Unsupervised Manifold Learning By Correlation Graph and Strongly Connected Components for Image Retrieval

IEEE International Conference on Image Processing (ICIP) 2014

Daniel Carlos Guimarães Pedronette and Ricardo da S. Torres

daniel@rc.unesp.br, rtorres@ic.unicamp.br
State University of São Paulo, University of Campinas
São Paulo, Brazil

October 28, 2014
Outline

1. Introduction
   - Content-Based Image Retrieval
   - Unsupervised Methods for Image Retrieval

2. Manifold Learning By Correlation Graph
   - Correlation Graph: Motivation
   - Correlation Graph
   - Strongly Connected Components
   - Correlation Graph Distance

3. Experimental Evaluation
   - Datasets and Descriptors
   - Impact of Parameters
   - Shape, Color, and Texture Descriptors
   - Object Retrieval
   - Comparison to Other Approaches

4. Conclusions
Content-Based Image Retrieval:

**Input:**
- Image collection
- Query image

**Objective:**
- To retrieve similar images according to visual properties
Content-Based Image Retrieval

**Motivation:**
- Huge growth of image collections:
  - People moved from consumers to producers!
- Image retrieval based on keywords ignores the visual content

_Photos = 500MM+ Uploaded & Shared Per Day, Growth Accelerating, on Trend to Rise 2x Y/Y..._

![Daily Number of Photos Uploaded & Shared on Select Platforms, 2005-2013YTD](chart.png)

Source: KPCB estimates based on publicly disclosed company data.
Limitations of CBIR Systems:

- "Semantic Gap": Gap between low-level features and high-level concepts
Recently, *Unsupervised Post-Processing* [25, 8, 26] approaches have been proposed:

- Aiming at improving effectiveness of image retrieval tasks.
- By reducing the Semantic Gap.

Unsupervised approaches use more *global affinity measures* instead of pairwise distance computations.

Exploiting the *global dataset structure* becomes a central problem in computer vision applications.
Contribution:
A novel Unsupervised Manifold Learning Algorithm based on the Correlation Graph and Strongly Connected Components (SCCs).

- The proposed algorithm computes a new distance which takes into account the **intrinsic geometry of the dataset manifold**.
Main ideas:

1. Constructing a graph representation of the dataset by exploiting the distance correlation between kNN constrained by a correlation threshold.
2. Strongly Connected Components (SCCs) of the graph are analyzed with the aim of discovering the *intrinsic geometry of the dataset manifold*;
3. A similarity score combines information from the Correlation Graph Adjacency and Strongly Connected Components;
4. A new *Correlation Graph Distance* is computed based on the similarity score.
Correlation Graph Motivation

Discussion:

- The edges defined by the Correlation Graph provide a very strong indication of similarity among images (specially for high correlation thresholds).

- However, although very precise, the edges include a very small neighborhood.

- We aim at expanding the similarity neighborhood, but still considering the geometry of the dataset manifold, by using SCCs.
Image Retrieval Model:

- Let $\mathcal{C} = \{img_1, img_2, \ldots, img_n\}$ be an image collection, where $n$ is the size of the collection.
- Let $\rho(i, j)$ denotes the distance between two images $img_i$ and $img_j$, according to a given image descriptor.
- Let $\tau_q = (img_1, img_2, \ldots, img_{n_s})$ be a ranked list, which can be defined as a permutation of the subset $\mathcal{C}_s \subset \mathcal{C}$.
  - The subset $\mathcal{C}_s$ contains the $n_s$ most similar images to query image $img_q$, such that and $|\mathcal{C}_s| = n_s$. 
Image Retrieval Model

- Query Image
- Distances
- Ranked Lists
Correlation Graph

Graph Definition:

- Given a directed graph $G = (V, E)$, the set of vertices $V$ is defined by the image collection $C$, such that each image is represented by a node and $V = C$.

- The edge set $E$ is defined considering the distances correlation among images at the top $n_s$ positions of each ranked list:
  - $E = \{(img_q, img_j) \mid \tau_q(j) \leq n_s \land cor(q,j) \geq t_c\}$,
  - $cor(q,j)$ is the correlation score between $img_q$ and $img_j$,
  - $t_c$ is the correlation threshold considered.
Correlation Graph: Motivation

Correlation Graph

Let $\mathcal{N}_k(q)$ be the set containing the $k$-nearest neighbors to given image $img_q$ and $\mathcal{N}_k(q, j) = \mathcal{N}_k(q) \cup \mathcal{N}_k(j)$.

