A Comparative Study of Basis Selection Techniques for SAR Automatic Target Recognition

Umamahesh Srinivas[†] Vishal Monga[†] Vahid Riasati[‡]

[†]Pennsylvania State University
University Park, PA

[‡]MacAulay-Brown Inc.
Dayton, OH

iPAL @ PENNSTATE
Information Processing and Algorithms Laboratory

MacB
MACAULAY-BROWN, INC.

IEEE Radarcon 2012

May 10, 2012

Automatic Target Recognition (ATR)

- Exploit imagery from diverse sensed sources for automatic target identification[1]
- Variety of sensors: synthetic aperture radar (SAR), inverse SAR (ISAR), forward looking infra-red (FLIR), hyperspectral
- Diverse scenarios: air-to-ground, air-to-air, surface-to-surface

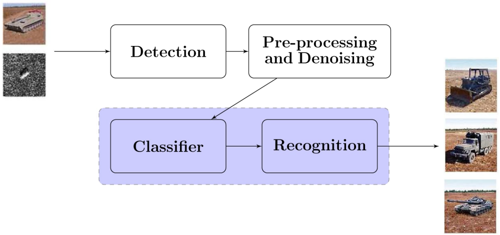


Figure: Schematic of ATR framework. The classification and recognition stages assign an input image/ feature to one of many target classes.

[1] Bhanu et al., IEEE AES Systems Magazine, 1993

Target classification

Two-stage framework:

1. Feature extraction from sensed imagery
   - Geometric feature-point descriptors[2]
   - Eigen-templates[3]
   - Transform domain coefficients: wavelets[4]

[2] Olson et al., IEEE Trans. Image Process., 1997

[3] Bhatnagar et al., IEEE ICASSP, 1998

[4] Casasent et al., Neural Networks, 2005

[5] Daniell et al., Optical Engineering, 1992

[6] Zhao et al., IEEE Trans. Aerosp. Electron. Syst., 2001

Target classification

Two-stage framework:

1. Feature extraction from sensed imagery
   - Geometric feature-point descriptors[2]
   - Eigen-templates[3]
   - Transform domain coefficients: wavelets[4]
2. Decision engine which performs class assignment
   - Linear and quadratic discriminant analysis
   - Neural networks[5]
   - Support vector machines (SVM)[6]

---

[2] Olson et al., IEEE Trans. Image Process., 1997

[3] Bhatnagar et al., IEEE ICASSP, 1998

[4] Casasent et al., Neural Networks, 2005

[5] Daniell et al., Optical Engineering, 1992

[6] Zhao et al., IEEE Trans. Aerosp. Electron. Syst., 2001

Recent research trends: Fusion

- Exploit complementary yet correlated information offered by different sets of features/classifiers
  - Ensemble classifiers[7]
  - Voting strategy[8]
  - Boosting[9]

---

[7] Rizvi and Nasrabadi, Applied Imagery Pattern Recognition Workshop, 2003

[8] Gomes et al., IEEE Radar Conf., 2008

[9] Sun et al., IEEE Trans. Aerosp. Electron. Syst., 2007

[10] Srinivas et al., IEEE Radar Conf., 2011

[11] Srinivas et al., IEEE Int. Conf. Image Processing, 2011

Recent research trends: Fusion

- Exploit complementary yet correlated information offered by different sets of features/classifiers
  - Ensemble classifiers[7]
  - Voting strategy[8]
  - Boosting[9]
  - Meta-classification[10]
  - Probabilistic graphical models for feature fusion[11] (boosting on graphs which model low-level features)

---

[7] Rizvi and Nasrabadi, Applied Imagery Pattern Recognition Workshop, 2003

[8] Gomes et al., IEEE Radar Conf., 2008

[9] Sun et al., IEEE Trans. Aerosp. Electron. Syst., 2007

[10] Srinivas et al., IEEE Radar Conf., 2011

[11] Srinivas et al., IEEE Int. Conf. Image Processing, 2011

Motivation: Feature extraction

- Feature extraction → projection to lower dimensional feature space
  1. Inherent low-dimensional space that captures image information with minimal redundancy[12]
  2. Computational benefits for real-time applications

