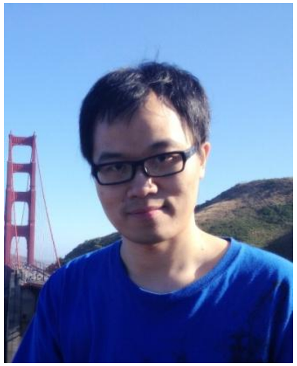
Feature Learning by Multidimensional Scaling and its Applications in Object Recognition

2013 26th SIBGRAPI Conference on Graphics, Patterns and Images

Presented by:

Kim L. Boyer kim@ecse.rpi.edu

Authors:
Quan Wang, Kim L. Boyer
Signal Analysis and Machine Perception Laboratory
Department of Electrical, Computer, and Systems Engineering
Rensselaer Polytechnic Institute

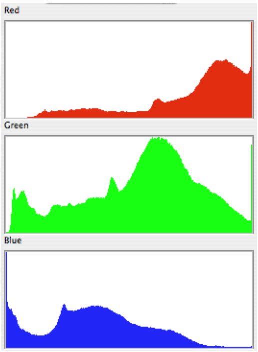


Review on Image Features

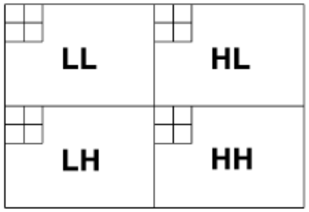
- How do we represent an image by a fixed-length feature vector?
  - Hand-designed features

Review on Image Features

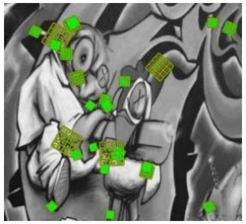
- How do we represent an image by a fixed-length feature vector?
  - Hand-designed features

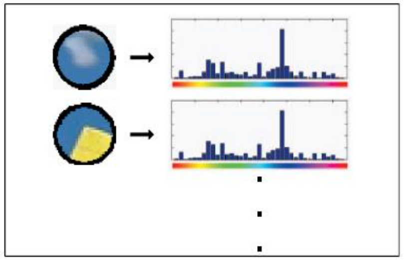


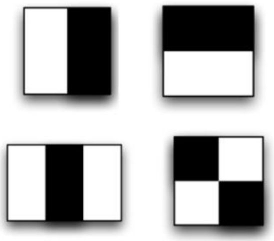
color histograms



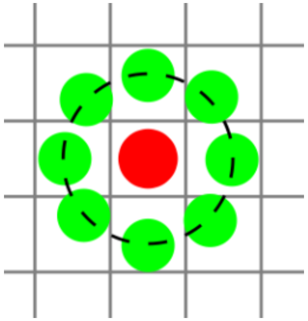
wavelet coefficients


SIFT


color-SIFT

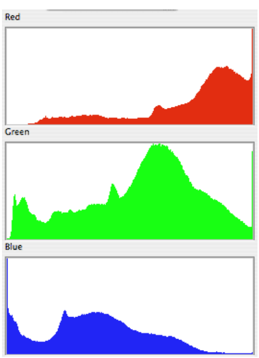

SURF


HOG


LBP

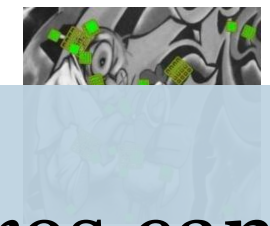
Review on Image Features

- How do we represent an image by a fixed-length feature vector?
  - Hand-designed features

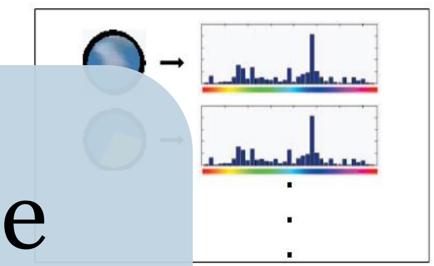


color histograms

wavelet coefficients



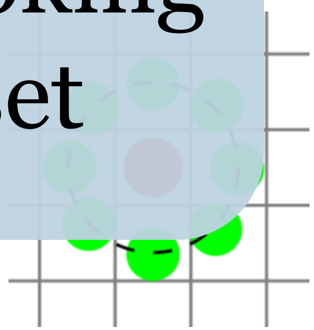
SIFT



color-SIFT

SURF

HOG



LBP

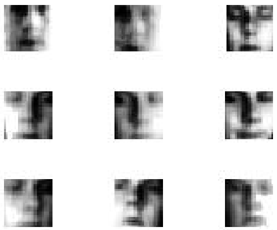
These features can be extracted without looking at the entire dataset

Review on Image Features

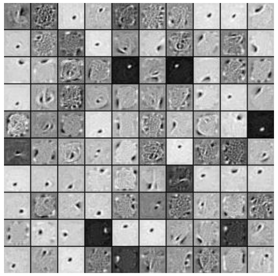
- How do we represent an image by a fixed-length feature vector?
  - Learning from a large dataset

Review on Image Features

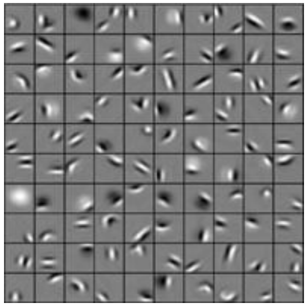
- How do we represent an image by a fixed-length feature vector?
  - Learning from a large dataset
    - Raw pixels based



PCA, kernel PCA



auto-encoders

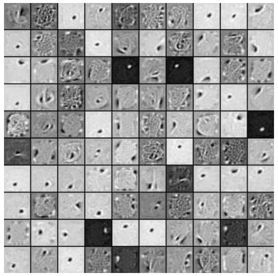


restricted Boltzmann machine (RBM)

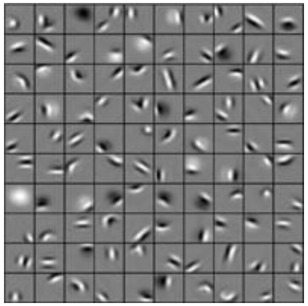
Review on Image Features

- How do we represent an image by a fixed-length feature vector?
  - Learning from a large dataset
    - Raw pixels based

PCA, kernel PCA

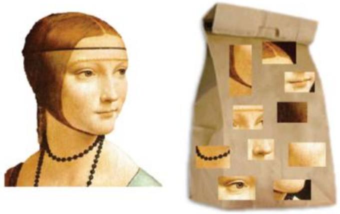


auto-encoders



restricted Boltzmann machine (RBM)

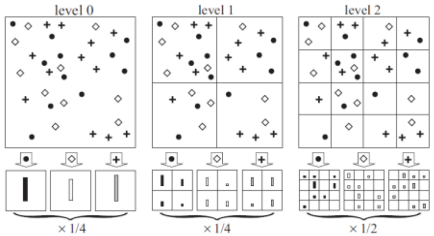
- Hand-designed feature based



bag-of-visual-words (BOV)

$$G_\lambda^X = \frac{1}{T}\sum_{t=1}^{T}\varphi_{FV}(x_t)$$

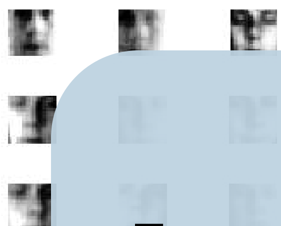
Fisher vector



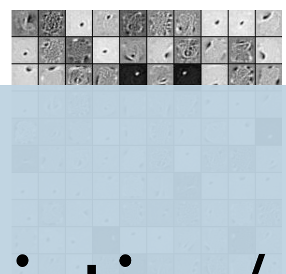
spatial pyramid matching (SPM)

Review on Image Features

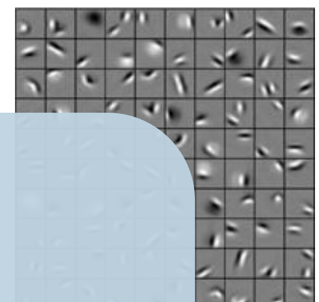
- How do we represent an image by a fixed-length feature vector?
  - Learning from a large dataset
    - Raw pixels based



PCA, kernel PCA



auto-encoder



Restricted Boltzmann machine (RBM)

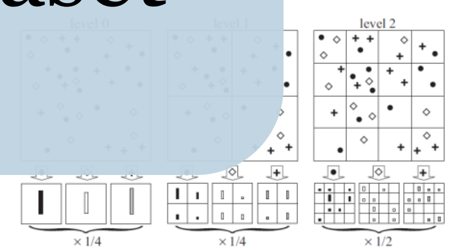
Learn statistics/patterns from the entire dataset

    - Hand-designed features based

bag-of-visual-words (BOV)

$$G_\lambda^X = \frac{1}{T}\sum_{t=1}^{T} \varphi_{FV}(x_t)$$

Fisher vector



spatial pyramid matching (SPM)

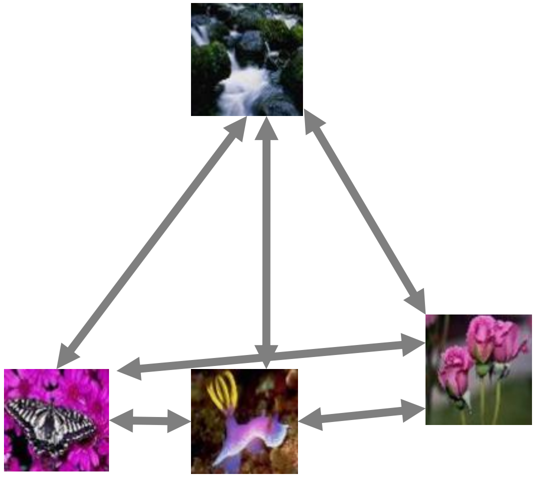
Other Possibilities?

- We already have:
  - Hand-designed features
  - Feature learning from raw pixels
  - Feature learning from hand-designed features

Other Possibilities?

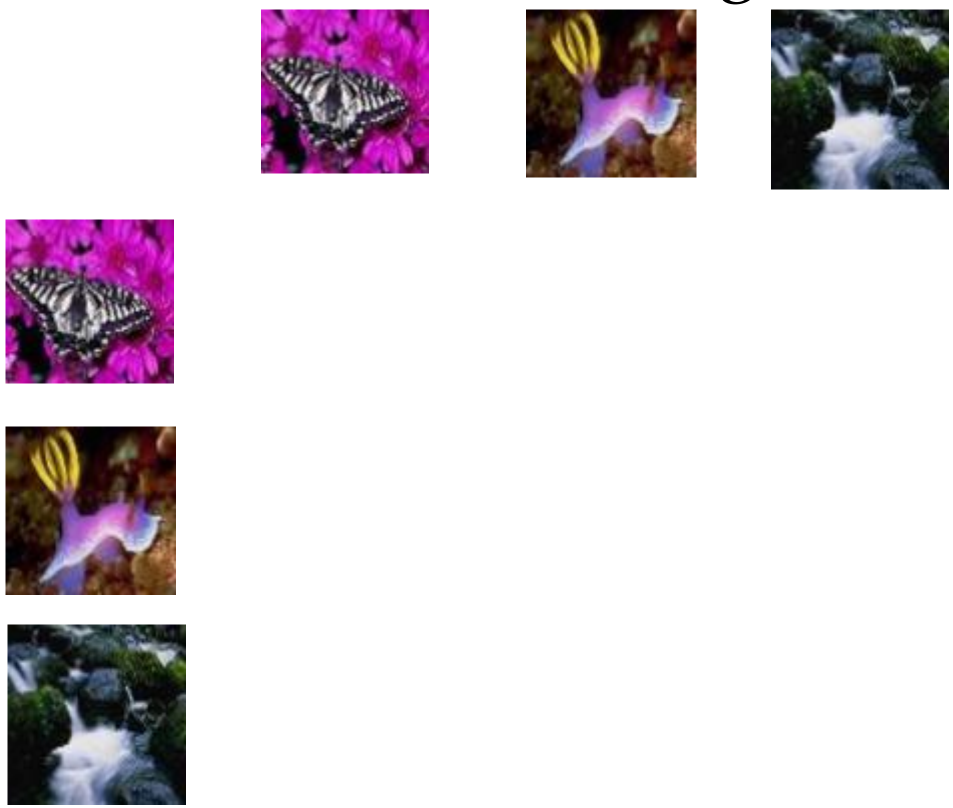
- We already have:
  - Hand-designed features
  - Feature learning from raw pixels
  - Feature learning from hand-designed features