Vectors $X$ and $Y$ contain the distances from images $img_q$, $img_j$ to $img_i \in \mathcal{N}_k(q, j)$:

- $X_i = \rho(q, i)$ and $Y_i = \rho(j, i)$

The correlation score $\text{cor}(q, j)$ is computed by the Pearson’s Correlation Coefficient, considering the distances to the $kNN$.

\[
\text{cor}(q, j) = \frac{\sum_{i=1}^{k_u}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{k_u}(X_i - \bar{X})^2} \sqrt{\sum_{i=1}^{k_u}(Y_i - \bar{Y})^2}}. \tag{1}
\]
Correlation Graph

\[ \tau_q(j) \leq n_s \]

**Graph Edge:** \( E(q, j) \)

**Union:** \( \bigcup \)

**Correlation Graph**

**Positions:**

\[
\begin{array}{c}
\tau_q \\
\tau_j
\end{array}
\]

\[
\begin{array}{c}
img_q \\
img_i \\
img_l \\
img_j \\
img_r
\end{array}
\]

\[
\begin{array}{c}
img_j \\
img_s \\
img_t \\
img_q \\
img_r
\end{array}
\]

**Pearson Correlation Coefficient:**

\[ \text{cor}(q, j) \geq t_c \]

**Distances to**

- \( \text{img}_q \)
- \( \text{img}_j \)

**Experimental Evaluation**

**Conclusions**

**Manifold Learning By Correlation Graph**

**Introduction**

**Strongly Connected Components**

**Correlation Graph:** Motivation

Daniel Carlos Guimarães Pedronette and Ricardo da S. Torres

Unsupervised Manifold Learning By Correlation Graph and Strongly Connected Components
Correlation Graph

Strongly Connected Components (SCCs)
The Strongly Connected Components of a directed graph are defined by subgraphs that are themselves strongly connected.
- Every vertex is reachable from every other vertex.

SCCs Computation
- Each SCC is defined as a set of images $S_i$, computed using Tarjan’s [22] Algorithm.
- The overall output of the algorithm is a set of SCCs $S = \{S_1, S_2, \ldots, S_m\}$
Correlation Graph

- Strongly Connected Components (SCCs):
  - Sets of similar images

![Diagram of Correlation Graph showing strongly connected components](image)
Correlation Graph Distance - Algorithm

Require: Correlation Graph $G = (V, E)$, Set of SCCs $S$
Ensure: Correlation Graph Similarity Score $W_{i,j}$

1: $t_c \leftarrow t_{start}$
2: while $t_c \leq 1$ do
3:   { Correlation Graph Adjacency }
4:   for all $img_q \in V$ do
5:       for all $img_i, img_j \in E(q)$ do
6:           $W_{i,j} \leftarrow W_{i,j} + t_c$
7:       end for
8:   end for
9:   { Strongly Connected Components }
10:  for all $S_c \in S$ do
11:      for all $img_i, img_j \in S_c$ do
12:          $W_{i,j} \leftarrow W_{i,j} + t_c$
13:      end for
14:  end for
15:  $t_c \leftarrow t_c + t_{inc}$
16: end while
The similarity score $W_{i,j}$ uses information from both Correlation Graph Adjacency and Strongly Connected Components (SCCs).

Based on the similarity score $W_{i,j}$, the Correlation Graph Distance $\rho_c(i,j)$ is computed:

$$\rho_c(i,j) = \frac{1}{1 + W_{i,j}}. \quad (2)$$
Example: Euclidean Distance

Two moon data set: Euclidean Distance.

Red neighbors
Blue neighbors
Example: Intermediary Correlation Graph Structures

Two moons data set: Correlation Graph at an intermediary threshold.