12 Jolliffe, Principal Component Analysis, Springer, 1986

Motivation: Feature extraction

- Feature extraction $\rightarrow$ projection to lower dimensional feature space
  1. Inherent low-dimensional space that captures image information with minimal redundancy[12]
  2. Computational benefits for real-time applications
- Optimization problem:

$$\boldsymbol{x} = \arg\min_{\hat{\boldsymbol{x}}} \|\boldsymbol{y} - \boldsymbol{A}\hat{\boldsymbol{x}}\|_2$$

  - $\boldsymbol{y}$: target image in $\mathbb{R}^m$
  - $\boldsymbol{x}$: corresponding feature vector in $\mathbb{R}^n, n < m$
  - $\boldsymbol{A}$: projection matrix in $\mathbb{R}^{m \times n} \rightarrow$ collection of n basis vectors, each in $\mathbb{R}^m$

[12] Jolliffe, Principal Component Analysis, Springer, 1986

Motivation: Feature extraction

- Feature extraction $\rightarrow$ projection to lower dimensional feature space
  1. Inherent low-dimensional space that captures image information with minimal redundancy[12]
  2. Computational benefits for real-time applications
- Optimization problem:

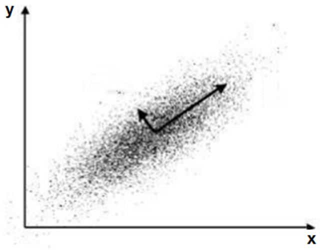
$$\boldsymbol{x} = \arg\min_{\hat{\boldsymbol{x}}} \|\boldsymbol{y} - \boldsymbol{A}\hat{\boldsymbol{x}}\|_2$$

  - $\boldsymbol{y}$: target image in $\mathbb{R}^m$
  - $\boldsymbol{x}$: corresponding feature vector in $\mathbb{R}^n, n < m$
  - $\boldsymbol{A}$: projection matrix in $\mathbb{R}^{m \times n}$ $\rightarrow$ collection of n basis vectors, each in $\mathbb{R}^m$
- How to choose $\boldsymbol{A}$?

[12] Jolliffe, Principal Component Analysis, Springer, 1986

Review: Principal Component Analysis (PCA)

- Statistical tool for dimensionality reduction via change of basis
- Modeling an observation of physical phenomena as a linear combination of basis vectors



- Eigenvectors of data covariance matrix form the projection basis
- Applications in image classification: eigenfaces for face recognition[13], eigen-templates for ATR[14]

[13] Turk and Pentland, IEEE Conf. CVPR, 1991

[14] Bhatnagar et al., IEEE ICASSP, 1998

Review: Singular Value Decomposition (SVD)

- Computational tool underlying PCA
- Data matrix $\boldsymbol{X} \in \mathbb{R}^{m \times N}$ can be factorized as:

$$\boldsymbol{X} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{V}^T = \sum_{i=1}^{r} \lambda_i \boldsymbol{u}_i \boldsymbol{v}_i^T$$

Review: Singular Value Decomposition (SVD)

- Computational tool underlying PCA
- Data matrix $\boldsymbol{X} \in \mathbb{R}^{m \times N}$ can be factorized as:

$$\boldsymbol{X} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{V}^T = \sum_{i=1}^{r} \lambda_i \boldsymbol{u}_i \boldsymbol{v}_i^T$$

- r: rank of $\boldsymbol{X}$
- $\lambda_1 \geq \lambda_2 \geq \ldots \lambda_r > 0$
- $\boldsymbol{U}^T\boldsymbol{U} = \boldsymbol{I}_m$, $\boldsymbol{V}^T\boldsymbol{V} = \boldsymbol{I}_N$

Review: Singular Value Decomposition (SVD)

- Computational tool underlying PCA
- Data matrix $\boldsymbol{X} \in \mathbb{R}^{m \times N}$ can be factorized as:

$$\boldsymbol{X} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{V}^T = \sum_{i=1}^{r} \lambda_i \boldsymbol{u}_i \boldsymbol{v}_i^T$$

- r: rank of $\boldsymbol{X}$
- $\lambda_1 \geq \lambda_2 \geq \ldots \lambda_r > 0$
- $\boldsymbol{U}^T\boldsymbol{U} = \boldsymbol{I}_m$, $\boldsymbol{V}^T\boldsymbol{V} = \boldsymbol{I}_N$
- Low-rank approximation:

$$\boldsymbol{X}_k = \sum_{i=1}^{k} \lambda_i \boldsymbol{u}_i \boldsymbol{v}_i^T$$

