Joint Optimization of Manifold Learning and Sparse Representations for Face and Gesture Analysis

Raymond Ptucha
rwpeec@rit.edu

Artificial Intelligence Seminar
Cornell University
March 1 , 2013

RIT Acknowledgements

Dissertation Advisor:
Dr. Andreas Savakis, Professor, Computer Engineering, RIT

Dissertation Committee:
Dr. Nathan Cahill, Associate Professor, School of Mathematical Sciences, RIT
Dr. Joe Geigel, Associate Professor, Computer Science, RIT
Dr. Andreas Savakis, Professor, Computer Engineering, RIT
Dr. Linwei Wang, Assistant Professor, GCCIS PhD, RIT

PhD Program Director:
Dr. Pengcheng Shi, Professor and Department Head of GCCIS PhD Program, RIT

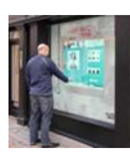
Motivation

- Facial understanding and gesture recognition are powerful enablers in intelligent vision systems.
- Potential applications include surveillance, security, entertainment, smart spaces, and human computer interfaces (HCI).
- Tomorrow's devices will need to embrace human subtleties while interacting with them in their natural conditions.

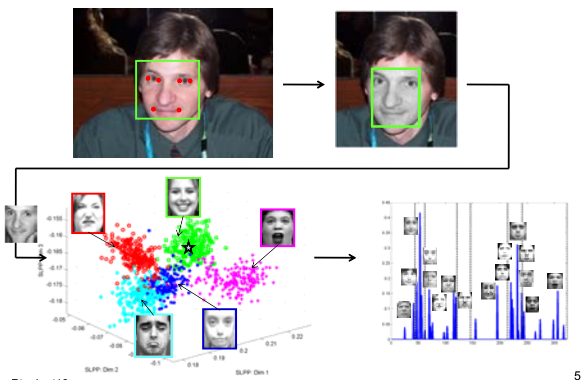
Interactive Digital Signage







Static Processing



A Few Milestones

- Yang [PAMI '07] used dimensionality reduction with SRs for classification purposes.
- Wright [PAMI '09] used SRs for best in class facial recognition.
- Zafeiriou [CVPR '10] used PCA and SR methods based on Wright for facial expression, but reported significant coefficient contamination.
- Ptucha [ICCV '11] used supervised manifold learning to minimize coefficient contamination.
- Jiang [CVPR '11,'12] used K-SVD to jointly optimize classification accuracy and more efficient dictionaries.

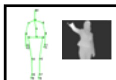
Agenda

- Introduction to Dimensionality Reduction
- Introduction to Sparse Representations
- Merging the two concepts into Manifold based Sparse Representations
- Optimizing the two concepts with LGE-KSVD
- Sample Results

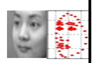
Hypothesis

- Methods based on manifold learning and sparse representations can achieve accurate, robust, and efficient classifiers for scene understanding.

Feature Extraction → **Feature Normalization** → **Manifold Learning** → **Sparse Representation** → **Classification Model** → **Temporal Filtering**



Dimensionality Reduction



- For the purpose of facial understanding, the dimensionality of a 26x20 ($\in \mathbf{R}^{520}$) pixel face image or a 82x2 ($\in \mathbf{R}^{164}$) set of ASM coordinates are artificially high.
- The high dimensionality space makes the facial understanding algorithms more complex than necessary.
- The set of 520 pixels (or 164 coordinates) actually are samples from a lower dimensional manifold that is embedded in a higher dimensional space.
- We would like to discover this lower dimensional manifold representation (to simplify our facial modeling)- a technique formally called manifold learning. [Cayton '05, Ghodsi '06]
- Given a set of inputs $x_1..x_n \in \mathbf{R}^D$, find a mapping $y_i = f(x_i)$, $y_1..y_n \in \mathbf{R}^d$; where $d < D$.

Ptucha '13 10

Locality Preserving Projections* (LPP)

[He '03]

- Given a set of input points $x_1..x_n \in \mathbf{R}^D$, find a mapping $y_i = A^T x_i$, where the resulting $y_1..y_n \in \mathbf{R}^d$; where $\cancel{d < D.}$ $d << D$.
  - Same algebra as PCA, if we kept the top d eigenvectors!
- Create a fully connected adjacency graph W. Assign high weights to close/similar nodes, and low weights to far/dissimilar nodes.
  - Mimic local neighborhood structure from input to projected space.
- LPP is a linear approximation to the nonlinear Laplacian Eigenmap and is solved via the generalized eigenvector problem:

  $X L X^T a = \lambda X D X^T a$
- Where:
  - D is a diagonal matrix whose values are the column sums of W,
  - L is the Laplacian matrix: $L = D\text{-}W$,
  - a is the resulting projection matrix (== "eigenvectors") , and
  - λ is the resulting vector importance (== "eigenvalues") .