- Now we propose a new group of features:
  - Feature learning from semantics-sensitive image distances



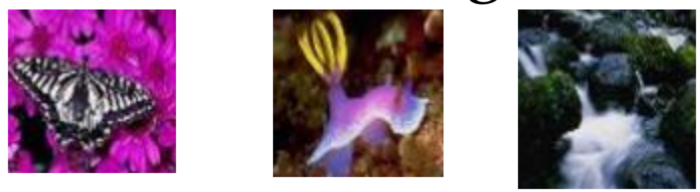
Basic Idea

1. Given a set of images, measure pair-wise semantics-sensitive image distances



Basic Idea

1. Given a set of images, measure pair-wise semantics-sensitive image distances



……

0	0.5	1.3
0.5	0	1.2
1.3	1.2	0

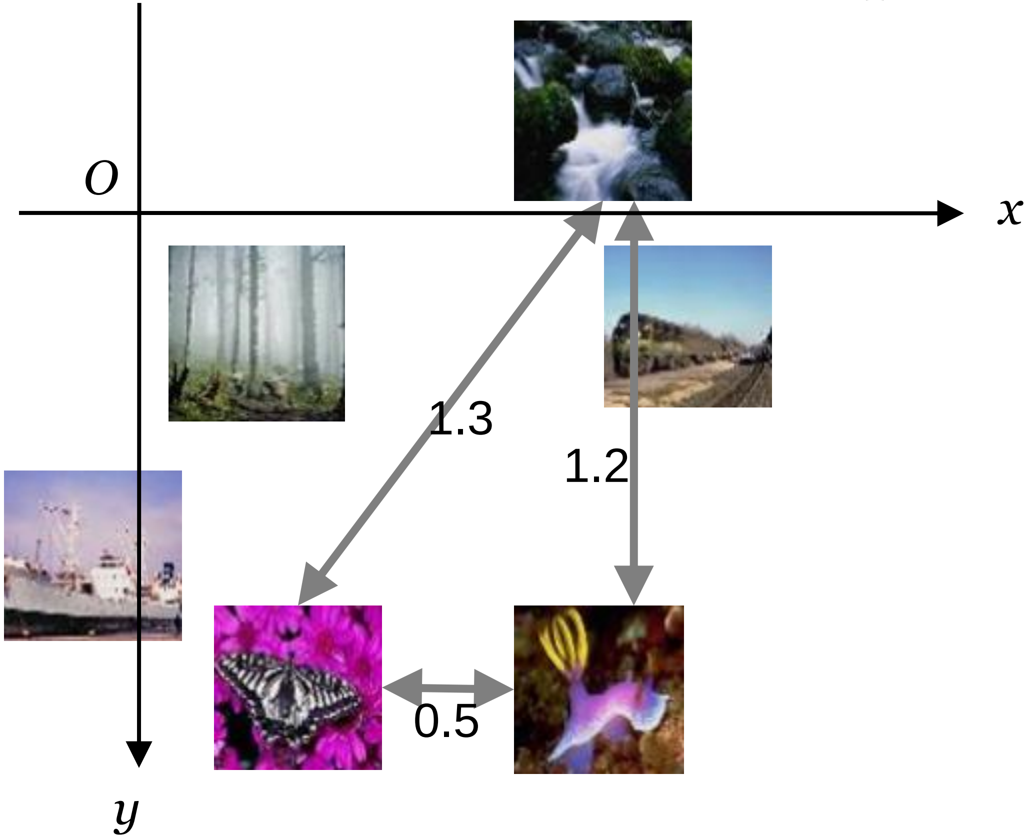
……

Basic Idea

2. Apply Multidimensional Scaling (MDS) on these distances, and encode all images in a low dimensional (*m*-d) Euclidean space

Basic Idea

2. Apply Multidimensional Scaling (MDS) on these distances, and encode all images in a low dimensional (*m*-d) Euclidean space



Basic Idea

3. We use the low dimensional representation of an image as its features, and call it *MDS codes*

Basic Idea

3. We use the low dimensional representation of an image as its features, and call it *MDS codes*
4. For an image classification application, MDS codes of training images are used to train classifiers

Basic Idea

3. We use the low dimensional representation of an image as its features, and call it *MDS codes*
4. For an image classification application, MDS codes of training images are used to train classifiers

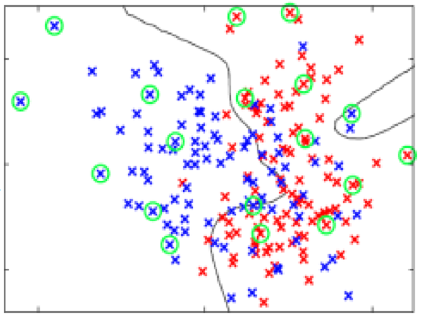
image1: $(-0.005, 0.000, \cdots)$

image2: $(-0.001, 0.013, \cdots)$

… … … …

MDS codes

SVM

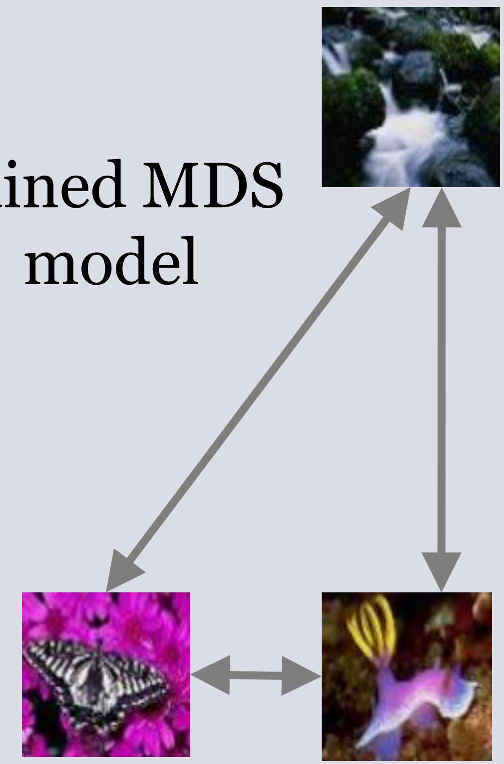


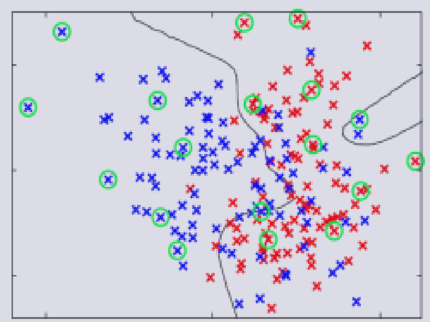
multi-class SVM classifier

Basic Idea

5. Given a new testing image, measure the distances from this image to training images to encode it, then apply trained classifiers

