Sparse Models in Image Understanding And Computer Vision

Jayaraman J. Thiagarajan

Arizona State University

Collaborators

Prof. Andreas Spanias

Karthikeyan Natesan Ramamurthy

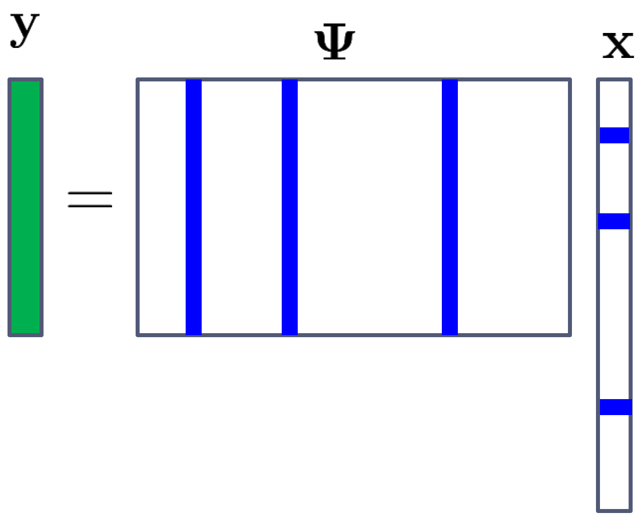
Sparsity

- Sparsity of a vector means that only a small number of its elements differ significantly from zero.
  - Modeling natural signals [Elad2010].
  - Supervised classification [Figueiredo 2003].
  - Understanding natural image statistics [Olshausen & Field 1996].
- Object recognition in the "cortex" [Mutch and Lowe 2008], [Riesenhuber and Poggio 2005]
  - Sparse, localized features
  - Hierarchical approach - Simple low-level features having are pooled and combined into complex, higher-level features.

Sparse methods lead to parsimonious (and interpretable) models in addition to being efficient for large scale learning.

Sparse Coding

- Solve an underdetermined (overcomplete) system of equations



Signal/data - $\mathbf{y} \in \mathbb{R}^M$

Dictionary - $\mathbf{\Psi} \in \mathbb{R}^{M \times K}$

Coefficient vector - $\mathbf{x} \in \mathbb{R}^K$

$$\min_{\mathbf{x}} \|\mathbf{x}\|_0 \quad \text{subject to} \quad \|\mathbf{y} - \mathbf{\Psi}\mathbf{x}\|_2 \leq \epsilon$$ → Combinatorial Complexity

$$\min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad \text{subject to} \quad \|\mathbf{y} - \mathbf{\Psi}\mathbf{x}\|_2 \leq \epsilon$$ → Basis Pursuit [Chen 2001]

- Algorithms - Matching Pursuit , Orthogonal Matching Pursuit, Iterated shrinkage, LARS, Feature sign search and many others.

Dictionary Design

- The dictionary Ψ can be constructed as a
  - Pre-defined set of basis functions
  - Union of orthonormal bases
  - Overcomplete set of features adapted to the data
- Algorithms – Conjugate gradient descent based methods, K-SVD, FOCUSS, MOD and many others…

- Adapting dictionaries to the training data – Generalization of data clustering [Thiagarajan et.al. 2010]
  - A data sample can be associated to more than one cluster and an activation value is computed for each cluster.
  - K-lines clustering – Special case of K-subspace clustering with K = 1 and the additional constraint that the subspace passes through origin.

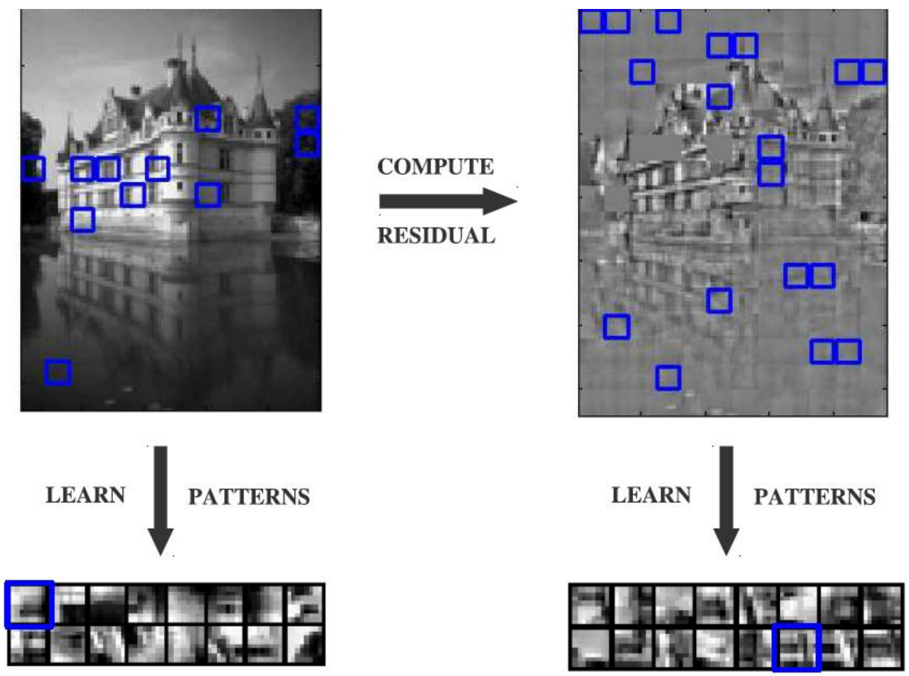
Learning Global Dictionaries

- Training a dictionary for every new set of training samples is not feasible.
- A learning algorithm is a map from the space of training samples to the hypothesis space of functional solutions.

