



**ΠΟΛΥΤΕΧΝΕΙΟ ΚΡΗΤΗΣ**  
**ΤΜΗΜΑ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ**  
**ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ**

Εργαστήριο Ψηφιακής Επεξεργασίας Σημάτων και  
Εικόνας

**ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ**

**Determination of Motion Models in Fish School**  
**Αναγνώριση Μοντέλων Κίνησης σε Σμήνος Ψαριών**

NTOKO ANEZA

**Επιβλέπων Καθηγητής:** Ζερβάκης Μιχάλης, Καθηγητής

**Εξεταστική επιτροπή:** Ζερβάκης Μιχάλης, Καθηγητής

Πετράκης Ευριπίδης, Καθηγητής

Παπανδρουλάκης Νίκος, Διευθυντής Ερευνών, Ι.Θ.Α.Β.Β.ΥΚ.

ΕΛ.ΚΕ.Θ.Ε

Χανιά, Δεκέμβριος 2018

## Abstract

The objective of the current thesis is the implementation, testing and validation of a video-surveillance system in fish farms that can be used in the future to analyze fish school behavior or their individual behavior. Such an application proves to be extremely useful in guiding the feeding procedure and inspect the quality of growth of fish populations in the expanding industry of fish farming. A crucial issue that we faced was the image enhancement of underwater images in videos, as they are characterized of intense noise, sudden light changes and blurring environment. Afterwards, the basic steps of our implementation involved the appropriate selection and combination of algorithms for detection, tracking and trajectory modeling of fish movements. The use of the Kalman Filter was important for the prediction and correction of fish motion. We further processed the trajectories after their extraction based on important observations on real videos. Finally, for the behavior analysis of fish we modeled and grouped trajectories into normal and abnormal classes, based on the uniformity of the trajectory slopes. As for the advantages of our system, we could say that future improvement is needed, but there is good indication that it can lead to a fully automatic system in order to be used under realistic conditions for real-time monitoring and interventions.

## Περίληψη

Αντικείμενο της παρούσας διπλωματικής είναι η υλοποίηση, η επιβεβαίωση και η αξιολόγηση ενός συστήματος επιτήρησης ψαριών σε ιχθυοκλωβούς η οποία στοχεύει στην ανάλυση της συμπεριφοράς κοπαδιού από ψάρια ή μεμονωμένων ατόμων αυτού. Μια τέτοια εφαρμογή αποδεικνύεται ιδιαίτερα χρήσιμη για την καθοδήγηση της τροφοδοσίας και τον έλεγχο ομαλής ανάπτυξης των ψαριών στον αναπτυσσόμενο κλάδο της ιχθυοκαλλιέργειας. Ένα κρίσιμο ζήτημα που χρειάστηκε να αντιμετωπίσουμε ήταν η βελτίωση των υποθαλάσσιων εικόνων στα βίντεο καθώς χαρακτηρίζονται από έντονο θόρυβο, ξαφνικές αλλαγές στη φωτεινότητα και θολό περιβάλλον. Έπειτα, βασικά βήματα στην υλοποίησή μας απετέλεσαν η επιλογή των κατάλληλων αλγορίθμων για την ανίχνευση, παρακολούθηση και μοντελοποίηση της τροχιάς των ψαριών. Βασικό εργαλείο για την υλοποίηση των παραπάνω ήταν η επιλογή του Kalman φίλτρου για την πρόβλεψη και εκτίμηση της κίνησης των ψαριών. Προβλήματα τα οποία παρουσιάστηκαν επιλύθηκαν σε ένα σημαντικό βαθμό μέσω της κατάλληλης επεξεργασίας των τροχιών μετά την εξαγωγή τους. Για την τελική ανάλυση της συμπεριφοράς των ψαριών, οι τροχιές μοντελοποιήθηκαν σε ομαλές και μη, με βάση την ομοιομορφία στην κλίση της τροχιάς. Στα πλεονεκτήματα του συστήματός μας μπορεί να καταγραφεί ότι, μέσω μελλοντικής βελτίωσής του, μπορεί να οδηγήσει σε ένα πλήρως αυτοματοποιημένο σύστημα με στόχο την εφαρμογή του σε ρεαλιστικές συνθήκες και πραγματικό χρόνο.

## Ευχαριστίες

Θα ήθελα να ευχαριστήσω:

Πρωτίστως τον επιβλέποντα καθηγητή μου κ. Ζερβάκη Μιχάλη για την πολύτιμη βοήθεια και καθοδήγησή του καθ' όλη τη διάρκεια εκπόνησης της διπλωματικής μου εργασίας καθώς επίσης και για την εμπιστοσύνη που μου έδειξε δίνοντας μου την δυνατότητα να εργαστώ στον τομέα που επιθυμούσα.

Επίσης τον καθηγητή κ. Πετράκη Ευριπίδη και τον Δρ. Παπανδρουλάκη Νίκο για την συνεισφορά τους ως μέλη της εξεταστικής επιτροπής.

Την υπεύθυνη του εργαστηρίου κυρία Μοιρογιώργου Ντίνα για τον πολύτιμο χρόνο που αφιέρωσε και την καθοδήγησή της από την αρχή μέχρι το τελικό στάδιο της διπλωματικής μου εργασίας καθώς και για την στήριξη και ενθάρρυνσή της.

Τους φίλους μου για όλες τις όμορφες εμπειρίες στα Χανιά και την ουσιαστική υποστήριξη.

Επίσης τον Βαγγέλη, την Αθηνά και τους συναδέλφους μου στο εργαστήριο DISPLAY για τη βοήθειά τους.

Τέλος οφείλω το μεγαλύτερο ευχαριστώ στην οικογένειά μου για την υλική και ψυχολογική υποστήριξη σε ολόκληρη την διάρκεια των σπουδών μου.

# Πίνακας Περιεχομένων

Abstract.....	2
Περίληψη.....	3
Ευχαριστίες.....	4
Κεφάλαιο 1: Εισαγωγή .....	9
1.1 Κίνητρο .....	9
1.2 Συναφείς εργασίες.....	10
1.3 Καινοτομία της εργασίας .....	17
1.4 Διάρθρωση διπλωματικής εργασίας .....	18
Κεφάλαιο 2: Βελτιστοποίηση και Κατάτμηση Εικόνας .....	19
2.1 Εισαγωγή.....	19
2.2 Βελτιστοποίηση Υποθαλάσσιας Εικόνας .....	20
2.2.1 Contrast Stretching.....	20
2.2.2 Histogram Equalization.....	22
2.2.3 Homomorphic Filtering.....	23
2.3 Κατάτμηση σε Υποθαλάσσιες Εικόνες .....	25
2.3.1 Thresholding .....	25
2.3.2 Active Contours .....	27
2.3.3 Απομάκρυνση ανεπιθύμητων περιοχών κατά το segmentation .....	32
Κεφάλαιο 3: Προτεινόμενη Μεθοδολογία.....	35
3.1 Εισαγωγή.....	35
3.2 Ανίχνευση Ψαριών.....	37
3.2.1 Αφαίρεση φόντου με Gaussian Mixture Models .....	39
3.2.2 Morphological Operations .....	43
3.2.3 Blob Analysis.....	45
3.2.4 Ακρίβεια Ανίχνευσης.....	48
3.3 Παρακολούθηση Κίνησης.....	51
3.3.1 Kalman Filtering .....	51
3.3.2 Ανάλυση της μεθοδολογίας μας .....	54
3.3.3 Ειδικές περιπτώσεις και προβλήματα .....	58
3.4 Επεξεργασία τροχιών .....	63
3.4.1 Επίλυση προβλημάτων.....	63

3.4.2 Τελικά αποτελέσματα tracking .....	69
Κεφάλαιο 4: Μοντελοποίηση τροχιάς .....	71
4.1 Εισαγωγή.....	71
4.2 Μοντελοποίηση με βάση την Εντροπία .....	74
Κεφάλαιο 5: Συμπεράσματα και μελλοντική δουλειά .....	79
Βιβλιογραφία.....	81

## Πίνακας Περιεχομένων Εικόνων

### Εικόνες Κεφάλαιο 2

Εικόνα 2.1 Αποτελέσματα contrast stretching σε RGB εικόνες.....	21
Εικόνα 2.2 Αποτέλεσμα contrast stretching σε gray scale εικόνα .....	21
Εικόνα 2.3 Αποτέλεσμα histogram equalization σε gray scale εικόνα.....	22
Εικόνα 2.4 Αποτέλεσμα homomorphic filtering σε gray scale εικόνα .....	24
Εικόνα 2.5 Αποτελέσματα μεθόδων enhancement και αντίστοιχα ιστογράμματα.....	25
Εικόνα 2.6 Σωστά αποτελέσματα threshold segmentation .....	26
Εικόνα 2.7 Εσφαλμένα αποτελέσματα threshold segmentation .....	27
Εικόνα 2.8 Αποτελέσματα segmentation based on active contour .....	31
Εικόνα 2.9 Επιτυχημένη απομάκρυνση διχτυών .....	33
Εικόνα 2.10 Αποτυχημένη απομάκρυνση διχτυών .....	34

### Εικόνες Κεφάλαιο 3

Εικόνα 3.1 Σχηματικό διάγραμμα της μεθοδολογίας.....	36
Εικόνα 3.2 Παραδείγματα από κοπάδια ψαριών .....	36
Εικόνα 3.3 Παραδείγματα από μικρό αριθμό από ψάρια .....	37
Εικόνα 3.4 Παραδείγματα από εικόνες με δύσκολες συνθήκες.....	38
Εικόνα 3.4.1 Dilation με rectangle structuring .....	44
Εικόνα 3.4.2 Erosion με rectangle structuring element .....	44
Εικόνα 3.4.3 Types of Structuring elements. ....	44
Εικόνα 3.5 Detection Video_1 .....	46
Εικόνα 3.6 Detection Video_2.....	46
Εικόνα 3.7 Detection Video_3.....	47

Εικόνα 3.8	Detection Video_6.....	47
Εικόνα 3.9	Detection σε κοπάδι ψαριών.....	48
Εικόνα 3.9.1	Frames από Video_1.....	51
Εικόνα 3.9.2	Frames από Video_2.....	51
Εικόνα 3.9.3	Frames από Video_3.....	52
Εικόνα 3.10	Trajectory visualization .....	56
Εικόνα 3.11	Παραδείγματα από σωστές τροχιές .....	57
Εικόνα 3.12	Επιτυχημένη πρόβλεψη Kalman.....	58
Εικόνα 3.13	Επικάλυψη για αρκετά frames μεταξύ δύο ψαριών .....	58
Εικόνα 3.14	Επικάλυψη σε πολλά ψάρια .....	59
Εικόνα 3.15	Αλλόκοτη κίνηση ψαριού.....	60
Εικόνα 3.16	Παράδειγμα εμφάνισης ίδιου ψαριού με διαφορετικά id .....	61
Εικόνα 3.17	Παράδειγμα επικάλυψης και διατήρηση ίδιας τροχιάς σε διαφορετικό ψάρι ...	62
Εικόνα 3.18	Παράδειγμα με σφάλματα στη τροχιά του ψαριού.....	63
Εικόνα 3.19	Σχηματικό διάγραμμα επεξεργασίας τροχιών .....	63
Εικόνα 3.20	Πληροφορίες βάση του bounding box.....	64
Εικόνα 3.21	Διόρθωση εσφαλμένης πληροφορίας στη τροχιά.....	65
Εικόνα 3.22	Διόρθωση εσφαλμένης πληροφορίας στη τροχιά.....	66
Εικόνα 3.23	Διαχωρισμός τροχιών με ίδιο id .....	67
Εικόνα 3.24	Διαχωρισμός τροχιών με ίδιο id .....	68

#### **Εικόνες Κεφάλαιο 4**

Εικόνα 4.1	Παραδείγματα τροχιών από ομαλή ανοδική πορεία.....	71
Εικόνα 4.2	Παραδείγματα τροχιών από ομαλή ευθεία πορεία .....	72
Εικόνα 4.3	Παραδείγματα τροχιών από ομαλή καθοδική πορεία.....	72
Εικόνα 4.4	Παράδειγμα μη ομαλής τροχιάς .....	73
Εικόνα 4.5	Παράδειγμα μη ομαλής τροχιάς .....	73
Εικόνα 4.6	Βελτίωση ιστογράμματος γωνιών .....	77
Εικόνα 4.7	Βελτίωση ιστογράμματος γωνιών .....	78

## Πίνακας Περιεχομένων Πινάκων

### Πίνακες Κεφάλαιο 3

Πίνακας 3.1	Υπολογισμός TP, TN, FP, FN .....	49
Πίνακας 3.2	Ακρίβεια Detection.....	49
Πίνακας 3.3	Ακρίβεια Tracking.....	69
Πίνακας 3.4	Ακρίβεια Tracking .....	70

### Πίνακες Κεφάλαιο 4

Πίνακας 4.1	Τιμές εντροπίας και κατεύθυνσης για ομαλές τροχιές .....	76
Πίνακας 4.2	Τιμές εντροπίας για μη ομαλές τροχιές .....	76



# Κεφάλαιο 1: Εισαγωγή

Στο κεφάλαιο αυτό γίνεται αναφορά ως προς το επιστημονικό πεδίο στο οποίο αναφέρεται η παρούσα διπλωματική εργασία καθώς και στο κίνητρο και τους στόχους της. Επίσης παρουσιάζονται συναφείς εργασίες τις σύγχρονης τεχνολογίας στην οποία βασιστήκαμε και ποια η καινοτομία της δικής μας. Τέλος παρουσιάζεται η διάρθρωσή της.

## 1.1 Κίνητρο

Η μηχανική όραση είναι ένας τομέας ιδιαίτερα ποικιλόμορφος και ενεργός στην επιστήμη, με τους ερευνητές της να παρουσιάζουν ενδιαφέρουσες και φιλόδοξες μεθοδολογίες που αντιμετωπίζουν προβλήματα του πραγματικού κόσμου. Ένα τέτοιο ζήτημα είναι η ανάπτυξη εφαρμογών που θα αυτοματοποιούν την διαδικασία ελέγχου και ανάλυσης της συμπεριφοράς των ψαριών στις ιχθυοκαλλιέργειες. Η ιχθυοκαλλιέργεια, αποτελεί σήμερα τον πιο αναπτυσσόμενο κλάδο ζωικής παραγωγής προϊόντων διατροφής. Σύμφωνα με τις προβλέψεις του Παγκόσμιου Οργανισμού Τροφίμων (FAO) και της Παγκόσμιας Τράπεζας μέχρι το 2030 πάνω από το 65% των αλιευτικών προϊόντων θα προέρχεται από την υδατοκαλλιέργεια. Στην Ελλάδα εκτιμάται ότι το 63% της εγχώριας παραγωγής αλιευτικών προϊόντων προήλθε από την υδατοκαλλιέργεια όπως αναφέρεται και στο Συνέδριο Ιχθυοκαλλιέργειας που πραγματοποιήθηκε τον Ιούνιο του 2018, [1]. Τα ψάρια και τα προϊόντα μεταποίησης αυτών αποτελούν έναν από τους πλέον εξωστρεφείς κλάδους της ελληνικής οικονομίας συμβάλλοντας θετικά στο εμπορικό ισοζύγιο της χώρας όπως αναφέρεται στο συνέδριο. Η διατροφή και η ποιότητα του νερού είναι οι πιο σημαντικοί παράγοντες για την φυσιολογική ανάπτυξη των ψαριών όπως περιγράφεται σε άρθρα της «Ελληνικής Κτηνιατρικής Εταιρίας», [2]. Μέχρι στιγμής οι τρόποι με τους οποίους πραγματοποιείται έλεγχος των πληθυσμών ψαριών σε μια μονάδα ιχθυοκαλλιέργειας με σκοπό τον έλεγχο της φυσιολογικής ή μη ανάπτυξης τους γίνεται με παρατηρήσεις, είτε με την φυσική παρουσία δυτών είτε με την καταγραφή βίντεο από υποβρύχιες κάμερες. Οι παρατηρήσεις αφορούν στο τρόπο κίνησης των ψαριών, στο χρώμα και την φυσιολογική μορφολογία τους, καθώς προβλήματα στη διατροφή ή την ποιότητα του νερού έχουν αντίκτυπο στα παραπάνω. Επίσης έχει σημασία η συμπεριφορά των ψαριών σε δεξαμενές ή κλουβιά, και το κοπάδιασμα τους είναι ένδειξη κανονικής συμπεριφοράς. Ωστόσο, η τεράστια ποσότητα των καταγεγραμμένων δεδομένων καθιστά την παρατήρηση παράξενων συμπεριφορών ψαριών πρακτικά αδύνατη. Είναι συνεπώς σημαντικό, εφόσον η εξέλιξη της τεχνολογίας μας το επιτρέπει, να υπάρξει οργανωμένος και αυτοματοποιημένος έλεγχος στην παραγωγή σε αυτό τον τομέα.

Στόχος της παρούσας διπλωματικής εργασίας είναι η δημιουργία ενός συστήματος για εξαγωγή και μοντελοποίηση της τροχιάς των ψαριών σε ρεαλιστικές συνθήκες. Βασικό σημείο είναι η αντιμετώπιση των προβλημάτων που έχουν οι υποθαλάσσιες εικόνες καθώς και η επιλογή κατάλληλων αλγορίθμων για την ανίχνευση και την παρακολούθηση πολλαπλών αντικειμένων που αποτελεί ιδιαίτερα δύσκολο ζήτημα. Η εφαρμογή μεθοδολογίας που ενσωματώνει κατάλληλους αλγορίθμους σε περιβάλλον ιχθυοτροφείου

θα μπορούσε να βοηθήσει στην εξαγωγή πολλών συμπερασμάτων για την συμπεριφορά των ψαριών και θα ήταν πολύ χρήσιμο για την καλύτερη χορήγηση τροφής και την ανάλυση των χαρακτηριστικών εκτροφής των ιχθυοπληθυσμών.

Τα βήματα που ακολουθούμε αφορούν στην βελτιστοποίηση της υποθαλάσσιας εικόνας, την κατάτμησή της με στόχο την ακριβή ανίχνευση των ψαριών και, τέλος, την εξαγωγή τροχιών είτε μεμονωμένων ψαριών είτε ομάδων με τελικό σκοπό την κατηγοριοποίηση των τροχιών με βάση το τι ορίζουμε εμείς ως ομαλές και το τι όχι.

## 1.2 Συναφείς εργασίες

Η παρακολούθηση (tracking) πολλαπλών αντικειμένων αποτελεί αντικείμενο μελέτης για πολλούς ερευνητές καθώς πρόκειται για ένα υψηλού επιπέδου πρόβλημα στη μηχανική όραση. Η εξαγωγή της τροχιάς των αντικειμένων αφορά στη συσχέτιση των ανιχνευμένων (detected) αντικειμένων από frame σε frame, με τέτοιο τρόπο ώστε να αναγνωρίζεται ότι πρόκειται για το ίδιο αντικείμενο σε μια αλληλουχία από frames σε ένα βίντεο. Ο κατάλληλος αλγόριθμος για το tracking επιλέγεται με βάση την εφαρμογή την οποία θέλουμε να αναπτύξουμε (automated surveillance, Motion-based recognition, vehicle navigation κτλ) καθώς και από τα χαρακτηριστικά των αντικειμένων και του βίντεο ( real time, Noise in images, non rigid or articulated nature of objects, scene illumination changes etc) όπως αναφέρεται και στις έρευνες «*A review on selected target tracking algorithms*», [3] και «*A survey on multiple object tracking algorithm*», [4].

Οι αλγόριθμοι αυτοί αποτελούνται από δύο μέρη. Το πρώτο αφορά στις μεθόδους για την αναπαράσταση και ανίχνευση των στόχων (Target Representation and Localization) και το δεύτερο αφορά στις μεθόδους για τη συσχέτιση τους σε κάθε frame (Filtering and Data Association).

Το Target Representation and Localization είναι μια bottom up διαδικασία. Η υπολογιστική ισχύς τέτοιων αλγορίθμων είναι συνήθως χαμηλή. Οι πιο γνωστές μέθοδοι αναλύονται παρακάτω:

- **Point tracking.** Τα αντικείμενα που ανιχνεύονται σε διαδοχικά frames αντιπροσωπεύονται από σημεία όπως το κέντρο (για παράδειγμα στο Blob Analysis) και η σύνδεση των σημείων βασίζεται στην προηγούμενη κατάσταση του αντικειμένου που μπορεί να περιλαμβάνει πληροφορίες για την θέση και κίνηση. Η μέθοδος διακρίνεται σε Deterministic και Statistical προσέγγιση.
- **Kernel-based tracking.** Εκτελείται υπολογίζοντας την κίνηση του αντικειμένου, που αντιπροσωπεύεται από μια αρχική περιοχή στο αντικείμενο, από το ένα frame στο άλλο. Η κίνηση είναι σε παραμετρική μορφή. Οι μέθοδοι Kernel χωρίζονται σε υποκατηγορίες βασισμένες στην εμφάνιση. Σε Template and Density-based Appearance Model και Multi-view Appearance Model.
- **Silhouette Tracking.** Παρέχει μια ακριβή περιγραφή του στόχου/ αντικειμένου. Στόχος είναι να βρει την περιοχή του αντικειμένου σε κάθε frame μέσω ενός μοντέλου αντικειμένου που δημιουργήθηκε χρησιμοποιώντας τα προηγούμενα

frames. Η εύρεση τροχιάς μέσω της σιλουέτας χωρίζεται σε δύο κατηγορίες, το Shape matching και το Contour tracking.

Το Filtering and Data Association, από την άλλη, είναι μια top-down διαδικασία που περιέχει την ενσωμάτωση προηγούμενης γνώσης σχετικά με τη σκηνή ή το αντικείμενο αντιμετωπίζοντας δυναμικές των αντικειμένων και αξιολογώντας διαφορετικές περιπτώσεις. Αυτές οι μέθοδοι επιτρέπουν το tracking για περίπλοκα αντικείμενα (μη σταθερό σχήμα) και με πιο περίπλοκες κινήσεις ή και επικαλύψεις μεταξύ τους. Η υπολογιστική ισχύς αυτών των αλγορίθμων είναι συνήθως αρκετά υψηλή. Ακολουθούν κάποιες από τις πιο γνωστές μεθόδους filtering:

- **Motion Models.** Τα μοντέλα κίνησης είναι ένας απλός αλγόριθμος πρόβλεψης και αρκετά συνηθισμένος για απλά συστήματα. Σκοπός είναι να προβλέψει την επόμενη θέση με βάση προηγούμενες παρατηρήσεις. Η εξίσωση που περιγράφει την μέθοδο:

$$P(t + 1) = P(t) + V(t) \quad (1.1)$$

Όπου  $P(t + 1)$  είναι η προβλεπόμενη επόμενη θέση την επόμενη χρονική στιγμή . Τα  $P(t)$  και  $V(t)$  είναι η παρούσα θέση και ταχύτητα αντίστοιχα. Υπάρχει η δυνατότητα χρήσης και τις επιτάχυνσης .

- **Kalman filter:** Το φίλτρο Kalman είναι ένα βέλτιστο αναδρομικό Bayesian φίλτρο για γραμμικές συναρτήσεις που υποβάλλονται σε Gaussian θόρυβο. Πρόκειται για ένα αναδρομικό φίλτρο πρόβλεψης, άρα λοιπόν υπολογίζει την κατάσταση ενός δυναμικού συστήματος αναδρομικά. Αυτό το στάδιο αποτελείται από δύο φάσεις, την πρόβλεψη (prediction) και τη διόρθωση (correction). Η φάση της πρόβλεψης προβάλλει μπροστά στο χρόνο την τρέχουσα κατάσταση και τη συνδιακύμανση σφάλματος, για την απόκτηση a priori προβλέψεων για το επόμενο χρονικό βήμα. Η φάση της διόρθωσης ενσωματώνει την τελευταία μέτρηση στο μοντέλο του συστήματος, για να πάρει τις a posteriori εκτιμήσεις της κατάστασης και της συνδιακύμανσης σφάλματος. Περαιτέρω ανάλυση του Kalman filter θα γίνει σε επόμενο κεφάλαιο καθώς είναι και η μέθοδος που επιλέξαμε για την υλοποίηση του αλγορίθμου μας.
- **Particle filter:** Το Particle φίλτρο χρησιμοποιεί μια Bayesian προσέγγιση για την εκτίμηση της κατάστασης για μη γραμμικές συναρτήσεις που υποβάλλονται σε μη Gaussian θόρυβο όπως αναφέρεται στο «*Particle Filters for Multiple Target Tracking*», [5]. Στο πλαίσιο αυτό, η συνάρτηση πυκνότητας πιθανότητας της κατάστασης εκτιμάται με βάση όλες τις διαθέσιμες πληροφορίες/μετρήσεις. Το particle filter προσεγγίζουν posterior πυκνότητα μέσω ενός συνόλου σωματιδίων (particles) και των σχετικών βαρών (weights) τους. Υπάρχουν δύο συστήματα, το ένα μοντελοποιεί την εξέλιξη του συστήματος στο χρόνο και το άλλο μοντελοποιεί

τις μετρήσεις για αποτελεσματικότερη εκτίμηση της επόμενης κατάστασης. Τα σωματίδια διαδίδονται με βάση το μοντέλο του συστήματος ενώ οι βαρύτητες υπολογίζονται αναδρομικά με βάση παρατηρήσεις στον θόρυβο που ελήφθησαν όπως αναφέρεται και την εργασία «*Tracking multiple objects with particle filtering*», [6]. Οι αλγόριθμοι Particle Filter εκτιμούν την posterior πυκνότητα ενσωματώνοντας τεχνικές δειγματοληψίας Monte Carlo. Θεωρούμε ένα δυναμικό σύστημα που εκπροσωπείται από την παρακάτω στοχαστική διαδικασία με συνάρτηση:

$$X_t = F_t(X_{t-1}, V_t) \quad , \text{όπου } (X_t) \in R^{n_x} \quad (1.2)$$

Σκοπός είναι να εκτιμήσουμε το  $(X_t)$  σε διακριτές τιμές με τη βοήθεια παρατηρήσεων του συστήματος που πραγματοποιούνται από τη στοχαστική διαδικασία που περιγράφεται από την συνάρτηση μετρήσεων:

$$Y_t = H_t(X_t, W_t) \quad , \text{όπου } (Y_t) \in R^{n_y} \quad (1.3)$$

Οι διαδικασίες  $(V_t) \in R^{n_v}$  και  $(W_t) \in R^{n_w}$  θεωρούνται λευκός θόρυβος. Οι  $F_t$  και  $H_t$  είναι μη γραμμικές. Το πρόβλημα ανάγεται στο να υπολογίσουμε για κάθε χρονική στιγμή  $t$  την πυκνότητα  $L_t$  του  $X_t$  δοσμένων των μετρήσεων  $L_t = p(X_t | Y_0=y_0, \dots, Y_t=y_t)$  μπορεί να εκτιμηθεί κάθε συνάρτηση της κατάστασης  $g(X_t)$  από το αποτέλεσμα  $\mathbb{E}(g(X_t) | Y_{0:t})$ . Το Recursive Bayesian filter λύνει ακριβώς αυτό το πρόβλημα σε δύο βήματα για κάθε χρόνο  $t$ .

Υποθέτοντας ότι γνωρίζουμε το  $L_{t-1}$ . Το **prediction** γίνεται με βάση την συνάρτηση:

$$p(X_t = x_t | Y_{0:t-1} = y_{0:t-1}) = \int_{R^{n_x}} p(X_t = x_t | X_{t-1} = x) L_{t-1}(x) dx \quad (1.4)$$

Το **correction** γίνεται με βάση τη παρακάτω συνάρτηση όπου το θεώρημα του Bayes μας επιτρέπει να υπολογίσουμε το  $L_t$  :

$$L_t(x_t) = \frac{p(Y_t = y_t | X_t = x_t) p(X_t = x_t | Y_{0:t-1} = y_{0:t-1})}{p(Y_t = y_t | Y_{0:t-1} = y_{0:t-1})} \quad (1.5)$$

Αν οι συναρτήσεις  $F_t$  και  $G_t$  είναι γραμμικές και οι  $W_t$  και  $V_t$  Gaussian θόρυβος τότε οι συναρτήσεις εκφράζουν το Kalman filter.

Εάν κοιτάξουμε το θέμα της παρούσας διπλωματικής εργασίας υπό το πρίσμα της εφαρμογής αυτή τη φορά, δηλαδή την παρακολούθηση συμπεριφοράς σε κοπάδι ψαριών υπό ρεαλιστικές συνθήκες, θα δούμε πως υπάρχουν στη βιβλιογραφία μελέτες που έχουν γίνει

πάνω στο θέμα αυτό αλλά κατά κύριο λόγο αφορούν σε συνθήκες ενυδρείου ή πρόκειται για την τροχιά σε ένα ή μικρό αριθμό από ψάρια. Παρακάτω παρουσιάζονται κάποιες από αυτές που αποτέλεσαν και βάση για την ανάπτυξη της δικής μας μεθοδολογίας.

Μια από τις πιο ολοκληρωμένες δουλειές που βρήκαμε για την εξαγωγή τροχιάς σε υποθαλάσσιες εικόνες είναι η εργασία «*ROV-based underwater vision system for intelligent fish ethology research.*», [7]. Πρόκειται για τον εντοπισμό και την εξαγωγή τροχιάς ενός ψαριού. Στην εργασία αυτή παρουσιάζεται ένα σύστημα ROV (remotely operated vehicles) για την καταγραφή της συμπεριφοράς των ψαριών στους ωκεανούς, έχοντας γνώση για την απόσταση της κάμερας από το βυθό της θάλασσας για την ταχύτητα και την κατεύθυνση με την οποία κινείται η κάμερα. Οι εικόνες δέχονται κάποια προ-επεξεργασία για να μειωθεί ο θόρυβος (wavelet decomposition), να διορθωθεί ο μη ομοιόμορφος φωτισμός, να ενισχυθούν οι αντιθέσεις και ταυτόχρονα να ακονιστούν οι ακμές (homomorphic filtering) με σκοπό την εξαγωγή χαρακτηριστικών. Το segmentation έγινε με curve evolution βασισμένο Mumford shah και το level set. Στη περίπτωση αυτή η τροχιά εξάγεται εφαρμόζοντας particle filtering στη διανομή του χρώματος της υποβρύχιας εικόνας. Γίνεται επίσης μια σύγκριση με την μέθοδο mean shift η οποία είχε καλά αποτελέσματα για real time επίβλεψη, κάτι το οποίο δεν ίσχυε για τις περιπτώσεις επικάλυψης και σε περιπτώσεις αλλαγών της εμφάνισης του ψαριού. Για το λόγο αυτό και το particle filtering ήταν πιο αποδοτικό. Τα αποτελέσματα προσομοίωσης είναι καλά και ως προς την αξιοπιστία και ως προς την αποτελεσματικότητα.

Στην εργασία «*Underwater Image Enhancement Using an Integrated Colour Model*», [8] προτείνεται ένα σύστημα ανίχνευσης της συμπεριφοράς ψαριών βασισμένο στην ανίχνευση του κεφαλιού του ψαριού. Το σχήμα και τα χαρακτηριστικά της εικόνας του ψαριού που προέρχονται από την κλίμακα του γκρι χρησιμοποιούνται για τον εντοπισμό της θέσης των ψαριών. Για κάθε ανιχνευμένη κεφαλή ψαριού, χρησιμοποιούμε τη γκρίζα κατανομή της περιοχής της κεφαλής για να υπολογίσουμε την κατεύθυνση του. Αυτό γίνεται συγκρίνοντας τα χαρακτηριστικά στα άκρα του ψαριού οπότε και φαίνεται ποια είναι η ουρά και ποιο το κεφάλι. Στη συνέχεια, η πληροφορία για τη θέση και την κατεύθυνση από την ανίχνευση ψαριών συνδυάζονται για να δημιουργήσουν μια συνάρτηση κόστους (cost function) των ψαριών με σκοπό την εκτίμηση του λάθους στο μοντέλου αναφορικά με την εκτίμηση της σχέσης μεταξύ τους. Η συνάρτηση κόστους και η μέθοδος συνολικής βελτιστοποίησης (global optimization method) χρησιμοποιούνται για να αντιστοιχηθούν οι τροχιές σε κάθε frame εικόνας δημιουργώντας την τελική τροχιά. Τα αποτελέσματα δείχνουν ότι η μέθοδος αυτή μπορεί να ανιχνεύσει με ακρίβεια τις πληροφορίες θέσης και κατεύθυνσης του κεφαλιού των ψαριών, και έχει καλή απόδοση παρακολούθησης για δεκάδες ψάρια. Στη συγκεκριμένη μέθοδο δεν χρησιμοποιείται κάποιος αλγόριθμος πρόβλεψης, παρόλα αυτά τα αποτελέσματα είναι πολύ καλά και σε περιπτώσεις επικαλύψεων. Αυτό οφείλεται σε σημαντικό βαθμό στις τεχνικές συνθήκες που τραβήχτηκαν τα βίντεο. Η παρακολούθηση έγινε με σταθερή κάμερα σε ψάρια που βρίσκονταν σε ενυδρείο με το ύψος του νερού να μην ξεπερνά τα 3cm βάθος και λευκό φόντο. Ο μεγάλος αριθμός των frame ανά δευτερόλεπτο (30 και 40) σε συνδυασμό με το μικρό μέγεθος των ψαριών δίνει μεγάλη ακρίβεια των κινήσεών τους.