trained MDS model

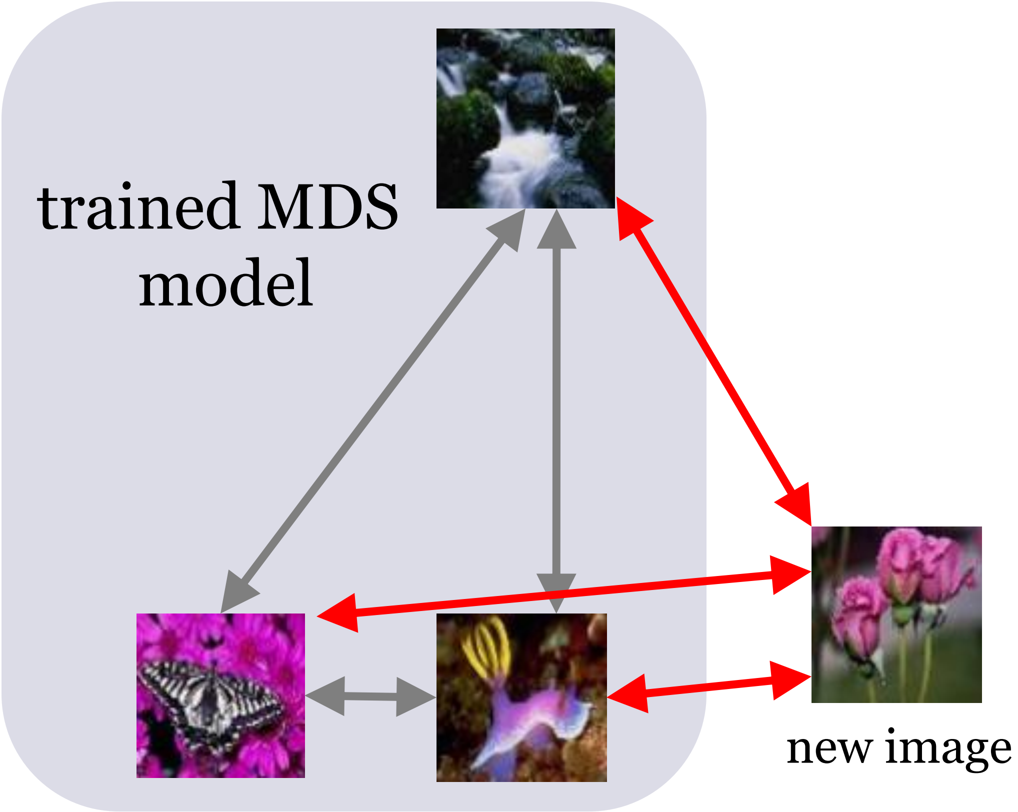


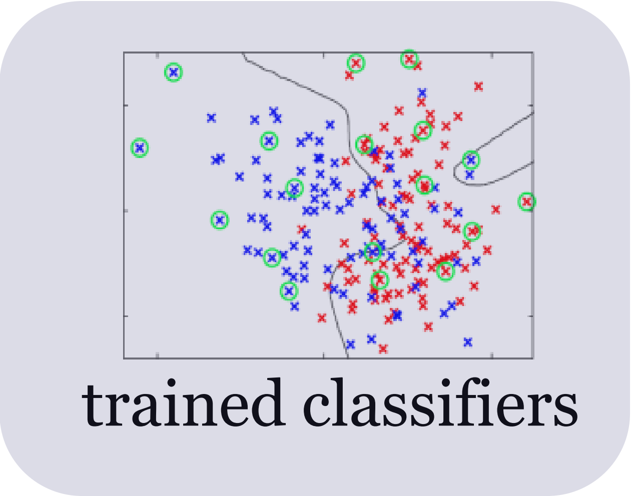


trained classifiers

Basic Idea

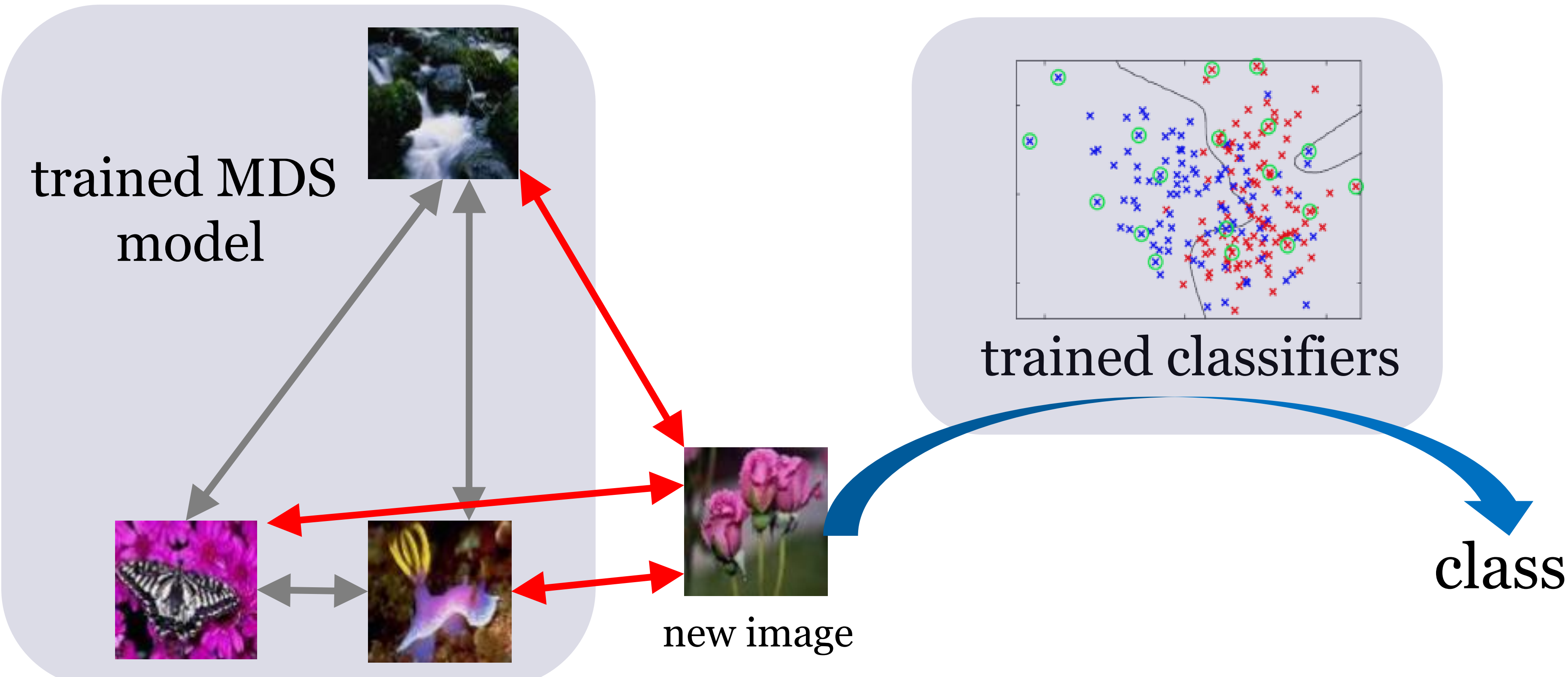
5. Given a new testing image, measure the distances from this image to training images to encode it, then apply trained classifiers





Basic Idea

5. Given a new testing image, measure the distances from this image to training images to encode it, then apply trained classifiers



Contributions of Our Work

1. Proposing the MDS feature learning scheme

Contributions of Our Work

1. Proposing the MDS feature learning scheme
2. The iterated Levenberg-Marquardt algorithm for efficient encoding

Contributions of Our Work

1. Proposing the MDS feature learning scheme

2. The iterated Levenberg-Marquardt algorithm for efficient encoding

3. Exploring MDS with different image distances:
   - IMage Euclidean Distance (IMED) [38]
   - Spatial Pyramid Matching (SPM) distance [24]
   - Integrated Region Matching (IRM) distance [48]

MDS Model Training: Concepts

- Given N images $\Omega = \{I_1, I_2, ..., I_N\}$

MDS Model Training: Concepts

- Given N images $\Omega = \{I_1, I_2, ..., I_N\}$
- Measure pair-wise image distances

$$d(I_i, I_j) : \Omega \times \Omega \to \mathbb{R}_{\geq 0}$$

MDS Model Training: Concepts

- Given N images $\Omega = \{I_1, I_2, ..., I_N\}$
- Measure pair-wise image distances

$$d(I_i, I_j) : \Omega \times \Omega \to \mathbb{R}_{\geq 0}$$

- Encode each image I_i to $\mathbf{x}_i \in \mathbb{R}^m$

MDS Model Training: Concepts

- Given N images $\Omega = \{I_1, I_2, ..., I_N\}$
- Measure pair-wise image distances

$$d(I_i, I_j) : \Omega \times \Omega \rightarrow \mathbb{R}_{\geq 0}$$

- Encode each image I_i to $\mathbf{x}_i \in \mathbb{R}^m$
- Representation error:

$$e_{ij} = d(I_i, I_j) - ||\mathbf{x}_i - \mathbf{x}_j||$$

MDS Model Training: Concepts

- Given N images $\Omega = \{I_1, I_2, ..., I_N\}$
- Measure pair-wise image distances

$$d(I_i, I_j) : \Omega \times \Omega \to \mathbb{R}_{\geq 0}$$

- Encode each image I_i to $\mathbf{x}_i \in \mathbb{R}^m$
- Representation error:

$$e_{ij} = d(I_i, I_j) - ||\mathbf{x}_i - \mathbf{x}_j||$$

- Raw stress (total cost to be minimized) [26]:

$$\text{Stress}^* = \sum_{1 \leq i < j \leq N} e_{ij}^2$$

MDS Model Training: Concepts

- Problem:

$$\mathbf{X}^* = \arg\min_{\mathbf{X}} \sum_{1 \le i < j \le N} \left(d(I_i, I_j) - ||\mathbf{x}_i - \mathbf{x}_j|| \right)^2$$

where $\mathbf{X} = (\mathbf{x}_1, \ldots, \mathbf{x}_N)^\mathsf{T}$

- Existing methods:
  - Iterative majorization algorithm (SMACOF) [30]
  - Variants of SMACOF [40]-[43]
- Our solution:
  - Iterated Levenberg-Marquardt algorithm (ILMA)

MDS Model Training: Concepts

The classical MDS problem

- Problem:

$$\mathbf{X}^* = \arg\min_{\mathbf{X}} \sum_{1 \le i < j \le N} \left(d(I_i, I_j) - \|\mathbf{x}_i - \mathbf{x}_j\| \right)^2$$

where $\mathbf{X} = (\mathbf{x}_1, \ldots, \mathbf{x}_N)^\mathsf{T}$

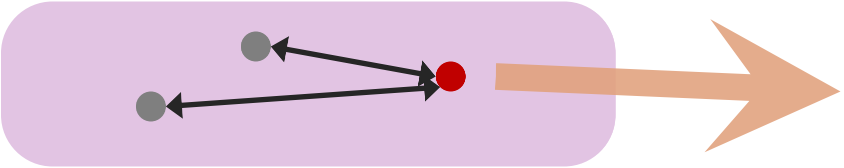
- Existing methods:
  - Iterative majorization algorithm (SMACOF) [30]
  - Variants of SMACOF [40]-[43]
- Our solution:
  - Iterated Levenberg-Marquardt algorithm (ILMA)

MDS Model Training: Algorithm

- Basic idea of ILMA:
  - Each time fix all $\mathbf{x}_i$'s except for one, and apply the standard Levenberg-Marquardt algorithm [33][34]

MDS Model Training: Algorithm

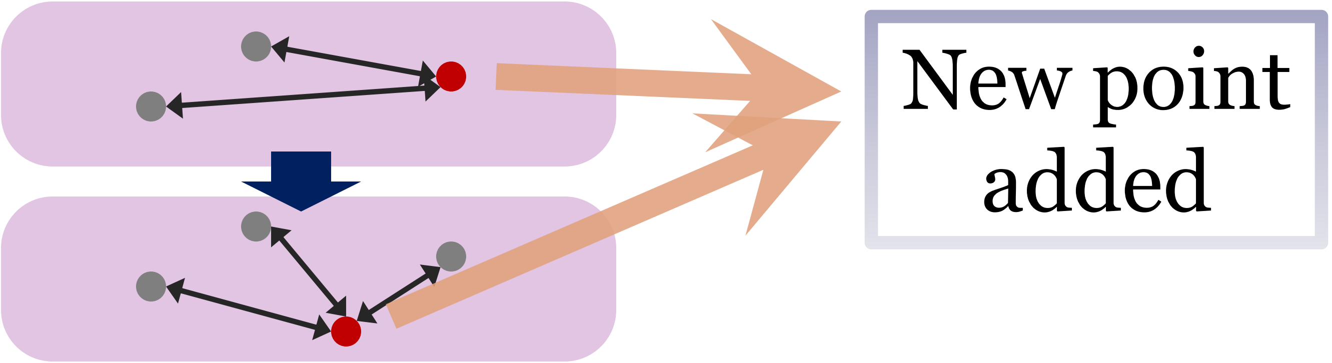
- Basic idea of ILMA:
  - Each time fix all $\mathbf{x}_i$'s except for one, and apply the standard Levenberg-Marquardt algorithm [33][34]



New point added

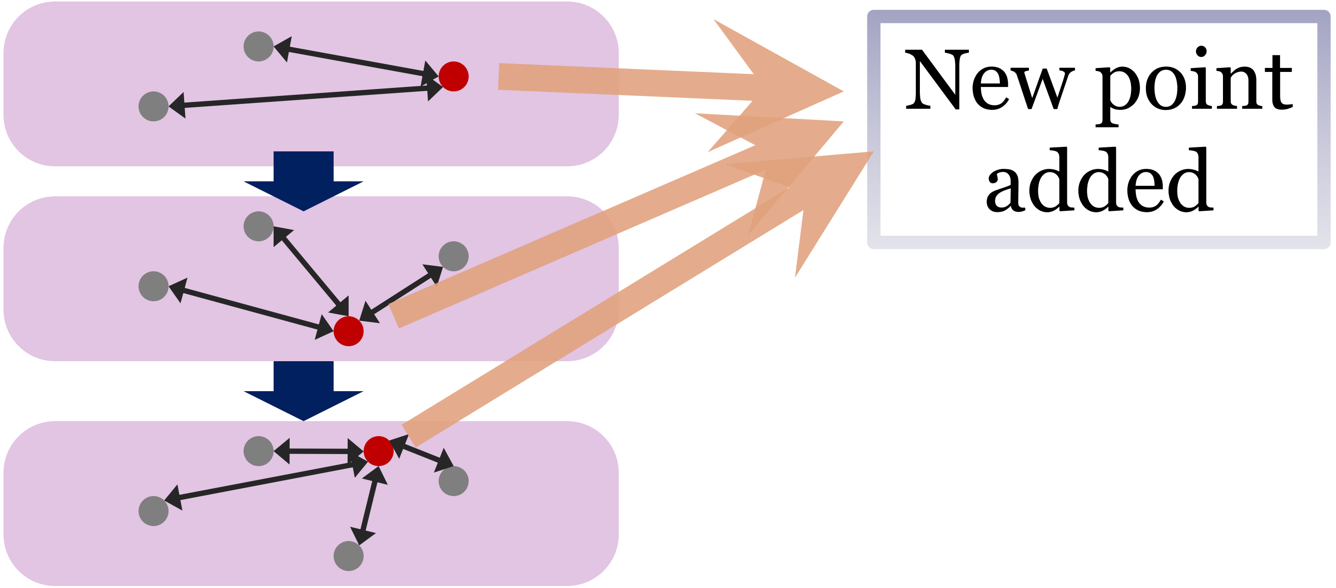
MDS Model Training: Algorithm

- Basic idea of ILMA:
  - Each time fix all $\mathbf{x}_i$'s except for one, and apply the standard Levenberg-Marquardt algorithm [33][34]



