Parsimonious Coding & Verification of Offline Handwritten Signatures

Presenter: ***Elias N. Zois***

Telecommunications and Signal Processing Laboratory, Technological and Educational Institute of Athens, Greece.

ezois@teiath.gr

Coauthors: Ilias Theodorakopoulos, Dimitrios Tsourounis & George Economou, Physics Dept., University of Patras, Greece

http://www.upcv.upatras.gr/

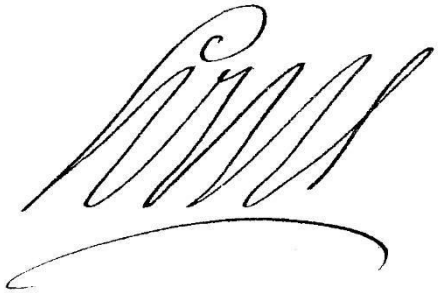
CVPR Biometrics Workshop 2017

21 July 2017

Contents

Signatures

- Signatures: A common (behavioral) way to declare our identity or our consent.
  - The outcome of the joint interaction between a person's specific motoric procedure and his/hers taught scripting customs.
- Taxonomy:
  - On-line
    - Sequence information
    - Considered to be easier
  - Off-line
    - No sequence information
    - Digital image – Gray level
- Applications: Forensics, Security, e-business, e.t.c.



Elizabeth
Elizabeth R
Diana.
June : 1997.
Joannes Paulus PP II
John D. Rockefeller.
Warren E Buffett
steven jobs
Bill Gates
Jeff P Bezos
Sergey Brin
Mark Zuckerberg, President
di leonardo da vinci
Michelagniolo
Picasso
William Shakspere
Mozart
Harry Houdini
A. Einstein
Bruce Lee
Marilyn Monroe
Muhammad Ali
DIEGO
(10)
Charlie Chaplin
John Lennon
Elvis Presley
Michael Jackson
G. Washington
John Hancock
Abraham Lincoln
John F. Kennedy
Hillary Rodham Clinton
Martin Luther King Jr
Winston S. Churchill
Harry Truman
M K Gandhi
M A Jinnah
Napoleon
محمود احمدی نژاد

