**MIHash: Online Hashing with Mutual Information**

Fatih Cakir*, Kun He*, Sarah Adel Bargal, Stan Sclaroff (*equal contribution)

Computer Science, Boston University  
{fcakir, hekun, sbargal, sclaroff}@cs.bu.edu

---

**Experiments: Online Hashing**

Datasets: CIFAR-10 (60K), Places205 (2.5M), LabelMe (22K)

OKH [IJCAI’13], AdaptHash [ICCV’15], SketchHash [CVPR’15], OSH [ICIP’15]

**Trigger Update on vs. off:** CIFAR-10, 32 bits, mAP

**Trigger Update on:** CIFAR-10, 32 bits, mAP

---

**Experiments: Batch Hashing**

Batch Hashing: CIFAR-10, mAP

**Setting 1:** 5K / 1K, VGG-F fc7.  **Setting 2:** 50K / 10K, VGG-F end-to-end.

**Methods:**
- SHK [CVPR’12]
- SDH [CVPR’15]
- FastHash [CVPR’14]
- VDSH [CVPR’16]
- DPSH [IJCAI’16]
- DTSH [ACCV’16]

**Quality Measure: Mutual Information**

- **Mutual Information:** decrease in entropy of distance distribution, when conditioned on neighborhood info.

\[ D_{k,\Phi}: X \to \{0, 1, \ldots, b\}, x \to d_k(x, \hat{x}) \]

\[ C_{k}: X \to \{0, 1\} \]

**Optimization approach:** continuous relaxation

- Differentiable histogram binning [Usnitova & Lempitsky, NIPS’16]
- Binary bits: sigmoid approximation

<br>

**Background**

Online Hashing: learn \( \Phi : X \to \mathcal{X}^b \) with streaming data.

Q: When do we update the hash table?  
A: Only when hashing quality improves!

<table>
<thead>
<tr>
<th>Quality Measure: Mutual Information</th>
</tr>
</thead>
<tbody>
<tr>
<td>- <strong>Mutual Information:</strong> decrease in entropy of distance distribution, when conditioned on neighborhood info.</td>
</tr>
<tr>
<td>- <strong>Quality measure:</strong> integrate over feature space</td>
</tr>
</tbody>
</table>

\[ Q(\Phi) = \int_{X} I(D_{k,\Phi}; C_{k}) p(\hat{x}) d\hat{x} \]

\[ = \int_{X} (H(C_{k}) - H(C_{k}|D_{k,\Phi})) p(\hat{x}) d\hat{x} \]

---

**MIHash: Gradient-based MI Optimization**

\[ \frac{\partial I(D_{k,\Phi} ; C_{k})}{\partial \Phi(x)} = \frac{\partial H(D_{k,\Phi})}{\partial \Phi(x)} - \frac{\partial H(D_{k,\Phi} ; C_{k})}{\partial \Phi(x)} \]

\[ = \sum_{i} \frac{\partial H(D_{k,\Phi})}{\partial p_{D_{i,\Phi}}} \cdot \frac{\partial p_{D_{i,\Phi}}}{\partial \Phi(x)} - \sum_{i} \frac{\partial H(D_{k,\Phi} ; C_{k})}{\partial p_{D_{i,\Phi}}} \cdot \frac{\partial p_{D_{i,\Phi}}}{\partial \Phi(x)} \]

**Trigger Update: Plug-in Module**

**Online Hashing:** learn \( \Phi : X \to \mathcal{X}^b \) with streaming data.

Q: When do we update the hash table?  
A: Only when hashing quality improves!

---

**Experiments: Batch Hashing**

Batch Hashing: CIFAR-10, mAP

**Setting 1:** 5K / 1K, VGG-F fc7.  **Setting 2:** 50K / 10K, VGG-F end-to-end.

**Methods:**
- SHK [CVPR’12]
- SDH [CVPR’15]
- FastHash [CVPR’14]
- VDSH [CVPR’16]
- DPSH [IJCAI’16]
- DTSH [ACCV’16]

**Experiments**

https://github.com/fcakir/mihash