Evaluating Dimensionality Reduction Techniques for Visual Category Recognition using Rényi Entropy

Ashish Gupta, Richard Bowden

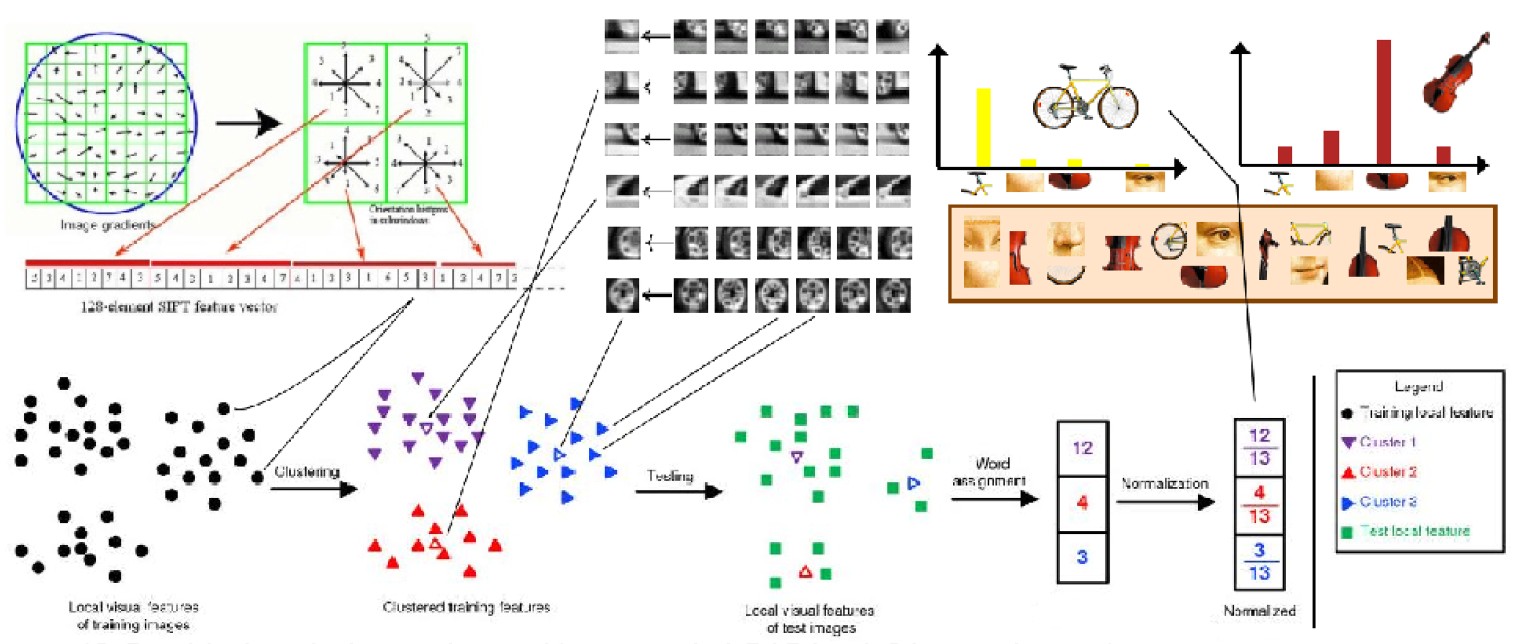
Center for Vision, Speech, and Signal Processing, University of Surrey, Guildford, United Kingdom

Introduction

Tradition linear and modern non-linear dimensionality reduction techniques are evaluated for their comparative benefit towards visual categorization. The evaluation metric utilizes Rényi entropy to measure preservation of data distribution structure upon projection to sub-space. The best technique should have superior entropic and computational time performance.

Visual Category Recognition

The fundamental component is the Bag-of-Features (BoF) model. It encodes high dimensional feature data by assigning features to nearest codebook vectors, which are most representative of the training data. A normalized histogram of training data is utilized as a feature vector for a classifier like a SVM.



* some parts of BoF model schematic above are borrowed from papers by L.Fei-Fei et al., D.Lowe et al., see references.

Dimensionality Reduction Techniques Evaluated

- Principal Component Analysis
- Linear Discriminant Analysis
- Multi-Dimensional Scaling
- Probabilistic PCA
- Factor Analysis
- Isomap
- Landmark Isomap
- Locally Linear Embedding
- Linear Local Tangent Space Alignment
- Diffusion Maps
- Kernel PCA
- Sym Stochastic Neighbor Embedding
- t-distributed SNE
- Neighborhood Preserving Embedding
- Locality Preserving Projection
- Stochastic Proximity Embedding