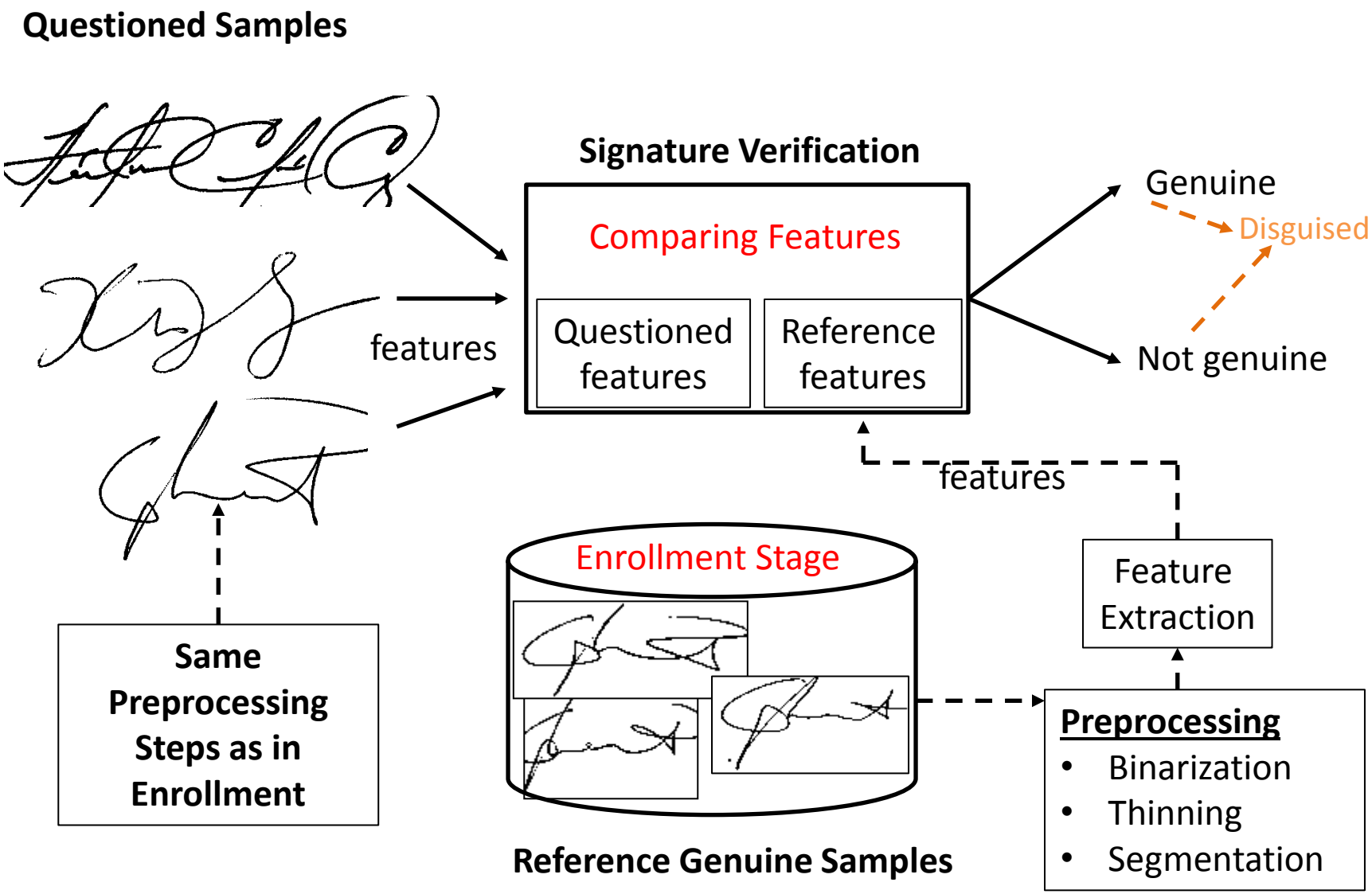
Terminology and Taxonomy

Derived from:

M. I. Malik, L. Alewijnse, M. Liwicki, M. Blumenstein: AFHA 2013

- **Genuine** = natural execution of signature
  - Regular
  - Unnatural situation = altered behavior by intrinsic (e.g., disease, drugs) or extrinsic (e.g., writing on a wall)
- **Forgery** (only non-specimen w. is trying to copy) & (specimen's signature is known)
  - Simple => Non-practiced => not much training => one attempt
  - Skilled => Practiced => after practicing, has to be specified
- **Disguise (intention vs. behavior) by specimen writer.**
  - Auto-simulated (intend for the signature to appear to be a forgery)
- **Fictitious signatures** = develop a new signature
  - Formerly called random signatures