* http://www.cad.zju.edu.cn/home/dengcai/Data/DimensionReduction.html

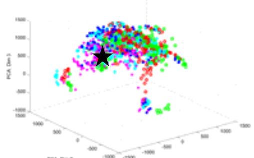
Ptucha '13 14

PCA vs. Supervised LPP

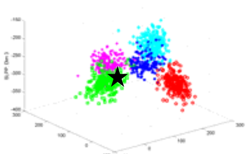
1072 samples, 26x20 pixel faces ($\mathbf{R}^{520} \rightarrow \mathbf{R}^3$)

[Lucey '10]

Top 3 dimms of PCA space.



Top 3 dimms of SLPP space.

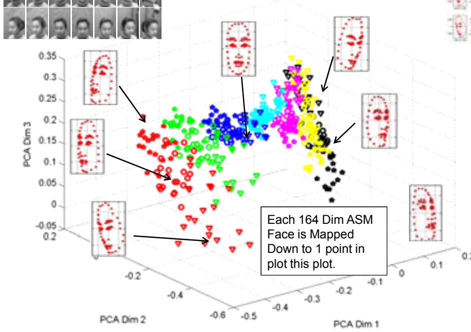


- PCA is good at dimensionality reduction, but assumes linearity.

Ptucha '13 19

Apply Dimensionality Reduction to Pose Training Set

21 subjects, each at 21 poses



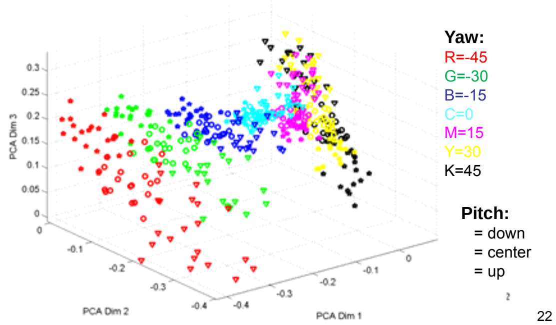
Yaw:
R=-45
G=-30
B=-15
C=0
M=15
Y=30
K=45

Pitch:
▽ = down
○ = center
* = up

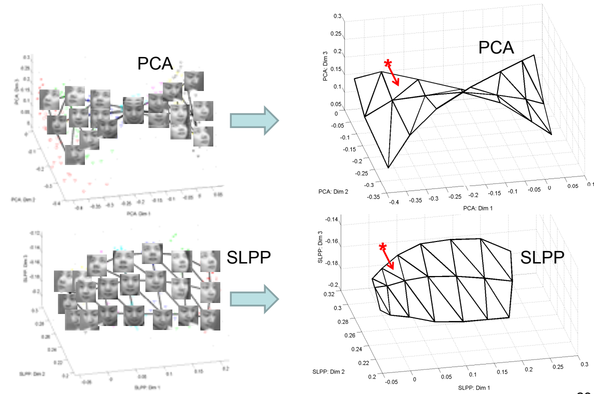
Ptucha '13 21

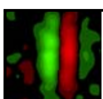
Apply Dimensionality Reduction to Pose Training Set

21 subjects, each at 21 poses



Model Manifold Surface





Actual simple cell response

Sparse Representations

- Inspired by studies of neurons in the visual cortex, the notion of Sparse Representations (SRs) has been proven applicable to a variety of scientific fields.
- For many input signals, such as natural images, only a small number of exemplars are needed to represent new test samples.
- SR gives state-of-the-art results for pattern recognition, noise reduction, super-resolution, tracking, …
- At the The First Facial Expression Recognition and Analysis Challenge (FERA2011) at FG'11:
  - 13/15 entrants used SVM, but 0/15 entrants used SR

[Images from DeAngelis, Ohzawa & Freeman, 1995]

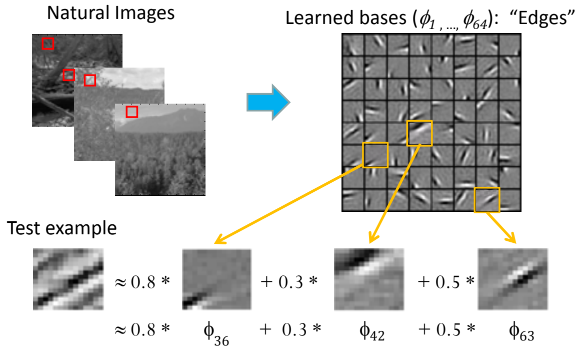
Sparse Representations

- Input: Images $x_1, x_2, \ldots, x_n$ (each $\in \mathbf{R}^{h \times w}$)
- Learn: Dictionary of bases $\phi_1, \phi_2, \ldots, \phi_k$ (each also $\in \mathbf{R}^{h \times w}$), so that each input x (and newly introduced test samples y) can be approximately decomposed as:

$$\hat{x} \approx \sum_{j=1}^{k} a_j \phi_j$$

s.t. a_j's are mostly zero ("sparse")

Sparse coding illustration [Ng ECCV '10]



[0, 0, …, 0, **0.8**, 0, …, 0, **0.3**, 0, …, 0, **0.5**, 0]
= [a_1, …, a_{64}] (feature representation)

