



**ΕΘΝΙΚΟ ΚΑΙ ΚΑΠΟΔΙΣΤΡΙΑΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ**

**ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ  
ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΤΗΛΕΠΙΚΟΙΝΩΝΙΩΝ  
ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ**

**ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ**

**Αποθρομβοποίηση εικόνας μέσω του αλγορίθμου k-SVD**

**Βασίλειος Ε. Μαρκούτης**

**Επιβλέπων: Σέργιος Θεοδωρίδης, Καθηγητής ΕΚΠΑ**

**ΑΘΗΝΑ**

**ΟΚΤΩΒΡΙΟΣ 2016**

## **ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ**

Αποθρομβοποίηση εικόνας μέσω του αλγορίθμου k-SVD

**Βασίλειος Ε. Μαρκούτης**

**A.M.: M1346**

**ΕΠΙΒΛΕΠΩΝ: Σέργιος Θεοδωρίδης, Καθηγητής ΕΚΠΑ**

## ΠΕΡΙΛΗΨΗ

*“Η επιστήμη είναι δεδομένα. Ακριβώς όπως τα σπίτια χτίζονται με τούβλα, έτσι και η επιστήμη χτίζεται με δεδομένα. Αλλά όπως ένας σωρός τούβλα δεν κάνει ένα σπίτι, έτσι και μια συλλογή δεδομένων δεν είναι απαραίτητα επιστήμη.”*  
Ανρί Πουανκαρέ

Οι εφαρμογές αλγορίθμων μηχανικής μάθησης σε κλασικά προβλήματα της θεωρίας Επεξεργασίας Σήματος αποτελεί μια περιοχή με τεράστιο ερευνητικό ενδιαφέρον στις σημερινές ημέρες. Στην εποχή των “Big Data”, η αναπαράσταση των τεράστιων όγκων δεδομένων με συναρτήσεις λίγων μεταβλητών αποτελεί μία μεγάλη πρόκληση για τους επιστήμονες της πληροφορικής και των μαθηματικών. Θέτοντας το θέμα στη γεωμετρική του σκοπιά, στόχος είναι η αναπαράσταση των δεδομένων από έναν χώρο πολλών διαστάσεων σε έναν χώρο με λίγες διαστάσεις χωρίς να χαθεί σημαντική ποσότητα πληροφορίας. Η εν λόγω διαδικασία, γνωστή ως μείωση της διάστασης (dimensionality reduction), αποτελεί εκ φύσεως ένα δύσκολο πρόβλημα.

Όσο και αν φαίνεται περίεργο, οι νόμοι της φύσης τείνουν να εφαρμοστούν ακόμη και στην επεξεργασία μεγάλων όγκων δεδομένων. Η φύση είναι “φειδωλή”, απορρίπτει τα περιττά στοιχεία και διατηρεί τα απολύτως αναγκαία. Επιστρέφοντας στον κόσμο των μαθηματικών και της πληροφορικής, το μαθηματικό εργαλείο που εκφράζει αυτή τη “λιτότητα” της φύσης είναι η αραιότητα (sparsity). Η έκφραση μιας μεγάλης ποσότητας δεδομένων με τη βοήθεια διανυσμάτων με λίγους συντελεστές και πολλά μηδενικά καθώς και με τη βοήθεια κατάλληλων μητρώων έχει προσελκύσει έντονο ερευνητικό ενδιαφέρον. Η “αραιή” έκφραση των δεδομένων αποκαλύπτει συγκεκριμένα μοτίβα στα δεδομένα τα οποία είναι δύσκολο να αποκαλυφθούν με κλασικούς αλγορίθμους συσχέτισης. Η αραιότητα δείχνει ότι σύνολα δεδομένων που φαίνονται εντελώς τυχαία, κρύβουν σχέσεις συσχέτισης.

Αυτό που επίσης είναι εντυπωσιακό, είναι ότι μπορούμε να εκμεταλλευτούμε αυτή την παρουσία αραιών μορφών προκειμένου να απορρίψουμε τα ανεπιθύμητα δεδομένα που συνιστούν την έννοια του θορύβου. Ο θόρυβος ενυπάρχει σε όλα τα προβλήματα του πραγματικού κόσμου. Η ανίχνευση και η απόρριψή του αποτελεί ένα αντίστροφο πρόβλημα που είναι αδύνατο να λυθεί με ακρίβεια. Εντούτοις, η ύπαρξη αραιών μοντέλων για την έκφραση συνόλων δεδομένων μπορεί να απορρίψει με τη χρήση κατάλληλων αλγορίθμων τη μεγαλύτερη ποσότητα θορύβου.

Η μηχανική μάθηση, από την άλλη πλευρά, συνίσταται σε αλγορίθμους που “εκπαιδεύουν” ένα σύστημα να εκτελέσει συγκεκριμένες διαδικασίες ανάλογα με τις παραμέτρους του προβλήματος. Υπό αυτό το πρίσμα, η αραιή έκφραση ενός συνόλου δεδομένων μπορεί να γίνει πιο αποτελεσματική αν το μητρώο που θα χρησιμοποιηθεί ως βάση για την αναπαράσταση των δεδομένων προσαρμοστεί στα ίδια τα δεδομένα με τη βοήθεια αλγορίθμων μηχανικής μάθησης. Η τεχνική αυτή, γνωστή ως εκμάθηση λεξικού (dictionary learning) αποτελεί ένα βήμα πιο μακριά στους αλγορίθμους επίλυσης αντίστροφων προβλημάτων.

Στην παρούσα εργασία, μελετάται η εφαρμογή ενός από τους διαδεδομένους αλγορίθμους εκμάθησης λεξικού, του k-SVD, στην αποθορυβοποίηση σήματος ψηφιακής εικόνας για διάφορα είδη θορύβου. Ο συνδυασμός του αλγορίθμου k-SVD καθώς και αλγορίθμων αναπαράστασης σήματος σε αραιή μορφή, μπορεί να πετύχει, κατά περίπτωση, αρκετά ικανοποιητικά αποτελέσματα. Η εφαρμογή των αλγορίθμων πάνω σε ψηφιακές εικόνες, γίνεται για καθαρά λόγους οπτικοποίησης των αποτελεσμάτων. Συνεπώς, μπορούν να

εφαρμοστούν για την αποθρομβοποίηση οποιουδήποτε σήματος κάτω από την κατάλληλη επιλογή παραμέτρων.

**ΘΕΜΑΤΙΚΗ ΠΕΡΙΟΧΗ:** Επεξεργασία Εικόνας

**ΛΕΞΕΙΣ ΚΛΕΙΔΙΑ:** υπερπλήρη λεξικά, αποθρομβοποίηση, μηχανική μάθηση, αναπαράσταση με αραιά διανύσματα, διάσπαση μητρώων



## ABSTRACT

*"Science is built up with facts, as a house is with stones. But a collection of facts is no more a science than a heap of stones is a house. "*  
*Henri Poincaré*

Machine learning algorithms can effectively apply on signal processing problems in a scientific area with wide research interest. One of the big data challenges is to represent large sets of data with mathematical functions consisted of a few variables. Viewing the geometric aspects of the problem, the goal is the representation of data from a higher to a lower subspace without the loss of a large amount of information. This procedure is called dimensionality reduction and it is a difficult problem.

Surprisingly, nature's laws tend to fit well on the big data processing. Nature rejects every unnecessary element expressing with this way its parsimony. In mathematical and information science world, this minimalism of nature corresponds to the concept of sparsity. The representation of large sets of data with sparse models reveals relationships between the elements of data and confirms that data on these sets are not randomly distributed.

The remarkable in sparse models, is that sparse representations can reject the useless data defined as noise. Noise is present in every real life problem. The detection and rejection of noise refers to an inverse problem which is infeasible to be solved precisely. However, suitable algorithms using sparse models can reject a large amount of noise on a set of data.

On the other hand, machine learning consists of algorithms which can train a system in order to execute specific procedures depending on the parameters of the problem. From this point of view, sparse representations of a large set of data can be more effective if the dictionary which is used for the sparse model can be resulted from the existing data after the appliance of machine learning algorithms. This technique is called dictionary learning and has introduced modern aspects to the solutions of inverse problems.

In the work at hand, the k-SVD algorithm, which is one of the most famous dictionary learning algorithms, is studied in image denoising problems for different kinds of noise. The k-SVD algorithm in combination with sparse representations algorithms can achieve quite competitive results in image denoising. Digital images are only used for the visualization of the results. Consequently, these algorithms can be used and in other types of signals by the appropriate choice of parameters.

**SUBJECT AREA:** Image Processing

**KEYWORDS:** overcomplete dictionaries, denoising, machine learning, sparsity, matrix decomposition

*Στη σύζυγό μου Μαρία, στην κόρη μου Ευαγγελία και στο γιο μου Επαμεινώνδα.*

## ΕΥΧΑΡΙΣΤΙΕΣ

Θα ήθελα να ευχαριστήσω τον καθηγητή κ. Σέργιο Θεοδωρίδη, επιβλέποντα της παρούσας εργασίας, για τη συμβολή του στην εκπόνηση αυτής. Είναι ένας καθηγητής που “εμπνέει” και αυτό δυστυχώς είναι σπάνιο στις μέρες που ζούμε. Επίσης θα ήθελα να ευχαριστήσω τον κ. Ιωάννη Κοψίνη για τη βοήθειά του, τις πολύτιμες συμβουλές του και τις ιδέες του πάνω στις οποίες στηρίχθηκε το μεγαλύτερο μέρος της παρούσας εργασίας. Τέλος, θα ήθελα να εκφράσω τις ευχαριστίες μου στη σύζυγό μου Μαρία, για την υπομονή που έδειξε όλο αυτό το χρονικό διάστημα.

## ΠΕΡΙΕΧΟΜΕΝΑ

|                                                                 |           |
|-----------------------------------------------------------------|-----------|
| <b>ΠΡΟΛΟΓΟΣ</b>                                                 | <b>15</b> |
| <b>1. ΑΝΤΙΣΤΡΟΦΑ ΠΡΟΒΛΗΜΑΤΑ</b>                                 | <b>16</b> |
| 1.1 Εισαγωγή                                                    | 16        |
| 1.2 Γραμμικά αντίστροφα προβλήματα                              | 16        |
| 1.3 Παλινδρόμηση Κορυφής (Ridge Regression)                     | 18        |
| <b>2. ΚΥΡΤΗ ΒΕΛΤΙΣΤΟΠΟΙΗΣΗ</b>                                  | <b>20</b> |
| 2.1 Το πρόβλημα                                                 | 20        |
| <b>3. ΔΙΑΣΠΑΣΗ ΜΗΤΡΩΟΥ ΜΕΣΩ ΤΩΝ ΙΔΙΑΖΟΥΣΩΝ ΤΙΜΩΝ ΤΟΥ</b>        | <b>22</b> |
| 3.1 Η διάσπαση SVD                                              | 22        |
| 3.2 Η μείωση της διάστασης                                      | 24        |
| <b>4. ΑΝΑΠΑΡΑΣΤΑΣΗ ΣΗΜΑΤΟΣ ΣΕ ΣΥΜΠΙΕΣΤΗ ΜΟΡΦΗ</b>               | <b>26</b> |
| 4.1 Εισαγωγή                                                    | 26        |
| 4.2 Διανυσματικοί χώροι και στάθμες διανυσμάτων                 | 26        |
| 4.3 Βάσεις                                                      | 28        |
| 4.4 Αραιά μοντέλα                                               | 30        |
| 4.5 Μοναδικότητα της αραιότερης λύσης                           | 31        |
| <b>5. Ο ΑΛΓΟΡΙΘΜΟΣ OMP</b>                                      | <b>34</b> |
| 5.1 Ο αλγόριθμος OMP                                            | 34        |
| 5.2 Υπερπλήρες λεξικό με βάση το δισδιάστατο μετασχηματισμό DCT | 35        |
| 5.3 Αποθορυβοποίηση εικόνας με βάση τον αλγόριθμο OMP           | 35        |
| <b>6. Ο ΑΛΓΟΡΙΘΜΟΣ K-SVD</b>                                    | <b>39</b> |
| 6.1 Εκμάθηση Λεξικού (Dictionary Learning)                      | 39        |
| 6.2 Ο αλγόριθμος k-SVD                                          | 41        |
| <b>7. Ο ΑΛΓΟΡΙΘΜΟΣ K-SVD ΣΤΗΝ ΑΠΟΘΟΥΒΟΠΟΙΗΣΗ ΣΗΜΑΤΟΣ</b>        | <b>45</b> |
| 7.1 Η μετρική PSNR                                              | 45        |
| 7.2 Η μέθοδος των επικαλυπτόμενων περιοχών                      | 45        |
| 7.3 Αποθορυβοποίηση σήματος με gaussian θόρυβο                  | 46        |

|     |                                                                             |    |
|-----|-----------------------------------------------------------------------------|----|
| 7.4 | Αποθορυβοποίηση σήματος με poisson θόρυβο . . . . .                         | 49 |
| 7.5 | Αποθορυβοποίηση σήματος με θόρυβο ομοιόμορφης κατανομής . . . . .           | 49 |
| 7.6 | Image inpainting . . . . .                                                  | 52 |
| 7.7 | Αποθορυβοποίηση σήματος με θόρυβο t-κατανομής . . . . .                     | 55 |
| 7.8 | Αποθορυβοποίηση σήματος με παλμικό θόρυβο (impulse noise) . . . . .         | 57 |
| 7.9 | Αποθορυβοποίηση σήματος με θόρυβο μικτού τύπου (mixed type noise) . . . . . | 61 |
| 8.  | ΣΥΜΠΕΡΑΣΜΑΤΑ . . . . .                                                      | 68 |
|     | ΠΙΝΑΚΑΣ ΟΡΟΛΟΓΙΑΣ . . . . .                                                 | 69 |
|     | ΣΥΝΤΜΗΣΕΙΣ, ΑΡΚΤΙΚΟΛΕΞΑ ΚΑΙ ΑΚΡΩΝΥΜΙΑ . . . . .                             | 70 |
|     | ΑΝΑΦΟΡΕΣ . . . . .                                                          | 71 |

## ΚΑΤΑΛΟΓΟΣ ΣΧΗΜΑΤΩΝ

|           |                                                                                                                                                                                                         |    |
|-----------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Σχήμα 1:  | Ένα μαθηματικό μοντέλο για τα αντίστροφα προβλήματα. . . . .                                                                                                                                            | 16 |
| Σχήμα 2:  | Το ευθύ πρόβλημα (Α) της εύρεσης του αθροίσματος κάποιων αριθμών και το αντίστροφο (Β) της εύρεσης κάποιων αριθμών που το άθροισμά τους είναι 27. Το δεύτερο επιδέχεται, φυσικά, πολλών λύσεων. . . . . | 17 |
| Σχήμα 3:  | Το μοντέλο γραμμικής παλινδρόμησης στη μία διάσταση. . . . .                                                                                                                                            | 18 |
| Σχήμα 4:  | Μία κυρτή συνάρτηση μιας διάστασης. Κάθε ευθύγραμμο τμήμα που ενώνει δύο σημεία της συνάρτησης κείται πάνω από το γράφημα της συνάρτησης . . . . .                                                      | 21 |
| Σχήμα 5:  | Η SVD παραγοντοποίηση ενός $m \times l$ μητρώου $X$ βαθμού . . . . .                                                                                                                                    | 23 |
| Σχήμα 6:  | Τα ιδιάζοντα διανύσματα $\mathbf{v}_1$ και $\mathbf{v}_2$ προκύπτουν από 1000 τυχαία σημεία μέσα σε μία έλλειψη. Τα δύο διανύσματα τείνουν να ταυτιστούν με τους δύο άξονες της έλλειψης . . . . .      | 23 |
| Σχήμα 7:  | Ένα σύνολο σημείων του τρισδιάστατου χώρου βρίσκεται “περίπου” πάνω σε μια καμπύλη κύκλου και επομένως η ενδογενής διάσταση του συνόλου προσεγγίζει τη μονάδα . . . . .                                 | 24 |
| Σχήμα 8:  | Η συνεισφορά ενός στοιχείου $x$ , ήτοι $ x ^p$ , στην τιμή της στάθμης $\ell_p$ του διανύσματος $\mathbf{x}$ , για διάφορες τιμές του $p$ . . . . .                                                     | 28 |
| Σχήμα 9:  | Αναπαράσταση ενός σήματος με τη βοήθεια ενός υπερπλήρους λεξικού και ενός αραιού διανύσματος . . . . .                                                                                                  | 29 |
| Σχήμα 10: | Αραιή αναπαράσταση μιας εικόνας. (α) Αρχική εικόνα. (β) Εικόνα που προκύπτει με τη διατήρηση του 1% των συντελεστών του μετασχηματισμού κυματιδίου . . . . .                                            | 30 |
| Σχήμα 11: | Το 2D-DCT υπερπλήρες λεξικό . . . . .                                                                                                                                                                   | 36 |
| Σχήμα 12: | Αραιή αναπαράσταση μία εικόνας με τη βοήθεια του 2D-DCT υπερπλήρους λεξικού 196 ατόμων και του αλγορίθμου OMP με 10 συντελεστές . . . . .                                                               | 36 |
| Σχήμα 13: | Αραιή αναπαράσταση μία εικόνας με τη βοήθεια του 2D-DCT υπερπλήρους λεξικού 196 ατόμων και του αλγορίθμου OMP με 3 συντελεστές . . . . .                                                                | 37 |
| Σχήμα 14: | Αποθρομβοποίηση μία εικόνας με τη βοήθεια του 2D-DCT υπερπλήρους λεξικού 196 ατόμων και του αλγορίθμου OMP με 5 συντελεστές . . . . .                                                                   | 37 |
| Σχήμα 15: | Αποθρομβοποίηση μία εικόνας με τη βοήθεια του 2D-DCT υπερπλήρους λεξικού 196 ατόμων και του αλγορίθμου OMP με 15 συντελεστές . . . . .                                                                  | 38 |

|           |                                                                                                                                              |    |
|-----------|----------------------------------------------------------------------------------------------------------------------------------------------|----|
| Σχήμα 16: | Το πρόβλημα Compressed Sensing (CS) . . . . .                                                                                                | 39 |
| Σχήμα 17: | Το πρόβλημα Dictionary Learning (DL) . . . . .                                                                                               | 39 |
| Σχήμα 18: | Μία σχηματική αναπαράσταση του αλγορίθμου k-SVD . . . . .                                                                                    | 43 |
| Σχήμα 19: | Η κατανομή του gaussian θορύβου διασποράς $\sigma=20$ πάνω στο σήμα μιας ψηφιακής εικόνας 8 bits . . . . .                                   | 46 |
| Σχήμα 20: | Το λεξικό που προκύπτει από την εφαρμογή του αλγορίθμου k-SVD σε εικόνα επηρεασμένη από gaussian θόρυβο . . . . .                            | 47 |
| Σχήμα 21: | Σχηματική αναπαράσταση της εφαρμογής του αλγορίθμου k-SVD στην περίπτωση αποθορυβοποίησης εικόνας επηρεασμένης από gaussian θόρυβο . . . . . | 48 |
| Σχήμα 22: | Η εφαρμογή του αλγορίθμου k-SVD σε μία εικόνα με gaussian θόρυβο διασποράς $\sigma=10$ . . . . .                                             | 48 |
| Σχήμα 23: | Η εφαρμογή του αλγορίθμου k-SVD σε μία εικόνα με gaussian θόρυβο διασποράς $\sigma=20$ . . . . .                                             | 49 |
| Σχήμα 24: | Η κατανομή του poisson θορύβου διασποράς $\sigma=20$ πάνω στο σήμα μιας ψηφιακής εικόνας 8 bits . . . . .                                    | 50 |
| Σχήμα 25: | Η εφαρμογή του αλγορίθμου k-SVD σε μία εικόνα με poisson θόρυβο(1) . . . . .                                                                 | 50 |
| Σχήμα 26: | Η εφαρμογή του αλγορίθμου k-SVD σε μία εικόνα με poisson θόρυβο(2) . . . . .                                                                 | 50 |
| Σχήμα 27: | Η κατανομή του θορύβου ομοιόμορφης κατανομής διασποράς πάνω στο σήμα μιας ψηφιακής εικόνας 8 bits . . . . .                                  | 51 |
| Σχήμα 28: | Η εφαρμογή του αλγορίθμου k-SVD σε μία εικόνα με θόρυβο ομοιόμορφης κατανομής $[-20,20]$ . . . . .                                           | 51 |
| Σχήμα 29: | Η εφαρμογή του αλγορίθμου k-SVD σε μία εικόνα με θόρυβο ομοιόμορφης κατανομής $[-40,40]$ . . . . .                                           | 52 |
| Σχήμα 30: | Η κατανομή θορύβου salt&pepper πάνω στο σήμα μιας ψηφιακής εικόνας 8 bits . . . . .                                                          | 53 |
| Σχήμα 31: | Σχηματική αναπαράσταση της εφαρμογής του αλγορίθμου OMP στην περίπτωση του image inpainting . . . . .                                        | 53 |
| Σχήμα 32: | Σχηματική αναπαράσταση της εφαρμογής του αλγορίθμου k-SVD στην περίπτωση του image inpainting . . . . .                                      | 54 |
| Σχήμα 33: | Η εφαρμογή του αλγορίθμου OMP σε μία εικόνα με 50% salt&pepper θόρυβο . . . . .                                                              | 54 |
| Σχήμα 34: | Η εφαρμογή του αλγορίθμου k-SVD σε μία εικόνα με 50% salt&pepper θόρυβο . . . . .                                                            | 55 |

|           |                                                                                                                        |    |
|-----------|------------------------------------------------------------------------------------------------------------------------|----|
| Σχήμα 35: | Η εφαρμογή του αλγορίθμου k-SVD σε μία εικόνα με 80% salt&pepper θόρυβο . . . . .                                      | 55 |
| Σχήμα 36: | Η κατανομή θορύβου t-κατανομής πάνω στο σήμα μιας ψηφιακής εικόνας 8 bits . . . . .                                    | 56 |
| Σχήμα 37: | Η γραφική απεικόνιση της t-κατανομής (κόκκινη καμπύλη) σε σχέση με την κανονική κατανομή (μπλέ καμπύλη) . . . . .      | 56 |
| Σχήμα 38: | Η εφαρμογή του αλγορίθμου k-SVD σε μία εικόνα με θόρυβο από t-κατανομή . . . . .                                       | 57 |
| Σχήμα 39: | Σχηματική αναπαράσταση της εφαρμογής του αλγορίθμου k-SVD στην περίπτωση της t-κατανομής θορύβου . . . . .             | 57 |
| Σχήμα 40: | Η εφαρμογή του τροποποιημένου αλγορίθμου k-SVD σε μία εικόνα με θόρυβο από t-κατανομή . . . . .                        | 58 |
| Σχήμα 41: | Σχηματική αναπαράσταση της εφαρμογής του αλγορίθμου OMP στην περίπτωση του παλμικού θορύβου κατηγορίας I . . . . .     | 59 |
| Σχήμα 42: | Η κατανομή παλμικού θορύβου κατηγορίας I πάνω στο σήμα μιας ψηφιακής εικόνας 8 bits . . . . .                          | 59 |
| Σχήμα 43: | Η εφαρμογή του αλγορίθμου του Σχήματος 41 σε μία εικόνα με παλμικό θόρυβο κατηγορίας I . . . . .                       | 60 |
| Σχήμα 44: | Η κατανομή παλμικού θορύβου κατηγορίας II πάνω στο σήμα μιας ψηφιακής εικόνας 8 bits . . . . .                         | 60 |
| Σχήμα 45: | Η εφαρμογή του αλγορίθμου του Σχήματος 41 σε μία εικόνα με παλμικό θόρυβο κατηγορίας II . . . . .                      | 60 |
| Σχήμα 46: | Η κατανομή του θορύβου μικτού τύπου 1 πάνω στο σήμα μιας ψηφιακής εικόνας 8 bits . . . . .                             | 62 |
| Σχήμα 47: | Σχηματική αναπαράσταση της εφαρμογής του αλγορίθμου του Σχήματος 47 στην περίπτωση του θορύβου μικτού τύπου 1. . . . . | 63 |
| Σχήμα 48: | Η εφαρμογή του αλγορίθμου του Σχήματος 47 σε μία εικόνα με θόρυβο μικτού τύπου 1 . . . . .                             | 63 |
| Σχήμα 49: | Η κατανομή του θορύβου μικτού τύπου 2 πάνω στο σήμα μιας ψηφιακής εικόνας 8 bits . . . . .                             | 63 |
| Σχήμα 50: | Η εφαρμογή του αλγορίθμου του Σχήματος 47 σε μία εικόνα με θόρυβο μικτού τύπου 2 . . . . .                             | 64 |
| Σχήμα 51: | Η κατανομή του θορύβου μικτού τύπου 3 πάνω στο σήμα μιας ψηφιακής εικόνας 8 bits . . . . .                             | 64 |
| Σχήμα 52: | Η εφαρμογή του αλγορίθμου k-SVD σε μία εικόνα με θόρυβο μικτού τύπου 3 . . . . .                                       | 65 |



|           |                                                                                            |    |
|-----------|--------------------------------------------------------------------------------------------|----|
| Σχήμα 53: | Η κατανομή του θορύβου μικτού τύπου 4 πάνω στο σήμα μιας ψηφιακής εικόνας 8 bits . . . . . | 65 |
| Σχήμα 54: | Η εφαρμογή του αλγορίθμου του Σχήματος 47 σε μία εικόνα με θόρυβο μικτού τύπου 4 . . . . . | 65 |
| Σχήμα 55: | Η κατανομή του θορύβου μικτού τύπου 5 πάνω στο σήμα μιας ψηφιακής εικόνας 8 bits . . . . . | 66 |
| Σχήμα 56: | Η εφαρμογή του αλγορίθμου του Σχήματος 47 σε μία εικόνα με θόρυβο μικτού τύπου 5 . . . . . | 66 |

## ΚΑΤΑΛΟΓΟΣ ΠΙΝΑΚΩΝ

|                                                                                                                                                                                     |    |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Πίνακας 1: Συγκριτικά αποτελέσματα του αλγορίθμου k-SVD σε σχέση με τους αλγορίθμους Kernel-based και BM3D για θορύβο ομοιόμορφης κατανομής . . . . .                               | 52 |
| Πίνακας 2: Συγκριτικά αποτελέσματα του αλγορίθμου του Σχήματος 41 σε σχέση με τον αλγόριθμο Kernel-based για παλμικό θόρυβο κατηγορίας I . . .                                      | 61 |
| Πίνακας 3: Συγκριτικά αποτελέσματα των αλγορίθμων των Σχημάτων 47 και 21 σε σχέση με τους αλγορίθμους Kernel-based Denoising και BM3D στην περίπτωση θορύβου μικτού τύπου . . . . . | 67 |

## **ΠΡΟΛΟΓΟΣ**

Η παρούσα διπλωματική εργασία εκπονήθηκε στην Αθήνα το έτος 2015 στα πλαίσια του Μεταπτυχιακού Προγράμματος Σπουδών “Πληροφορική και Τηλεπικοινωνίες” με κατεύθυνση “Επεξεργασία Σήματος για Επικοινωνίες και Πολυμέσα” του Τμήματος Πληροφορικής και Τηλεπικοινωνιών της Σχολής Θετικών Επιστημών του Εθνικού και Καποδιστριακού Πανεπιστημίου Αθηνών. Η χρηματοδότηση των διδάκτρων του προγράμματος έγινε από την Πολεμική Αεροπορία.

# 1. ΑΝΤΙΣΤΡΟΦΑ ΠΡΟΒΛΗΜΑΤΑ

## 1.1 Εισαγωγή

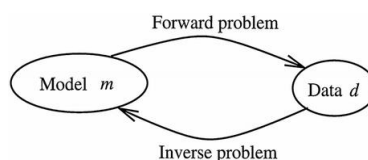
Τα αντίστροφα προβλήματα (inverse problems) αποτελούν αντικείμενο ευρείας μελέτης σε διάφορους κλάδους της επιστήμης λόγω των πολλών εφαρμογών που βρίσκουν στις ανθρώπινες δραστηριότητες. Η έννοια των αντίστροφων προβλημάτων έχει τις ρίζες του στον Victor Ambartsumian, έναν Αρμένιο φυσικό, ο οποίος τα συνάντησε στην έρευνά του, το 1929, πάνω σε εφαρμογές διαφορικών εξισώσεων σε προβλήματα της φυσικής. Ο Keller στο [12] περιγράφει τις βασικές αρχές αυτών των προβλημάτων.

**Ορισμός 1.1.** Δύο προβλήματα ονομάζονται **αντίστροφα (inverse)** αν η διατύπωση του ενός περιέχει τη λύση ή μέρος της λύσης του άλλου και το αντίστροφο.

**Παρατήρηση 1.1.** Στις περισσότερες περιπτώσεις το ένα από δύο προβλήματα έχει μελετηθεί περισσότερο και ονομάζεται **ευθύ (direct)** πρόβλημα ενώ το δεύτερο ονομάζεται **αντίστροφο (inverse)** πρόβλημα.

Για παράδειγμα, θεωρούμε το ζεύγος των προτάσεων “Βρες την απάντηση στην ερώτηση ποιά είναι η είναι η πρωτεύουσα της Ελλάδας” και “Βρες την ερώτηση στην απάντηση Αθήνα”. Είναι φανερό ότι στη διατύπωση της μιας πρότασης βρίσκεται η απάντηση στην άλλη πρόταση και το αντίστροφο. Επιπρόσθετα, το δεύτερο πρόβλημα επιδέχεται πολλών λύσεων γεγονός που συναντούμε συχνά στα αντίστροφα προβλήματα.

Η διατύπωση της έννοιας του αντίστροφου προβλήματος με τη βοήθεια ενός μαθηματικού μοντέλου μπορεί να δοθεί ως εξής όπως φαίνεται στο Σχήμα 1: “Δεδομένου ενός συνόλου παρατηρήσεων  $d$  και ενός τελεστή  $G$  που συνδέει τις παρατηρήσεις με τις παραμέτρους του μοντέλου, να βρεθούν οι παράμετροι  $m$  του μοντέλου ώστε  $G(m) = d$ .” Για πα-



Σχήμα 1: Ένα μαθηματικό μοντέλο για τα αντίστροφα προβλήματα.

ράδειγμα ένα αντίστροφο πρόβλημα στην άλγεβρα μητρώων είναι το εξής: “Δοθέντος ενός πραγματικού συμμετρικού μητρώου  $A$ , τάξης  $n$ , και  $n$  πραγματικών αριθμών  $\lambda_1, \lambda_2, \dots, \lambda_n$ , να βρεθεί ένα μητρώο  $D$  ώστε το άθροισμα των μητρώων  $A + D$  να έχει ως ιδιοτιμές τους αριθμούς  $\lambda_1, \lambda_2, \dots, \lambda_n$ ”. Το αντίστοιχο ευθύ πρόβλημα είναι η εύρεση των ιδιοτιμών ενός πραγματικού συμμετρικού μητρώου  $A + D$ .

