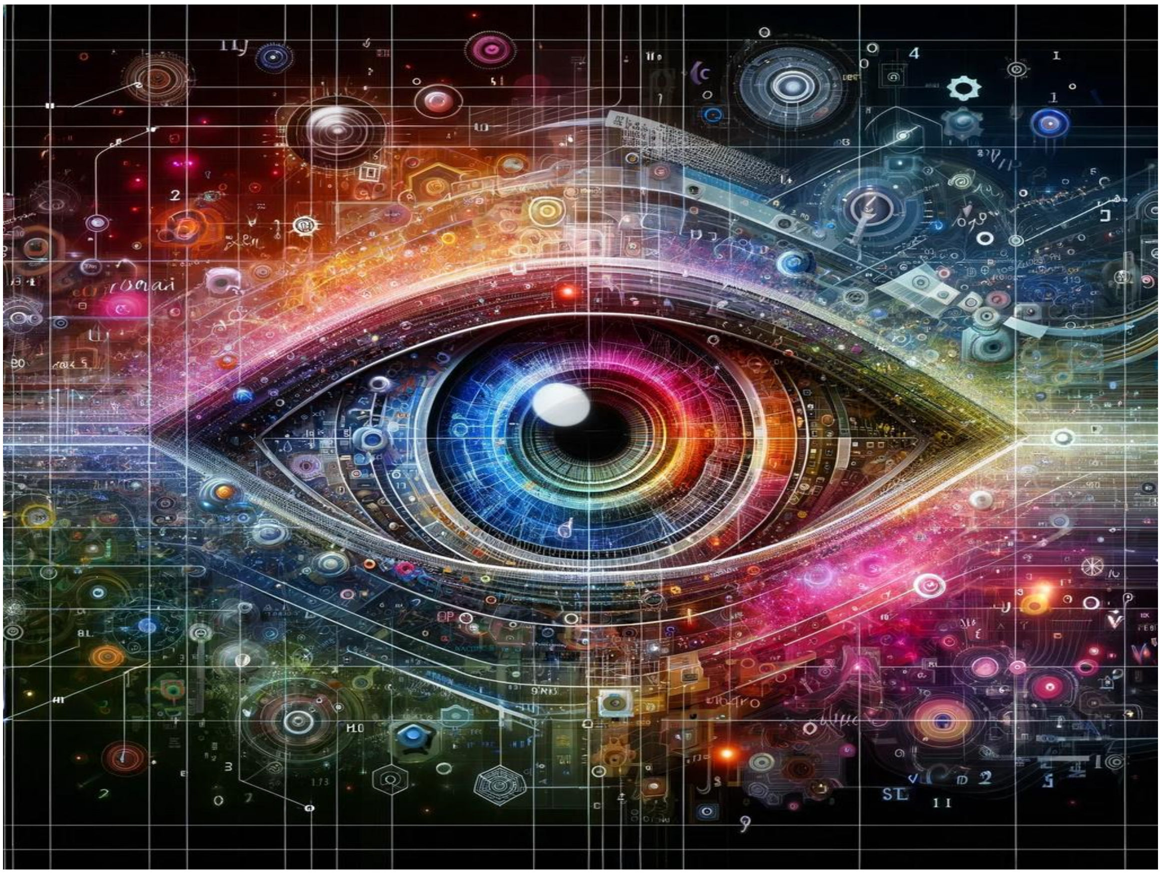
ΠΑΝΕΠΙΣΤΗΜΙΟ ΔΥΤΙΚΗΣ ΑΤΤΙΚΗΣ
ΣΧΟΛΗ ΜΗΧΑΝΙΚΩΝ
ΤΜΗΜΑ ΗΛΕΚΤΡΟΛΟΓΩΝ & ΗΛΕΚΤΡΟΝΙΚΩΝ ΜΗΧΑΝΙΚΩΝ

Διπλωματική Εργασία

Τεχνικές Υπολογιστικής Όρασης για την Εύρεση Αντιπροσώπων από ένα Σύνολο Δεδομένων



Φοιτητής: Λαμπρόπουλος Σταύρος
ΑΜ: 483212017028

Επιβλέπων Καθηγητής
Δρ. Ζώης Ηλίας
Αναπληρωτής Καθηγητής

ΑΘΗΝΑ-ΑΙΓΑΛΕΩ, ΣΕΠΤΕΜΒΡΙΟΣ 2024

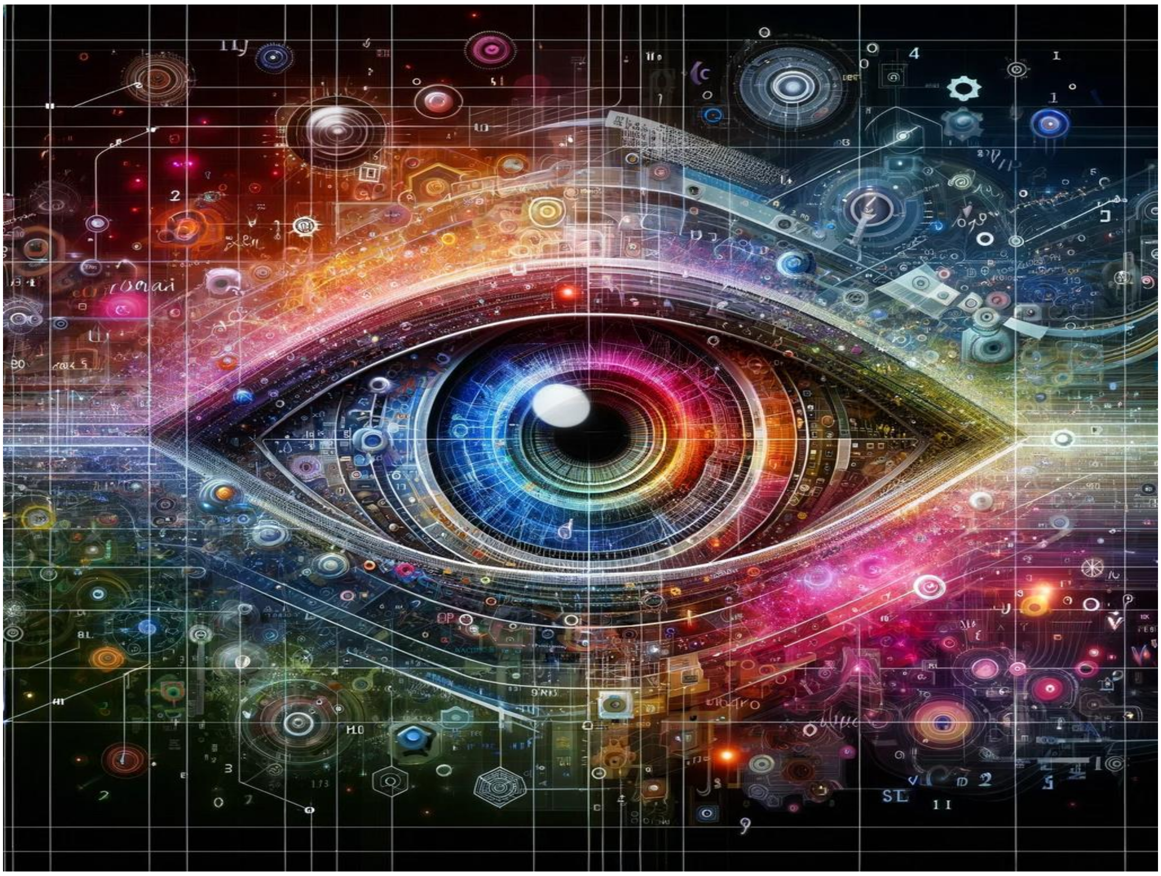
UNIVERSITY OF WEST ATTICA

FACULTY OF ENGINEERING

DEPARTMENT OF ELECTRICAL & ELECTRONICS ENGINEERING

Diploma Thesis

Computer Vision Algorithms for Representatives Selection of Big Data Set



Student: Lampropoulos Stavros

Registration Number: 483212017028

Supervisor

Dr. Elias Zois

Associate Professor

ATHENS-EGALEO, SEPTEMBER 2024

Η Διπλωματική Εργασία έγινε αποδεκτή και βαθμολογήθηκε από την εξής τριμελή επιτροπή:

Ζωής Ηλίας, Αναπληρωτής Καθηγητής	Ευάγγελος Ζέρβας, Καθηγητής	Δημήτριος Καλύβας, Καθηγητής
ILIAS ZOIS Digitally signed by ILIAS ZOIS Date: 2024.10.08 15:03:09 +03'00' (Υπογραφή)	EVANGELOS ZERVAS Digitally signed by EVANGELOS ZERVAS Date: 2024.10.08 13:46:22 +03'00' (Υπογραφή)	ΔΗΜΗΤΡΙΟΣ ΚΑΛΥΒΑΣ Digitally signed by ΔΗΜΗΤΡΙΟΣ ΚΑΛΥΒΑΣ Date: 2024.10.08 14:55:34 +03'00' (Υπογραφή)

ΔΗΛΩΣΗ ΣΥΓΓΡΑΦΕΑ ΔΙΠΛΩΜΑΤΙΚΗΣ ΕΡΓΑΣΙΑΣ

Ο κάτωθι υπογεγραμμένος Λαμπρόπουλος Σταύρος του Γεωργίου, με αριθμό μητρώου 483212017028 φοιτητής του Πανεπιστημίου Δυτικής Αττικής της Σχολής ΜΗΧΑΝΙΚΩΝ του Τμήματος ΗΛΕΚΤΡΟΛΟΓΩΝ ΚΑΙ ΗΛΕΚΤΡΟΝΙΚΩΝ ΜΗΧΑΝΙΚΩΝ,

δηλώνω υπεύθυνα ότι:

Είμαι συγγραφέας αυτής της διπλωματικής εργασίας και ότι κάθε βοήθεια την οποία είχα για την προετοιμασία της είναι πλήρως αναγνωρισμένη και αναφέρεται στην εργασία. Επίσης, οι όποιες πηγές από τις οποίες έκανα χρήση δεδομένων, ιδεών ή λέξεων, είτε ακριβώς είτε παραφρασμένες, αναφέρονται στο σύνολό τους, με πλήρη αναφορά στους συγγραφείς, τον εκδοτικό οίκο ή το περιοδικό, συμπεριλαμβανομένων και των πηγών που ενδεχομένως χρησιμοποιήθηκαν από το διαδίκτυο. Επίσης, βεβαιώνω ότι αυτή η εργασία έχει συγγραφεί από μένα αποκλειστικά και αποτελεί προϊόν πνευματικής ιδιοκτησίας τόσο δικής μου, όσο και του Ιδρύματος.

Παράβαση της ανωτέρω ακαδημαϊκής μου ευθύνης αποτελεί ουσιώδη λόγο για την ανάκληση του διπλώματός μου.

Ο Δηλών

Λαμπρόπουλος Σταυρός

Αφιερωμένο σε όσες αόρατες δυνάμεις με ώθησαν να φτάσω ως εδώ, στην περιέργεια που με συνόδευσε και στις προκλήσεις που μου έδωσαν κίνητρο να συνεχίσω.

Αρχικά, θέλω να ευχαριστήσω θερμά τον επιβλέποντα καθηγητή μου, τον κ. Ζώη Ηλία, για την καθοδήγηση του και την υπομονή που έδειξε κατά την διάρκεια διεκπεραίωσης της διπλωματικής μου εργασίας. Και τέλος, θα ήθελα να ευχαριστήσω την οικογένεια μου για τη στήριξη και την ενθάρρυνση που μου παρείχαν σε κάθε βήμα μου.

Περίληψη

Στην παρούσα διπλωματική εργασία διερευνάται η ανάλυση και εφαρμογή μεθόδων για την επιλογή αντιπροσωπευτικών υποσυνόλων δεδομένων από μεγάλα και πολυδιάστατα σύνολα δεδομένων. Η ανάγκη για αποδοτικές τεχνικές συμπύκνωσης δεδομένων έχει αυξηθεί σημαντικά τα τελευταία χρόνια λόγω της ανάπτυξης της τεχνητής νοημοσύνης και της μηχανικής μάθησης. Η παρούσα διπλωματική εργασία αυτή υλοποιεί και βασίζεται στη μεθοδολογία που παρουσιάστηκε στην δημοσίευση των Ehsan Elhamifar, Guillermo Sapiro και Rene Vidal, "See all by looking at a few: Sparse modeling for finding representative objects", εφαρμόζοντας τον αλγόριθμο SMRS (Sparse Modeling via Representative Selection). Ο SMRS αποτελεί ένα καινοτόμο εργαλείο για την εξαγωγή ενός μικρού, αλλά εξαιρετικά αντιπροσωπευτικού υποσυνόλου από μεγάλα δεδομένα. Πραγματοποιείται μια εκτενής ανάλυση της θεωρητικής βάσης του αλγορίθμου SMRS, ο οποίος στηρίζεται στην αραιή αναπαράσταση και τη μείωση διαστάσεων, προσφέροντας μια ισχυρή μέθοδο για την επιλογή αντιπροσώπων. Επιπλέον, γίνεται ανάπτυξη κώδικα σε περιβάλλον Matlab, ο οποίος χρησιμοποιείται για την εξαγωγή των πιο αντιπροσωπευτικών καρέ από βίντεο και εικόνες από σύνολα δεδομένων. Τα επιλεγμένα καρέ και εικόνες μπορούν να χρησιμοποιηθούν σε διάφορες εφαρμογές, όπως η συμπίεση δεδομένων, η βελτιστοποίηση αλγορίθμων μηχανικής μάθησης και η ανάλυση βίντεο. Τα αποτελέσματα της εργασίας αποδεικνύουν τη δυνατότητα του SMRS να εντοπίζει αποτελεσματικά αντιπροσωπευτικά καρέ από βίντεο και αντιπροσωπευτικές εικόνες από μεγάλες βιβλιοθήκες, ενώ παράλληλα παρέχονται προτάσεις για μελλοντικές βελτιώσεις και εφαρμογές του αλγορίθμου. Η παρούσα διπλωματική εργασία ακολουθεί τα βήματα της αρχικής μεθοδολογίας και συνεισφέρει σημαντικά στο πεδίο της υπολογιστικής όρασης και των τεχνικών μείωσης διαστάσεων, προτείνοντας μια πρακτική και καινοτόμο προσέγγιση για τη διαχείριση μεγάλων συνόλων δεδομένων. Τέλος, ανοίγει νέους δρόμους για την περαιτέρω βελτίωση της αποδοτικότητας του SMRS και τη χρήση του σε άλλους τομείς, όπως η βιοπληροφορική και η επεξεργασία σήματος.

Λέξεις – κλειδιά

Αραιή αναπαράσταση, εκμάθηση λεξικού, κυρτή βελτιστοποίηση, μείωση διαστάσεων, υπολογιστική όραση, αλγόριθμος SMRS.

Abstract

In the present thesis, we delve into the analysis and application of methods for selecting representative subsets of data from large and high-dimensional datasets are investigated. The need for efficient data compression techniques has significantly increased in recent years due to the development of artificial intelligence and machine learning. This thesis implements and is based on the methodology presented in the work of Ehsan Elhamifar, Guillermo Sapiro, and Rene Vidal, "See All by Looking at a Few: Sparse Modeling for Finding Representative Objects," applying the SMRS algorithm (Sparse Modeling via Representative Selection). SMRS is an innovative tool for extracting a small but highly representative subset from large datasets. An extensive analysis of the theoretical foundation of the SMRS algorithm is conducted, which relies on sparse representation and dimensionality reduction, offering a powerful method for representative selection. Additionally, code is developed in the Matlab environment, which is used for extracting the most representative frames from videos and images from datasets. The selected frames and images can be used in various applications, such as data compression, optimization of machine learning algorithms, and video analysis. The results of the thesis demonstrate the capability of SMRS to effectively identify representative frames from videos and representative images from large datasets, while also providing suggestions for future improvements and applications of the algorithm. This thesis follows the steps of the initial methodology and contributes significantly to the field of computer vision and dimensionality reduction techniques, proposing a practical and innovative approach for managing large datasets. Finally, it opens new avenues for further improvement of the efficiency of SMRS and its use in other fields, such as bioinformatics and signal processing

Keywords

Sparse representation, dictionary learning, convex optimization, dimensionality reduction, computer vision, SMRS algorithm

Περιεχόμενα

Κατάλογος Εικόνων

Αλφαβητικό Ευρετήριο

SMRS: Sparse Modeling Representative Selection

PCA: Principal Component Analysis

k-means: k-means clustering

ADMM: Alternating Direction Method of Multipliers

IEEE: The Institute for Electrical and Electronics Engineers

ΕΙΣΑΓΩΓΗ

Στη σημερινή εποχή, που βασίζεται στην ανάλυση και την κατανόηση τεράστιων όγκων πληροφορίας, η ικανότητα της αποτελεσματικής σύνοψης και ερμηνείας μεγάλων συνόλων δεδομένων είναι υψίστης σημασίας. Η παρούσα διπλωματική, επικεντρώνεται και υλοποιεί κατά ένα μεγάλο μέρος το καινοτόμο έργο που παρουσιάζεται στο επιστημονικό άρθρο "See All by Looking at A Few"[1] των Ehsan Elhamifar, Guillermo Sapiro και Rene Vidal, για την εύρεση αντιπροσώπων σε μεγάλα σύνολα δεδομένων (αλγόριθμος SMRS). Ο κώδικας μπορεί να αναζητηθεί στην διεύθυνση του GitHub https://github.com/StavLabr/Comp.Vis.Algo.-for-representatives-selection. Μέσω μιας περιεκτικής ανάλυσης και υλοποίησης, η παρούσα διπλωματική όχι μόνο ρίχνει φως στις καινοτόμες συνεισφορές των Elhamifar, Sapiro και Vidal, αλλά δίνει και τη δυνατότητα σε άλλους να εφαρμόσουν αυτές τις γνώσεις σε ένα ευρύ φάσμα τομέων. Αυτοί οι τομείς περιλαμβάνουν την υπολογιστική όραση, την επεξεργασία σήματος και τη βιοπληροφορική. Ο συνδυασμός της θεωρητικής εξερεύνησης και της πρακτικής εφαρμογής αποτελεί μια ισχυρή βάση για μελλοντική έρευνα και ανάπτυξη πάνω στις τεχνικές συμπύκνωσης δεδομένων που είναι όλο και πιο αναγκαίες κατά την εκπαίδευση των μοντέλων τεχνητής νοημοσύνης.

Αντικείμενο της διπλωματικής εργασίας

Το κύριο θέμα της διπλωματικής εργασίας είναι η παρουσίαση και η χρήση τεχνικών για την επιλογή ενός μικρού αλλά αντιπροσωπευτικού υποσυνόλου δεδομένων, το οποίο συχνά αναφέρεται ως "λεξικό", από μεγάλα σύνολα δεδομένων, χρησιμοποιώντας αραιή μοντελοποίηση. Τα τελευταία χρόνια, με την άνοδο της τεχνητής νοημοσύνης και την αυξανόμενη ανάγκη για εκπαίδευση νευρωνικών δικτύων, έχει δημιουργηθεί αυξημένη ζήτηση για τεχνικές εκπαίδευσης που εξοικονομούν πόρους και χρόνο κατά τη διαδικασία εκπαίδευσης των μοντέλων. Αυτές οι τεχνικές συμβάλλουν σημαντικά στην αποτελεσματική ανάλυση, συμπίεση και ομαδοποίηση μεγάλων συνόλων δεδομένων.

Σκοπός και στόχοι

Ο σκοπός της παρούσας διπλωματικής εργασίας είναι η ανάλυση και χρήση του αλγορίθμου SMRS (Sparse Modeling via Representative Selection) για την αποδοτική και ακριβή αναγνώριση και κατηγοριοποίηση στοιχείων σε πολυδιάστατα δεδομένα. Συγκεκριμένα, η εργασία εστιάζει στην εφαρμογή του SMRS αλγορίθμου για την εξαγωγή αντιπροσωπευτικών καρέ (Frames) από μεγάλα σύνολα δεδομένων βίντεο καθώς και για την επιλογή αντιπροσωπευτικών εικόνων από εκτενείς βιβλιοθήκες. Αυτή η προσέγγιση επιτρέπει την ανάδειξη των πιο χαρακτηριστικών δεδομένων, τα οποία συμπυκνώνουν την κύρια πληροφορία και δομή ολόκληρου του συνόλου. Η επιλογή των αντιπροσωπευτικών καρέ βοηθά στην μείωση της πολυπλοκότητας κατά την επεξεργασία και ανάλυση βίντεο, καθιστώντας δυνατή την καλύτερη διαχείριση των δεδομένων και την ενίσχυση της αποδοτικότητας σε εφαρμογές όπως η ανάλυση οπτικού περιεχομένου, η δημιουργία περιλήψεων βίντεο, καθώς και η βελτιστοποίηση συστημάτων μηχανικής μάθησης που επεξεργάζονται μεγάλα σύνολα δεδομένων. Επιπλέον, η επιλογή των αντιπροσωπευτικών εικόνων βοηθά στην εξοικονομήσει υπολογιστικών πόρων κατά την εκπαίδευση συστημάτων μηχανικής μάθησης. Η επιλογή του αλγορίθμου SMRS βασίζεται στην ικανότητά του να εντοπίζει σπάνιες αλλά σημαντικές πληροφορίες σε πολυδιάστατα δεδομένα, προσφέροντας μια ισχυρή μέθοδο για την αντιμετώπιση των προκλήσεων που προκύπτουν από την υψηλή διαστατικότητα και την πολυπλοκότητα των δεδομένων βίντεο και εικόνας.