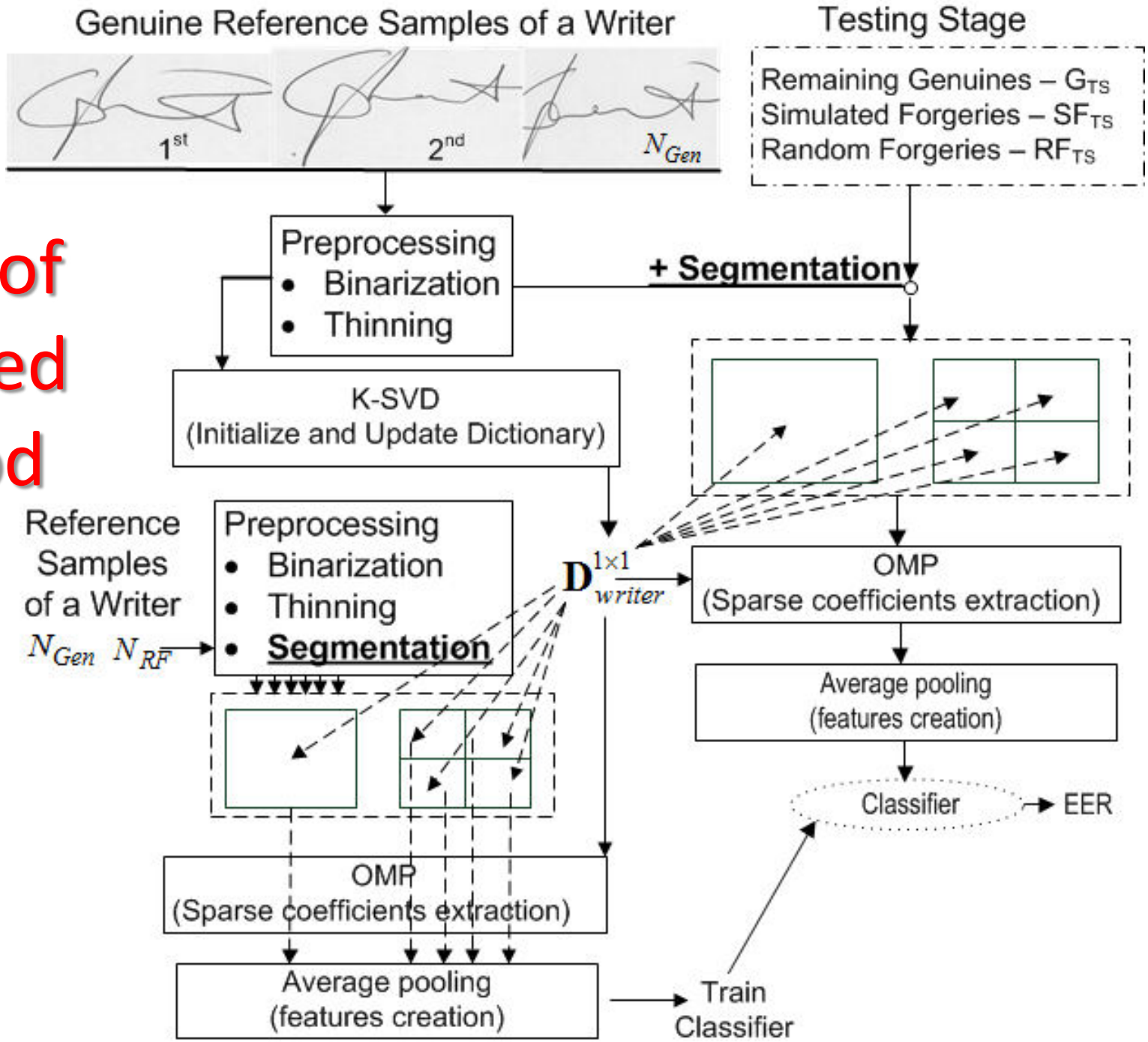
Generic SV system



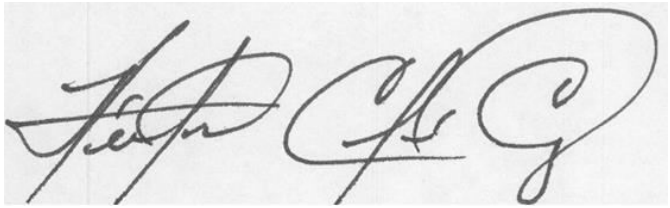
What we Propose

- Make use of Sparse Representation & Dictionary Learning techniques for creating the “characteristic profile space” of a writer.
  - Employ K-SVD.
  - Create the “atoms” of the writer i.e. the “**MODEL**”
  - Use a few (e.g. 5) samples to construct his/hers lexicon.
- Use this “lexicon” for any signature which claims that it originates from this specific writer
  - Use OMP to create sparse coefficients
  - Average pooling as feature extractor.