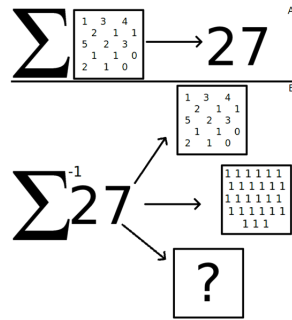
Στο Σχήμα 2 φαίνεται ένα απλό ζεύγος ευθέος και αντίστροφου προβλήματος. Αντίστροφα προβλήματα συναντούμε σε διάφορους κλάδους της επιστήμης όπως στην Επιστήμη των Υπολογιστών, στις Τηλεπικοινωνίες, στη Γεωφυσική και στην Αστρονομία.

## 1.2 Γραμμικά αντίστροφα προβλήματα

Με βάση τις βασικές αρχές των αντίστροφων προβλημάτων ένα γραμμικό αντίστροφο πρόβλημα ορίζεται ως εξής:

**Ορισμός 1.2.** Ένα γραμμικό πρόβλημα  $y = Ax$  είναι **αντίστροφο** όταν υπάρχουν δεδομένες οι παρατηρήσεις  $y$  και απαιτείται η εύρεση του  $x$  με δεδομένο το  $A$ .

Συσχετίζοντας τα παραπάνω στην επεξεργασία σημάτων, ο βασικός στόχος είναι η



**Σχήμα 2:** Το ευθύ πρόβλημα (A) της εύρεσης του αθροίσματος κάποιων αριθμών και το αντίστροφο (B) της εύρεσης κάποιων αριθμών που το άθροισμά τους είναι 27. Το δεύτερο επιδέχεται, φυσικά, πολλών λύσεων.

εξάλειψη του θορύβου και η ανασυγκρότηση, κατά το δυνατόν, του αρχικού σήματος. Η εξίσωση του κλασικού μοντέλου γραμμικής παλινδρόμησης παρουσία θορύβου [19] δίνεται από τη σχέση

$$y_n = \theta^T x_n + \eta_n \quad (1)$$

όπου  $y_n$  είναι η  $n$ -οστή παρατήρηση,  $\theta$  είναι το  $l \times 1$  διάνυσμα των άγνωστων παραμέτρων του προβλήματος,  $x_n$  είναι το  $l \times 1$  διάνυσμα της εισόδου  $n$  του προβλήματος και  $\eta_n$  είναι η άγνωστη τιμή του θορύβου (μηδενικής μέσης τιμής) για τη  $n$ -οστή παρατήρηση. Υποθέτοντας ότι έχουμε συγκεντρώσει  $N$  παρατηρήσεις, το προηγούμενο μοντέλο γράφεται

$$y = \theta^T X + \eta \quad (2)$$

όπου  $y$  είναι το  $N \times 1$  διάνυσμα των παρατηρήσεων,  $X$  είναι το  $N \times l$  μητρώο με γραμμές τα διανύσματα εισόδου  $x_n$  και  $\eta$  είναι το  $N \times 1$  διάνυσμα του θορύβου. Η παραπάνω εξίσωση παριστάνει ένα γραμμικό αντίστροφο πρόβλημα με άγνωστο το διάνυσμα  $\theta$  και η διαφορά σε σχέση με τον ορισμό 1.2, ότι στο διάνυσμα των παρατηρήσεων  $y$  επιδρά ο θόρυβος  $\eta$  του οποίου οι τιμές είναι άγνωστες.

Η εκτίμηση των ελαχίστων τετραγώνων (Least Squares Estimation) είναι ένας τρόπος επίλυσης του προβλήματος, ο οποίος λαμβάνει ως κριτήριο την Ευκλείδεια στάθμη (Euclidean norm) του διανύσματος σφάλματος  $e$  όπου

$$e = y - X\theta \quad (3)$$

Το πρόβλημα ελαχιστοποίησης

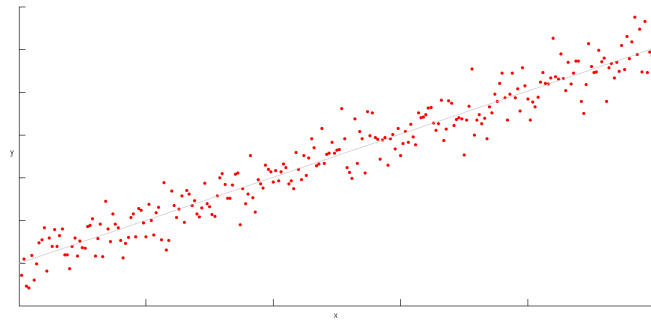
$$\hat{\theta}_{LS} = \arg \min_{\theta} \| y - X\theta \|^2 \quad (4)$$

έχει τη λύση

$$\hat{\theta}_{LS} = (X^T X)^{-1} X^T y \quad (5)$$

**Παρατήρηση 1.2.** Το μητρώο συντελεστής του  $y$  στη λύση του εκτιμητή ελάχιστων τετραγώνων,  $(X^T X)^{-1} X^T$ , ονομάζεται **Moore-Penrose ψευδοαντίστροφο (pseudoinverse)** μητρώο του μητρώου  $X$  και είναι ιδιαίτερα σημαντικό σε πολλά προβλήματα επεξεργασίας σήματος.

Στο Σχήμα 3 παριστάνεται ένα απλό μοντέλο γραμμικής παλινδρόμησης στη μία διάσταση ( $l = 1$ ) με επίδραση Gaussian θορύβου μηδενικής μέσης τιμής.



Σχήμα 3: Το μοντέλο γραμμικής παλινδρόμησης στη μία διάσταση.

### 1.3 Παλινδρόμηση Κορυφής (Ridge Regression)

Η λύση των ελαχίστων τετραγώνων στο πρόβλημα αποθορυβοποίησης δεν είναι πάντα η καλύτερη. Οι λόγοι είναι ότι στην εξίσωση (5) το μητρώο  $X^T X$  μπορεί να έχει πολύ μεγάλο αριθμό κατάστασης [20] και επομένως η λύση του εκτιμητή ελαχίστων τετραγώνων να είναι ασταθής.

**Ορισμός 1.3.** Ως **αριθμός κατάστασης (condition number)** ενός τετραγωνικού μητρώου  $A$  ονομάζεται το μέγεθος που εκφράζεται από την παρακάτω σχέση:

$$\kappa = \|A\| \|A\|^{-1}. \quad (6)$$

**Παρατήρηση 1.3.** Όταν ο αριθμός κατάστασης ενός μητρώου  $A$  πλησιάζει τη μονάδα, τότε το γραμμικό σύστημα με μητρώο συντελεστών το  $A$  έχει ευσταθή λύση. Αντίθετα όταν ο αριθμός κατάστασης είναι πολύ μεγαλύτερος της μονάδας τότε το σύστημα είναι ασταθές.

Επιπρόσθετα είναι δυνατόν στην προκύπτουσα λύση να προκύψει το φαινόμενο του **υπερταϊριάσματος (overfitting)** κάτι που οφείλεται στην αμεροληψία του εκτιμητή  $\hat{\theta}_{LS}$ . Προκειμένου να αποφευχθούν τα παραπάνω προβλήματα το πρόβλημα ελαχιστοποίησης (4) διαμορφώνεται σε

$$\hat{\theta}_R = \arg \min_{\theta} \{ \|\mathbf{y} - X\theta\|^2 + \lambda \|\theta\|^2 \} \quad (7)$$

όπου  $\lambda$  είναι μία παράμετρος που καθορίζει το βαθμό ομαλοποίησης (regularization) που επιθυμούμε να πετύχουμε στο μοντέλο της παλινδρόμησης. Το παραπάνω πρόβλημα ονομάζεται πρόβλημα **παλινδρόμησης κορυφής (ridge regression)** ή πρόβλημα **ομαλοποίησης Tikhonov (Tikhonov regularization)** προς τιμήν του μαθηματικού Andrey Tikhonov που πρώτος πρότεινε το παραπάνω μοντέλο βελτιστοποίησης. Ο εκτιμητής παλινδρόμησης κορυφής  $\hat{\theta}_R$  επιδέχεται αναλυτικής λύσης όπως και ο εκτιμητής ελαχίστων τετραγώνων η οποία είναι

$$\hat{\theta}_R = (X^T X + \lambda I)^{-1} X^T \mathbf{y} \quad (8)$$

όπου  $I$  είναι το μοναδιαίο μητρώο ανάλογης διάστασης. Ο εκτιμητής  $\hat{\theta}_R$  της παλινδρόμησης κορυφής προσφέρει ευσταθείς λύσεις στην περίπτωση που το μητρώο  $X^T X$ , λόγω του μεγάλου αριθμού κατάστασης, δίνει μία ασταθή λύση του εκτιμητή  $\hat{\theta}_{LS}$  μέσω της εξίσωσης (5). Η προηγούμενη πρόταση μπορεί να γίνει πιο αντιληπτή με ένα απλό αριθμητικό παράδειγμα.

**Παράδειγμα 1.1.** Ας υποθέσουμε ότι επιθυμούμε να εκτιμήσουμε την παράμετρο  $\theta$  του μοντέλου της οποίας η τιμή είναι  $\theta_0 = [1 \ 1]^T$ . Έστω ότι από τις δεδομένες εισόδους και παρατηρήσεις προκύπτει ότι το μητρώο  $X$  και το διάνυσμα  $y$  είναι ίσα με

$$X = \begin{pmatrix} 0.5 & 0.9938 \\ 0.1 & 0.23101 \\ 0.3 & 0.6 \\ 0.4 & 0.8 \\ 0.7 & 1.4 \end{pmatrix} \quad \text{και} \quad y = \begin{pmatrix} 1.51 \\ 0.31 \\ 0.9 \\ 1.2 \\ 2.1 \end{pmatrix}$$

Τότε η εξίσωση (5) δίνει ως λύση  $\theta_{LS} \simeq [2.51 \ 0.25]^T$  δηλαδή μία λύση πολύ μακριά από το  $\theta_0$ . Αυτό οφείλεται στο ότι ο αριθμός κατάστασης του μητρώου  $X^T X$  είναι πολύ μεγάλος [20],  $\kappa \simeq 2.5 * 10^4$ , και επομένως το γραμμικό σύστημα που βρίσκει τον εκτιμητή  $\theta_{LS}$  με μητρώο συντελεστών  $X^T X$  είναι ασταθές. Αντίθετα αν με τα ίδια δεδομένα υπολογίσουμε τον εκτιμητή παλινδρόμησης κορυφής  $\hat{\theta}_R$  με  $\lambda = 0.0008$  προκύπτει η λύση  $\theta_R \simeq [0.98 \ 1.01]^T$  που προσεγγίζει αρκετά την πραγματική τιμή  $\theta_0$ .

## 2. ΚΥΡΤΗ ΒΕΛΤΙΣΤΟΠΟΙΗΣΗ

### 2.1 Το πρόβλημα

Τα προβλήματα κυρτής βελτιστοποίησης αποτελούν μια ιδιαίτερη κλάση των προβλημάτων μαθηματικής βελτιστοποίησης και έχουν προσελκύσει έντονο ερευνητικό ενδιαφέρον τον τελευταίο καιρό. Ο λόγος είναι η ανάπτυξη αποδοτικών αλγορίθμων που επιλύουν τα προβλήματα κυρτής βελτιστοποίησης συνεισφέρει στην επίλυση και άλλων προβλημάτων βελτιστοποίησης που μπορούν να αναχθούν σε κυρτά προβλήματα. Οι παρακάτω ορισμοί σχετικά με τις έννοιες της κυρτής βελτιστοποίησης έχουν αντληθεί από το [4].

**Ορισμός 2.1.** Έστω ένα διάνυσμα  $\mathbf{x} = (x_0, x_1, \dots, x_{n-1})$  που ονομάζεται μεταβλητή βελτιστοποίησης (optimization variable) και μία συνάρτηση  $f_0 : \mathbb{R}^n \rightarrow \mathbb{R}$  που ονομάζεται αντικειμενική συνάρτηση (objective function). Έστω, επίσης, ένα σύνολο συναρτήσεων  $f_i : \mathbb{R}^n \rightarrow \mathbb{R}, i = 1, 2, \dots, m$  που καλούνται συναρτήσεις περιορισμού (constraint functions) και  $m$  σταθερές  $b_1, b_2, \dots, b_m$ . Ένα **πρόβλημα βελτιστοποίησης (optimization problem)** ορίζεται ως

$$\begin{aligned} &\text{ελαχιστοποίησε (minimize)} && f_0(\mathbf{x}) \\ &\text{υπό τις συνθήκες (subject to)} && f_i(\mathbf{x}) \leq b_i, \quad i = 1, 2, \dots, m \end{aligned} \quad (9)$$

Ένα διάνυσμα  $\mathbf{x}^*$  είναι λύση του προβλήματος (9) αν στο σημείο  $\mathbf{x}^*$  η αντικειμενική συνάρτηση έχει τη μικρότερη τιμή σε σχέση με την τιμή που αντιστοιχεί σε οποιοδήποτε άλλο διάνυσμα  $\mathbf{z}$  που ικανοποιεί τους περιορισμούς του προβλήματος, δηλαδή  $\forall \mathbf{z}$  με  $f_i(\mathbf{z}) \leq b_i, i = 1, 2, \dots, m$  ισχύει  $f_0(\mathbf{x}^*) \leq f_0(\mathbf{z})$ .

Αν η αντικειμενική συνάρτηση  $f_0$  και οι συναρτήσεις περιορισμού  $f_i : i = 1, 2, \dots, m$  είναι γραμμικές, δηλαδή για  $i = 0, 1, 2, \dots, m$  ισχύει

$$f_i(\alpha \mathbf{x} + \beta \mathbf{y}) = \alpha f_i(\mathbf{x}) + \beta f_i(\mathbf{y}) \quad (10)$$

για κάθε  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$  και για κάθε  $\alpha, \beta \in \mathbb{R}$ , τότε το πρόβλημα (9) ονομάζεται **πρόβλημα γραμμικής βελτιστοποίησης (linear optimization problem)**. Σε διαφορετική περίπτωση έχουμε ένα **πρόβλημα μη γραμμικής βελτιστοποίησης**.

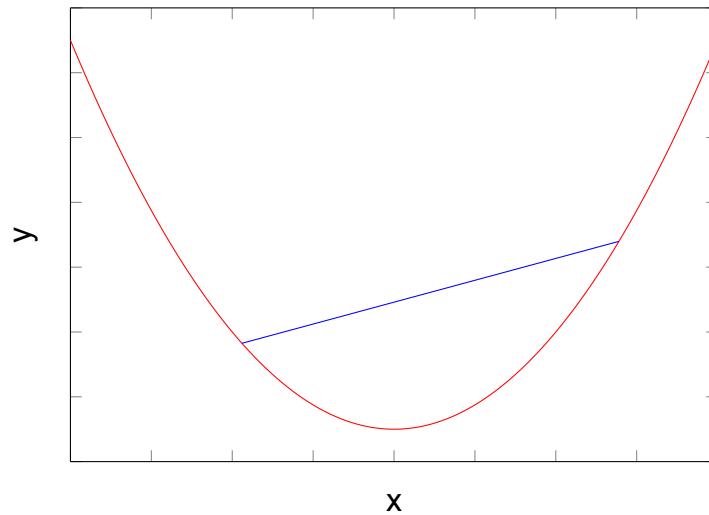
**Ορισμός 2.2.** Αν η αντικειμενική συνάρτηση  $f_0$  και οι συναρτήσεις περιορισμού  $f_i, i = 1, 2, \dots, m$  είναι κυρτές, δηλαδή ισχύει

$$f_i(\alpha \mathbf{x} + \beta \mathbf{y}) \leq \alpha f_i(\mathbf{x}) + \beta f_i(\mathbf{y}) \quad (11)$$

για κάθε  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$  και για κάθε  $\alpha, \beta \in \mathbb{R}$  με  $\alpha \geq 0, \beta \geq 0$  και  $\alpha + \beta = 1$  τότε το πρόβλημα (9) ονομάζεται **πρόβλημα κυρτής βελτιστοποίησης (convex optimization problem)**.

**Παρατήρηση 2.1.** Ένα πρόβλημα γραμμικής βελτιστοποίησης είναι ειδική περίπτωση των προβλημάτων κυρτής βελτιστοποίησης.





**Σχήμα 4: Μία κυρτή συνάρτηση μιας διάστασης. Κάθε ευθύγραμμο τμήμα που ενώνει δύο σημεία της συνάρτησης κείται πάνω από το γράφημα της συνάρτησης**

Στο Σχήμα 4 απεικονίζεται το γράφημα μιας κυρτής συνάρτησης με μονοδιάστατο όρισμα και ένα ευθύγραμμο τμήμα που ενώνει δύο σημεία της συνάρτησης. Κάθε τέτοιο ευθύγραμμο τμήμα βρίσκεται "πάνω" από το γράφημα της κυρτής συνάρτησης.

### 3. ΔΙΑΣΠΑΣΗ ΜΗΤΡΩΟΥ ΜΕΣΩ ΤΩΝ ΙΔΙΑΖΟΥΣΩΝ ΤΙΜΩΝ ΤΟΥ

#### 3.1 Η διάσπαση SVD

Η διάσπαση ενός μητρώου με τη βοήθεια των ιδιζουσών τιμών του αποτελεί από ένα από τα πιο ισχυρά εργαλεία της Γραμμικής Άλγεβρας που χρησιμοποιείται στους αλγορίθμους μηχανικής μάθησης. Η διάσπαση αυτή βασίζεται στις ιδιζουσες τιμές του προς διάσπαση μητρώου που ορίζονται όπως παρακάτω [2], [11], [13], [19]:

**Ορισμός 3.1.** Έστω ένα πραγματικό μητρώο  $X$  βαθμού  $r$  και  $\lambda_1, \lambda_2, \dots, \lambda_r$  οι ιδιοτιμές του μητρώου  $XX^T$  που είναι επίσης  $r$  βαθμού. Ως **ιδιζουσες τιμές (singular values)**  $\sigma_i$  του μητρώου  $X$  ονομάζονται οι τετραγωνικές ρίζες των (κοινών) μη μηδενικών ιδιοτιμών των μητρώων  $XX^T$  και  $X^T X$ , δηλαδή  $\sigma_i = \sqrt{\lambda_i}$  για  $i = 1, 2, \dots, r$ . Τα ιδιοδιανύσματα  $\mathbf{u}_i$  του μητρώου  $XX^T$  που αντιστοιχούν στις μη μηδενικές ιδιοτιμές του  $\lambda_i$  ονομάζονται **δεξιά ιδιζόντα διανύσματα (right singular vectors)** του μητρώου  $X$  και τα ιδιοδιανύσματα  $\mathbf{v}_i$  του μητρώου  $X^T X$  που αντιστοιχούν στις μη μηδενικές ιδιοτιμές του  $\lambda_i$  ονομάζονται **αριστερά ιδιζόντα διανύσματα (left singular vectors)** του μητρώου  $X$ .

**Θεώρημα 3.1. (Singular Value Decomposition-SVD)** Έστω ένα πραγματικό μητρώο  $X$  διαστάσεων  $m \times l$  με βαθμό  $r \leq \min(m, l)$ . Τότε, υπάρχουν ορθογώνια μητρώα  $U$  και  $V$  με διαστάσεις  $m \times m$  και  $l \times l$  αντίστοιχα έτσι ώστε

$$X = U \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} V^T \quad (12)$$

όπου το μητρώο  $D$  είναι ένα  $r \times r$  διαγώνιο μητρώο με στοιχεία κυρίας διαγωνίου  $d_{ii}$  τις ιδιζουσες τιμές του μητρώου  $X$ , δηλαδή  $d_{ii} = \sigma_i$  για  $i = 1, 2, \dots, r$  ενώ με  $0$  συμβολίζονται μητρώα μηδενικών στοιχείων και κατάλληλων διαστάσεων.

Είναι φανερό από το παραπάνω θεώρημα ότι μόνο οι  $r$  πρώτες στήλες των μητρώων  $U$  και  $V$  συμμετέχουν στον τύπο της παραγοντοποίησης του μητρώου  $X$ . Επομένως ο τύπος (12) μπορεί να γραφεί σε μία πιο "οικονομική" έκδοση ως

$$X = U_r D V_r^T = \sum_{i=1}^r \sigma_i \mathbf{u}_i \mathbf{v}_i^T, \quad (13)$$

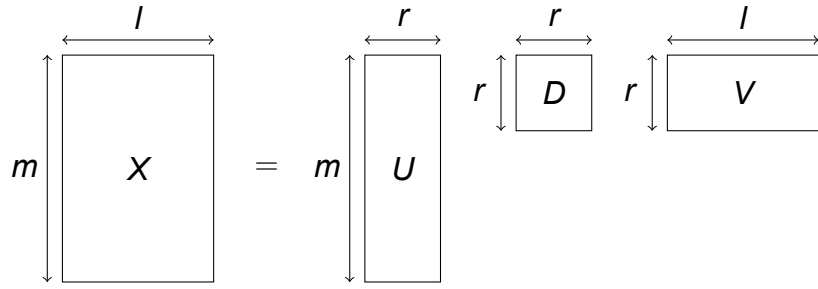
όπου το μητρώο  $U_r$  είναι το  $m \times r$  μητρώο με στήλες τα δεξιά ιδιζόντα διανύσματα του μητρώου  $X$  και το μητρώο  $V_r$  είναι το  $r \times l$  μητρώο με στήλες τα αριστερά ιδιζόντα διανύσματα του μητρώου  $X$ , δηλαδή

$$U_r = [\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_r] \quad \text{και} \quad V_r = [\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_r] \quad (14)$$

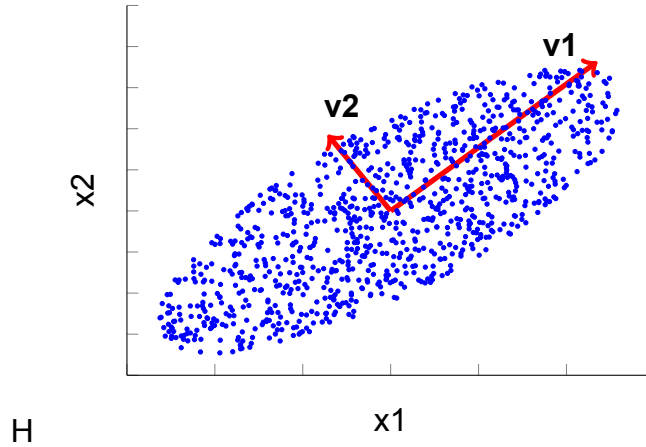
Στο Σχήμα 5 απεικονίζεται η παραγοντοποίηση ενός  $m \times l$  μητρώου  $X$   $r$  βαθμού με βάση τον τύπο (13).

Ένα από τα θεωρήματα που βασίζονται στην παραγοντοποίηση SVD και βρίσκει πολλές εφαρμογές σε προβλήματα μείωσης διαστάσεων της μηχανικής μάθησης είναι το θεώρημα *Eckart-Young-Mirsky*. Πριν τη διατύπωση του θεωρήματος δίνεται ο ορισμός δύο τύπων σταθμών ενός μητρώου. Οι στάθμες διανύσματος περιγράφονται εκτενέστερα σε επόμενη ενότητα.

**Ορισμός 3.2.** Έστω το πραγματικό μητρώο  $X$  διαστάσεων  $m \times l$  και βαθμού  $r$  με στοιχεία  $x_{ij}$ , ιδιζουσες τιμές  $\sigma_i$ , με  $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r$  και αντίστοιχα δεξιά και αριστερά ιδιζόντα διανύσματα  $\mathbf{u}_i$  και  $\mathbf{v}_i$ . Ως **Frobenius στάθμη (Frobenius norm)** του μητρώου  $X$ ,  $\|X\|_F$ ,



Σχήμα 5: Η SVD παραγοντοποίηση ενός  $m \times l$  μητρώου  $X$   $r$  βαθμού



Σχήμα 6: Τα ιδιάζοντα διανύσματα  $v_1$  και  $v_2$  προκύπτουν από 1000 τυχαία σημεία μέσα σε μία έλλειψη. Τα δύο διανύσματα τείνουν να ταυτιστούν με τους δύο άξονες της έλλειψης

ορίζεται η ποσότητα

$$\|X\|_F = \sqrt{\sum_{i=1}^m \sum_{j=1}^l |x_{ij}|^2} = \sqrt{\sum_{i=1}^r \sigma_i^2} \quad (15)$$

και ως **φασματική στάθμη (spectral norm)** του μητρώου  $X$ ,  $\|X\|_2$  η ποσότητα

$$\|X\|_2 = \sigma_1 \quad (16)$$

**Θεώρημα 3.2. (Eckart-Young-Mirsky)** Έστω το μητρώο  $X$  του προηγούμενου ορισμού. Το μητρώο  $\hat{X}$   $k$  βαθμού, με  $k < r$ , που αποτελεί τη βέλτιστη προσέγγιση  $k$  βαθμού του μητρώου  $X$  (**best  $k$ -rank approximation**) με κριτήριο τόσο τη Frobenius στάθμη όσο και τη φασματική στάθμη δίνεται από τον τύπο

$$\hat{X} = \sum_{i=1}^k \sigma_i u_i v_i^T \quad (17)$$

Τα σφάλματα προσέγγισης που προκύπτουν είναι αντίστοιχα για τις δύο στάθμες

$$\|X - \hat{X}\|_F = \sqrt{\sum_{i=k+1}^r \sigma_i^2} \quad \text{και} \quad \|X - \hat{X}\|_2 = \sigma_{k+1}. \quad (18)$$

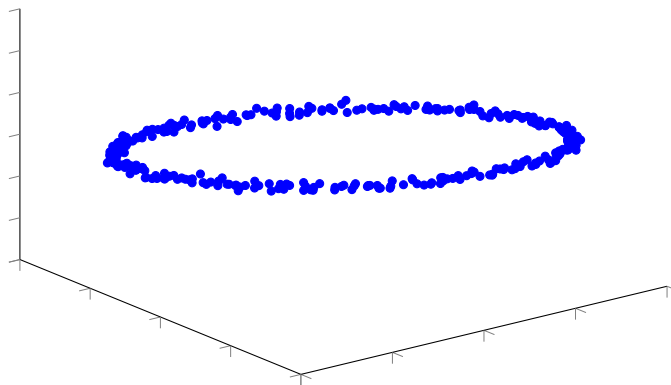
Το ερμηνεία της διάσπασης SVD είναι η “καρδιά” στην έννοια μείωση της διάστασης (*dimensionality reduction*) στο χώρο των δεδομένων ενός προβλήματος. Με τη μέθοδο

SVD επιτυγχάνεται η εύρεση των διαστάσεων στις οποίες τα δεδομένα παρουσιάζουν τη μεγαλύτερη διασπορά. Αυτό μπορεί να γίνει σαφές από το Σχήμα 6 στο οποίο απεικονίζονται τυχαία σημεία του επιπέδου τα οποία βρίσκονται μέσα σε μία έλλειψη. Τα δύο διανύσματα είναι τα ιδιάζοντα διανύσματα  $\mathbf{v}_1$  και  $\mathbf{v}_2$  που προέρχονται από την παραγοντοποίηση SVD του μητρώου των δεδομένων. Το διάνυσμα  $\mathbf{v}_1$  αντιστοιχεί στη μεγαλύτερη ιδιάζουσα τιμή που προκύπτει από τη διάσπαση. Αυτό σημαίνει ότι η διασπορά που θα προκύψει από τις προβολές όλων των σημείων πάνω στον άξονα του διανύσματος  $\mathbf{v}_1$  είναι η μέγιστη που μπορεί να προκύψει σε σχέση με οποιαδήποτε άλλη ευθεία του επιπέδου. Επομένως η βέλτιστη λύση για την αναπαράσταση των δεδομένων στη μία διάσταση είναι η ευθεία του διανύσματος  $\mathbf{v}_1$  η διαφορετικά η καλύτερη προσέγγιση βαθμού 1 για το μητρώο των δεδομένων  $X$  που έχει βαθμό 2 είναι το μητρώο  $\hat{X}$ , που έχει εκφυλιστεί σε διάνυσμα, και με βάση τον τύπο του Θεωρήματος (3.2) είναι  $\hat{X} = \sigma_1 \mathbf{u}_1 \mathbf{v}_1^T$ .

Αυτό που αξίζει να αναφερθεί επίσης είναι ότι η μείωση της διάστασης των δεδομένων δε σημαίνει απαραίτητα και απώλεια σημαντικής πληροφορίας. Η ταξινόμηση των διαστάσεων με βάση το μέγεθος διασποράς των δεδομένων και η απαλοιφή των διαστάσεων στις οποίες τα δεδομένα παρουσιάζουν τις μικρότερες διασπορές σε πραγματικά προβλήματα σημαίνει ότι με αυτόν τον τρόπο απαλείφονται οι διάφορες συνιστώσες του θορύβου που υπάρχει στα δεδομένα.

### 3.2 Η μείωση της διάστασης

Σε πολλές πρακτικές εφαρμογές, τα δεδομένα παρόλο βρίσκονται τοποθετημένα σε έναν χώρο μεγάλης διάστασης, η πραγματική τους διάσταση είναι κατά πολύ μικρότερη του χώρου στον οποίο εδράζουν. Για παράδειγμα, σε έναν τρισδιάστατο χώρο, τα δεδομένα μπορεί να βρίσκονται πάνω σε μία ευθεία γραμμή ή πάνω σε ένα κύκλο, όπως φαίνεται στο Σχήμα 7. Στην περίπτωση αυτή αρκεί η γνώση της τιμής μιας παραμέτρου προκειμένου να περιγράψουμε την πραγματική διάσταση των δεδομένων [19].



Σχήμα 7: Ένα σύνολο σημείων του τρισδιάστατου χώρου βρίσκεται “περίπου” πάνω σε μια καμπύλη κύκλου και επομένως η ενδογενής διάσταση του συνόλου προσεγγίζει τη μονάδα

**Ορισμός 3.3.** Ένα σύνολο  $\mathcal{X} \subset \mathbb{R}^L$  έχει **ενδογενή διάσταση (intrinsic dimensionality)**  $m \leq L$  εάν το σύνολο  $\mathcal{X}$  μπορεί να περιγραφεί με τη βοήθεια  $m$  παραμέτρων.

