



## **Πολυτεχνείο Κρήτης**

*Τμήμα Ηλεκτρονικών Μηχανικών και Μηχανικών  
Υπολογιστών*

Διπλωματική Εργασία με τίτλο :

### **“Αναγνώριση και Χαρακτηρισμός Κίνησης σε Κυκλοφοριακά Βίντεο”**

Τσεπαπαδάκης Μιχάλης

#### **Εξεταστική Επιτροπή**

**Επιβλέπων :** Ζερβάκης Μιχάλης, Καθηγητής

**Μέλη :** Καλαϊτζάκης Κωνσταντίνος, Καθηγητής  
Λαγουδάκης Μιχαήλ, Αναπληρωτής Καθηγητής

Χανιά, Απρίλιος 2014

## Περίληψη

Καθώς ο αριθμός των εφαρμογών που χρησιμοποιούν τεχνικές όρασης υπολογιστών έχει αυξηθεί πάρα πολύ, η ανίχνευση και ο εντοπισμός κινούμενων οχημάτων παραμένει μια πολύ δύσκολη και κρίσιμη διαδικασία στον τομέα της επιτήρησης της κυκλοφορίας. Στην παρούσα διπλωματική εργασία παρουσιάζεται ένα ολοκληρωμένο σύστημα επιτήρησης σε κυκλοφοριακά βίντεο για τον υπολογισμό της ταχύτητας των κινούμενων οχημάτων στο μετασχηματισμένο χώρο. Το σύστημά μας 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 κάποιων χαρακτηριστικών του. Το σύστημα επιστρέφει αρκετά καλά αποτελέσματα για όλα τα στάδια της επεξεργασίας και σε μεθόδους όπως ο εντοπισμός και η ταξινόμηση οχήματος παρατίθεται αναλυτικά η ακρίβειά τους σε σχέση με τα πραγματικά δεδομένα. Τέλος ο αλγόριθμος εντοπισμού οχήματος, που είναι το πιο βασικό και κρίσιμο στάδιο της επεξεργασίας, είναι δυνατό να βελτιωθεί ακόμα περισσότερο και προς αυτήν την κατεύθυνση μια μελλοντική εργασία θα μπορούσε να οδηγήσει σε ένα πιο πλήρες σύστημα παρακολούθησης της κυκλοφορίας ακόμα και σε πραγματικό χρόνο.

## Ευχαριστίες

Μου δίνεται η ευκαιρία, με την ολοκλήρωση της παρούσας διπλωματικής εργασίας να ευχαριστήσω όλα εκείνα τα πρόσωπα που με τίμησαν με τη στήριξη τους. Αρχικά, θα ήθελα να ευχαριστήσω θερμά τον επιβλέποντα καθηγητή μου κ. Ζερβάκη Μιχάλη που στήριξε την προσπάθειά μου, με καθοδήγησε και μου αφιέρωσε αρκετό από τον πολύτιμο χρόνο του, ώστε να καταφέρουμε να φτάσουμε ως εδώ.

Επίσης θα ήθελα να ευχαριστήσω την κ. Μοιρογιώργου Ντίνα για τη βοήθεια και την καθοδήγηση της στην προσπάθεια κατανόησης του θέματος της επεξεργασίας βίντεο, όπως και τους καθηγητές κ. Καλαϊτζάκη Κωνσταντίνο και κ. Λαγουδάκη Μιχαήλ για την συμμετοχή τους ως μέλη της εξεταστικής επιτροπής.

Τέλος, οφείλω ένα μεγάλο ευχαριστώ στην οικογένεια μου, που στηρίζει την κάθε μου επιλογή και είναι πάντα δίπλα μου.

## Περιεχόμενα

Περίληψη.....	1
---------------	---

Ευχαριστίες.....	2
------------------	---

### **Κεφάλαιο 1 : Εισαγωγή**

1.1 Κίνητρο.....	5
------------------	---

1.2 Συναφείς Εργασίες.....	6
----------------------------	---

1.3 Στόχοι.....	10
-----------------	----

1.4 Διάρθρωση Διπλωματικής Εργασίας.....	10
--	----

### **Κεφάλαιο 2 : Μετασχηματισμός Προοπτικής**

2.1 Εισαγωγή.....	13
-------------------	----

2.2 Μοντέλο Κάμερας.....	14
--------------------------	----

2.3 Προβολική Γεωμετρία.....	17
------------------------------	----

2.3.1 Προβολικός Μετασχηματισμός.....	17
---------------------------------------	----

2.3.2 Αφινικός Μετασχηματισμός.....	17
-------------------------------------	----

2.4 Υλοποίηση Μετασχηματισμού Προοπτικής.....	18
---	----

2.4.1 Επιλογή Χαρακτηριστικών Σημείων.....	23
--	----

### **Κεφάλαιο 3 : Ανίχνευση Κινούμενου Οχήματος**

3.1 Εισαγωγή.....	27
-------------------	----

3.2 Τεχνικές Μοντελοποίησης και Αφαίρεσης Φόντου.....	29
---	----

3.2.1 Frame Differencing.....	29
-------------------------------	----

<u>Αναγνώριση και Χαρακτηρισμός Κίνησης σε Κυκλοφοριακά Βίντεο</u>	4
--	---

3.2.2 Mean Filter.....	30
3.2.3 Approximate Median Filter.....	31
3.2.4 Gaussian Mixture Model.....	32
3.2.5 Αποτελέσματα και Σύγκριση Τεχνικών.....	34
3.3 Ανίχνευση και Αφαίρεση Σκιάς.....	44
3.3.1 Edge-based Αλγόριθμος Αφαίρεσης Σκιάς.....	45
3.3.2 Σύγκριση Αφαίρεσης Σκιάς σε Δύο Χώρους.....	51

## **Κεφάλαιο 4 : Εντοπισμός, Ταχύτητα και Ταξινόμηση Οχήματος** 54

4.1 Εισαγωγή.....	54
4.2 Εντοπισμός Οχήματος.....	56
4.2.1 Αλγόριθμος Kalman Filter.....	56
4.2.2 Υλοποίηση Αλγορίθμου.....	58
4.2.3 Αποτελέσματα Αλγορίθμου.....	60
4.2.4 Προβλήματα κατά τον Εντοπισμό Οχήματος.....	64
4.3 Υπολογισμός Ταχύτητας Οχήματος.....	67
4.4 Ταξινόμηση Οχήματος.....	73

## **Κεφάλαιο 5 : Συμπεράσματα και Μελλοντική Βελτίωση** 76

<b>Αναφορές</b>	<b>78</b>
-----------------	-----------

## **Κεφάλαιο 1 : Εισαγωγή**

Σε αυτό το κεφάλαιο γίνεται μια εισαγωγή, ως προς το υπόβαθρο του επιστημονικού πεδίου που καλύπτεται από την παρούσα διπλωματική εργασία και το κίνητρο της. Επίσης εδώ παρουσιάζονται οι στόχοι και η δομή της εργασίας.

### **1.1 Κίνητρο**

Η επεξεργασία και η εφαρμογή μεθόδων-αλγορίθμων σε βίντεο κυκλοφορίας αποτελεί ένα πολύ δύσκολο επιστημονικό πρόβλημα. Είναι όμως επίσης ένα πεδίο πολύ πρόσφορο για ανάπτυξη νέων, πολλά υποσχόμενων εφαρμογών και γι αυτό το λόγο, εδώ και πολλά χρόνια, έχει κεντρίσει το ενδιαφέρον πολλών ερευνητών αλλά και εταιρειών. Καθώς ο αριθμός των οχημάτων αυξάνεται, το ITS (Intelligent Transportation Systems) γίνεται όλο και πιο σημαντικό. Το ITS αποσκοπεί κυρίως στην απόκτηση ασφαλέστερων συνθηκών κυκλοφορίας, την άνεση στη μεταφορά και την αύξηση της αποτελεσματικότητας της οδικής κυκλοφορίας με τη βελτίωση της λειτουργικότητας των αυτοκινήτων και των δρόμων [1]. Για το σκοπό αυτό, τα συστήματα πληροφοριών για τις θέσεις και τα συστήματα προειδοποίησης για την ασφάλεια των οχημάτων έχουν ήδη εφαρμοστεί, ώστε να ενισχυθεί η ικανότητα του οδηγού να αντιλαμβάνεται τον περιβάλλοντα χώρο. Το κίνητρό μας είναι η μελέτη αλλά και η υλοποίηση μεθόδων ανίχνευσης και εντοπισμού κινούμενων οχημάτων.

Η ανίχνευση των κινούμενων οχημάτων είναι ένα από τα πιο βασικά βήματα για την περαιτέρω ανάλυση. Οι τεχνικές που χρησιμοποιούνται συνήθως για την ανίχνευση αντικειμένου είναι η τεχνική αφαίρεσης του φόντου (background) αλλά και τεχνικές που χρησιμοποιούν στατιστικά μοντέλα. Λόγω πιθανών αλλαγών στο περιβάλλον όπως είναι οι αλλαγές του φωτισμού και οι σκιές, η ανίχνευση οχήματος είναι ένα δύσκολο και σημαντικό πρόβλημα που πρέπει να αντιμετωπιστεί καλά από ένα ισχυρό σύστημα. Στην παρούσα διπλωματική εργασία επιλέξαμε διάφορες τεχνικές για την ανίχνευση οχήματος όπως είναι οι Frame Differencing, Approximate Median Filter, Gaussian Mixture Model (GMM).

Στη συνέχεια, η διαδικασία επεξεργασίας του βίντεο συνεχίζεται με το πιο βασικό βήμα, που είναι η προσπάθεια εντοπισμού όλων των κινούμενων οχημάτων. Η διαδικασία αυτή ορίζεται απλά ως η δημιουργία μιας χρονικής αντιστοιχίας μεταξύ

των οχημάτων που έχουν ανιχνευτεί από frame σε frame. Στη διαδικασία αυτήν προβλέπεται χρονική ταυτοποίηση των περιοχών και από αυτή μπορούν να εξαχθούν πληροφορίες σχετικά με τα οχήματα της περιοχής που παρακολουθείται όπως η πορεία, η ταχύτητα, το μήκος, το μέγεθος και η κατεύθυνση τους.

## 1.2 Συναφείς Εργασίες

Πολλοί ερευνητές έχουν ασχοληθεί με το μετασχηματισμό προοπτικής (perspective) καθώς είναι ένας μετασχηματισμός που προσφέρει πολλαπλά οφέλη στην επεξεργασία βίντεο και εικόνων από κάμερες κυκλοφορίας, όπως επίσης και από κάμερες που είναι τοποθετημένες πάνω στα οχήματα (onboard).

Στο [2] προτείνεται μια πρακτική μέθοδος βαθμονόμησης (calibration) για να λύσει το πρόβλημα του μετασχηματισμού προοπτικής. Χρησιμοποιεί τις πληροφορίες σχήματος μικρότερων τετραγωνικών προτύπων για τον υπολογισμό του πίνακα ομογραφίας. Είναι αρκετό να τα τοποθετήσει περίπου έτσι ώστε να καλύπτουν μια μεγάλη περιοχή της κάμερας ως σύνολο. Η προτεινόμενη μέθοδος είναι χρήσιμη ειδικά όταν η κάμερα οπισθοπορείας είναι εγκατεστημένη σε μια τοποθεσία, όπως ένα αυτοκινούμενο σώμα όπου είναι δύσκολο να διατηρηθεί ευρύς χώρος για τη βαθμονόμηση της κάμερας.

Οι M.Venkatesh, P.Vijayakumar στην έρευνα τους [3], χωρίζουν το μετασχηματισμό προοπτικής μετασχηματισμό σε τρία στάδια. Στο πρώτο στάδιο θα πρέπει να αναπαρασταθεί η εικόνα σ' ένα μετατοπισμένο σύστημα συντεταγμένων, έπειτα να εκτελεστεί η περιστροφή της και στη συνέχεια να προβληθεί η εικόνα σε ένα δισδιάστατο επίπεδο.

Στην εργασία [4], προτείνεται ο μετασχηματισμός top-view για τη μετατροπή συντεταγμένων της εικόνας. Ένας αλγόριθμος που ψάχνει ώστε να ταιριάζει τις παραμέτρους της κάμερας, υπολογίζει τις παραμέτρους που χρησιμοποιούνται για να μετατρέψουν τις συντεταγμένες από την αρχική εικόνα. Χρησιμοποιώντας αυτή την προσέγγιση, δεν είναι απαραίτητο να παρέχονται οποιεσδήποτε εσωτερικές και εξωτερικές παράμετροι προσανατολισμού της κάμερας. Το σχεδιασμένο σύστημα στάθμευσης αυτοκινήτων μπορεί να εγκατασταθεί στο πίσω μέρος του αυτοκινήτου, παρέχοντας στον οδηγό μια σαφέστερη εικόνα της περιοχής πίσω από το αυτοκίνητο.

Οι Qing Lin, Youngjun Han, Hernsoo Hahn [5], υλοποιούν μια μέθοδο αντίστροφης χαρτογράφησης προοπτικής (Inverse Perspective Mapping ή IPM) ώστε να δημιουργηθεί μια κάτοψη της αρχικής εικόνας. Η γνώση των παραμέτρων κάμερας απαιτείται για την εφαρμογή του μετασχηματισμού αυτού. Από μαθηματική άποψη, ο μετασχηματισμός IPM μπορεί να μοντελοποιηθεί σαν μία προβολή από ένα

τρισδιάστατο Ευκλείδειο χώρο  $W$ , που περιέχει τα στοιχεία  $(x,y,z) \in R^3$ , πάνω σε έναν επίπεδο (δισδιάστατο) υποχώρο του που συμβολίζεται με  $I$ , με στοιχεία  $(u,v) \in R^2$ . Η μετατροπή από το  $I$  έως το  $W$  γίνεται ως εξής :

$$x(u,v) = h \cot\{(\theta_0 - \partial_u) + u2\partial_u I(m-1)\} \cos\{(\gamma_0 - \partial_v) + v2\partial_v I(n-1) + 1\}$$

$$y(u,v) = 0$$

$$z(u,v) = h \cot\{(\theta_0 - \partial_u) + u2\partial_u I(m-1)\} \sin\{(\gamma_0 - \partial_v) + v2\partial_v I(n-1) + d\}$$

Με τη χρήση των παραπάνω εξισώσεων και τη συμπλήρωση των αντίστοιχων παραμέτρων της κάμερας, ο μετασχηματισμός μπορεί να πραγματοποιηθεί εύκολα.

Η αφαίρεση του φόντου είναι μια τεχνική που συνήθως χρησιμοποιείται για την εξαγωγή κίνησης σε στατικές εικόνες ή αλλιώς κατάτμηση (segmentation). Επιχειρεί να ανιχνεύσει κινούμενες περιοχές αφαιρώντας την τρέχουσα εικόνα από την εικόνα του στατικού φόντου, η οποία μπορεί να δημιουργηθεί βρίσκοντας τις μέσες τιμές των pixels των εικόνων κατά την πάροδο του χρόνου.

Στο [6] περιγράφεται ένας γρήγορος και ακριβής υπολογισμός του στατικού φόντου της ακολουθίας εικόνων, που είναι και το βασικό βήμα στο σύστημα ελέγχου της οθόνης. Ένας αλγόριθμος σε πραγματικό χρόνο δυναμικής εξαγωγής στατικού φόντου, προβάλλεται μέσω των χαρακτηριστικών του περιβάλλοντος και των κινούμενων αντικειμένων στατικού φόντου. Μεγάλες ποσότητες πειραματικών αποτελεσμάτων αποδεικνύουν ότι ο αλγόριθμος έχει ιδιότητες όπως ταχύτητα, απλότητα, εφαρμοσιμότητα και μια μικρή εξάρτηση από τις καιρικές συνθήκες.

Σύμφωνα με αυτήν την εργασία [7], ένα pixel στη θέση  $(x,y)$  στην τρέχουσα εικόνα  $I_t$  σημειώνεται ως μέρος του foreground εάν ικανοποιείται η παρακάτω εξίσωση

$$|I_t(x,y) - B_t(x,y)| > \tau$$

όπου  $\tau$  είναι ένα προκαθορισμένο κατώφλι. Η εικόνα του φόντου  $B_t$  ενημερώνεται από

$$B_{t+1} = aI_t + (1-a)B_t$$

Οι Rui Tan, Hong Huo, Jin Qian, Tao Fang, [8], στην εργασία αυτή προτείνουν μία τροποποιημένη online διαδικασία EM η οποία κατασκευάζει Adaptive-K Gaussian Mixture Model (AKGMM), στην οποία η διάσταση του χώρου παραμέτρων σε κάθε pixel μπορεί να αντανakλά προσαρμοστικά την πολυπλοκότητα του προτύπου στο pixel. Ένας ευρετικός κανόνας επιλογής συστατικών του φόντου αναπτύσσεται για να παρθεί η απόφαση ταξινόμησης pixel με βάση το προτεινόμενο μοντέλο.



Οι σκιές που ανιχνεύονται μαζί με τα κινούμενα οχήματα συνήθως προκαλούν προβλήματα στον εντοπισμό και στην αναγνώριση των οχημάτων με συνέπεια μεγάλη απώλεια χρήσιμης πληροφορίας σε εφαρμογές. Γι αυτό λοιπόν η τεχνική ανίχνευσης και αφαίρεσης σκιάς είναι ένα πολύ σημαντικό κομμάτι στην επεξεργασία κυκλοφοριακών βίντεο.

Στο [9] οι συγγραφείς προτείνουν 3 διαφορετικές μεθόδους για την αφαίρεση της κινούμενης σκιάς. Η πρώτη είναι η Dual-Pass Otsu μέθοδος [10], η οποία χρησιμοποιεί 2 κατώφλια ώστε να ξεχωρίσει τα “σκοτεινά” pixels που τα θεωρεί σκιές, από τα υπόλοιπα που θεωρούνται ως οχήματα. Η δεύτερη μέθοδος ονομάζεται Region Growing και είναι μια αυτοματοποιημένη διαδικασία στην οποία διαλέγονται pixels δίπλα από το όχημα, που έχει οριοθετοποιηθεί, στην κατεύθυνση της σκιάς. Για τη συγκεκριμένη μέθοδο χρειάζεται να είναι γνωστή η ακριβής ώρα που γίνεται η λήψη του βίντεο ώστε με βάση τη θέση του ήλιου για τη συγκεκριμένη ώρα, να βρεθεί η κατεύθυνση της σκιάς. Τέλος η τρίτη μέθοδος που παρουσιάζεται, ονομάζεται Edge Subtraction and Morphology. Σε αυτήν θεωρείται ότι οι περιοχές με τις πολλές ακμές είναι τα οχήματα και οι υπόλοιπες οι σκιές. Χρησιμοποιώντας κατάλληλες μορφολογικές λειτουργίες η μέθοδος αυτή επιστρέφει αρκετά ικανοποιητικά αποτελέσματα.

Οι Prati, A., Mikic, I., Grana C. και Trivedi M. στο [11] περιγράφουν διάφορες προσεγγίσεις ανίχνευσης σκιάς και συγκρίνουν δύο από αυτές. Ο στόχος αυτών των αλγορίθμων είναι να αποτραπεί ο λάθος χαρακτηρισμός κινούμενων σκιών ως κινούμενα αντικείμενα (ή μέρη τους), αποφεύγοντας έτσι τη συγχώνευση δύο ή περισσότερων αντικειμένων σε ένα και τη βελτίωση της ακρίβειας του εντοπισμού αντικειμένου. Τα σημαντικά χαρακτηριστικά των αλγορίθμων ανίχνευσης σκιάς και η αρχικοποίηση των παραμέτρων αναλύονται και συζητούνται.

Πολλοί ερευνητές έχουν δοκιμάσει διάφορες προσεγγίσεις για τις μεθόδους εντοπισμού οχήματος, υπολογισμού της ταχύτητας του αλλά και την ταξινόμηση που μπορεί να γίνει ανάμεσα στα οχήματα. Ο τρόπος με τον οποίο επιλέγεται κάθε φορά το ποια μέθοδος θα χρησιμοποιηθεί εξαρτάται από την εκάστοτε εφαρμογή. Μερικές από τις εργασίες στον τομέα αυτό είναι οι εξής :

Στην εργασία που προτείνεται στο [55], έχει μελετηθεί η τεχνική ταιριάσματος μπλοκ (block matching) για τον εντοπισμό αντικειμένων σε βίντεο κυκλοφορίας. Οι συγγραφείς προτείνουν την τεχνική ταιριάσματος μπλοκ για διάφορες αναλύσεις και πολυπλοκότητες.

Οι Yiwei Wang, John Doherty και Robert Van Dyck [56], περιγράφουν μια απλή μέθοδο εντοπισμού αντικειμένων με βάση κάποια χαρακτηριστικά τους. Η μέθοδος αρχικά υλοποιεί κατάτμηση στην εικόνα του foreground και του background για να βρεθούν αντικείμενα. Τότε τέσσερα διαφορετικά χαρακτηριστικά συγκεντρώνονται για κάθε αντικείμενο που βρέθηκε. Στη συνέχεια, για κάθε

διαδοχικό frame, οι μεταβολές στα χαρακτηριστικά αυτά υπολογίζονται για διάφορες πιθανές κατευθύνσεις κίνησης. Αυτή που πληροί ορισμένες προϋποθέσεις κατωφλίου επιλέγεται ως θέση του αντικειμένου στο επόμενο frame.

Οι A. Turolla, Λ. Marchesotti και CS Regazzoni παρουσιάζουν στην εργασία τους [57] ένα μοντέλο κάμερας που αποτελείται από πολλαπλές κάμερες. Χρησιμοποιούν χαρακτηριστικά των αντικειμένων που συγκεντρώθηκαν από δύο ή περισσότερες κάμερες που βρίσκονται σε διαφορετικές θέσεις. Τα χαρακτηριστικά αυτά συνδυάζονται κατόπιν για την εκτίμηση θέσης σε συστήματα παρακολούθησης βίντεο.

Οι ερευνητές Yoav Rosenberg και Michael Werman [58], προτείνουν έναν αλγόριθμο εντοπισμού αντικειμένου χρησιμοποιώντας κινούμενη κάμερα. Ο αλγόριθμος βασίζεται στη μοντελοποίηση της κίνησης. Στη μετατόπιση του κάθε σημείου ανατίθεται ένας πίνακας κατανομής πιθανοτήτων. Με βάση το μοντέλο, εκτελείται το βήμα εγγραφής της εικόνας. Η εικόνα που εγγράφηκε συγκρίνεται στη συνέχεια σε σχέση με το φόντο για να εντοπιστεί το κινούμενο αντικείμενο.

Οι ερευνητές Λάζαρος Γραμματικόπουλος, Γεώργιος Κάρρας, Έλλη Πέτσα στην εργασία τους [12] στοχεύουν στην αυτόματη εκτίμηση της ταχύτητας του οχήματος σε ακολουθίες βίντεο. Υποθέτοντας ότι το τμήμα του δρόμου που μελετάται είναι επίπεδο και ευθύ, το σημείο φυγής προς την κατεύθυνση του δρόμου εξάγεται αυτόματα από την αξιοποίηση του διαχωρισμού της λωρίδας. Έτσι, η προβολική παραμόρφωση του οδοστρώματος μπορεί να αφαιρεθεί επιτρέποντας αφινική διόρθωση. Ως εκ τούτου, δεδομένης μια γνωστή απόστασης στο έδαφος κατά μήκος του δρόμου, είναι δυνατή η μονοδιάστατη μέτρηση της θέσης οχημάτων. Τα οχήματα ανιχνεύονται αυτόματα και εντοπίζονται σε διαδοχικά frames. Πρώτον, η εικόνα του φόντου (άδειος δρόμος) έχει δημιουργηθεί από πολλά frames. Στη συνέχεια, η αφαίρεση της εικόνας φόντου από το τρέχον frame δυαδικοποιείται και χρησιμοποιούνται μορφολογικά φίλτρα για την ομαδοποίηση του οχήματος. Τα δεδομένα αναφοράς για την ταχύτητα του οχήματος προέρχονται από αυστηρό δισδιάστατο προβολικό μετασχηματισμό με βάση τα σημεία ελέγχου.

Σε αυτήν την εργασία, [13], δίνεται μια νέα προσέγγιση για την εκτίμηση της ταχύτητας, σε πραγματικό χρόνο, των πολλαπλών οχημάτων της κυκλοφορίας με την κατάλληλη ρύθμιση των παραμέτρων της κάμερας. Ο αλγόριθμος που περιγράφεται αποτελείται από πολλά σημαντικά βήματα. Πρώτα υπολογίζεται το στατικό φόντο. Στη συνέχεια τα νέα στοιχεία εξάγονται χρησιμοποιώντας την προσέγγιση ομοιότητας (Fuzzy Similarity Approach) μεταξύ των εκτιμώμενων pixels του φόντου και τα τρέχοντα pixels που περιέχουν και το foreground και το background. Τρίτον, οι λωρίδες κυκλοφορίας χωρίζονται σε δύο μέρη και για την κατεύθυνση των οχημάτων αλλά και για την παράλληλη επεξεργασία. Τέλος, οι ταχύτητες των οχημάτων υπολογίζονται από τον εκτιμητή Maximum a Posteriori Probability (MAP).

Οι Zhang Guohui, Ryan P. Avery και Yinhai Wang προτείνουν την ταξινόμηση των οχημάτων που εντοπίζονται με κριτήριο το μήκος τους, σε μικρά οχήματα (Small Vehicles ή SV) και σε μεγάλα οχήματα (Large Vehicles ή LV) [14]. Το μήκος του κάθε οχήματος υπολογίζεται από την εξίσωση

$$len = \sqrt{(e_x - s_x)^2 + (e_y - s_y)^2}$$

όπου  $s_x$ ,  $s_y$  είναι οι αρχικές συντεταγμένες και  $e_x$ ,  $e_y$  είναι οι τελικές συντεταγμένες του οχήματος. Το  $len$  αντιπροσωπεύει το μήκος του οχήματος σε pixels.

Στο [15], οι Surendra Gupte, Nikos Papanikolopoulos παραθέτουν αλγόριθμους για την ανίχνευση και την ταξινόμηση των οχημάτων σε ακολουθίες εικόνων από σκηνές κυκλοφορίας που καταγράφονται από σταθερή κάμερα. Η επεξεργασία γίνεται σε τρία επίπεδα : στις αρχικές εικόνες, στο επίπεδο κηλίδων (blobs) και στο επίπεδο του οχήματος. Τα οχήματα μοντελοποιούνται ως ορθογώνια κομμάτια με συγκεκριμένη δυναμική συμπεριφορά. Η προτεινόμενη μέθοδος βασίζεται στη δημιουργία αντιστοιχιών μεταξύ των κηλίδων και των οχημάτων, όπως τα οχήματα κινούνται μέσω της ακολουθίας εικόνων. Το σύστημα μπορεί να ταξινομήσει τα οχήματα σε δύο κατηγορίες, τα φορτηγά και τα μη φορτηγά, με βάση τις διαστάσεις των οχημάτων.

### 1.3 Στόχοι

Για την παρούσα διπλωματική εργασία τέθηκαν ορισμένοι στόχοι στο ξεκίνημά της. Στόχος της είναι να παρουσιάσει αλλά και να προτείνει ένα σύστημα που να υπολογίζει τις ταχύτητες των οχημάτων σε κυκλοφοριακά βίντεο, για κάθε frame που αυτά εμφανίζονται. Αυτό μπορεί να επιτευχθεί μέσω της σωστής ανίχνευσης και του σωστού εντοπισμού των κινούμενων οχημάτων και γι αυτό υλοποιήθηκε ένα σύστημα στην υψηλού επιπέδου τεχνική γλώσσα υπολογιστών Matlab. Επιπλέον στόχος, είναι η προσπάθεια του συστήματος να ταξινομήσει τα οχήματα χρησιμοποιώντας διάφορα χαρακτηριστικά τους ώστε να γίνει η ένταξή τους στην κατάλληλη κατηγορία. Το προτεινόμενο σύστημα που υλοποιείται στην παρούσα διπλωματική εργασία, παρατίθεται υπό τη μορφή διαγράμματος ροής στην Εικόνα 1.1 στο τέλος του κεφαλαίου.

