



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΔΙΑΤΜΗΜΑΤΙΚΟ ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ

«ΣΥΣΤΗΜΑΤΑ ΑΥΤΟΜΑΤΙΣΜΟΥ»

Μεταπτυχιακή Εργασία

Ταξινόμηση βλαβών ένσφαιρων εδράνων με εφαρμογή μεθόδων επεξεργασίας εικόνας και μηχανικής μάθησης

Ραβαζούλας Φώτιος

Επιβλέπων Καθηγητής: Ιωάννης Αντωνιάδης

ΑΘΗΝΑ 2019

ΠΕΡΙΕΧΟΜΕΝΑ

ΠΕΡΙΛΗΨΗ.....	2
ABSTRACT.....	3
ΚΕΦΑΛΑΙΟ 1 – ΕΙΣΑΓΩΓΗ.....	4
ΚΕΦΑΛΑΙΟ 2 – ΘΕΩΡΗΤΙΚΟ ΥΠΟΒΑΘΡΟ	7
2.1 Φασματική ανάλυση ανώτερης τάξης (Bispectrum).....	7
2.2 Μέθοδος SURF	16
2.2.1 Εισαγωγή.....	16
2.2.2 Εντοπισμός σημείων ενδιαφέροντος.....	17
2.2.3 Αναπαράσταση σε κλιμακούμενο χώρο (scale-space).....	18
2.2.4 Χωρικός εντοπισμός σημείων ενδιαφέροντος	19
2.2.5 Περιγραφή των σημείων ενδιαφέροντος.....	20
2.2.6 Δημιουργία διανύσματος περιγραφής SURF.....	21
2.3 Ομαδοποίηση K-MEANS	23
2.4 Μηχανές Διανυσματικής Υποστήριξης (SVM)	29
2.4.1 Βασικά Στοιχεία Θεωρίας Μηχανών Διανυσματικής Υποστήριξης (SVM).....	29
2.4.2 Γραμμικές Μηχανές Διανυσματικής Υποστήριξης.....	29
2.4.3 Ταξινόμηση Δεδομένων.....	34
ΚΕΦΑΛΑΙΟ 3 – ΠΕΡΙΓΡΑΦΗ ΜΕΘΟΔΟΛΟΓΙΑΣ.....	35
3.1 Εξαγωγή εικόνων bispectrum και διαχωρισμός σε ομάδες εκπαίδευσης και δοκιμών.....	36
3.2 Εξαγωγή χαρακτηριστικών SURF	37
3.3 Ομαδοποίηση χαρακτηριστικών με την μέθοδο K-Means και δημιουργία «οπτικού λεξιλογίου» ..	39
3.4 Κωδικοποίηση εικόνων με χρήση του «οπτικού λεξιλογίου» και δημιουργία ταξινομητή.....	41
ΚΕΦΑΛΑΙΟ 4 – ΠΕΙΡΑΜΑΤΙΚΑ ΑΠΟΤΕΛΕΣΜΑΤΑ	43
4.1 Προετοιμασία δεδομένων	43
4.2 Σετ 1 – Φθορά εσωτερικού δακτυλίου ρουλεμάν	45
4.3 Σετ 2 – Φθορά εξωτερικού δακτυλίου ρουλεμάν	51
ΚΕΦΑΛΑΙΟ 5 – ΣΥΜΠΕΡΑΣΜΑΤΑ ΚΑΙ ΜΕΛΛΟΝΤΙΚΗ ΕΡΓΑΣΙΑ.....	57
5.1 Συμπεράσματα	57
5.2 Μελλοντική εργασία.....	58
ΠΑΡΑΡΤΗΜΑ Α – ΚΩΔΙΚΕΣ MATLAB.....	59
ΒΙΒΛΙΟΓΡΑΦΙΑ	65

ΠΕΡΙΛΗΨΗ

Στην παρούσα εργασία γίνεται προσπάθεια ταξινόμησης βλαβών ένσφαιρων εδράνων με εφαρμογή μεθόδων επεξεργασίας εικόνας και μηχανικής μάθησης.

Η προτεινόμενη μεθοδολογία περιλαμβάνει μία σειρά συγκεκριμένων βημάτων με σκοπό τη δημιουργία ενός μοντέλου ταξινόμησης το οποίο θα είναι σε θέση να αναγνωρίζει την λειτουργική κατάσταση μίας περιστρεφόμενης μηχανής. Αρχικά δημιουργούνται απεικονίσεις των σημάτων κραδασμών της μηχανής εφαρμόζοντας στατιστική φασματική ανάλυση τρίτης τάξης (μέθοδος bispectrum). Στη συνέχεια γίνεται εξαγωγή χαρακτηριστικών από το σετ εικόνων με χρήση της μεθόδου SURF. Τα εξαγόμενα χαρακτηριστικά ομαδοποιούνται με τη μέθοδο K-means και τα κέντρα των ομάδων χρησιμοποιούνται για την κωδικοποίηση των εικόνων σε ιστογράμματα χαρακτηριστικών. Αυτά τα ιστογράμματα χρησιμοποιούνται για την εκπαίδευση μοντέλου ταξινόμηση μέσω μηχανών διανυσματικής υποστήριξης (SVM).

Η μεθοδολογία εφαρμόστηκε σε έδρανα με σφάλματα εσωτερικού και εξωτερικού δακτυλίου και έγινε αντιπαράθεση με σήματα σε κανονική κατάσταση λειτουργίας.

Οι πειραματικές δοκιμές έδειξαν ικανοποιητικά αποτελέσματα όσον αφορά την ικανότητα ταξινόμησης του μοντέλου. Περαιτέρω βελτιώσεις και τροποποιήσεις στα βήματα της μεθόδου, καθώς και δοκιμές σε διαφορετικούς τύπους σφαλμάτων και μηχανών δύναται να οδηγήσουν σε ακόμα καλύτερες αποδόσεις.

ABSTRACT

In the present paper, an attempt was made to classify ball bearing faults with a combination of image processing and machine learning methods.

The proposed methodology includes several steps in order to create a classifier that is able to determine the operational state of a rotating machine. First, the machine's vibration signals are converted into images through the use of third order statistical spectral analysis (Bispectrum). Next, interest features are extracted from the images with the SURF method. The extracted features are clustered with K-means and the cluster centers are used to encode the images into feature histograms. These histograms are, then, used to train the classifier with the use of support vector machines (SVM).

The method was applied on bearings with inner and outer race faults and were compared with signals under normal conditions.

The experiments showed satisfactory results regarding the accuracy of the classifier. Further improvements and modifications on the individual steps of the methodology, as well as application on different fault and machine types have the potential to lead to even better performance.

ΚΕΦΑΛΑΙΟ 1 – ΕΙΣΑΓΩΓΗ

Οι περιστρεφόμενες μηχανές αποτελούν αναπόσπαστο μέρος μιας πληθώρας συστημάτων της βιομηχανίας και η καλή κατάσταση λειτουργίας τους έχει αντίκτυπο όχι μόνο στην παραγωγή αλλά και στην ασφάλεια των εργαζομένων. Για αυτόν το λόγο, η έγκαιρη και αποτελεσματική διάγνωση βλαβών και συντήρηση έχει γίνει αντικείμενο έρευνας και μελέτης εδώ και πολλές δεκαετίες.

Η παρακολούθηση της κατάστασης τέτοιου είδους μηχανών, χρησιμοποιεί τις δονήσεις που παράγονται κατά την διάρκεια της λειτουργίας με χρήση κατάλληλων αισθητήρων. Ανάλογα με την κατάσταση και το φορτίο της μηχανής, οι συχνότητες και τα πλάτη αυτών των δονήσεων μεταβάλλονται. Με την χρήση κατάλληλων μεθόδων επεξεργασίας αυτών των σημάτων, μπορούν να εξαχθούν συμπεράσματα όχι μόνο για τον τύπο τυχόν υπάρχουσας βλάβης, αλλά και προβλέψεις για μελλοντικές βλάβες που μπορεί να εμφανιστούν. Σε πολλές βιομηχανίες εφαρμόζονται τέτοιου είδους τεχνικές για ελαχιστοποίηση του κόστους επισκευών και των ατυχημάτων.

Οι παραδοσιακές διαγνωστικές μέθοδοι αξιοποιούν ένα μεγάλο εύρος χαρακτηριστικών των σημάτων κραδασμών προκειμένου να επιτευχθεί όσο το δυνατόν υψηλότερη ακρίβεια. Ένα μεγάλο μέρος της βιβλιογραφίας ασχολείται με την εξαγωγή, ερμηνεία και συνδυασμό τέτοιων χαρακτηριστικών τόσο στο πεδίο του χρόνου όσο και στο πεδίο της συχνότητας. Παραδείγματα αποτελούν η μέση τιμή (mean value), η τιμή κορυφής (peak value), η τυπική απόκλιση (standard deviation), η διακύμανση (variance), η διασπορά (dispersion), η μέση τετραγωνική ρίζα (Root Mean Square RMS), ο βαθμός ασυμμετρίας (skewness) και η κύρτωση (kurtosis). Η επιλογή των κατάλληλων χαρακτηριστικών εξαρτάται από την εφαρμογή και το είδος των προς έλεγχο εξαρτημάτων.

Παρ' όλα αυτά, με τη ραγδαία ανάπτυξη των υπολογιστικών συστημάτων τα τελευταία χρόνια, έχει παρουσιασθεί ιδιαίτερο ενδιαφέρον για νέες διαγνωστικές μεθόδους, που αντλούν τεχνικές από ένα ευρύ φάσμα εφαρμογών, όπως γενετικοί αλγόριθμοι και νευρωνικά δίκτυα. Στις σύγχρονες αυτές μεθόδους, περιλαμβάνονται και υπολογιστικές μέθοδοι επεξεργασίας εικόνας. Καθώς η διαδικασία αναγνώρισης σφαλμάτων βασίζεται στο εντοπισμό συγκεκριμένων μοτίβων

σε ένα σήμα, παρουσιάζει πολλές ομοιότητες με τη διαδικασία αναγνώρισης εικόνας. Για το λόγο αυτό, οι υπολογιστικές μέθοδοι που εφαρμόζονται στο συγκεκριμένο πεδίο μπορούν να αποτελέσουν ισχυρά εργαλεία και στο τομέα διάγνωσης βλαβών.

Στην παρούσα εργασία, παρουσιάζεται μία μέθοδος που συνδυάζει τεχνικές επεξεργασίας εικόνας, εξαγωγής χαρακτηριστικών και μηχανικής μάθησης, με στόχο την αναγνώριση σφαλμάτων σε ένσφαιρους τριβείς.

Προκειμένου να μπορέσει να εφαρμοστεί οποιαδήποτε μέθοδος επεξεργασίας εικόνας, θα πρέπει αρχικά να μετατραπούν τα χρονικά σήματα κραδασμών σε εικόνες. Για το σκοπό αυτό θα αξιοποιηθεί φασματική ανάλυση ανώτερης τάξης (HOS), η οποία αποτελεί τυπική μέθοδο επεξεργασίας σήματος και προσφέρει μεγάλο πλήθος διαγνωστικών πληροφοριών. Συγκεκριμένα εφαρμόζεται στατιστική φασματική ανάλυση τρίτης τάξης (bi-spectrum), η οποία μπορεί να διατηρεί πληροφορίες για τη φάση του σήματος και έχει υψηλή αντίσταση στο θόρυβο. Οι εικόνες που δημιουργούνται από αυτήν την επεξεργασία θα χρησιμοποιηθούν ως βάση για τον αλγόριθμο αναγνώρισης.

Για την επεξεργασία των εικόνων, η μέθοδος SIFT (Scale Invariant Feature Transform) αποτελεί τη δημοφιλέστερη επιλογή για πρακτικές εφαρμογές αναγνώρισης και αντιστοίχισης χαρακτηριστικών, καθώς προσφέρει υψηλή ευρωστία και ακρίβεια. Παρ' όλα αυτά η μέθοδος SIFT έχει υψηλές απαιτήσεις όσον αφορά το χρόνο υπολογισμού, ιδιαίτερα σε εφαρμογές με μεγάλο όγκο δεδομένων. Η ανάγκη για καλύτερες αποδόσεις οδήγησε στην δημιουργία μιας τροποποιημένης εκδοχής της μεθόδου SIFT, την μέθοδο SURF (Speed-Up Robust Features), η οποία προσφέρει υψηλότερες ταχύτητες υπολογισμού και καλύτερη απόδοση. Η συγκεκριμένη μέθοδος είναι αυτή που επιλέχθηκε για τον εντοπισμό χαρακτηριστικών σημείων από τις εικόνες bispectrum.

Ο αλγόριθμος SURF εξάγει το κάθε χαρακτηριστικό μίας εικόνας ως ένα διάνυσμα περιγραφής 64 διαστάσεων. Προκειμένου να γίνει η κατηγοριοποίηση των εικόνων, όλα τα χαρακτηριστικά ομαδοποιούνται με τη μέθοδο K-means και δημιουργείται ένα «οπτικό λεξιλόγιο» μέσω του οποίου μπορούν να κωδικοποιηθούν όλες οι εικόνες. Η συγκεκριμένη τεχνική εφαρμόζεται για κατηγοριοποίηση εικόνων και έχει τις ρίζες της στο πεδίο αναγνώρισης φωνής και κειμένου, εξού και η ονομασία της.

Τέλος, για την κατηγοριοποίηση της κατάστασης της μηχανής, κατασκευάζεται ένα μοντέλο ταξινόμησης με χρήση μηχανών διανυσματικής υποστήριξης (SVM) το οποίο σε πρώτο στάδιο εκπαιδεύεται με ένα σετ σημάτων από διάφορες καταστάσεις της μηχανής και στη συνέχεια μπορεί να χρησιμοποιηθεί για κατηγοριοποίηση νέων σημάτων.

Για την αξιολόγηση της μεθόδου επιλέχθηκαν σήματα κραδασμών από ένσφαιρα έδρανα μίας τυπικής περιστρεφόμενης μηχανής σε κανονική κατάσταση λειτουργίας καθώς και σήματα με βλάβες εσωτερικού και εξωτερικού δακτυλίου.

ΚΕΦΑΛΑΙΟ 2 – ΘΕΩΡΗΤΙΚΟ ΥΠΟΒΑΘΡΟ

2.1 Φασματική ανάλυση ανώτερης τάξης (Bispectrum)

Το πλέον διαδεδομένο διαγνωστικό εργαλείο στην επεξεργασία των δυναμικών σημάτων είναι η πυκνότητα φάσματος ισχύος PSD. Το φάσμα ισχύος ενός περιοδικού σήματος είναι η συνάρτηση συχνότητας, όπου το συνολικό εμβαδόν κάτω από την καμπύλη αυτή παρέχει τη μέση ισχύ του σήματος. Επιπλέον, το φάσμα ισχύος είναι μία πραγματική ποσότητα και ως εκ τούτου δεν παρέχει καμία πληροφορία σχετικά με τη φάση του δυναμικού σήματος.

Η συνάρτηση της πυκνότητας φάσματος ισχύος για μία διακριτή χρονική ακολουθία $x(n)$ ορίζεται με την υποστήριξη του διακριτού μετασχηματισμού Fourier (DFT) ως εξής:

$$S_{xx} = E[X(k)X^*(k)] \quad (2.1)$$

όπου ο όρος k ορίζει μία διακριτή συχνотική μεταβλητή, $X(k)$ είναι ο μετασχηματισμός Fourier και ο αστερίσκος καθορίζει τον μιγαδικό συζυγή του όρου $X(k)$.

Το φάσμα ισχύος είναι δυνατόν να θεωρηθεί ως η στατιστική ροπή δεύτερης τάξης (second moment) του δυναμικού σήματος.

Τα τελευταία χρόνια, όμως, έχουν αποκτήσει περίοπτη θέση στη διαγνωστική βλαβών διαδικασίες που στηρίζονται στην επεξεργασία των ανώτερης τάξης στατιστικών χαρακτηριστικών των δυναμικών σημάτων και είναι ευρύτερα γνωστές ως στατιστικές μέθοδοι επεξεργασίας σήματος ανώτερης τάξης HOS (Higher Order Spectra, Higher Order Statistics).

Οι στατιστικές μέθοδοι επεξεργασίας σημάτων ανώτερης τάξης HOS είναι μαθηματική προέκταση των στατιστικών μεγεθών δεύτερης τάξης (second order measures), δηλαδή είναι διαδικασίες που διευρύνουν και γενικεύουν την ανάλυση των δυναμικών σημάτων με την ανάδειξη πληροφοριών που περικλείονται σε ανώτερης τάξης στατιστικές συνιστώσες του σήματος.

Επομένως, οι εν λόγω στατιστικές προσεγγίσεις βασίζονται στην αξιοποίηση των στατιστικών ροπών (moments) και συσσωρευτών (cumulants).

Το θεωρητικό αντικείμενο των στατιστικών προσεγγίσεων έχει αναπτυχθεί κάτω από ένα αυστηρό πλαίσιο, το οποίο απαιτεί μία σειρά από προϋποθέσεις που αφορούν το δυναμικό σήμα, το οποίο πρόκειται να εξετασθεί. Ένα τυχαίο σήμα $x(t)$ μπορεί να χαρακτηριστεί με αρκετούς τρόπους, όπως το πλάτος του και η περίοδος της κυματομορφής του. Η συνάρτηση πυκνότητας πιθανότητας PDF (probability density function) μίας χρονικής ακολουθίας εξασφαλίζει παροχή λεπτομερών πληροφοριών σχετικές με το πλάτος του σήματος. Έτσι, ένα δυναμικό σήμα είναι δυνατόν να περιγραφεί από την μορφή της καμπύλης της συνάρτησης πυκνότητας πιθανότητας PDF αυτού. Οι έμφυτες πληροφορίες της συνάρτησης πυκνότητας πιθανότητας PDF μπορούν να ποσοτικοποιηθούν με τη χρησιμοποίηση των στατιστικών ροπών και συσσωρευτών.

Η μέση τιμή (mean) ενός τυχαίου σήματος ορίζεται ως η αναμενόμενη τιμή (expected value) της τυχαίας μεταβλητής $x(t)$:

$$m_x(t) = E\{x(t)\} = \int_{-\infty}^{+\infty} x \cdot p_x(x, t) dx \quad (2.2)$$

όπου $p_x(x_1; t_1) = \frac{dP_x(x_1; t_1)}{dx_1}$ είναι η συνάρτηση πυκνότητας πιθανότητας, η οποία ορίζεται ως η παράγωγος της συνάρτησης κατανομής $P_x(x_1; t_1) = \text{prob}\{x(t_1) \leq x_1\}$, και ο όρος E ορίζει τον τελεστή της αναμενόμενης τιμής (expectation operator). Δηλαδή, η μέση τιμή m_x εντοπίζει το κέντρο βάρους της περιοχής κάτω από την καμπύλη της πυκνότητας πιθανότητας της τυχαίας μεταβλητής.

Ακολούθως, η μεταβλητότητα (variance) της τυχαίας μεταβλητής $x(t)$ ορίζεται από τη σχέση:

$$\sigma_x^2 = \text{Var}\{x\} = E\{(x - m_x)^2\} = \int_{-\infty}^{+\infty} (x - m_x)^2 \cdot p_x(x, t) dx \quad (2.3)$$

Η τετραγωνική ρίζα της μεταβλητότητας, δηλαδή σ_x , ονομάζεται τυπική απόκλιση (standard deviation) της τυχαίας μεταβλητής.

Η Ν-οστή ροπή του τυχαίου σήματος $x(t)$ ορίζεται από την ακόλουθη σχέση:

$$m_n = E\{x^n\} = \int_{-\infty}^{+\infty} x^n \cdot p_x(x, t) dx \quad (2.4)$$

Οι περισσότερες σημαντικές ροπές της τυχαίας μεταβλητής $x(t)$ είναι οι δύο πρώτες. Έτσι, θέτοντας $n=1$ στη σχέση 2.4, λαμβάνεται η πρώτη ροπή που είναι η μέση τιμή της τυχαίας μεταβλητής, όπως περιεγράφηκε στη σχέση 2.2, ενώ θέτοντας $n=2$ λαμβάνεται η δεύτερη ροπή που είναι η μέση τετραγωνική τιμή (mean square value) της τυχαίας μεταβλητής:

$$m_2 = E\{x^2\} = \int_{-\infty}^{+\infty} x^2 \cdot p_x(x, t) dx \quad (2.5)$$

Οι κεντρικές ροπές (central moments) είναι απλά οι ροπές της διαφοράς μεταξύ μιας τυχαίας μεταβλητής και της μέσης τιμής της, και ορίζονται ως εξής:

$$E\{(x - m_x)^n\} = \int_{-\infty}^{+\infty} (x - m_x)^n \cdot p_x(x, t) dx \quad (2.6)$$

Για $n=1$, η κεντρική ροπή είναι φυσικά μηδέν, ενώ για $n=2$ η δεύτερη κεντρική ροπή είναι η μεταβλητότητα που αναφέρθηκε παραπάνω.

Η τρίτη κεντρική ροπή ($n=3$) ονομάζεται λοξότητα (skewness) και αποτελεί το καλύτερο μέτρο σύνοψης της ασυμμετρίας μιας τυχαίας διαδικασίας. Στη διεθνή βιβλιογραφία συναντάται με την ακόλουθη ομαλοποιημένη σχέση:

$$skew = \frac{\frac{1}{N} \sum_{i=1}^N (x_i - m_x)^3}{\sigma_x^3} \quad (2.7)$$

Ένα μέτρο κύρτωσης ή ένα μέτρο της αιχμηρότητας που εμφανίζει η κατανομή μιας τυχαίας μεταβλητής αποτελεί η κεντρική ροπή τέταρτης τάξης (n=4), η οποία ονομάζεται κύρτωση (kurtosis) και εκφράζεται από την ακόλουθη ομαλοποιημένη σχέση:

$$Kurt = \frac{\frac{1}{N} \sum_{i=1}^N (x_i - m_x)^4}{\sigma_x^4} \quad (2.8)$$

Οι στατιστικές ροπές του τυχαίου σήματος συνδέονται με τη συνάρτηση πυκνότητας πιθανότητας PDF με την υποστήριξη της συνάρτησης παραγωγής ροπών (moment generating function), η οποία ορίζεται από την ακόλουθη σχέση:

$$\Phi(s) = \int_{-\infty}^{+\infty} p_x(x_1; t_1) \cdot e^{s \cdot x} dx \quad (2.9)$$

Η n-οστή ροπή του τυχαίου σήματος είναι δυνατόν να υπολογισθεί από την n-οστή παράγωγο της συνάρτησης παραγωγής ροπών $\Phi(s)$ για $s=0$ σύμφωνα με τη σχέση:

$$m_n = \left. \frac{d^n \Phi(s)}{ds^n} \right|_{s=0} \quad (2.10)$$

Ένας άλλος σημαντικός στατιστικός μέσος όρος είναι η χαρακτηριστική συνάρτηση (characteristic function), η οποία ισούται με το λογάριθμο της συνάρτησης παραγωγής ροπών $\Phi(s)$.