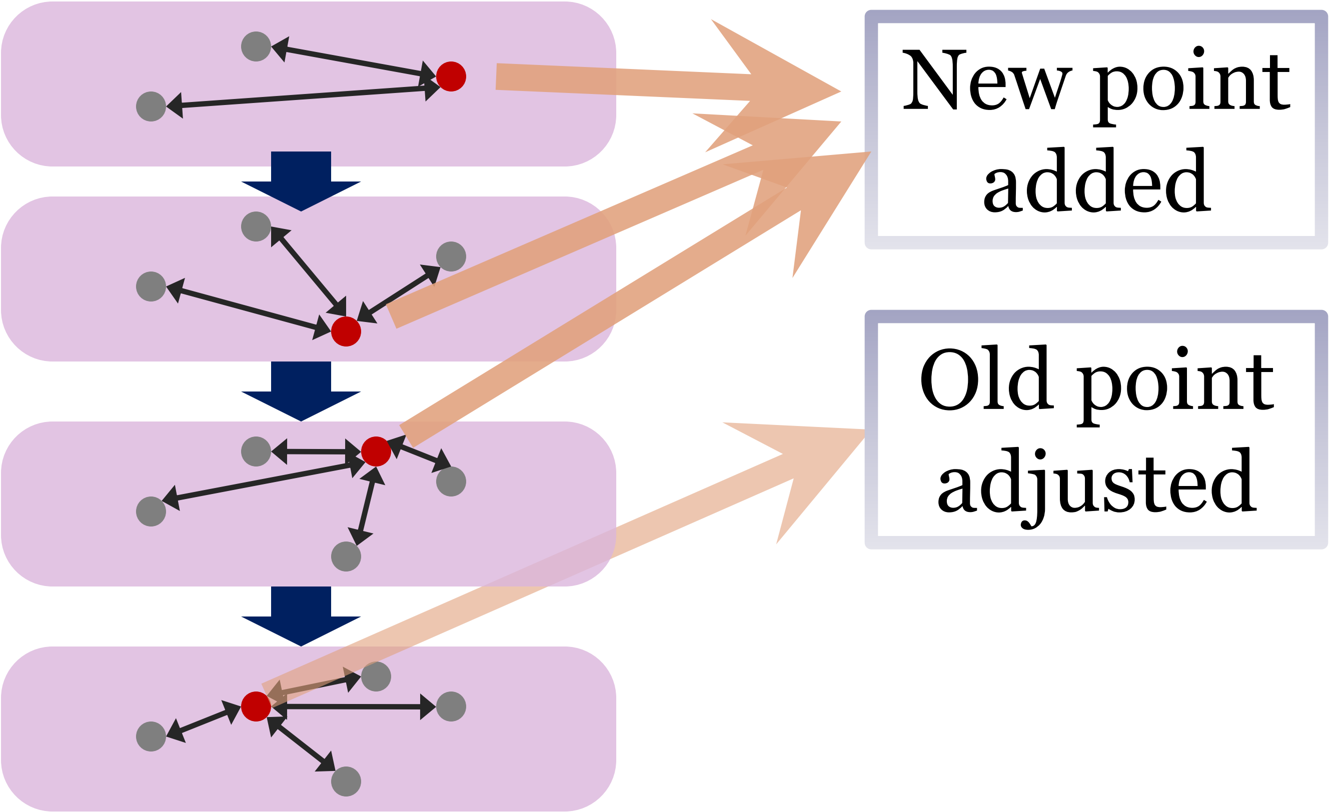
MDS Model Training: Algorithm

- Basic idea of ILMA:
  - Each time fix all $\mathbf{x}_i$'s except for one, and apply the standard Levenberg-Marquardt algorithm [33][34]



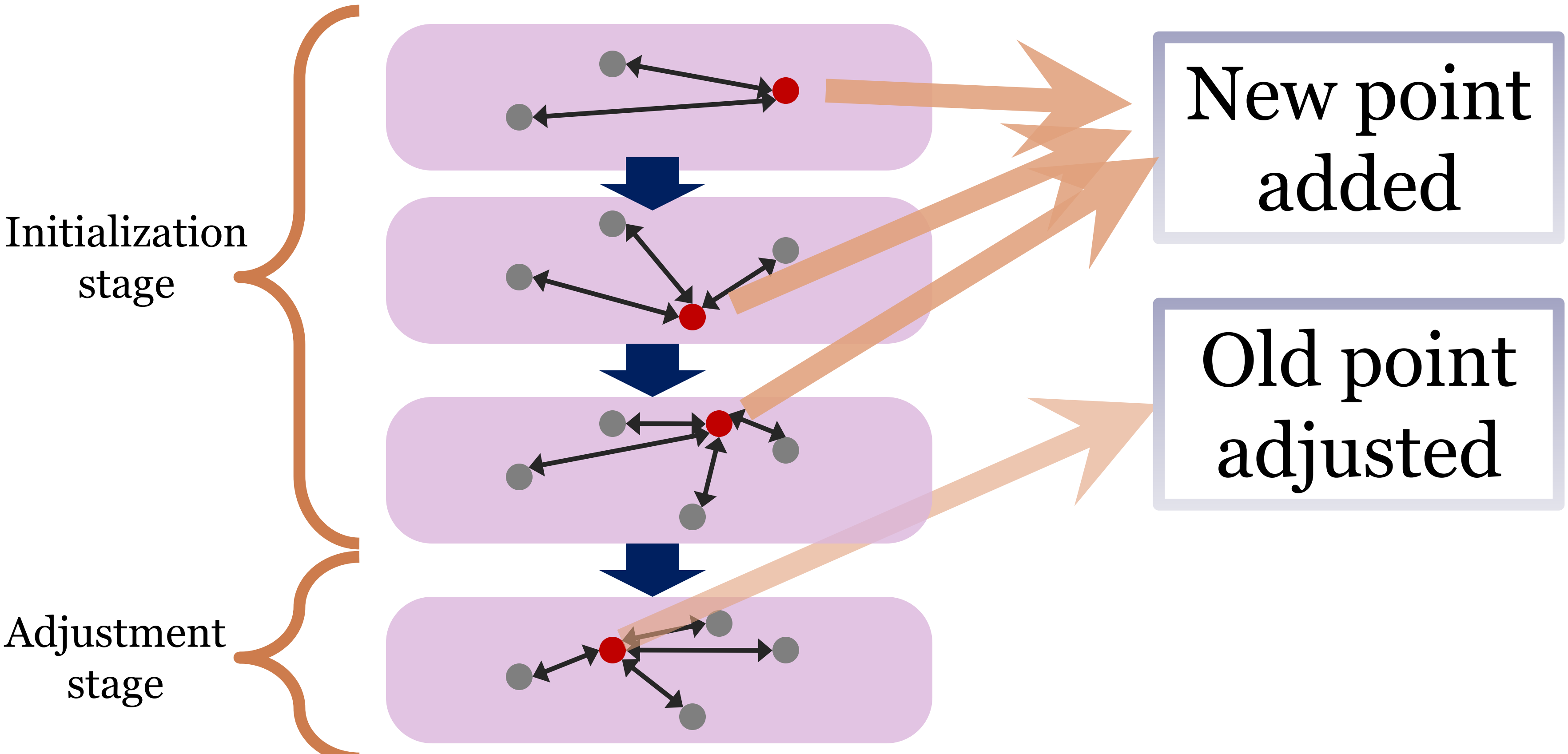
MDS Model Training: Algorithm

- Basic idea of ILMA:
  - Each time fix all $\mathbf{x}_i$'s except for one, and apply the standard Levenberg-Marquardt algorithm [33][34]



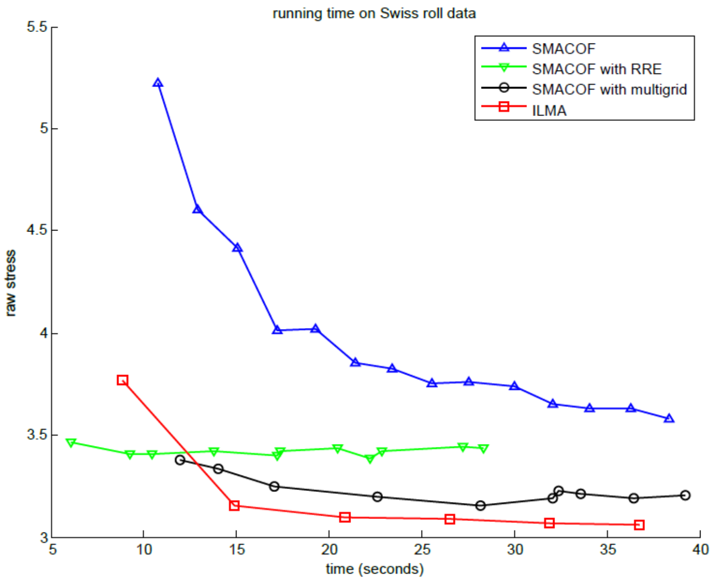
MDS Model Training: Algorithm

- Basic idea of ILMA:
  - Each time fix all $\mathbf{x}_i$'s except for one, and apply the standard Levenberg-Marquardt algorithm [33][34]

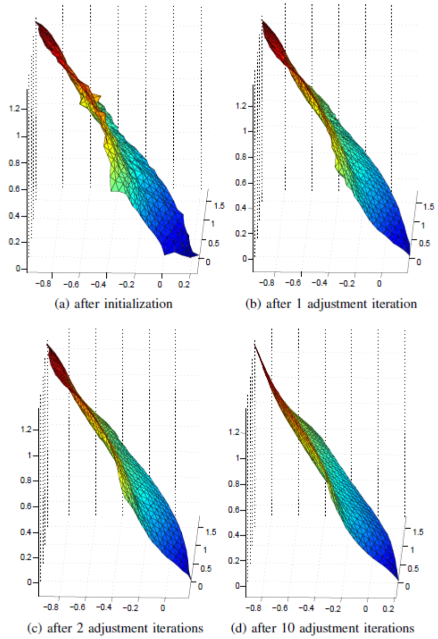


MDS Model Training: Algorithm

- Please see our paper for algorithm details
- ILMA runs faster and converges to smaller raw stress than many other solutions
  - Tested on the well-known Swiss roll flattening experiment [39][40]



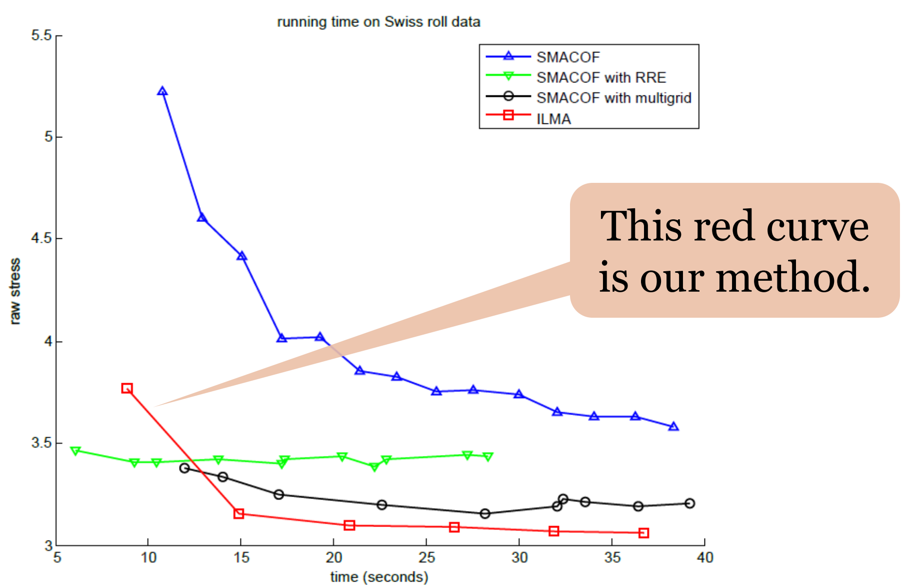
Running time comparison with other solutions



