Nonparametric Kernel Estimators for Image Classification

Barnabás Póczos, Liang Xiong, Dougal J. Sutherland, and Jeff Schneider
Auton Lab, School of Computer Science, Carnegie Mellon University – autonlab.org

Motivation

The “bag of features” approach to image classification, representing images as a set of local features over a grid, is very powerful and popular. But to use it for learning methods like SVMs, we need either a mapping into R^n (e.g. “bag of words” or histograms) or a kernel directly on these sets. We construct such a kernel by considering each set as a sample from an unknown probability distribution, then nonparametrically estimating divergences between the distributions.

Support Distribution Machines

Divergence Estimation

We estimate D with kth-nearest-neighbor distances:

- \( X_{1:n}, n \) samples from \( p \); \( Y_{1:m}, m \) samples from \( q \)
- \( \rho_k(i) \): the distance to the \( k \)th neighbor of \( X_i \in X_{1:n} \)
- \( \nu_k(i) \): the distance to the \( k \)th neighbor of \( X_i \in Y_{1:m} \)
- \( d \): dimension

\[
\hat{D}_{\alpha,\beta}(X_{1:n}, Y_{1:m}) = \frac{1}{n} \sum_{i=1}^{n} \rho_k^{\alpha d}(i) \nu_k^{\beta d}(i)
\]

Image Classification

We will now show experimental results on classifying images into categories, based on 384-dimensional color SIFT features after PCA dimensionality reduction. We compare to:

- Bag of words (BoW)
- BoW processed by pLSA [1]
- the Pyramid Matching Kernel (PMK) [2]
- Kernels based on Gaussians and GMMs:
  - with KL divergence [3]

Comparison to Bag of Words

BoW loses information in quantization, including correlations between codewords, and requires tuning the codebook size.

References


Acknowledgements

This work was funded in part by the National Science Foundation under grant NSF-IIS0910332 and the Department of Energy under grant DE-SC0010267.