Image Identification for Cultural Databases Using Local-Descriptors

Eduardo Valle
ETIS
Image Identification

- The users find an image in a book, thesis, post card, ...
Document Identification

- The users find an image in a book, thesis, post card, …

- The references are lacking, incorrect or imprecise
Image Identification

• The users find an image in a book, thesis, postcard, …

• The references are lacking, incorrect or imprecise

• The collection is too big
Image Identification

• The users find an image in a book, thesis, postcard, …

• The references are lacking, incorrect or imprecise

• The collection is too big

• The search can take weeks
Copyright Enforcement

- The users detain the copyright over a set of images
Copyright Enforcement

• The users detain the copyright over a set of images

• They find a collection which they suspect may contain unauthorized copies
Copyright Enforcement

• The users detain the copyright over a set of images

• They find a collection which they suspect may contain unauthorized copies

• How can they compute the intersection between those two very large image sets?
Document Identification

Cropping

Colour Changes

Rotation, Scale

Printing, Digitisation

Images C2RMF
Global Signature Method: CCV

Colour histogram:
- 32% black
- 68% white

CCV:
- 28% coherent black
- 4% isolated black
- 68% coherent white
- 0% coherent white

Original pommegrenade image C2RMF
Multiscale CCV

- In order to be robust to cropping, the CCV must be applied to regions.
Multiscale CCV

- A second grid in the regions in-between allows more precision
Multiscale CCV Summary

• Fails where CCV fails: changes in colour, illumination or coherence
  – Use hierarchic versions of more powerful global descriptors
• Works well if cropping is the main distortion
• Not so robust to occlusions
• Performance degrades if query has aspect ratio different from the regions
Local Descriptors: PoI
Local Descriptors: Invariance
Processing Chain

1. **Pre Processing**
2. **Point Detection**
3. **Points**
4. **Point Description**
5. **Vector Signature**
6. **Ranking**
7. **Geometric Consistency**
8. **Indexing kNN**
Processing Chain

- Original
- Processing
- Normal.
- Point Detection
- Points
- Description
- Vector Signature
- Ranking
- Geometric Consistency
- NN Found
- Indexing kNN
Processing Chain

Original Points → Processing → Normal. → Point Detection → Points → Description → Vector Signature → Indexing kNN → NN Found → Geometric Consistency → Ranking
Processing Chain

Original → Pre Processing → Normal → Point Detection → Points → Point Description

Point Description → Vector Signature

Vector Signature → Ranking → Geometric Consistency → NN Found → Indexing kNN
Eliminate spurious points of interested due to JPEG artifacts:
either smoothing or resizing the originals
Processing Chain

- Original
- Pre Processing
- Normal
- Point Detection
- Points
- Point Description
- Vector Signature
- Ranking
- Geometric Consistency
- NN Found
- Indexing kNN
Processing Chain

Criterion of significance: Is the best point distant enough from the second best point?

Underlying hypothesis: each point has at most one good match — this is false in our case
Processing Chain

1. Pre Processing
2. Normal.
3. Point Detection
4. Points
5. Point Description
6. Vector Signature
7. Geometric Consistency
8. NN Found
9. Indexing kNN
10. Ranking
Computer vision context: sometimes more than 90% of outliers!

Image identification: typically less than 10% of outliers.

Use RANSAC instead of generalised Hough Transform

Processing Chain

Original

Point Detection

Geometric Consistency

NN Found

Ranking

Indexing kNN

Vector Signature

Point Description
### Impact in Efficacy  Efficiency

<table>
<thead>
<tr>
<th>Query</th>
<th>CCV</th>
<th>Local Desc.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Original</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>Color</td>
<td>20%</td>
<td>100%</td>
</tr>
<tr>
<td>Dithering</td>
<td>0%</td>
<td>100%</td>
</tr>
<tr>
<td>Noise</td>
<td>45%</td>
<td>100%</td>
</tr>
<tr>
<td>Softening</td>
<td>90%</td>
<td>100%</td>
</tr>
<tr>
<td>Rotation</td>
<td>47%</td>
<td>98%</td>
</tr>
<tr>
<td>Extract</td>
<td>96%</td>
<td>100%</td>
</tr>
<tr>
<td>Skew</td>
<td>82%</td>
<td>95%</td>
</tr>
<tr>
<td>Occlusion</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>Frame</td>
<td>75%</td>
<td>100%</td>
</tr>
<tr>
<td>Incrustation</td>
<td>20%</td>
<td>98%</td>
</tr>
<tr>
<td>Compression</td>
<td>97%</td>
<td>100%</td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td><strong>61,0%</strong></td>
<td><strong>99,2%</strong></td>
</tr>
</tbody>
</table>