### 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 εργασία χωρίζεται στα

εξής πέντε κεφάλαια : *Εισαγωγή, Μετασχηματισμός Προοπτικής, Ανίχνευση Κινούμενου Οχήματος, Εντοπισμός, Ταχύτητα και Ταξινόμηση Οχήματος και Συμπεράσματα και Μελλοντική Βελτίωση.*

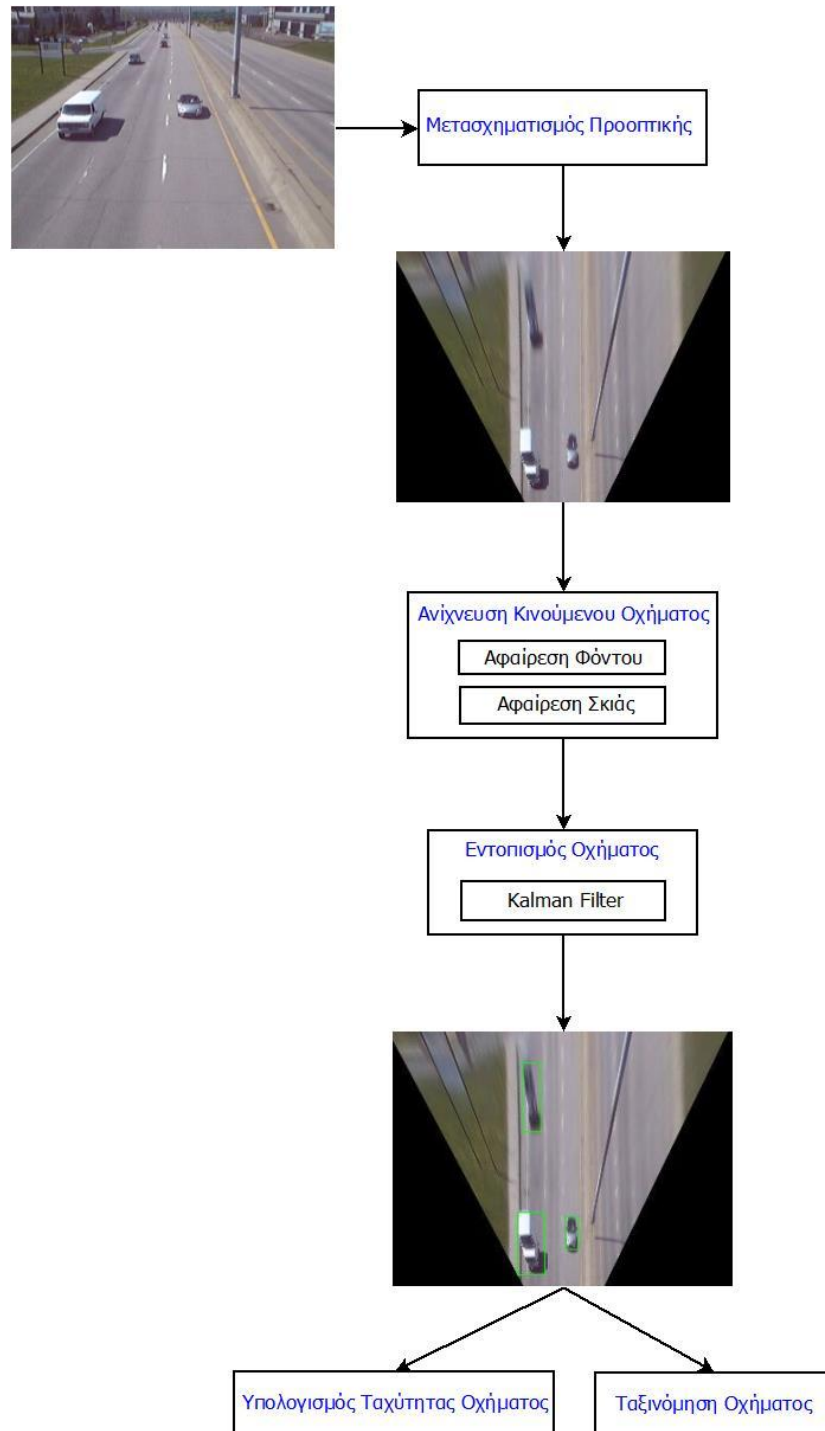
Το πρώτο κεφάλαιο παρέχει μια αρχική εισαγωγή στο θέμα της διπλωματικής εργασίας. Επίσης περιγράφονται το υπόβαθρο, το κίνητρο και οι στόχοι της.

Το δεύτερο κεφάλαιο παρουσιάζει αναλυτικά το μετασχηματισμό προοπτικής και το μοντέλο της κάμερας που συναντάται στα βίντεο τα οποία επεξεργαζόμαστε στην παρούσα διπλωματική εργασία. Ο μετασχηματισμός περιγράφεται συνοπτικά στο [16] και χρησιμοποιεί 4 χαρακτηριστικά σημεία από frames του βίντεο για να υπολογιστεί ο πίνακας μετασχηματισμού που χρειάζεται ώστε να αντιστοιχιστούν τα σημεία της αρχικής εικόνας στο μετασχηματισμένο πλέον χώρο. Η διαδικασία αυτή παρουσιάζεται πιο αναλυτικά, με τις απαραίτητες εξισώσεις, στα [17], [18], [19], [20]. Επιπλέον αυτό το κεφάλαιο ασχολείται με το μοντέλο της στατικής κάμερας που καταγράφει τα κυκλοφοριακά βίντεο που επεξεργαζόμαστε, δίνει δηλαδή το μαθηματικό μοντέλο [21], [18], [22] του επιπέδου της κάμερας και του δρόμου από τα βίντεο-εισόδους του συστήματος.

Στο τρίτο κεφάλαιο γίνεται η περιγραφή των μεθόδων που υλοποιούνται για το διαχωρισμό του background και του foreground. Έχουμε δηλαδή το διαχωρισμό των pixels που ανήκουν στο background και στο foreground αντίστοιχα, με αποτέλεσμα την ανίχνευση των κινούμενων οχημάτων σε κάθε frame. Μαζί με τη διαδικασία αυτήν, επιτυγχάνεται και η εξαγωγή του στατικού φόντου της ακολουθίας των frames από το βίντεο που επεξεργαζόμαστε κάθε φορά, πράγμα που μας χρησιμεύει σε πολλά στάδια της εργασίας μας. Εδώ περιγράφονται αναλυτικά οι τεχνικές μοντελοποίησης και αφαίρεσης του φόντου Frame Differencing, Approximate Median Filter, Mean Filter [23], [24], [25], [26] και Gaussian Mixture Model [27], [28], [29]. Το κεφάλαιο ασχολείται επίσης με το πολύ σημαντικό θέμα της ανίχνευσης και αφαίρεσης σκιάς. Εδώ περιγράφεται ένας Edge-Based αλγόριθμος αφαίρεσης σκιάς [30]. Επίσης στο τέλος συγκρίνονται τα αποτελέσματα που επιστρέφει ο αυτός ο αλγόριθμος όταν εφαρμοστεί στα αρχικά frames των βίντεο, σε σχέση με τα ίδια frames στο μετασχηματισμένο χώρο.

Στο τέταρτο κεφάλαιο περιγράφεται το τελικό στάδιο της συνολικής επεξεργασίας που γίνεται στα βίντεο, με σκοπό γίνει ο εντοπισμός σε κάθε frame όλων των οχημάτων του εκάστοτε βίντεο και έπειτα να βρεθεί η ταχύτητα του κάθε οχήματος, όπως επίσης και να γίνει η ταξινόμηση του ως μικρό (SV) ή ως μεγάλο όχημα (LV). Η μέθοδος υπολογισμού της ταχύτητας που εφαρμόζεται εδώ, υλοποιείται με βάση τη μετακίνηση του κέντρου βάρους (centroid) κάθε οχήματος [31], [32], [33] σε pixels και η ταχύτητα εκφράζεται σε km/h στη συνέχεια. Σε αυτό το κεφάλαιο παρουσιάζονται τα τελικά αποτελέσματα της παρούσας διπλωματικής εργασίας.

Στο πέμπτο και τελευταίο κεφάλαιο της διπλωματικής εργασίας παραθέτονται τα τελικά συμπεράσματα της εργασίας που έγινε, με βάση τα αποτελέσματα της και παρατίθεται ο σχολιασμός των στόχων που είχαν τεθεί εξ αρχής. Επίσης δίνονται προτάσεις για τη συνεισφορά που μπορεί να έχει η παρούσα διπλωματική εργασία σε ήδη υπάρχουσες εφαρμογές όπως επίσης και προτάσεις για μελλοντική βελτίωση.



**Εικόνα 1.1** Διάγραμμα ροής προτεινόμενου συστήματος

## **Κεφάλαιο 2 : Μετασχηματισμός Προοπτικής**

Στο κεφάλαιο αυτό παρουσιάζεται αρχικά το αναλυτικό μοντέλο της κάμερας που συναντάται στα βίντεο τα οποία επεξεργαζόμαστε κι αυτό γιατί είναι σημαντικό να γνωρίζουμε το μαθηματικό μοντέλο που υπάρχει πίσω από ένα απλό βίντεο κυκλοφορίας με οχήματα. Επίσης στο τέλος του κεφαλαίου παρουσιάζεται αναλυτικά η προτεινόμενη μέθοδος υλοποίησης του μετασχηματισμού προοπτικής, στην ακολουθία των frames που εξάγονται από τα βίντεο, που υλοποιείται στην παρούσα διπλωματική εργασία. Έμφαση δίνεται στην κατανόηση της θεωρίας που υπάρχει πίσω από το μετασχηματισμό αυτό και στις υποθέσεις που έχουν γίνει.

### **2.1 Εισαγωγή**

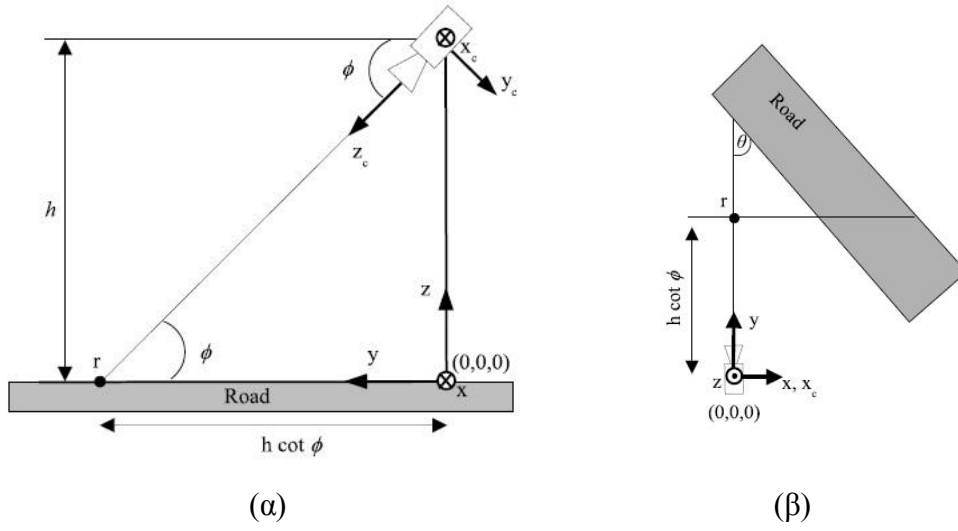
Τα κυκλοφοριακά βίντεο που χρησιμοποιούνται για τις ανάγκες της παρούσας διπλωματικής εργασίας έχουν ληφθεί από σταθερές κάμερες, οι οποίες είναι τοποθετημένες σε καθορισμένο ύψος πάνω από το επίπεδο του δρόμου. Σχεδόν σε όλες τις εφαρμογές επεξεργασίας κυκλοφοριακών βίντεο χρησιμοποιούνται τέτοιου είδους βίντεο, από σταθερές κάμερες. Υπάρχουν βέβαια και άλλες εφαρμογές στις οποίες χρησιμοποιούνται κινούμενες κάμερες, όμως κάτι τέτοιο δε θα μας απασχολήσει σε αυτήν την εργασία. Το μαθηματικό μοντέλο της κάμερας για τέτοιου είδους βίντεο, παρουσιάζεται αναλυτικά με όλες τις παραμέτρους της. Για τον ακριβή υπολογισμό των παραμέτρων αυτών είναι απαραίτητο να εφαρμοστεί η πολύπλοκη διαδικασία της βαθμονόμησης. Η βαθμονόμηση καμερών αναφέρεται στη διαδικασία του υπολογισμού ενός συνόλου παραμέτρων που απαιτούνται για τη λήψη μετρήσεων με μία κάμερα.

Η προβολική γεωμετρία είναι μία μη μετρική μορφή γεωμετρίας και ασχολείται κυρίως με τις ιδιότητες των γεωμετρικών σχημάτων υπό προβολή [34]. Κατά την κεντρική προβολή γεωμετρικών σχημάτων δε διατηρούνται μεγέθη όπως τα μήκη πλευρών και οι γωνίες μεταξύ πλευρών. Ευθείες που είναι παράλληλες προβάλλονται ως συγκλίνουσες σε σημείο, που ονομάζεται και ως σημείο φυγής (vanishing point [38]). Η προβολική γεωμετρία αποτελεί βασικό εργαλείο της μοντελοποίησης και της ανάλυσης προβλημάτων που έχουν σχέση με προβολές και περιλαμβάνει περισσότερους μετασχηματισμούς από την Ευκλείδεια γεωμετρία όπως τον προβολικό και τον αφινικό μετασχηματισμό, οι οποίοι είναι χρήσιμοι για να περιγραφούν πληρέστερα σχήματα μέσω κάμερας.

Ο μετασχηματισμός προοπτικής (ή αλλιώς top-view μετασχηματισμός) είναι ένας προβολικός μετασχηματισμός και ο λόγος που χρησιμοποιείται στην παρούσα διπλωματική εργασία είναι ότι σε αντίθεση με τον πραγματικό, στο μετασχηματισμένο χώρο ο υπολογισμός της ταχύτητας των οχημάτων παρουσιάζει σταθερότητα σε κάθε διαδοχικό frame και το μέγεθος των οχημάτων δεν αλλάζει δραματικά όσο αυτά απομακρύνονται (ή πλησιάζουν) από το φακό της κάμερας. Επίσης είναι δυνατό, σε συνδυασμό και με άλλες διαδικασίες, να μειωθούν λάθη στον εντοπισμό των οχημάτων που προκύπτουν από occlusions. Occlusion σημαίνει ότι υπάρχει κάτι που θέλουμε να δούμε αλλά δε μπορούμε για κάποιο λόγο, συνήθως λόγω επικάλυψης από άλλο αντικείμενο. Κάτι τέτοιο μπορεί να προκύψει στο σύστημά μας όταν το αντικείμενο που θέλουμε να εντοπίσουμε, στην προκειμένη περίπτωση ένα όχημα, “κρύβεται” από κάτι άλλο όπως για παράδειγμα ένα άλλο όχημα. Έτσι λοιπόν, για να αποφευχθούν αυτά τα ανεπιθύμητα αποτελέσματα οι ακολουθίες των frames του βίντεο μετασχηματίζονται σ’ έναν άλλο χώρο σύμφωνα με το μετασχηματισμό προοπτικής που παρουσιάζουμε σε αυτό το κεφάλαιο και η περαιτέρω επεξεργασία προχωράει στο μετασχηματισμένο χώρο.

## 2.2 Μοντέλο Κάμερας

Το μοντέλο της κάμερας και τα χαρακτηριστικά της ακολουθίας των frames από τα βίντεο που επεξεργάζεται το σύστημά μας περιγράφονται στα [21], [35]. Η κάμερα μπορεί να παρουσιαστεί με το τυποποιημένο μοντέλο pinhole κάμερας [36]. Το μοντέλο pinhole που δίνεται στο [37], περιλαμβάνει και τις εσωτερικές αλλά και τις εξωτερικές παραμέτρους. Ο δρόμος που “βλέπει” η κάμερα είναι ευθύς και επίπεδος. Οι υποθέσεις που γίνονται εδώ είναι ότι η κάμερα έχει γωνία roll (η γωνία από  $0^\circ$  έως  $360^\circ$  που αφορά την περιστροφή και την κλίση του άξονα  $X$ ) μηδέν, τετραγωνικά pixels, μηδενική λοξότητα (skew) και ένα αρχικό σημείο στο κέντρο της εικόνας. Με αυτές τις υποθέσεις, τρεις παράμετροι χρειάζονται για να απεικονιστούν τα Ευκλείδεια σημεία στο επίπεδο του δρόμου σε σημεία στο επίπεδο εικόνας : Η εστιακή απόσταση (focal length)  $f$  της κάμερας, το ύψος  $h$  που βρίσκεται η κάμερα πάνω απ’ το επίπεδο του δρόμου και η γωνία κλίσης  $\varphi$ . Επιπρόσθετα, η γωνία pan  $\theta$  είναι απαραίτητη για να ευθυγραμμίσει το δρόμο με την κατεύθυνση της ροής της κυκλοφορίας. Ως γωνία pan ορίζεται η οριζόντια γωνία ( $0^\circ..360^\circ$ ) που αφορά τον όρθιο άξονα  $Z$ . Ορίζονται δύο συστήματα συντεταγμένων, όπως φαίνονται στις Εικόνα 2.1 (α) και Εικόνα 2.1 (β).



**Εικόνα 2.1** (α) Αριστερή όψη της σκηνής, (β) Top-View όψη της σκηνής

Το σύστημα συντεταγμένων κάμερας είναι κεντραρισμένο πάνω στο επίπεδο εικόνας στο αρχικό σημείο, με  $x_c$  και  $y_c$  άξονες που είναι ευθυγραμμισμένοι με τις σειρές και τις στήλες της εικόνας, αντιστοίχως. Το πραγματικό σύστημα συντεταγμένων είναι κεντραρισμένο στο σημείο του επιπέδου του δρόμου ακριβώς κάτω από το κέντρο της προβολής. Ο άξονας  $x$  είναι παράλληλος προς τον άξονα  $x_c$ , ενώ ο άξονας  $y$  είναι κάθετος προς τον άξονα  $x$ , έτσι ώστε ο  $z$  άξονας δείχνει προς τα πάνω από το δρόμο. Υιοθετείται η σύμβαση ότι  $\phi=0$ , όταν η κάμερα έχει προσανατολισμό προς τον ορίζοντα (δηλαδή ο οπτικός  $z_c$  άξονας είναι παράλληλος προς τον άξονα  $y$ ), έτσι ώστε τα δύο frames συντεταγμένων να σχετίζονται με μια περιστροφή  $\phi+\pi/2$  γύρω από τον άξονα  $x$ . Η γωνία  $\phi$  ορίζεται ως η γωνία μεταξύ του επιπέδου  $yz$  και της κατεύθυνση του ταξιδιού. Η προβολή ενός σημείου  $(x,y,z)$  στο πραγματικό επίπεδο σε ένα σημείο  $(u,v)$  στην εικόνα μπορεί να εκφραστεί ως

$$p = Px = KRTx \quad (2.1)$$

όπου οι ομογενείς συντεταγμένες του σημείου στο πραγματικό επίπεδο και της προβολής του είναι  $x=[x \ y \ z \ 1]^T$  και  $p=[au \ av \ a]^T$ ,  $a \neq 0$ , αντίστοιχα και  $(u,v)$  είναι μια κλιμακωτή έκδοση του  $(x_c,y_c)$  σύμφωνα με την εστιακή απόσταση της κάμερας. Ο επόμενος πίνακας αποτυπώνει τις εσωτερικές παραμέτρους της κάμερας

$$K = \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (2.2)$$



$$R = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -\sin \varphi & -\cos \varphi \\ 0 & \cos \varphi & -\sin \varphi \end{bmatrix} \quad (2.3)$$

Ο πίνακας της Εξ. (2.3) αντιστοιχεί σε μία περιστροφή  $\varphi + \pi/2$  γύρω από τον άξονα  $x$  και ο παρακάτω πίνακας αποτυπώνει την κάθετη μετάφραση της κάμερας.

$$T = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -h \end{bmatrix} \quad (2.4)$$

Σημειώνεται εδώ ότι το  $f$  είναι σε pixel, το  $h$  σε μέτρα,  $0 < \varphi < \pi/2$  και  $0 \leq \theta \leq \pi$ . Επεκτείνοντας την Εξ. (2.1) παίρνουμε την έκφραση απεικόνισης για το μοντέλο :

$$\begin{bmatrix} au \\ av \\ a \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & -f \sin \varphi & -f \cos \varphi & fh \cos \varphi \\ 0 & \cos \varphi & -\sin \varphi & h \sin \varphi \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (2.5)$$

όπου  $P = KRT$  είναι ο  $3 \times 4$  πίνακας προβολής. Μετατρέποντας σε ανομοιογενές συντεταγμένες, βλέπουμε ότι ένα σημείο  $(x, y, 0)$  στο επίπεδο του δρόμου προβάλλεται πάνω στην εικόνα στο

$$u = \frac{au}{a} = \frac{fx \sec \varphi}{y + h \tan \varphi} \quad (2.6)$$

$$v = \frac{av}{a} = \frac{fh - fy \tan \varphi}{y + h \tan \varphi} \quad (2.7)$$

Αυτές είναι οι θεμελιώδεις εξισώσεις που χρησιμοποιούνται για να περιγράψουν την απεικόνιση από το επίπεδο του δρόμου προς το επίπεδο της εικόνας.

## 2.3 Προβολική Γεωμετρία

### 2.3.1 Προβολικός Μετασχηματισμός

Στη γεωμετρία, ένας προβολικός μετασχηματισμός μπορεί να είναι ένας συσχετισμός (correlation) ή ένας συγγραμμισμός (collineation) [39]. Εδώ περιγράφονται οι ιδιότητες του προβολικού επιπέδου  $P^2$  που παραμένουν χωρίς να μεταβληθούν κατά την εφαρμογή ενός συνόλου μετασχηματισμών που είναι γνωστοί και ως προβολικότητες [40]. Ως προβολικότητα ορίζεται η αντιστρέψιμη απεικόνιση  $h$  από το προβολικό επίπεδο  $P^2$  στον εαυτό του, τέτοια ώστε τρία σημεία  $x_1, x_2$  και  $x_3$  να βρίσκονται απ' ευθείας εάν και μόνον εάν συμβαίνει το ίδιο και για τα  $x_1', x_2'$  και  $x_3'$ . Όροι συνώνυμοι με την προβολικότητα είναι ο προβολικός μετασχηματισμός, ο συγγραμμισμός [41] και η ομογραφία (homography) [42]. Μία απεικόνιση  $h$  από τον  $P^2$  στον εαυτό του συνιστά προβολικότητα εάν και μόνον εάν υφίσταται ένας πίνακας μετασχηματισμού  $H$  [43]  $3 \times 3$  τέτοιος ώστε για κάθε σημείο  $x$  του  $P^2$  να ισχύει  $x' = Hx$ .

Ο επίπεδος προβολικός μετασχηματισμός είναι γραμμικός μετασχηματισμός που εφαρμόζεται σε ομογενή διανύσματα τριών στοιχείων και εκφράζεται από έναν πίνακα  $3 \times 3$  σύμφωνα με την παρακάτω εξίσωση

$$\begin{pmatrix} x_1' \\ x_2' \\ x_3' \end{pmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \Leftrightarrow x' = Hx \quad (2.8)$$

Εάν στην παραπάνω εξίσωση ο πίνακας  $H$  πολλαπλασιαστεί με μια τυχαία μη μηδενική σταθερά, ο προβολικός μετασχηματισμός παραμένει όπως είναι, χωρίς να επηρεαστεί. Στον προβολικό μετασχηματισμό υπάρχουν οκτώ βαθμοί ελευθερίας, καθώς μεταξύ των εννέα στοιχείων του μπορούν να σχηματιστούν οκτώ διαφορετικοί λόγοι. Ο προβολικός μετασχηματισμός μετατρέπει κάθε σχήμα σε ένα προβολικά ισοδύναμο σχήμα, αφήνοντας όλες τις προβολικές ιδιότητες του αμετάβλητες και αποτελεί μια γενίκευση του αφινικού μετασχηματισμού που περιγράφεται στην επόμενη υποενότητα.

### 2.3.2 Αφινικός Μετασχηματισμός

Η γενική αναπαράσταση του αφινικού μετασχηματισμού (ομοπαράλληλου) [18], [40], [44] σε μορφή πινάκων εκφράζεται ως :

$$\begin{bmatrix} x & y & 1 \end{bmatrix} = \begin{bmatrix} u & v & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & 0 \\ a_{21} & a_{22} & 0 \\ a_{31} & a_{32} & 0 \end{bmatrix} \quad (2.9)$$

Επομένως, ο αφινικός μετασχηματισμός χαρακτηρίζεται από έναν πίνακα μετασχηματισμού του οποίου η τελευταία στήλη είναι ίση με  $[0 \ 0 \ 1]^T$ . Αυτό αντιστοιχεί σε μια ορθογραφική ή παράλληλη ως προς το επίπεδο προβολή από το υν-επίπεδο στο xy-επίπεδο. Ως αποτέλεσμα, οι αφινικές αντιστοιχίες διατηρούν τις παράλληλες γραμμές, πράγμα που επιτρέπει να αποφευχθεί η συμπίκνωση των αξόνων κατά την εκτέλεση δισδιάστατων προβολών. Επιπλέον, τα σημεία που ισαπέχουν διατηρούνται. Ωστόσο, αυτός ο μετασχηματισμός δε μπορεί να σχεδιάσει τετράπλευρα σε τετράπλευρα, κάτι το οποίο υλοποιεί στο σύστημά μας ο μετασχηματισμός προοπτικής.

Για τους αφινικούς μετασχηματισμούς έχουμε τις εξής εξισώσεις :

$$x = \alpha_{11}u + \alpha_{21}v + \alpha_{31} \quad (2.10)$$

$$y = \alpha_{12}u + \alpha_{22}v + \alpha_{32} \quad (2.11)$$

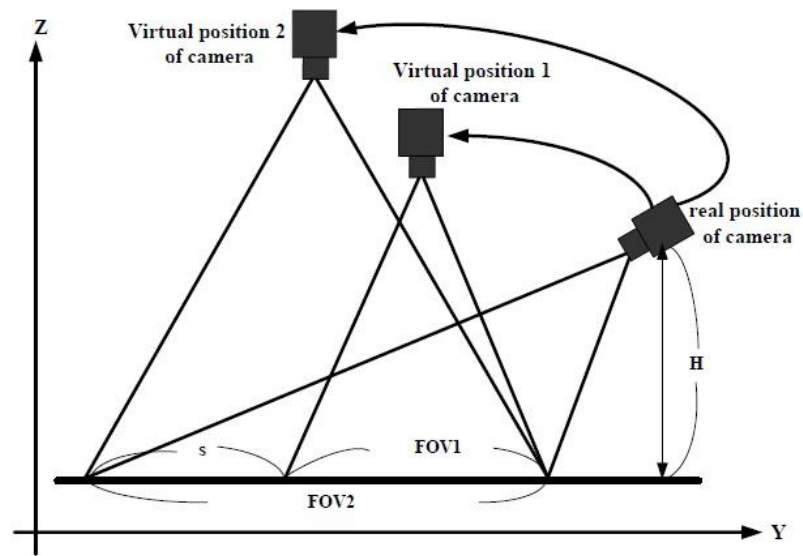
Παραδείγματα αφινικών μετασχηματισμών είναι οι μεταφράσεις, οι περιστροφές και οι κλιμακώσεις. Δεδομένου ότι το αποτέλεσμα που προκύπτει από τον αφινικό μετασχηματισμό είναι επίσης αφινικό, αυτό μπορεί να χρησιμοποιηθεί για να εκτελέσει ένα γενικό προσανατολισμό ενός συνόλου σημείων σχετικό με ένα αυθαίρετο σύστημα συντεταγμένων ενώ εξακολουθεί να διατηρείται μια ενιαία κλίμακα για τις ομοιογενείς συντεταγμένες. Αυτό είναι απαραίτητο για τη δημιουργία σύνθετων μετασχηματισμών.

## 2.4 Υλοποίηση Μετασχηματισμού Προοπτικής

Εφόσον μια κάμερα “συλλαμβάνει” μια εικόνα ενός τρισδιάστατου αντικειμένου [45], τότε, γενικά, θα υπάρχει μια αντιστοίχιση προοπτικής των σημείων του αντικειμένου προς τα αντίστοιχα σημεία της εικόνας. Ζεύγη σημείων που η απόσταση μεταξύ τους σχετικά με το αντικείμενο είναι ίδια, θα είναι πλησιέστερα ή μακρύτερα στην εικόνα ανάλογα με την απόστασή τους από την κάμερα και τον προσανατολισμό τους σε σχέση με το επίπεδο της εικόνας. Σε έναν μετασχηματισμό προοπτικής

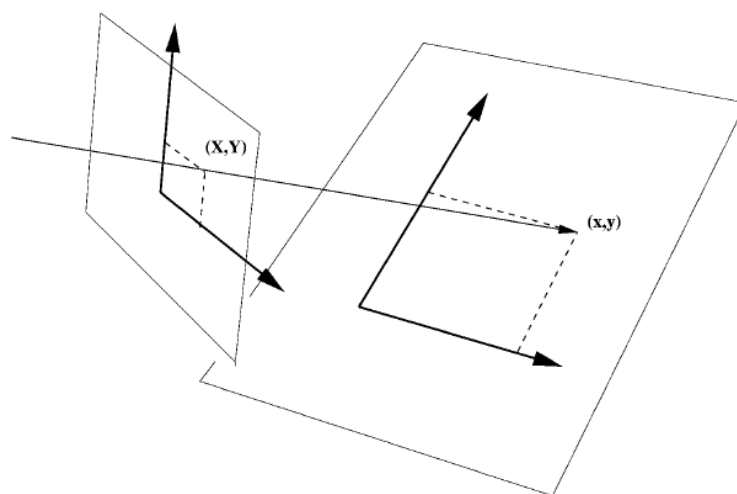
θεωρούμε τρισδιάστατες πραγματικές συντεταγμένες  $(X, Y, Z)$ , οι οποίες κατανέμονται τυχαία στον τρισδιάστατο χώρο και αντιστοιχήθηκαν στο επίπεδο εικόνας μίας κάμερας. Στην παρούσα διπλωματική εργασία η επεξεργασία του βίντεο και των frames επεκτείνεται στην κατασκευή ενός επίπεδου αναφοράς που είναι η ορθογραφική προβολή του δαπέδου. Αυτή η κατασκευή είναι ένας μετασχηματισμός προοπτικής, ο οποίος μπορεί να θεωρηθεί ως μια ειδική περίπτωση της προβολικής ομογραφίας [46]. Ο υπολογισμός της προβολικής ομογραφίας στηρίζεται γενικά στις αντιστοιχίες τεσσάρων τουλάχιστον σημείων της εικόνας με τέσσερα γνωστά σημεία του επιπέδου.