Οι στατιστικοί συσσωρευτές (cumulants) του τυχαίου σήματος είναι δυνατόν να προκύψουν από τον υπολογισμό των n παραγώγων της χαρακτηριστικής συνάρτησης για $s=0$:

$$c_n = \left. \frac{d^n \ln(\Phi(s))}{ds^n} \right|_{s=0} \quad (2.11)$$

Επίσης, οι τέσσερις πρώτοι στατιστικοί συσσωρευτές είναι δυνατόν να υπολογιστούν από τις τέσσερις πρώτες ροπές χρησιμοποιώντας τις παρακάτω σχέσεις:

$$c_1 = m_1$$

$$c_2 = m_2 - m_1^2 \quad (2.12)$$

$$c_3 = m_3 - 3 \cdot m_2 \cdot m_1 + 2 \cdot m_1^3$$

$$c_4 = m_4 - 3 \cdot m_2^2 - 4 \cdot m_3 \cdot m_1 + 12 \cdot m_2 \cdot m_1^2 - 6 \cdot m_1^4$$

Όταν η μέση τιμή του τυχαίου σήματος είναι μηδέν, οι συσσωρευτές της σχέσης 5.12 είναι:

$$c_1 = m_1 = 0$$

$$c_2 = m_2 \quad (2.13)$$

$$c_3 = m_3$$

$$c_4 = m_4 - 3 \cdot m_2^2$$

Η στατιστική φασματική ανάλυση τρίτης τάξης (Bispectrum/Bilinear Analysis) ταυτίζεται με το διπλό μετασχηματισμό Fourier του συσσωρευτή τρίτης τάξης του τυχαίου σήματος. Η γραφική αναπαράσταση της άνωθεν στατιστικής προσέγγισης αποτελεί το πρώτο και το πιο απλό από τα στατιστικά φάσματα ανώτερης τάξης (Bispectrum).

Είναι δυνατόν να εφαρμοσθεί στο κυβικό τυχαίο σήμα (cubed signal) ανάλογη διαδικασία με αυτήν που ακολουθείται για τον υπολογισμό του φάσματος ισχύος:

$$\int_{-\infty}^{+\infty} x^3(t) dt = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} E_{xxx}(f_1, f_2) df_1 df_2 \quad (2.14)$$

Έτσι, το στατιστικό φάσμα τρίτης τάξης (Bispectrum) μπορεί να εκφρασθεί από την ακόλουθη σχέση:

$$E_{xxx}(f_1, f_2) = X(f_1)^* \cdot X(f_2)^* \cdot X(f_1 + f_2) \quad (2.15)$$

Συνήθως, όμως, στη διεθνή βιβλιογραφία χρησιμοποιείται ο ακόλουθος ορισμός, ο οποίος πηγάζει από το διακριτό μετασχηματισμό Fourier του τυχαίου σήματος:

$$E_{xxx}(f_1, f_2) = X(f_1) \cdot X(f_2) \cdot X(f_1 + f_2)^* \quad (2.16)$$

Όπως έχει ήδη αναφερθεί το φάσμα ισχύος (power spectrum) αναλύει και ερμηνεύει την ροπή δεύτερης τάξης (διακύμανση) του τυχαίου σήματος στο πεδίο της συχνότητας. Με ανάλογη διαδικασία το στατιστικό φάσμα τρίτης τάξης (Bispectrum) αναλύει την ροπή τρίτης τάξης ή τη λοξότητα (skewness) της συνάρτησης πυκνότητας πιθανότητας PDF του τυχαίου σήματος.

Η ροπή τρίτης τάξης είναι μια συνάρτηση που υποστηρίζεται από δύο φασματικές μεταβλητές f_1 και f_2 . Σε αντίθεση με το φάσμα ισχύος, που εξετάζει τα χαρακτηριστικά μεγέθη κάθε φασματικής συνιστώσας ανεξάρτητα από τις υπόλοιπες, η ροπή τρίτης τάξης εντοπίζει και ερμηνεύει τις αλληλεπιδράσεις ανάμεσα στις φασματικές συνιστώσες του τυχαίου σήματος στις συχνότητες f_1 , f_2 και f_1+f_2 . Επιπλέον, είναι σημαντικό να αναφερθεί ότι η ροπή τρίτης τάξης είναι μια μιγαδική ποσότητα που περιλαμβάνει δύο φασματικές μεταβλητές και παρέχει πληροφορίες που αφορούν τη φάση του τυχαίου σήματος.

Ένα βασικό προνόμιο της εφαρμογής της στατιστικής φασματικής ανάλυσης ανώτερης τάξης HOS στην προβλεπτική συντήρηση στρεφόμενου μηχανολογικού εξοπλισμού πηγάζει από τις έμφυτες ιδιότητες που περικλείει μία τυχαία διαδικασία κανονικής κατανομής (Gauss distribution). Μία τυχαία διαδικασία κανονικής κατανομής χαρακτηρίζεται πλήρως από τις στατιστικές ροπές πρώτης και δεύτερης τάξης (μέση τιμή και διακύμανση). Αποδεικνύεται εύκολα πως όλοι οι συσσωρευτές ανώτερης τάξης μίας τυχαίας διαδικασίας κανονικής κατανομής είναι μηδενικοί. Έτσι, ο θόρυβος, ο οποίος ακολουθεί κανονική κατανομή, εξαφανίζεται μετά την εφαρμογή συσσωρευτών τρίτης και πλέον τάξης. Η χρήση της λέξης ‘εξαφανίζεται’ ίσως είναι υπερβολική για πραγματικές διαδικασίες, αλλά αναφέρεται για να τονίσει και να δώσει έμφαση σε αυτήν τη σημαντική ιδιότητα. Αυτό που συμβαίνει μετά την εφαρμογή των συσσωρευτών ανώτερης τάξης είναι η αύξηση του λόγου σήματος – θορύβου SNR.

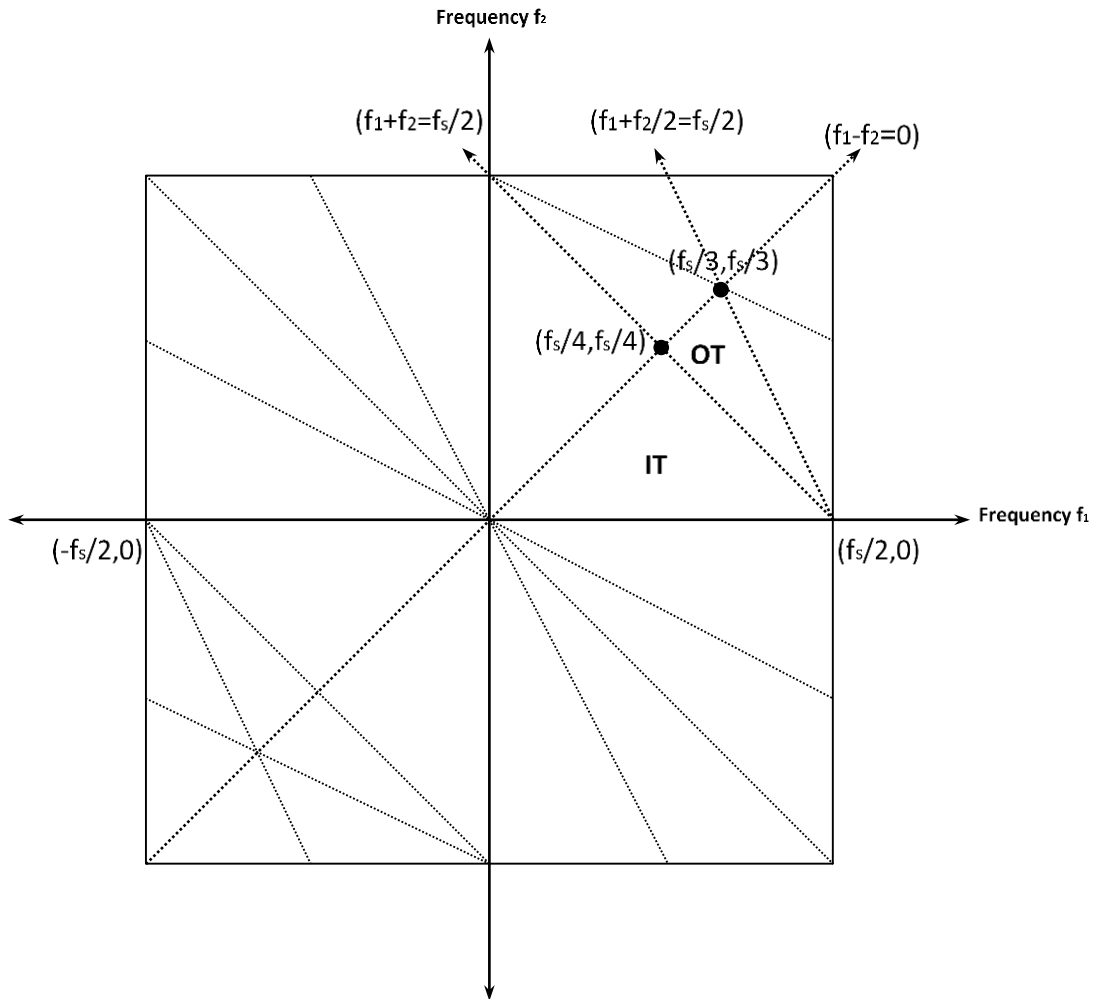
Στον πραγματικό κόσμο, όμως, τα περισσότερα σήματα παρουσιάζουν μη κανονική κατανομή (non-Gaussian distribution) λόγω της ύπαρξης μη γραμμικών φαινομένων. Έτσι, το τυχαίο σήμα που περικλείει επωφελείς πληροφορίες και ακολουθεί μη κανονική κατανομή αναδύεται μέσα από την κουρτίνα του θορύβου και αποκαλύπτεται μετά την εφαρμογή στατιστικών μεθόδων επεξεργασίας που βασίζονται σε συσσωρευτές ανώτερης τάξης. Αυτή η διεργασία αποτελεί το κλειδί της στατιστικής φασματικής ανάλυσης ανώτερης τάξης HOS.

Ο διακριτός μετασχηματισμός Fourier (DFT) και το φάσμα ισχύος περιέχουν περιττές πληροφορίες άνω της συχνότητας του Nyquist, η οποία ισούται με το ήμισυ της συχνότητας δειγματοληψίας f_s . Κατά παρόμοιο τρόπο, αναπτύσσονται συμμετρίες στο πεδίο (f_1, f_2) του στατιστικού φάσματος τρίτης τάξης, και για αυτό το λόγο δεν είναι απαραίτητος ο υπολογισμός της στατιστικής φασματικής ανάλυσης τρίτης τάξης $E_{xxx}(f_1, f_2)$ για όλα τα ζεύγη συχνοτήτων (f_1, f_2) . Η περιοχή του στατιστικού φάσματος τρίτης τάξης που περιλαμβάνει τις απαραίτητες και

χρήσιμες πληροφορίες ονομάζεται στοιχειώδες ή κύριο πεδίο PD (principal domain) (σχήμα 2-1) και ορίζεται από την ακόλουθη σχέση:

$$\left\{ f_1; 1: 0 \leq f_1 \leq \frac{f_s}{2}, f_2 \leq f_1, 2 \cdot f_1 + 1 = f_s \right\} \quad (2.17)$$

Για την ανάλυση των στοχαστικών σημάτων, το κύριο πεδίο PD είναι απαραίτητο να διαιρεθεί σε δύο τριγωνικές περιοχές: την εσωτερική IT (inner triangle) και την εξωτερική OT (outer triangle) περιοχή (σχήμα 2-1). Όλες οι σημαντικές πληροφορίες που ευρίσκονται στις περιοχές αυτές.



Σχήμα 2-1 : Το πεδίο της στατιστικής φασματικής ανάλυσης τρίτης τάξης και οι συμμετρίες που εμφανίζονται σε αυτό.

2.2 Μέθοδος SURF

2.2.1 Εισαγωγή

Η διαδικασία εντοπισμού κοινών σημείων μεταξύ δύο εικόνων που απεικονίζουν το ίδιο τοπίο ή αντικείμενο αποτελεί μέρος πολλών εφαρμογών μηχανικής όρασης όπως ταξινόμηση εικόνων, παραμετροποίηση καμερών και αναγνώριση αντικειμένων.

Η εύρεση τέτοιων διακριτών σημείων μπορεί να χωριστεί σε τρία βασικά βήματα. Πρώτον, επιλέγονται «σημεία ενδιαφέροντος» σε συγκεκριμένα σημεία της εικόνας, όπως γωνίες, τομές και σημεία έντονων διακυμάνσεων. Η σημαντικότερη ιδιότητα ενός «εντοπιστή» σημείων ενδιαφέροντος είναι η επαναληψιμότητα του. Η επαναληψιμότητα εκφράζει την αξιοπιστία του εντοπιστή στο να βρίσκει τα ίδια σημεία ενδιαφέροντος κάτω από διαφορετικές συνθήκες θέασης. Στη συνέχεια η γειτονιά του κάθε σημείου ενδιαφέροντος περιγράφεται από ένας διάνυσμα χαρακτηριστικών. Αυτό το «διάνυσμα περιγραφής» πρέπει να είναι ξεχωριστό και παράλληλα εύρωστο σε θόρυβο και γεωμετρικές και φωτομετρικές παραμορφώσεις. Τέλος, τα διανύσματα περιγραφής αντιστοιχίζονται μεταξύ διαφορετικών εικόνων. Η αντιστοίχιση βασίζεται στην απόσταση μεταξύ των διανυσμάτων, π.χ. απόσταση Mahalanobis ή Ευκλείδεια απόσταση. Η διάσταση του διανύσματος περιγραφής έχει άμεση επίδραση στο χρόνο υπολογισμού και διανύσματα χαμηλότερων διαστάσεων επιλέγονται για γρηγορότερη αντιστοίχιση σημείων ενδιαφέροντος. Παρ' όλα αυτά, διανύσματα λιγότερων διαστάσεων δεν έχουν την ίδια ευκρίνεια όσο τα αντίστοιχα διανύσματα περισσότερων διαστάσεων.

Η συγκεκριμένη μέθοδος αναπτύχθηκε από τους [Herbert Bay, Andreas Ess, Tinne Tuytelaars, Luc Van Gool] με στόχο να δημιουργηθεί ένας εντοπιστής και αντίστοιχο διάνυσμα περιγραφής τα οποία τα μπορούν να υπολογιστούν ταχύτερα από τις μέχρι τότε υπάρχουσες μεθόδους χωρίς να θυσιαστεί η απόδοσή. Για το λόγο αυτό έπρεπε να επιτευχθεί μία ισορροπία μεταξύ των παραπάνω απαιτήσεων, όπως απλοποίηση της μεθόδου εντοπισμού, διατηρώντας την ακρίβεια της, και μείωση των διαστάσεων του διανύσματος περιγραφής χωρίς να μειωθεί η ικανότητα περιγραφής. Η συγκεκριμένη μέθοδος ονομάζεται SURF (Speeded-Up Robust Features) και μπορεί να επιτύχει όχι μόνο υψηλότερη ταχύτητα υπολογισμού αλλά και καλύτερη επαναληψιμότητα.

2.2.2 Εντοπισμός σημείων ενδιαφέροντος

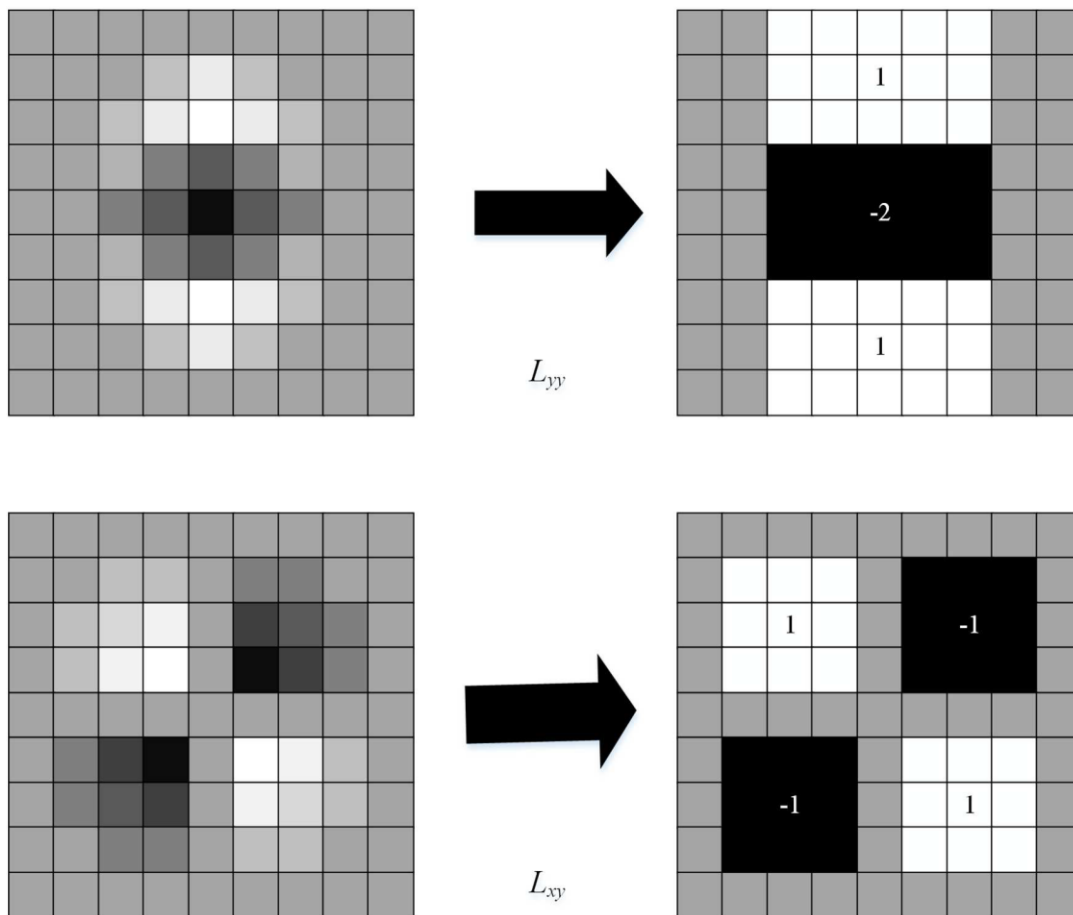
Για τον εντοπισμό των σημείων ενδιαφέροντος χρησιμοποιείται μία προσέγγιση του πίνακα Hessian. Το κάθε σημείο ενδιαφέροντος εντοπίζεται από το τοπικό μέγιστο του πίνακα.

Αν θεωρήσουμε ένα συγκεκριμένο σημείο (x,y) στη γραμμή x και στήλη y της εικόνας I , ο πίνακας Hessian $H(x,\sigma)$ στο x με φίλτρο Gaussian κλίμακας σ ορίζεται ως:

$$H(x, \sigma) = \begin{bmatrix} L_{xx}(x, \sigma) & L_{xy}(x, \sigma) \\ L_{xy}(x, \sigma) & L_{yy}(x, \sigma) \end{bmatrix} \quad (2.18)$$

Όπου $L_{xx}(x,\sigma)$, $L_{xy}(x,\sigma)$, $L_{yy}(x,\sigma)$ είναι η μερική παράγωγος δεύτερης τάξης και η διαδιάστατη σύγκλιση του σημείου (x,y) της εικόνας I , αντίστοιχα.

Η συνάρτηση Gauss μπορεί να εφαρμοστεί στη ανάλυση του χώρου κλίμακας, αλλά σε πρακτικές εφαρμογές, πρέπει να διακριτοποιηθεί και να αποκοπεί όπως φαίνεται στο σχήμα 2-2:



Σχήμα 2-2 : Προσεγγιστικές παράγωγοι 2^{ης} τάξης με φίλτρα κουτιού.

Αυτό έχει ως αποτέλεσμα σε απώλεια επαναληψιμότητας κατά την περιστροφή της εικόνας κατά περιττά πολλαπλάσια του $\pi/4$. Για βελτίωση της ακρίβειας της παραγώγου Gauss δεύτερης τάξης χρησιμοποιούνται φίλτρα τύπου «κουτιού» (**box filters**). Η επεξεργασία ακέραιων εικόνων με χρήση πλαισίου σύγκλισης (**convolution template**) μετά την προσέγγιση παρέχει ξεκάθαρο πλεονέκτημα, καθώς το πλαίσιο αποτελείται από ένα απλό παραλληλόγραμμο, του οποίου ο υπολογισμός είναι ανεξάρτητος του μεγέθους. Με αυτόν τον τρόπο η αποδοτικότητα των υπολογισμών επιταχύνεται σημαντικά.

