Colour similarity search in images

Norbert Fuhr

Outline

- Levels of image retrieval
- Human perception and colour spaces
- Colour histograms
- Colour layout
- Conclusion

Levels of image retrieval

- Syntactic level
  Which (sequences of) symbols occur in the image?
- Semantic level
  What's in the image?
- Pragmatic level
  What's the image about?

Pragmatic level

'meaning' of an image / topic illustrated by the image

New European regulations for lorry tolls

- Topics are rather subjective
- But pragmatic level important for many applications

Semantic level

Objects shown in the image (+ spatial relationships)

Example: Google Image search for 'kangaroo'

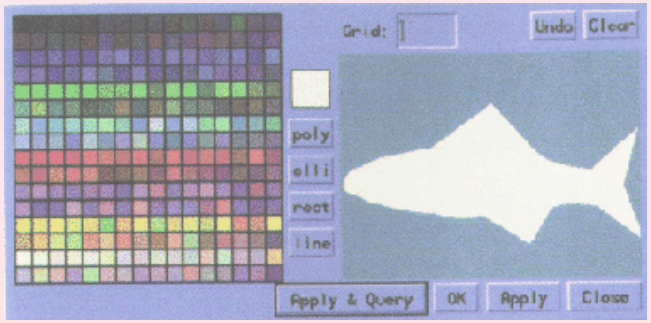
(based on text search in file name and image caption)

Automatic semantic indexing works only in limited domains

Syntactic level

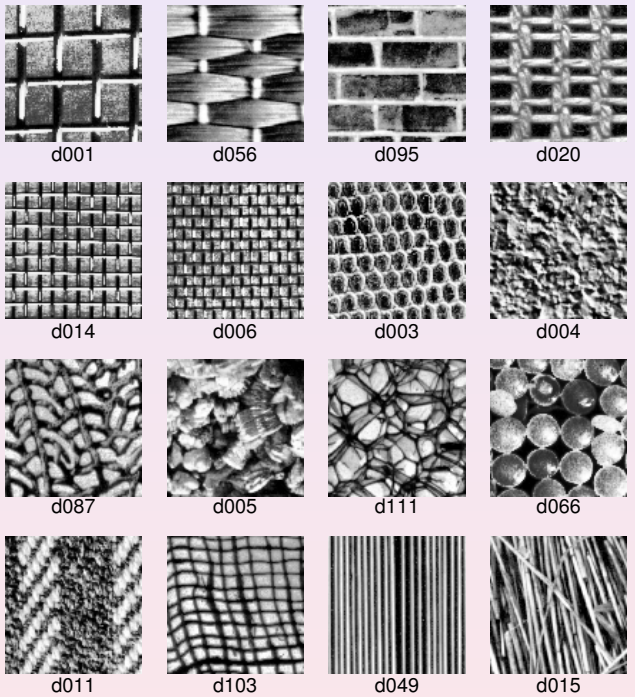
Image as matrix of pixels with colour values