Το οπτικό πεδίο μιας κάμερας FOV (Field Of View) ορίζει το πεδίο μιας πυραμίδας [4]. Το ποσό της σκηνής εντός αυτής της περιοχής χώρου πυραμίδας που θα προβάλλεται πάνω στο επίπεδο εικόνας της κάμερας εξαρτάται και από το FOV αλλά και από το έδαφος μπροστά από την κάμερα. Λόγω του ότι η κάμερα μπορεί να βρίσκεται σε διαφορετικό ύψος  $h$  ή/και με διαφορετική γωνία κλίσης  $\varphi$ , η πυραμίδα μπορεί να λοξεύει ανάλογα με τις συνθήκες στις οποίες εφαρμόζεται. Στην Εικόνα 2.2 παρατίθεται η απεικόνιση αυτών που αναφέρθηκαν παραπάνω. Το έδαφος που προβάλλεται στον άξονα  $Y$  χρησιμοποιείται για να υποδηλώσει το αποτελεσματικό FOV για διαφορετικές περιπτώσεις. Η Εικόνα 2.2 απεικονίζει την ίδια κάμερα που βρίσκεται σε τρεις διαφορετικές θέσεις, δείχνοντας το αποτελεσματικό FOV στην κατεύθυνση του  $Y$  άξονα. Η πραγματική θέση συμβολίζει την κάμερα που είναι εγκατεστημένη σε απόσταση  $h$  από το έδαφος. Εάν η κάμερα βρίσκεται στην εικονική θέση 1 της Εικόνας 2.2, το ύψος της κάμερας είναι υψηλότερο σε αυτήν την περίπτωση από το  $h$ , αλλά μικρότερο από εκείνο της εικονικής θέσης 2, πράγμα που μειώνει το αποτελεσματικό FOV της κάμερας στην εικονική θέση 1 και συμβολίζεται ως FOV1. Εάν η κάμερα βρίσκεται στην εικονική θέση 2, τότε το αποτελεσματικό FOV είναι ίδιο με αυτό της πραγματικής θέσης της κάμερας. Στην παρούσα διπλωματική εργασία, το FOV σχετίζεται με την επιλογή των τεσσάρων χαρακτηριστικών σημείων του μετασχηματισμού προοπτικής που αναλύεται παρακάτω. Όσο πιο κοντά στην πραγματική θέση της κάμερας επιλεγούν τα δύο πρώτα σημεία από τις άκρες του δρόμου, τόσο πιο περιορισμένο FOV θα έχει το αποτέλεσμα του μετασχηματισμού σε σχέση το πραγματικό. Στο σύστημά μας σε κάθε βίντεο που επεξεργαζόμαστε επιλέγουμε αυτά τα σημεία όσο το δυνατόν μακρύτερα από την πραγματική θέση της κάμερας έτσι ώστε να έχουμε, στο αποτέλεσμα που μας επιστρέφει ο μετασχηματισμός, FOV περίπου ίδιο με το πραγματικό. Η αναλυτική περιγραφή για την επιλογή των χαρακτηριστικών σημείων γίνεται στην υποενότητα 2.4.1.



**Εικόνα 2.2** Οπτικό πεδίο κάμερας (FOV)

Το αποτέλεσμα που επιτυγχάνεται με το μετασχηματισμό προοπτικής είναι ο σχεδιασμός ενός αυθαίρετου τετράπλευρου δύο διαστάσεων, σ' ένα άλλο. Η Εικόνα 2.3 δείχνει αυτό ακριβώς που θέλουμε να επιτύχουμε, με το πρώτο τετράπλευρο να αντιστοιχεί στην απεικόνιση του επιπέδου της εικόνας και το δεύτερο τετράπλευρο αυτήν στο πραγματικό επίπεδο.



**Εικόνα 2.3**

Για έναν προβολικό χώρο  $P^n$ , μια προβολική ομογραφία ορίζεται ως ένας πίνακας μετασχηματισμού  $H_{(n+1) \times (n+1)}$ . Η γενική αναπαράσταση ενός σημείου  $x$  που υποβάλλεται στο μετασχηματισμό προοπτικής [17], [18], [19], [46] είναι η εξής :

$$X=Hx \quad (2.12)$$

όπου το  $X$  δηλώνει το διάνυσμα με τις συντεταγμένες στο πραγματικό επίπεδο και  $x$  το διάνυσμα με τις συντεταγμένες στο επίπεδο εικόνας. Αναλυτικότερα η έκφραση γράφεται και ως :

$$\begin{bmatrix} XW \\ YW \\ W \end{bmatrix} = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (2.13)$$

όπου  $W=gx+hy+l$

Ξαναγράφουμε την εξίσωση με τρόπο που εκθέτει την πραγματική-μη γραμμική μορφή της ως

$$\begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} = \frac{\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}}{\begin{bmatrix} g & h & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}} \quad (2.14)$$

Όπως αναφέρθηκε και στην υποενότητα 2.3.1, ο μετασχηματισμός αυτός διατηρεί τις ευθείες γραμμές λόγω συγγραμμικότητας (δηλαδή όλα τα σημεία που βρίσκονται σε μια γραμμή αρχικά, εξακολουθούν να βρίσκονται σε μια γραμμή και μετά την μετατροπή), όπως επίσης και τις αναλογίες των αποστάσεων μεταξύ των σημείων που βρίσκονται σε μια ευθεία γραμμή.

Από την Εξ. (2.14) έχουμε

$$X = \frac{ax + by + c}{gx + hy + 1} \quad (2.15)$$

$$Y = \frac{dx + ey + f}{gx + hy + 1} \quad (2.16)$$

και πολλαπλασιάζοντας με τον παρονομαστή παίρνουμε

$$Y(gx + hy + 1) = dx + ey + f \Rightarrow gYx + hYy + Y = dx + ey + f \quad (2.17)$$

$$X(gx + hy + 1) = ax + by + c \Rightarrow gXx + hXy + X = ax + by + c \quad (2.18)$$

Απομονώνοντας τα  $X$  και  $Y$  και προσθέτοντας μηδενικούς όρους έχουμε

$$X = ax + by + c - 0d + 0e + 0f - Xxg - Xyh \quad (2.19)$$

$$Y = 0a + 0b + 0c + xd + yd + f - Yxg - Yyh \quad (2.20)$$

και βλέπουμε ότι αυτό τελικά είναι το γινόμενο ενός πίνακα μ' ένα διάνυσμα που έχει την εξής μορφή :

$$\begin{bmatrix} x & y & 1 & 0 & 0 & 0 & -Xx & -Xy \\ 0 & 0 & 0 & x & y & 1 & -Yx & -Yy \\ & & & & & & \cdot \\ & & & & & & \cdot \\ & & & & & & \cdot \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \\ e \\ f \\ g \\ h \end{bmatrix} = \begin{bmatrix} X \\ Y \\ \cdot \\ \cdot \\ \cdot \end{bmatrix} \quad (2.21)$$

Από την Εξ. (2.12) κάθε αντιστοιχία από το επίπεδο εικόνας στο πραγματικό παρέχει 2 εξισώσεις γραμμικές στα στοιχεία του πίνακα  $H$ . Για  $n$  αντιστοιχίες έχουμε ένα

σύστημα  $2n$  εξισώσεων και άρα λοιπόν στην περίπτωση μας που έχουμε την τιμή  $n=4$ , λόγω των 4 χαρακτηριστικών σημείων, υπάρχουν 8 εξισώσεις οι οποίες επαρκούν για να υπολογιστεί ο πίνακας  $H$ . Υπάρχουν 3 μέθοδοι υπολογισμού του  $H$  : α) η μη ομογενής γραμμική λύση, β) η ομογενής λύση και γ) η μη γραμμική γεωμετρική λύση. Από τη μέθοδο της ομογενής λύσης έχουμε ότι, γράφοντας τον πίνακα  $H$  στη διανυσματική μορφή  $h = (h_{11}, h_{12}, h_{13}, h_{21}, h_{22}, h_{23}, h_{31}, h_{32}, h_{33})^T$  η ομογενής Εξ. (2.12) για  $n$  σημεία γίνεται  $Ah=b$ , όπου έχουμε

$$\begin{bmatrix} x_1 & y_1 & 1 & 0 & 0 & 0 & -X_1x_1 & -X_1y_1 \\ 0 & 0 & 0 & x_1 & y_1 & 1 & -Y_1x_1 & -Y_1y_1 \\ x_2 & y_2 & 1 & 0 & 0 & 0 & -X_2x_2 & -X_2y_2 \\ 0 & 0 & 0 & x_2 & y_2 & 1 & -Y_2x_2 & -Y_2y_2 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ x_n & y_n & 1 & 0 & 0 & 0 & -X_nx_n & -X_ny_n \\ 0 & 0 & 0 & x_n & y_n & 1 & -Y_nx_n & -Y_ny_n \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \\ e \\ f \\ g \\ h \end{bmatrix} = \begin{bmatrix} X_1 \\ Y_1 \\ X_2 \\ Y_2 \\ \cdot \\ \cdot \\ \cdot \\ X_n \\ Y_n \end{bmatrix} \quad (2.22)$$

Η εξίσωση  $Ah=b$  μπορεί να λυθεί με διάφορες μεθόδους που ισοδυναμούν μ' έναν υπολογισμό ελαχίστων τετραγώνων του διανύσματος  $x$  που ικανοποιεί τη γραμμική σχέση μεταξύ του πίνακα  $A$  και του διανύσματος των συντεταγμένων εξόδου  $B$ . Ο πιο απλός τρόπος, αν και δεν είναι ο πιο σταθερός αριθμητικά, είναι να χρησιμοποιήσουμε τον ψευδοαντίστροφο

$$A\lambda=B \quad (2.23)$$

$$A^T A \lambda = A^T B \quad (2.24)$$

$$\lambda = (A^T A)^{-1} A^T B \quad (2.25)$$

#### 2.4.1 Επιλογή Χαρακτηριστικών Σημείων

Η επιλογή των χαρακτηριστικών σημείων για την υλοποίηση του μετασχηματισμού προοπτικής γίνεται ως εξής : Αναγνωρίζονται δύο γραμμές οι οποίες ενώ είναι παράλληλες στον πραγματικό κόσμο και αντιπροσωπεύονται από τέσσερα σημεία πάνω στην εικόνα, λόγω διαστρέβλωσης της προοπτικής, δε φαίνονται να είναι παράλληλες στις ακολουθίες των frames που επεξεργάζεται το σύστημά μας. Τα τέσσερα αυτά σημεία βρίσκονται στις άκρες του δρόμου όπως ακριβώς φαίνεται με κόκκινο χρώμα στις Εικόνα 2.5 και Εικόνα 2.8. Η επιλογή των συγκεκριμένων σημείων γίνεται ώστε να σχηματιστεί από αυτά ένα τετράπλευρο, μέσα στο οποίο



περιέχεται η περιοχή του δρόμου που μας ενδιαφέρει σε σχέση με το FOV της κάμερας. Οι περιοχές που περικλείονται από αυτά τα σημεία, αντιστοιχούν και στα δύο βίντεο στο κύριο ρεύμα της κυκλοφορίας που είναι στο οπτικό πεδίο της κάμερας. Τα οχήματα που κινούνται σε αυτό το ρεύμα μας ενδιαφέρει περισσότερο να εντοπίσουμε και να τα επεξεργαστούμε, καθώς για τα υπόλοιπα οχήματα του αντίθετου ρεύματος κυκλοφορίας η κάμερα δεν είναι τοποθετημένη μετωπικά προς αυτά. Αυτό δημιουργεί προβλήματα όπως φαίνεται χαρακτηριστικά στα οχήματα που κινούνται στη δεξιά λωρίδα της Εικόνας 2.9, καθώς μετά το μετασχηματισμό φαίνονται σαν να αλλάζουν κατεύθυνση και να είναι λοξά στη λωρίδα κυκλοφορίας τους λόγω του ότι τα οχήματα κινούνται στα πλάγια και όχι μετωπικά όπως αναφέρθηκε και πριν στην κάμερα. Στα επόμενα στάδια του συστήματος, εξαιτίας αυτού του φαινομένου δημιουργούνται συχνά επικαλύψεις οχημάτων σε αυτά τα ρεύματα κυκλοφορίας.

Άρα λοιπόν, τα τέσσερα σημεία επιλέγονται από το πρώτο frame (τα ίδια ακριβώς σημεία χρησιμοποιούνται και για τα υπόλοιπα) ώστε να μετασχηματιστούν όλα τα frames σ' έναν άλλο χώρο, διαδικασία που έχει ως αποτέλεσμα τις ίδιες εικόνες αλλά σαν να είχε τοποθετηθεί η κάμερα ακριβώς πάνω απ' το δρόμο και τα οχήματα (top-view). Με βάση τα σημεία που επιλέξαμε, βρίσκεται ο ομογενής πίνακας μετασχηματισμού  $H$  διαστάσεων  $3 \times 3$  και πραγματοποιεί το μετασχηματισμό για κάθε ένα σημείο της εικόνας, που περιλαμβάνεται στο τετράπλευρο, βρίσκοντας τη νέα θέση του στο μετασχηματισμένο χώρο σύμφωνα με τη Εξ. (2.12). Τα αποτελέσματα του μετασχηματισμού προοπτικής που επιστρέφει το σύστημά μας φαίνονται στις Εικόνα 2.4 και Εικόνα 2.7, όπου είναι τα αρχικά frames και στις Εικόνα 2.6 και Εικόνα 2.9 η αναπαράστασή τους στο μετασχηματισμένο χώρο.



**Εικόνα 2.4** Βίντεο #A frame 350



**Εικόνα 2.5** Το frame 350 με τα 4 χαρακτηριστικά σημεία



**Εικόνα 2.6** Το frame 350 αφού έχει υποστεί το μετασχηματισμό προοπτικής



**Εικόνα 2.7** Βίντεο #B frame 219



**Εικόνα 2.8** Το frame 219 με τα 4 χαρακτηριστικά σημεία



**Εικόνα 2.9** Το frame 219 αφού έχει υποστεί το μετασχηματισμό προοπτικής

## **Κεφάλαιο 3 : Ανίχνευση Κινούμενου Οχήματος**

Κάθε εφαρμογή που χρησιμοποιεί έξυπνη επεξεργασία βίντεο έχει και άλλες ανάγκες, οπότε λογικό είναι να απαιτεί και διαφορετική μεταχείριση. Ωστόσο υπάρχει κάτι που είναι κοινό, τα κινούμενα αντικείμενα (στην περίπτωσή μας είναι οχήματα). Σε αυτό το κεφάλαιο εισάγεται η έννοια των προσαρμοστικών μοντέλων φόντου για ακολουθίες εικόνων από βίντεο, περιγράφονται μέθοδοι για την μοντελοποίηση του φόντου και χρησιμοποιούνται τεχνικές αφαίρεσης του με στατιστικές προσεγγίσεις. Έτσι, η ανίχνευση των περιοχών που αντιστοιχούν σε κινούμενα οχήματα είναι από τα πρώτα βασικά στάδια της επεξεργασίας, δεδομένου ότι απλοποιεί την επεξεργασία για τα επόμενα βήματα. Λόγω των αλλαγών στις σκηνές κατά την πάροδο του χρόνου, όπως είναι ο ξαφνικός φωτισμός και οι καιρικές μεταβολές, όπως επίσης και επαναλαμβανόμενες κινήσεις (π.χ. φύλλα δέντρων που κινούνται όταν φυσάει άνεμος), η ανίχνευση κίνησης είναι ένα δύσκολο πρόβλημα για να αντιμετωπιστεί αξιόπιστα. Η τεχνική που χρησιμοποιούνται για την ανίχνευση των κινούμενων οχημάτων, είναι αυτή της αφαίρεσης του στατικού φόντου. Παραθέτουμε αναλυτικά τρεις μεθόδους για την ανίχνευση οχήματος. Οι δύο πρώτες μέθοδοι είναι σχετικά απλές και χρησιμοποιούν διαδοχικά frames για να ανιχνεύσουν τα κινούμενα οχήματα. Η τρίτη μέθοδος είναι πιο πολύπλοκη καθώς χρησιμοποιεί στατιστικές προσεγγίσεις για την ανίχνευση των κινούμενων οχημάτων. Δίνεται έμφαση στην κατανόηση της υποκείμενης θεωρίας της κάθε μεθόδου. Επιπρόσθετα σε αυτό το κεφάλαιο παρατίθεται μια αποτελεσματική μέθοδος ανίχνευσης και αφαίρεσης της σκιάς, ένα πρόβλημα που εμφανίζεται παράλληλα με την ανίχνευση των κινούμενων οχημάτων.

### **3.1 Εισαγωγή**

Στην επεξεργασία κυκλοφοριακών βίντεο, αναφερόμαστε στο μοντέλο του φόντου που δεν είναι τίποτα άλλο από μια εκτιμώμενη εικόνα ή αλλιώς τα στατιστικά στοιχεία του φόντου μιας σκηνής που απεικονίζεται σε μια ακολουθία εικόνων. Το μοντέλο του φόντου που χρησιμοποιείται για την ανίχνευση οχήματος σε ακολουθίες βίντεο, διαδραματίζει πολύ σημαντικό ρόλο στην επίτευξη του διαχωρισμού των αντικειμένων του foreground από τα αντίστοιχα του background.

Ο πιο απλός τρόπος για να βρεθεί το στατικό φόντο είναι χρησιμοποιώντας μία από τις εικόνες της ακολουθίας όταν δεν υπάρχουν σε αυτήν παρόντα οχήματα. Το foreground μπορεί να προσδιοριστεί με την τεχνική Frame Differencing, που περιγράφεται σε αυτό το κεφάλαιο, δηλαδή συγκρίνοντας κάθε pixel του τρέχοντος frame με την εικόνα του φόντου και αν η διαφορά είναι μικρότερη από κάποιο προκαθορισμένο κατώφλι, το συγκεκριμένο pixel θεωρείται ως pixel του στατικού φόντου. Μια τέτοια λύση μπορεί να είναι επαρκής σε ελεγχόμενα περιβάλλοντα, κάτι όμως που δε συμβαίνει σε αυθαίρετα περιβάλλοντα όπως είναι βίντεο από εξωτερικούς χώρους με συνθήκες φωτός που μεταβάλλονται με την πάροδο του χρόνου. Επίσης σε κάποιες περιπτώσεις μπορεί να είναι δύσκολο ή και αδύνατο να είμαστε σε θέση να βρούμε εικόνα της ακολουθίας χωρίς να υπάρχουν οχήματα σε αυτήν. Ως εκ τούτου, είναι επιθυμητό να υπάρχει ένα μοντέλο για το φόντο που να προσαρμόζεται ανάλογα με τις συνθήκες που επικρατούν κατά τη διάρκεια της λήψης του κάθε βίντεο.

Ένας αλγόριθμος ο οποίος υλοποιεί την αφαίρεση του φόντου, για να είναι αξιόπιστος και σταθερός θα πρέπει να μπορεί να χειρίζεται καταστάσεις όπως σταδιακές ή αιφνίδιες αλλαγές φωτισμού αλλά και επαναλαμβανόμενη κίνηση στο φόντο (π.χ. φύλλα δένδρων). Υπάρχουν αρκετές προσεγγίσεις πάνω στο συγκεκριμένο θέμα [6], [47], [23], [24]. Μια απλή προσέγγιση είναι, αφού βρούμε το στατικό φόντο για μια χρονική στιγμή  $t$  να το αφαιρέσουμε από το frame για τη στιγμή  $t$  ώστε έτσι να βρούμε την απόλυτη τιμή της διαφοράς τους. Στη συνέχεια εφαρμόζεται ένα προκαθορισμένο κατώφλι (threshold)  $Th$  σ' αυτήν την τιμή έτσι ώστε να εξαχθούν τα οχήματα που βρίσκονται στο foreground, όπως φαίνεται στη Εξ. (3.1).

$$| I(x,y,t) - B(x,y,t) | > Th \quad (3.1)$$

όπου  $I(x,y,t)$  είναι το frame τη χρονική στιγμή  $t$  και όπου  $B(x,y,t)$  το στατικό φόντο τη δεδομένη χρονική στιγμή.

Για την ανίχνευση του foreground και τον υπολογισμό του στατικού φόντου παρουσιάζονται σε αυτό το κεφάλαιο οι εξής μέθοδοι : Frame Differencing, Mean Filter, Approximate Median Filter, Gaussian Mixture Model. Επίσης οι προαναφερθείσες μέθοδοι συγκρίνονται ανάλογα με τα αποτελέσματα που επιστρέφει η κάθε μια και περιγράφεται ποιες και γιατί χρησιμοποιήθηκαν τελικά στο σύστημά μας.

## 3.2 Τεχνικές Μοντελοποίησης και Αφαίρεσης Φόντου

Η αφαίρεση του φόντου είναι μια ευρέως χρησιμοποιούμενη προσέγγιση για την ανίχνευση κινούμενων αντικειμένων σε βίντεο από στατικές κάμερες [48]. 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. Τα pixels εκείνα όπου η διαφορά είναι πάνω από ένα προκαθορισμένο κατώφλι, ταξινομούνται ως foreground. Στη συνέχεια εφαρμόζονται ορισμένες μορφολογικές διεργασίες, όπως η διάβρωση (dilation), η διαστολή (erosion) και το κλείσιμο (closing) που εκτελούνται για να μειώσουν τις επιπτώσεις του θορύβου και να ενισχυθούν οι περιοχές που ανιχνεύθηκαν. Η εικόνα του φόντου πρέπει να είναι μια αναπαράσταση της σκηνής χωρίς τα κινούμενα οχήματα 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.

Τα βασικά προβλήματα που παρουσιάζονται εδώ είναι το πώς οι περιοχές των οχημάτων διακρίνονται από το φόντο, το πώς το φόντο αλλάζει με την πάροδο του χρόνου και πώς οι περιοχές οχημάτων, μετά την κατάτμηση, υποβάλλονται σε εκ νέου επεξεργασία έτσι ώστε να απορριφθούν λανθασμένες ανιχνεύσεις. Αν και οι τεχνικές μοντελοποίησης και αφαίρεσης του φόντου αποδίδουν καλά όσον αφορά την εξαγωγή των περιοχών που βρίσκονται τα κινούμενα οχήματα αλλά και την εξαγωγή του στατικού φόντου, τα περισσότερα από τα σχετικά pixels των κινούμενων περιοχών είναι συνήθως ευαίσθητα σε δυναμικές αλλαγές, όταν για παράδειγμα στατικά αντικείμενα υπάρχουν στο φόντο (π.χ. ένα σταθμευμένο αυτοκίνητο κινείται έξω από το παρκινγκ) ή συμβαίνουν απότομες αλλαγές φωτισμού.

Υπάρχουν διαφορετικές προσεγγίσεις σε αυτό το βασικό πρόβλημα της αφαίρεσης του φόντου που έχουν να κάνουν με την ανίχνευση της περιοχής του foreground, με τον υπολογισμό του στατικού φόντου και με την περαιτέρω επεξεργασία. Παρακάτω παρουσιάζονται αναλυτικά όλες οι τεχνικές που χρησιμοποιήθηκαν στην παρούσα διπλωματική εργασία, με την παρουσίαση και του μαθηματικού μοντέλου που υλοποιήθηκε αλλά και τα αποτελέσματα της κάθε μιας τεχνικής ξεχωριστά.

### 3.2.1 Frame Differencing

Η μέθοδος Frame Differencing επιχειρεί να ανιχνεύσει κινούμενες περιοχές κάνοντας χρήση της διαφοράς pixel-by-pixel διαδοχικών frames σε μία ακολουθία βίντεο. Αυτή η μέθοδος είναι εξαιρετικά προσαρμοστική σε δυναμικές αλλαγές σκηνής, ωστόσο γενικά αποτυγχάνει στην ανίχνευση όλων των σχετικών pixels από μερικούς τύπους

κινούμενων οχημάτων. Επίσης, αυτή η μέθοδος αποτυγχάνει να ανιχνεύσει σταματημένα οχήματα που είναι παρόντα στη σκηνή. Έτσι λοιπόν άλλες μέθοδοι θα πρέπει να υλοποιηθούν προκειμένου να εντοπιστούν τα σταματημένα οχήματα για την επιτυχία της επεξεργασίας σε υψηλότερο επίπεδο.

Στη μέθοδο αυτή χρησιμοποιείται ως φόντο το ακριβώς προηγούμενο frame, δηλαδή τη χρονική στιγμή  $t-1$ , οπότε η Εξ. (3.1) που ανιχνεύει τα οχήματα που βρίσκονται στο foreground [23], [24], [25], [26], [49] παίρνει της εξής μορφή :

$$| I(x,y,t) - I(x,y,t-1) | > Th \quad (3.2)$$

Η συγκεκριμένη μέθοδος είναι πολύ ευαίσθητη στον καθορισμό του κατωφλίου  $Th$ , καθώς για χαμηλή τιμή του παραμένει στην εικόνα πληροφορία που δεν ανήκει στο foreground και για μεγάλη τιμή μπορεί να υπάρξει απώλεια πληροφορίας. Μετά από δοκιμές παρατηρήθηκε ότι θέτοντας το κατώφλι στην τιμή  $Th=25$ , επιτυγχάνεται η λιγότερη απώλεια πληροφορίας χωρίς αυτό να σημαίνει βέβαια ότι δεν υπάρχει καθόλου. Επίσης ανάλογα με χαρακτηριστικά όπως η ταχύτητα των οχημάτων, η δομή τους και το frame rate του βίντεο, η μέθοδος επιστρέφει καλά ή όχι και τόσο καλά αποτελέσματα. Αν το όχημα κινείται ήρεμα και με σχετικά μικρή ταχύτητα παρατηρείται μικρή αλλαγή, μικρότερη απ' το προκαθορισμένο κατώφλι, πράγμα που σημαίνει ότι είναι αδύνατο να ανιχνευθεί. Τα πράγματα γίνεται πολύ χειρότερα αν το όχημα κινείται πάρα πολύ αργά. Τότε η μέθοδος δεν επιστρέφει καν αποτέλεσμα. Αυξάνοντας την ευαισθησία (μειώνοντας δηλαδή την τιμή του κατωφλίου) θα λαμβάνουμε λάθος αποτελέσματα από μη κινούμενα οχήματα, όπως κάποιες σκιές και τον άνεμο. Προκειμένου να ξεπεραστούν οι ελλείψεις και τα προβλήματα της παραπάνω μεθόδου σε ορισμένες περιπτώσεις, μπορεί να χρησιμοποιηθεί μια μέθοδος αφαίρεσης τριών frame [50] όπως και ο αλγόριθμος που περιγράφεται στο [51], όπου αναπτύσσεται μια υβριδική μέθοδος που συνδυάζει τη μέθοδο Frame Differencing τριών frame με ένα προσαρμοστικό μοντέλο αφαίρεσης του φόντου. Ο υβριδικός αλγόριθμος εφαρμόζει με επιτυχία την κατάτμηση στις κινούμενες περιοχές χωρίς τα μειονεκτήματα που αναφέρθηκαν πριν, όμως η λύση των προβλημάτων που προκύπτουν σε αυτήν την τεχνική δεν είναι κάτι που θα μας απασχολήσει στην παρούσα εργασία.

### 3.2.2 Mean Filter

Η τεχνική Mean Filter [25] είναι μια μέθοδος μοντελοποίησης του φόντου και δεν ανιχνεύει τις περιοχές, σε αντίθεση με τις υπόλοιπες, που βρίσκονται τα οχήματα. Αυτό είναι δυνατό να συμβεί εάν χρησιμοποιηθεί η Εξ. (3.4), μετά την ολοκλήρωση της μοντελοποίησης, για κάθε ένα frame. Όμως στην παρούσα διπλωματική εργασία υλοποιήθηκε μόνο ο υπολογισμός του στατικού φόντου με αυτήν την τεχνική. Εδώ ως φόντο υπολογίζεται η μέση τιμή των προηγούμενων  $n$  frames, δηλαδή ως εξής :

$$B(x, y, t) = \frac{1}{n} \sum_{i=0}^{n-1} I(x, y, t-i) \quad (3.3)$$

Από το συνδυασμό ανάμεσα στις Εξ. (3.1) και Εξ. (3.3) έχουμε

$$\left| I(x, y, t) - \frac{1}{n} \sum_{i=0}^{n-1} I(x, y, t-i) \right| > Th \quad (3.4)$$

Αυτή είναι μια μέθοδος που επιστρέφει πολύ καλά αποτελέσματα καθώς σε κάθε στιγμή, η πιο πιθανή τιμή για κάθε pixel της εικόνας είναι η τιμή που έχει για να αναπαριστά το στατικό φόντο. Συνήθως για παραπάνω από τα μισά frames συμβαίνει αυτό, οπότε η μέση τιμή κάθε pixel για  $n$  frames θα είναι αυτή που θέλουμε. Όσο περισσότερα frames επιλεγθούν σαν είσοδοι (δηλαδή μεγαλύτερο  $n$ ), τόσο καλύτερα αποτελέσματα επιστρέφει όπως είναι φυσικό η τεχνική αυτή. Αποτελέσματα της μεθόδου αυτής, όπως και όλων των υπολοίπων, δίνονται στην υποενότητα 3.2.5.

### 3.2.3 Approximate Median Filter

Στη Median Filter τεχνική [25], [26], [48] χρησιμοποιείται ο διάμεσος των προηγούμενων  $n$  frames, υποθέτοντας ότι το φόντο είναι πιο πιθανό να εμφανιστεί σ' ένα frame. Παρόμοια με τη μέθοδο Mean Filter, η ιδέα είναι ότι το κάθε pixel μένει στην τιμή που έχει για το φόντο για παραπάνω από τα μισά frames. Οπότε

$$B(x, y, t) = \text{median}\{I(x, y, t-i)\} \quad (3.5)$$

⇓

$$\left| I(x, y, t) - \text{median}\{I(x, y, t-i)\} \right| > Th \quad (3.6)$$

Αυτή η τεχνική έχει αποδειχθεί ότι είναι πολύ σταθερή και έχει απόδοση που είναι συγκρίσιμη με μεθόδους μεγαλύτερης πολυπλοκότητας. Ωστόσο, η αποθήκευση και η επεξεργασία πολλών frames του βίντεο απαιτεί συχνά απαγορευτικά μεγάλη ποσότητα μνήμης. Γι αυτό, μια πιο αποτελεσματική τεχνική προτάθηκε από τους N. McFarlane και C. Schofield [52]. Η τεχνική αυτή ονομάζεται Approximate Median Filter και απ' όταν εμφανίστηκε έχει γνωρίσει ευρεία εφαρμογή στη βιβλιογραφία αφαίρεσης του φόντου [53]. Η τεχνική λειτουργεί ως εξής : αν ένα pixel στο τρέχον



frame έχει τιμή μεγαλύτερη από την αντίστοιχη τιμή pixel του φόντου, το pixel του φόντου αυξάνεται κατά 1. Ομοίως, εάν το τρέχον pixel έχει μικρότερη τιμή τότε μειώνεται κατά ένα. Με τον τρόπο αυτό, το φόντο συγκλίνει τελικά σε μια εκτίμηση, όπου τα μισά pixels εισόδου έχουν μεγαλύτερη τιμή από την τιμή του φόντου και τα μισά μικρότερη, που είναι ο διάμεσος. Έτσι για κάθε frame η τεχνική ανιχνεύει με πολύ μεγάλη ακρίβεια τα αντικείμενα που ανήκουν στο foreground.

