An Application of Graph Commute Times to Image Indexing

IGARSS 2008

[1,2]Régis Behmo, [1]Nikos Paragios, [2]Véronique Prinet
Presented by Mihai Datcu

[1]Ecole Centrale Paris
[2]Institute of Automation, Chinese Academy of Sciences

http://www.mas.ecp.fr/vision/Personnel/behmo/
regis.behmo@ecp.fr

July 9, 2008

Motivation

Objectives

Provide a novel type of image representation (i.e: global descriptor) that takes into account:

- The appearance of local regions of interest
- The spatial layout of these regions: regions that look like A are they usually close to regions that look like B?

Not a classification problem!

Motivation

Objectives

Provide a novel type of image representation (i.e: global descriptor) that takes into account:

- The appearance of local regions of interest
- The spatial layout of these regions: regions that look like A are they usually close to regions that look like B?

Not a classification problem!

Motivation

Objectives

Provide a novel type of image representation (i.e: global descriptor) that takes into account:

- The appearance of local regions of interest
- The spatial layout of these regions: regions that look like A are they usually close to regions that look like B?

Not a classification problem!

Motivation

Objectives

Provide a novel type of image representation (i.e: global descriptor) that takes into account:

- The appearance of local regions of interest
- The spatial layout of these regions: regions that look like A are they usually close to regions that look like B?

Not a classification problem!

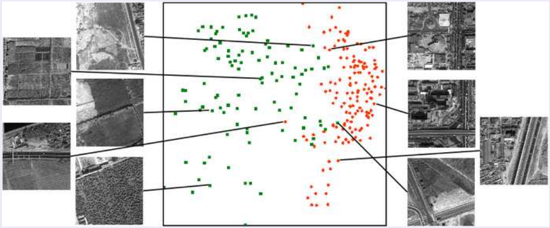
Motivation

Objectives

Provide a novel type of image representation (i.e: global descriptor) that takes into account:

- The appearance of local regions of interest
- The spatial layout of these regions: regions that look like A are they usually close to regions that look like B?

Not a classification problem!



Previous Work

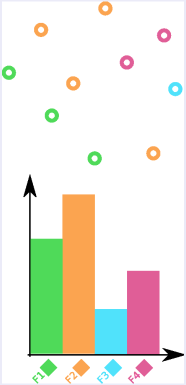
The Bag-of-Features (BoF) Representation (e.g: [Quelhas, Monay, Odobez, Gatica Perez, Tuytelaars 2007])

- Extraction of features of interest
- Descriptor quantization according to a descriptor codebook

Quality of a bag-of-features depends on:

- The feature extractor (Invariances: rotation / scale / affine...)
- The feature descriptor (Robustness to change VS Discriminative power)
- The descriptor codebook (Usually constructed by k-means)

The spatial layout information is lost in the BoF representation



Previous Work

Spectral graph properties

- Dimensionality reduction [Coifman & Lafon, 2006]
- Image segmentation, video tracking [Qiu & Hancock, 2007]
- Land development measure [Unsalan, Boyer, 2005]
- Satellite image content representation [Aksoy, 2006]

Previous Work

Spectral graph properties

- Dimensionality reduction [Coifman & Lafon, 2006]
- Image segmentation, video tracking [Qiu & Hancock, 2007]
- Land development measure [Unsalan, Boyer, 2005]
- Satellite image content representation [Aksoy, 2006]

Previous Work

Spectral graph properties

- Dimensionality reduction [Coifman & Lafon, 2006]
- Image segmentation, video tracking [Qiu & Hancock, 2007]
- Land development measure [Unsalan, Boyer, 2005]
- Satellite image content representation [Aksoy, 2006]

Previous Work

Spectral graph properties

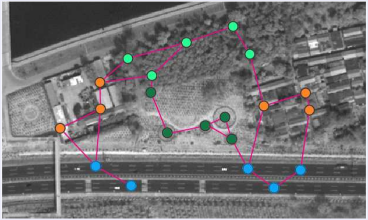
- Dimensionality reduction [Coifman & Lafon, 2006]
- Image segmentation, video tracking [Qiu & Hancock, 2007]
- Land development measure [Unsalan, Boyer, 2005]
- Satellite image content representation [Aksoy, 2006]

The image feature graph

General Idea

E.g: Image regions that look like roads are they usually located near image regions that look like residential areas?