Η προσέγγιση δεν επηρεάζει αρνητικά την απόδοση του πλαισίου σύγκλισης. Το πλαίσιο σύγκλισης του **σχήματος 13** είναι προσεγγίσεις μίας συνάρτησης Gauss με $\sigma=1.2$, η οποία είναι η μικρότερη κλίμακα για τους χάρτες απόκρισης των διακυμάνσεων (**blobs**). Το πλαίσιο προσέγγισης και η σύγκλιση ορίζονται ως D_{xx} , D_{yy} , D_{xy} , και χρησιμοποιούνται αντί για τα $L_{xx}(x,\sigma)$, $L_{xy}(x,\sigma)$, $L_{yy}(x,\sigma)$ για να εξαχθεί ο προσεγγιστικός πίνακας Hessian H .

$$|H| = D_{xx}(X)D_{yy}(X) - [wD_{xy}(X)]^2 \quad (2.19)$$

Όπου $w \approx 0.9$

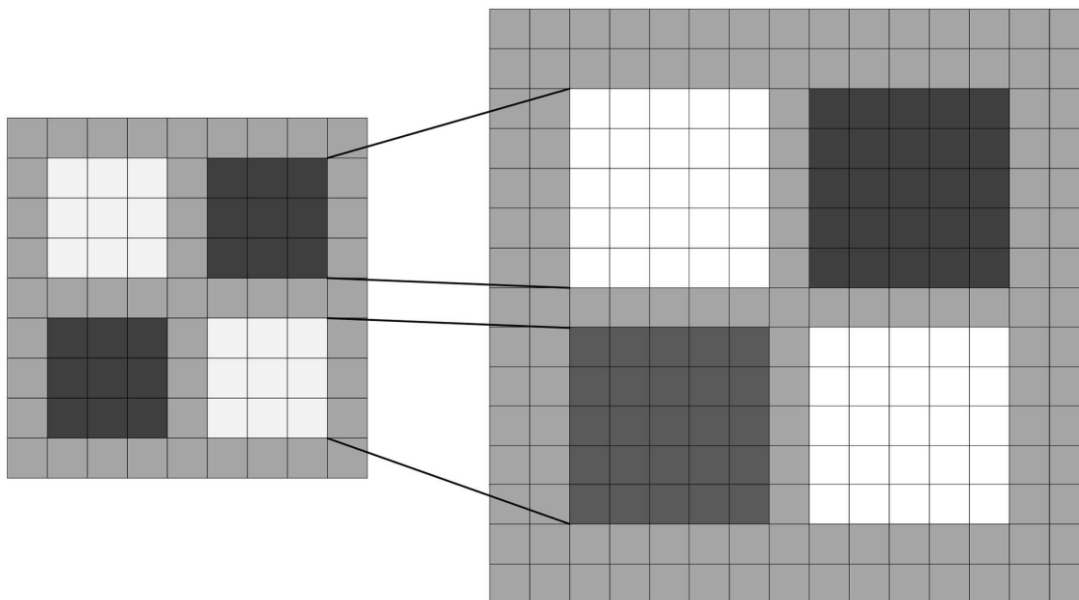
2.2.3 Αναπαράσταση σε κλιμακούμενο χώρο (scale-space)

Προκειμένου να γίνει ο εντοπισμός των σημείων ενδιαφέροντος με σχέση σμίκρυνσης-μεγέθυνσης, θα πρέπει ο εντοπιστής να έχει τη δυνατότητα να βρίσκει σημεία στην ίδια θέση αλλά σε διαφορετική κλίμακα. Ο κλιμακούμενος χώρος συνήθως αναπαρίσταται με μία πυραμίδα εικόνων. Οι εικόνες εξομαλύνονται ένα με φίλτρο Gauss και χρησιμοποιούνται για την επίτευξη υψηλότερου επιπέδου στη πυραμίδα. Αφαιρώντας όλα τα επίπεδα της πυραμίδας εξάγεται η διαφορά των Gaussians (Difference of Gaussians – DoG) όπου μπορούν να εντοπιστούν διακυμάνσεις και ακμές της εικόνας.

Σε αντίθεση με αυτή τη προσέγγιση, η μέθοδος SURF επεξεργάζεται την αρχική εικόνα με φίλτρα κουτιού διαφόρων μεγεθών. Χάρη στην εφαρμογή ακέραιων εικόνων, η ταχύτητα υπολογισμού μεταξύ φίλτρων κουτιού όλων των μεγεθών είναι ίδια.

Στην αριστερή στήλη του σχήματος 2-3 δυο πλαίσια προσέγγισης $9*9$ θεωρούνται ως το αρχικό πλαίσιο προσέγγισης (προσεγγιστικές παράγωγοι Gauss με $\sigma=1.2$). Το S αναφέρεται στη κλίμακα του πλαισίου προσέγγισης, όπου $s = \sigma = 1.2$. Ως αρχικό επίπεδο ορίζεται η σύγκλιση

μεταξύ των προσεγγιστικών πλαισίων της αρχικής κλίμακας της εικόνας. Τα επόμενα επίπεδα λαμβάνονται από τη σύγκλιση της αρχικής εικόνας και κλιμακωτά αυξανόμενου s . Προκειμένου να διατηρηθεί η ακεραιότητα και το κεντρικό pixel, τα γειτονικά πλαίσια διαφέρουν πάντα κατά περιττό αριθμό.



Σχήμα 2-3 : Φίλτρα D_{xy} για δύο διαδοχικές κλίμακες (9x9 και 15x15)

Κάθε οκτάβα αποτελείται από 4 πλαίσια. Στην πρώτη οκτάβα, τα γειτονικά πλαίσια διαφέρουν κατά 6 pixel. Στην δεύτερη οκτάβα κατά 8 pixel, στην τρίτη κατά 24 κ.λπ. Αν το μέγεθος ενός πλαισίου είναι $N*N$, η κλίμακα του πλαισίου είναι $S=1.2*9/N$. Υπολογίζοντας και καταγράφοντας την απόκριση σε κάθε σημείο, μπορούμε να αποτυπώσουμε την απόκριση σε διαφορετική κλίμακα. Με αυτόν τον τρόπο μπορεί να κατασκευαστεί ένας τρισδιάστατος κλιμακούμενος χώρος.

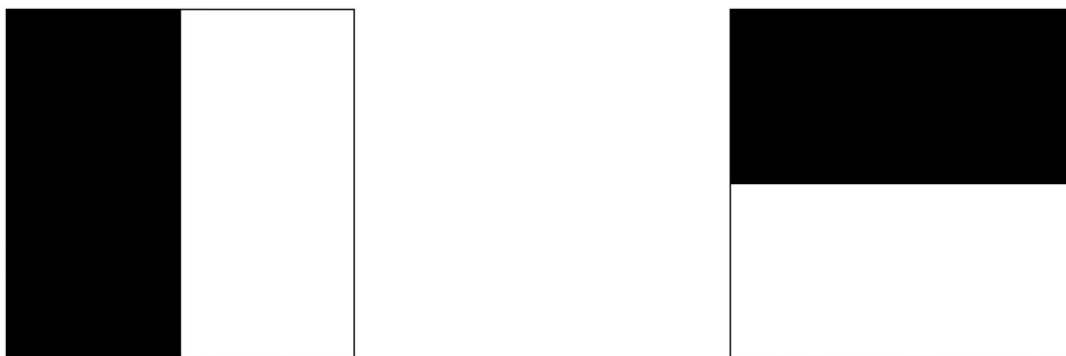
2.2.4 Χωρικός εντοπισμός σημείων ενδιαφέροντος

Αρχικά ερευνώνται ακρότατα της κλιμακούμενης εικόνας (x, y, σ) σύμφωνα με τον πίνακα Fast-Hessian. Τα σημεία που δεν αποτελούν μέγιστα απαλείφονται στην κυκλική γειτονιά $3x3x3$. Μόνο ακραία σημεία με διαφορά μεγαλύτερη ή μικρότερη του 26 συγκριτικά με την αντίστοιχη μία κλίμακας πάνω ή κάτω αντίστοιχα μπορούν να θεωρηθούν ως σημεία χαρακτηριστικών. Για τον εντοπισμό σημείων χαρακτηριστικών με ακρίβεια υπο-pixel εφαρμόζετε γραμμική παρεμβολή

μεταξύ του χώρου κλίμακας και του χώρου της εικόνας. Με αυτό τον τρόπο λαμβάνεται ένα ευσταθές χαρακτηριστικό σημείο και η κλίμακά του στο σημείο εντοπισμού.

2.2.5 Περιγραφή των σημείων ενδιαφέροντος

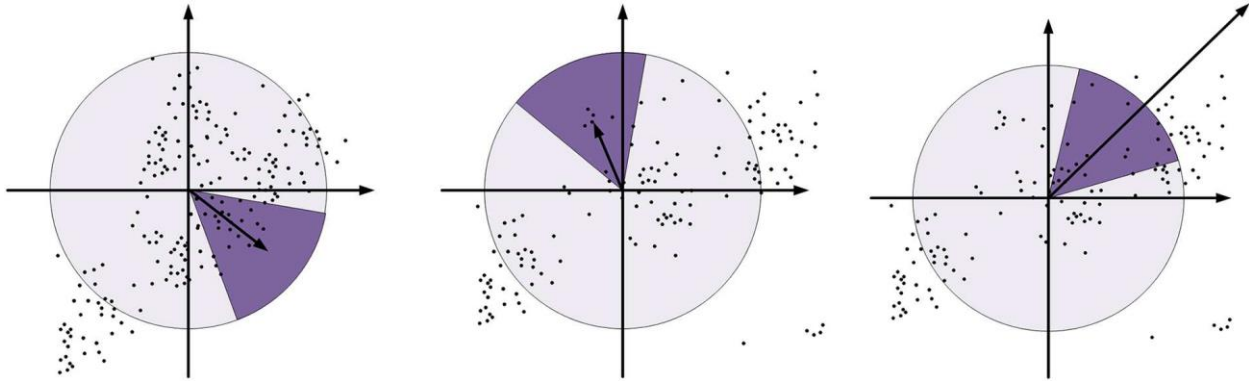
Προκειμένου το διάλυσμα περιγραφής να μην επηρεάζεται από περιστροφές, πρέπει να ληφθεί ο προσανατολισμός του σημείου. Για τον σκοπό αυτό, αρχικά κατασκευάζεται μία απόκριση wavelet της οποίας το κέντρο είναι το σημείο ενδιαφέροντος και η ακτίνα της $6s$ (όπου s είναι η κλίμακα του σημείου). Με εφαρμογή Haar wavelet μεγέθους $4s$ στην εικόνα, μπορεί να υπολογιστεί η απόκριση κατά τον άξονες x και y .



Σχήμα 2-4 : Φίλτρα Haar wavelet

Στο σχήμα 2-4 φαίνονται απεικονίσεις των Haar wavelet. Η αριστερή εικόνα χρησιμοποιείται για την απόκριση στον άξονα x και η αριστερή για τον άξονα y . Τα μαύρα τμήματα σηματοδοτούνται με -1 και τα λευκά με $+1$.

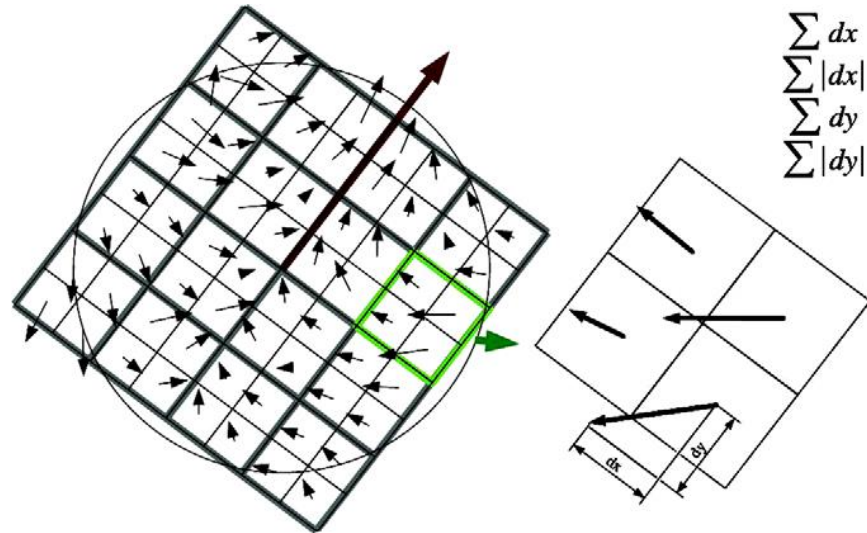
Η γειτονιά της απόκρισης wavelet βρίσκεται στην κατεύθυνση των x και y αξόνων. Στη συνέχεια αντιστοιχίζονται διαφορετικά βάρη Gauss σε κάθε απόκριση. Αθροίζοντας το Haar wavelet στις κατευθύνσεις των αξόνων x και y με εύρος 60° δημιουργείται ένα τοπικό διάλυσμα προσανατολισμού. Μόλις έχει ολοκληρωθεί η παραπάνω διαδικασία για ολόκληρο τον κύκλο, το μεγαλύτερο διάλυσμα επιλέγεται ως ο κύριος προσανατολισμός του σημείου. Η διαδικασία απεικονίζεται στο σχήμα 2-5.



Σχήμα 2-5 : Καθορισμός κύριου προσανατολισμού του χαρακτηριστικού σημείου

2.2.6 Δημιουργία διανύσματος περιγραφής SURF

Για την εξαγωγή του διανύσματος περιγραφής, αρχικά δημιουργείται ένα παράθυρο με κέντρο το σημείο ενδιαφέροντος και μήκος 20σ . Στη συνέχεια το παράθυρο χωρίζεται σε 4×4 υπο-περιοχές (σχήμα 2-6). Από κάθε υπο-περιοχή λαμβάνονται 5×5 σημεία δειγματοληψίας. Για κάθε ένα από αυτά τα σημεία λαμβάνονται οι αποκρίσεις wavelet κατά τους άξονες x και y και ορίζονται ως dx και dy αντίστοιχα. Σε κάθε υπο-περιοχή εφαρμόζεται ένα φίλτρο Gauss με κέντρο το σημείο ενδιαφέροντος και σχηματίζεται ένα ιδιοδιάνυσμα τεσσάρων διαστάσεων από τα αθροίσματα των dx , dy , $|dx|$ και $|dy|$ (Σdx , Σdy , $\Sigma |dx|$, $\Sigma |dy|$) όπως φαίνεται στο σχήμα 2-7. Για τις $4 \times 4 = 16$ υπο-περιοχές, έχουμε συνολικά $4 \times 16 = 64$ διαστάσεις του διανύσματος περιγραφής.



Σχήμα 2-6 : Αριστερά απεικονίζονται οι 4x4 υπο-περιοχές γύρω από το σημείο ενδιαφέροντος. Για κάθε υπο-περιοχή, οι αποκρίσεις wavelet λαμβάνονται από 5x5 δείγματα (για λόγους οπτικοποίησης, στο σχήμα έχουν ληφθεί δείγματα 2x2



Σχήμα 2-7 : Τιμές των ιδιοδιανυσμάτων σε διάφορα ήδη απεικονίσεων

2.3 Ομαδοποίηση K-MEANS

Η ομαδοποίηση K-means είναι ένας από τους πιο απλούς και πιο δημοφιλείς αλγορίθμους εκμάθησης χωρίς επίβλεψη για την επίλυση του προβλήματος ομαδοποίησης. Η ομαδοποίηση K-means δημιουργεί ένα συγκεκριμένο αριθμό ασυνεχών επίπεδων (μη ιεραρχικών) συστάδων. Δηλαδή, η συνάρτηση των K-means διαχωρίζει τις παρατηρήσεις που εξάγονται από τα δεδομένα σε k αποκλειστικές συστάδες, και επιστρέφει ένα διάνυσμα δεικτών, το οποίο υποδεικνύει σε ποια από τις k συστάδες έχει ανατεθεί κάθε σετ χαρακτηριστικών.

Το πρόβλημα της M-συσταδοποίησης στοχεύει στη δημιουργία διαμερισμάτων σε ένα σύνολο δεδομένων $X = \{x_1, \dots, x_N\}$, $x_n \in R^d$, σε M ασυνεχή υποσύνολα C_1, \dots, C_M , έτσι ώστε να βελτιστοποιείτε μια συνάρτηση F , που αναφέρεται ως κριτήριο ομαδοποίησης. Η πιο ευνόητη και συχνά χρησιμοποιούμενη συνάρτηση-κριτήριο είναι το άθροισμα των τετραγώνων της ευκλείδειας απόστασης ανάμεσα σε κάθε σημείο δεδομένων x_i και το κέντρο του συμπλέγματος m_k του υποσυνόλου που περιέχει το x_i . Το κριτήριο αυτό ονομάζεται σφάλμα ομαδοποίησης και εξαρτάται από τα κέντρα m_1, \dots, m_M :

$$F(m_1, \dots, m_M) = \sum_{i=1}^M \sum_{j=1}^{M_i} \|x_{ij} - m_i\|^2 \quad (2.20)$$

όπου M_i είναι ο αριθμός των σημείων της συστάδας C_i , x_{ij} είναι το j -th σημείο της i -th συστάδας και m_i είναι το κέντρο της i -th συστάδας το οποίο ορίζεται ως εξής:

$$m_i = \frac{1}{M_i} \sum_{j=1}^{M_i} x_{ij} \quad i = 1, \dots, M \quad (2.21)$$

Ένας ψευδοκώδικας για την εφαρμογή της ομαδοποίησης K-means έχει ως εξής:

(A) ο αλγόριθμος αρχίζει με μια αρχική κατάτμηση της βάσης δεδομένων σε ασυνεχή υποσύνολα και υπολογίζονται τα κέντρα βάρους των αρχικών συστάδων,

(B), στη συνέχεια, τα δεδομένα "επανατοποθετούνται" στη συστάδα με το πλησιέστερο κέντρο βάρους σε μια προσπάθεια να μειωθεί το τετραγωνικό σφάλμα. Ένα σημείο $x_i \in C_s$ στο βήμα της επανατοποθέτησης μπορεί να αλλάξει συστάδα στην οποία ανήκει $x_i \in C_t$ εάν

$\|x_i - m_i\| \leq \|x_i - m_j\|$. Στη συνέχεια τα κέντρα βάρους, των συστάδων, και το τετραγωνικό σφάλμα θα πρέπει να υπολογιστούν εκ νέου.

(C) Αυτή η διαδικασία επαναλαμβάνεται μέχρι το τετραγωνικό σφάλμα να μην μπορεί να μειωθεί περαιτέρω, που σημαίνει ότι κανένα δεδομένο, δεν είναι δεν μπορεί να αλλάξει συστάδα.

Έχει αποδειχθεί ότι η απόδοση της μεθόδου συνδέεται στενά με τις χρησιμοποιούμενες μετρική απόσταση. Μπορούν να επιλεγούν μια σειρά από διαφορετικά μέτρα απόστασης, ανάλογα με το είδος των δεδομένων που θέλουμε να ομαδοποιήσουμε.

Η Ευκλείδεια απόσταση μέτρα την απόσταση σε ευθεία γραμμή μεταξύ δύο σημείων. Κάθε κέντρο βάρους είναι ο μέσος όρος των σημείων της συστάδας. Αν έχουμε δύο γονίδια $A_i = \{x_{ik}\}$ και $A_j = \{x_{jk}\}$, $k = 1, \dots, n$ και $i, j \in \{1, \dots, n\}$, $i \neq j$ η Ευκλείδεια μεταξύ τους απόσταση και διατυπώνεται ως εξής:

$$r(A_i, A_j) = \sqrt{\sum_{k=1}^n (x_{ik} - x_{jk})^2} \quad (2.22)$$

όπου $x_i, x_j \in R$, είναι τα μετρούμενα επίπεδα έκφρασης. Η Ευκλείδεια απόσταση r μετρά τη διαφορά στα επιμέρους μεγέθη του κάθε γονιδίου. Η Ευκλείδεια απόσταση λαμβάνει τη διαφορά μεταξύ των δύο γονιδίων άμεσα. Θα πρέπει επομένως να χρησιμοποιούνται μόνο για τα δεδομένα που είναι κατάλληλα κανονικοποιημένα. Η Ευκλείδεια απόσταση περιέχει πληροφορίες τόσο για την κατεύθυνση και τάξη μεγέθους του κάθε διανύσματος από το κέντρο βάρους της συστάδας. Ως εκ τούτου, τα γονίδια μπορούν να θεωρηθούν όμοια με βάση την Ευκλείδεια απόσταση αν και είναι πολύ ανόμοια ως προς το σχήμα τους και αντίστροφα.

Η συσχέτιση Pearson μετρά την ομοιότητα μεταξύ δύο προφίλ. Κάθε κέντρο βάρους είναι ο μέσος όρος των σημείων της συστάδας, αφού τα σημεία αυτά κεντραριστούν και κανονικοποιηθούν με μέσο όρο μηδέν και τυπική απόκλιση ίση με μονάδα. Αν έχουμε δύο σειρές αριθμών $A_i = \{x_{ik}\}$ και $A_j = \{x_{jk}\}$, $k = 1, \dots, n$ και $i, j \in \{1, \dots, n\}$, $i \neq j$, η συσχέτιση του Pearson του (κέντρο), ορίζεται ως εξής:

$$r(A_i, A_j) = \frac{\sum_{k=1}^n (x_{ik} - m_i)(x_{jk} - m_j)}{\sqrt{\sum_{k=1}^n (x_{ik} - m_i)^2} \sqrt{\sum_{k=1}^n (x_{jk} - m_j)^2}} \quad (2.23)$$

όπου m_i και m_j είναι ο μέσος όρος των x_{ik} and x_{jk} , $k = 1, \dots, n$, αντιστοίχως. Ο συντελεστής συσχέτισης Pearson είναι πάντα μεταξύ -1 και 1, με τη μονάδα να δηλώνει ότι οι δύο σειρές είναι πανομοιότυπες, το μηδέν να δηλώνει ότι είναι εντελώς ασυσχέτιστες, και -1 δηλώνει ότι είναι τελείως αντίθετες.

Η απόσταση συσχέτισης στην ομαδοποίηση θεωρεί κάθε γονίδιο ως τυχαία μεταβλητή με n παρατηρήσεις και τα μέτρα την ομοιότητα μεταξύ των δύο γονιδίων, υπολογίζοντας τη γραμμική σχέση μεταξύ των κατανομών των δύο αντίστοιχων τυχαίων μεταβλητών. Ο συντελεστής συσχέτισης μένει αναλλοίωτος με γραμμικό μετασχηματισμό των δεδομένων. Η απόσταση συσχέτισης είναι ανεξάρτητη της ενίσχυσης της έκφρασης (παραβλέπει διαφορές στην τάξη μεγέθους) και λαμβάνει υπόψη τις τάσεις της αλλαγής. Ο συντελεστής συσχέτισης Pearson μπορεί να χρησιμοποιηθεί για ολοκληρωτικά και αναλογικά κλιμακούμενα δεδομένα. Το μόνο μειονέκτημα της χρήσης συσχέτισης είναι ότι είναι λίγο πιο αργή από την Ευκλείδεια απόσταση.