1. Contours



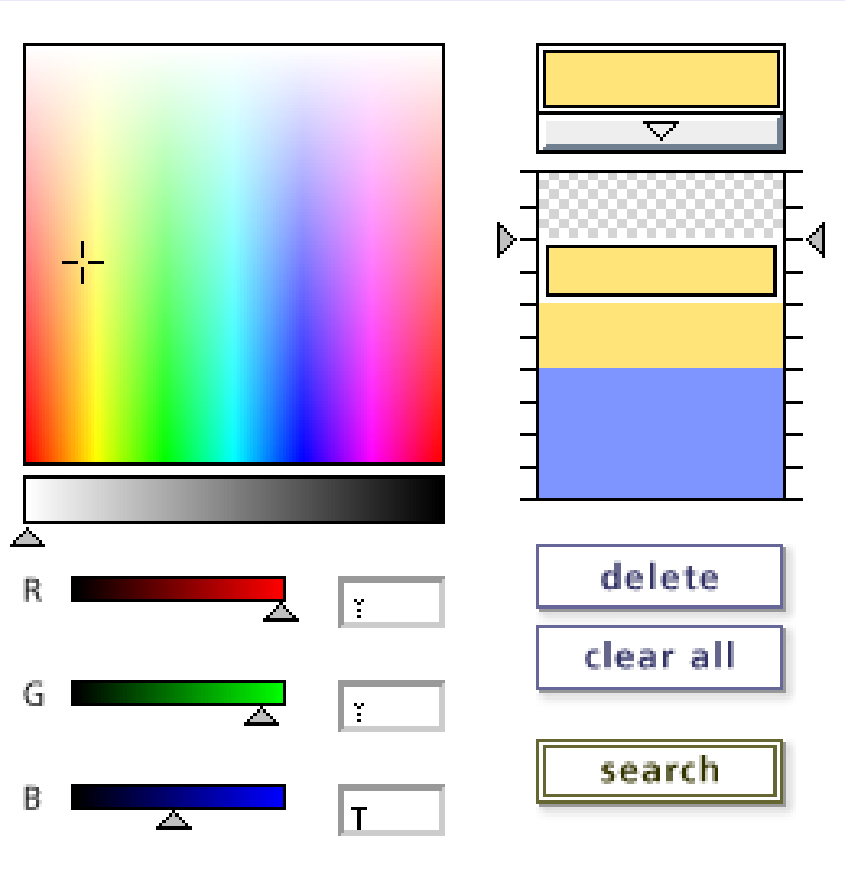
Syntactic level: 2. Textures

Textures:
patterns in the luminance band (greylevel image)
structural and/or statistical patterns



Syntactic level: 3. Colour

frequency / spatial distribution of pixel colours



1) Boats at Saintes-Maries

Gogh, Vincent van

2) Port

Derain, Andre
Circa 1905

Human perception and colour spaces

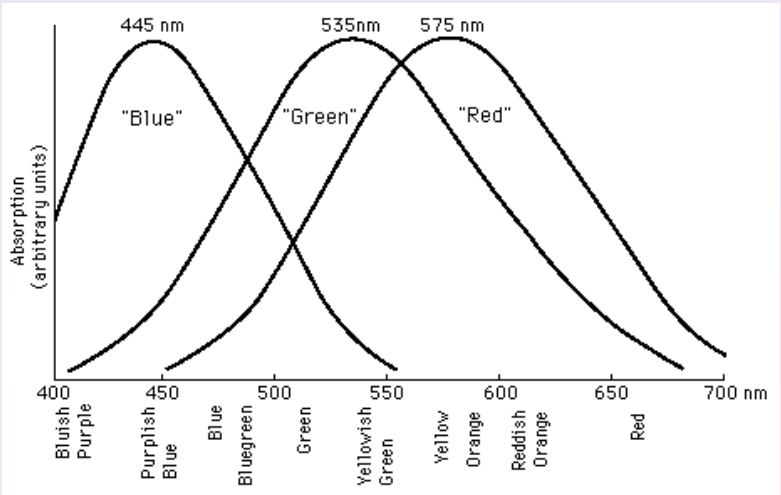
visible light:
$\lambda \in [380nm \ldots 780nm]$
([violet ... red])

retina:

- rods for brightness
- cones for chromaticity (colour)

three types of cones:

- red: $\lambda_x = 575nm$
- green: $\lambda_y = 535nm$
- blue: $\lambda_z = 445nm$



Response curves of the three types of cones

Colour Systems

$\varphi(\lambda)$: wavelength distribution of source light

k: normalization factor

$\bar{x}(\lambda)$, $\bar{y}(\lambda)$, $\bar{z}(\lambda)$: eye response functions