### 3.2.4 Gaussian Mixture Model

Οι παραπάνω τεχνικές αφαίρεσης του φόντου που αναλύθηκαν είναι αρκετά εύκολο να υλοποιηθούν και αρκετά γρήγορες. Όμως σε συστήματα όπου πρέπει να γίνεται εντοπισμός και αναγνώριση σε πραγματικό χρόνο (real-time tracking) δε μπορούν να χρησιμοποιηθούν οπότε εκεί πρέπει να εφαρμοστεί κάποια άλλη μέθοδος [47], [29]. Αυτό συμβαίνει γιατί το φόντο συνήθως ποικίλλει σε συστήματα πραγματικού χρόνου ανάλογα με τις καιρικές συνθήκες και την ώρα που περνάει, πράγμα που δείχνει ότι ακόμα και η χρησιμοποίηση ενός προκαθορισμένου, όπως στις προηγούμενες μεθόδους, κατωφλίου δεν οδηγεί σε σωστά αποτελέσματα. Γι αυτό, σε αυτές τις περιπτώσεις, χρειάζεται η ενημέρωση του φόντου μέσω ενός αλγορίθμου που υλοποιεί την αφαίρεση φόντου σε πραγματικό χρόνο. Μια τέτοια μέθοδος, που προτείνεται από τους Stauffer και Grimson [27], είναι η Gaussian Mixture Model. 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. Σε κάθε επανάληψη του αλγορίθμου που υλοποιεί η GMM, 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

$$\{X_1, \dots, X_t\} = \{I(x_0, y_0, i) : 1 \leq i \leq t\} \quad (3.7)$$

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

$$P(X_t) = \sum_{i=1}^K \omega_{i,t} * \eta(X_t, \mu_{i,t}, \Sigma_{i,t}) \quad (3.8)$$

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  $i^{\text{οστής}}$  Gaussian κατανομής στο μείγμα τη χρονική στιγμή  $t$ ,  $\mu_{i,t}$  είναι η μέση τιμή της  $i^{\text{οστής}}$  Gaussian κατανομής στο μείγμα τη χρονική στιγμή  $t$ ,  $\Sigma_{i,t}$  είναι ο πίνακας συνδιακύμανσης της  $i^{\text{οστής}}$  Gaussian κατανομής στο μείγμα τη χρονική στιγμή  $t$  και  $\eta$  είναι μια Gaussian συνάρτηση πυκνότητας πιθανότητας που υπολογίζεται σύμφωνα με την εξίσωση

$$\eta(X_t, \mu, \Sigma) = \frac{1}{(2\pi)^{\frac{n}{2}} |\Sigma|^{\frac{1}{2}}} e^{-\frac{1}{2}(X_t - \mu)^T \Sigma^{-1} (X_t - \mu)} \quad (3.9)$$

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.

$$|X_t - \mu_{i,t-1}| \leq D\sigma_{i,t-1} \quad (3.10)$$

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

$$\omega_{k,t} = (1-a)\omega_{k,t-1} + a(M_{k,t}) \quad (3.11)$$

όπου  $a$  είναι ο ρυθμός μάθησης και το  $M_{k,t}$  είναι 1 για την κατανομή που ταιριάζει και 0 για τις εναπομείναντες κατανομές. Μετά από αυτήν την προσέγγιση, τα βάρη κανονικοποιούνται.

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

$$\mu_t = (1 - \rho)\mu_{t-1} + \rho X_t \quad (3.12)$$

$$\sigma_t^2 = (1 - \rho)\sigma_{t-1}^2 + \rho(X_t - \mu_t)^T (X_t - \mu_t) \quad (3.13)$$

$$\text{όπου} \quad \rho = \alpha \eta(X_t | \mu_k, \sigma_k)$$

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 Εξ. (3.11) 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

$$B = \arg \min_b \left( \sum_{k=1}^b \omega_k > T \right) \quad (3.14)$$

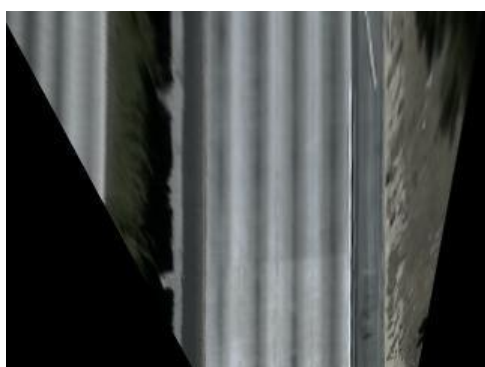
όπου  $T$  είναι το ελάχιστο τμήμα των δεδομένων pixel που πρέπει να ληφθούν υπόψη από το φόντο.

### 3.2.5 Αποτελέσματα και Σύγκριση Τεχνικών

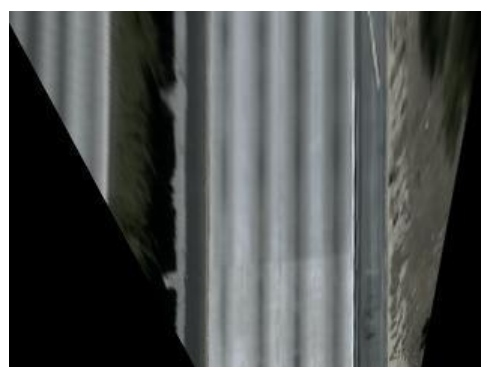
Οι τεχνικές μοντελοποίησης και αφαίρεσης φόντου Frame Differencing, Mean Filter και Approximate Median Filter έχουν τα πλεονεκτήματα [25] ότι είναι εξαιρετικά εύκολη και απλή η υλοποίηση τους, όλες έχουν καλή απόδοση και είναι αρκετά γρήγορες. Στα μειονεκτήματα απ' την άλλη μεριά έχουμε ότι η ακρίβεια των αποτελεσμάτων της τεχνικής Frame Differencing εξαρτάται από την ταχύτητα του κάθε οχήματος όπως και από το frame rate του βίντεο. Στις τεχνικές Mean Filter και Approximate Median Filter υπάρχει το πρόβλημα ότι έχουν σχετικά μεγάλες απαιτήσεις μνήμης για να τρέξουν. Επίσης το μεγαλύτερο μειονέκτημα και για τις 2 αυτές τεχνικές είναι η ύπαρξη και η χρησιμοποίηση ενός καθολικού κατωφλίου για όλα τα pixels της εικόνας.

Η τεχνική Gaussian Mixture Model [27] έχει, συγκριτικά με τις υπόλοιπες τεχνικές, περισσότερα και πιο ουσιώδη πλεονεκτήματα. Μπορεί να αντιμετωπίσει με επιτυχία τις αλλαγές φωτισμού, επαναλαμβανόμενες κινήσεις, ακαταστασία, την εισαγωγή ή την αφαίρεση αντικειμένων στη σκηνή και αργά κινούμενα αντικείμενα. Μεγάλο μειονέκτημα της από την άλλη είναι η ανάγκη αρχικοποίησης των Gaussian κατανομών, διαδικασία πάρα πολύ σημαντική. Τέλος το ότι υπάρχουν πολλές παράμετροι και οι τιμές τους θα πρέπει να επιλέγονται με έξυπνο τρόπο ώστε να είναι αποτελεσματική η μέθοδος, συγκαταλέγεται με τη σειρά του και αυτό στα μειονεκτήματά της.

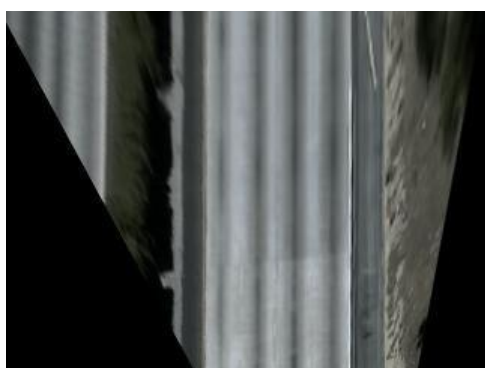
Τα αποτελέσματα που μας επιστρέφει η κάθε τεχνική παρατίθενται στις παρακάτω εικόνες. Αρχικά παρουσιάζεται το στατικό φόντο, ανά μέθοδο, για εικόνες-εισόδους όπως της Εικόνας 2.6 και της Εικόνας 2.9. Στη συνέχεια παρουσιάζονται στιγμιότυπα του βίντεο, μέσω των αντίστοιχων frames, όπου φαίνονται τα αποτελέσματα της κατάτμησης της κάθε τεχνικής και ο διαχωρισμός του foreground από το background.



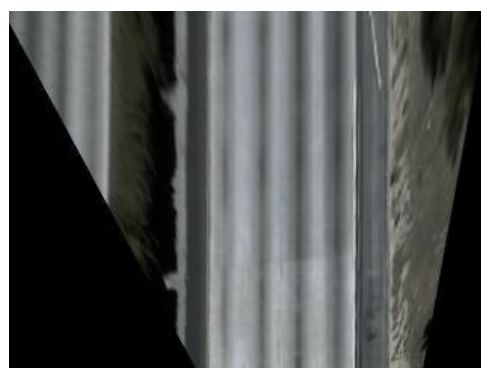
(α)



(β)



(γ)



(δ)

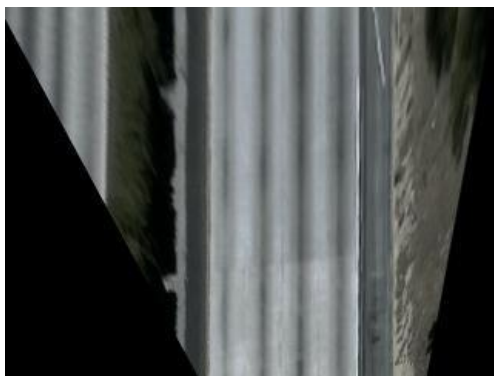


(ε)



(ζ)

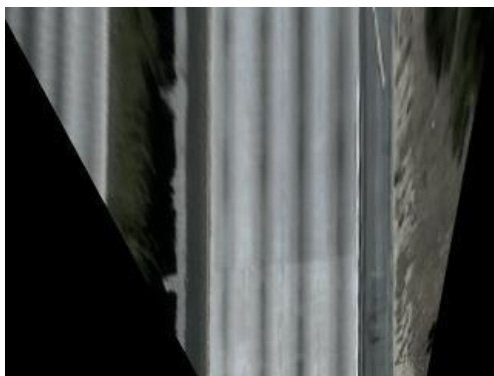
**Εικόνα 3.1** Mean Filter για το Βίντεο #A (α) 1-500 frames, (β) 1-100 frames, (γ) 101-200 frames, (δ) 201-300 frames, (ε) 301-400 frames, (ζ) 401-500 frames



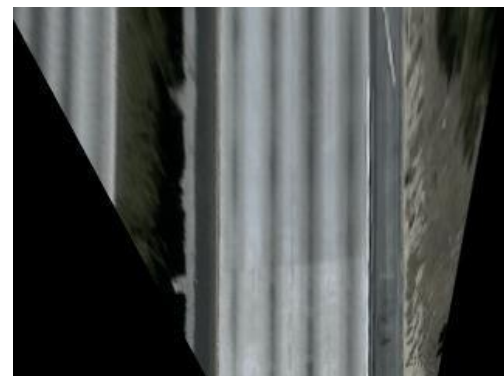
(α)



(β)



(γ)



(δ)

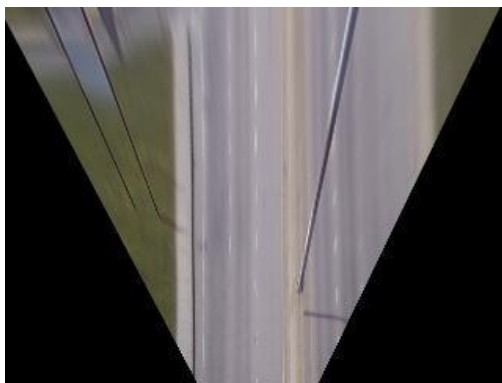


(ε)



(ζ)

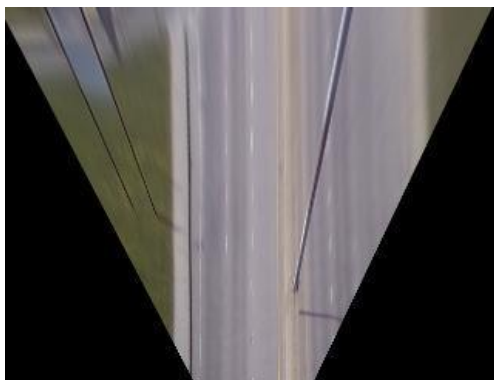
**Εικόνα 3.2** Approximate Median Filter για το Βίντεο #A (α) 1-500 frames, (β) 1-100 frames, (γ) 101-200 frames, (δ) 201-300 frames, (ε) 301-400 frames, (ζ) 401-500 frames



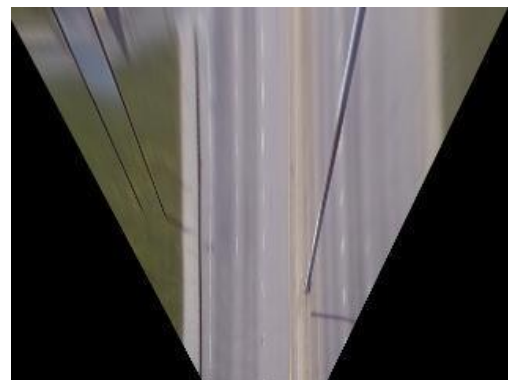
(α)



(β)



(γ)



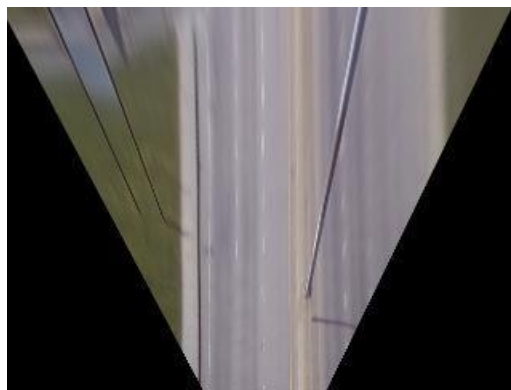
(δ)



(ε)

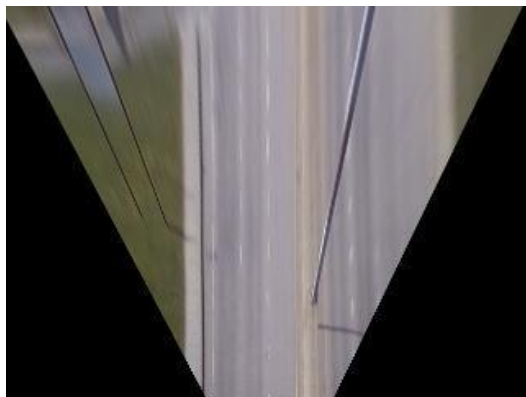


(ζ)

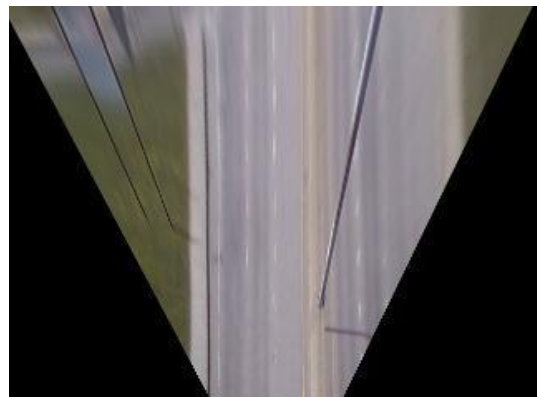


(η)

**Εικόνα 3.3** Mean Filter για το Βίντεο #B (α) 1-2227 frames, (β) 201-300 frames, (γ) 701-800 frames, (δ) 1101-1200 frames, (ε) 1601-1700 frames, (ζ) 2001-2100 frames, (η) 2127-2226 frames



(α)



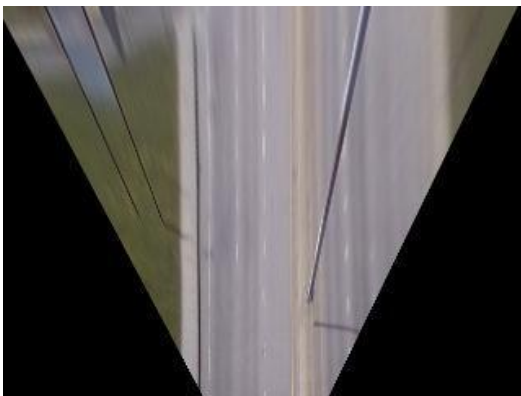
(β)



(γ)



(δ)



(ε)



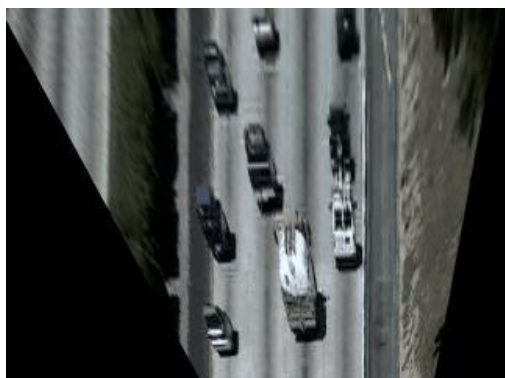
(ζ)



(η)

**Εικόνα 3.4** Approximate Median Filter για το Βίντεο #B (α) 1-2227 frames, (β) 201-300 frames, (γ) 701-800 frames, (δ) 1101-1200 frames, (ε) 1601-1700 frames, (ζ) 2001-2100 frames, (η) 2127-2226 frames





(α)



(β)



(γ)



(δ)



(ε)

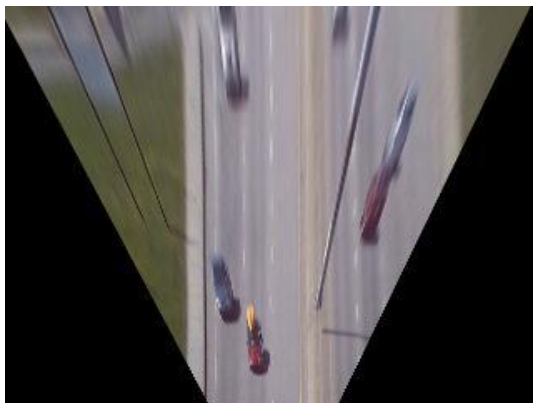


(ζ)

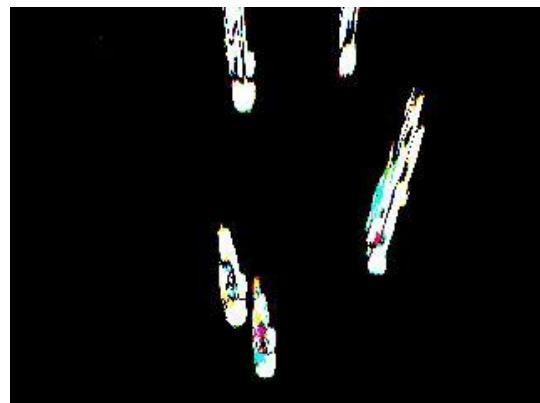


(η)

**Εικόνα 3.5** Βίντεο #A (α) Frame 33, (β) Frame Differencing, (γ) Approximate Median Filter, (δ) GMM, (ε) Εξαγωγή foreground από (β), (ζ) Εξαγωγή foreground από (γ), (η) Εξαγωγή foreground από (δ)



(α)



(β)



(γ)



(δ)



(ε)



(ζ)



(η)

**Εικόνα 3.6** Βίντεο #B (α) Frame 219, (β) Frame Differencing, (γ) Approximate Median Filter, (δ) GMM, (ε) Εξαγωγή foreground από (β), (ζ) Εξαγωγή foreground από (γ), (η) Εξαγωγή foreground από (δ)

Στις Εικόνα 3.1 έως Εικόνα 3.4, παρουσιάζονται τα αποτελέσματα της μοντελοποίησης του στατικού φόντου από τις τεχνικές Mean Filter και Approximate Median Filter. Από τις εικόνες αυτές μπορούμε να διαπιστώσουμε ότι και οι 2 τεχνικές που χρησιμοποιούνται, μοντελοποιούν άψογα το στατικό φόντο και επιστρέφουν πολύ καλά αποτελέσματα. Διαφορές ανάμεσα στις τεχνικές για το ίδιο βίντεο κατά τις ίδιες χρονικές περιόδους δεν παρατηρείται και γι αυτό το λόγο τα αποτελέσματα και των δύο τεχνικών μπορούν να χρησιμοποιηθούν εξίσου στην περαιτέρω επεξεργασία. Οι διαφορές που παρατηρούνται στο στατικό φόντο ανάμεσα στις διαφορετικές χρονικές περιόδους υπολογισμού του, για το ίδιο αρχικό βίντεο, είναι σχεδόν ανεπαίσθητες και οφείλονται σε μικρές αλλαγές του φωτισμού. Στη

συνέχεια στις Εικόνα 3.5 και Εικόνα 3.6 παρουσιάζονται τα αποτελέσματα της αφαίρεσης φόντου για τις τεχνικές Frame Differencing, Approximate Median Filter και GMM. Συγκρίνοντας το αποτέλεσμα από τις Εικόνα 3.5 (β)-(δ) αλλά και από τις Εικόνα 3.6 (β)-(δ), είναι ξεκάθαρο ότι η τεχνική Frame Differencing υστερεί σημαντικά έναντι των άλλων δύο. Όπως φαίνεται, ένα σημαντικό ελάττωμα αυτής της τεχνικής είναι ότι για οχήματα με ομοιόμορφα κατανεμημένες τιμές έντασης (όπως οι πλευρές αυτοκινήτων), τα εσωτερικά pixels ερμηνεύονται ως μέρος του φόντου. Ένα άλλο πρόβλημα είναι ότι τα οχήματα θα πρέπει να κινείται συνεχώς αλλιώς δεν ανιχνεύονται. Επίσης λόγω του προκαθορισμένου κατωφλίου που χρησιμοποιείται, χάνεται σημαντική πληροφορία από τα οχήματα που θέλουμε να ανιχνεύσουμε. Αυτό οδηγεί σε μη ακριβή κατάτμηση και θα προκαλέσει λάθη στον εντοπισμό στη συνέχεια, γι αυτό τα αποτελέσματα της συγκεκριμένης τεχνικής δε χρησιμοποιούνται στα επόμενα στάδια. Η τεχνική Approximate Median Filter αντίθετα, βλέποντας τις Εικόνα 3.5 (γ) και (ζ) όπως και τις Εικόνα 3.6 (γ) και (ζ), επιστρέφει καλύτερα αποτελέσματα από την προηγούμενη τεχνική, κατά το διαχωρισμό του συνόλου του οχήματος από το φόντο. Η τεχνική αυτή είναι μια πάρα πολύ καλή λύση για το σύστημά μας. Προσφέρει απόδοση κοντά σε αυτό που μπορεί να επιτευχθεί με μεθόδους υψηλότερης πολυπλοκότητας και δεν έχει πολύ μεγαλύτερο κόστος υλοποίησης και χώρου αποθήκευσης από τη Frame Differencing, που είναι η πιο ανέξοδη και εύκολη στην υλοποίηση από όλες τις τεχνικές που εξετάζονται. Βλέποντας τις Εικόνα 3.5 (δ) και (η) και Εικόνα 3.6 (δ) και (η), η τεχνική GMM αντιμετωπίζει αρκετά καλά το διαχωρισμό των οχημάτων αλλά και την εξαφάνιση του θορύβου στο φόντο σε σχέση με την τεχνική Frame Differencing. Ωστόσο, υπάρχουν και αρκετά σημεία στις παραπάνω εικόνες όπου η μέθοδος αποτυγχάνει, επιτρέποντας σε αρκετά pixels από το φόντο να διαρρεύσουν στο foreground. Τα σημεία αυτά αντιστοιχούν σε σχετικά γρήγορες αλλαγές στο φωτισμό. Συγκρίνοντας τα αποτελέσματα αυτά με τα αντίστοιχα της τεχνικής Approximate Median Filter, φαίνεται ότι το συγκεκριμένο πρόβλημα είναι πιο έντονο. Αυτό συμβαίνει επειδή το μοντέλο του φόντου δεν προσαρμόζεται αρκετά γρήγορα στην τεχνική GMM. Βέβαια αυτό δεν αποδεικνύει ότι η GMM τεχνική επιστρέφει λιγότερο σταθερά αποτελέσματα αναγκαστικά αλλά το ότι το πρόβλημα της με τις αλλαγές φωτισμού στα βίντεο δοκιμών, καταδεικνύει μία από τις βασικές προκλήσεις της συγκεκριμένης τεχνικής, που είναι η βέλτιστη επιλογή των παραμέτρων (Πίνακας 3.1) ανάλογα με την εκάστοτε εφαρμογή. Στην παρούσα διπλωματική εργασία για να έχει το σύστημά μας στη συνέχεια, όσο είναι δυνατόν, μεγαλύτερη επιτυχία στον εντοπισμό οχήματος, επιλέγεται να χρησιμοποιηθεί η τεχνική Approximate Median Filter καθώς επιστρέφει τα καλύτερα δυνατά αποτελέσματα για τα βίντεο που εξετάζονται από το σύστημά μας.

Σε αυτό το σημείο πρέπει να σημειωθεί ότι για την επεξεργασία που γίνεται στο σύστημά μας έχουν χρησιμοποιηθεί 2 βίντεο από σταθερές κάμερες καθώς ήταν αρκετά δύσκολο να βρεθούν πολλά διαφορετικά βίντεο που να έχουν ληφθεί από

τέτοιες κάμερες, όπως επίσης να έχουν και ικανοποιητική ανάλυση ώστε να επεξεργαστούν. Γι αυτό το λόγο τα 2 βίντεο που χρησιμοποιούνται, έχουν “τεμαχιστεί” σε μικρότερης διάρκειας βίντεο, μήκους 100 frames και 320x240 pixels το καθένα. Το πρώτο βίντεο προς επεξεργασία (Βίντεο #A = highwayII\_raw.avi) αποτελείται από 500 frames συνολικά ενώ το δεύτερο βίντεο (Βίντεο #B = HighwayIII\_Original\_sequence.avi) αποτελείται από 2227 frames. Από το πρώτο εξάγουμε 5 βίντεο (Βίντεο #A1-A5) και από το δεύτερο 6 βίντεο (Βίντεο #B1-B6) σε διάφορες χρονικές στιγμές. Με αυτόν τον τρόπο ουσιαστικά επιτυγχάνεται η επεξεργασία και η σύγκριση των αποτελεσμάτων από 11 βίντεο συνολικά αντί για τα 2 αρχικά. Από τη συγκεκριμένη επιλογή επεξεργασίας μπορούμε να διαπιστώσουμε, ειδικά στο δεύτερο βίντεο που είναι και μεγαλύτερο, κάποιες μικρές διαφορές στο στατικό φόντο κατά τις διαφορετικές χρονικές στιγμές, παρόλο που αναφέρονται στο ίδιο αρχικό βίντεο. Αυτό οφείλεται στο ότι με την πάροδο του χρόνου ο φωτισμός της εικόνας αλλάζει, όπως επίσης κάποιο αντικείμενο που αρχικά δεν υπήρχε μπορεί να γίνει μέρος του φόντου στη συνέχεια. Έτσι λοιπόν με αυτόν τον τρόπο αποδεικνύεται ότι το σύστημα μπορεί να επεκταθεί σε οποιοδήποτε βίντεο έχει ληφθεί από σταθερή κάμερα κυκλοφορίας και να δώσει σωστά αποτελέσματα.

Παράμετροι GMM
$K = 3$
$\alpha = 0.01$
$D = 2.5$

Πίνακας 3.1

### 3.3 Ανίχνευση και Αφαίρεση Σκιάς

Οι τεχνικές που περιγράφονται παραπάνω για την ανίχνευση κίνησης αποδίδουν καλά σε εσωτερικούς και εξωτερικούς χώρους και χρησιμοποιούνται αρκετά χρόνια σε παρόμοιες εφαρμογές. Ωστόσο, οι περισσότερες από αυτές τις τεχνικές είναι ευαίσθητες στις αλλαγές του φωτισμού που προκαλούνται συνήθως από τις σκιές. Έτσι οι τεχνικές ανίχνευσης κίνησης αποτυγχάνουν να ανιχνεύσουν μόνο τα οχήματα αλλά ανιχνεύουν παράλληλα και τις σκιές ως μέρος του οχήματος.