- Can the learning algorithm recover the underlying global dictionary "stably"?
  - A stable algorithm will depend only on the probability space to which the training samples belong.
- Given that the training error is small, can we ensure that the "Expected" test error is also small?
  - Need to obtain an upper bound for the difference between the empirical and expected errors.

Designing Global Dictionaries

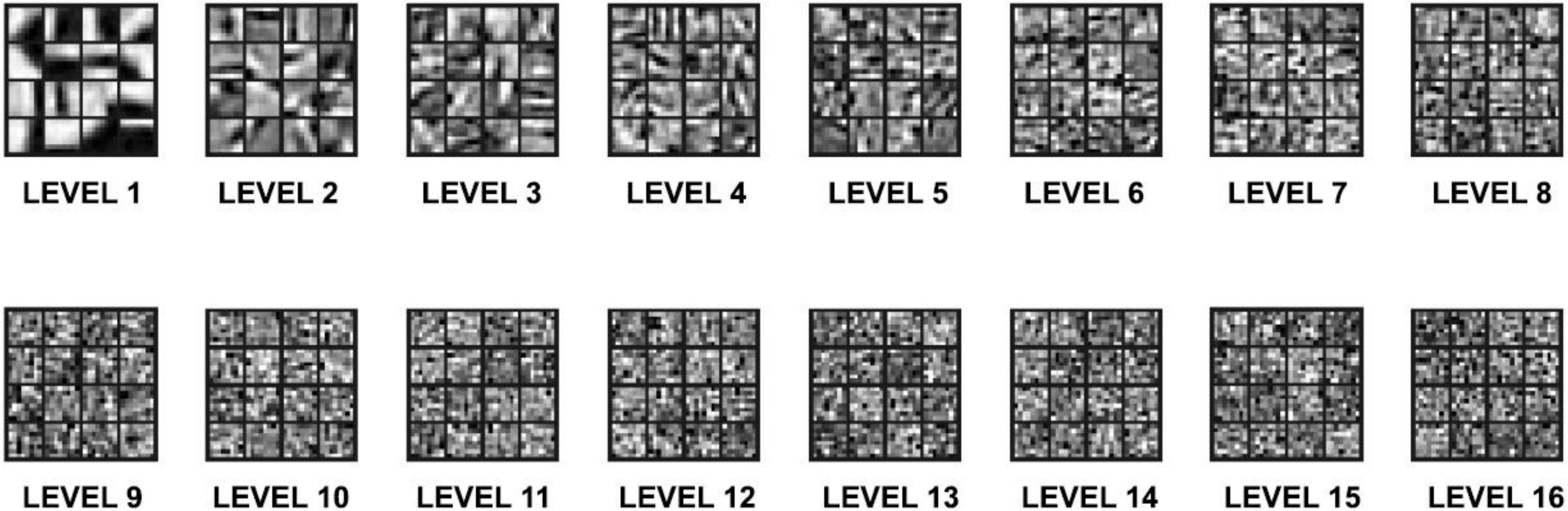
- Natural image patches exhibit redundancy and hence can be efficiently coded.
- Image patches contain either low dimensional geometric patterns or stochastic textures, or a combination of both.
  - Energy hierarchy in the patterns can be exploited.



Multilevel Dictionary Learning

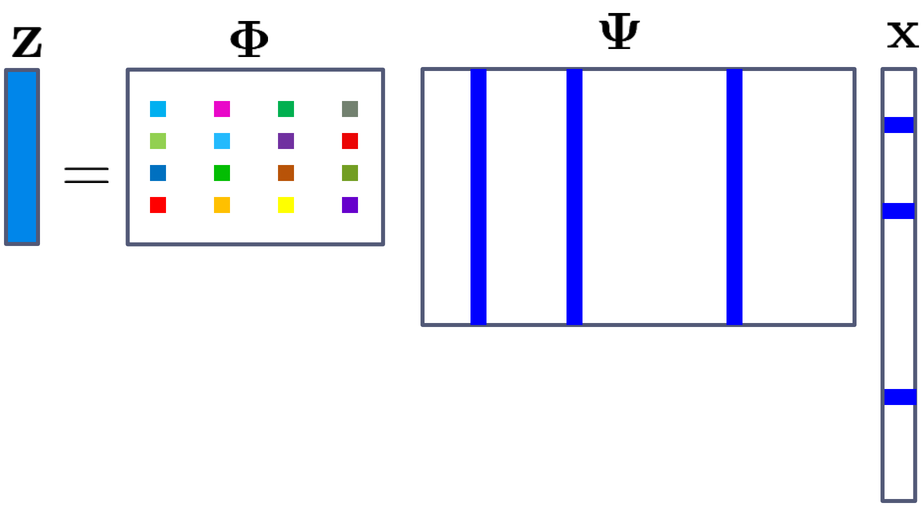
An Example Dictionary

- Geometric patterns in the first few levels and stochastic patterns in the last few levels.



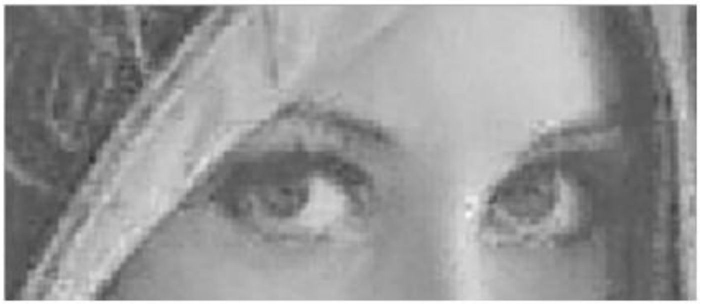
Compressed Sensing

- Recovery of images using compressed measurements
- Measurement system: Random (or) Optimized to the dictionary.