- For each feature, the "proximity" of another feature can be represented by an unoriented, weighted graph edge
- We obtain a weighted, unoriented feature graph in which:
  - Nodes are quantized features
  - Edges represent "proximity" relations



The image feature graph

Construction

The "proximity" between features is defined as a balanced mixture of

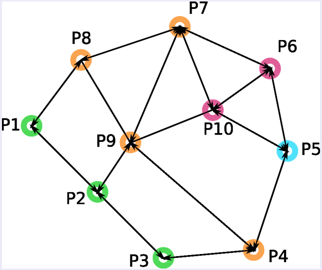
- Feature descriptor similarity
- Interest point spatial proximity

Strong edges connect interest points spatially close to each other and that bear a strong similarity

Parameters

The image feature graph

Representation

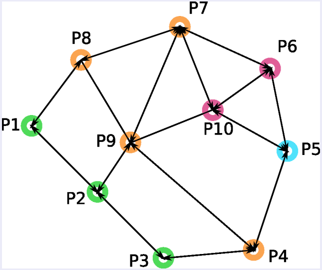


$$
N\left\{\overbrace{\begin{pmatrix} 0 & w_{1,2} & \cdots & w_{1,10} \\ w_{2,1} & 0 & \cdots & w_{2,10} \\ \vdots & & \ddots & \vdots \\ w_{10,1} & \cdots & w_{10,9} & 0 \end{pmatrix}}^{N}\right.
$$

- Graph transition matrix: does not reflect the graph structure
- Shortest path matrix: not robust to point addition/removal, graph variability
- *Graph commute times matrix*

The image feature graph

Representation

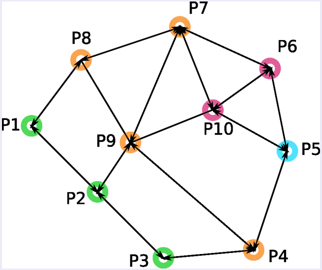


$$N\left\{\overbrace{\begin{pmatrix} 0 & w_{1,2} & \cdots & w_{1,10} \\ w_{2,1} & 0 & \cdots & w_{2,10} \\ \vdots & & \ddots & \vdots \\ w_{10,1} & \cdots & w_{10,9} & 0 \end{pmatrix}}^{N}\right.$$

- Graph transition matrix: does not reflect the graph structure
- Shortest path matrix: not robust to point addition/removal, graph variability
- *Graph commute times matrix*

The image feature graph

Representation



$$N\left\{\overbrace{\begin{pmatrix} 0 & w_{1,2} & \cdots & w_{1,10} \\ w_{2,1} & 0 & \cdots & w_{2,10} \\ \vdots & & \ddots & \vdots \\ w_{10,1} & \cdots & w_{10,9} & 0 \end{pmatrix}}^{N}\right.$$

- Graph transition matrix: does not reflect the graph structure
- Shortest path matrix: not robust to point addition/removal, graph variability
- *Graph commute times matrix*

Graph commute times

The commute time between nodes i and j is defined as the average number of steps of a random walk started at i that are required to reach j for the first time and then to come back to node i for the first time.