Η μετρική απόσταση συνημίτονου ορίζεται ως ένα μείον το συνημίτονο της γωνίας που περιλαμβάνονται μεταξύ των σημείων. Κάθε κέντρο βάρους είναι ο μέσος όρος των σημείων της συστάδας, μετά την κανονικοποίηση των σημείων αυτών σε μοναδιαίο ευκλείδειο μήκος. Δεδομένων δύο σειρών αριθμών A_i και A_j , $i, j \in \{1, \dots, n\}$, $i \neq j$, η μετρική απόσταση συνημίτονου ορίζεται ως εξής:

$$r(A_i, A_j) = 1 - \frac{\sum_{k=1}^n x_{ik} x_{jk}}{\sqrt{\sum_{k=1}^n x_{ik}^2} \sqrt{\sum_{k=1}^n x_{jk}^2}} \quad (2.24)$$

όπου

$$\frac{\sum_{k=1}^n x_{ik} x_{jk}}{\sqrt{\sum_{k=1}^n x_{ik}^2} \sqrt{\sum_{k=1}^n x_{jk}^2}} \quad (2.25)$$

είναι η αποκεντρωμένη συσχέτιση Pearson, το οποίο είναι ίσο με το συνημίτονο της γωνίας των δύο n-διαστάσεων διανυσμάτων A_i και A_j . Η Εξ. (2.25) είναι η ίδια συνάρτηση με την Εξ. (2.23), εκτός από το ότι θεωρεί ότι ο μέσος όρος είναι μηδέν, ακόμη και όταν δεν είναι. Η αποκεντρωμένη απόσταση συσχέτισης Pearson μετρά την ομοιότητα του σχήματος μεταξύ δύο προφίλ, αλλά μπορεί επίσης να εντοπίζει αντίστροφες συσχετίσεις. Ειδικότερα, συνδυάζοντας K-means με την μετρική απόσταση της Εξ. (2.25) μπορεί να οδηγήσει σε μη διαισθητικά κέντρα βάρους, επειδή μπορεί ομαδοποιήσει μη σχετιζόμενα αντικείμενα.

Η απόσταση Cityblock, γνωστή και ως απόσταση Manhattan, εξετάζει τις απόλυτες διαφορές μεταξύ των συντεταγμένων ενός ζεύγους σημείων. Κάθε κέντρο βάρους είναι ο μέσος όρος των σημείων της συστάδας. Με δεδομένα δύο γονίδια A_i και A_j , $i, j \in \{1, \dots, n\}$, $i \neq j$, η απόσταση Cityblock ορίζεται ως εξής:

$$r(A_i, A_j) = \frac{1}{n} \sqrt{\sum_{k=1}^n |x_{ik} - x_{jk}|} \quad (2.26)$$

Οι λόγοι πίσω από την δημοτικότητα του αλγορίθμου K-means είναι οι εξής:

(A) Είναι απλός και χωρίς επίβλεψη εκπαιδευόμενος αλγόριθμος. Τα K-means δεν περιορίζεται σε έλεγχο υπόθεσης με βάση κάποια προηγούμενη εκπαίδευση. Διερευνά τη δομή των δεδομένων με βάση τις ομοιότητες που παρουσιάζονται σε αυτά.

(B) Ο χρόνος που απαιτείται είναι $O(nML)$, όπου n είναι ο αριθμός των προτύπων, M είναι ο αριθμός των συστάδων και L είναι ο αριθμός των επαναλήψεων που λαμβάνονται από τον αλγόριθμο μέχρι να συγκλίνει. Οι παράμετροι M και L καθορίζονται εκ των προτέρων. Ως εκ

τούτου, η προσέγγιση K-means έχει γραμμική πολυπλοκότητα χρόνου ως προς το μέγεθος των δεδομένων.

(Γ) Η πολυπλοκότητα του χώρου είναι $O(m + n)$. Αποθηκεύονται μόνο τα δεδομένων και τα κέντρα βάρους. Ο αλγόριθμος απαιτεί πρόσθετο χώρο για να αποθηκεύσετε των πίνακα δεδομένων.

(Δ) Είναι ανεξάρτητος της διάταξης. Έτσι, για τις ίδιες αρχικές συνθήκες, δημιουργεί τον ίδιο διαχωρισμό των δεδομένων, ανεξάρτητα από τη σειρά με την οποία τα πρότυπα εισάγονται στον αλγόριθμο.

(Ε) Με μεγάλο αριθμό μεταβλητών, K-means μπορεί να είναι υπολογιστικά πιο γρήγορα.

(Στ) Μπορεί να παράγει πιο συμπαγείς συστάδες.

(Ζ) Μπορεί να χρησιμοποιηθεί για μια ευρεία ποικιλία τύπων δεδομένων.

(Η) Είναι σε θέση να βρει καθαρή υπο-ομάδα εάν έχει οριστεί αρκετά μεγάλος αριθμός συστάδων.

(Ι) Παρόλο που δεν υπάρχει καμία εγγύηση για την επίτευξη μιας ολικού ελάχιστου, η σύγκλιση του αλγορίθμου είναι εξασφαλισμένη.

Το μεγάλο μειονέκτημα της ομαδοποίησης K-means είναι η ευαισθησία του αλγορίθμου στην επιλογή των αρχικών συνθηκών. Μια τεχνική που χρησιμοποιείται για την αντιμετώπιση του προβλήματος της επιλογής αρχικής κέντρων είναι η πολλαπλές εκτελέσεις, που απορρέουν από την υψηλή ταχύτητα εκτέλεσης του αλγορίθμου, η καθεμία με διαφορετικό σύνολο τυχαία επιλεγμένων αρχικών κέντρων, και στη συνέχεια η επιλογή των συστάδων με το ελάχιστο σφάλμα ομαδοποίησης. Αυτή η στρατηγική μπορεί να μην λειτουργεί πολύ καλά, ανάλογα με το σύνολο των δεδομένων και τον αριθμό των συστάδων που έχουν ζητηθεί.

Λόγω του προβλήματος της χρήσης τυχαίων αρχικών κέντρων, το οποίο ακόμα και οι επαναλαμβανόμενες ίσως να μην υπερβούν, έχουν προταθεί άλλες τεχνικές για την αρχικοποίηση. Στη βιβλιογραφία, πολλές τεχνικές έχουν προταθεί πολλές τεχνικές για την επιλογή των καλύτερων αρχικές θέσεις των κέντρων βάρους, με σκοπό να επιτευχθεί μια υψηλής ποιότητας λύση στο τέλος της εκτέλεσης των K-Means. Ο Milligan (Milligan, 1980) εξέτασε την ισχυρή εξάρτηση των K-Means από τους αρχικούς όρους και υπέδειξε ότι μπορούν να προκύψουν καλές τελικές δομές τελικών συστάδων με την ιεραρχική μέθοδο του Ward (Ward, 1963). Οι Likas et al. πρότειναν μια βηματική προσέγγιση για την ομαδοποίηση, που ονομάζεται global K-Means (Likas, Vlasis & Verbeek, 2003). Τα αρχικά κέντρα παράγονται με μια μέθοδο δέντρων. Ο αλγόριθμος είναι ντετερμινιστικός, δεν εξαρτάται από οποιαδήποτε αρχική θέση για το κέντρο της συστάδας και δεν περιέχει καμία εμπειρική παράμετρο. Οι Bradley et al. παρουσιάζουν τον αλγόριθμο βελτιστοποίησης (Bradley & Fayyad, 1998) που λειτουργεί με μικρό αριθμό επιμέρους δειγμάτων από μια συγκεκριμένη βάση δεδομένων. Τα δεδομένα σε κάθε επιμέρους δείγμα είναι ομαδοποιημένα. Στη συνέχεια, όλα τα κέντρα όλων των επιμέρους δειγμάτων ομαδοποιούνται με K-Means. Τα κέντρα του κάθε επιμέρους δείγματος που χρησιμοποιούνται ως αρχικά κέντρα. Τα τελικά κέντρα που δίνουν ελάχιστη σφάλμα ομαδοποίησης στη συνέχεια χρησιμοποιούνται ως αρχικές συνθήκες στην ομαδοποίηση με K-Means του αρχικού συνόλου δεδομένων.

Οι προαναφερόμενοι αλγόριθμοι αρχικοποίησης είναι σε θέση να οδηγήσει στον υπολογισμό καλών αρχικών κέντρων. Δεν συνιστούν απλώς μεθόδους αρχικοποίησης. Πρόκειται για μεθόδους ομαδοποίησης καθαυτές και ορισμένες από αυτές χρησιμοποιούν τον αλγόριθμο K-Means ως μέρος των αλγορίθμων τους. Ως εκ τούτου, υποφέρουν από το ίδιο πρόβλημα με την προσέγγιση των K-Means.

2.4 Μηχανές Διανυσματικής Υποστήριξης (SVM)

2.4.1 Βασικά Στοιχεία Θεωρίας Μηχανών Διανυσματικής Υποστήριξης (SVM)

Η θεωρία των Μηχανών Διανυσματικής Υποστήριξης (Support Vector Machine - SVM) παρουσιάστηκε συστηματικά από τους Vapnik και Chervonenkis προς το τέλος της δεκαετίας του '60. Ωστόσο, δεν ήταν παρά μετά τα μέσα της δεκαετίας του 1990 που οι μέθοδοι που χρησιμοποιούνται για SVMs άρχισαν να ανακάμπτουν χάρη στην μεγάλη διαθεσιμότητα της υπολογιστικής δύναμης, που οδήγησε σε πολλές πρακτικές εφαρμογές. Η βασική ιδέα της SVM [Vapnik 1995], [Cristianini and Shawe-Taylor 2000Vapnik 1998] είναι να μετατρέπει το σήμα σε έναν χώρο χαρακτηριστικών ανώτερων διαστάσεων και να βρίσκει το βέλτιστο υπερεπίπεδο στο χώρο που μεγιστοποιεί το περιθώριο μεταξύ των τάξεων. Εν συντομία, το SVM λύνει ένα δυαδικό πρόβλημα όπου τα δεδομένα χωρίζονται από ένα υπερεπίπεδο. Το υπερεπίπεδο ορίζεται μέσω της χρήσης των Διανυσμάτων Υποστήριξης (Support Vectors), τα οποία αποτελούν ένα υποσύνολο των διαθέσιμων δεδομένων για τις δύο περιπτώσεις και καθορίζουν τα όρια μεταξύ των δύο κατηγοριών.

2.4.2 Γραμμικές Μηχανές Διανυσματικής Υποστήριξης

Αρχικά παρουσιάζεται η πιο απλή περίπτωση: γραμμικές μηχανές που εκπαιδεύονται με διαχωρίσιμα δεδομένα (όπως παρουσιάζεται παρακάτω, η ανάλυση της γενικής περίπτωσης των μη γραμμικών μηχανών που εκπαιδεύονται με μη διαχωρίσιμα δεδομένα καταλήγει σε ένα παρόμοιο πρόβλημα τετραγωνικού προγραμματισμού). Για τα δεδομένα εκπαίδευσης χρησιμοποιείται ο ακόλουθος συμβολισμός $\{\mathbf{x}_i, y_i\}, i = 1, K, l, y_i \in \{-1, 1\}, \mathbf{x}_i \in \mathbf{R}^d$.

Θεωρούνται υπερεπίπεδα τα οποία διαχωρίζουν τα θετικά από τα αρνητικά παραδείγματα (ένα υπερεπίπεδο διαχωρισμού). Τα σημεία \mathbf{x} που βρίσκονται στο υπερεπίπεδο ικανοποιούν την σχέση $\mathbf{w} \cdot \mathbf{x} + b = 0$, όπου το \mathbf{w} είναι κάθετο στο υπερεπίπεδο, $|b|/\|\mathbf{w}\|$ είναι η κατακόρυφη απόσταση του υπερεπιπέδου από την αρχή των αξόνων και $\|\mathbf{w}\|$ είναι η Ευκλείδεια Νόρμα του \mathbf{w} . Ορίζεται ως d_+ (d_-) ως η μικρότερη απόσταση του υπερεπιπέδου διαχωρισμού από το κοντινότερο θετικό (αρνητικό) παράδειγμα. Επίσης ορίζεται το “περιθώριο” του υπερεπιπέδου διαχωρισμού ίσο με $d_+ + d_-$. Για την γραμμικά διαχωρίσιμη περίπτωση, ο αλγόριθμος διανυσμάτων υποστήριξης απλά

αναζητεί το διαχωριστικό επίπεδο με το μέγιστο περιθώριο. Το γεγονός αυτό μπορεί να διατυπωθεί ως εξής: θεωρείται ότι όλα τα δεδομένα εκπαίδευσης ικανοποιούν τους παρακάτω περιορισμούς:

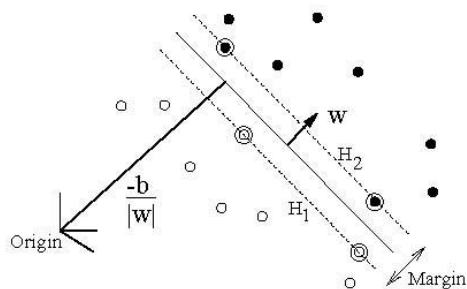
$$\mathbf{x}_i \cdot \mathbf{w} + b \geq +1 \text{ για } y_i = +1 \quad (2.27)$$

$$\mathbf{x}_i \cdot \mathbf{w} + b \leq -1 \text{ για } y_i = -1 \quad (2.28)$$

Οι δύο αυτές σχέσεις μπορούν να συνδυαστούν σε ένα σύνολο ανισοτήτων:

$$y_i (\mathbf{x}_i \cdot \mathbf{w} + b) - 1 \geq 0 \quad \forall i \quad (2.29)$$

Θεωρούνται τα σημεία για τα οποία ισχύει η ισότητα στη σχέση (2.28) (ζητώντας η ύπαρξη ενός τέτοιου τέτοιου σημείου είναι ισοδύναμη με την επιλογή της κλίμακας του \mathbf{w} και του b). Τα σημεία αυτά βρίσκονται στο υπερεπίπεδο $H_1: \mathbf{x}_i \cdot \mathbf{w} + b = 1$ με κάθετο \mathbf{w} και κατακόρυφη απόσταση από την αρχή των αξόνων $|1 - b| / \|\mathbf{w}\|$. Παρόμοια, τα σημεία για τα οποία ισχύει η Σχέση (2.28) βρίσκονται στο υπερεπίπεδο $H_2: \mathbf{x}_i \cdot \mathbf{w} + b = -1$, με κάθετο \mathbf{w} πάλι και κατακόρυφη απόσταση από την αρχή των αξόνων $|-1 - b| / \|\mathbf{w}\|$. Κατά συνέπεια $d_+ = d_- = 1 / \|\mathbf{w}\|$ και το περιθώριο είναι ίσο με $2 / \|\mathbf{w}\|$. Σημειώνεται ότι τα επίπεδα H_1 και H_2 είναι παράλληλα (έχουν την ίδια κάθετη) και ότι κανένα σημείο εκπαίδευσης δεν βρίσκεται μεταξύ τους. Κατά συνέπεια μπορεί να βρεθεί ένα ζεύγος υπερεπιπέδων που δίνει το μέγιστο περιθώριο ελαχιστοποιώντας το $\|\mathbf{w}\|^2$, υποκείμενο στους περιορισμούς (2.29).



Σχήμα 2-8 : Γραμμικά Υπερεπίπεδα Διαχωρισμού για την διαχωρίσιμη περίπτωση. Οι Μηχανές Υποστήριξης είναι κυκλωμένες.

Σημειώνεται ότι η λύση αυτού του προβλήματος βελτιστοποίησης είναι ολική. Το γεγονός αυτό αποτελεί ένα εξαιρετικό πλεονέκτημα συγκριτικά με τα MLPs ή τα RBF δίκτυα τα οποία μπορεί να έχουν πολλά τοπικά ελάχιστα και κατά συνέπεια συνήθως δεν είναι εγγυημένη η ολική βέλτιστη λύση.

Η λύση μιας τυπικής περίπτωσης δύο διαστάσεων αναμένεται να έχει την μορφή που παρουσιάζεται στο Σχήμα 2-8. Τα σημεία εκπαίδευσης για τα οποία ισχύει η ισότητα της σχέσης (2.29) (δηλαδή, εκείνα τα οποία καταλήγουν σε ένα από τα υπερεπίπεδα H_1 και H_2) και η αφαίρεση τους θα άλλαζε την λύση που βρέθηκε, ονομάζονται Διανύσματα Υποστήριξης και παρουσιάζονται στο Σχημα 2-8 με επιπλέον κύκλους.

Στη συνέχεια χρησιμοποιείται μια διατύπωση Lagrange του προβλήματος. Υπάρχουν δύο λόγοι για αυτό. Ο πρώτος λόγος είναι ότι οι περιορισμοί (2.29) θα αντικατασταθούν από περιορισμούς στους ίδιους τους πολλαπλασιαστές Lagrange, που είναι πιο εύκολα χειριζόμενοι. Ο δεύτερος λόγος είναι ότι στην αναδιατύπωση του προβλήματος, τα δεδομένα εκπαίδευσης θα εμφανίζονται μόνο (στους υπάρχοντες αλγορίθμους εκπαίδευσης και ελέγχου) στην μορφή εσωτερικών πηλίκων μεταξύ των διανυσμάτων. Η ιδιότητα αυτή είναι ζωτικής σημασίας, γιατί επιτρέπει την γενίκευση της διαδικασίας στην μη γραμμική περίπτωση.

Κατά συνέπεια, εισάγονται θετικοί πολλαπλασιαστές Lagrange $a_i, i = 1, K, l$, ένας για κάθε έναν από τους περιορισμούς ανισότητας (2.29). Υπενθυμίζεται ο κανόνας σύμφωνα με τον οποίο για περιορισμούς της μορφής $c_i \geq 0$, οι σχέσης περιορισμού πολλαπλασιάζονται με θετικούς

πολλαπλασιαστές Lagrange και αφαιρούνται από την αντικειμενική συνάρτηση για να σχηματιστεί η Lagrangian συνάρτηση. Για περιορισμούς ισότητας, οι πολλαπλασιαστές Lagrange δεν δεσμεύονται από περιορισμούς. Η Lagrangian συνάρτηση λαμβάνει την μορφή:

$$L_P = \frac{1}{2} \|w\|^2 - \sum_{i=1}^l a_i y_i (\mathbf{x}_i \cdot \mathbf{w} + b) + \sum_{i=1}^l a_i \quad (2.30)$$

Η L_P ελαχιστοποιείται ως προς \mathbf{w} , b και ταυτόχρονα απαιτείται οι παράγωγοι της L_P ως προς όλα τα a_i να εξαφανίζονται, υποβαλλόμενα όλα στους περιορισμούς $a_i \geq 0$ (το συγκεκριμένο σύνολο περιορισμών θα ονομάζεται C_1). Το πρόβλημα πλέον είναι ένα κυρτό πρόβλημα τετραγωνικού προγραμματισμού καθώς η αντικειμενική συνάρτηση είναι κυρτή και τα σημεία που ικανοποιούν τους περιορισμούς σχηματίζουν επίσης ένα κυρτό σύνολο (κάθε γραμμικός περιορισμός ορίζει έναν κυρτό σύνολο και ένα σύνολο N ταυτόχρονων γραμμικών περιορισμών ορίζει την τομή των N κυρτών συνόλων, το οποίο είναι επίσης ένα κυρτό σύνολο). Το γεγονός αυτό σημαίνει ότι μπορεί ισοδύναμα να επιλυθεί το ακόλουθο “διττό” (“dual”) πρόβλημα:

Μεγιστοποίηση της Lagrangian συνάρτησης L_P , υπό τους περιορισμούς ότι η παράγωγος της L_P ως προς το \mathbf{w} και το b εξαφανίζεται και υπόκειται επίσης στους περιορισμούς ότι $a_i \geq 0$ (το ιδιαίτερο αυτό σύνολο των περιορισμών θα αναφέρεται ως C_2). Η διττή αυτή διατύπωση του προβλήματος ονομάζεται διττό Wolfe [Fletcher 1987]. Παρουσιάζει την ιδιότητα, το μέγιστο της Lagrangian συνάρτησης L_P , υπό τους περιορισμούς C_2 , εμφανίζεται στις ίδιες τιμές των \mathbf{w} , b και a , όπου εμφανίζεται το ελάχιστο της Lagrangian συνάρτησης L_P , υπό τους περιορισμούς C_1 .

Η απαίτηση, να εξαφανίζεται η παράγωγος της Lagrangian συνάρτησης L_P ως αναφορά τα \mathbf{w} και b , δίνει τις παρακάτω συνθήκες:

$$\mathbf{w} = \sum_i a_i y_i \mathbf{x}_i \quad (2.31)$$

$$\sum_i a_i y_i = 0 \quad (2.32)$$

Εφόσον οι σχέσεις αυτές είναι περιορισμοί στην διττή διατύπωση, μπορούν να αντικατασταθούν στην σχέση 2-30:

$$L_D = \sum_i a_i - \frac{1}{2} \sum_{i,j} a_i a_j y_i y_j \mathbf{x}_i \cdot \mathbf{x}_j \quad (2.33)$$

Σημειώνεται ότι έχει δοθεί διαφορετικό σύμβολο στην συνάρτηση Lagrangian (P για το αρχικό (primal), D για διττό (dual)) για να τονιστεί ότι οι δύο διατυπώσεις είναι διαφορετικές: οι L_P και L_D προέρχονται από την ίδια αντικειμενική συνάρτηση αλλά με διαφορετικούς περιορισμούς και η λύση βρίσκεται ελαχιστοποιώντας την L_P ή μεγιστοποιώντας την L_D . Σημειώνεται επίσης ότι αν το πρόβλημα διατυπωθεί με $b=0$, που καταλήγει να απαιτεί όλα τα υπερεπίπεδα να περιλαμβάνουν την αρχή των αξόνων, ο περιορισμός (2.32) δεν εμφανίζεται. Αυτό αποτελεί έναν ήπιο περιορισμό για χώρους υψηλών διαστάσεων, καθώς μειώνει τον αριθμό των βαθμών ελευθερίας κατά έναν.

