



Πανεπιστήμιο Θεσσαλίας

Τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών

Διπλωματική Εργασία

“Ανάλυση δομής και κατάτμηση βίντεο δελτίων ειδήσεων”

“Structure analysis and segmentation of broadcast news video”

Δασκάλου Φώτιος-Χρυσοβαλάντης

**Επιβλέποντες:**

Ποταμιάνος Γεράσιμος,

Αναπληρωτής καθηγητής

Κατσαβουνίδης Ιωάννης,

Αναπληρωτής καθηγητής

Σεπτέμβριος 2014, Βόλος



## Ευχαριστίες

*Κατ' αρχήν θα ήθελα να ευχαριστήσω τον κ. Γεράσιμο Ποταμιάνο για την άριστη συνεργασία που είχαμε τον τελευταίο χρόνο. Χωρίς τις συμβουλές και την καθοδήγησή του δεν θα ήταν δυνατή η περάτωση της παρούσας διπλωματικής εργασίας.*

*Επίσης, θα ήθελα να ευχαριστήσω την οικογένειά μου και τους φίλους μου για τη στήριξή τους καθ' όλη την διάρκεια των προπτυχιακών σπουδών μου.*



## Περίληψη

Στόχος της διπλωματικής αυτής εργασίας είναι η παρουσίαση και εφαρμογή μιας ικανοποιητικής μεθόδου κατάτμησης δελτίων ειδήσεων. Συγκεκριμένα, ως βάση δεδομένων χρησιμοποιήθηκαν δελτία ειδήσεων του καναλιού Mega. Σημαντικό στοιχείο της έρευνας με σκοπό την κατάτμηση των δελτίων του καναλιού αυτού αποτελεί η σαφή κατανόηση της δομής τους. Στην παρούσα εργασία τα δελτία χωρίζονται σε δύο ειδών τμήματα. Κάθε τμήμα απεικονίζει είτε μια συζήτηση σε πάνελ είτε την παρουσίαση ενός ρεπορτάζ. Ειδικότερα, η ύπαρξη χαρακτηριστικών γραμμών και τίτλων ειδήσεων στα πάνελ του Mega οδήγησαν σε δύο διαφορετικές προσεγγίσεις κατάτμησης που στηρίχτηκαν σε οπτικά χαρακτηριστικά. Κατά την πρώτη προσέγγιση, προτείνεται μια μέθοδος που βασίζεται στην ανίχνευση προσώπων και γραμμών. Για το λόγο αυτό παρουσιάζεται μια προσπάθεια επίλυσης του προβλήματος της χαμηλής ταχύτητας στην περίπτωση που υπάρχουν μικρά πρόσωπα προς ανίχνευση. Κατά τη δεύτερη προσέγγιση, παρουσιάζεται μια πιο ορθόδοξη και ποιοτική μέθοδος κατάτμησης που βασίζεται στα χρωματικά ιστογράμματα των τίτλων ειδήσεων. Μετά την κατάτμηση ακολουθεί η σύνοψη των δελτίων ειδήσεων σε εικόνες-κλειδιά. Ως εικόνες-κλειδιά θεωρούνται εικόνες που φέρουν κάποιο τίτλο ειδήσεων, δηλαδή υψηλή πληροφορία. Παράλληλα, επιτυγχάνεται η εξαγωγή των λεζαντών με τους κύριους τίτλους ειδήσεων από τα πάνελ και τα ρεπορτάζ κάνοντας χρήση της ευκλείδειας απόστασης μεταξύ δύο εικόνων. Για την υλοποίηση της διπλωματικής χρησιμοποιήθηκε η βιβλιοθήκη OpenCV και συγκεκριμένα η διεπαφή C++.

# Abstract

This thesis aims to present and implement a satisfactory segmentation method for broadcast news videos. Specifically, Mega Channel broadcasts were used as our database. An important element of our research in order to segment such broadcasts is the clear understanding of their structure. In this project, the news videos are segmented into two types of segments. Panel conversations and news reports. More specifically, the existence of characteristic lines and news titles on Mega Channel panels has led to two different segmentation approaches, based on visual features. In the first approach, a method is proposed based on face and line detection. For this reason a method is presented, which can help solve the problem of low speed when detecting small faces. In the second approach a more orthodox and effective segmentation method is presented, based on the color histograms of news titles. The segmentation is followed by a summary of the news into key images. Key images are defined as images bearing a headline, i.e. high information. At the same time the extraction of the cues with the main news titles of the news is achieved, utilizing the Euclidean distance between two images. For the purposes of our research we used the OpenCV library and specifically its C++ interface.



# Περιεχόμενα

<b>Ευχαριστίες</b> .....	<b>3</b>
<b>Περίληψη</b> .....	<b>5</b>
<b>Abstract</b> .....	<b>6</b>
<b>Περιεχόμενα</b> .....	<b>8</b>
<b>Κατάλογος εικόνων</b> .....	<b>10</b>
<b>Κεφάλαιο 1: Εισαγωγή</b> .....	<b>13</b>
1.1 Σκοπός της διπλωματικής .....	13
1.2 Επισκόπηση της βιβλιογραφίας .....	13
1.3 Προσέγγιση και συνεισφορά της διπλωματικής .....	14
1.4 Δομή της διπλωματικής .....	15
<b>Κεφάλαιο 2: Η βιβλιοθήκη OpenCV και η ανίχνευση προσώπων και γραμμών</b> .....	<b>16</b>
2.1 Η βιβλιοθήκη OpenCV .....	16
2.2 Η ανίχνευση γραμμών μέσω της OpenCV .....	16
2.3 Η ανίχνευση προσώπων μέσω της OpenCV .....	19
<b>Κεφάλαιο 3: Ταχεία ανίχνευση προσώπων στα πάνελ</b> .....	<b>20</b>
3.1 Τα πάνελ του Mega και η ανίχνευση γραμμών .....	20
3.2 Η λύση του προβλήματος .....	22
<b>Κεφάλαιο 4: Μια πρώτη προσέγγιση κατάτμησης δελτίων ειδήσεων</b> .....	<b>25</b>



<b>Κεφάλαιο 5: Κατάτμηση των δελτίων ειδήσεων του Mega .....</b>	<b>27</b>
5.1 Εισαγωγή στις βασικές έννοιες κατάτμησης βίντεο .....	27
5.2 Κατάτμηση των δελτίων ειδήσεων του Mega .....	29
<b>Κεφάλαιο 6: Σύνοψη των δελτίων ειδήσεων του Mega .....</b>	<b>33</b>
<b>Κεφάλαιο 7: Εξαγωγή των τίτλων ειδήσεων .....</b>	<b>37</b>
7.1 Εξαγωγή των τίτλων ειδήσεων από τα πάνελ του Mega .....	37
7.2 Εξαγωγή των τίτλων ειδήσεων από τα ρεπορτάζ του Mega .....	39
<b>Κεφάλαιο 8: Πειραματικά αποτελέσματα .....</b>	<b>41</b>
8.1 Περιγραφή της βάσης δεδομένων .....	41
8.2 Το εργαλείο Anvil .....	42
8.3 Αποτελέσματα κατάτμησης, σύνοψης και εξαγωγής λεζαντών ...	43
8.4 Αποτελέσματα ταχείας ανίχνευσης προσώπων .....	49
<b>Κεφάλαιο 9: Συμπεράσματα .....</b>	<b>53</b>
9.1 Συνεισφορά της διπλωματικής εργασίας .....	53
9.2 Μελλοντικές ερευνητικές κατευθύνσεις .....	54
<b>Βιβλιογραφία .....</b>	<b>56</b>