- Red adjacency
- Red SCC
- Blue adjacency
- Blue SCC
- Other nodes
Example: Correlation Graph Distance

Two moon data set: Correlation Graph Distance.
## Experimental Evaluation

- Evaluation of impact of parameters
- 4 different datasets
- 13 CBIR descriptors
  - shape, color, and texture
- Statistical tests (t-tests)
- Comparison with state-of-the-art approaches

## Results
- Effectiveness gains up to +34.54%.
Impact of Parameter on Effectiveness

Impact of Parameters on Mean Average Precision (MAP) for ASC descriptor
Experimental Evaluation - Shape

<table>
<thead>
<tr>
<th>Descriptor</th>
<th>Dataset</th>
<th>Score (MAP)</th>
<th>Correlation Graph Distance</th>
<th>Gain</th>
<th>Statistical Significance 99%</th>
</tr>
</thead>
<tbody>
<tr>
<td>SS [4]</td>
<td>MPEG-7 [10]</td>
<td>37.67%</td>
<td>50.68%</td>
<td>+34.54%</td>
<td>•</td>
</tr>
<tr>
<td>BAS [1]</td>
<td>MPEG-7 [10]</td>
<td>71.52%</td>
<td>81.97%</td>
<td>+14.61%</td>
<td>•</td>
</tr>
<tr>
<td>IDSC [12]</td>
<td>MPEG-7 [10]</td>
<td>81.70%</td>
<td>89.39%</td>
<td>+9.41%</td>
<td>•</td>
</tr>
<tr>
<td>CFD [16]</td>
<td>MPEG-7 [10]</td>
<td>80.71%</td>
<td>91.93%</td>
<td>+13.90%</td>
<td>•</td>
</tr>
<tr>
<td>ASC [13]</td>
<td>MPEG-7 [10]</td>
<td>85.28%</td>
<td>92.53%</td>
<td>+7.25%</td>
<td>•</td>
</tr>
<tr>
<td>AIR [5]</td>
<td>MPEG-7 [10]</td>
<td>89.39%</td>
<td>97.98%</td>
<td>+9.61%</td>
<td>•</td>
</tr>
</tbody>
</table>

Positive gains ranging from +7.25% to +34.54%, considering MAP scores.
Experimental Evaluation - Shape

Shape Descriptors

Positive gains ranging from +6.90% to +29.28%, considering Bull’s Eye Score (Recall@40).

<table>
<thead>
<tr>
<th>Shape Descriptor</th>
<th>Bull’s Eye Score</th>
<th>Correlation Graph Distance</th>
<th>Gain</th>
</tr>
</thead>
<tbody>
<tr>
<td>SS [4]</td>
<td>43.99%</td>
<td>56.88%</td>
<td>+29.28%</td>
</tr>
<tr>
<td>BAS [1]</td>
<td>75.20%</td>
<td>86.52%</td>
<td>+15.05%</td>
</tr>
<tr>
<td>IDSC [12]</td>
<td>85.40%</td>
<td>92.20%</td>
<td>+7.80%</td>
</tr>
<tr>
<td>CFD [16]</td>
<td>84.43%</td>
<td>94.27%</td>
<td>+11.65%</td>
</tr>
<tr>
<td>ASC [13]</td>
<td>88.39%</td>
<td>95.22%</td>
<td>+7.73%</td>
</tr>
<tr>
<td>AIR [5]</td>
<td>93.67%</td>
<td>100%</td>
<td>+6.90%</td>
</tr>
</tbody>
</table>
Experimental Evaluation - Color

Color Descriptors

Positive gains ranging from +7.29% to +21.51%, considering MAP scores.

<table>
<thead>
<tr>
<th>Descriptor</th>
<th>Dataset</th>
<th>Score (MAP)</th>
<th>Correlation Graph Distance</th>
<th>Gain</th>
<th>Statistical Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>GCH [20]</td>
<td>Soccer [23]</td>
<td>32.24%</td>
<td>34.59%</td>
<td>+7.29%</td>
<td>●</td>
</tr>
<tr>
<td>ACC [6]</td>
<td>Soccer [23]</td>
<td>37.23%</td>
<td>45.24%</td>
<td>+21.51%</td>
<td>●</td>
</tr>
<tr>
<td>BIC [19]</td>
<td>Soccer [23]</td>
<td>39.26%</td>
<td>47.37%</td>
<td>+20.65%</td>
<td>●</td>
</tr>
</tbody>
</table>
Experimental Evaluation - Texture

Positive gains ranging from $+6.28\%$ to $+12.44\%$, considering MAP scores.