Στην εργασία «*A rule-based event detection system for real-life underwater domain*», [9] παρουσιάζεται ένα αυτόματο σύστημα ανίχνευσης συμβάντων για μοναχικές συμπεριφορές και συμπεριφορές ζευγαριών από ψάρια υπό πραγματικές υποθαλάσσιες συνθήκες καθώς γίνεται και αναγνώριση των ψαριών. Το σύστημα κάνει ανίχνευση και εξαγωγή της τροχιάς. Δεδομένου ότι η ανίχνευση είναι πολύ σημαντική χρησιμοποιήθηκαν 6 αλγόριθμοι αφαίρεσης φόντου οι οποίοι χρησιμοποιούνται ως weak classifiers για τον αλγόριθμο ενίσχυσης Ad Boost πετυχαίνοντας έτσι τα καλύτερα δυνατά αποτελέσματα. Η εξαγωγή της τροχιάς έγινε με Covariance Matrix το οποίο περιείχε πληροφορία και για την κίνηση και τα χαρακτηριστικά κάθε ψαριού κάτι το οποίο είναι χρήσιμο και αξιόπιστο και για την εξαγωγή τροχιάς αλλά και για την αναγνώριση του είδους των ψαριών. Γίνεται σύγκριση και με τους αλγορίθμους στο Particle Filtering και στο Mean Shift αποδεικνύοντας την ισχύ του Covariance Matrix. Στην εργασία αυτή επόμενο βήμα είναι η μοντελοποίηση των εξαγόμενων τροχιών με Hidden Markov Models. Μετά τη μοντελοποίηση γίνεται εκπαίδευση με το 70% των εξαγόμενων τροχιών για την εξαγωγή προτύπων όπου με βάση κάποια thresholds, ταξινομούνται σε συμπεριφορές μοναχικές ή σε ζευγάρια. Το υπόλοιπο 30% των τροχιών χρησιμοποιείται για την δοκιμή του συστήματος. Για την διαδικασία αυτή είναι σημαντικό να υπάρχουν αρκετά δεδομένα. Το σύστημα είχε 500,000 βίντεο με διάρκεια 10 λεπτών το καθένα επίσης έγιναν 31.221 ανιχνεύσεις και συνολικά που αντιστοιχούν σε 2.111 διαφορετικά ψάρια.

Στις εργασίες «*Fish School System Identification and Control Based on Artificial Neural Network*», [10] και «*Assessing fish abundance from underwater video using deep neural networks*», [11] γίνεται μια προσπάθεια αντιμετώπισης του προβλήματος της εξαγωγής τροχιάς με νευρωνικά δίκτυα. Στο [10], σκοπός είναι η εξαγωγή και μοντελοποίηση της τροχιάς μέσω των νευρωνικών δικτύων για να χρησιμοποιηθεί για την προσομοίωση της κίνησης σε ρομποτικά συστήματα ψαριών. Για να γίνει η ανίχνευση, τα δεδομένα από τα δύο βίντεο χωρίζονται στις R,G,B συντεταγμένες όπου το R και B προστίθενται εφόσον είναι οι συντεταγμένες που διευκολύνουν να εξαχθεί το ψάρι από το background. Με thresholding και την εφαρμογή κάποιων φίλτρων για την καλύτερη ανάδειξη των ψαριών γίνεται η εικόνα binary και με λευκό συμβολίζονται τα ψάρια. Ύστερα με το blob analysis βρίσκουμε τα κέντρα του κάθε ψαριού και με βάση την κίνηση γίνεται το tracking. Τα αποτελέσματα χρησιμοποιούνται για την εκπαίδευση των νευρωνικών δικτύων. Οι τεχνικές συνθήκες του βίντεο δίνουν την δυνατότητα για την εύκολη εξαγωγή της τρισδιάστατης τροχιάς καθώς έχουμε δυο σταθερές κάμερες που μαγνητοσκοπούν μόλις 5 ψάρια σε συνθήκες ενυδρείου με λευκό φόντο. Ενώ στο [11], ένα Region Proposal Network (RPN) που εισήχθη από έναν ανιχνευτή αντικειμένων τον Faster R-CNN συνδυάστηκε με τρία δίκτυα ταξινόμησης με σκοπό την ανίχνευση και την αναγνώριση των ειδών από ψάρια που λαμβάνονται από απομακρυσμένους υποβρύχιους σταθμούς. Το σύστημα καταφέρνει ακρίβεια 82,4% σε ιδιαίτερα δύσκολες συνθήκες.

Στις εργασίες «*Automatic Analysis of Fish Behaviors and Abnormality Detection*», [24] και «*Understanding fish behavior during typhoon events in real-life underwater environments*», [25] παρουσιάζονται ολοκληρωμένες δουλειές για τον εντοπισμό, την παρακολούθηση και την ανάλυση τροχιάς/συμπεριφοράς ψαριών σε

διαφορετικές συνθήκες. Η [24] αφορά τη διαφοροποίηση στην συμπεριφορά ομάδας ψαριών σε ένα ενυδρείο με καθαρό νερό και με τοξικό νερό. Γίνεται αυτόματη ταξινόμηση προτύπων συμπεριφοράς και εντοπισμός διαφοροποίησης με βάση την μέθοδο Spectral Clustering. Η εξαγωγή χαρακτηριστικών γίνεται με βάση το Fast Level Set, όπου ουσιαστικά κάθε ψάρι αντιπροσωπεύεται σαν μια περιοχή. Δοσμένων των περιοχών αυτών του κάθε ψαριού υπολογίζεται πίνακας ομοιοτήτων μεταξύ των περιοχών των ψαριών με Gaussian similarity function. Έπειτα υπολογίζεται ένα γράφημα Laplacian που υπολογίζει τις ιδιοτιμές και τα ιδιοδιανύσματα. Εν συνεχεία με τον K-means αλγόριθμο ομαδοποιούνται τα χαρακτηριστικά και εξάγονται πρότυπα τροχιάς με βάση την μικρότερη απόσταση που έχουν αυτά τα χαρακτηριστικά από frame σε frame. Ενώ η [25] αφορά τον προσδιορισμό τροχιάς/συμπεριφοράς ψαριών σε συνθήκες τυφώνα στο φυσικό τους περιβάλλον. Το detection γίνεται με 4 διαφορετικούς τρόπους. Gaussian Mixture Model, Adaptive Poisson Mixture, Intrinsic Model (IM) and Wave-Back (WB). Η παρακολούθηση της τροχιάς γίνεται με covariance matrix όπου κρατάμε χρήσιμες πληροφορίες όπως συντεταγμένες, RGB τιμές, μέση και τυπική απόκλιση του ιστογράμματος των τιμών του γκρι. Αυτό που διαφοροποιείται σε σχέση με άλλες μεθόδους είναι ότι εδώ δεν παρατηρούμε την Ευκλείδεια απόσταση μεταξύ των frame αλλά τις ομοιότητες στο covariance matrix. Υπολογίζονται μεγέθη όπως Difference of shape ratio between frames, Histogram difference, Direction smoothness, Speed smoothness, Texture difference, Temporal persistence. Οι τιμές αυτές επιστρέφουν αποτέλεσμα από 0-1 οπότε και βρίσκουμε το μέσο όρο τους. Έτσι υπολογίζεται η ακρίβεια στις τροχιές και αν πρόκειται για ομαλές ή όχι.

Πέραν των εργασιών που αφορούν σε παρόμοια θεματολογία με αυτή της παρούσας διπλωματικής εργασίας, εντοπίσαμε αρκετά ενδιαφέρουσες εργασίες με γενικότερο θέμα την μεμονωμένη παρακολούθηση πολλών κινούμενων αντικειμένων σε βίντεο. Γενικά, στα συστήματα πλοήγησης και τα συστήματα επιτήρησης, η παρακολούθηση αντικειμένων είναι ένα απαραίτητο πρώτο βήμα που πρέπει να υλοποιείται.

Ενδεικτικά, η εργασία «*Fluoroscopic tracking of multiple implanted fiducial markers using multiple object tracking*», [12] αφορά στην εφαρμογή αλγορίθμων παρακολούθησης σε βίντεο ιατρικού περιεχομένου. Παρουσιάζεται η παρακολούθηση κινητού όγκου μέσω της τοποθέτησης κάποιων δεικτών. Προτείνεται ένα σύστημα εντοπισμού των δεικτών που μπορεί να παρακολουθεί ταυτόχρονα πολλαπλούς δείκτες, χωρίς να προκαλεί σύγχυση, αρκετά ισχυρό για να συνεχίσει την παρακολούθηση ακόμη και όταν οι δείκτες κινούνται πίσω από την ανατομία του σκελετικού συστήματος και χάνονται για μικρό διάστημα. Οι δείκτες εντοπίζονται αρχικά χειροκίνητα και έπειτα με template matching γίνεται η ανίχνευσή τους σε κάθε frame. Έστερα για το tracking εφαρμόζεται ο linear prediction αλγόριθμος. Η μέθοδος έχει πολύ καλά αποτελέσματα ειδικά όταν συνδυάζονται και με πρότυπα αναπνοής του ασθενούς.

Στην εργασία «*Αναγνώριση και Χαρακτηρισμός Κίνησης σε Κυκλοφοριακά Βίντεο*», [13] παρουσιάζεται η ανίχνευση και ο εντοπισμός κινούμενων οχημάτων με σκοπό την επιτήρηση της κυκλοφορίας. Στην διπλωματική εργασία παρουσιάζεται ένα ολοκληρωμένο σύστημα επιτήρησης σε κυκλοφοριακά βίντεο για τον υπολογισμό της ταχύτητας των κινούμενων οχημάτων. Το σύστημά επεξεργάζεται βίντεο τα οποία έχουν ληφθεί από μια

στατική κάμερα και έχει τη δυνατότητα να ανιχνεύει, να εντοπίζει και να ταξινομεί οχήματα. Αρχικά τα frames του κάθε βίντεο υποβάλλονται σε έναν μετασχηματισμό προοπτικής και στη συνέχεια ακολουθούν, στο μετασχηματισμένο πλέον χώρο, κατά σειρά τα στάδια: της ανίχνευσης οχήματος από τρεις διαφορετικές τεχνικές οι οποίες συγκρίνονται, του εντοπισμού οχήματος από τον αλγόριθμο Kalman Filter, του υπολογισμού της ταχύτητας οχήματος βάση της μετακίνησης του κέντρου βάρους και τέλος της ταξινόμησης οχήματος βάση μορφολογικών χαρακτηριστικών του.

Σημαντική και ιδιαίτερα ποικιλόμορφη είναι η έρευνα που γίνεται για την παρακολούθηση κίνησης πεζών σε πλήθος ανθρώπων για την αξιολόγηση συμπεριφοράς τους όπως βλέπουμε στην εργασία «*Stable multi-target tracking in real-time surveillance video*», [15] επίσης στο «*Application of tracking technologies to the study of pedestrian spatial behavior*», [16] και στο «*Abnormal crowd behavior detection using social force model*», [17] για την αναγνώριση ανώμαλης συμπεριφοράς και την καταμέτρηση του πλήθους. Η παρακολούθηση αυτή αφορά είτε σε real time είτε σε εφαρμογές από αποθηκευμένα βίντεο. Στην εργασία «*Realtime multilevel crowd tracking using reciprocal velocity obstacles*», [18] παρουσιάζεται ένας αλγόριθμος για την παρακολούθηση πεζών σε πλήθος ανθρώπων με εφαρμογή σε πραγματικό χρόνο βασισμένο στο particle filter. Η προσέγγιση στη συγκεκριμένη εργασία αλλάζει δυναμικά τον αριθμό των σωματιδίων που κατανέμονται σε κάθε πεζό βάσει πολλαπλών μετρήσεων οι οποίες προέρχονται από ένα μη γραμμικό παραμετρικό μοντέλο κίνησης πολλαπλών παραγόντων το Reciprocal Velocity Obstacles «*Reciprocal n-body collision avoidance*», [19]. Η προσέγγιση αυτή αποσκοπεί στο να μειωθεί το υπολογιστικό κόστος κάτι το οποίο είναι πολύ σημαντικό για real time εφαρμογές.

Στην εργασία «*Unsupervised incremental learning for improved object detection in a video*», [20] παρουσιάζεται μια ενδιαφέρουσα μέθοδος που αποσκοπεί στη βελτίωση των αλγορίθμων παρακολούθησης σε πλήθος ανθρώπων. Το detection γίνεται μέσω ενός ανιχνευτή που έχει εκπαιδευτεί για την παρακολούθηση ανθρώπων σε διάφορες συνθήκες από δεδομένα σε βίντεο. Εδώ προτείνεται μια μέθοδος για την βελτιστοποίηση του ανιχνευτή μέσω του multiple instance learning (MIL). Ο αλγόριθμος εφαρμόζεται σε real time συνθήκες όπου εντοπίζονται οι εσφαλμένες ανιχνεύσεις και δίνονται σαν αποτελέσματα στο MIL. Ύστερα offline γίνεται επανεκπαίδευση του αλγορίθμου και ο ανιχνευτής έχει βελτιώσει τα αποτελέσματά του. Η μέθοδος αυτή δίνει πολύ καλά αποτελέσματα στο tracking του πλήθους.

Στην εργασία «*Kalman filter based multiple objects detection-tracking algorithm robust to occlusion*», [21] αναπτύσσεται μία μέθοδος tracking για ένα ή περισσότερα αντικείμενα. Εφόσον γίνεται η ανίχνευση των αντικειμένων με τη χρήση του Kalman filter και cost function επιτυγχάνεται το tracking όπως στην εργασία [13]. Η διαφοροποίηση στην εργασία αυτή βρίσκεται στο ότι χρησιμοποιούνται τα χαρακτηριστικά της περιοχής του αντικείμενου και της απόστασης. Όταν γίνεται επικάλυψη το μέγεθος των αντικειμένων αυξάνεται απότομα. Όταν εντοπιστεί μια τέτοια περίπτωση το cost function κρατά μόνο την πληροφορία της απόστασης. Με τον τρόπο αυτό είναι πιο αξιόπιστο το tracking. Η μέθοδος



αυτή έχει εφαρμοστεί σε πλήθος από ανθρώπους και σε σμήνος από πουλιά δίνοντας πολύ καλά αποτελέσματα και είναι κατάλληλη για real-time εφαρμογές.

Στην εργασία «*Multi-object tracking in non-stationary video using bacterial foraging swarms*», [22] παρουσιάζεται μια μέθοδος σε ιδιαίτερα δύσκολες συνθήκες με μη σταθερή κάμερα, σε real time και προσπαθεί να αντιμετωπίσει το πρόβλημα των επικαλύψεων. Η μέθοδος βασίζεται σε μεθόδους swarm intelligence που αφορά στον τρόπο σύνδεσης συλλογικοτήτων με σκοπό την αυτόματη ομαδοποίηση «*Swarm intelligence*», [23] και ειδικότερα Bacterial Foraging Swarms. Προσομοιώνοντας την κίνηση του E.coli βακτηρίου αρχικοποιεί έναν αριθμό από particles τα οποία απλώνονται στο χώρο αναζητώντας ‘τροφή’. Όπου τροφή θεωρείται ένα fitness function δοσμένο από το χρήστη, στην εργασία αυτή πρόκειται για ένα χρωματικό ιστόγραμμα. Σε κάθε αναζήτηση τα particles αξιολογούν τη θέση τους με βάση το gradient και εάν βρίσκονται πολύ μακριά αφαιρούνται, ενώ εκείνα που βρίσκονται πιο κοντά παράγουν και άλλα particles στα σημεία αυτά. Η μέθοδος αυτή αντιμετωπίζει την επικάλυψη μερική ή ολική για 4 αντικείμενα σε μη σταθερή κάμερα.

Με δεδομένο ότι ένα τέτοιο σύστημα δεν έχει υπάρξει ξανά με εφαρμογή σε περιβάλλον ιχθυοκαλλιέργειας, η εκτενής βιβλιογραφική έρευνα βόηθησε σε μεγάλο βαθμό στην υλοποίηση της συγκεκριμένης εργασίας. Από τη μια πλευρά μας βοήθησε στο να δούμε από πιο σφαιρική πλευρά το πρόβλημα και να μπορέσουμε να αντιληφθούμε τη συνολική δομή της εργασίας. Από την άλλη, στάδια της διπλωματικής μας όπως για παράδειγμα οι μεθοδολογίες για το detection αποτέλεσαν αντικείμενο μελέτης και σύγκρισης σε ολοκληρωμένες εργασίες με αποτέλεσμα τα συμπεράσματά τους να μας βοηθήσουν στην τελική επιλογή των δικών μας μεθόδων. Αναφέρεται αναλυτικά σε κάθε κεφάλαιο οποιοδήποτε κομμάτι της βιβλιογραφίας χρησιμοποιήθηκε.

### 1.3 Καινοτομία της εργασίας

Στην παρούσα διπλωματική παρουσιάζουμε ένα σύστημα παρακολούθησης με προοπτική τη δημιουργία πλήρως αυτοματοποιημένο συστήματος και με σκοπό την εξαγωγή συμπερασμάτων για την συμπεριφορά ψαριών σε ιχθυοκλωβούς. Σε γενικές γραμμές η καινοτομία της εργασίας μας εντοπίζεται στα ακόλουθα σημεία:

- Στην πρόταση ενός κατάλληλου συνδυασμού μεθοδολογιών για την εξαγωγή αποτελεσμάτων σε συνθήκες κίνησης σμήνους οργανισμών.
- Στην κατάλληλη παραμετροποίηση των αλγορίθμων ώστε να επιτρέπουν την ομαδοποίηση σε κοπάδι από την μια και την εξειδίκευση σε μεμονωμένες συμπεριφορές ψαριών.
- Στην βελτίωση ή την υλοποίηση δικών μας μεθοδολογιών στηριγμένων σε ειδικές παρατηρήσεις στο υπό εξέταση περιβάλλον.

Το σύστημα που προτείνεται περιλαμβάνει μεθόδους βελτιστοποίησης υποθαλάσσιων εικόνων και επεξεργασίας των με σκοπό την απομάκρυνση ανεπιθύμητων περιοχών που αφορούν σε παρακείμενα αντικείμενα, όπως τα δίχτυα. Αναφορικά με την κύρια μεθοδολογία εντοπισμού και παρακολούθησης, έγινε έρευνα βιβλιογραφίας με σκοπό την

επιλογή, συνδυασμό και την υλοποίηση μιας ροής διεργασιών που είναι κατάλληλη για εξαγωγή αποτελεσμάτων σε υποθαλάσσιες συνθήκες. Η εφαρμογή συγκεκριμένων παραμέτρων όσο αναφορά τις μορφολογικές ιδιότητες στις περιοχές εντοπισμού, με σκοπό τη μελέτη της μεθοδολογίας για κοπάδι και για μεμονωμένη συμπεριφορά, αποτελεί σημαντικό κομμάτι της εργασίας μας. Τα αποτελέσματα σε τέτοιες εφαρμογές χρήζουν περαιτέρω επεξεργασίας, την οποία και υλοποιήσαμε εκμεταλλευόμενοι χαρακτηριστικά που αφορούν στο μέγεθος του κάθε εντοπιζόμενου ψαριού και της εξαγόμενης τροχιάς του. Τέλος για την ανάλυση και ομαδοποίηση κίνησης σε ομαλές και μη συμπεριφορές, εκμεταλλευτήκαμε χαρακτηριστικά κλίσης στις τροχιές των ψαριών και εφαρμόσαμε την εντροπία σαν μέτρο για μοντελοποίηση της ομαλότητας της τροχιάς. Πρέπει να σημειωθεί ότι τα διαθέσιμα δεδομένα ήταν λίγα και δεν αφορούσαν την ποικιλία συμπεριφορών που αναμένονται σε ιχθυοκλωβούς. Για τον λόγο αυτό δημιουργήσαμε τεχνητά κατάλληλα βίντεο τα οποία και εξετάσαμε. Στην συνέχεια θα πρέπει να εξεταστεί η μεθοδολογία σε πολλαπλά πραγματικά βίντεο κάτω από διαφορετικές συνθήκες λήψης.

## 1.4 Διάρθρωση διπλωματικής εργασίας

Εδώ περιγράφεται ο τρόπος με τον οποίο έχει δομηθεί η διπλωματική εργασία με σκοπό να επιτευχθούν οι στόχοι που θέσαμε παραπάνω. Η διπλωματική χωρίζεται στα εξής πέντε κεφάλαια: Εισαγωγή, Βελτιστοποίηση και Κατάτμηση εικόνας, Προτεινόμενη Μεθοδολογία, Μοντελοποίηση τροχιάς και Συμπεράσματα και Μελλοντική δουλειά.

Το πρώτο κεφάλαιο παρέχει μια αρχική εισαγωγή στο θέμα της διπλωματικής και επίσης περιγράφονται το υπόβαθρο, το κίνητρο, η καινοτομία και οι στόχοι της.

Το δεύτερο κεφάλαιο αφορά στην αντιμετώπιση των προβλημάτων που εντοπίζονται στις υποθαλάσσιες εικόνες με σκοπό τη βελτιστοποίησή της με μια σειρά από μεθοδολογίες και αφορούν στην ανάδειξη χαρακτηριστικών της εικόνας όπως οι ακμές και το contrast με τις μεθοδολογίες Contrast Stretching, Histogram Equalization και Homomorphic Filtering. Η βελτίωση της εικόνας αποσκοπεί στο να μπορούν να εξαχθούν ευκολότερα πληροφορίες από αυτήν ώστε να έχουμε καλύτερα αποτελέσματα για το επόμενο στάδιο, την κατάτμησή της σε ομοιόμορφες περιοχές (image segmentation) με τελικό στόχο τον διαχωρισμό των ψαριών από το φόντο. Για την κατάτμηση επιλέγουμε δύο μεθόδους, το thresholding και το segmentation βασισμένο σε active contours. Πρόκειται ουσιαστικά για μια πρώτη εφαρμογή και έρευνα πάνω σε πραγματικά δεδομένα που μας δόθηκαν.

Στο τρίτο και πιο σημαντικό κεφάλαιο αναλύεται η μεθοδολογία που προτείνουμε για την παρακολούθηση των ψαριών. Τα δυο βασικά στάδια είναι η ανίχνευση των ψαριών (fish detection) και η παρακολούθηση των τροχιών τους (fish tracking). Για την ανίχνευση εκμεταλλευόμαστε την κίνηση των ψαριών και επιλέγουμε το αφαίρεση φόντου βασισμένο σε Gaussian Mixture Models και για την παρακολούθηση έναν αλγόριθμο πρόβλεψης που κάνει χρήση του Kalman Filtering. Επίσης στο κεφάλαιο αυτό παρουσιάζεται η επεξεργασία που κάνουμε στις τροχιές μετά την υλοποίηση του αλγορίθμου. Τέλος παρουσιάζονται τα

αποτελέσματα για την ακρίβεια στο για την παρακολούθηση πριν και μετά την επεξεργασία των τελικών τροχιών.

Το τέταρτο κεφάλαιο αφορά την μοντελοποίηση των εξαγόμενων τροχιών με σκοπό την αναγνώριση του είδους της κίνησής. Εξάγοντας τα ιστογράμματα των γωνιών σε κάθε τροχιά , βρίσκουμε την εντροπία στο καθένα και με βάση την τιμή της μπορούμε να έχουμε τον επιθυμητό διαχωρισμό όσο αφορά την ομαλή ή όχι τροχιά που ακολουθούν τα ψάρια.

Τέλος στο πέμπτο κεφάλαιο της διπλωματικής εργασίας παραθέτονται τα τελικά συμπεράσματα της εργασίας με βάση τα αποτελέσματα της και σχολιάζεται η επίτευξη των στόχων που είχαν τεθεί εξ αρχής. Επίσης δίνονται προτάσεις για μελλοντική βελτίωση.

## Κεφάλαιο 2: Βελτιστοποίηση και Κατάτμηση Εικόνας

### 2.1 Εισαγωγή

Το πρώτο μέρος αυτού του κεφαλαίου αφορά στη βελτιστοποίηση (enhancement) της υποθαλάσσιας εικόνας για την ευκολότερη εξαγωγή πληροφορίας. Οι υποθαλάσσιες εικόνες χαρακτηρίζονται από κακή ορατότητα και ασθενή αντίθεση των χρωμάτων, που προκαλούνται από τη σκέδαση και την αντανάκλαση του φωτός με την είσοδό του αισθητήρα κάμερας στο νερό καθώς και τις εναλλαγές στην πυκνότητα του νερού όσο αυξάνεται το βάθος τοποθέτησης της κάμερας. Μία επιπλέον αιτία για την κακή ποιότητα των υποθαλάσσιων εικόνων αποτελεί και η ύπαρξη μικροοργανισμών ή σκουπίδια που μπορεί να αιωρούνται στο νερό. Στη μελέτη «*Underwater Image Enhancement Using an Integrated Colour Model*», [8] γίνεται εκτενής ανάλυση των προβλημάτων που παρατηρούνται στις υποθαλάσσιες εικόνες και πώς αυτά προκαλούνται. Για την αντιμετώπισή τους έχουν αναπτυχθεί πολλές τεχνικές και μέθοδοι όπως αναφέρεται και στην εργασία «*Review of underwater image enhancement techniques*», [26]. Οι μεθοδολογίες αυτές είτε εφαρμόζονται στο πεδίο της συχνότητας είτε στο χωρικό πεδίο. Η βελτίωση της εικόνας στη δική μας εργασία αποσκοπεί στο να μπορούν να εξαχθούν ευκολότερα πληροφορίες από αυτήν ώστε να είναι χρήσιμη για τη μετέπειτα επεξεργασία της μέσω τεχνικών segmentation. Η βελτιστοποίηση αφορά στην ανάδειξη χαρακτηριστικών της εικόνας όπως οι ακμές και το contrast. Στη παρούσα εργασία εφαρμόσαμε κάποιες από τις πιο διαδεδομένες μεθόδους που εντοπίσαμε μέσα από την βιβλιογραφία και εφαρμόζονται και σε εργασίες όμοιες με τη δική μας «*ROV-based underwater vision system for intelligent fish ethology research*», [7]. Αυτές είναι το Contrast Stretching, το Histogram Equalization και το Homomorphic Filtering. Ο συνδυασμός τους έφερε πολύ καλά αποτελέσματα, τα οποία και παρουσιάζονται αναλυτικότερα παρακάτω στο πρώτο μέρος του κεφαλαίου αυτού.

Το δεύτερο μέρος του κεφαλαίου αναφέρεται στη κατάτμηση της εικόνας (segmentation). Κατάτμηση είναι η διαδικασία διαίρεσης της εικόνας σε ομοιόμορφες

περιοχές. Το σημείο μέχρι το οποίο η υποδιαίρεση συνεχίζεται εξαρτάται από το πρόβλημα το οποίο πρέπει να επιλυθεί. Αυτό σημαίνει ότι η κατάτμηση πρέπει να σταματά όταν τα αντικείμενα ενδιαφέροντος έχουν απομονωθεί. Οι μεθοδολογίες που εξετάσαμε με βάση και τα αποτελέσματα από την βελτιστοποίηση αφορούν στο thresholding και τα active contours.

## 2.2 Βελτιστοποίηση Υποθαλάσσιας Εικόνας

### 2.2.1 Contrast Stretching

Ο αλγόριθμος Contrast Stretching χρησιμοποιείται για τη βελτίωση του contrast στις εικόνες, στην αντίθεση και την ένταση (δηλαδή στο χρώμα) όπως αναφέρεται τις εργασίες «*Underwater Image Enhancement Using an Integrated Colour Model*», [8] και «*Review of underwater image enhancement techniques*», [26]. Αυτό επιτυγχάνεται με το ‘τέντωμα’ του εύρους των τιμών των χρωμάτων ή του γκρι αν πρόκειται για grayscale εικόνα, χρησιμοποιώντας έτσι όλες τις πιθανές τιμές. Με τη μέθοδο αυτή καταφέρνουμε το φωτεινότερο τμήμα να γίνει ακόμα πιο φωτεινό και το σκοτεινότερο τμήμα ακόμα πιο σκοτεινό αναδεικνύοντας, έτσι, τις αντιθέσεις στην ένταση. Ο αλγόριθμος χρησιμοποιεί γραμμική συνάρτηση κλιμάκωσης του κάθε pixel με βάση την παρακάτω συνάρτηση:

$$P_o = \frac{(P_i - c)(b - c)}{d - c} + a \quad (2.1)$$

Όπου  $P_o$  το τελικό pixel,  $P_i$  το pixel που εξετάζουμε,  $a$  η ελάχιστη τιμή του επιθυμητού εύρους,  $b$  η μέγιστη τιμή του επιθυμητού εύρους,  $c$  η ελάχιστη τιμή στο υπάρχων εύρος,  $d$  η μέγιστη τιμή στο υπάρχων εύρος. Στη περίπτωση μας το  $a=0$ ,  $b=1$ .

Στη μεθοδολογία μας χρησιμοποιούμε τη συγκεκριμένη μέθοδο για contrast stretching τόσο σε RGB εικόνες όσο και σε gray scale εικόνες. Τα αποτελέσματα φαίνονται στην Εικόνα 2.1 και Εικόνα 2.2.



(α)



(β)





(γ)



(δ)

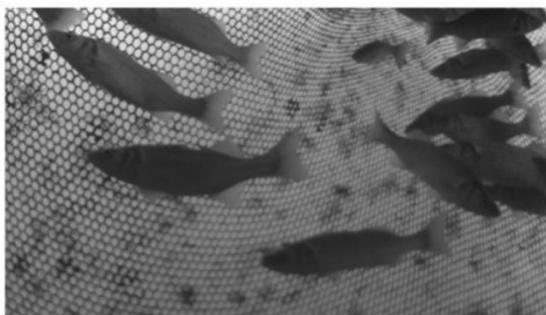


(ε)



(ζ)

**Εικόνα 2.1** (α), (γ), (ε) Αρχικές RGB εικόνες. (β), (δ), (ζ) Αποτελέσματα μετά το contrast stretching.



(α)



(β)

**Εικόνα 2.2** (α) Αρχική gray scale εικόνα. (β) Αποτέλεσμα μετά το contrast stretching.

### 2.2.2 Histogram Equalization

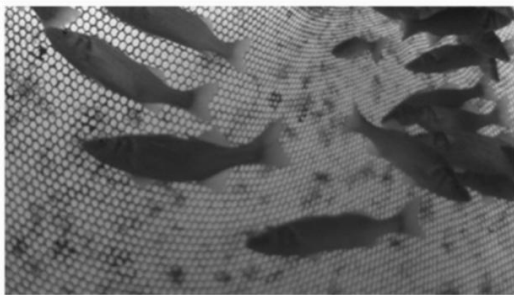
Το Histogram Equalization (εξισορρόπηση ιστογράμματος) είναι μια πολύ διαδεδομένη μέθοδος λόγω της απλότητας και της αποτελεσματικότητάς της σε κάθε είδους εικόνας όπως αναφέρεται και στις εργασίες «*Underwater Image Enhancement Using an Integrated Colour Model*», [8] και «*A dynamic histogram equalization for image contrast enhancement*», [27]. Το ιστόγραμμα αφορά σε στατιστική κατανομή της πιθανότητας για τα επίπεδα του γκρι σε μια ψηφιακή εικόνα. Η μέθοδος αυτή αυξάνει καθολικά το εύρος του ιστογράμματος διανέμοντας καλύτερα τις εντάσεις στην εικόνα. Αυτό επιτρέπει σε περιοχές με χαμηλότερη αντίθεση να αποκτήσουν μεγαλύτερη αντίθεση. Έστω ότι η εικόνα μας  $f(x,y)$  με διακριτές γκρι τιμές με εύρος  $[0,L-1]$ . Η συνάρτηση μετατροπής για κάθε επίπεδο  $C(r_k)$  είναι η εξής :

$$s_k = C(r_k) = \sum_{i=0}^k P(r_i) = \sum_{i=0}^k \frac{n_i}{n} \quad (2.2)$$

Όπου  $k=0,1,2,\dots,L-1$  και  $0 \leq 1$

Με  $n_i$  αναπαρίσταται ο αριθμός των pixel που έχουν grey level  $r_i$ ,  $n$  είναι ο συνολικός αριθμός των pixel στην εικόνα, το  $P(r_i)$  αναπαριστά την PDF για κάθε  $r_i$  και το  $s_k$  μπορεί να αναπαρασταθεί στο εύρος  $[0, L-1]$  αν το πολλαπλασιάσουμε με  $(L-1)$ .

Η εφαρμογή του histogram equalization συνδυαστικά με το contrast stretching δίνει πολύ καλά αποτελέσματα όπως ενδεικτικά φαίνεται στην Εικόνα 2.3.