Definition

- We define the random walk $(Y_n)_{0 \leq n}$ in the graph by: $P\left[Y_{n+1} = j | Y_n = i\right] = d_{ij} = \frac{w_{ij}}{\sum_{k=1}^{N} w_{ik}}$.
- The *first hitting time* $Q(i,j)$ between nodes i and j is the average number of steps of a random walk started at i that are required to reach j for the first time.$(Q(i,j) \neq Q(j,i))$
- The *commute time* is defined by $CT(i,j) = Q(i,j) + Q(j,i)$ (symmetric measure)

Graph commute times

Graph Laplacian

$$\mathcal{L}(i,j) = \begin{cases} 1 - \frac{w_{ii}}{d_i} & \text{if } i = j \\ \frac{-w_{ij}}{\sqrt{d_i d_j}} & \text{otherwise} \end{cases} \quad (1)$$

$$d_i = \sum_{k=1}^{N} w_{ik} \quad (2)$$

Results concerning the Laplacian [Chung & Yau 2000]

- Eigenvalues of $\mathcal{L}$: $0 = \lambda_1 < \lambda_2 \leq \cdots \lambda_N$
- Eigenvectors: $(\phi_k) 1 \leq N$: $\forall k, \phi_k = (\phi_k(1), \cdots \phi_k(N))$

$$CT(i,j) = \left(\sum_k d_k \right) \sum_{k=2}^{N} \frac{1}{\lambda_k} \left(\frac{\phi_k(i)}{\sqrt{d_i}} - \frac{\phi_k(j)}{\sqrt{d_j}} \right)^2 \quad (3)$$

Graph commute times

Results concerning the Laplacian [Chung & Yau 2000]

$$CT(i,j) = \left(\sum_k d_k\right) \sum_{k=2}^{N} \frac{1}{\lambda_k} \left(\frac{\phi_k(i)}{\sqrt{d_i}} - \frac{\phi_k(j)}{\sqrt{d_j}}\right)^2 \quad (4)$$

Graph embedding and dimensionality reduction

$$\chi_i = \frac{\sqrt{\sum_k d_k}}{d_i} \left(\frac{\phi_2(i)}{\lambda_2}, \dots \frac{\phi'_N(i)}{\lambda'_N}\right) \quad (5)$$

- $N' = N$: Commute time $=$ L^2 distance in an embedding space
- $N' < N$: Dimensionality reduction

The image collapsed graph

Collapsing the feature graph

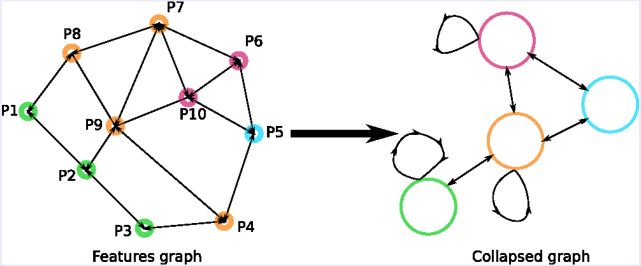
The commute time matrix of the feature graph cannot be used as a representation of image content:

- Variable number of points
- Unordered set of interest points

The feature graph is transformed to obtain the image *collapsed graph*. *The commute time matrix of the image collapsed graph is the image representation.*

Summary

Collapsed graph construction

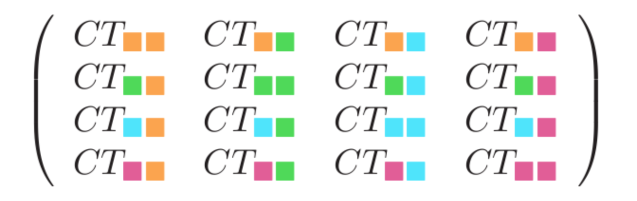


- Each node of the collapsed graph is a codebook entry
- Edges w' of the collapsed graph are defined as: $w'_{■■} = \sum w_{■■}$

The image collapsed graph

Commute time matrix of the collapsed graph

The commute time matrix of the image collapsed graph is the image representation.

