Wavelet theory and its applications to images retrieval

Aliaksandr Autayeu

Retrieval problem

- Rapid increase of availability of digital image producing devices

- Libraries of digital images became extremely large, even home photo albums are difficult to manage

Existing search methods

- **Attribute-based search**

  - Find pictures named *bear*.jpg
  - Find pictures from 20th of July.
  - Find pictures, shot for publication in "Nature"

- However:
- Attributes are loosely coupled with content
- Attributes do not represent visual features of the picture
- Attributes are "technical" and inconvenient to use

Existing search methods

- ## Keyword-based search

  - Find bears
  - Find sunsets with palm trees
  - Find one-storeyed standalone houses

- It is necessary to mark up the database with keywords. It is difficult and expensive work.

- Different descriptions of the same visual feature is natural "feature" of the language (synonyms)

Irrelevant results

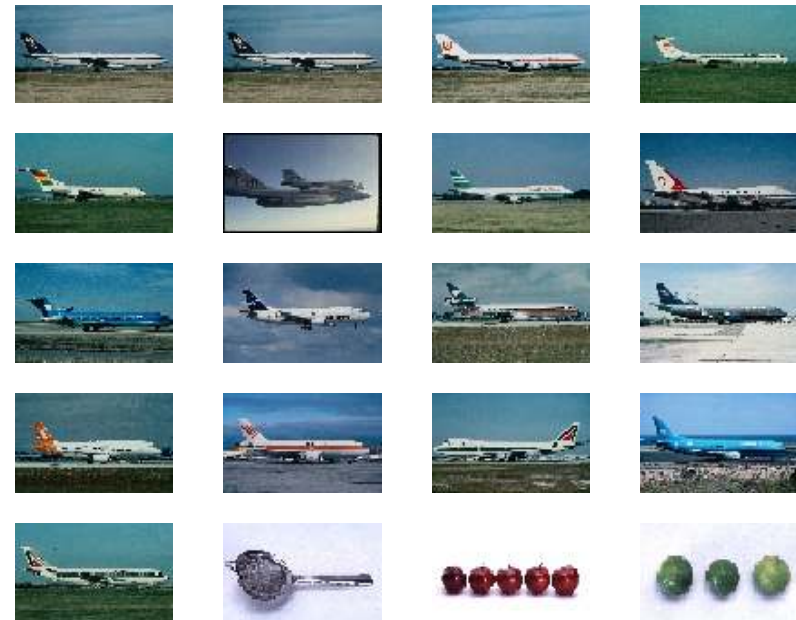- Result page for the simple query "round ancient building"

Content-based search

- «search by content»
- «query-by-example»
- «search by sketch»
- «similarity search».

Content based
image search:
powerful alternative
to traditional methods

Search results example.
Clustered database.
Top left image is query.

Algorithms

- Current algorithms are sharpen against global features, such as color and handle local features, such as shape, not very good.

Search results. Clustered database. Washington university algorithm. T. DeRose, D. Salesin.

Image searching systems

- HotBot (http://hotbot.lycos.com)
- NBCi (http://www.snap.com)
- Yahoo! Image Search (http://search.yahoo.com/images)
- Lycos multimedia searcher (http://multimedia.lycos.com)
- AltaVista Images Search Center (http://www.altavista.com/image)
- Google Image Search (http://images.google.com)
- PICSearch (http://www.picsearch.com/)

- QBIC (Query By Image Content). IBM Almaden Research Center, San Jose, CA
- Berkeley Digital Library Project. University of California, Berkeley
- Blobworld. Computer Science Division, University of California, Berkeley.
- Excalibur Visual RetrievalWare. Excalibur Technologies.
- VisualSEEk. Image and Advanced Television Lab, Columbia University, NY.
- WebSEEk. Image and Advanced Television Lab, Columbia University, NY

Washington University Algorithm

- **One of the first algorithms**

- **Efficient and original index design**

- **Handle shape not very good. Requires statistical tune-up.**

- **Metric**

$$\| Q - J \| = w_{0,0} \mid Q_{0,0} - J_{0,0} \mid + \sum_{i,j} w_{i,j} \mid Q1(i,j) - J1(i,j) \mid,$$