Μεθοδολογία

Στην παρούσα διπλωματική εργασία, η μεθοδολογία που ακολουθήθηκε για την υλοποίηση της περιλαμβάνει την ανάπτυξη κώδικα στο προγραμματιστικό περιβάλλον του Matlab για την εξαγωγή αντιπροσωπευτικών καρέ από βίντεο, καθώς και την εξαγωγή αντιπροσωπευτικών εικόνων από ένα dataset εικόνων. Αρχικά, ο κώδικας λαμβάνει ως είσοδο ένα βίντεο ανεξαρτήτως διάρκειας, στο οποίο πραγματοποιείται μια προ-επεξεργασία. Από την προ-επεξεργασία αυτή εξάγονται καρέ, τα οποία εισάγονται στον αλγόριθμο SMRS για τη δημιουργία του λεξικού μας, δηλαδή για να βρεθούν τα αντιπροσωπευτικά καρέ του βίντεο. Πιο συγκεκριμένα, κατά την προ-επεξεργασία του βίντεο, αποσκοπούμε στην εξαγωγή των καρέ, τα οποία μετατρέπουμε σε ασπρόμαυρα. Με αυτόν τον τρόπο, κάθε εικονοστοιχείο (pixel) του κάθε καρέ αντιπροσωπεύεται από μία μόνο τιμή φωτεινότητας (luminance) αντί για τρεις τιμές (κόκκινο, πράσινο, μπλε) που χρησιμοποιούνται στον χώρο RGB. Στη συνέχεια, κάθε καρέ μετατρέπεται σε μία μονοδιάστατη γραμμή δεδομένων, όπου τα εικονοστοιχεία διατάσσονται σειριακά, σχηματίζοντας έναν μεγάλο πίνακα, στον οποίο κάθε στήλη αναπαριστά ένα καρέ του βίντεο. Αυτή η μορφή επιτρέπει την αποδοτική επεξεργασία των δεδομένων από τον αλγόριθμο SMRS. Η επιλογή του αλγόριθμου SMRS έγινε λόγω της ικανότητάς του να επιλέγει αντιπροσωπευτικά καρέ με βάση την αραιή αναπαράσταση των δεδομένων, κάτι που είναι ιδιαίτερα χρήσιμο για την αποτελεσματική αναπαράσταση και ανάλυση μεγάλων βίντεο. Τα αντιπροσωπευτικά καρέ που προκύπτουν μπορούν να χρησιμοποιηθούν για τη συμπίεση δεδομένων ή για τη διευκόλυνση της επόμενης φάσης ανάλυσης, όπως η αναγνώριση προτύπων ή η κατηγοριοποίηση σκηνών. Παράλληλα, η μεθοδολογία αυτή εφαρμόστηκε και για την εξαγωγή αντιπροσωπευτικών εικόνων από ένα dataset εικόνων. Σε αυτή την περίπτωση, ο κώδικας διαβάζει όλες τις εικόνες από έναν προκαθορισμένο φάκελο και θεωρεί πως οι εικόνες είναι ήδη σε ασπρόμαυρη μορφή (grayscale). Για κάθε εικόνα, πραγματοποιείται μετατροπή της σε μονοδιάστατο διάνυσμα, όπου τα εικονοστοιχεία διατάσσονται σειριακά. Αυτά τα διανύσματα αποθηκεύονται σε έναν μεγάλο πίνακα δεδομένων, όπου κάθε στήλη του πίνακα αντιπροσωπεύει μία εικόνα. Με τον τρόπο αυτόν, οι εικόνες μπορούν να υποβληθούν σε επεξεργασία από τον αλγόριθμο SMRS, ο οποίος επιλέγει τις πιο αντιπροσωπευτικές εικόνες με βάση την αραιή αναπαράσταση των δεδομένων. Η διαδικασία αυτή είναι ιδιαίτερα αποδοτική όταν το dataset περιέχει πολλές εικόνες και χρειάζεται να μειωθεί ο όγκος των δεδομένων για περαιτέρω επεξεργασία ή ανάλυση. Οι αντιπροσωπευτικές εικόνες που προκύπτουν μπορούν να χρησιμοποιηθούν για τη συμπίεση δεδομένων, τη βελτίωση της απόδοσης αλγορίθμων μηχανικής μάθησης ή για την αναγνώριση μοτίβων σε μεγάλα σύνολα εικόνων. Συνολικά, η μεθοδολογία που ακολουθήθηκε, τόσο για τα βίντεο όσο και για τα dataset εικόνων, συνδυάζει τεχνικές προ-επεξεργασίας και επεξεργασίας δεδομένων που επιτρέπουν την εξαγωγή των πιο αντιπροσωπευτικών καρέ ή εικόνων, βελτιστοποιώντας την ανάλυση και τη διαχείριση μεγάλων συνόλων δεδομένων.

Καινοτομία

Η καινοτομία της διπλωματικής εργασίας έγκειται στη χρήση του αλγορίθμου SMRS για την εξαγωγή αντιπροσωπευτικών δειγμάτων από μεγάλα σύνολα δεδομένων, όπως βίντεο, κείμενα και εικόνες. Ο αλγόριθμος SMRS διακρίνεται για την ικανότητά του να επιλέγει ένα μικρό, αλλά εξαιρετικά αντιπροσωπευτικό υποσύνολο δεδομένων, διατηρώντας την αρχική τους μορφή και πληροφορία, χωρίς να απαιτείται μετασχηματισμός ή ομαδοποίηση. Αυτό τον καθιστά ιδανικό για εφαρμογές όπως η εξαγωγή αντιπροσωπευτικών καρέ από βίντεο, όπου η ακρίβεια στην αντιπροσωπευτικότητα είναι ζωτικής σημασίας

Δομή

Στο πρώτο κεφάλαιο της διπλωματικής εργασίας παρουσιάζονται οι βασικές εισαγωγικές έννοιες που αφορούν τη μείωση διαστάσεων. Αναλύεται η σημασία της μείωσης της διαστατικότητας σε μεγάλα σύνολα δεδομένων και εξετάζονται οι τεχνικές που χρησιμοποιούνται για αυτόν τον σκοπό. Ιδιαίτερη έμφαση δίνεται στη διαδικασία εύρεσης αντιπροσωπευτικών αντικειμένων σε μεγάλα σύνολα δεδομένων, η οποία διευκολύνει την ανάλυση, μειώνει την πολυπλοκότητα και διατηρεί τις πιο σημαντικές πληροφορίες. Παρουσιάζονται επίσης οι μέθοδοι PCA και K-means, με τα πλεονεκτήματα και μειονεκτήματά τους, όπως και βασικές έννοιες της γραμμικής άλγεβρας, που είναι απαραίτητες για την κατανόηση των μεθόδων μείωσης διαστάσεων. Στο δεύτερο κεφάλαιο, εμβαθύνουμε στην αραιή αναπαράσταση και τις μεθόδους κυρτής βελτιστοποίησης, οι οποίες είναι κεντρικές για την εκμάθηση λεξικού (dictionary learning). Γίνεται ανάλυση του ρόλου του Alternating Direction Method of Multipliers (ADMM) και του convex relaxation στη διαδικασία της βελτιστοποίησης, με εφαρμογές στην εύρεση αντιπροσωπευτικών δεδομένων. Στο τρίτο κεφάλαιο, παρουσιάζεται ο αλγόριθμος SMRS και η εφαρμογή του στην εξαγωγή αντιπροσωπευτικών καρέ από βίντεο καθώς και στην εξαγωγή αντιπροσωπευτικών εικόνων από ένα μεγάλο dataset. Αναλύονται οι βασικές αρχές του αλγορίθμου και παρέχεται η υλοποίηση του αλγορίθμου στο περιβάλλον Matlab, με παραδείγματα που επιδεικνύουν την αποδοτικότητά του στην επεξεργασία δεδομένων βίντεο και εικόνων. Στο τέταρτο κεφάλαιο, παρουσιάζονται τα αποτελέσματα που προέκυψαν από την εφαρμογή του αλγορίθμου SMRS και άλλων τεχνικών που αναλύθηκαν στα προηγούμενα κεφάλαια. Τα αποτελέσματα αποδεικνύουν τη δυνατότητα του αλγορίθμου να επιλέγει αποτελεσματικά αντιπροσωπευτικά καρέ από βίντεο και εικόνες από dataset. Τέλος, στο πέμπτο κεφάλαιο συνοψίζονται τα συμπεράσματα της έρευνας και προτείνονται κατευθύνσεις για περαιτέρω μελέτη και βελτιώσεις στον αλγόριθμο, με στόχο την ενίσχυση της αποδοτικότητάς του και την προσαρμογή του σε άλλα είδη δεδομένων.

1 ΚΕΦΑΛΑΙΟ 1ο : Εισαγωγικές Έννοιες

Σε αυτό το κεφάλαιο θα αναλυθεί η σημασία της μείωσης της διαστατικότητας σε μεγάλα σύνολα δεδομένων και θα παρουσιαστούν οι τεχνικές που χρησιμοποιούνται για τον σκοπό αυτό. Αρχικά, θα εξετάσουμε γιατί η μείωση της διαστατικότητας είναι κρίσιμη για την επεξεργασία μεγάλων δεδομένων, καθώς βοηθά στην αντιμετώπιση της υπερπροσαρμογής, στη βελτίωση της απόδοσης των αλγορίθμων και στη μείωση του υπολογιστικού κόστους. Ένα επιπλέον σημαντικό ζήτημα που θα εξεταστεί είναι η ανάγκη για την εύρεση αντιπροσωπευτικών αντικειμένων σε μεγάλα σύνολα δεδομένων. Στα μεγάλα σύνολα δεδομένων, η εύρεση τέτοιων αντιπροσώπων μπορεί να συμβάλει σημαντικά στη μείωση της πολυπλοκότητας των δεδομένων, διευκολύνοντας τη μετέπειτα ανάλυση και επιτρέποντας τη διατήρηση των πιο σημαντικών πληροφοριών. Αυτή η διαδικασία είναι ζωτικής σημασίας για την αποτελεσματική συμπύκνωση των δεδομένων και την εξαγωγή συμπερασμάτων που αντικατοπτρίζουν με ακρίβεια τη συνολική τους δομή και χαρακτηριστικά. Στη συνέχεια, θα αναλύσουμε δύο ευρέως χρησιμοποιούμενες τεχνικές μείωσης της διαστατικότητας: την Ανάλυση Κύριων Συνιστωσών (Principal Component Analysis - PCA) και την Ομαδοποίηση με K-Μέσους (K-means). Αν και αυτές οι τεχνικές έχουν σημαντικά πλεονεκτήματα, όπως η ικανότητα να μειώνουν τις διαστάσεις χωρίς μεγάλη απώλεια πληροφορίας και η απλότητά τους, παρουσιάζουν επίσης και ορισμένα μειονεκτήματα. Συγκεκριμένα, η PCA βασίζεται σε γραμμικές σχέσεις και μπορεί να μην αποδίδει καλά σε δεδομένα με πιο περίπλοκες δομές, ενώ η K-means μπορεί να δυσκολεύεται να εντοπίσει μη γραμμικά όρια ανάμεσα σε ομάδες δεδομένων. Επιπλέον, και οι δύο τεχνικές μπορεί να μην είναι ιδανικές για την εξαγωγή αντιπροσωπευτικών αντικειμένων που διατηρούν τη δομή των αρχικών δεδομένων. Για την κατανόηση και εφαρμογή πιο προχωρημένων μεθόδων μείωσης της διαστατικότητας είναι απαραίτητη η κατανόηση ορισμένων βασικών εννοιών της γραμμικής άλγεβρας. Στη συνέχεια του κεφαλαίου θα εξετάσουμε αυτές τις έννοιες, οι οποίες θα αποτελέσουν τη θεωρητική βάση για την ανάλυση πιο σύνθετων μεθόδων που θα παρουσιαστούν στα επόμενα κεφάλαια.

1.1 Σημασία και Έννοια της Μείωσης Διαστάσεων

Η μείωση διαστάσεων αποτελεί μια κρίσιμη διαδικασία στην ανάλυση δεδομένων και στη μηχανική μάθηση, όπου ο στόχος είναι η μείωση του αριθμού των τυχαίων μεταβλητών υπό εξέταση με την απόκτηση ενός συνόλου κύριων-βασικών μεταβλητών. Η μείωση των διαστάσεων διαδραματίζει σημαντικό ρόλο στη διαχείριση και επεξεργασία των δεδομένων υψηλών διαστάσεων, τα οποία είναι συχνά σε διάφορους τομείς όπως η επεξεργασία εικόνας, η εκπαίδευση νευρωνικών δικτύων και η ιατρική. Η σπουδαιότητα της μείωσης διαστάσεων έγκειται στην ικανότητά της να απλοποιεί τα μοντέλα, να μειώνει το υπολογιστικό κόστος και να βελτιώνει την απόδοση των αλγορίθμων μειώνοντας την κατάρα της διαστατικότητας [2]. Τα δεδομένα υψηλών διαστάσεων συχνά θέτουν σημαντικές προκλήσεις λόγω της κατάρας της διαστατικότητας [2], όπου ο όγκος του χώρου αυξάνεται εκθετικά με τον αριθμό των διαστάσεων. Αυτό έχει ως αποτέλεσμα τα δεδομένα να γίνονται αραιά, καθιστώντας δύσκολη την εύρεση ουσιαστικών προτύπων και σχέσεων. Οι τεχνικές μείωσης διαστάσεων βοηθούν στην υπέρβαση αυτών των προκλήσεων μετασχηματίζοντας τα δεδομένα σε έναν χώρο χαμηλότερων διαστάσεων διατηρώντας ταυτόχρονα τα ουσιώδη χαρακτηριστικά τους. Αυτός ο μετασχηματισμός όχι μόνο καθιστά τα δεδομένα πιο διαχειρίσιμα αλλά και βελτιώνει την αποδοτικότητα των αλγορίθμων μηχανικής μάθησης μειώνοντας το υπολογιστικό βάρος και τις απαιτήσεις αποθήκευσης [4]. Επιπλέον, η μείωση διαστάσεων συμβάλλει στη βελτίωση της επεξηγησιμότητας των δεδομένων. Στον τομέα της υπολογιστικής όρασης και της μηχανικής μάθησης, η επεξηγησιμότητα αναφέρεται στην ικανότητα κατανόησης και ερμηνείας των εσωτερικών μηχανισμών και των αποτελεσμάτων ενός μοντέλου. Μειώνοντας τον αριθμό των χαρακτηριστικών ή των μεταβλητών, καθίσταται ευκολότερο να αναλυθούν οι υποκείμενες σχέσεις και τα σημαντικότερα πρότυπα στα δεδομένα, γεγονός που διευκολύνει την ερμηνεία των αποτελεσμάτων και την εξαγωγή ουσιαστικών συμπερασμάτων. Τα δεδομένα υψηλών διαστάσεων είναι κατά κανόνα περίπλοκα και δύσκολα να οπτικοποιηθούν. Με τη μείωση του αριθμού των διαστάσεων, γίνεται ευκολότερη η οπτικοποίηση και η κατανόηση της υποκείμενης δομής των δεδομένων. Τεχνικές όπως η Ανάλυση Κύριων Συνιστωσών (PCA) [5] χρησιμοποιούνται συνήθως για αυτόν τον σκοπό. Αυτές οι μέθοδοι επιτρέπουν στους ερευνητές και τους αναλυτές να αποκτούν γνώση για τα δεδομένα αναγνωρίζοντας τα πιο σημαντικά χαρακτηριστικά και εξαλείφοντας τον θόρυβο, οδηγώντας σε πιο ακριβή και ερμηνεύσιμα μοντέλα [4]. Στο πλαίσιο της μελέτης της δημοσίευσης των E. Elhamifar, G. Sapiro and R. Vidal [1], η μείωση διαστάσεων είναι καθοριστική στο πλαίσιο του μοντέλου που προτείνουν. Η προσέγγισή τους αξιοποιεί την έννοια της αραιής αναπαράστασης για τον εντοπισμό ενός μικρού υποσυνόλου αντιπροσωπευτικών αντικειμένων από ένα μεγάλο σύνολο δεδομένων. Εστιάζοντας σε λίγα αντιπροσωπευτικά αντικείμενα, η μέθοδος μειώνει αποτελεσματικά τις διαστάσεις του προβλήματος, καθιστώντας εφικτή την υπολογιστική ανάλυση και ερμηνεία των δεδομένων μεγάλης κλίμακας. Αυτό όχι μόνο ενισχύει την αποδοτικότητα του αλγορίθμου αλλά και διασφαλίζει ότι τα επιλεγμένα αντιπροσωπευτικά αντικείμενα καταγράφουν τη βασική δομή των δεδομένων, διευκολύνοντας τη λήψη καλύτερων αποφάσεων και την ανάλυση [1]. Επιγραμματικά, η μείωση διαστάσεων είναι θεμελιώδης στη διαχείριση και επεξεργασία δεδομένων υψηλών διαστάσεων. Απλοποιεί τα μοντέλα, μειώνει το υπολογιστικό κόστος και βελτιώνει την ερμηνευσιμότητα των δεδομένων, καθιστώντας την ένα απαραίτητο εργαλείο στη σύγχρονη ανάλυση δεδομένων και στη μηχανική μάθηση. Το πλαίσιο του μοντέλου [1] που συζητείται στην διπλωματική εργασία αναδεικνύει τη σπουδαιότητα της μείωσης των διαστάσεων στον εντοπισμό αντιπροσωπευτικών αντικειμένων, δείχνοντας τις πρακτικές εφαρμογές και τα οφέλη της σε πραγματικά σενάρια.

1.2 Η Σπουδαιότητας της Εύρεσης Αντιπροσώπων σε Μεγάλα Σύνολα Δεδομένων

Στην εποχή των μεγάλων συνόλων δεδομένων, ο τεράστιος όγκος πληροφοριών που παράγεται σε διάφορους τομείς της κοινωνίας μας, παρουσιάζει τόσο ευκαιρίες όσο και προκλήσεις. Τα μεγάλα σύνολα δεδομένων, που περιλαμβάνουν ποικίλες και εκτεταμένες πληροφορίες, έχουν τη δυνατότητα να μας οδηγήσουν σε σημαντικές καινοτομίες [6]. Ωστόσο, η εξαγωγή ουσιαστικών προτύπων από αυτές τις τεράστιες δεξαμενές δεδομένων απαιτεί συχνά εξελιγμένες τεχνικές [7]. Μία τέτοια κρίσιμη τεχνική είναι η εύρεση αντιπροσώπων μέσα από το σύνολο των δεδομένων. Η εύρεση αντιπροσώπων επιτρέπει την περίληψη ολόκληρου του συνόλου των δεδομένων σε μια πιο διαχειρίσιμη μορφή. Με τον εντοπισμό βασικών παραδειγμάτων ή υποσυνόλων που αποτυπώνουν αποτελεσματικά τα ουσιώδη χαρακτηριστικά και τη μεταβλητότητα του συνόλου δεδομένων, οι αναλυτές μπορούν να εργαστούν με μια συμπυκνωμένη έκδοση που διατηρεί τις κρίσιμες πληροφορίες, διευκολύνοντας έτσι την εξερεύνηση και την κατανόηση [8]. Η εργασία με περιορισμένα σύνολα δεδομένων που αποτελούνται από αντιπροσώπους βελτιώνει σημαντικά την υπολογιστική αποδοτικότητα. Πολλοί αλγόριθμοι επεξεργασίας δεδομένων και μηχανικής μάθησης απαιτούν μεγάλη υπολογιστική ισχύ και η λειτουργία σε μικρότερα, αντιπροσωπευτικά υποσύνολα μειώνει τον χρόνο και την υπολογιστική ισχύ που απαιτείται [9]. Αυτή η αποδοτικότητα είναι κρίσιμη για εργασίες όπως η εκπαίδευση μοντέλων μηχανικής μάθησης, η εκτέλεση προσομοιώσεων και η διενέργεια αναλύσεων σε πραγματικό χρόνο [4]. Οι αντιπρόσωποι παρέχουν μια πιο καθαρή και ευανάγνωστη εικόνα των δεδομένων. Εστιάζοντας σε ένα επιλεγμένο υποσύνολο που ενσωματώνει τα κύρια χαρακτηριστικά του συνόλου δεδομένων, έτσι οι ενδιαφερόμενοι μπορούν να αποκτήσουν γνώσεις χωρίς να φορτώνονται από περιττές ή επαναλαμβανόμενες πληροφορίες [4]. Αυτή η αναγνωσιμότητα είναι ζωτικής σημασίας για τις διαδικασίες λήψης αποφάσεων, όπου η κατανόηση των βασικών προτύπων και τάσεων είναι απαραίτητη [10]. Επιπλέον, η οπτικοποίηση μεγάλων συνόλων δεδομένων μπορεί να είναι δύσκολη λόγω της υψηλής διαστατικότητας και του όγκου των δεδομένων. Οι αντιπρόσωποι απλοποιούν αυτή την εργασία μειώνοντας τον αριθμό των σημείων-δεδομένων που πρέπει να οπτικοποιηθούν, καθιστώντας εφικτή την δημιουργία καθαρών οπτικών αναπαραστάσεων [11]. Η αποτελεσματική οπτικοποίηση βοηθά στην καλύτερη επικοινωνία των ευρημάτων και στην πιο διαισθητική κατανόηση της δομής των δεδομένων [12]. Στη μηχανική μάθηση, η εκπαίδευση μοντέλων σε αντιπροσωπευτικά υποσύνολα μπορεί να βελτιώσει την ικανότητά τους να γενικεύουν σε νέα, άγνωστα δεδομένα [13]. Οι αντιπρόσωποι επιλέγονται για να αποτυπώνουν την ποικιλία του συνόλου δεδομένων, διασφαλίζοντας ότι τα μοντέλα που εκπαιδεύονται σε αυτά τα υποσύνολα εκτίθενται σε μια ευρεία γκάμα σεναρίων και προτύπων, που βελτιώνει την απόδοσή τους σε πραγματικά δεδομένα [14]. Η σημασία της εύρεσης αντιπροσώπων σε μεγάλα σύνολα δεδομένων δεν μπορεί να υπερεκτιμηθεί. Είναι μια θεμελιώδης τεχνική που βελτιώνει την αποδοτικότητα, την αναγνωσιμότητα και την ποιότητα της ανάλυσης δεδομένων [13]. Καθώς τα δεδομένα συνεχίζουν να αυξάνονται σε όγκο και πολυπλοκότητα, οι τεχνικές για την ταυτοποίηση των αντιπροσώπων θα γίνονται ολοένα και πιο ζωτικής σημασίας [13]. Με την αξιοποίηση αυτών των τεχνικών, μπορούμε να ξεκλειδώσουμε βαθύτερες γνώσεις και να προωθήσουμε την σε ένα πλήθος τομέων [14]. Στα επόμενα κεφάλαια, θα εξερευνήσουμε συγκεκριμένους αλγορίθμους που έχουν σχεδιαστεί για αυτόν τον σκοπό, συμπεριλαμβανομένου και κυρίως του αλγορίθμου SMRS (Sparse Modeling Representative Selection) που αναπτύχθηκε από τους E. Elhamifar, G. Sapiro and R. Vidal, ο οποίος προσφέρει μια ισχυρή προσέγγιση για την επιλογή αντιπροσώπων σε μεγάλα σύνολα δεδομένων [1].