Η εκπαίδευση των Διανυσμάτων Υποστήριξης (για την γραμμικά διαχωρίσιμη περίπτωση καταλήγει κατά συνέπεια στην μεγιστοποίηση της L_D ως προς τα a_i , υπό τους περιορισμούς (2.32) και την θετικότητα των a_i , ενώ η λύση δίνεται από την σχέση (2.31). Σημειώνεται ότι υπάρχει ένας πολλαπλασιαστής Lagrange a_i , για κάθε σημείο εκπαίδευσης. Στην λύση, τα σημεία εκείνα για τα οποία $a_i > 0$ καλούνται Διανύσματα Υποστήριξης και βρίσκονται σε ένα από τα υπερεπίπεδα H_1 , H_2 . Όλα τα υπόλοιπα σημεία εκπαίδευσης έχουν $a_i = 0$ και βρίσκονται πάνω στο υπερεπίπεδο H_1 ή στο H_2 (έτσι ώστε να ισχύει η ισότητα (2.29) ή σε εκείνη την πλευρά του H_1 ή του H_2 ώστε να ισχύει η αυστηρή ανισότητα της σχέσης (2.29)). Για τις μηχανές αυτές, τα Διανύσματα Υποστήριξης είναι τα κρίσιμα στοιχεία του συνόλου εκπαίδευσης και δίνουν το όνομα στην μηχανή μάθησης γιατί αποδεικνύονται πολύ σημαντικά για την ταξινόμηση όσον αφορά την γενίκευση και τους υπολογισμούς. Τοποθετούνται κοντά στο όριο απόφασης και εάν όλα τα υπόλοιπα σημεία εκπαίδευσης αφαιρεθούν (ή μετακινηθούν χωρίς όμως να τέμνουν το H_1 ή το H_2), η επανάληψη της εκπαίδευσης θα οδηγήσει στην εύρεση του ίδιου υπερεπιπέδου διαχωρισμού. Προσφέρουν μια αραιή λύση στο πρόβλημα της ταξινόμησης ενώ ο Vapnik παρουσίασε ένα απλό και ισχυρό αποτέλεσμα που συνδέει τον αριθμό των Διανυσμάτων Υποστήριξης και την δυνατότητα γενίκευσης του ταξινομητή [Vapnik 1995]. Εάν τα δεδομένα εκπαίδευσης περιλαμβάνουν M δείγματα που διαχωρίζονται από τον υπερεπίπεδο μέγιστου

περιθωρίου, το φράγμα της αναμενόμενης πιθανότητας του σφάλματος ελέγχου μπορεί να υπολογιστεί ως:

$$E[P_{error}] \leq \frac{\#SV}{M} \quad (2.34)$$

όπου #SV είναι ο αριθμός των SVs. Η σχέση αυτή δίνει έναν εύκολο τρόπο για την εκτίμηση της ικανότητας γενίκευσης του ταξινομητή με ένα φράγμα που είναι ανεξάρτητο από την διαστασιμότητα του χώρου εισόδου.

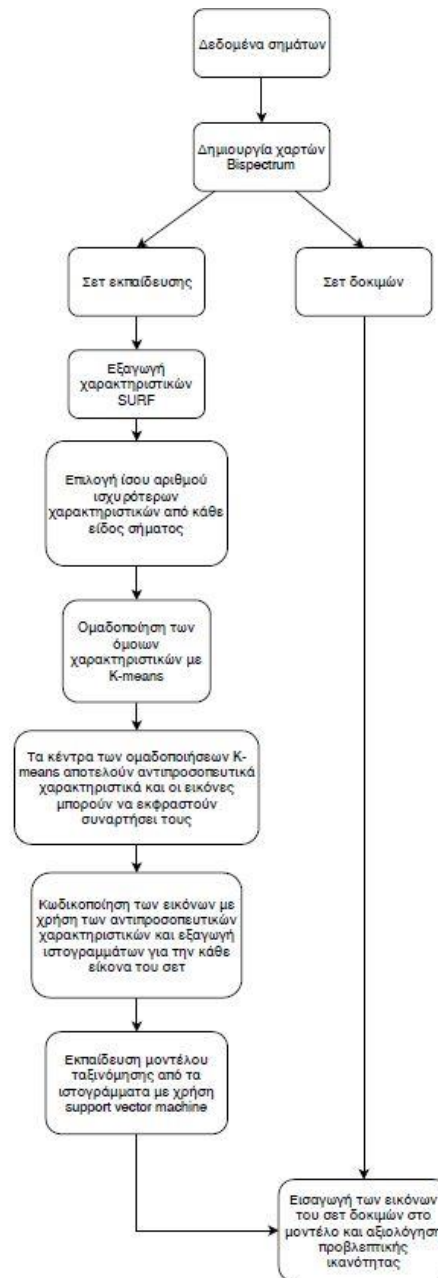
2.4.3 Ταξινόμηση Δεδομένων

Μετά την εκπαίδευση της Μηχανής Διανυσματικής Υποστήριξης, κάθε ελεγχόμενο στοιχείο καθορίζεται σε ποια πλευρά του ορίου απόφασης (δηλαδή του υπερεπιπέδου που βρίσκεται στην μέση απόσταση μεταξύ των H1 και H2 και παράλληλο με αυτά) βρίσκεται και ταξινομείται στην Κλάση εκείνη σύμφωνα με τον υπολογισμό της παρακάτω σχέσης:

$$f(x) = \text{sgn} \left(\sum_{i=1}^{N_S} a_i y_i \mathbf{x}_i \cdot \mathbf{x} + b \right) \quad (2.35)$$

ΚΕΦΑΛΑΙΟ 3 – ΠΕΡΙΓΡΑΦΗ ΜΕΘΟΔΟΛΟΓΙΑΣ

Στο παρόν κεφάλαιο γίνεται περιγραφή των βημάτων που θα εφαρμοστούν για την δημιουργία του μοντέλου ταξινόμησης των σημάτων. Το σχήμα 3-1 απεικονίζει τη ροή της διαδικασίας.

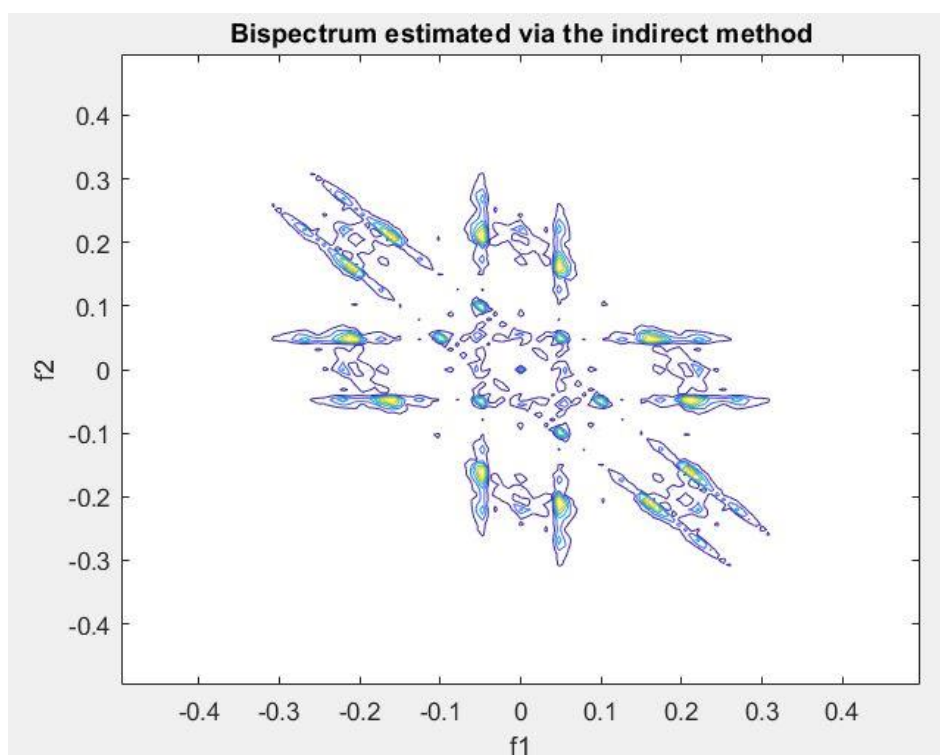


Σχήμα 3-1 : Διάγραμμα ροής μεθοδολογίας

3.1 Εξαγωγή εικόνων bispectrum και διαχωρισμός σε ομάδες εκπαίδευσης και δοκιμών

Για την εξαγωγή των εικόνων bispectrum θα χρησιμοποιείται η συνάρτηση *[bispeci]*. Οι εικόνες που εξάγονται αποθηκεύονται σε αντίστοιχους φακέλους προκειμένου να επεξεργαστούν στα επόμενα βήματα. Οι άξονες των εικόνων αποτελούνται από συχότητες που εξήχθησαν από μετασχηματισμό Fourier των σημάτων, μήκους $nfft=127$. Οι συχότητες έχουν κανονικοποιηθεί από -0.5 έως 0.5.

Το σχήμα 3-2 απεικονίζει την εξαγόμενη εικόνα bispectrum ενός τυχαίου σήματος.



Σχήμα 3-2 : Παράδειγμα απεικόνισης bispectrum σήματος

Προκειμένου να επιτευχθεί η εκπαίδευση και η αξιολόγηση του αλγορίθμου ταξινόμησης, οι εικόνες χωρίζονται σε δύο σετ με αναλογία 60 – 40 με χρήση της εντολής *[splitEachLabel]*. Το σετ «εκπαίδευσης» θα χρησιμοποιηθεί για την εξαγωγή κυρίαρχων χαρακτηριστικών με τα οποία θα γίνει ο σχεδιασμός του αλγόριθμου ταξινόμησης. Το σετ «δοκιμών» θα χρησιμοποιηθεί για να ελεγχθεί η ακρίβεια του αλγορίθμου.

3.2 Εξαγωγή χαρακτηριστικών SURF

Για την εξαγωγή των χαρακτηριστικών κάθε εικόνας θα χρειαστεί αρχικά να εντοπιστούν τα σημεία ενδιαφέροντος που περιλαμβάνει. Αυτά τα σημεία μπορεί να αποτελούν διακυμάνσεις φωτεινότητας, σχηματισμοί, γωνίες ή σημεία έντονου χρωματισμού.

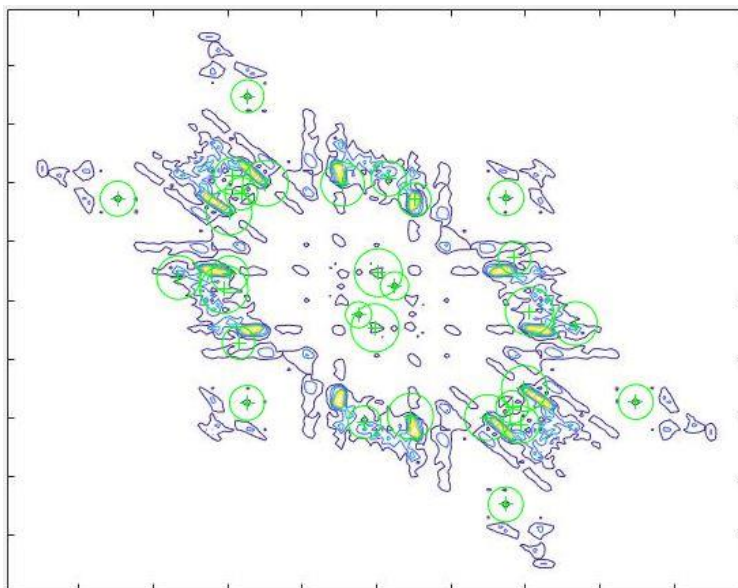
Η συνάρτηση εντοπισμού λαμβάνει ως είσοδο αποκλειστικά εικόνες στην κλίμακα του γκρι. Με αυτόν τον τρόπο όλες οι εικόνες μπορούν να εκφραστούν ως πίνακες δύο διαστάσεων. Ο εντοπισμός γίνεται ελέγχοντας το κάθε στοιχείο του πίνακα και καθορίζει αν αυτό αποτελεί σημείο ενδιαφέροντος βασισμένο στην θεωρία SURF και τις αντίστοιχες παραμέτρους.

Η συνάρτηση εντοπισμού σημείων ενδιαφέροντος [*detectSURFFeatures*] έχει τις εξής ρυθμίσεις:

- Όριο ισχυρότερων χαρακτηριστικών ('MetricThreshold'): 1000
- Αριθμός οκτάβων ('NumOctaves'): 3
- Αριθμός κλιμάκων ανά οκτάβα ('NumScaleLevels'): 4

Καθώς η συνάρτηση εντοπισμού των σημείων μπορεί να επιστρέψει πολύ μεγάλο αριθμό σημείων, κάποια από τα οποία μπορεί να μην είναι δυνατόν να χρησιμοποιηθούν, η συνάρτηση περιλαμβάνει ένα όριο κάτω από το οποίο, τα χαρακτηριστικά αγνοούνται. Όσο μικρότερη είναι η παράμετρος 'MetricThreshold', τόσο περισσότερα χαρακτηριστικά θα εντοπιστούν στην εικόνα.

Το σχήμα 3-3 απεικονίζει κάποια από τα εξαγόμενα χαρακτηριστικά ενός τυχαίου σήματος.



Σχήμα 3-3 : Απεικόνιση κάποιων από τα ισχυρότερα χαρακτηριστικά ενός τυχαίου σχήματος

Αυτά τα σημεία θα πρέπει στη συνέχεια να εκφραστούν ως διανύσματα περιγραφής, προκειμένου να χρησιμοποιηθούν στη συνέχεια από το μοντέλο ταξινόμησης.

Η συνάρτηση εξαγωγής των διανυσμάτων περιγραφής [*extractFeatures*] για το κάθε σημείο έχει τις εξής ρυθμίσεις:

- Μέθοδος εξαγωγής διανύσματος περιγραφής ('Method'): SURF
- Μήκος διανύσματος περιγραφής ('FeatureSize'): 64
- Δείκτης μη περιστροφικής μεταβολής ('Upright'): true

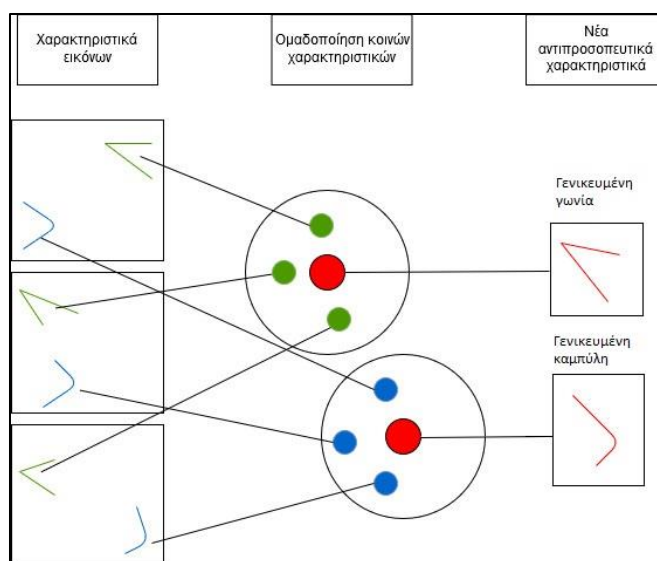
Όταν η τρίτη επιλογή είναι αληθής, δεν περιλαμβάνεται ο προσανατολισμός του κάθε σημείου στο αντίστοιχο διάνυσμα περιγραφής. Στη παρούσα εφαρμογή οι εικόνες που χρησιμοποιούνται δεν υπόκεινται σε περιστροφή και για αυτό το λόγο δεν θα συμπεριληφθεί ο προσανατολισμός των σημείων στην εξαγωγή των διανυσμάτων περιγραφής.

3.3 Ομαδοποίηση χαρακτηριστικών με την μέθοδο K-Means και δημιουργία «οπτικού λεξιλογίου»

Προκειμένου να απομονωθούν χαρακτηριστικά που οφείλονται σε τυχαίους παράγοντες, επιλέγεται το 80% των ισχυρότερων χαρακτηριστικών από την κάθε κατηγορία σημάτων (εσφαλμένα/κανονικά). Για να επιτευχθεί πιο εύρωστη ομαδοποίηση στο επόμενο στάδιο, επιλέγεται ίσος αριθμός ισχυρότερων χαρακτηριστικών και από τις δύο κατηγορίες.

Τα εξαγόμενα χαρακτηριστικά βρίσκονται κατανεμημένα σε χώρο 64 διαστάσεων βασισμένα στα διανύσματα περιγραφής. Τα χαρακτηριστικά που βρίσκονται σε κοντινή «απόσταση» μεταξύ τους, περιγράφουν παρεμφερή σημεία των εικόνων του σετ.

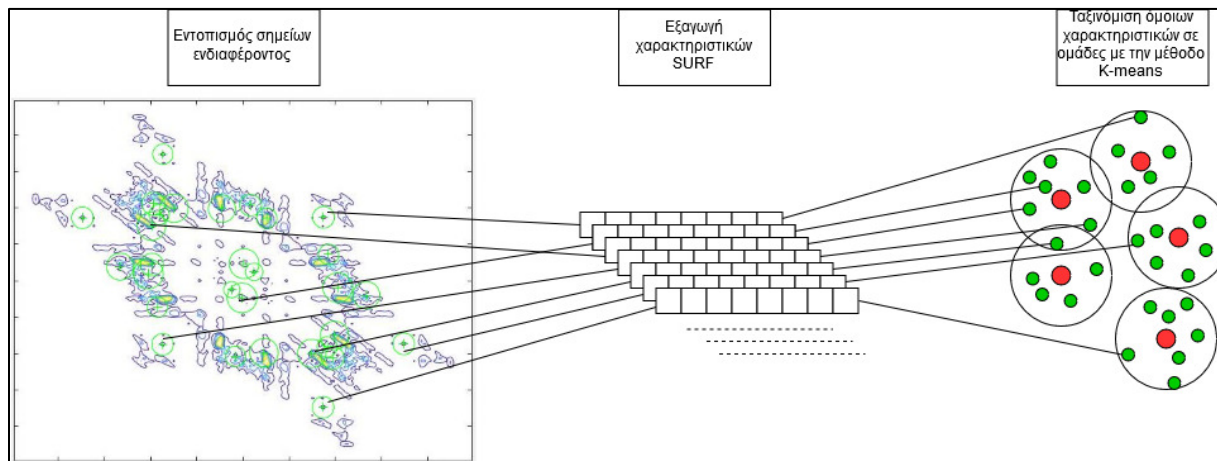
Στο σχήμα 3-4 φαίνεται ένα θεωρητικό παράδειγμα του τρόπου ομαδοποίησης.



Σχήμα 3-4 : Θεωρητικό παράδειγμα ομαδοποίησης «όμοιων» χαρακτηριστικών

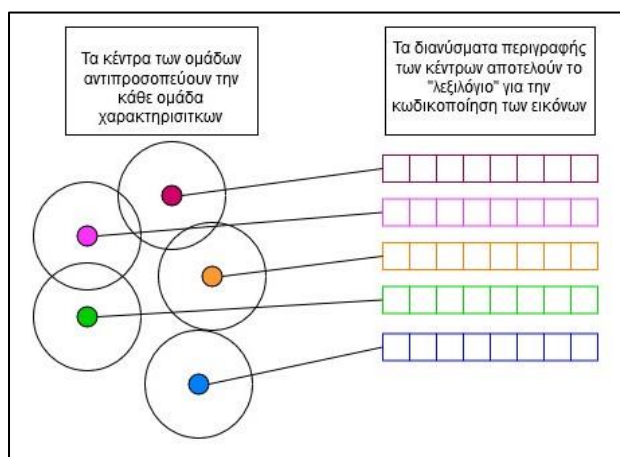
Σε αυτό το παράδειγμα, όλα τα διανύσματα περιγραφής που αφορούν «γωνίες» βρίσκονται σε κοντινότερη απόσταση μεταξύ τους στον χώρο περιγραφής σε σύγκριση με τα διανύσματα που περιγράφουν «καμπύλες». Με αυτό τον τρόπο μπορούν να δημιουργηθούν δύο ομάδες που αφορούν το κάθε χαρακτηριστικό. Το κέντρο αυτών των ομάδων θα αποτελεί επίσης ένα γενικευμένο χαρακτηριστικό που αντιπροσωπεύει ολόκληρη την ομάδα.

Η ομαδοποίηση αυτή μπορεί να γίνει χρησιμοποιώντας την μέθοδο K-means. Το κέντρο της κάθε ομαδοποίησης στον χώρο 64 διαστάσεων αποτελεί το κυρίαρχο χαρακτηριστικό της ομάδας.



Σχήμα 3-5 : Ταξινόμηση εξαγόμενων χαρακτηριστικών SURF με ομάδες K-means

Στο τέλος της διαδικασίας ομαδοποίησης με K-means έχουμε ένα σύνολο κυρίαρχων χαρακτηριστικών. Αυτά τα χαρακτηριστικά αποτελούν ένα «λεξιλόγιο» οπτικών στοιχείων τα οποία μπορεί να χρησιμοποιηθεί για την κωδικοποίηση όλων των εικόνων. Το κάθε ένα από αυτά τα στοιχεία μπορεί να θεωρηθεί ως μία «οπτική λέξη».



Σχήμα 3-6 : Δημιουργία «οπτικών λέξεων» από τα κέντρα των ομάδων K-means

Με χρήση του αυτής της μεθόδου η κάθε εικόνα θα μπορεί να περιγράφεται όχι μόνο από τα μοναδικά χαρακτηριστικά SURF που την αποτελούν, αλλά και ως συνάρτηση του «λεξιλογίου» που έχει δημιουργηθεί. Η κωδικοποίηση των εικόνων συναρτήσει του λεξιλογίου θα γίνει στο επόμενο στάδιο.

