Introduction

- What is CBIR?
- Setbacks in CBIR
- Suggested Solutions
  - What is relevance feedback?
What is CBIR?

- Using low-level features to retrieve images from a database.
- Content-based image retrieval is recommended for a more accurate search.
Setbacks in CBIR

- Time
- Space
- Semantic vs. Content
Suggested Solutions

- Use both high- and low-level features to increase accuracy
- Relevance feedback
What is Relevance Feedback?

- User feedback
- Updated based on each particular user's preference
Related Work

- Google
- Conference paper approach
Related Work

- Google’s new image search
Train data

- Set up semantic network
- Find direct relationships
- Find indirect relationships

Test the data

- Use low-level features
- Update the system after so many queries
1. **Train**

A. Get the access matrix
   
i. How frequently images are accessed, and which images access other images

B. Find all direct relationships from access matrix

C. Find all indirect relationships from direct relationship matrix

D. Store the highest relationship between direct and indirect in closest relationship matrix
2. Test

A. Use a set of images in database as queries.

B. For every query, compute its ‘distance’ from other images in the set based on low-level features using equations (4) and (5) in paper.

   i.  Equation (4): \( W_1(i) = (1 - \frac{|b_i(o_1) - b_q(o_1)|}{b_q(o_1)}) \)

   ii. Equation (5): \( W_{t+1}(i) = W_t(i) \ a_{q,i} \ (1 - \frac{|b_i(o_{t+1}) - b_q(o_{t+1})|}{b_q(o_{t+1})}) \)

C. Store similarity score in edge weight matrix

D. Every so many queries, update the access matrix, direct relationship matrix, indirect relationship matrix, and highest relationship matrix.
Experimental Results

- Image Database
- The Numbers
Experimental Results - Image Database

- Flickr database
- Image features extracted previous to project
- Features include:
  - Color
  - Edge
  - Texture
Experimental Results – The Numbers

- **Accuracy**
  - ~11%
    - Only 11% of the time, images that had similar low-level features were in the same category.

- **Speed**
  - Avg. ~6.5 minutes
    - 10% of images used for training
    - 1800 images used in testing
Conclusions

- My implementation does not provide sufficient accuracy.
- Speed was improved
- Space is still an issue
- If time had permitted, I believe that the accuracy would have been improved substantially.
  - When the matrices are updated more often, there is more current, more accurate information.
References