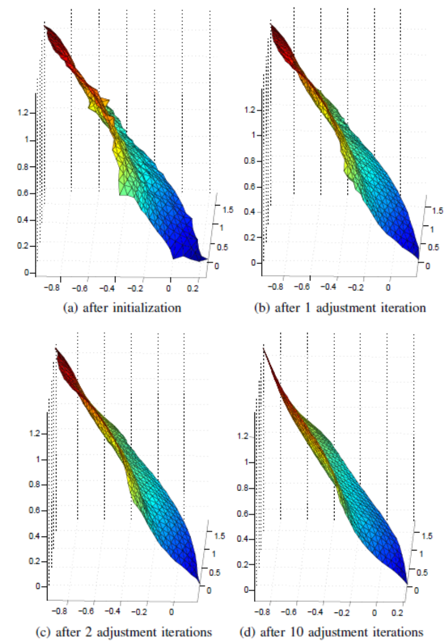
Swiss roll flattening results

MDS Model Training: Algorithm

- Please see our paper for algorithm details
- ILMA runs faster and converges to smaller raw stress than many other solutions
  - Tested on the well-known Swiss roll flattening experiment [39][40]



Running time comparison with other solutions



Swiss roll flattening results

New Image Encoding

- After MDS model training, we need to be able to encode a new image $\tilde{I}$ to $\widetilde{\mathbf{x}}$:

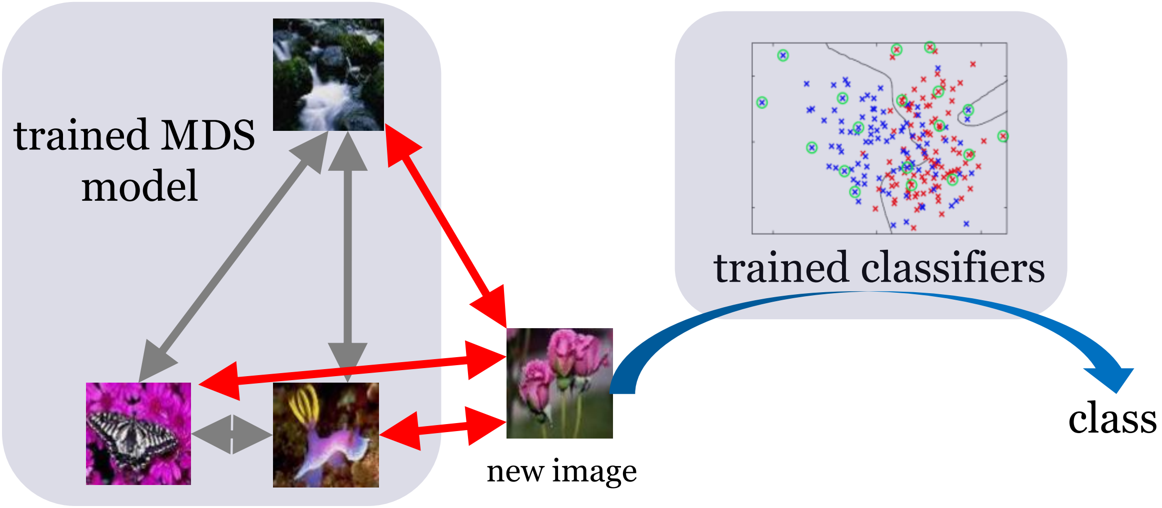
$$\min_{\widetilde{\mathbf{x}}} \sum_{I_i \in \Omega_{\mathrm{train}}} \left(||\widetilde{\mathbf{x}} - \mathbf{x}_i|| - d(\widetilde{I}, I_i) \right)^2$$

New Image Encoding

- After MDS model training, we need to be able to encode a new image $\tilde{I}$ to $\tilde{\mathbf{x}}$:

$$\min_{\tilde{\mathbf{x}}} \sum_{I_i \in \Omega_{\mathtt{train}}} \left(\|\tilde{\mathbf{x}} - \mathbf{x}_i\| - d(\tilde{I}, I_i) \right)^2$$

- This can be solved by the standard Levenberg-Marquardt algorithm [33][34]



Semantics-Sensitive Image Distances

- We have talked about:
  - Feature learning
  - MDS model training
  - Encoding new image
- One thing left – what are we talking about when we say "distance"?
  - A metric on set $\Omega=(I_1,I_2,\ldots,I_N)$ in the strict sense?

Semantics-Sensitive Image Distances

- We have talked about:
  - Feature learning
  - MDS model training
  - Encoding new image
- One thing left – what are we talking about when we say "distance"?
  - A metric on set $\Omega=(I_1,I_2,\ldots,I_N)$ in the strict sense?
    - **No**, subadditivity triangle inequality does not necessarily hold.

$$d(x, z) \leq d(x, y) + d(y, z)$$

Semantics-Sensitive Image Distances

- We have talked about:
  - Feature learning
  - MDS model training
  - Encoding new image
- One thing left – what are we talking about when we say "distance"?
  - A metric on set $\Omega=(I_1,I_2,\ldots,I_N)$ in the strict sense?
    - **No**, subadditivity triangle inequality does not necessarily hold.

$$\cancel{d(x, z) \leq d(x, y) + d(y, z)}$$

  - Just a dissimilarity measurement!

Semantics-Sensitive Image Distances

- IMage Euclidean Distance (IMED) [38]
  - Traditional pixel-wise Euclidean distance on a smoothed version of two images
  - Low level, not much semantics information
- Spatial pyramid matching (SPM) distance [24]
  - Based on pyramid matching kernel [25]
  - Well applied to image classification
  - Highly semantics-sensitive
- Integrated region matching (IRM) distance [48]
  - Well applied to content-based image retrieval (CBIR)
  - Highly semantics-sensitive

- We will skip the descriptions of these image distances
- Please refer to original work [38][24][48]

Experiment 1: Car Recognition

- **Data:** UIUC car dataset [44]
  - (550 car, 500 non-car)

- **Task:** 2-way (binary) classification on fixed-length feature vectors

- **Classifier:** RBF kernel SVM [46][47]

- **Validation:** 5-fold cross validation

Experiment 1: Car Recognition

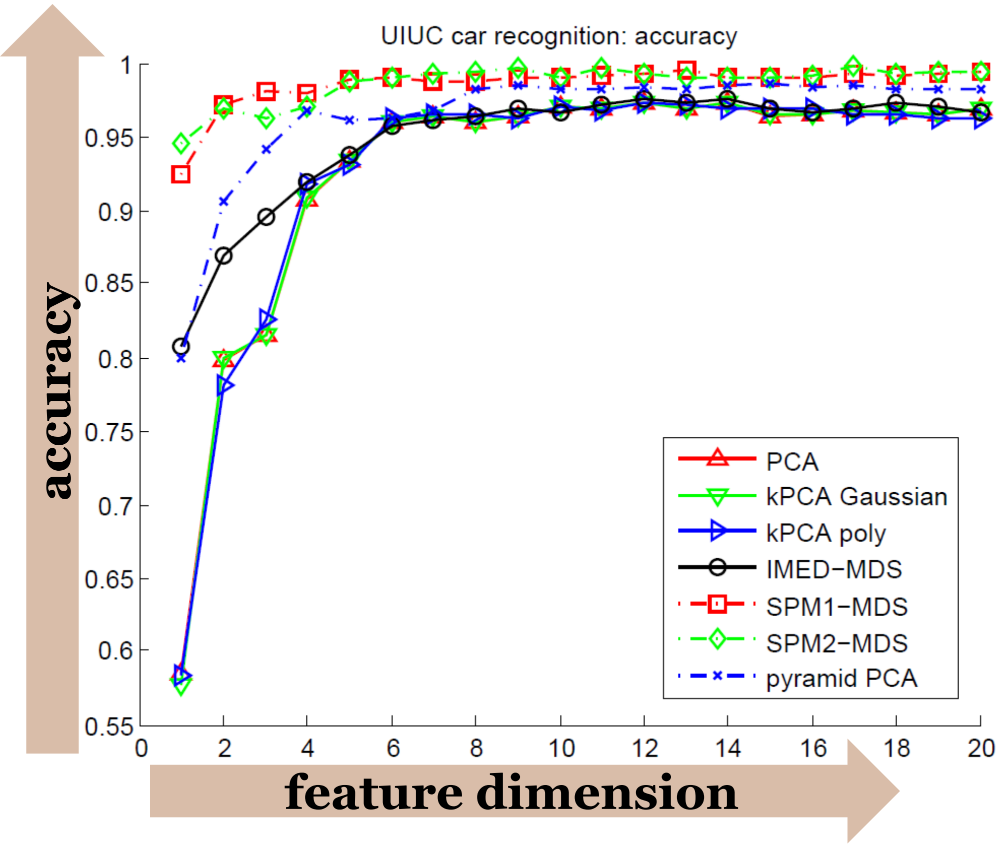
- Methods employed:
  1. PCA/kernel PCA on raw intensities
  2. MDS on IMage Euclidean Distances (IMED-MDS)
  3. MDS on Spatial Pyramid Matching distances (SPM-MDS)
  4. PCA on spatial pyramid vectors

Experiment 1: Car Recognition

- Methods employed:
  1. PCA/kernel PCA on raw intensities
  2. MDS on IMage Euclidean Distances (IMED-MDS)
  3. MDS on Spatial Pyramid Matching distances (SPM-MDS)
  4. PCA on spatial pyramid vectors

Make this comparison to see if MDS brings new information beyond pyramid kernels

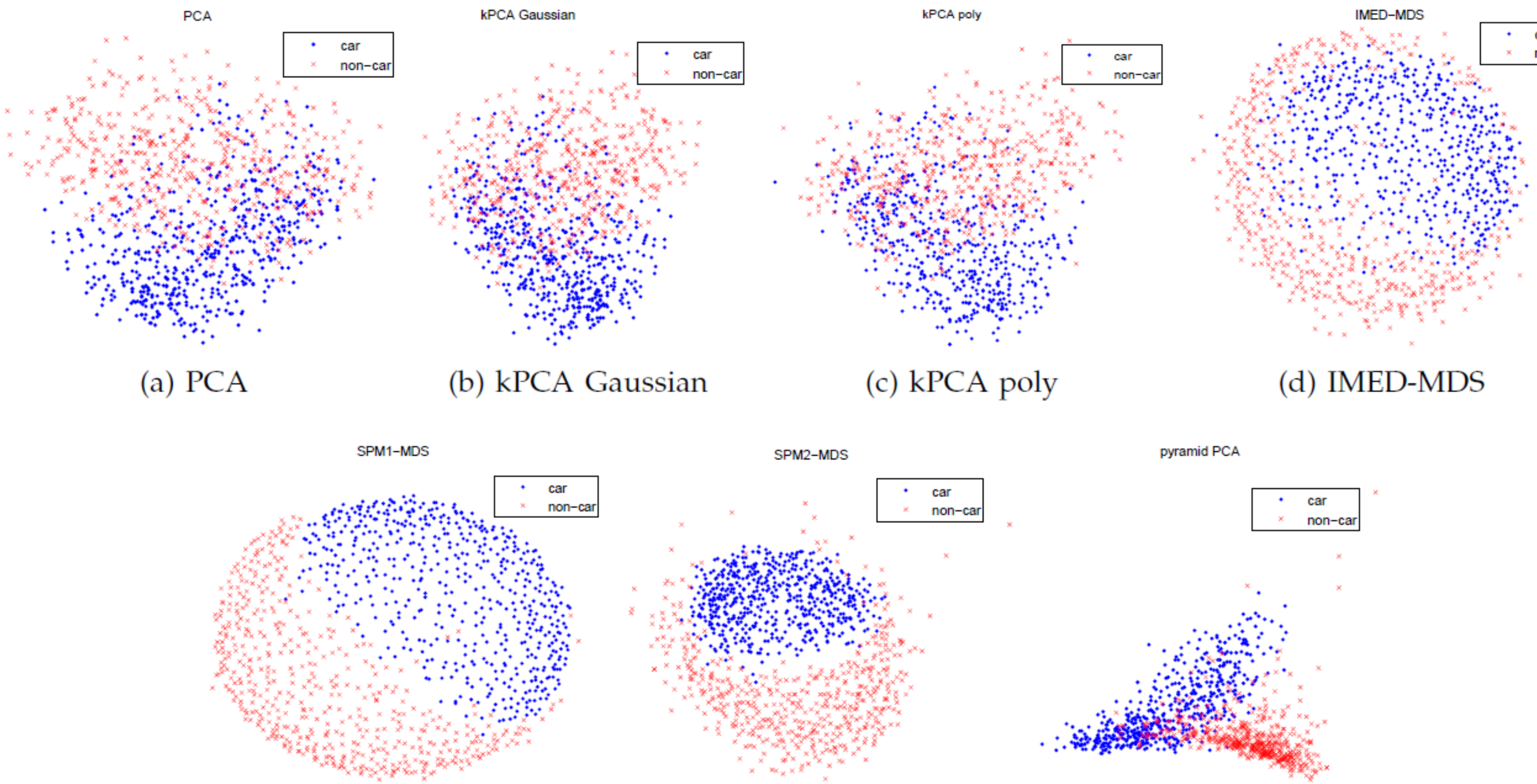
Experiment 1: Car Recognition



- SPM-MDS outperforms all other methods
- SPM-MDS wins pyramid PCA, thus **MDS encodes semantics beyond pyramid kernels**
- IMED-MDS wins PCA / kernel PCA
- (Note: SPM1-MDS and SPM2-MDS use different scaling functions)

