analysis**

Table 1. Experimental results

Queries	Concept-based queries		Non-concept-based queries	
	Precision	Recall	Precision	Recall
Query 1	80%	59%	30%	70%
Query 2	65%	68%	60%	55%
Query 3	55%	70%	40%	65%
Query 4	65%	65%	50%	57%
Query 5	75%	64%	80%	42%
Query 6	75%	60%	75%	50%
Query 7	80%	57%	70%	53%
.....				
Query 98	85%	55%	40%	62%
Query 99	87%	50%	55%	56%
Average	80%	85%	55%	70%

Experimental Results *(continue...)*



Conclusion



- Visual and textual content-based retrieval system is presented.
- The system has two major components : extraction and queries.
- In queries, the user can use both visual and textual features to specify the full semantics of the required images.