Review: Singular Value Decomposition (SVD)

- Computational tool underlying PCA
- Data matrix $\boldsymbol{X} \in \mathbb{R}^{m \times N}$ can be factorized as:

$$\boldsymbol{X} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{V}^T = \sum_{i=1}^{r} \lambda_i \boldsymbol{u}_i \boldsymbol{v}_i^T$$

- r: rank of $\boldsymbol{X}$
- $\lambda_1 \geq \lambda_2 \geq \ldots \lambda_r > 0$
- $\boldsymbol{U}^T\boldsymbol{U} = \boldsymbol{I}_m$, $\boldsymbol{V}^T\boldsymbol{V} = \boldsymbol{I}_N$
- Low-rank approximation:

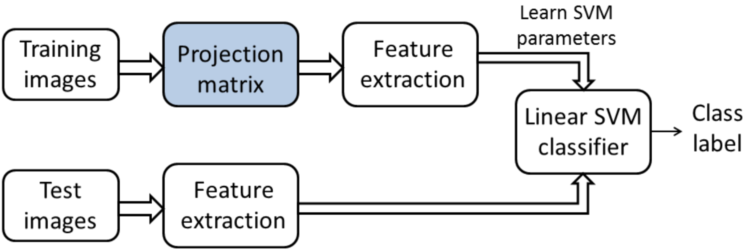
$$\boldsymbol{X}_k = \sum_{i=1}^{k} \lambda_i \boldsymbol{u}_i \boldsymbol{v}_i^T$$

- Of all k-rank approximations, $\boldsymbol{X}_k$ is optimal

$$\boldsymbol{X}_k = \arg \min_{\text{rank}(\tilde{\boldsymbol{X}})=k} \|\boldsymbol{X} - \tilde{\boldsymbol{X}}\|_F$$

- Robustness to noise

Contribution of our work



- Feature projection
  1. Non-negative matrix approximation (NNMA)
  2. Oriented principal component analysis (OPCA)
- Nature of training basis
  1. Shared basis
  2. Class-specific basis

Alternatives to PCA for SAR ATR: Rationale

- Underlying generative model $\rightarrow$ linear combination of basis functions with element-wise non-negative components
- $\boldsymbol{U}$ and $\boldsymbol{V}$ have both positive and negative elements in general $\rightarrow$ interpretation of basis vectors difficult
- Orthogonality of PCA basis vectors unnatural for ATR problem

Alternatives to PCA for SAR ATR: Rationale

- Underlying generative model → linear combination of basis functions with element-wise non-negative components
- $\boldsymbol{U}$ and $\boldsymbol{V}$ have both positive and negative elements in general → interpretation of basis vectors difficult
- Orthogonality of PCA basis vectors unnatural for ATR problem

→ NNMA

Alternatives to PCA for SAR ATR: Rationale

- Underlying generative model $\rightarrow$ linear combination of basis functions with element-wise non-negative components
- $\boldsymbol{U}$ and $\boldsymbol{V}$ have both positive and negative elements in general $\rightarrow$ interpretation of basis vectors difficult
- Orthogonality of PCA basis vectors unnatural for ATR problem

$\rightarrow$ NNMA

- First few principal components sufficient for classification when inter-class variations are dominant
- Incorporate class-specific information

Alternatives to PCA for SAR ATR: Rationale

- Underlying generative model $\rightarrow$ linear combination of basis functions with element-wise non-negative components
- $\boldsymbol{U}$ and $\boldsymbol{V}$ have both positive and negative elements in general $\rightarrow$ interpretation of basis vectors difficult
- Orthogonality of PCA basis vectors unnatural for ATR problem

$\rightarrow$ NNMA

- First few principal components sufficient for classification when inter-class variations are dominant
- Incorporate class-specific information

$\rightarrow$ OPCA

$\rightarrow$ Class-specific basis representations

Non-negative Matrix Approximation (NNMA)

- Follows from non-negative matrix factorization (NMF) technique[15]

$$\boldsymbol{X} \approx \boldsymbol{W}\boldsymbol{H}; \quad \boldsymbol{W}, \boldsymbol{H} \geq \boldsymbol{0}$$

