Shape matching and correspondence through global energy minimization

ANR SATTIC

Christian Wolf

Université de Lyon, CNRS
INSA-Lyon, LIRIS, UMR5205, F-69622, France
http://liris.cnrs.fr/christian.wolf

March 10th, 2009
Plan

1. Shape matching and graph matching

2. Defining correspondence quality through a global energy

3. Ta, Wolf, Lavoué and Baskurt, 3D Object detection and viewpoint selection in sketch images using local patch-based Zernike moments

4. Outlook: global optimization (Stage PFE Nellie Cardot & Matthieu Barralon)
Plan

1. Shape matching and graph matching

2. Defining correspondence quality through a global energy

3. Ta, Wolf, Lavoué and Baskurt, 3D Object detection and viewpoint selection in sketch images using local patch-based Zernike moments

4. Outlook: global optimization (Stage PFE Nellie Cardot & Matthieu Barralon)
Matching sets of local primitives

Primitives can be interest points, patches, contour fragments, ... 

Matching using different types of criteria:
- feature distances between primitives
- geometric coherence
- topological coherence
Geometric constraints

- Distances are preserved (or not too different)
- Rotation angles are coherent (not too different) in a spatial neighborhood
Defining the neighborhood relationship

Different possibilities:

- Thresholding the Euclidean distance
- Using the nearest spatial neighbors
- In the case of non-overlapping primitives extracted on a regular grid: through the underlying natural spatial neighborhood
- Delaunay triangulation

Notation

\[ i \sim j : \quad i \text{ and } j \text{ are neighbors} \]
### Graphmatching-like algorithms

**Symbol** | **Description** | **Type**
--- | --- | ---
$p_i$ | The spatial position of the patch in the image | observed
$f_i$ | The feature vector extracted at the patch in the image | observed
$x_i$ | The assigned patch in the scene image | hidden
Plan

1. Shape matching and graph matching

2. Defining correspondence quality through a global energy

3. Ta, Wolf, Lavoué and Baskurt, 3D Object detection and viewpoint selection in sketch images using local patch-based Zernike moments

4. Outlook: global optimization (Stage PFE Nellie Cardot & Matthieu Barralon)
Shape/graph matching  
Ta/Wolf/Lavoué/Baskurt  
Outlook: global optimization

Global energy function

\[ \hat{x} = \arg \min_x \lambda_1 \sum_i d_{zd}(f_i^m, f_{x_i}^s) + \lambda_2 \sum_{i \sim j} D_e(d_e(p_i^m, p_j^m), d_e(p_{x_i}^s, p_{x_j}^s)) + \lambda_3 \sum_{i \sim j} D_p(path(x_i, x_j)) + \lambda_4 \sum_{i \sim j} D_a(d_{za}(f_i^m, f_{x_i}^s), d_{za}(f_j^m, f_{x_j}^s)) \] (1)

\begin{align*}
    d_e(\cdot, \cdot) & \quad \text{Euclidean distance} \\
    d_{zd}(\cdot, \cdot) & \quad \text{Feature distance (e.g. Zernike distance)} \\
    d_{za}(\cdot, \cdot) & \quad \text{Rotation angle} \\
    path(x_i, x_j) & \quad \text{Length of the shortest path between } x_i \text{ and } x_j \\
    D_p(\cdot) & \quad \text{a monotonically decreasing function} \\
    D_e & \quad \text{a distance between two Euclidean distances} \\
    D_a(\cdot, \cdot) & \quad \text{a distance between two angles s.t. } (D_a(0, \epsilon) = D_a(2\pi, \epsilon)) \\
    \lambda_i & \quad \text{parameters}
\end{align*}
Additional uniqueness criteria

Alternative and additional uniqueness constraint (each target primitive in the scene may receive a single assignment only) [Torresani, Kolmogorov and Rother, ECCV 2008] :

$$\forall k = 1..S : \sum_{i=1}^{M} \delta_{x_i,k} \leq 1$$ (2)

$$\delta_{a,b} = \begin{cases} 
1 & \text{if } a = b \\
0 & \text{else} 
\end{cases}$$

$$\implies$$ Constraint satisfaction problem

Currently NOT pursued.
Plan

1. Shape matching and graph matching
2. Defining correspondence quality through a global energy
3. Ta, Wolf, Lavoué and Baskurt, 3D Object detection and viewpoint selection in sketch images using local patch-based Zernike moments
4. Outlook : global optimization (Stage PFE Nellie Cardot & Matthieu Barralon)
CBMI 2009 Paper

- Work of Anh-Phuong TA, PhD student at LIRIS since February 2008.
- Industrial contract (CIFRE) with Pinka Productions, Annecy.
- Storyboards: sketch images
- Multiple 2D view images per 3D object

**Library of 3D models**

**offline**

- Preprocessing
- Extraction of overlapping multi-scale patches
- Calculation of Zernike moments

**online**

- Preprocessing
- Extraction of overlapping multi-scale patches
- Calculation of Zernike moments
- Matching
- Verify the neighborhood consistency properties
- Detect and find optimal view
Storyboards: sketch images

Primitives
- Interest points do not work (unstable)
- Contour fragments (k-as, Ferrari et al., PAMI 2008): depend too much on the contour polygonalization
- Retained solution: patches extracted on a regular grid (10,000 positions)

Features
- Zernike moments and rotation angle retrieval

\[ d_{zd}(f^m_i, f^s_j) \] Zernike distance between patches
\[ d_{za}(f^m_i, f^s_j) \] Rotation angle between patches

[ J. Revaud, G. Lavoué and A. Baskurt, IEEE-Tr. PAMI 2009 ]
The global minimum is not searched!