Sketch of Proposed Method

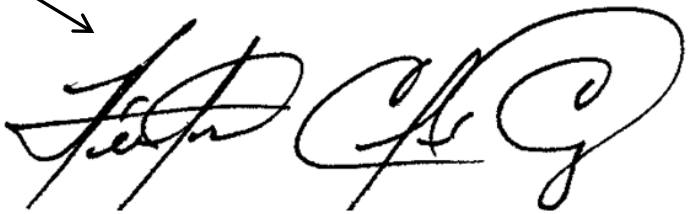


Preprocessing



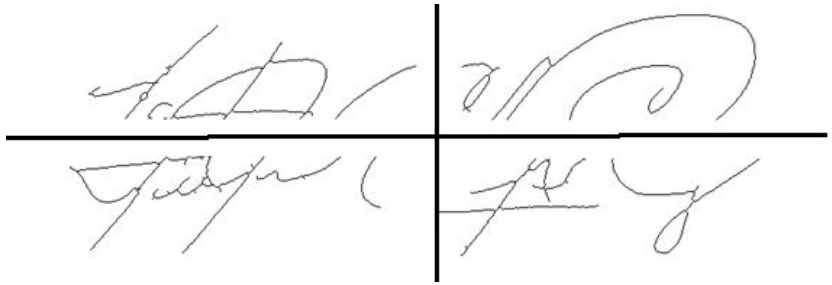
Original Image

- Binarization