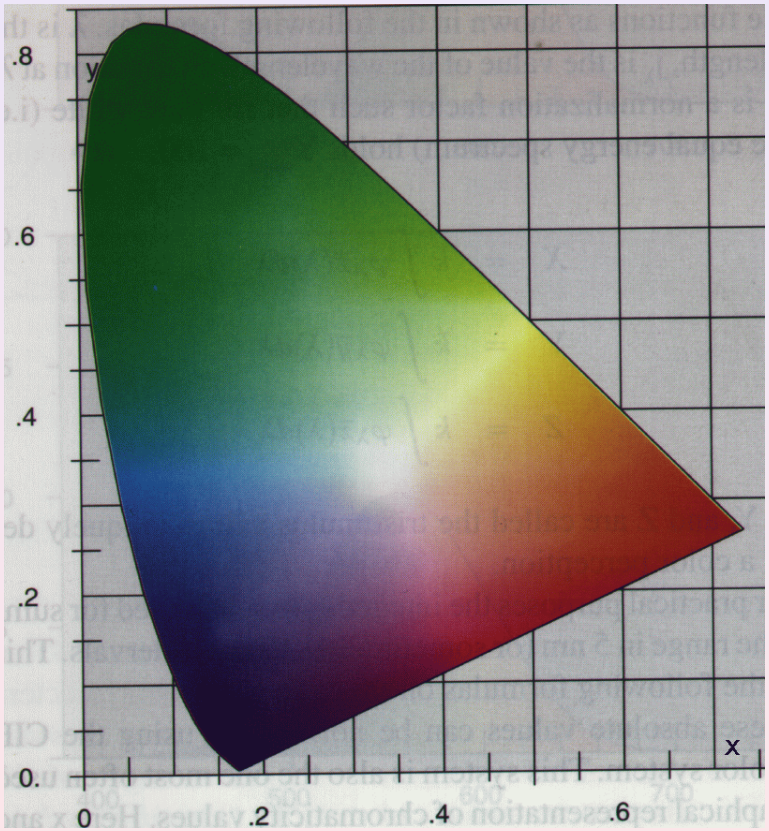
X, Y, Z: perceived colour intensity

$$
\begin{aligned}
X &= k \int \varphi(\lambda)\bar{x}(\lambda)d\lambda \\
Y &= k \int \varphi(\lambda)\bar{y}(\lambda)d\lambda \\
Z &= k \int \varphi(\lambda)\bar{z}(\lambda)d\lambda
\end{aligned}
$$

The CIE Yxy colour system

CIE Yxy colour system

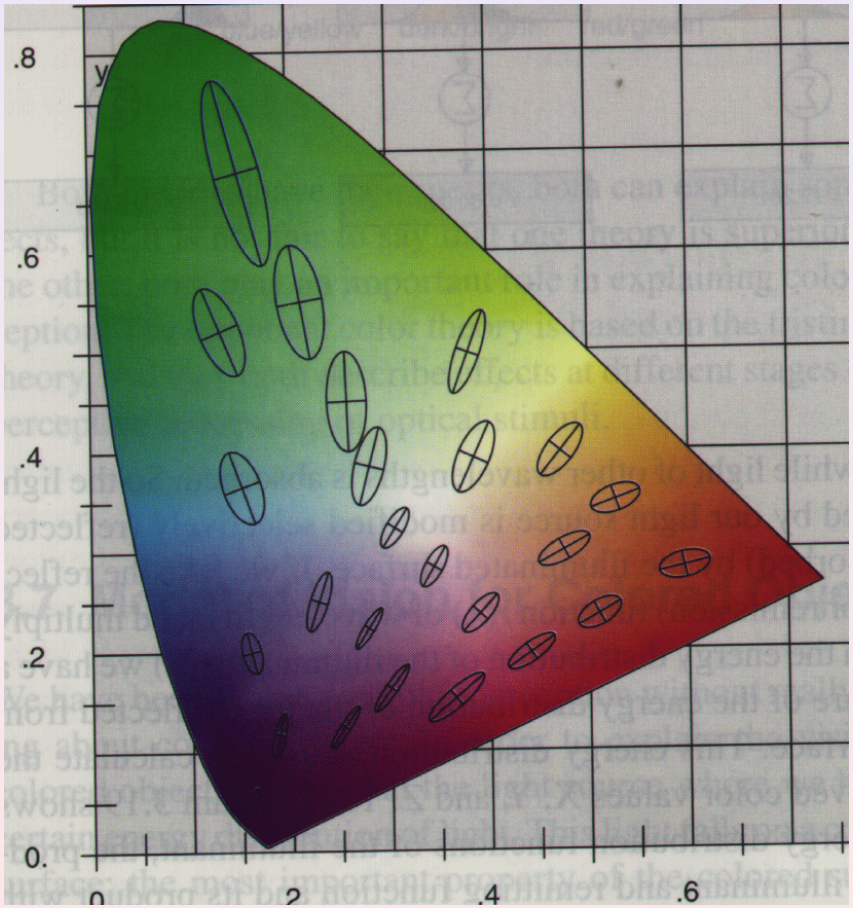
$$x = \frac{X}{X+Y+Z}$$
$$y = \frac{Y}{X+Y+Z}$$
$$z = \frac{Z}{X+Y+Z} = 1 - x - y$$