Sparse Representations

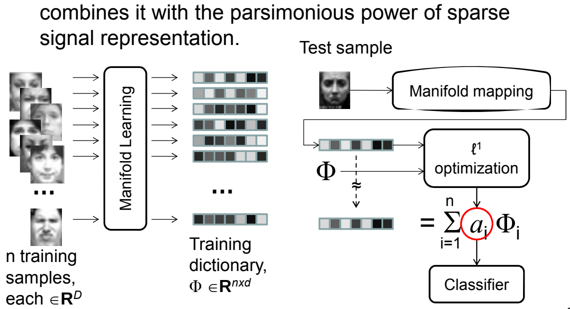
- Given y and Φ, the objective of SRs is to identify the smallest number of nonzero coefficients $a \in \mathbf{R}^n$ such that:
  $y \approx \hat{y} = \Phi\, a$.
- The solution is equivalent to the Lasso regression:

$$\hat{a} = \min\left\{ \left\|y - \Phi a\right\|^2 + \lambda \left\|a\right\|_1 \right\}$$

  where $\|a\|_1 = \Sigma\, |a|$.
- Although not differentiable like a ridge regression, the ℓ^1 minimization problem can be efficiently solved using convex optimization algorithms. [Donoho '06, Candes '06]
- Some of the fastest approaches include several variants of Least Angle Regression with lasSo (LARS). [Efron '04]

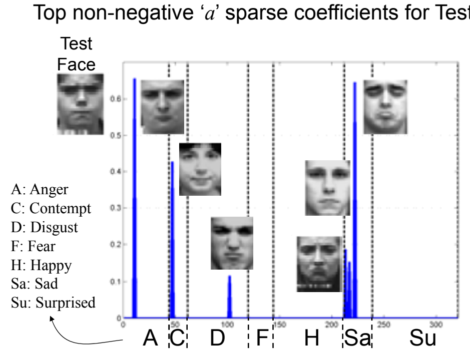
Putting it Together

- Manifold based Sparse Representations (MSR) exploit the discriminative behavior of manifold learning, and combines it with the parsimonious power of sparse signal representation.



Sparse Coefficients

Top non-negative 'a' sparse coefficients for Test "sad" Face.



Interesting…but ..how do we turn this into a classifier?

- Max peak?
- Max non-zero coefficients?
- Max Energy?

Reconstruction Error

- A reconstruction error classifier generally outperforms other methods. [Yang '07, Wright '09]
- Estimate the class, c^* of a query sample y by comparing the reconstruction error inquired when only the reconstruction coefficients a_c corresponding to a specific class c are selected.

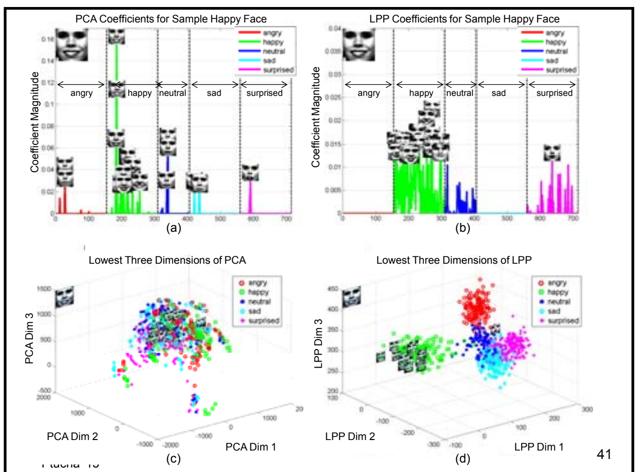
$$c^* = \arg\min_{c=1\ldots z} \|y - \Phi\, \underbrace{a_c}\|_2$$

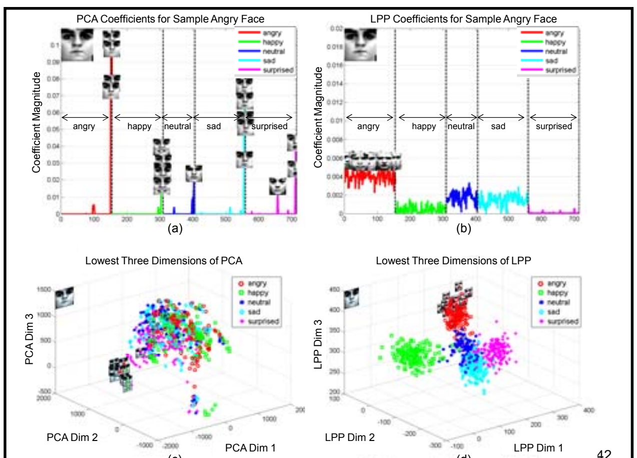
Use non-zero coefficients from all classes to estimate, $y \approx \Phi\, a$

Use non-zero coefficients from each class

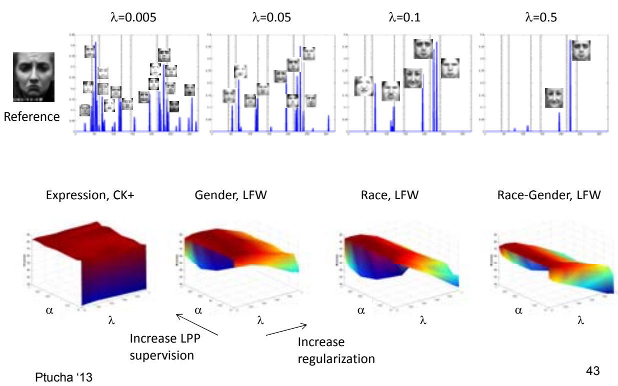
Coefficient Contamination

- Applying the reconstruction error is not a straightforward process for natural images.
- For example, facial identity of the person is often confused with facial expression.
- The usage of semi-supervised manifold learning encourages clustering of sample images in accordance with classification labels.