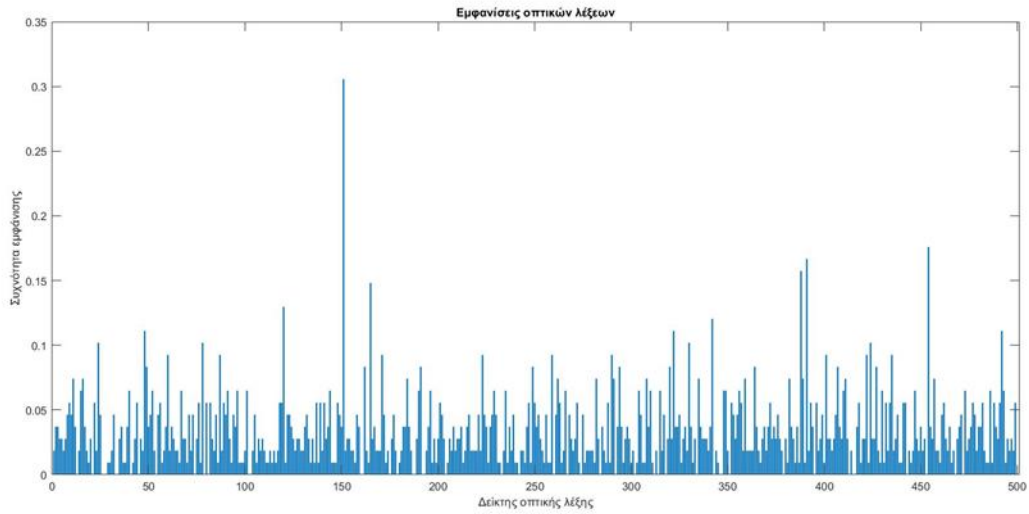
Η ομαδοποίηση που περιεγράφηκε σε αυτό το βήμα καθώς και η επιλογή των σημαντικότερων χαρακτηριστικών μπορούν να επιτευχθούν με τη συνάρτηση *[bagOfFeatures]*. Οι ρυθμίσεις της συνάρτησης είναι οι εξής:

- Αριθμός στοιχείων του «λεξιλογίου» / ομάδων K-means ('VocabularySize'): 500
- Ποσοστό ισχυρότερων χαρακτηριστικών (βήμα 4) ('StrongestFeatures'): 0.8
- Συνάρτηση εξαγωγής χαρακτηριστικών ('CustomExtractor'): extractorFcn (η συνάρτηση για εξαγωγή χαρακτηριστικών SURF που περιεγράφηκε στο βήμα 3)

3.4 Κωδικοποίηση εικόνων με χρήση του «οπτικού λεξιλογίου» και δημιουργία ταξινομητή

Η κωδικοποίησή των εικόνων του σετ εκπαίδευσης μέσω του λεξιλογίου που έχουμε κατασκευάσει γίνεται συγκρίνοντας το κάθε χαρακτηριστικό μίας εικόνας με τα κυρίαρχα χαρακτηριστικά ή «λέξεις» του λεξιλογίου.

Για τη σύγκριση χρησιμοποιείται η μέθοδος nearest neighbor και κάθε χαρακτηριστικό της εικόνας αντιστοιχίζεται σε μία «λέξη». Η συχνότητες εμφάνισης των «λέξεων» για κάθε εικόνα αποθηκεύονται σε ένα ιστόγραμμα. Αυτό το ιστόγραμμα αποτελεί ένα διάνυσμα περιγραφής της εικόνας συναρτήσει του λεξιλογίου.



Σχήμα 3-7 : Παράδειγμα ιστογράμματος εικόνας bispectrum

Η ίδια διαδικασία επαναλαμβάνεται για όλες τις εικόνες του σετ εκπαίδευσης και τα ιστογράμματα κατηγοριοποιούνται ανάλογα με την κατηγορία των σημάτων τους.

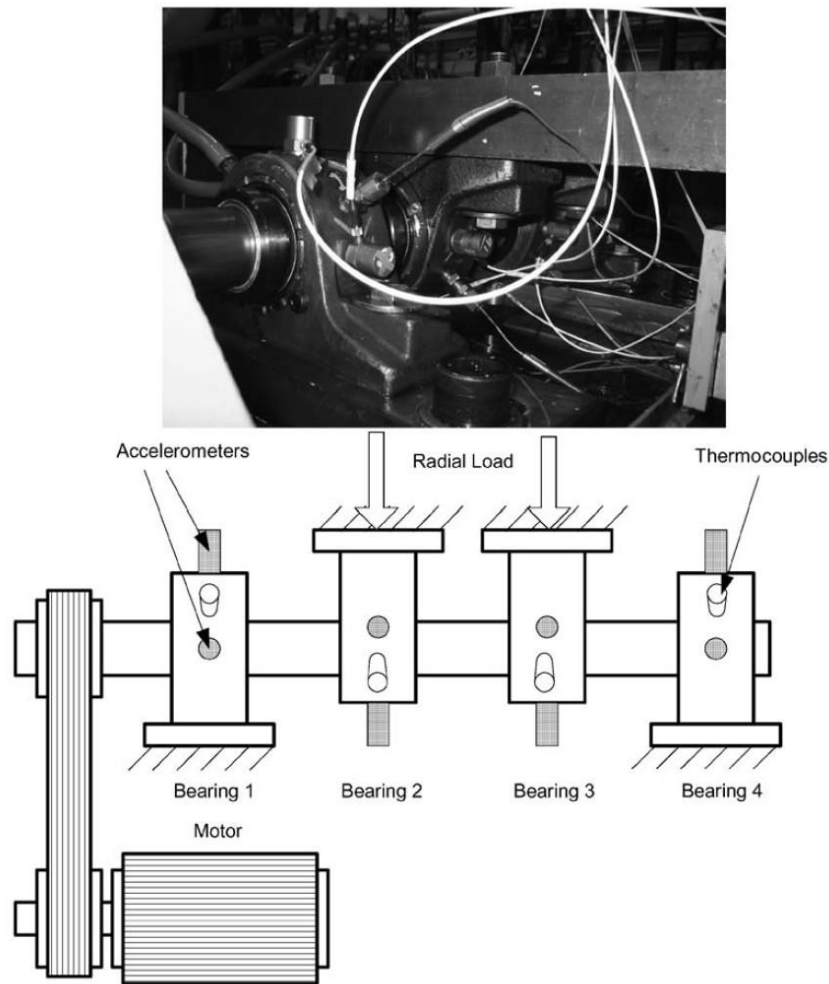
Με χρήση της συνάρτησης `[trainImageCategoryClassifier]` οι εικόνες κωδικοποιούνται και παράλληλα εκπαιδεύεται ο αλγόριθμος ταξινόμησης με χρήση support vector machine.

Μόλις δημιουργηθεί ο αλγόριθμος ταξινόμησης, οποιαδήποτε νέα εικόνα μπορεί να ταξινομηθεί με χρήση της εντολής `[predict]`.

ΚΕΦΑΛΑΙΟ 4 – ΠΕΙΡΑΜΑΤΙΚΑ ΑΠΟΤΕΛΕΣΜΑΤΑ

4.1 Προετοιμασία δεδομένων

Τέσσερα έδρανα είναι τοποθετημένα κατά μήκος του άξονα. Η ταχύτητα περιστροφής παραμένει σταθερή στις 2000RPM. Ακτινικό φορτίο 6000 lb ασκείται στον άξονα και τα έδρανα μέσω ενός μηχανισμού ελατηρίου. Σε κάθε έδρανο εγκαθίστανται δύο επιταχυνσιόμετρα υψηλής ευαισθησίας τύπου PCB 353B33 (ένα οριζόντιο κατά τον άξονα X και ένα κατακόρυφο κατά τον άξονα Y, καταλήγοντας σε οκτώ επιταχυνσιόμετρα συνολικά). Το σχήμα 4-1 παρουσιάζει την πειραματική διάταξη και τα σημεία τοποθέτησης των αισθητήρων. Όλες οι βλάβες εμφανίστηκαν μετά τον προβλεπόμενο κύκλο ζωής των εδράνων, ο οποίος ήταν ένα εκατομμύριο περιστροφές.



Σχήμα 4-1 : Πειραματική διάταξη

Τα δεδομένα είναι οργανωμένα ως εξής:

Κάθε αρχείο δεδομένων αποτελείται από 8 στήλες. Κάθε στήλη αντιστοιχεί στο σήμα ενός επιταχυνσιόμετρου όπως φαίνεται παρακάτω.

Στήλη	1	2	3	4	5	6	7	8
Έδρανο	1	1	2	2	3	3	4	4
Επιταχυνσιόμετρο	X	Y	X	Y	X	Y	X	Y

Το όνομα του κάθε αρχείου καθορίζει πότε συλλέχθηκαν τα δεδομένα. Σήματα διάρκειας 1 δευτερολέπτου συλλέχθηκαν κάθε 20 λεπτά για το πρώτο σετ δεδομένων και κάθε 10 λεπτά για το δεύτερο σετ. Κάθε γραμμή στα αρχεία δεδομένων αποτελεί ένα χρονικό σημείο της μέτρησης. Ο ρυθμός δειγματοληψίας ήταν 20 kHz.

Σετ 1:

Διάρκεια μετρήσεων: 19/12/2003 11:16:18 – 30/12/2003 23:55:42

Δειγματοληψία: 20 kHz

Περιγραφή: Δεδομένα κραδασμών συλλέχθηκαν κάθε 20 λεπτά από κάρτα NI DAQ 6062E. Στο τέλος της δοκιμής παρατηρήθηκε μία ρωγμή κοντά στο έδρανο 3. Τα τελευταία αρχεία εμφανίζουν αισθητή αλλαγή στο σήμα του εδράνου 3.

Σετ 2:

Διάρκεια μετρήσεων: 12/02/2004 10:32:39 – 19/02/2004 06:22:39

Δειγματοληψία: 20 kHz

Περιγραφή: Δεδομένα κραδασμών συλλέχθηκαν κάθε 20 λεπτά από κάρτα NI DAQ 6062E. Στο τέλος της δοκιμής παρατηρήθηκε σφάλμα στο έδρανο 1.

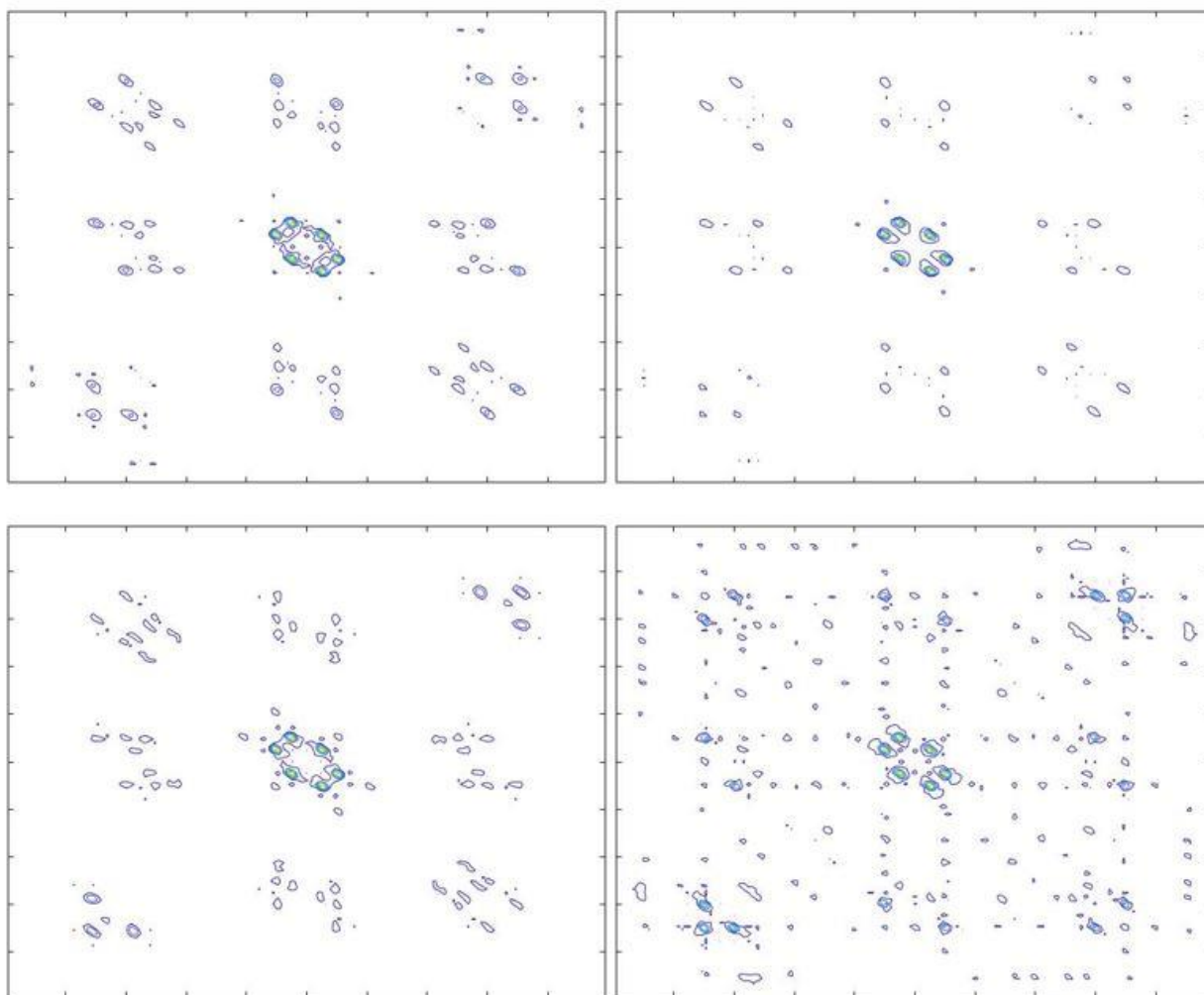
4.2 Σετ 1 – Φθορά εσωτερικού δακτυλίου ρουλεμάν

Από το 1^ο σετ χρησιμοποιήθηκαν 62 αρχεία μετρήσεων εκ των οποίων τα 40 αποτελούν δεδομένα κανονικής λειτουργίας και τα 22 δεδομένα λειτουργίας με φθαρμένο τον εσωτερικό δακτύλιο του 3^{ου} ρουλεμάν (στήλη 5)

Βήμα 1- Εξαγωγή γάρτη bispectrum από τα αρχεία δεδομένων.

Οι εικόνες που αποθηκεύονται στους φακέλους δεν περιλαμβάνουν τους άξονες, καθώς θα επηρέαζαν την εξαγωγή χαρακτηριστικών SURF.

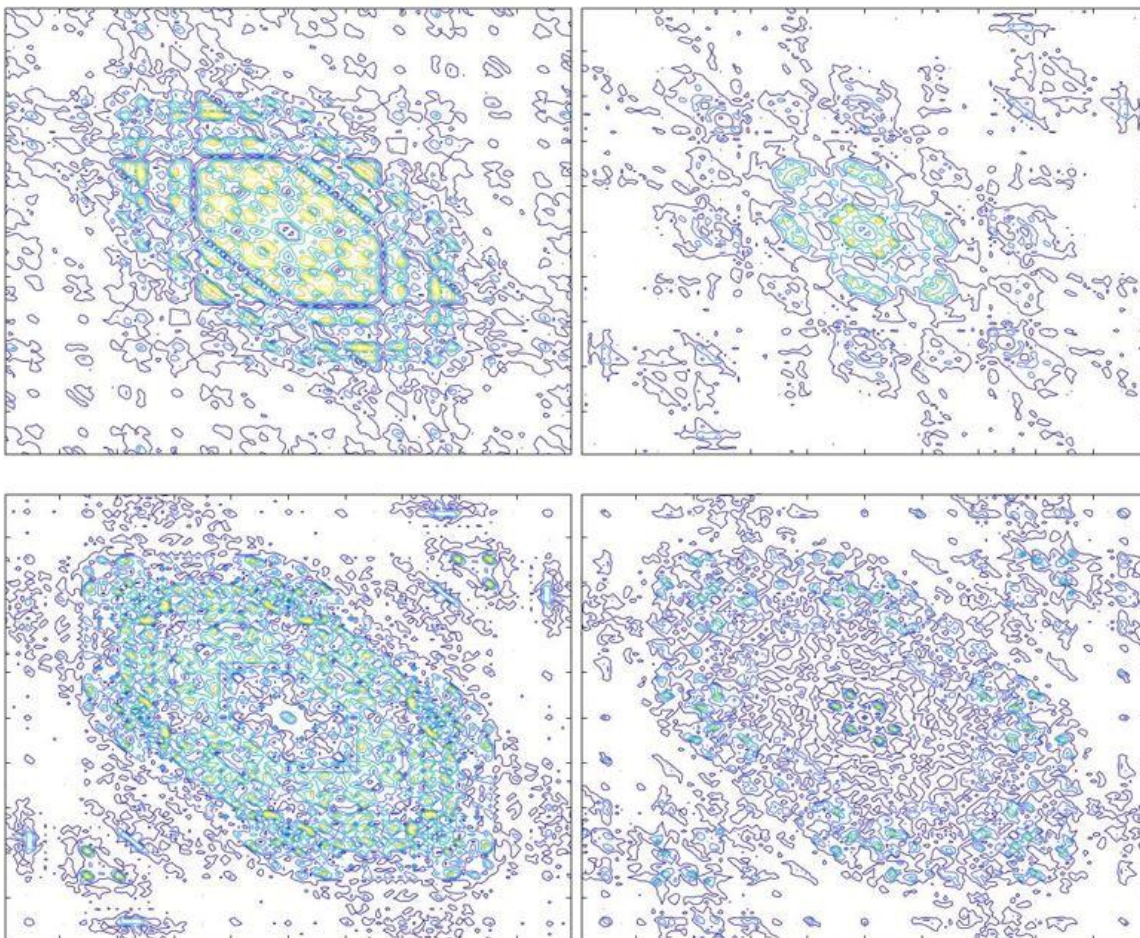
Στο σχήμα 4-2 φαίνονται παραδείγματα από τις εικόνες bispectrum σημάτων υπό κανονική λειτουργία όπως έχουν αποθηκευτεί στους φακέλους:



Σχήμα 4-2 : Εικόνες bispectrum σε κανονική λειτουργία (Σετ 1)

Παρατηρούμε πως η κύρια δραστηριότητα των σημάτων σε κανονική λειτουργία εμφανίζεται κυρίως στο κέντρο της εικόνας, καθώς και στα άκρα του εξαγώνου που σχηματίζουν οι συχνότητες. Λόγω εξωγενών παραγόντων όπως θόρυβος ή εξωτερικών κραδασμών, οι εικόνες μπορεί να διαφέρουν μεταξύ τους.

Το σχήμα 4-3 απεικονίζει παραδείγματα από τις εικόνες bispectrum σημάτων υπό λειτουργία με σφάλμα εσωτερικού δακτυλίου όπως έχουν αποθηκευτεί στους φακέλους:



Σχήμα 4-3 : Εικόνες bispectrum σε λειτουργία με σφάλμα εσωτερικού δακτυλίου (Σετ 1)

Οι εικόνες των σημάτων με τη φθορά του εσωτερικού δακτυλίου φαίνεται να εμφανίζουν έντονη δραστηριότητα σε εξαγωνική μορφή που καλύπτει μεγάλο μέρος του παραθύρου παρατήρησης.

Σε σύγκριση με τις εικόνες των κανονικών σημάτων είναι εμφανής η διαφορά μεταξύ των δύο καταστάσεων. Αυτά τα κοινά χαρακτηριστικά που εμφανίζουν οι δύο ομάδες εικόνων θα πρέπει να εξαχθούν και στη συνέχεια να μεταφραστούν σε μορφή που θα μπορεί να αξιοποιήσει ο αλγόριθμος ταξινόμησης.

Βήμα 2- Διαχωρισμός εικόνων σε ομάδες εκπαίδευσης και δοκιμών

Μετά την εξαγωγή των απεικονίσεων bispectrum έχουμε 40 εικόνες που αντιστοιχούν σε κανονική λειτουργία και 22 σήματα που αντιστοιχούν σε λειτουργία με σφάλμα εσωτερικού δακτυλίου.

Το σετ «εκπαίδευσης» αποτελείται από 24 σήματα κανονικής λειτουργίας και 13 σήματα εσφαλμένης λειτουργίας.

Το σετ «δοκιμών» αποτελείται από 16 σήματα κανονικής λειτουργίας και 9 σήματα εσφαλμένης λειτουργίας.

Βήμα 3 - Εξαγωγή χαρακτηριστικών SURF από το σετ εκπαίδευσης

Στο τέλος της διαδικασίας εξαγωγής έχουν εξαχθεί συνολικά 74641 χαρακτηριστικά από 37 εικόνες. Το κάθε χαρακτηριστικό περιγράφεται από ένα διάνυσμα 64 διαστάσεων.

Βήμα 4 - Επιλογή ισχυρότερων χαρακτηριστικών και εξισορρόπηση μεταξύ των κατηγοριών

Μεταξύ των δύο κατηγοριών, τα εσφαλμένα σήματα έχουν τον μικρότερο αριθμό χαρακτηριστικών (25747). Για να επιτευχθεί πιο εύρωστη ομαδοποίηση στο επόμενο στάδιο, επιλέγονται τα 25747 ισχυρότερα χαρακτηριστικά και από την κατηγορία των κανονικών σημάτων ώστε ο αριθμός των χαρακτηριστικών να είναι ίδιος και για τις δύο κατηγορίες.