Experiment 1: Car Recognition

- 2-d feature scattering plot

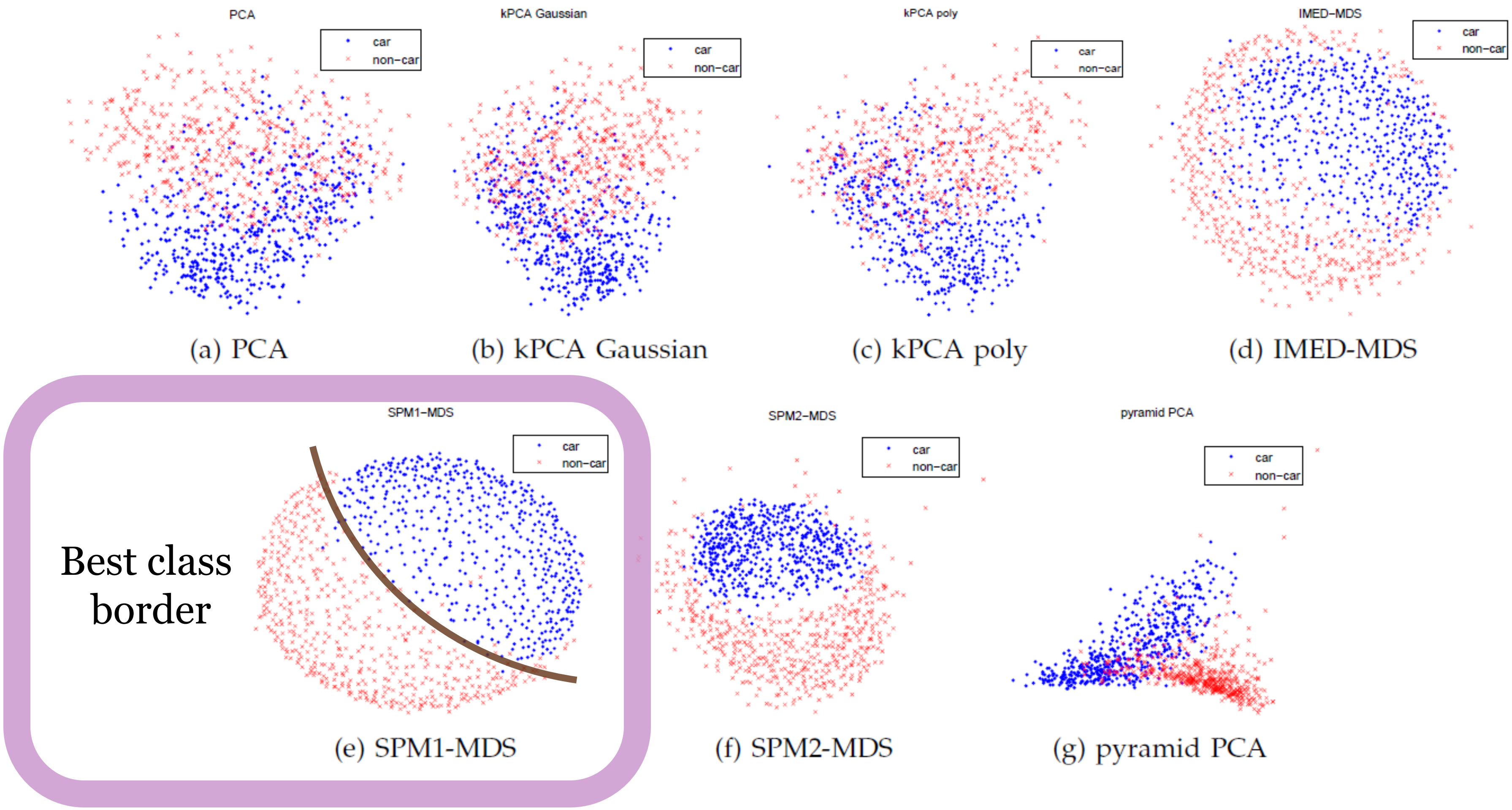


(a) PCA (b) kPCA Gaussian (c) kPCA poly (d) IMED-MDS

(e) SPM1-MDS (f) SPM2-MDS (g) pyramid PCA

Experiment 1: Car Recognition

- 2-d feature scattering plot



(a) PCA (b) kPCA Gaussian (c) kPCA poly (d) IMED-MDS

(e) SPM1-MDS (f) SPM2-MDS (g) pyramid PCA

Experiment 2: Multi-Class Object Recognition

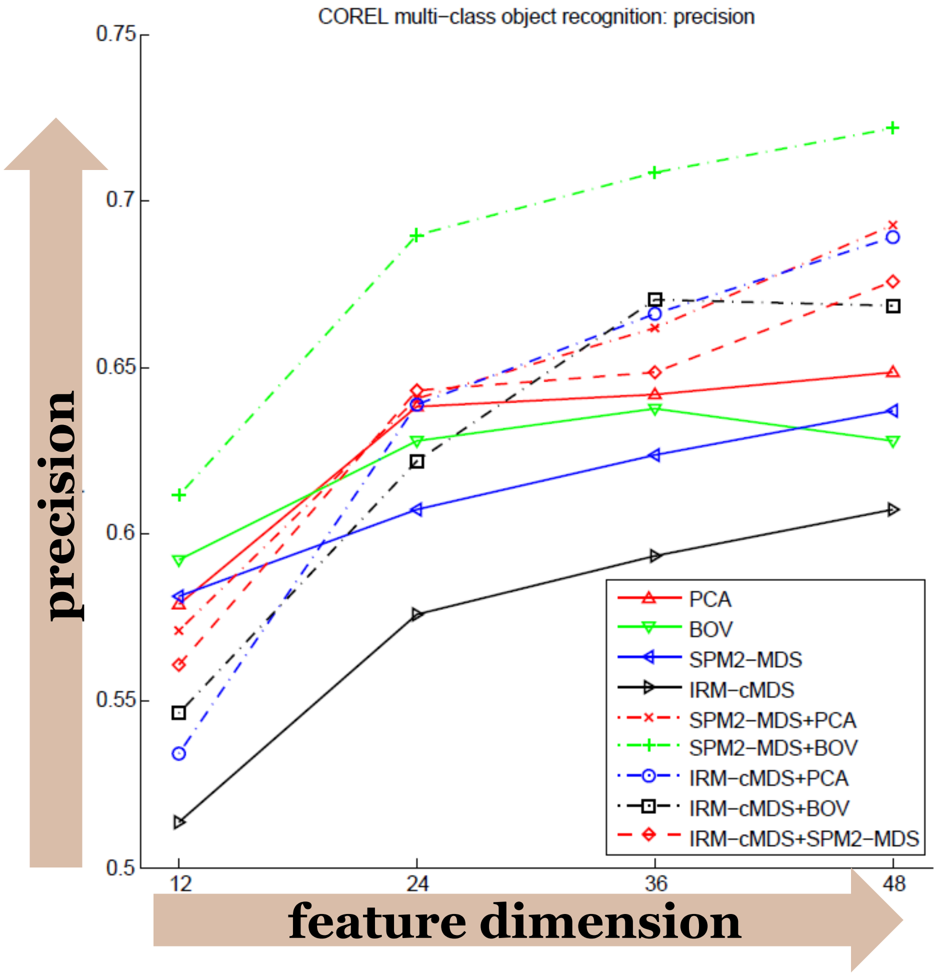
- **Data:** 12 categories from COREL dataset [48][50]
- **Task:** 12-way classification on fixed-length feature vectors
- **Classifier:** RBF kernel SVM [46][47]
- **Validation:** 5-fold cross validation

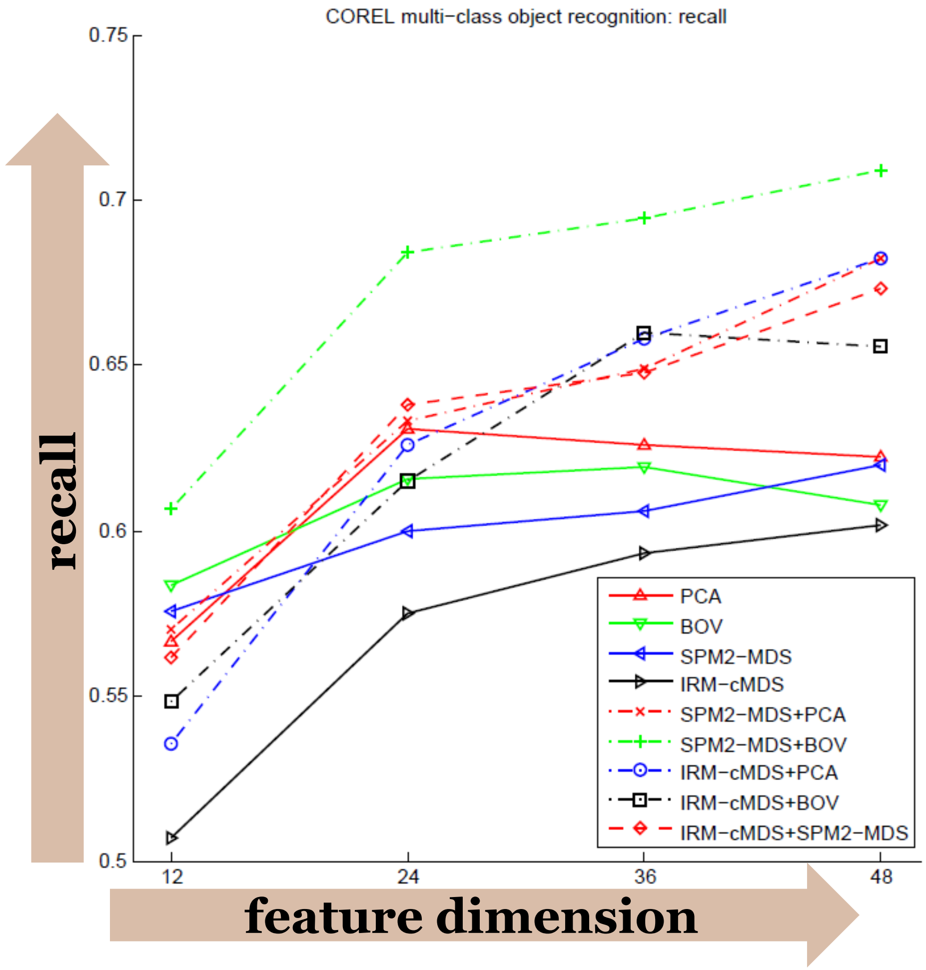
Experiment 2: Multi-Class Object Recognition