# Κατάλογος εικόνων

2.1: Παράδειγμα γκρίζας εικόνας	17
2.2: Ανίχνευση ακμών στην εικόνα 2.1	17
2.3: Σύγκριση του μετασχηματισμού Hough και του αλγόριθμου LSWMS	18
3.1: Τα τρία κυρίαρχα πάνελ του Mega και οι χαρακτηριστικές γραμμές τους	21
3.2: Εικόνα όπου συνυπάρχει πάνελ και ρεπορτάζ	22
5.1: Μια τυχαία εικόνα	27
5.2: Το χρωματικό ιστόγραμμα της εικόνας 5.1	28
5.3: Μια εικόνα που απεικονίζει ένα πάνελ	29
5.4: Η κομμένη λεζάντα με τον τίτλο ειδήσεων της εικόνας 5.3	29
5.5: Η κατάτμηση ενός δελτίου ειδήσεων σε πάνελ και ρεπορτάζ	32
6.1: Δύο λεζάντες ειδήσεων από ένα τυχαίο ρεπορτάζ του Mega	33
6.2: Η εικόνα βάσης για τα ρεπορτάζ του Mega	34
6.3: Λεζάντα όπου αναφέρεται ο ρεπόρτερ	34
6.4: Λεζάντα όπου αναφέρεται ο ομιλητής	34
8.1: Εικόνα από το εργαλείο Anvil	42
8.2: Εικόνες-κλειδιά	44
8.3: Η εναλλαγή των τίτλων ειδήσεων στα πάνελ του Mega	46

8.4: Η εναλλαγή των τίτλων ειδήσεων στα ρεπορτάζ του Mega	48
8.5: Στιγμιότυπο της ταχείας ανίχνευσης προσώπων	51
8.6: Η μέση ταχύτητα και ακρίβεια για 6 ομιλητές	52
8.7: Η μέση ταχύτητα και ακρίβεια για 5 ομιλητές	52
8.8: Η μέση ταχύτητα και ακρίβεια για 2 ομιλητές	52
Πίνακας 8.1: Η μέση ταχύτητα και η ακρίβεια της ανίχνευσης προσώπων σε κάθε πάνελ ξεχωριστά χρησιμοποιώντας ανίχνευση γραμμών	50



## Κεφάλαιο 1: Εισαγωγή

### 1.1 Σκοπός της διπλωματικής

Η αυτόματη κατάτμηση βίντεο σε τμήματα με διαφορετικό περιεχόμενο αποτελεί αντικείμενο έρευνας εδώ και χρόνια στον τομέα της επεξεργασίας πολυμέσων. Η χρησιμοποίηση δελτίων ειδήσεων για το σκοπό αυτό ενδείκνυται αφενός μεν διότι τα δελτία ειδήσεων έχουν σαφή δομή και αφετέρου διότι με την επίτευξη της κατάτμησης δελτίων ειδήσεων γίνεται εφικτή η ύπαρξη ψηφιακών ειδησεογραφικών βιβλιοθηκών και η ανάκτηση πληροφορίας με βάση το περιεχόμενο. Η παρούσα διπλωματική εργασία μπορεί να χωριστεί νοητικά σε δύο ενότητες, όπου κάθε μία ακολουθεί μια εντελώς διαφορετική προσέγγιση όσον αφορά την κατάτμηση των δελτίων ειδήσεων και συγκεκριμένα μιας βάσης δεδομένων του καναλιού Mega [1]. Ουσιαστικά, απώτερος σκοπός είναι ο διαχωρισμός των βίντεο σε τμήματα ώστε το κάθε τμήμα (shot), το οποίο ουσιαστικά είναι μια σειρά από εικόνες (frames), να διαφέρει σημασιολογικά ως προς το περιεχόμενο από τα υπόλοιπα τμήματα.

### 1.2 Επισκόπηση της βιβλιογραφίας

Μια από τις πιο διαδεδομένες μεθόδους κατάτμησης δελτίων ειδήσεων αποτελεί ο αλγόριθμος anchor shot detection (ASD). Ένας υπάρχων τέτοιος αλγόριθμος [2] χρησιμοποιεί μια μέθοδο ομαδοποίησης για την εξατομίκευση υποψήφιων anchor shots και στη συνέχεια εφαρμόζει μια τεχνική κλάδεματος δύο σταδίων με σκοπό τη μείωση του αριθμού των εσφαλμένα ανιχνευμένων anchor shots. Η ομαδοποίηση και το κλάδεμα (pruning) πραγματοποιούνται χωρίς επίβλεψη (unsupervised).

Για την κατάτμηση δελτίων ειδήσεων αποτελεί επίσης συχνό φαινόμενο η χρήση διάφορων χαρακτηριστικών όπως λεξικών χαρακτηριστικών, ακουστικών χαρακτηριστικών και χαρακτηριστικών βίντεο. Ουσιαστικά ερευνάται η αποτελεσματικότητα ενός αριθμού πολυμέσων που χρησιμεύουν ως πιθανοί δείκτες των ορίων ενός τμήματος (shot) [3] [4]. Για παράδειγμα, η επιθυμητή κατάτμηση μπορεί να πραγματοποιηθεί με:

- μια σταθμισμένη μέθοδο ψηφοφορίας βασισμένη σε λεζάντες κειμένου [5]
- οπτικά χαρακτηριστικά και πληροφορίες κειμένου [6] [7]
- ροή του βίντεο και πληροφορίες κειμένου [8]
- ήχο και χρώματα [9]
- οπτικοακουστικά χαρακτηριστικά και πληροφορίες κειμένου [10]

Τα διάφορα οπτικοακουστικά και λεξικά χαρακτηριστικά μπορούν να ενοποιηθούν με τη χρήση conditional random fields (CRFs) με σκοπό την κατάτμηση του βίντεο [11]. Ουσιαστικά, τα χαρακτηριστικά εξάγονται σε μια ακολουθία από υποψήφιες θέσεις ορίων (boundary candidate positions) πάνω στο δελτίο ειδήσεων. Μια γραμμική αλυσίδα CRF χρησιμοποιείται ώστε να ανιχνευθεί κάθε υποψήφιος ως boundary/non boundary ετικέτα με βάση τα χαρακτηριστικά. Με τον τρόπο αυτό οι σχέσεις μεταξύ των ετικετών και η συνάφεια των πληροφοριών των χαρακτηριστικών συλλαμβάνονται αποδοτικά.

Τέλος, λόγω της προσπάθειας να αποκτηθεί πληροφορία με βάση το περιεχόμενο, τα τελευταία χρόνια έχει γίνει σημαντική έρευνα όσον αφορά το boundary shot detection. Ο διαχωρισμός των δελτίων ειδήσεων σε τμήματα έχει οδηγήσει στη κατασκευή πολλών μετρικών για τη μέτρηση της ομοιότητας μεταξύ των γειτονικών εικόνων ενός βίντεο, η οποία βασίζεται σε διάφορα διαθέσιμα χαρακτηριστικά του βίντεο. Ωστόσο, η χρήση πολλών χαρακτηριστικών μειώνει την απόδοση του boundary shot detection. Για κάθε απόφαση του boundary shot detection μπορεί να επιτευχθεί μείωση των χαρακτηριστικών με τη χρήση του rough-fuzzy operator [12].

### 1.3 Προσέγγιση και συνεισφορά της διπλωματικής

Η ανάλυση της δομής ενός δελτίου ειδήσεων και η εύρεση ευδιάκριτων χαρακτηριστικών αποτελεί ουσιαστικά το θεμέλιο λίθο κατά την προσπάθεια της κατάτμησης του. Ειδικότερα, στα δελτία ειδήσεων αποτελεί συχνό φαινόμενο η χρήση στατικής μορφής τηλεπαραθύρων, τα οποία χωρίζονται μεταξύ τους με χαρακτηριστικές γραμμές. Επιπλέον, οι τίτλοι ειδήσεων ενός δελτίου εμφανίζονται συνήθως σε συγκεκριμένο σημείο της οθόνης και με καθορισμένα χρώματα. Η παρούσα εργασία προσεγγίζει το πρόβλημα της κατάτμησης χρησιμοποιώντας τα παραπάνω οπτικά χαρακτηριστικά.

