The University of Auckland

CBIR: Colour Features

COMPSCI.708.S1.C

A/P Georgy Gimel'farb

The University of Auckland

Other Opponent Colour Spaces

- **YUV** – PAL Phase Alternating Line TV (most of the European countries, some Asian countries, Australia, and New Zealand)
- Luminance (**Y**) and chrominance (**U**,**V**):

$$\mathbf{Y} = 0.299\mathbf{R} + 0.587\mathbf{G} + 0.114\mathbf{B}$$

$$\mathbf{U} = 0.492(\mathbf{B}-\mathbf{Y}) = -0.147\mathbf{R} - 0.289\mathbf{G} + 0.436\mathbf{B}$$

$$\mathbf{V} = 0.877(\mathbf{R}-\mathbf{Y}) = 0.615\mathbf{R} - 0.515\mathbf{G} - 0.100\mathbf{B}$$

The University of Auckland

Other Opponent Colour Spaces

- **YIQ** – NTSC National Television Systems Committee TV (USA, Canada, and Japan) with the luminance **Y** and chrominance **IQ** (equal to **UV** rotated by 33°):

  $\mathbf{Y} = 0.299\mathbf{R} + 0.587\mathbf{G} + 0.114\mathbf{B}$

  $\mathbf{I} = -0.545\mathbf{U} + 0.839\mathbf{V} = 0.596\mathbf{R} - 0.275\mathbf{G} - 0.321\mathbf{B}$

  $\mathbf{Q} = 0.839\mathbf{U} + 0.545\mathbf{V} = 0.212\mathbf{R} - 0.523\mathbf{G} + 0.311\mathbf{B}$

- $\mathbf{YD_bD_r}$ – SECAM (Sequential Couleur a Memoire) TV (France, Russia, Eastern Europe) is the scaled **YUV**:

  $$\mathbf{D}_b = 3.059\mathbf{U} \text{ and } \mathbf{D}_r = -2.169\mathbf{V}$$

The University of Auckland

Other Opponent Colour Spaces

- **YC_bC_r** – JPEG and MPEG coding standards: non-negative chrominance components by scaling and shifting the **YUV** co-ordinates:

  $\mathbf{Y} = 0.257\mathbf{R} + 0.504\mathbf{G} + 0.098\mathbf{B} + 16$

  $\mathbf{C_b} = -0.148\mathbf{R} - 0.291\mathbf{G} + 0.439\mathbf{B} + 128$

  $\mathbf{C_r} = 0.439\mathbf{R} - 0.368\mathbf{G} - 0.071\mathbf{B} + 128$

- CIE uniform **Lab** (colour fax) and **Luv** luminance - chrominance colour spaces

Other Opponent Colour Spaces

- CIE **Lab** and **Luv** colour spaces compress well in the case of pictorial images:

RGB

L (luminance)

a

b

HMMD colour space

- **Hue-Min-Max-Difference**: a new space in MPEG-7
  - Used in the colour structure descriptor (CSD)
  - MPEG-7 supports also the greyscale (intensity only) space, RGB, HSV, and YC_rC_b colour spaces
  - **Hue**: as in HSV space for MPEG-7
  - **Min / Max**: $\max\{R, G, B\}$ and $\min\{R, G, B\}$
  - **Difference**: $\max\{R, G, B\} - \min\{R, G, B\}$
    - Only 3 of the 4 components define the HMMD colour space
    - **Intensity**: $(\min\{R,G,B\} + \max\{R,G,B\})/2$
    - **Chromaticity** relates to the difference component

Vector Quantisation

- The whole 3D colour space is partitioned into K disjoint subsets, one per code word $\mathbf{c}_k$; $k=1,\ldots,K$
- All the colours of one subset are represented by, or quantised to, the same code word $\mathbf{c}_k$
- In a perceptually good palette $\mathbf{C} = \{\mathbf{c}_1,\mathbf{c}_2,\ldots,\mathbf{c}_K\}$, each subset contains visually similar colours
- *Multidimensional clustering* of colours: K-means, fuzzy K-means, EM (Expectation - Maximisation)

