Fuzzy rough c-means for image colour quantisation

Gerald Schaefer\textsuperscript{1}, Qinghua Hu\textsuperscript{2}, Huiyu Zhou\textsuperscript{3} and James F. Peters\textsuperscript{4}

\textsuperscript{1}Loughborough University, U.K.  
\textsuperscript{2}Harbin Institute of Technology, China  
\textsuperscript{3}Brunel University, U.K.  
\textsuperscript{4}University of Manitoba, Canada
Outline

• The colour quantisation problem.

• Colour quantisation algorithms.
  – Popularity, median cut, octree, Neuquant.

• Clustering based colour quantisation.
  – Hard c-means.
  – Fuzzy c-means.
  – Rough c-means.
  – Fuzzy rough c-means.

• Experimental results.

• Summary
Colour images

- High quality colour images can contain many different colours.
- Variety of colour allows for good image quality (smooth shading, etc.).
- True colour images have a colour resolution of 24 bits, 8 bits for each channel
  - i.e. each pixel is represented by three 8-bit numbers (one for red, green, and blue each).
- $2^{24}$ (~16.8 mio) possible colours.
- Not all of them are used in an image.
Colour quantisation

• Sometimes it is useful to use fewer colours.
  – For devices with limited hardware, e.g. mobile devices.
  – Image compression.
  – Image retrieval.
  – Image analysis and pre-processing.

• How can an image be displayed with fewer colours?

• Colour quantisation: Select a subset of colours and map each pixel to one of them

• Image coded as palette (codebook) and indices to palette.
Colour quantisation (2)

True colour image

[RGB RGB RGB RGB …]

Colour quantised image (16 colours)

[ I I I I …]

Palette
Colour quantisation (3)

true colour

256 colours

16 colours

4 colours

2 colours
Colour quantisation (4)

• Image quality depends directly on choice of colour in the colour palette.

• Finding good codebook of $n$ colours is crucial for resulting image quality.

• Aim: finding a codebook so that resulting image quality is maximised.

• But: problem of finding optimal codebook is $np$-complete!

• Heuristic and statistical approaches.
Colour quantisation algorithms – Popularity algorithm

- Take the $n$ most frequent colours.
Colour quantisation algorithms – Median cut algorithm
Median cut quantisation
Colour quantisation algorithms – Octree quantisation

- Successively subdivide RGB cube to build octree.
- Merge subtrees so as to maintain best possible image quality.
Octree quantisation
Colour quantisation algorithms – Neuquant

- 1-dimensional self-organising map is built.
- Learning algorithm to adapt to image in order to build palette.
Neuquant quantisation
Clustering for colour quantisation

- Colour quantisation can also be seen as a clustering problem.

- Pixels = samples.
- Cluster centres = palette entries.

- Clustering is $np$-complete.
C-means / k-means

- **Objective function:**
  \[ E = \sum_{j=1}^{C} \sum_{i=1}^{N} ||x_i^{(j)} - c_j||^2 \]

- **Idea:** iteratively approximate cluster centres.
C-means (2)

- Algorithm:

  1. Initialise cluster centres.
  2. Map each pixels to closest cluster.
  3. Recalculate cluster centres (centroids).
  4. Repeat 2-3 until convergence.
Fuzzy c-means

- Similar to hard c-means but allows for partial membership of pixels to clusters.

Objective function:

$$E = \sum_{j=1}^{C} \sum_{i=1}^{N} \mu_{ij}^k \|x_i^{(j)} - c_j\|^2$$

Fuzzy membership function:

$$\mu_{ij} = \frac{1}{\sum_{m=1}^{C} \left( \frac{||x_j - c_i||}{||x_j - c_m||} \right)^{2/(k-1)}}$$
Fuzzy c-means (2)

- Algorithm

1. Initialise cluster centres.

2. Compute fuzzy memberships functions
   \[ \mu_{ij} = \frac{1}{\sum_{m=1}^{C} \left( \frac{||x_j - c_i||}{||x_j - c_m||} \right)^{2/(k-1)}} \]

3. Compute cluster centres
   \[ c_i = \frac{\sum_{j=1}^{N} \mu_{ij}^k x_j}{\sum_{j=1}^{N} \mu_{ij}^k} \]

4. Repeat 2-3 until convergence.
Rough c-means

- Each cluster has two approximations
  - Lower approximation
  - Upper approximation

- Samples can fall in lower approximation or boundary area.