Perform sparse recovery using observations (Z)

KSVD (PSNR = 30.45 dB)



Proposed (PSNR = 32.58 dB)

Using 25% Noisy Measurements

Sparse Coding in Recognition

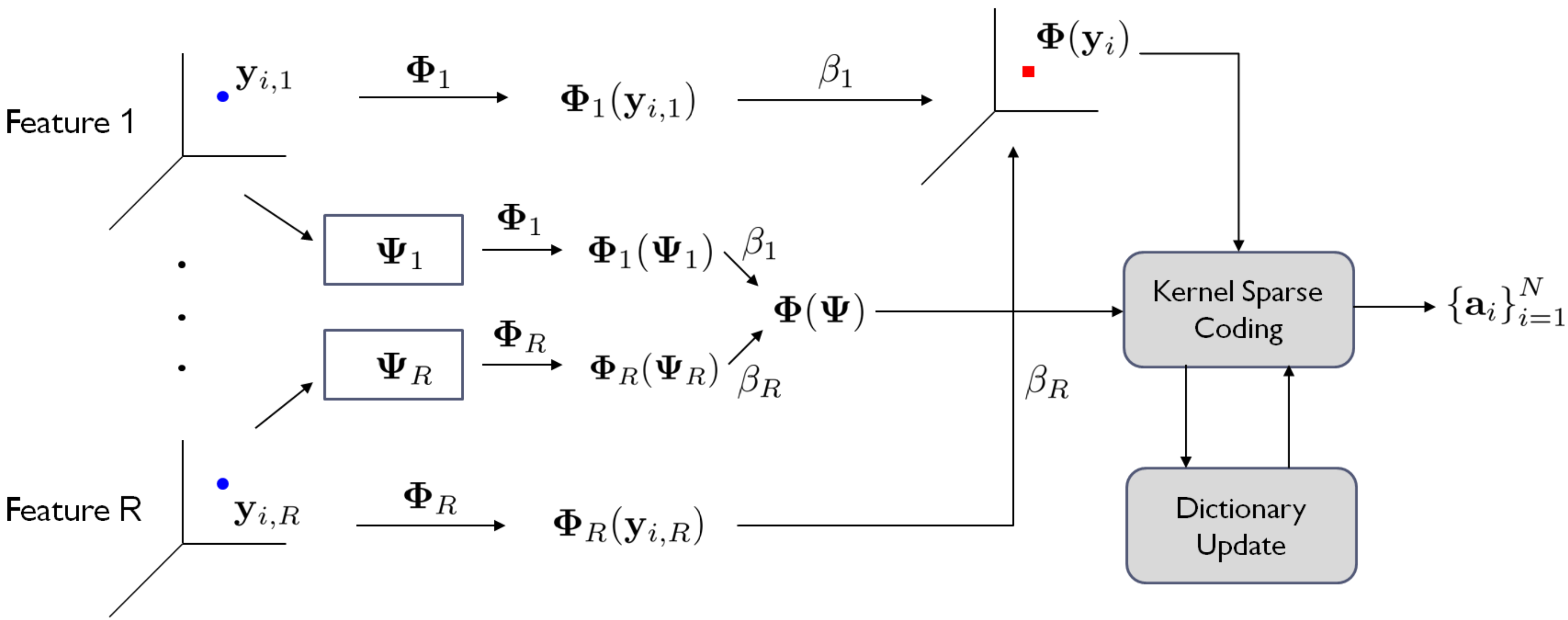
- Challenges:
  - No single descriptor can describe the whole dataset.
  - Diverse nature and high dimensionality of the descriptors – vectors, histograms, matrices and tensors.

- Proposed solution:
  - Employ kernel methods to learn models using the similarities between data samples.
  - Perform sparse coding in the feature space obtained by fusing multiple kernels (MKSR).
  - Low-dimensional compact representation for recognition.

- Learning dictionaries in the ensemble feature space.