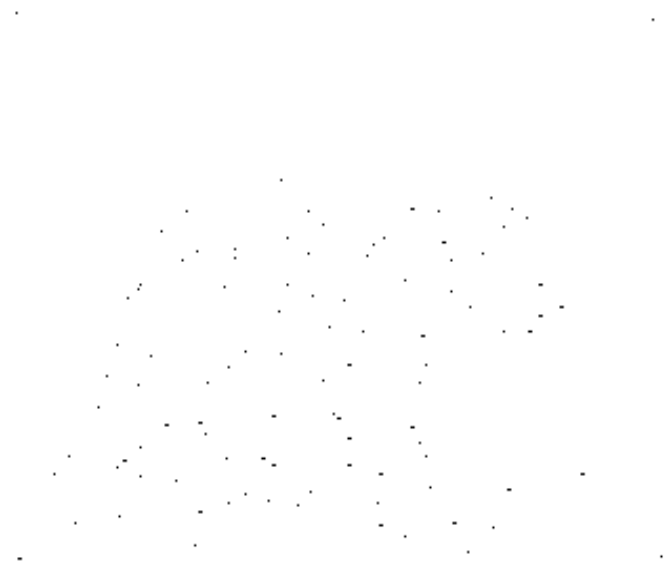where Q – query, J – image from database

Salient points algorithm

- Uses salient points idea, uses 2D wavelet decomposition and multiresolution presentation of the image for extraction of these points
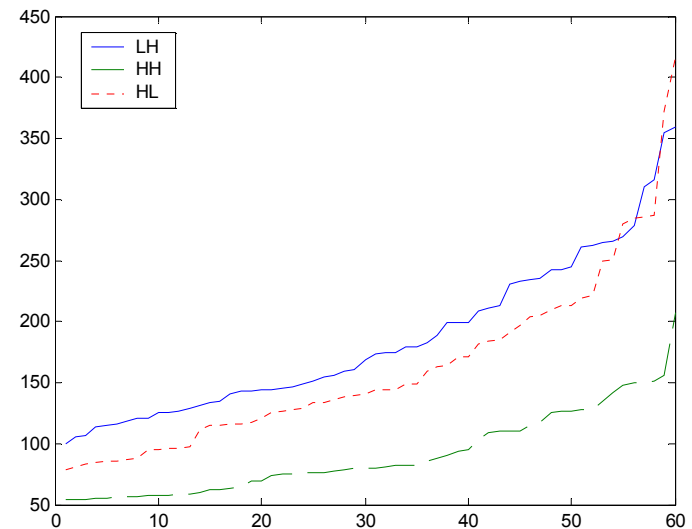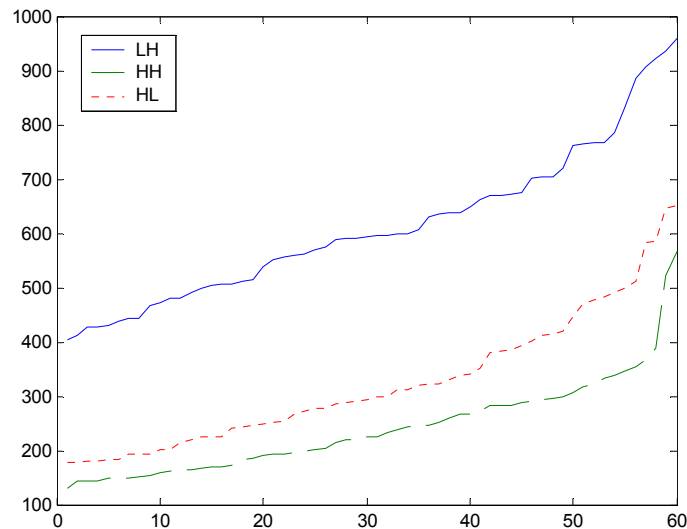
Original image   Extracted points

Extraction subband and points quality

Saliency values can be calculated along one decomposition subband.
Here is represented graphs of saliency values of the points,
extracted from two different images, using all three available subbands.
Visual difference between subbands illustrated below.