1.3 Χρήσιμες Έννοιες της Γραμμική Άλγεβρα

Η γραμμική άλγεβρα είναι ένας θεμελιώδης κλάδος των μαθηματικών και είναι καίρια για την κατανόηση και τη διαχείριση δεδομένων σε μια πληθώρα επιστημονικών πεδίων. Ειδικότερα, η γραμμική άλγεβρα μας προσφέρει μια ισχυρή εργαλειοθήκη για τη μοντελοποίηση και την επίλυση προβλημάτων σε πεδία όπως η φυσική, η πληροφορική και πολλά άλλα. Η σημασία της γραμμικής άλγεβρας έγκειται στην ικανότητά της να απλοποιεί περίπλοκα προβλήματα σε διαχειρίσιμα μαθηματικά μοντέλα. Οι διανυσματικοί χώροι και οι υποχώροι διανυσματικών χώρων αποτελούν τη ραχοκοκαλιά αυτού του κλάδου, παρέχοντας ένα πλαίσιο για την εξερεύνηση των ιδιοτήτων των γραμμικών συστημάτων μέσω εννοιών όπως η βάση και η διάσταση. Αυτές οι έννοιες επιτρέπουν την περιγραφή των χώρων ως προς τις κατευθύνσεις και τα μεγέθη τους, δίνοντάς μας μια βαθύτερη κατανόηση των γεωμετρικών και αλγεβρικών ιδιοτήτων των γραμμικών εξισώσεων. Οι γραμμικοί μετασχηματισμοί και οι πίνακες επίσης μας επιτρέπουν την αναπαράσταση και την επεξεργασία γραμμικών συστημάτων με συμπαγή και αποτελεσματικό τρόπο. Ο πυρήνας (null space) και η εικόνα (column space) ενός γραμμικού μετασχηματισμού παίζουν κρίσιμο ρόλο στην κατανόηση της συμπεριφοράς των γραμμικών απεικονίσεων, συμπεριλαμβανομένης της αντιστρεψιμότητας και των λύσεων των συσχετιζόμενων γραμμικών εξισώσεων. Περαιτέρω, η μελέτη των ιδιοτιμών και των ιδιοδιανυσμάτων εμπλουτίζει αυτό το μαθηματικό τοπίο, αποκαλύπτοντας τις εγγενείς ιδιότητες των γραμμικών μετασχηματισμών που παραμένουν αμετάβλητες υπό ορισμένες συνθήκες. Αυτές οι έννοιες είναι καθοριστικές σε διάφορες εφαρμογές, από την ανάλυση της σταθερότητας σε δυναμικά συστήματα έως τεχνικές μείωσης της διαστατικότητας των δεδομένων όπως η Ανάλυση Κυρίων Συνιστωσών (PCA). Επιπλέον, οι νόρμες, οι οποίες χρησιμοποιούνται για τη μέτρηση του μεγέθους ή της απόστασης μεταξύ διανυσμάτων και πινάκων, χρησιμοποιούνται εκτενώς σε διάφορους τομείς των μαθηματικών και της επιστήμης των υπολογιστών, καθώς παρέχουν έναν αυστηρό τρόπο ποσοτικοποίησης του "μεγέθους" ή της "κλίμακας" ενός διανύσματος ή πίνακα. Η κατανόηση και η εφαρμογή των νορμών είναι κρίσιμη για την ανάλυση και την επίλυση προβλημάτων που αφορούν την επεξεργασία δεδομένων, τη βελτιστοποίηση και τη μοντελοποίηση συστημάτων. Εν ολίγοις, η παρουσίαση και η εμβάθυνση στις ανωτέρω αναφερθείσες έννοιες είναι αναγκαία για την κατανόηση και εμβάθυνση στις έννοιες της αραιής αναπαράστασης, της κυρτής βελτιστοποίησης κ.λπ., που χρησιμοποιούνται στην παρούσα διπλωματική εργασία.

1.3.1 Διανυσματικοί Χώροι και Υπόχωροι

Ένας διανυσματικός χώρος είναι ένα σύνολο διανυσμάτων μαζί με δύο βασικές πράξεις: την πρόσθεση διανυσμάτων και τον πολλαπλασιασμό ενός διανύσματος με έναν αριθμό από ένα πεδίο αριθμών, όπως οι πραγματικοί ή οι μιγαδικοί αριθμοί. Οι δύο αυτές πράξεις πρέπει να ικανοποιούν τις ακόλουθες ιδιότητες:

Πρόσθεση:

$$+: V \times V \rightarrow V \text{ Εξ. } 1.1$$

Για κάθε διάνυσμα u, v, w που ανήκουν στον χώρο, ισχύει η προσεταιρίστηκη ιδιότητα $[\boldsymbol{u} + (\boldsymbol{v} + \boldsymbol{w}) = (\boldsymbol{u} + \boldsymbol{v}) + \boldsymbol{w}],$ η αντιμεταθετική ιδιότητα $[\boldsymbol{u} + \boldsymbol{v} = \boldsymbol{v} + \boldsymbol{u}]$, το ουδετερο στοιχειο της πρόσθεσης $[υπάρχει\ ένα\ στοιχείο\ 0 \in V, που\ ονομάζεται\ μηδενικό\ διάνυσμα, τέτοιο\ ώστε\ \boldsymbol{v} + \boldsymbol{0} = \boldsymbol{v}\ για\ κάθε\ v \in V]$ και το αντίθετο στοιχείο της πρόσθεσης $[Για\ κάθε\ v \in V, υπάρχει\ ένα\ στοιχείο - v \in V, τέτοιο\ ώστε\ \boldsymbol{v} + (-\boldsymbol{v}) = \boldsymbol{0}]$.

Πολλαπλασιασμός με αριθμό:

$$\cdot : F \times V \rightarrow V \text{ Εξ. } 1.2$$

Το σύμβολο F αναφέρεται σε ένα σώμα (ή πεδίο αριθμών), το οποίο μπορεί να είναι, για παράδειγμα, το σύνολο των πραγματικών αριθμών R, το σύνολο των μιγαδικών αριθμών C, ή κάποιο άλλο πεδίο αριθμών. Στον διανυσματικό χώρο, τα διανύσματα μπορούν να πολλαπλασιαστούν με αριθμούς από το πεδίο αυτό, διατηρώντας τις ιδιότητες που αναφέρονται παρακάτω.

Για κάθε αριθμό a,b και κάθε διάνυσμα u, v ισχύει η συμβατότητα του βαθμωτού πολλαπλασιασμού με τον πολλαπλασιασμό στοιχείων των σωμάτων $[\boldsymbol{\alpha}(\boldsymbol{\beta} \cdot \boldsymbol{v}) = (\boldsymbol{a} \cdot \boldsymbol{b})\boldsymbol{v}]$, το ουδέτερο στοιχείο του βαθμωτού πολλαπλασιασμού $[\boldsymbol{1} \cdot \boldsymbol{v} = \boldsymbol{v}]$, η επιμεριστική ιδιότητα σε σχέση με την προσθήκη διανύσματος $[\boldsymbol{\alpha} \cdot (\boldsymbol{u} + \boldsymbol{v}) = \boldsymbol{a} \cdot \boldsymbol{u} + \boldsymbol{a} \cdot \boldsymbol{v}]$ και επιμεριστική ιδιότητα σε σχέση με την προσθήκη στοιχείων του σώματος $[(\boldsymbol{a} + \boldsymbol{b}) \cdot \boldsymbol{v} = \boldsymbol{a} \cdot \boldsymbol{v} + \boldsymbol{b} \cdot \boldsymbol{v}]$[15].

Ένας υπόχωρος είναι ένα υποσύνολο ενός διανυσματικού χώρου, το οποίο είναι το ίδιο διανυσματικός χώρος με τις ίδιες πράξεις [15]. Δηλαδή, το υποσύνολο πρέπει να ικανοποιεί τις συνθήκες, το μηδενικό διάνυσμα του αρχικού χώρου πρέπει να ανήκει στο υποσύνολο, αν δύο διανύσματα ανήκουν στο υποσύνολο, τότε και το άθροισμά τους πρέπει να ανήκει στο υποσύνολο (κλειστότητα ως προς την πρόσθεση), και αν ένας αριθμός πολλαπλασιάζεται με ένα διάνυσμα που ανήκει στο υποσύνολο, τότε το αποτέλεσμα πρέπει να ανήκει στο υποσύνολο (κλειστότητα ως προς τον πολλαπλασιασμό).

Ο μηδενοχώρος (null space) είναι ένας ειδικός υπόχωρος που αποτελείται μόνο από το μηδενικό διάνυσμα[15]. Συμβολίζεται συνήθως με $\{\mathbf{0}\}$ και ισχύει ότι ο μηδενοχώρος είναι πάντα ένας υποχώρος οποιουδήποτε διανυσματικού χώρου, καθώς περιέχει το μηδενικό διάνυσμα και είναι κλειστός ως προς τις πράξεις της πρόσθεσης και του πολλαπλασιασμού με αριθμό και η διάσταση του μηδενοχώρου είναι 0, καθώς δεν περιέχει κανένα γραμμικά ανεξάρτητο διάνυσμα εκτός από το ίδιο το μηδενικό διάνυσμα, το οποίο δεν θεωρείται μέρος μιας βάσης [15],[16].

Η βάση ενός διανυσματικού χώρου είναι ένα σύνολο από γραμμικά ανεξάρτητα διανύσματα που παράγουν ολόκληρο το χώρο. Δηλαδή, κάθε διάνυσμα στο χώρο μπορεί να εκφραστεί μοναδικά ως γραμμικός συνδυασμός των διανυσμάτων της βάσης:

$$v = a_1 v_1 + a_2 v_2 + \ldots\ldots + a_n v_n \;\; \text{Εξ. } 1.3$$

όπου $\boldsymbol{a_1}, \boldsymbol{a_2}, \ldots\ldots, \boldsymbol{a_n}$ είναι αριθμοί από το πεδίο των αριθμών και $\boldsymbol{v_1}, \boldsymbol{v_2}, \ldots\ldots, \boldsymbol{v_n}$ είναι τα διανύσματα της βάσης.

Η διάσταση ενός διανυσματικού χώρου είναι ο αριθμός των διανυσμάτων στη βάση του. Αν ο χώρος έχει πεπερασμένη διάσταση, τότε η διάστασή του είναι ένας θετικός ακέραιος αριθμός. Αν δεν υπάρχει πεπερασμένο σύνολο διανυσμάτων που να παράγει όλο το χώρο, τότε ο χώρος έχει άπειρη διάσταση [16].

Μερικές σημαντικές ιδιότητες που αφορούν τη βάση και τη διάσταση είναι ότι όλες οι βάσεις ενός διανυσματικού χώρου έχουν το ίδιο πλήθος διανυσμάτων, δηλαδή η διάσταση του διανυσματικού χώρου είναι μοναδική, αν ένας χώρος έχει διάσταση n και ένας υπόχωρος του έχει διάσταση m, τότε ισχύει ότι $\boldsymbol{m} \leq \boldsymbol{n}$ και η διάσταση του υποχώρου μπορεί να είναι οποιοσδήποτε ακέραιος από 0 έως n και επιπλέον κάθε διανυσματικός χώρος είναι ισόμορφος με έναν χώρο $\mathbb{F}^n$, δηλαδή υπάρχει ένας αμφιμονοσήμαντος και γραμμικά ανεξάρτητος μετασχηματισμός που διατηρεί τις πράξεις της πρόσθεσης και του πολλαπλασιασμού [15].

1.3.2 Γραμμικοί Μετασχηματισμοί και Πίνακες

Ένας γραμμικός μετασχηματισμός $T: \mathbb{R}^n \to \mathbb{R}^m$ ορίζεται ως μια απεικόνιση που διατηρεί τις πράξεις της πρόσθεσης και του πολλαπλασιασμού με αριθμό [16]. Αυτό σημαίνει ότι για κάθε $u, v \in \mathbb{R}^n$ και για κάθε $a \in \mathbb{R}$, ισχύει:

$$T(u + v) = T(u) + T(v) \;\; \text{Εξ. } 1.4$$

$$T(au) = aT(u) \;\; \text{Εξ. } 1.5$$

Επιπλέον, οι γραμμικοί μετασχηματισμοί μπορούν να αναπαρασταθούν με τη χρήση πινάκων. Ειδικότερα, για έναν γραμμικό μετασχηματισμό $T: \mathbb{R}^n \to \mathbb{R}^m$ υπάρχει ένας πίνακας $A \in \mathbb{R}^{m \times n}$ τέτοιος ώστε:

$$T(x) = Ax \;\; \text{Εξ. } 1.6$$

Όπου x είναι ένα διάνυσμα στο $\mathbb{R}^n$ [16],[17]

Οι γραμμικοί μετασχηματισμοί και οι αντίστοιχοι πίνακες έχουν αρκετές σημαντικές ιδιότητες που είναι κρίσιμες για την ανάλυση και την εφαρμογή τους:

• Πυρήνας (Kernel) ή Μηδενοχώρος (Null Space): Ο πυρήνας ή μηδενοχώρος ενός γραμμικού μετασχηματισμού T είναι το σύνολο των διανυσμάτων x που απεικονίζονται στο μηδενικό διάνυσμα 0. Δηλαδή:

$$\ker(T) = \{x\epsilon\mathbb{R}^n: T(x) = 0\} \text{ Εξ. } 1.7$$

Αντίστοιχα, για τον πίνακα A, ο μηδενοχώρος είναι:

$$null(A) = \{x\epsilon\mathbb{R}^n: Ax = 0\} \text{ Εξ. } 1.8$$

Ο μηδενοχώρος είναι ένας γραμμικός υπόχωρος του $\mathbb{R}^n$ και παίζει κεντρικό ρόλο στη μελέτη της λύσης συστημάτων γραμμικών εξισώσεων [15],[17]. Ειδικότερα, αν ο μηδενοχώρος αποτελείται μόνο από το μηδενικό διάνυσμα, ο γραμμικός μετασχηματισμός T είναι μονοσήμαντος.

• Εικόνα (Image) ή Πεδίο Τιμών (Range): Η εικόνα ενός γραμμικού μετασχηματισμού T είναι το σύνολο των διανυσμάτων που μπορεί να προκύψουν ως αποτέλεσμα του μετασχηματισμού. Δηλαδή:

$$Im(T) = \{T(x): x\epsilon\mathbb{R}^n\} \text{ Εξ. } 1.9$$

Αντίστοιχα, για τον πίνακα A, το πεδίο τιμών είναι:

$$range(A) = \{Ax: x\epsilon\mathbb{R}^n\} \text{ Εξ. } 1.10$$

Η εικόνα είναι ένας γραμμικός υπόχωρος του $\mathbb{R}^m$.

• Διαστάσεις και Θεώρημα Διάστασης: Το Θεώρημα Διάστασης (ή Θεώρημα Κατάταξης) ορίζει ότι η διάσταση του χώρου από τον οποίο ξεκινά ο μετασχηματισμός είναι ίση με το άθροισμα των διαστάσεων του πυρήνα και της εικόνας:

$$\dim(\mathbb{R}^n) = dim\big(ker(T)\big) + dim\big(Im(T)\big) \text{ Εξ. } 1.11$$

Αυτό το θεώρημα υποδεικνύει ότι ο αριθμός των γραμμικά ανεξάρτητων διανυσμάτων που απεικονίζονται στο μηδέν (μηδενοχώρος) και ο αριθμός των γραμμικά ανεξάρτητων διανυσμάτων που απεικονίζονται σε μη μηδενικές τιμές (εικόνα) καλύπτουν συνολικά όλο τον χώρο $\mathbb{R}^n$ [17].

• Ιδιότιμες τιμές και ιδιοδιανύσματα: Ένα διάνυσμα v είναι ιδιοδιάνυσμα του γραμμικού μετασχηματισμού T αν υπάρχει ένα $\lambda \in \mathbb{R}$ τέτοιο ώστε:

$$T(v) = \lambda v \text{ Εξ. } 1.12$$

Το λ καλείται ιδιοτιμή του T που αντιστοιχεί στο ιδιοδιάνυσμα v. Οι ιδιοτιμές και τα ιδιοδιανύσματα έχουν σημαντική γεωμετρική σημασία, καθώς συνδέονται με την κλίση και τη συστολή ή επιμήκυνση που εφαρμόζει ο μετασχηματισμός στις διάφορες κατευθύνσεις του χώρου [17], [19].

Οι γραμμικοί μετασχηματισμοί μπορούν να ερμηνευτούν γεωμετρικά ως παραμορφώσεις του χώρου. Για παράδειγμα:

• Κλιμάκωση: Ένας γραμμικός μετασχηματισμός μπορεί να κλιμακώνει έναν χώρο, δηλαδή να μεγεθύνει ή να μειώσει τις αποστάσεις στον χώρο κατά έναν σταθερό παράγοντα [20].

• Περιστροφή: Ένας γραμμικός μετασχηματισμός μπορεί να περιστρέφει έναν χώρο γύρω από κάποιο σημείο ή άξονα [20]

• Διάτμηση: Ο μετασχηματισμός μπορεί να διαμορφώσει έναν χώρο μέσω διάτμησης, όπου οι ευθείες γραμμές που περνούν από την αρχή των αξόνων παραμένουν ευθείες, αλλά η γωνία μεταξύ τους μεταβάλλεται [20].

Οι πίνακες που αντιστοιχούν σε γραμμικούς μετασχηματισμούς έχουν τις δικές τους ιδιότητες, που είναι κρίσιμες στη μελέτη τους:

• Αναστρέψιμοι Πίνακες: Ένας πίνακας A είναι αναστρέψιμος αν υπάρχει πίνακας A^{-1} τέτοιος ώστε:

$$AA^{-1} = A^{-1}A = I \quad \text{Εξ. } 1.13$$

όπου I είναι ο μοναδιαίος πίνακας. Ένας πίνακας είναι αναστρέψιμος αν και μόνο αν οι γραμμές ή οι στήλες του είναι γραμμικά ανεξάρτητες, που σημαίνει ότι ο μηδενοχώρος του είναι ο μηδενικός χώρος {0} [19].