Multiple Kernel Sparse Representations

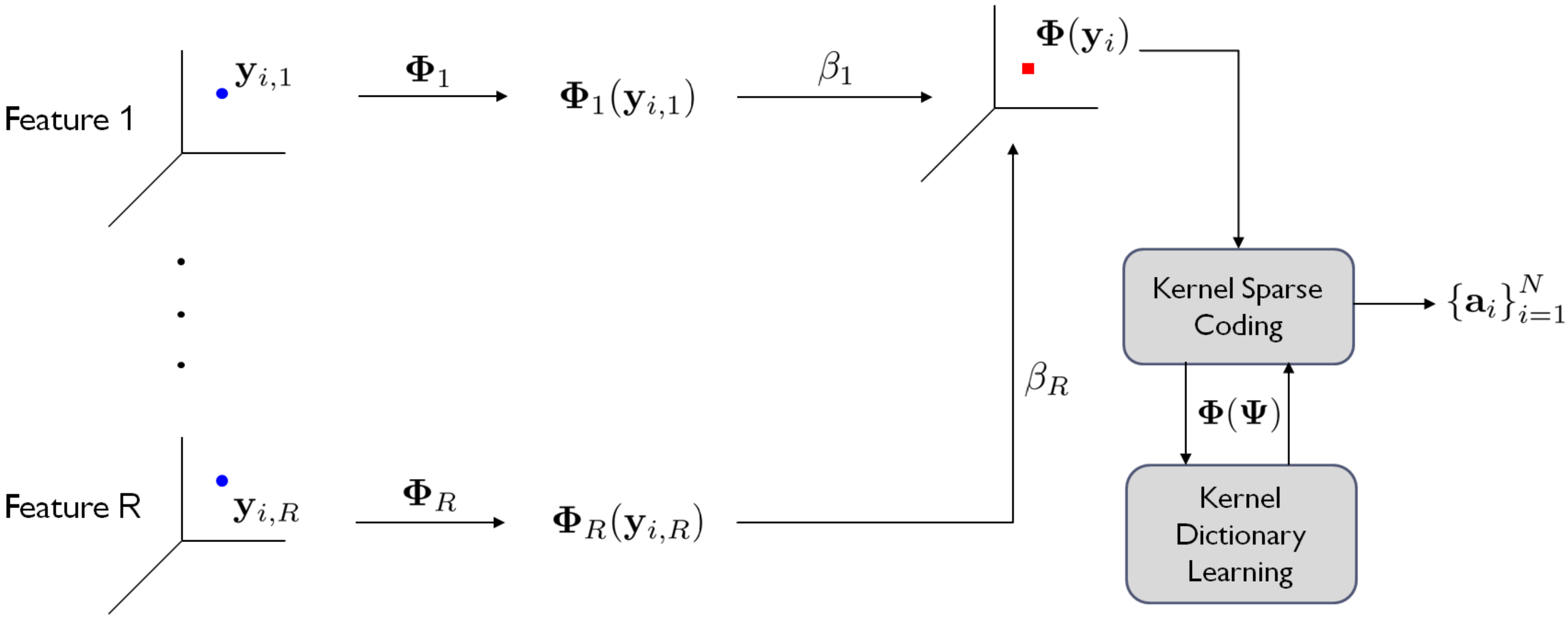
Approach I



- Learn a separate dictionary for each descriptor and obtain ensemble kernel matrices for sparse coding.
- Complexity: O(MK).

Multiple Kernel Sparse Representations

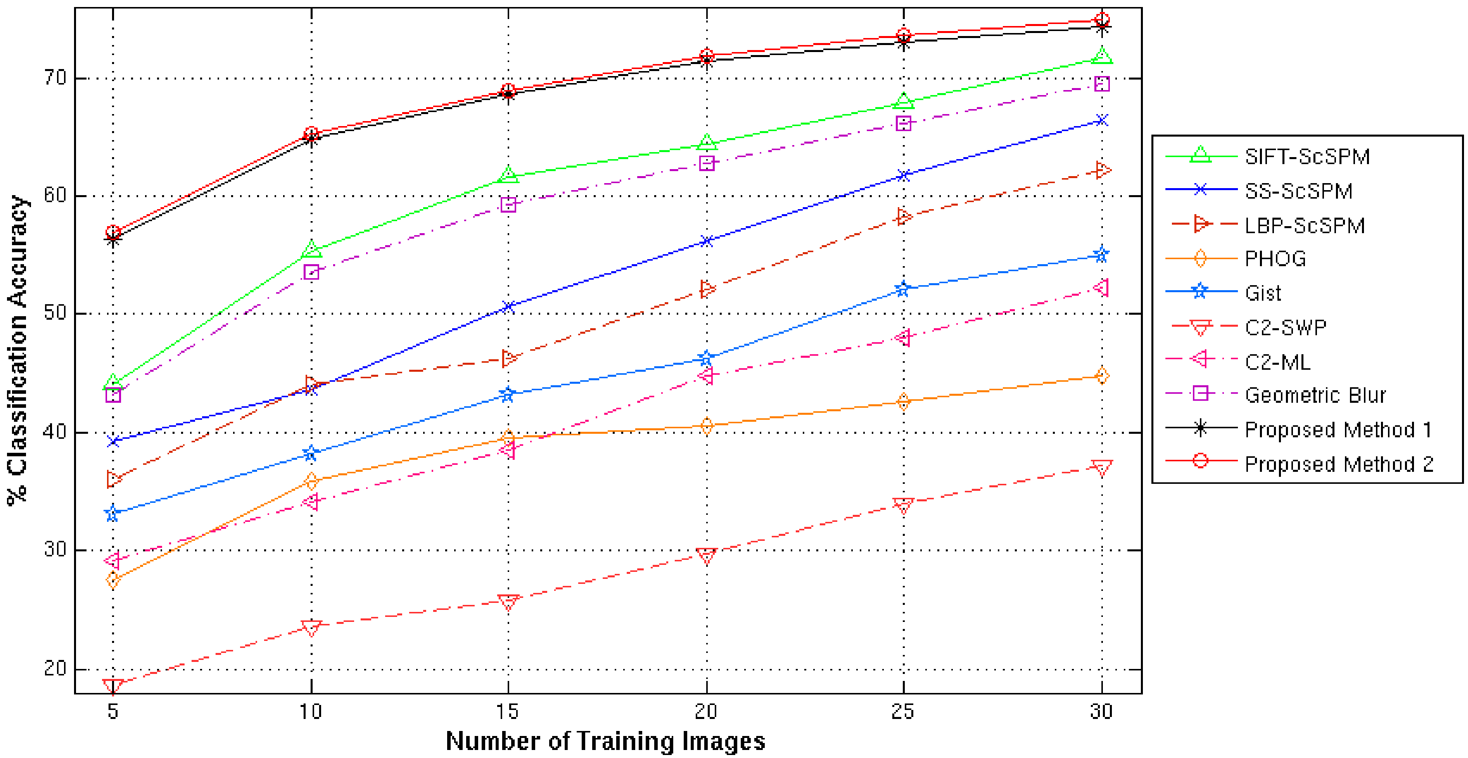
Approach 2



- Perform kernel dictionary learning using the ensemble kernel matrix directly.
- Complexity: O(MN).

Object Recognition Performance

- Caltech-101/Caltech-256 Object Datasets
  - SIFT, Self similarity, LBP, Gist, PHOG, Geometric Blur, C2-SWP, C2-ML.
  - Linear SVM for classifying the MKSR codes.