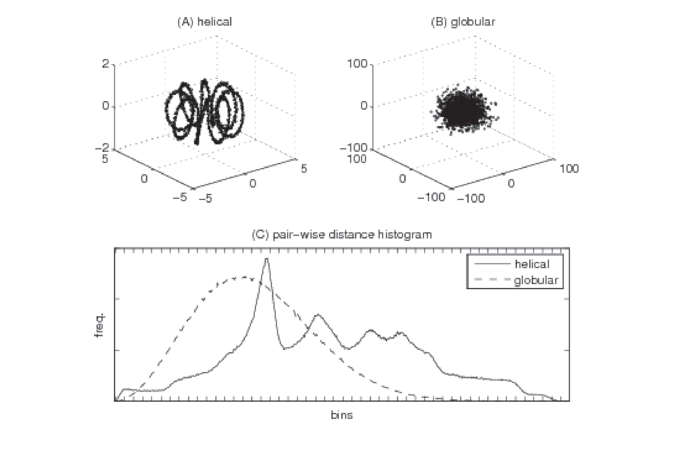
Feature Descriptor

The feature descriptor utilized is the Scale Invariant Feature Transform (SIFT). It is a 128 dimensional affine invariant feature descriptor which perform remarkably well for image matching under significant affine variation and background clutter.

Data set

The dataset used is one of the most popular and benchmark datasets Caltech-101. There are a total of 9146 images, split between 101 different object categories, as well as an additional background/clutter category.

Evaluation Criterion: Rényi Entropy



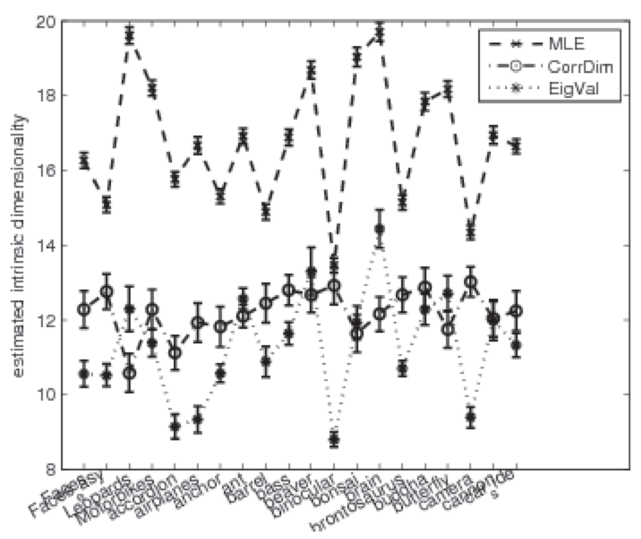
$H_\alpha(\text{helical}) > H_\alpha(\text{globular})$

Rényi entropy H_α: is a family of functionals for quantifying the diversity in a data distribution; admits a generic closed form expression for distributions belonging to the same exponential family; varies monotonically with information and so can be used interchangeably with it.

$$H_\alpha(p) = \frac{1}{1-\alpha} ln \sum_{i=1}^{N} p_i^\alpha$$

Performance metric is the rényi entropy of the pair-wise distance distribution of feature vectors. Data distributions with greater 'structure' have superior H_α of pair-wise distance distribution.

Experiment 1: Intrinsic Dimensionality Estimation

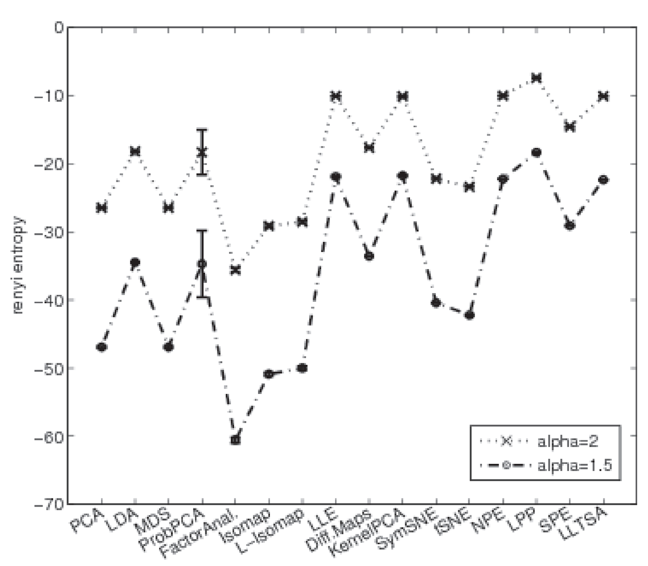


Estimation methods:

- Eigenvalue: is based on PCA, number of eigenvalues greater than a given threshold.
- Correlation dimension: is a fractal based technique to estimate the dimension of the underlying dynamic system.
- Maximum likelihood estimate: It estimates the intrinsic dimensions based on the distance of each data point from its k neighbors.

The average intrinsic dimensionality for Caltech-101 dataset was computed to be 13 dimensions.

