Automatic shape expansion with verification to improve 3D retrieval, classification and matching

EG'13 workshop on 3D object retrieval

Jan Knopp, Mukta Prasad, Luc Van Gool

KU Leuven, ETH Zurich
Problem: relevance in retrieval

Retrieval is a problem of finding relevant examples in the database.

Sketch based retrieval
Eitz et al.'12

Template based retrieval
Ovsjanikov et al.'12

Context based retrieval
Fisher et al.'10
Proposed: verification & expansion

- Bag-of-words: results are similar, but don't belong to the category

Query
Proposed: verification & expansion

- Bag-of-words: results are similar, but don't belong to the category

- Verification: relevant results are found.
Proposed: verification & expansion

• Bag-of-words: results are similar, but don't belong to the category.

• Verification: relevant results are found.

• Expansion: get more relevant results.
• Improvement based on local features i.e. Spin images (Joachin'03), **BoW** (Ovsjanikov et al.'08, Toldo et al.'09). BoW does not capture spatial layout of the object, **ssBoW** (Bronstein et al.'11)
Related work

- Improvement based on local features i.e. Spin images (Joachin'03), BoW (Ovsjanikov et al.'08, Toldo et al.'09). BoW does not capture spatial layout of the object, ssBoW (Bronstein et al.'11)

- **2D retrieval: local verification/statistics**
  - Estimated homography (Philbin et al.'04)
  - LDA (Quin et al.'11), Nearest neigh. (Philbin et al.'08)

- **Expansion/Metric Learning** based on verified results using methods above (Chum et al.'07, Philbin et al.'10, Bronstein et al.'10)
Related work: structure

• **Verification driven by representation:** ISM (Leibe'03, Knopp et al.'10), constellations of object parts (Felzenszwalb et al.05, Prasad et al.'11)

• **Verification driven by correspondence:** Fuzzy correspondence(Kim et al.'12), isometric matching (Ovsjanikov et al.'11). If one correspondence is observed, there should be more in expected configuration.

One correspondence $\rightarrow$ many
Related work: structure

- **Verification driven by representation:** ISM (Leibe'03, Knopp et al.'10), constellations of object parts (Felzenszwalb et al.05, Prasad et al.'11)

- **Verification driven by correspondence:** Fuzzy correspondence (Kim et al.'12), isometric matching (Ovsjanikov et al.'11). If one correspondence is observed, there should be more in expected configuration.

- We use local configuration of features to verify similarity and learn more information about shape.
Initial search (vanilla BoW)

Shape → features extraction → Descriptors → quantization

Visual Vocabulary = set of cluster centres (visual words)

(Ovsjanikov et al.'08, Toldo et al.'09, Knopp et al.'10, Ovsjanikov et al.'11)
Initial search (vanilla BoW)

Shape → features extraction → Descriptors

vector of visual words occurrences

quantization

Visual Vocabulary = set of cluster centres (visual words)

Vector of visual words occurrences

(Ovsjanikov et al.'08, Toldo et al.'09, Knopp et al.'10, Ovsjanikov et al.'11)
Verification

Query

Initial results

Query

Verified
Distance between two shapes is represented as Modified Hausdorff Distance.

$$d_{MHD}(q, r) = \frac{1}{F} \sum_{f \in \mathcal{F}} \min_{g \in \mathcal{G}} \{ \text{dist}(f, g) \} + \frac{1}{G} \sum_{g \in \mathcal{G}} \min_{f \in \mathcal{F}} \{ \text{dist}(g, f) \}$$

Finding $\min$ can be faster by searching a smaller set of potential correspondences.
Correspondence between parts can be checked by checking spatial consistency (w.r.t object center).

\[
\text{dist}_{W2} (f, g) = \frac{(\mathbf{o}_F f - \mathbf{o}_G g)^2}{\sigma_x^2}
\]
We use a short-range correspondence verification between two pairs of correspondences on the two objects (thus not relying on object centers any more).

\[
\text{dist}(f, g) = 1 - \frac{1}{M} \sum_{\{i,j\} \in M} \exp\left(-\min(\overline{fi}, \overline{gj})^2 / \sigma_{\alpha}^2\right) \cdot \exp\left(-\left(\overline{fi} - \overline{gj}\right)^2 / \sigma_{\beta}^2\right)
\]

- distance \(ij\) to \(fg\)
- configuration \(ij\) w.r.t. \(fg\)
Expansion

- **Average Query Expansion:**
  Create new BoW vector from query and positively verified shapes.

- **Use also negative examples:**
  Use positives as AQE, but also decrease the effect of visual words that were found as negative.

- **Distance to spiting plane:**
Expansion

• **Average Query Expansion:**
  Create new BoW vector from query and positively verified shapes.

• **Use also negative examples:**
  Use positives as AQE, but also decrease the effect of visual words that were found as negative.

• **Distance to spitting plane:**
• **Average Query Expansion:**
  Create new BoW vector from query and positively verified shapes.

• **Use also negative examples:**
  Use positives as AQE, but also decrease the effect of visual words that were found as negative.