Συνοπτικά, όσον αφορά την κατάτμηση, η παρούσα εργασία αποδεικνύει ότι:

- Μέσω της ανίχνευσης γραμμών, δηλαδή τηλεπαραθύρων, σε συνδυασμό με την ανίχνευση προσώπων είναι δυνατή η κατάτμηση των ειδήσεων του Mega σε πάνελ και ρεπορτάζ.
- Μέσω της σύγκρισης χρωματικών ιστογραμμάτων και την έξυπνη χρήση του γεγονότος ότι οι τίτλοι ειδήσεων στα δελτία του Mega εμφανίζονται σε συγκεκριμένη θέση στην οθόνη, τα δελτία ειδήσεων μπορούν να κατατμηθούν σε πάνελ και ρεπορτάζ εύκολα, γρήγορα και αποτελεσματικά.

## 1.4 Δομή της διπλωματικής

Η έρευνα της διπλωματικής αυτής δεν περιορίζεται μόνο στο πρόβλημα της κατάτμησης των δελτίων ειδήσεων. Αυτό συμβαίνει διότι κατά την προσέγγιση του προβλήματος είτε ανέκυψαν επιμέρους προβλήματα που έπρεπε να λυθούν είτε κρίθηκε ωφέλιμο να δοθούν απαντήσεις σε παρεμφερή ζητήματα. Στη συνέχεια πραγματοποιείται μια παρουσίαση της δομής της διπλωματικής με σκοπό την κατανόηση του περιεχομένου της.

Πέρα από το πρώτο αυτό κεφάλαιο της εισαγωγής, η πρώτη ενότητα ξεκινάει με το δεύτερο κεφάλαιο, το οποίο ουσιαστικά αποτελεί μια εισαγωγή στη βιβλιοθήκη OpenCV και την αναγνώριση προσώπων και γραμμών. Επίσης, στο τρίτο κεφάλαιο παρουσιάζεται μια προσπάθεια αύξησης της ταχύτητας της ανίχνευσης προσώπων στα πάνελ του Mega. Επιπλέον, στο τέταρτο και τελευταίο κεφάλαιο της πρώτης ενότητας προτείνεται μια μέθοδος κατάτμησης των δελτίων ειδήσεων με τη χρήση της ανίχνευσης προσώπων και γραμμών.

Η δεύτερη ενότητα ξεκινάει με το πέμπτο κεφάλαιο όπου πραγματοποιείται κατάτμηση των δελτίων ειδήσεων του Mega με μια πιο ορθόδοξη και αποτελεσματική μέθοδο που βασίζεται στα χρώματα των τίτλων ειδήσεων. Επίσης, στο έκτο κεφάλαιο, μετά το πέρας της κατάτμησης των δελτίων παρουσιάζεται ένας τρόπος σύνοψής τους σε εικόνες. Η δεύτερη ενότητα τελειώνει με το έβδομο κεφάλαιο όπου εξάγονται οι λεζάντες με τους κύριους τίτλους ειδήσεων από τα πάνελ και τα ρεπορτάζ του Mega.

Τα πειραματικά αποτελέσματα της εργασίας παρουσιάζονται αναλυτικά στο όγδοο κεφάλαιο. Τέλος, στο ένατο και τελευταίο κεφάλαιο πραγματοποιείται μια σύνοψη των συμπερασμάτων της διπλωματικής και γίνεται αναφορά σε μελλοντικές ερευνητικές κατευθύνσεις.

## Κεφάλαιο 2: Η βιβλιοθήκη OpenCV και η ανίχνευση προσώπων και γραμμών

### 2.1 Η βιβλιοθήκη OpenCV

Η βιβλιοθήκη OpenCV αποτελεί το πιο διαδεδομένο εργαλείο ελεύθερου λογισμικού στην όραση υπολογιστών. Σχεδιάστηκε με σκοπό την επίτευξη υψηλής υπολογιστικής απόδοσης και τη δημιουργία εφαρμογών πραγματικού χρόνου. Είναι γραμμένη σε C++ και περιέχει διεπαφές σε C/C++, Python και Java. Συγκεκριμένα, στην παρούσα εργασία χρησιμοποιείται η διεπαφή C++, καθώς η χρήση της είναι περισσότερο διαδεδομένη σε σχέση με τις υπόλοιπες. Επιπλέον, η ανίχνευση προσώπων, η ανίχνευση γραμμών και η σύγκριση χρωματικών ιστογραμμάτων εφαρμόζονται με τη χρήση κατάλληλων συναρτήσεων της βιβλιοθήκης αυτής.

### 2.2 Η ανίχνευση γραμμών μέσω της OpenCV

Όσον αφορά την ανίχνευση γραμμών [13] η OpenCV παρέχει δύο συναρτήσεις που εκτελούν το γνωστό μετασχηματισμό γραμμών του Hough:

Τον κλασσικό μετασχηματισμό του Hough (Standard Hough Transform) και μια πιο αποδοτική εκτέλεσή του, τον πιθανοτικό μετασχηματισμό του Hough (Probabilistic Hough Line Transform), τον οποίο και χρησιμοποιούμε. Ωστόσο, για να εφαρμοστεί ο μετασχηματισμός αυτός πρέπει πρώτα μια εικόνα να μετατραπεί σε γκρίζα εικόνα (grayscale image, βλέπε εικόνα 2.1) και στη συνέχεια να εφαρμοστεί ανίχνευση ακμών (π.χ. Canny edge detection [14], βλέπε εικόνα 2.2). Γκρίζες εικόνες είναι οι ψηφιακές ασπρόμαυρες εικόνες που αποτελούνται αποκλειστικά από αποχρώσεις του γκρι.





Εικόνα 2.1: Παράδειγμα γκρίζας εικόνας



Εικόνα 2.2: Ανίχνευση ακμών στην εικόνα 2.1

Στο σημείο αυτό πρέπει να τονιστεί για όσους ενδιαφέρονται περισσότερο για την ανίχνευση γραμμών ότι ο μετασχηματισμός του Hough δημιουργεί πολλές γραμμές σε περιοχές με υψηλό θόρυβο. Αν και στη συγκεκριμένη εργασία δεν αποτέλεσε πρόβλημα, μια καλή προσπάθεια για την επίλυση αυτού του προβλήματος αποτελεί ο αλγόριθμος LSWMS (Line Segment detection using Weighted Mean Shift) [15] [16].



Εικόνες 2.3: Στις εικόνες αριστερά η ανίχνευση γραμμών πραγματοποιείται με το μετασχηματισμό του Hough. Στα δεξιά με τον αλγόριθμο LSWMS. [15]

## 2.3 Η ανίχνευση προσώπων μέσω της OpenCV

Όταν αναφερόμαστε στην επίλυση του προβλήματος της ανίχνευσης προσώπων πρέπει να λαμβάνουμε υπ' όψιν όχι μόνο την επίδοση του ανιχνευτή αλλά και την ταχύτητά του. Με τον όρο ταχύτητα εννοούμε κατά πόσο η επεξεργασία μιας εικόνας από τον ανιχνευτή για την εύρεση τυχόν προσώπων πραγματοποιείται γρήγορα χωρίς να απαιτείται μεγάλη υπολογιστική ισχύ. Με τον όρο επίδοση εννοούμε κατά πόσο ο ανιχνευτής βρίσκει τα υπάρχοντα πρόσωπα καθώς επίσης κατά πόσο βρίσκει πρόσωπα σε σημεία της εικόνας όπου δεν υπάρχουν πρόσωπα (false face detection).