(α)



(β)

**Εικόνα 2.3** (α) Αρχική gray scale εικόνα. (β) Αποτέλεσμα μετά το histogram equalization.

### 2.2.3 Homomorphic Filtering

Τέλος, η τρίτη μέθοδος που εφαρμόσαμε για την ενίσχυση της εικόνας αφορά στο Homomorphic filtering «*ROV-based underwater vision system for intelligent fish ethology research*», [7] και «*Review of underwater image enhancement techniques*», [26]. Αποτελεί μία μέθοδο που αφορά στη βελτίωση στο πεδίο της συχνότητας. Αυτό που καταφέρνουμε με τη μέθοδο αυτή είναι να εξομαλυνθεί η φωτεινότητα, να αυξηθεί η αντίθεση των χρωμάτων και να ακονιστούν οι ακμές. Μια εικόνα μπορεί να περιγραφεί ως το γινόμενο του φωτός και του τρόπου που τα αντικείμενα το ανακλούν. Στις υποθαλάσσιες εικόνες που μελετάμε το φως είναι σχετικά ομοιόμορφο παντού στην εικόνα καθώς η πηγή φωτός είναι ο ήλιος. Αυτό το οποίο μεταβάλλεται και ξεχωρίζει τα αντικείμενα από το φόντο είναι ο τρόπος που το ανακλούν. Οπότε μεταφέροντας την εικόνα στο πεδίο της συχνότητας οι διακυμάνσεις αυτές μεταφράζονται σε υψηλές συχνότητες. Εφαρμόζοντας υψι-περατό φίλτρο πετυχαίνουμε οι ακμές και οι αντιθέσεις τις εικόνες να γίνουν πιο έντονες. Επίσης η μέθοδος αυτή μπορεί να δώσει καλά αποτελέσματα και στην αφαίρεση του θορύβου που προκαλείται από αυτό. Αρχικά η εικόνα παρουσιάζεται ως παράγωγο της φωτεινότητας ( $i$ ) και της ανάκλασης ( $r$ ).

$$f(x, y) = i(x, y)r(x, y) \quad (2.3)$$

Για να πετύχουμε το διαχωρισμό πρώτα τοποθετούμε το μοντέλο σε φυσικό λογαριθμικό πεδίο και μετά παίρνουμε το μετασχηματισμό Fourier

$$\begin{aligned} g &= \ln f = \ln i + \ln r \\ F\{g(x, y)\} &= F\{\ln i(x, y)\} + F\{\ln r(x, y)\} \\ G(u, v) &= I(u, v) + R(u, v) \end{aligned} \quad (2.4)$$

Τη μεταφέρουμε στο πεδίο της συχνότητας όπου ,εφαρμόζοντας ένα ύψι-περατό φίλτρο, ενισχύονται οι ακμές της εικόνας.

$$\begin{aligned} S(u, v) &= H(u, v)G(u, v) = H(u, v)(I(u, v) + R(u, v)) \\ s(u, v) &= F^{-1}\{S(u, v)\} \\ &= F^{-1}\{H(u, v)I(u, v)\} + F^{-1}\{H(u, v)R(u, v)\} \\ &= i'(x, y) + r'(x, y) \end{aligned} \quad (2.5)$$

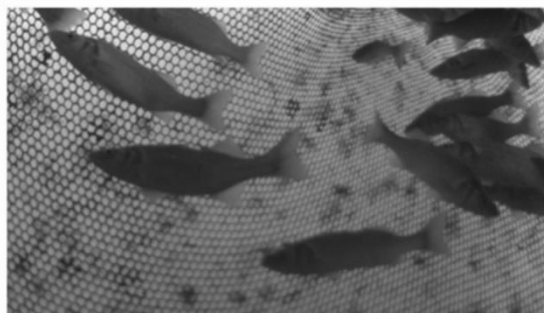
Το  $H(u, v)$  είναι ένα high pass φίλτρο που μειώνει τη συνεισφορά των χαμηλών συχνοτήτων και ενισχύει τη συμβολή των μεσαίων και υψηλών συχνοτήτων 'ακονίζοντας' τις ακμές των ψαριών στην εικόνα.

Για να επιστρέψουμε στο χωρικό πεδίο εφαρμόζουμε αντίστροφο μετασχηματισμό,

$$S'(x, y) = \exp(s(x, y)) = \exp(i'(x, y)) * \exp(r'(x, y)) = i''(x, y) * r''(x, y) \quad (2.6)$$

οπότε τα  $i''$ ,  $r''$  είναι η φωτεινότητα και η αντανάκλαση της νέας εικόνας.

Όταν το homomorphic filtering εφαρμόζεται ως συνέχεια των δυο πρώτων μεθόδων που περιγράφηκαν, δηλαδή το contrast stretching και το histogram equalization, μας δίνει βέλτιστο αποτέλεσμα όπως φαίνεται το τελικό αποτέλεσμα στην Εικόνα 2.4. Επίσης στην Εικόνα 2.5 φαίνεται το αποτέλεσμα της κάθε εικόνας και το αντίστοιχο ιστόγραμμα.



(α)

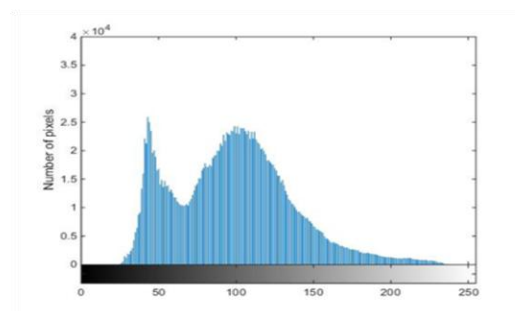


(β)

**Εικόνα 2.4** (α) Αρχική gray scale εικόνα. (β) Αποτέλεσμα μετά το homomorphic filtering.



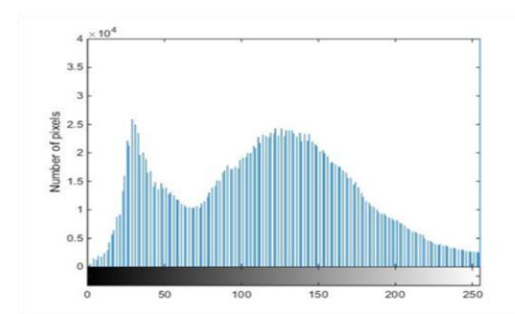
(α)



(β)



(γ)

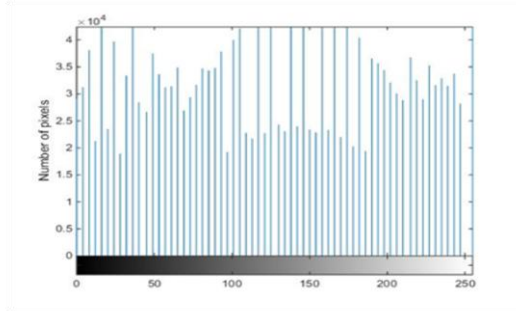


(δ)





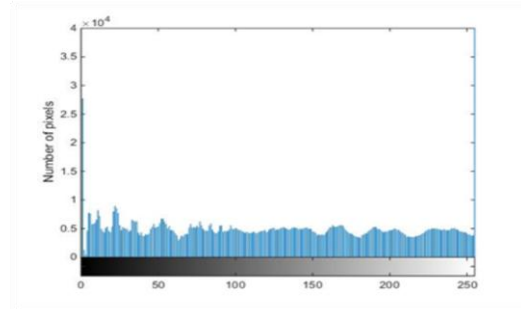
(ε)



(ζ)



(η)



(θ)

**Εικόνα 2.5** (α) Αρχική gray scale εικόνα. (β) Ιστόγραμμα gray scale εικόνας. (γ) Contrast stretching εικόνας, (δ) Ιστόγραμμα contrast stretching εικόνας, (ε) Histogram equalization εικόνας (γ), (ζ) Ιστόγραμμα histogram equalization εικόνας (γ), (η) Homomorphic filtering εικόνας (ζ), (θ) Ιστόγραμμα Homomorphic filtering εικόνας (ζ).

## 2.3 Κατάτμηση σε Υποθαλάσσιες Εικόνες

Όπως αναφέραμε και παραπάνω, η κατάτμηση της εικόνας αφορά στη διαδικασία κατά την οποία η αρχική εικόνα διαιρείται σε όμοιες περιοχές (τμήματα), βάσει κοινών κριτηρίων.

### 2.3.1 Thresholding

Με την βελτιστοποίηση εικόνας που έχει προηγηθεί έχουμε σαν αποτέλεσμα grayscale εικόνες με έντονες αντιθέσεις κι εναλλαγές στη φωτεινότητα. Εκμεταλλευόμενοι αυτά τα δύο στοιχεία, η πρώτη μέθοδος που δοκιμάσαμε για το segmentation, αφορά στην χρήση ορίου (thresholding) με βάση τη διαβάθμιση του γκρι. Γνωρίζοντας ότι μια grayscale εικόνα αναπαριστά τη φωτεινότητα της εικόνας, θεωρήσαμε ότι ένα κατώφλι στη φωτεινότητά της μπορεί να διαχωρίσει τα αντικείμενα από το φόντο που τα περιβάλλει. Η μέθοδος αυτή είναι απλή, γρήγορη και είναι η παλαιότερη μέθοδος κατάτμησης εικόνας που χρησιμοποιήθηκε. Έστω ότι με  $f$  αναπαριστάται η φωτεινότητα για κάθε pixel και  $T$  είναι το όριο, τότε ισχύει ότι:

- $f(i, j) \geq T$  τότε ανήκει στα αντικείμενα
- Αλλιώς ανήκει στο φόντο

Παρατηρώντας τα αποτελέσματα είδαμε ότι η τιμή έντασης της φωτεινότητας που αντιστοιχεί στα ψάρια είναι από 100 και πάνω. Οπότε το όριο στην περίπτωση μας επιλέχθηκε να είναι  $T=0.4$ . Πειραματικά αποδείχτηκε σωστή για τις περισσότερες περιπτώσεις που εξετάσαμε, συγκριτικά με άλλα thresholds. Αποτελέσματα της μεθόδου φαίνονται στην Εικόνα 2.6.



(α)



(β)



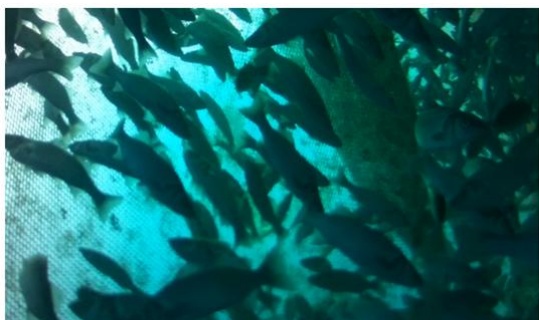
(γ)



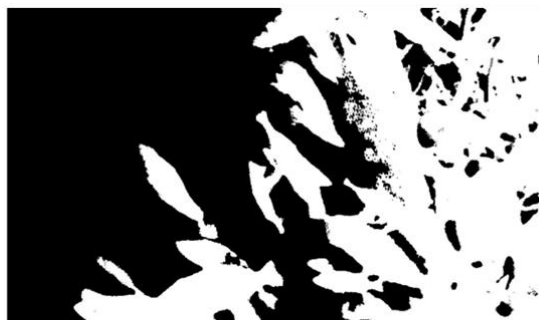
(δ)

**Εικόνα 2.6** (α) Αρχική εικόνα frame=112. (β) Αποτέλεσμα μετά το enhancement. (γ) Αποτέλεσμα segmentation με thresholding. (δ) Τελικό αποτέλεσμα.

Παρόλα αυτά, τα αποτελέσματα δεν είναι ικανά να εντοπίσουν τα ψάρια σε κάθε περίπτωση, όπως επίσης σε βίντεο με διαφορετικές συνθήκες όπως χαρακτηριστικά φαίνεται στην Εικόνα 2.7, οπότε προχωρήσαμε σε μια πιο περίπλοκη μέθοδο κατάτμησης βασισμένη στα active contours.



(α)



(β)



(γ)



(δ)

**Εικόνα 2.7** (α) Αρχική εικόνα frame=242. (β) Αποτέλεσμα threshold. (γ) Αρχική εικόνα frame=223. (δ) Αποτέλεσμα threshold.

### 2.3.2 Active Contours

Τα active contours (ενεργά περιγράμματα) χρησιμοποιούνται για να χωρίσουν την εικόνα σε δυο περιοχές, όπου η μία αφορά στα αντικείμενα και η άλλη στο φόντο. Η μέθοδος αυτή είναι κατάλληλη για να διαχωρίσει αντικείμενα που είναι ομογενή και διαχωρίσιμα από το background. Η μέθοδος που επιλέξαμε βασίζεται στην τεχνική της εξέλιξης καμπύλης, στην Mumford και Shah συνάρτηση για την κατάτμηση και στη μέθοδο των level sets. Βασική ιδέα του αλγορίθμου είναι ότι δοσμένης μιας συνάρτησης ενέργειας η οποία έχει προκύψει από ένα minimal partition problem, δηλαδή κατάτμηση εικόνας σε ομοιόμορφες περιοχές βάση Mumford Shah κριτηρίου, προσπαθούμε να την ελαχιστοποιήσουμε. Η καμπύλη μας αναπαρίσταται και εξελίσσεται βάση mean- curvature flow με χρήση των level sets και σταματά στα επιθυμητά όρια. Τα όρια αυτά δε βασίζονται στο gradient της εικόνας όπως σε κλασσικές μεθόδους active contour αλλά στην κατάτμηση που έχει ήδη γίνει. Η μέθοδος αυτή προτάθηκε και εξελίχθηκε από τους Chan Vese στις εργασίες τους «An active contour model without edges», [28] και εξελίχθηκε από τους ίδιους στην εργασία «A multiphase level set framework for image segmentation using the Mumford and Shah model» [29] για το λόγο αυτό ο αλγόριθμος πήρε και την ονομασία τους. Την χρήση του αλγορίθμου την συναντάμε επίσης και σε συνθήκες υποθαλάσσιας εικόνας για εντοπισμό ψαριών όπως στην εργασία «ROV-based underwater vision system for intelligent

*fish ethology research*», [7] και στην εργασία «*Automatic Analysis of Fish Behaviors and Abnormality Detection*», [24].

Ξεκινώντας λοιπόν από μια αρχική καμπύλη στην εικόνα μας, σκοπός είναι η καμπύλη αυτή να περικυκλώσει τα αντικείμενα σε έναν αριθμό επαναλήψεων έως ότου δεν μπορεί να εξελιχθεί παραπάνω. Αυτό επιτυγχάνεται μέσω της ελαχιστοποίησης μιας κατάλληλα σχεδιασμένης συνάρτησης ενέργειας. Τα αντικείμενα στην εικόνα ύστερα από την προ επεξεργασία είναι smoothed περιοχές.

Έστω ότι η αρχική εικόνα μας ανήκει στο χώρο  $\Omega$ . Το πρόβλημα του segmentation όπως εκφράστηκε από τους Mumford Shah στην εργασία «*Optimal approximations by piecewise smooth functions and associated variational problems*», [30] αφορά στον τρόπο με τον οποίο θα μπορέσουμε να χωρίσουμε το  $\Omega$  σε υποσύνολα ώστε να μας δώσει ένα βέλτιστο piecewise smoothed αποτέλεσμα εικόνας  $u$  σε σχέση με την αρχική. Έτσι στα υποσύνολα της  $u$  η διαφορά είναι ομαλή και μεταξύ των ορίων η διαφορά είναι έντονη ή υπάρχει ασυνέχεια.

Για να δωθεί λύση σε αυτό το πρόβλημα στην παραπάνω εργασία, [30] προτείνεται η ακόλουθη συνάρτηση ελαχιστοποίησης:

$$F(u, C) = \int_{\Omega} (u - u_0)^2 dx dy + \mu \int_{\Omega/C} |\nabla u|^2 dx dy + \nu/C, \quad (2.7)$$

όπου  $\mu, \nu > 0$  σταθερές παράμετροι για να δώσουν διαφορετικά βάρη στην ενέργεια της εικόνας και  $/C/$  είναι το μήκος της καμπύλης μας.

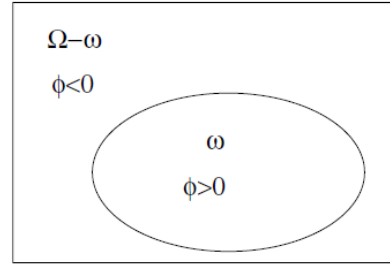
Ελαχιστοποιώντας την παραπάνω συνάρτηση έχουμε σαν αποτέλεσμα την εικόνα  $u=c_i$  όπου  $c_i$  είναι μια σταθερά που δηλώνει τη μέση τιμή για κάθε περιοχή  $\Omega_i$  της  $u_0$ . Αυτό καλείται minimal partition problem. Στη περίπτωση μας εφαρμόζεται μια ακόμα πιο απλή μορφή η οποία αφορά την τμηματοποίηση σε δυο περιοχές, εντός και εκτός της καμπύλης.

Αρα έχουμε :

$$u = \begin{cases} \text{average}(u_0), & \text{εντός της } C \\ \text{average}(u_0), & \text{εκτός της } C \end{cases} \quad (2.8)$$

Για την αναπαράσταση των καμπυλών και την επίλυση της συνάρτησης στην εργασία «*Fronts propagating with curvature-dependent speed: Algorithms based on Hamilton-Jacobi formulation*», [31] μέθοδο levels sets. Η καμπύλη  $C$  βρίσκεται σε περιοχή  $\omega \in \Omega$  με  $C=\partial\omega$  και αναπαριστάται ως μηδενικό level set μέσω μιας Lipschitz συνεχούς συνάρτησης  $\phi: \Omega \rightarrow \mathbb{R}$  ώστε:

- $\phi(x, y) > 0$  ανήκει στο  $\omega$ ,
- $\phi(x, y) < 0$  ανήκει στο  $\Omega \setminus \omega$
- $\phi(x, y) = 0$  ανήκει στη καμπύλη  $C$



(2.9)

Με αυτό τον τρόπο μπορούν να προσδιοριστούν γεωμετρικές ποσότητες, καθώς και ιδιότητες της  $C$ . Το μήκος  $|C|$  και το εμβαδόν  $|\omega|$  της καμπύλης της προσδιορίζονται μέσω της γνωστής συνάρτησης Heavicide,  $H$  και της μονοδιάστατης συνάρτησης Dirac  $\delta_0$ , που ορίζονται ως εξής:

$$H(z) = \begin{cases} 1, & \text{για } z \geq 0 \\ 0, & \text{για } z < 0 \end{cases}, \quad \delta_0(z) = \frac{d}{dx} H(x) \quad (2.10)$$

Το μήκος  $|C|$  και το εμβαδόν  $|\omega|$  ορίζονται ως εξής:

$$|C| = \int_{\Omega} |\nabla H(\phi)| = \int_{\Omega} \delta_0(\phi) |\nabla \phi|, \quad |\omega| = \int_{\Omega} H(\phi) dx dy \quad (2.11)$$

Οπότε έστω  $c_1, c_2$  δυο άγνωστες σταθερές που αναφέρονται στη μέση τιμή εσωτερικά και εξωτερικά της καμπύλης αντίστοιχα. Δηλώνοντας επίσης με  $\Omega_1 = \omega$  και  $\Omega_2 = \Omega \setminus \omega$ , προτίνεται η ελαχιστοποίηση της ακόλουθης συνάρτησης ενέργειας συνδιαστικά με την αναπαράσταση level sets σε σχέση με τα  $c_1, c_2$  και  $C = \{(x, y) | \phi = 0\}$

$$\begin{aligned} F_2(c_1, c_2, \phi) = & \int_{\Omega} (u_0(x, y) - c_1)^2 H(\phi) dx dy \\ & + \int_{\Omega} (u_0(x, y) - c_2)^2 (1 - H(\phi)) dx dy \\ & + v \int_{\Omega} |\nabla H(\phi)| \end{aligned} \quad (2.12)$$

Με δεδομένο ότι  $H_\varepsilon$  και  $\delta_\varepsilon$  είναι κάθε  $C^1$  προσεγγιστική και κανονικοποιημένη μορφή της Heaviside συνάρτησης  $H$  και της μονοδιάστατης Dirac συνάρτησης  $\delta_0$  αντίστοιχα, καθώς το



$\varepsilon$  τείνει στο 0 και με  $H'_\varepsilon = \delta_\varepsilon$  και ελαχιστοποιώντας την ενέργεια  $F(\Phi, c_1, c_2)$  έχουμε  $\Phi(0, x, y) = \Phi_0(x, y)$ :

$$c_1(\varphi) = \frac{\int_{\Omega} u_0(x, y) H_\varepsilon(\varphi(t, x, y)) dx dy}{\int_{\Omega} H_\varepsilon(\varphi(t, x, y)) dx dy} \quad (2.13)$$

$$c_2(\varphi) = \frac{\int_{\Omega} u_0(x, y) (1 - H_\varepsilon(\varphi(t, x, y))) dx dy}{\int_{\Omega} (1 - H_\varepsilon(\varphi(t, x, y))) dx dy} \quad (2.14)$$

Κρατώντας σταθερά τα  $c_1, c_2$ , ελαχιστοποιούμε την ενέργεια  $F(\Phi, c_1, c_2)$  με βάση το  $\Phi$  και ορίζουμε την Euler-Lagrange συνάρτηση για το  $\Phi$ :

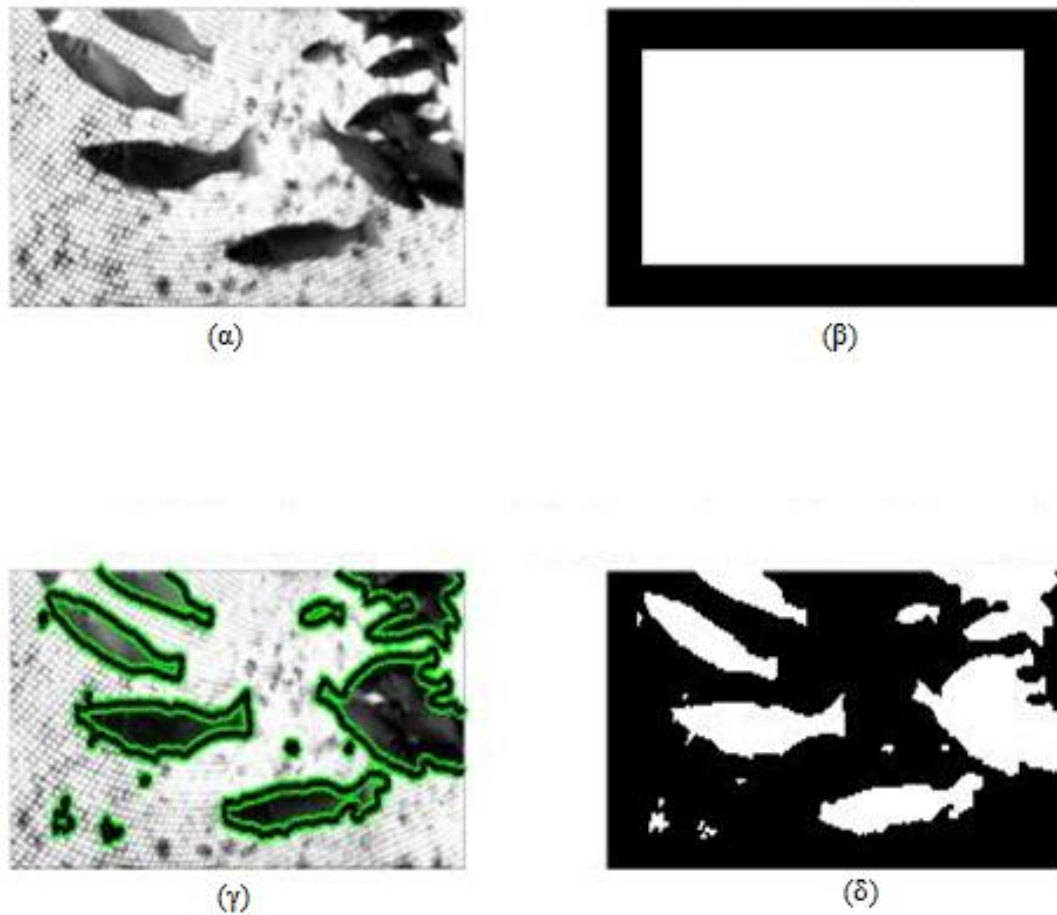
$$\frac{\partial \varphi}{\partial t} = \delta_\varepsilon(\varphi) \left[ v \operatorname{div} \left( \frac{\nabla \varphi}{|\nabla \varphi|} \right) - (u_0 - c_1)^2 + (u_0 - c_2)^2 \right] \quad (2.15)$$

Ο αλγόριθμος τερματίζει όταν η καμπύλη που περιγράφεται από τα  $c_1, c_2$  και αφορά στη φωτεινότητα μέσα και έξω από αυτήν παραμένει αμετάβλητη σε διαδοχικές επαναλήψεις.

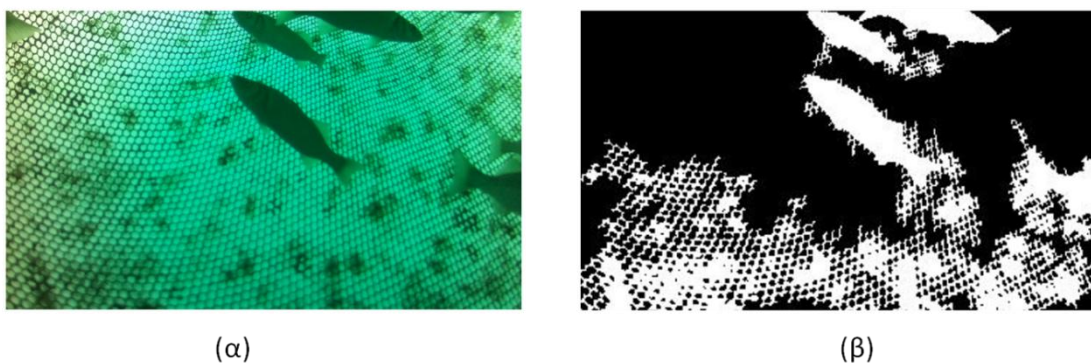
Στα θετικά της μεθόδου είναι ότι έχει πολύ καλά αποτελέσματα στις εναλλαγές φωτεινότητας. Επίσης χωρίζει την εικόνα σε δυο περιοχές οπότε είναι εύκολο να την κάνουμε binary και τέλος δεν χρειάζεται κάποια προεργασία για αφαίρεση θορύβου καθώς έχει πολύ καλά αποτελέσματα σε εικόνες ακόμα και με πολύ έντονο θόρυβο καθώς στις επαναλήψεις η εικόνα μας γίνεται πιο smoothed. Βάση της προεργασίας που έχει γίνει στο enhancement, είναι εύκολο με ένα όριο να εντοπίσουμε ποια είναι τα ψάρια και ποιο το φόντο. Στα αρνητικά μπορούμε να πούμε ότι έχει σημασία η αρχικοποίηση της καμπύλης να είναι σε σημείο που πιάνει και ψάρια και φόντο αλλιώς μια καμπύλη που ξεκινά σε περιοχή μέσα στο ψάρι θα σταματήσει στα όρια αυτού. Επίσης ο αλγόριθμος έχει μεγάλη υπολογιστική ισχύ και είναι πολύ ευαίσθητος στον εντοπισμό αντικειμένων που δεν θέλουμε να τα εντοπίσουμε (δίχτυα, φυτά). Στην Εικόνα 2.9 φαίνεται αποτέλεσμα του αλγορίθμου χωρίς πρότερη επεξεργασία της εικόνας με σκοπό να δούμε τι προβλήματα έχει γενικά.

Οι δύο παράμετροι που κληθήκαμε να ορίσουμε αφορούν στο είδος της καμπύλης και στον αριθμό των επαναλήψεων. Για το είδος της καμπύλης αρχικοποιήσαμε ένα τετράγωνο που απέχει 100 pixels από κάθε πλευρά της εικόνας. Είναι μια αρκετά μεγάλη περιοχή στο εύρος της εικόνας που συνδυαστικά με τον αριθμό των επαναλήψεων εντοπίζει περιοχές και μέσα και έξω από αυτό. Αναφορικά με τον αριθμό των επαναλήψεων όπου επιλέξαμε τις 1500 βάση πειραμάτων για να καλύψουμε κάθε ενδεχόμενο ώστε να εξαχθεί μια εικόνα με ικανοποιητικό αποτέλεσμα συνδυάζοντας όσο το δυνατόν μικρότερη

πολυπλοκότητα και χρόνο εκτέλεσης. Κατά μέσο όρο θα μπορούσαμε να πούμε ότι στις δικές μας εικόνες 1000-1200 επαναλήψεις ήταν αρκετές. Όταν η καμπύλη σταματά να εξελίσσεται ο αλγόριθμος τερματίζει. Ενδεικτικό παράδειγμα φαίνεται στην Εικόνα 2.8.



**Εικόνα 2.8** (α) Αρχική εικόνα μετά το enhancement. (β) Αρχικοποιούμε μια καμπύλη τετράγωνο που απέχει 100 pixels από κάθε πλευρά της εικόνας μας. (γ) Με πράσινο φαίνεται το τελικό αποτέλεσμα της καμπύλης μας. (δ) Τελικό αποτέλεσμα segmentation, binary εικόνα.



**Εικόνα 2.9** Εσφαλμένα αποτελέσματα active contours. (α)αρχικό frame(β) τελικό frame

### 2.3.3 Απομάκρυνση ανεπιθύμητων περιοχών κατά το segmentation

Τα αποτελέσματα του segmentation δεν είναι τέτοια που να επαρκούν για την παρούσα εφαρμογή. Πολλές εικόνες κατατέμνονται με τέτοιο τρόπο ώστε να παραμένουν μέρη διχτυών ή σκουπίδια που υπάρχουν στην εικόνα μας. Αυτά αποτελούν πληροφορίες που πρέπει να απομακρυνθούν καθώς δημιουργούν πρόβλημα στον εντοπισμό των ψαριών. Μάλιστα αποτελεί ένα από τα πιθανά προβλήματα που θα χρειαστεί να επιλύσουμε σε βίντεο ιχθυοκαλλιεργειών. Οπότε, για να αντιμετωπίσουμε τέτοια προβλήματα στην εικόνα μας, εφαρμόσαμε δυο απλές τεχνικές βασισμένες στο blob analysis\*. Η πρώτη αφορά στην αφαίρεση πολύ μικρών αντικειμένων. Εφόσον έχουμε σαν αποτέλεσμα μια binary εικόνα, με blob analysis απομακρύνουμε πολύ μικρά αντικείμενα. Αυτό είχε σαν αποτέλεσμα να παραμείνουν στην εικόνα μας μόνο τα ψάρια όπως φαίνεται στην Εικόνα 2.6 (δ).

- Σε κάποιες περιπτώσεις, το παραπάνω δεν ήταν αρκετό, οπότε εκμεταλλευτήκαμε τον τρόπο που εμφανίζονται τα δίχτυα στην binary εικόνα μας. Παρατηρούμε ότι τα ψάρια εμφανίζονται σαν ομογενείς περιοχές ενώ αντιθέτως τα δίχτυα έχουν πολλές οπές. Επαναλαμβάνοντας το blob analysis διώξαμε τέτοιες περιοχές όπως φαίνεται στην Εικόνα 2.10. Βέβαια αυτό είχε σαν αποτέλεσμα σε κάποιες εικόνες να διώξουμε και ψάρια, ειδικά σε περιπτώσεις μερικών επικαλύψεων που σχηματίζονται οπές μεταξύ τους όπως στην Εικόνα 2.11.



(α)



(β)





(γ)

**Εικόνα 2.10** (α) Αρχική εικόνα. (β) Binary αποτέλεσμα εικόνας, με τους κόκκινους κύκλους δείχνουμε τα δίχτυα που χάνονται από την επεξεργασία. (γ) Τελικό αποτέλεσμα.



(α)



(β)



(γ)

**Εικόνα 2.11** (α) Αρχική εικόνα. (β) Binary αποτέλεσμα εικόνας, με τους κόκκινους κύκλους δείχνουμε τα ψάρια που χάνονται από την επεξεργασία. (γ) Τελικό αποτέλεσμα.

Αξίζει να σημειωθεί ότι τα frames που χρησιμοποιήθηκαν για βελτιστοποίηση και κατάτμηση εικόνας είναι από βίντεο πραγματικού ιχθυοτροφείου. Έγινε σε αυτά μία πρώτη μελέτη των συνθηκών, η οποία παρουσιάστηκε στο τρέχον κεφάλαιο, αλλά γενικώς το συγκεκριμένο βίντεο κρίθηκε ακατάλληλο για τους σκοπούς της παρούσας διπλωματικής κι επομένως για τα επόμενα βήματα της δουλειάς μας καθώς παρουσιάζει ιδιαίτερα δύσκολες συνθήκες, χωρίς πρότερο σχεδιασμό λήψης με κύριο στόχο την εξαγωγή τροχιάς των ψαριών. Συγκεκριμένα τραβήχτηκαν από δύτη με έντονη κίνηση της κάμερας και καμία γνώση για τη θέση της κάμερας σε σχέση με τα ψάρια. Επίσης το frame rate είναι πολύ μικρό, 6 frames ανά δευτερόλεπτο, με αποτέλεσμα η μετακίνηση των στόχων στην εικόνα να είναι πολύ μεγάλη από frame σε frame κάτι το οποίο προκαλεί πρόβλημα ειδικότερα σε περιπτώσεις επικάλυψης ψαριών. Τέλος, σε τέτοιου είδους βίντεο, οι υποθαλάσσιες εικόνες έχουν πολύ μεγάλο αριθμό από ψάρια και πρόκειται για βίντεο που είναι τραβηγμένα με τέτοιο τρόπο ώστε να υπάρχει δυσκολία να διαχωριστούν ακόμα και με το μάτι σε κάποιες περιπτώσεις. Σημαντικό για την έρευνα σε περιπτώσεις ιχθυοκαλλιέργειας είναι να εξαχθούν δεδομένα με τέτοιο τρόπο που θα μπορούν να δώσουν διακριτά αποτελέσματα. Για το λόγο αυτό, τα βίντεο που χρησιμοποιήθηκαν για τον αλγόριθμό μας, βρέθηκαν στο διαδίκτυο κατόπιν αναζήτησης που έγινε με κύριο γνώμονα την κάλυψη όλων των ενδεχόμενων περιπτώσεων συνθηκών και, παράλληλα, ο αριθμός από ψάρια να ποικίλει ώστε τελικά να πλησιάσουμε όσο γίνεται συνθήκες ιχθυοκαλλιέργειας.