Colour similarity

Perceived visual distance
(McAdams ellipses)

$\longrightarrow$ need for colour space with uniform distances



Colour Spaces

Non-uniform colour spaces:

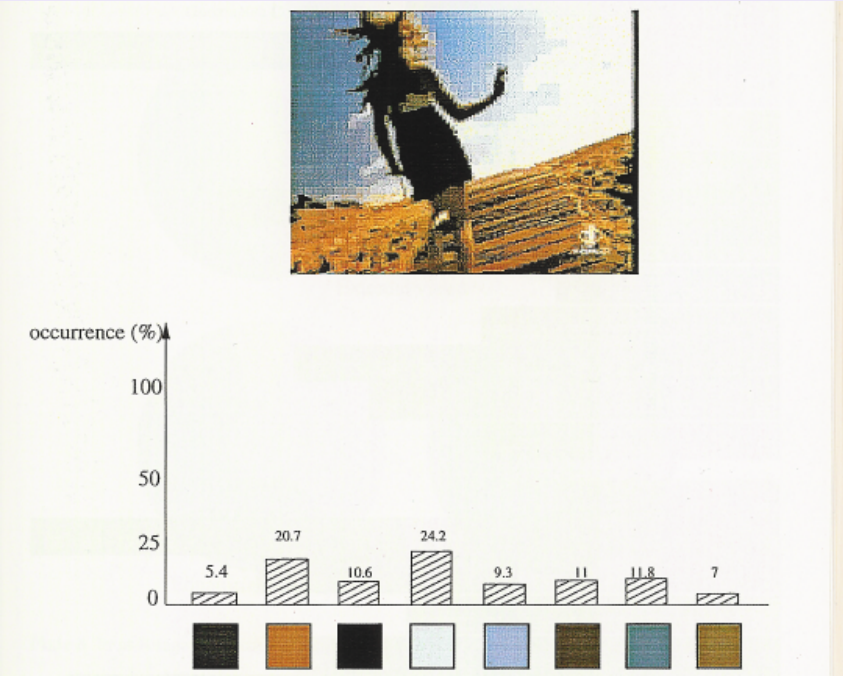
RGB	red, green, blue — used for displays
CMY(k)	cyan, magenta, yellow (black) — used for printers
YCrCb	used in JPEG digital image standard

Uniform colour spaces

MTM	approximation of the (perceptually uniform) Munsell colour space
$L^*a^*b^*$	recommended by CIE for near-daylight colour
$L^*u^*v^*$	recommended by CIE for monitor displays and additive light-source conditions

Mappings between colour spaces: $L^*u^*v^*$ and $L^*a^*b^*$ values can be derived from Yxy values by nonlinear transformations

Colour histograms



Histogram: $D = (d_1, \ldots, d_n)^T$ (with $n =$# colours)

Problems:

- Dealing with large numbers of colours
- Similarity function for histograms

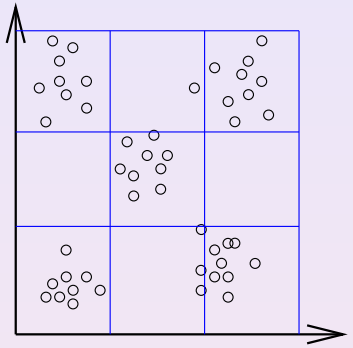
Size of the histogram

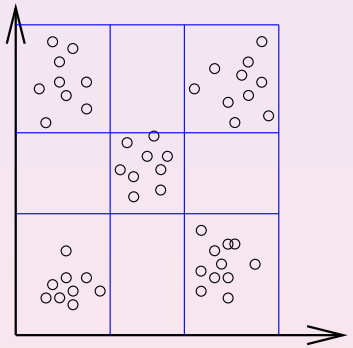
Colour models: 3-dimensional colour space
8 bits per colour channel: $= 2^{24}$ possible colours
$\rightarrow$ quantization necessary (for reducing # bins)