<table>
<thead>
<tr>
<th>Descriptor</th>
<th>Dataset</th>
<th>Score (MAP)</th>
<th>Correlation Graph Distance</th>
<th>Gain</th>
<th>Statistical Significance 99%</th>
</tr>
</thead>
<tbody>
<tr>
<td>CCOM [9]</td>
<td>Brodatz [3]</td>
<td>57.57%</td>
<td>64.73%</td>
<td>$+12.44%$</td>
<td>●</td>
</tr>
<tr>
<td>LAS [21]</td>
<td>Brodatz [3]</td>
<td>75.15%</td>
<td>79.87%</td>
<td>$+6.28%$</td>
<td>●</td>
</tr>
</tbody>
</table>
## Experimental Evaluation - Object Retrieval

### Object Retrieval - Color Descriptors

Positive gains ranging from +4.39% to +18.10%, considering MAP scores.

<table>
<thead>
<tr>
<th>Descriptor</th>
<th>Dataset</th>
<th>Score (MAP)</th>
<th>Correlation Graph Distance</th>
<th>Gain</th>
</tr>
</thead>
<tbody>
<tr>
<td>BIC [19]</td>
<td>ETH-80  [11]</td>
<td>49.72%</td>
<td>54.20%</td>
<td>+9.01%</td>
</tr>
</tbody>
</table>
### Comparison to State-of-the-Art

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Descriptor(s)</th>
<th>Bull’s Eye Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mutual kNN Graph [8]</td>
<td>IDSC [12]</td>
<td>93.40%</td>
</tr>
<tr>
<td>RL-Sim [18]</td>
<td>ASC [13]</td>
<td>94.69%</td>
</tr>
<tr>
<td><strong>Correlation Graph Distance</strong></td>
<td><strong>ASC [13]</strong></td>
<td><strong>95.22%</strong></td>
</tr>
<tr>
<td>Tensor Product Graph [26]</td>
<td>ASC [13]</td>
<td>96.47%</td>
</tr>
<tr>
<td><strong>Correlation Graph Distance</strong></td>
<td><strong>AIR [5]</strong></td>
<td><strong>100%</strong></td>
</tr>
</tbody>
</table>
Conclusions

Contributions:

- A novel manifold learning approach is presented using the distance correlation for representing the dataset.
- The use of Strongly Connected Components (SCCs) for discovering the intrinsic geometry of the dataset manifold.
- Experimental results demonstrated the high effectiveness of the proposed method in several image retrieval tasks.
Future Work

- Investigation of distance fusion approaches for descriptors combination.
- Investigation of rank correlation measures for construction the Correlation Graph.
The authors are grateful to:

- São Paulo Research Foundation - FAPESP (grants 2013/08645-0 and 2013/50169-1)
- CNPq (grants 306580/2012-8 and 484254/2012-0)
- CAPES
- AMD
- Microsoft Research.
Thank you for your attention!

Questions?
References I

BAS: a perceptual shape descriptor based on the beam angle statistics.

Shape matching and object recognition using shape contexts.

*Textures: A Photographic Album for Artists and Designers.*
Dover, 1966.

Contour Salience Descriptors for Effective Image Retrieval and Analysis.

Articulation-invariant representation of non-planar shapes.

Image indexing using color correlograms.

Unsupervised metric learning by self-smoothing operator.
References II


Multiresolution gray-scale and rotation invariant texture classification with local binary patterns.

Shape retrieval using contour features and distance optimization.

Exploiting pairwise recommendation and clustering strategies for image re-ranking.

Image re-ranking and rank aggregation based on similarity of ranked lists.

A compact and efficient image retrieval approach based on border/interior pixel classification.

Color indexing.

Texture recognition and image retrieval using gradient indexing.
References IV

Depth first search and linear graph algorithms.

Coloring local feature extraction.
In *ECCV*.

Learning context-sensitive similarity by shortest path propagation.

Locally constrained diffusion process on locally densified distance spaces with applications to shape retrieval.

Affinity learning with diffusion on tensor product graph.