Όσον αφορά την ανίχνευση προσώπων, παρότι δίνεται η δυνατότητα στο χρήστη να φτιάξει ένα δικό του ανιχνευτή, η OpenCV παρέχει έτοιμους δύο σειριακούς ταξινομητές [17], τον LBP (Local Binary Patterns) και τον Haar σε μορφή .xml αρχείων που επιτυγχάνουν αρκετά ποιοτική ανίχνευση προσώπων. Προτείνεται στους επίδοξους χρήστες αυτού του εργαλείου να χρησιμοποιήσουν έναν από αυτούς τους ταξινομητές καθώς η δημιουργία ποιοτικών ανιχνευτών είναι αρκετά επίπονη και χρονοβόρα διαδικασία. Επίσης, η ταχύτητα και η επίδοση του ανιχνευτή διαφέρει ανάλογα με την επιλογή του ταξινομητή. Για παράδειγμα, ο LBP ταξινομητής είναι αρκετά πιο γρήγορος σε σχέση με τον Haar, ωστόσο έχει χειρότερη επίδοση καθώς βρίσκει λιγότερα πρόσωπα και εμφανίζει μεγαλύτερο false face detection, γι' αυτό και η χρήση του είναι λιγότερο συνηθισμένη.

Η τεχνολογία έχει επιτύχει μεγάλη πρόοδο στον τομέα της όρασης υπολογιστών τα τελευταία χρόνια. Ένα από τα πιο δημοφιλή ζητήματα που προσπάθησε να λύσει είναι το ζήτημα της ανίχνευσης προσώπων. Ωστόσο, παρ' όλες τις βελτιώσεις που επιτεύχθηκαν η ανίχνευση προσώπων έχει ένα βασικό μειονέκτημα. Έχει μεγάλες υπολογιστικές απαιτήσεις, δηλαδή είναι μια αργή διαδικασία. Συγκεκριμένα, πολύ μικρά πρόσωπα απαιτούν πολλούς κύκλους μηχανής για να ανιχνευθούν. Σε ένα δελτίο ειδήσεων μάλιστα όπως αυτό του Mega υπάρχουν πάνελ των 5 ή και 6 προσώπων. Τα πρόσωπα αυτά είναι ιδιαίτερα μικρά και απαιτούν τόσο μεγάλη υπολογιστική ισχύ για να ανιχνευθούν που είναι αδύνατο να ανιχνεύονται πρόσωπα και ταυτόχρονα το βίντεο να παίζει σε πραγματικό χρόνο. Ουσιαστικά, όσο αυξάνεται η ποιότητα της ανίχνευσης, δηλαδή όσο πιο μικρά πρόσωπα είναι ικανός να ανιχνεύει ένας ανιχνευτής, τόσο πιο χρονοβόρα γίνεται η διαδικασία. Από την άλλη, εάν θελήσουμε να αυξήσουμε την ταχύτητα τότε πρέπει να θυσιάσουμε την ποιότητα της ανίχνευσης, δηλαδή ο ανιχνευτής δεν θα είναι ικανός να ανιχνεύει πολύ μικρά πρόσωπα.

## Κεφάλαιο 3: Ταχεία ανίχνευση προσώπων στα πάνελ

### 3.1 Τα πάνελ του Mega και η ανίχνευση γραμμών

Με απώτερο σκοπό την κατανόηση της επίλυσης του προβλήματος της χαμηλής ταχύτητας του ανιχνευτή όταν ανιχνεύονται μικρά πρόσωπα στα πάνελ των δελτίων ειδήσεων του Mega, είναι απαραίτητη η κατανόηση της δομής των πάνελ αυτών. Συγκεκριμένα, στα δελτία ειδήσεων του Mega χρησιμοποιούνται 3 βασικά πάνελ. Ένα πάνελ για 2 ομιλητές, ένα πάνελ για 5 ομιλητές και ένα πάνελ για 6 ομιλητές. Ο αριθμός των τηλεπαραθύρων που εμφανίζονται είναι όμοιος με τον αριθμό των υπαρχόντων ομιλητών.

Μείζονος σημασίας χαρακτηριστικό των τηλεοπτικών παραθύρων είναι η ύπαρξη συγκεκριμένων γραμμών, οι οποίες είναι χαρακτηριστικές ανάλογα με τον αριθμό των εμφανιζόμενων ομιλητών. Για παράδειγμα, στην περίπτωση όπου συζητάνε δύο ομιλητές υπάρχουν δύο τηλεπαραθύρα. Εάν πραγματοποιηθεί ανίχνευση γραμμών σε μια τέτοια εικόνα θα ανιχνευθεί μια χαρακτηριστική κάθετη γραμμή περίπου στη μέση της εικόνας. Συγκεκριμένα, στα βίντεο της βάσης δεδομένων που έχουμε στην κατοχή μας, το μέγεθος της κάθε εικόνας είναι 352 x 288 εικονοστοιχεία (pixels). Ο ανιχνευτής ανιχνεύει την κάθετη γραμμή  $x = 172$ , η οποία βρίσκεται σχεδόν στη μέση της εικόνας ( $352/2 = 176$ ). Στις εικόνες 3.1 γίνονται εύκολα αντιληπτές οι χαρακτηριστικές γραμμές που υπάρχουν σε κάθε πάνελ.



**Εικόνες 3.1:** Τα τρία κυρίαρχα πάνελ του Mega και οι χαρακτηριστικές γραμμές τους

Στα δελτία ειδήσεων του Mega ορισμένες φορές, αν και δεν αποτελεί συχνό φαινόμενο, συνδυάζεται η ύπαρξη τηλεπαραθύρων με την παρουσίαση ενός ρεπορτάζ, όπως φαίνεται στην εικόνα 3.2.



Εικόνα 3.2: Εικόνα όπου συνυπάρχει πάνελ και ρεπορτάζ

Ουσιαστικά το πάνελ αυτό είναι παρόμοιο με το πάνελ των 5 ομιλητών, απλά η οριζόντια χαρακτηριστική γραμμή του βρίσκεται λίγο πιο πάνω από τη μέση της εικόνας.

### 3.2 Η λύση του προβλήματος

Η ύπαρξη των παραπάνω χαρακτηριστικών γραμμών παρέχει μια πολύ σημαντική πληροφορία. Εάν σε μια εικόνα πραγματοποιηθεί ανίχνευση γραμμών και βρεθεί κάποια από αυτές τις χαρακτηριστικές γραμμές τότε κατά πάσα πιθανότητα η εικόνα αυτή απεικονίζει ένα πάνελ 2, 5 ή 6 ατόμων ανάλογα με τη γραμμή που βρέθηκε. Επίσης, στα πάνελ οι ομιλητές δεν κινούνται ιδιαίτερα. Οπότε, στην περίπτωση όπου στα τηλεπαράθυρα ανιχνεύονται μικρά πρόσωπα, επειδή η ανίχνευσή τους είναι μια χρονικά απαιτητική διαδικασία, αντί να πραγματοποιηθεί ανίχνευση προσώπων σε κάθε εικόνα του βίντεο απαιτώντας μεγάλη υπολογιστική ισχύ, μπορεί απλά να πραγματοποιηθεί ανίχνευση γραμμών σε κάθε εικόνα και ανίχνευση προσώπων ανά κάποιες εικόνες (π.χ. ανά 50 εικόνες). Η ανίχνευση γραμμών είναι αρκετά πιο γρήγορη διαδικασία σε σύγκριση με την ανίχνευση προσώπων [18]. Με αυτόν τον τρόπο επιτυγχάνεται ποιοτική ανίχνευση προσώπων χωρίς να απαιτείται τόσο μεγάλη υπολογιστική ισχύ.

Στοχεύοντας στην αύξηση της ταχύτητας της ανίχνευσης προσώπων στα δελτία ειδήσεων του Mega προτείνεται ο παρακάτω αλγόριθμος. Ο όρος αυστηρή ανίχνευση προσώπων αναφέρεται στο γεγονός ότι ο ανιχνευτής ψάχνει για μικρά πρόσωπα ενώ ο όρος χαλαρή ανίχνευση προσώπων αναφέρεται στο γεγονός ότι ο ανιχνευτής δεν ψάχνει για πολύ μικρά πρόσωπα.