Σε εφαρμογές που ασχολούνται με την ανάλυση και την επεξεργασία κυκλοφοριακών βίντεο, όπως είναι και η παρούσα διπλωματική εργασία, η ύπαρξη σκιών στα βίντεο μπορεί να προκαλέσει σοβαρά προβλήματα, κατά το στάδιο της

επεξεργασίας που γίνεται ο εντοπισμός των κινούμενων οχημάτων. Προβλήματα προκύπτουν λόγω εσφαλμένης ταξινόμησης των σημείων σκιάς ως foreground. Οι σκιές μπορούν να προκαλέσουν συγχώνευση οχημάτων, παραμόρφωση στο σχήμα του οχήματος ακόμα και απώλεια ολόκληρου του οχήματος (λόγω σκιάς πάνω από ένα άλλο όχημα). Οι δυσκολίες που συνδέονται με την ανίχνευση σκιάς προκύπτουν από το ότι οι σκιές και τα οχήματα μοιράζονται δύο σημαντικά οπτικά χαρακτηριστικά. Καταρχάς, τα σημεία σκιάς ανιχνεύονται ως σημεία του foreground δεδομένου ότι διαφέρουν σημαντικά από αυτά του φόντου. Δεύτερον, οι σκιές έχουν την ίδια κίνηση όπως και τα οχήματα που τις προκαλούν. Για να αντιμετωπιστούν λοιπόν αυτά τα προβλήματα που προκύπτουν, χρησιμοποιούνται οι αλγόριθμοι ανίχνευσης και αφαίρεσης σκιάς. Όπως και για την αφαίρεση του φόντου, έτσι και εδώ για την ανίχνευση και αφαίρεση σκιάς υπάρχουν διάφορες μέθοδοι που τις υλοποιούν. Οι σκιές αποτελούνται από 2 συστατικά : την self σκιά και την cast σκιά. Οι σκιές που αναφέρονται ως cast, είναι σκιές που πέφτουν από το όχημα που μας ενδιαφέρει σε άλλα οχήματα ή στο δρόμο και είναι οι σκιές που έρχονται στο μυαλό όταν αναφέρεται ο όρος σκιά. Η σκιά που αναφέρεται ως self, είναι η σκιά που μπορεί να πέσει σ' ένα αντικείμενο από τον εαυτό του όταν κρύβεται από την πηγή φωτισμού. Η διάκριση μεταξύ αυτών των τύπων σκιών είναι σημαντική για την αναγνώριση οχημάτων. Μια επιτυχημένη μέθοδος αφαίρεσης σκιάς αποσκοπεί στο να αφαιρέσει τις cast σκιές, καθώς αναγνωρίζει τις self σκιές ως μέρος του αντικειμένου ενδιαφέροντος και ως εκ τούτου τις διατηρεί. Δυστυχώς, τα παρόμοια χαρακτηριστικά μεταξύ cast και self σκιών συχνά μπερδεύουν την ικανότητα ενός υπολογιστή να τις διακρίνει μεταξύ τους. Στη μέθοδο που παρατίθεται παρακάτω, θεωρείται ως δεδομένο ότι υπάρχουν εικόνες στις οποίες περιέχονται οι περιοχές με τα οχήματα και τις σκιές τους, όπως επίσης και μια εικόνα με το στατικό φόντο. Και τα δύο αυτά δεδομένα υπάρχουν στο σύστημά από την επεξεργασία που παρουσιάστηκε και υλοποιήθηκε στην ενότητα 3.2.

### 3.3.1 Edge-based Αλγόριθμος Αφαίρεσης Σκιάς

Στην παρούσα διπλωματική εργασία προτείνεται μια μέθοδος αφαίρεσης σκιάς edge-based [30]. Ο αλγόριθμος εφαρμόζει την ανίχνευση σκιάς και αποφασίζει αν υπάρχουν σκιές σε μια εικόνα. Προφανώς για να έχει νόημα η διαδικασία αυτή, έχει σημασία το πότε και υπό τι συνθήκες εφαρμόζουμε την επεξεργασία στα βίντεο και στις εικόνες μας. Για παράδειγμα κατά τη διάρκεια της νύχτας ή κατά τη διάρκεια μια βροχερής μέρας δεν έχει νόημα να εφαρμοστεί η ανίχνευση σκιάς.

Σε αυτήν τη μέθοδο οι πληροφορίες για τις ακμές (edges), όπως προδίδει και η ονομασία της, παίζουν σημαντικό ρόλο για την ανίχνευση των σκιών. Εφαρμόζεται η τεχνική ανίχνευσης ακμών Sobel [54] σε μια επιλεγμένη περιοχή του foreground αντί για ολόκληρη την εικόνα. Σκοπός της μεθόδου που περιγράφεται, είναι να αφαιρεθεί

η οριοθετημένη περιοχή της σκιάς και παράλληλα να διατηρηθούν οι ακμές του αντικειμένου. Τα 3 βήματα που ακολουθούνται ώστε να επιτευχθεί η διαδικασία αυτή είναι τα εξής :

### Βήμα 1

Ο υπολογισμός της οριοθέτησης του foreground είναι το πρώτο βήμα για την αφαίρεση της ακμής της σκιάς. Ως  $C_t$  (Εικόνα 3.7 (β) και Εικόνα 3.8 (β)) ορίζεται το foreground που περιέχει τα οχήματα του εκάστοτε frame και ως  $E_t$  (Εικόνα 3.7 (γ) και Εικόνα 3.8 (γ)) η ανίχνευση ακμών για το  $C_t$ . Είναι εύκολο να βρεθεί μια εικόνα ακμών του foreground που ορίζεται ως  $EC_t$ , επειδή το  $C_t$  είναι μια δυαδική εικόνα με τιμές 0 και 1. Το  $EC_t$  μπορεί να βρεθεί χρησιμοποιώντας πρώτα τη μορφολογική διαδικασία της διάβρωσης για τα  $C_t$  και  $B_1$  και στη συνέχεια εκτελώντας την αφαίρεση μεταξύ του αποτελέσματος από το  $C_t$

$$EC_t = C_t - (C_t \odot B_1) \quad (3.15)$$

όπου  $B_1$  είναι ένα δομικό στοιχείο που περιέχει 1 μεγέθους 3x3 pixels και ο  $\odot$  τελεστής υποδηλώνει τη διαδικασία της διάβρωσης.

Για να αφαιρεθούν οι ακμές από το  $E_t$  πρέπει να βρεθεί μια εικόνα που να περιέχει την οριοθετημένη περιοχή του κάθε οχήματος (μαζί με τη σκιά του), που έχει υποστεί τη μορφολογική διαδικασία της διαστολής και έχει πάχος μεγαλύτερο από ένα pixel. Αυτή η εικόνα ορίζεται από την Εξ. (3.16) και φαίνεται στις Εικόνα 3.7 (δ) και Εικόνα 3.8 (δ)

$$DB_t(x, y) = \begin{cases} 1, & DE_t(x, y) = 1 \text{ and } C_t(x, y) = 1 \\ 0, & \text{otherwise} \end{cases} \quad (3.16)$$

### Βήμα 2

Το  $DE_t$  δίνεται από

$$DE_t = EC_t \oplus B_2 \quad (3.17)$$

όπου το  $DE_t$  δηλώνει το  $EC_t$  που έχει υποστεί τη διαδικασία της διαστολής και όπου  $B_2$  είναι ένα δομικό στοιχείο που περιέχει 1 μεγέθους 3x3 pixels.

### Βήμα 3

Τα σημεία ακμών του foreground έξω από το  $DB_t$  θεωρούνται ως αρχικές εσωτερικές ακμές που ορίζονται από την Εξ. (3.18) και φαίνεται στις Εικόνα 3.7 (ε) και Εικόνα 3.8 (ε)

$$IE_t(x, y) = \begin{cases} 1, & E_t(x, y) = 1 \text{ and } DB_t(x, y) = 1 \\ 0, & \text{otherwise} \end{cases} \quad (3.18)$$

Μετά την παραπάνω επεξεργασία, όλα τα σημεία ακμών στην περιοχή σκιάς έχουν σχεδόν εξαλειφθεί. Ωστόσο, ένα στρώμα από λάθος ακμές λόγω θορύβου πρέπει να αφαιρεθούν επειδή έχουν μεγάλη επίδραση στην ανίχνευση σκιάς. Η διαδικασία βελτίωσης αποτελείται από οριζόντιες και κάθετες λειτουργίες που ορίζονται ως  $HR_t$  (Εικόνα 3.7 (ζ) και Εικόνα 3.8 (ζ)) και  $VR_t$  (Εικόνα 3.7 (η) και Εικόνα 3.8 (η)) αντίστοιχα. Οι δύο αυτές λειτουργίες εφαρμόζονται στο  $IE_t$ . Για την οριζόντια λειτουργία έχουμε :

Αρχικά η περιοχή μέσα στα πρώτα και τα τελευταία σημεία ακμής σε κάθε γραμμή ορίζεται με την τιμή 1 και δημιουργεί μια μάσκα  $IH_t$ . Για να το πετύχουμε αυτό, χρησιμοποιούμε πλαίσια οριοθέτησης (bounding boxes) στο  $C_t$  έτσι ώστε οτιδήποτε υπάρχει έξω από αυτά τα όρια να μη λαμβάνεται υπόψη. Στη συνέχεια αφαιρείται η μικρή περιοχή θορύβου της μάσκας  $IH_t$ . Για την κάθετη λειτουργία  $VR_t$  κάνουμε ακριβώς το ίδιο αλλά για κάθε στήλη σε αυτήν την περίπτωση. Μετά από αυτές τις 2 λειτουργίες οι ακμές της περιοχής του foreground και στα δύο,  $HR_t$  και  $VR_t$ , είναι οι εσωτερικές ακμές των κινούμενων οχημάτων και συμβολίζεται ως

$$RE_t(x, y) = \begin{cases} 1, & HR_t(x, y) = 1 \text{ and} \\ & VR_t(x, y) = 1 \text{ and} \\ & E_t(x, y) = 1 \\ 0, & \text{otherwise} \end{cases} \quad (3.19)$$

Το αποτέλεσμα φαίνεται στις Εικόνα 3.7 (θ) και Εικόνα 3.8 (θ).

Αν μία περιοχή foreground περιέχει περισσότερα από ένα οχήματα, οι σκιές από τα διαφορετικά αυτά οχήματα ενδεχομένως να συνδέονται μεταξύ τους στην περιοχή αυτή. Για να αφαιρεθούν αποτελεσματικά οι ανεπιθύμητες σκιές, θα πρέπει να προσδιοριστούν ποιά από τα σημεία ακμής του  $RE_t$  ανήκουν στο ίδιο όχημα. Η ταξινόμηση της περιοχής foreground υλοποιείται ως εξής :

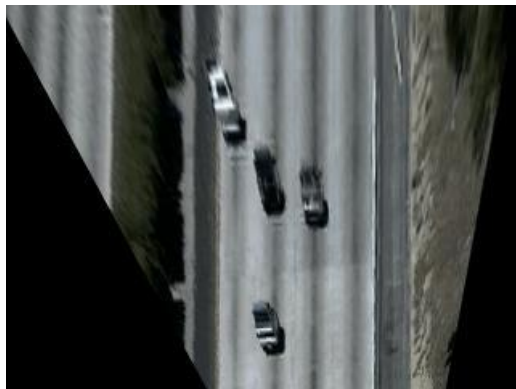
Η οριζόντια και η κάθετη λειτουργία, οι οποίες περιγράφηκαν προηγουμένως, εφαρμόζονται και εδώ, αυτή τη φορά στο  $RE_t$ . Τα αποτελέσματα της οριζόντιας και της κάθετης λειτουργίας στο  $RE_t$  ορίζονται ως  $HRE_t$  και  $VRE_t$  αντίστοιχα.



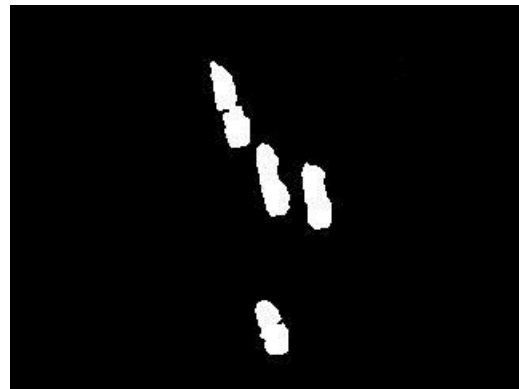
Το τελευταίο στάδιο της μεθόδου αυτής είναι το στάδιο της κατασκευής των κινούμενων οχημάτων. Η κατασκευή περιγράφεται στην Εξ. (3.20) και το αποτέλεσμα φαίνεται στις Εικόνα 3.7 (ι) και Εικόνα 3.8 (ι).

$$IF_t(x, y) = \begin{cases} 1, & HRE_t(x, y) = 1 \text{ or } VRE_t(x, y) = 1 \\ 0, & \text{otherwise} \end{cases} \quad (3.20)$$

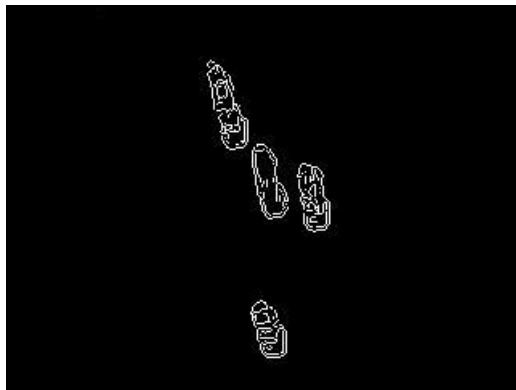
Από τα  $HRE_t$  και  $VRE_t$  του προηγούμενου βήματος και με τη χρήση της λογικής λειτουργίας OR καταλήγουμε στον πίνακα  $IF_t$  ο οποίος περιέχει τα αντικείμενα που μας έχει αφήσει η υλοποίηση των παραπάνω βημάτων, έχοντας εξαλείψει ανεπιθύμητα κομμάτια που περιείχαν σκιές και θα προκαλούσαν σημαντικά προβλήματα στον εντοπισμό των οχημάτων στη συνέχεια. Για το τελικό αποτέλεσμα (Εικόνα 3.7 (κ) και Εικόνα 3.8 (κ)) στο μετασχηματισμένο χώρο χρησιμοποιείται το στατικό φόντο της επεξεργασίας για κάθε βίντεο, που περιγράφηκε στην ενότητα 3.2, ώστε να αντικατασταθούν τα pixels εκείνα στα οποία βρέθηκε η ύπαρξη σκιάς.



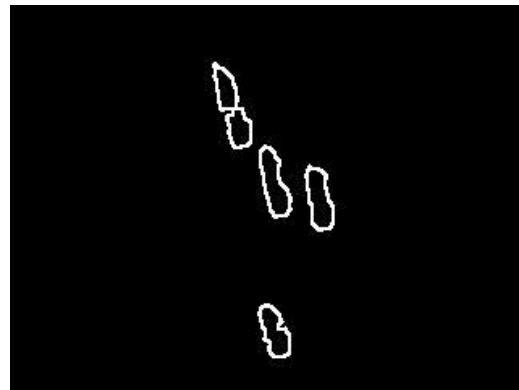
(α)



(β)



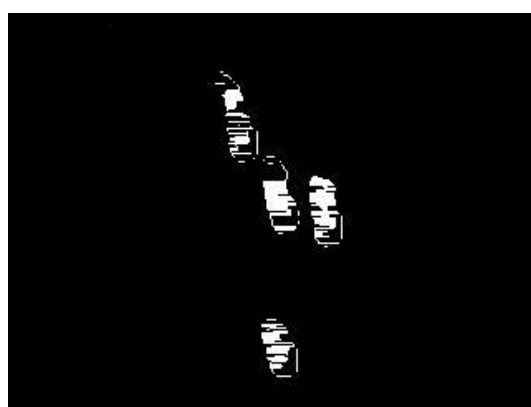
(γ)



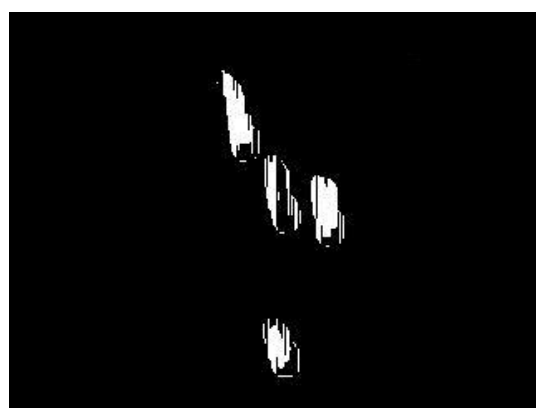
(δ)



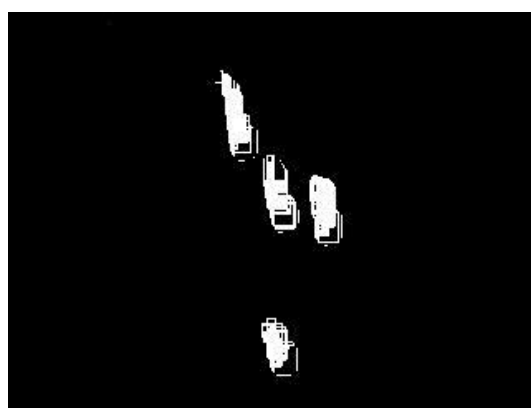
(ε)



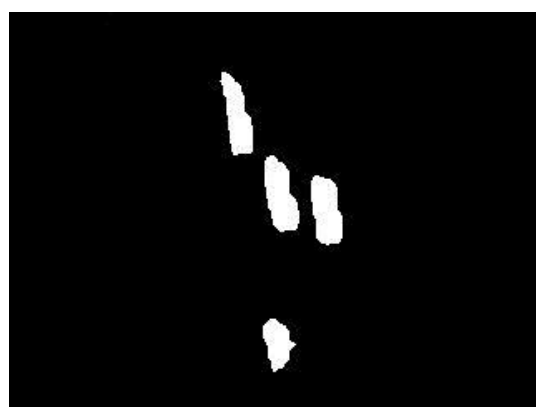
(ζ)



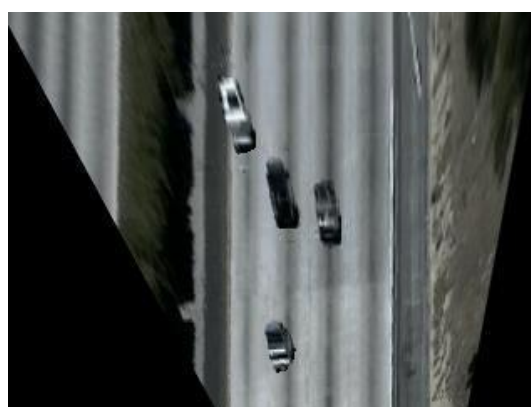
(η)



(θ)

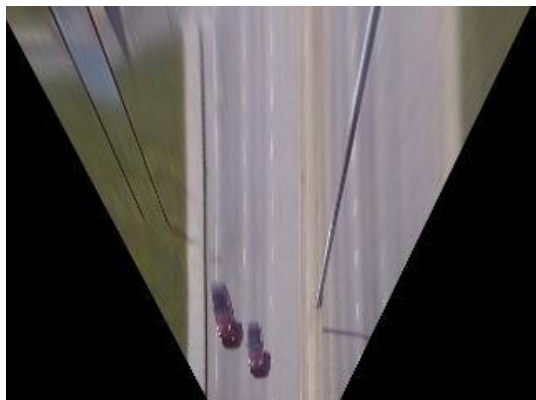


(ι)



(κ)

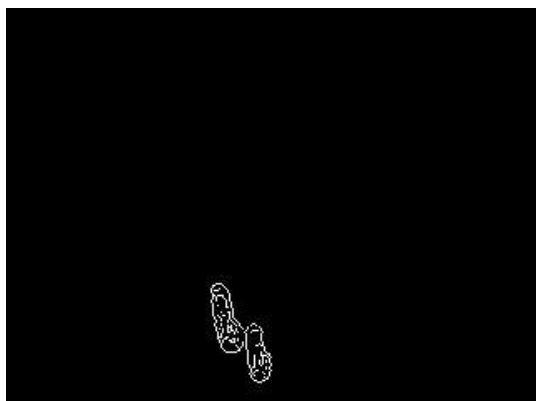
**Εικόνα 3.7** Βίντεο #A4 (α) Frame 350, (β)  $C_t$ , (γ)  $E_t$ , (δ)  $DB_t$ , (ε)  $IE_t$ , (ζ)  $HR_t$ , (η)  $VR_t$ , (θ)  $RE_t$ , (ι)  $IF_t$ , (κ) Αποτέλεσμα Edge-Based μεθόδου για το frame 350



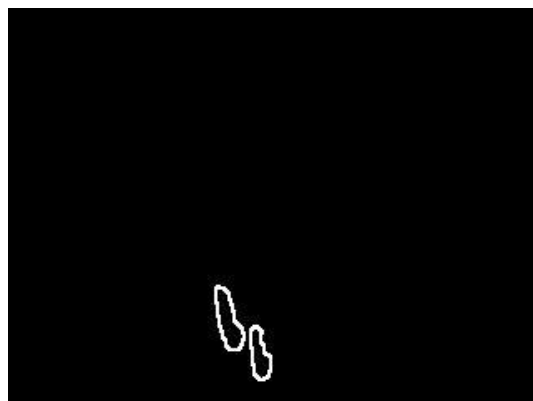
(α)



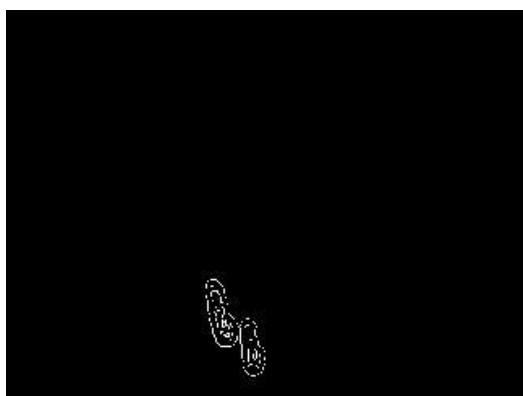
(β)



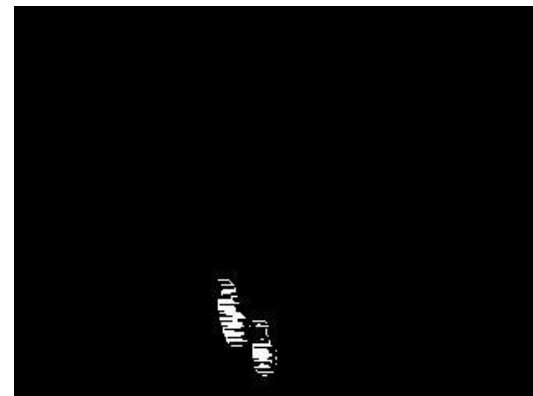
(γ)



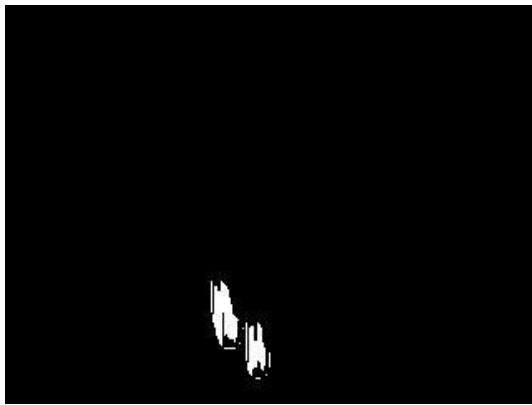
(δ)



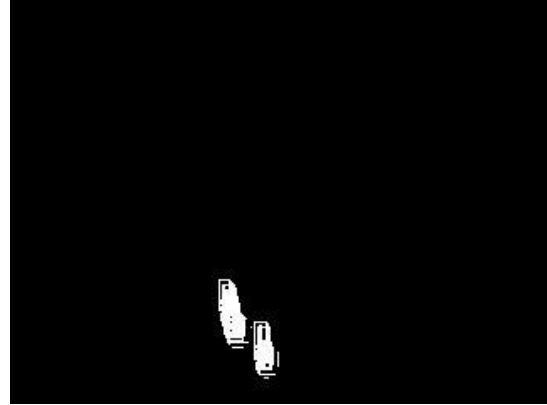
(ε)



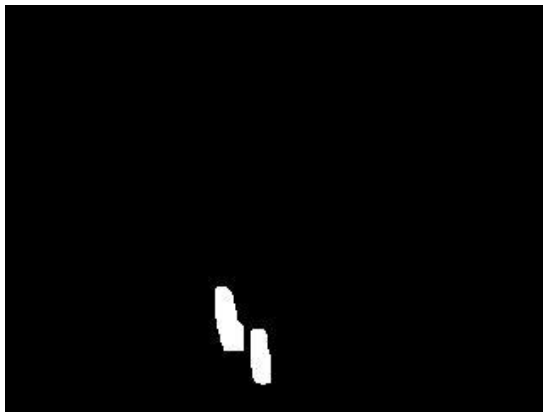
(ζ)



(η)



(θ)



(ι)



(κ)

**Εικόνα 3.8** Βίντεο #B2 (α) Frame 749, (β)  $C_t$ , (γ)  $E_t$ , (δ)  $DB_t$ , (ε)  $IE_t$ , (ζ)  $HR_t$ , (η)  $VR_t$ , (θ)  $RE_t$ , (ι)  $IF_t$ , (κ) Αποτέλεσμα Edge-Based μεθόδου για το frame 749

### 3.3.2 Σύγκριση Αφαίρεσης Σκιάς σε Δύο Χώρους

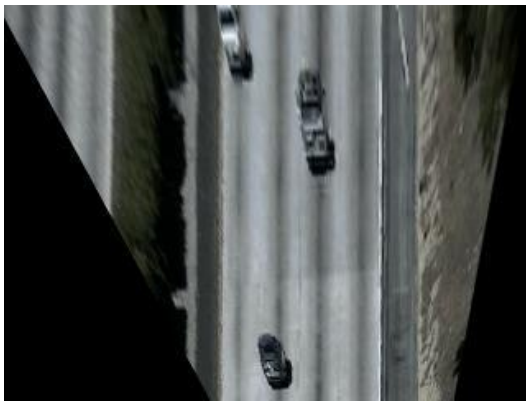
Βλέποντας τις Εικόνα 3.7 (κ) και Εικόνα 3.8 (κ) παρατηρούμε ότι η μέθοδος αφαίρεσης σκιάς που χρησιμοποιήθηκε, επιστρέφει αρκετά καλά αποτελέσματα και διώχνει τις ανεπιθύμητες σκιές κατά πολύ μεγάλο ποσοστό, διατηρώντας παράλληλα τα οχήματα χωρίς να υπάρχει σημαντική απώλεια πληροφορίας. Σε αυτήν εδώ την υποενότητα παρουσιάζονται τα αποτελέσματα της ίδιας μεθόδου, όταν αυτή εφαρμοστεί στα αρχικά frames των βίντεο, πριν αυτά υποστούν το μετασχηματισμό προοπτικής που αναλύθηκε στο δεύτερο κεφάλαιο. Έτσι θα μπορέσει να συγκριθεί αλλά και να σχολιαστεί η αποτελεσματικότητα της Edge-Based μεθόδου αφαίρεσης σκιάς στον πραγματικό και στο μετασχηματισμένο χώρο.



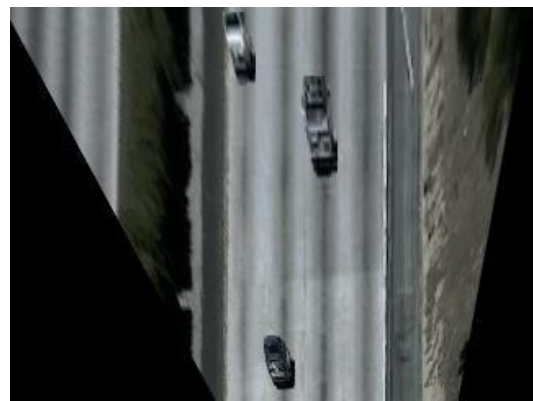
(α)



(β)



(γ)



(δ)

**Εικόνα 3.9** Βίντεο #A3 (α) Frame 293, (β) Αφαίρεση σκιάς για το (α), (γ) Το frame 293 μετά το μετασχηματισμό προοπτικής, (δ) Αφαίρεση σκιάς για το (γ)



(α)



(β)



(γ)



(δ)

**Εικόνα 3.10** Βίντεο #B1 (α) Frame 202, (β) Αφαίρεση σκιάς για το (α), (γ) Το frame 202 μετά το μετασχηματισμό προοπτικής, (δ) Αφαίρεση σκιάς για το (γ)

Σύμφωνα με τα αποτελέσματα που επιστρέφει η Edge-Based μέθοδος αφαίρεσης σκιάς, όπως αυτά που παραθέτονται στις Εικόνα 3.9 και Εικόνα 3.10, μπορεί εύκολα κανείς να συμπεράνει ότι είναι μια μέθοδος που μπορεί να διαχειριστεί σκιές ανεξάρτητα από την είσοδο. Στο σύστημά μας αντιμετωπίστηκαν με αρκετή επιτυχία οι περισσότερες περιοχές που περιείχαν σκιές, διαχωρίζοντας τα οχήματα από αυτές. Η μέθοδος εφαρμόστηκε στο σύστημά μας στα 11 βίντεο του δείγματος πειραμάτων που αποτελούνταν από frames στον πραγματικό χώρο και στα 11 βίντεο του ίδιου δείγματος που αποτελούνταν από frames στο μετασχηματισμένο χώρο. Στις δύο αυτές περιπτώσεις αφαιρέθηκαν σκιές, ανεξάρτητα από τον αριθμό των οχημάτων που εμφανίζονταν σε κάθε frame. Όμως όπως φαίνεται και από τις παραπάνω εικόνες, υπάρχουν και περιπτώσεις για τις οποίες σε κάποια οχήματα δεν αφαιρέθηκε η σκιά στο σύνολό της. Άρα λοιπόν παρόλο που η μέθοδος μπορεί να εφαρμοστεί εξίσου και στους δύο χώρους (πραγματικό και μετασχηματισμένο), χρήζει ακόμα αρκετής βελτίωσης. Από την επιτυχία της μεθόδου αφαίρεσης σκιάς, για εφαρμογές παρόμοιες με αυτήν που περιγράφεται στην παρούσα διπλωματική επεξεργασία, εξαρτάται σε μεγάλο βαθμό και η επιτυχία του σταδίου εντοπισμού των κινούμενων οχημάτων που περιγράφεται στο επόμενο κεφάλαιο. Η επιλογή της συγκεκριμένης μεθόδου έγινε μετά από αρκετές δοκιμές αντίστοιχων μεθόδων που απέτυχαν πλήρως να αντιμετωπίσουν τις σκιές και θα δημιουργούσαν μεγαλύτερα προβλήματα στα επόμενα στάδια της επεξεργασίας βίντεο.

## **Κεφάλαιο 4 : Εντοπισμός, Ταχύτητα και Ταξινόμηση Οχήματος**

Σε αυτό το κεφάλαιο παρουσιάζονται τα τελικά αποτελέσματα της παρούσας διπλωματικής εργασίας. Περιγράφεται η μέθοδος εντοπισμού οχήματος Kalman Filter που χρησιμοποιήθηκε, πάνω στην οποία βασίζονται και τα τελικά αποτελέσματα, παρουσιάζοντας παράλληλα τα προβλήματα που προκύπτουν από μη σωστό εντοπισμό. Επίσης γίνεται προσπάθεια, όπου είναι εφικτό, να δοθούν λύσεις στα προβλήματα αυτά. Περιγράφεται επίσης η μέθοδος που υλοποιήθηκε για τον υπολογισμό της ταχύτητας των οχημάτων σε κάθε frame και τέλος γίνεται ταξινόμηση των οχημάτων που εντοπίστηκαν σε δύο κατηγορίες με βάση δύο από τα χαρακτηριστικά τους.