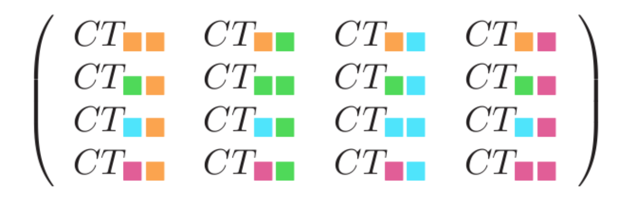


- Problem: huge dimensionality ($\approx K^2/2$, with K = feature codebook size $\approx 10^3$)
- Solution: spectral embedding (again)

The image collapsed graph

Commute time matrix of the collapsed graph

The commute time matrix of the image collapsed graph is the image representation.



- Problem: huge dimensionality ($\approx K^2/2$, with K $=$ feature codebook size $\approx 10^3$)
- Solution: spectral embedding (again)

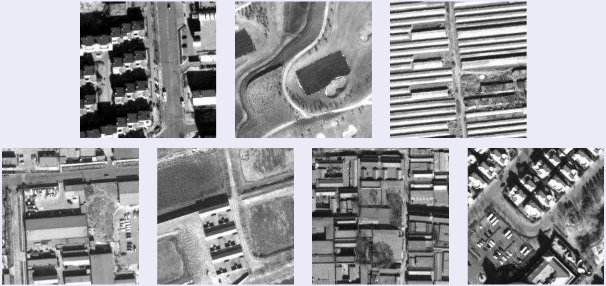
Method Summary

- Extraction of quantized features of interest
- Construction of feature graph
- Construction of collapsed graph
- Computation of the commute time matrix of the collapsed graph
  $\Longrightarrow$ image representation
- Dimensionality reduction of the representation

Performance evaluation

Dataset

- 0.6m Quickbird panchromatic images of Beijing province (China)
- 878 subimages of size 200×200
- 7 classes: (1) big buildings, (2) golf fields, (3) greenhouses, (4) small industry, (5) fields, (6) dense urban, (7) residential area.

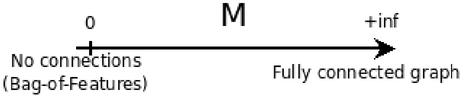


Performance evaluation

Parameters

- Speeded-Up Robust Features (SURF)
- Codebook of size $K = 500$
- Image representations embedded in a space of dimension 20
- 1 VS 1 AdaBoost classifier

Performance evaluation as a function of parameter M (graph connectivity)



	Perf. $M = 0$	Opt. val. M	**Perf. Diff.**
Big buildings	88.24	4	+2.92%
Golf field	92.31	5	+3.84%
Greenhouses	74.55	1	+10.90%
Small industry	75.41	8	+9.84%
Fields	42.41	10	+38.36%
Dense urban	97.48	0	+0.00%
Residential area	91.21	4	+4.39%

Conclusions

- Novel image description method based on spectral properties of graphs inferred from the image content
- Description is based on the appearance of local interest points as well as their layout with respect to one another
- Difficulties posed by many graph structures are overcome by the graph collapse
- Improvement over orderless bag-of-features for most classes
- Unfortunately, there is no single set of parameters that is optimal for all classes

Future work

- Improve feature graph construction
- Modeling the feature graph connections is the right thing to do

Conclusions

- Novel image description method based on spectral properties of graphs inferred from the image content
- Description is based on the appearance of local interest points as well as their layout with respect to one another
- Difficulties posed by many graph structures are overcome by the graph collapse
- Improvement over orderless bag-of-features for most classes
- Unfortunately, there is no single set of parameters that is optimal for all classes

Future work

- Improve feature graph construction
- Modeling the feature graph connections is the right thing to do

Conclusions

- Novel image description method based on spectral properties of graphs inferred from the image content
- Description is based on the appearance of local interest points as well as their layout with respect to one another
- Difficulties posed by many graph structures are overcome by the graph collapse
- Improvement over orderless bag-of-features for most classes
- Unfortunately, there is no single set of parameters that is optimal for all classes

Future work

- Improve feature graph construction
- Modeling the feature graph connections is the right thing to do