• Ιδιοτιμές και Ιδιοδιανύσματα Πίνακα: Οι ιδιοτιμές και τα ιδιοδιανύσματα ενός πίνακα A είναι οι λύσεις της εξίσωσης:

$$Av = \lambda v \quad \text{Εξ. } 1.14$$

Οι γραμμικοί μετασχηματισμοί και οι πίνακες είναι θεμελιώδεις στην κατανόηση πιο προχωρημένων εννοιών, όπως η κυρτή βελτιστοποίηση και η αραιή αναπαράσταση. Στην αραιή αναπαράσταση, ένας πίνακας που παριστάνει τον γραμμικό μετασχηματισμό μπορεί να χρησιμοποιηθεί για να επιβληθούν ιδιότητες αραιότητας στις λύσεις ενός συστήματος εξισώσεων [21]. Ειδικότερα, ο μηδενοχώρος του πίνακα παίζει σημαντικό ρόλο στην αναγνώριση της αραιότητας, καθώς επηρεάζει τον αριθμό και τη δομή των διανυσμάτων που μπορούν να αναπαρασταθούν αραιά. Η γεωμετρική ερμηνεία των γραμμικών μετασχηματισμών είναι επίσης σημαντική στην κατανόηση της επίδρασης των περιορισμών βελτιστοποίησης και στην εκμάθηση λεξικού [22].

1.3.3 Ιδιοτιμές και Ιδιοδιανύσματα

Σε ένα διανυσματικό χώρο V πάνω από ένα σώμα F (όπως το R ή το C), ένα ιδιοδιάνυσμα και η αντίστοιχη ιδιοτιμή ορίζονται στο πλαίσιο ενός γραμμικού τελεστή $A: V \to V$ [16].

Έστω A ένας γραμμικός τελεστής που δρα σε έναν διανυσματικό χώρο V. Ένα μη μηδενικό διάνυσμα $v \in V$ ονομάζεται ιδιοδιάνυσμα του A αν υπάρχει ένας αριθμός $\lambda \in F$ τέτοιος ώστε:

$$Av = \lambda v \quad \text{Εξ. } 1.15$$

Ο αριθμός λ ονομάζεται ιδιοτιμή του A που αντιστοιχεί στο ιδιοδιάνυσμα v [15].

Χαρακτηριστική Εξίσωση: Η εύρεση των ιδιοτιμών μπορεί να πραγματοποιηθεί λύνοντας την εξίσωση του χαρακτηριστικού πολυωνύμου. Το χαρακτηριστικό πολυώνυμο p(λ) ενός γραμμικού τελεστή A σε έναν πεπερασμένης διάστασης διανυσματικό χώρο V δίνεται από την εξίσωση:

$$p(\lambda) = \det(A - \lambda I) = 0 \text{ Εξ. } 1.16$$

όπου I είναι ο μοναδιαίος τελεστής. Οι λύσεις αυτής της εξίσωσης είναι οι ιδιοτιμές του A [17].

Γραμμική Ανεξαρτησία Ιδιοδιανυσμάτων: Τα ιδιοδιανύσματα που αντιστοιχούν σε διαφορετικές ιδιοτιμές είναι γραμμικά ανεξάρτητα. Αυτό σημαίνει ότι αν $\lambda_1 + \lambda_2 + \ldots\ldots + \lambda_n$ είναι διαφορετικές ιδιοτιμές και $v_1 + v_2 + \ldots\ldots + v_n$ τα αντίστοιχα ιδιοδιανύσματα, τότε το σύνολο $\{v_1 + v_2 + \ldots\ldots + v_n\}$ είναι γραμμικά ανεξάρτητο.

Διαγωνιοποίηση: Ένας τελεστής A είναι διαγωνιοποιήσιμος αν υπάρχει μια βάση του V που αποτελείται από ιδιοδιανύσματα του A. Στην περίπτωση αυτή, ο A μπορεί να γραφεί ως:

$$A = P\Lambda P^{-1} \text{ Εξ. } 1.17$$

όπου P είναι ο πίνακας που οι στήλες του είναι τα ιδιοδιανύσματα και Λ είναι ο διαγώνιος πίνακας με τις αντίστοιχες ιδιοτιμές [19].

Πολλαπλότητα Ιδιοτιμών: Για κάθε ιδιοτιμή λ, η αλγεβρική πολλαπλότητα είναι ο βαθμός του λ ως ρίζας του χαρακτηριστικού πολυωνύμου, ενώ η γεωμετρική πολλαπλότητα είναι η διάσταση του διανυσματικού υπόχωρου που αντιστοιχεί στην λ.

1.3.4 Νόρμες

Οι μαθηματικές νόρμες είναι θεμελιώδεις έννοιες της γραμμικής άλγεβρας και της θεωρίας των διανυσμάτων, οι οποίες χρησιμοποιούνται για τη μέτρηση του μεγέθους ή της απόστασης σε διανυσματικούς χώρους. Οι νόρμες παίζουν σημαντικό ρόλο σε πολλές περιοχές της εφαρμοσμένης μαθηματικής ανάλυσης, όπως η αραιή αναπαράσταση, η κυρτή βελτιστοποίηση και η εκμάθηση λεξικών, τα οποία συνιστούν βασικά εργαλεία για την ανάλυση δεδομένων και την επιλογή αντιπροσώπων από σύνολα δεδομένων.

Μία νόρμα $\|\cdot\|$ σε έναν διανυσματικό χώρο $\mathbb{R}^n$ είναι μια συνάρτηση που ικανοποιεί τις ακόλουθες ιδιότητες για κάθε $x, y \in \mathbb{R}^n$ και για κάθε $a \in \mathbb{R}$:

- Μη αρνητικότητα: $\|x\| \geq 0$, με ισότητα $\|x\| = 0$ αν και μόνο αν $x = 0$.
- Ομοιογένεια: $\|ax\| = |a|\|x\|$ [1].
- Ανισότητα τριγώνου: $\|x + y\| \leq \|x\| + \|y\|$ [20].

Αυτές οι ιδιότητες εξασφαλίζουν ότι η νόρμα μπορεί να χρησιμοποιηθεί ως μέτρο του "μεγέθους" ή της "απόστασης" στο χώρο των διανυσμάτων.

Υπάρχουν διάφορα είδη νόρμων που χρησιμοποιούνται συνήθως:

• Η νόρμα l_1: $\|x\|_1 = \sum_{i=1}^{n} |x_i|$. Αυτή η νόρμα είναι συχνά χρησιμοποιείται στην αραιή αναπαράσταση, καθώς προάγει τη σπανιότητα, επιλέγοντας λύσεις με μικρό αριθμό μη μηδενικών στοιχείων [24].

• Η νόρμα l_2 $(ευκλείδεια\ νόρμα)$: $\|x\|_2 = \left(\sum_{i=1}^{n} {x_i}^2\right)^{1/2}$ Αυτή η νόρμα μετρά την απόσταση του διανύσματος από την αρχή και χρησιμοποιείται εκτενώς σε προβλήματα κυρτής βελτιστοποίησης [24].

• Η νόρμα l_∞: $\|x\|_\infty = \max_i |x_i|$ Αυτή η νόρμα λαμβάνει το μέγιστο απόλυτο στοιχείο του διανύσματος και είναι χρήσιμη σε διάφορα προβλήματα που απαιτούν έλεγχο μέγιστης απόκλισης [16].

Στα προβλήματα κυρτής βελτιστοποίησης, οι νόρμες χρησιμοποιούνται για να καθοριστεί το μέτρο της απόστασης στον χώρο των παραμέτρων και να επιβληθούν περιορισμοί. Ένα παράδειγμα είναι το πρόβλημα:

$$\min_x \|Ax - b\|_2^2 + \lambda \|x\|_1 \quad \text{Εξ. } 1.18$$

Όπου $\min_x \|Ax - b\|_2^2$ είναι η συνάρτηση κόστους (loss function) και $\|x\|_1$ είναι ο κανονικοποιητικός όρος (regularization term) που επιβάλλει σπανιότητα στη λύση x [22].

Η χρήση της νόρμας l_1 είναι κεντρική στην αραιή αναπαράσταση, καθώς επιτρέπει την εύρεση λύσεων με λιγότερα μη μηδενικά στοιχεία. Ένα βασικό πρόβλημα είναι το ακόλουθο:

$$\min_x \|Dx - s\|_2^2 \; s.t. \; \|x\|_1 \leq r \quad \text{Εξ. } 1.19$$

Όπου D είναι ένα λεξικό που περιέχει τα βασικά στοιχεία (atoms) και s είναι το σήμα που προσπαθούμε να αναπαραστήσουμε αραιά [25].

Στην εκμάθηση λεξικού (dictionary learning), οι νόρμες χρησιμοποιούνται για την εύρεση των κατάλληλων λεξικών D και των αναπαραστάσεων X που ικανοποιούν την συνθήκη:

$$\min_{D,X} \|S - DX\|_F^2 + \lambda \|x\|_1 \quad \text{Εξ. } 1.20$$

Όπου $\|\cdot\|_F$ είναι η νόρμα Frobenius, η οποία γενικεύει τη νόρμα l_2 για πίνακες [26].

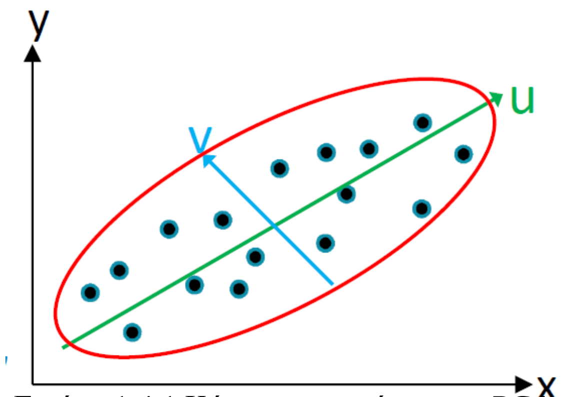
Οι μαθηματικές νόρμες αποτελούν θεμελιώδη εργαλεία στην ανάλυση και επεξεργασία δεδομένων, παρέχοντας τα αναγκαία μέσα για την διατύπωση και επίλυση προβλημάτων όπως η αραιή αναπαράσταση, η κυρτή βελτιστοποίηση και η εκμάθηση λεξικού. Η κατανόηση των διαφόρων τύπων νόρμας και των ιδιοτήτων τους είναι κρίσιμη για την ανάπτυξη αποτελεσματικών αλγορίθμων και μεθόδων, όπως ο αλγόριθμος SMRS, για την επιλογή αντιπροσώπων σε σύνολα δεδομένων.

1.4 Ανάλυση Κύριων Συνιστωσών (PCA) και Αλγόριθμος K-means

Στον χώρο της ανάλυσης δεδομένων, της μηχανικής μάθησης και της υπολογιστικής όρασης, οι τεχνικές μείωσης διαστάσεων και ομαδοποίησης παίζουν καθοριστικό ρόλο στην κατανόηση και ερμηνεία μεγάλων και πολύπλοκων συνόλων δεδομένων. Δύο από τις πιο διαδεδομένες μεθόδους σε αυτούς τους τομείς είναι η Ανάλυση Κύριων Συνιστωσών (PCA) και ο αλγόριθμος K-means. Παρά τη δημοφιλία και τη χρησιμότητά τους, αμφότερες οι μέθοδοι παρουσιάζουν ορισμένες αδυναμίες, οι οποίες περιορίζουν την αποτελεσματικότητά τους σε συγκεκριμένα σενάρια. Η PCA είναι μια μέθοδος μείωσης διαστάσεων που χρησιμοποιείται για την απλοποίηση πολυδιάστατων δεδομένων, διατηρώντας ταυτόχρονα όσο το δυνατόν περισσότερη πληροφορία. Ωστόσο, η γραμμικότητα της μεθόδου και η πιθανότητα απώλειας πληροφορίας αποτελούν σημαντικά εμπόδια σε πιο πολύπλοκα δεδομένα. Από την άλλη, ο αλγόριθμος K-means είναι ένας από τους πιο γνωστούς αλγόριθμους ομαδοποίησης, που χρησιμοποιείται ευρέως για τη διαχωριστική κατάταξη των δεδομένων σε ομάδες (clusters). Παρ' όλα αυτά, ο K-means εξαρτάται έντονα από την επιλογή των αρχικών κεντροειδών και από τον προκαθορισμένο αριθμό ομάδων (clusters), γεγονός που μπορεί να οδηγήσει σε αναποτελεσματικά ή τοπικά βέλτιστα αποτελέσματα. Σε αυτό το μέρος της παρούσας διπλωματικής εργασίας, θα εξετάσουμε τις εφαρμογές και τις αδυναμίες αυτών των δύο σημαντικών αλγορίθμων, αναδεικνύοντας τα πεδία στα οποία αποδίδουν καλύτερα αλλά και τα σημεία στα οποία υστερούν. Η ανάλυση αυτή θα μας προετοιμάσει για τη διερεύνηση πιο εξελιγμένων μεθόδων που μπορούν να αντιμετωπίσουν τις περιορισμένες δυνατότητες αυτών των κλασικών αλγορίθμων.

1.4.1 Ανάλυση Κύριων Συνιστωσών (PCA)

Η Ανάλυση Κύριων Συνιστωσών (PCA) είναι μια ισχυρή τεχνική που χρησιμοποιείται ευρέως στην ανάλυση δεδομένων, κυρίως για τη μείωση των διαστάσεων, την οπτικοποίηση και την εξάλειψη του θορύβου από τα δεδομένα. Η μέθοδος αυτή είναι ιδιαίτερα σημαντική σε προβλήματα υψηλών διαστάσεων, όπως αυτά που συναντώνται στην ανάλυση εικόνων και βίντεο, όπου τα δεδομένα μπορεί να είναι πολυδιάστατα και γεμάτα θόρυβο[27].



Εικόνα 1.4.1 Κύριες συνιστώσες της PCA

Η PCA λειτουργεί μετατρέποντας το σύνολο των δεδομένων σε έναν νέο χώρο συντεταγμένων, όπου οι κύριες συνιστώσες ευθυγραμμίζονται με τις μεγαλύτερες διακυμάνσεις στα δεδομένα όπως φαίνεται στην εικόνα 1.4.1 [27]. Αρχικά, τα δεδομένα κανονικοποιούνται σύμφωνα με την σχέση:

$$X_{std} = \frac{X - \mu}{\sigma} \quad \text{Εξ. } 1.21$$

όπου X είναι ο αρχικός πίνακας δεδομένων, μ είναι ο διάνυσμα μέσης τιμής, και σ είναι ο διάνυσμα τυπικής απόκλισης. Στη συνέχεια, υπολογίζεται ο πίνακας συνδιακύμανσης, ο οποίος καταγράφει τη διακύμανση και τις γραμμικές σχέσεις μεταξύ των χαρακτηριστικών. Ο πίνακας αυτός ορίζεται ως

$$C = \frac{1}{n-1} {X_{std}}^T X_{std} \quad \text{Εξ. } 1.22$$

όπου n είναι ο αριθμός των δειγμάτων. Μετά τον υπολογισμό του πίνακα συνδιακύμανσης, ακολουθεί η ανάλυση των ιδιοτιμών και ιδιοδιανυσμάτων. Ο πίνακας συνδιακύμανσης διασπάται για να βρεθούν οι ιδιοτιμές και τα αντίστοιχα ιδιοδιανύσματα του, σύμφωνα με την εξίσωση

$$Cv = \lambda v \quad \text{Εξ. } 1.23$$

όπου λ αντιπροσωπεύει μια ιδιοτιμή και v αντιπροσωπεύει το αντίστοιχο ιδιοδιάνυσμα [27].
Οι ιδιοτιμές και τα αντίστοιχα ιδιοδιανύσματα ταξινομούνται κατά φθίνουσα σειρά των ιδιοτιμών. Το ιδιοδιάνυσμα με την υψηλότερη ιδιοτιμή είναι η κύρια συνιστώσα με τη μεγαλύτερη διακύμανση. Έπειτα, ακολουθεί η επιλογή των κύριων συνιστωσών, όπου επιλέγονται τα κορυφαία k ιδιοδιανύσματα για να σχηματίσουν έναν νέο πίνακα W. Αυτός ο πίνακας χρησιμοποιείται για τη μετατροπή των δεδομένων στο νέο χωρικό σύστημα. Ο νέος πίνακας δεδομένων υπολογίζεται ως:

$$Z = X_{std} W \quad \text{Εξ. } 1.24$$

όπου Z είναι το μετασχηματισμένο σύνολο δεδομένων στο νέο χωρικό σύστημα. Η ερμηνεία των αποτελεσμάτων του PCA είναι σημαντική για την κατανόηση της ανάλυσης. Οι κύριες συνιστώσες είναι οι κατευθύνσεις στις οποίες τα δεδομένα διαφέρουν περισσότερο. Η πρώτη κύρια συνιστώσα αντιπροσωπεύει τη μεγαλύτερη διακύμανση, η δεύτερη τη δεύτερη μεγαλύτερη διακύμανση, και ούτω καθεξής. Οι ιδιοτιμές αντιπροσωπεύουν το ποσό της διακύμανσης που καταλαμβάνει κάθε κύρια συνιστώσα. Υψηλότερες ιδιοτιμές υποδεικνύουν μεγαλύτερη διακύμανση. Επιλέγοντας ένα υποσύνολο των κύριων συνιστωσών (αυτών με τις υψηλότερες ιδιοτιμές), μπορούμε να μειώσουμε τον αριθμό των διαστάσεων, ενώ διατηρούμε τη μεγαλύτερη διακύμανση των δεδομένων [27]. Παρόλο που η PCA είναι ένα εξαιρετικά χρήσιμο εργαλείο, έχει τους έξεις περιορισμούς:

- *Γραμμικότητα:* Η PCA μπορεί να εντοπίσει μόνο γραμμικές σχέσεις στα δεδομένα. Σε περιπτώσεις όπου τα δεδομένα έχουν πιο σύνθετες, μη γραμμικές σχέσεις, η αποτελεσματικότητα της PCA μειώνεται σημαντικά.
- *Απώλεια πληροφορίας:* Αν και η PCA προσφέρει συμπίεση των δεδομένων, μπορεί να χαθεί σημαντική πληροφορία κατά τη μείωση των διαστάσεων, ειδικά όταν τα δεδομένα δεν περιέχουν σαφώς διαχωρισμένες διακυμάνσεις.

• *Ευαισθησία στον θόρυβο:* Η PCA μπορεί να επηρεαστεί από τον θόρυβο των δεδομένων, καθώς οι διακυμάνσεις μικρής κλίμακας ενδέχεται να επηρεάσουν τη διαδικασία ανάλυσης και να οδηγήσουν σε λιγότερο αποδοτικά αποτελέσματα.

Η ανάγκη για αλγορίθμους που μπορούν να αντιμετωπίσουν την πολυπλοκότητα και τη μη γραμμικότητα των δεδομένων, ιδιαίτερα σε περιπτώσεις όπως η ανάλυση βίντεο ή μεγάλων συνόλων εικόνων, είναι εμφανής. Οι πιο εξελιγμένες τεχνικές που αντιμετωπίζουν περιορισμούς όπως η σπανιότητα (sparsity) και παρέχουν πιο αποδοτική επιλογή αντιπροσώπων (representative selection) είναι κρίσιμες για την ακριβέστερη ανάλυση και κατανόηση των δεδομένων.

1.4.2 Ο Αλγόριθμός K-means

Ο αλγόριθμος k-means είναι ένας από τους πιο γνωστούς και απλούς αλγορίθμους ομαδοποίησης που ανήκουν στην κατηγορία της μη επιβλεπόμενης μάθησης [28]. Η δημοτικότητά του προκύπτει από την ευκολία στην εφαρμογή του και τη γραμμική πολυπλοκότητά του, η οποία είναι της τάξης $O(n)$, όπου n είναι ο αριθμός των στοιχείων του dataset [29]. Για να ομαδοποιηθεί ένα σύνολο δεδομένων μέσω του k-means, πρέπει εκ των προτέρων να καθοριστεί ο αριθμός των ομάδων (clusters) k [28]. Ο αλγόριθμος ξεκινά με την επιλογή k αρχικών κεντροειδών (centroids), ένα για κάθε cluster [30]. Η επιλογή αυτών των αρχικών σημείων παίζει κρίσιμο ρόλο, καθώς διαφορετικές αρχικές θέσεις μπορούν να οδηγήσουν σε διαφορετικά αποτελέσματα [31]. Η αρχική κατανομή των centroids επηρεάζει σημαντικά την έκβαση του αλγορίθμου, γι' αυτό είναι συνήθως προτιμότερο τα centroids να τοποθετούνται όσο το δυνατόν πιο απομακρυσμένα μεταξύ τους [32]. Στη συνέχεια, κάθε στοιχείο του dataset αποδίδεται στο κοντινότερο centroid [31]. Με αυτό τον τρόπο ολοκληρώνεται το πρώτο στάδιο της ομαδοποίησης, και έχει παραχθεί μία αρχική ταξινόμηση των δεδομένων. Το επόμενο βήμα είναι ο υπολογισμός νέων centroids, τα οποία αντιπροσωπεύουν το κέντρο βάρους κάθε cluster [29]. Ακολουθεί ξανά η ανάθεση κάθε στοιχείου στο κοντινότερο centroid, και η διαδικασία επαναλαμβάνεται. Σε κάθε επανάληψη, τα centroids ανανεώνονται και τα στοιχεία ανατίθενται εκ νέου στα clusters. Η διαδικασία ολοκληρώνεται όταν δεν παρατηρούνται πλέον μετακινήσεις των στοιχείων μεταξύ των clusters. Τελικά, τα δεδομένα ομαδοποιούνται σε k clusters [28]

Ο αλγόριθμος στοχεύει να ελαχιστοποιήσει μία *αντικειμενική συνάρτηση*, την λεγόμενη συνάρτηση τετραγωνικού λάθους που ορίζεται ως εξής:

$$J = \sum_{j=1}^{k} \sum_{i=1}^{n} \left\| x_i{}^{(j)} - c_j \right\|^2 \quad \text{Εξ. } 1.25$$

, όπου $\left\|x_i^{(j)} - c_j\right\|^2$ είναι ένα μέτρο απόστασης που χρησιμοποιείται για να μετρά την απόσταση κάθε στοιχείου $x_i^{(j)}$ από το centroid c_j του κάθε cluster. Όπου n το σύνολο των στοιχείων του συνόλου δεδομένων [30].