### **4.1 Εισαγωγή**

Ο εντοπισμός αντικειμένων είναι μια πολύ σημαντική και αρκετά δύσκολη διαδικασία στον τομέα της όρασης υπολογιστών. Ο στόχος της είναι ο καθορισμός μιας αντιστοιχίας μεταξύ των αντικειμένων ή τμημάτων των αντικειμένων σε διαδοχικά frames όπως επίσης και να εξαχθούν πληροφορίες σχετικά με αντικείμενα, όπως η πορεία, η ταχύτητα και η κατεύθυνση. Είναι ένα κρίσιμο κομμάτι των έξυπνων συστημάτων επιτήρησης δεδομένου ότι χωρίς τη διαδικασία εντοπισμού αντικειμένων, οι εφαρμογές δε θα μπορούσαν να εξάγουν συνεκτικές χρονικές πληροφορίες σχετικά με την κίνηση των αντικειμένων και δε θα ήταν δυνατή η περαιτέρω ανάλυση της συμπεριφοράς τους σε υψηλότερο επίπεδο. Από την άλλη, η μη ακριβής κατάτμηση στην εικόνα που περιέχει τα αντικείμενα λόγω σκιάς, ανάκλασης και occlusion καθιστά τη διαδικασία ένα δύσκολο πρόβλημα. Έτσι, η διαχείριση των κινούμενων σκιών στο στάδιο της ανίχνευσης κίνησης και η αντιμετώπιση του φαινομένου occlusion, τόσο κατά τη διαδικασία της κατάτμησης όσο και κατά τη διαδικασία του εντοπισμού είναι σημαντική για την αποτελεσματικότητα του συστήματος.

Ο εντοπισμός αντικειμένου σε βίντεο μπορεί να κατηγοριοποιηθεί ανάλογα με τις ανάγκες της εκάστοτε εφαρμογής που χρησιμοποιείται ή σύμφωνα με τις μεθόδους που χρησιμοποιούνται για τη λύση του. Ο εντοπισμός ολόκληρων των

αντικειμένων είναι γενικά επαρκής για βίντεο σε εξωτερικά περιβάλλοντα, ενώ αντίθετα για κάποια βίντεο σε εσωτερικούς χώρους ο εντοπισμός κομματιών του αντικειμένου είναι απαραίτητος όπως και σε ορισμένες εφαρμογές υψηλότερου επίπεδου συμπεριφοράς και κατανόησης.

Το στάδιο του εντοπισμού οχήματος, ακολουθεί το στάδιο υπολογισμού της ταχύτητας τους. Τα έξυπνα συστήματα μεταφορών, ITS, γίνονται ολοένα και πιο σημαντικά λόγω των πλεονεκτημάτων τους όπως το να σώζουν ζωές, χρήματα και χρόνο. Η μεγάλη αύξηση της κυκλοφοριακής συμφόρησης στις πόλεις έχει γίνει ένα πολύ κρίσιμο πρόβλημα και ο αριθμός των ατυχημάτων που οφείλονται σε υπερβολική ταχύτητα συνεχίζει να αυξάνεται. Η απόκτηση πληροφοριών για την κίνηση σε κύριους δρόμους ή αυτοκινητόδρομους είναι επιτακτική ανάγκη, καθώς παρέχει στους μηχανικούς κυκλοφορίας πολύ γόνιμες πληροφορίες. Τέτοιες είναι η πυκνότητα της κυκλοφορίας (ο αριθμός του συνόλου των οχημάτων σε μια συγκεκριμένη περιοχή για μια δεδομένη στιγμή) και οι ταχύτητες των οχημάτων, που αποτελούν και τα βασικά μέρη του ITS καθώς οι πληροφορίες αυτές χρησιμοποιούνται για τη διαχείριση και τον έλεγχο της κυκλοφορίας. Η παρακολούθηση της ταχύτητας των οχημάτων είναι σημαντική για την επιβολή των νόμων ορίου ταχύτητας στους δρόμους. Υποδεικνύει επίσης τις συνθήκες κυκλοφορίας του τμήματος του δρόμου ώστε να αποτρέψει τους οδηγούς από υπερβολική ταχύτητα, με σκοπό έτσι τη μείωση των ατυχημάτων.

Τέλος σε αυτό το κεφάλαιο παρουσιάζεται μια απλή μέθοδος ταξινόμησης οχήματος [59]. Μια από τις σημαντικότερες διαδικασίες του χαρακτηρισμού κίνησης είναι η αναγνώριση του τύπου και η ταξινόμηση του οχήματος. Για τον προσδιορισμό του τύπου του οχήματος, το βασικό πρόβλημα έγκειται στην επιλογή αποτελεσματικών χαρακτηριστικών του οχήματος και μιας μεθόδου αναγνώρισης προτύπου. Ένα βίντεο, όπως αυτά που επεξεργάζεται το σύστημά μας, μπορεί να περιέχει μεγάλη ποικιλία από αντικείμενα, όπως ανθρώπους, οχήματα, ζώα και φυσικά φαινόμενα (βροχή, χιόνι). Ωστόσο, ο κύριος στόχος σε μια εφαρμογή επιτήρησης της κυκλοφορίας είναι τα οχήματα. Λόγω των σημαντικών διαφορών στο μέγεθος και το βάρος μεταξύ ενός μεγάλου οχήματος και ενός μικρού οχήματος, η ταξινόμηση των οχημάτων είναι θεμελιώδους σημασίας για τη λειτουργία της κυκλοφορίας και το σχεδιασμό των μεταφορών. Η γεωμετρική σχεδίαση ενός δρόμου επηρεάζεται από τα διαφορετικά χαρακτηριστικά κίνησης των μεγάλων οχημάτων, λόγω του μεγάλου βάρους τους. Το μεγάλο βάρος αυτών των οχημάτων είναι επίσης σημαντικό για το σχεδιασμό και τη συντήρηση του οδοστρώματος, καθώς ο όγκος φορτηγών επηρεάζουν τόσο τη ζωή του οδοστρώματος όσο και τις παραμέτρους σχεδιασμού του.



## 4.2 Εντοπισμός Οχήματος

Ο στόχος της μεθόδου εντοπισμού οχήματος, όπως αναφέρθηκε και παραπάνω, δεν είναι άλλος από τον καθορισμό μιας αντιστοιχίας μεταξύ των αντικειμένων ή τμημάτων των οχημάτων που εμφανίζονται σε διαδοχικά frames του βίντεο που επεξεργάζεται το σύστημα. Στην απλούστερη μορφή της, η μέθοδος μπορεί να οριστεί ως το πρόβλημα εκτίμησης της τροχιάς ενός οχήματος στο επίπεδο της εικόνας καθώς αυτό κινείται. Με άλλα λόγια, εκχωρούνται σταθερές ετικέτες στα εντοπισμένα οχήματα στα διαδοχικά frames που αυτά εμφανίζονται και μέχρι να αφήσουν το πεδίο που καταγράφει η κάμερα. Επίσης, ανάλογα με τη μέθοδο που χρησιμοποιείται, μπορούν ακόμα εξαχθούν επιπρόσθετες πληροφορίες για κάθε όχημα, όπως η κατεύθυνση, το εμβαδό, το μήκος ή το σχήμα του. Στην παρούσα διπλωματική εργασία, υλοποιήθηκε, με τη βοήθεια της τεχνικής γλώσσας υπολογιστών Matlab, ο αλγόριθμος Kalman Filter [60], [61]. Ένα φίλτρο Kalman χρησιμοποιείται στη διαδικασία εντοπισμού για να προβλέψει τις θέσεις των οχημάτων στα μελλοντικά frames του βίντεο. Τα πλεονεκτήματα της συμπερίληψης του φίλτρου Kalman στη διαδικασία εντοπισμού είναι :

- Παρέχει την καλύτερη εκτίμηση θέσης ώστε να γίνει έλεγχος εκεί για τα οχήματα στο επόμενο frame του βίντεο και ως εκ τούτου βελτιώνει την ταχύτητα επανανίχνευσης.
- Μειώνει την περιοχή αναζήτησης για εκ νέου ανίχνευση ενός οχήματος και έτσι μειώνεται ο χρόνος επεξεργασίας.
- Μπορεί να μειώσει τον αριθμό των εσφαλμένων ανιχνεύσεων δεδομένου ότι η περιοχή της εικόνας που δεν περιέχει όχημα αποκλείεται από την αναζήτηση.

### 4.2.1 Αλγόριθμος Kalman Filter

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

$$x_k = F_{k-1}x_{k-1} + w_{k-1} \quad (4.1)$$

$$y_k = H_k x_k + v_k \quad (4.2)$$

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 :

$$w_k : (0, Q_k) \quad (4.3)$$

$$v_k : (0, R_k) \quad (4.4)$$

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

- Πρόβλεψη (prediction)
- 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. 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 συνοψίζεται οι δύο φάσεις του αλγορίθμου. Η εξίσωση της φάσης πρόβλεψης για τον υπολογισμό της a priori κατάστασης εκτίμησης  $\hat{x}_k^-$  και η συνδιακύμανση σφάλματος  $P_k^-$  δίνονται από τις :

$$\hat{x}_k^- = F_{k-1} \hat{x}_{k-1}^+ \quad (4.5)$$

$$P_k^- = F_{k-1} P_{k-1}^+ F_{k-1}^T \quad (4.6)$$

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 :

$$\hat{x}_k^+ = \hat{x}_k^- + K_k (y_k - H_k \hat{x}_k^-) \quad (4.7)$$

$$P_k^+ = (I - K_k H_k) P_k^- \quad (4.8)$$

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

$$K_k = \frac{P_k^- H_k^T}{H_k P_k^- H_k^T + R_k} \quad (4.9)$$

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.

Ο αλγόριθμος που υλοποιείται στην παρούσα διπλωματική εργασία, χρησιμοποιεί ένα φίλτρο Kalman για να προβλέψει την κίνηση ενός οχήματος και τις αλλαγές στο μέγεθος του στο επίπεδο της εικόνας. Η κίνηση των οχημάτων μεταξύ δύο διαδοχικών frames του βίντεο είναι μικρή. Έτσι λοιπόν, αυτή η παραδοχή είναι αρκετή για να μοντελοποιηθεί μια τέτοια κίνηση, ως κίνηση με σταθερή ταχύτητα [62]. Οι παραδοχές επιτρέπουν στο σύστημα να μοντελοποιηθεί χρησιμοποιώντας το γραμμικό φίλτρο Kalman διακριτού χρόνου που περιγράφεται σε αυτήν την υποενότητα. Το φίλτρο Kalman μοντελοποιεί την κατανομή πιθανότητας της κάθε κατάστασης ως μια Gaussian κατανομή. Αυτό σημαίνει ότι κάθε φίλτρο Kalman μπορεί να εντοπίσει μόνο ένα όχημα 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 οχήματα σε κάθε frame, ένα ξεχωριστό φίλτρο Kalman αρχικοποιείται για κάθε διαφορετικό όχημα στη λίστα με τα οχήματα προς εντοπισμό. Η a priori εκτίμηση του φίλτρου Kalman υποδεικνύει τη θέση και το μέγεθος της περιοχής όπου ένα όχημα θα μπορούσε ενδεχομένως να εμφανιστεί στο επόμενο frame του βίντεο. Οι πληροφορίες αυτές χρησιμοποιούνται από τη διαδικασία εντοπισμού για να περιοριστεί η αναζήτηση χώρου για την εκ νέου ανίχνευση του οχήματος. Μόλις ολοκληρωθεί η εκ νέου ανίχνευση, τα νέα δεδομένα των μετρήσεων θα πρέπει να προσαρμοστούν στο μοντέλο του συστήματος. Η a posteriori εκτίμηση υπολογίζεται στη συνέχεια και χρησιμοποιείται ως η καλύτερη εκτίμηση για τη θέση και το μέγεθος του οχήματος.

#### 4.2.2 Υλοποίηση Αλγορίθμου

Για την υλοποίηση του εντοπισμού οχήματος με τον αλγόριθμο Kalman Filter στην παρούσα διπλωματική εργασία χρησιμοποιήθηκε το παράδειγμα που υπάρχει στο εγχειρίδιο της τεχνικής γλώσσας υπολογιστών Matlab για τον εντοπισμό πολλαπλών αντικειμένων [63]. Η αντιστοίχιση των ανιχνεύσεων, σε διαδοχικά frames, για το ίδιο όχημα βασίζεται αποκλειστικά στην κίνηση. Η κίνηση αυτή του κάθε οχήματος υπολογίζεται από ένα φίλτρο Kalman. Το φίλτρο χρησιμοποιείται για να προβλέψει την τοποθεσία του οχήματος σε κάθε frame και για να καθορίσει την πιθανότητα κάθε ανίχνευσης που αντιστοιχίζεται σε κάθε όχημα. Σε κάθε frame, κάποιες ανιχνεύσεις μπορεί να αντιστοιχιστούν σε οχήματα, ενώ άλλες ανιχνεύσεις και οχήματα μπορεί να παραμείνουν χωρίς αντιστοίχιση. Τα οχήματα που έχουν αντιστοιχισθεί, ενημερώνονται χρησιμοποιώντας τις αντίστοιχες ανιχνεύσεις. Για κάθε όχημα

αποθηκεύεται ο αριθμός των διαδοχικών frames για τα οποία παρέμεινε χωρίς αντιστοίχιση. Εάν ο αριθμός αυτός υπερβαίνει ένα συγκεκριμένο κατώφλι, που για το σύστημά μας η τιμή του ορίστηκε να είναι τα 3 διαδοχικά frames, το σύστημα υποθέτει ότι το όχημα έφυγε από το οπτικό πεδίο και το διαγράφει.

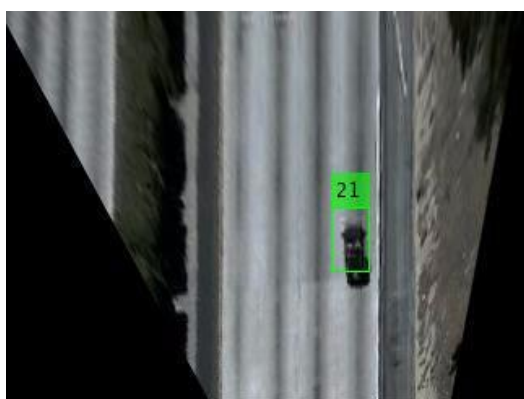
Το σύστημα έχει σαν είσοδο το βίντεο, frame προς frame, που επεξεργαζόμαστε και τα αποτελέσματα της διαδικασίας ανίχνευσης οχήματος, που παρουσιάστηκαν αναλυτικά στο τρίτο κεφάλαιο, μετά και από την αφαίρεση των σκιών που βρέθηκαν. Αρχικά δημιουργείται ένας πίνακας όπου κάθε θέση του είναι μια δομή που αντιπροσωπεύει ένα κινούμενο όχημα στο βίντεο. Ο σκοπός της δομής είναι να διατηρήσει την κατάσταση ενός οχήματος που προσπαθεί να εντοπιστεί από τον αλγόριθμο. Η δομή αυτή αποτελείται από πληροφορίες όπως είναι το ξεχωριστό id που δίνεται σε κάθε ένα από τα οχήματα που εντοπίζονται, το μήκος του και το εμβαδό του σε pixels, όπως επίσης και κάποιες επιπρόσθετες πληροφορίες που βοηθούν στην προσπάθεια για σωστό εντοπισμό (μια μεταβλητή για παράδειγμα, που κρατάει την τιμή που δείχνει στο σύστημα για πόσα frames εμφανίζεται το συγκεκριμένο όχημα ή για πόσα συνεχόμενα frames δεν εντοπίστηκε καθόλου). Έπειτα βρίσκονται τα κέντρα βάρους των οχημάτων που ανιχνεύθηκαν ανά frame και στη συνέχεια χρησιμοποιείται το φίλτρο Kalman ώστε να προβλεφθούν τα κέντρα βάρους των αντίστοιχων οχημάτων και στα επόμενα frames. Το κέντρο βάρους [64] ενός επίπεδου σχήματος ή ενός σχήματος δύο διαστάσεων, είναι η θέση του αριθμητικού μέσου όλων των σημείων του σχήματος. Ο ορισμός επεκτείνεται σε οποιοδήποτε αντικείμενο στο χώρο  $n$ -διαστάσεων : το κέντρο βάρους του είναι η μέση θέση όλων των σημείων σε όλες τις κατευθύνσεις συντεταγμένων. Χρησιμοποιώντας το φίλτρο Kalman, προβλέπεται το κέντρο βάρους του κάθε οχήματος στο τρέχον frame και ενημερώνεται το κουτί οριοθέτησης (bounding box) που το περιέχει, αναλόγως.

Η αντιστοίχιση ανίχνευσης οχήματος στο τρέχον frame με τα υπάρχοντα οχήματα που έχουν εντοπιστεί, γίνεται με ελαχιστοποίηση του κόστους. Ο υπολογισμός του κόστους της αντιστοίχισης κάθε ανίχνευσης σε κάθε όχημα γίνεται χρησιμοποιώντας τη μέθοδο της απόστασης. Το κόστος λαμβάνει υπόψη την Ευκλείδεια απόσταση ανάμεσα στο κέντρο βάρους που έχει προβλεφθεί για το όχημα και του κέντρου βάρους της ανίχνευσης. Τα αποτελέσματα αποθηκεύονται σε έναν πίνακα με διαστάσεις  $M \times N$ , όπου  $M$  είναι ο αριθμός των οχημάτων και  $N$  είναι ο αριθμός των ανιχνεύσεων. Ο αλγόριθμος διαγράφει οχήματα που δε βρέθηκαν για συνεχόμενα frames. Επόμενο βήμα είναι η ενημέρωση κάθε αντιστοιχισμένου οχήματος με την αντίστοιχη ανίχνευση, ώστε να διορθωθεί η εκτίμηση της τοποθεσίας του. Στη συνέχεια, αποθηκεύεται το κουτί οριοθέτησης των οχημάτων και ενημερώνονται επίσης οι υπόλοιπες μεταβλητές της δομής οχήματος. Τέλος οχήματα που δεν έχουν αντιστοιχισθεί με ανιχνεύσεις για 3 διαδοχικά frames διαγράφονται από το σύστημα όπως αναφέρθηκε και παραπάνω.

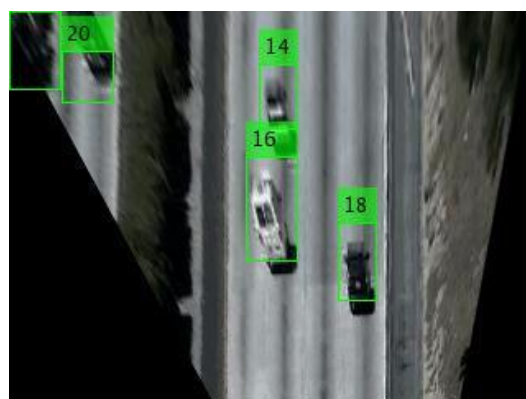
Ο εντοπισμός οχήματος σε αυτήν τη μέθοδο βασίζεται αποκλειστικά στην κίνηση, με την παραδοχή ότι όλα τα οχήματα κινούνται σε ευθεία γραμμή με σταθερή ταχύτητα. Όταν η κίνηση ενός οχήματος αποκλίνει σημαντικά από το μοντέλο αυτό, η μέθοδος προκαλεί σφάλματα εντοπισμού. Αναλυτικά τα αποτελέσματα που επιστρέφει ο αλγόριθμος που υλοποιήθηκε, παρουσιάζονται στην αμέσως επόμενη υποενότητα του κεφαλαίου αυτού. Επίσης στην ίδια υποενότητα παρουσιάζονται, μέσω των αποτελεσμάτων που επιστρέφει ο αλγόριθμος, τα προβλήματα που προέκυψαν κατά τον εντοπισμό και περιγράφονται αναλυτικά, μαζί με την προσπάθεια αντιμετώπισης τους, στην υποενότητα 4.2.4.

#### 4.2.3 Αποτελέσματα Αλγορίθμου

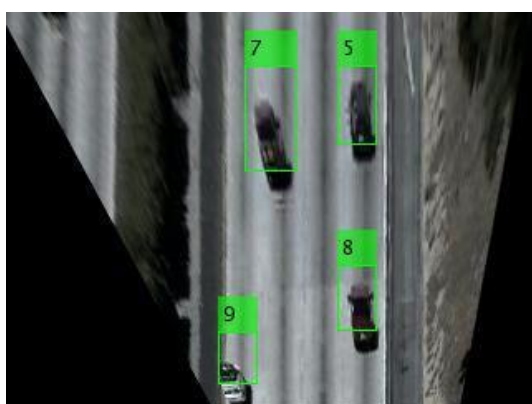
Στις εικόνες που ακολουθούν παρουσιάζεται η έξοδος της μεθόδου εντοπισμού οχήματος που υλοποιήθηκε, για τα 11 βίντεο-εισόδους του συστήματός μας.



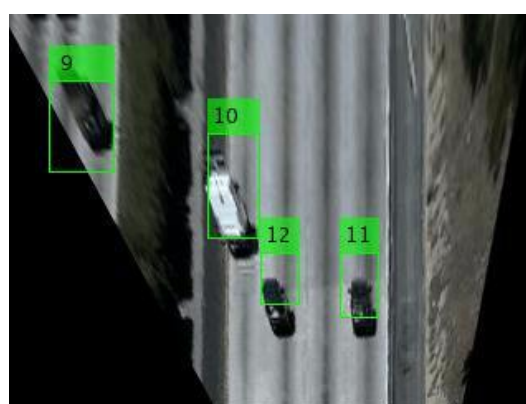
(α)



(β)



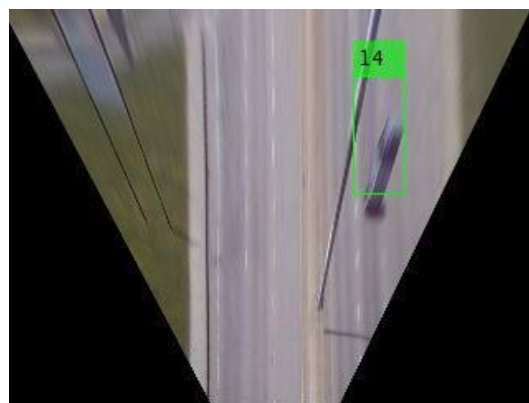
(γ)



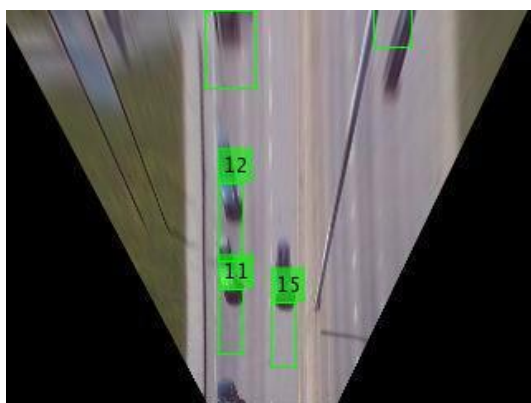
(δ)



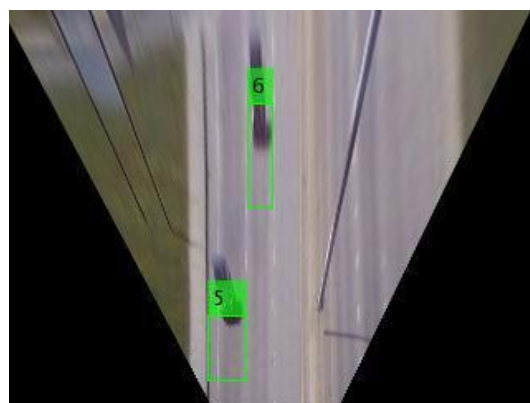
(ε)



(ζ)



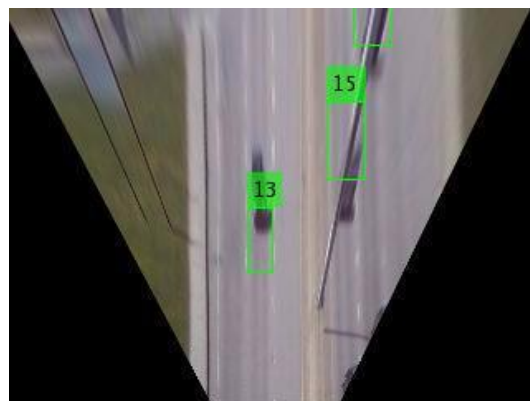
(η)



(θ)



(ι)



(κ)



(λ)

**Εικόνα 4.1** (α) Βίντεο #A1 frame 98, (β) Βίντεο #A2 frame 188, (γ) Βίντεο #A3 frame 243, (δ) Βίντεο #A4 frame 383, (ε) Βίντεο #A5 frame 461, (ζ) Βίντεο #B1 frame 270, (η) Βίντεο #B2 frame 785, (θ) Βίντεο #B3 frame 1136, (ι) Βίντεο #B4 frame 1656, (κ) Βίντεο #B5 frame 2070, (λ) Βίντεο #B6 frame 2131

Στις Εικόνα 4.1 (α)-(λ) βλέπουμε το στιγμιότυπα από τα αποτελέσματα που επιστρέφει ο αλγόριθμος Kalman Filter. Για κάθε ξεχωριστό όχημα που εντοπίζεται, όπως φαίνεται και από τις παραπάνω εικόνες, σχεδιάζεται ένα κουτί οριοθέτησης μαζί με ένα id και αυτά το ακολουθούν έως ότου εξαφανιστεί από το οπτικό πεδίο της κάμερας. Ο αλγόριθμος φαίνεται ότι δουλεύει αρκετά καλά και εντοπίζει όσα οχήματα κινούνται σε κάθε frame. Στις Εικόνα 4.1 (ζ)-(λ) παρουσιάζονται τα αποτελέσματα του αλγορίθμου για τα Βίντεο #B1-B6 και σε σχέση με τα υπόλοιπα πέντε βίντεο των Εικόνα 4.1 (α)-(ε) η ανάλυση της εικόνας δεν είναι καλή, με αποτέλεσμα τα οχήματα που απομακρύνονται από την κάμερα στις εικόνες αυτές να μη φαίνονται καλά όπως και αυτά στη δεξιά λωρίδα να φαίνονται λοξά. Το φαινόμενο αυτό παρατηρείται εξαιτίας της τοποθέτησης τη στατικής κάμερας στον πραγματικό χώρο, αλλά και της επιλογής των χαρακτηριστικών σημείων του μετασχηματισμού προοπτικής καθώς ο δρόμος και τα οχήματα ουσιαστικά “τεντώνονται” στα pixels που πρέπει να αναπαρασταθούν και η ανάλυση χαλάει. Οι παράμετροι που επηρεάζουν την ανάλυση της εικόνας στο μετασχηματισμένο χώρο είναι η κλίση και η γωνία που είναι τοποθετημένη η κάμερα όπως και η απόστασή της από το έδαφος, καθώς όσο μεγαλύτερο είναι το ύψος στο οποίο τοποθετείται, τόσο μεγαλύτερο είναι και το οπτικό της πεδίο με αποτέλεσμα όταν υποστούν τα αρχικά frames το μετασχηματισμό προοπτικής να μη μπορεί να απεικονιστεί ο χώρος που βρίσκεται μακριά από αυτή με καλή ανάλυση και να παραμορφώνεται ελαφρά. Στην Εικόνα 4.1 (λ) επίσης, παρατηρείται ότι έχει εντοπιστεί ένα όχημα πάνω και αριστερά στην εικόνα που φαίνεται να μην κινείται σε κανένα από τα δύο ρεύματα κυκλοφορίας του αυτοκινητόδρομου. Αυτό συμβαίνει διότι υπάρχει ένας παράδρομος σε εκείνο το

σημείο που βρίσκεται στο οπτικό πεδίο της κάμερας, με αποτέλεσμα να εντοπίζονται και τα οχήματα κινούνται σε αυτόν. Τέλος, παρατηρούμε ότι το κουτί οριοθέτησης τοποθετείται λίγο πιο μπροστά σε σχέση με την παρούσα θέση του οχήματος. Αυτό συμβαίνει γιατί ο αλγόριθμος το τοποθετεί στην επόμενη θέση που έχει προβλεφθεί ότι θα βρίσκεται το όχημα στο αμέσως επόμενο και όχι στο παρών frame.

Στους παρακάτω πίνακες (Πίνακας 4.1 και Πίνακας 4.2) παρουσιάζονται ποσοτικά αποτελέσματα της μεθόδου εντοπισμού οχήματος, με είσοδο τα 11 βίντεο που επεξεργάζεται το σύστημά μας και υπολογίζεται η ακρίβεια του για κάθε ένα ξεχωριστά.

Είσοδος	Βίντεο #A1	Βίντεο #A2	Βίντεο #A3	Βίντεο #A4	Βίντεο #A5
<b>Frames</b>	100	100	100	100	100
<b>Αριθμός Οχημάτων</b>	22	16	17	15	19
<b>Αριθμός Εντοπισμένων Οχημάτων</b>	21	20	15	14	21
<b>Ακρίβεια</b>	95.45%	75%	88.26%	93.33%	89.47%

**Πίνακας 4.1**

Είσοδος	Βίντεο #B1	Βίντεο #B2	Βίντεο #B3	Βίντεο #B4	Βίντεο #B5	Βίντεο #B6
<b>Frames</b>	100	100	100	100	100	100
<b>Αριθμός Οχημάτων</b>	27	23	19	23	28	36
<b>Αριθμός Εντοπισμένων Οχημάτων</b>	20	18	18	22	24	24
<b>Ακρίβεια</b>	74.07%	78.26%	94.74%	95.65%	85.71%	66.67%

**Πίνακας 4.2**



Η ακρίβεια υπολογίζεται ως εξής :

$$\text{Ακρίβεια} = \frac{\text{Αριθμός Οχημάτων} - \text{Error}}{\text{Αριθμός Οχημάτων}} 100\% \quad (4.10)$$

όπου Error είναι ο αριθμός των οχημάτων που μετρήθηκαν από λάθος, συμπεριλαμβάνοντας αυτά που δεν εντοπίστηκαν και δε μετρήθηκαν καθόλου αλλά και αυτά που μετρήθηκαν παραπάνω από μια φορά. Ο Αριθμός Οχημάτων είναι ο πραγματικός αριθμός των οχημάτων που μετρήθηκαν ανά βίντεο.