Object Recognition Performance

Algorithm	5	10	15	20	25	30
Spatial Pyramid Matching	-	-	56.4	-	-	64.6
Sparse Coding + SPM	-	-	67	-	-	73.2
LLC + SPM	51.15	59.77	65.43	67.74	70.16	73.44
LC-KSVD	54	63.1	67.7	70.5	72.3	73.6
Multiple Kernel SC (Approach 1)	**56.34**	**64.81**	**68.56**	**71.4**	**73.07**	**74.29**
Multiple Kernel SC (Approach 2)	**56.9**	**65.3**	**68.94**	**71.83**	**73.61**	**74.88**

Algorithm	15	30	45	60
Sparse Coding + SPM	27.73	34.02	37.46	40.14
LLC + SPM	34.46	41.19	45.91	47.68
Multiple Kernel SC (Approach 1)	**36.46**	**43.12**	**46.24**	**48.26**
Multiple Kernel SC (Approach 2)	**37.19**	**43.81**	**46.92**	**48.87**

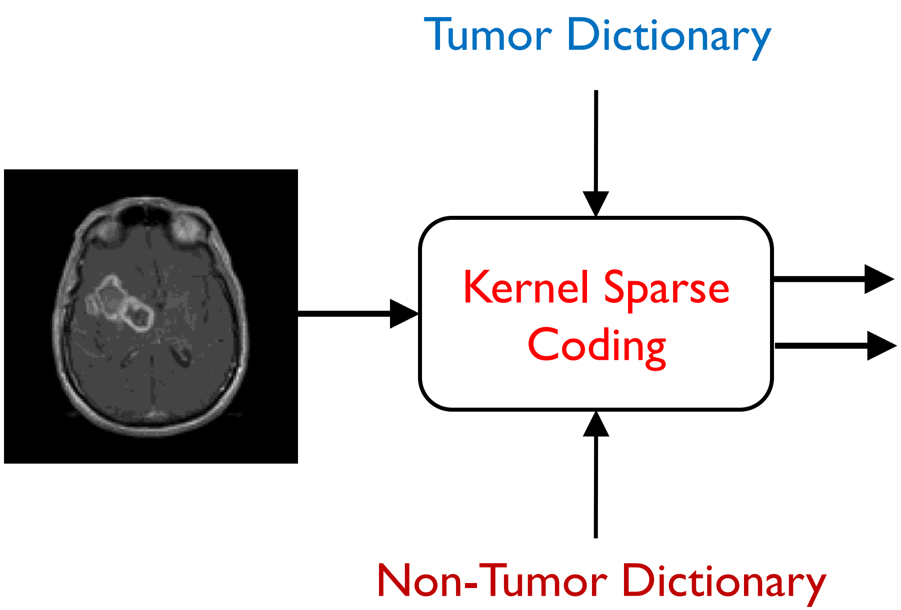
Tumor Segmentation

- Robust method to automatically segment a medical image into its constituent heterogeneous regions.
  - Active and necrotic tumor components from T1-weighted contrast enhanced MR images.

- Challenges:
  - Variability in size, shape and location.
  - Similarity in intensities of normal and abnormal brain tissue regions.
  - Intensity variations of identical tissues across volumes.
  - Avoid overestimation.

Kernel Coding for Segmentation

- Sparse coding typically applied to image patches (or) feature vectors.
  - Trivial to obtain sparse codes for pixel intensities.
  - Proposed solution: Perform coding using kernel similarities.



$$E_T = \|\Phi(y_i) - \Phi(\mathbf{D}_T)x_i^T\|_2$$

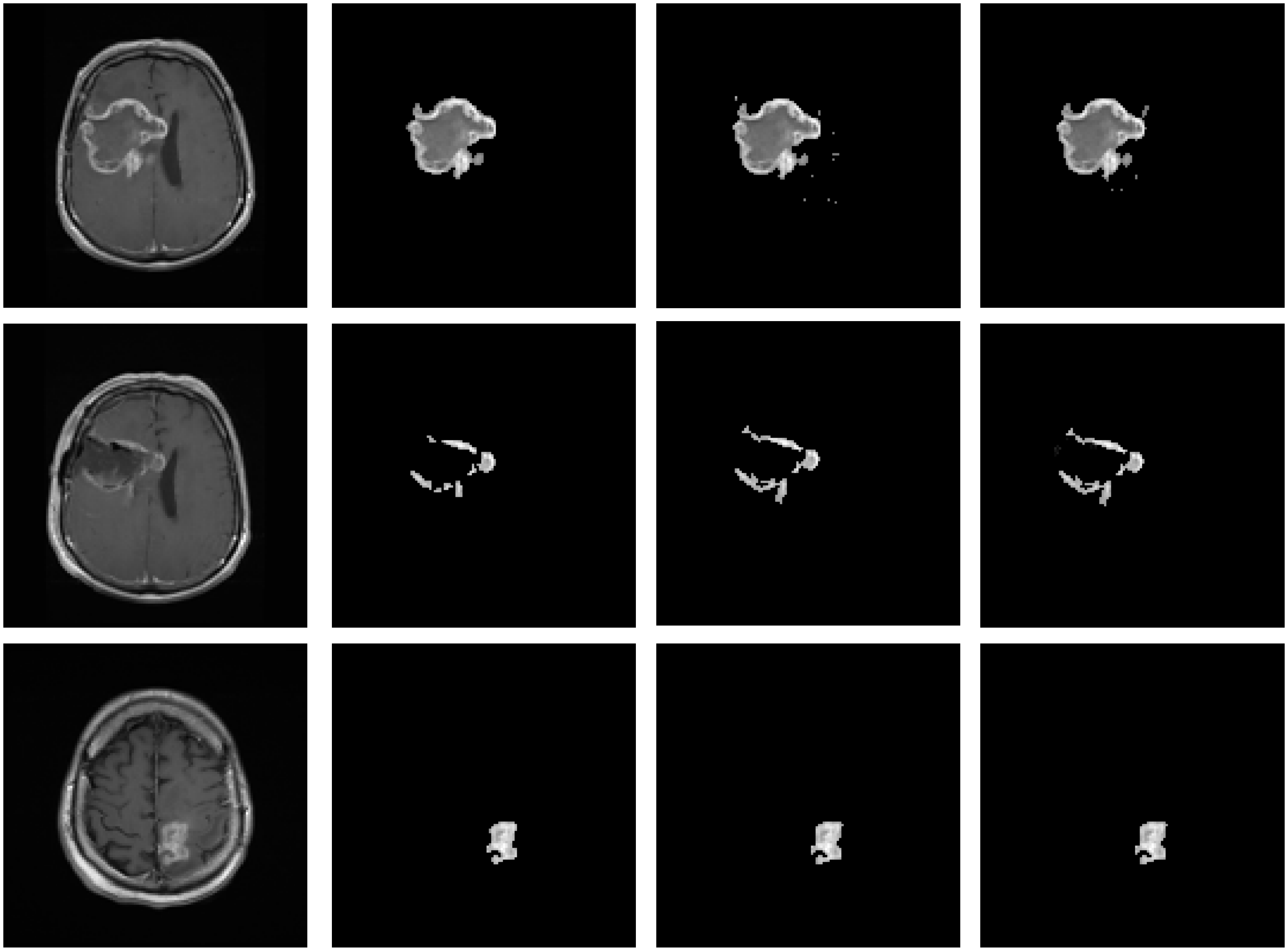
$$y_i = \begin{cases} \mathit{Tumor}, & \text{if } E_N - E_T \geq \epsilon, \\ \mathit{Non\text{-}tumor}, & \text{otherwise.} \end{cases}$$