- Ready interpretation of $\boldsymbol{W}$ as additive basis
- Intuitively motivated by SAR imaging physics (non-negativity)
- Dimensionality reduction: k-rank non-negative matrix approximation

[15] Lee and Seung, Nature, 1999

Non-negative Matrix Approximation (NNMA)

- Follows from non-negative matrix factorization (NMF) technique[15]

$$\boldsymbol{X} \approx \boldsymbol{W}\boldsymbol{H}; \quad \boldsymbol{W}, \boldsymbol{H} \geq \mathbf{0}$$

- Ready interpretation of $\boldsymbol{W}$ as additive basis
- Intuitively motivated by SAR imaging physics (non-negativity)
- Dimensionality reduction: k-rank non-negative matrix approximation

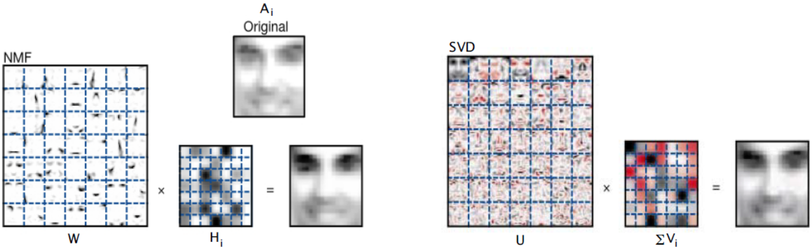


Figure: Illustration: NMF vs. PCA for image representation.

[15] Lee and Seung, Nature, 1999

Non-negative Matrix Approximation (NNMA)

Properties:

- Basis vectors $\boldsymbol{w}_i$ not orthogonal by design
- Sparsity of $\boldsymbol{W}, \boldsymbol{H}$ can be enforced additionally
- $\boldsymbol{W}, \boldsymbol{H}$ not unique

Advantages over SVD/PCA for ATR:

- Easy interpretation of basis vectors
- No restriction of orthogonality

Non-negative Matrix Approximation (NNMA)

Alternating Least Squares[16]:

$$\min_{\boldsymbol{W},\boldsymbol{H}} \quad \|\boldsymbol{X} - \boldsymbol{W}\boldsymbol{H}\|_F^2$$
$$\text{s.t.} \quad \boldsymbol{W}, \boldsymbol{H} \geq \mathbf{0}$$

- Not jointly convex in $\boldsymbol{W}, \boldsymbol{H}$ (separably convex however)

[16] Paatero and Tapper, 1994

[17] Lee and Seung, 2000

Non-negative Matrix Approximation (NNMA)

Alternating Least Squares[16]:

$$\min_{\boldsymbol{W},\boldsymbol{H}} \quad \|\boldsymbol{X} - \boldsymbol{W}\boldsymbol{H}\|_F^2$$

$$\text{s.t.} \quad \boldsymbol{W}, \boldsymbol{H} \geq \mathbf{0}$$

- Not jointly convex in $\boldsymbol{W}, \boldsymbol{H}$ (separably convex however)

Alternate formulation: Divergence update[17]

$$\min_{\boldsymbol{W},\boldsymbol{H}} D(\boldsymbol{X}\|\boldsymbol{W}\boldsymbol{H}) = \sum_{i,j} \left(\boldsymbol{X}_{ij} \log \frac{\boldsymbol{X}_{ij}}{[\boldsymbol{W}\boldsymbol{H}]_{ij}} - \boldsymbol{X}_{ij} + [\boldsymbol{W}\boldsymbol{H}]_{ij} \right)$$

$$\text{s.t.} \quad \boldsymbol{W}, \boldsymbol{H} \geq \mathbf{0}$$

[16] Paatero and Tapper, 1994

[17] Lee and Seung, 2000

Non-negative Matrix Approximation (NNMA)

Alternating Least Squares[16]:

$$\begin{aligned} \min_{\boldsymbol{W},\boldsymbol{H}} \quad & \|\boldsymbol{X} - \boldsymbol{W}\boldsymbol{H}\|_F^2 \\ \text{s.t.} \quad & \boldsymbol{W}, \boldsymbol{H} \geq \boldsymbol{0} \end{aligned}$$