1. Patch assignment using first order terms only (minimal feature distance)
2. Clique by clique evaluation of second order terms
3. Selection of a single node having cliques with maximum coherence
Two step algorithm

1. Patch assignment according to minimum feature distance (first order terms):

\[ \forall i : x_i = \arg \min_x d_{zd}(f^m_i, f^s_x) \]

2. Verification of geometric coherence (second order terms):

\[ \alpha_i = \sum_{i \sim j} \exp \left\{ -\frac{[d_e(p^m_i, p^m_j) - d_e(p^s_{x_i}, p^s_{x_j})]^2}{2\sigma^2_{de}} \right\} \]

\[ \times \exp \left\{ -\frac{[d_{za}(f^m_i, f^s_{x_i}) - d_{za}(f^m_j, f^s_{x_j})]^2}{2\sigma^2_{za}} \right\} \]
Properties

Advantages
- Fast!
- Implicitely deals with outliers and occlusion: wrongly assigned patches are ignored
- Excellent results!

Disadvantages
- Patch assignment on distance only may fail, especially in case of multiple objects of same type in the same scene image
(Recall for 100% precision)

Global method:
- tents: 50
- trailers: 67
- bushes: 30
- trees: 03

(Revaud/Lavoué/Baskurt)
Shown results are from Ferrari et al.
Dataset of natural images

Receiver operation characteristic (ROC): recall and precision for different detection thresholds

Results by Ferrari et al.

Christian Wolf - http://liris.cnrs.fr/christian.wolf

Correspondence & global energy minimization
Recall and precision for different detection quality criteria

![Graphs showing recall and precision](image-url)
Plan

1. Shape matching and graph matching

2. Defining correspondence quality through a global energy

3. Ta, Wolf, Lavoué and Baskurt, 3D Object detection and viewpoint selection in sketch images using local patch-based Zernike moments

4. Outlook: global optimization (Stage PFE Nellie Cardot & Matthieu Barralon)
Discrete optimization

Global minimization of the energy function with:

- Dynamic programming on a chain or tree structured graph (exact!)
- Graph cuts and $\alpha$-expansion move
- Simulated annealing (slow . . .)
- Belief propagation (slow . . .)
- Junction tree algorithm (exact, but even slower . . .)
Dynamic programming

The model patches are chain structured (linear ordering):

Exact solution with dynamic programming

Outliers and occlusion must be dealt with!
Graph cuts

Fast tool for the minimizing of energy functions:

\[ E(x) = \sum_i E_1(x_i) + \sum_{i\sim j} E_2(x_i, x_j) \]

- In its simplest form, only two labels are allowed (=two scene patches)
- Second order terms must satisfy the submodularity condition

\[ E_2(0, 0) + E_2(1, 1) \leq E_2(1, 0) + E_2(0, 1) \]

Neither of these conditions is satisfied in our case!
Graph cuts and $\alpha$-expansion move

$x \leftarrow$ initialization through first order terms only

repeat

\[ \alpha \leftarrow \text{random} \in \{1 \ldots S\} \]

Create a subproblem with binary variables $y_i$ such that:

\[
y_i = 0 \iff x_i \text{ remains unchanged} \]
\[
y_i = 1 \iff x_i \leftarrow \alpha
\]

Solve the problem using graph cuts and assign the $x_i$ accordingly

until convergence/happy/conference deadline

Extensions: $\alpha_i$ different for each node, such that

- $\alpha_i$ corresponds to the most likely label change for node $i$.
- the number of non-submodular terms is decreased
Graph cuts and $\alpha$-expansion move

$x \leftarrow$ initialization through first order terms only

repeat

\[ \alpha \leftarrow \text{random} \in \{1 \ldots S\} \]

Create a subproblem with binary variables $y_i$ such that:

\[ y_i = 0 \iff x_i \text{ remains unchanged} \]

\[ y_i = 1 \iff x_i \leftarrow \alpha \]

Solve the problem using graph cuts and assign the $x_i$ accordingly

until convergence/happy/conference deadline

Extensions: $\alpha_i$ different for each node, such that

- $\alpha_i$ corresponds to the most likely label change for node $i$.
- the number of non-submodular terms is decreased
Conclusion

- Lots of things to do . . .
- Dynamic programming and graph cuts
- Automatic learning of the parameters
- Manage outliers and occlusion
- Uniqueness criterion $\Rightarrow$ constraint satisfaction problem (collaboration wanted)
- Combination with RANSAC type algorithms
- Non-rigid transformations?