Enhancing Sketch-Based Image Retrieval by Re-Ranking and Relevance Feedback

Heechan Shin

CS688
Student paper presentation

“Enhancing Sketch-Based Image Retrieval by Re-Ranking and Relevance Feedback” (IEEE TIP 16)
Contents

• Problems & Related work

• Solution
  • Image Grouping
  • Visual Feature Verification
  • Contour-Based Relevance Feedback

• Experimental Result

• Conclusion
Problems

• Sketch Based Image Retrieval (SBIR)

What a user want to find

What a user queries

How to measure the relevance of an image and a query sketch?
Problems

• To solve the problem...
  • Contour matching
  • Local feature matching

Angular Radial Partitioning (ARP)


Related work

• Angular Radial Partitioning (ARP)
Related work

- Edgel index (Edgel: edge pixel)
Problems

• Sketch should be fairly close to the image.

• Irrelevant image may be retrieved.

Re-ranking and finding relevant images are important!
Solution

• Contribution
  • **Optimizing module** with the search result of any SBIR framework
Solution

- **Image Grouping**
  - Fining more relevant images

- **RVFV**
  - Removing irrelevant images

- **CBRF**
  - Making new queries to find relevant images using contours
Solution

- **Image Grouping**
  - Fining more relevant images

- **RVFV**
  - Removing irrelevant images

- **CBRF**
  - Making new queries to find relevant images using contours
• Relevant Images Grouping for Relevant Feedback

- Initial result (size N)
- Select images (size R) (R < N)
- Find near-duplicated images using existing image matching approach (ex, binary edge-SIFT)
- Cluster near-duplicated images (size of cluster K) (K <= R)

Relevant group
Rank high
Solution

- Re-ranking via Visual Feature Verification (RVFV)

Ranked images (size N)

Top ranked image (Standard image, $I_s$)

Calculate similarity score to $I_s$

Similarity score $S_k = SM (I_s, I_k)$, $k = 1 \sim N$
(ex. $S_1 = SM (I_s, I_1) = 1.0$)

Select top M images

Re-ranked images according to $S_k$ (size N)
Solution

• Similarity score $S_k$
  • $d_A$: SIFT descriptor of image A

• L2 norm of two descriptor $|d_A - d_B|^2 = 2 - \sum d_A^l d_B^l$
  since $|d_A|^2 + |d_B|^2 = 2$, $SIM (d_A, d_B) = \sum d_A^l d_B^l$

• $SIM (I_A, I_B) = \sum_m SIM (d_A, d_B) W(m)$
  
  $W$ here,
  $m$ $m^{th}$ SIFT pair between $I_A$ and $I_B$
  $W(m)$ is weight
Solution

• Contour-Based Relevance Feedback

Re-ranked images (size M) → Create contour from image (size M) → Relevant Feedback Score

\[ S_{RF}(k) : \sum_{m=1}^{M} S_D(m, k) \times S_N(m) ; \quad k = 1, \ldots, T \]

Where T is size of entire image data
Solution

• Contour-Based Relevance Feedback
  • Relevant Feedback Score
    \[ S_{RF}(k) : \sum_{m=1}^{M} S_{ID} (m,k) \times S_N (m) ; \quad k = 1, \ldots, T \]
    
    \[ W \text{ here } T \text{ is size of entire image data} \]
  • \( S_N (m) \) : Initial score of image \( m \)
  • \( S_D (m,k) \) : Score after first RVFV of image \( k \), when a query is contour of image \( m \)

  • Final score \( S(k) = (1 - w) \times S_N (k) + w \times S_{RF}(k) ; \quad k = 1, \ldots, T \)

  • With \( S(k) \), we have new ranked list
Solution

• **Image Grouping**
  - Fining more relevant images

• **RVFV**
  - Removing irrelevant images

• **CBRF**
  - Making new queries to find relevant images using contours
Solution
Experimental Result

• Experimental setting
  • Dataset
    • SBIR_100K Dataset: 1,240 images for 31 sketches and 100,000 noise images
    • Authors' own Dataset: from Google keyword search 296,562 images with 68,647 sketch-describable images + 523 sketches
Experimental Result

- Result 1. Performance Evaluation

Result of authors’ dataset

Result of SBIR_100K dataset
### Experimental Result

- **Result 2. Computational cost**

<table>
<thead>
<tr>
<th></th>
<th>Initial SBIR</th>
<th></th>
<th>ours</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Clustering</td>
<td>RVFV1</td>
<td>CBRF</td>
<td>RVFV2</td>
<td>Total</td>
<td></td>
</tr>
<tr>
<td>Edgel</td>
<td>9.77</td>
<td>0.017</td>
<td>0.73</td>
<td>0.14</td>
<td>0.41</td>
<td>11.06</td>
<td>+1.28s</td>
</tr>
<tr>
<td>ARP</td>
<td>0.64</td>
<td>0.015</td>
<td>0.53</td>
<td>0.10</td>
<td>0.26</td>
<td>1.55</td>
<td>+0.91s</td>
</tr>
</tbody>
</table>
Conclusion

- Image Grouping
  - Find which images are more relevant
- Re-ranking via Visual Feature Verification (RVFV)
  - Filter out irrelevant images
- Contour-Based Relevance Feedback (CBRF)
  - Explore deeply to retrieve what does not be found with original SBIR
- Improved result with low time cost