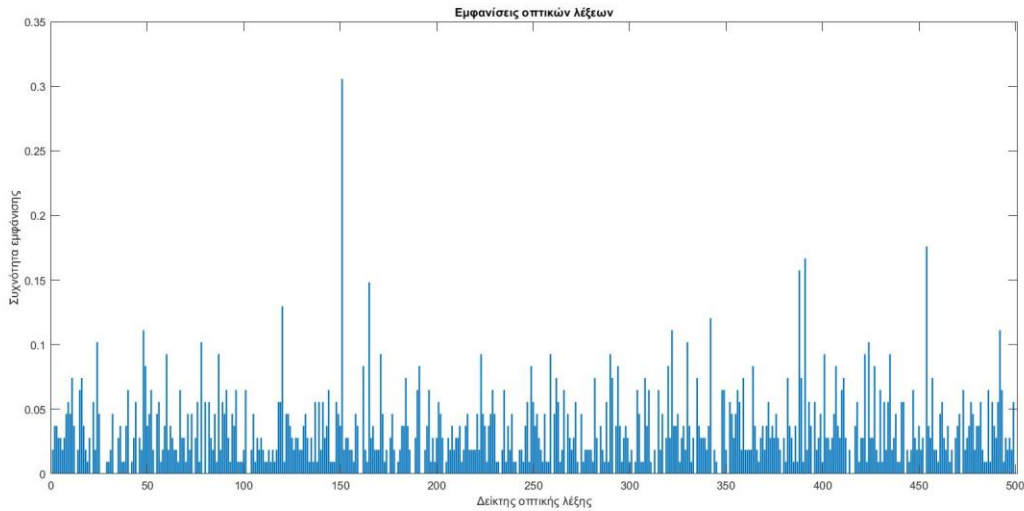
Ο τελικός αριθμός χαρακτηριστικών που θα χρησιμοποιηθούν είναι 51494 (25747 για κάθε κατηγορία).

Βήμα 5 - Ομαδοποίηση χαρακτηριστικών και καθορισμός του «οπτικού λεξιλογίου»

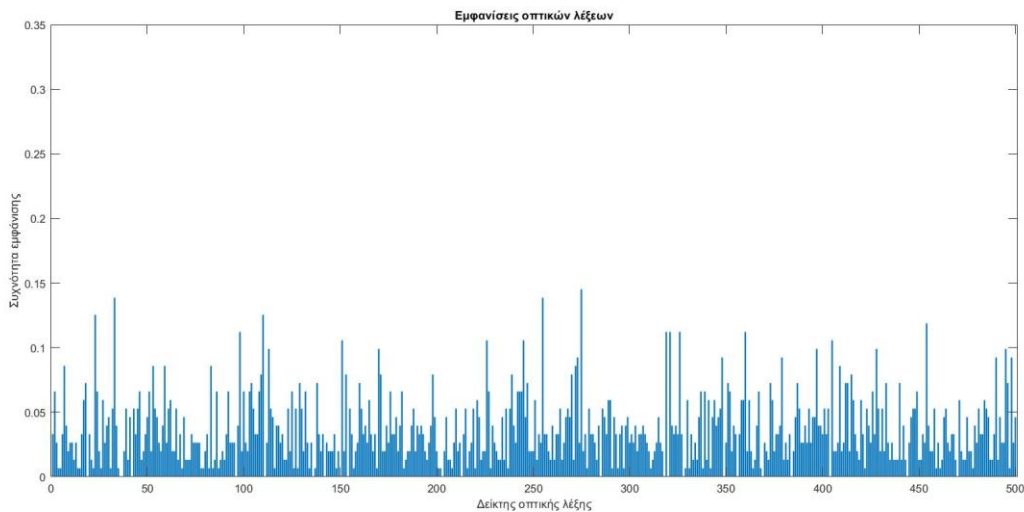
Στο συγκεκριμένο βήμα θα γίνει ομαδοποίηση των χαρακτηριστικών σε ομάδες των 10 και 500 λέξεων προκειμένου να ελεγχθεί η επίδραση του αριθμού των ομάδων στην ακρίβεια του ταξινομητή.

Βήμα 6 – Κωδικοποίηση των εικόνων με χρήση του «οπτικού λεξιλογίου» και εκπαίδευση αλγορίθμου ταξινόμησης

Στο σχήμα 4-4 απεικονίζεται το ιστόγραμμα κατανομής της κάθε οπτικής λέξης για ένα τυχαίο σήμα σε κανονική λειτουργία και στο σχήμα 4-5 το αντίστοιχο ιστόγραμμα για τυχαίο σήμα με σφάλμα εσωτερικού δακτυλίου:



Σχήμα 4-4 : Παράδειγμα ιστογράμματος σήματος σε κανονική λειτουργία



Σχήμα 4-5 : Παράδειγμα ιστογράμματος σήματος με σφάλμα στον εσωτερικό δακτύλιο

Όπως φαίνεται και από τις εικόνες bispectrum των σημάτων, σε κανονική λειτουργία, κάποια χαρακτηριστικά (κυρίως στο κέντρο της εικόνων) εμφανίζονται εντονότερα από άλλα και για αυτό το λόγο παρατηρούνται εντονότερες ακμές στο ιστόγραμμα περιγραφής. Το ιστόγραμμα του εσφαλμένου σήματος παρουσιάζει πιο ομοιόμορφη κατανομή, καθώς οι συχνότητες είναι διασκορπισμένες σχεδόν σε ολόκληρη την εικόνα.

Βήμα 7 – Αξιολόγηση αλγορίθμου ταξινόμησης

Για τον έλεγχο της ακρίβειας του αλγορίθμου θα χρησιμοποιηθούν οι εικόνες του σετ δοκιμών

Πραγματική κατηγορία σήματος	Πρόβλεψη με αριθμό ομάδων K-means = 500	Πρόβλεψη με αριθμό ομάδων K-means = 10
Εσφαλμένο	Εσφαλμένο	Εσφαλμένο
Εσφαλμένο	Εσφαλμένο	Εσφαλμένο
Εσφαλμένο	Εσφαλμένο	Εσφαλμένο
Εσφαλμένο	Εσφαλμένο	Εσφαλμένο
Εσφαλμένο	Εσφαλμένο	Εσφαλμένο
Εσφαλμένο	Εσφαλμένο	Εσφαλμένο
Εσφαλμένο	Εσφαλμένο	Εσφαλμένο
Εσφαλμένο	Εσφαλμένο	Εσφαλμένο
Εσφαλμένο	Κανονικό	Κανονικό
Κανονικό	Κανονικό	Κανονικό
Κανονικό	Κανονικό	Κανονικό
Κανονικό	Κανονικό	Κανονικό
Κανονικό	Κανονικό	Κανονικό
Κανονικό	Κανονικό	Κανονικό
Κανονικό	Εσφαλμένο	Εσφαλμένο
Κανονικό	Κανονικό	Κανονικό
Κανονικό	Κανονικό	Εσφαλμένο
Κανονικό	Κανονικό	Κανονικό
Κανονικό	Κανονικό	Κανονικό
Κανονικό	Κανονικό	Κανονικό
Κανονικό	Κανονικό	Κανονικό
Κανονικό	Κανονικό	Κανονικό
Κανονικό	Κανονικό	Κανονικό
Κανονικό	Εσφαλμένο	Εσφαλμένο
Κανονικό	Κανονικό	Κανονικό

Στην περίπτωση του 1^{ου} σετ, βλέπουμε πως ο αριθμός των ομάδων K-means δεν επηρεάζει ιδιαίτερα την απόδοση του μοντέλου. Παρ' όλα αυτά, αν ο αριθμός των ομάδων μειωθεί πολύ, παρατηρείται μείωση της ακρίβειας πρόβλεψης

Για τα εσφαλμένα σήματα ο αλγόριθμος είχε ακρίβεια $8/9 = 88.88\%$ ανεξαρτήτως αριθμού ομάδων K-means

Για τα φυσιολογικά σήματα ο αλγόριθμος είχε ακρίβεια $14/16=87.5\%$ για υψηλούς αριθμούς ομάδων K-means και $13/16 = 81\%$ για μικρό αριθμό ομάδων

Συνολική ακρίβεια $22/25 = 88\%$ και 84% αντίστοιχα.

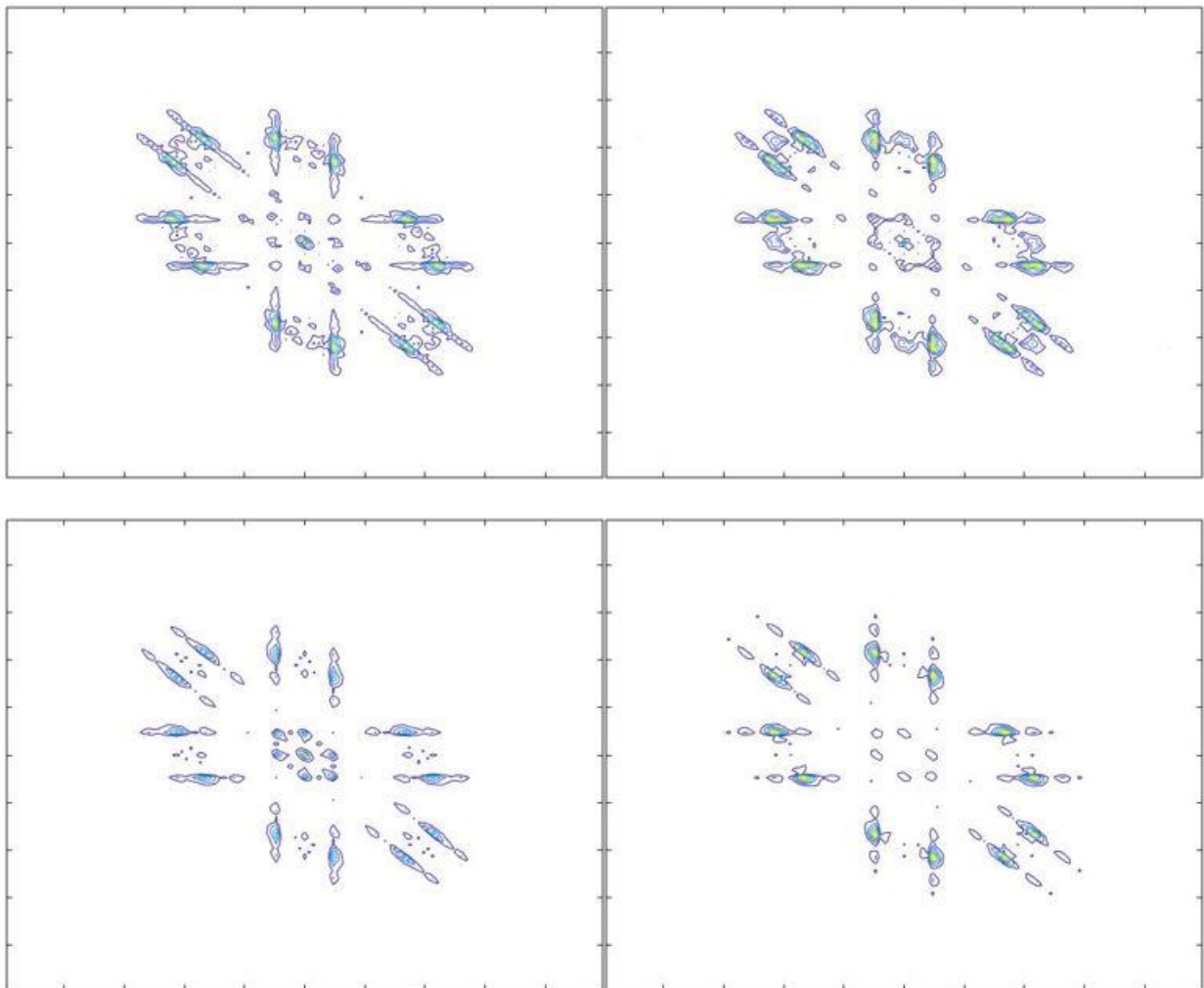
4.3 Σετ 2 – Φθορά εξωτερικού δακτυλίου ρουλεμάν

Από το 2^ο σετ χρησιμοποιήθηκαν 61 αρχεία μετρήσεων εκ των οποίων τα 45 αποτελούν δεδομένα κανονικής λειτουργίας και τα 16 δεδομένα λειτουργίας με φθαρμένο τον εξωτερικό δακτύλιο του 1^{ου} ρουλεμάν (στήλη 1)

Βήμα 1- Εξαγωγή γάρτη bispectrum από τα αρχεία δεδομένων.

Οι εικόνες που αποθηκεύονται στους φακέλους δεν περιλαμβάνουν τους άξονες, καθώς θα επηρέαζαν την εξαγωγή χαρακτηριστικών SURF.

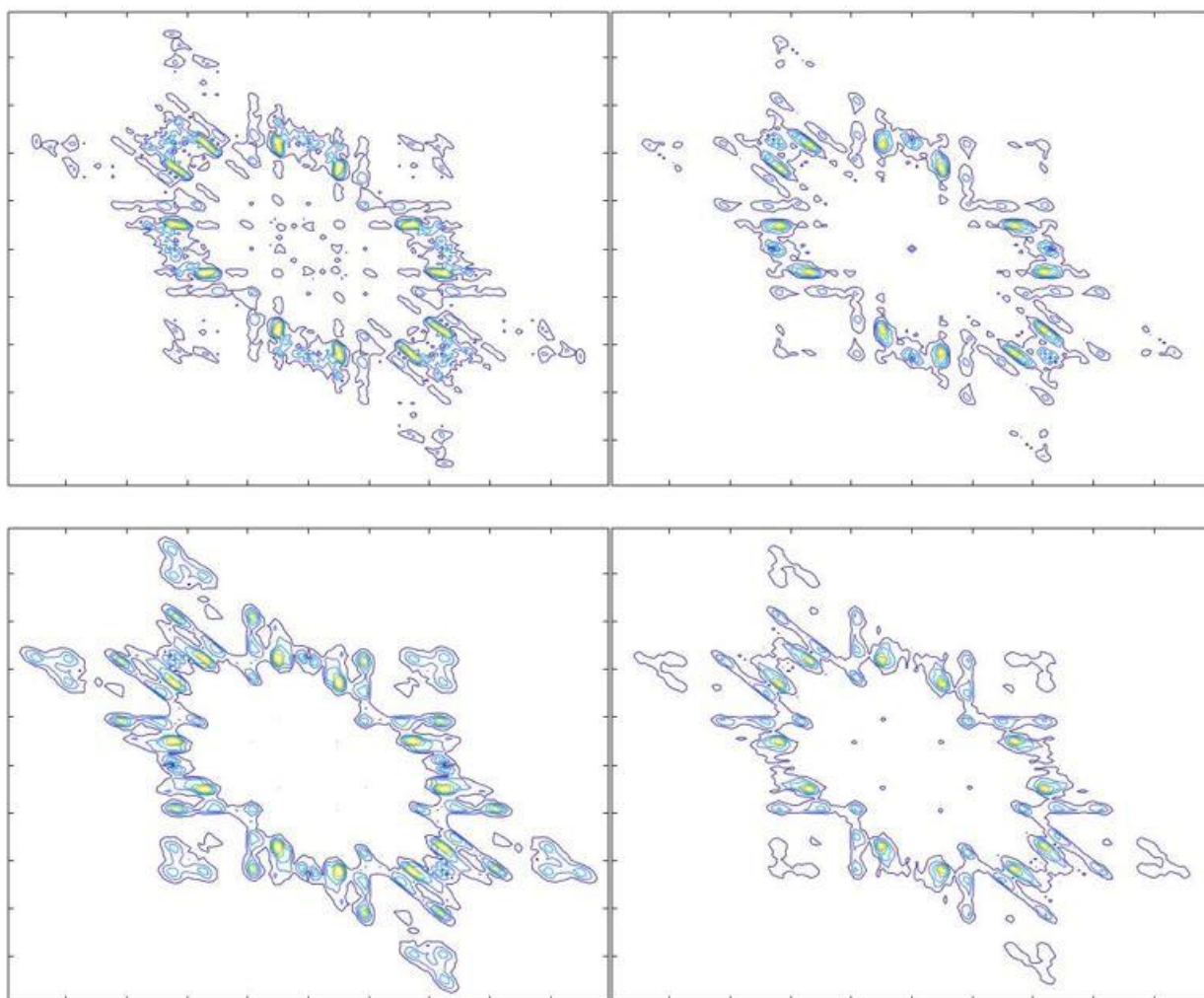
Στο σχήμα 4-6 φαίνονται παραδείγματα από τις εικόνες bispectrum σημάτων υπό κανονική λειτουργία όπως έχουν αποθηκευτεί στους φακέλους :



Σχήμα 4-6 : Εικόνες bispectrum σε κανονική λειτουργία (Σετ 2)

Παρατηρούμε πως η κύρια δραστηριότητα των σημάτων σε κανονική λειτουργία εμφανίζεται κυρίως στο κέντρο της εικόνας, καθώς ακτινικά στα άκρα του εξάγωνου που σχηματίζουν οι συχνότητες. Λόγω εξωγενών παραγόντων όπως θόρυβος ή εξωτερικών κραδασμών, οι εικόνες μπορεί να διαφέρουν μεταξύ τους.

Οι παρακάτω εικόνες αποτελούν παραδείγματα από τους χάρτες bispectrum σημάτων υπό λειτουργία με σφάλμα εξωτερικού δακτυλίου:



Σχήμα 4-7 : Εικόνες bispectrum σε λειτουργία με σφάλμα εξωτερικού δακτυλίου (Σετ 2)

Οι εικόνες των σημάτων με τη φθορά του εσωτερικού δακτυλίου φαίνεται να εμφανίζουν εντονότερη δραστηριότητα περιμετρικά του εξάγωνου και λιγότερη δραστηριότητα στο κέντρο της εικόνας.

Σε σύγκριση με τις εικόνες των κανονικών σημάτων είναι εμφανής η διαφορά μεταξύ των δύο καταστάσεων. Αυτά τα κοινά χαρακτηριστικά που εμφανίζουν οι δύο ομάδες εικόνων θα πρέπει να εξαχθούν και στη συνέχεια να μεταφραστούν σε μορφή που θα μπορεί να αξιοποιήσει ο αλγόριθμος ταξινόμησης.

Βήμα 2- Διαχωρισμός εικόνων σε ομάδες εκπαίδευσης και δοκιμών

Μετά την εξαγωγή των απεικονίσεων bispectrum έχουμε 45 εικόνες που αντιστοιχούν σε κανονική λειτουργία και 16 σήματα που αντιστοιχούν σε λειτουργία με σφάλμα εξωτερικού δακτυλίου.

Το σετ «εκπαίδευσης» αποτελείται από 27 σήματα κανονικής λειτουργίας και 10 σήματα εσφαλμένης λειτουργίας.

Το σετ «δοκιμών» αποτελείται από 18 σήματα κανονικής λειτουργίας και 6 σήματα εσφαλμένης λειτουργίας.

Βήμα 3 - Εξαγωγή χαρακτηριστικών SURF από το σετ εκπαίδευσης

Στο τέλος της διαδικασίας έχουν εξαχθεί συνολικά 22192 χαρακτηριστικά από 37 εικόνες.

Το κάθε χαρακτηριστικό περιγράφεται από ένα διάνυσμα 64 διαστάσεων.

Βήμα 4 - Επιλογή ισχυρότερων χαρακτηριστικών και εξισορρόπηση μεταξύ των κατηγοριών

Μεταξύ των δύο κατηγοριών, τα εσφαλμένα σήματα έχουν τον μικρότερο αριθμό χαρακτηριστικών (7206). Για να επιτευχθεί πιο εύρωστη ομαδοποίηση στο επόμενο στάδιο, επιλέγονται τα 7206 ισχυρότερα χαρακτηριστικά και από την κατηγορία των κανονικών σημάτων ώστε ο αριθμός των χαρακτηριστικών να είναι ίδιος και για τις δύο κατηγορίες.

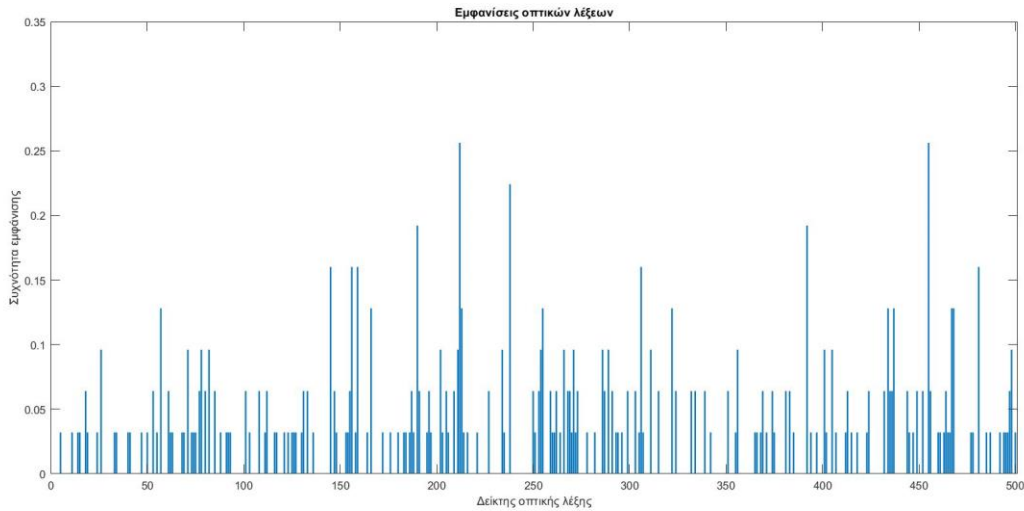
Ο τελικός αριθμός χαρακτηριστικών που θα χρησιμοποιηθούν είναι 14412 (7206 για κάθε κατηγορία).