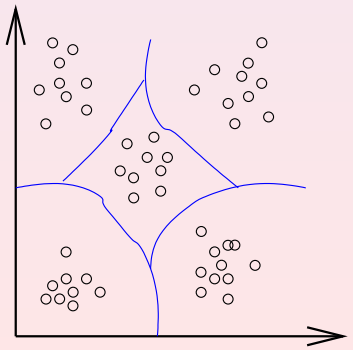
Colour histogram:
b_i bins for ith colour dimension, $i = 1, 2, 3$
$\rightarrow$ histogram = vector with $n = b_1 \cdot b_2 \cdot b_3$ values

Quantization methods

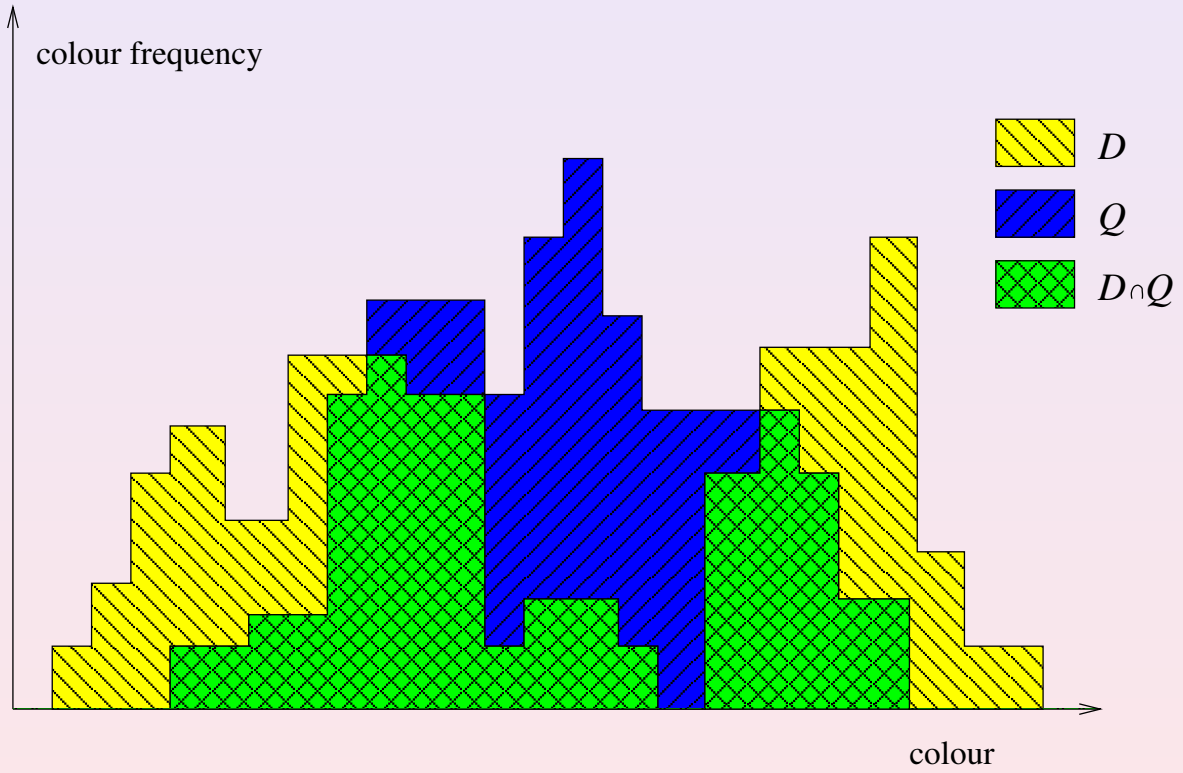
- **uniform:** divide each axis into intervals of equal length
- **product vector:** minimize mean-squared error for each dimension
- **LGB:** minimize mean-squared error resulting from quantization subdivide 3D colour space into n subspaces s.th. resulting error is minimized