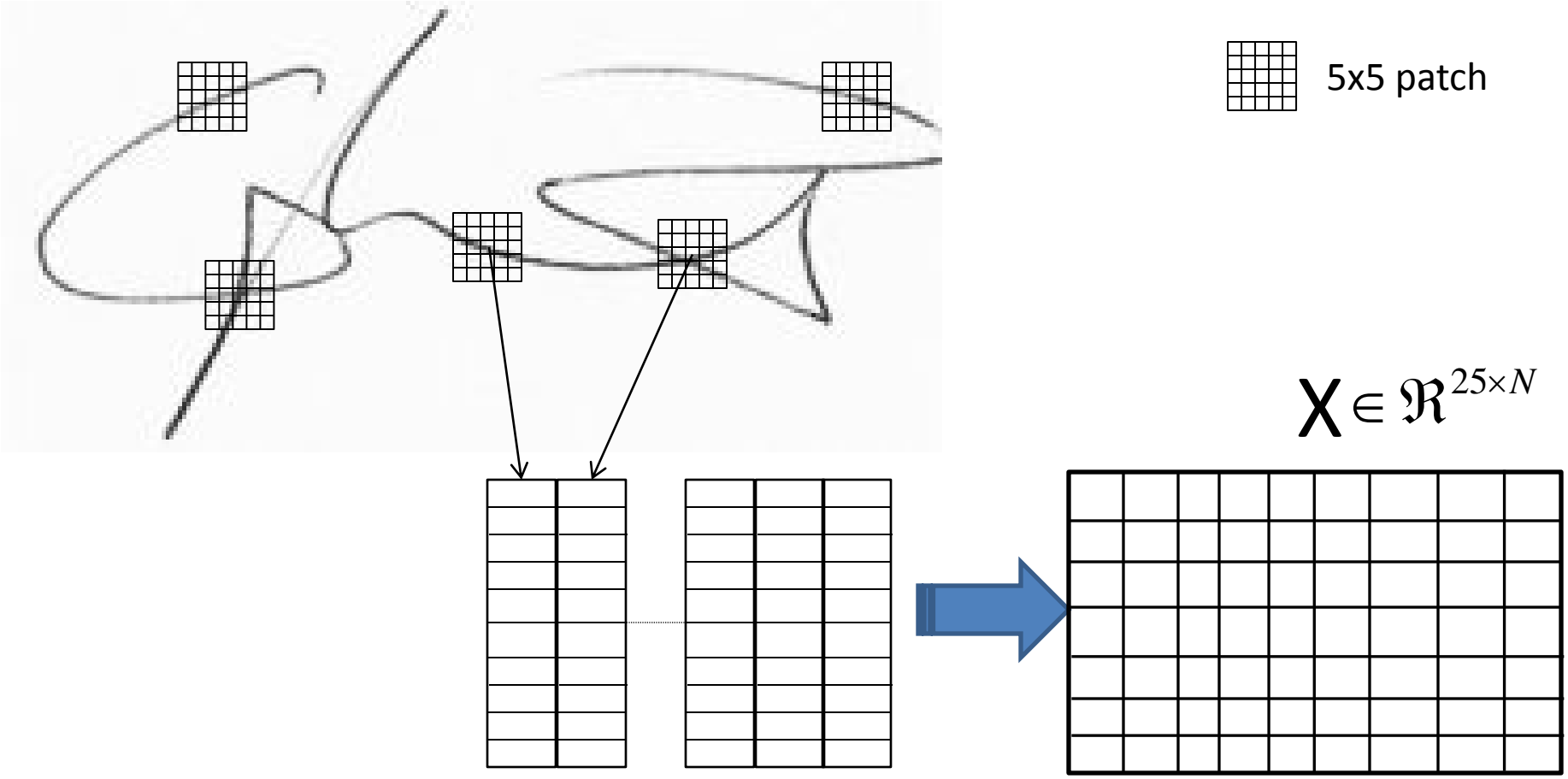


- Thinning

- Equimass Segmenting

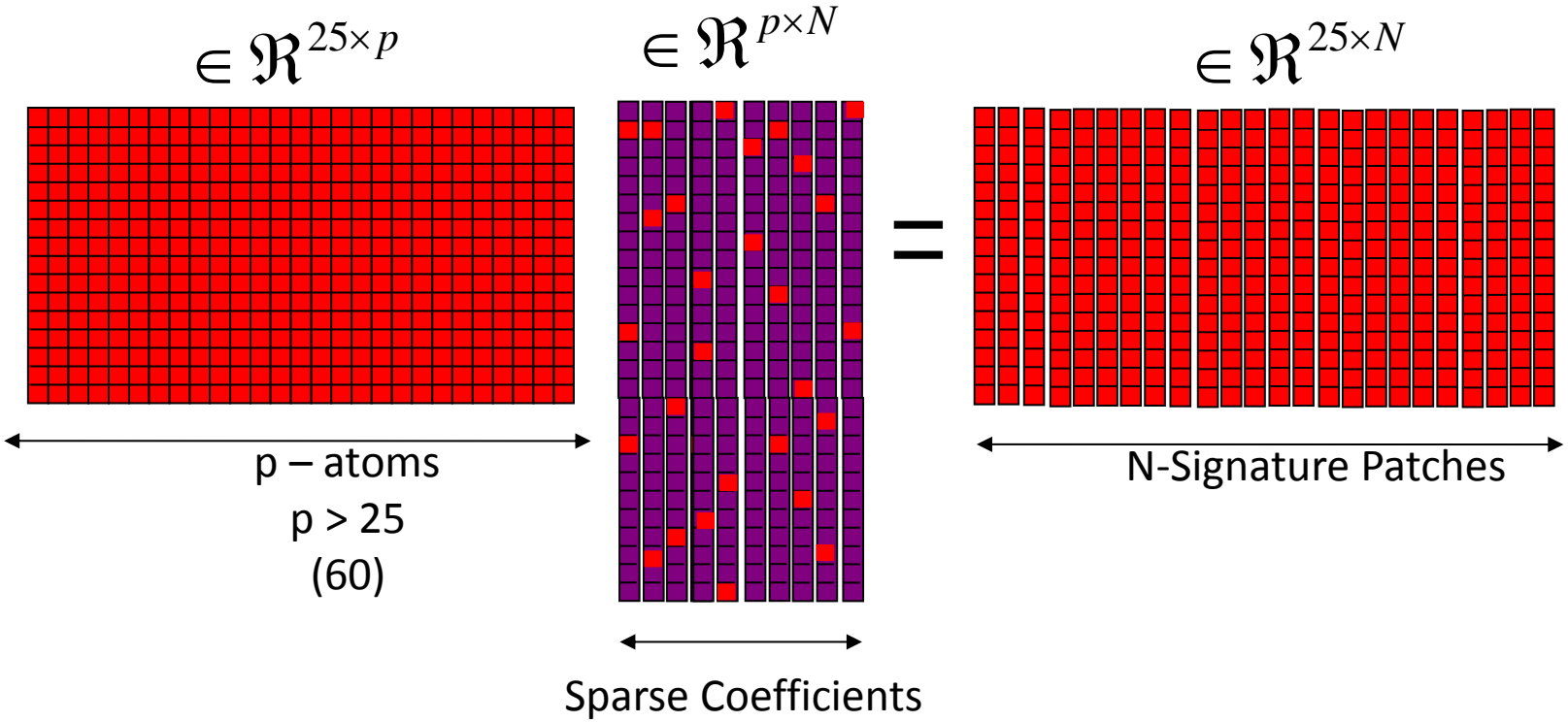


Create patch matrix X

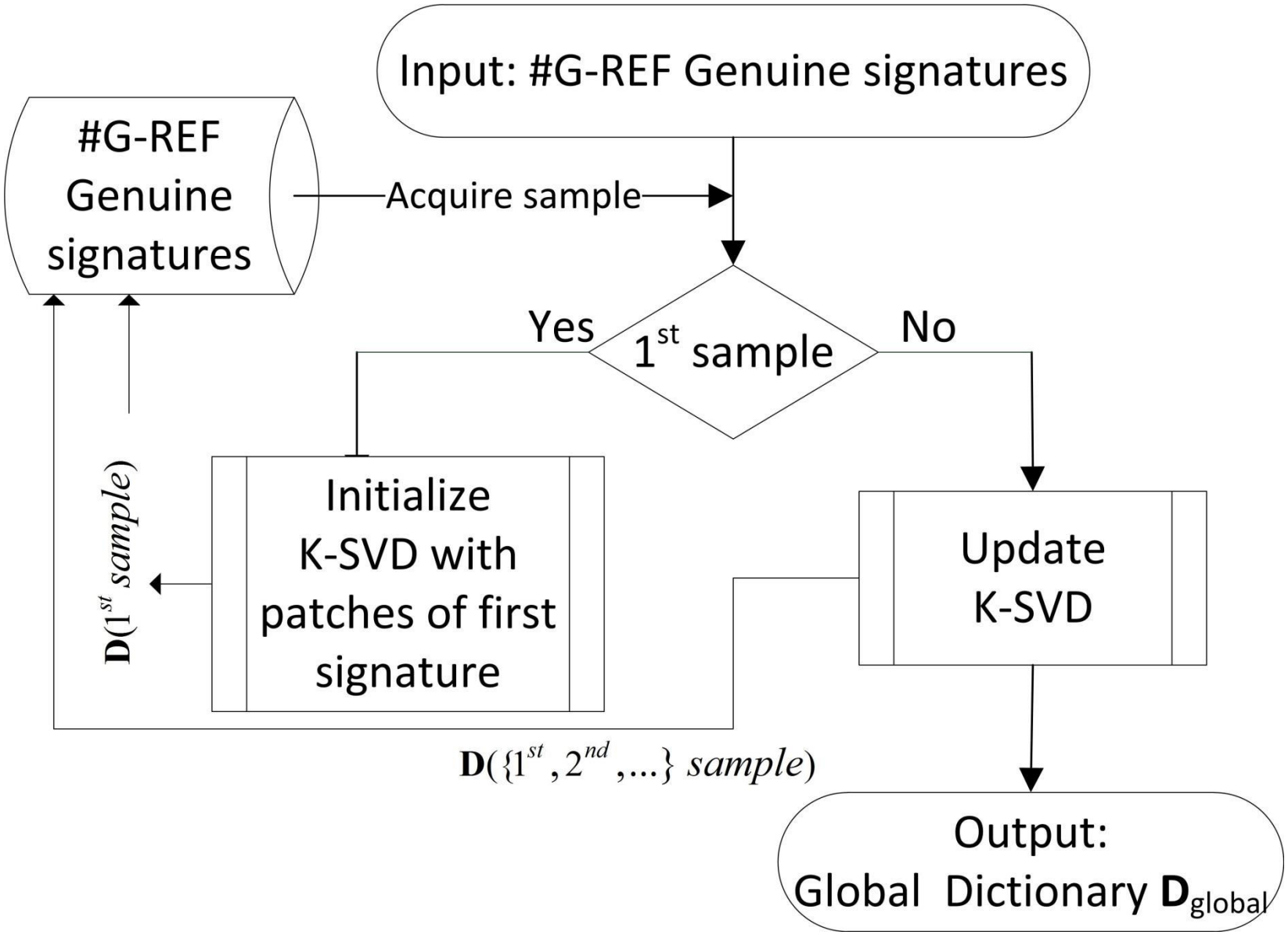


