Approximate Correspondences in High Dimensions

Kristen Grauman\textsuperscript{1,2} and Trevor Darrell\textsuperscript{1}
\textsuperscript{1}CSAIL, Massachusetts Institute of Technology
\textsuperscript{2}Department of Computer Sciences, University of Texas-Austin

Problem
The correspondence between sets of local feature vectors is often a good measure of similarity, but it is computationally expensive.

Optimal partial match
\[ X = \{x_1, \ldots, x_n\}, \quad Y = \{y_1, \ldots, y_n\} \]
\[ \pi^* = \arg \min_{\pi} \sum_{i \in X} |x_i - y_{\pi(i)}|_2 \]

The Pyramid Match
\cite{Grauman and Darrell, ICCV 2005}

In \(O(n^2)\) time, approximate the optimal partial matching cost: use multi-resolution histograms to count matches that are possible within a discrete set of distances.

Pyramid match cost:
\[ C_{PM}(\Psi(X), \Psi(Y)) = \sum_{i=1}^{L} \frac{1}{2} (T_i(H_i(X), H_i(Y)) - T_i(H_i(X), \hat{H}_i(Y))) \]

\[ \Psi(X) = [H_0(X), \ldots, H_L(X)] \]

Weighting options:
1. \( w_{ij} = \text{diameter of cell } i,j \)
2. \( w_{ij} = d_{ij}(X) + d_{ij}(Y) \)

admits a Mercer kernel
input-specific upper bound

The Vocabulary-Guided Pyramid Match

Uniform bins
Vocabulary-specific bins

Uniformly shaped bins result in decreased matching accuracy for high-dimensional features...

\[ C_{VG-FM}(\Psi(X), \Psi(Y)) \]

\[ = \sum_{i=0}^{L-1} \sum_{j=1}^{k_i} \min(n_{ij}(X), n_{ij}(Y)) - \sum_{i=0}^{L-1} \sum_{j=1}^{k_i} \min(e(n_{ij}(X)), e(n_{ij}(Y))) \]

\[ \text{Number of new matches for } j\text{th bin at } i\text{th level} \]

\[ \text{Number of matches in bin } i,j \]

\[ \text{Number of matches in bin } i,j\text{'s children} \]

Weighting options:
1. \( w_{ij} = \text{diameter of cell } i,j \)
2. \( w_{ij} = d_{ij}(X) + d_{ij}(Y) \)

admits a Mercer kernel
input-specific upper bound

The Vocabulary-guided Pyramid Match

\[ C_{VG}(\Psi(X), \Psi(Y)) = \sum_{i,j} \frac{1}{n_{ij}} \sum_{k \in \text{histograms in level } i} \sum_{l \in \text{bins in level } j} \min(c_k, c_l) \]

\[ \text{Number of new matches for } \Psi(X) \text{ at } i\text{th level} \]

\[ \Psi(X) = \{H_0(X), \ldots, H_L(X)\} \]

Future work
- Sub-linear time PM hashing (ongoing)
- Learning weights on pyramid bins
- Beyond geometric vocabularies
- Distortion bounds for the VG-PM?