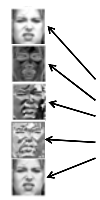


Supervision & Regularization



Region and Pixel Processing

- It is quite conceivable that different regions of the face [Kumar '08] may benefit from different types of pixel processing.
  - Each pixel processing↔facial region combination is a valid feature input to the statistical inference model.

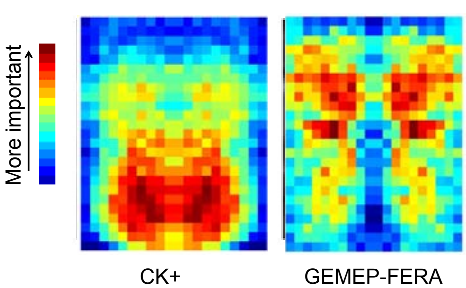


	Fullimg	Face	Mustache	Mouth	EyeReg
Lum	91.4	88.5	82.1	**83.3**	61.0
Mag	88.7	**92.0**	80.1	76.7	**65.3**
Phase	88.9	87.8	70.4	75.4	61.4
LBP	87.7	85 8	63.1	71.1	46.4
Gabor	**91.9**	87.8	**85.5**	79.1	63.9

MSR accuracies on CK expression dataset

Posed vs. Natural Datasets

MSR enables evaluation of any region of the face

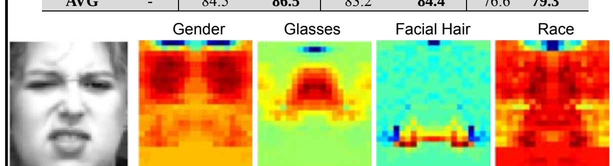


*Correlates well with [Pfister ICCV2011]

MSR Used On Other Facial Attributes

LFW Classification Accuracy

	No. class	SVM no occl.	MSR no occl.	SVM mouth occl.	MSR mouth occl.	SVM eye occl.	MSR eye occl.
Gender	2	89.6	**90.8**	89.8	**90.3**	80.5	**80.8**
Glasses	4	85.0	**87.9**	84.3	**85.0**	71.8	**79.6**
Hair	7	86.9	**87.7**	80.8	**85.6**	87.3	**87.4**
Race	5	85.1	**87.5**	**85.0**	84.3	78.7	**82.0**
Mixed	10	75.9	**78.5**	76.2	**76.6**	64.6	**66.5**
AVG	-	84.5	**86.5**	83.2	**84.4**	76.6	**79.3**



Optimization of Dimensionality Reduction and Sparse Representations

- Sparsity Preserving Projections [Qiao'09] uses (unsupervised) sparse coefficients instead of Laplacian for dimensionality reduction
- Global SR Projections [Lai '09], Discriminative Sparse Coding [Zang '11], and Graph Regularized Sparse Coding [Zheng '11] create variations of joint objective function (DR and SR)
- Supervised LPP [Cai '11] modifies LPP to have (unsupervised) Laplacian and (supervised) LDA properties.
- LC-KSVD [Jiang '11] forces (unsupervised) sparse terms to be (supervised) discriminative and jointly learns a (supervised) classifier

LGE-KSVD

- Each of the previous methods introduce a new dimensionality reduction technique or a new SR technique.
- What lacks is a unified approach that optimizes dimensionality reduction projection matrix *U* with dictionary Φ, and sparse coefficients $\hat{a}$.
- The next few slides will present such a method called LGE-KSVD, for the optimization and infusion of Linear extension of Graph Embedding with K-SVD dictionary learning.
  - Note: LGE is a broader category of linear dimensionality reduction methods which use adjacency matrix *W* to describe neighbor to neighbor topology (includes LDA, LPP, and NPE).

LGE-KSVD

- Classification frameworks based on SR concepts have been found to suffer from:
  1. Coefficient contamination that compromises classification accuracy; and
  2. Computational inefficiencies due to high dimensional features and large dictionaries.
- LGE-KSVD uses:
  - Semi-supervised dimensionality reduction to address both limitations.
  - K-SVD dictionary learning to not only make the dictionaries more efficient, but yield higher classification accuracies.