Modeling a Writer:
Dictionary Learning – Sparse Representation

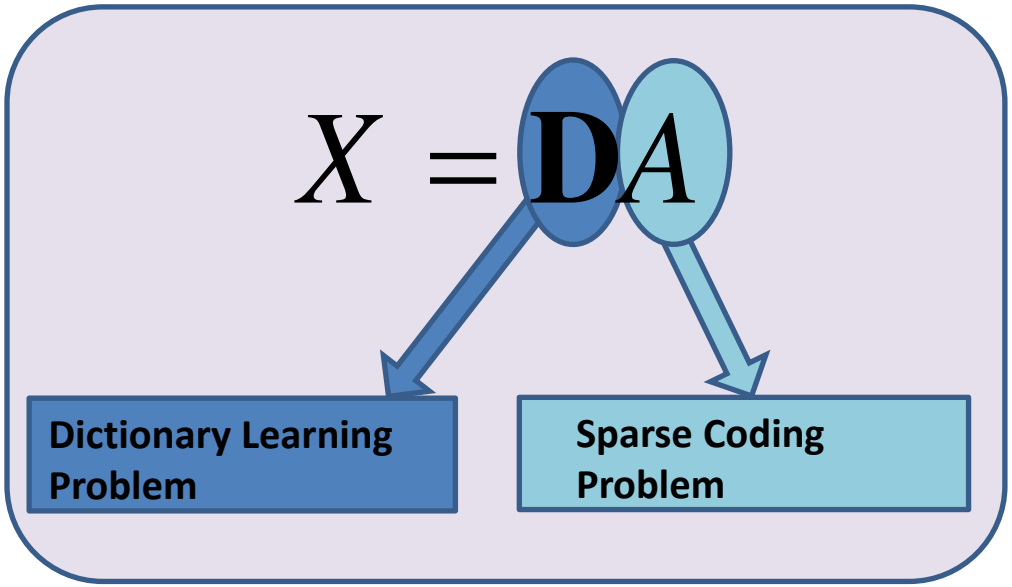
$$\mathbf{D} \times \mathbf{A} = \mathbf{X}$$