\*Εκτενείς αναφορά στο blob analysis γίνεται στο Κεφάλαιο 3, (3.2.3 Blob Analysis)

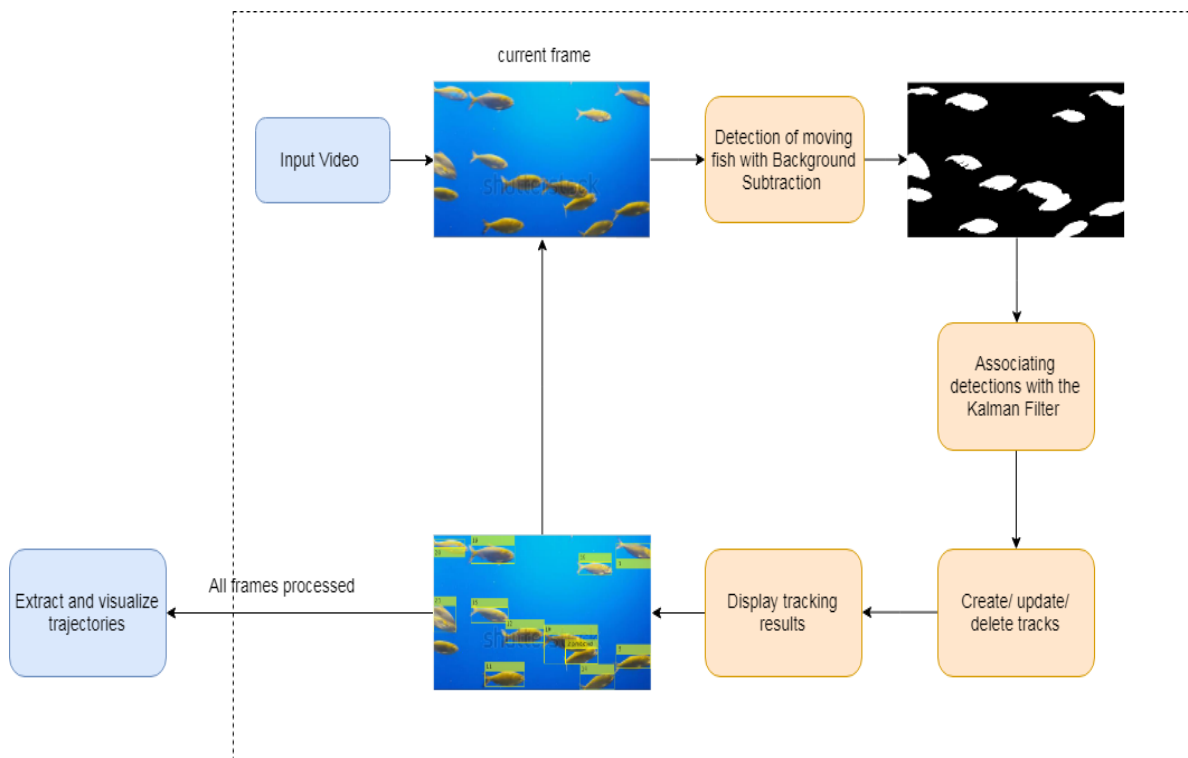
## Κεφάλαιο 3: Προτεινόμενη Μεθοδολογία

### 3.1 Εισαγωγή

Στο κεφάλαιο αυτό αναλύεται η μεθοδολογία που προτείνουμε για την παρακολούθηση ψαριών σε υποθαλάσσια βίντεο. Κύριος σκοπός της εργασίας μας είναι η εξαγωγή αποτελεσμάτων μέσω της μεθοδολογίας μας ώστε να είναι δυνατή η κατανόηση της συμπεριφοράς των ψαριών. Τα δύο βασικά στάδια της μεθοδολογίας είναι η ανίχνευση (detection) των ψαριών και η παρακολούθηση της κίνησης τους (tracking). Έπειτα ιδιαίτερο βάρος δίνεται στην εξαγωγή και επεξεργασία των τροχιών. Τα κύρια βήματα της υλοποίησης μας φαίνονται στο σχηματικό διάγραμμα στην Εικόνα 3.1 και θα αναλυθούν παρακάτω.

Έχοντας σαν βάση την έρευνα που κάναμε όσο αφορά τη συμπεριφορά των ψαριών η οποία περιλαμβάνει γενικότερη αναζήτηση στο διαδίκτυο, άρθρα σε σχέση με την συμπεριφορά των ψαριών από την Ελληνική Κτηνιατρική Εταιρία, [2] και το βιβλίο «*Some Aspects of the Schooling Behaviour of Fish*», [32], θεωρήσαμε ότι ένας τρόπος για καλύτερη εξαγωγή συμπερασμάτων θα ήταν η ομαδοποίηση και μελέτη της δουλειάς μας να αναλυθεί σε δύο κατευθύνσεις. Η πρώτη αφορά τη συμπεριφορά κοπαδιού ψαριών και η δεύτερη τη μεμονωμένη συμπεριφορά των ψαριών. Οι δυο αυτές πλευρές περιγράφονται και στο βιβλίο «*Some Aspects of the Schooling Behaviour of Fish*», [32]. Συμπεριφορά κοπαδιού σημαίνει ότι τα ψάρια κινούνται συντονισμένα και συγχρονισμένα, έτσι ώστε να φαίνεται το κοπάδι σαν ενιαίος οργανισμός που ακολουθεί τη δική του λογική κινήσεων, χαρακτηριστικά παραδείγματα φαίνονται στην Εικόνα 3.2. Τα ψάρια ειδικότερα όταν βρίσκονται στο φυσικό τους περιβάλλον, στη θάλασσα, κινούνται σε κοπάδια με σκοπό την προστασία από την επίθεση θηρευτών. Η συμπεριφορά αυτή, εφόσον πρόκειται για ψάρια του ίδιου είδους, παρατηρείται και σε συνθήκες ιχθυοτροφείου. Οπότε η μελέτη της συμπεριφοράς του κοπαδιού είναι σημαντική και ειδικότερα αν παρατηρηθεί ότι κάποιο ή κάποια από αυτά αποκλίνουν πολύ από το κοπάδι. Όσο αφορά τη δεύτερη πλευρά, δηλαδή τη μελέτη της συμπεριφοράς του κάθε ψαριού ξεχωριστά, σκοπός είναι να εξαχθούν οι τροχιές σε καθένα από αυτά και να δούμε τις διαφοροποιήσεις που μπορεί να υπάρχουν. Στην περίπτωση αυτή έχουμε να κάνουμε με μικρό ή μεγάλο αριθμό από ψάρια με κύριο πρόβλημα τις επικαλύψεις μεταξύ τους. Στην Εικόνα 3.3 φαίνονται κάποια χαρακτηριστικά frames που αφορούν αυτή τη περίπτωση.

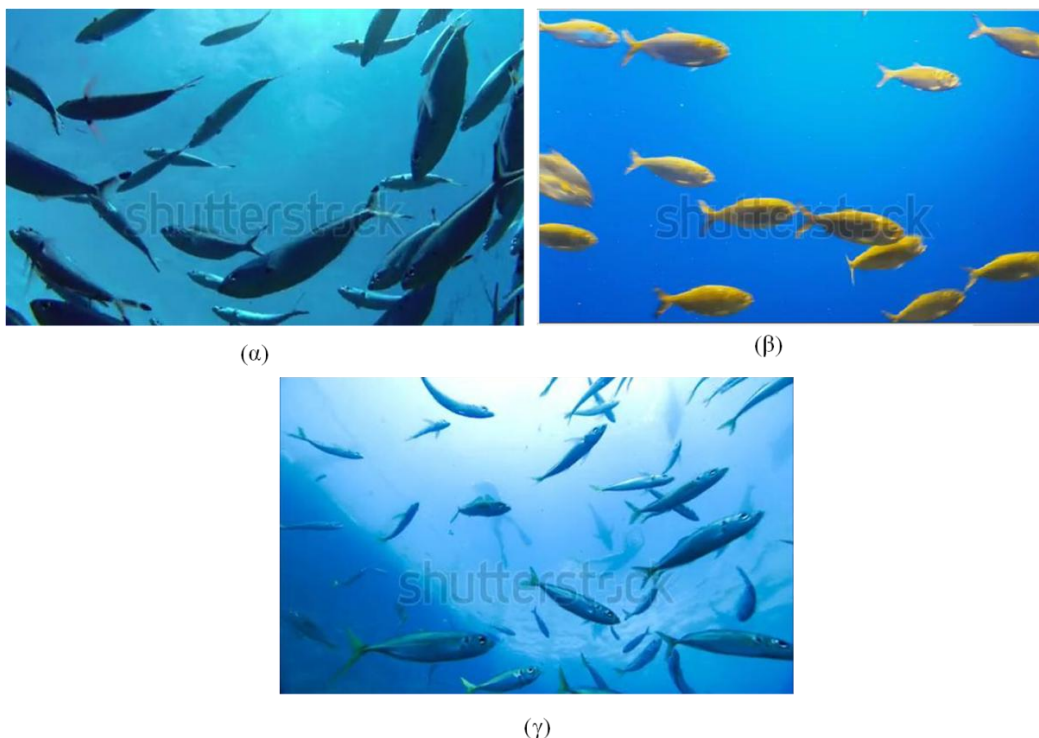
Η μεθοδολογία μας εφαρμόζεται και στις δυο περιπτώσεις με τον ίδιο τρόπο αλλάζοντας τις παραμέτρους του αλγορίθμου μας. Για την κατάτμηση και ανίχνευση των ψαριών εστιάσαμε στη κίνηση τους επιλέγοντας την αφαίρεση φόντου (Background Subtraction) βασισμένο σε Μίγμα Γκαουσιανών Μοντέλων (Gaussian Mixture Models) και κατάλληλη παραμετροποίηση στο μορφολογικό φιλτράρισμα (Morphological Operations) και το Blob Analysis. Για την παρακολούθηση και την εξαγωγή τροχιάς επιλέξαμε μεθοδολογία βασισμένη στον αλγόριθμο πρόβλεψης Kalman filter.



**Εικόνα 3.1** Σχηματικό διάγραμμα της μεθοδολογίας μας για detection και tracking.



**Εικόνα 3.2** Παραδείγματα από κοπάδια ψαριών.



**Εικόνα 3.3** Frames από τα βίντεο που επεξεργαστήκαμε για την ατομική συμπεριφορά των ψαριών: (α) frame από Video\_1. (β) frame από Video\_2. (γ) frame από Video\_3

### 3.2 Ανίχνευση Ψαριών

Στη μηχανική όραση η ανίχνευση αντικειμένων (object detection) αναφέρεται στη διαδικασία εξαγωγής αντικειμένων από την εικόνα και συνήθως αναφέρεται σε επιλεγμένες περιοχές ενδιαφέροντος για την εκάστοτε εφαρμογή. Υπάρχουν διάφορες μέθοδοι ανίχνευσης και το κριτήριο επιλογής της σωστής εξαρτάται από τα δεδομένα και το σκοπό που έχει η κάθε εφαρμογή. Στη δική μας περίπτωση ως αντικείμενα θεωρούνται τα ψάρια και θέλουμε από τη διαδικασία αυτή να μπορούμε σε κάθε frame να ανιχνεύουμε τα ψάρια είτε ως μεμονωμένες οντότητες είτε ως μάζα. Σε μια υποθαλάσσια εικόνα με μεγάλο αριθμό από ψάρια τα προβλήματα που συνήθως εντοπίζουμε και αποτελούν κριτήριο για την επιλογή της μεθοδολογίας μας, όπως είχαμε αναφέρει και στο 2<sup>ο</sup> Κεφάλαιο, (Εικόνα 3.4), αφορούν σε:

- **Εναλλαγές στη φωτεινότητα.** Κατά τη διάρκεια της μέρας υπάρχουν αλλαγές στη φωτεινότητα που είτε γίνονται σταδιακά και αφορούν τη θέση του ήλιου ή πιο ξαφνικά, παραδείγματος χάρη λόγω κάποιου σύννεφου που μπορεί να καλύψει για λίγο τον ήλιο.

- **Προβλήματα περιβάλλοντος.** Συνήθως το φόντο στις υποθαλάσσιες εικόνες δεν είναι απόλυτα σταθερό. Επαναλαμβανόμενοι κυματισμοί λόγω καιρικών συνθηκών, μικρά σκουπίδια που μπορεί να υπάρχουν στο περιβάλλον, η ποιότητα του νερού που αλλού μπορεί να είναι πιο θολό αλλού πιο διαυγές, είναι ζητήματα που η μεθοδολογία μας καλείται να αντιμετωπίσει για να συμπεριληφθούν στο φόντο.



- **Καμουφλάζ.** Τα ψάρια μπορεί να έχουν παρόμοια χρώματα με το περιβάλλον οπότε είναι δύσκολο να τα εντοπίσεις ακόμα και με το μάτι.
- **Θόρυβος.** Τα βίντεο μπορεί να έχουν θόρυβο που επηρεάζει τα αποτελέσματά μας.



(α)



(β)



(γ)



(δ)

**Εικόνα 3.4** Παραδείγματα από εικόνες σε δύσκολες συνθήκες: (α) Καμουφλαρισμένα ψάρια. (β) Αντανάκλαση του ήλιου πάνω στα ψάρια. (γ) Κυματισμοί και εναλλαγές στη φωτεινότητα λόγω του ήλιου. (δ) Μεγάλος αριθμός από ψάρια, δεν έχει διαφορά με το background.

Για την αντιμετώπιση αυτών των προβλημάτων έγινε μια προσπάθεια επίλυσης με τις μεθόδους enhancement και segmentation που παρουσιάζονται στο 2<sup>ο</sup> Κεφάλαιο της διπλωματικής. Τα αποτελέσματα ήταν ικανοποιητικά για κάποιες εικόνες παρόλα αυτά σε ακολουθία από πολλά frames και σε δοκιμές σε διαφορετικά βίντεο με διαφορετικές συνθήκες δεν μπορούσαν να δώσουν αποτελέσματα ικανά ώστε να υλοποιηθεί το επόμενο στάδιο του tracking. Σε αυτό παίζει σημαντικό ρόλο η μεγάλη εναλλαγή φωτισμών κατά τη διάρκεια ενός βίντεο και οι διαφορές στο περιβάλλον σε διαφορετικά βίντεο. Επίσης η πολυπλοκότητα και ο χρόνος εκτέλεσης της μεθοδολογίας active contour το καθιστά δύσκολο να δώσει αποτελέσματα και σε real time εφαρμογές που όπως είπαμε, είναι κάτι που δεν είναι μέρος της συγκεκριμένης εργασίας αλλά είναι κάτι που μας αφορά για ενδεχόμενη εξέλιξη δουλειάς μας. Για τους λόγους αυτούς δοκιμάσαμε και τελικά επιλέξαμε έναν αλγόριθμο που βασίζεται αποκλειστικά στην κίνηση κινούμενων αντικειμένων. Ως γνωστόν τα ψάρια βρίσκονται σε συνεχή κίνηση καθώς ξεκουράζονται ή βρίσκονται σε ακινησία σπάνια, οπότε είναι ένα γνώρισμα που εκμεταλλευόμαστε για την επιλογή της

μεθόδου μας. Η μέθοδος που επιλέγουμε για το διαχωρισμό των ψαριών από το φόντο είναι το Background Subtraction (Αφαίρεση Φόντου) βασισμένο σε Gaussian Mixture Models (Μίγμα Γκαουσιανών Μοντέλων) που θα αναλυθεί στην επόμενη υποενότητα. Ακολουθεί επεξεργασία των αποτελεσμάτων με μορφολογικούς τελεστές και τέλος, με το Blob Analysis, γίνεται η τελική ανίχνευση των ψαριών. Παρακάτω εξηγούνται συνοπτικά τα βήματα που ακολουθούμε.

### Βήματα αλγορίθμου για το Detection

#### Βήμα 1<sup>ο</sup>

Εισαγωγή βίντεο.

#### Βήμα 2<sup>ο</sup>

Εκπαίδευση GMM για 30 frames και απόκτηση εικόνας φόντου.

#### Βήμα 3<sup>ο</sup>

Εισαγωγή τρέχουσας εικόνας και αφαίρεσης της από το background, δίνοντας μας ένα binary αποτέλεσμα.

#### Βήμα 4<sup>ο</sup>

Εφαρμογή morphological operations βάση της προτεινόμενης ομαδοποίησης.

#### Βήμα 5<sup>ο</sup>

Εφαρμογή blob analysis βάση της προτεινόμενης ομαδοποίησης.

#### Βήμα 6<sup>ο</sup>

Τελικά αποτελέσματα detection. Οι τιμές '0' αναφέρονται στο background και οι τιμές '1' αναφέρονται στο foreground δηλαδή τα εντοπισμένα ψάρια.

### 3.2.1 Αφαίρεση φόντου με Gaussian Mixture Models

Η αφαίρεση φόντου (Background Subtraction) είναι μια ευρέως χρησιμοποιούμενη προσέγγιση για την ανίχνευση κινούμενων αντικειμένων σε βίντεο από στατικές κάμερες όπως γίνεται αναφορά και στις δημοσιεύσεις [33,34,35]. Το σκεπτικό της προσέγγισης είναι ότι ανιχνεύει την κίνηση των αντικειμένων από τη διαφορά μεταξύ του τρέχοντος frame και ενός frame αναφοράς, που αποκαλείται εικόνα φόντου ή αλλιώς μοντέλο φόντου. Η εικόνα του φόντου πρέπει να είναι μια αναπαράσταση της σκηνής χωρίς τα κινούμενα ψάρια και πρέπει επίσης να είναι συνεχώς ενημερωμένη, έτσι ώστε να προσαρμόζεται στις μεταβαλλόμενες συνθήκες που επικρατούν σ' ένα εξωτερικό περιβάλλον. Περισσότερο σύνθετα μοντέλα, όπως είναι οι στατιστικές μέθοδοι, έχουν διευρύνει την τεχνική της αφαίρεσης φόντου πέρα από την κυριολεκτική της έννοια. Ένας background αλγόριθμος δέχεται σαν είσοδο ένα βίντεο και επιστρέφει κάθε εικόνα με δυαδικές τιμές όπου το 1 αντιστοιχεί στο foreground και το 0 στο background. Η διαδικασία αυτή μπορεί να εφαρμοστεί και σε real time εφαρμογές καθώς γίνεται επεξεργασία και έχουμε αποτελέσματα σε κάθε frame.

Για τη μοντελοποίηση του φόντου επιλέγουμε τη μέθοδο Gaussian Mixture Models. Για τέτοιων ειδών εργασίες έχουν αναπτυχθεί και άλλες μέθοδοι όπως αναφέρονται και στην εργασία «*Background subtraction techniques: A review*», [34], άλλες πιο απλές και

άλλες πιο σύνθετες. Στην εργασία «Αναγνώριση και Χαρακτηρισμός Κίνησης σε Κυκλοφοριακά Βίντεο», [13] γίνεται μια σύγκριση μεταξύ πιο απλών μεθόδων και της Gaussian Mixture Model. Οι μέθοδοι αυτές είναι η Frame Differencing, Mean Filter και η Approximate Median Filter.

- Frame Differencing. Το φόντο εκτιμάται από το προηγούμενο frame και δίνεται από τον τύπο:

$$B(x, y, t) = I(x, y, t - 1)$$

$$|I(x, y, t) - I(x, y, t - 1)| > Th$$
(3.1.1)

- Mean Filter. Το φόντο εκτιμάται από τον μέσο όρο των προηγούμενων n frames και δίνεται από τον τύπο:

$$B(x, y, t) = \frac{1}{n} \sum_{i=0}^{n-1} I(x, y, t - i)$$

$$\left| I(x, y, t) - \frac{1}{n} \sum_{i=0}^{n-1} I(x, y, t - i) \right| > Th$$
(3.1.2)

- Median filter. Με την παραδοχή ότι το background είναι πιο πιθανό να εμφανιστεί σε μια σκηνή χρησιμοποιούμε το median των n frames εικόνα και δίνεται από τον τύπο:

$$B(x, y, t) = \text{median}\{I(x, y, t - i)\}$$

$$|I(x, y, t) - \text{median}\{I(x, y, t - i)\}| > Th$$
(3.1.3)

Όπου  $B(x, y, t)$  είναι το background,  $I(x, y, t)$  είναι η τρέχουσα εικόνα και το  $i \in \{0, \dots, n - 1\}$ .

Και οι τρεις αλγόριθμοι αφορούν σε προσεγγίσεις που είναι αρκετά γρήγορες παρόλα αυτά το global threshold τα κάνει ανεπαρκή και αναποτελεσματικά σε δύσκολες συνθήκες του πραγματικού κόσμου. Η μέθοδος μοντελοποίησης του φόντου με βάση Gaussian Mixture Models αφορά σε στατιστική προσέγγιση του Background Subtraction. Μπορεί να αντιμετωπίσει με επιτυχία τις αλλαγές φωτισμού, επαναλαμβανόμενες κινήσεις, ακαταστασία, την εισαγωγή ή την αφαίρεση αντικειμένων στη σκηνή και αργά κινούμενα αντικείμενα. Σύμφωνα με αυτή τη μέθοδο, οι τιμές ενός συγκεκριμένου pixel μοντελοποιούνται σαν ένα μείγμα προσαρμοστικών Gaussian κατανομών. Χρησιμοποιείται ένα μείγμα γιατί σ' ένα pixel εμφανίζονται πολλαπλές επιφάνειες και οι κατανομές πρέπει να



είναι προσαρμοστικές λόγω του ότι οι συνθήκες φωτισμού αλλάζουν. Αυτό συμβαίνει γιατί το φόντο συνήθως ποικίλλει σε συστήματα πραγματικού χώρου ανάλογα με τις καιρικές συνθήκες που επηρεάζουν την διαύγεια και την ώρα που περνάει, πράγμα που δείχνει ότι ακόμα και η χρησιμοποίηση ενός προκαθορισμένου, όπως στις προηγούμενες μεθόδους (thresholding), κατωφλίου δεν οδηγεί σε σωστά αποτελέσματα. Σε κάθε επανάληψη του αλγορίθμου που υλοποιεί η Gaussian Mixture Models, οι Gaussian κατανομές αξιολογούνται χρησιμοποιώντας μια απλή διαδικασία ώστε να προσδιοριστεί ποιά από αυτές είναι η πιο πιθανή να αντιστοιχεί στο φόντο.

Τα pixels εκείνα που δεν ταιριάζουν με τις Gaussian κατανομές του φόντου κατατάσσονται ως στοιχεία που ανήκουν στο foreground. Σε κάθε χρονική στιγμή  $t$ , αυτό που είναι γνωστό για κάθε pixel  $(x_0, y_0)$  είναι το ιστορικό του.

$$\{X_1, \dots, X_t\} = \{I(x_0, y_0, i): 1 \leq i \leq t\} \quad (3.1)$$

όπου  $I$  είναι η ακολουθία των frames.

Αυτό το ιστορικό του κάθε pixel,  $\{X_1, \dots, X_t\}$ , μοντελοποιείται σαν ένα μείγμα από  $K$  Gaussian κατανομές. Η πιθανότητα παρατήρησης της τιμής του τρέχοντος pixel είναι:

$$P(X_t) = \sum_{i=1}^K \omega_{i,t} * \eta(X_t, \mu_{i,t}, \Sigma_{i,t}) \quad (3.2)$$

όπου το  $K$  είναι ο αριθμός των κατανομών με την τιμή του να καθορίζεται από τη διαθέσιμη μνήμη και την υπολογιστική ισχύ και παίρνει τιμές από 3 έως 5 συνήθως,  $\omega_{i,t}$  είναι μια εκτίμηση του βάρους της  $i$ -οστής Gaussian κατανομής στο μείγμα τη χρονική στιγμή  $t$ ,  $\mu_{i,t}$  είναι η μέση τιμή της  $i$ -οστής Gaussian κατανομής στο μείγμα τη χρονική στιγμή  $t$ ,  $\Sigma_{i,t}$  είναι ο πίνακας συνδιακύμανσης της  $i$ -οστής Gaussian κατανομής στο μείγμα τη χρονική στιγμή  $t$  και  $\eta$  είναι μια Gaussian συνάρτηση πυκνότητας πιθανότητας που υπολογίζεται σύμφωνα με την εξίσωση:

$$\eta(X_t, \mu, \Sigma) = \frac{1}{(2\pi)^{\frac{n}{2}} |\Sigma|^{\frac{1}{2}}} e^{-\frac{1}{2}(X_t - \mu)^T \Sigma^{-1} (X_t - \mu)} \quad (3.3)$$

Η απόφαση για την τιμή του  $K$  εξαρτάται από τη διαθέσιμη μνήμη και την υπολογιστική ισχύ.

Κατά την έναρξη του συστήματος, οι  $K$  κατανομές Gaussian για ένα pixel αρχικοποιούνται με προκαθορισμένη μέση τιμή, υψηλή διακύμανση και χαμηλή τιμή βάρους. Κάθε νέα τιμή pixel,  $X_t$ , ελέγχεται με βάση τις υπάρχουσες Gaussian κατανομές, έως ότου βρεθεί μια αντιστοιχία. Μια αντιστοιχία ορίζεται ως μια τιμή ενός pixel εντός της τυπικής απόκλισης ( $D=2.5$ ) μιας κατανομής.

$$|X_t - \mu_{i,t-1}| \leq D\sigma_{i,t-1} \quad (3.4)$$

Τα προγενέστερα βάρη των  $K$  κατανομών τη χρονική στιγμή  $t$ ,  $\omega_{k,t}$  αναπροσαρμόζονται ως εξής:

$$\omega_{k,t} = (1 - \alpha)\omega_{k,t-1} + \alpha(M_{k,t}) \quad (3.5)$$

όπου  $\alpha$  είναι ο ρυθμός μάθησης και το  $M_{k,t}$  είναι 1 για την κατανομή που ταιριάζει και 0 για τις εναπομείναντες κατανομές.

Μετά από αυτό το στάδιο τα προηγούμενα βάρη των κατανομών κανονικοποιούνται και οι παράμετροι της Gaussian κατανομής η οποία ταιριάζει με τη νέα παρατήρηση ενημερώνεται ως εξής:

$$\begin{aligned} \mu_t &= (1 - \rho)\mu_{t-1} + \rho X_t \\ \sigma_t^2 &= (1 - \rho)\sigma_{t-1}^2 + \rho(X_t - \mu_t)^T(X_t - \mu_t) \end{aligned} \quad (3.6)$$

όπου

$$\rho = \alpha\eta(X_t | \mu_k, \sigma_k)$$

Εάν δεν βρεθεί καμία αντιστοιχία για το νέο pixel, η κατανομή Gaussian με τη μικρότερη πιθανότητα αντικαθίσταται με μια νέα κατανομή με την τρέχουσα τιμή pixel ως μέση τιμή της, μια αρχικά υψηλή διακύμανση και χαμηλή τιμή βάρους.

Οι κατανομές ταξινομούνται από την τιμή  $\omega/\sigma$ . Αυτή η ταξινομημένη λίστα κατανομών αντικατοπτρίζει τις πιο πιθανές κατανομές φόντου από πάνω προς τα κάτω, δεδομένου ότι από την  $\omega_{k,t}$  οι διαδικασίες pixel του φόντου κάνουν την αντίστοιχη Gaussian κατανομή να έχει μεγαλύτερη τιμή βάρους και μικρότερη διακύμανση. Στη συνέχεια οι πρώτες  $B$  κατανομές επιλέγονται ως φόντο:

$$B = \arg \min_b \left( \sum_{k=1}^b \omega_k > T \right) \quad (3.7)$$

όπου  $T$  είναι το ελάχιστο τμήμα των δεδομένων pixel που πρέπει να ληφθούν υπόψη από το φόντο.

$$\begin{aligned} B_r(x, y) &= \mu_{N,t}(x, y) \\ D_c &= \begin{cases} 1 & \text{if } |I_r(x, y) - B_r(x, y)| > T \\ 0 & \text{else} \end{cases} \end{aligned} \quad (3.8)$$

Τελικά τα pixel  $D_c = 1$  είναι τα κινούμενα κομμάτια στην εικόνα που τα παίρνουμε από την αφαίρεση της τρέχουσας εικόνας  $I_r$  με το μοντελοποιημένο φόντο  $B_r$ .

Μεγάλη σημασία για τη λειτουργικότητα του αλγορίθμου μας έχουν οι παράμετροι που επιλέγουμε για την αρχικοποίηση των Gaussian Μοντέλων. Η ιδιαιτερότητα με τα δικά

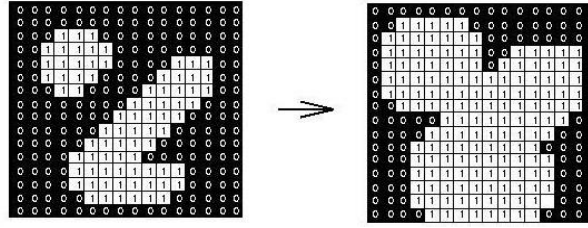
μας βίντεο έχει να κάνει με το ότι πρόκειται για βίντεο με μικρό αριθμό από frames οπότε ενώ ο default αριθμός είναι τα 150 frames για μεγάλα βίντεο που συνήθως εφαρμόζεται ο αλγόριθμος εμείς έχουμε  $N=30$ . Επίσης ο ρυθμός μάθησης ενώ ο default είναι 0,005 εμείς τον επιλέγουμε με βάση τα frames που χρειάζονται για εκπαίδευση και προκύπτει προσεγγιστικά από τον τύπο  $\alpha = \frac{1}{N}$  οπότε  $\alpha = 0,03$ . Τέλος επιλέγουμε το  $K=3$  πειραματικά. Δοκιμάζοντας για  $K=4$  και  $K=5$  δεν υπήρξε κάποια ουσιαστική διαφορά στα αποτελέσματά μας οπότε και δεν υπήρχε λόγος για να κάνουμε πιο περίπλοκο το σύστημα.

Οι αλγόριθμοι Kernel Density Estimation and Visual Background Extractor είναι δίνουν επίσης καλά αποτελέσματα σε παρόμοια λογική με το GMM όπως αναφέρεται και στην εργασία «*Fish detection for species recognition.*», [35] και χρήζουν περαιτέρω εξερεύνησης. Επιλέγουμε όμως την GMM γιατί θα μπορέσει να δώσει ικανοποιητικά αποτελέσματα καθώς έχει εφαρμογή σε πληθώρα από διαφορετικά βίντεο, διαφορετικών συνθηκών και είναι ευαίσθητη στην κίνηση συγκριτικά με τις υπόλοιπες τεχνικές.

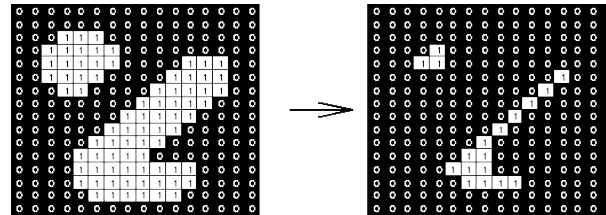
### 3.2.2 Morphological Operations

Από την παραπάνω διαδικασία έχουμε σαν αποτέλεσμα μια ακριβή κατάτμηση των εικόνων σε δυαδική αναπαράσταση όπου με την τιμή '0' και μαύρο χρώμα απεικονίζεται το background και με την τιμή '1' και λευκό χρώμα απεικονίζεται το foreground. Λόγω του ότι οι υποθαλάσσιες εικόνες έχουν αρκετό θόρυβο που εντάσσεται λανθασμένα στο foreground χρειάζεται να γίνει μια επεξεργασία στα δεδομένα που προκύπτουν από το background subtraction με σκοπό να μειωθεί ο θόρυβος αλλά και να γίνει μια τέτοια επεξεργασία που θα βοηθήσει σε καλύτερη εξαγωγή συμπερασμάτων όπως βλέπουμε και στην εργασία "*Tracking a Moving Objects Using Foreground Detector and Improved Morphological Filter.*" , [36] και χρησιμοποιείται και στην διπλωματική εργασία «*Fish detection for species recognition.*», [35]. Η επεξεργασία που προτείνουμε στο συγκεκριμένο στάδιο αφορά στο μορφολογικό φιλτράρισμα. Το μορφολογικό φιλτράρισμα απλοποιεί την κατατετμημένη δυαδική εικόνα με σκοπό την ακριβέστερη ανίχνευση των ψαριών και διαχωρίζεται στις εξής διαδικασίες: το morphological opening, morphological closing και filling the holes.

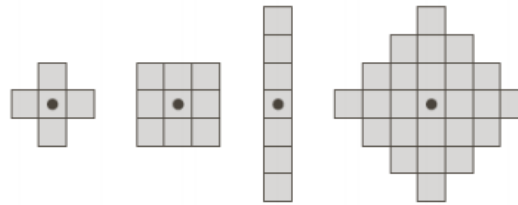
Για να μπορέσουν να εξηγηθούν οι παραπάνω διαδικασίες είναι χρήσιμο να γίνει μια αναφορά στις έννοιες erosion, dilation και structure element, όπως αναφέρονται και στο σύγγραμμα «*Ψηφιακή επεξεργασία εικόνας*», [22]. Το dilation (Εικόνα 3.2.1) είναι μια μεθοδολογία που διευρύνει, επεκτείνει τα αντικείμενα και έχει σαν αποτέλεσμα το γέμισμα μικρών οπών και τη συνένωση κομμένων κομματιών. Από την άλλη το erosion (Εικόνα 3.2.2) είναι μια μεθοδολογία που οδηγεί στη σμίκρυνσή τους περιορίζοντας ή και εξαφανίζοντας τελείως μικρά αντικείμενα. Ο τρόπος με τον οποίο επιτυγχάνονται αυτές οι αλλαγές στη δυαδική εικόνα έχει να κάνει με τη δομή (structuring element) που θα επιλέξουμε. Όπου δομή είναι ένα παράθυρο με μέγεθος που ορίζουμε εμείς και συγκεκριμένο σχεδιασμό όπως φαίνεται στην (Εικόνα 3.2.3.)



