A color image segmentation approach for content-based image retrieval

Authors : Mustafa Ozden, Ediz Polat

Source : Pattern Recognition 40 (2007) 1318-1325

Speaker : 鄭中川

Overview of the proposed approach

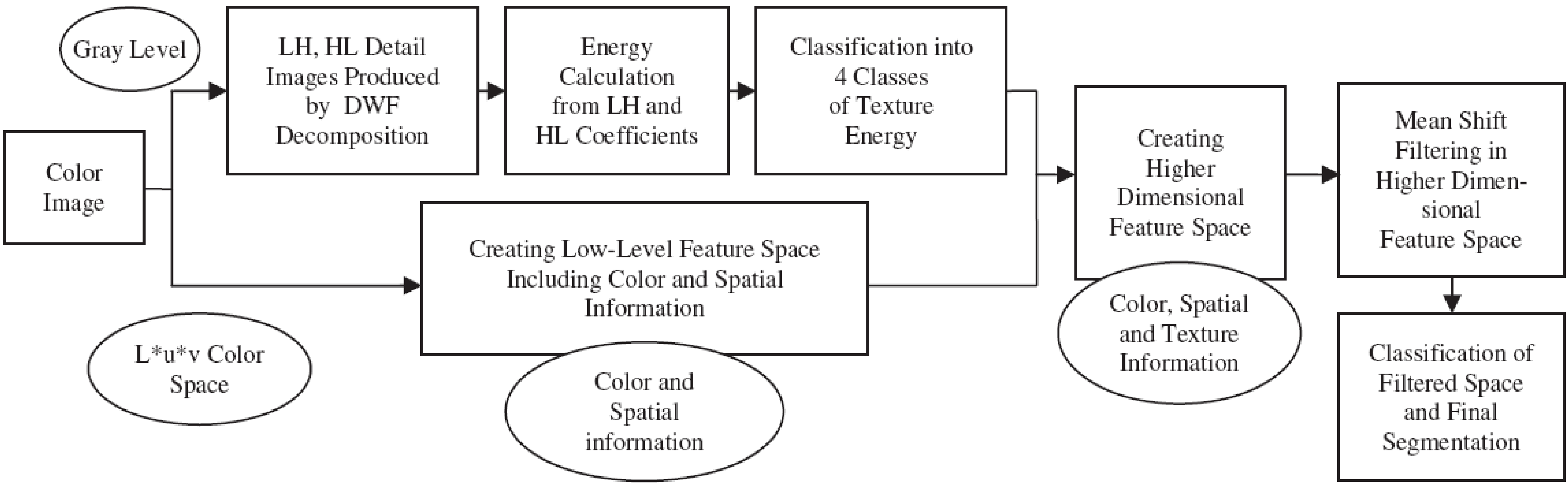

Fig. 3. Overview of the proposed approach.

Wavelet frames for texture characterization

1. DWF : discrete wavelet frames, Haar or Daubechies??
2. Energy : the square of the coefficients
3. median filter
4. LH & HL => 4 texture categories

smooth (not enough energy in any orientation)
vertical (dominant energy in vertical direction)
horizontal (dominant energy in horizontal direction)
complex (no dominant orientation)

Wavelet frames for texture characterization (cont.)

4.1 The energy of LH & HL => K-means => 2 clusters (0, 1)
4.2 A pixel is classified as smooth if its category is 0 in both LH and HL sub-bands.
4.3 vertical => LH : 0, HL : 1
4.4 horizontal = > LH : 1, HL : 0
4.5 complex = > LH : 1, HL : 1

Wavelet frames for texture characterization (cont.)