$$E_N = \|\Phi(y_i) - \Phi(\mathbf{D}_N)x_i^N\|_2$$

Tumor Identification

- Need to identify locally connected segments
  - Segmentation algorithms typically consider pixel locations in addition to intensities.
- Incorporation of locality information
  - Approach 1: Perform spectral clustering only on the pixels determined as tumor based on kernel codes.
  - Approach 2: Include the locality information as part of the kernel
- Ensemble kernel can be constructed as the Hadamard product of intensity and locality kernels.
  - Tumor region can be identified using linear SVM.
- Complexity reduction can be achieved by allowing user to initialize the tumor region.

Experiment Results



Experiment Results

Image	Acc (%)	CR	Acc (%)	CR
Slice 1	93	0.9	94	0.93
Slice 2	96	0.95	97	0.95
Slice 3	92	0.9	95	0.94
Slice 4	90	0.86	90	0.86
Slice 5	94	0.85	94	0.87
Slice 6	92	0.82	92	0.81
Slice 7	94	0.76	95	0.72
Slice 8	98	0.95	98	0.95
Slice 9	98	0.92	98	0.92
Slice 10	92	0.81	92	0.84

Image Retrieval

- Local descriptors from small patches – Object recognition.
- For general image retrieval tasks – Typical to consider a heterogeneous combination of multiple features.
- Assumption: Well annotated tags available for a sample set.
- Is it possible to use this supervised information in coding?

Locality in Sparse Models

- Moving away from VQ – Relative importance of the different bases are not considered.
- Sparse Coding – Lesser reconstruction error, but loses correlation between the codes.
- Consistency - Similar features must have similar codes.
- Salient patterns in the neighborhood – Local linear model.
- By adding a suitable regularization for locality, sparse coding can provide improved recognition performance.

$$\min_{\mathbf{x}} \sum_{k=1}^{K} w(k)|x_k| \quad \text{subject to} \quad \|\mathbf{y} - \mathbf{\Psi x}\|_2 \leq \epsilon$$

$$w(k) = \|\mathbf{y} - \psi_k\|_2^2 \qquad\qquad w(k) = \|\mathbf{y} - (\mathbf{y}^T \psi_k)\psi_k\|_2^2$$