Ο αλγόριθμος που υλοποιήθηκε παρουσιάζει αρκετά υψηλή ακρίβεια εντοπισμού κινούμενων οχημάτων, σύμφωνα με τα αποτελέσματα που παρουσιάζονται στους Πίνακα 4.1 και Πίνακα 4.2. Για τα πέντε βίντεο του Πίνακα 4.1 η μέση ακρίβεια είναι 88.3%, ενώ για τα έξι βίντεο του Πίνακα 4.2 είναι 82.52%. Επίσης υπολογίζεται η συνολική μέση ακρίβεια όλων των βίντεο που εξετάστηκαν στο σύστημα και έχει τιμή 85.15%. Στην πραγματικότητα, το σύστημά μας μπορεί να ανιχνεύσει και να εντοπίσει σχεδόν κάθε όχημα που κινείται, στα βίντεο που αναλύονται. Βέβαια όπως φαίνεται και από τις σχετικά χαμηλές τιμές της ακρίβειας του αλγορίθμου για τα Βίντεο #A2, Βίντεο #B1 και Βίντεο #B6, κατά την εφαρμογή του εντοπισμού οχήματος προκύπτουν προβλήματα κάτω υπό συγκεκριμένες συνθήκες. Οι συνθήκες αυτές αλλά και η προσπάθεια επίλυσης αυτών των σφαλμάτων εντοπισμού περιγράφονται στην υποενότητα 4.2.4.

#### 4.2.4 Προβλήματα κατά τον Εντοπισμό Οχήματος

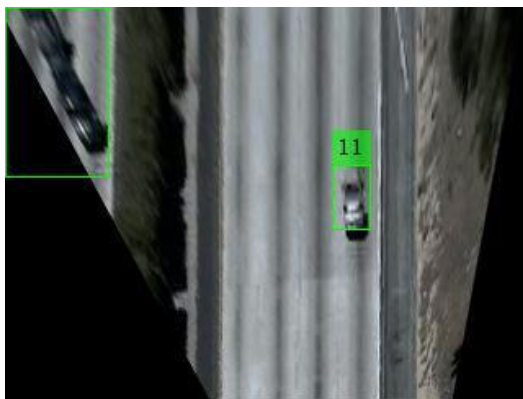
Στην όραση υπολογιστών μία απλή εικόνα, δηλαδή ένα frame ενός βίντεο για το σύστημά μας, δε δίνει καμία πληροφορία για τον εντοπισμό των οχημάτων που υπάρχουν σε αυτό. Η ανίχνευση και στη συνέχεια ο εντοπισμός ενός οχήματος από τον υπολογιστή είναι μία απαιτητική και δύσκολη διαδικασία που χρειάζεται να περάσει από πολλά στάδια όπως έχει περιγραφεί έως εδώ. Κατά τη διαδικασία της ανίχνευσης οχήματος, ανεξάρτητα από την τεχνική που χρησιμοποιείται, είναι σίγουρο ότι θα συναντηθούν αρκετά προβλήματα που όπως εύκολα γίνεται αντιληπτό, επηρεάζουν και τη διαδικασία εντοπισμού οχήματος. Ορισμένα από αυτά τα προβλήματα είναι αρκετά απλά, με συνέπεια η επίλυσή τους να επιτυγχάνεται εύκολα. Παρόλα αυτά, για κάποια άλλα απαιτούνται πιο δύσκολες και σύνθετες διαδικασίες που δεν έχουν σε όλες τις περιπτώσεις αποτέλεσμα διότι σε εφαρμογές και συστήματα, όπως αυτό που προτείνει η παρούσα διπλωματική εργασία, υπάρχουν ειδικές συνθήκες που καθιστούν αυτά τα προβλήματα μη επιλύσιμα. Τα προβλήματα που δυσχεραίνουν τη διαδικασία εντοπισμού οχήματος είναι :

- Ύπαρξη θορύβου κατά τη διαδικασία της αφαίρεσης του φόντου.

- Οι σκιές που ακολουθούν κάθε όχημα (τα βίντεο που χρησιμοποιούνται ως είσοδοι στο σύστημά μας έχουν ληφθεί κατά τη διάρκεια της ημέρας με την παρουσία ηλιοφάνειας).
- Εναλλαγές στο φωτισμό με αποτέλεσμα να δημιουργούνται διακυμάνσεις στην ένταση των pixels.
- Μερική ή ολική επικάλυψη οχημάτων μεταξύ τους.
- Σύνθετη μορφή κίνησης οχήματος.
- Ανάθεση ίδιου id σε διαφορετικά οχήματα

Ο θόρυβος και ο φωτισμός αντιμετωπίζονται σχετικά εύκολα, με τη χρήση φίλτρων για την εξομάλυνση του φωτισμού και τη μείωση του θορύβου. Όμως οι σκιές, η επικάλυψη οχημάτων και οχήματα που η κίνησή τους στο δρόμο είναι σύνθετη, δυσκολεύουν αρκετά τη διαδικασία εντοπισμού οχήματος. Όσον αφορά το πρόβλημα ύπαρξης σκιών, αναλύθηκε ο αλγόριθμος που χρησιμοποιήθηκε στην ενότητα 3.3 του τρίτου κεφαλαίου της διπλωματικής εργασίας. Επίσης ο αλγόριθμος αυτός βοήθησε ώστε ορισμένα φαινόμενα επικάλυψης οχημάτων από το στάδιο της ανίχνευσης να αντιμετωπιστούν στο στάδιο του εντοπισμού οχήματος. Στο πρόβλημα της επικάλυψης οχημάτων μπορεί υπό προϋποθέσεις να δώσει λύση και ο μετασχηματισμός προοπτικής, όπως αναφέρθηκε στο δεύτερο κεφάλαιο, όμως δυστυχώς σε κάποιες περιπτώσεις δεν αρκεί. Χαρακτηριστικό παράδειγμα είναι η πολύ χαμηλή ακρίβεια των Βίντεο #A2, Βίντεο #B1 και Βίντεο #B6 (Πίνακας 4.1-4.2) που οφείλεται σε λάθος εντοπισμούς λόγω επικάλυψης οχημάτων, καθώς παραπάνω από τα μισά οχήματα σε αυτά τα βίντεο δεν κινούνται στα ρεύματα κυκλοφορίας που έχει μέτωπο η κάμερα αλλά στα πλαϊνά για τα οποία όπως έχει αναφερθεί από το δεύτερο κεφάλαιο ο μετασχηματισμός προοπτικής σε αυτά το ρεύματα δε λειτουργεί ιδανικά. Σχεδόν όλοι οι αλγόριθμοι εντοπισμού, έτσι και ο Kalman Filter που υλοποιήθηκε στην παρούσα διπλωματική εργασία, υποθέτουν για σωστό εντοπισμό ότι η κίνηση των οχημάτων είναι ομαλή. Ομαλή χαρακτηρίζεται η κίνηση που δεν παρουσιάζει απότομες αλλαγές στην κατεύθυνση και στην ταχύτητα του οχήματος. Τέλος παρατηρήθηκαν περιπτώσεις όπου τη στιγμή που ένα εντοπισμένο όχημα αφήνει το οπτικό πεδίο της κάμερας, ένα νέο όχημα με παρόμοια χαρακτηριστικά εντοπίζεται ακριβώς τότε και παίρνει το ίδιο id αντί για ένα νέο (όπως το id=18 των Εικόνα 4.3 (α)-(β)). Αυτό έχει ως αποτέλεσμα να μη θεωρείται από το σύστημα ως νέο όχημα και επίσης να παρουσιάζεται λάθος στην απεικόνιση της ταχύτητάς του όπως θα δούμε στη συνέχεια, καθώς για το συγκεκριμένο id το

σύστημα υπολογίζει αρχικά την ταχύτητα για το πρώτο όχημα και στη συνέχεια για το δεύτερο στο οποίο δόθηκε από σφάλμα το ίδιο id. Για τα προβλήματα που προαναφέρθηκαν λοιπόν, σε κάποιες περιπτώσεις η προσπάθεια αντιμετώπισης τους δεν ήταν επιτυχημένη και γι αυτό το λόγο άλλωστε, η ακρίβεια του συστήματος που παρουσιάστηκε στην προηγούμενη υποενότητα δεν είναι σε απόλυτο ποσοστό επιτυχίας αλλά παρουσιάζει και αποτυχίες εντοπισμού οχημάτων. Στις εικόνες που ακολουθούν, παρουσιάζονται παραδείγματα λάθος εντοπισμού του συστήματός μας.



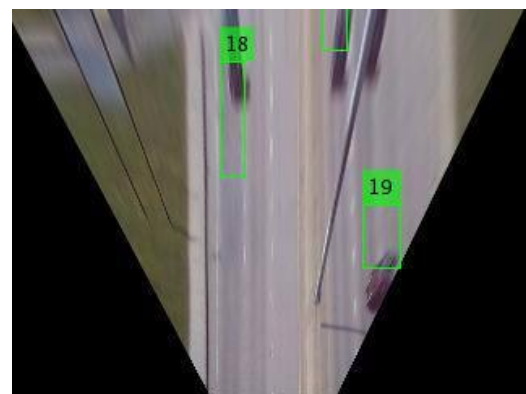
(α) Βίντεο #A2 frame 161



(β) Βίντεο #B6 frame 2171

**Εικόνα 4.2** Παραδείγματα λάθος εντοπισμού λόγω επικάλυψης οχημάτων

(α) Βίντεο #B5 frame 2078



(β) Βίντεο #B5 frame 2084

**Εικόνα 4.3** Παράδειγμα διπλής ανάθεσης id

### 4.3 Υπολογισμός Ταχύτητας Οχήματος

Το ένα από τα δύο τελικά στάδια της επεξεργασίας βίντεο στο σύστημά μας, είναι ο υπολογισμός της ταχύτητας με την οποία ταξιδεύουν τα οχήματα που εντοπίζονται όπως φαίνεται και από το διάγραμμα ροής της Εικόνας 1.1. Η γνώση της ταχύτητας των οχημάτων είναι εξαιρετικά χρήσιμη για πολλές εφαρμογές που ασχολούνται με την επεξεργασία βίντεο από κάμερες κυκλοφορίας καθώς αξιοποιώντας κατάλληλα τη γνώση αυτή, είναι δυνατό να εξαχθούν αρκετές επιπλέον πληροφορίες για τις συνθήκες κυκλοφορίας που επικρατούν στο δρόμο που είναι τοποθετημένη η κάμερα.

Στην παρούσα διπλωματική εργασία υλοποιείται ο υπολογισμός της ταχύτητας πολλαπλών οχημάτων σε διαδοχικά frames, βρίσκοντας τις αποστάσεις που διανύουν σε σχέση με το χρόνο [32], [33], [65], [66]. Κατά τη διαδικασία εντοπισμού οχήματος, βρίσκονται τα κέντρα βάρους των οχημάτων κάθε χρονική στιγμή (Εικόνα 4.4 (α)) και αυτή η διαδικασία τα αποθηκεύει σε έναν πίνακα ώστε να είναι διαθέσιμα στο σύστημα από αυτό το στάδιο και έπειτα. Η ταχύτητα οχήματος υπολογίζεται χρησιμοποιώντας τη θέση του οχήματος σε κάθε frame. Η απόσταση που διανύει το κάθε όχημα καθορίζεται από τη μετακίνηση του κέντρου βάρους του και υπολογίζεται χρησιμοποιώντας τη μαθηματική φόρμουλα της Ευκλείδειας απόστασης. Έχοντας υπολογίσει την απόσταση που ταξιδεύει το κάθε όχημα ανάμεσα σε δύο διαδοχικά frames και θεωρώντας πάντα ως γνωστό το frame rate του εκάστοτε βίντεο, είναι εφικτός ο υπολογισμός της ταχύτητας των οχημάτων. Από τη στιγμή που όπως προαναφέρθηκε υπάρχουν αποθηκευμένα σε έναν πίνακα τα κέντρα βάρους του κάθε οχήματος ξεχωριστά, υπολογίζεται η ταχύτητά του σύμφωνα με την Εξ. (4.11).

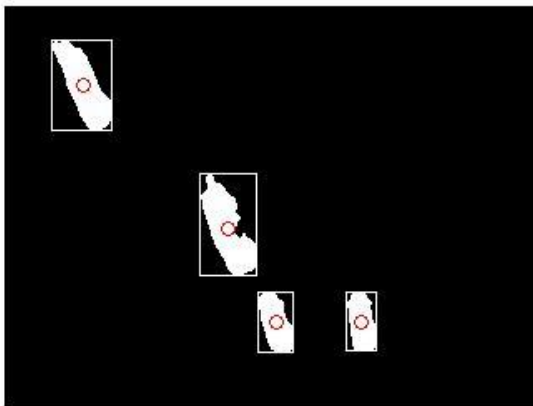
$$Velocity = ((distTravelled * (mean\_veh\_length / vehicleLength)) * frame\_rate) * kilometers \quad (4.11)$$

όπου *distTravelled* είναι η απόσταση, σε pixels, που διανύει το κέντρο βάρους από το ένα frame στο επόμενο. Η απόσταση αυτή υπολογίζεται από τον τύπο

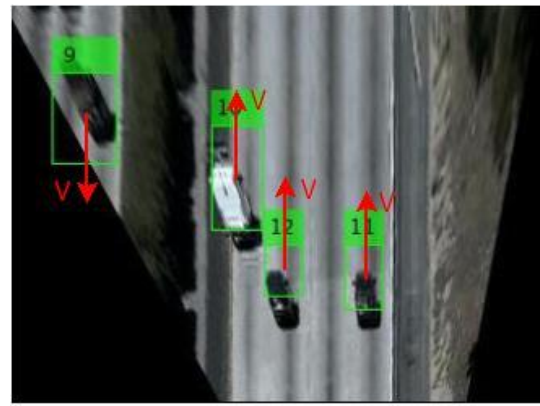
$$distTravelled = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \quad (4.12)$$

όπου (*x1*, *y1*) και (*x2*, *y2*) είναι οι συντεταγμένες του κέντρου βάρους στα frame (*i-1*) και frame (*i*) αντίστοιχα. Ως *mean\_veh\_length* θεωρείται το μέσο μήκος ενός επιβατικού οχήματος (περίπου 4m) και χρησιμοποιείται εδώ για να γίνει κανονικοποίηση του μήκους των οχημάτων από pixels σε μέτρα, απαραίτητο βήμα ώστε οι ταχύτητες των οχημάτων τελικά να εκφραστούν όχι σε pixels/s, αλλά σε km/h. Το *vehicleLength* είναι το μήκος του οχήματος σε pixels, η τιμή του *frame\_rate* δίνεται από το βίντεο και όπου *kilometers* ορίζεται η τιμή 3.6 και χρησιμοποιείται για τη μετατροπή του τελικού αποτελέσματος από m/s σε km/h.

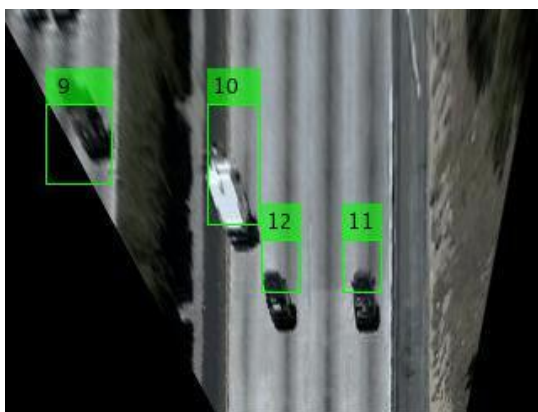
Το αποτέλεσμα της Εξ. (4.11) είναι ένας πίνακας που περιέχει σε κάθε γραμμή του το id του οχήματος και σε κάθε στήλη τις τιμές της ταχύτητας που υπολογίστηκαν, για όσα διαδοχικά frames εντοπίστηκε το καθένα από το σύστημα. Οι τιμές της ταχύτητας κάθε οχήματος για τα πέντε τελευταία frames στα οποία εντοπίζεται, πριν φύγει από το οπτικό πεδίο της κάμερας, μηδενίζονται για την αποφυγή λάθος υπολογισμού της καθώς το όχημα χάνεται. Για κάθε όχημα που έχει εντοπιστεί σωστά, αναμένεται η ταχύτητά του από frame σε frame να παραμένει περίπου σταθερή, χωρίς μεγάλες αυξομειώσεις. Στις παρακάτω εικόνες παρουσιάζονται τα αποτελέσματα της διαδικασίας υπολογισμού ταχύτητας οχήματος.



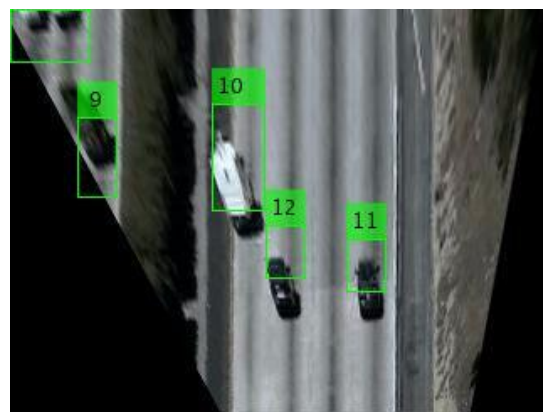
(α)



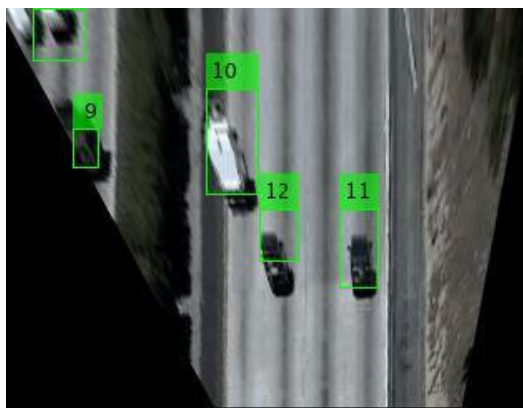
(β)



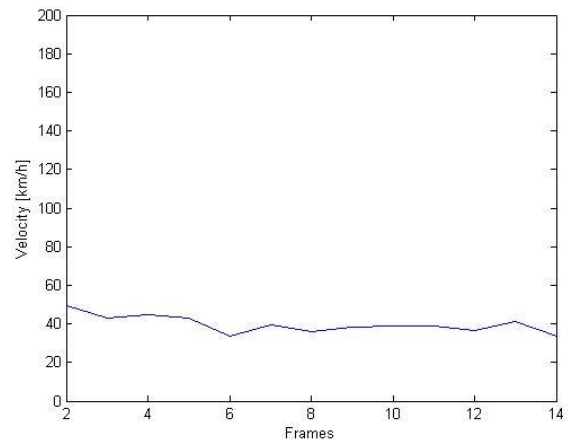
(γ)



(δ)

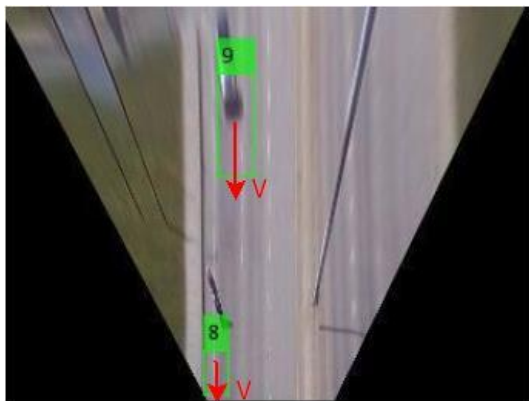


(ε)

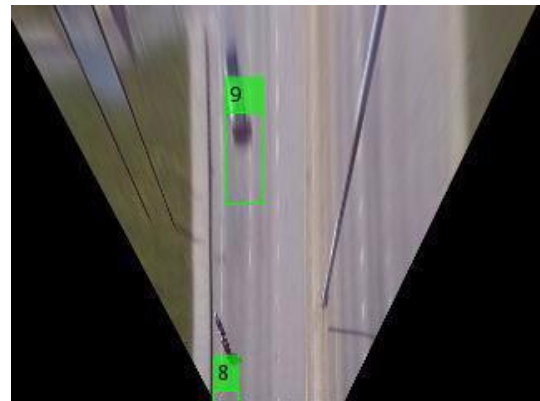


(ζ)

**Εικόνα 4.4** (α) Απεικόνιση κέντρου βάρους οχημάτων, (β)-(ε) Βίντεο #A4 διαδοχικά frames 383-386, (ζ) Ταχύτητα οχήματος id=11 στο Βίντεο #A4



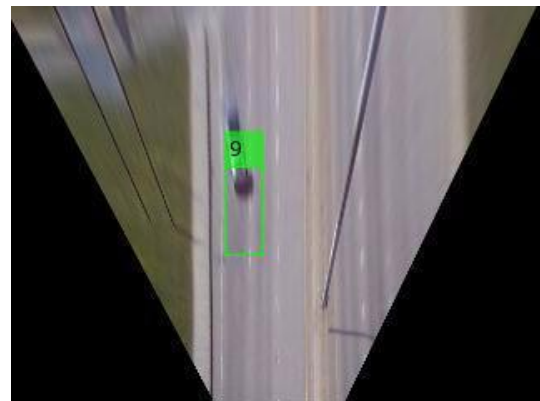
(α)



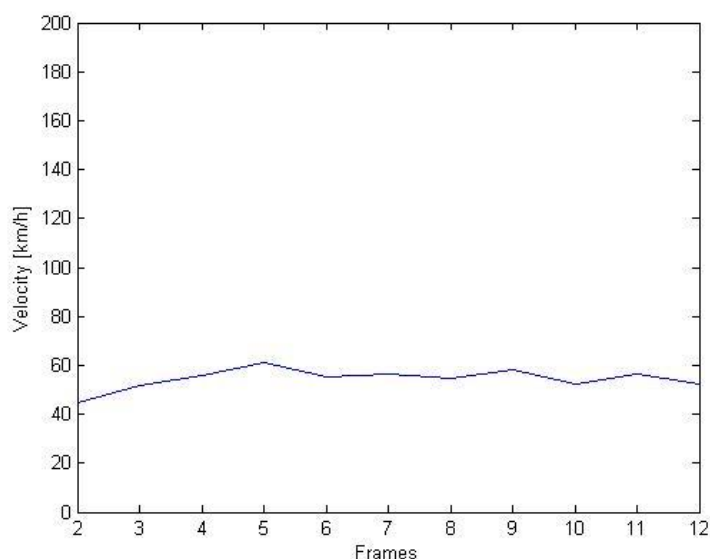
(β)



(γ)



(δ)



(ε)

**Εικόνα 4.5** (α)-(δ) Βίντεο #B3 διαδοχικά frames 1158-1161, (ε) Ταχύτητα οχήματος id=9 στο Βίντεο #B3

Στις Εικόνα 4.4 (β) και Εικόνα 4.5 (α) φαίνονται τα οχήματα μαζί με τα διανύσματα της ταχύτητας τους με κατεύθυνση. Όπως φαίνεται από τα αποτελέσματα (Εικόνα 4.4 (ζ) και Εικόνα 4.5 (ε)) που επιστρέφει η μέθοδος υπολογισμού της ταχύτητας οχήματος που υλοποιήθηκε στο σύστημά μας, για περιπτώσεις όπου ο αλγόριθμος εντοπισμού οχήματος λειτουργεί σωστά, ο υπολογισμός της ταχύτητας γίνεται επιτυχώς. Τα αποτελέσματα αποδεικνύουν αυτό που αναμενόταν, δηλαδή οι ταχύτητες των οχημάτων να είναι περίπου σταθερές για τα διαδοχικά frames που εντοπίζονται. Η διαδικασία υπολογισμού της ταχύτητας είναι πολύ σημαντικό να είναι επιτυχής στο σύστημα που προτείνεται από την παρούσα διπλωματική εργασία, καθώς αποτελεί το τελικό στάδιο του χαρακτηρισμού κίνησης στα κυκλοφοριακά βίντεο που επεξεργαζόμαστε και έναν από τους βασικούς στόχους που είχαμε θέσει εξ αρχής. Εδώ πρέπει να σημειωθεί ότι οι τιμές της ταχύτητας των οχημάτων αποτελούν εκτίμηση της πραγματικής τους ταχύτητας καθώς η μετατροπή από pixels/s σε km/h βασίζεται στην κανονικοποίηση του μήκους τους εφόσον δεν είναι διαθέσιμες οι παράμετροι της κάμερας, χωρίς αυτό να σημαίνει βέβαια ότι η εκτίμηση αυτή απέχει πολύ από την πραγματική ταχύτητα με την οποία ταξιδεύουν τα οχήματα. Ένας ακόμα παράγοντας που έχει συμβολή στην τιμή των ταχυτήτων που υπολογίζονται, είναι το frame rate που συνοδεύει το κάθε βίντεο. Αν η τιμή του frame rate δεν είναι ακριβής, τότε οι τιμές της ταχύτητας που υπολογίζονται θα επηρεάζονται. Για μεγαλύτερη ακρίβεια στον υπολογισμό της ταχύτητας απαιτείται η πιο σύνθετη διαδικασία της βαθμονόμησης της κάμερας όπως αναφέρθηκε και στο

δεύτερο κεφάλαιο, ώστε να βρεθούν οι παράμετροι όπως το ύψος στο οποίο είναι τοποθετημένη πάνω από το έδαφος και η γωνία κλίσης της, για να υπολογιστεί η ακριβής απόσταση που διανύει το κάθε όχημα σε μέτρα.

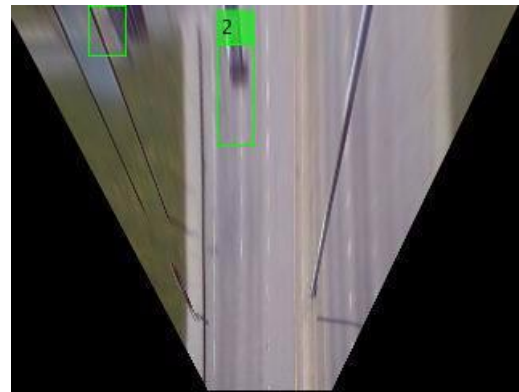
Δυστυχώς, όπως αναφέρθηκε παραπάνω η διαδικασία υπολογισμού της ταχύτητας οχήματος εξαρτάται απόλυτα από τα αποτελέσματα που επιστρέφει στο σύστημα ο αλγόριθμος εντοπισμού οχήματος. Συνεπώς για να έχει επιτυχία ο υπολογισμός ταχύτητας για κάθε όχημα θα πρέπει η μέθοδος εντοπισμός του οχήματος να έχει ιδανικά ακρίβεια 100%, κάτι που είναι πρακτικά πάρα πολύ δύσκολο. Στις υποενότητες 4.2.3 και 4.2.4 του παρόντος κεφαλαίου παρουσιάστηκαν αναλυτικά, η ακρίβεια του αλγορίθμου εντοπισμού οχήματος που χρησιμοποιήθηκε στην παρούσα διπλωματική εργασία και τα προβλήματα που προκύπτουν στον εντοπισμό των οχημάτων. Υπάρχουν αρκετές περιπτώσεις στο σύστημα στις οποίες απέτυχε ο υπολογισμός ταχύτητας κάποιου οχήματος είτε λόγω μη εντοπισμού του, είτε λόγω επικάλυψης του με άλλο όχημα οπότε υπολογίζεται μια τιμή της ταχύτητας και για τα δύο οχήματα σαν να ήταν ένα ενιαίο, είτε λόγω ανάθεσης του ίδιου id σε δύο διαφορετικά οχήματα. Στην τελευταία περίπτωση μπορεί να παρατηρηθεί ότι παρόλο που ο εντοπισμός παρουσίασε σφάλμα, αναθέτοντας το ίδιο id σε δύο διαφορετικά οχήματα, ο μέθοδος υπολογισμού της ταχύτητας που υλοποιήσαμε δουλεύει σωστά υπολογίζοντας τις ταχύτητες και τον δύο αυτών οχημάτων. Αρχικά το σύστημά μας εντοπίζει τον ποδηλάτη (Εικόνα 4.6 (α)) αναθέτοντας του το id=2 και την ταχύτητα με την οποία κινείται, η οποία απεικονίζεται στα 11 πρώτα frames της Εικόνας 4.6 (γ). Στην Εικόνα 4.6 (β) φαίνεται ένα διαφορετικό όχημα 3 frames αργότερα από αυτό της Εικόνας 4.6 (α), το οποίο λόγω σφάλματος δεν αναγνωρίζεται ως νέο όχημα του και του δίνεται το id=2. Ο ποδηλάτης σταματάει σε αυτή τη χρονική στιγμή να εντοπίζεται σαν κινούμενο όχημα από το σύστημά μας διότι κινείται πολύ αργά και αυτή η εναλλαγή του id έχει ως αποτέλεσμα το λάθος υπολογισμό της ταχύτητας για το όχημα με id=2 όπως απεικονίζεται στο frame 12 της Εικόνας 4.6 (γ). Παρόλα αυτά, η ταχύτητα του νέου πλέον οχήματος με το συγκεκριμένο id υπολογίζεται σωστά από το frame 13 και έπειτα. Στην εικόνα αυτή με κόκκινο χρώμα απεικονίζεται η τιμή της ταχύτητας που επιστρέφει το σύστημα και με μπλε χρώμα η βελτίωση της απεικόνισης της ταχύτητας με τη χρήση του Median φίλτρου. Το Median φίλτρο είναι μια μη γραμμική, ψηφιακή τεχνική φιλτραρίσματος που χρησιμοποιείται συχνά για την αφαίρεση του θορύβου. Εδώ χρησιμοποιείται καθαρά για λόγους που έχουν να κάνουν με την καλύτερη απεικόνιση των τιμών της ταχύτητας και την πιο ομαλή μετάβαση από το σφάλμα στο frame 12 (Εικόνα 4.6 (γ) κόκκινο χρώμα). Τελικά σε αυτήν την περίπτωση, όπως φαίνεται στην Εικόνα 4.6 (γ) με μπλε χρώμα, το σύστημά μας υπολογίζει την ταχύτητα με την οποία κινείται ο ποδηλάτης αρχικά. Στα 11 πρώτα frames υπολογίζεται η τιμή της και όπως παρατηρείται είναι σταθερή και χαμηλή, κάτι απολύτως λογικό για ποδήλατο. Στα αμέσως επόμενα frames παρατηρείται το σφάλμα στον υπολογισμό της ταχύτητας



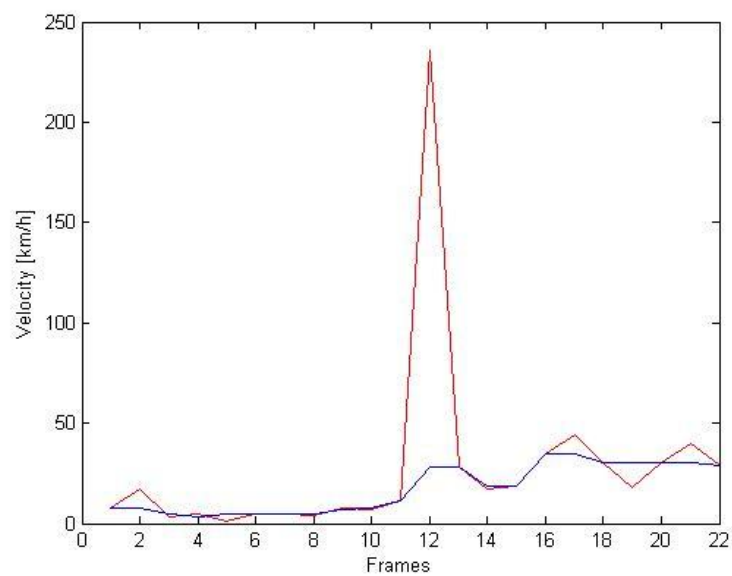
λόγω σφάλματος στο εντοπισμό, αλλά και η εκ νέου σταθεροποίηση στην τιμή της ταχύτητας που υπολογίζεται πλέον για το κινούμενο όχημα.