**Εικόνα 3.4.1** Dilation με rectangle structuring element με παράθυρο: [3, 3]



**Εικόνα 3.4.2** Erosion με rectangle structuring element με παράθυρο: [3, 3]



**Εικόνα 3.4.3** Types of Structuring elements.

Οι δυο αυτές λειτουργίες μπορούν να συνδυαστούν σε πιο περίπλοκες λειτουργίες με σκοπό καλύτερα αποτελέσματα. Αυτές που χρησιμοποιούμε εμείς είναι το opening και το closing. Το opening αναφέρεται σε erosion ακολουθούμενο από dilation. Με τον τρόπο αυτό αφαιρούμε μικρά αντικείμενα που κυρίως αναφέρονται σε θόρυβο.

$$A \circ B = (A \ominus B) \oplus B \quad (3.9)$$

Το closing αναφέρεται σε dilation ακολουθούμενο από erosion και είναι χρήσιμο να γεμίσουν τρύπες που υπάρχουν και να διευρύνει γενικά τα όρια του αντικειμένου.

$$A \cdot B = (A \oplus B) \ominus B \quad (3.10)$$

Όπου A είναι η δυαδική μας εικόνα και B είναι το επιλεγμένο structure element.

Στη μεθοδολογία μας εφαρμόζουμε και τις δυο αυτές λειτουργίες με σημαντική διαφορά το structure element που επιλέγουμε με βάση τις δυο ομάδες που εξετάζουμε. Όταν εξετάζουμε κοπάδι ψαριών κάνουμε opening και closing των αντικειμένων με δομή το

τετράγωνο (rectangle). Ο λόγος είναι ότι θέλουμε να δούμε την κίνηση της μάζας και όχι τις μεμονωμένες κινήσεις τους οπότε εφαρμόζουμε τέτοιες παραμέτρους που τα αντικείμενα διαστέλλονται και ενώνονται σε μια μάζα, εκτός εάν κάποιο ψάρι αποκλίνει πολύ από αυτήν οπότε και δεν επηρεάζεται από τις μορφολογικές διαδικασίες (Εικόνα 3.9). Ενώ για την κίνηση του κάθε ψαριού μας ενδιαφέρει τα ψάρια να μην πιάνουν μεγάλο χώρο και όσο γίνεται να μην ενώνονται μεταξύ τους πέρα από περιπτώσεις occlusion που δεν μπορεί να διορθωθεί με τέτοιες τεχνικές. Οπότε κάνουμε opening και closing των αντικειμένων με δομή την ευθεία (line), (Εικόνα 3.7). Τέλος κάνουμε filling the holes με σκοπό να κλείσουν όσες τρύπες έχουν απομείνει στα ψάρια. Οι δομές που επιλέξαμε τελικά, προέκυψαν πειραματικά.

### 3.2.3 Blob Analysis

Επόμενο στάδιο είναι ο ακριβής εντοπισμός/ανίχνευση των αντικειμένων ο οποίος γίνεται με μια απλή μεθοδολογία, το Blob Analysis. Πρόκειται ουσιαστικά για μια μέθοδο που αναγνωρίζει τις ομάδες από λευκά pixel, blobs και τα διαχωρίζει μεταξύ τους με ένα boundary box παρέχοντάς μας πληροφορίες για το κέντρο βάρους του καθενός και την περιοχή που καλύπτει στην εικόνα. Με τον τρόπο αυτό μπορούμε να ξέρουμε πόσα αντικείμενα έχουμε στην εικόνα, ποιά η θέση τους και να εξαγάγουμε χρήσιμες πληροφορίες που θα μας χρειαστούν για το επόμενο στάδιο του tracking. Επίσης, με το blob analysis έχουμε τη δυνατότητα κατά τον εντοπισμό να θέσουμε όριο στο μέγεθος των blobs που επιτρέπεται να συμπεριληφθούν στο detection. Στο σημείο αυτό γίνεται και πάλι διαχωρισμός στο όριο που θα βάλουμε με βάση την ομάδα που εξετάζουμε, το κοπάδι δηλαδή ή την ξεχωριστή συμπεριφορά των ψαριών. Εφόσον το όριο αυτό αναφέρεται στη περιοχή που καλύπτει στην εικόνα και αφορά pixels είναι σημαντικό να έχουμε σταθερό και ίδιο μέγεθος για κάθε εικόνα σε όλα τα βίντεο, όπως επίσης σταθερή απόσταση και το ζουμ της κάμερας. Με βάση αυτές τις πληροφορίες μπορούμε να κάνουμε και τις αντίστοιχες παραμετροποιήσεις. Όταν εξετάζουμε τη μάζα των ψαριών το όριο αυτό είναι πολύ μικρό και είναι στα **50 pixels**. Εφόσον το κοπάδι είναι απομακρυσμένο και με τα μορφολογικά φίλτρα που έχουμε εφαρμόσει εμφανίζεται ενιαία μάζα, οπότε οποιαδήποτε σημαντική απόκλιση από αυτήν, που μπορεί να είναι ακόμα και ένα ψάρι, αυτό πρέπει να εντοπιστεί. Οπότε, πειραματικά εξετάζοντας τις περιπτώσεις είδαμε ότι ένα τέτοιο όριο μπορεί να μας δώσει αποτελέσματα σε συγκεκριμένες συνθήκες εικόνας (Εικόνα 3.8, Εικόνα 3.9). Όσον αφορά στη μελέτη της μεμονωμένης κίνησης των ψαριών το όριο αυτό δεκαπλασιάζεται και είναι στα **500 pixels**. Εφόσον πρόκειται για συνθήκες με πολλά ψάρια τα οποία βρίσκονται πιο κοντά στην εικόνα, προσπαθούμε όσα απομακρυσμένα να μην τα συμπεριλάβουμε στην ανίχνευση γιατί το πιθανότερο είναι όπως παρατηρήσαμε στα βίντεο να εξαφανιστούν σε λίγα pixels ή να βρίσκονται για αρκετά frames χαμένα. Οπότε και δεν υπάρχει λόγος να γίνει η ανίχνευσή τους καθώς δεν βοηθά στην εξαγωγή κάποιου συμπεράσματος (Εικόνα 3.5, Εικόνα 3.6, Εικόνα 3.7).



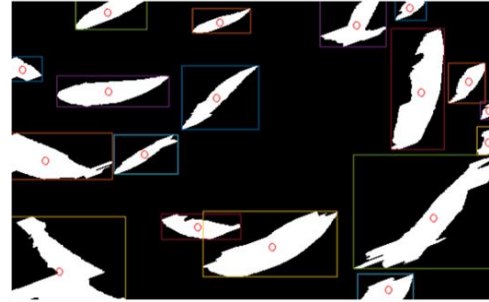
(α)



(β)



(γ)



(δ)

**Εικόνα 3.5** (α) Frame 126 από Video\_1. (β) Αποτέλεσμα από το background subtraction στο ίδιο frame. (γ) Αποτέλεσμα από το morphological operation. (δ) Αποτέλεσμα από το blob analysis.



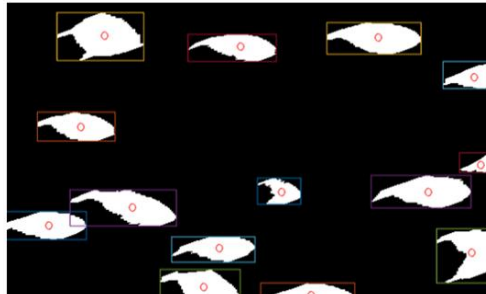
(α)



(β)



(γ)



(δ)

**Εικόνα 3.6** (α) Frame 141 από Video\_2. (β) Αποτέλεσμα από το background subtraction στο ίδιο frame. (γ) Αποτέλεσμα από το morphological operation. (δ) Αποτέλεσμα από το blob analysis.





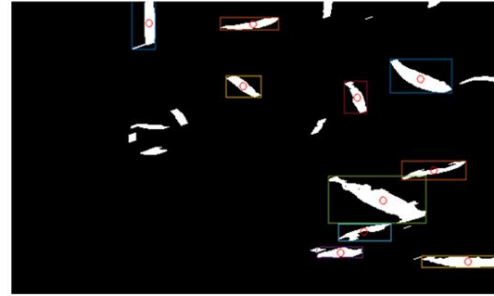
(α)



(β)



(γ)



(δ)

**Εικόνα 3.7** (α) Frame 113 από Video\_3. (β) Αποτέλεσμα από το background subtraction στο ίδιο frame. (γ) Αποτέλεσμα από το morphological operation. (δ) Αποτέλεσμα από το blob analysis.



(α)



(β)

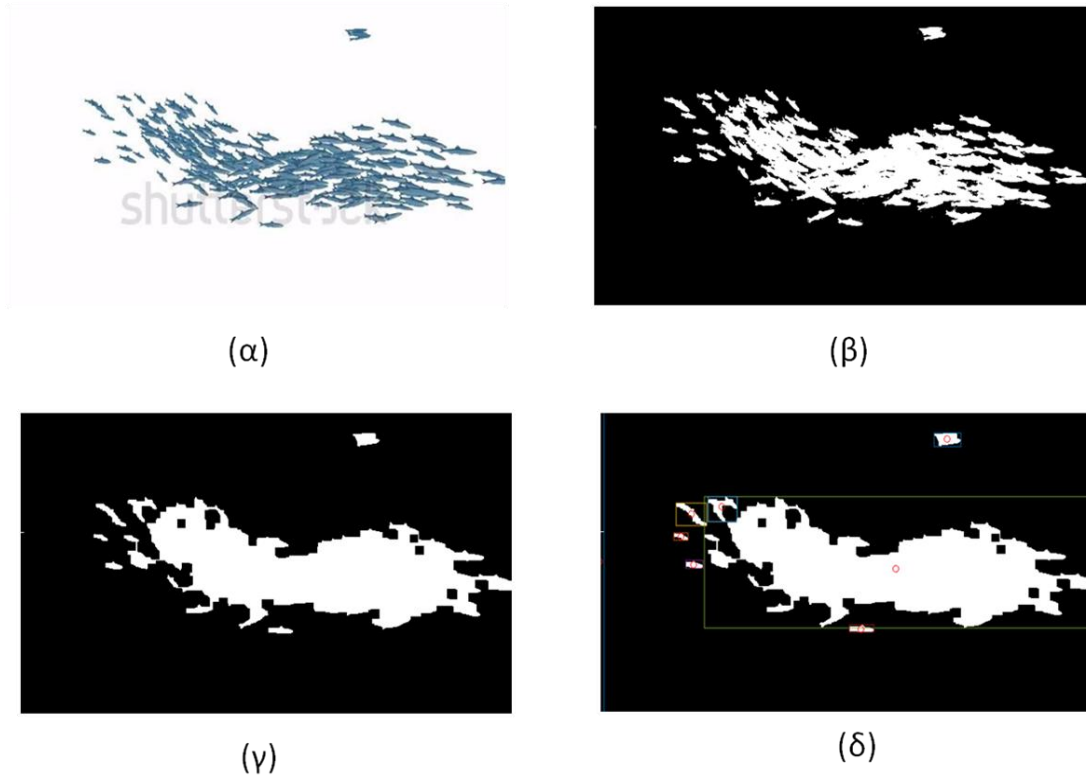


(γ)



(δ)

**Εικόνα 3.8** (α) Frame 103 από Video\_6. (β) Αποτέλεσμα από το background subtraction στο ίδιο frame. (γ) Αποτέλεσμα από το morphological operation. (δ) Αποτέλεσμα από το blob analysis.



**Εικόνα 3.9** Ομάδα ψαριών με ψάρι να αποκλίνει από την ομάδα: (α) Frame 110 από Video\_4. (β) Αποτέλεσμα από το background subtraction στο ίδιο frame. (γ) Αποτέλεσμα από το morphological operation. (δ) Αποτέλεσμα από το blob analysis.

### 3.2.4 Ακρίβεια Ανίχνευσης

Παρακάτω παρουσιάζεται η ακρίβεια των αποτελεσμάτων μας για το detection βάση των ακόλουθων εξισώσεων που συναντάμε και στην εργασία «Fish detection for species recognition.»,[35]:

$$P = \frac{TP}{TP + FP} \quad (3.11)$$

Precision (P): Προσδιορίζει πόσο αποδοτικό είναι το detection. Υψηλή τιμή του P δηλώνει ότι ο αλγόριθμος επιστρέφει περισσότερα σωστά detection από ότι λανθασμένα.

$$R = \frac{TP}{TP + FN} \quad (3.12)$$

Recall (R): Προσδιορίζει πόσα ψάρια έχουν εντοπιστεί σωστά.

$$A = 2 \frac{P * R}{P + R} \quad (3.13)$$

Accuracy score (A) : Προσδιορίζει την τελική ακρίβεια .

Τα TP,TN,FP,FN προσδιορίζονται βάση του Πίνακα 3.1

	Fish was detected	Fish was not detected
Frame contains fish	True Positive	False Negative
Frame does not contain fish	False Positive	True Negative

**Πίνακας 3.1** Υπολογισμός TP,TN,FP,FN.

Βάση αυτών των τιμών μετράμε την ακρίβεια των αποτελεσμάτων από το detection. Τα τελικά αποτελέσματα φαίνονται στον Πίνακα 3.2. Σημαντικό είναι να αναφέρουμε ότι παρόλο που σε βίντεο με μεγάλο αριθμό ψαριών δεν εντοπίζονται όλα, μας ενδιαφέρει κυρίως αυτά που θα εντοπιστούν να αντιστοιχούν όντως σε ψάρια. Από όλο το σύνολο των 11 βίντεο που εξετάσαμε σταθήκαμε σε 3 βίντεο για να εξάγουμε την ακρίβεια στο detection βάση των παραπάνω τύπων. Ο λόγος είναι ότι τα βίντεο αυτά συνδυάζουν τις πιο δύσκολες συνθήκες που έχουμε αναφέρει και παραπάνω. Σε ιδανικές συνθήκες με βίντεο από μικρό αριθμό ψαριών και σταθερό background χωρίς επικαλύψεις η ακρίβεια έχει πολύ καλά αποτελέσματα. Για τα παρακάτω αποτελέσματα έχουμε πάρει ένα δείγμα 30 τυχαίων frame σε κάθε βίντεο. Στις Εικόνες 3.10, 3.11, 3.12 φαίνονται ενδεικτικά frames των βίντεο.

Βίντεο	30/ Σύνολο Frames	P	R	A
Video_1	30/440	95%	75%	83%
Video_2	30/325	99%	94%	96%
Video_3	30/636	94%	69%	79%
Συγκεντρωτικά Αποτελέσματα	90/1401	96%	79%	87%

**Πίνακας 3.2** Πίνακας ακρίβειας detection.



**Εικόνα 3.9.1** Frames από Video\_1



**Εικόνα 3.9.2** Frames από Video\_2



**Εικόνα 3.9.3** Frames από Video\_3

### 3.3 Παρακολούθηση Κίνησης

Η παρακολούθηση της κίνησης (tracking) και η εξαγωγή των τροχιών αποτελεί το σημαντικότερο βήμα στη μεθοδολογία μας. Αφορά στην εκτίμηση της τροχιάς των ψαριών στο επίπεδο της εικόνας καθώς αυτά κινούνται. Για να εξάγουμε συμπεράσματα για τη συμπεριφορά των ψαριών πρέπει να μπορούμε να μελετάμε τις τροχιές τους, είτε αφορούν σε ολόκληρο το κοπάδι είτε σε μεμονωμένα ψάρια, σε επίπεδο αριθμητικής αναπαράστασης, ώστε να μπορούμε να εξάγουμε και αυτόματα συμπεράσματα. Εφόσον έχει γίνει η ανίχνευση των ψαριών σε κάθε frame, σκοπός είναι να γίνει η αντιστοίχιση από frame σε frame ώστε να είναι διακριτό ότι πρόκειται για το ίδιο ψάρι (ή κοπάδι) κάθε φορά. Στη περίπτωση μας έχουμε να κάνουμε με ιδιαίτερα δύσκολες συνθήκες καθώς πρόκειται για μεγάλο και μη σταθερό αριθμό από ψάρια που υπόκεινται σε πολλές επικαλύψεις. Αυτό είναι ένα πρόβλημα που έχει απασχολήσει αρκετά τους ερευνητές έχοντας αναπτύξει διάφορες μεθοδολογίες όπως αναφέρθηκε και στο 1ο Κεφάλαιο. Παρακάτω γίνεται εκτενής αναφορά στο αλγόριθμο πρόβλεψης Kalman και αναλύεται η μεθοδολογία που ακολουθούμε σε αυτή την εργασία. Επίσης περιγράφεται η επεξεργασία των τροχιών για βελτίωση των αποτελεσμάτων και τέλος παρουσιάζονται τα τελικά αποτελέσματα.

#### 3.3.1 Kalman Filtering

Η μέθοδος Kalman Filtering όπως αναφέρεται και στην εργασία "*Kalman filter tracking*", [40] είναι ένας από τους πιο διαδεδομένους και πιο αποτελεσματικούς αλγόριθμους σε ζητήματα tracking τον οποίο αναπτύξαμε μέσω της τεχνικής γλώσσας προγραμματισμού Matlab «*Kalman Filtering Implementation with Matlab*», [41]. Η επιλογή



του φίλτρου έγινε βάση της μικρής πολυπλοκότητάς του και των βέλτιστων αποτελεσμάτων που δίνει σε σχέση με άλλους αλγόριθμους πρόβλεψης όπως αναφέρεται στην έρευνα «*The Kalman Filter and Related Algorithms: A Literature Review*», [42]. Το φίλτρο Kalman χρησιμοποιείται σε γραμμικά, διακριτού χρόνου συστήματα με σταθερή ταχύτητα. Στο σημείο αυτό είναι σημαντικό να αναφέρουμε ότι μπορεί η κίνηση των ψαριών ορισμένες φορές να είναι απρόβλεπτη και όχι γραμμική αλλά με την παραδοχή ότι τα βίντεό μας έχουν μεγάλο frame rate (30 fps) και το ότι το φίλτρο λειτουργεί αναδρομικά με κύρια βαρύτητα την προηγούμενη κατάσταση μας δίνει τη δυνατότητα να μοντελοποιήσουμε μια τέτοια κίνηση, ως κίνηση με σταθερή ταχύτητα και να μην έχουμε προβλήματα με την επιλογή του φίλτρου Kalman σαν αλγόριθμο πρόβλεψης, κάτι το οποίο αναφέρεται και έχει εφαρμοστεί και στην εργασία «*Using the Kalman filter to track human interactive motion modeling and initialization of the Kalman filter for translational motion*», [43] για εντοπισμό κίνησης πολλών ανθρώπων. Σκοπός του φίλτρου είναι να προβλέψει στα μελλοντικά frames την επόμενη θέση των ψαριών με βάση προγενέστερη γνώση. Κάτι το οποίο αποδεικνύεται πολύ χρήσιμο για περιπτώσεις επικαλύψεων και για περιπτώσεις που το detection δεν είναι απόλυτα ακριβές σε κάθε frame. Τα πλεονεκτήματα στη διαδικασία ένταξης του φίλτρου στην ανίχνευση είναι ότι :

- Μειώνει την περιοχή αναζήτησης για εκ νέου ανίχνευση ενός ψαριού και έτσι μειώνεται ο χρόνος επεξεργασίας, κάτι που είναι πολύ σημαντικό σε real time εφαρμογές.
- Μπορεί να μειώσει τον αριθμό των εσφαλμένων ανιχνεύσεων δεδομένου ότι η περιοχή της εικόνας που δεν περιέχει ψάρια αποκλείεται από την αναζήτηση.

Ο τρόπος με τον οποίο λειτουργεί το φίλτρο Kalman αφορά στη μοντελοποίηση της κατανομής πιθανότητας κάθε κατάστασης ως μιας Gaussian κατανομής. Αυτό σημαίνει ότι κάθε φίλτρο Kalman μπορεί να εντοπίσει μόνο ένα ψάρι σε κάθε frame. Όμως αυτό δεν είναι πρακτικό για σχεδόν καμία πραγματική εφαρμογή και γι αυτόν το λόγο, για να ξεπεραστεί αυτός ο περιορισμός και να εντοπίζει ο αλγόριθμος πολλαπλά ψάρια σε κάθε frame, ένα ξεχωριστό φίλτρο Kalman αρχικοποιείται για κάθε διαφορετικό ψάρι στη λίστα με τα ψάρια προς εντοπισμό. Η *a priori* εκτίμηση του φίλτρου Kalman υποδεικνύει τη θέση και το μέγεθος της περιοχής όπου ένα ψάρι θα μπορούσε ενδεχομένως να εμφανιστεί στο επόμενο frame του βίντεο. Οι πληροφορίες αυτές χρησιμοποιούνται από τη διαδικασία εντοπισμού και συσχέτισης των ψαριών ώστε να γνωρίζουμε ότι αναφέρεται στο ίδιο ψάρι κάθε φορά. Μόλις ολοκληρωθεί η εκ νέου ανίχνευση, τα νέα δεδομένα των μετρήσεων θα πρέπει να προσαρμοστούν στο μοντέλο του συστήματος. Η *aposteriori* εκτίμηση υπολογίζεται στη συνέχεια και χρησιμοποιείται ως η καλύτερη εκτίμηση για τη θέση και το μέγεθος του ψαριού.

Τα μοντέλα διαδικασίας και μέτρησης ενός γραμμικού συστήματος διακριτού χρόνου μπορεί να οριστούν από τις ακόλουθες εξισώσεις:

$$\begin{aligned}x_k &= F_{k-1} + w_{k-1} \\ y_k &= H_k x_k + v_k\end{aligned}\tag{3.14}$$



όπου  $x_k$  και  $y_k$  είναι τα διανύσματα κατάστασης και μέτρησης κατά τη χρονική στιγμή  $k$ . Τα  $F_k$  και  $H_k$  είναι οι πίνακες μετάβασης και μέτρησης αντίστοιχα. Τα  $w_k$  και  $v_k$  είναι η διαδικασία και η μέτρηση του θορύβου. Υποτίθεται ότι είναι ανεξάρτητα με λευκό Gaussian θόρυβο και πίνακες συνδιακύμανσης  $Q_k$  και  $R_k$  αντίστοιχα:

$$w_k: (0, Q_k)$$

$$v_k: (0, R_k)$$

(3.15)

Το φίλτρο Kalman είναι ένα αναδρομικό φίλτρο πρόβλεψης, άρα λοιπόν υπολογίζει την κατάσταση τις δυναμικού συστήματος αναδρομικά. Αυτό το στάδιο αποτελείται από δύο φάσεις:

- Πρόβλεψη (prediction)
- Διόρθωση (correction)

Η φάση τις πρόβλεψης προβάλλει μπροστά στο χρόνο την τρέχουσα κατάσταση και τη συνδιακύμανση σφάλματος, για την απόκτηση a priori προβλέψεων για το επόμενο χρονικό βήμα. Η φάση της διόρθωσης ενσωματώνει την τελευταία μέτρηση στο μοντέλο του συστήματος, για να πάρει τις a posteriori εκτιμήσεις της κατάστασης και της συνδιακύμανσης σφάλματος. Αυτή η διαδικασία επαναλαμβάνεται για κάθε χρονικό βήμα, με την κατάσταση του προηγούμενου χρονικά βήματος ως αρχική τιμή. Γι αυτό το φίλτρο Kalman αποκαλείται ως ένα αναδρομικό φίλτρο. Στις παρακάτω εξισώσεις συνοψίζονται οι δύο φάσεις του αλγορίθμου. Η εξίσωση τις φάσης πρόβλεψης για τον υπολογισμό της a priori κατάστασης εκτίμησης  $\hat{x}_k^-$  και η συνδιακύμανση σφάλματος  $P_k^-$  δίνονται από τις:

$$\hat{x}_k^- = F_{k-1} \hat{x}_{k-1}^+$$

$$\hat{P}_k^- = F_{k-1} P_{k-1}^+ F_{k-1}^T$$

(3.15)

όπου ο δείκτης υποδεικνύει το χρονικό βήμα, ενώ οι εκθέτες “-” και “+” υποδεικνύουν τις a priori και a posteriori εκτιμήσεις αντίστοιχα. Το  $\hat{x}$  υποδηλώνει την εκτίμηση για το  $x$ .

Μετά από την απόκτηση της πιο πρόσφατης μέτρησης  $y_k$  η a posteriori εκτίμηση  $\hat{x}_k^+$  και η αντίστοιχη συνδιακύμανση σφάλματος  $P_k^+$ , υπολογίζονται χρησιμοποιώντας τις παρακάτω εξισώσεις της φάσης διόρθωσης :

$$\hat{x}_k^+ = \hat{x}_k^- + K_k (y_k - H_k \hat{x}_k^-)$$

$$\hat{P}_k^+ = (I - K_k H_k) P_k^- \quad (3.16)$$

όπου το  $K_k$  είναι το κέρδος του φίλτρου Kalman και δίνεται από :

$$K_k = \frac{P_k^- H_k^T}{H_k P_k^- H_k^T + R_k} \quad (3.17)$$

Η αναδρομή για τις φάσεις πρόβλεψης και διόρθωσης επαναλαμβάνεται για κάθε χρονικό βήμα μέτρησης. Ο στόχος του φίλτρου Kalman είναι να παρέχει την καλύτερη εκτίμηση για την κατάσταση του συστήματος με βάση την τρέχουσα διαθέσιμη γνώση στο μοντέλο του συστήματος και τα πιο πρόσφατα δεδομένα των μετρήσεων. Στην Εικόνα 3.12 φαίνεται ενδεικτικό παράδειγμα για ψάρι το οποίο χάνεται για λίγα frames και με επιτυχία εντοπίζεται ξανά με το ίδιο id λόγω της πρόβλεψης του Kalman Filter.

### 3.3.2 Ανάλυση της μεθοδολογίας μας

Η βασική λογική του αλγορίθμου που υλοποιήσαμε βασίζεται σε παράδειγμα που υπάρχει στο εγχειρίδιο της Matlab για tracking σε πολλαπλά αντικείμενα, προσαρμόζοντας τον βεβαία στα δικά μας δεδομένα και κάνοντας τις απαραίτητες αλλαγές. Εφόσον πρόκειται για επεξεργασία βίντεο ο αλγόριθμος λειτουργεί αναδρομικά για κάθε εικόνα και βασικό ζητούμενο είναι το πως θα γίνουν οι αντιστοιχίσεις μεταξύ των ανιχνευμένων ψαριών. Αρχικό βήμα αποτελεί η δημιουργία και αρχικοποίηση πίνακα δομών με πληροφορίες που αποθηκεύουμε για κάθε ψάρι που θεωρείται tracked. Παρακάτω ακολουθούν τα βασικά βήματα του αλγορίθμου για κάθε αναδρομή του και ύστερα περιγράφονται τα σημαντικά σημεία του αλγορίθμου μας.

#### Βήματα αλγορίθμου για το Tracking

##### Βήμα 1°

Εισαγωγή binary εικόνας από το detection.

##### Βήμα 2°

Ενημέρωση δομής για κάθε detected ψάρι με χρήσιμες πληροφορίες για την κατάστασή του.

##### Βήμα 3°

Πρόβλεψη για την επόμενη θέση των detected ψαριών με βάση το Kalman filter.

##### Βήμα 4°

Δημιουργία πίνακα κόστους για αντιστοίχιση detected και tracked ψαριών.

##### Βήμα 5°

Ενημέρωση πίνακα για αντιστοιχισμένα tracked ψάρια. Διόρθωση της θέσης βάση του Kalman filter και ενημέρωση της δομής του.

##### Βήμα 6°

Ενημέρωση πίνακα και δομών για μη αντιστοιχισμένα tracked ψάρια.

##### Βήμα 7°

Διαγραφή tracks που δεν εντοπίζονται για πολλά συνεχόμενα frames ή tracks που

ξεπερνάνε συγκεκριμένο αριθμό από detections σε όλη τους την τροχιά.

#### **Βήμα 8<sup>ο</sup>**

Δημιουργία νέα νέων tracks από μη αντιστοιχισμένα detections.

#### **Βήμα 9<sup>ο</sup>**

Εμφάνιση αποτελεσμάτων bounding boxes μόνο των tracked ψαριών.

Πιο αναλυτικά, το σύστημα δέχεται σαν είσοδο το βίντεο. Επεξεργαζόμαστε frame προς frame το βίντεο και σε πρώτο στάδιο γίνεται το detection όπως αναλύθηκε παραπάνω. Έπειτα δημιουργούμε έναν πίνακα όπου κάθε θέση του είναι μια δομή που αντιπροσωπεύει κάθε κινούμενο ψάρι στο βίντεο. Ο σκοπός της δομής είναι να διατηρήσει την κατάσταση ενός ψαριού που προσπαθεί να εντοπιστεί από τον αλγόριθμο. Η δομή αυτή αποτελείται από μεταβλητές που αποθηκεύουν πληροφορίες για το κάθε detection . Αυτές είναι:

- id : Ξεχωριστή ταυτότητα για κάθε ψάρι που εντοπίζεται.
- bbox : Το bounding box που το περικλείει.
- KalmanFilter : Ένα αντικείμενο Kalman Filter που χρησιμοποιείται για το tracking για κάθε ξεχωριστό id ψαριού.
- Age : Ο αριθμός από frames από τη στιγμή που έγινε tracked.
- totalVisibleNum : Ο συνολικός αριθμός που το tracker ψάρι είναι "ορατό" και εντοπίζεται.
- totalInvisibleNum : Ο αριθμός από συνεχόμενα frames που το tracked ψάρι δεν εντοπίζεται.

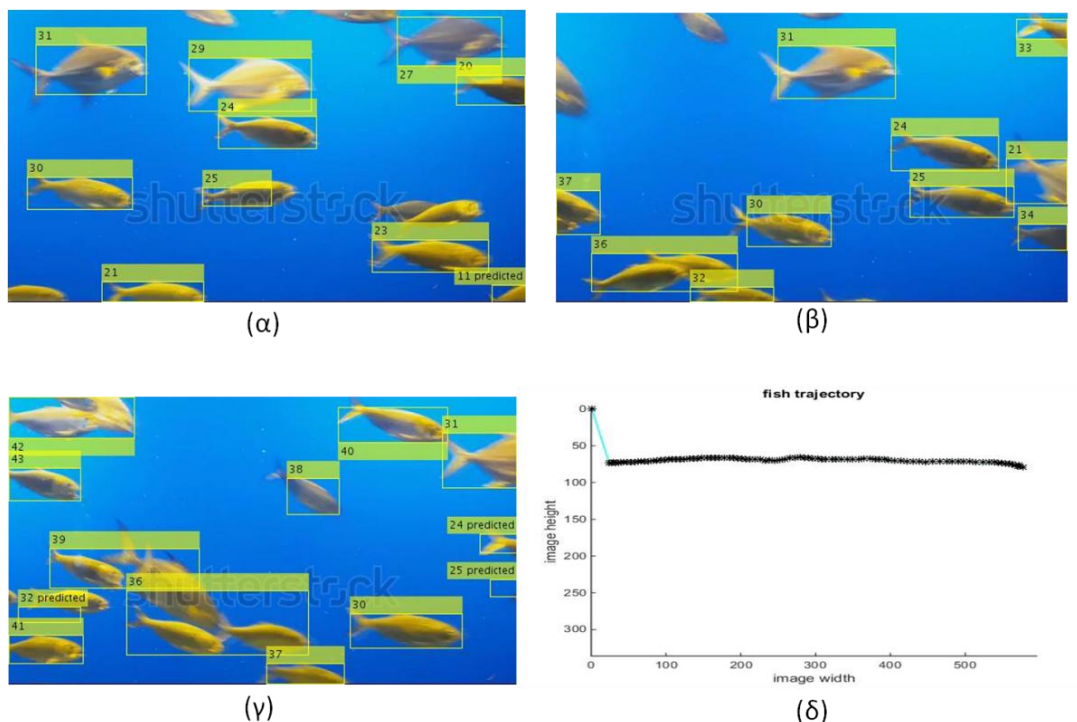
Έπειτα από το bbox βρίσκουμε το κέντρο βάρους για κάθε ψάρι που ανιχνεύθηκε ανά frame. Στη συνέχεια χρησιμοποιείται το φίλτρο Kalman ώστε να προβλεφθεί το κέντρο βάρους των αντιστοιχισμένων ψαριών στο επόμενο frame και να ενημερωθεί το bbox που το περιέχει αναλόγως. Το κέντρο βάρους ενός επίπεδου σχήματος ή ενός σχήματος δύο διαστάσεων, είναι η θέση του αριθμητικού μέσου όρου όλων των σημείων του σχήματος. Σε κάθε frame έχουμε ψάρια που προκύπτουν από την ανίχνευση και ψάρια που προκύπτουν από την πρόβλεψη. Σημαντικό βήμα είναι η αντιστοίχιση στο ίδιο ψάρι και η δημιουργία νέων id για νέες τροχιές. Η αντιστοίχιση ανίχνευσης ψαριού στο τρέχον frame με τα υπάρχοντα ψάρια που έχουν εντοπιστεί, γίνεται με ελαχιστοποίηση του κόστους. Ο υπολογισμός του κόστους της αντιστοίχισης κάθε ανίχνευσης σε κάθε ψάρι γίνεται χρησιμοποιώντας τη μέθοδο της Ευκλείδειας απόστασης. Το κόστος λαμβάνει υπόψη την Ευκλείδεια απόσταση ανάμεσα στο κέντρο βάρους που έχει προβλεφθεί για το ψάρι και του κέντρου βάρους της ανίχνευσης. Τα αποτελέσματα αποθηκεύονται σε έναν πίνακα cost matrix με διαστάσεις  $M*N$ , όπου  $M$  είναι ο αριθμός των προβλέψεων και  $N$  είναι ο αριθμός των ανιχνεύσεων. Επειδή ο αριθμός των ψαριών δεν είναι σταθερός γιατί κάθε φορά νέα ψάρια εμφανίζονται ενώ άλλα διαγράφονται εφαρμόζουμε τον αλγόριθμο Munkres Version of the Hungarian Algorithm «*Solving large-scale assignment problems by Kuhn-Munkres algorithm*», [39]. Σκοπός του είναι με απλές μαθηματικές πράξεις στον πίνακα να εξάγει τις αντιστοιχίσεις με το ελάχιστο κόστος. Από τη διαδικασία αυτή ενημερώνονται οι ακόλουθες περιπτώσεις:

➤ Αντιστοιχισμένα tracks. Πρόκειται για ενημέρωση της τροχιάς με το ίδιο id εφόσον αυτό συνεχίζει να εντοπίζεται στην εικόνα.