---

### *Αλγόριθμος 1: Ταχεία ανίχνευση προσώπων*

---

1. Πραγματοποίησε ανίχνευση γραμμών σε κάθε εικόνα του βίντεο.
2. Αν η ανίχνευση γραμμών βρει τη χαρακτηριστική γραμμή που δηλώνει την ύπαρξη δύο τηλεπαραθύρων κάνε χαλαρή ανίχνευση προσώπων ανά 20 εικόνες.
3. Αν η ανίχνευση γραμμών βρει τις χαρακτηριστικές γραμμές που δηλώνουν την ύπαρξη 5 ή 6 τηλεπαραθύρων κάνε αυστηρή ανίχνευση προσώπων ανά 50 εικόνες.
4. Αν η ανίχνευση γραμμών δεν βρει κάποια από τις χαρακτηριστικές γραμμές κάνε χαλαρή ανίχνευση προσώπων σε κάθε εικόνα.
5. Μέτρα το χρόνο που απαιτεί η επεξεργασία κάθε εικόνας.

---

Ως επίλογος του κεφαλαίου αυτού είναι σημαντικό να γίνουν δύο παρατηρήσεις όσον αφορά την επίδοση του ανιχνευτή προσώπων. Πρώτον, όπως αναφέρθηκε παραπάνω, ένα από τα βασικά πλεονεκτήματα του LBP ταξινομητή σε σχέση με τον Haar είναι η ταχύτητά του, ωστόσο υστερεί όσον αφορά την απόδοσή του. Εφόσον πλέον η ανίχνευση προσώπων πραγματοποιείται αυστηρά σε πολύ λίγες εικόνες του βίντεο, θα προτιμηθεί σαφώς η χρήση του Haar ταξινομητή για την επίτευξη καλύτερων αποτελεσμάτων, αφού το μειονέκτημα της ταχύτητας έχει ήδη καλυφθεί.

Δεύτερον, η επιλογή του σειριακού ταξινομητή παρουσιάζει ακόμα μια σημαντική παράμετρο. Στην υπάρχουσα εργασία χρησιμοποιήθηκε κατάλληλο .xml αρχείο ώστε να ανιχνεύονται πρόσωπα που βρίσκονται στην πρόσθια όψη (frontal faces). Καθοριστικό ρόλο φυσικά στην επιλογή αυτή διαδραμάτισε το γεγονός ότι στα δελτία ειδήσεων τα πρόσωπα εμφανίζονται με αυτόν τον τρόπο. Ωστόσο, όπως είναι ευνόητο, ο χρησιμοποιούμενος ανιχνευτής δεν είναι ικανός να ανιχνεύσει πρόσωπα που εμφανίζονται για παράδειγμα σε στάση προφίλ.



## Κεφάλαιο 4: Μια πρώτη προσέγγιση κατάτμησης δελτίων ειδήσεων

Πρέπει να καταστεί σαφές ότι στο κεφάλαιο αυτό παρουσιάζεται ένας κάπως ανορθόδοξος και μη συνηθισμένος τρόπος κατάτμησης βίντεο. Αντίθετα, στο επόμενο κεφάλαιο πραγματοποιείται μια ενδεέστερη έρευνα όσον αφορά την κατάτμηση βίντεο. Ωστόσο, στο προηγούμενο κεφάλαιο έγινε μια προσπάθεια να αυξηθεί η ταχύτητα της ανίχνευσης προσώπων στα δελτία ειδήσεων του Mega. Με τον τρόπο αυτό έχουμε φτάσει πολύ κοντά στην κατάτμηση ενός δελτίου ειδήσεων. Ουσιαστικά, ανιχνεύοντας πρόσωπα και χαρακτηριστικές γραμμές, ο αλγόριθμος του κεφαλαίου 3.2 μπορεί να αποφανθεί για την ύπαρξη ή μη ύπαρξη τηλεπαραθύρων.

Κατ' αρχάς πρέπει να αποσαφηνιστεί ο τρόπος με τον οποίο θα κατατμηθούν τα δελτία ειδήσεων. Γενικότερα, ένα δελτίο ειδήσεων έχει την εξής δομή: Ειδήσεις - Αθλητικά - Καιρός. Ωστόσο, καθώς η βάση δεδομένων που βρίσκεται στην κατοχή μας απαρτίζεται από δελτία ειδήσεων χωρίς την ύπαρξη αθλητικών ειδήσεων καθώς και του καιρού, η κατάτμηση βίντεο θα επικεντρωθεί καθαρά στις ειδήσεις του Mega. Συγκεκριμένα, τα τμήματα των δελτίων ειδήσεων που εξάγονται μπορούν να λάβουν τριών ειδών ετικέτες: Πάνελ, ρεπορτάζ, συνδυασμός πάνελ και ρεπορτάζ (βλέπε εικόνα 3.2).

Η βασική λογική με την οποία είναι δυνατόν να κατατμηθεί ένα δελτίο ειδήσεων του Mega χρησιμοποιώντας ανίχνευση προσώπων και γραμμών είναι η εξής:

Κατ' αρχήν θεωρείται ότι έχει εφαρμοστεί ο αλγόριθμος του κεφαλαίου 3.2. Επίσης, γίνονται οι εξής λογικές παραδοχές. Κάθε εικόνα του βίντεο λαμβάνει μια ετικέτα. Στην περίπτωση όπου ανιχνευτεί η κάθετη χαρακτηριστική γραμμή περίπου στη μέση μιας εικόνας και δύο πρόσωπα τότε η εικόνα αυτή λαμβάνει την ετικέτα πάνελ. Το ίδιο συμβαίνει αν ανιχνευτεί κάποια από τις χαρακτηριστικές γραμμές για το πάνελ των 6 ομιλητών και 6 πρόσωπα. Εάν πάλι ανιχνευτεί η χαρακτηριστική γραμμή που δηλώνει την ύπαρξη πάνελ για 5 ομιλητές υπάρχουν δύο περιπτώσεις. Είτε η εικόνα απεικονίζει ένα πάνελ, οπότε αναμένουμε να ανιχνευτούν 5 πρόσωπα, είτε απεικονίζει πάνελ και ρεπορτάζ ταυτόχρονα οπότε αναμένουμε να ανιχνευτούν 2 ή περισσότερα πρόσωπα. Σε κάθε άλλη περίπτωση η εικόνα θεωρείται ότι αποτελεί μέρος ενός ρεπορτάζ, οπότε λαμβάνει την ανάλογη ετικέτα.

Ο αλγόριθμος που παρουσιάστηκε στο κεφάλαιο 3.2 είναι εύκολο να προσαρμοστεί στα παραπάνω δεδομένα. Ουσιαστικά, το μόνο που χρειάζεται να προστεθεί είναι η μέτρηση του αριθμού των προσώπων που ανιχνεύονται στην περίπτωση ύπαρξης πάνελ και η απονομή κατάλληλων ετικετών στις εικόνες του βίντεο.

Όπως είναι φυσιολογικό η μέθοδος αυτή δεν είναι 100% επιτυχής. Κάθε εικόνα του βίντεο δεν είναι σίγουρο ότι θα λάβει τη σωστή ετικέτα (ρεπορτάζ, πάνελ ή ρεπορτάζ και πάνελ). Για να μειωθεί το σφάλμα πρέπει ο κώδικας να είναι πιο ανεκτικός στον αριθμό των προσώπων που ανιχνεύονται, ώστε να αποφευχθεί το σφάλμα του ανιχνευτή προσώπων. Για παράδειγμα, είναι πιθανόν ο ανιχνευτής προσώπων να ανιχνεύσει σε κάποια εικόνα τα 5 από τα 6 υπάρχοντα πρόσωπα. Οπότε, στην περίπτωση αυτή, αν ανιχνευτεί μια χαρακτηριστική γραμμή που δηλώνει την ύπαρξη 6 ομιλητών και 5 πρόσωπα, θα πρέπει να δοθεί στην εικόνα η σωστή ετικέτα: πάνελ.