- Not jointly convex in $\boldsymbol{W}, \boldsymbol{H}$ (separably convex however)

Alternate formulation: Divergence update[17]

$$\begin{aligned} \min_{\boldsymbol{W},\boldsymbol{H}} D(\boldsymbol{X}||\boldsymbol{W}\boldsymbol{H}) \quad &= \quad \sum_{i,j} \left(\boldsymbol{X}_{ij} \log \frac{\boldsymbol{X}_{ij}}{[\boldsymbol{W}\boldsymbol{H}]_{ij}} - \boldsymbol{X}_{ij} + [\boldsymbol{W}\boldsymbol{H}]_{ij} \right) \\ \text{s.t.} \quad & \quad \boldsymbol{W}, \boldsymbol{H} \geq \boldsymbol{0} \end{aligned}$$

Feature extraction (corresponding to target vector $\boldsymbol{y}$):

$$\boldsymbol{h} = \min_{h} \|\boldsymbol{y} - \boldsymbol{W}\boldsymbol{h}\|_2, \text{ s.t. } \boldsymbol{h} \geq 0$$

[16] Paatero and Tapper, 1994

[17] Lee and Seung, 2000

Oriented Principal Component Analysis (OPCA)[18]

- Generalization of PCA for binary classification

- Maximizes the signal-to-noise ratio between a pair of stochastic signals $\boldsymbol{u}, \boldsymbol{v}$:

$$J_{\mathsf{OPCA}}(\boldsymbol{w}) = \frac{\boldsymbol{w}^T \boldsymbol{R}_u \boldsymbol{w}}{\boldsymbol{w}^T \boldsymbol{R}_v \boldsymbol{w}},$$

  where $\boldsymbol{R}_u = E\{\boldsymbol{u}\boldsymbol{u}^T\}, \boldsymbol{R}_v = E\{\boldsymbol{v}\boldsymbol{v}^T\}$

[18] Diamantaras and Kung, 1996

Oriented Principal Component Analysis (OPCA)[18]

- Generalization of PCA for binary classification
- Maximizes the signal-to-noise ratio between a pair of stochastic signals $\boldsymbol{u}, \boldsymbol{v}$:

$$J_{\text{OPCA}}(\boldsymbol{w}) = \frac{\boldsymbol{w}^T \boldsymbol{R}_u \boldsymbol{w}}{\boldsymbol{w}^T \boldsymbol{R}_v \boldsymbol{w}},$$

  where $\boldsymbol{R}_u = E\{\boldsymbol{u}\boldsymbol{u}^T\}, \boldsymbol{R}_v = E\{\boldsymbol{v}\boldsymbol{v}^T\}$

- Maximizer $\boldsymbol{w} = \boldsymbol{e}_1$ of $J_{\text{OPCA}} \rightarrow$ principal oriented component; generalized eigenvector of $[\boldsymbol{R}_u, \boldsymbol{R}_v]$

18 Diamantaras and Kung, 1996

Oriented Principal Component Analysis (OPCA)[18]

- Generalization of PCA for binary classification
- Maximizes the signal-to-noise ratio between a pair of stochastic signals $\boldsymbol{u}, \boldsymbol{v}$:

$$J_{\mathsf{OPCA}}(\boldsymbol{w}) = \frac{\boldsymbol{w}^T \boldsymbol{R}_u \boldsymbol{w}}{\boldsymbol{w}^T \boldsymbol{R}_v \boldsymbol{w}},$$

  where $\boldsymbol{R}_u = E\{\boldsymbol{u}\boldsymbol{u}^T\}, \boldsymbol{R}_v = E\{\boldsymbol{v}\boldsymbol{v}^T\}$

- Maximizer $\boldsymbol{w} = \boldsymbol{e}_1$ of $J_{\mathsf{OPCA}} \rightarrow$ principal oriented component; generalized eigenvector of $[\boldsymbol{R}_u, \boldsymbol{R}_v]$
- Oriented components $\boldsymbol{e}_2, \boldsymbol{e}_3, \ldots, \boldsymbol{e}_m$: maximize J_{OPCA} subject to

$$\boldsymbol{e}_i^T \boldsymbol{R}_u \boldsymbol{e}_j = \boldsymbol{e}_i^T \boldsymbol{R}_v \boldsymbol{e}_j = 0, i \neq j.$$