Vector Quantisation Algorithm

- Doubling the number of cluster centres until a prescribed limit
- Each iteration t: $K_t = 2^t$ centres $\boldsymbol{C}_t = \{\mathbf{c}_{k,t}\text{:}\ k = 1, \ldots, K_t\}$
  - $\mathrm{T} = 0$ - a single centre $\boldsymbol{c}_{1,0}$ averaging colour vector over an image
  - At each next iteration t each previous centre $\mathbf{c}_{k,t-1}$; $k = 1, \ldots, K_{t-1}$, splits into the two new centres:
    - Provisional splitting into $\mathbf{c}_{\mathrm{pr}:\,k,t}$ and $\mathbf{c}_{\mathrm{pr}:\,K_{t-1}+k,t}$ by multiplying $\mathbf{c}_{k,t-1}$by the fixed factors $(1+w)$ and $(1-w)$, respectively (w - a fixed value), or shifting it to and from the most distant vector in the cluster, etc
    - Assigning each colour vector in the image to the closest cluster centre
    - Forming the new centres $\mathbf{c}_{k,t}$ and $\mathbf{c}_{K_{t-1}+k,t}$ by averaging the colour vectors assigned to each such cluster

The University of Auckland

Colour Descriptors

- ***Colour histogram***: the distribution of colours within a whole image or within a specified region
  - invariance to rotation, translation, and scaling
  - but: it captures no semantic information
- ***Colour set***: by thresholding the colour histogram
- ***Colour moments***: the mean colour, variance, etc.
- ***Colour correlogram***: changes of spatial correlation of pairs of colours with distance

The University of Auckland

Colour Histogram

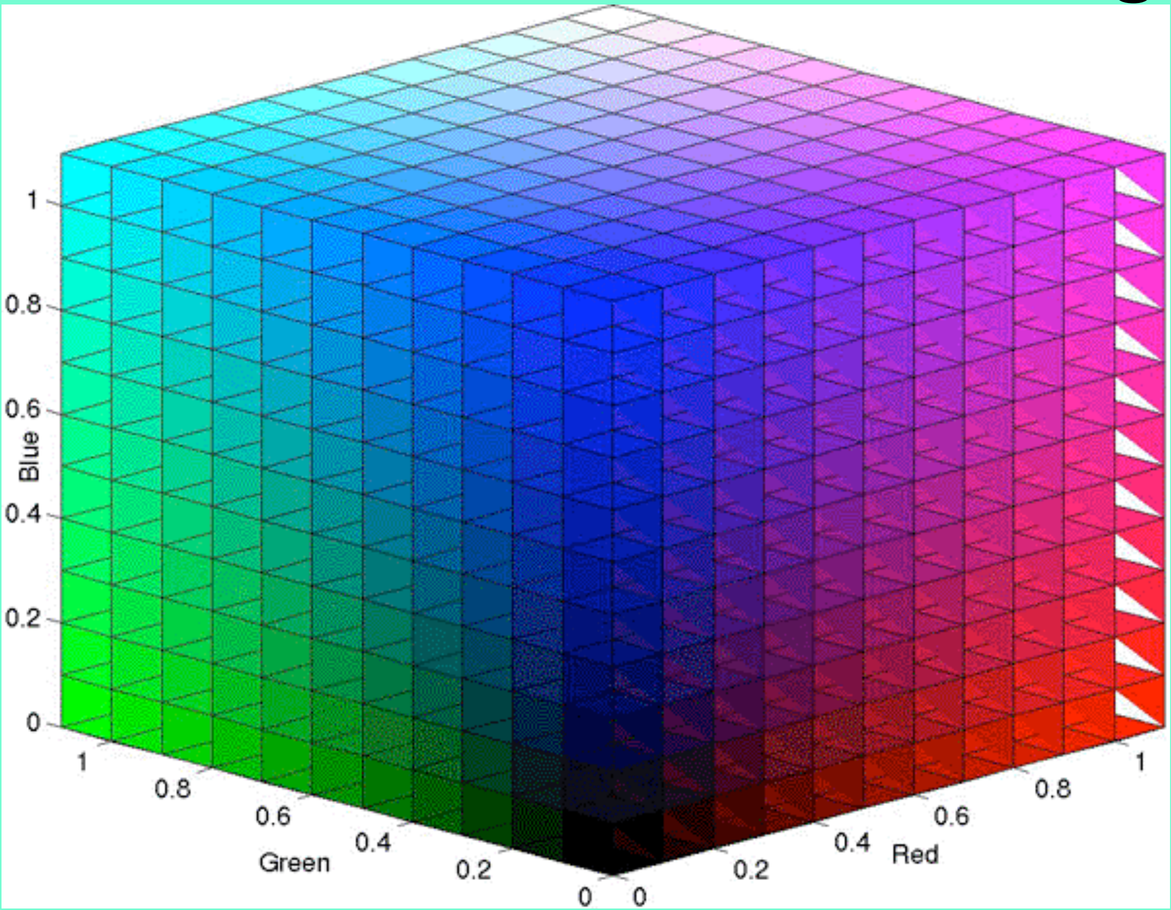
- Distribution of colours within a whole image or a specified region
- Invariant to rotation, translation, and scaling of an object
- Does not capture semantic information: two images with similar colour histograms can possess totally different contents
- Quantised HSI (HSV) colour space is typically used to make the search partially invariant to illumination and object viewpoints
  - In the HSI (HSV) space, an Euclidean or similar component-wise distance between the colours specifies colour similarity quite well
  - The YUV colour space is also often used since it is standard for the MPEG video compression