K-SVD

- K-SVD [Aharon '06] was introduced as a means to learn an over-complete but small dictionary.

$$\{\hat{\Phi}, \hat{a}\} = \min \|x - \Phi a\|_2^2 \quad s.t. \|a\|_0 \le \delta$$

- K-SVD is an iterative technique, where at each iteration, training samples are first sparsely coded using the current dictionary estimate, and then dictionary elements are updated one at a time while keeping others fixed.
- Each new dictionary element is a linear combination of training samples.
- [Rubinstein '08] implemented an efficient implementation of K-SVD using Batch Orthogonal Matching Pursuit (http://www.cs.technion.ac.il/~ronrubin/software.html)

Classification of K-SVD Sparse Coefficients

- Because dictionary elements from K-SVD are a linear combination of input samples, we cannot use the minimum reconstruction error.
- Alternatively we can pass SR coefficients into any regression or machine learning classifier.
- Define *H* as ground truth (GT) matrix, $H \in \mathbf{R}^{k \times n}$.
  - Each column of *H* corresponds to a GT sample. The k^{th} position is 1 if y_i belongs to class k_j, otherwise 0.
- Coefficients a from each training sample are stored in matrix *A*, $A \in \mathbf{R}^{m \times n}$.
- Then solve for coefficient transformation matrix C.

$$\hat{C} = \min \left\| H - C^T A \right\|_2^2 \quad \Rightarrow \quad C = \left(A\ A^T\right)^{-1} A\ H^T$$

Ptucha '13 58

LGE-KSVD Objective Function

- Combining LGE dimensionality reduction with K-SVD minimization functions, we get:

$$\left\{\hat{U}, \hat{\Phi}, \hat{a}\right\} = \min\left\{ \left\| X^T U - \Phi a \right\|_2^2 + \frac{U^T X\ L\ X^T U}{U^T X\ D\ X^T U} \right\} \quad s.t. \left\| a \right\|_0 \le \delta$$

X: input data
U: dim. reduc. matrix
Φ: dictionary
a: sparse coeffs

K-SVD in low dimensional space

LGE dimensionality reduction objective function.

- The above equation is neither directly solvable nor convex.

Ptucha '13 59

LGE-KSVD Objective Function

$$\left\{\hat{U}, \hat{\Phi}, \hat{a}\right\} = \min\left\{ \left\| X^T U - \Phi a \right\|_2^2 + \frac{U^T X\ L\ X^T U}{U^T X\ D\ X^T U} \right\} \quad s.t. \left\| a \right\|_0 \le \delta$$

- We learn a dictionary of *m* atoms, $m \le n$.
- It can be shown that there is an implicit transformation *T*, $\Phi = T X^T U$, where the rank of *T* is greater than the rank of *U*.
- The solution is to use K-SVD to iteratively solve for a, then Φ; then wrap this entire procedure with an update procedure on *U*.

$$\hat{U} = \min \left\| X^T U - A^T \Phi^T \right\|_2^2 \quad \Rightarrow \quad U = \left(X\ X^T\right)^{-1} X\ A^T \Phi^T$$

Ptucha '13 60

Training Procedure for LGE-KSVD

WHILE ε has not converged or $\varepsilon > \tau$
  IF firstIteration
    1a. Calculate *U* using LGE
    1b. Calculate *U* using $U=(XX^T)^{-1}XA^T\Phi^T$
  ENDIF
  2. Calculate low dimensional samples $Y^T=X^TU$
  3. Initializes the *m* samples of Φ randomly from the *n* low dimensional training samples
  4. Calculate $\{A,\Phi\}$ using modified K-SVD, substituting *Y* for *X*.
  5. Calculate *C* using $C=(AA^T)^{-1}AH^T$
  6. Calculate verification set error, $\varepsilon = ||H - C^TA||_2^2$
ENDWHILE

Ptucha '13 61

Testing Procedure for LGE-KSVD

- Given a test sample *x*, along with *U*, Φ, and *C*:

1. Calculate low dimensional sample $y=x^T U$.
2. Use Φ and *y* to calculate sparse coefficients a, using a pursuit* algorithm.
3. Use *C* along with a to estimate class label vector $l \in \mathbf{R}^{k x 1}$ where the maximum value of l is used as a class predictor. $\hat{l} = \max_{i=1:k}\left(l = C^T \hat{a}\right)$

*such as SLEP, http://www.public.asu.edu/~jye02/Software/SLEP

Modified K-SVD

- K-SVD enforces sparsity by fixing the support of each atom in the iteration process to a subset of training samples.
- The addition of supervision injects classification smarts into K-SVD, but still fixes atom support.
- We propose to use semi-supervised LGE adjacency matrix *W* to regulate the support of each dictionary element.

Modified K-SVD

- The support of each dictionary element *j* may:
  - Expand: Modify the support of element *j* by adding (union) all training entries *similar* to element *j*.
  - Contract: Modify the support of element *j* by removing (intersection) training entries *not similar* to element *j*.
  - Redefine: Set the support of element *j* to be only training samples *similar* to element *j*.
  - Fixed: Maintain the support of element *j*, as in the K-SVD algorithm.
- *Similar* is defined in terms of the LGE adjacency matrix

Results: CK+ Expression Dataset

7 static facial expressions, 68 AAM points, 164 training and 163 testing samples

Method	*d* (dimension)	*m* (# dictionary atoms)	*% Accuracy*
PCA	62	-	82.2
LDA	6	-	89.6
LPP	62	-	83.4
NPE	24	-	80.4
SPP	48	-	87.7
K-SVD	136	63	79.1
LC-KSVD1	136	63	79.1
LC-KSVD2	136	63	75.5
SRC	136	164	43.6
MSR	62	164	75.5
LGE based K-SVD	62	63	**92.0**