- Methods employed:
  1. PCA on color+HOG+LBP [5][6]
  2. Bag-of-visual-words (BOV) [19]-[21] on color+HOG+LBP
  3. MDS on SPM distances (SPM-MDS)
  4. MDS on IRM distances (IRM-MDS)
  5. Combined features
     - *E.g.* 12-d PCA + 12-d BOV = 24-d combined

- For method details, see our paper and references

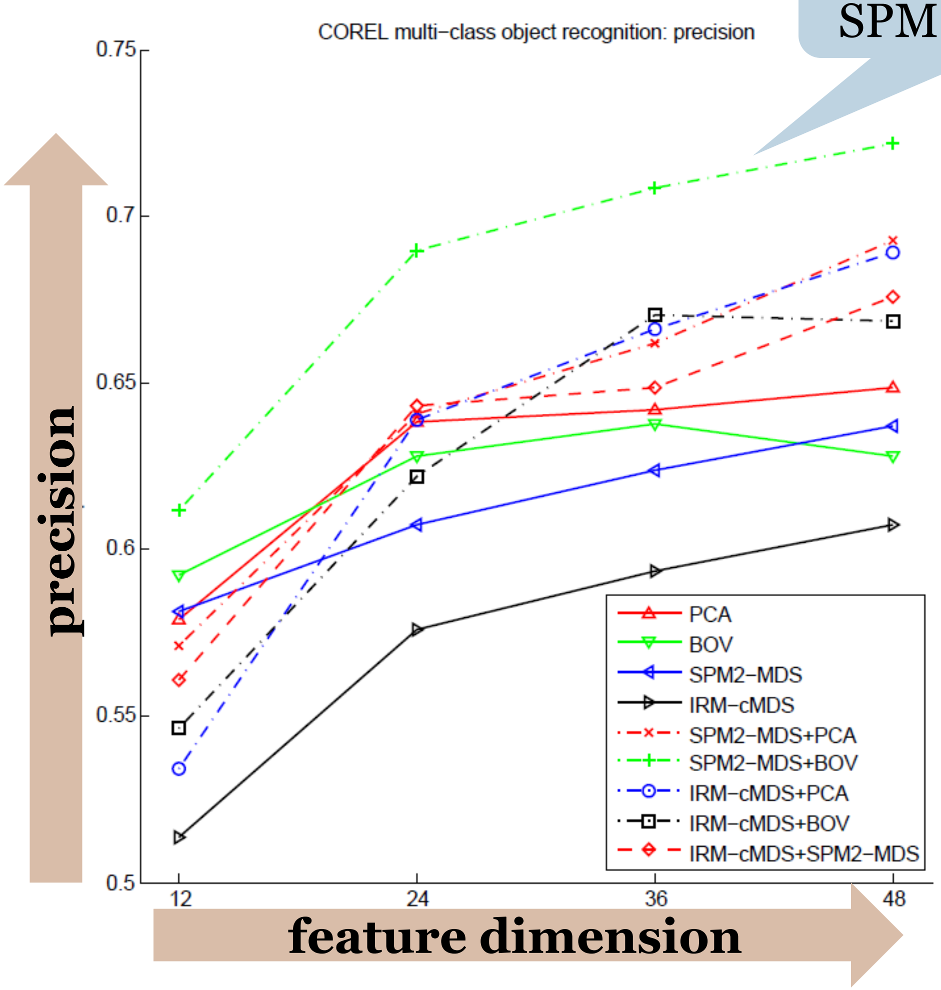
Experiment 2: Multi-Class Object Recognition

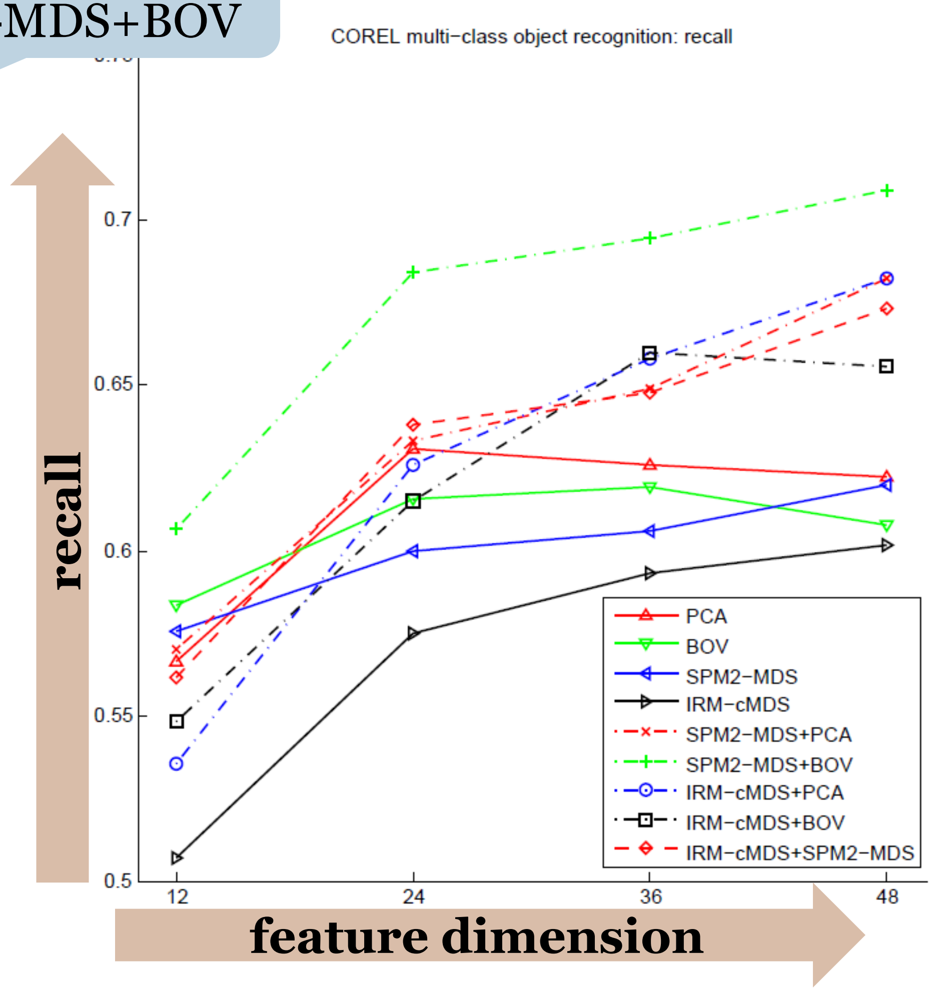




Experiment 2: Multi-Class Object Recognition

Winner is SPM-MDS+BOV





Experiment 2: Multi-Class Object Recognition

- Conclusions:
  - Without combining, PCA and bag-of-visual-words on color+HOG+LBP outperform MDS features 

Experiment 2: Multi-Class Object Recognition

- Conclusions:
  - Without combining, PCA and bag-of-visual-words on color+HOG+LBP outperform MDS features 
  - SPM-MDS combined with bag-of-visual-words significantly outperforms all other methods 

Experiment 2: Multi-Class Object Recognition

- Conclusions:
  - Without combining, PCA and bag-of-visual-words on color+HOG+LBP outperform MDS features
  - SPM-MDS combined with bag-of-visual-words significantly outperforms all other methods
  - We conclude: **SPM-MDS features capture semantics information from images that are not captured by other methods, such as PCA or bag-of-visual-words**

Experiment 2: Multi-Class Object Recognition

- Classification confusion matrix of SPM-MDS + BOV

	castle	bonsai	ship	train	flower	mushroom	forests	waterfall	butterfly	fish	wolf	woman
castle	67.0	0.0	2.6	12.6	0.0	2.6	1.0	2.0	0.0	0.0	1.7	1.8
bonsai	3.2	91.5	0.0	0.0	0.0	0.7	0.0	0.0	0.0	4.0	2.2	0.0
ship	2.4	1.8	88.8	2.5	0.0	0.0	0.8	0.0	0.0	0.0	1.4	0.0
train	11.6	0.9	3.4	80.1	0.0	0.0	0.8	6.0	0.0	0.9	2.2	0.0
flower	0.0	0.0	0.0	0.0	65.6	7.0	4.0	0.7	13.9	7.7	0.0	6.6
mushroom	0.0	0.0	0.0	0.0	4.1	54.8	3.5	3.9	12.9	9.3	3.3	1.6
forests	10.2	1.0	1.2	1.1	1.8	2.9	79.1	1.0	6.2	1.1	1.7	0.0
waterfall	2.3	0.0	2.2	0.0	0.0	4.3	1.0	72.3	2.6	4.4	10.5	4.1
butterfly	1.4	0.0	0.0	3.8	11.3	9.6	5.1	0.7	49.9	11.8	2.0	2.6
fish	0.0	2.0	0.8	0.0	12.2	8.0	1.2	4.2	9.2	59.9	0.0	0.0
wolf	1.8	1.9	1.0	0.0	0.0	6.6	2.8	7.8	1.0	0.9	74.9	1.0
woman	0.0	0.9	0.0	0.0	5.1	3.6	0.8	1.5	4.4	0.0	0.0	82.3

- SPM-MDS running time:
    - Training MDS model on 960 images takes **20 min**
    - Encoding one new image takes **0.3 s** (including feature extraction)

Final Discussions

- We proposed a new feature learning scheme:
  - Feature learning from semantics-sensitive image distances

Final Discussions

- We proposed a new feature learning scheme:
  - Feature learning from semantics-sensitive image distances
- It works well for object recognition tasks:
  - MDS on IMED, SPM distances, IRM distances
  - Experiments with UIUC car data, COREL images

Final Discussions

- We proposed a new feature learning scheme:
  - Feature learning from semantics-sensitive image distances
- It works well for object recognition tasks:
  - MDS on IMED, SPM distances, IRM distances
  - Experiments with UIUC car data, COREL images
- We also introduced an efficient MDS algorithm:
  - Iterated Levenberg-Marquardt algorithm (ILMA)
  - Code can be downloaded at: https://sites.google.com/site/mdsfeature/

Final Discussions

- There is more to do:
  - MDS model training on large dataset is still slow. Can we make it parallel? (on multi-core, GPU, …)
  - Are there other semantics-sensitive image distances that can be employed?
  - Other applications apart from object recognition?
    - Style classification
    - Affective image classification
    - Aesthetics analysis
    - Emotion/sentiment detection
    - Face beautification rating
    - ……

Final Discussions

- There is more to do:
  - MDS model training on large dataset is still slow. Can we make it parallel? (on multi-core, GPU, …)
  - Are there other semantics-sensitive image distances that can be employed?
  - Other applications apart from object recognition?
    - Style classification
    - Affective image classification
    - Aesthetics analysis
    - Emotion/sentiment detection
    - Face beautification rating
    - ……

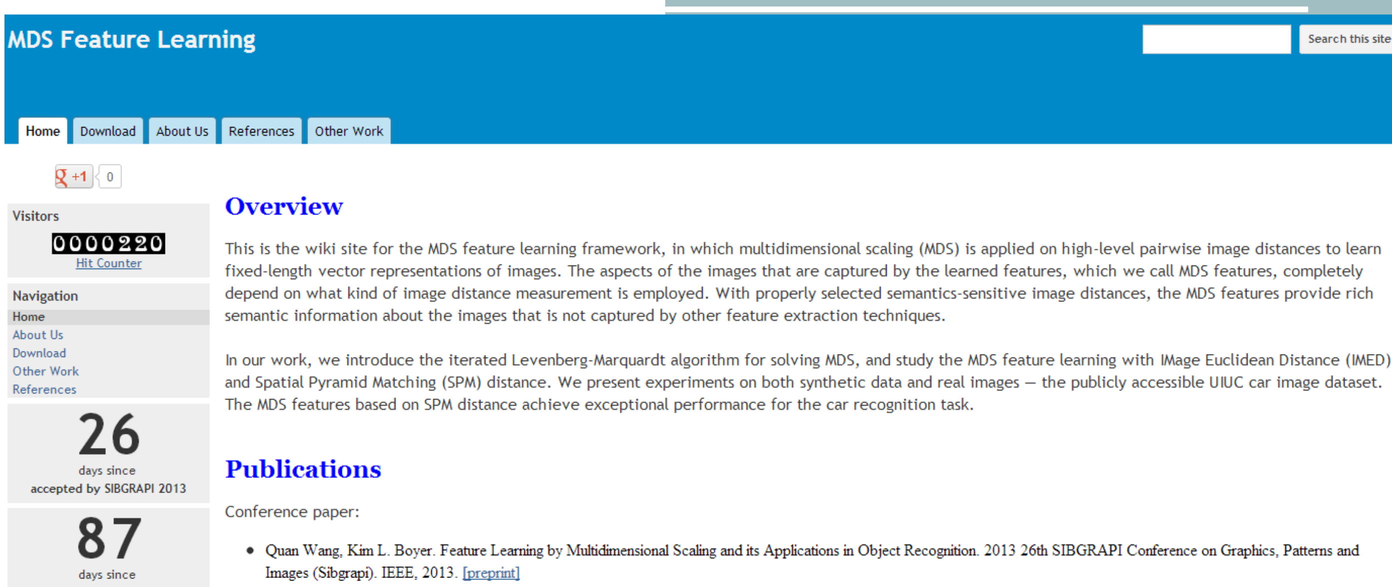
All depends on how you can define a distance!