- Identical to PCA if $\boldsymbol{v}$ is white noise

[18] Diamantaras and Kung, 1996

Oriented Principal Component Analysis (OPCA)[18]

- Generalization of PCA for binary classification
- Maximizes the signal-to-noise ratio between a pair of stochastic signals $\boldsymbol{u}, \boldsymbol{v}$:

$$J_{\text{OPCA}}(\boldsymbol{w}) = \frac{\boldsymbol{w}^T \boldsymbol{R}_u \boldsymbol{w}}{\boldsymbol{w}^T \boldsymbol{R}_v \boldsymbol{w}},$$

  where $\boldsymbol{R}_u = E\{\boldsymbol{u}\boldsymbol{u}^T\}, \boldsymbol{R}_v = E\{\boldsymbol{v}\boldsymbol{v}^T\}$

- Maximizer $\boldsymbol{w} = \boldsymbol{e}_1$ of $J_{\text{OPCA}} \rightarrow$ principal oriented component; generalized eigenvector of $[\boldsymbol{R}_u, \boldsymbol{R}_v]$
- Oriented components $\boldsymbol{e}_2, \boldsymbol{e}_3, \ldots, \boldsymbol{e}_m$: maximize J_{OPCA} subject to

$$\boldsymbol{e}_i^T \boldsymbol{R}_u \boldsymbol{e}_j = \boldsymbol{e}_i^T \boldsymbol{R}_v \boldsymbol{e}_j = 0, i \neq j.$$

- Identical to PCA if $\boldsymbol{v}$ is white noise
- Application to SAR ATR: signal $\boldsymbol{v}$ chosen from complementary class

[18] Diamantaras and Kung, 1996

Oriented Principal Component Analysis (OPCA)

Class-specific basis:

- OPCA inherently designed for binary classification
- K-class scenario: solve K different binary problems
- For the i-th such problem:
  - $\boldsymbol{R}_u$: sample covariance matrix of training images from class i
  - $\boldsymbol{R}_v$: sample covariance matrix of representative training images chosen from all other classes

Shared training vs. class-specific training

Shared basis:

- Data matrix contains training from all classes
- Assumption: inter-class variations dominant compared to intra-class

Shared training vs. class-specific training

Shared basis:

- Data matrix contains training from all classes
- Assumption: inter-class variations dominant compared to intra-class

Class-specific basis:

- Separate projection matrix $\boldsymbol{A}_i$ for each class $i = 1, \ldots, K$
- Projection matrix $\boldsymbol{A} = [\boldsymbol{A}_1 \ \boldsymbol{A}_2 \ \ldots \boldsymbol{A}_K]$

Shared training vs. class-specific training

Shared basis:

- Data matrix contains training from all classes
- Assumption: inter-class variations dominant compared to intra-class

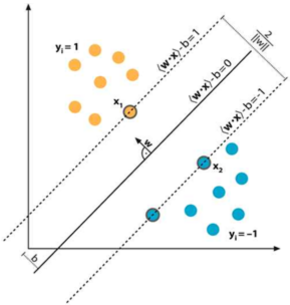
Class-specific basis:

- Separate projection matrix $\boldsymbol{A}_i$ for each class $i = 1, \ldots, K$
- Projection matrix $\boldsymbol{A} = [\boldsymbol{A}_1 \ \boldsymbol{A}_2 \ \ldots \boldsymbol{A}_K]$
- More discriminative than shared basis
- Sensitive to scenario of inadequate training
- Possibly higher feature dimension $\rightarrow$ computational cost

Classifier: Support Vector Machine (SVM)[21]

$$f(\boldsymbol{x}) = \sum_{i=1}^{N} \alpha_i y_i K(\boldsymbol{s}_i, \boldsymbol{x}) + b$$

- Widely used in ATR problems[19,20]