Παρακάτω φαίνονται συνοπτικά τα βήματα του αλγορίθμου k-means:

Είσοδος:

$D = \{x_1, x_2, \ldots\ldots, x_n\}$ \\\\ *Σύνολο στοιχείων*

k \\\\ *Αριθμός επιθυμητών clusters*

Έξοδος:

k \\\\ *Σύνολο clusters*

k-Means αλγόριθμος:

Ανέθεσε τιμές στα αρχικά centroids $c_1, c_2, \ldots\ldots, c_n$;

Επανέλαβε

Ανέθεσε καθε x_i *στο cluster με του οποίου το centroid η απόσταση ειναι η μικροτερη*;

Υπολόγισε νεα centroids για καθε cluster;

Μέχρι να συναντηθεί το κριτήριο σύγκλισης;

Αν και μπορεί να αποδειχθεί ότι ο αλγόριθμος πάντα τερματίζει, αξίζει να τονιστεί ότι δεν καταφέρνει πάντα να βρίσκει τη βέλτιστη λύση. Ο αλγόριθμος επηρεάζεται σημαντικά από τα αρχικά centroids. Για αυτό πολλές φορές συνίσταται η εκτέλεση του πολλές φορές μέχρι να μειωθεί η επίδραση αυτή. Έστω ότι υπάρχουν n διανύσματα τα $\boldsymbol{x_1}, \boldsymbol{x_2}, \ldots\ldots, \boldsymbol{x_n}$ και όλα είναι της ίδια διάστασης. Ακόμη είναι γνωστό ότι όλα εμπίπτουν σε k συμπαγή clusters, για k < n. Έστω m_i είναι το μέσο διάνυσμα του i cluster. Εφόσον τα clusters είναι σαφώς διαχωρισμένα μεταξύ τους, μπορεί να χρησιμοποιηθεί σαν μέτρο απόστασης μεταξύ των στοιχείων η Ευκλείδεια απόσταση. Αυτό σημαίνει ότι σε κάθε βήμα θα λέγεται: το στοιχείο x ανήκει στο cluster i, εάν η Ευκλείδεια απόσταση του από το centroid του i cluster είναι η μικρότερη σε σχέση με όλες τις άλλες αποστάσεις του από τα centroids των άλλων clusters. Έτσι βρίσκονται οι Ευκλείδειες αποστάσεις για όλα τα στοιχεία και κάθε ένα από αυτά ανατίθεται στο cluster από του οποίου το centroid απέχει λιγότερο (δηλαδή η Ευκλείδεια απόσταση είναι η μικρότερη). Στην συνέχεια υπολογίζονται τα νέα centroids και μετά πάλι οι Ευκλείδειες αποστάσεις όλων των στοιχείων για τα νέα centroids. Γίνονται οι κατάλληλες μετακινήσεις στοιχείων και η ίδια διαδικασία επαναλαμβάνεται μέχρι κανένα στοιχείο να μην μετακινείται σε άλλο cluster, δηλαδή τα clusters να μένουν αμετάβλητα. Παρόλο που ο K-means είναι ένας αποδοτικός αλγόριθμος, παρουσιάζει αρκετές αδυναμίες που καθιστούν τη χρήση του περιορισμένη σε ορισμένες εφαρμογές:

- *Εξάρτηση από τα αρχικά centroids:* Τα αρχικά centroids επιλέγονται τυχαία, και αυτό μπορεί να οδηγήσει σε τοπικά βέλτιστα σημεία, αποτρέποντας τον αλγόριθμο από το να βρει τη βέλτιστη λύση για την ομαδοποίηση των δεδομένων.

• *Καθορισμός του k:* Ο αριθμός των clusters πρέπει να είναι προκαθορισμένος, γεγονός που μπορεί να είναι δύσκολο να εκτιμηθεί σωστά χωρίς προηγούμενη γνώση των δεδομένων.

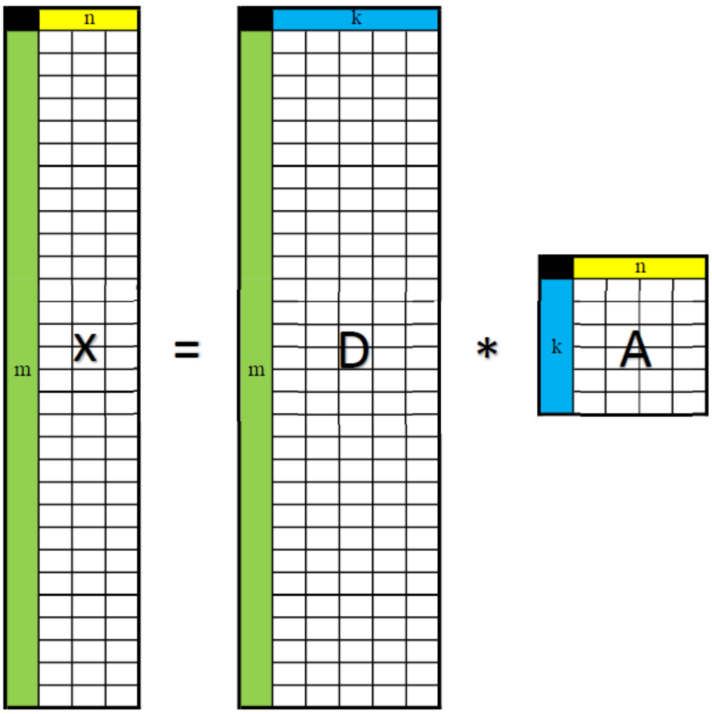
• *Σφαιρική γεωμετρία των clusters:* Ο K-means λειτουργεί καλύτερα όταν τα clusters έχουν σφαιρικό σχήμα και παρόμοιο μέγεθος. Αν τα δεδομένα έχουν ελλειπτικό σχήμα ή διαφορετική πυκνότητα, ο αλγόριθμος μπορεί να αποδώσει κακά αποτελέσματα.

Λόγω αυτών των αδυναμιών, συχνά προκύπτει η ανάγκη για πιο εξελιγμένες μεθόδους, που μπορούν να αντιμετωπίσουν περιορισμούς όπως η σπανιότητα (sparsity) των δεδομένων, να εντοπίσουν αντιπροσώπους πιο αποτελεσματικά και να λειτουργούν καλά σε πολύπλοκα datasets εικόνων και βίντεο. Αυτές οι τεχνικές στοχεύουν στην καλύτερη απόδοση σε περιβάλλοντα όπου τα δεδομένα είναι πολυδιάστατα και τα clusters είναι λιγότερο σαφώς καθορισμένα.

2 ΚΕΦΑΛΑΙΟ 2ο : Αραιή Αναπαράσταση και Κυρτή Βελτιστοποίηση

Στο παρόν κεφάλαιο, θα εξετάσουμε τις θεμελιώδεις έννοιες και τεχνικές που αποτελούν τη βάση του αλγορίθμου SMRS, ο οποίος έχει σχεδιαστεί για την εύρεση αντιπροσώπων σε μεγάλα σύνολα δεδομένων. Ειδικότερα, θα επικεντρωθούμε στις έννοιες της αραιής αναπαράστασης (sparse representation), της εκμάθησης λεξικού (dictionary learning) και της κυρτής βελτιστοποίησης (convex optimization), οι οποίες αποτελούν βασικές έννοιες στη διαδικασία αυτή. Η αραιή αναπαράσταση αναφέρεται στη διαδικασία αναπαράστασης δεδομένων χρησιμοποιώντας τον ελάχιστο δυνατό αριθμό μη μηδενικών συντελεστών από ένα λεξικό, επιτρέποντας έτσι την καλύτερη κατανόηση και συμπίεση των δεδομένων. Το λεξικό αποτελεί ένα σύνολο βασικών συνιστωσών (ή atoms) που χρησιμοποιούνται για την αναπαράσταση των δεδομένων. Η εκμάθηση λεξικού (dictionary learning) είναι η διαδικασία κατά την οποία οι βασικές συνιστώσες μαθαίνονται από τα δεδομένα, ώστε να επιτυγχάνεται η αραιή αναπαράστασή τους. Η κυρτή βελτιστοποίηση (convex optimization) είναι η μαθηματική διαδικασία επίλυσης ενός προβλήματος βελτιστοποίησης όπου η συνάρτηση κόστους είναι κυρτή, διασφαλίζοντας έτσι την ύπαρξη ενός μοναδικού βέλτιστου σημείου. Αυτές οι τεχνικές συνδυάζονται στον αλγόριθμο SMRS που χρησιμοποιούμε στην παρούσα διπλωματική εργασία για την αποτελεσματική επιλογή ενός υποσυνόλου δεδομένων που είναι αντιπροσωπευτικό ολόκληρου του συνόλου. Με την ανάλυση αυτών των εννοιών, θα κατανοήσουμε καλύτερα πώς ο αλγόριθμος εκμεταλλεύεται τις ιδιότητες της αραίωσης και της κυρτότητας για την επίτευξη των στόχων του.

2.1 Αραιή Αναπαράσταση και εκμάθηση λεξικού



Εικόνα 2.1 Αραιή αναπαράσταση και εκμάθηση λεξικού

Στην Εικόνα 2.1 απεικονίζεται με γραφικό τρόπο η διαδικασία της αραιής αναπαράστασης για την εύρεση αντιπροσώπων από σύνολα δεδομένων. Στο παράδειγμα αυτό, ο πίνακας X είναι n στηλών και m γραμμών και αντιπροσωπεύει ένα σύνολο δεδομένων από το οποίο θέλουμε να επιλέξουμε αντιπροσωπευτικά δείγματα. Το λεξικό D με τη σειρά του, αποτελείται από k στήλες (ή αλλιώς atoms) και m γραμμές αντίστοιχα και αντιπροσωπεύει το υποσύνολο των πιο αντιπροσωπευτικών δεδομένων που επιλέγονται από το σύνολο X. Αντί να είναι καινούργια κατασκευασμένα δεδομένα, όπως σε διάφορους άλλους αλγορίθμους που προαναφέραμε στην παρούσα διπλωματική εργασία, το D προκύπτει από τα ίδια τα δεδομένα, επιλέγοντας τις στήλες που είναι πιο χαρακτηριστικές και μπορούν να περιγράψουν τα υπόλοιπα δεδομένα. Ο πίνακας A, είναι ο πίνακας των αραιών συντελεστών δηλαδή αντιστοιχεί στους συντελεστές που υποδεικνύουν πώς κάθε στήλη του X μπορεί να γραφτεί ως γραμμικός συνδυασμός των αντιπροσώπων που περιέχονται στον D. Ο πίνακας αυτός είναι αραιός, καθώς μόνο λίγοι αντιπρόσωποι από το D χρειάζονται για να αναπαραστήσουν το κάθε δείγμα του X. Το πλήθος των γραμμών k του A είναι ίσο με το πλήθος των στηλών k του λεξικού D και το πλήθος των στηλών n του A είναι ίσο με το πλήθος των στηλών n του X . Ο πίνακας X αναπαρίσταται ως ένας γραμμικός συνδυασμός στοιχείων αντιπροσώπων από το D, με αραιούς συντελεστές που δίνονται από το A. Ο πίνακας D επιλέγει τα αντιπροσωπευτικά δείγματα από το X, ενώ ο πίνακας A δείχνει πώς κάθε δείγμα X αναπαρίσταται μέσω του υποσυνόλου αυτών των αντιπροσώπων. Επιπλέον, ο πίνακας A είναι αραιός (sparse) δηλαδή έχει πολλά μηδενικά για να δείχνει ότι κάθε δείγμα από το X αναπαρίσταται μόνο από λίγους αντιπροσώπους από το D.

2.1.1 Αραιή Αναπαράσταση

Ειδικότερα, η αραιή αναπαράσταση (sparse representation) είναι μια τεχνική που βοηθά στην επίλυση προβλημάτων υπολογιστικής όρασης [33]. Με τον όρο αραιή αναπαράσταση εννοείται η αναπαράσταση ενός σήματος με όσο το δυνατόν λιγότερους σημαντικούς μη μηδενικούς συντελεστές, δηλαδή συντελεστές που συμβάλλουν ουσιαστικά στη δημιουργία του σήματος [34], [35]. Η συγκεκριμένη τεχνική είναι από τις πιο δημοφιλείς στο πεδίο της επεξεργασίας εικόνας, της υπολογιστικής όρασης και της μηχανικής μάθησης. Η αραιή αναπαράσταση είναι σημαντική σε πολλές εφαρμογές, όπως η συμπίεση δεδομένων, η ανίχνευση και αναγνώριση προτύπων, και η επεξεργασία εικόνων. Επιτρέπει την αντιμετώπιση του προβλήματος της υψηλής διάστασης και της πληθώρας δεδομένων, μειώνοντας την πολυπλοκότητα της αναπαράστασης και εξάγοντας τα κυρίαρχα χαρακτηριστικά [36]. Η θεωρία της αραιής αναπαράστασης επικεντρώνεται στον τρόπο εύρεσης των κατάλληλων βασικών συνιστωσών (atoms) και στον τρόπο αναπαράστασης των δεδομένων με βάση αυτές τις συνιστώσες. Συνολικά, η αραιή αναπαράσταση αποτελεί ένα σημαντικό εργαλείο στον τομέα της επεξεργασίας σήματος και εικόνας, επιτρέποντας την αποτελεσματική αναπαράσταση και εξαγωγή χρήσιμης πληροφορίας από πολύπλοκα δεδομένα [36]. Η αραιή αναπαράσταση στον αλγόριθμο SMRS που χρησιμοποιείται στο πειραματικό μέρος της παρούσας διπλωματικής εργασίας αναφέρεται στη διαδικασία επιλογής αντιπροσωπευτικών δειγμάτων από ένα σύνολο δεδομένων με τρόπο που να διατηρείται η ποικιλότητα και η πληρότητα της αρχικής πληροφορίας. Ο αλγόριθμος SMRS επιδιώκει να εντοπίσει ένα υποσύνολο από τα δεδομένα που μπορεί να αναπαραστήσει το σύνολο με τρόπο όσο το δυνατόν πιο αραιό, εξασφαλίζοντας παράλληλα τη διατήρηση της δομής των δεδομένων [13], [38].

Η συνάρτηση κόστους που χρησιμοποιεί ο αλγόριθμος SMRS για την επιλογή των αντιπροσώπων περιγράφεται ως εξής:

$$\min \lambda \|C\|_{1,q} + \frac{1}{2}\|Y - DC\|_F^2 \;\; s.t. 1^T C = 1^T \quad \text{Εξ. 2.1}$$

Όπου Y είναι ο πίνακας των αρχικών δεδομένων, C είναι ο πίνακας συντελεστών αναπαράστασης που δείχνει πόσο κάθε δείγμα Y μπορεί να αναπαρασταθεί από τα υπόλοιπα, λ είναι η παράμετρος κανονικοποίησης δηλαδή είναι ο συντελεστής που ρυθμίζει την συμπεριφορά του όρου $\|C\|_{1,q}$ δηλαδη ελέγχει τη συνεισφορά του όρου αραιότητας στην συνολική συνάρτηση κόστους . Όπου $\|C\|_{1,q}$ είναι ο ορος που περιγραφει την αραιοτητα και είναι η $l_{1,q}$-νορμα του πινακα συντελεστων C [39]. Συγκεκριμένα, η $l_{1,q}$-νορμα είναι ο συνδιασμος της l_1-νορμας, η οποια για έναν πινακα C υπολογίζεται ως το άθροισμα των απόλυτων τιμών όλων των στοιχείων του πίνακα ($|x|_1 = \sum_{r=1}^{n}|x_r|$) και ουσιαστικα “τιμωρει” τις μεγαλες τιμες των συντελεστων και ωθει πολλους από αυτους να γινουν μηδενικοι ή πολλοι μικροί, και της l_q-νορμας οπου όταν συνδιαζονται (δηλαδη στην μορφη $l_{1,q}$-νορμα) το q μπορεί να ρυθμίσει πως θα επιβληθεί η αραιότητα σε ομάδες στοιχείων του πίνακα C, εξασφαλίζονται έτσι ότι ορισμένες στήλες (ή γραμμές) του πίνακα C θα έχουν περισσότερα μη μηδενικά στοιχεία από άλλες [40]. Ο όρος $\frac{1}{2}\|Y - DC\|_F^2$ εκφραζει το σφαλμα ανακατασκευης του αρχικου πινακα δεδομενων Y.

2.1.2 Εκμάθηση Λεξικού

Η διαδικασία της εκμάθησης λεξικού (Dictionary Learning) αναφέρεται σε μια μη εποπτευόμενη μέθοδο μάθησης που αποσκοπεί στην εξαγωγή ενός "ιδανικού" λεξικού για την αραιή αναπαράσταση δεδομένων [41]. Στην αραιή αναπαράσταση, το λεξικό αποτελείται από ένα σύνολο "βασικών" στοιχείων (atoms) τα οποία χρησιμοποιούνται για την αναπαράσταση των δεδομένων [26]. Κάθε δείγμα δεδομένων μπορεί να εκφραστεί ως ένας γραμμικός συνδυασμός αυτών των βασικών στοιχείων, με την προϋπόθεση ότι η αναπαράσταση θα είναι αραιή, δηλαδή θα περιλαμβάνει μόνο λίγα μη μηδενικά στοιχεία [42]. Συνεπώς, το λεξικό ορίζει τον χώρο στον οποίο τα δεδομένα μας "ζουν" και τον τρόπο με τον οποίο αναπαρίστανται [41]. Η εκμάθηση λεξικού περιλαμβάνει συνήθως την αναζήτηση του πιο κατάλληλου λεξικού που προσαρμόζεται στα δεδομένα μέσω της ελαχιστοποίησης κάποιου κριτηρίου, όπως το σφάλμα ανακατασκευής ή το σφάλμα αραιής αναπαράστασης, χρησιμοποιώντας αλγορίθμους βελτιστοποίησης [26], [43]. Η εκμάθηση λεξικού εφαρμόζεται σε πολλούς τομείς, όπως στην επεξεργασία εικόνας και σήματος, κειμένου, ήχου, αναγνώριση προτύπων, αποθορυβοποίηση και άλλα, επιτρέποντας μια αποτελεσματική και συμπαγή αναπαράσταση δεδομένων, κάτι που είναι ιδιαίτερα χρήσιμο για πολύπλοκα και υψηλής διάστασης δεδομένα [13], [41].

Η πρώτη φάση για την επίλυση του προβλήματος της αραιής αναπαράστασης περιλαμβάνει την δημιουργία του λεξικού [26]. Αυτό το λεξικό μπορεί είτε να είναι ήδη προκαθορισμένο, περιλαμβάνοντας ένα σύνολο από διανύσματα που έχουν επιλεγεί εκ των προτέρων, είτε να προκύπτει μέσω μιας διαδικασίας εκμάθησης, όπου το περιεχόμενό του προσαρμόζεται στα διαθέσιμα σήματα [42]. Η χρήση ενός προκαθορισμένου λεξικού είναι απλή και μπορεί να οδηγήσει σε ταχύτερους αλγορίθμους, όπως ο γρήγορος μετασχηματισμός Fourier [24]. Ωστόσο, η αποδοτικότητά του εξαρτάται από το κατά πόσο τα δεδομένα ταιριάζουν με το λεξικό [42]. Από την άλλη πλευρά, ένα λεξικό που εκπαιδεύεται από τα ίδια τα δεδομένα μπορεί να προσαρμοστεί καλύτερα στις ιδιαιτερότητες των σημάτων, ξεπερνώντας τους περιορισμούς των προκαθορισμένων και γενικών λεξικών [41], [42]. Η επιλογή του κατάλληλου λεξικού είναι κρίσιμη για την επίτευξη αραιών αναπαραστάσεων και εξαρτάται από τις ιδιότητες των δεδομένων και τους στόχους της αναπαράστασης [43].