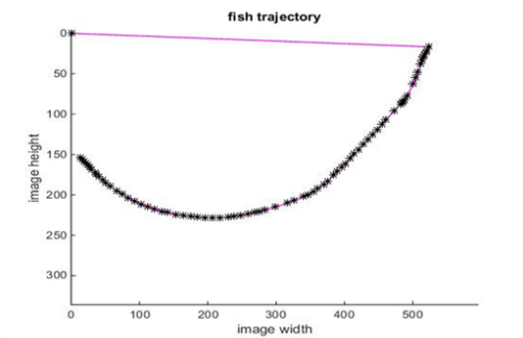
➤ Μη αντιστοιχισμένα tracks. Πρόκειται για ενημέρωση τροχιών με κάποιο id τα οποία ενώ έχουν γίνει tracked δεν εντοπίζονται σε αυτό το frame. Αν αυτό συνεχίζει για πάνω από 20 frames συνεχόμενα ή ο λόγος που το ψάρι είναι ορατό σε σχέση με το συνολικό αριθμό των detection είναι μικρότερος του 0,6 τότε σημαίνει ότι το ψάρι εξαφανίζεται από την εικόνα οπότε και διαγράφουμε το id.

➤ Μη αντιστοιχισμένα detections. Πρόκειται για detections που ξεκινάνε νέα tracks. Για να θεωρηθούν βέβαια σαν τροχιές πρέπει να υπερβούν το όριο των 8 συνεχόμενων frames.

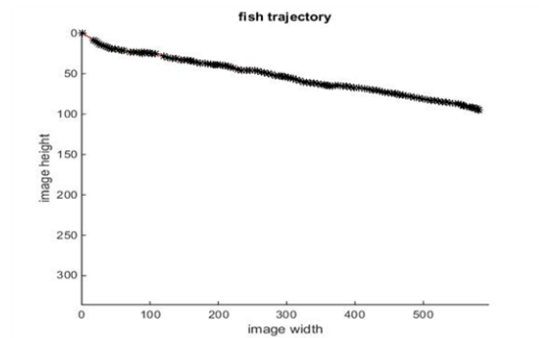
Τελικό βήμα είναι η οπτικοποίηση των αποτελεσμάτων. Για την εξαγωγή των τροχιών εμφανίζουμε στο επίπεδο τις εικόνες τα αποτελέσματα ξεχωριστά για κάθε id όπου απεικονίζουμε τις θέσεις των κέντρων στις διαστάσεις της εικόνας για κάθε frame. Με “\*” φαίνονται τα σημεία της τροχιάς ενώ η έγχρωμη ενιαία γραμμή που ξεκινά από την αρχή των αξόνων μας δείχνει που ξεκινά η τροχιά με αποτέλεσμα να καταλαβαίνουμε και την κατεύθυνση του ψαριού. Με τον τρόπο αυτό ελέγξαμε την ορθότητα των αποτελεσμάτων μας. Στην Εικόνα 3.10 φαίνονται κάποια ενδεικτικά frames για το πώς υπολογίζεται η τελική τροχιά για συγκεκριμένο id, και στην Εικόνα 3.11 φαίνονται τροχιές που έχουμε εξάγει με επιτυχία.



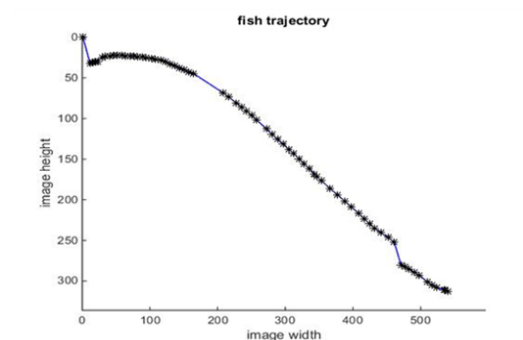
**Εικόνα 3.10** (α),(β),(γ) Αντιστοιχούν στα frame 201,239,279 του Video\_1. (δ) Τροχιά που αντιστοιχεί στο id=31 για όλη τη διάρκεια του βίντεο.



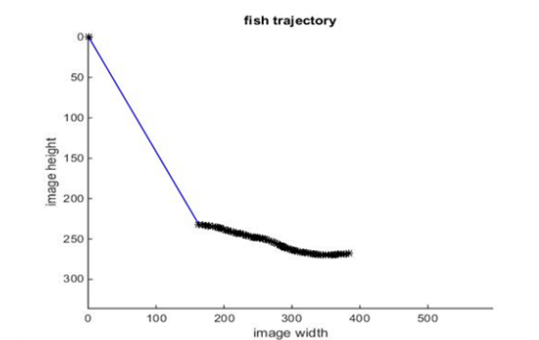
(α)



(β)

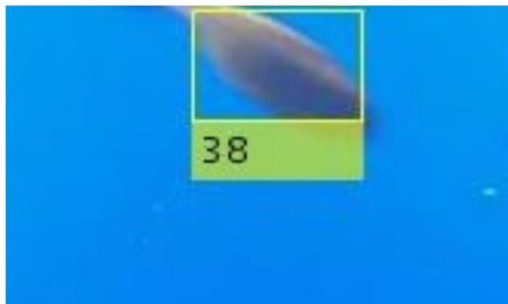


(γ)

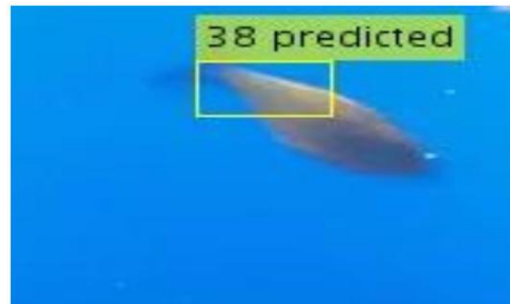


(δ)

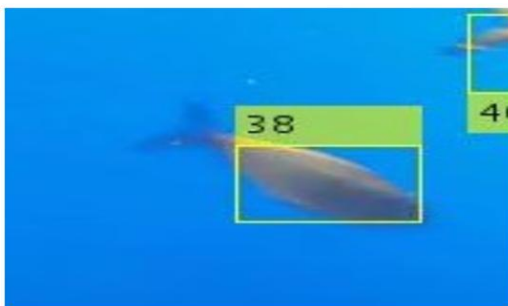
**Εικόνα 3.11** Παραδείγματα από σωστές τροχιές: (α) id=1 από video\_1.(β) id=42 από video\_2 (γ) id=20 από video\_1 (δ) id=2 από video\_2



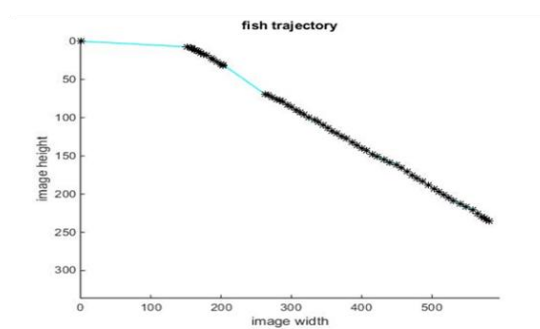
(α)



(β)



(γ)



(δ)

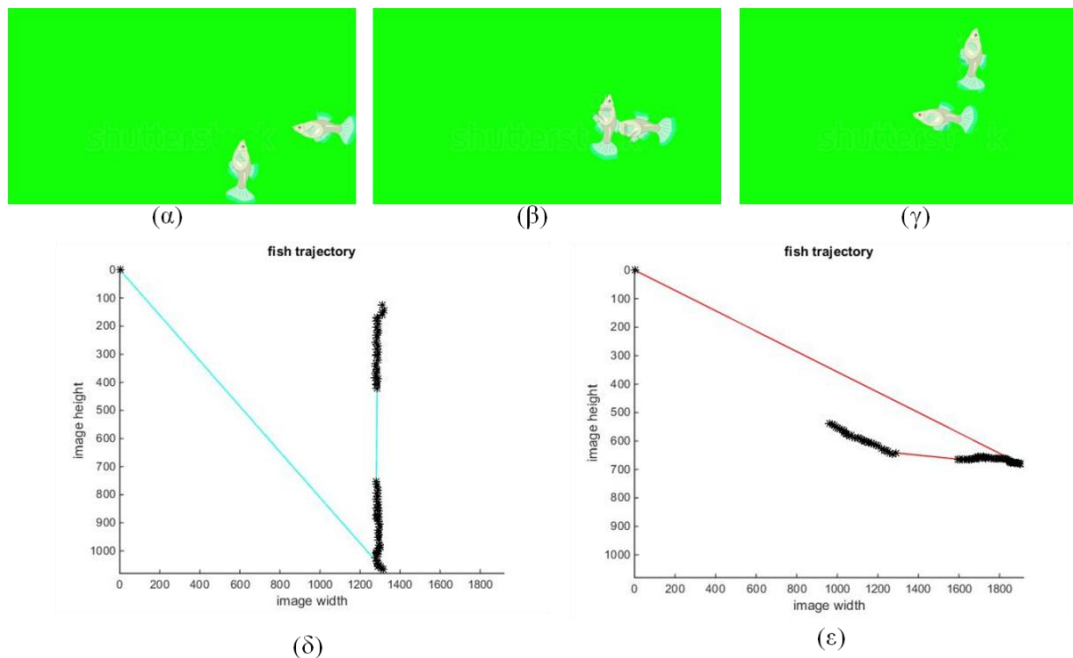
**Εικόνα 3.12** (α) Ο αλγόριθμος εντοπίζει το ψάρι με id=38. (β) το ψάρι χάνετε κατά το detection για 7 frames. (γ) Γίνεται ξανά εντοπισμός του ψαριού και συνεχίζει με το ίδιο id.(δ) Ολόκληρη η τροχιά του ψαριού.

### 3.3.3 Ειδικές περιπτώσεις και προβλήματα

Για να μελετήσουμε και να μπορέσουμε να εξάγουμε καλύτερα συμπεράσματα για τη λειτουργία της μεθοδολογίας μας, κρίθηκε σκόπιμο να εξετάσουμε κάποιες ειδικές περιπτώσεις που αφορούν κυρίως στις επικαλύψεις μεταξύ των ψαριών, στον αριθμό των ψαριών που εμφανίζονται και στην εναλλαγή κατεύθυνσης και ταχύτητας ενός ψαριού. Η προσέγγιση γίνεται στα ψάρια και την κίνηση τους ενώ οι συνθήκες είναι ιδανικές όσον αφορά στο φόντο και τη σταθερότητα της κάμερας.

#### 1<sup>η</sup> περίπτωση

Για να εξετάσουμε την επικάλυψη δημιουργήσαμε βίντεο με δύο μόνο ψάρια τα οποία και επικαλύπτονται πλήρως για 27 frames και μετά συνεχίζουν κανονικά την πορεία τους. Παρατηρούμε ότι θέτοντας με 100 τον αριθμό που συνεχίζει το prediction χωρίς να διαγραφεί, τα ids μετά την επικάλυψη ξαναβρίσκουν τα ίδια ψάρια όπως φαίνεται και στην Εικόνα 3.13.

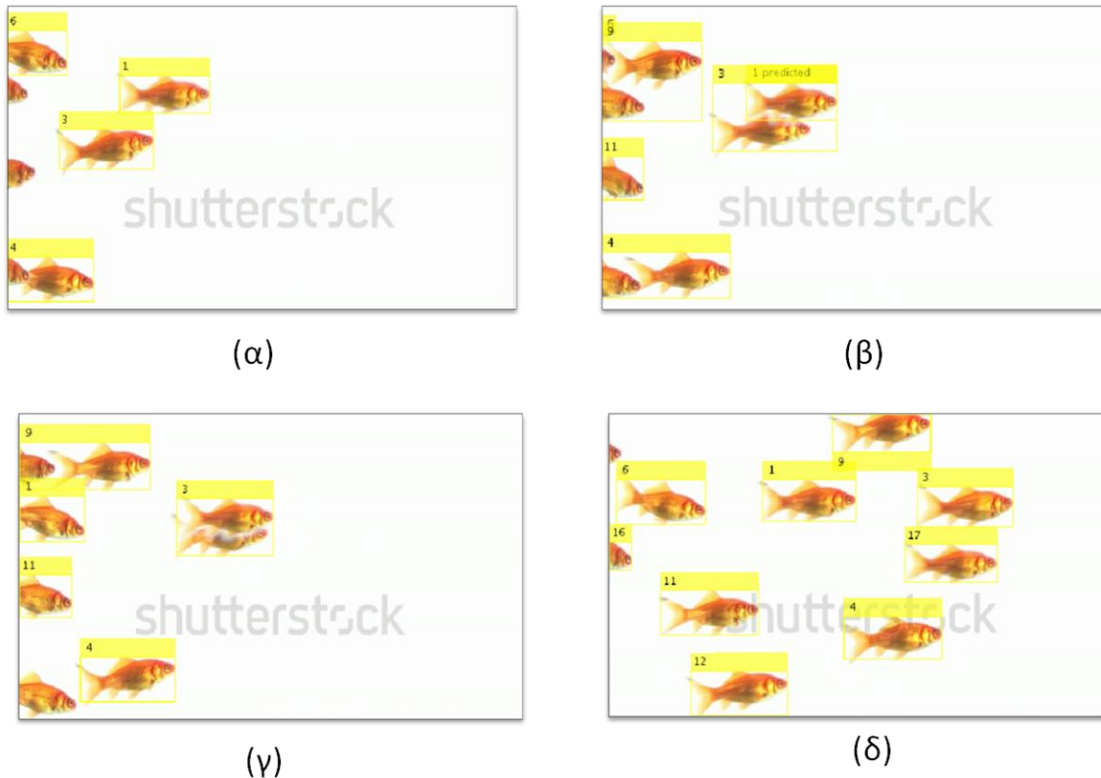


**Εικόνα 3.13** (α),(β),(γ) δείχνουν την επικάλυψη μεταξύ των ψαριών. (δ),(ε) δείχνουν τις τροχιές με id=1 και id=2 αντίστοιχα.



## 2<sup>η</sup> περίπτωση

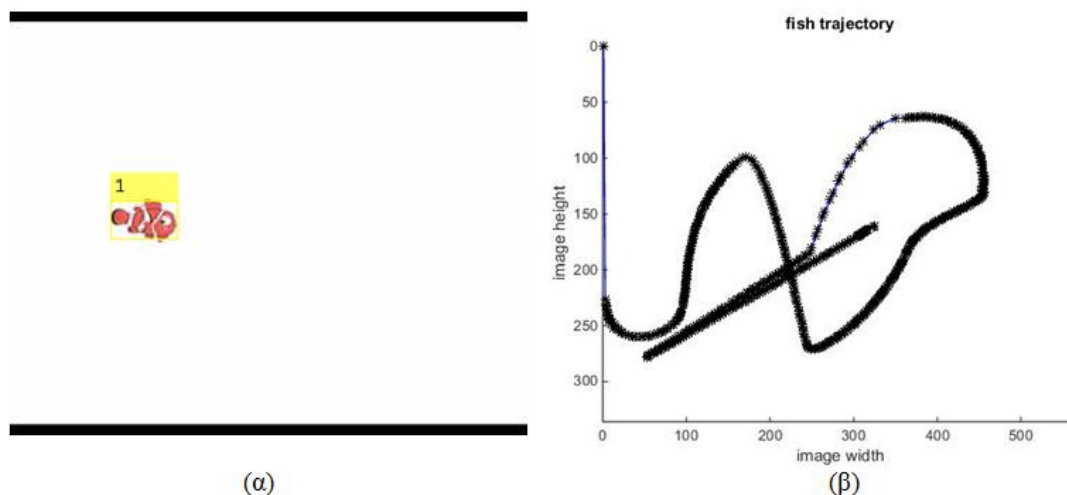
Θέλοντας να εξετάσουμε την επικάλυψη όταν έχουμε πάνω από δύο ψάρια στο βίντεο δημιουργήσαμε ένα βίντεο όπου έχουμε 10 ψάρια με σταθερή πορεία και 2 από αυτά επικαλύπτονται για 40 frames. Αν θέσουμε τη μεταβλητή πάλι με 100 frames για τη συνέχεια του prediction τότε, όταν εμφανίζεται νέο ψάρι στην εικόνα, το id με το prediction πάει σε αυτό. Αυτή η παρατήρηση είναι πολύ σημαντική καθώς είναι ένα από τα σημαντικότερα προβλήματα σημαντικότερα προβλήματα όσο αναφορά την επικάλυψη. Στην Εικόνα 3.14 φαίνονται τα ενδεικτικά frames της περίπτωσης αυτής.



**Εικόνα 3.14** (α) Τα id=1 και id=3 (β) Τα id =1 και id=3 επικαλύπτονται πλήρως και το id=1 συνεχίζει ως πρόβλεψη. (γ) Το id=1 αντιστοιχεί πλέον σε νέο ψάρι που έχει εμφανιστεί. (δ) το id=1 πλέον μετά την επικάλυψη συνεχίζει με νέο id=17.

### 3<sup>η</sup> περίπτωση

Δημιουργήσαμε βίντεο με ψάρι που η κίνηση του είναι αρκετά αλλόκοτη και έχει εναλλαγές στη ταχύτητα και στην κατεύθυνση. Η τροχιά αυτή αποτέλεσε κύριο μέτρο σύγκρισης για τη μοντελοποίηση και εξαγωγή συμπερασμάτων για τις ομαλές και μη τροχιές που θα δούμε σε επόμενο κεφάλαιο. Στην Εικόνα 3.15 φαίνεται η τροχιά του ψαριού.



**Εικόνα 3.15** (α) frame 150 του ψαριού με id=1.(β) Η τροχιά του για 600 frames .

Συμπερασματικά να πούμε ότι ο αλγόριθμος σε πολύ άπλες συνθήκες έχει βέλτιστα αποτελέσματα. Οι δοκιμές στα παραπάνω βίντεο βοήθησαν στο να εξάγουμε συμπεράσματα για τον αλγόριθμο μας και να μπορέσουμε να δούμε τι είδους σφάλματα μπορεί να προκύψουν σε χειρότερες συνθήκες. Επόμενο βήμα είναι η εφαρμογή του αλγορίθμου σε πιο σύνθετες συνθήκες που θα κλιθούμε να αντιμετωπίσουμε και σε περιβάλλον ιχθυοκαλλιέργειας.

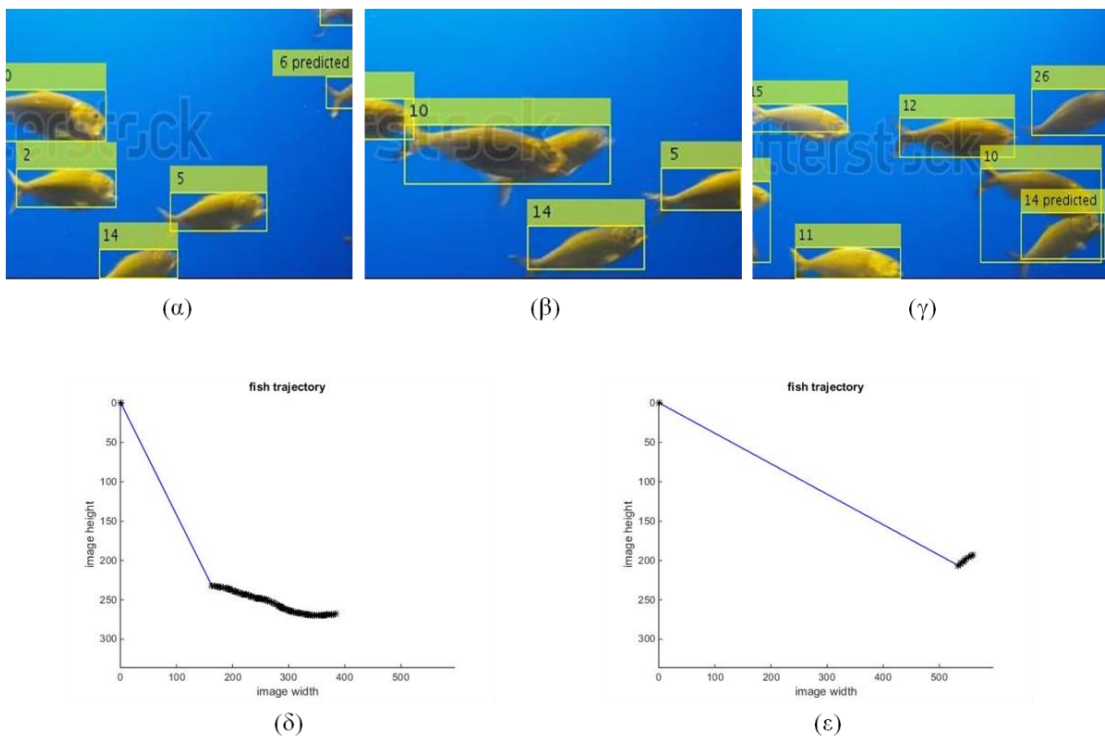
Είναι σημαντικό όμως να αναφερθεί ότι εξετάσαμε τη μεθοδολογία μας σε τρία (3) βίντεο όπου είχαμε να αντιμετωπίσουμε εναλλαγές σε φωτισμό, μεγάλο αριθμό από ψάρια και πολλές επικαλύψεις προσπαθώντας να καλύψουμε όσο πιο καλά γίνεται την πραγματική εικόνα της εφαρμογής. Στην Εικόνα 3.3 φαίνονται ενδεικτικά frames από αυτά τα βίντεο.

Δεδομένου ότι η παρακολούθηση της τροχιάς πολλών στόχων ταυτόχρονα είναι μια διαδικασία αρκετά δύσκολη και οι συνθήκες στις οποίες εξάγουμε αποτελέσματα και καλούμαστε να μελετήσουμε τη συμπεριφορά των ψαριών είναι ιδιαίτερα δύσκολες, η στοχοθέτηση της παρούσας διπλωματικής εργασίας ήταν να εξαχθούν ακριβή συμπεράσματα ως προς τις κατευθύνσεις της μεθοδολογίας που πρέπει να ακολουθήσουμε. Κοιτώντας λοιπόν τα αποτελέσματα μετά και την εξαγωγή των τροχιών, παρατηρούμε ότι παραμένουν κάποια ζητήματα που πρέπει να διερευνηθούν περαιτέρω και να επιλυθούν όπως:

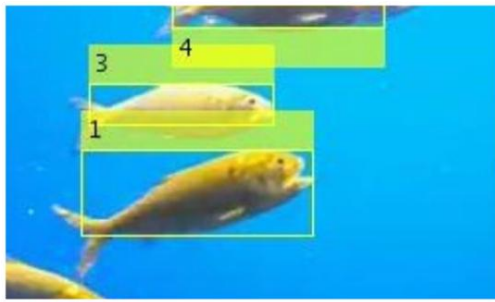
- Πολλά ψάρια που χάνονται για κάποιο διάστημα στην εικόνα εμφανίζονται με διαφορετικό id μετά. Αυτό είναι πρόβλημα για τη μελέτη της συμπεριφοράς ενός ψαριού καθώς μπορούμε να έχουμε μερική εκτίμηση για την κίνηση του. Είναι κάτι που επιλέξαμε να μη διορθώσουμε καθώς δεν αποτελεί μεγάλο πρόβλημα

αφού η ομαλή ή όχι συμπεριφορά θα υπάρχει πριν και μετά την επικάλυψη (Εικόνα 3.16). Σε αντίθετη περίπτωση θα ‘θυσιάσαμε’ μέρος της επεξεργαστικής ισχύος.

- Ένα πρόβλημα που συναντάμε συχνά αφορά σε τροχιές από διαφορετικά ψάρια που εμφανίζονται με το ίδιο id λόγω σφαλμάτων στο tracking (Εικόνα 3.17).
- Να εμφανίζεται και η περίοδος επικάλυψης ως μέρος της τροχιάς του ψαριού (Εικόνα 3.17).
- Τιμές που δεν αντιστοιχούν στην τροχιά, είτε από σφάλματα στην ανίχνευση, είτε από επικαλύψεις (Εικόνα 3.18).



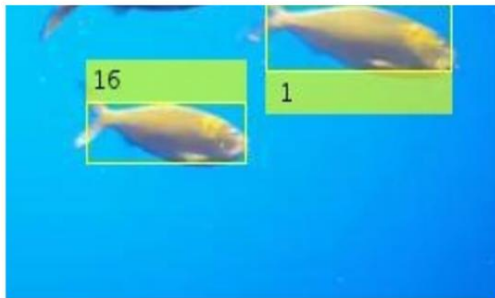
**Εικόνα 3.16** Παράδειγμα εμφάνισης ίδιου ψαριού με διαφορετικά id: (α) το ψάρι γίνεται tracked με id=2. (β) Γίνεται επικάλυψη με το ψάρι με id=10. (γ) Γίνεται ξανά tracked με διαφορετικό id, το id=26. (δ),(ε) φαίνονται οι τροχιές για id=2 και id=26 αντίστοιχα.



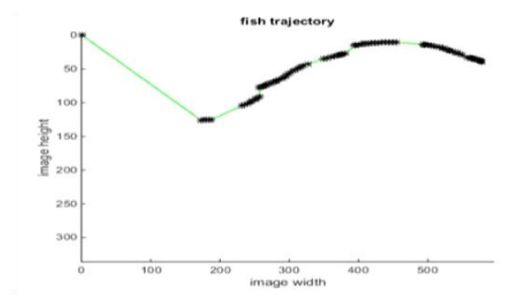
(α)



(β)



(γ)

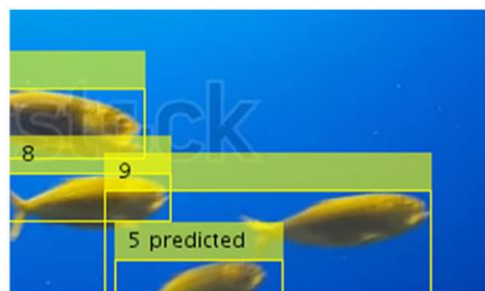


(δ)

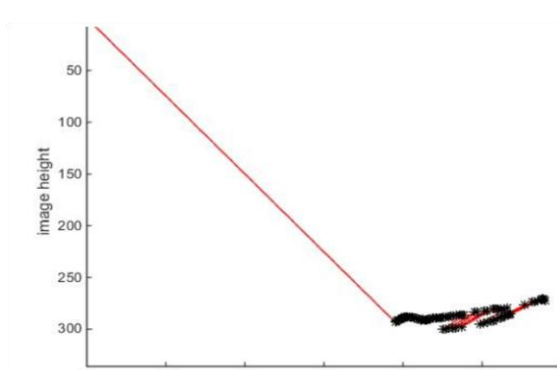
**Εικόνα 3.17** Παράδειγμα επικάλυψης και διατήρηση ίδιας τροχιάς σε διαφορετικό ψάρι:(α),(β) παρατηρούμε το ψάρι με id=1 επικαλύπτεται με το ψάρι με id=3.(γ) συνεχίζει με id=1 διαφορετικό ψάρι εφόσον το προηγούμενο χάνεται από την εικόνα. (δ) Τροχιά για το id=1 όπου περιέχει τιμές από τα 2 ψάρια και από την επικάλυψη μεταξύ τους.



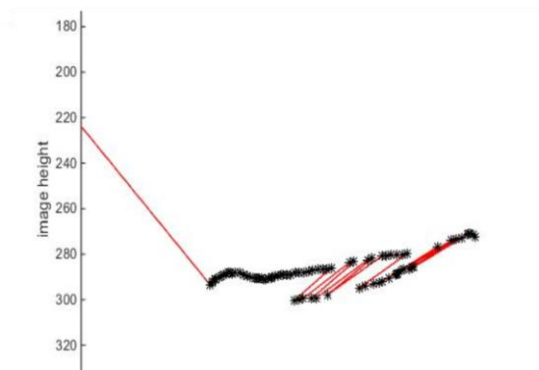
(α)



(β)



(γ)



(δ)

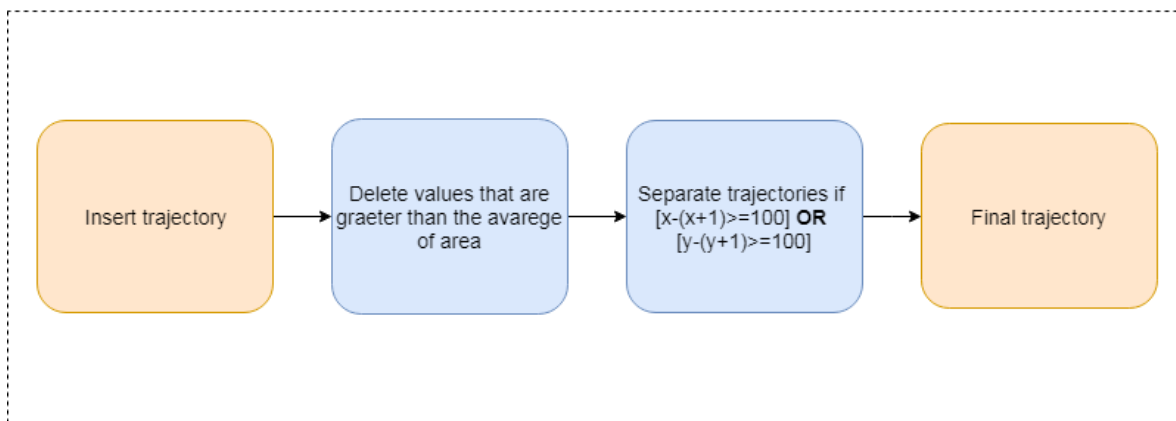
**Εικόνα 3.18** Παράδειγμα από σφάλματα στο detection λόγω στιγμιαίων επικαλύψεων στη τροχιά του ψαριού. (α),(β) Τα frames που γίνονται στιγμιαίες επικαλύψεις. (γ) Η τελική τροχιά μαζί με τα σφάλματα. (δ) Μεγέθυνση τροχιάς.

### 3.4 Επεξεργασία τροχιών

Η προσέγγισή μας στην επίλυση των παραπάνω προβλημάτων γίνεται με βάση την επεξεργασία του κάθε id ξεχωριστά. Σκοπός είναι να διορθωθούν σφάλματα τα οποία με βάση την κίνηση των ψαριών δεν δικαιολογούνται. Είναι σημαντικό να κατανοηθεί ότι με αυτή την επεξεργασία των τροχιών δεν θέλουμε να διώξουμε τροχιές που δεν είναι φυσιολογικές, αλλά να μπορέσουμε να διορθώσουμε τα σφάλματα που προκύπτουν στην ανίχνευση και τον εντοπισμό.

#### 3.4.1 Επίλυση προβλημάτων

Η επίλυση αυτών των σφαλμάτων γίνεται με την επεξεργασία των τροχιών. Τροχιά θεωρείται ένας πίνακας με συντεταγμένες του κέντρου βάρους για κάθε ξεχωριστό id και για την ακολουθία των frames στα οποία εντοπίζεται το συγκεκριμένο id όπως είδαμε και παραπάνω (Εικόνα 3.10). Τις τροχιές που δεν είναι σωστές εξ αρχής, τις επεξεργαζόμαστε όπως φαίνεται στο σχηματικό διάγραμμα Εικόνα 3.19. Στο σημείο αυτό αξίζει να αναφερθεί ότι η επεξεργασία αυτή δε γίνεται σε όλες τις τροχιές αλλά μόνο σε εκείνες που έχουν σφάλματα καθώς χάνεται πολύτιμη πληροφορία αν εφαρμοστεί καθολικά.