Histogram similarity



Similarity metrics

L_1 and L_2 (Euclidean distance):

$$H_1(Q, D) = \sum_{i=1}^{n} |q_i - d_i|$$

$$H_2(Q, D) = \sqrt{\sum_{i=1}^{n} (q_i - d_i)^2}$$

$\longrightarrow$ poor retrieval results

Histogram intersection

$$H_{\min}(Q, D) = \frac{\sum_{i=1}^{n} \min(q_i, d_i)}{\sum_{i=1}^{n} d_i}$$

$\longrightarrow$ gives better retrieval results

Considering colour similarity

Colour similarity matrix:

$$\mathbf{A} = [a_{ij}], \qquad i = 1, \ldots, n; \quad j = 1, \ldots, n$$

where a_{ij} gives similarity between colours i and j
based on Euclidean distance d_{ij} in a uniform space,
with $a_{ij} = 1 - d_{ij}$

Improved colour histogram similarity:

$$H_s(Q, D) = (Q - D)^T \mathbf{A}(Q - D) = \sum_{i=1}^{n} \sum_{j=1}^{n} (q_i - d_i) a_{ij} (q_j - d_j)$$

Colour histogram search: Example

1) Boats at Saintes-Maries

Gogh, Vincent van
1888

2) Port

Derain, Andre
Circa 1905

3) Lake

Hodler, Ferdinand
1910-1912s

4) The Road to Asgaard: Adirondacks

Kent, Rockwell
1960

5) View of the Rocks on the Third Island of Cyclops

6) Palace in Athens

Klenze, Leo von
1835

Colour layout

Consider spatial distribution of colours

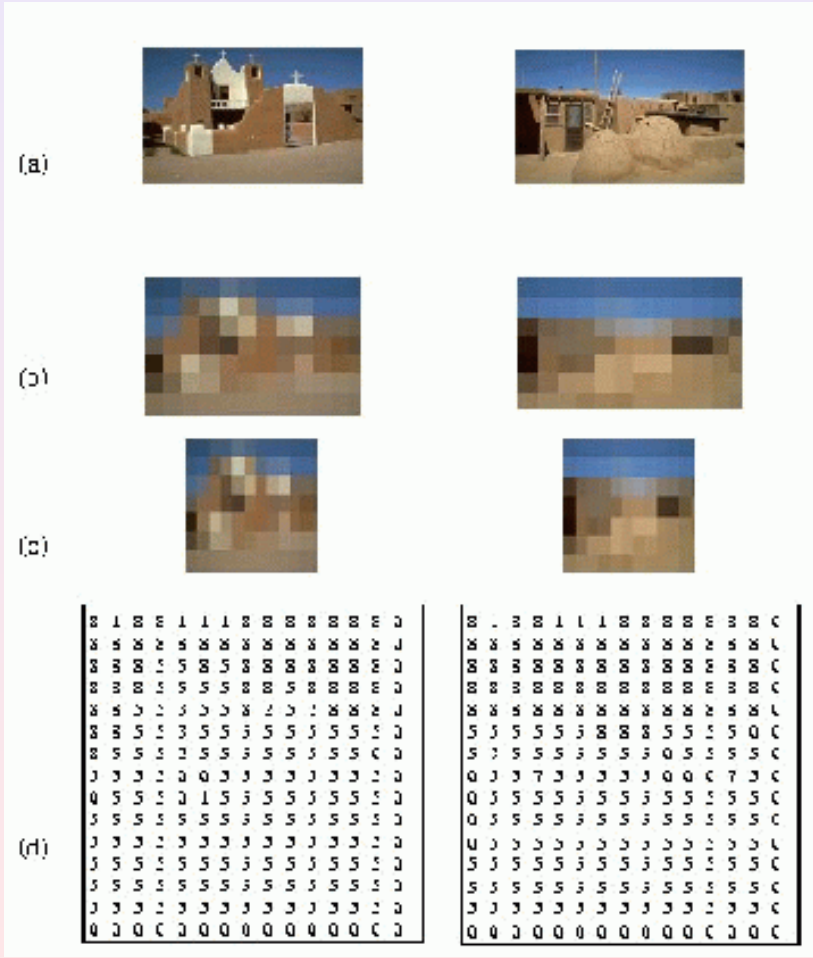
"fuzzy images": reduce resolution similarity matching based on 15×15 grid

(a) map image colours onto colour bins

(b) reduce resolution to 15×15 grid

(c) scale to square pattern

(d) construct matrix of colour group numbers



Problems in colour layout search



Problems:

missing **invariance** to scaling, rotation, reflection and translation

Colour cooccurrence descriptors

c_i – colour of pixel i
c_j – colour of pixel j
d_{ij} – euclidian distance between pixel i and j
represent image as matrix

$$\mathbf{W}(c_i, c_j, d_{ij})$$

W frequency of cooccurrence of colours c_i and c_j at distance d_{ij}

Cooccurrence Descriptors: Example

1	2	1
2	3	2
1	2	1

Example image
(color values)

c_i	c_j	d_{ij}	w
1	1	2	4
1	1	3	2
1	2	1	8
1	2	8	8
1	3	1	4
2	3	1	4

representation is invariant to rotation, reflection and translation!