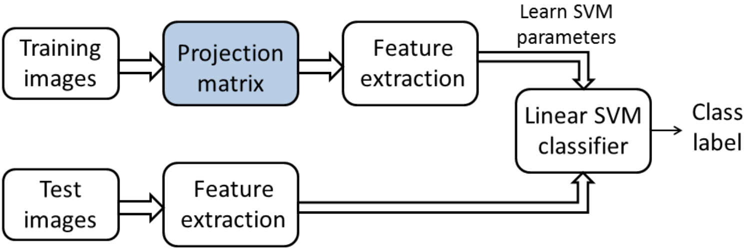


[19] Zhao and Principe, IEEE Trans. Aerosp. Electron. Syst., 2001

[20] Casasent and Wang, Neural Networks, 2005

[21] Vapnik, The nature of statistical learning theory, 1995

Overall classification framework



- Shared training basis: PCA and NNMA
- Class-specific training basis: PCA, NNMA, OPCA
- Linear SVM: representative of state-of-the-art classifiers

Experimental set-up

- MSTAR database: one-foot resolution X-band SAR images
- Five target classes
  1. T-72 tanks
  2. BMP-2 infantry fighting vehicles
  3. BTR-70 armored personnel carriers
  4. ZIL131 trucks
  5. D7 tractors

Target class	Serial number	# Training images	# Test images
BMP-2	SN_C21	233	196
	SN_9563	233	195
	SN_9566	232	196
BTR-70	SN_C71	233	196
T-72	SN_132	232	196
	SN_812	231	195
	SN_S7	228	191
ZIL131	-	299	274
D7	-	299	274

Table: Target classes in the experiment.

Experimental set-up

- Training images: 17° depression angle
- Test images: 15° depression angle
- Images cropped to 64×64 pixels (i.e. vectorized data in $\mathbb{R}^{4096}$)
- Pose: varies from 0° to 360°
- Number of basis vectors: 750

Results: Classification performance

Table: Confusion matrix: Shared PCA basis.

Class	BMP-2	BTR-70	T-72	ZIL131	D7
BMP-2	**0.84**	0.06	0.04	0.02	0.04
BTR-70	0.05	**0.87**	0.03	0.02	0.03
T-72	0.03	0.07	**0.83**	0.03	0.04
ZIL131	0.05	0.03	0.02	**0.84**	0.06
D7	0.06	0.02	0.04	0.06	**0.82**

Table: Confusion matrix: Shared NNMA basis.

Class	BMP-2	BTR-70	T-72	ZIL131	D7
BMP-2	**0.86**	0.05	0.02	0.05	0.02
BTR-70	0.07	**0.88**	0.04	0.01	0.0
T-72	0.03	0.04	**0.86**	0.02	0.05
ZIL131	0.01	0.06	0.05	**0.87**	0.01
D7	0.04	0.02	0.06	0.04	**0.84**

Results: Classification performance

Table: Class-specific PCA basis.

Class	BMP-2	BTR-70	T-72	ZIL131	D7
BMP-2	**0.86**	0.05	0.04	0.02	0.03
BTR-70	0.04	**0.88**	0.04	0.03	0.01
T-72	0.04	0.05	**0.85**	0.02	0.04
ZIL131	0.02	0.02	0.06	**0.86**	0.04
D7	0.01	0.01	0.07	0.06	**0.85**

Table: Class-specific NNMA basis.

Class	BMP-2	BTR-70	T-72	ZIL131	D7
BMP-2	**0.88**	0.05	0.02	0.01	0.04
BTR-70	0.03	**0.90**	0.02	0.03	0.02
T-72	0.02	0.05	**0.87**	0.04	0.02
ZIL131	0.04	0.02	0.03	**0.89**	0.02
D7	0.02	0.03	0.04	0.04	**0.87**

Table: Class-specific OPCA basis.

Class	BMP-2	BTR-70	T-72	ZIL131	D7
BMP-2	**0.91**	0.03	0.04	0.01	0.01
BTR-70	0.04	**0.91**	0.01	0.03	0.01
T-72	0.02	0.05	**0.88**	0.02	0.03
ZIL131	0.03	0.01	0.03	**0.90**	0.03
D7	0.03	0.03	0.02	0.03	**0.89**

Conclusions

- Proposed alternatives to PCA-based feature extraction in ATR problems
  - NNMA: Non-negativity motivated by underlying SAR image physics
  - OPCA: Captures inter-class variability better

Conclusions

- Proposed alternatives to PCA-based feature extraction in ATR problems
  - NNMA: Non-negativity motivated by underlying SAR image physics
  - OPCA: Captures inter-class variability better
- Future work:
  - NNMA/OPCA features for meta-classification.

Thank You

Questions?