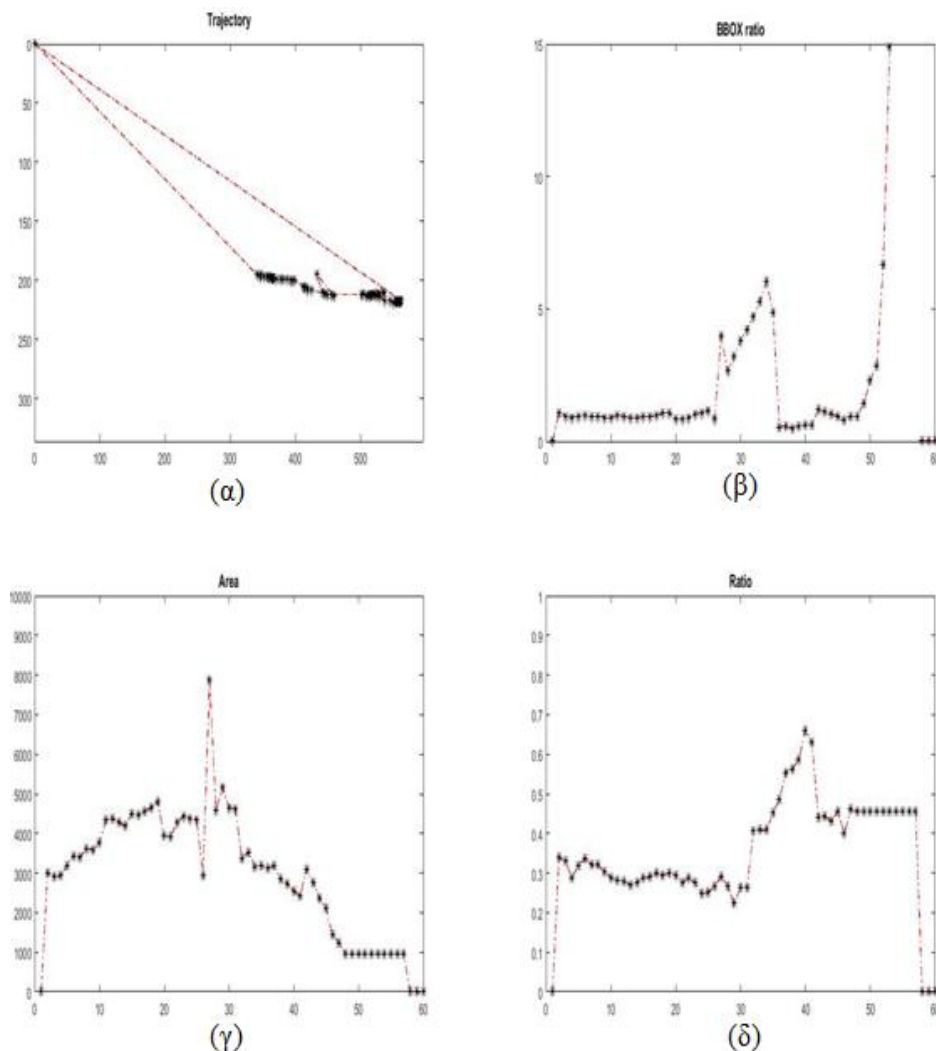
**Εικόνα 3.19** Σχηματικό διάγραμμα επεξεργασίας τροχιών.

Για να μπορέσουμε να διορθώσουμε αυτά τα προβλήματα βασιστήκαμε στην εξαγωγή πληροφοριών από το blob analysis. Αυτές οι πληροφορίες αφορούν στο εμβαδόν του bounding box, την αναλογία ύψους και πλάτους και την αναλογία άσπρων και μαύρων τιμών του bounding box, σε κάθε τιμή της τροχιάς (Εικόνα 3.20). Παρατηρώντας τα αποτελέσματα σε μια σειρά από δοκιμές είδαμε ότι το εμβαδόν του bounding box μπορεί να μας δώσει πολύ χρήσιμη πληροφορία καθώς σε περιπτώσεις επικαλύψεων το μέγεθός του

αλλάζει πολύ απότομα. Οπότε, θέτοντας σαν όριο το μέσο όρο του εμβαδού για όλη την τροχιά, αφαιρούμε τις τιμές εκείνες που είναι μεγαλύτερες από αυτό. Στις Εικόνες 3.21 και 3.22 φαίνονται τέτοιες περιπτώσεις. Αξίζει να παρατηρήσουμε τη μεγάλη διαφορά που έχει το εμβαδόν για τις τιμές επικάλυψης στην εικόνα (3.21.(β), 3.22.(β)) και πως εξομαλύνονται μετά (εικόνες 3.21.(δ), 3.22.(δ)).

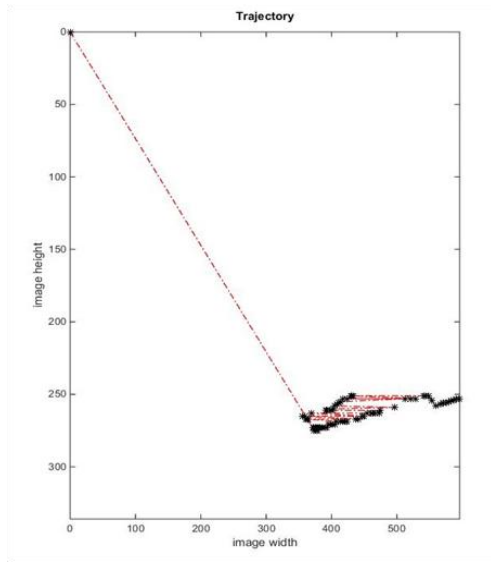
Επίσης το πρόβλημα που αφορά στην ανάθεση του ίδιου id σε διαφορετικό ψάρι μπορέσαμε να το λύσουμε πολύ εύκολα με την παρατήρηση της ακολουθίας των τιμών της τροχιάς. Αν μία από τις δύο συντεταγμένες της τροχιάς, σε σχέση με την προηγούμενη τιμή της, ξεπερνά την τιμή 100 τότε διαχωρίζουμε αυτές τις τροχιές όπως φαίνεται στον ακόλουθο τύπο με x,y τις συντεταγμένες της τροχιάς (Εικόνα 3.23, Εικόνα 3.24):

$$[x-(x+1)>=100] \text{ OR } [y-(y+1)>=100] \quad (3.18)$$

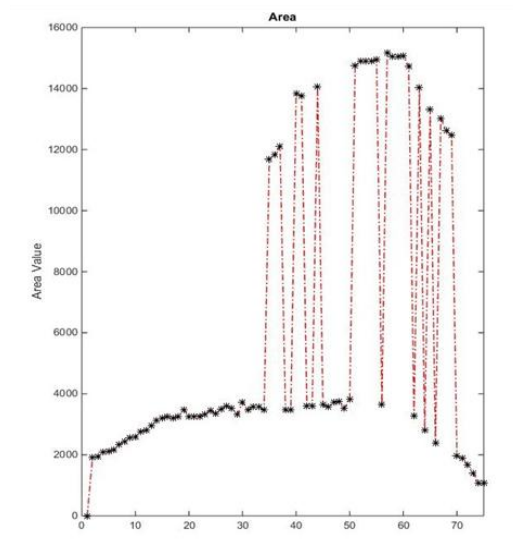


**Εικόνα 3.20** (α) Τροχιά ψαριού. (β) Αναλογία άσπρων και μαύρων pixels σε ένα bbox για κάθε σημείο της τροχιάς. (γ) Εμβαδόν bbox για κάθε σημείο της τροχιάς. (δ) Αναλογία ύψους-πλάτους του bbox για κάθε σημείο της τροχιάς.

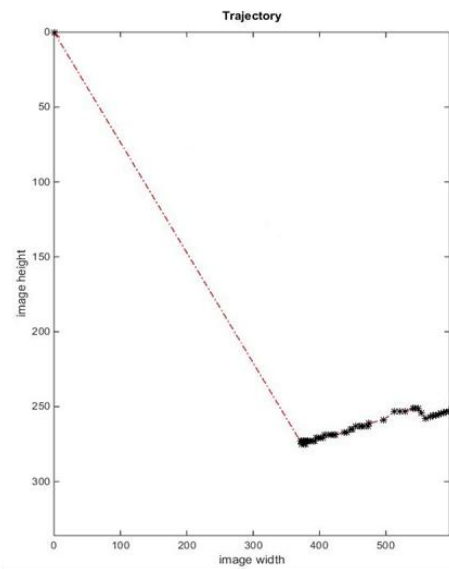




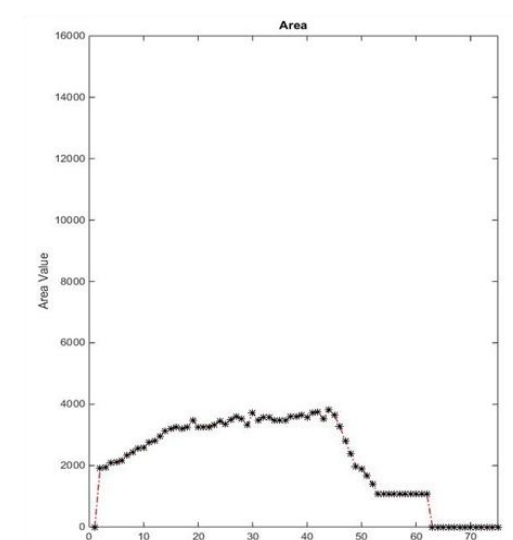
(α)



(β)

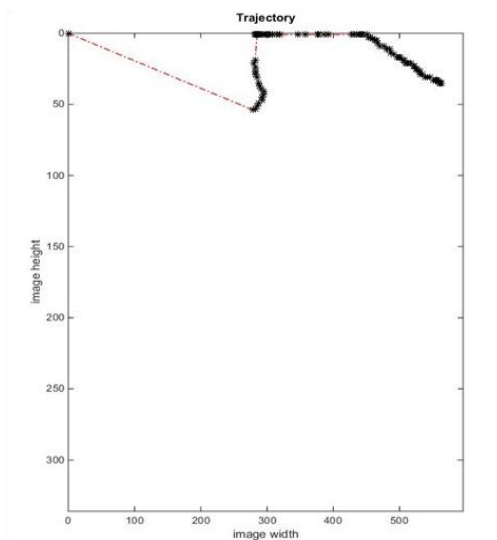


(γ)

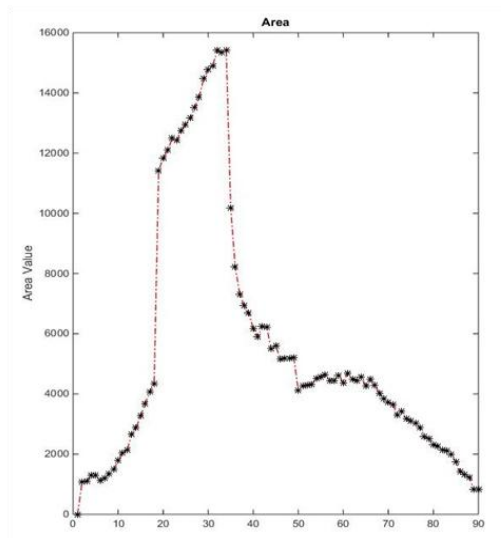


(δ)

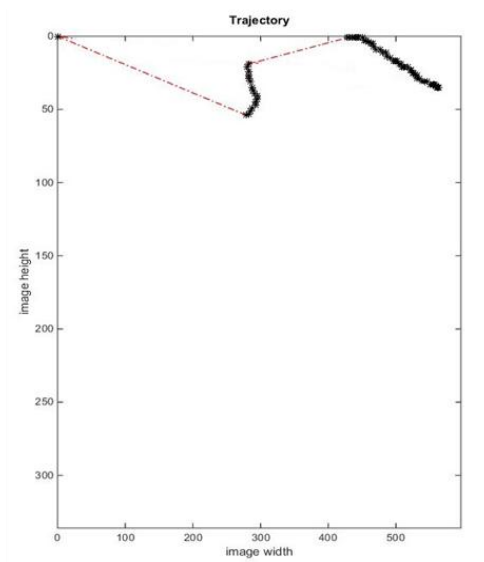
**Εικόνα 3.21** (α) Αρχική εσφαλμένη τροχιά για id=1 που περιέχει και την επικάλυψη (β) Μέγεθος εμβαδού σε κάθε τιμή της τροχιάς (γ) Τροχιά μετά την επεξεργασία (δ) Μέγεθος εμβαδού μετά την επεξεργασία.



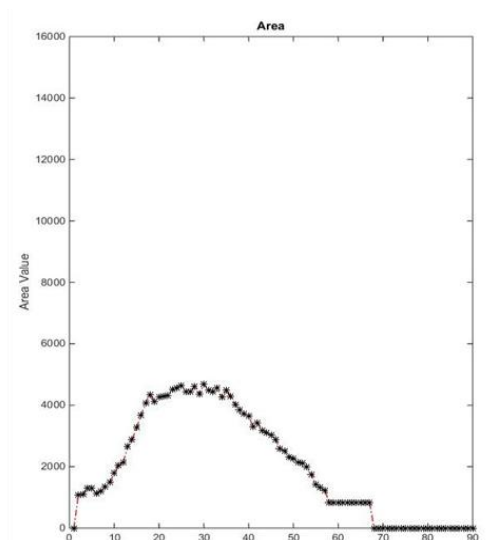
(α)



(β)

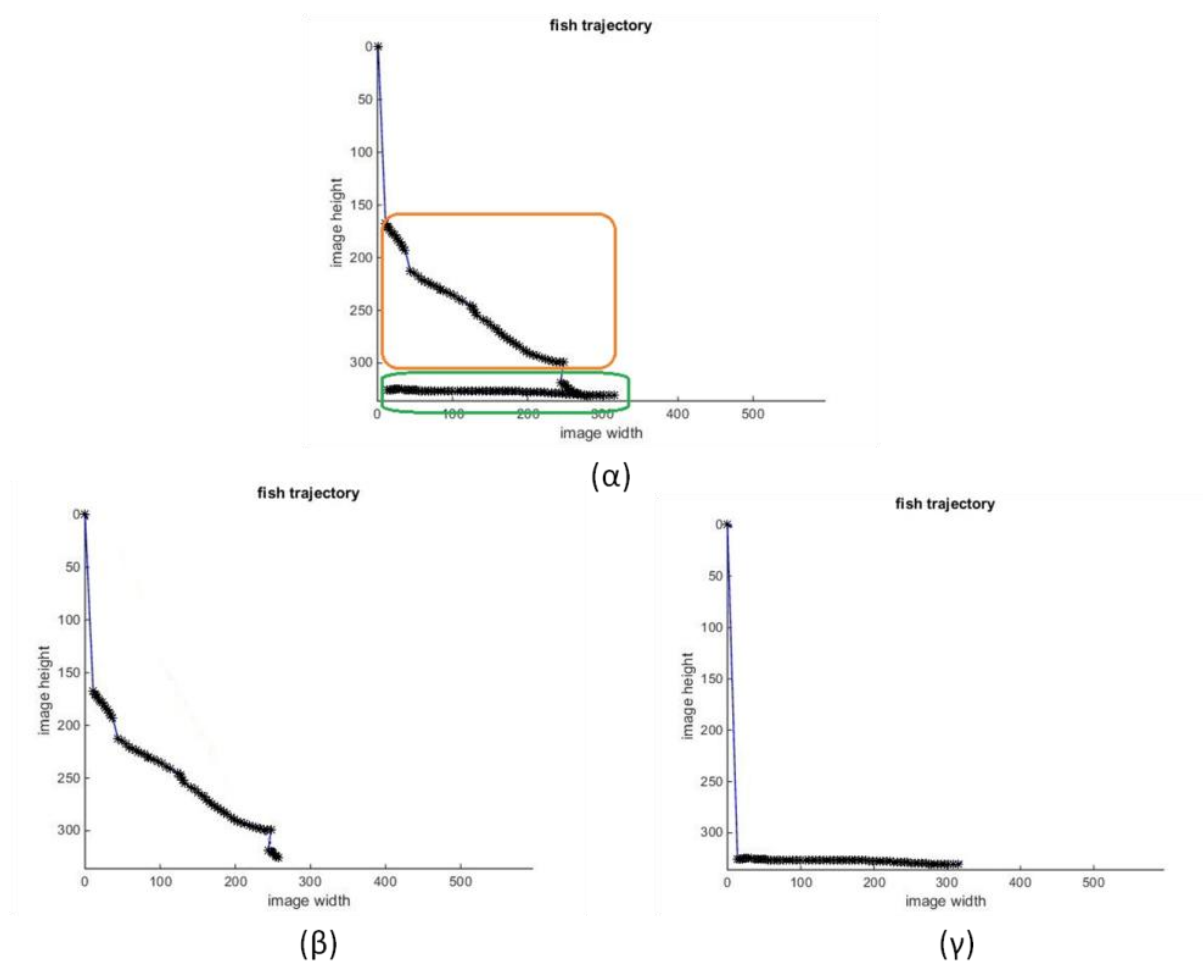


(γ)

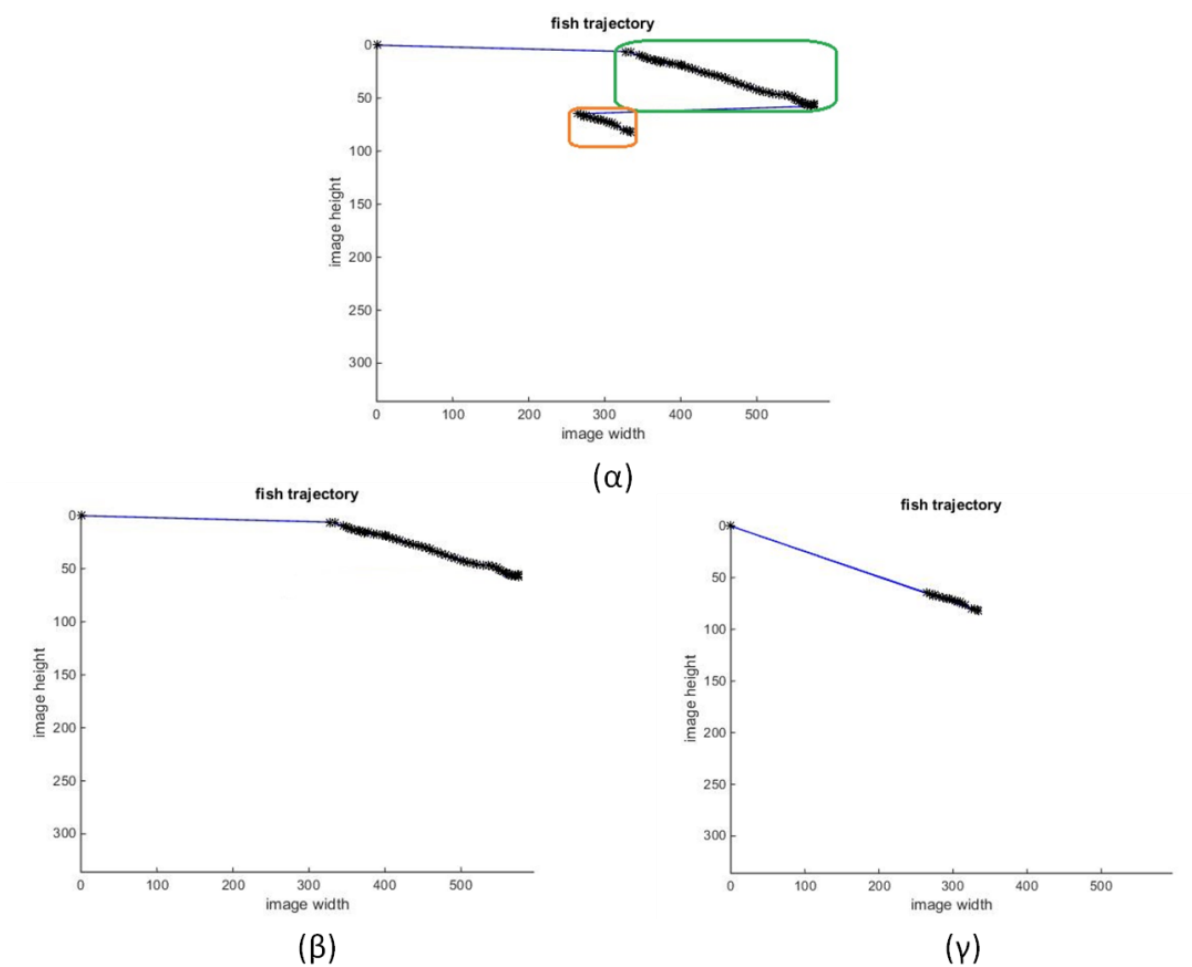


(δ)

**Εικόνα 3.22** (α) Αρχική εσφαλμένη τροχιά (β) Μέγεθος εμβαδού σε κάθε τιμή της τροχιάς (γ) Τροχιά μετά την επεξεργασία (δ) Μέγεθος εμβαδού μετά την επεξεργασία



**Εικόνα 3.23** (α) Αρχική τροχιά για id=40 στο πράσινο και πορτοκαλί πλαίσιο φαίνονται οι δύο τροχιές. (β),(γ) Διαχωρισμένες τροχιές.



**Εικόνα 3.24** (α) Αρχική τροχιά για id=29 στο πράσινο και πορτοκαλί πλαίσιο φαίνονται οι δύο τροχιές. (β),(γ) Διαχωρισμένες τροχιές.

### 3.4.2 Τελικά αποτελέσματα tracking

Στην ενότητα αυτή γίνεται μια προσπάθεια να προσδιορίσουμε την ακρίβεια του tracking πριν και μετά την επεξεργασία των τροχιών ώστε να δούμε τελικά τι αποτελέσματα έχει ο αλγόριθμός μας και ποια η βελτίωση που καταφέρνουμε να έχουμε μετά την επεξεργασία των τροχιών. Γενικότερα η ακρίβεια στο tracking είναι δύσκολο να προσδιοριστεί όπως αναφέρεται και στην εργασία «*Understanding fish behavior during typhoon events in real-life underwater environments*», [25] εφόσον μια τροχιά μπορεί στην αρχή να είναι σωστή και μετά εσφαλμένη. Οπότε με βάση το ζητούμενο της εργασίας μας ορίζουμε την ακρίβεια όσον αφορά το ποσοστό των τροχιών που είναι σωστές και αναφέρονται σε ένα συγκεκριμένο id σε σχέση με αυτές που εντοπίζει συνολικά όπως φαίνεται στον ακόλουθο τύπο.

$$Traj_{Accuracy} = \frac{Num_{true\_traj}}{Num_{traj}} * 100\% \quad (3.19)$$

Όπου  $Num_{traj}$  είναι ο συνολικός αριθμός των τροχιών που αντιστοιχούν σε ψάρια παρόλο που περιέχουν σφάλματα λόγω επικαλύψεων ή άλλων προβλημάτων και μετρήθηκαν ανά βίντεο και  $Num_{true\_traj}$  είναι ο αριθμός των id ψαριών που μετρήθηκαν και μας δίνουν σωστά αποτελέσματα. Σωστό και ακριβές θεωρείται ένα αποτέλεσμα όταν οι τιμές που αντιστοιχούν στη τροχιά ενός id αντιστοιχούν σε ψάρι ανεξάρτητα αν αυτό εντοπιστεί και με διαφορετικά id. Παρακάτω παρουσιάζονται τα αποτελέσματα του αλγορίθμου μας. Στον Πίνακα 3.3 φαίνονται τα αρχικά αποτελέσματα, ενώ στον Πίνακα 3.4 φαίνονται τα αποτελέσματα μετά την επεξεργασία των τροχιών.

Βίντεο	Frames	Όλες οι τροχιές	Σωστές Τροχιές	Accuracy
Video_1	440	37	15	40%
Video_2	325	50	33	66%
Video_3	636	47	14	30%
Video_4	378	14	12	85%

**Πίνακας 3.3** Ακρίβεια tracking πριν την επεξεργασία των τροχιών.

Βίντεο	Frames	Όλες οι τροχιές	Σωστές Τροχιές	Accuracy
Video_1	440	37	27	59%
Video_2	325	50	47	94%
Video_3	636	47	34	72%
Video_4	378	14	14	100%

**Πίνακας 3.4** Ακρίβεια tracking μετά την επεξεργασία των τροχιών.

Παρατηρούμε ότι με την επεξεργασία των τροχιών καταφέρνουμε να έχουμε αρκετά βελτιωμένα αποτελέσματα. Βέβαια, όπως είναι φυσικό, στα Video\_1 και Video\_3 έχουμε τα χειρότερα αποτελέσματα καθώς όπως είπαμε και παραπάνω είναι τα βίντεο με τις πιο δύσκολες συνθήκες. Ορισμένα σφάλματα δεν διορθώνονται παρόλη την επεξεργασία. Στα Video\_2 και Video\_4 τα αποτελέσματα είναι πολύ καλά, κάτι το οποίο οφείλεται στο μικρό αριθμό από ψάρια και τις λίγες επικαλύψεις. Να τονίσουμε ξανά ότι η ακρίβεια υπολογίζεται σε σχέση με την απομόνωση των τροχιών που εντοπίζονται συνολικά.

Συμπερασματικά να πούμε ότι με την μεθοδολογία μας μέχρι αυτό το στάδιο καταφέρνουμε να εξάγουμε τις τροχιές είτε αφορούν σε κοπάδι ψαριών είτε στη μεμονωμένη συμπεριφορά τους. Η χρήση του Kalman φίλτρου με την πρόβλεψη και διόρθωση βοήθησε στο να είναι τα αποτελέσματα αυτά όσο το δυνατόν πιο ομαλά. Με την επεξεργασία των τροχιών μετά την εξαγωγή τους καταφέρνουμε καλύτερα αποτελέσματα και να διορθώσουμε προβλήματα του αλγορίθμου όπως αναλύθηκε παραπάνω. Οπότε όσο αφορά την συμπεριφορά κοπαδιού, αυτό που μπορούμε να εντοπίσουμε, αφορά στο εάν όλα τα ψάρια ακολουθούν την κίνηση του κοπαδιού ή κάποια από αυτά έχουν σημαντική απόκλιση. Αυτό το καταφέρνουμε με την σύγκριση στην περιοχή του bounding box. Από την άλλη οι τροχιές οι οποίες αναφέρονται σε μεμονωμένα ψάρια εξηγούμε στο επόμενο κεφάλαιο τον τρόπο με τις οποίες τις μοντελοποιούμε ώστε να τις χαρακτηρίσουμε ως ομαλές ή όχι.

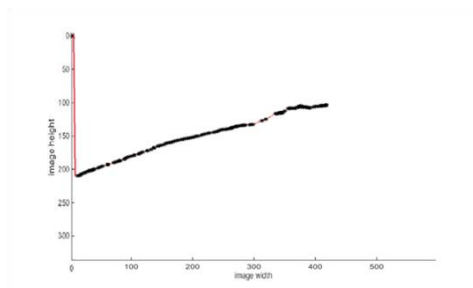


## Κεφάλαιο 4: Μοντελοποίηση τροχιάς

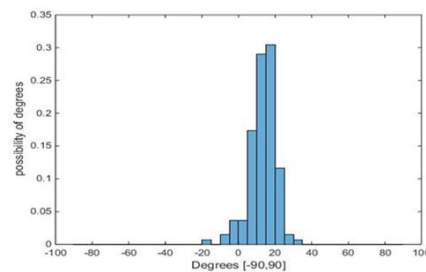
### 4.1 Εισαγωγή

Στο κεφάλαιο αυτό αναλύεται το τελικό στάδιο της εργασίας μας που αφορά στη μοντελοποίηση των τροχιών με σκοπό την εξαγωγή συμπερασμάτων για την κίνηση των ψαριών. Στόχος για μια αυτοματοποιημένη εφαρμογή για την αναγνώριση της κίνησης είναι το σύστημα να μπορεί να εξάγει αποτελέσματα και παρατηρήσεις που έχουμε με το μάτι ή ακόμα καλύτερα και αυτές που δεν αντιλαμβάνεται εύκολα ένας παρατηρητής να μπορεί να τις εξάγει αυτόματα. Για να μπορεί το σύστημα μας να έχει τέτοια λειτουργικότητα, χρειάζεται μαθηματική αναπαράσταση των τροχιών και κάποιες μετρικές που να μπορούν να μας αποκαλύπτουν τι είδους τροχιά έχουμε και αν ή κάθε μία εντάσσεται στις ομαλές ή όχι περιπτώσεις. Εδώ χρειάζεται να σημειωθεί ότι το να χαρακτηριστεί ομαλή ή μη ομαλή μια τροχιά δεν εντάσσεται στα πλαίσια της συγκεκριμένης διπλωματικής (με την έννοια του ορισθέντος θέματος), ανά περίπτωση μπορεί να σημαίνει και κάτι διαφορετικό ως προς το συμπέρασμα. Η ομαδοποίηση που μπορούμε να κάνουμε εμείς βάση παρατηρήσεων στα αποτελέσματα είναι:

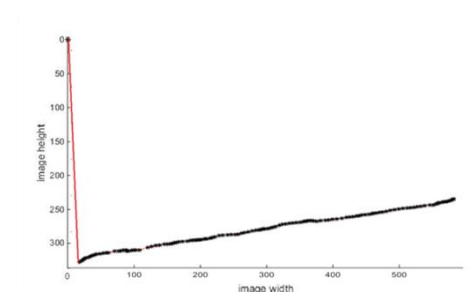
- **Ομαλές** θεωρούμε τις τροχιές που η πορεία τους δεν παρουσιάζει πολλές εναλλαγές στην κλίση. (Εικόνες: 4.1(α),(γ) , 4.2(α),(γ) , 4.3(α),(γ)).
- **Μη ομαλές** αυτές που έχουν πολλές εναλλαγές στην κλίση με αποτέλεσμα να δημιουργούνται πολλές καμπυλώσεις στο διάγραμμα τις τροχιές τους. (Εικόνες 4.4(α),(γ) , 4.5(α),(γ)).



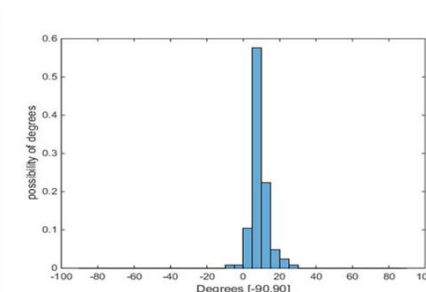
(α)



(β)

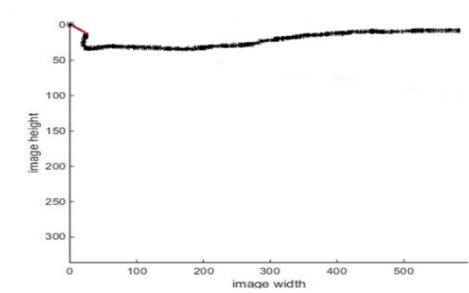


(γ)

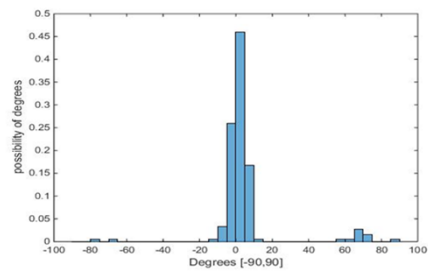


(δ)

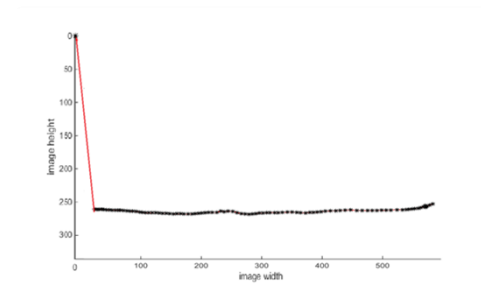
**Εικόνα 4.1** (α) Ομαλή, ανοδική πορεία id=15 από Video\_2. (β) Ιστόγραμμα γωνιών της (α). (γ) Ομαλή, ανοδική πορεία id=20 από Video\_2. (δ) Ιστόγραμμα γωνιών της (γ)



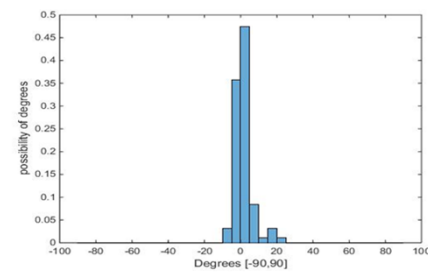
(α)



(β)

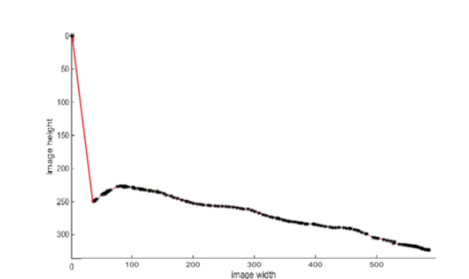


(γ)

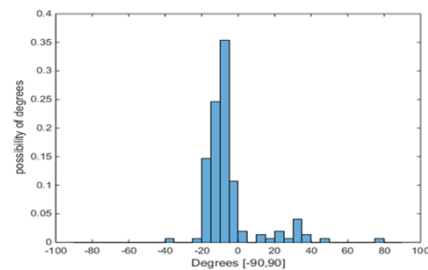


(δ)

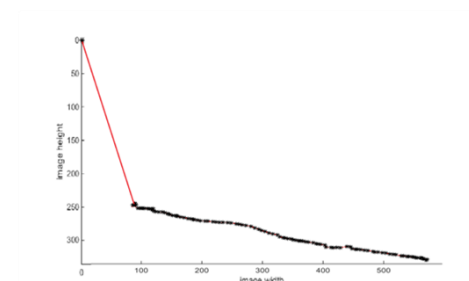
**Εικόνα 4.2** (α)Ομαλή, ευθεία πορεία id=11 από Video\_2.(β) Ιστόγραμμα γωνιών της (α).(γ) Ομαλή, ευθεία πορεία id=31 από Video\_2 (δ) Ιστόγραμμα γωνιών της (γ)



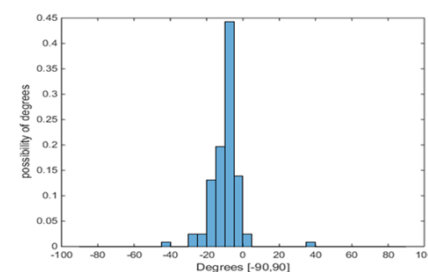
(α)



(β)

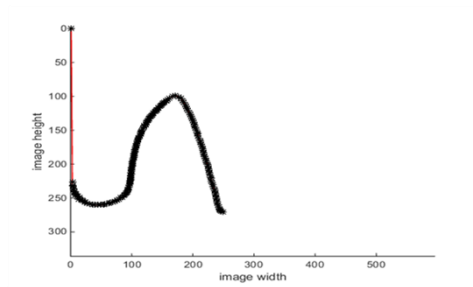


(γ)

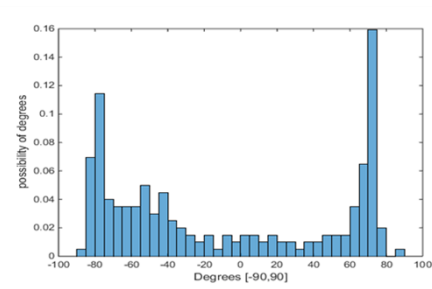


(δ)

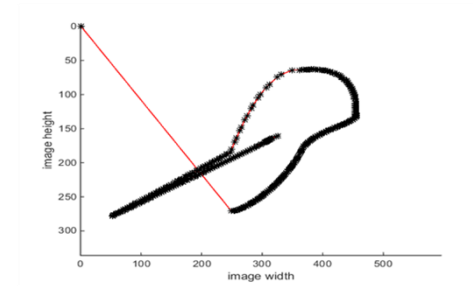
**Εικόνα 4.3** (α)Ομαλή, καθοδική πορεία id=9 από Video\_4.(β) Ιστόγραμμα γωνιών της (α).(γ) Ομαλή, καθοδική πορεία id=13 από Video\_4 (δ) Ιστόγραμμα γωνιών της (γ)



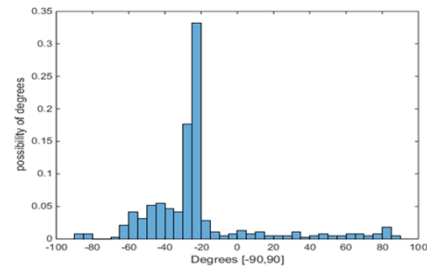
(α)



(β)

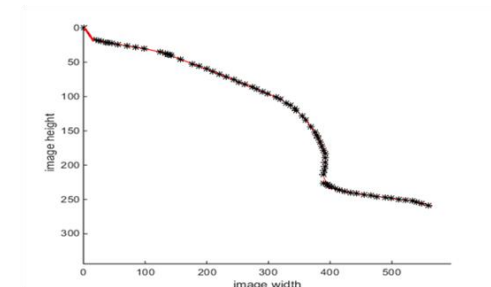


(γ)

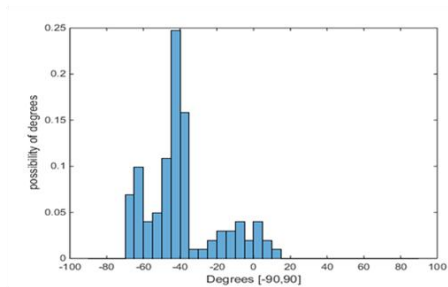


(δ)

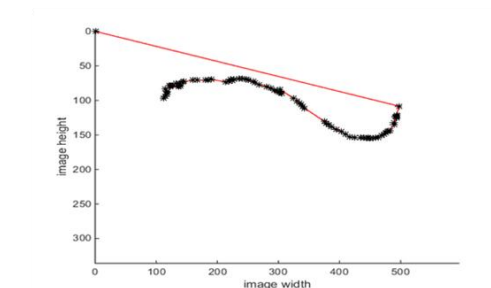
**Εικόνα 4.4** (α) Μη ομαλή πορεία id=1 από Video\_5. (β) Ιστόγραμμα γωνιών της (α). (γ) Μη ομαλή πορεία id=1 από Video\_5 (δ) Ιστόγραμμα γωνιών της (γ).



