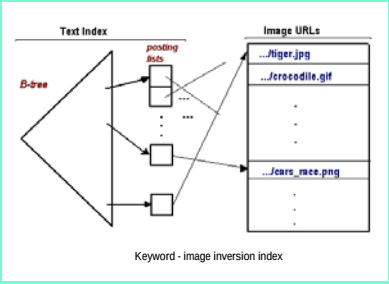
Text Based Search

- At present, most of multimedia search engines (Google Image, AltaVista Photo Finder), and all first generation visual information retrieval (VIR) systems, are textual…
  - But both Google and Microsoft Research are now competing in the development of Internet-oriented image retrieval tools
- Textual retrieval relies strongly on **metadata** (keywords, text strings, full scripts) containing in images and videos
  - An ***inverted file index*** describing the multimedia content allows for obtaining fast query response

Text Based Search



Keyword - image inversion index

Text Based Search

- Another indexing techniques:
  - partitioning multimedia content into ***categories***, which can be browsed through for images that match **keywords**
  - using the ***text*** embedded around multimedia content as a way to identify its content
- But keywords and texts relate only implicitly to image / video / audio content

Content Based Search

- **Content** or **semantics-based search**: *retrieval criteria and queries are specified in terms of **computable** data features related to semantic content of a multimedia object*
- Most content-based CBIR systems allow for searching the visual database contents in several different ways, either alone or combined

Content Based Search

1. **General interactive browsing:** with no specific idea about the desired images or video clips:
   - ***category (subject) search*** to retrieve an arbitrary image that represents a specific class (the search may be simplified by *clustering of visually similar images* into groups and *navigation through a subject hierarchy* to get to the target subject)
   - ***search by association*** having no specific aim and iteratively refined by the user's relevance feedback

Content Based Search

2. **Search to illustrate a story or document** or for arbitrary pictures of expected aesthetical value
3. **Search for a specific image**, or Query-by-X where "X" can be:
   - an image example,
   - a visual sketch,
   - a specification of visual features, or
   - a keyword or complete text

Query-by-X

- *An image example* or examples (QBE)
- *A visual sketch* drawn by a user
- Direct *specification of visual features* (e.g., colour, texture, shape, and motion properties), which appeals to more technical users
- *A keyword or complete text* entered by the user to search for previously annotated visual information

Content Based Search

- QBE framework:
  - Extract colour, texture, and other features of each image in the database and store as metadata
  - Extract the same features of the query image and present as metadata
  - Match the query metadata to each image metadata in the database of indexed images
  - Output retrieval results based on matching scores

Search by Similarity

- *Range search* finding all data items (e.g., images) with features within given ranges
- *k-Nearest-Neigbour search* (kNN) finding the k most similar data items to a query template.
- *Within-distance search*, or $\boldsymbol{A}$*-cut* (*alpha-cut*), finding all data items with a similarity score better than a given threshold $\boldsymbol{A}$, or at a distance less than $\boldsymbol{A}$ from a query template

Queries to CBIR Systems

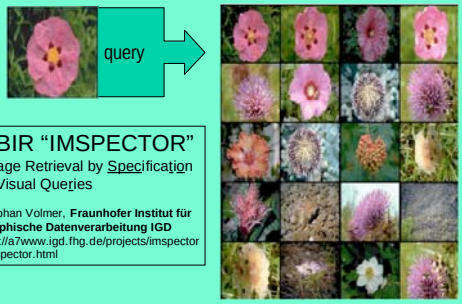
- The “user-friendly“ querying has to maximally correspond to the user's cognition and intuition
- A text-based paradigm is mostly invalid when visual information is sought
- It is difficult to exactly describe the semantic meaning of the image content, especially when the user only has a "fuzzy" idea of what should be looked for

Query by Image Example (QBE)

More flexible description of the image content and more versatile queries than the texts permit:

- it is easier for a user to show examples of what should be looked for rather than verbally describe it
- the query becomes more compact, natural, and convenient for the user
- the results can be more easily evaluated by the user through a direct comparison with the initial query

Query by Image Example (QBE)



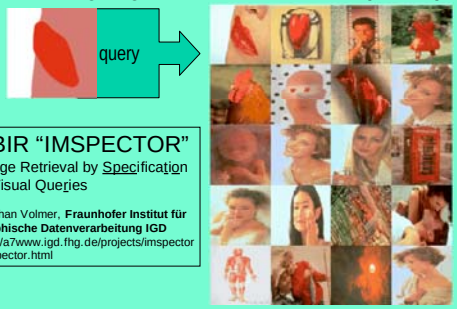
CBIR “IMSPECTOR”
Image Retrieval by Specification of Visual Queries

Stephan Volmer, **Fraunhofer Institut für Graphische Datenverarbeitung IGD**
http://a7www.igd.fhg.de/projects/imspector/imspector.html

Query by User's Sketch (QBS)

- Retrieval of images looking like a given sketch
- Both QBE and QBS use low-level features (colour, texture, etc) rather than the meaning, or semantics (objects, their relationships, events, etc)
- Users are interested in the application-dependent semantics
  - It makes no sense to search for pictures of George W. Bush, using colour and simple texture features that describe nearest neighbouring pixels, whereas in medical imaging these features are very important

Query by User's Sketch (QBS)



CBIR "IMSPECTOR"
Image Retrieval by Specification of Visual Queries

Stephan Volmer, **Fraunhofer Institut für Graphische Datenverarbeitung IGD**
http://a7www.igd.fhg.de/projects/imspector/imspector.html

Querying by Multiple Examples

- Natural extension of the QBE (it is unlikely to find an example with exactly all the features needed)
- A querying system accepts *multiple* examples from the user, gathers them into a joint *pseudo-example*, and responds with the best matches to this latter combined "image"
- The QBE-based CBIR complements the text-based querying, rather than tends to replace it

CBIR: Relevance Feedback

- Due to subjective choices and incompleteness of examples, the CBIR may not return the correct data at the very first querying step
- Dynamic dialog between the user and the CBIR system stimulates on-line learning and facilitates the target information search:
  - Relevance feedback: retrieved images, classified by the user into positive and negative matches, are used to reformulate the query

CBIR "IKONA": Relevance Feedback

N.Boujemaa and M. Ferecatu: Mining Multimedia Documents by Visual Content, ERCIM News No.47, October 2001 http://www.ercim.org/publication/Ercim_News/enw47/boujemaa.html

QBE Paradigm

- Human aspect: comfortable and intuitive querying
- Computational aspect: detailed automated analysis of the content of the query images
- Cognitive aspect: explicit knowledge representation for the search domain
  - Syntactic (literal) criteria of similarity and equality between pixels or features of images
  - Perceptual similarity / equality criteria
  - Physical, geometric and topological rules describing equality of and differences between 3D objects on images, etc