The University of Auckland

Colour Histogram

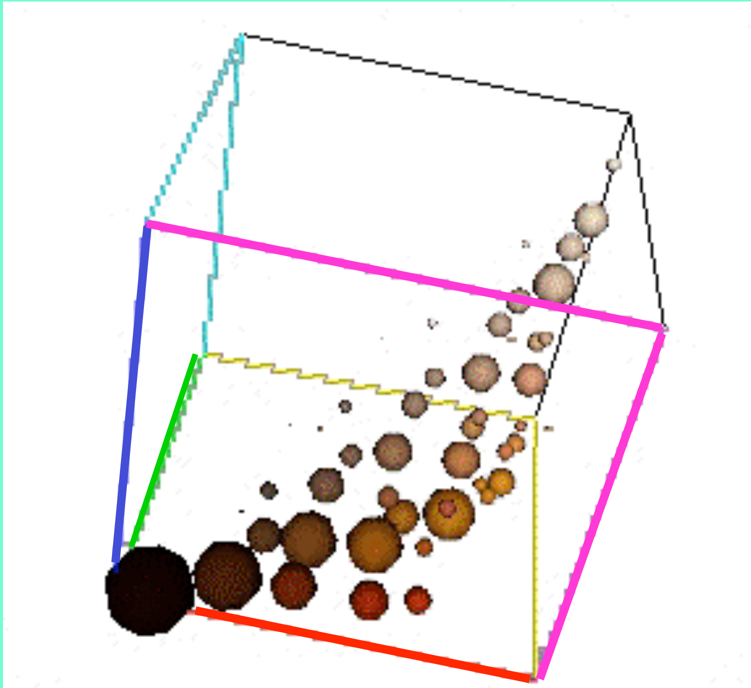
- Colour histogram $\mathbf{h}(\text{image})=(h_k(\text{image}): k=1,\ldots,K)$ is a K-dimensional vector
  - each component $h_k(\text{image})$: the relative number of pixels of colour $\mathbf{c}_k$ in the image (i.e. the fraction of pixels being most similar to the corresponding representative colour)
  - To built the colour histogram, the image colours should be transformed to an appropriate colour space and quantised according to a particular codebook of the size K
  - Histograms for image regions roughly account for spatial colour distribution