Modeling a Writer



Modeling a Writer



$$\min_{\mathbf{D},\mathbf{A}} \left\{ \left\| \mathbf{X} - \mathbf{D}\mathbf{A} \right\|_F^2 \right\}$$

$$\text{s.t.} \quad \left\| \mathbf{a}_i \right\|_0 \le T_0 \,, \quad i = 1, \ldots, M$$

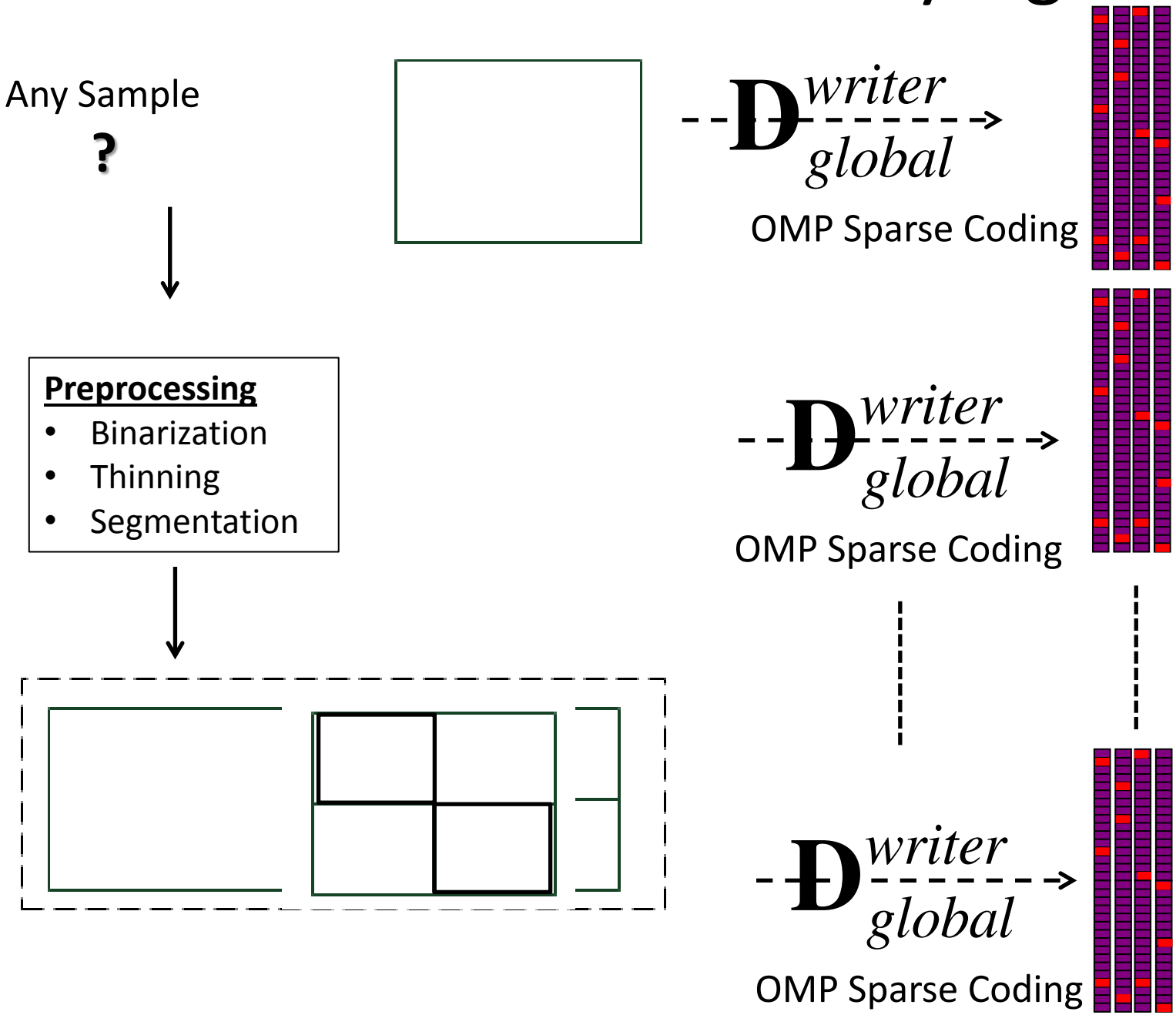
K-SVD for Dictionary Learning

$$\mathbf{D}_{global}^{writer}$$

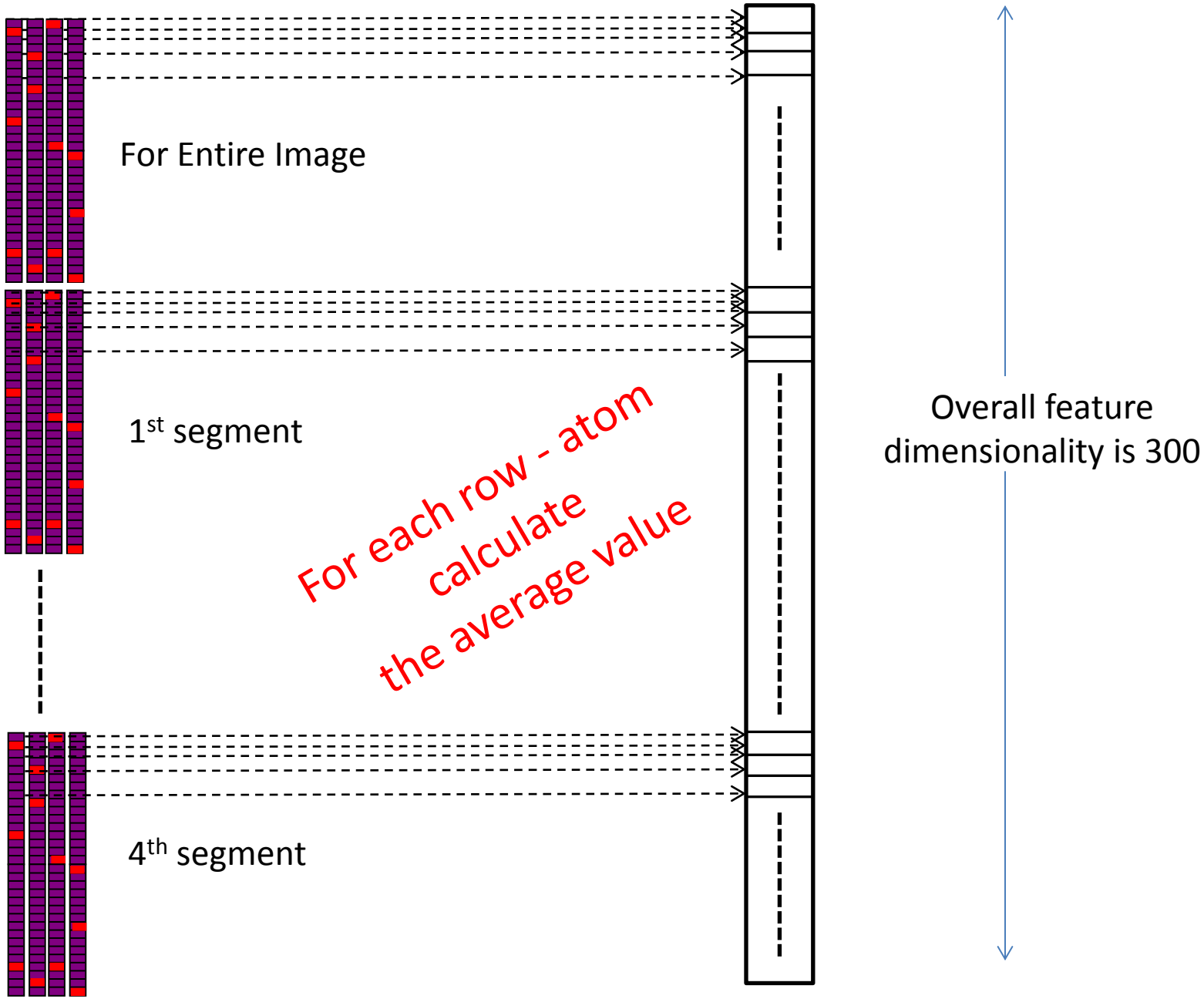
Operating Parameters

- T_0=3: How many non- zero elements for each a_i
- Patch size = 25 (5x5)
- Number of Atoms (p) = 60, p>25.
- Number of iterations = 50
- Use of K-SVDbox & OMPBox toolbox
  - *by Ron Rubinstein, Computer Science Department Technion -- Israel Institute of Technology* *http://www.cs.technion.ac.il/~ronrubin/software.html*

Feature Extraction for Any Signature



Feature Extraction for Any Signature



Experimental Procedure - Verification

- **Datasets**
  - CEDAR: 55 writers, 24 genuine, 24 forgeries
  - MCYT: 75 writers, 15 genuine, 15 forgeries
  - GPDS300, 300 writers, 30 genuine, 24 forgeries
- **Training set: for each writer**
  - 5 genuine reference samples
  - 1 genuine sample from 10 other writers = 10 RF samples
- **Testing set**
  - The remaining genuine samples
  - The remaining skilled forgery samples
  - Genuine samples from other writers **which did not participate during training**