Τέλος, πρέπει να αποσαφηνιστεί ότι ακόμα και αν εφαρμοστεί μεγαλύτερη ανεκτικότητα όσον αφορά το σφάλμα του ανιχνευτή προσώπων και πάλι δεν είναι δυνατόν να υπάρξει 100% επιτυχία. Αυτό σημαίνει δηλαδή ότι, για παράδειγμα, στις 5000 συνεχόμενες εικόνες όπου εμφανίζεται ένα πάνελ μπορεί κάποιες, αν και λίγες από αυτές, να έχουν λάβει λανθασμένη ετικέτα.

Η ύπαρξη αυτού του σφάλματος δημιουργεί προβλήματα στην επίτευξη μιας σωστής κατάτμησης βίντεο. Το πρόγραμμα δηλαδή, στην περίπτωση όπου κάποιες εικόνες έχουν λάβει λάθος ετικέτα, ίσως θεωρήσει ότι ξεκινάει ένα νέο τμήμα. Η λύση αυτού του προβλήματος είναι απλή. Το πρόγραμμα θα πρέπει να βρίσκει έναν αριθμό από συνεχόμενες εικόνες με την ίδια ετικέτα ώστε να θεωρήσει ότι ξεκινάει ένα νέο τμήμα.

Ωστόσο, μια τέτοια προσέγγιση για να κατατμηθούν μεγάλα βίντεο 50 λεπτών όπως είναι τα δελτία ειδήσεων είναι ιδιαίτερα χρονοβόρα, αφού ακόμα και αν χρησιμοποιηθεί ο ταχύς αλγόριθμος που περιγράφηκε στο κεφάλαιο 3.2, η επεξεργασία παραμένει αργή. Παρακάτω παρουσιάζεται μια πιο κλασσική και γρήγορη προσέγγιση επίλυσης του προβλήματος της κατάτμησης ενός βίντεο.

## Κεφάλαιο 5: Κατάτμηση των δελτίων ειδήσεων του Mega

### 5.1 Εισαγωγή στις βασικές έννοιες κατάτμησης βίντεο

#### Ανίχνευση ορίων τμημάτων (boundary shot detection)

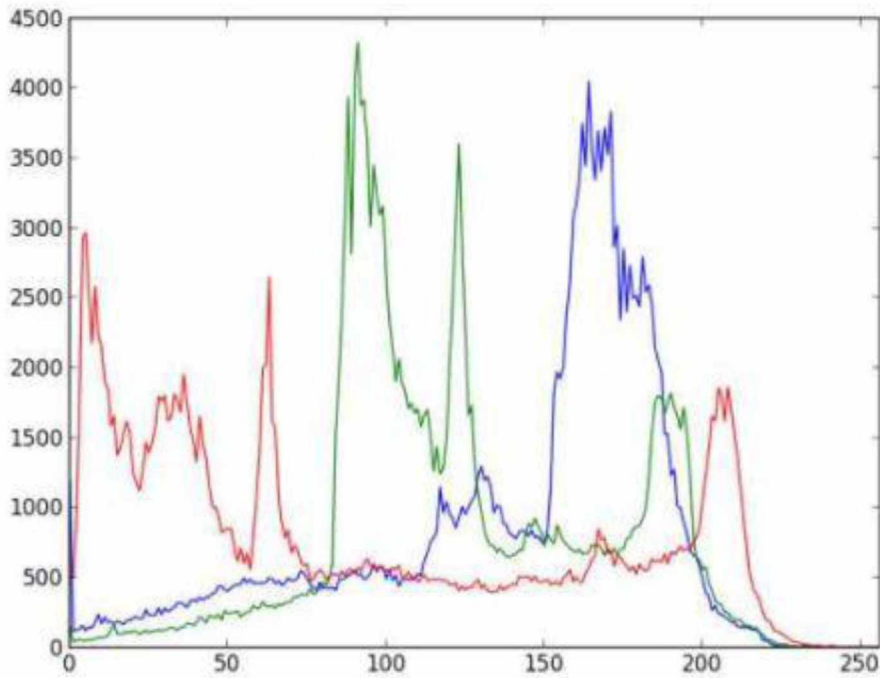
Boundary shot detection ονομάζεται η αυτόματη κατάτμηση ενός ψηφιακού βίντεο σε τμήματα που ονομάζονται shots. Κάθε τμήμα είναι ουσιαστικά μια σειρά συνεχόμενων εικόνων με την ίδια θεματολογία. Μια πολύ καλή ανάλυση και σύγκριση των βασικών μεθόδων που χρησιμοποιούνται για την πραγματοποίηση boundary shot detection πραγματοποιείται στη δημοσίευση “Comparison of video shot boundary detection techniques” [19]. Μία από αυτές τις μεθόδους και μάλιστα ιδιαίτερα κλασσική είναι η χρήση χρωματικού ιστογράμματος, την οποία θα περιγράψουμε ευθύς αμέσως.

#### Χρωματικό ιστόγραμμα

Χρωματικό ιστόγραμμα (color histogram) ή πιο απλά ιστόγραμμα ονομάζεται η απεικόνιση της κατανομής των χρωμάτων μιας εικόνας. Για παράδειγμα, η εικόνα 5.2 απεικονίζει το ιστόγραμμα μιας εικόνας στο τρισδιάστατο χώρο RGB (red-green-blue).



**Εικόνα 5.1: Μια τυχαία εικόνα**



Εικόνα 5.2: Το χρωματικό ιστόγραμμα της εικόνας 5.1

Η OpenCV δίνει τη δυνατότητα στο χρήστη να συγκρίνει τα χρωματικά ιστογράμματα δύο εικόνων μέσω της συνάρτησης `compareHist`. Η συσχέτιση των ιστογραμμάτων αυτών κυμαίνεται από  $-1$  έως  $1$ . Όσο πιο κοντά στη μονάδα είναι το αποτέλεσμα της σύγκρισης δύο εικόνων μέσω της `compareHist` τόσο περισσότερο χρωματικά όμοιες είναι οι δύο εικόνες.

## Σημαντικότητα (Saliency)

Στην περίπτωση της όρασης υπολογιστών ως saliency μπορεί να θεωρηθεί οτιδήποτε ξεχωρίζει σε σχέση με το σύνολο, οτιδήποτε ελκύει την προσοχή του ανθρώπινου ματιού. Για παράδειγμα saliency μπορεί να προκύψει από μια μαύρη εικόνα με μια άσπρη κουκίδα στη μέση. Στην περίπτωση της δικής μας βάσης δεδομένων, παρατηρήσαμε ότι τα δελτία ειδήσεων του Mega και συγκεκριμένα τα πάνελ έχουν συγκεκριμένη μορφή. Όπως είναι ευρέως αντιληπτό, κατά τη διάρκεια της συζήτησης ενός θέματος στα τηλεπαράθυρα του Mega, κάτω από τα τηλεπαράθυρα με άσπρο χρώμα και σε κόκκινο φόντο εμφανίζονται οι τίτλοι ειδήσεων του Mega που αναφέρονται στο θέμα της συζήτησης των ομιλητών. Η διαπίστωση αυτή συνέβαλε σημαντικά στη σύλληψη του τρόπου με τον οποίο μπορούν να κατατμηθούν τα δελτία ειδήσεων του Mega.

## 5.2 Κατάτμηση των δελτίων ειδήσεων του Mega

Σκοπός της κατάτμησης του κεφαλαίου αυτού είναι ο διαχωρισμός των δελτίων ειδήσεων σε τμήματα, όπου κάθε τμήμα θα λαμβάνει την ετικέτα πάνελ ή ρεπορτάζ. Δηλαδή σε τμήματα όπου είτε θα γίνεται συζήτηση διάφορων ομιλητών σε τηλεπαράθυρα είτε θα παρουσιάζεται κάποιο ρεπορτάζ. Ένας τρόπος να ανιχνευτεί η ύπαρξη πάνελ είναι να επαληθευτεί εάν υπάρχει κάποιος τίτλος ειδήσεων κάτω από τα τηλεπαράθυρα. Όπως αναφέρθηκε προηγουμένως, η ύπαρξη τίτλων ειδήσεων στα τηλεπαράθυρα του Mega αποτελεί ένα διακριτό χαρακτηριστικό τους. Με τη βοήθεια της OpenCV παίρνουμε μια τυχαία εικόνα από ένα δελτίο ειδήσεων όπου εμφανίζονται τηλεπαράθυρα και “κόβουμε” την είδηση που υπάρχει στο κάτω μέρος της εικόνας.