LH-subband

- LH-subband (60 points). Points stick to horizontal edges.

HL-subband

- HL-subband (60 points). Points stick to vertical edges.

HH-subband

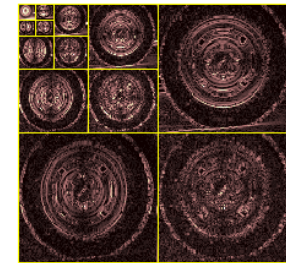- HH (60 points). Points stick to horizontal and vertical edges.

Metric

- Composite metric was used, which allows combination of the several features.

- i=1 shape
- i=2 color
- i=3 texture (not implemented)
- etc

$$D_j = \sum_i W_i S_j(f_i)$$

Query processing overview

- Decomposition
- Salient points
- Shape
- Color
- Metric

Multiresolution decomposition example
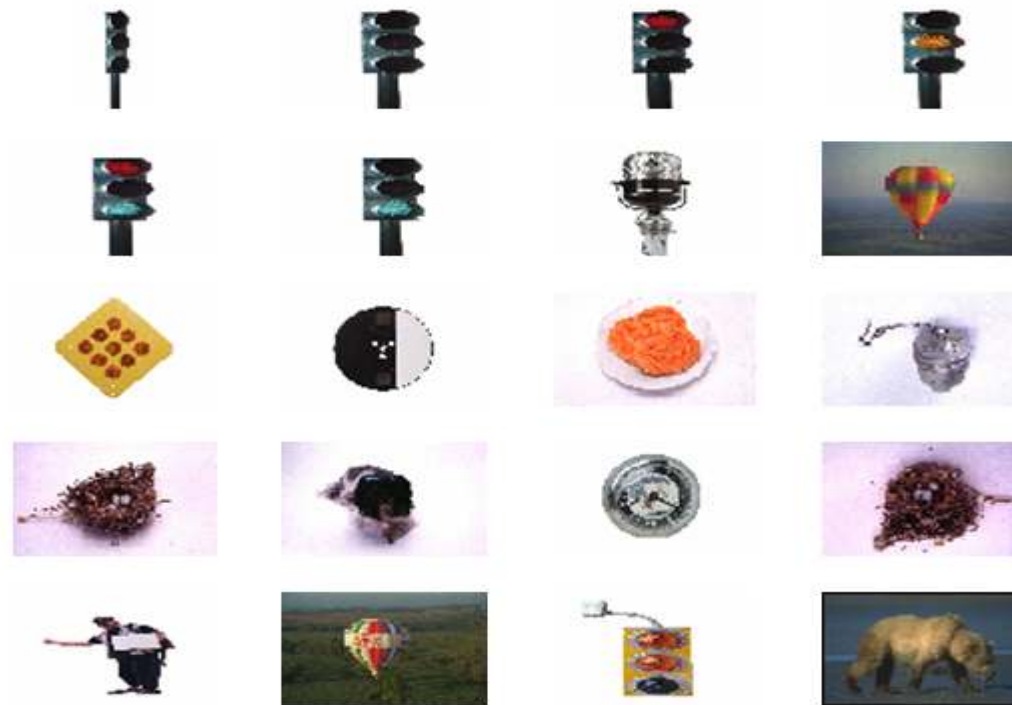

Salient points extraction example

Query example

- Clustered database (~800 images), algorithm correctly found all images of the target cluster (air liners) and correctly arranged them.

Query example

- Clustered database (~800 images), algorithm correctly found all images of the target cluster (traffic lights) and correctly arranged them.

Results

- Image searching algorithm, which handles several features, including shape

- Salient points extraction algorithm was investigated and expanded

Applications

- Internet image searching system
- Add-on for an existing document searching system, to allow search by illustrations.
- Intranet and desktop image-retrieval systems.