Supervised Coding

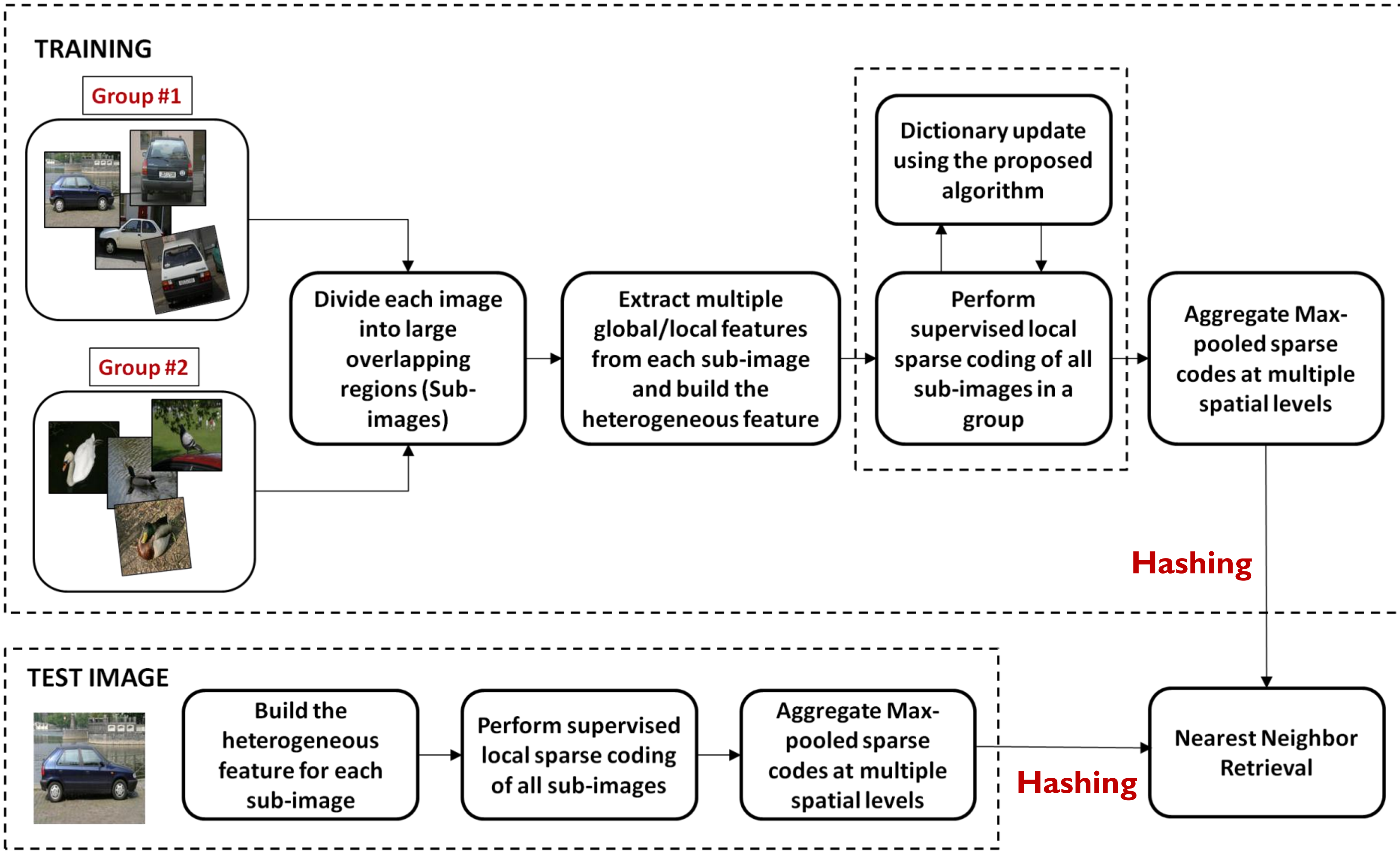
- Using heterogeneous features from large regions of an image.
- Provides enough variability to understand the interactions between different entities.
- Learning a dictionary for sparse coding these features – Bag of Visual Phrases.

- Can be further improved by performing supervised coding
  - Simultaneous sparse coding of features within a group (tag/label).

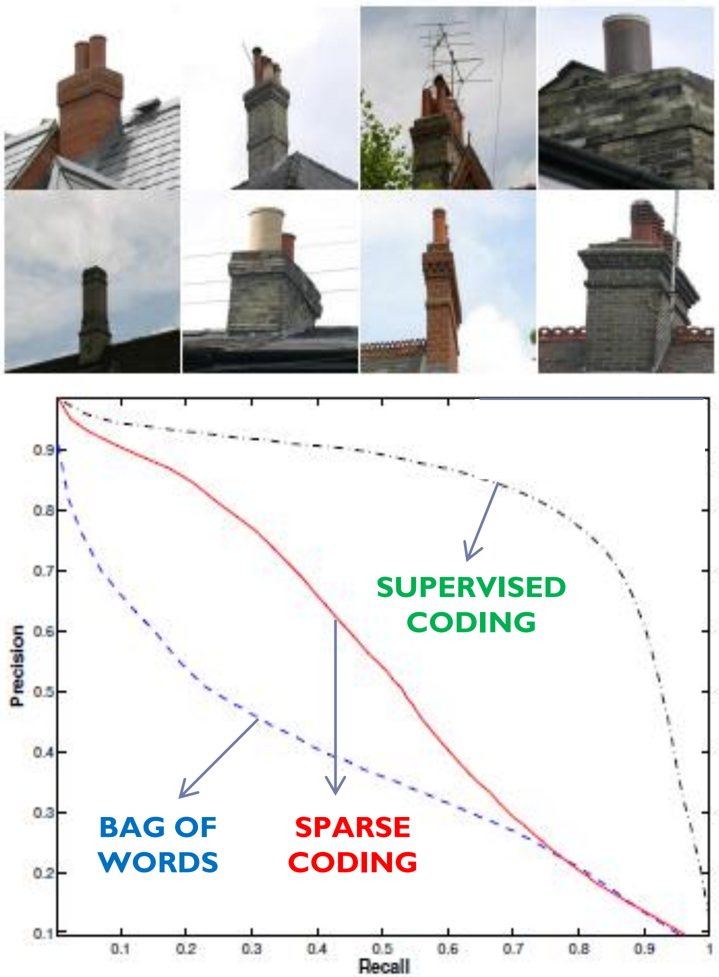
$$\hat{\mathbf{X}}^{(g)} = \min_{\mathbf{X}} \|\mathbf{Y}^{(g)} - \Psi \mathbf{W}_{(g)}^{-1} \mathbf{X}\|_F^2 \text{ s.t. } \|\mathbf{X}\|_{row-0} \leq L$$