Από μαθηματικής σκοπιάς, ο όρος «λεξικό» στην αραιή αναπαράσταση αναφέρεται σε έναν πίνακα $D \in R^{m\times k}$, δηλαδή έναν πίνακα που ανήκει σε έναν χώρο διαστάσεων m x k [26]. Οι στήλες αυτού του πίνακα, γνωστές ως atoms d_j (οπου j = 1,2,….,k), αποτελούν τα διανύσματα αναφοράς μέσα σε αυτόν τον χώρο [26], [42]. Στην γενική περίπτωση, το λεξικό είναι υπερπλήρες, δηλαδή m < k, κάτι που σημαίνει ότι το λεξικό περιέχει περισσότερα atoms από τις διαστάσεις του χώρου m στον οποίο ζουν τα δεδομένα [26]. Αυτό επιτρέπει την αναπαράσταση ενός δεδομένου σημείου y_i ως γραμμικό συνδυασμό ενός μικρού αριθμού από αυτά τα atoms [42], δηλαδή:

$$y_i = Dx_i = \sum_{j=1}^{k} x_{ij} d_j \quad \text{Εξ. } 2.2$$

Όπου x_i είναι ένας αραιός διανυσματικός συντελεστής (συνήθως με λίγες μη μηδενικές τιμές) [42].

Ωστόσο, στον αλγόριθμο SMRS, το λεξικό D δεν χρειάζεται να είναι υπερπλήρες [13]. Αντιθέτως, είναι υποπλήρες και αποτελείται από τα ίδια τα δεδομένα $Y \in R^{m\times N}$, όπου κάθε στήλη του πίνακα Y αντιπροσωπεύει ένα δεδομένων y_i [24]. Σε αυτή την περίπτωση, η αναπαράσταση ενός δεδομένου ως γραμμικός συνδυασμός των υπόλοιπων δεδομένων εκφράζεται ως:

$$y_i = Yc_i \quad \text{Εξ. } 2.3$$

όπου c_i είναι ο αντίστοιχος συντελεστής αναπαράστασης που υποδεικνύει την συμβολή κάθε σημείου δεδομένων y_j στην αναπαράσταση του y_i [13].

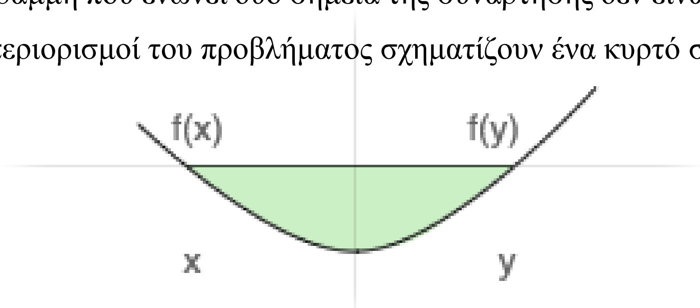
Επιπλέον στην περίπτωση που τα δεδομένα Y χρησιμοποιούνται ως λεξικό, τα atoms είναι απλά τα ίδια τα σημεία δεδομένων άρα ισχύει:

$$Y = YC \quad \text{Εξ. } 2.4$$

Όπου $C \in R^{N\times N}$ είναι ο πίνακας συντελεστών που καθορίζει ποια σημεία δεδομένων επιλέγονται ως αντιπρόσωποι (representatives) [24]. Ο περιορισμός που επιβάλλεται είναι ότι το C έχει λίγες μη μηδενικές σειρές, που αντιστοιχούν στους αντιπροσώπους [13], [24].

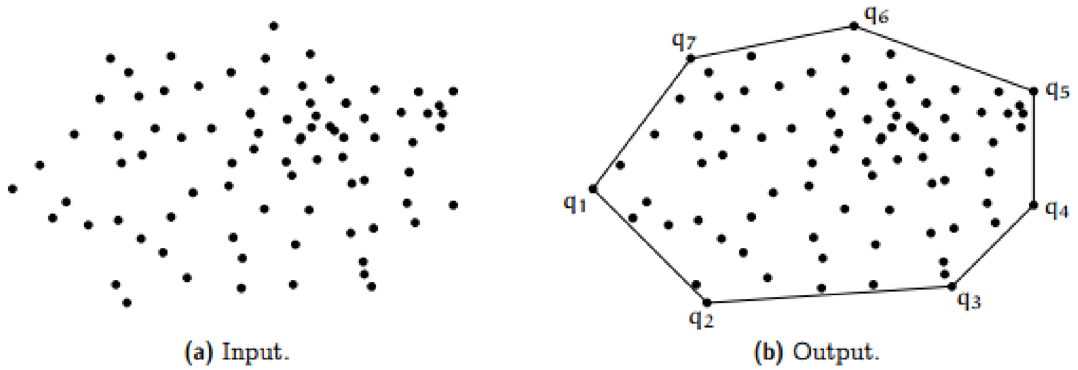
2.2 Κυρτή Βελτιστοποίηση (Convex Optimization)

Η κυρτή βελτιστοποίηση είναι ένας κλάδος της μαθηματικής βελτιστοποίησης και επικεντρώνεται σε προβλήματα όπου έχουν ως στόχο την ελαχιστοποίηση (ή μεγιστοποίηση) μιας κυρτής συνάρτησης. Ένα πρόβλημα βελτιστοποίησης θεωρείται κυρτό όταν η συνάρτηση είναι κυρτή, δηλαδή κάθε γραμμή που ενώνει δύο σημεία της συνάρτησης δεν είναι ποτέ χαμηλότερη από τη συνάρτηση και οι περιορισμοί του προβλήματος σχηματίζουν ένα κυρτό σύνολο [22], [44].



Εικόνα 2.2.1 Κυρτή συνάρτηση

Η κυρτή βελτιστοποίηση μας δίνει την δυνατότητα αποτελεσματικής εύρεσης της λύσης με τη χρήση τεχνικών όπως ο Gradient Descent και ο Alternating Direction Method of Multipliers (ADMM) [22]. Ο ADMM είναι ιδιαίτερα αποτελεσματικός για προβλήματα που μπορούν να διαχωριστούν σε επιμέρους προβλήματα, κάτι που τον καθιστά εξαιρετικά χρήσιμο για τη αραιή αναπαράσταση, όπου πρέπει να γίνει βελτιστοποίηση με πολλούς περιορισμούς [45]. Στην ανάλυση της κυρτής βελτιστοποίησης, μια κεντρική έννοια είναι το κυρτό περίβλημα (convex hull) των στηλών ενός πίνακα δεδομένων.



Εικόνα 2.2.2 Κυρτό περίβλημα

Έστω H το κυρτό περίβλημα των στηλών ενός πίνακα δεδομένων Y. Η βελτιστοποίηση για την επιλογή αντιπροσώπων επιδιώκει να βρει τις κορυφές του κυρτού περιβλήματος που περιγράφουν καλύτερα το σύνολο των δεδομένων. {Αυτό το θέμα αναλύεται μέσω δύο βασικών θεωρημάτων, τα οποία παρέχουν τις βάσεις για την επιλογή αντιπροσώπων με χρήση κυρτής βελτιστοποίησης}. Σύμφωνα με το πρώτο θεώρημα [1], [13], οι μη μηδενικές γραμμές της λύσης του προβλήματος βελτιστοποίησης αντιστοιχούν στις κορυφές του κυρτού περιβλήματος H. Συγκεκριμένα, η βέλτιστη λύση C^* μπορεί να εκφραστεί ως:

$$C^* = \Gamma \begin{bmatrix} I_k & \Delta \\ 0 & 0 \end{bmatrix} \text{ Εξ. 2.5}$$

όπου I είναι ο k-διάστατος μοναδιαίος πίνακας, Δ είναι ένας πίνακας με στοιχεία στο διάστημα [0, 1], και Γ είναι ένας πίνακας μεταθέσεων. Αυτό υποδεικνύει ότι η λύση της βελτιστοποίησης είναι στην πραγματικότητα μια συνδυαστική επιλογή των κορυφών του κυρτού περιβλήματος, οι οποίες λειτουργούν ως αντιπρόσωποι του συνόλου δεδομένων. Επεκτείνοντας αυτή την ιδέα, το δεύτερο θεώρημα [1], [13] εξετάζει την περίπτωση όπου τα δεδομένα βρίσκονται σε έναν (k-1)-διάστατο υπόχωρο του $\mathbb{R}^m$. Η λύση του προβλήματος βελτιστοποίησης σε αυτή την περίπτωση οδηγεί σε ένα υποσύνολο των κορυφών που περιλαμβάνουν τον υπόχωρο όπου βρίσκονται τα δεδομένα. Και σε αυτή την περίπτωση, η βέλτιστη λύση μπορεί να εκφραστεί με παρόμοιο τρόπο, υποδεικνύοντας ότι η επιλογή των αντιπροσώπων μπορεί να επιτευχθεί με ακρίβεια, διατηρώντας τη γεωμετρική δομή του υπόχωρου. Οι τεχνικές κυρτής βελτιστοποίησης που βασίζονται στα παραπάνω θεωρήματα επιτρέπουν την επιλογή αντιπροσώπων με τέτοιο τρόπο ώστε κάθε σημείο δεδομένων να μπορεί να εκφραστεί ως γραμμικός συνδυασμός των επιλεγμένων αντιπροσώπων διατηρώντας παράλληλα τη δομή του χώρου των δεδομένων [46]. Εδώ, ο ADMM έρχεται να ενισχύσει τη διαδικασία, καθώς είναι ικανός να διαχειρίζεται την επίλυση μεγάλων προβλημάτων κυρτής βελτιστοποίησης. Χρησιμοποιώντας τον ADMM, μπορούμε να διασπάσουμε την αρχική βελτιστοποίηση σε μικρότερα, πιο διαχειρίσιμα επιμέρους προβλήματα, τα οποία επιλύονται παράλληλα. Αυτό επιτρέπει γρήγορη σύγκλιση και αποτελεσματική λύση ακόμα και σε μεγάλα σύνολα δεδομένων [45]. Αυτή η τεχνική είναι ιδιαίτερα χρήσιμη σε εφαρμογές όπως η ανάλυση αραιών σημάτων και οι διάφορες μορφές μηχανικής μάθησης, όπου τα δεδομένα συχνά βρίσκονται σε υπόχωρους ενός υψηλότερης διάστασης χώρου [46]. Οι επιλεγμένοι αντιπρόσωποι δεν μειώνουν μόνο την πολυπλοκότητα του προβλήματος, αλλά και διασφαλίζουν την ακριβή αναπαράσταση των δεδομένων. Ο αλγόριθμος SMRS που χρησιμοποιείται στο πειραματικό μέρος της παρούσας διπλωματικής εργασίας γενικά επιλεγεί αντιπροσώπους από ένα σύνολο δεδομένων και ειδικότερα χρησιμοποιεί τις θεωρητικές βάσεις της κυρτής βελτιστοποίησης όπως περιγράφονται από τα παραπάνω θεωρήματα για να επιλέξει αντιπροσώπους που βρίσκονται στο κυρτό περίβλημα των δεδομένων. Συγκεκριμένα, ο αλγόριθμος επιλύει ένα πρόβλημα βελτιστοποίησης το οποίο επιδιώκει να βρει μια αραιή λύση με ελάχιστο αριθμό μη μηδενικών στοιχείων [48]. Σε αυτό το σημείο, ο ADMM ενσωματώνεται για να βελτιστοποιήσει τη διαδικασία επίλυσης του προβλήματος, βελτιώνοντας την απόδοση και την ταχύτητα της σύγκλισης [45]. Η λύση του SMRS έχει την ίδια μορφή με αυτή που περιγράφεται στα θεωρήματα, όπου οι επιλεγμένοι αντιπρόσωποι σχηματίζουν ένα υποσύνολο των κορυφών του κυρτού περιβλήματος των δεδομένων. Αυτή η διαδικασία είναι εξαιρετικά αποδοτική για τη μείωση της διάστασης και την αναπαράσταση μεγάλων συνόλων δεδομένων με μικρότερο αριθμό αντιπροσώπων [47], [49].Η αποδοτικότητα και η ακρίβεια του SMRS στην επιλογή αντιπροσώπων τον καθιστούν έναν από τους πιο αποτελεσματικούς αλγόριθμους για σπάνια αναπαράσταση, επιτρέποντας την ανάλυση μεγάλων και πολύπλοκων συνόλων δεδομένων με μια πιο κατανοητή και συμπαγή μορφή [47].

3 ΚΕΦΑΛΑΙΟ 3ο : Πειραματική διαδικασία

Ο αλγόριθμος SMRS (Sparse Modeling via Representative Selection) αποτελεί μια καινοτόμο μέθοδο για την επιλογή αντιπροσωπευτικών καρέ από βίντεο, επιτρέποντας την αποδοτική και ακριβή αναπαράσταση του περιεχομένου. Βασισμένος στην ιδέα της αραιής μοντελοποίησης, ο SMRS στοχεύει στην επιλογή ενός υποσυνόλου από καρέ που να αντιπροσωπεύει κατά τον καλύτερο δυνατό τρόπο το συνολικό περιεχόμενο του βίντεο, μειώνοντας έτσι τον όγκο πληροφορίας του. Αυτή η προσέγγιση είναι εξαιρετικά σημαντική σε περιπτώσεις όπου ο χρήστης χρειάζεται μια γρήγορη επισκόπηση ή επιθυμεί να εστιάσει στα πιο σημαντικά μέρη ενός βίντεο χωρίς να χρειάζεται να παρακολουθήσει ολόκληρο το υλικό. Η διαδικασία επιλογής βασίζεται στη λογική ότι τα αντιπροσωπευτικά καρέ θα πρέπει να καλύπτουν την ποικιλία των σκηνών και των αλλαγών που συμβαίνουν στο βίντεο, ενώ ταυτόχρονα να διατηρούν μια συμπαγή και συνεκτική αναπαράσταση του περιεχομένου. Ο αλγόριθμος επιτυγχάνει αυτόν τον στόχο μέσω της μοντελοποίησης του προβλήματος ως ένα πρόβλημα επιλογής υποσυνόλου, όπου το σύνολο των επιλεγμένων καρέ αναπαριστά το βίντεο με τον πιο αραιό αλλά και αντιπροσωπευτικό τρόπο. Η αραιότητα αυτή επιτρέπει στον αλγόριθμο να επικεντρώνεται στα πιο σημαντικά καρέ, μειώνοντας τον αριθμό των καρέ που πρέπει να επεξεργαστεί ή να αποθηκεύσει κανείς, ενώ παράλληλα διατηρεί την πληροφορία που είναι κρίσιμη για την κατανόηση του περιεχομένου. Στην προσέγγιση αυτή, ο ρόλος της κυρτής βελτιστοποίησης είναι κρίσιμος, καθώς επιτρέπει την επίλυση του προβλήματος επιλογής υποσυνόλου με τρόπο που εξασφαλίζει την εύρεση της παγκοσμίου βέλτιστης λύσης. Αυτό συμβάλλει στην ακριβέστερη και αποδοτικότερη επιλογή των αντιπροσωπευτικών καρέ, διασφαλίζοντας ότι το τελικό αποτέλεσμα καλύπτει με πληρότητα και ακρίβεια το περιεχόμενο του βίντεο. Παράλληλα, η εκμάθηση λεξικού παίζει σημαντικό ρόλο στη βελτιστοποίηση της αραιής αναπαράστασης του βίντεο. Μέσω της εκμάθησης λεξικού, ο αλγόριθμος μπορεί να δημιουργήσει ένα προσαρμοσμένο λεξικό από τα καρέ του βίντεο, το οποίο στη συνέχεια χρησιμοποιείται για την επιλογή των πιο αντιπροσωπευτικών καρέ. Αυτό το λεξικό λειτουργεί ως βάση για την αραιή μοντελοποίηση, βελτιώνοντας την ποιότητα της αναπαράστασης και μειώνοντας την πληροφορία που πρέπει να επεξεργαστεί. Στο πειραματικό μέρος της παρούσας διπλωματικής εργασίας, ο αλγόριθμος SMRS θα χρησιμοποιηθεί για την εξαγωγή αντιπροσωπευτικών καρέ από δύο διαφορετικά βίντεο. Μέσα από τη διαδικασία αυτή, θα αναλυθεί η αποτελεσματικότητα του αλγορίθμου στην επιλογή καρέ που αναπαριστούν με ακρίβεια το περιεχόμενο, εξετάζοντας παραμέτρους όπως η πληρότητα της αναπαράστασης, η ποιότητα των επιλεγμένων καρέ καθώς και η ταχύτητα της διαδικασίας εξαγωγής. Τα αποτελέσματα που θα προκύψουν αναμένεται να επιβεβαιώσουν την αξία του SMRS ως εργαλείου για την αποδοτική επεξεργασία και ανάλυση βίντεο, καθιστώντας τον έναν σημαντικό αλγόριθμο στην περιοχή της υπολογιστικής όρασης και της ανάλυσης πολυμέσων.

3.1 Χρήση προγραμματιστικού περιβάλλοντος Matlab

Το Matlab είναι ένα εργαλείο προσαρμοσμένο στις ανάγκες μηχανικών και επιστημόνων. Επιλέχθηκε ως προγραμματιστικό περιβάλλον για την ανάπτυξη του κώδικα της παρούσας διπλωματικής εργασίας, καθώς προσφέρει ενσωματωμένα εργαλεία που καλύπτουν όλες τις απαιτήσεις του έργου. Το Matlab διαθέτει βιβλιοθήκες για την επεξεργασία βίντεο και εικόνων, κάτι που διευκολύνει την υλοποίηση του αλγορίθμου. Η χρήση του είναι ιδανική για τον προγραμματισμό, αφού η μαθηματική του προσέγγιση επιτρέπει εύκολους υπολογισμούς με αλγεβρικούς πίνακες, οι οποίοι είναι θεμελιώδεις για την επεξεργασία εικόνας και σήματος. Εναλλακτικά προγραμματιστικά περιβάλλοντα θα απαιτούσαν περισσότερο χρόνο για την ενσωμάτωση των απαραίτητων βιβλιοθηκών και την επικύρωση του κώδικα. Επιπλέον, το Matlab επιτρέπει την αυτόματη μετατροπή του κώδικα σε C/C++ μέσω του MATLAB Coder, διευκολύνοντας την εφαρμογή του αλγορίθμου σε άλλες πλατφόρμες.

3.2 Προ-επεξεργασία δεδομένων

Γενικά, η διαδικασία για την εφαρμογή του αλγορίθμου SMRS σε δεδομένα (είτε πρόκειται για βίντεο, εικόνες ή άλλους τύπους δεδομένων) περιλαμβάνει τα παρακάτω βασικά βήματα:

- Αναπαράσταση δεδομένων ως διανύσματα: Τα δεδομένα (όπως καρέ από βίντεο ή εικόνες) πρέπει να αναπαρίστανται ως διανύσματα. Κάθε δείγμα (π.χ., κάθε καρέ ή εικόνα) μετατρέπεται σε μία μονοδιάστατη στήλη διανύσματος.
- Σχηματισμός πίνακα δεδομένων: Ο πίνακας δεδομένων, που εισάγεται στον αλγόριθμο SMRS, αποτελείται από στήλες, όπου κάθε στήλη είναι το διάνυσμα ενός δείγματος. Έτσι, αν υπάρχουν N δείγματα, ο πίνακας θα έχει N στήλες και κάθε στήλη θα έχει μήκος ίσο με τον αριθμό των χαρακτηριστικών κάθε δείγματος (π.χ. ο αριθμός των pixel σε μια εικόνα).

3.2.1 Επεξεργασία Βίντεο

Για την υλοποίηση και επίδειξη της χρησιμότητας των ανωτέρω μεθόδων, έγινε χρήση του βίντεο Society Raffles για την εξαγωγή των πιο αντιπροσωπευτικών καρέ (δηλαδή των αντιπροσώπων) από ολόκληρο το βίντεο. Το βίντεο χωρίζεται σε μεμονωμένα καρέ χρησιμοποιώντας μια βιβλιοθήκη όπως το VideoReader για να διαβαστούν τα καρέ από το αρχείο βίντεο. Καθώς εφαρμόζεται ο αλγόριθμος SMRS κρίνεται αναγκαίο να μετατρέπει κάθε καρέ από έγχρωμο σε ασπρόμαυρο (grayscale). Αυτό απλοποιεί τα δεδομένα μας από το βίντεο μειώνοντας τη διάσταση από τα τρία κανάλια χρώματος (κόκκινο, πράσινο, μπλε) σε μία. Στην συνέχεια, κάθε καρέ μετατρέπεται σε διάνυσμα, δηλαδή η δισδιάστατη εικόνα (πλάτος × ύψος) μετατρέπεται σε μία μονοδιάστατη στήλη διανύσματος. Αυτά τα διανύσματα θα είναι οι στήλες του πίνακα εισόδου στον αλγόριθμο SMRS. Ο τελικός πίνακας εισόδου αποτελείται από στήλες, όπου κάθε στήλη αντιπροσωπεύει ένα καρέ βίντεο ως διάνυσμα. Έτσι, κάθε στήλη του πίνακα περιέχει τα pixel του αντίστοιχου καρέ. Ο αριθμός των καρέ στο βίντεο θα αντιστοιχεί στις στήλες του πίνακα και το μήκος κάθε διανύσματος (αριθμός pixel σε ένα καρέ) καθορίζει τη διάσταση του πίνακα.

