**ABSTRACT**

We propose a method to discover family photos from group photos using discriminative subgraphs. We represent an image as a graph with social contexts such as age, gender, and face position. We consider frequent subgraphs from all group photos as features for classification. Numerous subgraphs, however, result in the high dimensions, some of which are not discriminative. To address this issue, we adopt a state-of-the-art frequent subgraph mining technique to remove non-discriminative subgraphs. Our method shows approximately 4%~6% higher classification accuracy in lower feature dimensions compared to the previous method.

**APPROACH**

**Overview**

Group photos with Annotation → Face Graph construction → Discriminative Frequent Subgraphs → Train set → Subgraph Frequency → Non-Family Evaluation

**Discriminative Frequent Subgraph Mining**

- **DFS Lexicographic Order**
  - canonical comparison of subgraphs
  - minimal DFS code for building & comparing
  - maximal Discriminative quality

**Discriminative Quality**

Class: A, B,

One Feature (a frequent subgraph mined): X,

AX₁ and BX₁: containing the feature,

AX₀ and BX₀: not containing the feature

$q(X) = (AX₁ - BX₁ + AX₀ - BX₀)$

Subgraph set U. The $q(·)$ of any supergraph $T$ of $X$ ($T ⊆ S$) cannot exceed the bound $\text{Maximal } D(·)$:

$$\text{Maximal } D(·) = \sum \text{max}(AX₁ - BX₁ + AX₀ - BX₀)$$

**Algorithm**

Input: Graph set $G$, optional threshold $\alpha$

Output: Set of discriminative frequent subgraphs $S$  

1: $S = \emptyset$
2: $S = \text{best subgraph for } q(S) \bigcup \{G \}$ // gSpan call
3: if $q(S) > q(S')$, then
4: $S' = S \bigcup \{G \}$ // $S$ is an improvement
5: goto 2
6: return $S$

**EVALUATION**

- Family vs. non-Family classification using SVM-linear
- More than 1,073 photos than previous one
- Comprehensive family types

<table>
<thead>
<tr>
<th>Feature Dimension</th>
<th>Duplicate Subgraphs (not)</th>
<th>gSpan</th>
<th>Ours</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>10 (5)</td>
<td>71.82%</td>
<td>71.80%</td>
</tr>
<tr>
<td>74</td>
<td>24 (50)</td>
<td>74.11%</td>
<td>80.67%</td>
</tr>
<tr>
<td>106</td>
<td>26 (80)</td>
<td>76.73%</td>
<td>81.44%</td>
</tr>
<tr>
<td>212</td>
<td></td>
<td>77.76%</td>
<td></td>
</tr>
<tr>
<td>1175</td>
<td></td>
<td>77.50%</td>
<td></td>
</tr>
<tr>
<td>1961</td>
<td></td>
<td>77.28%</td>
<td></td>
</tr>
<tr>
<td>5104</td>
<td></td>
<td>76.28%</td>
<td></td>
</tr>
</tbody>
</table>

**Discovering Family Photo using Discriminative Frequent Subgraph**

Changmin Choi†, YoonSeok Lee‡, Sung-Eui Yoon‡

Scalable Graphics/Geometric Algorithm Lab. (sglab.kaist.ac.kr)

†Division of Web Science and Technology, KAIST, ‡Department of Computer Science, KAIST