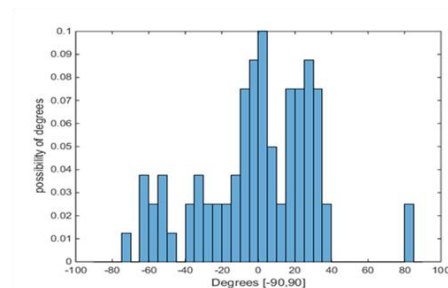
(α)



(β)



(γ)



(δ)

**Εικόνα 4.5** (α) Μη ομαλή πορεία id=2 από Video\_1. (β) Ιστόγραμμα γωνιών της (α). (γ) Μη ομαλή πορεία id=51 από Video\_1 (δ) Ιστόγραμμα γωνιών της (γ).

## 4.2 Μοντελοποίηση με βάση την Εντροπία

Έχοντας λοιπόν σαν αποτέλεσμα τις τελικές τροχιές που αντιστοιχούν στα ψάρια βάση της προηγούμενης μεθοδολογίας, προτείνουμε μια μορφή ανάλυσης όπως παρουσιάζεται με τα παρακάτω βήματα.

### Βήματα αλγορίθμου για την μοντελοποίηση της τροχιάς

#### Βήμα 1<sup>ο</sup>

Εισαγωγή δεδομένων για μια τροχιά.

#### Βήμα 2<sup>ο</sup>

Εύρεση γωνιών μεταξύ των σημείων της τροχιάς.

#### Βήμα 3<sup>ο</sup>

Δημιουργία ιστογράμματος γωνιών.

#### Βήμα 4<sup>ο</sup>

Κανονικοποίηση του ιστογράμματος ως προς την πιθανότητα εμφάνισης της κάθε γωνίας.

#### Βήμα 5<sup>ο</sup>

Εύρεση εντροπίας στο κανονικοποιημένο ιστόγραμμα.

#### Βήμα 6<sup>ο</sup>

Ένταξη σε κάποια από τις δυο ομάδες βάση ορίου στην τιμή της εντροπίας.

Βάση παρατηρήσεων θεωρήσαμε ότι το να βρούμε τη γωνία μεταξύ των σημείων της τροχιάς θα μας έδινε πληροφορία για τον τρόπο που κινείται το ψάρι. Ειδικά, αν πρόκειται για ομαλή τροχιά τότε από το μέσο όρο των γωνιών μπορούμε να βρούμε και την κατεύθυνση του ψαριού. Σε αρχικό στάδιο βρήκαμε τις γωνίες και το ιστόγραμμα από σημείο σε σημείο της τροχιάς μας όπως φαίνεται στις Εικόνες 4.6(β),(γ) και 4.7(β),(γ). Η γωνία μεταξύ δυο σημείων βρίσκεται από την αντίστροφη εφαπτομένη (ή τόξο εφαπτομένης) του ορθογώνιου τριγώνου που σχηματίζεται μεταξύ τους. Για να βρούμε πιο αντιπροσωπευτικά και εξομαλυσμένα αποτελέσματα, υπολογίσαμε το μέσο όρο των γωνιών από τρεις αποστάσεις σε σχέση με το σημείο που εξετάζουμε κάθε φορά, δηλαδή για τις επόμενες 3, 5, και 7 τιμές τις τροχιές. Παρακάτω παρουσιάζονται οι σχέσεις όπως τις υλοποιήσαμε. Με  $K$  θεωρούμε τον πίνακα με τα σημεία της τροχιάς, και  $x, y$  οι συντεταγμένες για το συγκεκριμένο σημείο της τροχιάς που εξετάζουμε:

$$\tan 1 = \frac{K[x, y + 3] - K[x, y]}{K[x + 3, y] - K[x, y]}, \quad \tan 2 = \frac{K[x, y + 5] - K[x, y]}{K[x + 5, y] - K[x, y]}, \quad \tan 3 = \frac{K[x, y + 7] - K[x, y]}{K[x + 7, y] - K[x, y]} \quad (4.1)$$

$$d = \frac{\tan 1 + \tan 2 + \tan 3}{3} \quad (4.2)$$

Τελικά η γωνία  $\theta$  είναι :  $\theta = \tan^{-1} d$  σε rad

$$\text{Την μετατρέπουμε σε μοίρες και τελικά έχουμε: } \theta' = \frac{\theta * 180}{\pi} \quad (4.3)$$

Η αλλαγή αυτή βελτίωσε κατά πολύ τα αποτελέσματά μας όπως φαίνεται και στις εικόνες 4.6(δ),(ε) και 4.7(δ),(ε). Επόμενο βήμα είναι η κανονικοποίηση των τιμών του ιστογράμματος ως προς την πιθανότητα που έχει κάθε γωνία να εμφανιστεί. Ο λόγος που το κάνουμε αυτό είναι για να μπορούμε να συγκρίνουμε από κοινή βάση κάθε τροχιά. Για τη σύγκριση αυτή επιλέγουμε την εφαρμογή της εντροπίας στο ιστόγραμμα των γωνιών. Η έννοια της εντροπίας πληροφοριών παρουσιάστηκε πρώτη φορά από τον Shannon, «*A mathematical theory of communication*», [44]. Παρατηρώντας τα αποτελέσματα, τροχιές που ακολουθούν μια κατεύθυνση και ομαλή πορεία έχουν τη μεγαλύτερη συγκέντρωση των τιμών τους σε συγκεκριμένες γωνίες. Ενώ αντίθετα σε τροχιές με εναλλαγές στην πορεία τους υπάρχει μεγάλη διασπορά στο ιστόγραμμα γωνιών. Αυτό που περιμένουμε από την μέση εντροπία (μέση αβεβαιότητα) του ιστογράμματος είναι από τη τιμή της να ‘καταλαβαίνει’ το σύστημα για τι είδους κίνηση πρόκειται, η μέθοδος αυτή προτείνεται στην εργασία «*Using complexity measures of movement for automatically detecting movement types of unknown GPS trajectories*», [45]. Δηλαδή, μια υψηλότερη τιμή εντροπίας υποδηλώνει ότι τα διαδοχικά δεδομένα είναι μια τυχαία σειρά, οπότε μεγαλύτερη διασπορά, ενώ μια μικρότερη τιμή συνεπάγεται λιγότερη πολυπλοκότητα και μεγαλύτερη κανονικότητα (προβλέψιμο μοτίβο) στα διαδοχικά δεδομένα. Η εντροπία υπολογίζεται από τον παρακάτω τύπο:

$$E = - \sum_{i=1}^n p(x_i) * \log_2 p(x_i) \quad (4.4)$$

Όπου  $p(x_i)$ , είναι η πιθανότητα εμφάνισης της κάθε γωνίας για  $n$  τιμές. Προφανώς οι μηδενικές τιμές δεν συμπεριλαμβάνονται καθώς ο λογάριθμος του μηδέν δεν ορίζεται. Η εντροπία ποσοτικοποιεί την αβεβαιότητα που εμπλέκεται στην πρόβλεψη της τιμής μίας τυχαίας μεταβλητής. Αν είναι ισοπίθανες τότε η τιμή αυτή είναι ίση με αυτόν τον αριθμό. Αν έχω μόνο μια τιμή τότε η τιμή αυτή είναι μηδέν. Η ιδιότητα που εκμεταλλευόμαστε αφορά στην παρατήρηση πολλών διαφορετικών γωνιών σε συνδυασμό με τις πιθανότητες εμφάνισής τους, κάτι που μας δίνει περισσότερη πληροφορία, άρα χρειαζόμαστε και πολλά περισσότερα bits για αναπαράσταση.

Παρακάτω υπολογίζεται η εντροπία των ιστογραμμάτων για ενδεικτικά παραδείγματα ομαλών τροχιών Πίνακας 4.1 και ενδεικτικές μη ομαλές Πίνακας 4.2. Όπως βλέπουμε και στα αποτελέσματα, επιβεβαιώνεται η θεωρία καθώς οι ομαλές τροχιές έχουν τιμές κοντά στο 2 άρα και  $2^2=4$  bits ενώ οι μη ομαλές κοντά στο 4 άρα  $2^4=16$  bits. Είναι αρκετά σημαντική διαφορά, ώστε βάζοντας ένα όριο μπορούμε να έχουμε σημαντική διάκριση στη μορφή κίνησης.

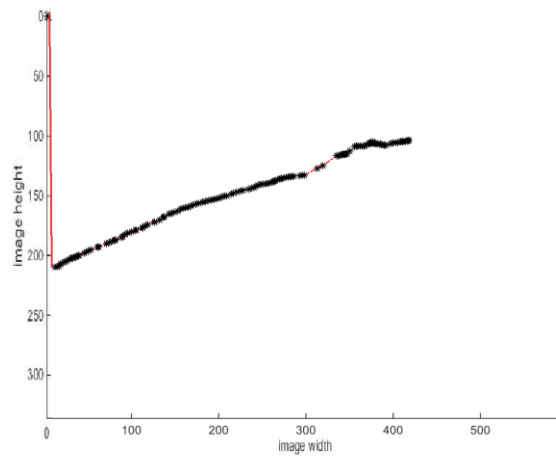
Ιστόγραμμα	Εντροπία (E)	Average of degrees
Εικόνα 4.1 (β)	2,46	12,3°
Εικόνα 4.1 (δ)	1,78	8,11 °
Εικόνα 4.2 (β)	2,13	2,7 °
Εικόνα 4.2 (δ)	1,79	1,32 °
Εικόνα 4.3 (β)	2,64	-3,29 °
Εικόνα 4.3(δ)	2,27	-9 °

**Πίνακας 4.1** Τιμές εντροπίας και κατεύθυνσης για ομαλές τροχιές.

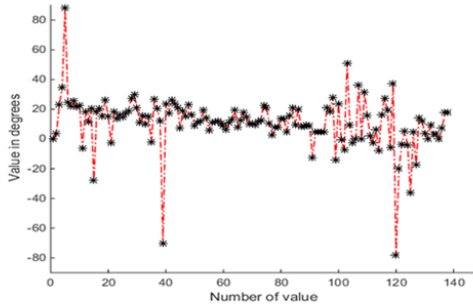
Ιστόγραμμα	Εντροπία(E)
Εικόνα 4.4 (β)	4,52
Εικόνα 4.4 (δ)	3,65
Εικόνα 4.5 (β)	3,77
Εικόνα 4.5(δ)	4,22

**Πίνακας 4.2** Τιμές εντροπίας για μη ομαλές τροχιές.

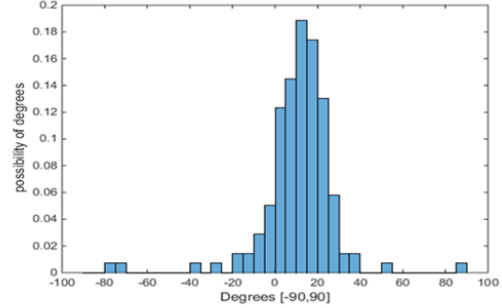




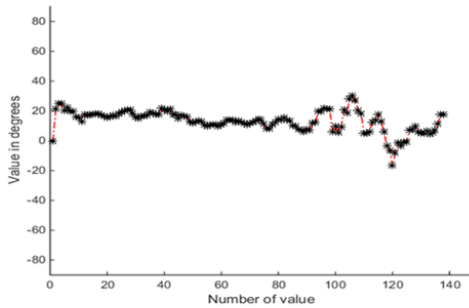
(α)



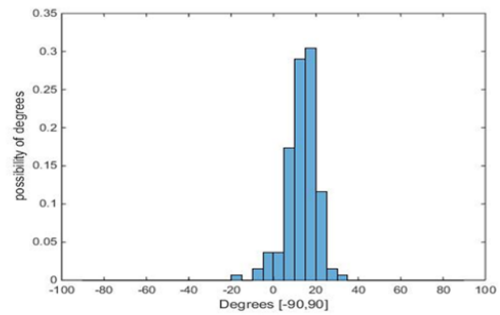
(β)



(γ)

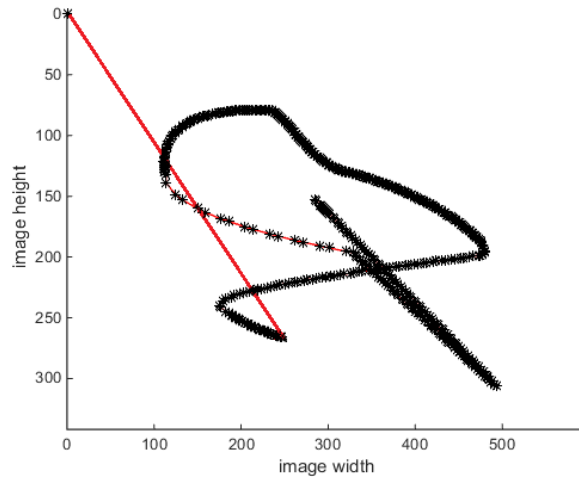


(δ)

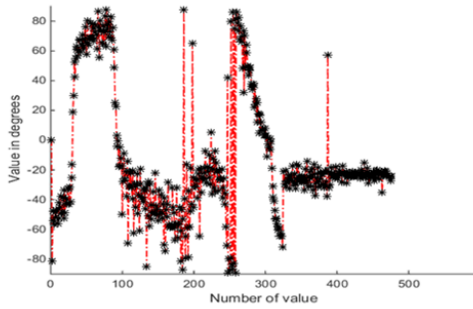


(ε)

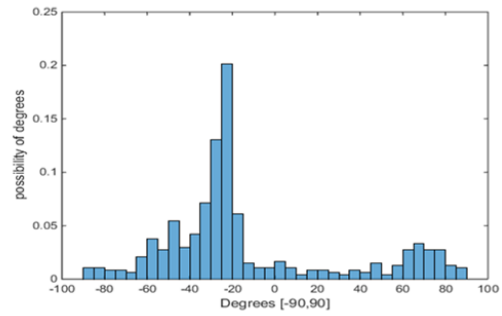
**Εικόνα 4.6** (α) Τροχιά ψαριού με id=15 από Video\_2. (β) Διάγραμμα γωνιών για κάθε σημείο της τροχιάς με βάση το επόμενο σημείο. (γ) Ιστόγραμμα γωνιών του (β). (δ) Διάγραμμα γωνιών με βάση τον μέσο όρο τριών σημείων. (ε) Ιστόγραμμα του (δ).



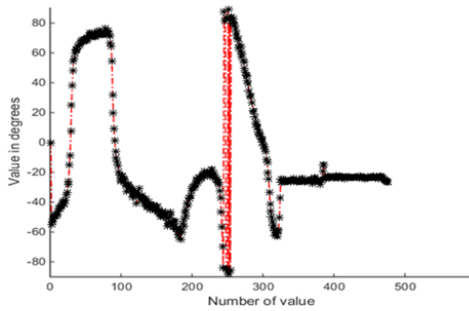
(α)



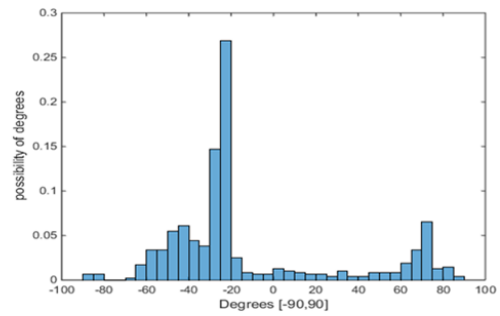
(β)



(γ)



(δ)



(ε)

**Εικόνα 4.7** (α) Τροχιά ψαριού με id=15 από Video\_2. (β) Διάγραμμα γωνιών για κάθε σημείο της τροχιάς με βάση το επόμενο σημείο. (γ) Ιστόγραμμα γωνιών του (β). (δ) Διάγραμμα γωνιών με βάση τον μέσο όρο τριών σημείων. (ε) Ιστόγραμμα του (δ).

## Κεφάλαιο 5: Συμπεράσματα και μελλοντική δουλειά

Στην παρούσα διπλωματική εργασία παρουσιάσαμε ένα ημιαυτόματο σύστημα ανίχνευσης, παρακολούθησης και ανάλυσης της συμπεριφοράς κοπαδιού και ατομικής συμπεριφοράς ψαριών.

Σε αρχικό στάδιο ασχοληθήκαμε με την ανάλυση των προβλημάτων που συναντάμε σε υποθαλάσσιες εικόνες και την αντιμετώπιση και βελτίωση των εικόνων. Στη συνέχεια, βάση της προηγούμενης επεξεργασίας, αναλύσαμε δυο μεθόδους κατάτμησης των εικόνων, το thresholding και τα active contours. Παρότι είχαμε πολύ καλά αποτελέσματα οι μέθοδοι αυτές είναι ευαίσθητοι σε δραστικές αλλαγές στη φωτεινότητα οπότε και σε μια σειρά από πολλά frames δεν υπάρχει αξιοπιστία ως προς το συνολικό αποτέλεσμα. Θετικό βέβαια είναι ότι η μέθοδος των ενεργών περιγραμμάτων μπορεί να δώσει πολύ ακριβή περιγραφή των κατατετμημένων περιοχών οπότε με βελτιώσεις και έχοντας μεγαλύτερη αξιοπιστία θα μπορούσε να είναι βασική μέθοδος κατάτμησης για όλο το σύστημα. Η μέθοδος κατάτμησης που επιλέγουμε στη παρούσα διπλωματική αφορά το Background Subtraction βασισμένο σε Gaussian Mixture Models. Με τη μέθοδο αυτή πετυχαίνουμε πολύ καλά αποτελέσματα αντιμετωπίζοντας εναλλαγές στο φωτισμό και όλα τα υπόλοιπα προβλήματα που έχει μια υποθαλάσσια εικόνα. Στο σημείο αυτό θα ήταν δόκιμο να γίνει σύγκριση και με άλλες μεθόδους background subtraction όπως αναφέρονται και στην εργασία «*Fish detection for species recognition*» [35] αλλά και να εφαρμοστεί και σε διαφορετικές και ρεαλιστικές συνθήκες.

Η ανίχνευση γίνεται με το blob analysis όπου από τις πληροφορίες που παίρνουμε και αφορούν στο κέντρο μάζας και το μέγεθος προχωράμε στο επόμενο στάδιο που είναι η παρακολούθηση της πορείας των ψαριών. Το στάδιο αυτό αποτελεί τη ραχοκοκαλιά κάθε συστήματος παρακολούθησης καθώς από αυτό εξαρτώνται σε μεγάλο βαθμό τα τελικά αποτελέσματα. Η χρήση του φίλτρου Kalman εξυπηρέτησε στο να μπορέσουμε να αντιστοιχίσουμε τα ψάρια με το ίδιο id. Πιστεύουμε ότι μπορούμε να έχουμε πιο βελτιωμένα αποτελέσματα αν για την αντιστοίχιση χρησιμοποιηθούν και άλλα χαρακτηριστικά των ψαριών όπως το σχήμα, το μέγεθος και η ταχύτητα. Με τον τρόπο αυτό θα μπορέσουμε να αντιμετωπίσουμε και προβλήματα επικαλύψεων σε κάποιο βαθμό, ώστε μετά τις επικαλύψεις να ξαναβρίσκουμε τα ίδια ψάρια και να μπορούμε να έχουμε πιο αντιπροσωπευτικά αποτελέσματα για την τροχιά τους.

Επόμενο στάδιο είναι η επεξεργασία των τροχιών που βελτιώνει σε μεγάλο βαθμό τα αποτελέσματά μας δίνοντας τις τελικές τροχιές προς εξαγωγή συμπερασμάτων. Στο στάδιο αυτό υπάρχουν και πιο περίπλοκες μέθοδοι, για πιο ακριβή ανάλυση με σκοπό την ταξινόμηση και την αναγνώριση κινήσεων όπως είναι οι αλυσίδες Markov που προτείνονται και στην εργασία «*trajectory-based activity classification and recognition using hidden Markov models*», [46].

Σε κάθε στάδιο της προτεινόμενης μεθόδου υπάρχουν περιθώρια βελτίωσης. Σε γενικό πλαίσιο, λόγω της έλλειψης ρεαλιστικών δεδομένων στη συγκεκριμένη εργασία,

κινήθηκαμε με άξονα την αντιμετώπιση ενδεχόμενων προβλημάτων και έρευνα σε ειδικές περιπτώσεις δημιουργώντας και δικά μας βίντεο. Σκοπός είναι σε μελλοντική δουλειά ο συνδυασμός των μεθόδων βελτιστοποίησης της εικόνας και η δημιουργία ενός τελείως αυτοματοποιημένου συστήματος που θα έχει αποτελέσματα και θα εφαρμόζεται σε πραγματικές συνθήκες ιχθυοκαλλιέργειας. Επίσης σημαντικό θα είναι το σύστημα αυτό να μπορεί να έχει εφαρμογή και σε real time εφαρμογές εξάγοντας ακόμα περισσότερες πληροφορίες όπως ο αριθμός και τα είδη των ψαριών. Κάτι τέτοιο βέβαια προϋποθέτει να υπάρχουν κατάλληλα βίντεο από πραγματικές συνθήκες και στενή συνεργασία με τους υπευθύνους της μονάδας ιχθυοκαλλιέργειας καθώς η διαμόρφωση των συνθηκών για να πετύχουμε ένα τέτοιο αποτέλεσμα πρέπει να προέρχεται από την εμπειρία των αρμοδίων. Αυτό αποτελεί μεγάλη πρόκληση και θα μπορέσει να λύσει τα χέρια στη βιομηχανία της ιχθυοκαλλιέργειας στα ευρύτερα θέματα αυτοματοποίησης της διαδικασίας παρακολούθησης της ευζωίας των ψαριών.

## Βιβλιογραφία

- [1] Συνέδριο Ιχθυοκαλλιέργειας 2018. (<https://aquaculture.gr>)
- [2] Ελληνική Κτηνιατρική Εταιρία (<http://hvms.gr>)
- [3] Kaawaase, K. S., et al. "A review on selected target tracking algorithms." *Information Technology Journal* 10.4 (2011): 691-702.
- [4] Fan, Litong, et al. "A survey on multiple object tracking algorithm." *Information and Automation (ICIA)*, 2016 IEEE International Conference on. IEEE, 2016.
- [5] Jinan, Rooji, and Tara Raveendran. "Particle Filters for Multiple Target Tracking." *Procedia Technology* 24 (2016): 980-987.
- [6] Hue, Carine, J-P. Le Cadre, and Patrick Pérez. "Tracking multiple objects with particle filtering." *IEEE transactions on aerospace and electronic systems* 38.3 (2002): 791-812.
- [7] Nian, Rui, et al. "ROV-based underwater vision system for intelligent fish ethology research." *International Journal of Advanced Robotic Systems* 10.9 (2013): 326.
- [8] Iqbal, Kashif, et al. "Underwater Image Enhancement Using an Integrated Colour Model." *IAENG International Journal of Computer Science* 34.2 (2007).
- [9] Spampinato, Concetto, et al. "A rule-based event detection system for real-life underwater domain." *Machine vision and applications* 25.1 (2014): 99-117.
- [10] Majeed, Ammar Ibrahim, and Abduladhem Abdulkareem Ali. "Fish School System Identification and Control Based on Artificial Neural Network." *GSTF Journal of BioSciences (JBio)* 3.1 (2018).
- [11] Mandal, Ranju, et al. "Assessing fish abundance from underwater video using deep neural networks." *2018 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2018.
- [12] Tang, Xiaoli, Greg C. Sharp, and Steve B. Jiang. "Fluoroscopic tracking of multiple implanted fiducial markers using multiple object tracking." *Physics in Medicine & Biology* 52.14 (2007): 4081.
- [13] Τσεπαπαδάκης Μιχαήλ. "Αναγνώριση και Χαρακτηρισμός Κίνησης σε Κυκλοφοριακά Βίντεο". Πολυτεχνείο Κρήτης 2014.
- [14] Bera, Aniket, and Dinesh Manocha. "Realtime multilevel crowd tracking using reciprocal velocity obstacles." *Pattern Recognition (ICPR)*, 2014 22nd International Conference on. IEEE, 2014.
- [15] B. Benfold and I. Reid, "Stable multi-target tracking in real-time surveillance video," in CVPR, 2011.
- [16] N. Shoval and M. Isaacson, "Application of tracking technologies to the study of pedestrian spatial behavior\*," *The Professional Geographer*, 2006.
- [17] R. Mehran, A. Oyama, and M. Shah, "Abnormal crowd behavior detection using social force model," in CVPR, 2009.
- [18] Bera, Aniket, and Dinesh Manocha. "Realtime multilevel crowd tracking using reciprocal velocity obstacles." *Pattern Recognition (ICPR)*, 2014 22nd International Conference on. IEEE, 2014.
- [19] J. Van Den Berg, S. J. Guy, M. Lin, and D. Manocha, "Reciprocal n-body collision avoidance," in *Robotics Research*, 2011.
- [20] Sharma, Pramod, Chang Huang, and Ram Nevatia. "Unsupervised incremental learning for improved object detection in a video." *Computer Vision and Pattern Recognition (CVPR)*, 2012 IEEE Conference on. IEEE, 2012.

- [21] Jeong, Jong-Min, Tae-Sung Yoon, and Jin-Bae Park. "Kalman filter based multiple objects detection-tracking algorithm robust to occlusion." *SICE Annual Conference (SICE), 2014 Proceedings of the*. IEEE, 2014.
- [22] Nguyen, Hoang Thanh, and Bir Bhanu. "Multi-object tracking in non-stationary video using bacterial foraging swarms." *Image Processing (ICIP), 2009 16th IEEE International Conference on*. IEEE, 2009.
- [23] Kennedy, James. "Swarm intelligence." *Handbook of nature-inspired and innovative computing*. Springer, Boston, MA, 2006. 187-219.
- [24] Thida, Myo, How-Lung Eng, and Boon Fong Chew. "Automatic Analysis of Fish Behaviors and Abnormality Detection." *MVA*. 2009.
- [25] Spampinato, Concetto, et al. "Understanding fish behavior during typhoon events in real-life underwater environments." *Multimedia Tools and Applications* 70.1 (2014): 199-236.
- [26] Bharal, Sonam. "Review of underwater image enhancement techniques." *International Research Journal of Engineering and Technology* 2.3 (2015): 340-344.
- [27] Abdullah-Al-Wadud, Mohammad, et al. "A dynamic histogram equalization for image contrast enhancement." *IEEE Transactions on Consumer Electronics* 53.2 (2007).
- [28] Chan, Tony, and Luminita Vese. "An active contour model without edges." *International Conference on Scale-Space Theories in Computer Vision*. Springer, Berlin, Heidelberg, 1999.
- [29] Vese, Luminita A., and Tony F. Chan. "A multiphase level set framework for image segmentation using the Mumford and Shah model." *International journal of computer vision* 50.3 (2002): 271-293.
- [30] Mumford, David, and Jayant Shah. "Optimal approximations by piecewise smooth functions and associated variational problems." *Communications on pure and applied mathematics* 42.5 (1989): 577-685.
- [31] Osher, S. and Sethian, J.A. 1988. "Fronts propagating with curvature-dependent speed: Algorithms based on Hamilton-Jacobi formulation". *JCP*, 79:12–49
- [32] Keenleyside, Miles HA. "Some aspects of the schooling behaviour of fish." *Behaviour* 8.1 (1955): 183-247
- [33] Wang, Kejun, et al. "Target detection algorithm based on Gaussian mixture background subtraction model." *Proceedings of the 2015 Chinese intelligent automation conference*. Springer, Berlin, Heidelberg, 2015.
- [34] M. Piccardi, "Background subtraction techniques: A review", in Proc. IEEE Int. Conf. Systems, Man and Cybernetics, pp. 3099–3104, 2004
- [35] Shevchenko, Violetta. "Fish detection for species recognition." (2017).
- [36] Saidou, Diallo Mamadou. "Tracking a Moving Objects Using Foreground Detector and Improved Morphological Filter." *Open Access Library Journal* 4.12 (2017): 1.
- [37] Gonzalez, Rafael C. Ψηφιακή επεξεργασία εικόνας / Rafael C. Gonzalez, Richard E. Woods • μετάφραση Αθανάσιος Ι. Μάργαρης • επιμέλεια Στέφανος Κόλλιας.
- [38] K. Markus, "Using the Kalman filter to track human interactive motion modeling and initialization of the Kalman filter for translational motion", Tech. rep., University of Dortmund, Germany, 1997.
- [39] Hong, C., et al. "Solving large-scale assignment problems by Kuhn-Munkres algorithm." *2nd Int. Conf. Adv. Mech. Eng. Ind. Informatics (AMEII 2016), no. Ameii*. 2016.
- [40] Ali, Nasser H., and Ghassan M. Hassan. "Kalman filter tracking." *International Journal of Computer Applications* 89.9 (2014).
- [41] Rachel Kleinbauer, *Kalman Filtering Implementation with Matlab*, Study Report, University Stuttgart, Helsinki, November 2004.
- [42] Montella, C. "The Kalman Filter and Related Algorithms: A Literature Review." *Lehigh University* (2011).



- [43] K. Markus, "Using the Kalman filter to track human interactive motion modeling and initialization of the Kalman filter for translational motion", Tech. rep., University of Dortmund, Germany, 1997.
- [44] Shannon, Claude Elwood. "A mathematical theory of communication." *Bell system technical journal* 27.3 (1948): 379-423.
- [45] Li, Xun. "Using complexity measures of movement for automatically detecting movement types of unknown GPS trajectories." *American Journal of Geographic Information System* 3.2 (2014): 63-74.
- [46] Bashir, Faisal I., Ashfaq A. Khokhar, and Dan Schonfeld. "Object trajectory-based activity classification and recognition using hidden Markov models." *IEEE transactions on Image Processing* 16.7 (2007): 1912-1919.