3.2.2 Επεξεργασία Εικόνων

Επιπλέον, για την εφαρμογή του αλγορίθμου SMRS σε ένα dataset εικόνων η διαδικασία είναι παρόμοια με αυτή που χρησιμοποιείται για βίντεο, με μερικές διαφορές που σχετίζονται με το γεγονός ότι χρησιμοποιούνται μεμονωμένες εικόνες αντί για καρέ βίντεο. Αρχικά οι εικόνες που χρησιμοποιήθηκαν στην παρούσα διπλωματική εργασία πάρθηκαν από το JPEG dataset του MNIST (Scolianni,2018) από το Kaggle [50] το οποίο περιέχει χειρόγραφα ψηφία του αριθμού τρία σε μορφή jpeg. Η επιλογή αυτού έγινε λόγω της ευρείας χρήσης του σε πειράματα μηχανικής μάθησης και ως εκ τούτου θεωρήθηκε ως μια καλή ευκαιρία για δοκιμή του αλγορίθμου SMRS για την εύρεση αντιπροσωπευτικών εικόνων. Στην προκείμενη περίπτωση οι εικόνες είναι ήδη σε ασπρόμαυρη μορφή (grayscale) οπότε δεν είναι αναγκαία η μετατροπή τους σε μια διάσταση. Ακολούθως, κάθε εικόνα μετατρέπεται σε ένα μονοδιάστατο διάνυσμα (στήλη), όπου τα pixel της εικόνας αναδιατάσσονται από 2D (ύψος × πλάτος) σε 1D (διάνυσμα pixel). Αυτά τα διανύσματα θα αποτελέσουν τις στήλες του πίνακα δεδομένων. Για τον πίνακα δεδομένων, κάθε εικόνα αντιστοιχεί σε μία στήλη του πίνακα εισόδου Y. Ο αριθμός των εικόνων N στο dataset θα αντιστοιχεί στον αριθμό των στηλών, και το μήκος κάθε στήλης (ο αριθμός των στοιχείων στο διάνυσμα) θα είναι ίσο με τον αριθμό των pixel στην εικόνα.

3.3 Αλγόριθμος SMRS

Γενικά ο αλγόριθμος Sparse Modeling via Representative Selection (SMRS) αποτελεί μια εξελιγμένη τεχνική για την επιλογή αντιπροσωπευτικών δειγμάτων από μεγάλα σύνολα δεδομένων. Στόχος του είναι να επιλέξει ένα μικρό υποσύνολο δειγμάτων, το οποίο μπορεί να αναπαραστήσει επαρκώς ολόκληρο το σύνολο δεδομένων με τη μέγιστη ακρίβεια και την ελάχιστη απώλεια πληροφορίας. Στη θεωρητική του βάση, ο SMRS εκμεταλλεύεται την έννοια της αραιής αναπαράστασης για να βρει έναν μικρό αριθμό γραμμικά ανεξάρτητων δειγμάτων που δημιουργούν έναν υπόχωρο, ο οποίος περιγράφει το ολόκληρο το σύνολο δεδομένων. Εδώ ακριβώς είναι που ο αλγόριθμος SMRS διαπρέπει διότι απομονώνει τα σημαντικά καρέ (ή εικόνες) που είναι αναγκαία για την περιγραφή του πλήρους συνόλου δεδομένων. Ειδικότερα, η εφαρμογή του SMRS σε δεδομένα βίντεο επιτρέπει την αυτόματη εξαγωγή των πιο σημαντικών καρέ από ένα βίντεο, κάτι που μπορεί να βρει χρήσεις σε πληθώρα εφαρμογών, όπως η συμπίεση βίντεο, η ανάλυση σκηνών και η αναγνώριση μοτίβων. Για παράδειγμα, σε ένα βίντεο παρακολούθησης, ο SMRS μπορεί να εντοπίσει τα πιο σημαντικά καρέ όπου συμβαίνει κίνηση ή δραστηριότητα, αποφεύγοντας τα καρέ που είναι στατικά ή δεν περιέχουν σημαντικές πληροφορίες. Παρόμοια, στην επεξεργασία εικόνων μπορεί να χρησιμοποιηθεί για την εξαγωγή των πιο σημαντικών εικόνων, μειώνοντας έτσι το υπολογιστικό κόστος σε εφαρμογές αναγνώρισης εικόνων και ανάλυσης περιεχομένου. Ο αλγόριθμος SMRS προσφέρει επίσης τη δυνατότητα να χρησιμοποιηθεί είτε απευθείας στα αρχικά δεδομένα, είτε μετά από την χρήση της PCA, βελτιώνοντας την αποδοτικότητα και την ταχύτητα της ανάλυσης. Αυτό παρέχει στον χρήστη ευελιξία στον τρόπο που προσεγγίζει την ανάλυση των δεδομένων, προσαρμόζοντας τον αλγόριθμο στις ανάγκες της κάθε εφαρμογής. Συνολικά, ο αλγόριθμος SMRS είναι ένα ισχυρό εργαλείο για την ανάλυση δεδομένων υψηλών διαστάσεων, ιδιαίτερα σε εφαρμογές βίντεο και εικόνας, όπου η αντιπροσωπευτικότητα είναι καίριας σημασίας για την εξαγωγή αξιόπιστων και χρήσιμων πληροφοριών με ελάχιστο υπολογιστικό κόστος. Παρακάτω φαίνονται τα βήματα του αλγορίθμου SMRS:

Είσοδος:

$Y \epsilon \mathbb{R}^{DxN}: Data\ matrix$ \\ $Πίνακας\ δεδομένων$

$οπου\ D\ η\ διάσταση\ των\ δεδομένων\ και\ N\ το\ πλήθος\ των\ δειγμάτων$

$\lambda: regularization\ parameter$ \\ $Δείκτης\ κανονικοποίησης$

Έξοδος:

$repInd$ \\ $Δείκτες\ των\ αντιπροσώπων$

$C \epsilon \mathbb{R}^{NxN}$ \\ $Πίνακας\ συντελεστών,$

$όπου\ οι\ μη\ μηδενικές\ γραμμές\ αντιστοιχούν\ στους\ αντιπροσώπους$

SMRS αλγόριθμος:

$Y = Y - \frac{1}{N}\sum_{i=1}^{N} Y_i$ \\ $Κανονικοποίηση\ δεδόμενων$

\\ $Υπολογισμός\ Πίνακα\ συντελεστών$

$\operatorname*{Min}_{C} \|Y - YC\|_F^2 + \lambda \|C\|_{1,q}\ οπου\ \|C\|_{1,q} = \sum_{i=1}^{N} \|C_i\|_q$

\\ $Ενημέρωση\ C$

$Z = A(C + \lambda)\ \ όπου\ A = (\mu_1 Y^T Y + \mu_2 I)^{-1}\ με\ \mu_1 \&\ \mu_2\ penalty\ parameters$ $και\ I\ ο\ μοναδιαίος\ πίνακας\ \ )$

$C = shrink\left(Z, \frac{1}{\lambda}, q\right)$

$\|Z - C\|_F^2 < \varepsilon\ όπου\ \varepsilon\ ειναι\ το\ κατώφλι\ σύγκλισης$

\\ $Επιλογή\ αντιπροσώπων$

$Οι\ αντιπρόσωποι\ ειναι\ οι\ γραμμές\ του\ C\ με\ τις\ μεγαλύτερες\ τιμές$:

$$r(i) = \|C_i\|_q$$

\\ $Αφαίρεση\ κοντινών\ αντιπροσώπων$

$d(i,j) = \left\|Y_i - Y_j\right\|_2$

\\ $Επιστροφή\ αποτελεσμάτων$

Η shrink εκτελεί ανάλογα την τιμή του q τις νόρμες l_1, l_2, l_∞(δηλαδή $q = 1 \rightarrow l_1, q = 2 \rightarrow l_2, q = \infty \rightarrow\ l_\infty$)

Στην συνέχεια εξηγούνται οι βασικές συναρτήσεις που έχουν δημιουργηθεί στο προγραμματιστικό περιβάλλον του Matlab για την υλοποίηση του αλγορίθμου SMRS.

3.3.1 Μέθοδος ADMM

Η συνάρτηση almLasso_mat_func υλοποιεί τον αλγόριθμο βελτιστοποίησης ADMM για την επίλυση του προβλήματος ελαχιστοποίησης της κανονικοποίησης L_1/L_q. Στόχος της ελαχιστοποίησης είναι η επιβολή της αραιότητας στην αναπαράσταση του πίνακα εισόδου Y. Η συνάρτηση στοχεύει στην ελαχιστοποίηση ενός συνδυασμού όρων - ένας όρος μετρά το σφάλμα ανακατασκευής (το πόσο καλά μπορούν να αναπαρασταθούν τα αρχικά δεδομένα από λίγα κύρια καρέ ή εικόνες) και ο άλλος όρος επιβάλλει σπανιότητα στον πίνακα συντελεστών C. Η L_1 - νόρμα (άθροισμα των απόλυτων τιμών) χρησιμοποιείται για να προωθήσει τη σπανιότητα στον πίνακα C, γεγονός που διασφαλίζει ότι οι περισσότεροι συντελεστές είναι μηδενικοί, οδηγώντας στον εντοπισμό μόνο λίγων αντιπροσωπευτικών δεδομένων. Η L_q - νόρμα χρησιμοποιείται για να μετρηθεί το συνολικό μέγεθος των συντελεστών. Ανάλογα με την τιμή του q, αλλάζει η συμπεριφορά της ποινής: για παράδειγμα, η L_2 - νόρμα επιβάλλει αυστηρότερη ποινή στις μεγαλύτερες τιμές σε σύγκριση με την L_1. Σε κάθε επανάληψη, ο αλγόριθμος εναλλάσσεται μεταξύ της ενημέρωσης διαφόρων μεταβλητών (Z, C, Lambda2) που αντιπροσωπεύουν τη διαδικασία βελτιστοποίησης. Ο πίνακας Z ενημερώνεται για να ελαχιστοποιηθεί το σφάλμα μεταξύ των αρχικών δεδομένων και της χαμηλής διάστασης προσέγγισής τους, ενώ ο πίνακας C ενημερώνεται για να διατηρηθεί η σπανιότητα. Οι πολλαπλασιαστές Lagrange (Lambda2) βοηθούν στη διατήρηση των περιορισμών που επιβάλλει το πρόβλημα, διασφαλίζοντας ότι οι ενημερώσεις παραμένουν συνεπείς. Μετά από κάθε επανάληψη, ο αλγόριθμος υπολογίζει το σφάλμα (err1, err2) για να καθορίσει εάν οι ενημερώσεις έχουν ελαχιστοποιήσει επαρκώς τη συνάρτηση κόστους. Αν το σφάλμα είναι κάτω από ένα προκαθορισμένο όριο (thr), οι επαναλήψεις σταματούν. Διαφορετικά, η διαδικασία συνεχίζεται μέχρι να επιτευχθεί ο μέγιστος αριθμός επαναλήψεων (maxIter). Αυτή η μέθοδος ADMM είναι απαραίτητη γιατί χειρίζεται αποτελεσματικά την ελαχιστοποίηση των L_1/L_q, που είναι δύσκολο να βελτιστοποιηθούν άμεσα. Η ADMM διασπά το πρόβλημα σε μικρότερα υποπροβλήματα που είναι ευκολότερα στην επίλυση.

3.3.2 Εντοπισμός Αντιπροσωπευτικών Καρέ

Η συνάρτηση findRep εντοπίζει ποια καρέ του βίντεο (ή εικόνες) θα πρέπει να επιλεγούν ως αντιπρόσωποι αναλύοντας τους αραιούς συντελεστές που προκύπτουν από τη συνάρτηση almLasso_mat_func. Ο βασικός στόχος εδώ είναι η χρήση του αραιού προτύπου του πίνακα C για την επιλογή βασικών καρέ. Για κάθε στήλη στον πίνακα C, η συνάρτηση υπολογίζει την L_q-νόρμα των συντελεστών. Αυτή η νόρμα αντιπροσωπεύει τη σημασία κάθε καρέ στη συνολική αναπαράσταση του βίντεο (ή εικόνας στο συνολικό dataset). Η συνάρτηση στη συνέχεια ταξινομεί τις νόρμες σε φθίνουσα σειρά. Προοδευτικά συσσωρεύει τις συνολικές τιμές των νόρμων έως ότου ένα καθορισμένο ποσοστό (thr, π.χ. 99%) της συνολικής συμβολής έχει καλυφθεί. Τα καρέ που συνεισφέρουν περισσότερο επιλέγονται ως αντιπροσωπευτικά. Η λογική πίσω από αυτό είναι ότι μόνο λίγα καρέ πρέπει να είναι αρκετά για να αναπαραστήσουν ολόκληρο το βίντεο, εάν το βίντεο εμφανίζει πλεονασμό (π.χ. πολλά παρόμοια καρέ). Αυτή η συνάρτηση επιτρέπει στον χρήστη να ελέγξει πόσα καρέ επιλέγονται προσαρμόζοντας το κατώφλι thr.

3.3.3 Αφαίρεση Περιττών Αντιπροσωπευτικών Καρέ

Ακόμη και μετά την επιλογή των πιο σημαντικών καρέ μέσω της συνάρτησης findRep, μπορεί να υπάρχει ακόμα πλεονασμός μεταξύ των επιλεγμένων αντιπροσώπων. Για παράδειγμα, δύο διαδοχικά καρέ (εικόνες) μπορεί να είναι πολύ παρόμοια. Η συνάρτηση rmRep ασχολείται με αυτό το ζήτημα αφαιρώντας τα πλεονάζοντα καρέ (εικόνες), βελτιστοποιώντας περαιτέρω το αντιπροσωπευτικό σύνολο. Η συνάρτηση υπολογίζει την απόσταση μεταξύ των επιλεγμένων αντιπροσωπευτικών καρέ (στήλες του πίνακα Y που αντιστοιχούν στους δείκτες sInd). Η απόσταση μετράτε ως η νόρμα της διαφοράς μεταξύ δύο καρέ. Η συνάρτηση ταξινομεί τις αποστάσεις με φθίνουσα σειρά. Προοδευτικά αφαιρεί καρέ που είναι πολύ παρόμοια μεταξύ τους (δηλαδή, των οποίων η απόσταση είναι κάτω από ένα συγκεκριμένο κατώφλι thr). Αφαιρώντας αυτά τα πλεονάζοντα καρέ, το τελικό σύνολο των αντιπροσωπευτικών καρέ είναι τόσο συνοπτικό όσο και διαφοροποιημένο. Αυτό το βήμα διασφαλίζει ότι τα επιλεγμένα καρέ δεν είναι μόνο αντιπροσωπευτικά, αλλά και αρκετά διαφορετικά μεταξύ τους, αποφεύγοντας την υπερβολική αναπαράσταση οποιασδήποτε σκηνής.

3.3.4 Τελεστής Μείωσης

Η συνάρτηση shrinkL1Lq είναι ένα κρίσιμο μέρος της μεθόδου ADMM, καθώς επιβάλλει τη σπανιότητα στη λύση. Ο τελεστής μείωσης εφαρμόζει ουσιαστικά μια αποκοπή στους συντελεστές στον πίνακα C, δηλαδή μηδενίζει τις μικρότερες τιμές και συρρικνώνει τις μεγαλύτερες. Αυτή η λειτουργία είναι κρίσιμη για την προώθηση της σπανιότητας. Για κάθε συντελεστή στον πίνακα C, η συνάρτηση τον συγκρίνει με ένα κατώφλι lambda. Αν ο συντελεστής είναι μικρότερος από το κατώφλι, τίθεται σε μηδενικό. Αν είναι μεγαλύτερος, μειώνεται κατά την τιμή του κατωφλίου. Αυτή η διαδικασία "συρρικνώνει" τους συντελεστές προς το μηδέν. Ανάλογα με την τιμή του q, η συνάρτηση εφαρμόζει διαφορετική στρατηγική μείωσης. Αυτή η διαδικασία μείωσης είναι ο βασικός μηχανισμός που διασφαλίζει ότι οι περισσότεροι συντελεστές στον πίνακα C παραμένουν μηδενικοί, επιτρέποντας στον αλγόριθμο να επιλέξει μόνο λίγα αντιπροσωπευτικά καρέ ενώ αγνοεί τα υπόλοιπα.

3.3.5 Υπολογισμός της Παραμέτρου Κανονικοποίησης

Η συνάρτηση computeLambda_mat υπολογίζει την παράμετρο κανονικοποίησης lambda. Η παράμετρος κανονικοποίησης ελέγχει την ισορροπία μεταξύ της ακρίβειας ανακατασκευής και της σπανιότητας. Η παράμετρος lambda καθορίζει πόσο έντονα επιβάλλεται ο περιορισμός σπανιότητας: οι μεγαλύτερες τιμές του lambda οδηγούν σε πιο αραιές λύσεις, ενώ οι μικρότερες τιμές επιτρέπουν να εκπροσωπούνται περισσότερα καρέ. Με τον δυναμικό υπολογισμό του lambda με βάση τα δεδομένα εισόδου, ο αλγόριθμος προσαρμόζει την κανονικοποίηση ώστε να ταιριάζει στις ιδιαιτερότητες του συγκεκριμένου βίντεο (ή dataset εικόνων). Αυτό σημαίνει ότι η διαδικασία προσαρμόζεται αυτόματα στις απαιτήσεις των δεδομένων, είτε αυτά περιέχουν έντονες αλλαγές σκηνών είτε πιο στατικές ακολουθίες καρέ.

4 ΚΕΦΑΛΑΙΟ 4º : Αποτελέσματα

4.1 Αποτελέσματα Εξαγωγής Αντιπροσωπευτικών Καρέ από Βίντεο

Όπως έχει αναφερθεί ο κώδικας εφαρμόστηκε αρχικά στο βίντεο Society Raffles με συνολική διάρκεια 48.15 δευτερολέπτων και περιλάμβανε 1443 καρέ (frames), τα οποία φαίνονται ενδεικτικά στην εικόνα 4.1.

Εικόνα 4.1.1 Μερικά Καρέ από το βίντεο Society Raffles

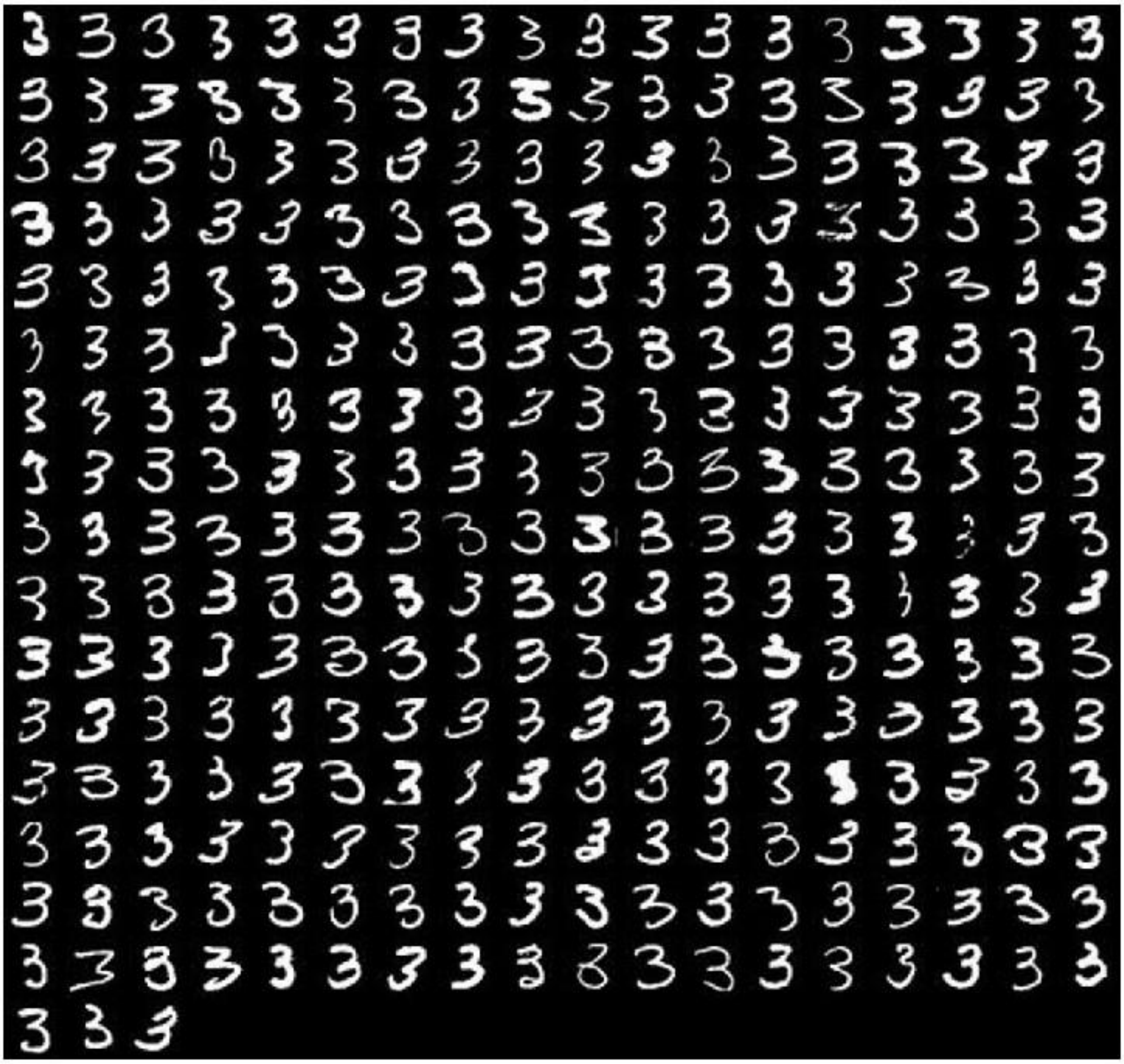
Η εφαρμογή του αλγορίθμου SMRS στο βίντεο από το σύνολο των 1443 καρέ (frames), κατάφερε να επιλέξει 14 αντιπροσωπευτικά καρέ, που αποτύπωσαν με ακρίβεια την πιο σημαντική οπτική πληροφορία από το βίντεο. Τα επιλεγμένα καρέ δεν ήταν απλώς τυχαία δείγματα από το βίντεο, αλλά αποδείχθηκαν εκείνα που περιείχαν τα κύρια συμβάντα ή αλλαγές στη σκηνή. Τα 14 αντιπροσωπευτικά καρέ φαίνονται στην εικόνα 4.2.