Έστω για παράδειγμα ότι τα δεδομένα του συνόλου  $\mathcal{X}$  μπορούν να περιγραφούν με τη βοήθεια συναρτήσεων  $m$  μεταβλητών ως  $\mathbf{x} = \mathbf{g}(u_1, u_2, \dots, u_m)$ ,  $u_i \in \mathbb{R}, i = 1, 2, \dots, m$ . Η γεωμετρική ερμηνεία της ενδογενούς διάστασης είναι ότι το διάνυσμα δεδομένων βρίσκεται πάνω σε μία καμπύλη της οποίας η μορφή δίνεται με τη βοήθεια της συνάρτησης  $\mathbf{g} : \mathbb{R}^m \rightarrow \mathbb{R}^L$ . Για παράδειγμα, μία τέτοια συνάρτηση είναι η  $\mathbf{g}(u) = [C \cos u, C \sin u]$  όπου

$C$  είναι μία σταθερά και  $u$  η παράμετρος της συνάρτησης. Στο συγκεκριμένο παράδειγμα τα σημεία του τρισδιάστατου χώρου που αναπαριστούν το διάνυσμα των δεδομένων βρίσκονται “προσεγγιστικά” πάνω σε έναν κύκλο όπως φαίνεται και στο Σχήμα 7. Επομένως η ενδογενής διάσταση του συνόλου των σημείων προσεγγίζει τη μονάδα καθώς τα σημεία δε βρίσκονται ακριβώς πάνω σε μία καμπύλη κύκλου.

Εξετάζοντας τη στατιστική πλευρά της ενδογενούς διάστασης, ένα σύνολο δεδομένων που η ενδογενής διάστασή του προσεγγίζει τη μονάδα παρουσιάζει μεγάλο βαθμό συσχέτισης. Η γνώση επομένως του τοπολογικού χώρου (manifold) στον οποίο εδράζουν τα δεδομένα μας δίνει τη δυνατότητα να αναπαραστήσουμε μεγάλες ποσότητες δεδομένων (big data) σε χώρους πολύ μικρότερων διαστάσεων.

## 4. ΑΝΑΠΑΡΑΣΤΑΣΗ ΣΗΜΑΤΟΣ ΣΕ ΣΥΜΠΙΕΣΤΗ ΜΟΡΦΗ

### 4.1 Εισαγωγή

Τα τελευταία χρόνια, οι τεχνικές αναπαράστασης ενός σήματος σε συμπίεστη μορφή (*compressed sensing*) έχουν προσελκύσει το ενδιαφέρον πολλών ερευνητών των εφαρμοσμένων μαθηματικών και της μηχανικής των υπολογιστών. Βασική αρχή των τεχνικών αυτών είναι η αναπαράσταση πολλών σημάτων με τη βοήθεια κατάλληλων μικρών διανυσμάτων και ενός κατάλληλου μητρώου που χρησιμοποιείται ως βάση ή λεξικό για την αναπαράσταση του σήματος [8]. Για την ανακατασκευή του σήματος με τη βοήθεια μετρήσεων χρησιμοποιούνται στη συνέχεια τεχνικές μη γραμμικής βελτιστοποίησης (*nonlinear optimization*).

### 4.2 Διανυσματικοί χώροι και στάθμες διανυσμάτων

Οι διανυσματικοί χώροι χρησιμοποιούνται κατά κύριο λόγο για την αναπαράσταση σημάτων στην επεξεργασία δεδομένων. Ο λόγος είναι ότι στα περισσότερα φυσικά συστήματα υπάρχει γραμμικότητα έτσι ώστε τα σήματα τα οποία παράγουν να αναπαρίστανται ως διανύσματα πάνω σε διανυσματικούς χώρους. Οι διανυσματικοί χώροι αποτελούν ένα εργαλείο για τους μηχανικούς της επεξεργασίας σημάτων καθώς μπορεί να γίνει οπτικοποίηση κάποιων αποτελεσμάτων επεξεργασίας με βάση τις γεωμετρικές αναπαραστάσεις των διανυσματικών χώρων στον τρισδιάστατο χώρο που είναι αντιληπτός από τον ανθρώπινο νου. Επιπρόσθετα οι διανυσματικές αναπαραστάσεις προσφέρουν μετρικές συγκρίσεων όπως είναι οι στάθμες, οι αποστάσεις και οι γωνίες.

Οι στάθμες διανυσμάτων  $\ell_p$  αποτελούν μια εύχρηστη μετρική παραμέτρων του μεγέθους ενός σήματος [6].

**Ορισμός 4.1.** Έστω ένα διάνυσμα  $\mathbf{x}$  με  $n$  στοιχεία  $x_i$ ,  $i = 1, 2, \dots, n$ , που ορίζεται σε έναν Ευκλείδειο χώρο  $\mathbb{R}^n$ . Ονομάζουμε  $\ell_p$  **στάθμη (norm)** του διανύσματος  $\mathbf{x}$  για κάποιο  $p \in [1, \infty)$  την ποσότητα

$$\ell_p = \|\mathbf{x}\|_p = \left( \sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}}, \quad p \in [1, \infty) \quad (19)$$

Επιπρόσθετα, για  $p \rightarrow \infty$  ορίζεται η  $\ell_\infty$  στάθμη διανύσματος ως

$$\ell_\infty = \|\mathbf{x}\|_\infty = \max_{i=1,2,\dots,n} |x_i|. \quad (20)$$

Οι  $\ell_p$  στάθμες διανυσμάτων ικανοποιούν τις 4 ιδιότητες που έχουν οι συναρτήσεις στάθμης δηλαδή:

1.  $\|\mathbf{x}\|_p \geq 0$ .
2.  $\|\mathbf{x}\|_p = 0 \iff \mathbf{x} = \mathbf{0}$ .
3.  $\|\alpha\mathbf{x}\|_p = |\alpha| \|\mathbf{x}\|_p \quad \forall \alpha \in \mathbb{R}$ .
4.  $\|\mathbf{x}_1 + \mathbf{x}_2\|_p \leq \|\mathbf{x}_1\|_p + \|\mathbf{x}_2\|_p$ .

Για  $p = 1$  η  $\ell_1$  στάθμη του διανύσματος  $\mathbf{x}$  είναι το άθροισμα των απολύτων τιμών των

στοιχείων του διανύσματος, ήτοι,

$$\ell_1 = \|\mathbf{x}\|_0 = \sum_{i=1}^n |x_i| \quad (21)$$

ενώ για  $p = 2$  η  $\ell_2$  στάθμη είναι η Ευκλείδεια στάθμη. Οι  $\ell_p$  στάθμες διανυσμάτων μπορούν να επεκταθούν και για αριθμούς  $p \in (0, 1)$ . Σε αυτή την περίπτωση οι στάθμες  $\ell_p$  που δημιουργούνται δεν ικανοποιούν την τέταρτη ιδιότητα των συναρτήσεων στάθμης (τριγωνική ανισότητα) και για αυτό ο όρος “στάθμη” σε αυτή την περίπτωση χρησιμοποιείται καταχρηστικά. Για  $p \rightarrow 0$  παίρνουμε την  $\ell_0$  “στάθμη” για ένα διάνυσμα  $\mathbf{x}$  ως εξής:

$$\ell_0 = \|\mathbf{x}\|_0 = \sum_{i=1}^n \chi_{(0,\infty)}(|x_i|), \quad (22)$$

όπου η συνάρτηση  $\chi_A$  είναι η χαρακτηριστική συνάρτηση ενός συνόλου  $A$  και για ένα στοιχείο  $x$  ορίζεται ως εξής:

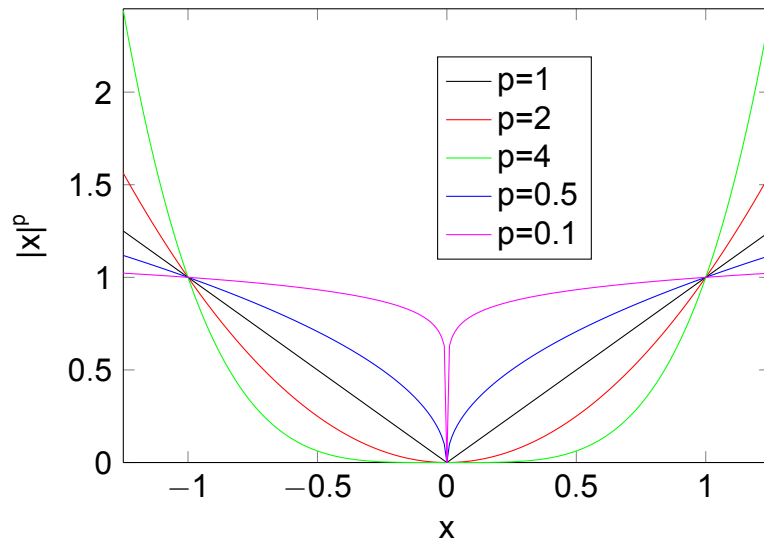
$$\chi_A(x) := \begin{cases} 0, & x \in A \\ 1, & x \notin A. \end{cases} \quad (23)$$

Δηλαδή, η  $\ell_0$  ενός διανύσματος  $\mathbf{x}$  εκφράζει τον αριθμό των μη μηδενικών στοιχείων του  $\mathbf{x}$ . Η  $\ell_0$  “στάθμη”, σε αντίθεση με τις “στάθμες” του διαστήματος  $(0, 1)$ , ικανοποιεί την ιδιότητα της τριγωνικής ανισότητας αλλά δεν ικανοποιεί την τρίτη ιδιότητα της θετικά ομογενούς (positively homogeneous) συνάρτησης.

Στο Σχήμα 8 απεικονίζεται το μέγεθος συνεισφοράς ενός στοιχείου ενός διανύσματος στην τιμή της στάθμης του διανύσματος σε συνάρτηση με την τιμή του στοιχείου και για διάφορες στάθμες. Αυτό που γίνεται φανερό είναι ότι για τιμές του στοιχείου  $x$ , με  $|x| > 1$ , όσο μεγαλύτερη είναι η απόλυτη τιμή  $|x|$  και όσο μεγαλύτερος είναι ο αριθμός  $p$  της στάθμης, για  $p \geq 1$ , τόσο μεγαλύτερη είναι η συνεισφορά του στοιχείου  $x$  στην τιμή της στάθμης. Αντίθετα, για στοιχεία  $x$  με  $|x| < 1$ , όσο μεγαλύτερη είναι η απόλυτη τιμή  $|x|$  και όσο μεγαλύτερος είναι ο αριθμός  $p$  της στάθμης, για  $p \geq 1$ , τόσο μικρότερη είναι η συνεισφορά του στοιχείου  $x$  στην τιμή της στάθμης. Από όλες τις στάθμες  $\ell_p$  με  $p \geq 1$ , μόνο η στάθμη  $\ell_1$  διατηρεί μια σχετική αναλογική συνεισφορά στο μέγεθος συνεισφοράς κάποιου στοιχείου στην τιμή της στάθμης, δηλαδή λαμβάνει υπόψη της και τα “μικρά στοιχεία” του διανύσματος.

Με βάση τα παραπάνω, μία αντικατάσταση της στάθμης  $\ell_2$  από τη στάθμη  $\ell_1$  στο πρόβλημα ελαχιστοποίησης (7) της παλινδρόμησης κορυφής θα έχει ως αποτέλεσμα ένα αραιό διάνυσμα, δηλαδή ένα διάνυσμα με πολλά μηδενικά, καθώς η στάθμη  $\ell_1$  σε έναν επαναληπτικό αλγόριθμο θα διατηρήσει εκείνα τα στοιχεία του διανύσματος παραμέτρων που βελτιώνουν ουσιαστικά την προσαρμογή (fitting) των δεδομένων στο πρόβλημα και θα εκμηδενίσει τα υπόλοιπα στοιχεία-παραμέτρους. Το συμπέρασμα αυτό αποτελεί βάση της αναπαράστασης σήματος σε συμπιεστή μορφή (compressed sensing) και γενικότερα της μείωσης της διάστασης (dimensionality reduction) ενός προβλήματος μηχανικής μάθησης.

Επίσης από το Σχήμα 8 είναι σαφές ότι για  $p < 1$  οι γραφικές αναπαραστάσεις των “σταθμών” χάνουν την κυρτότητά τους, κάτι που ήταν αναμενόμενο αφού δεν είναι πραγματικές στάθμες. Η έλλειψη κυρτότητας σε αυτές τις στάθμες, τις κάνει πιο δύσχρηστες σε προβλήματα βελτιστοποίησης.



Σχήμα 8: Η συνεισφορά ενός στοιχείου  $x$ , ήτοι  $|x|^p$ , στην τιμή της στάθμης  $\ell_p$  του διανύσματος  $\mathbf{x}$ , για διάφορες τιμές του  $p$

### 4.3 Βάσεις

Οι αναπαραστάσεις σημάτων με τη βοήθεια βάσεων αποτελεί μια τεχνική που χρησιμοποιείται ευρέως όταν το πλήθος των δεδομένων είναι μεγάλο.

**Ορισμός 4.2.** Έστω ο χώρος των πραγματικών αριθμών  $n$ -διαστάσεων  $\mathbb{R}^n$ . Ένα σύνολο γραμμικά ανεξάρτητων διανυσμάτων  $\{\boldsymbol{\varphi}_i\}_{i=1}^n$  ονομάζεται **βάση (basis)** αν για κάθε διάνυσμα  $\mathbf{x}$  του χώρου υπάρχουν μοναδικοί πραγματικοί αριθμοί  $c_i, i = 1, 2, \dots, n$  ώστε να ισχύει

$$\mathbf{x} = \sum_{i=1}^n c_i \boldsymbol{\varphi}_i. \quad (24)$$

**Παρατήρηση 4.1.** Η βάση στην επεξεργασία σήματος και στη μηχανική μάθηση συναντάται με τον όρο **λεξικό (dictionary)** ενώ τα διανύσματα  $\boldsymbol{\varphi}_i$  που είναι στοιχεία της βάσης αναφέρονται ως **άτομα (atoms)**.

Αν θεωρήσουμε το  $n \times n$  μητρώο  $\Phi$  με στήλες τα διανύσματα  $\boldsymbol{\varphi}_i$  και  $\mathbf{c}$  το  $n \times 1$  διάνυσμα με στοιχεία  $c_i$  τότε η προηγούμενη εξίσωση (24) μπορεί να γραφεί

$$\mathbf{x} = \Phi \mathbf{c} \quad (25)$$

Μία ειδική κατηγορία των βάσεων είναι οι **ορθοκανονικές βάσεις**. Μία βάση είναι  $\{\boldsymbol{\varphi}_i\}_{i=1}^n$  είναι ορθοκανονική όταν ισχύει

$$\boldsymbol{\varphi}_i \boldsymbol{\varphi}_j^T = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases} \quad (26)$$

ή διαφορετικά

$$\Phi^T \Phi = \mathbf{I} \quad (27)$$

Επομένως για μια ορθοκανονική βάση θα ισχύει

$$\mathbf{c} = \Phi^T \mathbf{x} \quad (28)$$



Η ιδέα της βάσης μπορεί να επεκταθεί για σύνολα γραμμικά εξαρτημένων διανυσμάτων. Στην περίπτωση αυτή έννοια της βάσης αντικαθίσταται από την έννοια του πλαισίου.

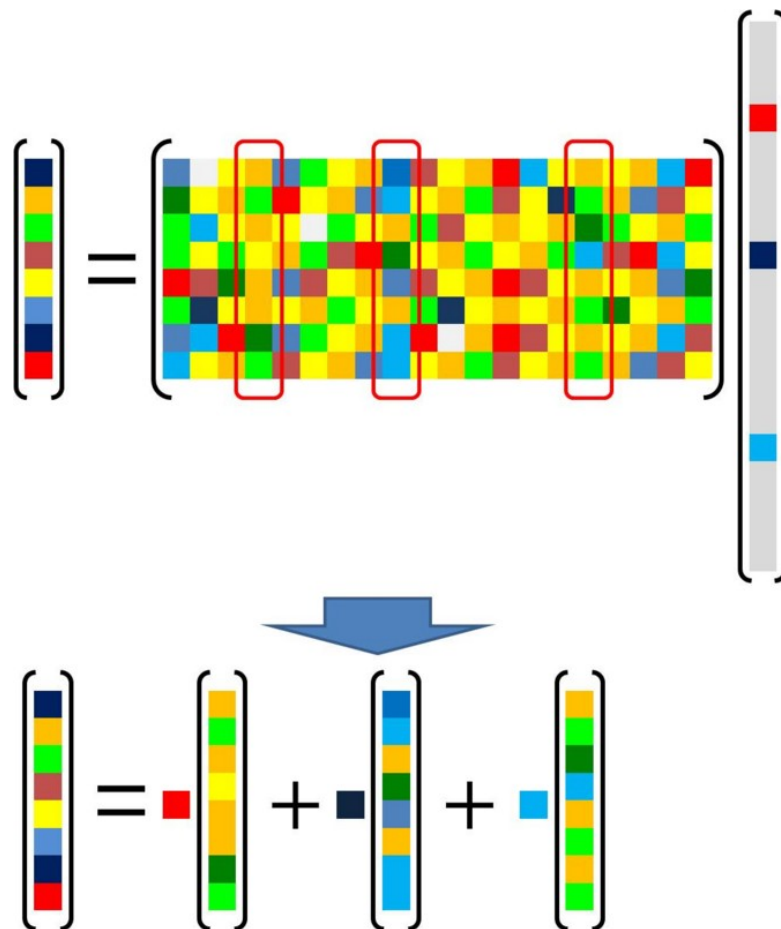
**Ορισμός 4.3.** Έστω ο  $d$ -διάστατος χώρος των πραγματικών αριθμών  $\mathbb{R}^d$ . Ένα σύνολο γραμμικά ανεξάρτητων διανυσμάτων  $\{\boldsymbol{\varphi}_i\}_{i=1}^n$ , με  $n > d$ , ονομάζεται **πλαίσιο (frame)** αν για το  $d \times n$  μητρώο  $\Phi$  με στήλες τα διανύσματα  $\boldsymbol{\varphi}_i$  και για κάθε διάνυσμα  $\mathbf{x}$  του χώρου  $\mathbb{R}^d$  υπάρχουν πραγματικές σταθερές  $A$  και  $B$  με  $0 < A \leq B$  έτσι ώστε

$$A\|\mathbf{x}\|_2^2 \leq \|\Phi^T \mathbf{x}\|_2^2 \leq B\|\mathbf{x}\|_2^2 \quad (29)$$

**Παρατήρηση 4.2.** Το πλαίσιο στην επεξεργασία σήματος και στη μηχανική μάθηση συναντάται με τον όρο **υπερπλήρες λεξικό (overcomplete dictionary)**.

Τα πλαίσια μπορούν να παράγουν καλύτερες αναπαραστάσεις των σημάτων σε σχέση με τις βάσεις εξαιτίας της ύπαρξης περισσότερων ατόμων που μπορούν να χρησιμοποιηθούν ως μονάδες αναπαράστασης. Στην περίπτωση των πλαισίων η εξίσωση (25) με άγνωστο το διάνυσμα  $\mathbf{c}$  των συντελεστών είναι ένα γραμμικό σύστημα με  $n$  αγνώστους και  $d$  εξισώσεις με  $d < n$  (underdetermined system) με άπειρες λύσεις. Στην επεξεργασία σήματος συνήθως ο στόχος είναι η εύρεση του αραιότερου διανύσματος, δηλαδή του διανύσματος με τα περισσότερα μηδενικά.

Στο Σχήμα 9 φαίνεται η βασική ιδέα της αναπαράστασης ενός σήματος με τη βοήθεια ενός υπερπλήρους λεξικού και ενός αραιού διανύσματος.



**Σχήμα 9:** Αναπαράσταση ενός σήματος με τη βοήθεια ενός υπερπλήρους λεξικού και ενός αραιού διανύσματος

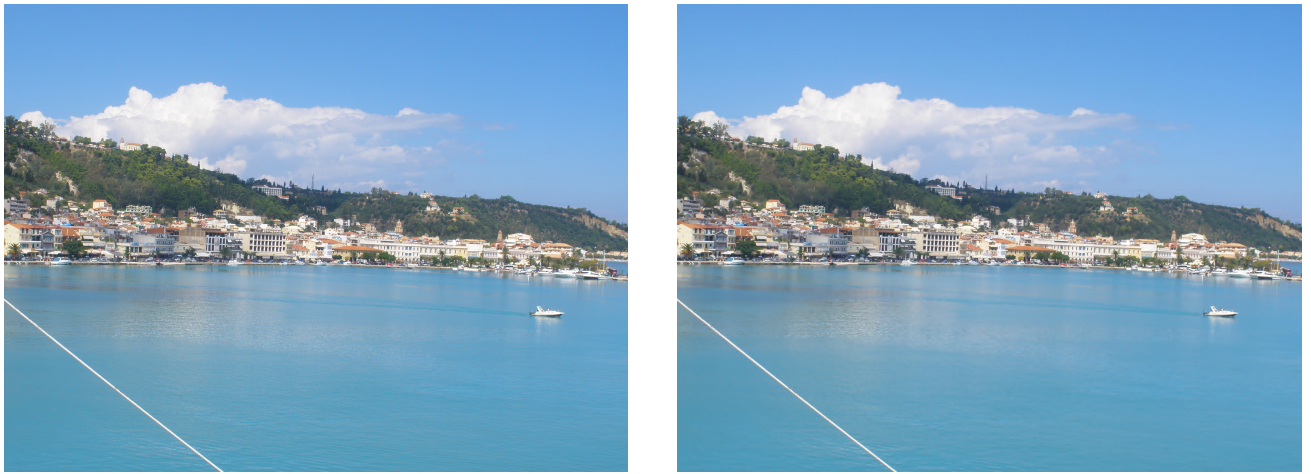
#### 4.4 Αραιά μοντέλα

Ένα σήμα μπορεί τις περισσότερες φορές να αναπαρασταθεί με πολύ καλή προσέγγιση ως γραμμικός συνδυασμός μικρού αριθμού διανυσμάτων μιας βάσης.

**Ορισμός 4.4.** Ένα σήμα το οποίο αναπαρίσταται με ένα διάνυσμα  $\mathbf{x}$  είναι **k-αραιό (k-sparse)** όταν το πλήθος των μη μηδενικών του στοιχείων είναι το πολύ  $k$  δηλαδή ισχύει

$$\|\mathbf{x}\|_0 \leq k. \quad (30)$$

**Παρατήρηση 4.3.** Αυτό που πρακτικά είναι το ζητούμενο δεν είναι η εύρεση αραιών σημάτων αλλά η προσέγγιση αυτών με τη βοήθεια των λεξικών που αναφέρθηκαν στην προηγούμενη ενότητα. Έτσι ένα σήμα  $\mathbf{x}$  θα είναι k-αραιό όταν μπορεί να αναπαρασταθεί με μία εξίσωση της μορφής (25) όπου το διάνυσμα  $\mathbf{c}$  είναι k-αραιό.



Σχήμα 10: Αραιή αναπαράσταση μιας εικόνας. (α) Αρχική εικόνα. (β) Εικόνα που προκύπτει με τη διατήρηση του 1% των συντελεστών του μετασχηματισμού κυματιδίου

Ας υποθέσουμε ότι επιθυμούμε να επιλύσουμε το ακόλουθο σύστημα των γραμμικών εξισώσεων

$$\mathbf{y} = \mathbf{X}\boldsymbol{\theta} \quad (31)$$

που περιγράφει ένα σύνθετο μοντέλο εισόδων-παρατηρήσεων στην επεξεργασία σημάτων. Το  $\mathbf{y}$  είναι το  $N \times 1$  διάνυσμα των παρατηρήσεων,  $\mathbf{X}$  είναι το  $N \times I$  μητρώο εισόδου, το οποίο υποθέτουμε ότι έχει βαθμό  $\text{rank}(\mathbf{X}) = N$ , και  $\boldsymbol{\theta}$  είναι το  $I \times 1$  άγνωστο διάνυσμα παραμέτρων. Επειδή  $N < I$  το παραπάνω γραμμικό σύστημα των εξισώσεων έχει μικρότερο αριθμό εξισώσεων από αγνώστους (underdetermined system) και επιδέχεται άπειρο πλήθος λύσεων. Εξετάζοντας από γεωμετρική άποψη το παραπάνω πρόβλημα, το σύνολο των πιθανών λύσεων του βρίσκεται στην τομή των  $N$  υπερεπιπέδων του  $I$ -διάστατου χώρου, δηλαδή

$$y_n = \mathbf{x}_n^T \boldsymbol{\theta} \quad n = 1, 2, \dots, N \quad \boldsymbol{\theta} \in \mathbb{R}^I. \quad (32)$$

Με δεδομένο ότι το μητρώο  $\mathbf{X}$  είναι βαθμού  $N$  τα  $N$  υπερεπίπεδα που περιγράφονται από την παραπάνω εξίσωση τέμνονται και η τομή τους είναι ένα υπερεπίπεδο διάστασης  $I - N$ . Επομένως το σύνολο όλων των πιθανών λύσεων του συστήματος (32) βρίσκεται σε ένα υπερεπίπεδο διάστασης  $I - N$ . Για την αλγεβρική ερμηνεία των λύσεων αυτών απαιτείται ο ακόλουθος ορισμός:

**Ορισμός 4.5.** Έστω το  $N \times I$  μητρώο  $X$ . Ορίζουμε ως **null space** και συμβολίζουμε με  $null(X)$  το γραμμικό υπόχωρο που ορίζουν τα διανύσματα  $\mathbf{z}$  για τα οποία ισχύει

$$null(X) = \{\mathbf{z} \in \mathbb{R}^I : X\mathbf{z} = \mathbf{0}\} \quad (33)$$

Υποθέτουμε ότι  $\boldsymbol{\theta}_0$  είναι μία λύση του συστήματος (32) και ότι  $\Theta$  είναι το σύνολο των λύσεων. Επομένως  $\forall \boldsymbol{\theta} \in \Theta$  θα ισχύει

$$X(\boldsymbol{\theta} - \boldsymbol{\theta}_0) = \mathbf{0} \Rightarrow \boldsymbol{\theta} - \boldsymbol{\theta}_0 \in null(X) \Rightarrow \Theta = \boldsymbol{\theta}_0 + null(X). \quad (34)$$

Αλλά ο null space του  $N \times I$  μητρώου  $X$ ,  $N$  βαθμού, με  $I > N$  είναι ένας γραμμικός υπόχωρος διάστασης  $I - N$ . Η επιλογή, λοιπόν, μιας συγκεκριμένης λύσης του συστήματος (32) ή διαφορετικά ενός συγκεκριμένου σημείου του υποχώρου  $\Theta$  απαιτεί επιπλέον περιορισμό ή περιορισμούς. Ο περιορισμός που χρησιμοποιούμε είναι οι στάθμες διανυσμάτων. Με αυτόν τον τρόπο αναζητείται μία λύση του συστήματος (32) η οποία ικανοποιεί ένα κριτήριο ελαχιστοποίησης συγκεκριμένης στάθμης. Η γενική μορφή του προκύπτοντος προβλήματος βελτιστοποίησης θα είναι

$$\begin{aligned} \min_{\boldsymbol{\theta} \in \mathbb{R}^I} \quad & \|\boldsymbol{\theta}\|_p \\ \text{s.t.} \quad & y_n = \mathbf{x}_n^T \boldsymbol{\theta} \quad n = 1, 2, \dots, N \end{aligned} \quad (35)$$

Στην περίπτωση που ο στόχος μας είναι ένα διάνυσμα παραμέτρων όσο το δυνατόν πιο αραιό, είναι προφανές ότι η στάθμη η οποία πρέπει να χρησιμοποιηθεί είναι η  $\ell_0$ . Η ελαχιστοποίηση αυτής της στάθμης εξασφαλίζει ότι η λύση που θα προκύψει θα δώσει το βέλτιστο διάνυσμα παραμέτρων  $\boldsymbol{\theta}$  με κριτήριο τον αριθμό των μηδενικών του στοιχείων. Το συγκεκριμένο πρόβλημα βελτιστοποίησης είναι δύσκολο να επιλυθεί και απαιτεί μέθοδο εξαντλητικής αναζήτησης (NP-hard problem). Η μοναδικότητα της λύσης του συγκεκριμένου προβλήματος αναλύεται στην επόμενη ενότητα.

#### 4.5 Μοναδικότητα της αραιότερης λύσης

Για την εύρεση κατάλληλων συνθηκών που εξασφαλίζουν μοναδικότητα στη λύση του προβλήματος βελτιστοποίησης (35) απαιτείται ο ακόλουθος ορισμός:

**Ορισμός 4.6.** Έστω το  $N \times I$  μητρώο  $X$  με  $I \geq N$  και βαθμό  $N$ . Ως **spark** του μητρώου  $X$  ορίζεται ένας αριθμός που συμβολίζεται με  $spark(X)$  και ισούται με το μικρότερο πλήθος των γραμμικά εξαρτημένων στηλών του μητρώου  $X$ .

Με βάση τον παραπάνω ορισμό οποιεσδήποτε  $m$  στήλες του μητρώου  $X$  με  $m < spark(X)$  είναι γραμμικά ανεξάρτητες.

**Παρατήρηση 4.4.** Η εύρεση του **spark** ενός μητρώου δεν μπορεί να υπολογιστεί με κάποιον αποδοτικό αλγόριθμο και απαιτεί μέθοδο εξαντλητικής αναζήτησης.

Το ακόλουθο λήμμα συνδέει το **spark** ενός μητρώου με το πρόβλημα βελτιστοποίησης (35).

**Λήμμα 4.1.** Αν για το σύστημα γραμμικών εξισώσεων (31) υπάρχει μία λύση  $\boldsymbol{\theta}_0$  για την

οποία ισχύει

$$\|\boldsymbol{\theta}\|_0 < \frac{1}{2} \text{spark}(X), \quad (36)$$

τότε το  $\boldsymbol{\theta}_0$  είναι η λύση του προβλήματος βελτιστοποίησης (35) δηλαδή είναι η αραιότερη λύση του συστήματος.

Από τη στιγμή που η εύρεση του  $\text{spark}$  ενός μητρώου είναι ένα δύσκολο πρόβλημα, έχει βρεθεί μια άλλη μετρική που σχετίζεται με την εύρεση του ελαχιστοποιητή  $\ell_0$  για το πρόβλημα (35). Η μετρική αυτή είναι η αμοιβαία συνοχή (mutual coherence) ενός μητρώου.