Results: CK+ Expression Dataset

7 static facial expressions, 60x51 images, 164 training and 163 testing samples

Method	d (dimension)	m (# dictionary atoms)	% Accuracy
PCA	162	-	82.8
LDA	6	-	**86.5**
LPP	163	-	84.7
NPE	71	-	84.0
SPP	80	-	77.9
K-SVD	3060	63	84.0
LC-KSVD1	3060	63	85.9
LC-KSVD2	3060	63	84.7
SRC	500	164	71.8
MSR	163	164	79.1
LGE based K-SVD	163	63	**86.5**

Results: YaleB Recognition Dataset

38 subjects, 192x168 static images reduced to 504 dimensions via random projections, 1216 training and 1198 testing samples

Method	d (dimension)	m (# dictionary atoms)	% Accuracy
PCA	477	-	89.1
LDA	37	-	90.3
LPP	477	-	89.3
NPE	271	-	91.2
SPP	288	-	88.7
K-SVD	504	570	93.2
LC-KSVD1	504	570	93.7
LC-KSVD2	504	570	93.4
SRC	504	1216	86.1
MSR	477	1216	**96.5**
LGE based K-SVD	477	570	95.3

Results: GEMEP-FERA Emotion

5 class, two 24x20 MHI static images per video, 155 training and 134 testing samples

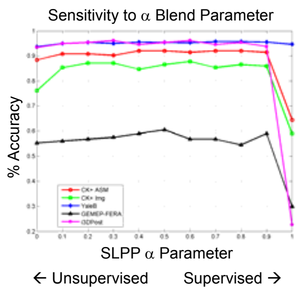
Method	d (dimension)	m (# dictionary atoms)	% Accuracy
PCA	154	-	55.2
LDA	4	-	55.2
LPP	154	-	55.2
NPE	66	-	56.7
SPP	75	-	52.2
K-SVD	1920	75	51.5
LC-KSVD1	1920	75	53.7
LC-KSVD2	1920	75	51.5
SRC	500	155	57.5
MSR	154	155	56.0
LGE based K-SVD	154	75	**60.5**

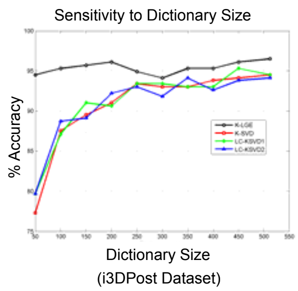
Results: i3DPost Multi-View Activity

12 class, 125 MHI sequences per video, PCA reduced 767 dimensions per video, 512 training and 256 testing samples

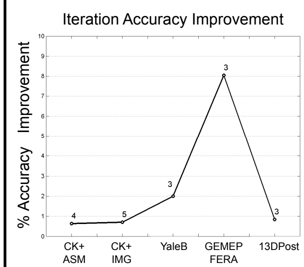
Method	d (dimension)	m (# dictionary atoms)	% Accuracy
PCA	510	-	94.9
LDA	510	-	94.5
LPP	510	-	**96.1**
NPE	224	-	94.9
SPP	241	-	91.0
K-SVD	767	450	94.1
LC-KSVD1	767	450	95.3
LC-KSVD2	767	450	93.8
SRC	767	512	88.7
MSR	510	512	95.3
LGE based K-SVD	510	450	**96.1**

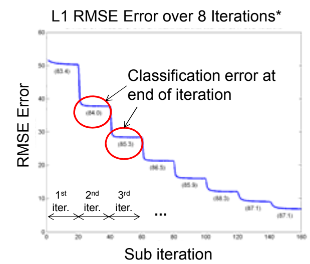
LGE-KSVD Analysis





LGE-KSVD Analysis



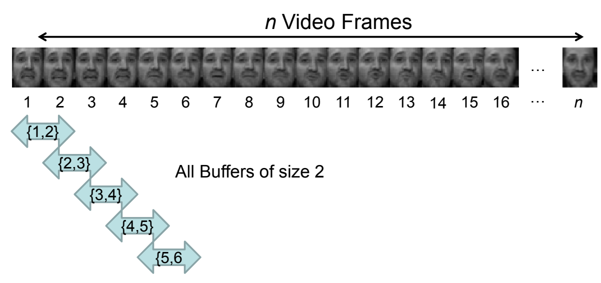


*Each iteration has 20 K-SVD sub iterations
CK+ w/ images

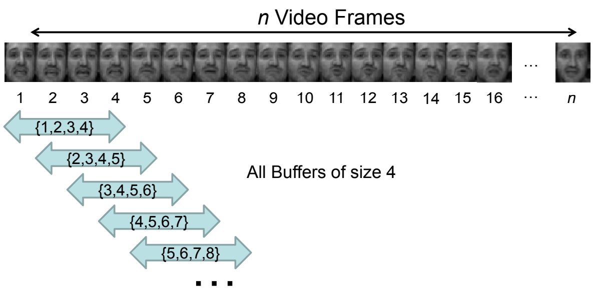
Temporal Processing

- Communication between humans naturally contains temporal signature.
  - Rolling of eyes, waving of hand, wink, etc.
- Previous studies adopted both sparse and dense optical flow techniques and contrast to static methods.
- Facial expressions and gestures can occur at any point in time and are variable in length.
- We define sliding temporal windows, W^{θ}_{l}, each of duration θ frames, $l=1..m$ sliding windows.