Εικόνα 5.3: Μια εικόνα που απεικονίζει ένα πάνελ από κάποιο τυχαίο δελτίο ειδήσεων του Mega

ΣΑΛΟΣ ΑΠΟ ΤΙΣ ΑΠΟΚΑΛΥΨΕΙΣ ΓΙΑ ΤΑ ΟΜΟΛΟΓΑ

Εικόνα 5.4: Η κομμένη λεζάντα με τον τίτλο ειδήσεων της εικόνας 5.3

Έστω ότι η εικόνα 5.4 ονομάζεται εικόνα βάσης. Αρχικά, ένα δελτίο ειδήσεων ξεκινάει να παίζει με τη βοήθεια της OpenCV. Ωστόσο, δεν μας ενδιαφέρουν ολόκληρες οι εικόνες (frames) του βίντεο, καθώς αυτές “κόβονται” με τον ίδιο ακριβώς τρόπο που κόπηκε η εικόνα βάσης από την εικόνα 5.3. Η εκάστοτε κομμένη εικόνα του βίντεο πρέπει να εξεταστεί αν φέρει κάποιο τίτλο ειδήσεων, γι’ αυτό ονομάζεται εξεταζόμενη εικόνα. Για να πραγματοποιηθεί ο έλεγχος αυτός, συγκρίνεται το χρωματικό ιστογράμμα της εικόνας βάσης με το χρωματικό ιστογράμμα της εξεταζόμενης εικόνας μέσω της συνάρτησης `compareHist`. Εάν τα χρωματικά ιστογράμματα έχουν υψηλό βαθμό συσχέτισης (για παράδειγμα πάνω από 0.85), τότε θεωρείται ότι η εξεταζόμενη εικόνα όντως φέρει κάποιο τίτλο ειδήσεων, οπότε η αρχική εικόνα αποτελεί μέρος ενός πάνελ. Στην αντίθετη περίπτωση αποτελεί μέρος ενός ρεπορτάζ. Η περίπτωση όπου παρατηρείται η ύπαρξη πάνελ και ρεπορτάζ ταυτόχρονα, γεγονός που αναφέρθηκε και στο τέλος του κεφαλαίου 3.1, θεωρείται υποπερίπτωση της ύπαρξης πάνελ, καθώς η μορφή τους είναι παρόμοια.

Οπότε, μέσω τη σύγκρισης χρωματικών ιστογραμμάτων είναι δυνατόν να βρεθούν ποιες εικόνες ενός δελτίου ειδήσεων του Mega αποτελούν μέρος ενός πάνελ και ποιες μέρος ενός ρεπορτάζ. Ωστόσο, είναι λογικό ορισμένες εικόνες να έχουν λάβει λανθασμένη ετικέτα. Για το λόγο αυτό, για να χαρακτηριστεί ένα τμήμα ως πάνελ ή ρεπορτάζ πρέπει να βρεθεί ένας αριθμός από διαδοχικές εικόνες με την ίδια ετικέτα. Για παράδειγμα, για να δοθεί σε ένα τμήμα η ετικέτα πάνελ θα πρέπει να βρεθεί ένας αριθμός από διαδοχικές εικόνες που φέρουν κάποιο τίτλο ειδήσεων, δηλαδή εξεταζόμενες εικόνες που είναι χρωματικά όμοιες με την εικόνα βάσης. Αυτό συμβαίνει έτσι ώστε αν κάποια εικόνα του τμήματος δεν περάσει για κάποιο λόγο το κατώφλι που έχει τεθεί όσον αφορά τη χρωματική ομοιότητα, να μη δημιουργήσει πρόβλημα στην κατάτμηση του βίντεο, δηλαδή να μη θεωρήσει το πρόγραμμα ότι ξεκινάει ένα νέο τμήμα με ετικέτα ρεπορτάζ. Το ίδιο και για την αντίθετη περίπτωση. Εάν δηλαδή για κάποιο λόγο κάποια ή κάποιες από τις εξεταζόμενες εικόνες ενός ρεπορτάζ θεωρηθούν από το πρόγραμμά μας χρωματικά όμοιες με την εικόνα βάσης, το πρόγραμμά μας δεν πρέπει να θεωρήσει ότι ξεκινάει ένα νέο τμήμα με ετικέτα πάνελ. Συγκεκριμένα, η εργασία αυτή χρησιμοποιεί 350 διαδοχικές εικόνες σαν μια τέτοια δικλείδα ασφαλείας.

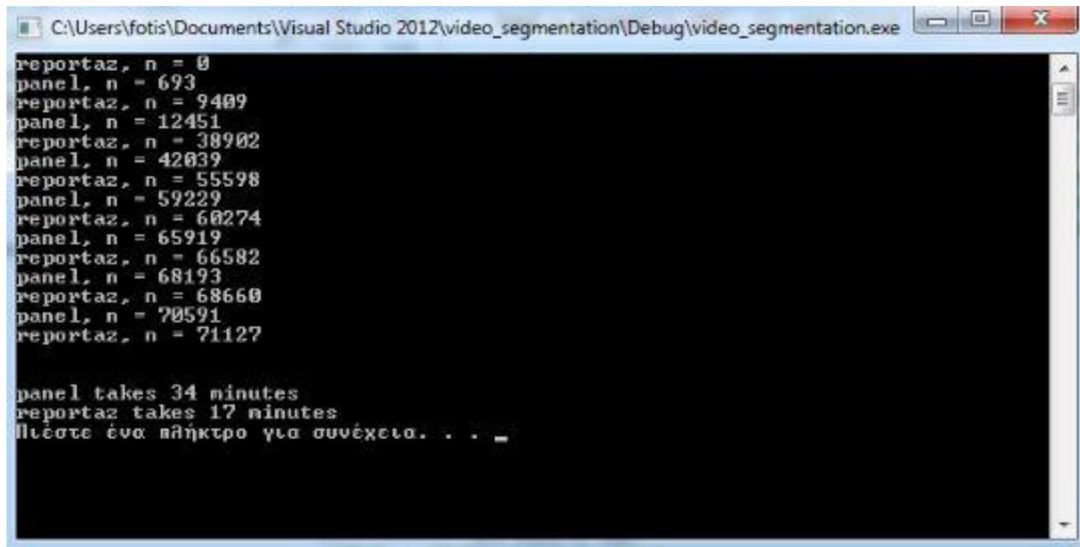
Ο παρακάτω αλγόριθμος περιγράφει τον τρόπο με τον οποίο επιτεύχθηκε η κατάτμηση των δελτίων ειδήσεων του Mega και μάλιστα με ιδιαίτερα ικανοποιητικά αποτελέσματα.

