It Takes More Than Local Image Features: A Closer Look At Feature Context

Dushyant Mehta
dmetha@mpi-inf.mpg.de

---

**Object Retrieval:** Get me all Walruses!!

**Classification:** Is that a Whale? Dolphin?

---

**Why Dense SIFT fares better over Keypoint SIFT?**

- **Is it # samples ??** Partly
- **What else ??** SIFT Orientation?

- Experiment: Compare (b) and (c)
  - Absolutely Aligned Keypoint SIFT performs significantly better than Relatively Aligned Keypoint SIFT !!
  - Price Paid: Lost Invariance to In-plane rotations

Want the best of both worlds? Let's take a look at the inner workings of Absolutely Aligned Keypoint SIFT

---

**Establishing Explicit Relative Context**

- Get Keypoints in log neighborhood per Keypoint (= 80 x log(scale))
- Bin orientation of neighbor Keypoints relative to own orientation
- Count #Keypoints in quadrants of local reference frame
- Append together to get weak context

**Option #1:** Quantize and bin appearance (SIFT) and context orthogonally (2D Histogram)

- Larger Descriptor Size !!

**Option #2:** Push context information into appearance descriptor and then quantize

---

**Implicit Relative Orientation Context In Absolutely Aligned Keypoint SIFT**

- Modes of variation in orientation bins capture weak relative orientation information

---

**BoW of Feature**

<table>
<thead>
<tr>
<th>Classification Accuracy (15 Scenes Dataset)</th>
<th>Invariance to In-plane Rotation</th>
<th>Invariance to Backgnd Clutter (Hypothesis)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dense SIFT</td>
<td>79% (400°)</td>
<td>X</td>
</tr>
<tr>
<td>Rel. Aligned Keypoint SIFT</td>
<td>59% (400°)</td>
<td>✓</td>
</tr>
<tr>
<td>Abs. Aligned Keypoint SIFT</td>
<td>68% (400°)</td>
<td>✓</td>
</tr>
<tr>
<td>R.A. Keypoint SIFT + Context</td>
<td>68% (400° x 20)**</td>
<td>✓</td>
</tr>
<tr>
<td>R.A. Keypoint SIFT + Dir. Cont.</td>
<td>68% (400° x 20)**</td>
<td>✓</td>
</tr>
</tbody>
</table>

* SIFT Dictionary Size
** Context Dictionary Size

---

**Directional Pooling For Context**

- Try to avoid influence of clutter in context

- Establish Relative Orientation Context by considering neighbor keypoints one local quadrant at a time

- 4x features !! Can be mined for discriminative features [1]

---

**Apply to Convolutional Nets**

- Activation Paths in CNN For Rotated Images

---