Electrical, Computer, and Systems Engineering at Rensselaer

Thank you for your interest!

Welcome to our project wiki site:

https://sites.google.com/site/mdsfeature/



MDS Feature Learning

Search this site

Home | Download | About Us | References | Other Work

+1 0

Visitors

0000220

Hit Counter

Navigation

Home

About Us

Download

Other Work

References

26

days since

accepted by SIBGRAPI 2013

87

days since

submission to SIBGRAPI 2013

Overview

This is the wiki site for the MDS feature learning framework, in which multidimensional scaling (MDS) is applied on high-level pairwise image distances to learn fixed-length vector representations of images. The aspects of the images that are captured by the learned features, which we call MDS features, completely depend on what kind of image distance measurement is employed. With properly selected semantics-sensitive image distances, the MDS features provide rich semantic information about the images that is not captured by other feature extraction techniques.

In our work, we introduce the iterated Levenberg-Marquardt algorithm for solving MDS, and study the MDS feature learning with IMage Euclidean Distance (IMED) and Spatial Pyramid Matching (SPM) distance. We present experiments on both synthetic data and real images — the publicly accessible UIUC car image dataset. The MDS features based on SPM distance achieve exceptional performance for the car recognition task.

Publications

Conference paper:

- Quan Wang, Kim L. Boyer. Feature Learning by Multidimensional Scaling and its Applications in Object Recognition. 2013 26th SIBGRAPI Conference on Graphics, Patterns and Images (Sibgrapi). IEEE, 2013. [preprint]

References

- [1] I. Daubechies et al., Ten lectures on wavelets. SIAM, 1992, vol. 61.
- [2] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," International Journal of Computer Vision, vol. 60, no. 2, pp. 91–110, 2004.
- [3] K. E. A. Van de Sande, T. Gevers, and C. G. M. Snoek, "Evaluating color descriptors for object and scene recognition," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 32, no. 9, pp. 1582–1596, 2010.
- [4] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, "Speeded-up robust features (surf)," Computer vision and image understanding, vol. 110, no. 3, pp. 346–359, 2008.
- [5] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 1, June 2005, pp. 886–893.
- [6] T. Ojala, M. Pietik¨ainen, and D. Harwood, "A comparative study of texture measures with classification based on featured distributions," Pattern recognition, vol. 29, no. 1, pp. 51–59, 1996.
- [7] B. A. Olshausen and D. J. Field, "Emergence of simple-cell receptive field properties by learning a sparse code for natural images," Nature, vol. 381, no. 12, pp. 607–609, 1996.
- [8] H. Lee, A. Battle, R. Raina, and A. Y. Ng, "Efficient sparse coding algorithms," in Advances in Neural Information Processing Systems 19, 2007, pp. 801–808.
- [9] P. Vincent, H. Larochelle, Y. Bengio, and P.-A. Manzagol, "Extracting and composing robust features with denoising autoencoders," in Proceedings of the 25th International Conference on Machine learning. ACM, 2008, pp. 1096–1103.
- [10] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," Proceedings of the IEEE, vol. 86, no. 11, pp. 2278–2324, 1998.
- [11] G. Hinton and R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," Science, vol. 313, no. 5786, pp. 504–507, 2006.
- [12] R. Salakhutdinov and G. E. Hinton, "Deep boltzmann machines," in Proceedings of the International Conference on Artificial Intelligence and Statistics, vol. 5, no. 2. MIT Press Cambridge, MA, 2009, pp. 448–455.
- [13] L. J. P. van der Maaten, E. O. Postma, and H. J. van den Herik, "Dimensionality Reduction: A Comparative Review," Tilburg University, Tech. Rep., 2009.
- [14] K. Pearson, "On lines and planes of closest fit to systems of points in space," Philosophical Magazine, vol. 2, pp. 559–572, 1901.
- [15] S. Mika, B. Sch¨olkopf, A. Smola, K.-R. M¨uller, M. Scholz, and G. R¨atsch, "Kernel pca and de-noising in feature spaces," in Advances in Neural Information Processing Systems 11. MIT Press, 1999, pp. 536–542.
- [16] M. Turk and A. Pentland, "Eigenfaces for recognition," Journal of Cognitive Neuroscience, vol. 3, no. 1, pp. 71–86, 1991.
- [17] P. N. Belhumeur, J. a. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. fisherfaces: Recognition using class specific linear projection," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 19, no. 7, pp. 711–720, Jul. 1997.
- [18] Q. Le, M. Ranzato, R. Monga, M. Devin, K. Chen, G. Corrado, J. Dean, and A. Ng., "Building high-level features using large scale unsupervised learning," in International Conference on Machine Learning, June 2012.

References

- [19] J. Sivic and A. Zisserman, “Video google: A text retrieval approach to object matching in videos,” in Ninth IEEE International Conference on Computer Vision, 2003. IEEE, 2003, pp. 1470–1477.
- [20] G. Csurka, C. Dance, L. Fan, J. Willamowski, and C. Bray, “Visual categorization with bags of keypoints,” in Workshop on statistical learning in computer vision, ECCV, vol. 1, 2004, p. 22.
- [21] L. Fei-Fei and P. Perona, “A bayesian hierarchical model for learning natural scene categories,” in IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005, vol. 2. IEEE, 2005, pp. 524–531.
- [22] F. Perronnin and C. Dance, “Fisher kernels on visual vocabularies for image categorization,” in IEEE Conference on Computer Vision and Pattern Recognition, 2007. IEEE, 2007, pp. 1–8.
- [23] F. Perronnin, J. S´anchez, and T. Mensink, “Improving the fisher kernel for large-scale image classification,” Computer Vision–ECCV 2010, pp. 143–156, 2010.
- [24] S. Lazebnik, C. Schmid, and J. Ponce, “Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories,” in IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2006, vol. 2. IEEE, 2006, pp. 2169–2178.
- [25] K. Grauman and T. Darrell, “The pyramid match kernel: Discriminative classification with sets of image features,” in Tenth IEEE International Conference on Computer Vision, 2005, vol. 2. IEEE, 2005, pp. 1458–1465.
- [26] I. Borg and P. J. F. Groenen, Modern Multidimensional Scaling: Theory and Applications (Springer Series in Statistics), 2nd ed. Springer, 2005.
- [27] E. L. Schwartz, A. Shaw, and E. Wolfson, “A numerical solution to the generalized mapmaker’s problem: flattening nonconvex polyhedral surfaces,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 11, no. 9, pp. 1005–1008, 1989.
- [28] A. Elad and R. Kimmel, “On bending invariant signatures for surfaces,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 25, no. 10, pp. 1285–1295, 2003.
- [29] J. Kruskal, “Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis,” Psychometrika, vol. 29, pp. 1–27, 1964.
- [30] J. de Leeuw, “Applications of convex analysis to multidimensional scaling,” in Recent Developments in Statistics, J. Barra, F. Brodeau, G. Romier, and B. V. Cutsem, Eds. Amsterdam: North Holland Publishing Company, 1977, pp. 133–146.
- [31] C. K. Williams, “On a connection between kernel pca and metric multidimensional scaling,” Machine Learning, vol. 46, no. 1, pp. 11–19, 2002.
- [32] J. Besag, “On the Statistical Analysis of Dirty Pictures,” Journal of the Royal Statistical Society. Series B (Methodological), vol. 48, no. 3, pp. 259–302, 1986.
- [33] K. Levenberg, “A method for the solution of certain non-linear problems in least squares,” Quarterly of Applied Mathematics, vol. 2, pp. 164–168, 1944.
- [34] D. Marquardt, “An algorithm for least-squares estimation of nonlinear parameters,” Journal of the Society for Industrial and Applied Mathematics, vol. 11, no. 2, pp. 431–441, 1963.

References

- [35] R. Datta, D. Joshi, J. Li, and J. Z. Wang, "Image retrieval: Ideas, influences, and trends of the new age," ACM Computing Surveys (CSUR), vol. 40, no. 2, p. 5, 2008.
- [36] A. W. M. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, "Content-based image retrieval at the end of the early years," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22, no. 12, pp. 1349–1380, 2000.
- [37] Y. Rubner, C. Tomasi, and L. J. Guibas, "The earth mover's distance as a metric for image retrieval," International Journal of Computer Vision, vol. 40, no. 2, pp. 99–121, 2000.
- [38] L. Wang, Y. Zhang, and J. Feng, "On the euclidean distance of images," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 27, no. 8, pp. 1334–1339, Aug. 2005.
- [39] S. T. Roweis and L. K. Saul, "Nonlinear dimensionality reduction by locally linear embedding," Science, vol. 290, no. 5500, pp. 2323–2326, 2000.
- [40] M. M. Bronstein, A. M. Bronstein, R. Kimmel, and I. Yavneh, "Multigrid multidimensional scaling," Numerical linear algebra with applications, vol. 13, no. 2-3, pp. 149–171, 2006.
- [41] G. Rosman, A. M. Bronstein, M. M. Bronstein, A. Sidi, and R. Kimmel, "Fast multidimensional scaling using vector extrapolation," SIAM J. Sci. Comput, vol. 2, 2008.
- [42] G. Rosman, A. M. Bronstein, M. M. Bronstein, and R. Kimmel, "Topologically constrained isometric embedding," in Human Motion. Springer, 2008, pp. 243–262.
- [43] A. M. Bronstein, M. Bronstein, M. M. Bronstein, and R. Kimmel, Numerical geometry of non-rigid shapes. Springer, 2008.
- [44] S. Agarwal, A. Awan, and D. Roth, "Learning to detect objects in images via a sparse, part-based representation," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 26, no. 11, pp. 1475–1490, Nov. 2004.
- [45] Q. Wang, "Kernel principal component analysis and its applications in face recognition and active shape models," arXiv preprint arXiv:1207.3538, 2012.
- [46] C. Cortes and V. Vapnik, "Support-vector networks," Machine Learning, vol. 20, pp. 273–297, 1995.
- [47] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," ACM Transactions on Intelligent Systems and Technology, vol. 2, pp. 27:1–27:27, 2011.
- [48] J. Wang, J. Li, and G. Wiederhold, "Simplicity: semantics-sensitive integrated matching for picture libraries," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 23, no. 9, pp. 947–963, Sep. 2001.
- [49] 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 Computer Vision and Pattern Recognition Workshop, 2004. IEEE, 2004, pp. 178–178.
- [50] D. Tao, X. Tang, X. Li, and X. Wu, "Asymmetric bagging and random subspace for support vector machines-based relevance feedback in image retrieval," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 28, no. 7, pp. 1088–1099, July 2006.