Matrix W stored as set of elements:

$$E_k \in \{(i_k, w_k) | \exists w_k = \mathbf{W}(c_i, c_j, d_{ij}) \neq 0 \wedge i_k = f(c_i, c_j, d_{ij})\}$$

(i_k – element index)

Image retrieval

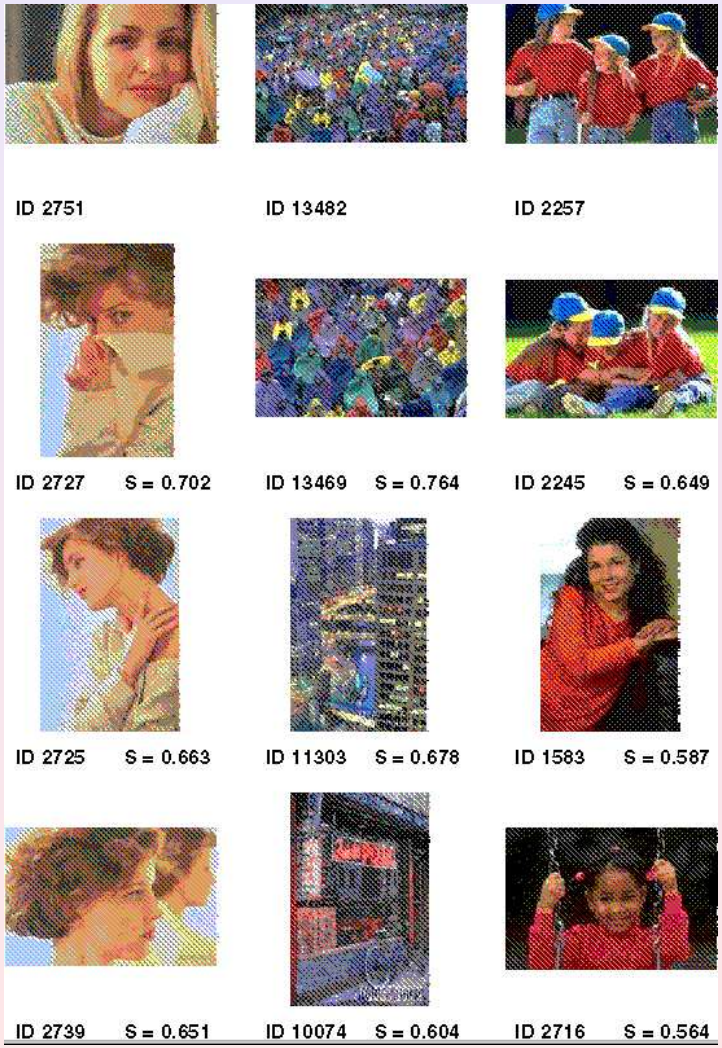
T_q – query image descriptor
T_t – target image descriptor

distance measure:

$$L(T_q, T_t) = \frac{\sum_{E_k \in T_q \cap T_t} |w_k^q - w_k^t|}{\sum_{E_k \in T_q} w_k + \sum_{E_k \in T_t} w_k}$$

similarity measure:

$$S(T_q, T_t) = 1 - L(T_q, T_t)$$



Conclusion

- Image retrieval can be performed at the syntactic, semantic and pragmatic level
- At the syntactic level, one can search for similar colours, contours and textures
- Colour search should be based on a (perceptually) uniform colour space
- Colour histogram similarity requires reduction of the number of colours
- Similarity metrices should be based on the intersection of histograms and/or the similarity of colours
- Colour layout search should be invariant to image transformation such as rotation, reflection, translation and scaling