- Samples in lower approximation definitely belong to the cluster.

- Samples in boundary area may belong to the cluster (or another one in whose boundary area it also resides).
Rough c-means

- Algorithm

1. Randomly assign samples to lower approximations.

2. Compute cluster means as weighted average of samples in lower approximation and samples in boundary area.

3. Assign samples to approximations. If the difference between the distance to the closest mean and the distance to the other cluster exceeds the threshold, assign to upper approximation, otherwise assign to lower approximation.

4. Repeat 2-3 until convergence.
**Fuzzy rough c-means**

- Hard c-means always assigns one sample to one cluster.
- Fuzzy c-means allows partial membership to several clusters.
- Rough c-means assigns samples with insufficient information to boundary region of multiple clusters, samples in lower approximations to one cluster.
- In our approach we combine these ideas.
- As with rough c-means we work with lower approximation, upper approximation and boundary region.
- If sample is in lower approximation it definitely belongs to the cluster (membership=1)
- If sample is in boundary region are assigned fuzzy memberships.
Fuzzy rough c-means

- Algorithm

1. Initialisation.
2. Compute fuzzy memberships functions
3. Compute cluster centres
4. Assign samples to approximations.
5. Repeat 2-4 until convergence.
Experimental results

- Test dataset: 6 images commonly used in CQ literature:
Experimental results (2)

- Original image
- Popularity algorithm
- Median cut
- Octree
- Neuquant
- Fuzzy rough c-means
Experimental results (3)

- popularity algorithm
- median cut
- octree
- Neuquant
- Fuzzy rough c-means
Experimental results (4)

popularity algorithm

median cut

octree

Neuquant

SWASA – S-CIELAB
Experimental results (5)

- popularity algorithm
- median cut
- octree
- Neuquant
- Fuzzy rough c-means
Experimental results (6)

- Image quality in terms of peak signal-to-noise-ratio (PSNR).
  - Function of MSE (mean squared error) which is the objective function

<table>
<thead>
<tr>
<th></th>
<th>Lenna</th>
<th>Peppers</th>
<th>Mandrill</th>
<th>Sailboat</th>
<th>Pool</th>
<th>Airplane</th>
<th>average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Popularity algorithm</td>
<td>22.24</td>
<td>18.56</td>
<td>18.00</td>
<td>8.73</td>
<td>19.87</td>
<td>15.91</td>
<td>17.22</td>
</tr>
<tr>
<td>Median cut</td>
<td>23.79</td>
<td>24.10</td>
<td>21.52</td>
<td>22.01</td>
<td>24.57</td>
<td>24.32</td>
<td>23.39</td>
</tr>
<tr>
<td>Octree</td>
<td>27.45</td>
<td>25.80</td>
<td>24.21</td>
<td>26.04</td>
<td>29.39</td>
<td>28.77</td>
<td>26.94</td>
</tr>
<tr>
<td>Neuquant</td>
<td>27.82</td>
<td>26.04</td>
<td>24.59</td>
<td>26.81</td>
<td>27.08</td>
<td>28.24</td>
<td>26.73</td>
</tr>
</tbody>
</table>
Conclusions

• Colour quantisation is an \textit{np}-complete problem of identifying the optimal colours in an image for reduced colour reproduction.

• Colour quantisation can also be seen as a clustering problem.

• We presented a fuzzy rough \(c\)-means clustering approach for colour quantisation.
  – Pixels belong either to lower approximation of a cluster or to boundary region between clusters.
  – Fuzzy memberships are employed but only in boundary region (membership in lower approximation = 1).

• Our method was shown to outperform common colour quantisation algorithms.