Colour Bins for an RGB Histogram



11×11×11 RGB colour space (from http://www.owlnet.rice.edu/~elec301/Projects02/artSpy/color.html)

Colour Histogram



6×6×6 histogram (sphere-size-coded bin values)

From ij-plugins.sourceforge.net/ij-vtk/color-space/

Matching Colour Histograms

- **QBE**: the database image compared to the query has to be re-quantised by finding for every pixel the closest colour in the query codebook
- Distance between two normalised histograms:

$$D(\mathbf{h},\mathbf{h}') = \begin{cases} \sum_{k=1}^{K} | h_k - h'_k | \\ \sum_{k=1}^{K} \left(h_k - h'_k \right)^2 \\ 1 - \sum_{k=1}^{K} \min\left\{ h_k, h'_k \right\} \equiv 0.5 \sum_{k=1}^{K} \left| h_k - h'_k \right| \end{cases}$$

Matching Colour Histograms

The University of Auckland

- <u>Quadratic-form metrics</u>
  - comparing all colour pairs to account for cross-relations of different colour clusters:

$$D(\mathbf{h},\mathbf{h}') = (\mathbf{h} - \mathbf{h}')^{\mathsf{T}}\mathbf{A}(\mathbf{h} - \mathbf{h}')$$

  $\mathbf{A} = [a_{ij}]$ is the positive definite symmetric $K \times K$ matrix

  Its components $a_{ij} = a_{ji}$ specify the dissimilarity between the codewords $\mathbf{c}_i$ and $\mathbf{c}_j$ for the histogram components with indices i and j

  Covariance matrix
- *Special case*: Mahalanobis distance: $\mathbf{A} = \Sigma^{-1}$

The University of Auckland

MPEG-7 Colour Descriptors

Dominant Colour Descriptor

- YC_rC_b
- Y only (monochrome)
- RGB
- HSV
- HMMD

Scalable Colour Descriptor (SCD)

- HSV
- Group of frames / pictures histogram (GoP)

Colour Structure Descriptor (CSD)

- HMMD

Colour Layout Descriptor

- YC_rC_b

- Generic colour histogram accurately captures the colour distribution for image search and retrieval but has too many independent characteristics to choose, e.g. colour space, quantisation in that space, and quantisation of the histogram values
- For the interoperability between different MPEG-7 descriptors: the histogram-based descriptors are limited to the SCD and the CSD

Scalable Colour Descriptor

- HSV colour space uniformly quantised to 16 … 256 bins
  - Histogram values: truncated to 11 bits; non-linearly mapped to 4 bits

bins	16	32	64	128	256
H	4	8	8	8	16
S	2	2	2	4	4
V	2	2	4	4	4

- Haar transform coding: $a,b \Rightarrow a+b,\ a-b$
  - Repetitive HT: 256 bins - 128, 64, … bins; sign-alternate high-pass coefficients: a truncated integer representation
- Similarity matching of the SCD histograms: the absolute distance
  - The same matching for the Haar coefficients
  - Fast coarse-to-fine matching
- **GoP descriptor**: extending the SCD to a collection of images

The University of Auckland

Colour Structure Descriptor

- HMMD colour space + an 8 x 8 structuring element
  - Colour structure histogram $\mathbf{h}_{\mathrm{cs}}=(h_{\mathrm{cs},k} : k=0,1,...,K-1)$ has in each bin k the number of structuring elements in the image containing one or more pixels with colour $\mathbf{c}_k$
  - Bin values: normalised to the range [0.0, 1.0] and quantised to 8 bits / bin
  - 184-bin CSD is approximated with 120, 64, or 32 bins
  - Subsampling of the images larger than 2^{16} pixels: the 8 x 8 structuring element with the spatial extent $E \times E$:

    $$E = 8S;\ S = 2^p;\ p = \max\{0, \mathrm{int}\{0.5\log_2(WH)\} - 8\}$$

    where W and H denote the width and height of an image, respectively, and $\mathrm{int}\{z\}$ is the closest to z integer