Success rate of methods for each kind of query.
## Impact in Efficacy

### Efficiency

<table>
<thead>
<tr>
<th>Query</th>
<th>CCV</th>
<th>Local Desc.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Original</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>Color</td>
<td>20%</td>
<td>100%</td>
</tr>
<tr>
<td>Dithering</td>
<td>0%</td>
<td>100%</td>
</tr>
<tr>
<td>Noise</td>
<td>45%</td>
<td>100%</td>
</tr>
<tr>
<td>Softening</td>
<td>90%</td>
<td>100%</td>
</tr>
<tr>
<td>Rotation</td>
<td>47%</td>
<td>98%</td>
</tr>
<tr>
<td>Extract</td>
<td>96%</td>
<td>100%</td>
</tr>
<tr>
<td>Skew</td>
<td>82%</td>
<td>95%</td>
</tr>
<tr>
<td>Occlusion</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>Frame</td>
<td>75%</td>
<td>100%</td>
</tr>
<tr>
<td>Incrustation</td>
<td>20%</td>
<td>98%</td>
</tr>
<tr>
<td>Compression</td>
<td>97%</td>
<td>100%</td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td><strong>61,0%</strong></td>
<td><strong>99,2%</strong></td>
</tr>
</tbody>
</table>

Success rate of methods for each kind of query.

Original images, Arquivo Público Mineiro
## Impact in Efficacy

### Query | CCV | Local Desc.
--- | --- | ---
Original | 100% | 100%
Color | 20% | 100%
Dithering | 0% | 100%
Noise | 0% | 100%
Softening | 90% | 100%
Rotation | 47% | 98%
Extract | 96% | 100%
Skew | 82% | 98%
Occlusion | 100% | 100%
Frame | 75% | 100%
Incrustation | 20% | 98%
Compression | 97% | 100%
**Average** | **61.0%** | **99.2%**

Success rate of methods for each kind of query.

**Success rate:**
- A few seconds
- Several minutes
Processing Chain

Original -> Processing Normalisation -> Normalisation -> Point Detection -> Points -> Point Description

Before Point Description:
- Vector Signature
- Geometric Consistency
- NN Found

After Point Detection:
- Ranking
Processing Chain

Original → Processing Normalisation

Points → Point Description

NN Found → Consistency → Most Time Consuming

Indexing kNN →

- Geometric Consistency
  - kNN
  - Vector Signature

- Indexing
  - Points
  - kNN
Brute Force Solution

• Inspect each point of the database and keep the $k$ nearest to the query point

• Linear time consumption: $O(n)$ comparisons

✦ Best solution for exact research when $\text{dimensions} > 12$ !!

✦ Often beats approximate methods for dimensions over 30
Approximate Nearest Neighbours
Solution Families

• Tree-based
  – Partition of Space
    • Kd-trees, …
  – Partition of Data
    • Ball-trees, M-trees, …

• Projection-on-axis based
  – Friedman‘75, Nene‘97, MedRank…

• Hashing Based
  – Locality Sensitive Hashing, …
Current Work

• Create a "SIFT-friendly" nearest neighbour algorithm

• Reduce the dimensionality without losing descriptiveness
  – PCA, kernel-PCA...

• Reduce the number of points (in the base and in the query)
Summarising

• In the paper
  – Direct comparison between typical image processing and computer vision method
    • Compelling superiority of the computer vision approach
  – Adaptations for employing SIFT for image identification

• Current work
  – Make nearest neighbour search "SIFT friendly"...
  – ...and make SIFT "nearest neighbour friendly"
Image Identification for Cultural Databases Using Local-Descriptors

Eduardo Valle
ETIS
You are free:

® to share — to copy, distribute, transmit the work;

Under the following conditions:

ℹ attribution — you must attribute the work to its author (but not in any way that suggests that he endorses you or your use of the work);

<pair> noncommercial — you may not use this work for commercial purposes;

※ no derivative works — you may not alter, transform, or build upon this work.

For any reuse or distribution, you must make clear to others the license terms of this work. The best way to do this is with a link to this web page:
http://creativecommons.org/licenses/by-nc-nd/3.0/

Any of the above conditions can be waived if you get permission from the copyright holder.

Nothing in this license impairs or restricts the author's moral rights.

The images indicated licensed from their respective owners and cannot be reused out of this context without their explicit agreement.