Examine Video In Variable Size Rolling Frame Buffers

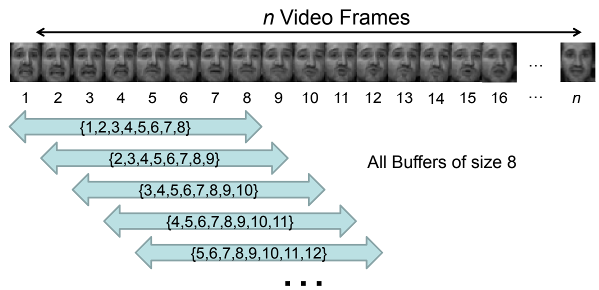


Examine Video In Variable Size Rolling Frame Buffers



Ptucha '13 79

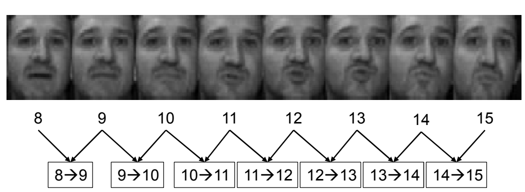
Examine Video In Variable Size Rolling Frame Buffers



Ptucha '13 80

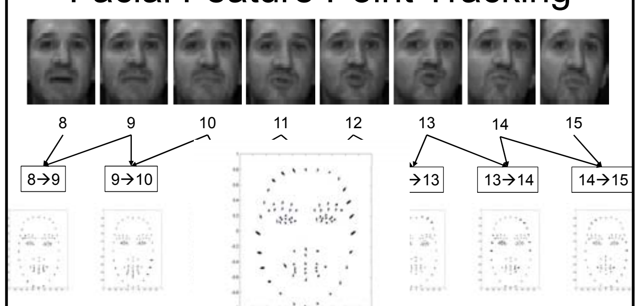
Analysis Example

- Lets say, we are looking at window widths of 8.
- Our first position center is frame 12.
- We then look at 7 motion trajectories:



Ptucha '13 81

Facial Feature Point Tracking



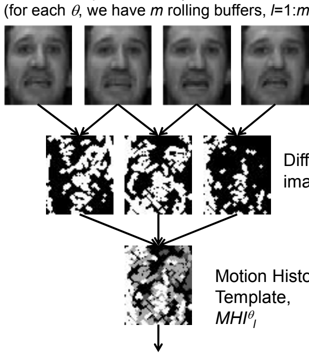
Similarly, can compute point tracking from current frame the mean frame.

Ptucha '13 82

Motion History Images

[Bobick ‘01][Koelstra ‘10]

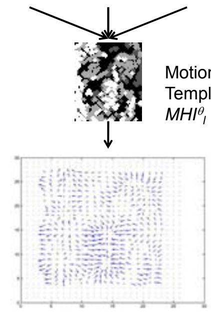
Example buffer W^{θ}_{l} of size θ=4
(for each θ, we have m rolling buffers, l=1:m)



Difference images

Motion History Template, MHI^{θ}_{l}

Motion History Images (Cont’d)



Motion History Template, MHI^{θ}_{l}

Pixels point towards recent movement

Δx and Δy of each vector passed into classifier

Summary

- Face and gesture understanding problems can be reliably solved in unconstrained scenes using SRs.
- The usage of semi-supervised LPP before SR clusters by classification task, avoiding coefficient contamination.
- The usage of K-SVD dictionary learning makes the dictionaries more compact and results in higher classification accuracies.
- If the training dictionary is not over complete, SR methods have trouble generalizing test samples from training dictionary exemplars.

References (1 of 2)

