Flower Image Database Construction and its Retrieval

S. Nakamura   M. Sawada   Y. Aoki   P. Hartono   S. Hashimoto
{shingo, mie, aoki, hartono, shuji}@shalab.phys.waseda.ac.jp
Dept. of Applied Physics, School of Science and Engineering
Waseda University, Tokyo, Japan

Presented at
FCV2001: 7th Korea-Japan Joint Workshop on Computer Vision
Chung-Ang University in Seoul
February 5-6, 2001

Motivation
• Text-based retrieval has been already widely practiced, in example in WWW
• The demand for image retrieval is increasing
• The developments of content-based image data base and its retrieval methods are important

We focus on flower images as retrieval images
- Retrieval of flower image by textual key-words is very difficult
- Flower images taken in natural environment have complex background

Objectives
• Developing Content-based flower-image database
• Dealing with images taken in natural environments
• Developing a novel method of image retrieval from the database

Organization of Presentation
• Database Generation
• Retrieval System
• Experimental Results
• Conclusion and Future Works

Classification of HSV Space
• Classify whole HSV space into the 9 color classes:
  White, Red, Yellow, Light Blue, Blue, Purple: Color of flower leaf
  Dark, Green, Brown: Shadow, Green Leaves, Trees, etc.

Flower database construction
• Including flower color and shape information as flower image features
• Being used for retrieval by image contents

Index Database
• Including several text items
• Being used for retrieval at interactive Q and A style and explanation about flower for user

Image Database

Classification of HSV space into nice class
Flower Area Extraction

- **Step 1**: Exclusion of Background Color
  - Green: Leaves and Stalks
  - Brown: ground and trunks of trees
  - Dark: shadow

- **Step 2**: Find out dominant flower color
  - Find out the color to which the center of gravity of pixels is the nearest to the center of image

- **Step 3**: Flower’s area
  - If the number of pixels in the color space is enough for image size, we certainly regard the color as flower color

- **Step 4**: Noise Reduction
  - Expansion and contraction by 4 pixels

Features included in Image Database

- **1. Color Histogram**
  - Value-histogram of R, G and B of pixels that belong to Flower Area extracted by above technique
  - Total of histogram is equal to 1 in each RGB histogram for normalization

- **2. Shape Diagram**
  - Value is the distance from the center of gravity of the Flower Area for each angle
  - Angle as abscissa is quantized to 256 steps and total of diagram is equal to 1

Properties included in index database

- **Number of properties**: 15
  1. Flower Name
  2. Alias
  3. Family name
  4. Generic name
  5. Botanical name
  6. Blooming season
  7. Phase
  8. Distribution
  9. Diameter of flower
  10. Count of flower leaf
  11. Height
  12. Kind of leaf shape
  13. State of leaf
  14. Main color
  15. Sub Color

Retrieval System

- **Input key image**
  - If it is not good enough to process, user can select rectangle region to process.

- **Retrieval by image database**
  - Retrieve several candidates images by image features matching

- **Retrieval by index database**
  - Inquire with dialog window about the most meaningful property to user
  - Using user’s answers for reduction candidates retrieved by image database
Evaluation of similarity between 2 histograms

- Similarity of between two $i$-th image in the database and the key-image is defined as

$$S_i = \min \left( \sum \sum (f_i(x) - g_C(x))^2 \right)$$

- $f_i(x)$: density of $x=0$ to $255$ in color $C=\{R,G,B\}$ of $i$-th image
- $g_C(x)$: density of $x$ in color area $C$ of key-image
- $\delta$: toleration margin to accommodate environmental influence
  (If it moves whole range in the case of shape histogram)

Interactive Q & A step

- Step1 Four properties are classified into 6 classes

<table>
<thead>
<tr>
<th>Class</th>
<th>Description</th>
<th>Index No.</th>
<th>Height (cm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Making</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>2</td>
<td>Making</td>
<td>2</td>
<td>10</td>
</tr>
<tr>
<td>3</td>
<td>Making</td>
<td>3</td>
<td>15</td>
</tr>
<tr>
<td>4</td>
<td>Making</td>
<td>4</td>
<td>20</td>
</tr>
<tr>
<td>5</td>
<td>Making</td>
<td>5</td>
<td>25</td>
</tr>
<tr>
<td>6</td>
<td>Making</td>
<td>6</td>
<td>about 30</td>
</tr>
</tbody>
</table>

- Step2 Count number of flowers belonging to each class from images retrieved by image database

- Step3 With the dialog window, system asks user the property whose variance is the largest

   - Number of flower leaves
   - Height
   - Diameter
   - Blooming season

How many flower leaves does the searched flower have?

Interactive Q & A step

- Step1 Four properties are classified into 6 classes

<table>
<thead>
<tr>
<th>Class</th>
<th>Description</th>
<th>Index No.</th>
<th>Height (cm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Making</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>2</td>
<td>Making</td>
<td>2</td>
<td>10</td>
</tr>
<tr>
<td>3</td>
<td>Making</td>
<td>3</td>
<td>15</td>
</tr>
<tr>
<td>4</td>
<td>Making</td>
<td>4</td>
<td>20</td>
</tr>
<tr>
<td>5</td>
<td>Making</td>
<td>5</td>
<td>25</td>
</tr>
<tr>
<td>6</td>
<td>Making</td>
<td>6</td>
<td>about 30</td>
</tr>
</tbody>
</table>

Experiments

- Image database has constructed by 1,229 flower images
- Index database has constructed by 230 kinds of flower
- Top 30 images by retrieval of image content are presented to user

Result of retrieval by image content

(a) The flower color of key-image is mainly purple
(b) The flower color of key-image is mainly yellow

Fig.5 The top 10 results of retrieval by color feature histograms

Result of retrieval by image content

(c) The flower color of key-image is mainly constructed plural
(d) The candidates are not good

Fig.6 The not-well results of retrieval by image database
Reduction of candidates by Q & A

Key-image

Candidates by color feature retrieval

Reduction

Conclusion

• Pull up flower images whose color is similar by inputting key flower image
• Easily reduction images by asking the most meaningful property of flower

• Future work
  – Add more flower images to flower database
  – Evaluation speed and usability of retrieval system
  – Improvement for flowers whose color is complex
  – Applications in another image database