(α)



(β)



(γ)

**Εικόνα 4.6** (α) Όχημα με id=2 στο Βίντεο #B2 frame 714, (β) Όχημα με id=2 στο Βίντεο #B2 frame 718, (γ) Απεικόνιση ταχύτητας 2 οχημάτων στα οποία από σφάλμα ανατέθηκε το ίδιο id

## 4.4 Ταξινόμηση Οχήματος

Παράλληλα με τη διαδικασία υπολογισμού της ταχύτητας οχήματος, υλοποιείται και αυτή της ταξινόμησης των οχημάτων. Η διαδικασία αυτή αποτελεί το τελικό στάδιο, μαζί με αυτό του υπολογισμού ταχύτητας οχήματος που παρουσιάστηκε στην προηγούμενη ενότητα, του συστήματος για την επεξεργασία βίντεο από κάμερες κυκλοφορίας που υλοποιείται στην παρούσα διπλωματική εργασία. Η ταξινόμηση των οχημάτων είναι εξαιρετικά χρήσιμη και σημαντική για την κυκλοφοριακή λειτουργία, το σχεδιασμό του οδοστρώματος και των μεταφορών, λόγω των σημαντικών διαφορών στις επιδόσεις, στο μέγεθος και στο βάρος μεταξύ των μικρών και των μεγάλων οχημάτων.

Στην παρούσα διπλωματική εργασία υλοποιείται μια απλή μέθοδος ταξινόμησης οχήματος [59] ελαφρώς τροποποιημένη, η οποία λόγω της ευκολίας και της απλότητας της υλοποίησης είναι υπολογιστικά ανέξοδη, πράγμα που την καθιστά εφαρμόσιμη σε εφαρμογές πραγματικού χρόνου που απαιτούν γρήγορες μεθόδους επεξεργασίας και ανάλυσης δεδομένων. Η μέθοδός μας βασίζεται στο μήκος και το εμβαδόν των οχημάτων σε pixels, θεωρώντας αυτά τα χαρακτηριστικά ικανά για να ταξινομήσουν σωστά κάθε όχημα. Κατά το στάδιο του εντοπισμού οχήματος, βρίσκονται τα μήκη και το εμβαδόν των οχημάτων κάθε χρονική στιγμή και οι πληροφορίες αυτές αποθηκεύονται σε πίνακες, ώστε να είναι διαθέσιμες στο σύστημα για τα επόμενα στάδια όπως αυτό της ταξινόμησης οχήματος. Λόγω του ότι η μέθοδος ταξινόμησης που επιλέχθηκε γίνεται με βάση το μήκος και το εμβαδόν οχήματος, οι σκιές είναι ένα πρόβλημα που προκύπτει καθώς αν δεν αφαιρεθούν τότε λογίζονται ως μέρος των οχημάτων, πράγμα που οδηγεί σε λάθος ταξινόμηση. Στο σύστημά μας το στάδιο της ταξινόμησης ακολουθεί εκείνο της αφαίρεσης σκιάς που παρουσιάστηκε στο τρίτο κεφάλαιο, οπότε το πρόβλημα αντιμετωπίζεται.

Η προτεινόμενη μέθοδος ταξινόμησης οχήματος χρησιμοποιεί το μήκος του οχήματος εκφρασμένο σε pixels και όχι το πραγματικό μήκος του σε μέτρα. Εφόσον ο σκοπός εδώ είναι μόνο η κατηγοριοποίηση των οχημάτων και τίποτα άλλο, δεν είναι απαραίτητη η γνώση ή η προσπάθεια υπολογισμού του πραγματικού μήκους του. Το μήκος του κάθε οχήματος και το εμβαδόν του σε pixels συγκρίνονται με δύο κατώφλια για να προσδιοριστεί εάν ανήκει το όχημα στην κατηγορία SV ή την κατηγορία LV. Τα κατώφλια που χρησιμοποιούνται για την ταξινόμηση δε γίνεται να είναι προκαθορισμένα, καθώς τα οχήματα στην κάμερα δείχνουν διαφορετικά από ότι είναι στην πραγματικότητα λόγω της απόστασης τους από αυτήν, το ύψος και τη γωνία τοποθέτησης της. Γι αυτόν το λόγο επιλέγουμε το μέσο μήκος και το μέσο εμβαδόν των δέκα πρώτων οχημάτων, ως ένα καλό δείγμα και μέτρο σύγκρισης, που εντοπίζονται σε κάθε βίντεο ως κατώφλια της ταξινόμησης. Το όχημα με μήκος και εμβαδόν μεγαλύτερο από το 50% του μέσου μήκους και του μέσου εμβαδού, ταξινομείται ως LV αλλιώς εάν είναι μικρότερο τότε ταξινομείται στην κατηγορία SV. Τα αποτελέσματα από την εφαρμογή της μεθόδου που περιγράφηκε, ακολουθούν

στους παρακάτω πίνακες. Τα οχήματα που επιχειρεί η μέθοδος να ταξινομήσει είναι προφανώς τα οχήματα που επιστρέφει ο αλγόριθμος εντοπισμού και οι αριθμοί στους πίνακες αναφέρονται σε αυτά, καθώς εκείνα που δεν εντοπίστηκαν είναι αδύνατο να ταξινομηθούν.

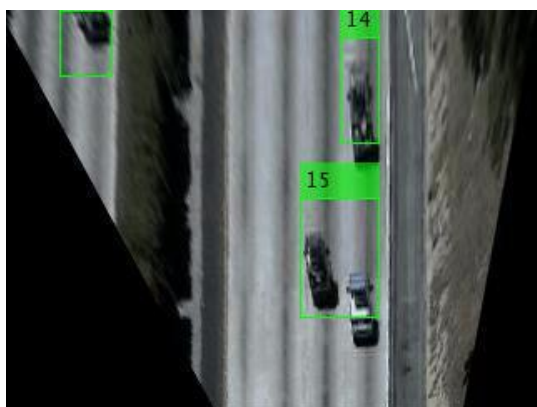
Είσοδος	Βίντεο #A1	Βίντεο #A2	Βίντεο #A3	Βίντεο #A4	Βίντεο #A5
<b>Frames</b>	100	100	100	100	100
<b>Αριθμός SV</b>	19	20	15	14	21
<b>Ταξινόμηση ως SV</b>	19	19	15	14	17
<b>Error SV</b>	0	1	0	0	4
<b>Αριθμός LV</b>	2	0	0	0	0
<b>Ταξινόμηση ως LV</b>	2	1	0	0	4
<b>Error LV</b>	0	1	0	0	4
<b>Ακρίβεια Ταξινόμησης</b>	100%	90%	100%	100%	61.90%

Πίνακας 4.3

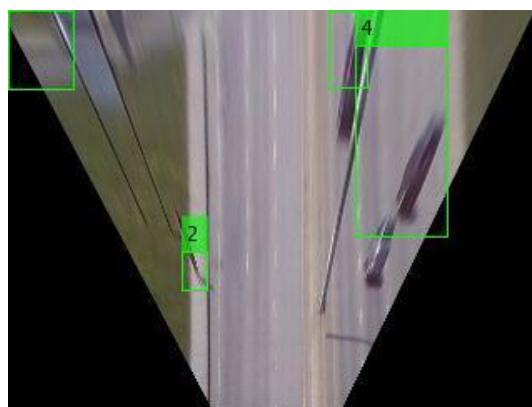
Είσοδος	Βίντεο #B1	Βίντεο #B2	Βίντεο #B3	Βίντεο #B4	Βίντεο #B5	Βίντεο #B6
<b>Frames</b>	100	100	100	100	100	100
<b>Αριθμός SV</b>	19	17	17	22	24	23
<b>Ταξινόμηση ως SV</b>	19	15	18	22	24	21
<b>Error SV</b>	0	2	1	0	0	2
<b>Αριθμός LV</b>	1	1	1	0	0	1
<b>Ταξινόμηση ως LV</b>	1	3	0	0	0	3
<b>Error LV</b>	0	2	1	0	0	2
<b>Ακρίβεια Ταξινόμησης</b>	100%	77.78%	88.89%	100%	100%	83.33%

Πίνακας 4.4

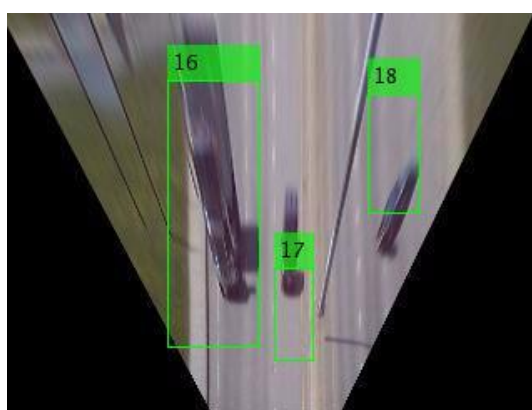
Σύμφωνα με τους Πίνακα 4.3 και Πίνακα 4.4, η ταξινόμηση έχει αρκετά μεγάλη επιτυχία εκτός από κάποιες συγκεκριμένες περιπτώσεις, για τις οποίες επιστρέφει λάθος αποτελέσματα. Για τα πέντε βίντεο του Πίνακα 4.3 η μέση ακρίβεια είναι 90.38%, ενώ για τα έξι βίντεο του Πίνακα 4.4 η αντίστοιχη τιμή είναι 91.67%. Επίσης υπολογίζεται η συνολική μέση ακρίβεια ταξινόμησης οχήματος, όλων των βίντεο που εξετάστηκαν στο σύστημα και έχει τιμή 91.08%, κάτι που σημαίνει ότι εάν ο εντοπισμός του οχήματος είναι σωστός τότε και η ταξινόμηση του οχήματος θα γίνει σωστά. Λάθος στην ταξινόμηση, όπως συμβαίνει και για την ταχύτητα του οχήματος όπως είδαμε στην προηγούμενη ενότητα, ισοδυναμεί τις περισσότερες φορές με λάθος του αλγορίθμου εντοπισμού των οχημάτων. Τα περισσότερα σφάλματα παρατηρήθηκαν στα Βίντεο #A5 και Βίντεο #B2, καθώς σε αυτά ο αλγόριθμος εντοπισμού αντιμετωπίζει προβλήματα επικάλυψης οχημάτων όπως φαίνεται και στις παρακάτω εικόνες από τα βίντεο αυτά.



(α)



(β)



(γ)

**Εικόνα 4.7** (α) Παράδειγμα λάθους ταξινόμησης οχήματος με id=15 ως LV, (β) Παράδειγμα λάθους ταξινόμησης οχήματος με id=4 ως LV, (γ) Σωστή ταξινόμηση οχήματος με id=16 ως LV

## **Κεφάλαιο 5 : Συμπεράσματα και Μελλοντική Βελτίωση**

Σε αυτή τη διπλωματική εργασία παρουσιάσαμε ένα ολοκληρωμένο σύστημα αναγνώρισης και παρακολούθησης της κίνησης σε κυκλοφοριακά βίντεο. Αρχικά υλοποιήσαμε ένα μετασχηματισμό προοπτικής και στη συνέχεια τρεις μεθόδους ανίχνευσης αντικειμένων συγκρίνοντας την ποιότητα των αποτελεσμάτων τους. Καμία μέθοδος ανίχνευσης αντικειμένων δεν επιστρέφει πάντα τέλεια αποτελέσματα, έτσι λοιπόν και η μέθοδος που χρησιμοποιήσαμε χρειάζεται με τη σειρά της αρκετές βελτιώσεις ώστε να ανταποκρίνεται καλύτερα σε δύσκολες συνθήκες όπως είναι οι δραστικές αλλαγές φωτισμού και οι επικαλύψεις οχημάτων. Στη συνέχεια υλοποιήθηκε μια μέθοδος αφαίρεσης σκιάς που βοήθησε αρκετά στο στάδιο του εντοπισμού αντικειμένου, που πραγματοποιήθηκε με τη βοήθεια του αλγορίθμου Kalman Filter. Το στάδιο αυτό χρειάζεται ακόμα αρκετή βελτίωση καθώς αποτελεί τη ραχοκοκαλιά ολόκληρης της επεξεργασίας και οποιοδήποτε λάθος στον εντοπισμό οχήματος είναι κρίσιμο, αφού από αυτή τη διαδικασία εξαρτάται η επιτυχία των συστημάτων παρακολούθησης της κυκλοφορίας. Το σύστημά μας προϋποθέτοντας σωστό εντοπισμό οχήματος, είναι ικανό να υπολογίσει μια εκτίμηση για την ταχύτητα των οχημάτων όπως επίσης και να τα ταξινομήσει, κατορθώνοντας έτσι να επιτύχει, άλλον σε μεγαλύτερο και άλλον σε μικρότερο βαθμό, όλους τους στόχους που είχαν τεθεί εξ αρχής. Το πιο σημαντικό εδώ, είναι ότι το σύστημα που προτείνεται μας δίνει ελπιδοφόρα αποτελέσματα και μπορεί να χρησιμοποιηθεί τόσο ως μέρος κάποιου μεγαλύτερου συστήματος παρακολούθησης της κυκλοφορίας όσο και ως βάση για περαιτέρω έρευνα.

Κάθε στάδιο του προτεινόμενου συστήματος χρήζει βελτίωσης στο μέλλον. Ο μετασχηματισμός προοπτικής μπορεί χρησιμοποιώντας περισσότερα και καλύτερα επιλεγμένα χαρακτηριστικά σημεία, να δώσει πιο καθαρή εικόνα εκείνων των οχημάτων που βρίσκονται σε απόσταση από την κάμερα και η μέθοδος βαθμονόμησης της κάμερας εάν υλοποιηθεί, να δώσει στο σύστημα πολύ χρήσιμες παραμέτρους της κάμερας. Ο εντοπισμός οχήματος μπορεί να επεκταθεί ώστε να ανταποκρίνεται σωστά σε οχήματα που εκτελούν μη ομαλές κινήσεις και δεν κινούνται με σταθερή ταχύτητα. Επίσης σε μελλοντική εργασία μπορεί να δοθεί έμφαση στην υλοποίηση ενός συστήματος με όσο το δυνατόν μικρότερη πολυπλοκότητα και φυσικά απώτερος σκοπός είναι όλα αυτά να βρουν εφαρμογή σε ένα σύστημα πραγματικού χρόνου, πράγμα το οποίο αποτελεί μεγάλη πρόκληση.

Το εύρος του φάσματος των εφαρμογών στις οποίες απευθύνεται το προτεινόμενο σύστημα είναι πολύ μεγάλο, καθώς ο έλεγχος της κυκλοφορίας γίνεται πλέον από αυτοματοποιημένες εφαρμογές. Υπάρχουν συστήματα καταγραφής παραβάσεων των ορίων ταχύτητας στο οποία θα μπορούσε μέρος του συστήματός μας να χρησιμοποιηθεί, όπως επίσης και σε μεγάλους αυτοκινητόδρομους θα μπορούσε να χρησιμοποιηθεί ώστε να λειτουργήσει προειδοποιητικά για λόγους ασφαλείας (π.χ. συμφόρηση ενός συγκεκριμένου κομματιού του δρόμου). Με κατάλληλες τροποποιήσεις το σύστημα αυτό μπορεί να βρει εφαρμογή και σε συστήματα παρακολούθησης χώρων, είτε εσωτερικών είτε εξωτερικών αρκεί να υπάρχει σταθερή κάμερα που να καταγράφει την εικόνα. Επίσης ένας ακόμα πιθανός τομέας όπου το σύστημά μας είναι δυνατόν να χρησιμοποιηθεί είναι οι ολοένα και αναπτυσσόμενες ρομποτικές εφαρμογές, καθώς θα μπορούσε για παράδειγμα να αποτελέσει τμήμα ενός ρομπότ για τη συλλογή πληροφοριών από το εξωτερικό περιβάλλον.

## Βιβλιογραφία

- [1] C. Little, “*The intelligent vehicle initiative: Advancing ‘Human-Centered’ smart vehicles*” Public Roads Mag., vol. 61, no. 2, pp. 18-25, Sept./Oct. 1997.
- [2] H. Kano, K. Asari, Y. Ishii, H. Hongo, “*Precise top view image generation without global metric information*”, IEICE Transaction on Information and System, vol. E91-D, pp. 1893-1898, 2008.
- [3] M.Venkatesh, P.Vijayakumar, *A Simple Bird's Eye View Transformation Technique*, International Journal of Scientific & Engineering Research, Volume 3, Issue 5, May-2012.
- [4] Lin C.-C., Wang M.-S., *A vision based top-view transformation model for a vehicle parking assistant*, Sensors 2012, 12, 4431–4446.
- [5] Qing Lin, Youngjun Han, Hernsoo Hahn, *Vision-Based Obstacle Avoidance in Sidewalk Environment Using Top-View Transform and Optical-Flow*, Journal of Measurement Science and Instrumentation, Vol.2 No.4, pp.341-345, August 2011.
- [6] Yong Fan, Zhengyu Zhang, *A Fast Algorithm of Dynamic Background Extraction*, Journal of Communication and Computer, Volume 3, Jul. 2006.
- [7] J.Heikkila and O.Silven, *A real time system for monitoring of cyclists and pedestrians*, In proc. of second IEEE Workshop on visual Surveillance, pages 74-81, Fort Collins, Colorado, June 1999.
- [8] Tan R, Huo H, Qian J, Fang T., *Traffic video segmentation using adaptive-k gaussian mixture model*, ” The International Workshop on Intelligent Computing, IWICPAS 2006, China, 125-134, August 2006.
- [9] Ryan P. Avery, Guohui Zhang, Yinhai Wang and Nancy Nihan, *An Investigation into Shadow Removal from Traffic Images*, Transportation Research Record: Journal of the Transportation Research Board. No. 2000, pp 70-77, 2007.
- [10] Otsu N., *A Threshold Selection Method from Gray-Level Histograms*, IEEE Transactions on Systems, Man and Cybernetics, Vol. 9, No. 1, pp. 62-66, 1979.

- [11] Prati, A., Mikic, I., Grana C., Trivedi M., "*Shadow detection algorithms for traffic flow analysis: a comparative study*", Intelligent Transportation Systems, 2001. Proceedings. 2001 IEEE, vol., no., pp.340,345, 2001.
- [12] Lazaros Grammatikopoulos, George Karras, Elli Petsa, *Automatic Estimation of Vehicle Speed from Uncalibrated Video Sequences*, in *Proceedings of International Symposium on Modern Technologies, Education and Professional Practice in Geodesy and Related Fields*, pp. 332–338, Sofia, Bulgaria, November 2005.
- [13] Azhar Hussain, Kashif Shahzad, Chunming Tang, *Real Time Speed Estimation of Vehicles*, World Academy of Science, Engineering & Technology, Issue 61, p726, 2012.
- [14] Zhang Guohui, Ryan P. Avery, Yinhai Wang, "A Video-Based Vehicle Detection and Classification System for Real-Time Traffic Data Collection Using Uncalibrated Video Cameras", Transportation Research Record: Journal of the Transportation Research Board, No. 1993, pp 138-147, 2007.
- [15] Surendra Gupte, Nikos Papanikolopoulos, *Algorithms for Vehicle Classification*, University of Minnesota Dept. of Computer Science and Engineering, Final Report 1999-2000, July 2000.
- [16] Chang Liu, George Chen, Yingdong Ma, Yajun Ding, Jinwu Wu, *Accurate Vehicle Tracking by Shadow Removal in the Transformed Space*, IASTED International Conference, Signal and Image Processing and Applications(SIPA 2011), 2011 Crete.
- [17] G. Wolberg and S. Zokai, *Image registration for perspective deformation recovery*, In SPIE Conf. on Automatic Target Recognition X, Orlando, Florida, April 2000.
- [18] [http://www.ldv.ei.tum.de/uploads/media/Vorlesung\\_3.2\\_SpatialTransformations.pdf](http://www.ldv.ei.tum.de/uploads/media/Vorlesung_3.2_SpatialTransformations.pdf)
- [19] <http://alumni.media.mit.edu/~cwren/interpolator/>
- [20] Shiho Tanaka, Kenichi Yamada, Toshio Ito, and Takenao Ohkawa, "Vehicle Detection Based on Perspective Transformation Using Rear-View Camera", International Journal of Vehicular Technology, Vol. 2011, 2011.
- [21] N. K. Kanhere and S. T. Birchfield, "A taxonomy and analysis of camera calibration methods for traffic monitoring applications", IEEE Trans. Intell. Transp. Syst., vol. 11, no. 2, pp.441-452, 2010.



- [22] Lili Huang, *Roadside camera calibration and its application in length-based vehicle classification*, in Proc. 2nd Int. Asia Conf. Inf. CAR, vol. 2, pp. 329–332, 2010.
- [23] <http://www-staff.it.uts.edu.au/~massimo/BackgroundSubtractionReview-Piccardi.pdf>
- [24] [http://www.slideshare.net/ravi5raj\\_88/background-subtraction](http://www.slideshare.net/ravi5raj_88/background-subtraction)
- [25] [http://www.cs.utexas.edu/~grauman/courses/fall2009/slides/lecture9\\_background.pdf](http://www.cs.utexas.edu/~grauman/courses/fall2009/slides/lecture9_background.pdf)
- [26] Marwa abd el Azeem Marzouk, *Modified Background Subtraction Algorithm for Motion Detection in Surveillance Systems*, AMARABAC, Journal of American Academy for Sciences and Technology, Volume 1, Number 2, pp. 112-123, 2010.
- [27] C. Stauffer, W. E. L. Grimson, “*Adaptive background mixture models for real-time tracking*”, in Proc. IEEE Computer Soc. Conf. Computer Vision and Pattern Recognition, pp. 246–252, 1999.
- [28] Rui Tan, Hong Huo, Jin Qian, Tao Fang, *Traffic Video Segmentation Using Adaptive-K Gaussian Mixture Model*, The International Workshop on Intelligent Computing, IWICPAS 2006, China, August 2006.
- [29] Douglas A. Reynolds, *Gaussian Mixture Models*, Encyclopedia of Biometrics 2009: 659-663.
- [30] Mei Xiao, Chong-Zhao Han, Lei Zhang, *Moving Shadow Detection and Removal for Traffic Sequences*, International Journal of Automation and Computing, Volume 4, Issue 1, pp 38-46, January 2007.
- [31] Karishma Malhotra, G.S. Sekhon, *Consecutive Frame Scanning Algorithm for Multivehicle Velocity Detection*, International journal of Computer Science & Communication Vol. 3, No. 1, pp. 207-210, January-June 2012.
- [32] A.Sai Suneel, *Person or Object Tracking and velocity Estimation in Real Time Videos*, International Journal Publications of Problems and Applications in Engineering Research, pp. 292-299, CSEA – 2012.
- [33] T. Malavika, M. Poornima, “*Moving object detection and velocity estimation using MATLAB*”, International Journal of Engineering Research & Technology (IJERT), Vol.2 - Issue 1, January 2013.
- [34] [http://en.wikipedia.org/wiki/Projective\\_geometry](http://en.wikipedia.org/wiki/Projective_geometry)

[35] Lili Huang, "*Roadside Camera Calibration and its Application in Length-based Vehicle Classification*", 2nd International Asia Conference on Informatics in Control, Automation and Robotics, pp.329-332, 2010.

[36] [http://en.wikipedia.org/wiki/Pinhole\\_camera](http://en.wikipedia.org/wiki/Pinhole_camera)

[37] Zengyou Zhang, "*A Flexible New Technique for Camera Calibration*", Microsoft Research Technical Report MSR-TR-98-71, 1998.

[38] [http://en.wikipedia.org/wiki/Vanishing\\_point](http://en.wikipedia.org/wiki/Vanishing_point)

[39] [http://en.wikipedia.org/wiki/Projective\\_transformation](http://en.wikipedia.org/wiki/Projective_transformation)

[40] Grammatikopoulos L., *Geometric Information from Single Images in Photogrammetry and Computer Vision*, Ph.D. Dissertation, Department of Surveying, School of Rural and Agricultural Surveying, National Technical University, Athens Greece, December 2007.

[41] <http://en.wikipedia.org/wiki/Collineation>

[42] <http://en.wikipedia.org/wiki/Homography>

[43] [http://en.wikipedia.org/wiki/Transformation\\_matrix](http://en.wikipedia.org/wiki/Transformation_matrix)

[44] [http://en.wikipedia.org/wiki/Affine\\_transformation](http://en.wikipedia.org/wiki/Affine_transformation)

[45] Solomon C., Breckon T., *Fundamentals of Digital Image Processing: A Practical Approach with Examples in Matlab*, pp. 180-184.

[46] Makantasis K., Doulamis A., *3D Measures Computed in Monocular Camera System and SVM-based Classifier for Humans Fall Detection*, TMC Academic Journal, 7(2): 1-14, 2013.

[47] Alan M. M<sup>c</sup>Ivor, *Background Subtraction Techniques*, In Proc. of Image and Vision Computing, Auckland, New Zealand, 2000.

[48] M. Piccardi, "*Background subtraction techniques: A review*", in Proc. IEEE Int. Conf. Systems, Man and Cybernetics, pp. 3099–3104, 2004.

[49] A. Lipton, H. Fujiyoshi, R. Patil, "*Moving target detection and classification from real-time video*", Proc. IEEE Workshop Applications of Computer Vision (WACV), pp. 8-14, Oct. 1998.

- [50] L. Wang, W. Hu, T. Tan, *Recent developments in human motion analysis*, Pattern Recognition, 36(3) : 585–601, March 2003.
- [51] R. T. Collins, Alan J. Lipton, Takeo Kanade, *A system for video surveillance and monitoring*, VSAM final Report, Technical report CMU-RI-TR-00-12, Robotics Institute, Carnegie Mellon University, May 2000.
- [52] N. McFarlane and C. Schofield, “*Segmentation and tracking of piglets in images*”, Machine Vision and Applications 8(3), pp. 187-193, 1995.
- [53] [http://www.eetimes.com/document.asp?doc\\_id=1275604&page\\_number=1](http://www.eetimes.com/document.asp?doc_id=1275604&page_number=1)
- [54] <http://homepages.inf.ed.ac.uk/rbf/HIPR2/sobel.htm>
- [55] A. Gyaourova, C. Kamath, S. and C. Cheung, *Block matching object Tracking*, LLNL Technical report, October 2003.
- [56] Y. Wang, J. Doherty and R. Van Dyck, *Moving Object Tracking in video*, Proc. Conference on Information Sciences and Systems, Princeton, NJ, March 2000.
- [57] A. Turolla, L. Marchesotti, C.S. Regazzoni, *Multicamera object tracking in video surveillance applications*, Target Tracking 2004: Algorithms and Applications, IEE, vol., no., pp.85,90, 23-24 March 2004.
- [58] Y. Rosenberg, M. Werman, *Real-Time Object Tracking from a Moving Video Camera: A software approach on a PC*, Proceedings of the 4th IEEE Workshop on Applications of Computer Vision (WACV'98), p.238, October 19-21, 1998.
- [59] Y. Yang, Y. Ming, Y. Gang, Z. Yandong, *Length-Based Vehicle Classification in Multi-lane Traffic Flow*, Transactions of Tianjin University, Vol.17, pp. 362-368, 2011.
- [60] Teoh S. and Bräunl T., *A Reliability Point and Kalman Filter-based Vehicle Tracking Technique*, International Conference on Intelligent Systems.
- [61] Rachel Kleinbauer, *Kalman Filtering Implementation with Matlab*, Study Report, University Stuttgart, Helsinki, November 2004.
- [62] 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.

[63] <http://www.mathworks.com/help/vision/examples/motion-based-multiple-object-tracking.html>

[64] <http://en.wikipedia.org/wiki/Centroid>

[65] M. Jyothi Kiran and K.S. Roy, *A Video Surveillance System for Speed Detection in Vehicles*, International Journal of Engineering Trends and Technology (IJETT), Volume 4, pp. 1437-1441, May 2013.

[66] H. A. Rahim, U. U. Sheikh, R. B. Ahmad and A. S. M. Zain, *Vehicle Velocity Estimation for Traffic Surveillance System*, World Academy of Science, Engineering and Technology, September 2010.