**Ορισμός 4.7.** Έστω το  $N \times I$  μητρώο  $X$  με στήλες  $\mathbf{x}_i$ ,  $i = 1, 2, \dots, I$ . Τότε ορίζουμε ως **αμοιβαία συνοχή (mutual coherence)** του μητρώου  $X$  και συμβολίζουμε με  $\mu(X)$  την ποσότητα

$$\mu(X) = \max_{1 \leq i < j \leq I} \frac{|\mathbf{x}_i^T \mathbf{x}_j|}{\|\mathbf{x}_i\| \|\mathbf{x}_j\|} \quad (37)$$

**Παρατήρηση 4.5.** Για το mutual coherence ενός μητρώου  $X$  ισχύει πάντα  $0 \leq \mu(X) \leq 1$ . Για ένα  $N \times I$  μητρώο  $X$  με  $I > N$  ισχύει πάντα

$$\sqrt{\frac{I-N}{N(I-1)}} \leq \mu(X) \leq 1. \quad (38)$$

Το παραπάνω κάτω φράγμα είναι γνωστό και ως *Welch bound* και για  $I \gg N$  ισχύει  $\mu(X) \geq \frac{1}{\sqrt{N}}$ .

Το ακόλουθο λήμμα συνδέει το  $\text{spark}$  με το mutual coherence ενός μητρώου.

**Λήμμα 4.2.** Έστω ένα  $N \times I$  μητρώο  $X$ . Τότε θα ισχύει

$$\text{spark}(X) \geq 1 + \frac{1}{\mu(X)} \quad (39)$$

Βασιζόμενοι στο προηγούμενο λήμμα και στο λήμμα 4.1 το επόμενο θεώρημα είναι προφανές.

**Θεώρημα 4.1.** Αν για το σύστημα γραμμικών εξισώσεων (31) υπάρχει μία λύση  $\boldsymbol{\theta}_0$  για την οποία ισχύει

$$\|\boldsymbol{\theta}\|_0 < \frac{1}{2} \left(1 + \frac{1}{\mu(X)}\right), \quad (40)$$

τότε το  $\boldsymbol{\theta}_0$  είναι η λύση του προβλήματος βελτιστοποίησης (35) δηλαδή είναι η αραιότερη λύση του συστήματος.

Με δεδομένο ότι η επίλυση του προβλήματος βελτιστοποίησης (35) είναι ένα NP-hard πρόβλημα, η εύρεση κάποιων φραγμάτων είναι ένας τρόπος αντιμετώπισής του. Αυτό που έχει ιδιαίτερη σημασία και συνοψίζεται στο ακόλουθο θεώρημα είναι ότι κάτω από ορισμέ-

νες συνθήκες είναι δυνατή η ελαχιστοποίηση της στάθμης  $\ell_1$  αντί για τη στάθμη  $\ell_0$  στο πρόβλημα να μην επηρεάσει την προσδοκώμενη λύση που είναι το αραιότερο διάνυσμα παραμέτρων.

**Θεώρημα 4.2.** Έστω το γραμμικό σύστημα εξισώσεων (31) όπου  $X$  ένα μητρώο  $N \times I$  με  $I > N$  και βαθμό  $N$ . Αν το σύστημα έχει λύση η οποία ικανοποιεί τη συνθήκη

$$\|\theta\|_0 < \frac{1}{2} \left(1 + \frac{1}{\mu(X)}\right) \quad (41)$$

τότε αυτή είναι η λύση του προβλήματος βελτιστοποίησης (35) ακόμη και αν στη θέση της στάθμης  $\ell_0$  ελαχιστοποιήσουμε τη στάθμη  $\ell_1$ .

**Παρατήρηση 4.6.** Το παραπάνω θεώρημα που διατυπώθηκε το 2003 ανεξάρτητα από τους Dono και Grib υπήρξε το θεωρητικό επιστέγασμα των μέχρι τότε εμπειρικών αποτελεσμάτων ότι δηλαδή σε πολλές περιπτώσεις ο  $\ell_0$  ελαχιστοποιητής ταυτίζεται με τον  $\ell_1$  ελαχιστοποιητή.

Η ισοδυναμία των  $\ell_0$  και  $\ell_1$  ελαχιστοποιητών θεμελιώνεται με την *Ιδιότητα Περιορισμένης Ισομετρίας (Restricted Isometry Property-RIP)* η οποία δημοσιεύθηκε το 2006 από τον Candes [5].

**Ορισμός 4.8. (Restricted Isometry Property-RIP)** Για κάθε ακέραιο  $k = 1, 2, \dots, n$  και για ένα  $N \times I$  μητρώο  $X$  ορίζεται η **παράμετρος ισομετρίας (isometry parameter)**  $\delta_k$  έτσι ώστε η συνθήκη

$$(1 - \delta_k) \|\theta\|_2^2 \leq \|X\theta\|_2^2 \leq (1 + \delta_k) \|\theta\|_2^2 \quad (42)$$

να ισχύει για όλα τα  $k$ -αραιά διανύσματα  $\theta$ .

Η RIP χρησιμοποιείται στο παρακάτω Θεώρημα

**Θεώρημα 4.3.** Έστω το πρόβλημα βελτιστοποίησης

$$\begin{aligned} \min_{\theta \in \mathbb{R}^I} \quad & \|\theta\|_1 \\ \text{s.t.} \quad & y_n = \mathbf{x}_n^T \theta \quad n = 1, 2, \dots, N \end{aligned} \quad (43)$$

Έστω, επίσης, για την παράμετρο ισομετρίας  $\delta_{2k}$  για το μητρώο  $X = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N]$  για κάποιο  $k$  ισχύει  $\delta_{2k} < \sqrt{2} - 1$ . Τότε, η λύση  $\theta^*$  του  $\ell_1$  ελαχιστοποιητή του παραπάνω προβλήματος βελτιστοποίησης ικανοποιεί τις δύο ακόλουθες συνθήκες:

- $\|\theta - \theta^*\|_1 \leq C_0 \|\theta - \theta_k\|_1$
- $\|\theta - \theta^*\|_2 \leq C_0 k^{-\frac{1}{2}} \|\theta - \theta_k\|_1$

για κάποια σταθερά  $C_0$ . Το διάνυσμα  $\theta$  είναι το αραιό διάνυσμα-στόχος του προβλήματος βελτιστοποίησης ενώ το διάνυσμα  $\theta_k$  είναι το διάνυσμα που προκύπτει από το διάνυσμα  $\theta$  αν διατηρήσουμε τα  $k$  μεγαλύτερα στοιχεία του και μηδενίσουμε τα υπόλοιπα.

## 5. Ο ΑΛΓΟΡΙΘΜΟΣ OMP

### 5.1 Ο αλγόριθμος OMP

Όπως προαναφέρθηκε, από τη στιγμή που η εύρεση του ελαχιστοποιητή της  $\ell_0$  είναι ένα NP-hard πρόβλημα έχουν αναπτυχθεί διάφοροι “άπληστοι” αλγόριθμοι (greedy algorithms) που δίνουν ικανοποιητικά αποτελέσματα στη λύση του προβλήματος. Οι άπληστοι αλγόριθμοι βασίζονται στην επαναληπτική προσέγγιση των συντελεστών της αραιής αναπαράστασης του σήματος.

Ένας αρκετά γνωστός αλγόριθμος είναι ο αλγόριθμος OMP (Orthogonal Matching Pursuit), που βασίζεται στη μέθοδο basis pursuit που αναλύεται στο [7], και προτάθηκε το 1993 από τους Pati, Rezaifar και Krishnaprasad [16]. Ο συγκεκριμένος αλγόριθμος μπορεί να χρησιμοποιηθεί σε ένα από τα δύο στάδια του αλγορίθμου k-SVD που αναπτύσσεται στη συνέχεια και αποτελεί τον κύριο άξονα της παρούσας εργασίας. Το πρόβλημα που “προσπαθεί” να επιλύσει ο αλγόριθμος OMP είναι το γνωστό πρόβλημα βελτιστοποίησης

$$\begin{aligned} \min \quad & \|\mathbf{x}\|_0 \\ \text{s.t.} \quad & \mathbf{Ax} = \mathbf{b} \end{aligned} \quad (44)$$

Ο αλγόριθμος OMP ξεκινά βρίσκοντας τη στήλη του μητρώου  $\mathbf{A}$  που συσχετίζεται περισσότερο με το σήμα  $\mathbf{b}$ . Το βήμα αυτό επαναλαμβάνεται συσχετίζοντας τις στήλες του μητρώου  $\mathbf{A}$  με την απόκλιση εκτίμησης που προκύπτει από τη διαφορά του αρχικού σήματος  $\mathbf{b}$  και της προσέγγισης αυτού από το γινόμενο  $\mathbf{Ax}$ . Το κριτήριο τερματισμού του αλγορίθμου μπορεί να είναι είτε ο αριθμός των μη μηδενικών συντελεστών του ζητούμενου αραιού διάνυσματος  $\mathbf{x}$  είτε ο βαθμός προσέγγισης του σήματος  $\mathbf{b}$  από το γινόμενο  $\mathbf{Ax}$ . Με δεδομένο το μητρώο  $\mathbf{A}$ , το διάνυσμα  $\mathbf{b}$  και ένα κατώφλι σφάλματος  $\varepsilon_0$  ο αλγόριθμος OMP διακρίνεται σε δύο φάσεις:

#### Αλγόριθμος OMP:

**Αρχικοποίηση.** Αρχικοποίησε τον αριθμό των επαναλήψεων  $k = 0$  και θέσε:

- Αρχική λύση  $\mathbf{x}^0 = 0$ .
- Αρχικό διάνυσμα απόκλισης εκτίμησης (residual)  $\mathbf{r}^0 = \mathbf{b} - \mathbf{Ax}$ .
- Αρχικό “στήριγμα” της λύσης (solution support)  $\mathcal{S}^0 = \mathbf{x}^0$ .

**Επαναληπτικό σχήμα.** Αύξησε το  $k$  κατά 1 και εκτέλεσε τα παρακάτω βήματα:

- Υπολόγισε τα σφάλματα  $\varepsilon(j) = \min_{z_j} \|\mathbf{a}_j z_j - \mathbf{r}^{k-1}\|_2^2$  για όλα τα  $j$  χρησιμοποιώντας την καλύτερη επιλογή  $z_j^* = \mathbf{a}_j^T \mathbf{r}^{k-1} / \|\mathbf{a}_j\|_2^2$ .
- Βρες έναν ελαχιστοποιητή  $j_0$  του σφάλματος  $\varepsilon(j)$  έτσι ώστε  $\forall j \notin \mathcal{S}^{k-1}$  να ισχύει  $\varepsilon(j_0) \leq \varepsilon(j)$  και ανανέωσε το “στήριγμα”  $\mathcal{S}^k = \mathcal{S}^{k-1} \cup \{j_0\}$ .
- Υπολόγισε τον ελαχιστοποιητή  $\mathbf{x}^k$  της ποσότητας  $\|\mathbf{Ax} - \mathbf{b}\|_2^2$  ως προς το στήριγμα  $\mathcal{S}^k$ .
- Υπολόγισε το νέο διάνυσμα απόκλισης εκτίμησης  $\mathbf{r}^k = \mathbf{b} - \mathbf{Ax}^k$ .
- Αν  $\|\mathbf{r}^k\|_2 < \varepsilon_0$  τερμάτισε τον αλγόριθμο αλλιώς συνέχισε στην επόμενη επανάληψη.

**Παρατήρηση 5.1.** Η υπολογιστική πολυπλοκότητα του αλγορίθμου OMP είναι  $\mathcal{O}(k \log n)$  όπου  $k$  είναι ο βαθμός αραιότητας του διάνυσματος  $\mathbf{x}$  και  $n$  είναι η διάσταση του σήματος  $\mathbf{b}$ . Αυτό σημαίνει ότι με μεγάλη πιθανότητα-λόγω της απληστίας του-ο αλγόριθμος OMP βρίσκει τη βέλτιστη  $k$ -αραιή αναπαράσταση του  $n$ -διάστατου σήματος  $\mathbf{b}$  σε  $k \log n$  το πολύ

βήματα.

**Παρατήρηση 5.2.** Εκτός από τον αλγόριθμο OMP, ένας άλλος γνωστός αλγόριθμος για την επίλυση του συγκεκριμένου προβλήματος είναι ο αλγόριθμος LASSO (Least Absolute Shrinkage and Selection Operator) τον οποίο δε θα χρησιμοποιήσουμε στην παρούσα εργασία.

## 5.2 Υπερπλήρες λεξικό με βάση το δισδιάστατο μετασχηματισμό DCT

Η έννοια του υπερπλήρους λεξικού περιγράφηκε στην ενότητα 4.3. Στην περίπτωση των σημάτων ψηφιακών εικόνων το πιο σύνηθες λεξικό που χρησιμοποιείται για την αραιή αναπαράστασή του είναι το δισδιάστατο υπερπλήρες λεξικό **Διακριτού Μετασχηματισμού Συνημιτόνου (Discrete Cosine Transform)** ή εν συντομία **2D-DCT**. Το εν λόγω λεξικό είναι ένα μητρώο διαστάσεων  $p \times \sqrt{m}$  όπου  $p \times p$  είναι το μέγεθος των blocks-patches στην οποία διαιρούμε την εικόνα και χρησιμοποιείται στην τεχνική των overlapped patches, μέθοδος που περιγράφεται σε επόμενη ενότητα. Επίσης  $m$  είναι ο αριθμός των στοιχείων (atoms) από τα οποία αποτελείται το υπερπλήρες λεξικό. Επειδή το λεξικό είναι υπερπλήρες πρέπει να ισχύει  $m > p^2$ . Προκειμένου να σχηματίσουμε το 2D-DCT υπερπλήρες λεξικό ακολουθούμε τα παρακάτω βήματα [19]:

- Δημιούργησε ένα μητρώο  $\tilde{D}$  με διανύσματα στήλες τα  $\mathbf{d}_i$  με  $\mathbf{d}_i = [d_{i,1}, d_{i,2}, \dots, d_{i,p}]$  και  $i = 0, 1, \dots, \sqrt{m} - 1$  με

$$d_{i,t+1} = \cos\left(\frac{t\pi i}{\sqrt{m}}\right), t = 0, 1, \dots, p - 1. \quad (45)$$

- Για όλες τις στήλες εκτός από την πρώτη βρες το μέσο όρο τους και αφάιρεσε τον από κάθε στοιχείο της στήλης.
- Κανονικοποίησε κάθε στήλη έτσι ώστε κάθε διάνυσμα στήλη να έχει μοναδιαίο μέτρο. Προκύπτει ένα μητρώο  $D$  διαστάσεων  $p \times \sqrt{m}$  που είναι το μητρώο μετασχηματισμού DCT.
- Χρησιμοποίησε το γινόμενο Kronecker  $\otimes$  προκειμένου να εξάγουμε το τελικό μητρώο  $\psi = D \otimes D$  που είναι το 2D-DCT υπερπλήρες λεξικό διάστασης  $p^2 \times m$ .

Στο Σχήμα 11 φαίνονται τα 196 άτομα του 2D-DCT υπερπλήρους λεξικού.

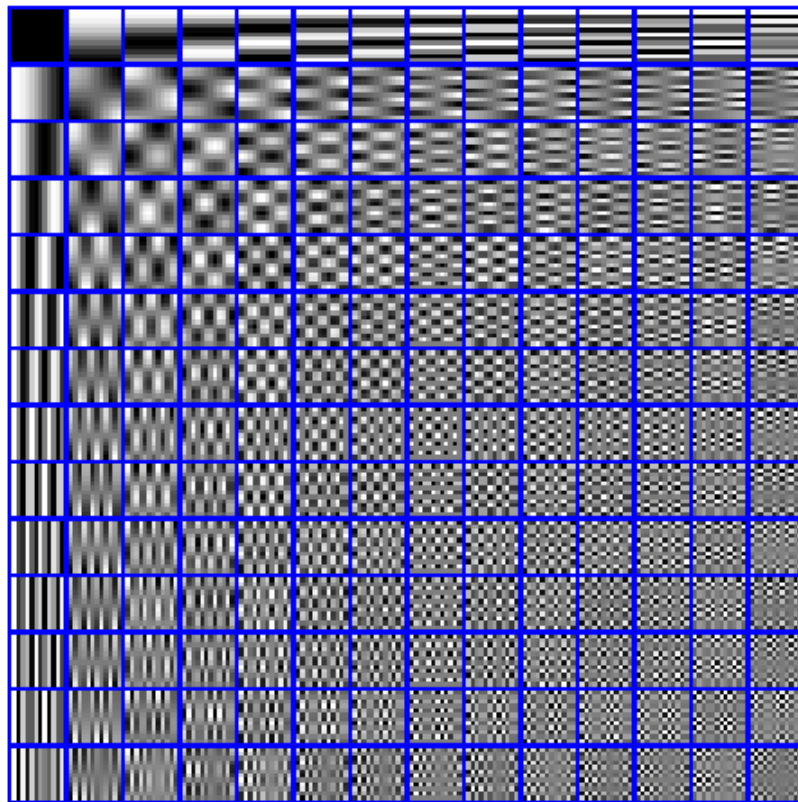
Στα Σχήμα 12 και 13 φαίνεται η αραιή αναπαράσταση μία εικόνας με τη βοήθεια του 2D-DCT υπερπλήρους λεξικού του Σχήματος 11.

## 5.3 Αποθορυβοποίηση εικόνας με βάση τον αλγόριθμο OMP

Στην προηγούμενη ενότητα, είδαμε ότι με τη βοήθεια του αλγορίθμου OMP και βάση ένα 2D-DCT υπερπλήρες λεξικό μπορούμε να αναπαραστήσουμε μία εικόνα με πολύ λιγότερα δεδομένα χωρίς να χάνουμε ιδιαίτερα σε ποιότητα. Επομένως ο αλγόριθμος OMP είναι ένα χρήσιμο εργαλείο στη συμπίεση εικόνας. Αυτό που μπορούμε επίσης να επιτύχουμε με τα ίδια συστατικά είναι η αποθορυβοποίηση μιας ψηφιακής εικόνας.

Παίρνοντας ως αρχικό σήμα μία ενθόρυβη εικόνα, μπορούμε να περιορίσουμε μέρος του θορύβου εφόσον αναπαραστήσουμε το μητρώο της εικόνας ως γινόμενο αραιών διανυσμάτων και ενός σταθερού κατάλληλου λεξικού (fixed dictionary). Το σταθερό αυτό λεξικό στην περίπτωσή μας είναι το 2D-DCT υπερπλήρες λεξικό. Η απόρριψη μέρους του θορύβου εξηγείται ως εξής: Ο αλγόριθμος OMP προσπαθεί να εντοπίσει τα πλέον κατάλ-





Σχήμα 11: Το 2D-DCT υπερπλήρες λεξικό



Original image



Sparse representation of the image with 10 coefficients OMP algorithm

Σχήμα 12: Αραιή αναπαράσταση μία εικόνας με τη βοήθεια του 2D-DCT υπερπλήρους λεξικού 196 ατόμων και του αλγορίθμου OMP με 10 συντελεστές





Original image



Sparse representation of the image with 3 coefficients OMP algorithm

**Σχήμα 13:** Αραιή αναπαράσταση μία εικόνας με τη βοήθεια του 2D-DCT υπερπλήρους λεξικού 196 ατόμων και του αλγορίθμου OMP με 3 συντελεστές



Original image



Noisy image with gaussian noise  $\sigma=20$ , PSNR=22.19



Denoised image with 5 coefficients OMP algorithm PSNR=30.23

**Σχήμα 14:** Αποθρομβοποίηση μία εικόνας με τη βοήθεια του 2D-DCT υπερπλήρους λεξικού 196 ατόμων και του αλγορίθμου OMP με 5 συντελεστές

ληλα άτομα του λεξικού προκειμένου να αναπαραστήσει το αρχικό σήμα εικόνας. Επομένως, θεωρώντας ότι η εικόνα χωρίς θόρυβο είναι αποτέλεσμα σύνθεσης κάποιων από τα άτομα του λεξικού τα οποία δεν περιέχουν θόρυβο, πραγματοποιείται μία προσέγγιση της ζητούμενης εικόνας στην οποία ο θόρυβος έχει μερικώς εξαφανιστεί.

Το προηγούμενο σκεπτικό φαίνεται υλοποιημένο στα Σχήματα 14 και 15, στα οποία απεικονίζεται η αποθρομβοποίηση μίας εικόνας, στην οποία έχει επιδράσει gaussian θόρυβος διασποράς  $\sigma = 20$ , με τη βοήθεια του αλγορίθμου OMP και ενός 2D-DCT υπερπλήρους λεξικού 196 ατόμων διάστασης  $12 \times 12$  το καθένα. Ο αριθμός των συντελεστών που έχει χρησιμοποιηθεί για την αραιή αναπαράσταση της εικόνας είναι 5 και 15 αντίστοιχα.

**Παρατήρηση 5.3.** Η εφαρμογή του αλγορίθμου OMP σε ένα πρόβλημα αποθρομβοποίησης σήματος γίνεται με τη βοήθεια μικρού αριθμού συντελεστών των αραιών διανυσμάτων. Για παράδειγμα, στο Σχήμα 14 έχουν χρησιμοποιηθεί 5 συντελεστές για την αποθρομβοποίηση της εικόνας με αποτέλεσμα καλύτερο σε σχέση με το 15 όπου έχουν χρησιμοποιηθεί 15 συντελεστές για την εφαρμογή του αλγορίθμου OMP. Αυτό μπορεί να εξηγηθεί ως φαινόμενο υπερταυρίσματος (overfitting), αφού ο θόρυβος με την αύξηση των



Original image



Noisy image with gaussian noise  $\sigma=20$ , PSNR=22.16



Denoised image with 15 coefficients OMP algorithm PSNR=28.24

**Σχήμα 15: Αποθορυβοποίηση μία εικόνας με τη βοήθεια του 2D-DCT υπερπλήρους λεξικού 196 ατόμων και του αλγορίθμου OMP με 15 συντελεστές**

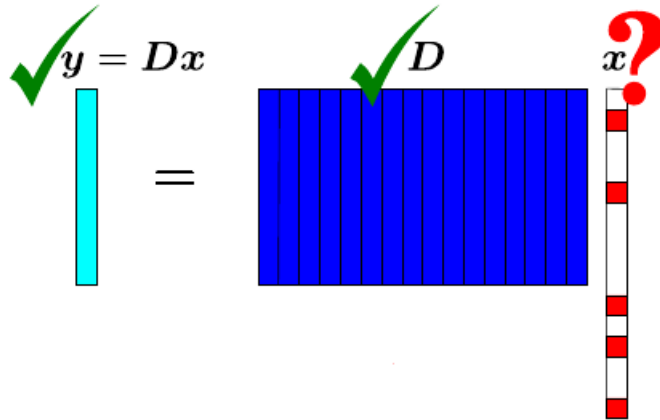
*συντελεστών “καταφέρνει” να προσαρμοστεί στο προκύπτον σήμα με αποτέλεσμα να μην απορρίπτεται.*

**Παρατήρηση 5.4.** Στην αραιή αναπαράσταση μιας εικόνας με τη βοήθεια ενός 2D-DCT υπερπλήρους λεξικού, όπως φαίνεται στα Σχήματα 12 και 13 η αύξηση των συντελεστών του αλγορίθμου OMP οδηγεί σε καλύτερα αποτελέσματα σε αντίθεση με την αποθορυβοποίηση όπου η αύξηση των συντελεστών (από ένα σημείο και έπειτα) οδηγεί σε μη δυνατότητα απόρριψης του θορύβου από τον αλγόριθμο.

## 6. Ο ΑΛΓΟΡΙΘΜΟΣ K-SVD

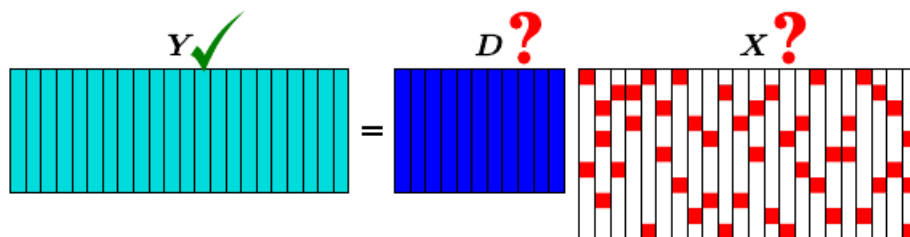
### 6.1 Εκμάθηση Λεξικού (Dictionary Learning)

Στο πρόβλημα αναπαράστασης σήματος σε συμπιεστή (αραιή) μορφή (Compressed Sensing-CS), το ζητούμενο είναι η εύρεση των συντελεστών του αραιού διανύσματος  $\mathbf{x}$  με δεδομένο το αρχικό σήμα  $\mathbf{y}$  και ένα κατάλληλο υπερπλήρες λεξικό  $D$  όπως είναι το 2D-DCT υπερπλήρες λεξικό. Η απεικόνιση του προβλήματος φαίνεται στο Σχήμα 16.



Σχήμα 16: Το πρόβλημα Compressed Sensing (CS)

Πολλές φορές, όμως, η ύπαρξη ενός σταθερού λεξικού (fixed dictionary) δημιουργεί προβλήματα στην αναπαράσταση του σήματος καθώς είναι αδύνατο ένα λεξικό να καλύψει ένα σύνολο ομοειδών σημάτων. Το γεγονός αυτό δημιούργησε την ανάγκη κατασκευής λεξικών που προέρχονται από τα δεδομένα του κατά περίπτωση προβλήματος δηλαδή από το αρχικό δεδομένο σήμα. Η μη χρήση σταθερού λεξικού δημιουργεί στο πρόβλημα CS έναν επιπλέον άγνωστο. Το νέο πρόβλημα, που ονομάζεται εκμάθηση λεξικού (Dictionary Learning-DL), απεικονίζεται στο Σχήμα 17 και συνίσταται στην εύρεση ενός κατάλληλου λεξικού  $D$  και ενός αραιού διανύσματος  $\mathbf{x}$  προκειμένου να επιτευχθεί η βέλτιστη αραιή αναπαράσταση του σήματος  $\mathbf{y}$  [17], [6].



Σχήμα 17: Το πρόβλημα Dictionary Learning (DL)

Πιο συγκεκριμένα, υποθέτουμε ότι έχουμε  $n$   $N$  – διστατα διανύσματα  $\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n$  που είναι το σύνολο των αρχικών μας σημάτων. Το ζητούμενο είναι η εύρεση ενός μητρώου-λεξικού  $D$  διάστασης  $N \times \ell$  και  $n$   $k$ -αραιών  $\ell$ -διάστατων διανυσμάτων  $\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \dots, \boldsymbol{\theta}_n$  έτσι ώστε να ισχύει

$$\mathbf{y}_i \approx D\boldsymbol{\theta}_i, \quad i = 1, 2, \dots, n. \quad (46)$$

Χρησιμοποιώντας ως μετρική ποσοτικοποίησης του σφάλματος τη στάθμη Frobenius,

το πρόβλημα DL μπορεί να γραφεί ως πρόβλημα βελτιστοποίησης σύμφωνα με την εξίσωση (47).

$$\begin{aligned} \min_{D, \theta_1, \theta_2, \dots, \theta_n} \quad & \sum_{i=1}^n \|\mathbf{y}_i - D\theta_i\|_F^2 \\ \text{s.t.} \quad & \|\theta_i\|_0 \leq k, i = 1, 2, \dots, n. \end{aligned} \quad (47)$$

Κάθε στήλη  $\mathbf{d}_i$  του μητρώου-λεξικού  $D$  εκφράζει ένα άτομο του. Συνεπώς το λεξικό  $D$  αποτελείται από  $\ell$  άτομα διάστασης  $N$  το καθένα. Επιπρόσθετα, κάθε στήλη  $\mathbf{d}_i$  του μητρώου-λεξικού  $D$  κανονικοποιείται στη μοναδιαία Ευκλείδεια στάθμη, ήτοι  $\|\mathbf{d}_i\|_2$ , προκειμένου με αυτόν τον τρόπο το λεξικό  $D$  και τα αραιά διανύσματα  $\theta_i$  να εξάγονται με μοναδικό τρόπο με βάση έναν αλγόριθμο DL. Για ένα υπερπλήρες λεξικό θα ισχύει  $\ell > N$  ενώ συνήθως ο αριθμός των μη μηδενικών συντελεστών των αραιών διανυσμάτων  $\theta_i$  είναι πολύ μικρότερος από τη διάσταση των σημάτων, ήτοι  $k \ll N$ .

**Παρατήρηση 6.1.** Στην ειδική περίπτωση που ο βαθμός αραιότητας των διανυσμάτων  $\theta_i$  είναι  $k = 1$ , το πρόβλημα DL ανάγεται σε ένα πρόβλημα συσταδοποίησης (clustering). Στην περίπτωση αυτή, ο αλγόριθμος DL προσπαθεί να κατασκευάσει ένα κατάλληλο λεξικό  $D$  και να βρει το καταλληλότερο άτομο του λεξικού προκειμένου να αναπαραστήσει κάθε σήμα  $\mathbf{y}_i$  ως γινόμενο ενός συντελεστή και ενός ατόμου.

**Παρατήρηση 6.2.** Για την ποσοτικοποίηση του σφάλματος στην εξίσωση (47) μπορούμε να χρησιμοποιήσουμε και άλλες μετρικές, όπως για παράδειγμα τη στάθμη  $\ell_1$ .

Το πρόβλημα βελτιστοποίησης (47) μπορεί να γραφεί με τη βοήθεια μητρώων θεωρώντας ως  $Y = [\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n]$  το μητρώο σημάτων και ως  $\Theta = [\theta_1, \theta_2, \dots, \theta_n]$  το μητρώο των αραιών διανυσμάτων-συντελεστών. Στην περίπτωση αυτή το πρόβλημα γράφεται ως

$$\begin{aligned} \min_{A, Z} \quad & \|X - AZ\|_F^2 \\ \text{s.t.} \quad & \|\mathbf{z}_n\|_0 \leq k, n = 1, 2, \dots, N, \end{aligned} \quad (48)$$

και ανάγεται σε ένα πρόβλημα παραγοντοποίησης μητρώων (matrix factorization problem).