Verification

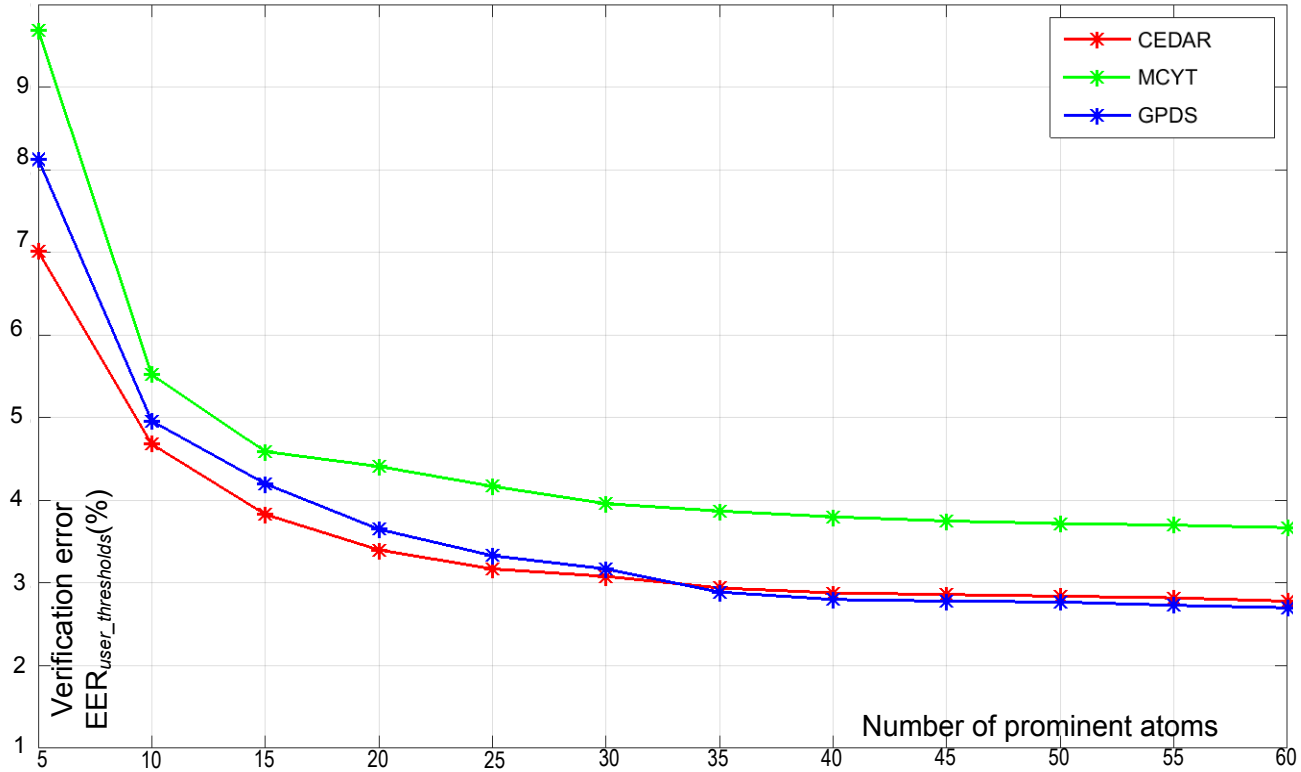
10
Repetitions

- **Classifier**
  - SVM with RBF kernel
  - Training set is used for selecting optimal SVM parameters with AUC metric (cross validation).
  - LibSVM library
- **Testing Stage metrics**
  - A function of the score that the SVM outputs.
  - **Test:** Remaining genuine Vs. skilled forgeries
    - FAR, FRR, EER.
    - @ EER threshold calculate random forgery error rate.
    - Per user (writer) EER and global EER
    - Results with hard decisions also: FAR & FRR

Results

Dataset	FRR	FAR_{random}	$FAR_{skilled}$	$EER_{global\_threshold}$	$EER_{user\_threshold}$
CEDAR	7.32	0.34	6.83	7.58	2.78
MCYT-75	7.32	0.35	10.4	8.43	3.67
GPDS300	6.91	0.33	6.22	7.21	2.70

← Hard Decisions →



Comparisons

First Author	Dataset	Method	# Genuine Sigs. for training	EER
Kumar R.	CEDAR	Signature Morphology	24	11.6
Kumar R.	CEDAR	Surroundness	24	8.33
Kalera	CEDAR	Gradient, Structural and Concavity	16	21.9
Zois	CEDAR	Partially Ordered Sets	5	4.12
Proposed	**CEDAR**	**K-SVD dictionary learning – OMP**	**5**	**2.78**
Vargas	MCYT75	Local binary patterns (LBP)	5	11.3
Zois	MCYT75	Partially Ordered Sets	5	6.02
Alonso-Fernandez	MCYT75	Slant and envelope	5	22.4
Fierrez-Aguilar	MCYT75	Global and local slant	5	11.0
Wen	MCYT75	Invariant ring peripheral	5	15.0
Yin Ooi	MCYT75	Discrete Radon transform	5	13.9
Soleimani	MCYT75	Deep multitask metric	5	13.44
Proposed	**MCYT75**	**K-SVD dictionary learning – OMP**	**5**	**3.67**
Hu	GPDS 150	LBP & HOG & GLCM	10	7.66
Yilmaz	GPDS 160	Local binary patterns	12	9.64
Yilmaz	GPDS 160	LBP & HOG	12	9.67
Pirlo	GPDS 300	Cosine similarity	12	7.20
Pirlo	GPDS 300	Optical flow	6	4.60
Parodi	GPDS 300	Circular Grid	13	4.21
Kumar	GPDS 300	Surroundness	24	13.8
Zois	GPDS 300	Partially Ordered Sets	5	5.48
Hafemann	GPDS 300	Deep Convolutional N. N.	5	4.53
Hafemann	GPDS 300	Deep Convolutional N. N.	14	3.47
Proposed:	**GPDS 300**	**K-SVD dictionary learning – OMP**	**5**	**2.70**

Conclusions

- Offline Signature Verification Problem
- Using Dictionary Learning and Sparse Representation for building a writer's model and analyze any other signature which claims the writer's identity.
- Propose Average pooling as feature extractor
- Testing in three popular datasets provides state of the art results

Many thanks to our… minions…

- Experiments were made with:
- FTS Dual Xeon Server equiped with:
  - 24 threads
  - 100GB RAM

Thank you