Βήμα 5 - Ομαδοποίηση χαρακτηριστικών και καθορισμός του «οπτικού λεξιλογίου»

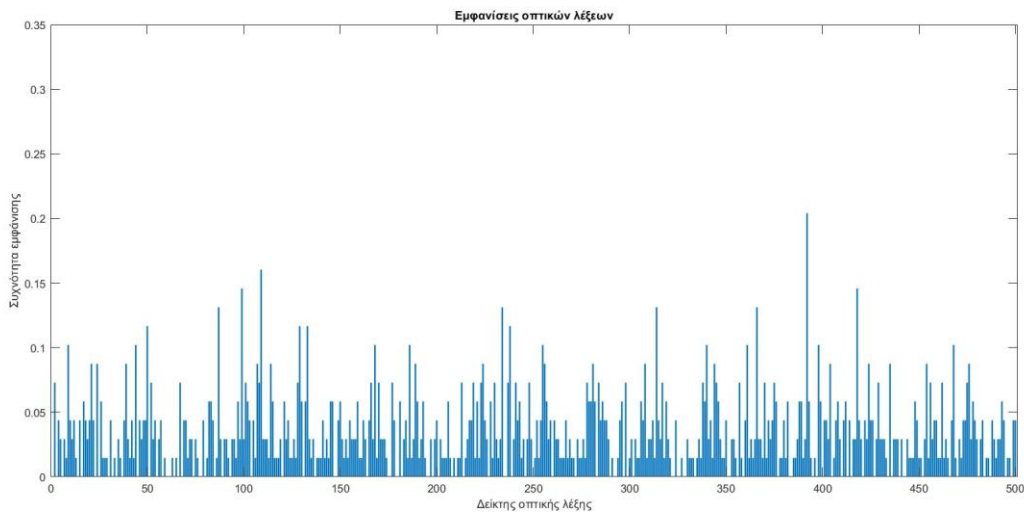
Στο συγκεκριμένο βήμα θα γίνει ομαδοποίηση των χαρακτηριστικών σε ομάδες των 50, 500 και 5000 λέξεων προκειμένου να ελεγχθεί η επίδραση του αριθμού των ομάδων στην ακρίβεια του ταξινομητή.

Βήμα 6 – Κωδικοποίηση των εικόνων με χρήση του «οπτικού λεξιλογίου» και εκπαίδευση αλγορίθμου ταξινόμησης

Στο σχήμα 4-8 απεικονίζεται το ιστόγραμμα κατανομής της κάθε οπτικής λέξης για ένα τυχαίο σήμα σε κανονική λειτουργία και στο σχήμα 4-9 το αντίστοιχο ιστόγραμμα για τυχαίο σήμα με σφάλμα εξωτερικού δακτυλίου:



Σχήμα 4-8 : Παράδειγμα ιστογράμματος σήματος σε κανονική λειτουργία



Σχήμα 4-9 : Παράδειγμα ιστογράμματος σήματος με σφάλμα στον εσωτερικό δακτύλιο

Συγκρίνοντας τα ιστογράμματα των σημάτων παρατηρούμε πως το ιστόγραμμα στην κανονική λειτουργία παρουσιάζει υψηλότερες τιμές σε ορισμένα χαρακτηριστικά καθώς και κάποια χαρακτηριστικά τα οποία δεν εμφανίζονται καθόλου. Συγκριτικά, το ιστόγραμμα της εσφαλμένης λειτουργίας εμφανίζει πιο ομοιόμορφη κατανομή χαρακτηριστικών. Αυτή η παρατήρηση συμφωνεί με τη μορφή των εικόνων bispectrum, καθώς οι συχνότητες στην εσφαλμένη λειτουργία εμφανίζονται σε ένα συνεχές εξαγωνικό σχήμα, ενώ στην κανονική λειτουργία παρουσιάζονται κενά σε διάφορα σημεία της εικόνας.

Βήμα 7 – Αξιολόγηση αλγορίθμου ταξινόμησης

Για τον έλεγχο της ακρίβειας του αλγορίθμου θα χρησιμοποιηθούν οι εικόνες του σετ δοκιμών

Πραγματική κατηγορία σήματος	Πρόβλεψη με αριθμό ομάδων K-means = 500	Πρόβλεψη με αριθμό ομάδων K-means = 50	Πρόβλεψη με αριθμό ομάδων K-means = 5000
Εσφαλμένο	Εσφαλμένο	Εσφαλμένο	Εσφαλμένο
Εσφαλμένο	Εσφαλμένο	Κανονικό	Κανονικό
Εσφαλμένο	Εσφαλμένο	Εσφαλμένο	Εσφαλμένο
Εσφαλμένο	Εσφαλμένο	Εσφαλμένο	Εσφαλμένο
Εσφαλμένο	Εσφαλμένο	Εσφαλμένο	Εσφαλμένο
Εσφαλμένο	Κανονικό	Κανονικό	Κανονικό
Κανονικό	Κανονικό	Εσφαλμένο	Κανονικό
Κανονικό	Κανονικό	Κανονικό	Κανονικό
Κανονικό	Κανονικό	Κανονικό	Κανονικό
Κανονικό	Κανονικό	Κανονικό	Κανονικό
Κανονικό	Κανονικό	Κανονικό	Κανονικό
Κανονικό	Κανονικό	Κανονικό	Κανονικό
Κανονικό	Κανονικό	Κανονικό	Κανονικό
Κανονικό	Κανονικό	Κανονικό	Κανονικό
Κανονικό	Κανονικό	Κανονικό	Κανονικό
Κανονικό	Κανονικό	Κανονικό	Κανονικό
Κανονικό	Κανονικό	Κανονικό	Κανονικό
Κανονικό	Κανονικό	Κανονικό	Κανονικό
Κανονικό	Κανονικό	Εσφαλμένο	Κανονικό
Κανονικό	Κανονικό	Εσφαλμένο	Κανονικό
Κανονικό	Κανονικό	Εσφαλμένο	Κανονικό
Κανονικό	Κανονικό	Κανονικό	Κανονικό
Κανονικό	Κανονικό	Κανονικό	Κανονικό
Κανονικό	Κανονικό	Κανονικό	Κανονικό
Κανονικό	Κανονικό	Κανονικό	Κανονικό

Για το 2^ο σετ, το εύρος του αριθμού ομάδων K-means στα οποία παρουσιάζεται μέγιστη ακρίβεια είναι μικρότερο από ότι στο 1^ο σετ.

Για τα εσφαλμένα σήματα ο αλγόριθμος είχε ακρίβεια:

- Για ομάδες K-means = 500 : $5/6 = 83.33\%$
- Για ομάδες K-means = 50 : $4/6 = 66.66\%$
- Για ομάδες K-means = 5000 : $4/6 = 66.66\%$

Για τα φυσιολογικά σήματα ο αλγόριθμος είχε ακρίβεια $18/18=100\%$ για αριθμούς ομάδων = 500 και = 5000 και ακρίβεια $14/18 = 77.77\%$ για αριθμό ομάδων = 50.

Συνολική ακρίβεια :

- K-means 500 : $23/24 = 95.83\%$
- K-means 50 : $19/24 = 79.17\%$
- K-means 5000 : $22/24 = 91.67\%$

ΚΕΦΑΛΑΙΟ 5 – ΣΥΜΠΕΡΑΣΜΑΤΑ ΚΑΙ ΜΕΛΛΟΝΤΙΚΗ ΕΡΓΑΣΙΑ

5.1 Συμπεράσματα

Μετά την αξιολόγηση των μοντέλων ταξινόμησης με τα σετ δοκιμών, είναι εμφανές πως η συγκεκριμένη μεθοδολογία έχει προοπτικές για εφαρμογή στο πεδίο της διάγνωσης βλαβών. Τον σημαντικότερο, ίσως, ρόλο στην όλη διαδικασία κατέχει η επιλογή της καλύτερης μεθόδου απεικόνισης των σημάτων, καθώς οι εικόνες αυτές είναι που θα χρησιμοποιηθούν για την εξαγωγή των χαρακτηριστικών. Ο κύριος στόχος στο συγκεκριμένο βήμα είναι τα σήματα που ανήκουν στην ίδια κατάσταση λειτουργίας να είναι όσο το δυνατόν πιο όμοια μεταξύ τους και παράλληλα να διαφέρουν από τα σήματα άλλων καταστάσεων. Αυτός ο διαχωρισμός μπορεί να εντοπιστεί και οπτικά ως ένα βαθμό, αλλά η εισαγωγή των εικόνων στο μοντέλο ταξινόμησης είναι ο μόνος σίγουρος τρόπος για να αξιολογηθεί η επίδραση των εικόνων στο τελικό αποτέλεσμα.

Κατά την εξαγωγή των χαρακτηριστικών SURF επιλέχθηκαν τυπικές παράμετροι του αλγόριθμου, ώστε να υπάρχει επαρκής αριθμός διανυσμάτων περιγραφής. Κατά την ομαδοποίηση των χαρακτηριστικών με K-means, επιλέχθηκαν διάφοροι αριθμοί ομάδων με το 500 να δίνει τα πιο ικανοποιητικά αποτελέσματα και στις δύο κατηγορίες σφαλμάτων. Καθώς το πλήθος των σημάτων τόσο του σετ εκπαίδευσης όσο και του σετ δοκιμών ήταν σχετικά μικρό, οι αλλαγές στις παραμέτρους έδιναν σχετικά μικρές διακυμάνσεις στα αποτελέσματα με μόνο κάποιες ακραίες τιμές να έχουν σημαντική επίδραση.

5.2 Μελλοντική εργασία

Μελλοντικές βελτιώσεις και τροποποιήσεις των επιμέρους βημάτων της μεθοδολογίας μπορούν να αυξήσουν ακόμα περισσότερο την ακρίβεια του μοντέλου.

Όσον αφορά την μέθοδο απεικόνισης των σημάτων, υπάρχουν αρκετές εναλλακτικές όπως, μέθοδος bicoherence και κυκλοστάσιμη ανάλυση (cyclostationary analysis) και κυρτογράμματα (kurtogram). Με εφαρμογή της κατάλληλης μεθόδου είναι δυνατόν να υπάρχει καλύτερος διαχωρισμός των σημάτων και κατά συνέπεια, υψηλότερη ακρίβεια ταξινόμησης.

Στο βήμα της επεξεργασίας των εξαγόμενων χαρακτηριστικών SURF, κύριος στόχος είναι η μείωση του μεγάλου όγκου των δεδομένων χωρίς να χαθεί η πληροφορία που περιέχουν. Για το σκοπό αυτό μπορεί εναλλακτικά να χρησιμοποιηθεί principal component analysis (PCA) ή μέθοδος t-SNE για μείωση των διαστάσεων των διανυσμάτων περιγραφής.

Για την εκπαίδευση του ταξινομήτη υπάρχουν επίσης πολλές διαφορετικές μέθοδοι που μπορούν να εφαρμοστούν όπως KNN ή νευρωνικά δίκτυα.

Προκειμένου να ελεγχθεί σε ακόμα μεγαλύτερο βαθμό η αξιοπιστία της μεθόδου θα χρειαστεί να γίνουν δοκιμές με μεγαλύτερες ομάδες δεδομένων καθώς και δοκιμές σε σήματα από διαφορετικές μηχανές και διαφορετικά ήδη σφαλμάτων / καταστάσεις λειτουργίας.

ΠΑΡΑΡΤΗΜΑ Α – ΚΩΔΙΚΕΣ MATLAB

Κώδικας μετατροπής δεδομένων σημάτων σε εικόνες bispectrum και τοποθέτηση σε αντίστοιχους φακέλους

```
% This code reads the signals and creates bispectrum map images
to appropriate files
% Determine appropriate data file location.
myFolder = 'Set01'; %Or 'Set02'
% Check if file exists
if ~isdir(myFolder)
    errorMessage = sprintf('Error: The following folder does not
exist:\n%s', myFolder);
    uiwait(warndlg(errorMessage));
    return;
end
% Create a list of all files in the folder
filePattern = fullfile(myFolder, '*.txt');
theFiles = dir(filePattern);
for k = 1: length(theFiles) % Read signals
    baseFileName = theFiles(k).name;
    fullFileName = fullfile(myFolder, baseFileName);
    fprintf(1, 'Now reading %s\n', fullFileName);
    signal = load(fullFileName);'
    signal=signal(:,1);%<<<<<<<CHOOSE COLUMN or (:,5)
    %Creation of bispectrum map
    figure (1)
    [Bspec,waxis] = bispeci
(signal,50,length(signal),50,'biased',127,1);
    I =getframe; %Extract image
    %Files are saved to separate files depending on the type of
    %signal(faulty/normal).Empty files must already exist.
    if contains(fullFileName,"N.txt")%normal signal namings
        %Determine the file where the normal images will be saved
        ImageFolder = 'Images_Set01\Normal'; %Or 'Images_Set02\Normal'
        imagename=sprintf('Normal%d.jpg',k);
        ImageFullpath=fullfile(ImageFolder,imagename);
        imwrite(I.cdata,ImageFullpath);
    elseif contains(fullFileName,"F.txt")%faulty signal namings
        %Determine the file where the faulty images will be saved
        ImageFolder = 'Images_Set01\Faulty'; %Or 'Images_Set02\Faulty'
        imagename=sprintf('Fault%d.jpg',k);
        ImageFullpath=fullfile(ImageFolder,imagename);
        imwrite(I.cdata,ImageFullpath);
    end
end
end
```

%At the end of the code the images will be saved to the appropriate files

Κώδικας εξαγωγής χάρτη bispectrum με την έμμεση μέθοδο.

```
function [Bspec,waxis] = bispeci (y,nlag,nsamp, overlap,flag,
nfft, wind)
%BISPECI Bispectrum estimation using the indirect method.
% [Bspec,waxis] = bispeci (y,nlag,segsamp,overlap,flag,nfft,
wind)
% y - data vector or time-series
% nlag - number of lags to compute [must be specified]
% segsamp - samples per segment [default: row dimension of
y]
% overlap - percentage overlap [default = 0]
% flag - 'biased' or 'unbiased' [default is 'unbiased']
% nfft - FFT length to use [default = 128]
% wind - window function to apply:
% if wind=0, the Parzen window is applied (default);
% otherwise the hexagonal window with unity values is
applied.
% Bspec - estimated bispectrum it is an nfft x nfft array
% with origin at the center, and axes pointing down and to
the right
% waxis - frequency-domain axis associated with the
bispectrum.
% - the i-th row (or column) of Bspec corresponds to
f1 (or f2)
% value of waxis(i).

% Copyright (c) 1991-2001 by United Signals & Systems, Inc.
% $Revision: 1.7 $
% A. Swami January 20, 1993.

% RESTRICTED RIGHTS LEGEND
% Use, duplication, or disclosure by the Government is subject
to
% restrictions as set forth in subparagraph (c) (1) (ii) of the
% Rights in Technical Data and Computer Software clause of DFARS
% 252.227-7013.
% Manufacturer: United Signals & Systems, Inc., P.O. Box 2374,
% Culver City, California 90231.
%
% This material may be reproduced by or for the U.S. Government
pursuant
% to the copyright license under the clause at DFARS 252.227-
7013.
```

```

% ----- parameter checks -----

[ly, nrecs] = size (y);
if (ly == 1) y=y(:);   ly = nrecs; nrecs = 1;       end
if (exist('overlap') ~= 1)   overlap = 0;         end
overlap = min(99, max(overlap,0));
if (nrecs > 1)                overlap = 0;         end
if (exist('nsamp') ~= 1)      nsamp   = ly;        end
if (nsamp > ly | nsamp <= 0) nsamp   = ly;        end
if (exist('flag') ~= 1)      flag    = 'biased';   end
if (flag(1:1) ~= 'b')       flag    = 'unbiased'; end
if (exist('nfft') ~= 1)     nfft    = 128;        end
if (nfft <= 0)              nfft    = 128;        end
if (exist('wind') ~= 1)     wind    = 0;          end

nlag = min(nlag, nsamp-1);
if (nfft < 2*nlag+1)   nfft = 2^nextpow2(nsamp); end

% ----- create the lag window -----
Bspec = zeros(nfft,nfft) ;

if (wind == 0)
    indx = (1:nlag)';
    window = [1; sin(pi*indx/nlag) ./ (pi*indx/nlag)];
else
    window = ones(nlag+1,1);
end
window = [window; zeros(nlag,1)];

% ----- cumulants in non-redundant region -----
% -----
% define cum(i,j) = E conj(x(n)) x(n+i) x(n+j)
% for a complex process, we only have cum(i,j) = cum(j,i)
%
overlap = fix(nsamp * overlap / 100);
nadvance = nsamp - overlap;
nrecord = fix ( (ly*nrecs - overlap) / nadvance );

c3 = zeros(nlag+1,nlag+1);
ind = [1:nsamp]';
for k=1:nrecord,
    x = y(ind); x = x - mean(x);
    ind = ind + nadvance;
    for j=0:nlag
        z = x(1:nsamp-j) .* x(j+1:nsamp);
    end
end

```

```

        for i=j:nlag
            sum = z(1:nsamp-i)' * x(i+1:nsamp);
            if (flag(1:1) == 'b'), sum = sum/nsamp;
            else, sum = sum / (nsamp-i);
            end
            c3(i+1,j+1) = c3(i+1,j+1) + sum;
        end
    end
end
c3 = c3 / nrecord;

% cumulants elsewhere by symmetry -----
-----
c3 = c3 + tril(c3,-1)';           % complete I quadrant
c31 = c3(2:nlag+1,2:nlag+1);
c32 = zeros(nlag,nlag);  c33 = c32;  c34 = c32;
for i=1:nlag,
    x = c31(i:nlag,i);
    c32(nlag+1-i,1:nlag+1-i) = x';
    c34(1:nlag+1-i,nlag+1-i) = x;
    if (i < nlag)
        x = flipud(x(2:length(x)));
        c33 = c33 + diag(x,i) + diag(x,-i);
    end
end
c33 = c33 + diag(c3(1,nlag+1:-1:2));
cmat = [ [c33, c32, zeros(nlag,1)]; [ [c34; zeros(1,nlag)]
, c3 ] ];

% ----- apply lag-domain window -----
-----

wcmat = cmat;
if (wind ~= -1)
    indx = [-nlag:nlag]';
    for k=-nlag:nlag
        wcmat(:,k+nlag+1) = cmat(:,k+nlag+1) ...
            .* window(abs(indx-k)+1) .* window(abs(indx)+1) ...
            * window(abs(k)+1);
    end
end

% ----- compute 2d-fft, and shift and rotate for proper
orientation -----

Bspec = fft2(wcmat, nfft, nfft);

```



```

        Bspec = fftshift(Bspec);           % axes d and r; orig
at ctr

        if (rem(nfft,2) == 0)
            waxis = [-nfft/2:(nfft/2-1)]/nfft;
        else
            waxis = [-(nfft-1)/2:(nfft-1)/2]/nfft;
        end

%     hold off, clf

%     contour(abs(Bspec),4,waxis,waxis),
contour(waxis,waxis,abs(Bspec),6),
title('Bispectrum estimated via the indirect method')
xlabel('f1'), ylabel('f2')
set(gcf, 'Name', 'Hosa BISPECI')
return

```

Κώδικας εκπαίδευσης μοντέλου ταξινόμησης:

```

% This code reads the images, extracts SURF features, clusters
similar
% features together with k-means and then uses the cluster
centers to train
% a classifier
rng ('default');
%Create imageDatastore object that includes all images with
corresponding
%labels based on file names
imds
=imageDatastore('Images_Set01','IncludeSubfolders',true,'LabelSo
urce','foldernames');

%Separate images into training and testing sets with 60-40 ratio
[trainingSet, validationSet] = splitEachLabel(imds,
0.6,'randomize');

%Declaration of custom function that will be used for SURF point
extraction
extractorFcn = @SURFBOW;

%The next command uses the custom function to extract surfpoints
from all
%images in the training set. Next, it selects an equal number of
the
%strongest from each set to allow for better clustering.

```

```
%With the use of Kmeans, clusters are created from all chosen
features.The
%center of each cluster will later be used to train the
classifier
bag =
bagOfFeatures(trainingSet,'CustomExtractor',extractorFcn,'Vocabu
larySize',500);

%The next command trains an image classifier using the new
generalized
%features generated from the previous command.
%For each image in the training set, every SURFfeature vector is
compared
%with the kmeans centers and the closest center is determined.
%After doing this with all features, a histogram of each Kmeans
center
%occurrence is created.
%The histograms of all images in the data set will be used for
the
%classifier. The classifier can then classify new image
histograms by using
%support vector machine
categoryClassifier =
trainImageCategoryClassifier(trainingSet,bag);
%This final command evaluates the classifier by comparing the
predicted classification with the actual signal type for each
image in the testing set
confMatrix = evaluate(categoryClassifier, validationSet);
```

Συνάρτηση εξαγωγής χαρακτηριστικών SURF

```
%Function to detect and extract surfpoints
function [features,featureMetrics]=SURFBOW(I)
grayImage = rgb2gray(I);

%Surfpoint extraction
multiscaleSURFPoints =
detectSURFFeatures(grayImage,'MetricThreshold',100);
features = extractFeatures(grayImage,
multiscaleSURFPoints,'Upright',true);
[featureMetrics] = multiscaleSURFPoints.Metric;
```

ΒΙΒΛΙΟΓΡΑΦΙΑ

- [1] Chen Lu, Yang Wang, Minvydas Ragulskis and Yujie Cheng, “*Fault Diagnosis for Rotating Machinery: A Method based on Image Processing*” (2016)
- [2] Herbert Bay, Andreas Ess, Tinne Tuytelaars and Luc Van Gool, “*Speeded-Up Robust Features*” (2008)
- [3] Edouard Oyallon and Julien Rabin, “*An Analysis of the SURF Method*” (2015)
- [4] Alessandro Rivola, “*Applications of Higher Order Spectra to the Machine Condition Monitoring*” (2000)
- [5] Lingli Jiang, Yilun Liu, Xuejun Li and Siwen Tang, “*Using bispectral distribution as a feature for rotating machinery fault diagnosis*” (2011)
- [6] Miguel E. Iglesias-Martinez, Jose Alfonso Antonino-Daviu, Pedro Fernandez de Cordoba and J. Alberto Conejero, “*Rotor Fault Detection in Induction Motors Based on Time-Frequency Analysis Using the Bispectrum and the Autocovariance of Stray Flux Signals*” (2019)
- [7] Paul S. Bradley and Usama M. Fayaad, “*Refining Initial Points for K-Means Clustering*” (1998)
- [8] Aristidis Likas, Nikos Vlassis and Jakob J. Verbeek, “*The global K-Means clustering algorithm*” (2003)
- [9] Stephen O’ Hara and Bruce A. Draper, “*Introduction to the bag of Features Paradigm for Image Classification and Retrieval*” (2010)
- [10] Nello Cristianini and John Shawe-Taylor , “*An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*” (2000)