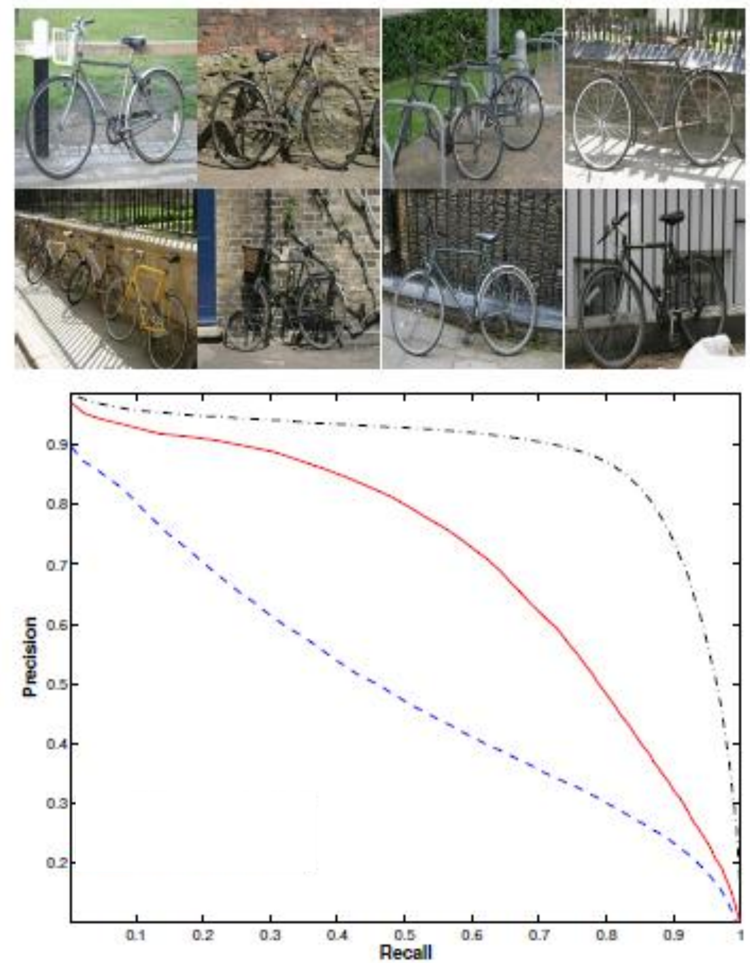
  - One image can be part of several groups.
  - Design dictionaries to "optimize" this representation.

Algorithm


TRAINING
Group #1
Group #2
Divide each image into large overlapping regions (Sub-images)
Extract multiple global/local features from each sub-image and build the heterogeneous feature
Dictionary update using the proposed algorithm
Perform supervised local sparse coding of all sub-images in a group
Aggregate Max-pooled sparse codes at multiple spatial levels
Hashing
TEST IMAGE
Build the heterogeneous feature for each sub-image
Perform supervised local sparse coding of all sub-images
Aggregate Max-pooled sparse codes at multiple spatial levels
Hashing
Nearest Neighbor Retrieval

Simulations





Other Research Problems

- Sparse coding on Riemannian manifolds for activity recognition
- Dictionary Learning with graph embedding constraints
- Discriminative clustering in ambient and feature spaces
- Combined sparse representations
  - Derived conditions for unique recovery using convex and greedy methods.
- Wavelet domain statistical models for template learning
  - Fast image registration using non-stationary GMRF templates.
- Example based coding for Image recovery
- Shift-invariant sparse representations
- Transform domain features for ion-channel signal classification

References

[1] M. Figueiredo., "Adaptive sparseness for supervised learning," IEEE PAMI, 25(9):1150–1159, September 2003.

[2] B. Olshausen and D. Field. "Emergence of simple-cell receptive field properties by learning a sparse code for natural images," Nature, 381:607–609, 1996

[3] M. Riesenhuber and T. Poggio, "Hierarchical models of object recognition in cortex.," Nature Neuroscience, 2(11):1019–1025, 1999.

[4] J J. Mutch and D. G. Lowe, "Multiclass object recognition with sparse, localized features," In CVPR, pages 11–18, New York, June 2006.

[5] M. Elad, Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing. Springer, 2010.

[6] Z. He et.al., "K-hyperline clustering learning for sparse component analysis," Signal Processing, vol. 89, pp. 1011-1022, 2009.

[7] A. Rakhlin and A. Caponnetto, "Stability of K-means clustering," In NIPS, vol. 19. Cambridge, MA: MIT Press, 2007.

[8] S. Lazebnik et.al., "Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories," In CVPR, 2006.

[9] Roberto Rigamonti et.al., "Are Sparse Representations Really Relevant for Image Classification ?," In CVPR 2011.

[10] D. L. Donoho and M. Elad, "Optimally sparse representation in general (non-orthogonal) dictionaries via l1minimization," In PNAS, vol. 100, no. 5, pp. 2197-2202, March 2003.

Selected Publications

[1] J. J. Thiagarajan et.al., "Optimality and stability of the K-hyperline clustering algorithm," Pattern Recognition Letters, vol. 32, no. 9, pp. 1299-1304, 2010.

[2] J. J. Thiagarajan et.al., "Learning stable multilevel dictionaries for sparse representation of images", IEEE PAMI, 2012.

[3] J. J. Thiagarajan, K. N. Ramamurthy and A. Spanias, "Local sparse coding for image classification and retrieval", Pattern Recognition Letters, 2012.

[4] J. J. Thiagarajan and A. Spanias, "Multiple kernel sparse representations for object recognition," IEEE Transactions on Image Processing (Under review)

[5] J. J. Thiagarajan et.al., "Supervised local sparse coding of sub-image features for image retrieval," IEEE ICIP 2012.

[6] P. Sattigeri et.al., "Implementation of a fast image coding and retrieval system using a GPU", IEEE ESPA, 2012.

[7] J. J. Thiagarajan and A. Spanias, "Learning dictionaries for local sparse coding in image classification," Asilomar 2011 (Nominated for the Best Student Paper award).

[8] K. N. Ramamurthy, J. J. Thiagarajan and A. Spanias, "Improved sparse coding using manifold projections," Proc. of IEEE ICIP, 2011.

[9] J. J. Thiagarajan, et.al., "Multilevel dictionary learning for sparse representation of images," in Proc. of IEEE DSP Workshop, Sedona, 2011 (Nominated for the Best Student Paper award).

[10] K. N. Ramamurthy et.al., "Fast image registration using non-stationary Gauss Markov random field templates," Proc. of IEEE ICIP, 2009.