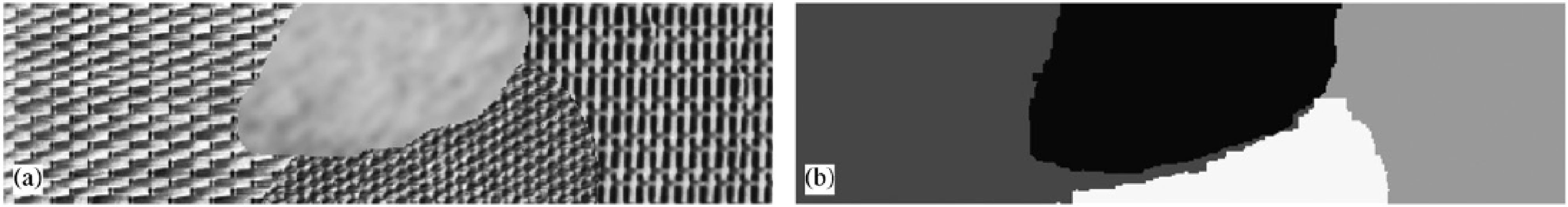

Fig. 2. Illustration of classifying different textured regions: (a) a textured image containing vertical, horizontal, smooth and complex textures created using Brodatz textures; (b) classification result using wavelet frame decomposition with median energy in a local window.

Mean-shift filtering in higher dimensional space and segmentation

Given n data points x_i, $i = 1, \ldots, n$ in the d-dimensional space R^d, the kernel density estimation at the location x can be calculated by

$$\hat{f}(x) = \frac{1}{nh_i^d} \sum_{i=1}^{n} K\left(\frac{x - x_i}{h}\right)$$

$$K(x) = c_{k,d}\, k(\|x\|^2) \qquad k(x) = \begin{cases} 1 - x & 0 \le x \le 1 \\ 0 & x \ge 1 \end{cases}$$

$$\hat{f}(x) = \frac{c_{k,d}}{nh_i^d} \sum_{i=1}^{n} k\left(\left\|\frac{x - x_i}{h}\right\|^2\right)$$

Mean-shift filtering in higher dimensional space and segmentation (cont.)

The modes are : $\nabla \hat{f}(x)=0$

$$\nabla \hat{f}(x)=\frac{2c_{k,d}}{nh_i^{d+2}}\sum_{i=1}^{n}(x_i-x)g(\|\frac{x-x_i}{h}\|^2)$$

$$=\frac{2c_{k,d}}{nh_i^{d+2}}[\sum_{i=1}^{n}g(\|\frac{x-x_i}{h}\|^2)][\frac{\sum_{i=1}^{n}x_i g(\|\frac{x-x_i}{h}\|^2)}{\sum_{i=1}^{n}g(\|\frac{x-x_i}{h}\|^2)}-x]$$

$g(x)=-k'(x)$

Mean-shift filtering in higher dimensional space and segmentation (cont.)

$$m_h(x) = \frac{\sum_{i=1}^{n} x_i g\left(\left\|\frac{x - x_i}{h}\right\|^2\right)}{\sum_{i=1}^{n} g\left(\left\|\frac{x - x_i}{h}\right\|^2\right)} - x$$

1. The initial position of the kernel can be chosen as one of the data points x_i => X^t
2. compute $m_h(X^t)$
3. $X^{t+1} = X^t + m_h(X^t)$

Mean-shift filtering in higher dimensional space and segmentation (cont.)

feature vector

1. f_r = {L*, u*, v*, x, y} (color range)
2. f_s = {T, x, y} (spatial)

$$K_{h_s, h_r}(x) = \frac{c}{(h_s)^p (h_r)^q} k_s\left(\left\|\left(\frac{x_s}{h_s}\right)\right\|^2\right) k_r\left(\left\|\left(\frac{x_r}{h_r}\right)\right\|^2\right)$$

$$\hat{f}(x) = \frac{c}{n(h_s)^p (h_r)^q} \sum_{i=1}^{n} k_s\left(\left\|\left(\frac{x_s - x_{s,i}}{h_s}\right)\right\|^2\right) k_r\left(\left\|\left(\frac{x_r - x_{r,i}}{h_r}\right)\right\|^2\right)$$