Dominant Colour Descriptor

- Representative colours from each image instead of being fixed in the colour space:

$$\text{DCD} = \{\{(\mathbf{c}_i, p_i, v_i): i=1,\ldots,n\}, s\}; \quad p_1 + \ldots + p_n = 1$$

  - n: the total number of colour clusters in the image region: $1 \le n \le 8$; typically, 3 - 4 colours describe satisfactorily the colour region
  - **Attribute triple**: 3D colour vector $\mathbf{c}_i$, its percentage p_i, and its variance v_i
  - s: the overall colour spatial coherency
  - **Colour clustering** minimises the weighted scatter D_i in each cluster C_i:

$$D_i = \sum_{\substack{(x,y):\\ \mathbf{c}(x,y)\in C_i}} w_{x,y} \left\| \mathbf{c}(x,y) - \mathbf{c}_i \right\|^2 \Rightarrow \mathbf{c}_i = \sum_{\substack{(x,y):\\ \mathbf{c}(x,y)\in C_i}} w_{x,y}\mathbf{c}(x,y) \Bigg/ \sum_{\substack{(x,y):\\ \mathbf{c}(x,y)\in C_i}} w_{x,y}$$

K-means-like algoritm

perceptual weight depending on local pixel statistics

Dominant Colour Descriptor

- **Spatial coherency**: overall spatial homogeneity of the dominant colours in the image
  - Spatial coherency of a dominant colour: the normalised average number of connected pixels of this colour (computed using a 3×3 mask)
  - Overall coherency: a linear combination of the individual spatial ones weighed with the corresponding percentages p_i
- **Dissimilarity** between DCDs with no optional variance/coherency:

$$\mathrm{DCD}_1=\{(\mathbf{c}_{1i}, p_{1i}) : i=1,\ldots,n_1\};\ \mathrm{DCD}_2=\{(\mathbf{c}_{2i}, p_{2i}) : i=1,\ldots,n_2\}$$

$$D^2\left(\mathrm{DCD}_1,\mathrm{DCD}_2\right)=\sum_{i=1}^{n_1} p_{1i}^2+\sum_{j=1}^{n_2} p_{2j}^2-2\sum_{i=1}^{n_1}\sum_{j=1}^{n_2} a_{1i,2j}\, p_{1i}\, p_{2j}$$

Similarity coefficient depending on the Euclidean distance between the two colours

The University of Auckland

Colour Layout Descriptor (CLD)

- Representative colours on an 8×8 grid ⇒ a discrete cosine transform (DCT) ⇒ encoding of the resulting coefficients
  - Input image ⇒ into 8×8 blocks ⇒ average colours in the YC_rC_g space
  - Average colours ⇒ a series of 8×8 DCT coefficients (independently for Y, C_r, and C_g components) ⇒ a few low-frequency coefficients by zigzag scanning and quantisation:

$$\text{CLD} = \{(\Psi_{Y,i}, \Psi_{C_r,i}, \Psi_{C_g,i}): i = 1, \ldots, \nu\}$$

    - $\Psi_{\ldots,i}$ - the i-th DCT coefficient of the corresponding colour component
    - ν - the number of the coefficients (generally different for each component)

$$D(\text{CLD},\text{CLD}') = \sqrt{\sum_{i=1}^{\nu} w_{Y,i}\left(\Psi_{Y,i} - \Psi'_{Y,i}\right)^2} + \sqrt{\sum_{i=1}^{\nu} w_{C_r,i}\left(\Psi_{C_r,i} - \Psi'_{C_r,i}\right)^2} + \sqrt{\sum_{i=1}^{\nu} w_{C_g,i}\left(\Psi_{C_g,i} - \Psi'_{C_g,i}\right)^2}$$

Larger weights - the to lower frequency coefficients