Επίσης, υπάρχουν περιπτώσεις που στο πρόβλημα βελτιστοποίησης (47) στόχος μας είναι όσο το δυνατόν πιο αραιή αναπαράσταση των σημάτων με δεδομένο ένα κατώφλι σφάλματος  $\varepsilon$ . Στην περίπτωση αυτή το πρόβλημα γράφεται ως

$$\begin{aligned} \min_{D, \theta_1, \theta_2, \dots, \theta_n} \quad & \sum_{i=1}^n \|\mathbf{y}_i - D\theta_i\|_F^2 \\ \text{s.t.} \quad & \|\mathbf{y}_i - D\theta_i\|^2 \leq \varepsilon, i = 1, 2, \dots, n. \end{aligned} \quad (49)$$

Οι δύο μορφές του προβλήματος DL, όπως περιγράφεται στις εξισώσεις (47) και (49), μπορούν να συγχωνευθούν σε μία αν αντί των κατωφλίων αραιότητας  $k$  και σφάλματος  $\varepsilon$  χρησιμοποιήσουμε μία παράμετρο κανονικοποίησης  $\lambda$  που εκφράζει το βάρος που επιθυμούμε να δώσουμε στην αραιότητα και στο σφάλμα προσέγγισης. Το πρόβλημα DL με τη βοήθεια αυτής της παραμέτρου γράφεται ως

$$\min_{D, \theta_1, \theta_2, \dots, \theta_n} \sum_{i=1}^n \|\mathbf{y}_i - D\theta_i\|_F^2 + \lambda \|\theta_i\|_0 \quad (50)$$

Όσο μεγαλύτερη είναι η παράμετρος κανονικοποίησης  $\lambda$ , τόσο μεγαλύτερη είναι η βαρύτητα που δίνεται στην αραιότητα μέσω της στάθμης  $\ell_0$  και επομένως η αναπαράσταση του σήματος είναι κατά το δυνατόν αραιή σε βάρος του προσεγγιστικού σφάλματος  $\varepsilon$ . Αντίθετα, όσο μικρότερη είναι η παράμετρος κανονικοποίησης  $\lambda$ , τόσο πιο μικρό είναι το προσεγγιστικό σφάλμα εις βάρος της αραιότητας.

Η ύπαρξη της στάθμης  $\ell_0$  στο πρόβλημα DL έχει ως αποτέλεσμα την ανάγκη ανάπτυξης μεθόδων εξαντλητικής αναζήτησης προκειμένου να βρεθεί η λύση του προβλήματος βελτιστοποίησης (πρόβλημα κλάσης NP). Σε αρκετές περιπτώσεις η στάθμη  $\ell_0$  αντικαθίσταται από τη στάθμη  $\ell_1$  καθώς όπως έχει επισημανθεί στην ενότητα η  $\ell_1$  στάθμη είναι η πιο κοντινή στάθμη στη στάθμη  $\ell_0$  που διατηρεί τον κυρτό χαρακτήρα στο πρόβλημα. Στην περίπτωση αυτή το πρόβλημα βελτιστοποίησης (50) γράφεται ως

$$\min_{D, \theta_1, \theta_2, \dots, \theta_n} \sum_{i=1}^n \|\mathbf{y}_i - D\theta_i\|_F^2 + \lambda \|\theta_i\|_1 \quad (51)$$

και είναι ένα κυρτό πρόβλημα του οποίου η λύση προσεγγίζει τη βέλτιστη λύση του προβλήματος DL.

Παρόλο που τόσο το πρόβλημα CS όσο και το πρόβλημα DL συνίστανται στο παραπάνω πρόβλημα βελτιστοποίησης με ένα ή δύο αγνώστους ο τελικός τους στόχος είναι διαφορετικός. Στο πρόβλημα CS ο στόχος είναι, έπειτα από την προσεκτική επιλογή του σταθερού λεξικού  $D$ , η εύρεση των καταλληλότερων αραιών διανυσμάτων, για την αναπαράσταση του αρχικού σήματος. Στο πρόβλημα DL, το αρχικό σήμα χρησιμοποιείται ως δεδομένο εκπαίδευσης, με στόχο την επιλογή των καταλληλότερων ατόμων για τη δημιουργία του λεξικού  $D$ . Επίσης, το αραιό διάνυσμα που υπολογίζεται από έναν αλγόριθμο DL, είναι μία έκφραση προκειμένου να απεικονιστεί η βέλτιστη αραιή γραμμική αναπαράσταση του σήματος από το λεξικό.

## 6.2 Ο αλγόριθμος k-SVD

Ο αλγόριθμος k-SVD προτάθηκε το 2006 από τους Aharon, Elad και Bruckstein [1] και έχει γίνει ιδιαίτερα δημοφιλής σε προβλήματα επεξεργασίας σήματος όπως είναι η αποθορυβοποίηση και η συμπίεση ενός σήματος. Ο στόχος του αλγορίθμου είναι η αναπαράσταση ενός σήματος ως γινόμενο ενός αραιού διανύσματος και ενός υπερπλήρους λεξικού με όσο το δυνατόν πιο ακριβή τρόπο. Η διαφορά, σε σχέση με τους αλγορίθμους LASSO και OMP είναι ότι το υπερπλήρες λεξικό δεν είναι σταθερό (fixed) αλλά μεταβάλλεται σε κάθε βήμα του αλγορίθμου μαζί με το αραιό διάνυσμα συντελεστών.

Στην ουσία ο αλγόριθμος k-SVD προσπαθεί σε κάθε του βήμα να βρει το αραιότερο διάνυσμα και το καταλληλότερο υπερπλήρες λεξικό ώστε να αναπαραστήσει ένα σήμα. Πιο συγκεκριμένα, έστω ότι υπάρχουν  $N$  δεδομένα  $l$ -διάστατα σήματα  $\mathbf{x}_n$ ,  $n = 1, 2, \dots, N$ , καθένα από τα οποία επιθυμούμε να αναπαραστήσουμε με τη μορφή γινομένου ως

$$\mathbf{x}_n = A\mathbf{z}_n \quad (52)$$

όπου  $A$  είναι το μητρώο που παριστάνει το υπερπλήρες λεξικό διάστασης  $l \times m$  και το  $z$  είναι το αραιό διάνυσμα συντελεστής. Ο πραγματικός στόχος δεν είναι η ακριβής αναπαράσταση του σήματος με τη βοήθεια του παραπάνω γινομένου αλλά η καλύτερη δυνατή προσέγγιση αυτού υπό τον περιορισμό της αραιότητας του διανύσματος συντελεστή. Έτσι το πρόβλημα βελτιστοποίησης που ανακύπτει είναι το

$$\begin{aligned} \min_{A, Z} \quad & \|X - AZ\|_F^2 \\ \text{s.t.} \quad & \|z_n\|_0 \leq T_0, n = 1, 2, \dots, N, \end{aligned} \quad (53)$$

όπου το μητρώο  $X$  είναι το  $l \times N$  μητρώο που έχει ως στήλες τα διανύσματα που αναπαριστούν τα σήματα  $\mathbf{x}_n$ , ήτοι  $X = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N\}$ , το μητρώο  $Z$  είναι το  $m \times N$  μητρώο με στήλες τα αραιά διανύσματα συντελεστές, ήτοι  $Z = \{z_1, z_2, \dots, z_N\}$ , και  $T_0$  είναι το κατώφλι που προσδιορίζει το βαθμό αραιότητας κάθε διανύσματος.

Προκειμένου να επιλυθεί το παραπάνω πρόβλημα ελαχιστοποίησης, ο κλασικός τρόπος αντιμετώπισης είναι η διαίρεση του κάθε βήματος του αλγορίθμου σε δύο ανεξάρτητα στάδια. Στο πρώτο στάδιο που ονομάζεται αραιή κωδικοποίηση (sparse coding) ο αλγόριθμος διατηρεί σταθερό το υπερπλήρες λεξικό  $A$  και βρίσκει μέσω ενός αλγορίθμου όπως ο OMP ή ο LASSO το βέλτιστο αραιό διάνυσμα συντελεστή  $z_n$  για κάθε σήμα  $\mathbf{x}_n$ . Στο δεύτερο στάδιο, που ονομάζεται ενημέρωση λεξικού (codebook update) διατηρούνται σταθεροί οι συντελεστές  $z_n$  και αναζητείται η βέλτιστη λύση του για τις στήλες του λεξικού  $A$ . Τα δύο στάδια του αλγορίθμου περιγράφονται πιο αναλυτικά ως εξής:

1. *Αραιή κωδικοποίηση.* Θεώρησε το λεξικό  $A$  σταθερό, όπως έχει διαμορφωθεί με βάση την τελευταία επανάληψη του αλγορίθμου και επίλυσε το παρακάτω πρόβλημα βελτιστοποίησης:

$$\begin{aligned} \min_Z \quad & \|X - AZ\|_F^2 \\ \text{s.t.} \quad & \|z_n\|_0 \leq T_0, n = 1, 2, \dots, N, \end{aligned} \quad (54)$$

το οποίο είναι ισοδύναμο με την επίλυση  $N$  προβλημάτων βελτιστοποίησης της μορφής

$$\begin{aligned} \min_{z_n} \quad & \|\mathbf{x}_n - Az_n\|^2 \\ \text{s.t.} \quad & \|z_n\|_0 \leq T_0, n = 1, 2, \dots, N. \end{aligned} \quad (55)$$

Ισοδύναμα, εφόσον ο περιορισμός δεν είναι το κατώφλι αραιότητας  $T_0$  αλλά το προσεγγιστικό σφάλμα  $\varepsilon$  στη στάθμη της διαφοράς του σήματος  $\mathbf{x}_n$  από την εκτίμησή του  $Az_n$  τότε το παραπάνω πρόβλημα βελτιστοποίησης μπορεί να γραφεί ως

$$\begin{aligned} \min_{z_n} \quad & \|z_n\|_0 \\ \text{s.t.} \quad & \|\mathbf{x}_n - Az_n\|^2 < \varepsilon, n = 1, 2, \dots, N. \end{aligned} \quad (56)$$

2. *Ενημέρωση λεξικού.* Με δεδομένο ότι στο πρώτο στάδιο του αλγορίθμου έχουν εξαχθεί τα βέλτιστα αραιά διανύσματα συντελεστές  $z_n$ , ο στόχος σε αυτό το στάδιο είναι η βελτιστοποίηση κάθε στήλης του λεξικού  $A$ . Υποθέτουμε ότι βρισκόμαστε στο βήμα ενημέρωσης της στήλης  $\mathbf{a}_k$  του μητρώου  $A$  με στόχο να ελαχιστοποιήσουμε την ποσότητα  $\|X - AZ\|_F^2$ . Το γινόμενο των μητρώων  $AZ$  μπορεί να γραφεί ως άθροισμα μητρώων βαθμού ένα ως εξής:

$$AZ = \sum_{i=1}^m \alpha_i \mathbf{z}_i^T. \quad (57)$$

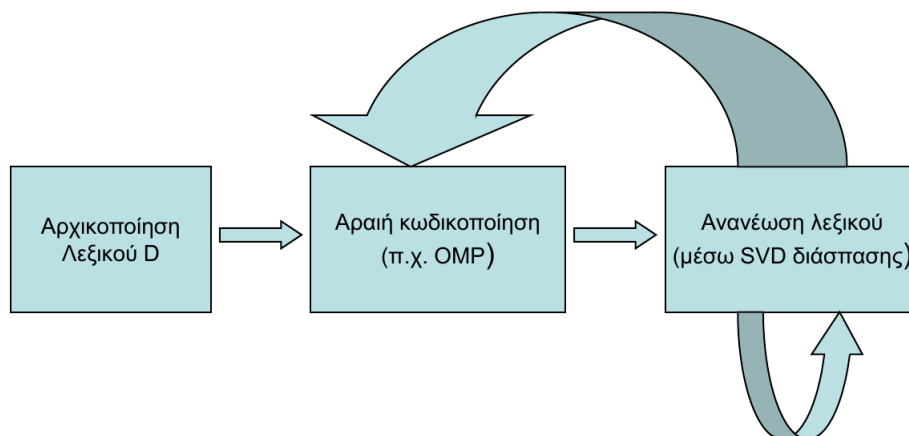
όπου τα  $\mathbf{z}_i^T, i = 1, 2, \dots, m$  είναι οι γραμμές του μητρώου  $Z$ .

Συνοπτικά, κάθε επαναληπτικό βήμα του αλγορίθμου περιλαμβάνει τα ακόλουθα υπολογιστικά βήματα:

#### Αλγόριθμος k-SVD:

- Αρχικοποίησε το μητρώο  $A^{(0)}$  με στήλες κανονικοποιημένες στη μοναδιαία  $\ell_2$  στάθμη.
- Θέσε  $i = 1$ .
- **Στάδιο 1.** Επίλυσε το πρόβλημα βελτιστοποίησης (53) για την εύρεση των βέλτιστων διανυσμάτων αραιής κωδικοποίησης  $\mathbf{z}_n, n = 1, 2, \dots, N$  με τη χρήση οποιουδήποτε κατάλληλου αλγορίθμου (OMP, LASSO, κ.α.).
- **Στάδιο 2.** Για κάθε στήλη  $k = 1, 2, \dots, m$  του μητρώου  $A^{(t-1)}$  ανανέωσε (update) με βάση τα παρακάτω:
  - Βρες τις θέσεις των μη μηδενικών στοιχείων της  $k$ -οστής γραμμής του μητρώου  $Z$  που έχει υπολογιστεί από το **Στάδιο 1**.
  - Επέλεξε τις στήλες του μητρώου  $X$  που αντιστοιχούν στις θέσεις των μη μηδενικών στοιχείων της  $k$ -οστής γραμμής του μητρώου  $Z$  και κατασκεύασε με αυτές ένα μικρότερης τάξης μητρώο σφάλματος  $\tilde{E}_k$ .
  - Κάνε διάσπαση ιδιαζουσών τιμών (SVD) στο μητρώο  $\tilde{E}_k$  ως  $\tilde{E}_k = UDV^T$ .
  - Ανανέωσε (update) την  $k$ -οστή στήλη του μητρώου  $A^{(i)}$  έτσι ώστε να είναι ίδια με το ιδιοδιάνυσμα που αντιστοιχεί στη μεγαλύτερη ιδιάζουσα τιμή,  $\alpha_k^{(i)} = \mathbf{u}_1$ .
  - Ανανέωσε το μητρώο  $Z$ , ενσωματώνοντας στις θέσεις των μη μηδενικών της  $k$ -οστής γραμμής του μητρώου, τις τιμές  $D(1, 1)\mathbf{z}_1^T$ .
  - Σταμάτα εφόσον ικανοποιείται το κριτήριο τερματισμού.
  - Αλλιώς θέσε  $i = i + 1$  και συνέχισε.

Μία σχηματική αναπαράσταση του αλγορίθμου k-SVD φαίνεται στο Σχήμα 18.



Σχήμα 18: Μία σχηματική αναπαράσταση του αλγορίθμου k-SVD

**Παρατήρηση 6.3.** Ο αλγόριθμος  $k$ -SVD έχει πολλά πλεονεκτήματα. Είναι εύκολος στην υλοποίηση, γρήγορος και συγκλίνει στις περισσότερες περιπτώσεις στη βέλτιστη λύση. Η επιλογή του λεξικού αρχικοποίησης του αλγορίθμου είναι κρίσιμη για τη σύγκλιση του αλγορίθμου στη βέλτιστη λύση.

**Παρατήρηση 6.4.** Η σύγκλιση του αλγορίθμου  $k$ -SVD εξαρτάται από την αποτελεσματικότητα του αλγορίθμου που χρησιμοποιείται στο *sparse coding step*.

**Παρατήρηση 6.5.** Ο αλγόριθμος  $k$ -SVD αποτελεί γενίκευση του αλγορίθμου  $k$ -means που αποτελεί κλασικό αλγόριθμο προβλημάτων συσταδοποίησης (*clustering*). Στην ουσία, ο αλγόριθμος  $k$ -means είναι ειδική περίπτωση του αλγορίθμου  $k$ -SVD αν στο πρόβλημα βελτιστοποίησης (53) θέσουμε ως κατώφλι αραιότητας  $T_0 = 1$ .



## 7. Ο ΑΛΓΟΡΙΘΜΟΣ K-SVD ΣΤΗΝ ΑΠΟΘΟΡΥΒΟΠΟΙΗΣΗ ΣΗΜΑΤΟΣ

### 7.1 Η μετρική PSNR

Το κύριο μέρος της εργασίας παρατίθεται σε αυτό το κεφάλαιο και αφορά την εφαρμογή του αλγορίθμου k-SVD προκειμένου να “αποθορυβοποιήσουμε” ένα σήμα. Για λόγους οπτικοποίησης χρησιμοποιούμε σήματα εικόνων σε grayscale μορφή με αναπαράσταση χρωμάτων 8 bits. Πάνω στις εικόνες “επιδρούμε” διάφορα ήδη θορύβων προκειμένου να μελετήσουμε την αποτελεσματικότητα του αλγορίθμου k-SVD. Για κάθε είδος θορύβου που μελετάμε παραθέτουμε 3 εικόνες, την αρχική, την ενθόρυβη και την τελική εικόνα μετά την εφαρμογή του αλγορίθμου. Η μετρική που χρησιμοποιείται προκειμένου να “ποσοτικοποιήσουμε” την ποιότητα των αποτελεσμάτων είναι η αναλογία **Peak Signal-to-Noise Ratio (PSNR)** η οποία σε dB ορίζεται ως εξής:

$$PSNR_{dB} = 10 \log_{10} \left( \frac{MAX_I^2}{MSE} \right) \quad (58)$$

όπου η ποσότητα **MSE** αντιπροσωπεύει το **Μέσο Τετραγωνικό Σφάλμα (Mean Squared Error)** του προκύπτοντος σήματος-εικόνας σε σχέση με το αρχικό σήμα. Πιο συγκεκριμένα αν  $I$  είναι η αρχική εικόνα και  $K$  είναι η εικόνα που αποτελεί την προσέγγιση της αρχικής τότε το MSE δίνεται από τον τύπο:

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [I(i,j) - K(i,j)]^2 \quad (59)$$

Επιπρόσθετα, το  $MAX_I$  στην εξίσωση (58) αντιπροσωπεύει τη μέγιστη τιμή που μπορεί να πάρει η ένταση του pixel της εικόνας  $I$ . Έτσι στην περίπτωση μας, οι 8-bits εικόνες συνεπάγονται ότι  $MAX_I = 255$ .

### 7.2 Η μέθοδος των επικαλυπτόμενων περιοχών

Για την εφαρμογή ενός αλγορίθμου DL στην περίπτωση της αποθορυβοποίησης εικόνας χρησιμοποιείται η τεχνική των **επικαλυπτόμενων περιοχών (overlapped patches)** όπως περιγράφεται στο [19].

Με βάση τη μέθοδο αυτή από μία  $n \times n$  εικόνα εξάγουμε όλες τις  $m \times m$  περιοχές της εικόνας με  $m \ll n$ . Μία τέτοια  $n \times n$  εικόνα έχει  $(n - m + 1)^2$  overlapped patches διάστασης  $m \times m$ . Έτσι, από μία εικόνα  $256 \times 256$  μπορούμε να εξάγουμε  $(256 - 12 + 1)^2 = 60025$  overlapped patches διάστασης  $12 \times 12$  το καθένα. Κάθε patch στη συνέχεια μετατρέπεται σε διάνυσμα στήλη με  $m^2$  στοιχεία. Τοποθετώντας αυτά τα διανύσματα με τη σειρά δημιουργείται ένα μητρώο διαστάσεων  $m^2 \times (n - m + 1)^2$  το οποίο αποτελεί το μητρώο που αντιπροσωπεύει το αρχικό ενθόρυβο σήμα. Προκειμένου να ανακατασκευάσουμε το σήμα της εικόνας, δηλαδή ένα μητρώο διαστάσεων  $n \times n$ , υπολογίζουμε για κάθε εικονοστοιχείο (pixel) το μέσο όρο της έντασης. Ένας τρόπος για να επιτευχθεί αυτό είναι η κατασκευή ενός μητρώου  $W$  ίδιων διαστάσεων με αυτό της εικόνας, δηλαδή  $n \times n$ , του οποίου κάθε στοιχείο  $(i,j)$  εκφράζει τον αριθμό των patches στα οποία συμμετέχει κάθε pixel. Για παράδειγμα, για μία  $8 \times 8$  εικόνα και για μέγεθος patch  $2 \times 2$  το μητρώο  $W$  θα έχει ως εξής:

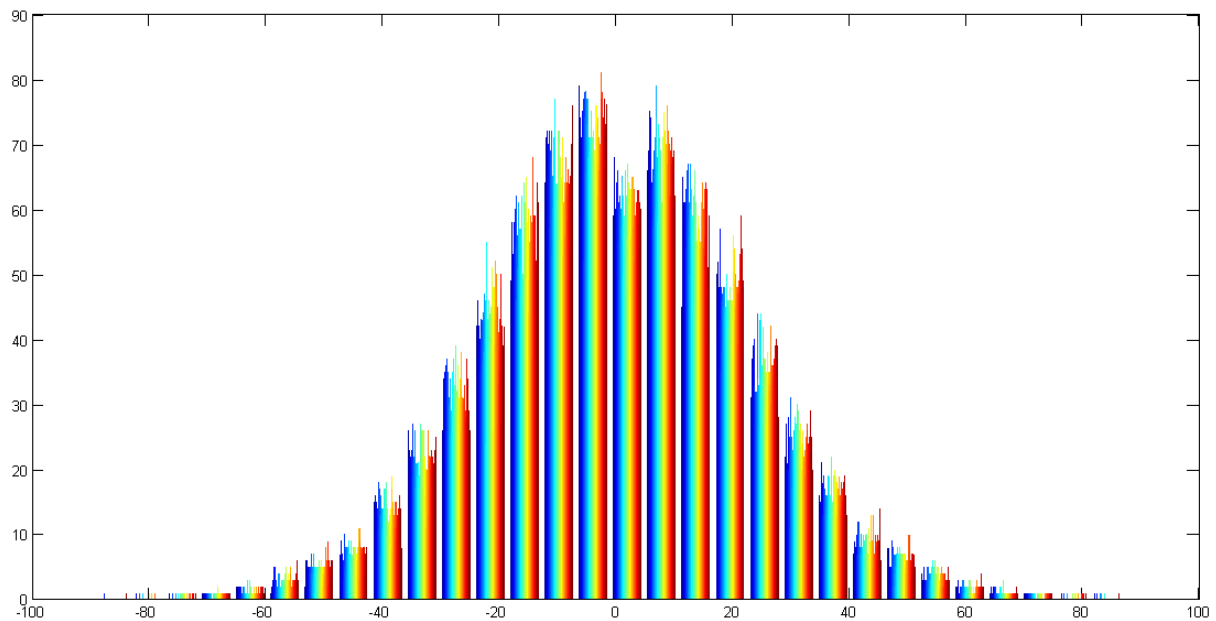
$$W = \begin{pmatrix} 1 & 2 & \dots & 2 & 1 \\ 2 & 4 & \dots & 4 & 2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 2 & 4 & \dots & 4 & 2 \\ 1 & 2 & \dots & 2 & 1 \end{pmatrix}. \quad (60)$$

### 7.3 Αποθορυβοποίηση σήματος με gaussian θόρυβο

Υποθέτουμε ότι σε μία εικόνα επιδρά προσθετικά θόρυβος ο οποίος ακολουθεί κανονική κατανομή. Στα μοντέλα που χρησιμοποιούμε ο gaussian θόρυβος είναι μηδενικής μέσης τιμής δηλαδή με πυκνότητα πιθανότητας όπως φαίνεται παρακάτω.

$$p_N(z) = \frac{1}{\sigma\sqrt{2\pi}} \exp^{-\frac{z^2}{2\sigma^2}} \quad (61)$$

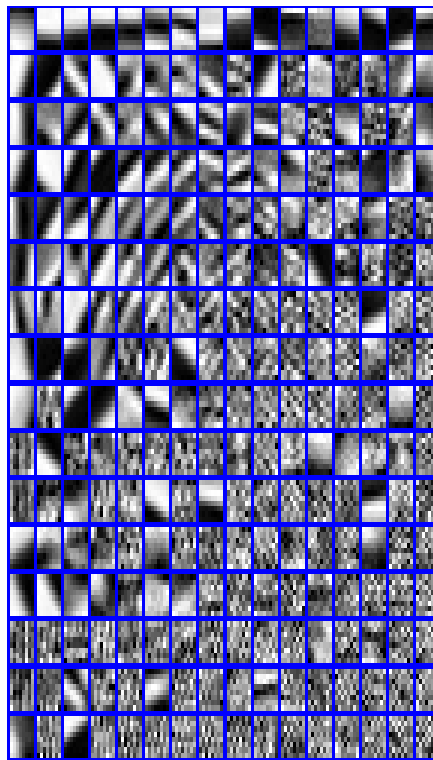
Στο Σχήμα 19 φαίνεται η κατανομή gaussian θορύβου διασποράς  $\sigma = 20$  που έχει επιδράσει προσθετικά πάνω σε μία εικόνα  $512 \times 512$  pixels της οποίας κάθε pixel αναπαρίσταται με τη βοήθεια 8 bits δηλαδή παίρνει τιμές από 0 έως 255. Ο στόχος είναι η ανίχνευση και η απαλοιφή του θορύβου με δεδομένο το σήμα της ενθόρυβης εικόνας.



**Σχήμα 19:** Η κατανομή του gaussian θορύβου διασποράς  $\sigma=20$  πάνω στο σήμα μιας ψηφιακής εικόνας 8 bits

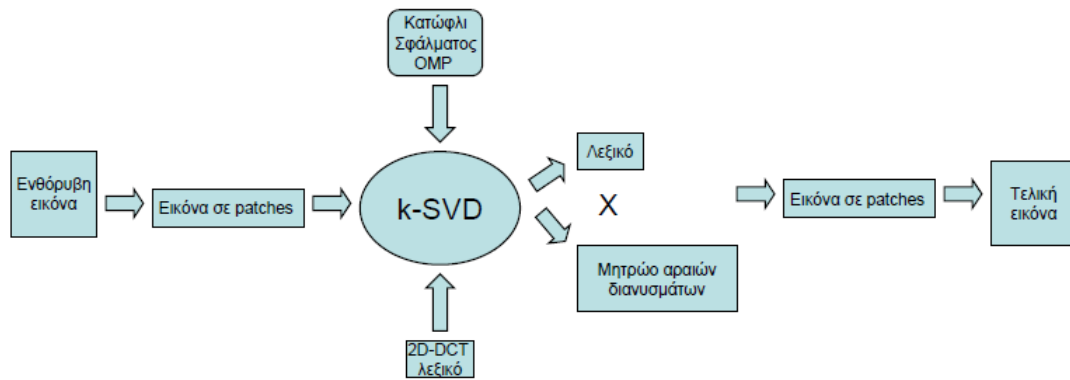
Προκειμένου να εφαρμόσουμε τον αλγόριθμο k-SVD μετασχηματίζουμε το δεδομένο σήμα της ενθόρυβης εικόνας σύμφωνα με τη μέθοδο που περιγράφηκε στην ενότητα 7.2. Το αρχικό σήμα, που στην περίπτωσή μας, είναι ένα μητρώο  $512 \times 512$  μετατρέπεται με τη βοήθεια της τεχνικής των overlapped patches με μέγεθος patch  $8 \times 8$  σε ένα μητρώο διαστάσεων  $64 \times 255025$ . Κάθε στήλη αυτού του μητρώου αναπαριστάει ένα patch της εικόνας όπου τα στοιχεία του  $8 \times 8$  μητρώου patch έχουν αναδιαταχθεί κατά στήλες προκειμένου να σχηματίσουν ένα διάνυσμα στήλη 64 στοιχείων. Εν συνεχεία, εφαρμόζουμε

τον αλγόριθμο k-SVD προκειμένου να αναπαραστήσουμε κάθε διάνυσμα patch ως γινόμενο ενός αραιού διανύσματος και ενός λεξικού το οποίο εξάγεται από το αρχικό ενθόρυβο σήμα. Ως αρχικό λεξικό χρησιμοποιούμε ένα υπερπλήρες λεξικό του δισδιάστατου μετασχηματισμού DCT όπως περιγράφηκε στην ενότητα 5.2. Η έξοδος του αλγορίθμου έχει ως αποτέλεσμα ένα  $64 \times 256$  μητρώο που αποτελεί το λεξικό, που αποτελείται από 256 patches όπως φαίνεται στο Σχήμα 20, και ένα  $256 \times 255025$  μητρώο του οποίου κάθε στήλη αναπαριστά ένα αραιό διάνυσμα. Ο πολλαπλασιασμός του λεξικού με κάθε αραιό διάνυσμα εξάγει τα νέα διανύσματα τα οποία με αναδιάταξη των στοιχείων τους μετατρέπονται σε  $8 \times 8$  patches. Τέλος, με τη βοήθεια ενός μητρώου βαρών όμοιου με αυτό της εξίσωσης (60) εξάγουμε την τιμή της έντασης για κάθε pixel της νέας εικόνας ως το μέσο όρο των τιμών του pixel στα διαφορετικά patches στα οποία συμμετέχει. Στο Σχήμα 21 φαίνεται μία σχηματική αναπαράσταση της παραπάνω διαδικασίας.



**Σχήμα 20:** Το λεξικό που προκύπτει από την εφαρμογή του αλγορίθμου k-SVD σε εικόνα επηρεασμένη από gaussian θόρυβο

Το κρίσιμο σημείο είναι η επιλογή της κατάλληλης παραμέτρου αραιότητας του αλγορίθμου k-SVD προκειμένου να γίνει όσο το δυνατόν μεγαλύτερη απόρριψη θορύβου. Στη συγκεκριμένη περίπτωση του gaussian θορύβου και έπειτα από εμπειρικές μελέτες [10], [14] τα καλύτερα αποτελέσματα επιτυγχάνονται με την επιλογή του κατωφλίου σφάλματος ως  $\varepsilon = 1.15 * m * \sigma$  όπου  $m$  είναι η διάσταση του πλαισίου και  $\sigma$  είναι η διασπορά του gaussian θορύβου. Η ρίζα του συγκεκριμένου κατωφλίου σφάλματος συναντάται και ως ακτίνα του αλγορίθμου OMP (radius OMP) που χρησιμοποιείται στο sparse coding step του αλγορίθμου k-SVD. Η συγκεκριμένη αυτή τιμή του κατωφλίου σφάλματος δίνει αρκετά μικρές τιμές στο κατώφλι αραιότητας που πλησιάζουν τη μονάδα. Η τιμή του κατωφλίου αραιότητας δεν εξαρτάται μόνο από το κατώφλι σφάλματος αλλά και από τη διάσταση του patch και το πλήθος των ατόμων του λεξικού. Για παράδειγμα, για διάσταση patch  $m = 8$  και πλήθος ατόμων λεξικού  $K = 256$ , ο αλγόριθμος k-SVD για την αποθορυβοποίηση μιας  $512 \times 512$  εικόνας που έχει επηρεαστεί από gaussian θόρυβο διασποράς  $\sigma = 20$  έχει ως αποτέλεσμα τιμή κατωφλίου αραιότητας  $T = 1,57$ . Επομένως κάθε αραιό διάνυσμα συντελεστής διάστασης  $256 \times 1$  έχει κατά μέσο όρο 1,57 μη μηδενικά στοιχεία.