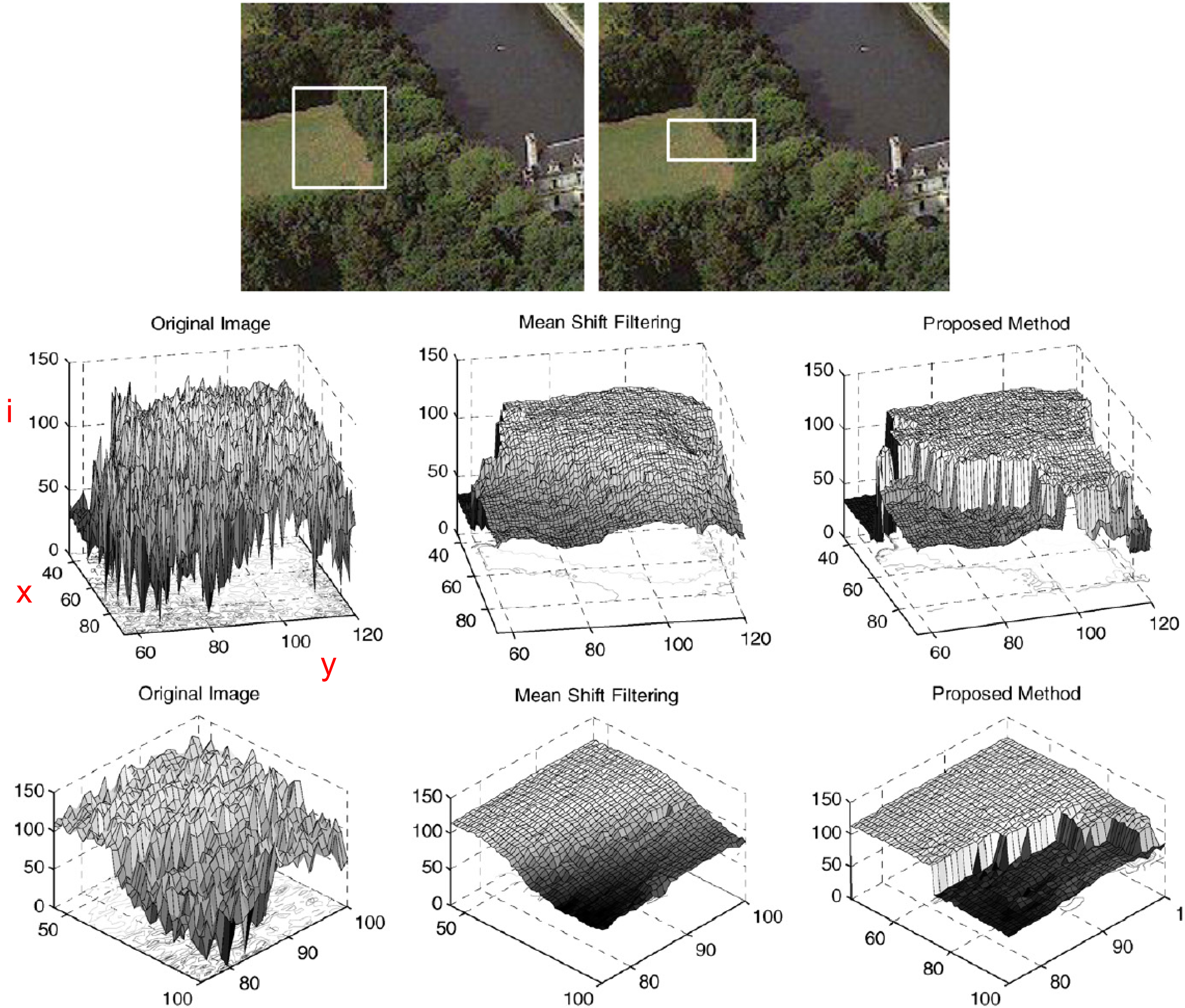

Fig. 5. Three-dimensional visualization of results. *First row*: original images with two different windows that the both standard mean shift and proposed filtering are applied. *Second and third row*: the proposed method provides well-separated regions which result in better segmentation.