- [Aharon ‘06] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Transactions on Signal Processing,* vol. 54, pp. 4311-22, 2006.
- [Brunelli ‘93] R. Brunelli and T. Poggio, "Face recognition: features versus templates," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 15, pp. 1042-1052, 1993.
- [Bobick ‘01] A. F. Bobick and J. W. Davis, "The recognition of human movement using temporal templates," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 23, pp. 257-67, 2001.
- [Cai ‘11] Xian-Fa Cai et al., “Enhanced Supervised Locality Preserving Projections for Face Recognition”, ICML, 2011.
- [Candes ‘06] E. J. Candes, J. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *IEEE Transactions on Information Theory,* vol. 52, pp. 489-509, 2006.
- [Cayton ‘05] L. Cayton. (2005, Algorithms for manifold learning.
- [Chew ‘11] S. Chew, P. Lucey, S. Lucey, J. Saragih, J. Cohn, and S. Sridharan, "Person-Independent Facial Expression Detection Using Constrained Local Models," in *Automatic Face and Gesture Recognition*, Santa Barbara, CA, USA, 2011.
- [Cootes ‘01] T. F. Cootes, G. J. Edwards, and C. J. Taylor, "Active appearance models," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 23, pp. 681-5, 2001.
- [Donoho ‘06] D. L. Donoho, M. Elad, and V. N. Temlyakov, "Stable recovery of sparse overcomplete representations in the presence of noise," *IEEE Transactions on Information Theory,* vol. 52, pp. 6-18, 2006.
- [Efron ‘04] B. Efron, T. Hastie, I. Johnstone, and R. Tibshirani, "Least Angle Regression," *Ann. Statist.,* vol. 32, pp. 407--499, 2004.
- [Ghodsi ‘06] A. Ghodsi. (2006, Dimensionality Reduction A Short Tutorial.
- [He ‘03] X. He and P. Niyogi, "Locality Preserving Projections," in *Advances in Neural Information Processing Systems 16*, Vancouver, Canada, 2003.
- [Jiang ‘11] Zhuolin Jiang et al., “Learning a Discriminative Dictionary for Sparse Coding via Label Consistent K-SVD”, ICML, 2011.
- [Kanade ‘00] T. Kanade, J. F. Cohn, and T. Yingli, "Comprehensive database for facial expression analysis," in *Proceedings of the Fourth International Conference on Automatic Face and Gesture Recognition, 28-30 March 2000*, Los Alamitos, CA, USA, 2000, pp. 46-53.
- [Kumar ‘08] N. Kumar, P. Belhumeur, and S. Nayar, "FaceTracer: a search engine for large collections of images with faces," in *Computer Vision. 10th European Conference on Computer Vision, ECCV 2008, 12-18 Oct. 2008*, Berlin, Germany, 2008, pp. 340-53.
- [Lai ‘09] Zhihui Lai et al., “Global Sparse Representation Projections for Feature Extraction and Classification”, CCPR, 2009.
- [Koelstra ‘10] S. Koelstra, M. Pantic, and I. Patras, "A dynamic texture-based approach to recognition of facial actions and their temporal models," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 32, pp. 1940-54, 2010.

References (2 of 2)

- [Lucey ‘10] P. Lucey, J. F. Cohn, T. Kanade, J. Saragih, Z. Ambadar, and I. Matthews, "The Extended Cohn-Kanade Dataset (CK+): A complete dataset for action unit and emotion-specified expression," in *2010 CVPR, 13-18 June 2010*, Los Alamitos, CA, USA, 2010, p. 8 pp.
- [Matthews ‘04] J. Matthews and S. Baker, "Active appearance models revisited," *International Journal of Computer Vision,* vol. 60, pp. 135-164, 2004.
- [Murphy ‘09] E. Murphy-Chutorian and M. M. Trivedi, "Head pose estimation in computer vision: A survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 31, pp. 607-626, 2009.
- [Olshausen ‘97] B. A. Olshausen and D. J. Field, "Sparse coding with an overcomplete basis set: a strategy employed by V1?," *Vision Research,* vol. 37, pp. 3311-25, 1997.
- [Qiao ‘09] Lishan Qiao et al., “Sparsity Preserving Projections with Applications to Face Recognition”, PR, 2009.
- [Roweis ‘00] S. T. Roweis and L. K. Saul, "Nonlinear dimensionality reduction by locally linear embedding," *Science,* vol. 290, pp. 2323-6, 2000.
- [Rubinstein ‘08] R. Rubinstein, M. Zibulevsky, and M. Elad, "Efficient Implementationof the K-SVD Algorithm using Batch Orthogonal Matching Pursuit," Technion, Computer Science Dept., Haifa, Israel, 2008.
- [Sherrah ‘01] J. Sherrah, S. Gong, and E. J. Ong, "Face distributions in similarity space under varying head pose," *Image and Vision Computing,* vol. 19, pp. 807-819, 2001.
- [Tenanbaum ‘00] J. B. Tenenbaum, V. de Silva, and J. C. Langford, "A global geometric framework for nonlinear dimensionality reduction," *Science,* vol. 290, pp. 2319-23, 2000.
- [Valstar ‘11] M. Valstar, B. Jiang, M. Mehu, M. Pantic, and K. R. Scherer, "The First Facial Expression Recognition and Analysis Challenge," in *Face and Gesture Recognition*, Santa Barbara, CA, 2011.
- [Viola ‘01] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *Computer Vision and Pattern Recognition*, 2001, pp. I-511-I-518 vol.1.
- [Wright ‘09] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and M. Yi, "Robust face recognition via sparse representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 31, pp. 210-27, 2009.
- [Zafeiriou ‘10] S. Zafeiriou and M. Petrou, "Sparse representations for facial expressions recognition via l1 optimization," in *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Workshops, CVPRW 2010, June 13, 2010 - June 18, 2010*, San Francisco, CA, United states, 2010, pp. 32-39.
- [Zang ‘11] Fei Zang et al., “Discriminative Learning by Sparse Representation for Classification”, Jour. Of Neurocomputing, 2011.
- [Zheng ‘11] Miao Zheng et al., “Grpah Regularized Sparse Coding for Image Representation”, Trans on Imag Proc., 2011.
- [Zhi ‘09] Z. Ruicong and R. Qiuqi, "Discriminant sparse nonnegative matrix factorization," in *2009 IEEE International Conference on Multimedia and Expo (ICME), 28 June-3 July 2009*, Piscataway, NJ, USA, 2009, pp. 570-3.

Thank you!!

Ray Ptucha
rwpeec@rit.edu
raymond.ptucha@kodak.com