**Σχήμα 21:** Σχηματική αναπαράσταση της εφαρμογής του αλγορίθμου k-SVD στην περίπτωση αποθρομβοποίησης εικόνας επηρεασμένης από gaussian θόρυβο



Original image



Noisy image with gaussian noise  $\sigma=10$ , PSNR=26.16



Denised image with K-SVD PSNR=34.59

**Σχήμα 22:** Η εφαρμογή του αλγορίθμου k-SVD σε μία εικόνα με gaussian θόρυβο διασποράς  $\sigma=10$

Ο αλγόριθμος k-SVD δίνει πολύ καλά αποτελέσματα όταν εφαρμοστεί πάνω σε ένα σήμα που επηρεάζεται από gaussian θόρυβο. Απαραίτητη προϋπόθεση για την εξαγωγή βέλτιστου αποτελέσματος στην ανάκτηση του αρχικού σήματος είναι η εκ των προτέρων γνώση της διασποράς του gaussian θορύβου. Η παράμετρος της διασποράς χρησιμοποιείται στο sparse coding step του αλγορίθμου και εκφράζει το μέτρο της αραιότητας που θέλουμε να επιτύχουμε στην αναπαράσταση του σήματος με τη βοήθεια του υπερπλήρους λεξικού. Στα Σχήματα 22 και 23 φαίνεται η εφαρμογή του αλγορίθμου πάνω σε μία εικόνα η οποία έχει επηρεαστεί από gaussian θόρυβο διασποράς  $\sigma=10$  και  $\sigma=20$  αντίστοιχα.

**Παρατήρηση 7.1.** Ο αλγόριθμος k-SVD έχει πολύ καλή επίδοση πάνω σε κανονικές κατανομές θορύβου. Το γεγονός αυτό μπορεί να χρησιμοποιηθεί προκειμένου να εφαρμοστεί ο αλγόριθμος και σε διαφορετικές περιπτώσεις θορύβου όπως είναι ο θόρυβος poisson ή ο θόρυβος με ομοιόμορφη κατανομή χωρίς τα αποτελέσματα να είναι εξίσου καλά αλλά ικανοποιητικά σε ένα βαθμό.

**Παρατήρηση 7.2.** Προκειμένου η τιμή της παραμέτρου στο sparse coding step του αλγορίθμου να είναι κατάλληλη, είναι αναγκαία η εκ των προτέρων γνώση της διασποράς του gaussian θορύβου από τον οποίο έχει επηρεαστεί το αρχικό σήμα της εικόνας.



Original image

Noisy image with gaussian noise  $\sigma=20$ , PSNR=22.21

Denoised image with k-SVD PSNR=30.94

**Σχήμα 23:** Η εφαρμογή του αλγορίθμου k-SVD σε μία εικόνα με gaussian θόρυβο διασποράς  $\sigma=20$

#### 7.4 Αποθορυβοποίηση σήματος με poisson θόρυβο

Στην περίπτωση του θορύβου poisson κάθε pixel της εικόνας επηρεάζεται από θόρυβο κατανομής poisson διασποράς ίσης με την τιμή της έντασης του συγκεκριμένου pixel. Στην ουσία, για τη μοντελοποίηση του θορύβου παράγεται ένας τυχαίος αριθμός από την κατανομή poisson με παράμετρο  $\lambda$  ίση με την τιμή της έντασης του pixel, ο οποίος προστίθεται στην τιμή του pixel. Στο Σχήμα 24 φαίνεται η κατανομή του θορύβου poisson που έχει επιδράσει προσθετικά πάνω σε μία εικόνα  $512 \times 512$  pixels της οποίας κάθε pixel αναπαρίσταται με τη βοήθεια 8 bits.

Για την απαλοιφή του θορύβου poisson, ο αλγόριθμος k-SVD εφαρμόζεται με τον ίδιο ακριβώς τρόπο όπως και στην περίπτωση του gaussian θορύβου. Η διαφορά στον αλγόριθμο σε σχέση με αυτόν της προηγούμενης ενότητας είναι ότι χρησιμοποιείται ως παράμετρος διασποράς η διασπορά της κατανομής του θορύβου poisson. Έτσι, αν ο θόρυβος ακολουθεί κατανομή poisson που δίνεται από τον τύπο

$$f(k; \lambda) = \frac{\lambda^k e^{-\lambda}}{k!} \quad (62)$$

τότε η διασπορά του θορύβου είναι  $\sigma = \lambda$  και χρησιμοποιείται ως παράμετρος στο sparse coding step του αλγορίθμου. Επειδή η παράμετρος  $\lambda$  είναι διαφορετική για κάθε pixel της εικόνας, υπολογίζουμε τη μέση τιμή της έντασης όλων των pixels της ενθόρυβης εικόνας και στη συνέχεια χρησιμοποιούμε αυτήν την τιμή ως παράμετρο στον αλγόριθμο.

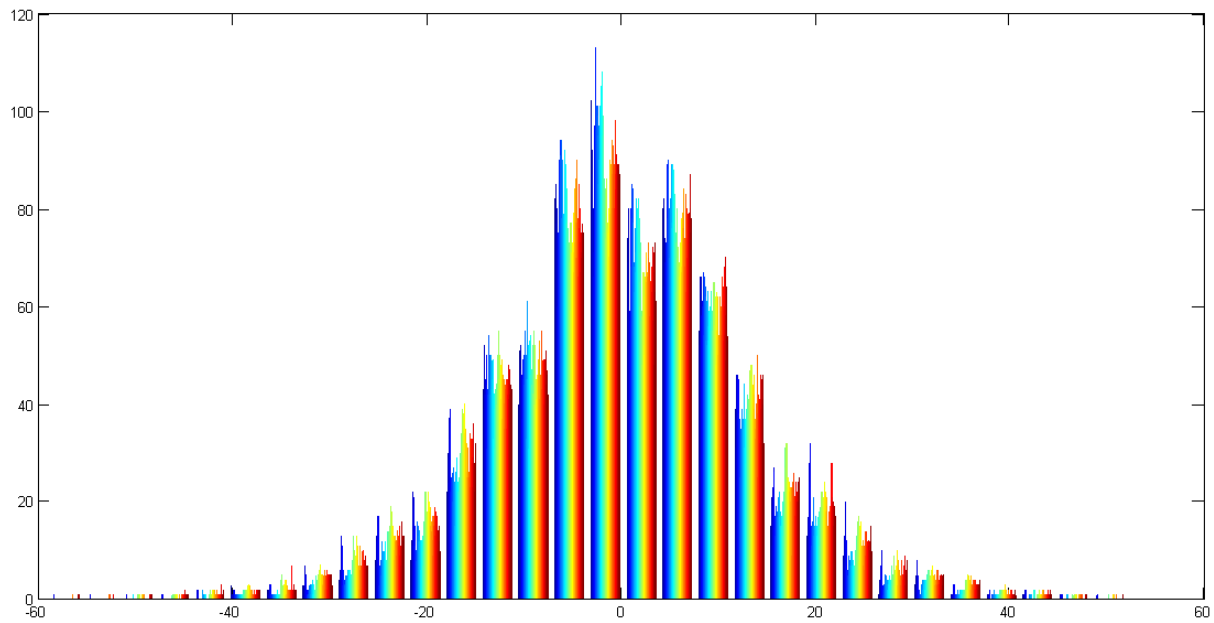
Στα Σχήματα 25 και 26 φαίνεται η εφαρμογή του αλγορίθμου πάνω σε δύο εικόνες οι οποίες έχουν επηρεαστεί από θόρυβο poisson.

**Παρατήρηση 7.3.** Ο αλγόριθμος k-SVD, όπως εφαρμόζεται παραπάνω, εκλαμβάνει το θόρυβο poisson ως gaussian θόρυβο με διασπορά που υπολογίζεται με βάση τις τιμές των pixels της ενθόρυβης εικόνας. Η θεώρηση αυτή δεν απέχει πολύ από την πραγματικότητα αν συγκρίνουμε τις μορφές κατανομών του θορύβου στα Σχήματα 19 και 24.

#### 7.5 Αποθορυβοποίηση σήματος με θόρυβο ομοιόμορφης κατανομής

Στην περίπτωση που ο θόρυβος του σήματος προέρχεται από ομοιόμορφη κατανομή, τότε χρησιμοποιούμε τον ίδιο αλγόριθμο όπως και στις δύο προηγούμενες περιπτώσεις. Έτσι, αν ο θόρυβος ακολουθεί ομοιόμορφη κατανομή σε ένα διάστημα  $[a, b]$  με τύπο

$$f(x) = \frac{1}{b - a + 1} \quad (63)$$



**Σχήμα 24:** Η κατανομή του poisson θορύβου διασποράς  $\sigma=20$  πάνω στο σήμα μιας ψηφιακής εικόνας 8 bits



Original image



Noisy image with poisson noise,PSNR=27.05



Denoised image with k-SVD PSNR=32.73

**Σχήμα 25:** Η εφαρμογή του αλγορίθμου k-SVD σε μία εικόνα με poisson θόρυβο(1)



Original image



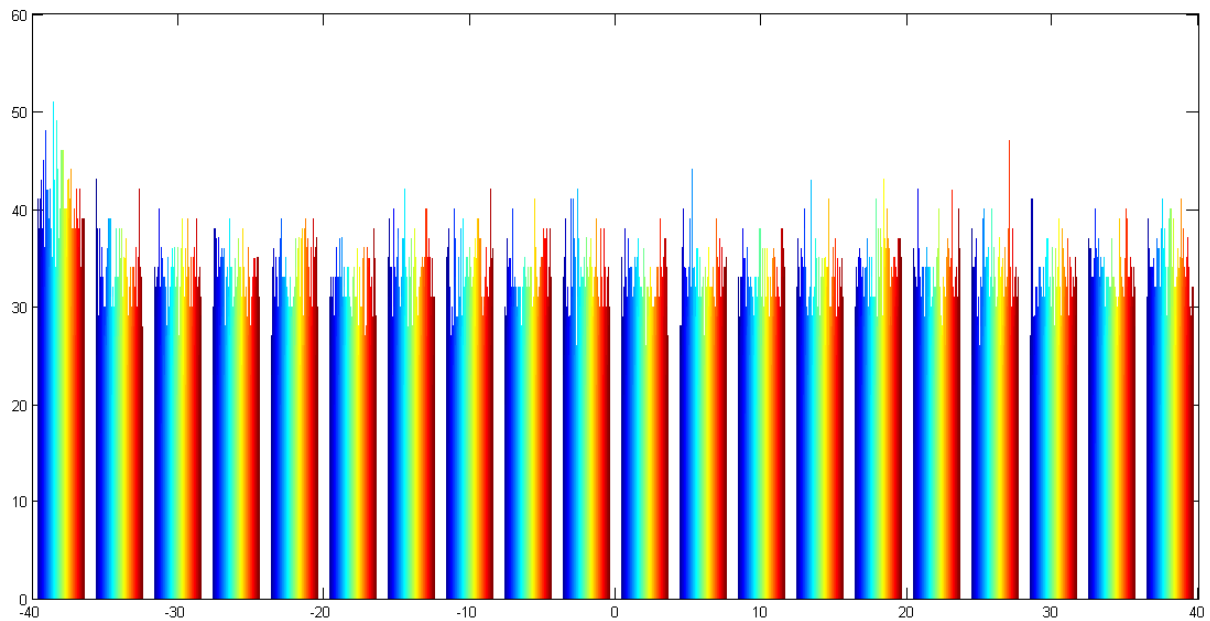
Noisy image with poisson noise,PSNR=27.23



Denoised image with k-SVD PSNR=34.36

**Σχήμα 26:** Η εφαρμογή του αλγορίθμου k-SVD σε μία εικόνα με poisson θόρυβο(2)





**Σχήμα 27:** Η κατανομή του θορύβου ομοιόμορφης κατανομής διασποράς πάνω στο σήμα μιας ψηφιακής εικόνας 8 bits



Original image



Noisy image with uniform noise [-20,20], PSNR=26.7



Denoised image with k-SVD PSNR=33.71

**Σχήμα 28:** Η εφαρμογή του αλγορίθμου k-SVD σε μία εικόνα με θόρυβο ομοιόμορφης κατανομής [-20,20]

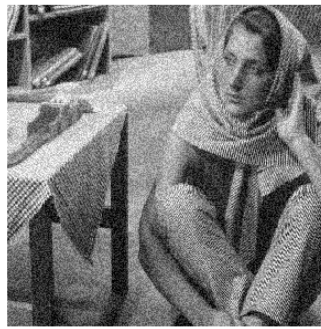
τότε η διασπορά του θορύβου είναι  $\sigma = \frac{(b - a + 1)^2 - 1}{12}$  η οποία χρησιμοποιείται ως παράμετρος στο sparse coding step του αλγορίθμου. Στα Σχήματα 28 και 29 φαίνεται η εφαρμογή του αλγορίθμου πάνω σε μία εικόνα η οποία έχει επηρεαστεί από θόρυβο ομοιόμορφης κατανομής στα διαστήματα [-20,20] και [-40,40] αντίστοιχα.

**Παρατήρηση 7.4.** Ο αλγόριθμος k-SVD, όπως εφαρμόζεται παραπάνω, εκλαμβάνει το θόρυβο ομοιόμορφης κατανομής ως gaussian θόρυβο με διασπορά που υπολογίζεται με βάση τις τιμές των pixels της ενθόρυβης εικόνας. Η θεώρηση αυτή δίνει αρκετά καλά αποτελέσματα αν και η κατανομή του θορύβου με βάση τα Σχήματα 19 και 27 φαίνεται τελείως διαφορετική.

Στον πίνακα 1 φαίνονται τα συγκριτικά αποτελέσματα του προτεινόμενου αλγορίθμου σε σχέση με τα αποτελέσματα του Kernel-based αλγορίθμου και του αλγορίθμου BM3D όπως περιγράφονται στο [3].



Original image



Noisy image with uniform noise [-40,40], PSNR=20.83



Denoised image with K-SVD PSNR=30.15

**Σχήμα 29: Η εφαρμογή του αλγορίθμου k-SVD σε μία εικόνα με θόρυβο ομοιόμορφης κατανομής [-40,40]**

| Image   | Noise    | noisy PSNR (dB) | PSNR (dB) με Kernel denoising [3] | PSNR (dB) με BM3D [3] | PSNR (dB) με k-SVD |
|---------|----------|-----------------|-----------------------------------|-----------------------|--------------------|
| Lena    | $\pm 20$ | 26.70           | 33.00                             | 34.99                 | 34.82              |
|         | $\pm 30$ | 23.25           | 30.81                             | 33.51                 | 33.14              |
|         | $\pm 40$ | 20.79           | 29.41                             | 32.01                 | 31.82              |
| Peppers | $\pm 20$ | 26.73           | 31.78                             | 34.42                 | 34.22              |
|         | $\pm 30$ | 23.31           | 30.28                             | 32.94                 | 32.78              |
|         | $\pm 40$ | 20.89           | 28.81                             | 31.57                 | 31.72              |
| Barbara | $\pm 20$ | 26.71           | 27.16                             | 34.13                 | 33.73              |
|         | $\pm 30$ | 23.24           | 25.95                             | 32.66                 | 31.69              |
|         | $\pm 40$ | 20.82           | 24.19                             | 30.90                 | 30.18              |
| Boat    | $\pm 20$ | 26.73           | 31.15                             | 34.09                 | 32.85              |
|         | $\pm 30$ | 23.28           | 29.38                             | 32.11                 | 30.95              |
|         | $\pm 40$ | 20.85           | 27.47                             | 30.40                 | 29.62              |

**Πίνακας 1: Συγκριτικά αποτελέσματα του αλγορίθμου k-SVD σε σχέση με τους αλγορίθμους Kernel-based και BM3D για θορύβο ομοιόμορφης κατανομής**

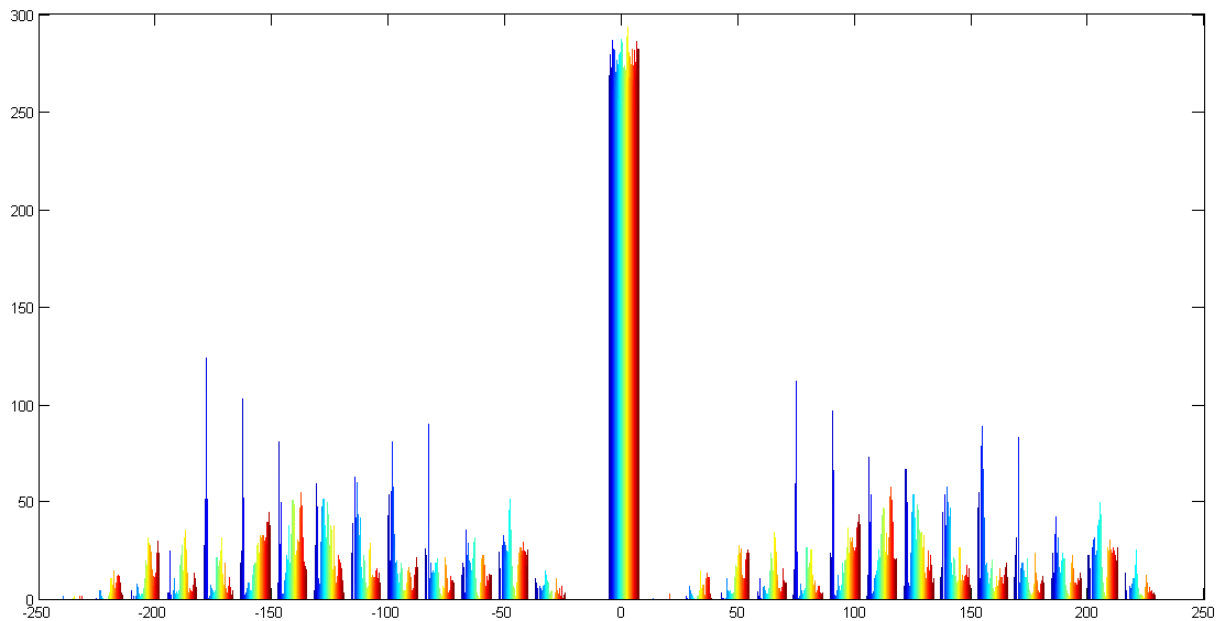
## 7.6 Image inpainting

Μέχρι τώρα, ο αλγόριθμος k-SVD επιδιώκει να διορθώσει τις τιμές ενός σήματος στο οποίο έχει επιδράσει θόρυβος. Υπάρχουν όμως περιπτώσεις που οι τιμές του σήματος δεν μας είναι γνωστές σε όλο το εύρος μελέτης. Προκειμένου να βρούμε τις ελλείπουσες τιμές χρησιμοποιούμε αλγορίθμους που έχουν παρόμοιο στόχο με αυτόν της αριθμητικής παρεμβολής, δηλαδή με βάση τις δεδομένες προσεγγίζουν τις ελλείπουσες τιμές του σήματος. Στην περίπτωση μιας ψηφιακής εικόνας η προσέγγιση της τιμής της έντασης στις θέσεις των pixels που δεν υπάρχει τιμή ονομάζεται **image inpainting**.

Μία κλασική περίπτωση θορύβου που συνδέεται με το πρόβλημα image inpainting είναι ο θόρυβος τύπου salt&pepper. Ένα ποσοστό pixels της εικόνας καταστρέφεται και αντικαθίσταται με τις ακραίες τιμές της έντασης των pixels, δηλαδή για μια grayscale εικόνα των 8 bits τα κατεστραμμένα pixels έχουν τιμή 0 ή 255. Στο Σχήμα 30 φαίνεται η κατανομή του salt&pepper θορύβου που έχει επιδράσει πάνω σε μία εικόνα  $512 \times 512$  pixels της οποίας κάθε pixel αναπαρίσταται με τη βοήθεια 8 bits.

Στην περίπτωση αυτή και βασιζόμενοι στο [15] εφαρμόζουμε τον αλγόριθμο OMP προκειμένου να ανακτήσουμε μία προσέγγιση του αρχικού σήματος. Η διαφορά σε σχέση

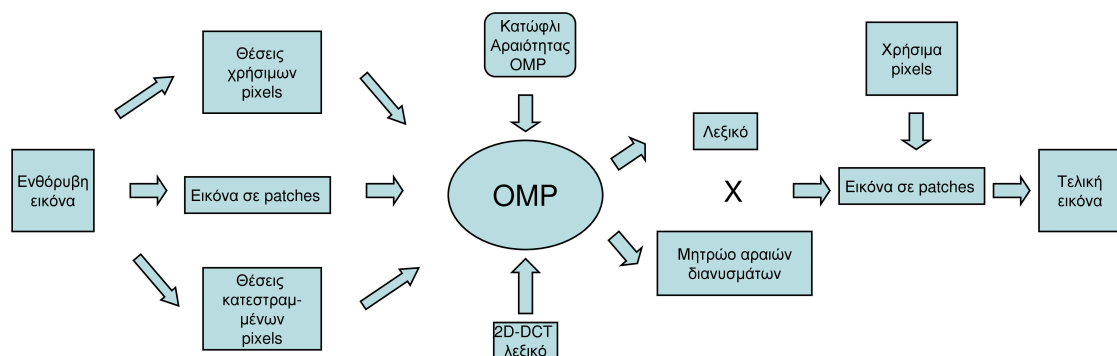




**Σχήμα 30:** Η κατανομή θορύβου salt&pepper πάνω στο σήμα μιας ψηφιακής εικόνας 8 bits

με τις προηγούμενες περιπτώσεις θορύβου είναι ουσιώδης. Στις προηγούμενες ενότητες τα pixels της εικόνας που έχουν επηρεαστεί από θόρυβο είναι άγνωστα. Στην περίπτωση όμως του θορύβου salt&pepper τα pixels τα οποία είναι προϊόν θορύβου είναι γνωστά. Επομένως κατά το image inpainting είναι γνωστό ένας μέρος του σήματος και στόχος είναι ο προσδιορισμός των τιμών του σήματος στις θέσεις όπου η τιμή του σήματος είναι άγνωστη.

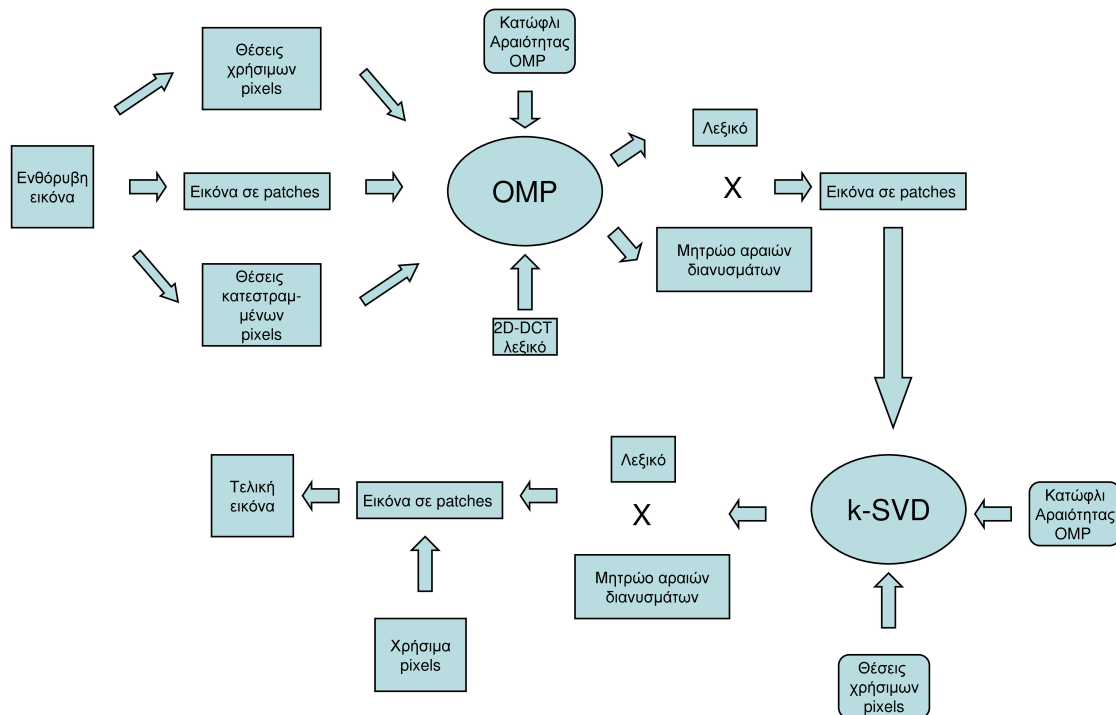
Σχηματικά η διαδικασία που ακολουθείται φαίνεται στο Σχήμα 31. Η εκ των προτέρων γνώση μέρους του σήματος δημιουργεί δύο κατηγορίες θέσεων εικονοστοιχείων: τις θέσεις με γνωστή την τιμή των pixels και τις θέσεις που η τιμή των pixels είναι παντελώς άγνωστη. Η μάσκα αυτή των άγνωστων pixels χρησιμοποιείται ως είσοδος στον αλγόριθμο OMP προκειμένου να προσδιοριστεί με τη βοήθεια αραιών διανυσμάτων η προσέγγιση της αρχικής εικόνας. Ως λεξικό αρχικοποίησης χρησιμοποιείται το 2D-DCT υπερπλήρες λεξικό. Το κατώφλι αραιότητας που χρησιμοποιείται για την επίτευξη βέλτιστων αποτελεσμάτων είναι περίπου 25 συντελεστές σε κάθε αραιό διάνυσμα.



**Σχήμα 31:** Σχηματική αναπαράσταση της εφαρμογής του αλγορίθμου OMP στην περίπτωση του image inpainting

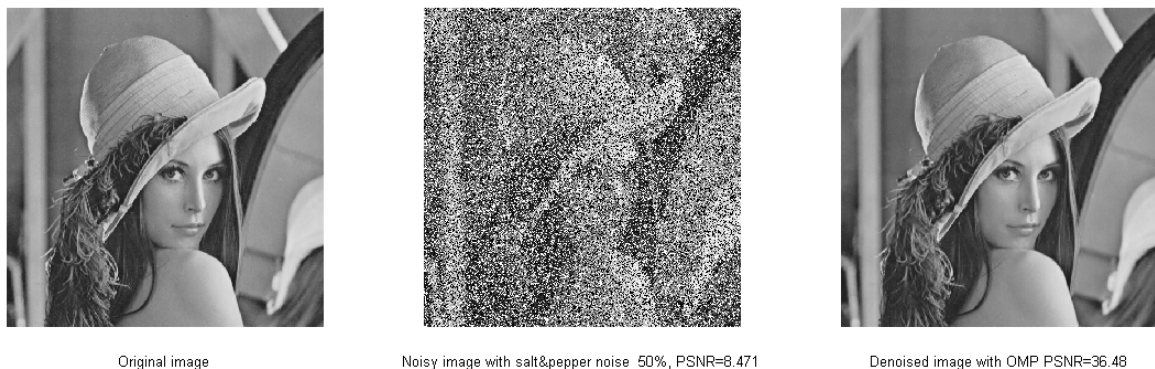
Η εφαρμογή του αλγορίθμου k-SVD στην περίπτωση του image inpainting μπορεί να

γίνει μετά την εφαρμογή του αλγορίθμου OMP. Και σε αυτήν την περίπτωση ως είσοδος στον αλγόριθμο χρησιμοποιείται η μάσκα των άγνωστων pixels. Σχηματικά η διαδικασία που ακολουθείται φαίνεται στο Σχήμα 32. Εντούτοις, η προσθήκη του αλγορίθμου k-SVD στο image inpainting βελτιώνει ελάχιστα το PSNR της προκύπτουσας εικόνας σε σχέση με το PSNR που προκύπτει από την αποκλειστική χρήση του αλγορίθμου OMP. Επιπρόσθετα, η προσθήκη του αλγορίθμου k-SVD καθυστερεί παρά πολύ το συνολικό αλγόριθμο. Στο Σχήμα 34 φαίνεται το αποτέλεσμα του image inpainting με τη βοήθεια της παραπάνω τροποποιημένης έκδοσης του αλγορίθμου. Το PSNR που προκύπτει είναι ελαφρώς μεγαλύτερο σε σχέση με το PSNR του Σχήματος 33 που έχει προκύψει με την αποκλειστική χρήση του αλγορίθμου OMP.



**Σχήμα 32:** Σχηματική αναπαράσταση της εφαρμογής του αλγορίθμου k-SVD στην περίπτωση του image inpainting

Στο Σχήμα 35 φαίνεται το αποτέλεσμα της επίδρασης του αλγορίθμου k-SVD πάνω σε μία εικόνα στην οποία έχουμε επίδραση θορύβου τύπου salt&pepper με ποσοστό απωλεσμένων pixels 80%.



**Σχήμα 33:** Η εφαρμογή του αλγορίθμου OMP σε μία εικόνα με 50% salt&pepper θόρυβο



Original image



Noisy image with salt&pepper noise 50%, PSNR=8.471

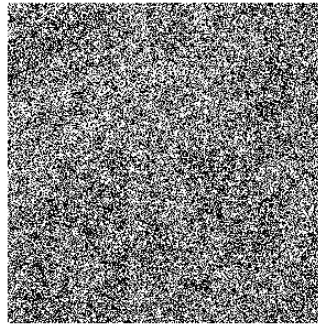


Denoised image with modified K-SVD PSNR=36.58

**Σχήμα 34:** Η εφαρμογή του αλγορίθμου k-SVD σε μία εικόνα με 50% salt&pepper θόρυβο



Original image



Noisy image with salt&pepper noise 80%, PSNR=6.27



Denoised image with modified K-SVD PSNR=27.1

**Σχήμα 35:** Η εφαρμογή του αλγορίθμου k-SVD σε μία εικόνα με 80% salt&pepper θόρυβο

**Παρατήρηση 7.5.** Από τα αποτελέσματα είναι φανερό ότι ο αλγόριθμος OMP δίνει εντυπωσιακά αποτελέσματα στην περίπτωση του *image inpainting*. Το γεγονός αυτό, εκμεταλλευόμεστε προκειμένου να πετύχουμε αποθορυβοποίηση σημάτων στα οποία έχει επιδράσει θόρυβος μικτού τύπου (*mixed type noise*) καθώς και θόρυβος που έχει ως αποτέλεσμα δημιουργία ακραίων τιμών στην εικόνα (*long tails noise*).