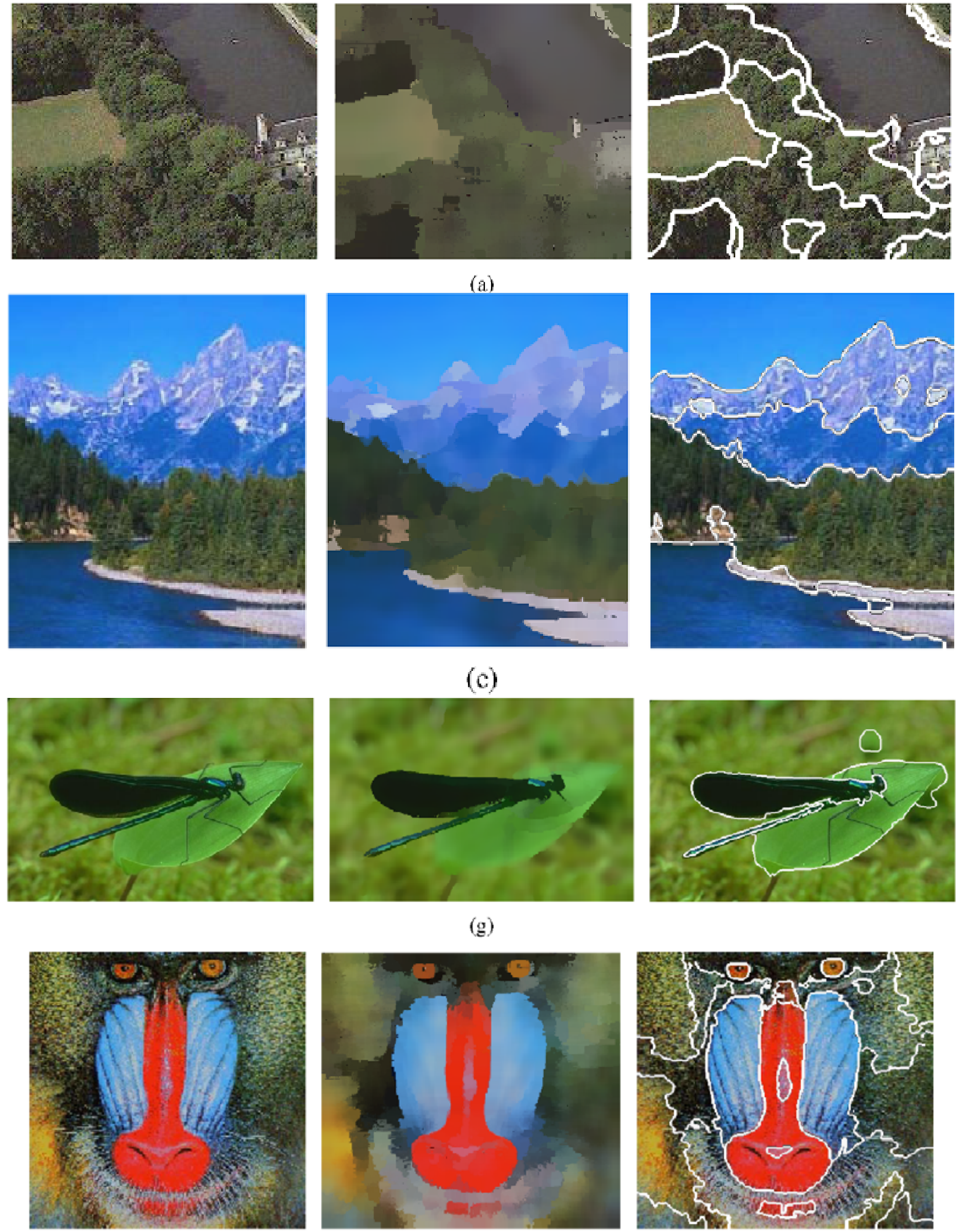

Fig. 6

Experiment results

Table 1
Comparison of number of misclassified pixels using mean shift and proposed algorithms

Image	Number of misclassified pixels		Error (%)	
	Mean shift	Proposed algorithm	Mean shift	Proposed algorithm
Fig. 2(a)	3988	2891	13.00	9.60
Fig. 6(a)	542	340	9.40	5.90
Fig. 6(b)	250	123	9.16	4.50
Fig. 6(c)	360	270	17.02	12.70
Fig. 6(d)	55	50	0.85	0.77
Fig. 6(e)	5548	1360	33.40	8.20
Fig. 6(g)	155	155	3.67	3.67
Fig. 6(h)	592	30	7.50	0.68

$$\text{Error} = \frac{\text{Number of misclassified pixels}}{\text{Total number of pixels in region}}.$$