Εικόνα 4.1.2 Αντιπροσωπευτικά καρέ από το βίντεο Society Raffles

Επιπλέον, η ανάλυση των αντιπροσωπευτικών καρέ μπορεί να αξιοποιηθεί σε μελλοντικές εφαρμογές, όπως η συμπίεση βίντεο, η αναγνώριση σκηνών, καθώς και η αναγνώριση αντικειμένων ή συμβάντων.

4.2 Αποτελέσματα Εξαγωγής Αντιπροσωπευτικών Εικόνων από Dataset

Στη δεύτερη εφαρμογή του κώδικα, χρησιμοποιήθηκε το JPEG dataset του MNIST (Scolianni,2018) από το Kaggle [50] που περιλαμβάνει 4351 εικόνες του χειρόγραφου αριθμού "3". Ο στόχος ήταν να εντοπιστούν οι πιο σημαντικές και αντιπροσωπευτικές εικόνες από το σύνολο, ώστε να δημιουργηθεί ένα υποσύνολο το οποίο διατηρεί την πολυπλοκότητα και τις διαφοροποιήσεις του αριθμού "3" χωρίς επαναλήψεις ή πλεονασμούς. Αυτό το dataset περιλάμβανε διάφορες παραλλαγές του αριθμού "3", όπως διαφορές στη γραμματοσειρά, το πάχος των γραμμών, και άλλες οπτικές παραλλαγές που εισάγουν διαφοροποίηση στις εικόνες, μερικές από αυτές τις εικόνες φαίνονται στην εικόνα 4.2.1



.

Εικόνα 4.2.1 Μερικές εικόνες του χειρόγραφου αριθμού τρία από το JPEG dataset του MNIST

Ο αλγόριθμος SMRS εφαρμόστηκε στο dataset, και από τις 4351 εικόνες, επέλεξε 33 αντιπροσωπευτικές εικόνες. Αυτές οι 33 εικόνες χαρακτηρίζονται από το γεγονός ότι περιέχουν τις πιο σημαντικές παραλλαγές του αριθμού "3", καθιστώντας τις αντιπροσωπευτικές για το σύνολο του dataset. Η επιλογή έγινε με βάση την αρχή της σπανιότητας, όπου οι εικόνες που περιέχουν μοναδικά χαρακτηριστικά ή σημαντικές διαφοροποιήσεις επιλέχθηκαν ως αντιπροσωπευτικές. Οι εικόνες που ήταν πλεονάζουσες ή περιείχαν μικρές διαφοροποιήσεις απορρίφθηκαν, ελαχιστοποιώντας έτσι το μέγεθος του τελικού dataset. Οι αντιπροσωπευτικές εικόνες φαίνονται στην εικόνα 4.2.2.

Εικόνα 4.2.2 Αντιπροσωπευτικές εικόνες του χειρόγραφου αριθμού τρία

Είναι σημαντικό να σημειωθεί ότι οι αντιπρόσωποι περιλαμβάνουν διαφορετικές παραλλαγές γραφής του αριθμού "3" το οποίο είναι σημαντικό για εφαρμογές μηχανικής μάθησης όπου η επεξεργασία μεγάλων όγκων δεδομένων μπορεί να είναι χρονοβόρα και κοστοβόρα. Με την επιλογή των πιο σημαντικών εικόνων, ο αλγόριθμος SMRS μειώνει τον αριθμό των δεδομένων που χρειάζεται να χρησιμοποιηθούν για εκπαίδευση, βελτιώνοντας την αποδοτικότητα και την ταχύτητα των αλγορίθμων εκπαίδευσης. Η μείωση του μεγέθους του dataset χωρίς απώλεια της πληροφορίας είναι εξαιρετικά χρήσιμη για εφαρμογές όπως η ταξινόμηση εικόνων, η αναγνώριση χαρακτήρων και η ανάλυση γραμματοσειρών, όπου οι διαφοροποιήσεις μεταξύ των εικόνων παίζουν σημαντικό ρόλο. Τα αποτελέσματα από την εφαρμογή του SMRS σε αυτές τις δύο διαφορετικές περιπτώσεις δείχνουν την ευελιξία και την ισχύ του αλγορίθμου στην ανάλυση δεδομένων υψηλών διαστάσεων. Στην περίπτωση του βίντεο, ο αλγόριθμος κατάφερε να εντοπίσει τα πιο σημαντικά καρέ, προσφέροντας μια συνοπτική αναπαράσταση του βίντεο με βάση την πληροφορία που περιείχαν τα καρέ. Στην περίπτωση του dataset εικόνων του χειρόγραφου αριθμού "3", ο SMRS επέλεξε τις πιο αντιπροσωπευτικές εικόνες, μειώνοντας το μέγεθος του dataset χωρίς να χάσει την ποικιλία και τις διαφοροποιήσεις του αριθμού.

5 ΚΕΦΑΛΑΙΟ 5^{ο} : Συμπεράσματα και Προτάσεις

Ο αλγόριθμος SMRS αποδεικνύεται εξαιρετικά αποτελεσματικός στην εξαγωγή αντιπροσωπευτικών δειγμάτων από σύνολα δεδομένων υψηλών διαστάσεων, όπως βίντεο και εικόνες. Μέσω της εφαρμογής του σε διαφορετικά είδη δεδομένων, γίνεται σαφές ότι η ικανότητα του αλγορίθμου να εντοπίζει αντιπροσώπους που ελαχιστοποιούν τον πλεονασμό και διατηρούν την πληροφορία είναι κρίσιμη για εφαρμογές όπως η επεξεργασία βίντεο, η ανάλυση εικόνων, η συμπίεση δεδομένων, αλλά και για τη μείωση διαστάσεων των δεδομένων.

5.1 Πλεονεκτήματα και Κύριες Εφαρμογές

Η εύρεση αντιπροσώπων από μεγάλα σύνολα δεδομένων αποτελεί τη βασική εφαρμογή του SMRS. Ο αλγόριθμος διασφαλίζει ότι το τελικό σύνολο των αντιπροσωπευτικών δειγμάτων περιλαμβάνει όλα τα κρίσιμα χαρακτηριστικά, αποφεύγοντας τα επαναλαμβανόμενα ή πλεονάζοντα δεδομένα. Αυτό είναι ιδιαίτερα χρήσιμο σε εφαρμογές επεξεργασίας βίντεο, όπου η εξαγωγή των πιο σημαντικών καρέ μπορεί να βοηθήσει σε πληθώρα εφαρμογών, από τη συμπίεση βίντεο έως την ανάλυση σκηνών και αντικειμένων. Παρόμοια, στην ανάλυση εικόνων, ο SMRS μπορεί να εντοπίσει τις πιο σημαντικές παραλλαγές μέσα σε σύνολα εικόνων, προσφέροντας μια συνοπτική και ακριβή αναπαράσταση των δεδομένων. Επιπλέον, η μείωση των διαστάσεων είναι μια εξίσου σημαντική εφαρμογή του SMRS. Με την επιλογή ενός μικρού αριθμού αντιπροσωπευτικών δειγμάτων, ο αλγόριθμος μειώνει τον αριθμό των διαστάσεων των δεδομένων χωρίς να χάνεται η πληροφορία. Αυτό έχει κρίσιμη σημασία για εφαρμογές που σχετίζονται με μεγάλα και πολύπλοκα σύνολα δεδομένων, όπως στην μηχανική μάθηση και στα συστήματα μεγάλων δεδομένων (big data).

5.2 Προτάσεις για Μελλοντικές Χρήσεις

Μερικές προτάσεις για τη μελλοντική χρήση του SMRS περιλαμβάνουν τομείς όπως η μηχανική μάθηση και τα συστήματα ανάλυσης μεγάλων δεδομένων (big data) που προαναφέρθηκαν. Στην μηχανική μάθηση, η εξαγωγή αντιπροσωπευτικών δειγμάτων σε συνδυασμό με τη μείωση διαστάσεων μπορεί να βελτιώσει την αποδοτικότητα των αλγορίθμων, επιτρέποντας την εκπαίδευση μοντέλων με λιγότερα αλλά πιο σημαντικά δεδομένα μειώνοντας έτσι το κόστος και τον χρόνο εκπαίδευσης. Επιπλέον, σε τομείς όπως η γενετική ανάλυση, όπου η διαχείριση μεγάλων συνόλων δεδομένων είναι κρίσιμη, η επιλογή αντιπροσωπευτικών δειγμάτων θα μπορούσε να επιταχύνει τη διαδικασία εξέτασης χωρίς να μειώνεται η ακρίβεια. Επίσης, παρόλο που ο SMRS αποδεικνύεται αποδοτικός στην εξαγωγή αντιπροσωπευτικών δειγμάτων και στη μείωση διαστάσεων, θα μπορούσε να γίνει πιο αποτελεσματικός σε εφαρμογές που απαιτούν ανάλυση σε πραγματικό χρόνο. Για παράδειγμα, σε συστήματα real-time παρακολούθησης βίντεο, όπου η επεξεργασία δεδομένων πρέπει να είναι άμεση, η περαιτέρω βελτιστοποίηση του αλγορίθμου για γρηγορότερη επεξεργασία θα ήταν ιδιαίτερα χρήσιμη. Βελτιώσεις στους αλγορίθμους βελτιστοποίησης που χρησιμοποιούνται από τον SMRS, όπως η ταχύτερη σύγκλιση των μεθόδων ADMM θα μπορούσαν να προσφέρουν σημαντική επιτάχυνση της ανάλυσης των δεδομένων. Επιπλέον, η ικανότητα του SMRS να προσαρμόζεται σε διαφορετικά είδη δεδομένων είναι ένα από τα βασικά του πλεονεκτήματα, αλλά υπάρχουν περιθώρια για ακόμα καλύτερη προσαρμογή. Για παράδειγμα, σε εφαρμογές που αφορούν πολυδιάστατα δεδομένα (όπως δεδομένα ιατρικής απεικόνισης ή δεδομένα αισθητήρων), ο αλγόριθμος θα μπορούσε να βελτιωθεί περαιτέρω ώστε να αναγνωρίζει καλύτερα τις διαφοροποιήσεις μεταξύ των διαφορετικών τύπων δεδομένων και να επιλέγει τα πιο σημαντικά χαρακτηριστικά για κάθε περίπτωση. Αυτό θα μπορούσε να επιτευχθεί μέσω μιας πιο εξειδικευμένης επεξεργασίας, όπου ο SMRS θα ενσωματώνει πληροφορίες που αφορούν τα ιδιαίτερα χαρακτηριστικά κάθε τύπου δεδομένων. Καταληκτικά, ο αλγόριθμος SMRS έχει ήδη αποδείξει την αποτελεσματικότητά του στην ανάλυση δεδομένων υψηλών διαστάσεων, ιδιαίτερα σε εφαρμογές που σχετίζονται με την εξαγωγή αντιπροσωπευτικών καρέ ή εικόνων. Ωστόσο, οι δυνατότητές του δεν σταματούν εδώ. Με την κατάλληλη βελτίωση στην αποδοτικότητα και την καλύτερη προσαρμογή του σε διαφορετικά είδη δεδομένων, ο SMRS μπορεί να βρει ακόμα περισσότερες χρήσεις σε σύγχρονα πεδία, προσφέροντας μια ισχυρή λύση στην επεξεργασία και ανάλυση μεγάλων όγκων δεδομένων.

Βιβλιογραφία – Αναφορές - Διαδικτυακές Πηγές

[1] E. Elhamifar, G. Sapiro and R. Vidal, "See all by looking at a few: Sparse modeling for finding representative objects," 2012 IEEE Conference on Computer Vision and Pattern Recognition, Providence, RI, USA, 2012, pp. 1600-1607, doi: 10.1109/CVPR.2012.6247852

[2] I. T. Jolliffe, Principal Component Analysis, Springer-Verlag, 1986.

[3] J. A. Lee and M. Verleysen, "Nonlinear Dimensionality Reduction," Springer, 2007.

[4] C. M. Bishop, Pattern Recognition and Machine Learning, Springer, 2006.

[5] J. Shlens, "A Tutorial on Principal Component Analysis," arXiv preprint, 2014.

[6] I. A. T. Hashem, I. Yaqoob, N. B. Anuar, S. Mokhtar, A. Gani, and S. U. Khan, "The rise of 'big data' on cloud computing: Review and open research issues," Information Systems, vol. 47, pp. 98-115, Jan. 2015.

[7] Y. Bengio, A. Courville, and P. Vincent, "Representation learning: A review and new perspectives," IEEE Trans. Pattern Anal. Mach. Intell., vol. 35, no. 8, pp. 1798–1828, Aug. 2013.

[8] D. Shulman and G. Lewis, "Computational challenges of big data," IEEE Internet Comput., vol. 17, no. 6, pp. 10–13, Nov./Dec. 2013.

[9] M. Götz, J. V. Gael, and S. Feuerriegel, "Efficient Data Representation with Representative Points," IEEE Trans. Knowl. Data Eng., vol. 29, no. 8, pp. 1543–1555, Aug. 2017.

[10] T. Hastie, R. Tibshirani, and J. Friedman, The Elements of Statistical Learning: Data Mining, Inference, and Prediction, 2nd ed. New York: Springer, 2009.

[11] D. Donoho, "Compressed sensing," IEEE Trans. Inf. Theory, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[12] A. Mnih and G. E. Hinton, "A scalable hierarchical distributed language model," in Advances in Neural Information Processing Systems, vol. 21, pp. 1081–1088, 2008.

[13] E. Elhamifar and R. Vidal, "Sparse subspace clustering: Algorithm, theory, and applications," IEEE Trans. Pattern Anal. Mach. Intell., vol. 35, no. 11, pp. 2765–2781, Nov. 2013.

[14] E. Elhamifar, G. Sapiro, and R. Vidal, "Sparse modeling for representative selection," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 34, no. 12, pp. 2431-2444, Dec. 2012.

[15] S. Axler, Linear Algebra Done Right, 3rd ed. New York, NY, USA: Springer, 2015.

[16] G. Strang, Linear Algebra and Its Applications, 4th ed. Belmont, CA, USA: Thomson, 2006

[17] D. C. Lay, S. R. Lay, and J. McDonald, Linear Algebra and Its Applications, 5th ed. Boston, MA: Pearson, 2015

[18] L. Hogben, Handbook of Linear Algebra, 2nd ed. Boca Raton, FL: CRC Press, 2014.

[19] G. Golub and C. Van Loan, Matrix Computations, 4th ed. Baltimore, MD: Johns Hopkins University Press, 2013.

[20] R. A. Horn and C. R. Johnson, Matrix Analysis, 2nd ed. Cambridge, UK: Cambridge University Press, 2012.

[21] Y. C. Eldar and G. Kutyniok, Compressed Sensing: Theory and Applications, Cambridge, UK: Cambridge University Press, 2012.

[22] S. Boyd and L. Vandenberghe, Convex Optimization, Cambridge, UK: Cambridge University Press, 2004.

[23] M. Elad, Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing. New York, NY, USA: Springer, 2010

[24] S. Mallat, A Wavelet Tour of Signal Processing: The Sparse Way, 3rd ed. Academic Press, 2008

[25] B. K. Natarajan, "Sparse approximate solutions to linear systems," SIAM Journal on Computing, vol. 24, no. 2, pp. 227-234, 1995

[26] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," IEEE Transactions on Signal Processing, vol. 54, no. 11, pp. 4311-4322, 2006

[27] Η. Ζώης, 'Διαφάνειες διαλέξεων του μαθήματος Ψηφιακής Επεξεργασίας Εικόνας και Αναγνώρισης Προτύπων', Τμήμα Ηλεκτρολόγων και Ηλεκτρονικών Μηχανικών, Πανεπιστήμιο Δυτικής Αττικής, 2018

[28] J. MacQueen, "Some Methods for Classification and Analysis of Multivariate Observations," Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, University of California Press, 1967, pp. 281-297

[29] D. Arthur and S. Vassilvitskii, "k-means++: The Advantages of Careful Seeding," Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms, New Orleans, Louisiana, 2007, pp. 1027-1035

[30] T. Kanungo et al., "An Efficient k-Means Clustering Algorithm: Analysis and Implementation," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, no. 7, pp. 881-892, July 2002.

[31] C. Ding and X. He, "k-means Clustering via Principal Component Analysis," Proceedings of the Twenty-first International Conference on Machine Learning, Banff, Alberta, Canada, 2004.

[32] S. Lloyd, "Least Squares Quantization in PCM," IEEE Transactions on Information Theory, vol. 28, no. 2, pp. 129-137, Mar. 1982

[33] Z. Zhang, Y. Xu, J. Yang, X. Li, και D. Zhang, 'A Survey of Sparse Representation: Algorithms and Applications', IEEE Access, τ. 3, σσ. 490–530, 2015.

[34] J. M. Giron-Sierra, Digital Signal Processing with Matlab Examples, Volume 3. στο Signals and Communication Technology. Singapore: Springer Singapore, 2017.

[35] E. N. Zois, D. Tsourounis, I. Theodorakopoulos, A. L. Kesidis, και G. Economou, ‘A Comprehensive Study of Sparse Representation Techniques for Offline Signature Verification’, IEEE Trans. Biom. Behav. Identity Sci., τ. 1, τχ. 1, σσ. 68–81, Ιανουαρίου 2019

[36] J. Wright, A. Yang, A. Ganesh, S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 31, no. 2, pp. 210–227, Feb. 2009

[37] E. Candes, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," IEEE Transactions on Information Theory, vol. 52, no. 2, pp. 489–509, Feb. 2006

[38] E. Elhamifar and R. Vidal, "Sparse manifold clustering and embedding," Advances in Neural Information Processing Systems, vol. 24, pp. 55–63, 2011

[39] S. Foucart and H. Rauhut, A Mathematical Introduction to Compressive Sensing, Birkhäuser, 2013

[40] R. Gribonval and P. Vandergheynst, "Sparse and redundant signal representations: A compressed sensing perspective," Proceedings of the IEEE, vol. 98, no. 6, pp. 944–958, Jun. 2010

[41] J. Mairal, F. Bach, J. Ponce, and G. Sapiro, "Online Learning for Matrix Factorization and Sparse Coding," Journal of Machine Learning Research, vol. 11, pp. 19-60, 2010

[42] I. Tosic and P. Frossard, "Dictionary Learning," IEEE Signal Processing Magazine, vol. 28, no. 2, pp. 27-38, Mar. 2011

[43] R. Rubinstein, A. M. Bruckstein, and M. Elad, "Dictionaries for Sparse Representation Modeling," Proceedings of the IEEE, vol. 98, no. 6, pp. 1045-1057, Jun. 2010

[44] A. Ben-Tal and A. Nemirovski, Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications, MPS-SIAM Series on Optimization, 2001

[45] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers," Foundations and Trends in Machine Learning, vol. 3, no. 1, pp. 1-122, 2011

[46] G. Liu, Z. Lin, S. Yan, J. Sun, Y. Yu, and Y. Ma, "Robust Recovery of Subspace Structures by Low-Rank Representation," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 35, no. 1, pp. 171-184, Jan. 2013

[47] Z. Zhang, J. Wang, D. Zhao, and X. Tang, "Robust Subspace Segmentation via Low-Rank Representation," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2837-2844, 2012

[48] C. You, D. P. Robinson, and R. Vidal, "Scalable Sparse Subspace Clustering by Orthogonal Matching Pursuit," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 40, no. 12, pp. 3356-3371, Dec. 2018.

[49] X. Yang, M. Yang, and K. Huang, "Subspace Recovery via Sparse Regularized Nonnegative Matrix Factorization," Proceedings of the IEEE International Conference on Computer Vision (ICCV), pp. 2760-2767, 2015

[50] S. Colianni, "MNIST as JPEG", Kaggle, Available: https://www.kaggle.com/datasets/scolianni/mnistasjpg?resource=download

Παράρτημα Κώδικα

Ο κώδικας μπορεί να αναζητηθεί στην ακόλουθη ηλεκτρονική διεύθυνση github: https://github.com/StavLabr/Comp.Vis.Algo.-for-representatives-selection.