## 7.7 Αποθορυβοποίηση σήματος με θόρυβο t-κατανομής

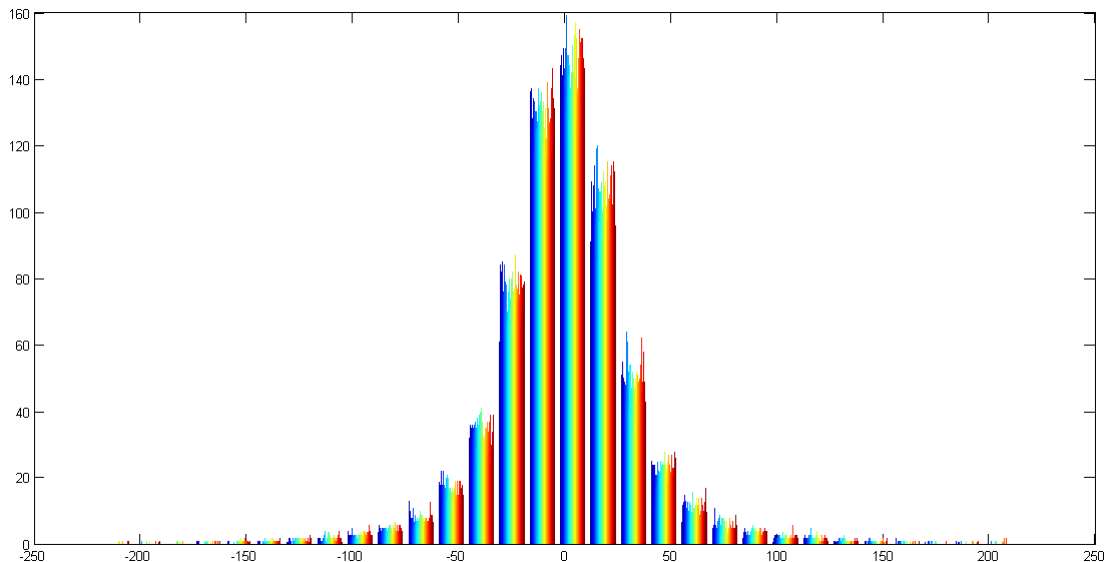
Η κατανομή t είναι μια κατανομή μακρών ουρών. Ο τύπος της πυκνότητας πιθανότητας της κατανομής δίνεται από τη σχέση:

$$f(t) = \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\nu\pi} \Gamma(\frac{\nu}{2})} \left(1 + \frac{t^2}{\nu}\right)^{-\frac{\nu+1}{2}}, \quad (64)$$

όπου  $\nu$  είναι οι βαθμοί ελευθερίας και  $\Gamma$  είναι η συνάρτηση γάμμα.

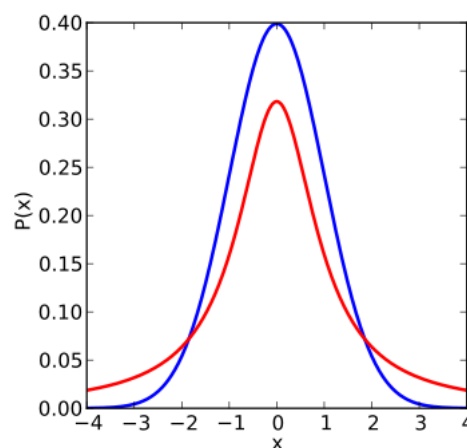
Όταν επιδράσει θόρυβος κατανομής t πάνω σε μία εικόνα υπάρχουν περισσότερα pixels με ακραίες τιμές σε σχέση με μία ενθόρυβη εικόνα στην οποία έχει επιδράσει gaussian θόρυβος. Η κατανομή ενός τέτοιου τύπου θορύβου πάνω σε μία ψηφιακή εικόνα φαίνεται στο Σχήμα 36 ενώ στο Σχήμα 37 φαίνεται η διαφορά ανάμεσα στην κατανομή-t και στη gaussian κατανομή.

Εάν εφαρμόσουμε τον αλγόριθμο k-SVD σε αναλογία με τις προηγούμενες περιπτώσεις θορύβων (gaussian, poisson, uniform) θα προκύψει το αποτέλεσμα του Σχήματος 38. Λόγω της κατανομής μακρών ουρών, ο αλγόριθμος που συμπεριφέρεται βέλτιστα στην περίπτωση της gaussian κατανομής θορύβου, δεν καταφέρνει να απορρίψει τις ακραίες



**Σχήμα 36:** Η κατανομή θορύβου t-κατανομής πάνω στο σήμα μιας ψηφιακής εικόνας 8 bits

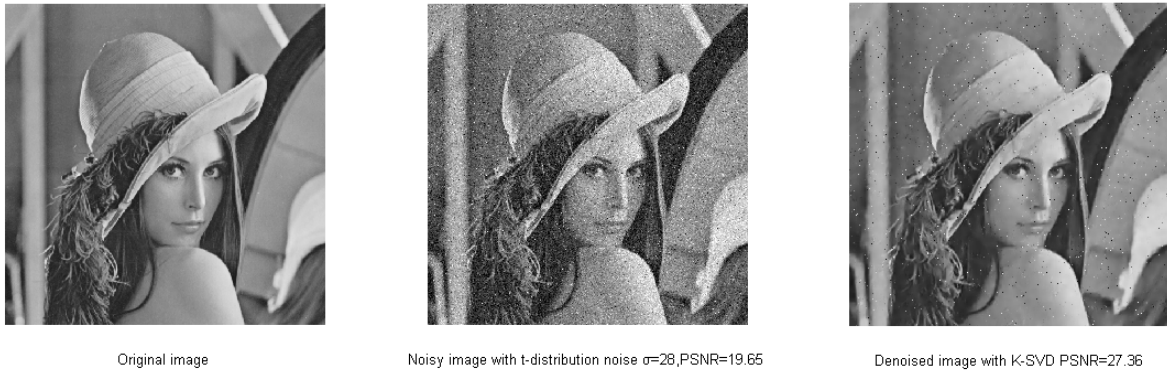
τιμές του θορύβου οι οποίες και παραμένουν στην εικόνα ως outliers.



**Σχήμα 37:** Η γραφική απεικόνιση της t-κατανομής (κόκκινη καμπύλη) σε σχέση με την κανονική κατανομή (μπλέ καμπύλη)

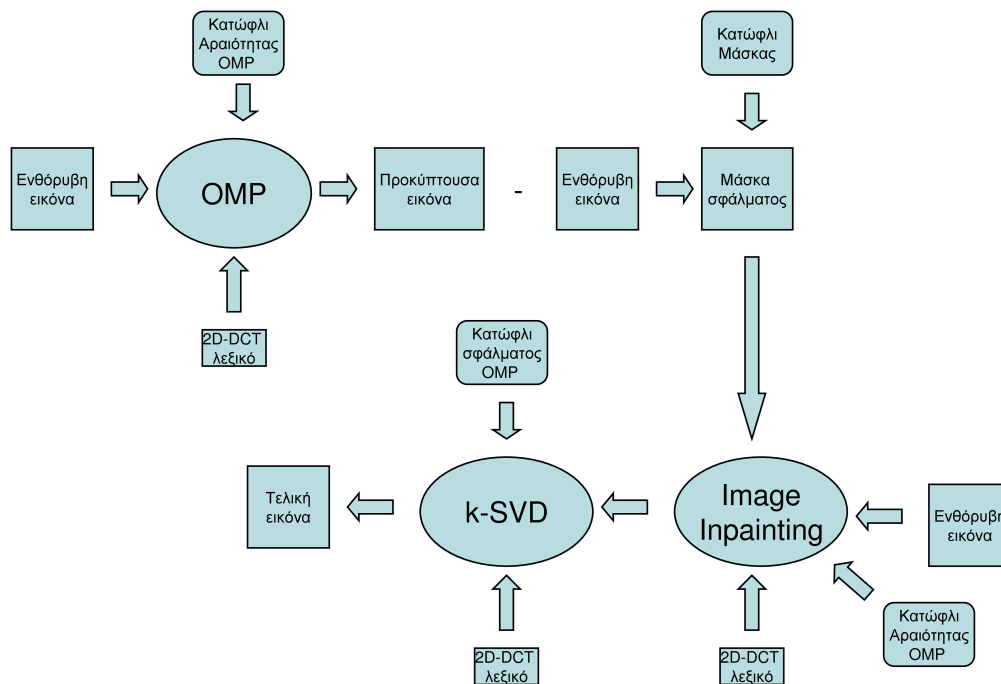
Επομένως στην περίπτωση αυτή δε μας αρκεί η εφαρμογή του κλασικού αλγορίθμου k-SVD για την αποθορυβοποίηση του σήματος που χρησιμοποιήθηκε στη gaussian περίπτωση θορύβου καθώς ο αλγόριθμος δεν μπορεί να απαλείψει τις ακραίες τιμές του θορύβου. Αυτό που εφαρμόζεται είναι ένας συνδυασμός του αλγορίθμου k-SVD των περιπτώσεων ενθόρυβης εικόνας με gaussian θόρυβο καθώς και του image inpainting.

Ο αλγόριθμος στην περίπτωση αυτή λειτουργεί ως εξής: Αρχικά εφαρμόζουμε τον αλγόριθμο OMP πάνω στην ενθόρυβη εικόνα προκειμένου να αποκτήσουμε μία πρώτη προσέγγιση της εικόνας. Το εξαγόμενο αποτέλεσμα αφαιρείται από το αρχικό ενθόρυβο σήμα και προκύπτει μία μάσκα. Στόχος είναι ο προσδιορισμός των θέσεων των εικονοστοιχείων στα οποία οι τιμές του θορύβου είναι ακραίες και επομένως είναι αδύνατη η απαλοιφή τους με τη βοήθεια του κλασικού αλγορίθμου k-SVD που εφαρμόζεται στη gaussian κατανομή θορύβου. Επιλέγοντας ένα κατώφλι, εξάγουμε τη μάσκα με τις συγκεκριμένες θέσεις των εικονοστοιχείων την οποία χρησιμοποιούμε χρησιμοποιούμε ως είσοδο προκειμένου να εφαρμόσουμε τον αλγόριθμο της προηγούμενης ενότητας image inpainting στην αρχική



**Σχήμα 38:** Η εφαρμογή του αλγορίθμου k-SVD σε μία εικόνα με θόρυβο από t-κατανομή

ενθόρυβη εικόνα. Τέλος, στην προκύπτουσα εικόνα μετά το inpainting εφαρμόζουμε τον κλασικό αλγόριθμο k-SVD της gaussian περίπτωσης. Σχηματικά, η προηγούμενη διαδικασία παρουσιάζεται με ένα μπλοκ διάγραμμα στο Σχήμα 39.



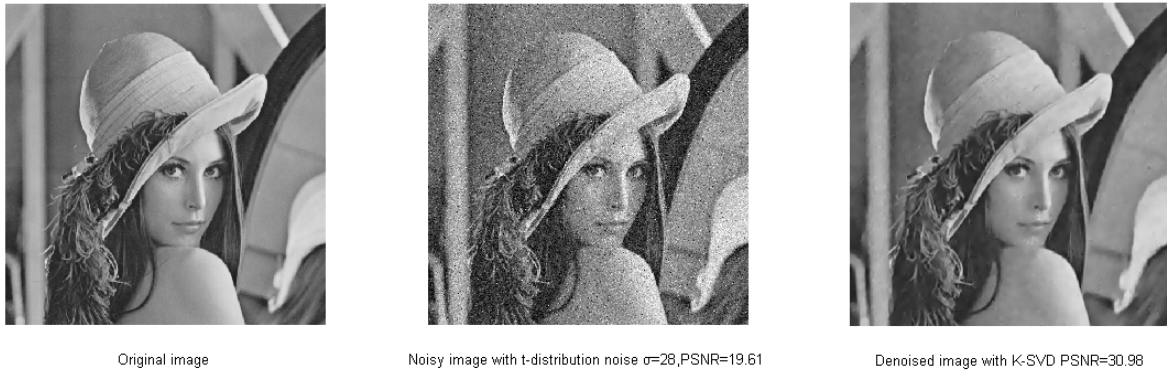
**Σχήμα 39:** Σχηματική αναπαράσταση της εφαρμογής του αλγορίθμου k-SVD στην περίπτωση της t-κατανομής θορύβου

Τα αποτελέσματα αυτής της εφαρμογής του παραπάνω αλγορίθμου φαίνονται στο Σχήμα 40. Οι ακραίες τιμές θορύβου που παρατηρούνται στο Σχήμα 38 έχουν πλέον απορριφθεί και το PSNR είναι πολύ βελτιωμένο.

## 7.8 Αποθορυβοποίηση σήματος με παλμικό θόρυβο (impulse noise)

Ο παλμικός θόρυβος διακρίνεται σε δύο κατηγορίες.

- Στην κατηγορία I ένα ποσοστό των pixels της εικόνας αλλοιώνεται με θόρυβο που παίρνει μόνο τις ακραίες τιμές από ένα διάστημα. Η πυκνότητα πιθανότητας της κατηγορίας αυτού του θορύβου δίνεται από τον τύπο:



**Σχήμα 40: Η εφαρμογή του τροποποιημένου αλγορίθμου k-SVD σε μία εικόνα με θόρυβο από t-κατανομή**

$$p(z) = \begin{cases} p, & \text{αν } z \geq \alpha, \text{ ή } z \leq -\alpha \\ 1 - 2p, & \text{αλλιώς.} \end{cases} \quad (65)$$

όπου  $0 \leq p \leq \frac{1}{2}$  και  $\alpha > 0$ . Με βάση την παραπάνω κατανομή το 200p% περίπου των εικονοστοιχείων αλλοιώνεται με προσθετικό θόρυβο ακραίων τιμών  $\pm\alpha$  ενώ τα υπόλοιπα pixels της εικόνας παραμένουν αναλλοίωτα.

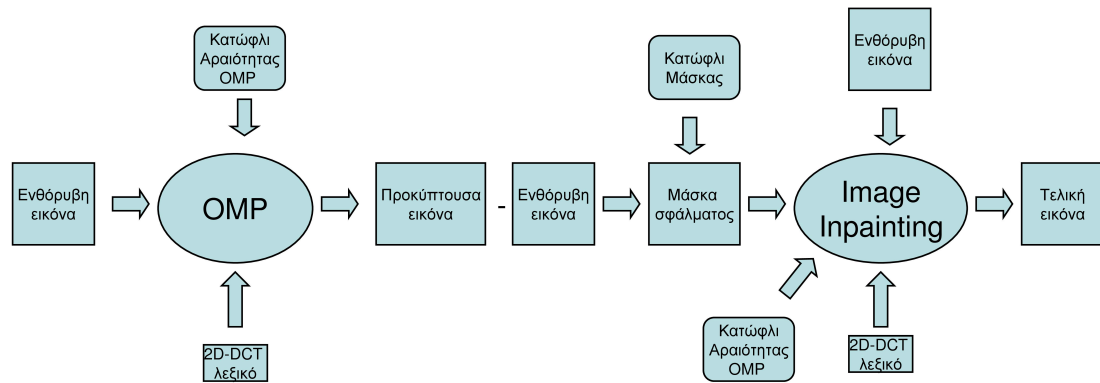
- Στην κατηγορία II ένα ποσοστό των pixels της εικόνας αλλοιώνεται με προσθετικό θόρυβο ομοιόμορφης κατανομής. Η πυκνότητα πιθανότητας της κατηγορίας αυτού του θορύβου δίνεται από τον τύπο:

$$p(z) = \begin{cases} p, & \text{αν } -\alpha \leq z \leq \alpha \\ 1 - 2p, & \text{αλλιώς.} \end{cases} \quad (66)$$

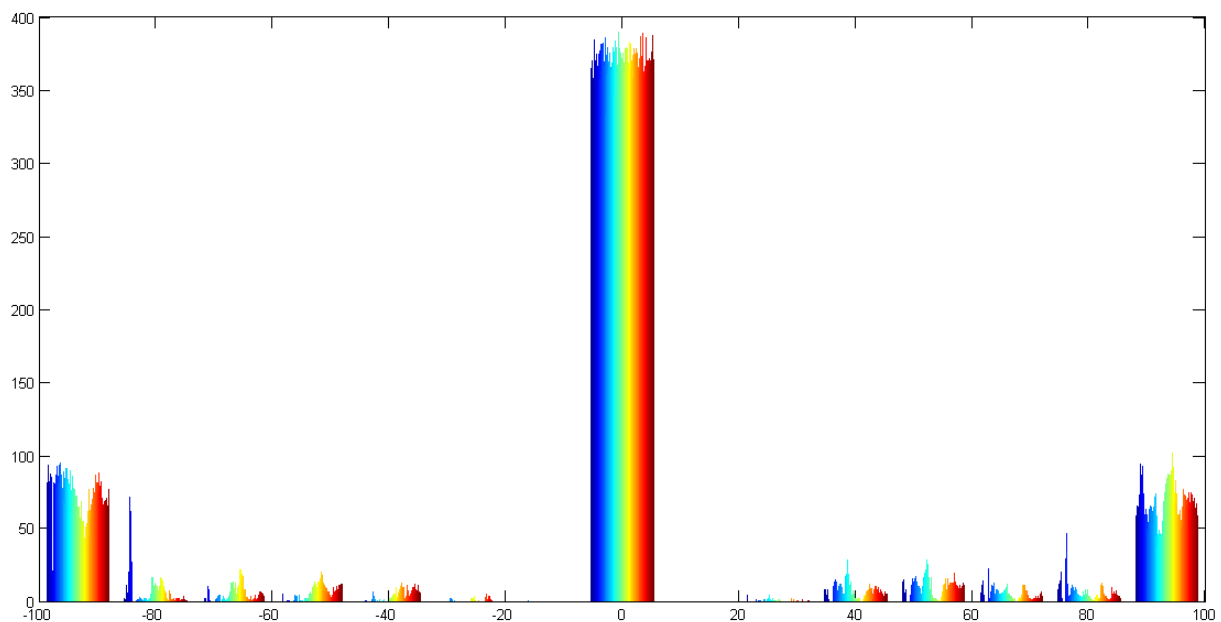
Με βάση την παραπάνω κατανομή το 200p% περίπου των εικονοστοιχείων αλλοιώνεται με προσθετικό θόρυβο ομοιόμορφης κατανομής από το διάστημα  $[-\alpha, \alpha]$  ενώ τα υπόλοιπα pixels της εικόνας παραμένουν αναλλοίωτα.

Στο Σχήμα 42 φαίνεται η κατανομή παλμικού θορύβου κατηγορίας I στο 30% περίπου των pixels μιας εικόνας με προσθετικές τιμές θορύβου  $\pm 100$ . Για την απόρριψη του θορύβου χρησιμοποιούμε τον αλγόριθμο image inpainting που αναπτύχθηκε στην ενότητα 7.6. Στην περίπτωση αυτή, όπως και στην περίπτωση του θορύβου t-κατανομής, δε γνωρίζουμε τις θέσεις των pixels στις οποίες πρέπει να γίνει το inpainting. Για το λόγο αυτό στην αρχική ενθόρυβη εικόνα εφαρμόζουμε τον αλγόριθμο OMP προκειμένου να αποκτήσουμε μία αρχική προσέγγιση της ζητούμενης εικόνας. Στη συνέχεια με ένα κατάλληλο κατώφλι προσδιορίζουμε τη μάσκα των pixels στα οποία θα γίνει το inpainting. Μετά την εφαρμογή του αλγορίθμου image inpainting εξαγάγουμε την τελική εικόνα διατηρώντας τις ίδιες τιμές των pixels στις θέσεις που είναι εκτός μάσκας. Τα εικονοστοιχεία αυτών των θέσεων, εφόσον με την πρώτη προσέγγιση του αλγορίθμου OMP δε θεωρηθούν ως pixels με θόρυβο, καταγράφονται από τον αλγόριθμο ως χρήσιμη πληροφορία και διατηρούνται αναλλοίωτα στην τελική εικόνα.





**Σχήμα 41: Σχηματική αναπαράσταση της εφαρμογής του αλγορίθμου OMP στην περίπτωση του παλμικού θορύβου κατηγορίας I**



**Σχήμα 42: Η κατανομή παλμικού θορύβου κατηγορίας I πάνω στο σήμα μιας ψηφιακής εικόνας 8 bits**

Στο μπλοκ διάγραμμα του Σχήματος 41 παρουσιάζεται η παραπάνω διαδικασία για την αποθρομβοποίηση εικόνων επηρεασμένων από παλμικό θόρυβο κατηγορίας I.

Το αποτέλεσμα του παραπάνω αλγορίθμου σε μία ψηφιακή εικόνα επηρεασμένη από παλμικό θόρυβο κατηγορίας I φαίνεται στο Σχήμα 43.

Στον πίνακα 2 φαίνονται τα συγκριτικά αποτελέσματα του προτεινόμενου αλγορίθμου σε σχέση με τα αποτελέσματα του [3].

Όμοια, στο Σχήμα 44 φαίνεται η κατανομή παλμικού θορύβου κατηγορίας II στο 30% περίπου των pixels μιας εικόνας με προσθετικές τιμές θορύβου στο διάστημα  $[-128, 128]$ . Ο αλγόριθμος που χρησιμοποιείται για την απόρριψη του θορύβου είναι ίδιος με αυτόν του Σχήματος 41 που χρησιμοποιήθηκε για τον παλμικό θόρυβο κατηγορίας I με τη διάφορα ότι το κατώφλι για την εξαγωγή της μάσκας πρέπει να είναι μικρότερο.

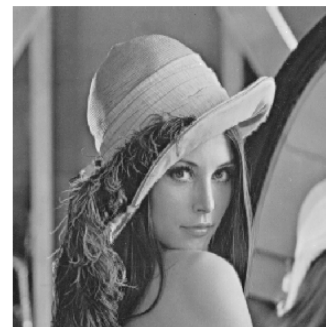
Το αποτέλεσμα του αλγορίθμου σε μία ψηφιακή εικόνα επηρεασμένη από παλμικό θόρυβο κατηγορίας II φαίνεται στο Σχήμα 43.



Original image

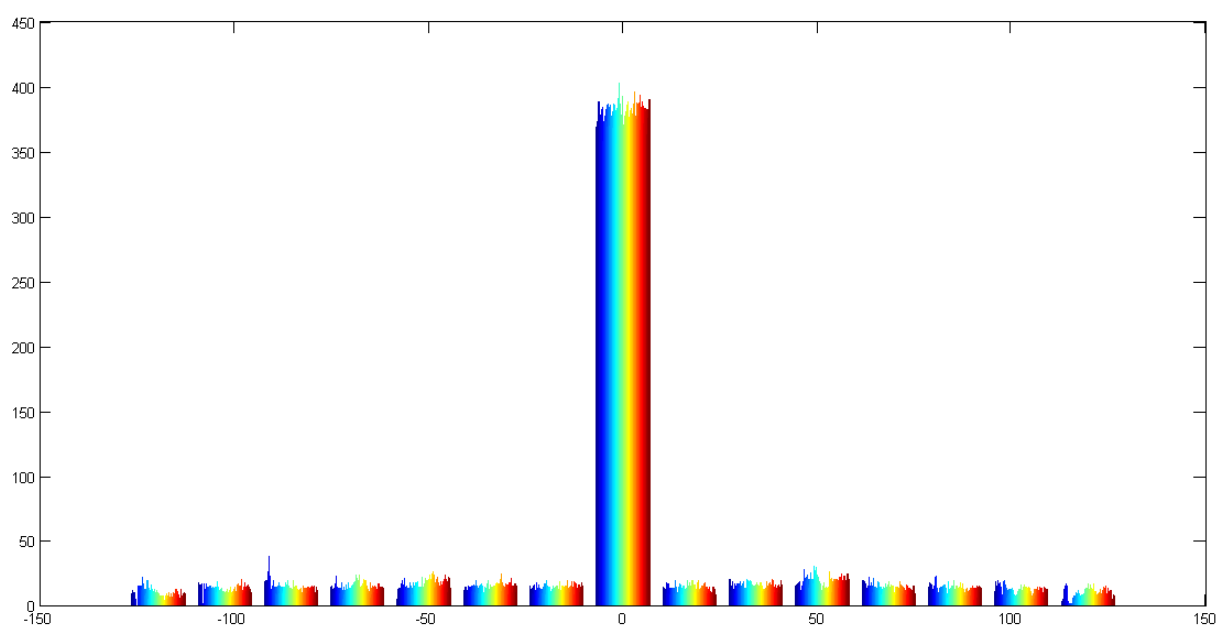


Noisy image with impulse noise type I [-100,100] 30%, PSNR=14.04



Denoised image with OMP PSNR=33.77

**Σχήμα 43: Η εφαρμογή του αλγορίθμου του Σχήματος 41 σε μία εικόνα με παλμικό θόρυβο κατηγορίας I**



**Σχήμα 44: Η κατανομή παλμικού θορύβου κατηγορίας II πάνω στο σήμα μιας ψηφιακής εικόνας 8 bits**



Original image



Noisy image with impulse noise type II [-128,128] 30%, PSNR=16.65



Denoised image with OMP PSNR=30.13

**Σχήμα 45: Η εφαρμογή του αλγορίθμου του Σχήματος 41 σε μία εικόνα με παλμικό θόρυβο κατηγορίας II**



| Image   | Noise | noisy PSNR (dB) | PSNR (dB) με Kernel denoising [3] | PSNR (dB) με τον Αλγόριθμο του Σχήματος 41 |
|---------|-------|-----------------|-----------------------------------|--------------------------------------------|
| Lena    | 20%   | 15.80           | 34.34                             | 35.00                                      |
|         | 30%   | 14.04           | 32.10                             | 33.80                                      |
|         | 40%   | 12.79           | 30.31                             | 33.05                                      |
|         | 50%   | 11.82           | 28.57                             | 31.35                                      |
| Peppers | 20%   | 15.93           | 32.27                             | 30.32                                      |
|         | 30%   | 14.16           | 30.48                             | 30.17                                      |
|         | 40%   | 12.91           | 29.19                             | 29.56                                      |
|         | 50%   | 11.96           | 27.88                             | 28.08                                      |
| Barbara | 20%   | 15.97           | 26.37                             | 31.62                                      |
|         | 30%   | 14.26           | 25.11                             | 30.07                                      |
|         | 40%   | 13.00           | 24.02                             | 29.37                                      |
|         | 50%   | 12.02           | 23.09                             | 26.90                                      |
| Boat    | 20%   | 15.68           | 31.04                             | 30.88                                      |
|         | 30%   | 13.91           | 28.86                             | 29.76                                      |
|         | 40%   | 12.65           | 27.62                             | 28.51                                      |
|         | 50%   | 11.69           | 25.88                             | 26.49                                      |

Πίνακας 2: Συγκριτικά αποτελέσματα του αλγορίθμου του Σχήματος 41 σε σχέση με τον αλγόριθμο Kernel-based για παλμικό θόρυβο κατηγορίας I

**Παρατήρηση 7.6.** Από τα αποτελέσματα για την αποθορυβοποίηση εικόνων επηρεασμένων από παλμικό θόρυβο κατηγορίας I και II είναι φανερό ότι ο αλγόριθμος του Σχήματος 41 επιτυγχάνει καλύτερα αποτελέσματα στην πρώτη περίπτωση όπου ο θόρυβος έχει περισσότερες ακραίες τιμές. Αυτό συμβαίνει γιατί σε αυτή την περίπτωση εκμεταλλευόμαστε την αποτελεσματικότητα αντικατάστασης των επηρεασμένων ενθόρυβων εικονοστοιχείων με τη βοήθεια του αλγορίθμου *image inpainting*.

## 7.9 Αποθορυβοποίηση σήματος με θόρυβο μικτού τύπου (mixed type noise)

Στην περίπτωση που σε ένα σήμα επιδρούν περισσότεροι από έναν τύποι θορύβου, έχουμε θόρυβο μικτού τύπου (mixed type noise). Στην παρούσα εργασία μελετούμε 5 μικτού τύπου κατηγορίες θορύβου ως εξής:

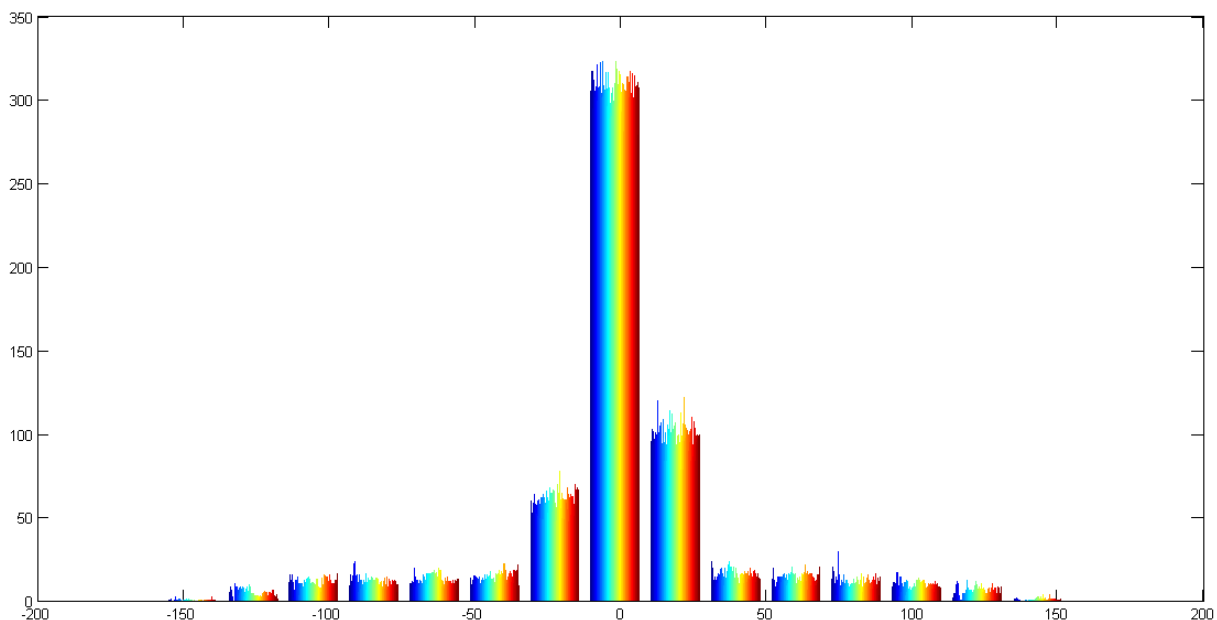
- μικτού τύπου 1: 20% παλμικός θόρυβος τύπου II με  $\alpha = 128$  και gaussian θόρυβος με  $\sigma = 10$ .
- μικτού τύπου 2: 30% παλμικός θόρυβος τύπου II με  $\alpha = 128$  και gaussian θόρυβος με  $\sigma = 20$ .
- μικτού τύπου 3: θόρυβος ομοιόμορφης κατανομής στο διάστημα  $[-10, 10]$  και gaussian θόρυβος με  $\sigma = 10$ .
- μικτού τύπου 4: θόρυβος ομοιόμορφης κατανομής στο διάστημα  $[-10, 10]$  και 10% παλμικός θόρυβος τύπου II με  $\alpha = 128$ .
- μικτού τύπου 5: θόρυβος ομοιόμορφης κατανομής στο διάστημα  $[-10, 10]$  και 10% παλμικός θόρυβος τύπου II με  $\alpha = 128$  και gaussian θόρυβος με  $\sigma = 10$ .