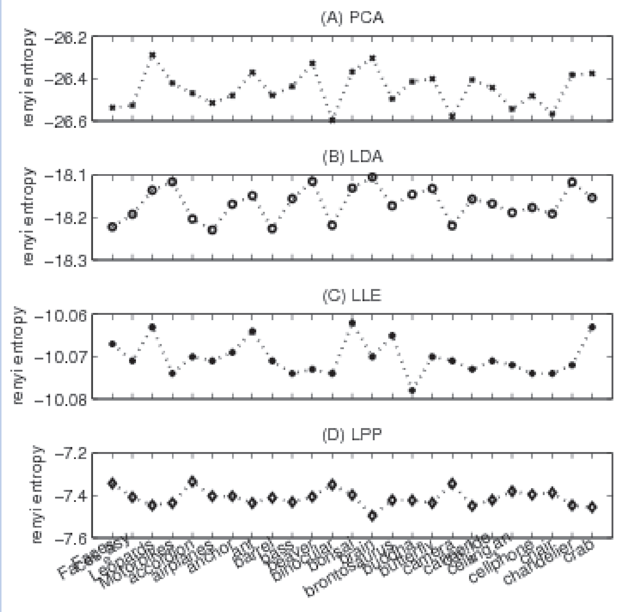
Experiment 2: Entropic performance of the techniques



The rényi entropy H_α for various techniques for parameter values of $\alpha = 1.5$ and $\alpha = 2$. The error bars show the standard deviation in the entropy measure across all the categories. The small variance in entropy measured across categories demonstrates that the choice of dimensionality reduction technique is independent of visual category.

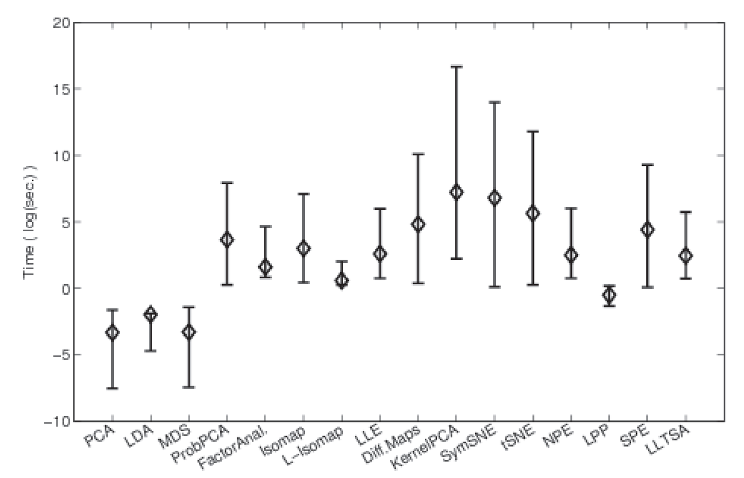
The best performance is achieved by Locality Preserving Projection (LPP). The non-linear techniques in general out perform the linear techniques.

Experiment 3: Correlation between technique and category



No category had a consistently better or worse performance for all techniques. The rényi entropy for PCA,LDA,LLE,and LPP for the first 25 categories of Caltech-101 shows no consistent correlation between technique and visual category. Therefore, the technique which performs best for these 101 categories is likely to remain the best as dataset size increases to more categories.

Experiment 4: Computational cost of techniques



The computational time for all the techniques is displayed on a log scale. The error bars denote the variation across categories, showing the minimum and maximum time utilized. The traditional linear methods PCA, LDA, and MDS are the fastest. Importantly, LPP with the best entropic performance is also faster than other non-linear methods.

Summary

The comparatively best dimensionality reduction technique amongst those evaluated is Locality Preserving Projection. It has superior entropic performance as well as reasonable computational cost. The results could be explained by the ability of this technique to preserve local clusters of visually semantically related feature vectors. Overall, the results highlight the efficacy of non-linear methods over the prevalent linear methods, at a cost of higher computational time. Future work would include testing Locality Preserving Projection on other datasets to ascertain consistent superior performance, and assessing classification performance along with the entropic measure as evaluation criteria.

References

1. L. Fei-Fei, R. Fergus and P. Perona, "Learning generative visual models from few training examples: an incremental Bayesian approach tested on 101 object categories", in CVPR'04, Workshop on Generative-Model Based Vision. 2004.
2. G. Csurka, et. al., "Visual categorization with bags of keypoints", in Proc. of ECCV Int'l Workshop on Statistical Learning in Computer Vision, 2004.
3. D.G. Lowe, "Distinctive Image Features from Scale Invariant Keypoints", Int'l Journal Computer Vision, vol. 60, no. 2, pp. 91-110, 2004.

Acknowledgement

This work is supported by the EU project Dicta-Sign (FP7/2007-2013) under Grant No. 231135 and the EPSRC project "Making Sense".

Center for Vision, Speech, and Signal Processing - University of Surrey - Guildford, United Kingdom

Mail: a.gupta@surrey.ac.uk

WWW: http://www.ee.surrey.ac.uk/cvssp