• **Distance to spitting plane:**
  
  ![Diagram](image)

  Query:

  Verified results:

  Improve:
• **LDA** (Philbin et al.'08)
  Topics as mixtures of visual words.

• **DMLL: Distance metric learning for large margins**
  (Philbin et al.'10)
  Decrease distance between related features, while increase distance between features that are not in correspondence.

• **KRN: K-Reciprocal Nearest Neigh.** (Quin et al.'11)
  Q is similar to A if A is similar to Q.
**Evaluation: retrieval**

- **Precision**: the number of correct results divided by the number of all retrieved results.

- **Recall**: the fraction of all relevant documents that is retrieved.

PR is evaluated for the first retrieved shape, then for the first two retrieved shapes and so on. The performance is the area under the PR curve.

<table>
<thead>
<tr>
<th>Tosca</th>
<th>Princeton</th>
<th>SHREC’09</th>
</tr>
</thead>
<tbody>
<tr>
<td>method</td>
<td>12 classes, 474 shapes</td>
<td>~1.8K shapes, ~0.9K training</td>
</tr>
<tr>
<td>W2</td>
<td>0.624, 0.635, 0.649, 0.634</td>
<td>0.283, 0.337, 0.339, 0.326</td>
</tr>
<tr>
<td>WN-global</td>
<td>0.635, 0.713, 0.718, 0.671</td>
<td>0.284, 0.335, 0.332, 0.321</td>
</tr>
<tr>
<td>WN</td>
<td>0.629, 0.624, 0.678, 0.672</td>
<td>0.284, 0.336, 0.337, 0.323</td>
</tr>
<tr>
<td>KRN</td>
<td>0.626, 0.667, 0.684, 0.668</td>
<td>0.283, 0.322, 0.322, 0.308</td>
</tr>
<tr>
<td>LDA</td>
<td>0.627</td>
<td>0.344</td>
</tr>
<tr>
<td>DMLLM</td>
<td>0.632</td>
<td>0.272</td>
</tr>
<tr>
<td>Vanilla</td>
<td>0.622</td>
<td>0.282</td>
</tr>
</tbody>
</table>
Evaluation: classification

Query

Class = ?

Training database
Evaluation: classification

Query

Class = ?

Training database

- **BoW-SVM**
  our implementation of (Toldo et al.'09)

- **knn**
  class is the most occurring one in first k-similar shapes.
Evaluation: classification

<table>
<thead>
<tr>
<th>method</th>
<th>TOSCA</th>
<th>SHREC’09</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>knn</td>
<td>SVM</td>
</tr>
<tr>
<td>ST: standard</td>
<td>89.4%</td>
<td>79.0%</td>
</tr>
<tr>
<td>OE: one-example</td>
<td>47.3%</td>
<td>54.6%</td>
</tr>
<tr>
<td>OE+QE: one-example+QE</td>
<td>63.2%</td>
<td>63.2%</td>
</tr>
</tbody>
</table>

Query

Training database
Evaluation: classification

Keep only one shape per example.

Query

Training database

<table>
<thead>
<tr>
<th>Method</th>
<th>TOSCA knn</th>
<th>TOSCA SVM</th>
<th>SHREC’09 knn</th>
<th>SHREC’09 SVM</th>
</tr>
</thead>
<tbody>
<tr>
<td>ST: standard</td>
<td>89.4%</td>
<td>79.0%</td>
<td>50.0%</td>
<td>60.0%</td>
</tr>
<tr>
<td>OE: one-example</td>
<td>47.3%</td>
<td>54.6%</td>
<td>45.0%</td>
<td>40.0%</td>
</tr>
<tr>
<td>OE+QE: one-example+QE</td>
<td>63.2%</td>
<td>63.2%</td>
<td>45.0%</td>
<td>65.0%</td>
</tr>
</tbody>
</table>
Evaluation: classification

Use the rest of the database for expansion.

<table>
<thead>
<tr>
<th>method</th>
<th>TOSCA</th>
<th>SHREC’09</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>knn</td>
<td>SVM</td>
</tr>
<tr>
<td>ST: standard</td>
<td>89.4%</td>
<td>79.0%</td>
</tr>
<tr>
<td>OE: one-example</td>
<td>47.3%</td>
<td>54.6%</td>
</tr>
<tr>
<td>OE+QE: one-example+QE</td>
<td>63.2%</td>
<td>63.2%</td>
</tr>
</tbody>
</table>

Query

Training database

For expansion
Evaluation: matching

Augmentation of shape descriptors.
Evaluation: matching

Same matching algorithm with the original set of features (top) and enriched (bottom).
Evaluation: matching

Same matching algorithm with the original set of features (top) and enriched (bottom).
Open questions/limitations

• Verification threshold is estimated globally. Class dependent? Object dependent?

• Are all parts of negative data really negative?

• Non-rigid state-of-the-art?
Conclusion

• Improving shape search by weak verification (takes less than 1sec in Matlab) and expansion.

• Local vs. Global.

• Each part of the pipeline can be replaced by different algorithm according to the application.

• Can be directly used for shape search, classification, matching.
Thank you for your attention!