Σε όλες τις παραπάνω περιπτώσεις ο αλγόριθμος k-SVD χρησιμοποιείται συνδυαστικά προκειμένου να απαλείψει τις ακραίες τιμές του θορύβου αλλά και αυτές που διασπείρονται με ομοιόμορφη ή κανονική κατανομή πάνω στην εικόνα.

#### • Περίπτωση 1

Στην περίπτωση του θορύβου μικτού τύπου 1, η κατανομή του πάνω σε μία ψηφιακή εικόνα φαίνεται στο Σχήμα 46. Ο αλγόριθμος που χρησιμοποιείται είναι όμοιος με τον αλγόριθμο που χρησιμοποιήθηκε για την εξάλειψη του θορύβου t-κατανομής και του παλμικού θορύβου δηλαδή συνδυασμός του αλγορίθμου για το image inpainting και του αλγορίθμου k-SVD. Η διαδικασία της εξαγωγής της μάσκας των ενθόρυβων εικονοστοιχείων, του image inpainting και εν συνέχεια της εφαρμογής του αλγορίθμου k-SVD μπορεί να επαναληφθεί περισσότερες από μία φορές για επίτευξη μεγαλύτερου PSNR.

Ως κατώφλι σφάλματος στο sparse coding step του αλγορίθμου k-SVD χρησιμοποιείται η διασπορά που προκύπτει από τη διαφορά ανάμεσα στην ενθόρυβη εικόνα και στην προσέγγισή της με την εφαρμογή του αλγορίθμου OMP. Το μπλοκ διάγραμμα του παραπάνω αλγορίθμου φαίνεται στο Σχήμα 47.



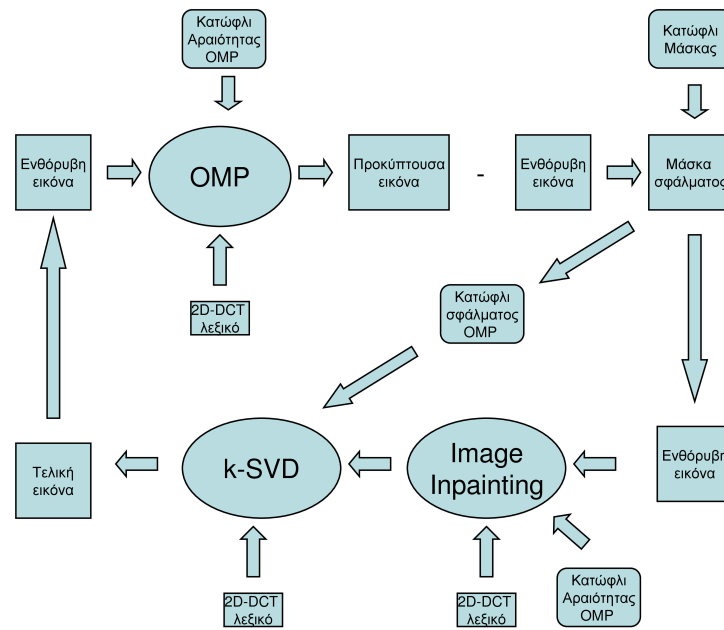
**Σχήμα 46:** Η κατανομή του θορύβου μικτού τύπου 1 πάνω στο σήμα μιας ψηφιακής εικόνας 8 bits

Η εφαρμογή του αλγορίθμου του Σχήματος 47 σε μια ψηφιακή εικόνα, που έχει επηρεαστεί από θόρυβο μικτού τύπου 1, φαίνεται στο Σχήμα 48.

#### • Περίπτωση 2

Στην περίπτωση του θορύβου μικτού τύπου 2, η κατανομή του πάνω σε μία ψηφιακή εικόνα φαίνεται στο Σχήμα 49. Η περίπτωση αυτού του θορύβου είναι ίδια με την προηγούμενη με τη διαφορά ότι ο παλμικός θόρυβος επηρεάζει περισσότερα εικονοστοιχεία ενώ ο gaussian θόρυβος έχει μεγαλύτερη διασπορά. Ο αλγόριθμος που χρησιμοποιείται είναι ίδιος με αυτόν της προηγούμενης περίπτωσης με τη διαφορά ότι το κατώφλι που χρησιμοποιείται για την εξαγωγή των ενθόρυβων pixels πρέπει να έχει πιο μεγάλη τιμή.

Η εφαρμογή του αλγορίθμου του Σχήματος 47 σε μια ψηφιακή εικόνα επηρεασμένη από θόρυβο μικτού τύπου 2 φαίνεται στο Σχήμα 50.



**Σχήμα 47: Σχηματική αναπαράσταση της εφαρμογής του αλγορίθμου του Σχήματος 47 στην περίπτωση του θορύβου μικτού τύπου 1.**



Original image

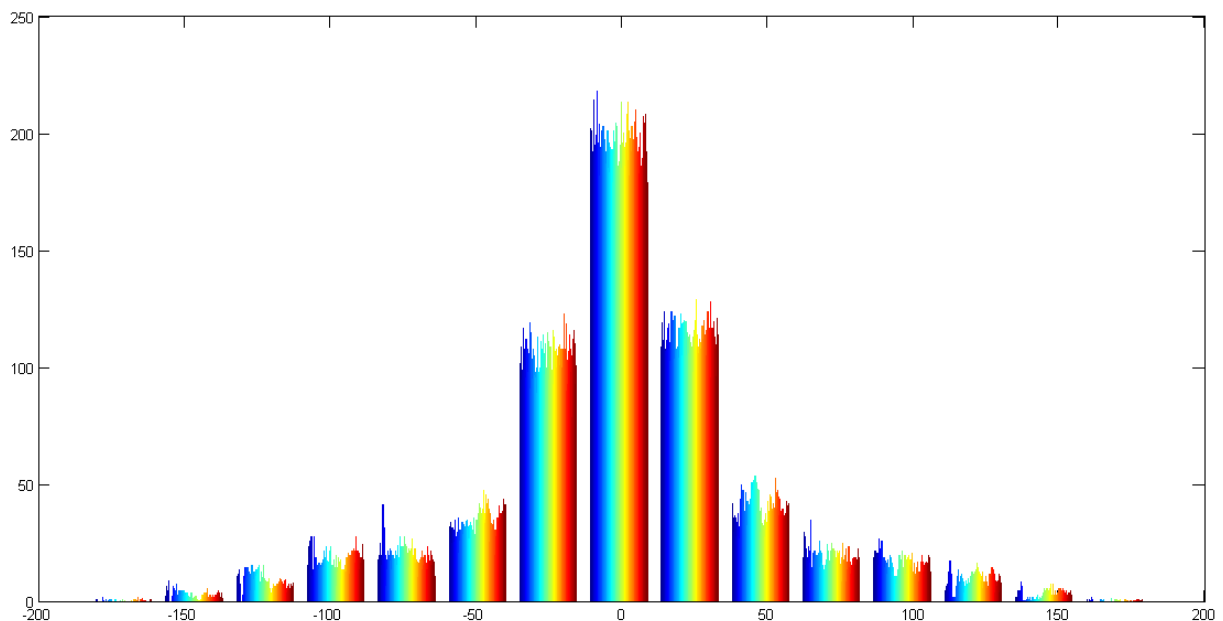


Noisy image with mixed noise type I, PSNR=18.02

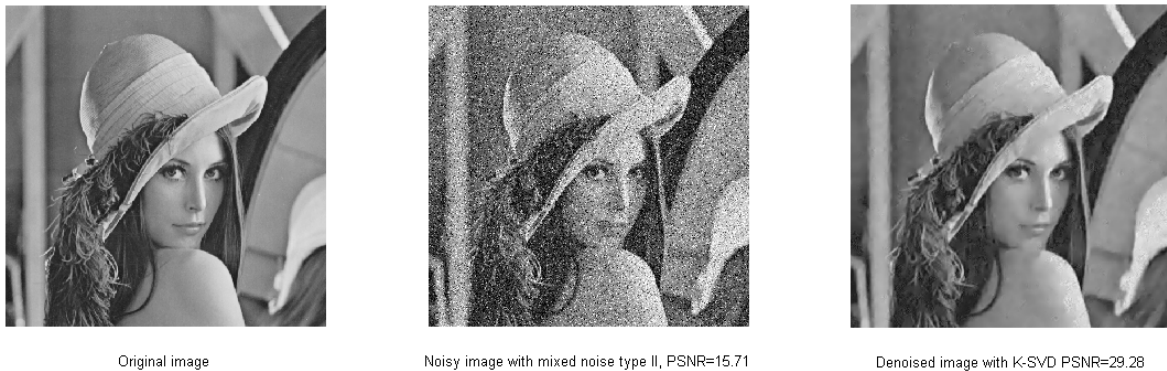


Denoised image with K-SVD PSNR=32.79

**Σχήμα 48: Η εφαρμογή του αλγορίθμου του Σχήματος 47 σε μία εικόνα με θόρυβο μικτού τύπου 1**



**Σχήμα 49: Η κατανομή του θορύβου μικτού τύπου 2 πάνω στο σήμα μιας ψηφιακής εικόνας 8 bits**

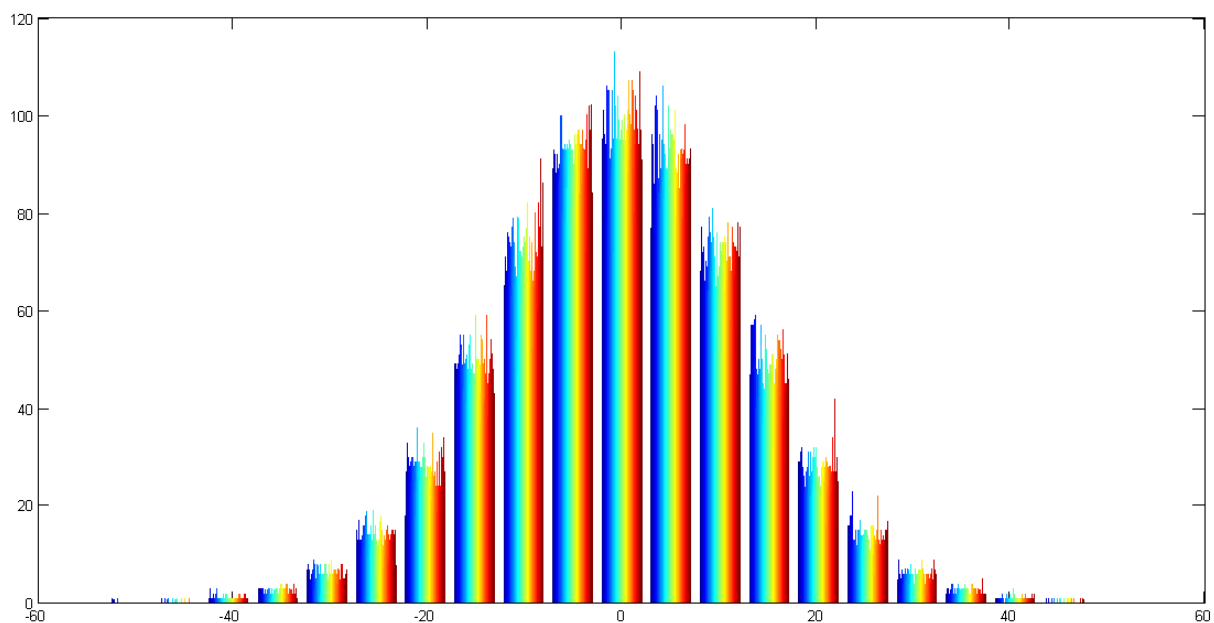


**Σχήμα 50:** Η εφαρμογή του αλγορίθμου του Σχήματος 47 σε μία εικόνα με θόρυβο μικτού τύπου 2

### • Περίπτωση 3

Στην περίπτωση του θορύβου μικτού τύπου 3, η κατανομή του πάνω σε μία ψηφιακή εικόνα φαίνεται στο Σχήμα 51. Ο αλγόριθμος που χρησιμοποιείται για την απόρριψη του συγκεκριμένου τύπου θορύβου είναι κλασικός αλγόριθμος k-SVD του Σχήματος 21 που χρησιμοποιήθηκε τόσο στην περίπτωση του gaussian θορύβου όσο και στην περίπτωση του θορύβου ομοιόμορφης κατανομής. Η διασπορά του θορύβου που χρησιμοποιείται στο κατώφλι σφάλματος OMP, στο sparse coding step του αλγορίθμου υπολογίζεται από τη μίξη των διασπορών των θορύβων γκαουσιανής και ομοιόμορφης κατανομής.

Η εφαρμογή του αλγορίθμου k-SVD σε μια ψηφιακή εικόνα επηρεασμένη από θόρυβο μικτού τύπου 3 φαίνεται στο Σχήμα 52.



**Σχήμα 51:** Η κατανομή του θορύβου μικτού τύπου 3 πάνω στο σήμα μιας ψηφιακής εικόνας 8 bits

### • Περίπτωση 4

Στην περίπτωση του θορύβου μικτού τύπου 4, η κατανομή του πάνω σε μία ψηφιακή εικόνα φαίνεται στο Σχήμα 53. Ο αλγόριθμος που χρησιμοποιείται για την απόρριψη του συγκεκριμένου θορύβου είναι όμοιος με αυτόν των περιπτώσεων 1 και 2 όπως περιγράφεται στο Σχήμα 47 με κατώφλι μάσκας προσαρμοσμένο στο συγκεκριμένο θόρυβο.



Original image

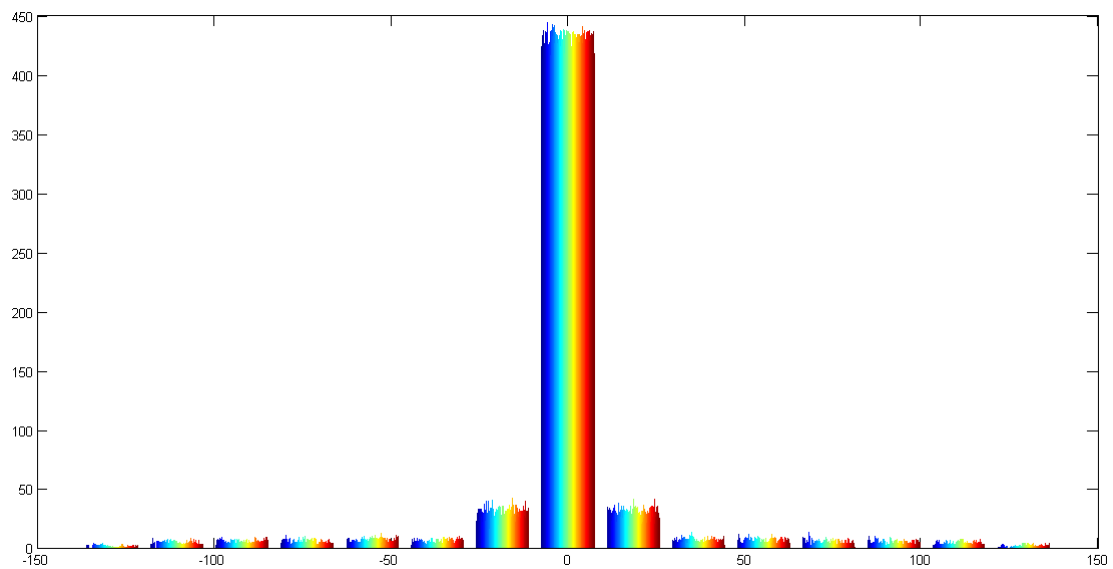


Noisy image with mixed noise type III, PSNR=26.8



Denoised image with K-SVD PSNR=34.78

**Σχήμα 52: Η εφαρμογή του αλγορίθμου k-SVD σε μία εικόνα με θόρυβο μικτού τύπου 3**



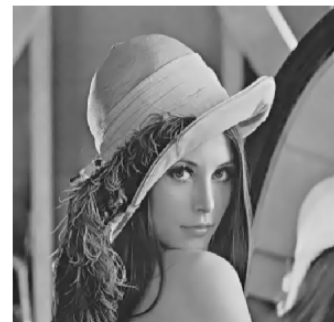
**Σχήμα 53: Η κατανομή του θορύβου μικτού τύπου 4 πάνω στο σήμα μιας ψηφιακής εικόνας 8 bits**



Original image

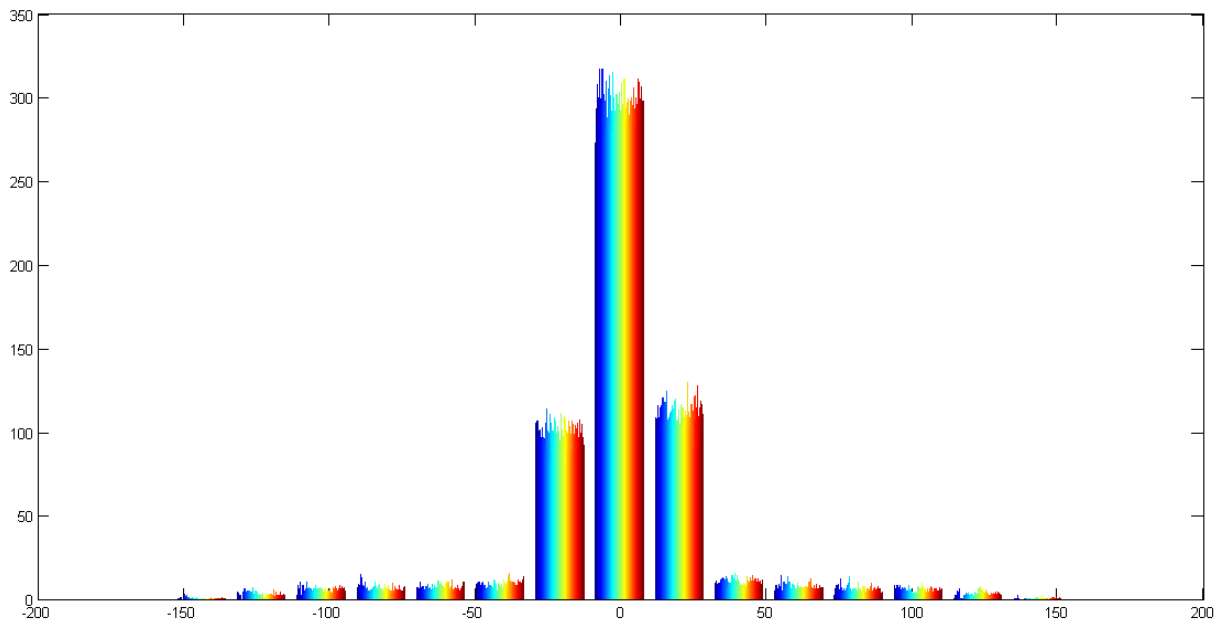


Noisy image with mixed noise type IV, PSNR=21.08



Denoised image with K-SVD PSNR=34.37

**Σχήμα 54: Η εφαρμογή του αλγορίθμου του Σχήματος 47 σε μία εικόνα με θόρυβο μικτού τύπου 4**



**Σχήμα 55:** Η κατανομή του θορύβου μικτού τύπου 5 πάνω στο σήμα μιας ψηφιακής εικόνας 8 bits

Η εφαρμογή του αλγορίθμου του σε μια ψηφιακή εικόνα επηρεασμένη από θόρυβο μικτού τύπου 4 φαίνεται στο Σχήμα 54.

#### • Περίπτωση 5

Στην περίπτωση του θορύβου μικτού τύπου 5, η κατανομή του πάνω σε μία ψηφιακή εικόνα φαίνεται στο Σχήμα 55. Και σε αυτή την περίπτωση εφαρμόζουμε με κατάλληλη επιλογή κατωφλίου μάσκας τον αλγόριθμο που περιγράφεται στο μπλοκ διάγραμμα του Σχήματος 47.

Η εφαρμογή του αλγορίθμου σε μια ψηφιακή εικόνα επηρεασμένη από θόρυβο μικτού τύπου 5 φαίνεται στο Σχήμα 56.



Original image



Noisy image with mixed noise type V, PSNR=20.37



Denoised image with k-SVD PSNR=32.83

**Σχήμα 56:** Η εφαρμογή του αλγορίθμου του Σχήματος 47 σε μία εικόνα με θόρυβο μικτού τύπου 5

**Παρατήρηση 7.7.** Αυτό που αξίζει να σημειωθεί είναι ότι στην παραπάνω περίπτωση ο αλγόριθμος δεν χρειάζεται να "γνωρίζει" εκ των προτέρων τις στατιστικές ιδιότητες του θορύβου, παρά μόνο το γεγονός ότι πρόκειται για θόρυβο μικτού τύπου. Η προσεγγιστική διασπορά θορύβου προκύπτει θεωρώντας ως πραγματικό σήμα το σήμα που προκύπτει με την εφαρμογή του αλγορίθμου OMP πάνω στο αρχικό ενθόρυβο σήμα της εικόνας.

| Image       | Noise   | noisy PSNR (dB) | PSNR (dB) με Kernel denoising [3] | PSNR (dB) με BM3D [3] | PSNR (dB) με τους αλγορίθμους των Σχημάτων 47 και 21 |
|-------------|---------|-----------------|-----------------------------------|-----------------------|------------------------------------------------------|
| <b>Lena</b> | mixed 1 | 18.02           | 32.27                             | 30.65                 | 32.98                                                |
|             | mixed 2 | 15.70           | 29.20                             | 29.10                 | 29.05                                                |
|             | mixed 3 | 26.80           | 33.42                             | 34.10                 | 34.80                                                |
|             | mixed 4 | 21.09           | 34.48                             | 32.14                 | 34.35                                                |
|             | mixed 5 | 20.33           | 32.74                             | 31.77                 | 32.90                                                |
| <b>Boat</b> | mixed 1 | 17.93           | 30.27                             | 29.21                 | 29.49                                                |
|             | mixed 2 | 15.67           | 26.90                             | 27.60                 | 27.13                                                |
|             | mixed 3 | 26.84           | 31.56                             | 32.57                 | 32.84                                                |
|             | mixed 4 | 20.99           | 32.28                             | 30.50                 | 30.03                                                |
|             | mixed 5 | 20.34           | 30.70                             | 30.10                 | 29.15                                                |

**Πίνακας 3:** Συγκριτικά αποτελέσματα των αλγορίθμων των Σχημάτων 47 και 21 σε σχέση με τους αλγορίθμους Kernel-based Denoising και BM3D στην περίπτωση θορύβου μικτού τύπου

Τα αποτελέσματα των παραπάνω αλγορίθμων συγκρινόμενα με τα αποτελέσματα στο [3] φαίνονται στον Πίνακα 3. Από τα αποτελέσματα είναι φανερό ότι οι προτεινόμενοι αλγόριθμοι είναι ανταγωνίσιμοι και εξάγουν αρκετά ικανοποιητικά αποτελέσματα.



## 8. ΣΥΜΠΕΡΑΣΜΑΤΑ

Από την ανάλυση που αναπτύχθηκε στο προηγούμενο κεφάλαιο, είναι φανερό ότι ο αλγόριθμος k-SVD με τη βοήθεια του αλγορίθμου OMP και κάτω από την κατάλληλη επιλογή παραμέτρων μπορεί να αποδώσει ικανοποιητικά αποτελέσματα στην αποθορυβοποίηση ψηφιακών εικόνων.

Προκειμένου να επιτύχουμε καλά αποτελέσματα, εκμεταλλευόμαστε κυρίως δύο στοιχεία:

- Τα εξαιρετικά αποτελέσματα του αλγορίθμου k-SVD στην αποθορυβοποίηση σήματος στο οποίο έχει επιδράσει προσθετικά gaussian θόρυβος.
- Τη "δύναμη" του αλγορίθμου OMP στο image inpainting.

Υπό αυτό το πρίσμα και κάτω από κατάλληλη επιλογή παραμέτρων τα αποτελέσματα της αποθορυβοποίησης εικόνων ειδικότερα στις περιπτώσεις του παλμικού θορύβου και του μικτού τύπου θορύβου είναι αρκετά ικανοποιητικά. Αυτό που αξίζει να σημειωθεί είναι ότι οι αλγόριθμοι που αναπτύχθηκαν μπορούν να εφαρμοστούν χωρίς να είναι εκ προτέρων γνωστές οι παράμετροι της κατανομής του θορύβου παρά μόνο το είδος της στατιστικής κατανομής του θορύβου.

Διάφορες ερευνητικές ιδέες μπορούν να εξαχθούν προκειμένου να μελετηθεί και να βελτιστοποιηθεί η αποτελεσματικότητα του αλγορίθμου k-SVD στην αποθορυβοποίηση σήματος όπως για παράδειγμα:

- Η εφαρμογή ενός διαφορετικού αλγορίθμου από τον OMP, τόσο στο sparse coding step του αλγορίθμου k-SVD όσο και στο ανεξάρτητο κομμάτι που απαιτείται αναπαράσταση σήματος σε αραιή μορφή.
- Η κατασκευή αλγορίθμων αποθορυβοποίησης σήματος με τη βοήθεια του αλγορίθμου k-SVD, που δε θα απαιτούν την εκ των προτέρων γνώση των στατιστικών ιδιοτήτων του θορύβου του σήματος.
- Η χρήση διαφορετικού τύπου λεξικού από το 2-D DCT υπερπλήρες λεξικό προκειμένου να βελτιωθεί η αποτελεσματικότητα των αλγορίθμων με την επίτευξη μεγαλύτερων τιμών PSNR.

Σε γενικότερη βάση, η εκμάθηση λεξικού αποτελεί μία σχετικά καινούρια τεχνική στον τομέα της μηχανικής μάθησης που αφήνει αρκετές προοπτικές βελτίωσης στη εξαγωγή της χρήσιμης πληροφορίας από ένα μεγάλο πλήθος δεδομένων. Η όλη ιδέα της μεθόδου διανοίγει ορίζοντες για φιλοσοφικές προεκτάσεις περί της μινιμαλιστικής συμπεριφοράς της φύσης και επαληθεύει, ακόμη και στο σύγχρονο κόσμο των "big data", τη θεώρηση του Αριστοτέλη ότι "η φύση συμπεριφέρεται με τον πιο απλό πιθανό τρόπο".



**ΠΙΝΑΚΑΣ ΟΡΟΛΟΓΙΑΣ**

|                        |                          |
|------------------------|--------------------------|
| αραιότητα              | sparsity                 |
| αποθορυβοποίηση        | denoising                |
| μηχανική μάθηση        | machine learning         |
| εκμάθηση λεξικού       | dictionary learning      |
| αντίστροφο πρόβλημα    | inverse problem          |
| στάθμη                 | norm                     |
| κυρτή βελτιστοποίηση   | convex optimization      |
| ψευδοαντίστροφο μητρώο | pseudoinverse matrix     |
| ενδογενής διάσταση     | intrinsic dimensionality |
| μείωση διάστασης       | dimensional reduction    |
| ιδιάζουσες τιμές       | singular values          |
| υπερπλήρες λεξικό      | overcomplete dictionary  |
| πλαίσιο                | frame                    |
| άτομο                  | atom                     |

## ΣΥΝΤΜΗΣΕΙΣ, ΑΡΚΤΙΚΟΛΕΞΑ ΚΑΙ ΑΚΡΩΝΥΜΙΑ

|       |                                                 |
|-------|-------------------------------------------------|
| CS    | Compressed Sensing                              |
| DCT   | Discrete Cosine Transform                       |
| DL    | Dictionary Learning                             |
| LASSO | Least Absolute Shrinkage and Selection Operator |
| LSE   | Least Squares Estimator                         |
| MSE   | Mean Squared Error                              |
| NP    | Non-deterministic polynomial time               |
| OMP   | Orthogonal Matching Pursuit                     |
| PSNR  | Peak Signal to Noise Ratio                      |
| RIP   | Restricted Isometry Property                    |
| SVD   | Singular Value Decomposition                    |

## ΑΝΑΦΟΡΕΣ

- [1] Aharon M., Elad M., Bruckstein A., *k-SVD: An Algorithm for Designing Overcomplete Dictionaries for Sparse Representation*, IEEE Transactions on Signal Processing **54** (11):4311–4322, 2006.
- [2] Baker K., *Singular Value Decomposition Tutorial*, 2013, <https://www.ling.ohio-state.edu/>.
- [3] Bouboulis P., Slavakis K., Theodoridis S.: *Adaptive Kernel-based Image Denoising employing Semi-Parametric Regularization*, IEEE Transactions on Signal Processing **19**(6):1465–1479, 2010.
- [4] Boyd S., Vandenberghe L., *Convex Optimization*, Cambridge University Press 2004.
- [5] Candes E., *The restricted isometry property and its implications for compressed sensing*, Comptes Rendus Mathematique, **346**(9), 589–592, 2008.
- [6] Chen G., Needell D., *Compressed Sensing and Dictionary Learning*, Proceedings of Symposia in Applied Mathematics, **73**, 2016.
- [7] Chen S., Donoho D., Saunders M., *Atomic decomposition by basis pursuit*, SIAM Rev., **43**(1), 129–159, 2001.
- [8] Davenport M., Du arte M., Eldar Y., Kutyniok G., *Introduction to Compressed Sensing Compressed Sensing: Theory and Applications*, Cambridge University Press, 2012.
- [9] Elad M., *Sparse and Redundant Representations*, Springer, 2010.
- [10] Elad M., Aharon M., *Image Denoising Via Sparse and Redundant Representations Over Learned Dictionaries*, IEEE Transactions on Image Processing, **15**(12), 3736–3745, 2006.
- [11] Gallier J., *Fundamentals of Linear Algebra and Optimization*, Course Notes, 2013, <http://www.cis.upenn.edu/>.
- [12] Keller J.B., *Inverse Problems*, J. Math. Anal. Appl. **83**(2), 107–118, 1976.
- [13] Leskovec J., Rajaraman A., Ullman J., *Mining of Massive Datasets*, Cambridge University Press, 2014.
- [14] Mairal J., Elad M., Sapiro G., *Sparse learned representations for image restoration*, IASC Yokohama, 2008.
- [15] Mairal J., Elad M., Sapiro G., *Sparse representation for Color Image Restoration*, IEEE Transactions on Image Processing, **17**(1), 53–69, 2008.
- [16] Pati Y., Rezailifar R., Krishnaprasad P., *Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition Signals, Systems and Computers*, Conference Record of The Twenty-Seventh Asilomar Conference on Pacific Grove, CA, **1**, 40–44, 1993.
- [17] Schnass K., *A personal introduction to theoretical dictionary learning*, Internationale Mathematische Nachrichten Nr. 228 , 5–15, 2015.
- [18] Shen Y., Han B., Braverman E., *Image inpainting using directional tensor product complex tight framelets*, Mathematics Subject Classification, 2010.
- [19] Theodoridis S., *Machine Learning: A Bayesian and Optimization Perspective*, Academic Press, 2015.
- [20] Βραχάτης Μ.Ν., *Αριθμητική Ανάλυση*, Ελληνικά Γράμματα, 2002.