---

### *Αλγόριθμος 2: Κατάτμηση δελτίων ειδήσεων*

---

1. Κόψε ένα τίτλο ειδήσεων και δημιούργησε την εικόνα βάσης.
  2. Σε κάθε επανάληψη, δηλαδή για κάθε εικόνα του βίντεο, δημιούργησε την εξεταζόμενη εικόνα.
  3. Σύγκρινε το χρωματικό ιστόγραμμα της εικόνας βάσης με αυτό της εξεταζόμενης εικόνας.
  4. Εάν τα ιστογράμματα δεν είναι όμοια για 350 συνεχόμενες εικόνες, θεώρησε ότι πριν από 350 εικόνες ξεκίνησε ένα νέο τμήμα με ετικέτα ρεπορτάζ. Αποθήκευσε σε έναν πίνακα (πρώτος πίνακας) σε ποια εικόνα ξεκινάει το τμήμα και σε ένα δεύτερο πίνακα την ετικέτα του.
  5. Εάν τα ιστογράμματα είναι όμοια για 350 συνεχόμενες εικόνες, θεώρησε ότι πριν από 350 εικόνες ξεκίνησε ένα νέο τμήμα με ετικέτα πάνελ. Αποθήκευσε στον πρώτο πίνακα σε ποια εικόνα ξεκινάει το τμήμα και στο δεύτερο πίνακα την ετικέτα του.
  6. Πήγαινε στο βήμα 2 έως ότου τελειώσουν όλες οι εικόνες του βίντεο.
  7. Υπολόγισε σε πόσα λεπτά από τη συνολική διάρκεια του δελτίου ειδήσεων γίνεται ρεπορτάζ και σε πόσα συζήτηση σε τηλεπαράθυρα.
-



```
C:\Users\fofis\Documents\Visual Studio 2012\video_segmentation\Debug\video_segmentation.exe
reportaz, n = 0
panel, n = 693
reportaz, n = 9409
panel, n = 12451
reportaz, n = 38902
panel, n = 42039
reportaz, n = 55598
panel, n = 59229
reportaz, n = 60274
panel, n = 65919
reportaz, n = 66582
panel, n = 68193
reportaz, n = 68660
panel, n = 70591
reportaz, n = 71127

panel takes 34 minutes
reportaz takes 17 minutes
Πιέστε ένα πλήκτρο για συνέχεια. . . _
```

**Εικόνα 5.5:** Η κατάτμηση ενός δελτίου ειδήσεων σε πάνελ και ρεπορτάζ. Η μεταβλητή  $n$  δηλώνει τον αριθμό της εικόνας (frame) όπου ξεκινάει ένα νέο τμήμα.

Στην εικόνα 5.5 εμφανίζονται οι ετικέτες και η πρώτη εικόνα (frame) του κάθε τμήματος για το δελτίο ειδήσεων του Mega που προβλήθηκε στις 23 Μαρτίου του 2007. Στη συνέχεια εμφανίζεται ο συνολικός αριθμός των λεπτών όπου το δελτίο ειδήσεων δείχνει κάποιο πάνελ ή παίζει ένα ρεπορτάζ. Όπως παρατηρείται στο δελτίο αυτό είναι συχνότερη η χρήση πάνελ. Εάν πιεστεί οποιοδήποτε πλήκτρο το πρόγραμμα συνεχίζει στη σύνοψη του δελτίου ειδήσεων σε εικόνες-κλειδιά. Στη διαδικασία αυτή θα αναφερθούμε ενδεότερα στη συνέχεια.



## Κεφάλαιο 6: Σύνοψη των δελτίων ειδήσεων του Mega

Η σύνοψη (summarization) ενός βίντεο σε εικόνες-κλειδιά είναι μια διαδικασία που ακολουθεί συνήθως μετά την κατάτμησή του. Αυτό πρακτικά σημαίνει ότι για την επίτευξη της σύνοψης των δελτίων ειδήσεων θεωρείτε ότι έχουν πραγματοποιηθεί τα 7 βήματα του αλγορίθμου που περιγράφηκε στο κεφάλαιο 5.2. Στην εργασία μας, η ποιότητα της σύνοψης είναι στενά συνδεδεμένη με την ποιότητα της κατάτμησης. Εάν έχει γίνει σωστή κατάτμηση αυτό σημαίνει ότι έχουμε και ποιοτική σύνοψη. Το πλεονέκτημα της έρευνάς μας είναι η επίτευξη ιδιαίτερα ικανοποιητικών αποτελεσμάτων χωρίς να χρησιμοποιούμε περίπλοκες μεθόδους.

Η σύνοψη ουσιαστικά ενός δελτίου ειδήσεων, αλλά και ενός βίντεο γενικότερα, επιτυγχάνεται διαλέγοντας συγκεκριμένες εικόνες-κλειδιά από κάθε τμήμα-shot του βίντεο. Για παράδειγμα, σε μια ταινία οι εικόνες-κλειδιά, πιο γνωστά και ως screenshots, μπορεί να είναι εικόνες με ιδιαίτερη συναισθηματική φόρτιση. Στην περίπτωση των δελτίων ειδήσεων, διαλέξαμε ως εικόνες-κλειδιά εικόνες που εμφανίζουν κάποιο τίτλο ειδήσεων, διότι φέρουν υψηλή πληροφορία.

Στο κεφάλαιο 5.2 πραγματοποιήθηκε εκτενής αναφορά στον τρόπο με τον οποίο μπορούν να ανιχνευτούν οι τίτλοι ειδήσεων στα πάνελ του Mega. Καθώς αυτοί εμφανίζονται σε μια συγκεκριμένη θέση στις εικόνες (frames) των βίντεο, η χρωματική σύγκριση με έναν τυχαία παρμένο τίτλο ειδήσεων ως εικόνα βάσης δίνει τη ζητούμενη λύση στο πρόβλημα. Ωστόσο, στην περίπτωση των ρεπορτάζ δεν είναι δυνατόν να συμβεί το ίδιο. Αυτό οφείλεται στο γεγονός ότι, δυστυχώς, αν και οι τίτλοι ειδήσεων των ρεπορτάζ του Mega εμφανίζονται και αυτοί σε συγκεκριμένο σημείο, διαφέρουν αρκετά χρωματικά. Για παράδειγμα, αν συγκριθούν τα χρωματικά ιστογράμματα των εικόνων 6.1, το αποτέλεσμα που δίνει η OpenCV είναι 0.53, το οποίο είναι αρκετά μακριά από τη μονάδα.



Εικόνες 6.1: Δύο λεζάντες ειδήσεων από ένα τυχαίο ρεπορτάζ του Mega

Για το λόγο αυτό ως εικόνα βάσης για τα ρεπορτάζ του Mega χρησιμοποιείται η εικόνα 6.2, η οποία είναι χρωματικά όμοια σε όλους τους τίτλους ειδήσεων:



**Εικόνα 6.2: Η εικόνα βάσης για τα ρεπορτάζ του Mega**

Οπότε με την ίδια λογική, εάν συγκριθεί χρωματικά η εικόνα βάσης με τις αντίστοιχες εικόνες προς εξέταση, τότε μπορεί να εξακριβωθεί αν κάποια εικόνα ενός τμήματος με ετικέτα ρεπορτάζ περιέχει κάποιο τίτλο ειδήσεων.

Δυστυχώς, η εύρεση των τίτλων ειδήσεων στα ρεπορτάζ του Mega έχει ένα βασικό μειονέκτημα. Πέρα από τους τίτλους ειδήσεων ανιχνεύονται επίσης λεζάντες που απεικονίζουν τους ρεπόρτερ και τους ομιλητές του εκάστοτε ρεπορτάζ, καθώς η αναφορά τους γίνεται με την ίδια μορφή. Οι εικόνες 6.3 και 6.4 είναι διαφωτιστικές:



**Εικόνα 6.3: Λεζάντα όπου αναφέρεται ο ρεπόρτερ**



**Εικόνα 6.4: Λεζάντα όπου αναφέρεται ο ομιλητής**

Ξεκινώντας τη διαδικασία της σύνοψης, χρησιμοποιείται η βιβλιοθήκη OpenCV ώστε να ξεκινήσει το ίδιο βίντεο με αυτό που κατατμήθηκε προηγουμένως. Όπως αναφέρθηκε και στο κεφάλαιο 5.2, κατά τη διάρκεια της κατάτμησης έχουν αποθηκευτεί σε ένα πίνακα οι εικόνες στις οποίες αρχίζει ένα ρεπορτάζ και οι εικόνες στις οποίες αρχίζει ένα πάνελ. Επιπλέον, έχει δημιουργηθεί ένας ίδιων διαστάσεων πίνακας όπου έχουν αποθηκευτεί οι ετικέτες των τμημάτων του δελτίου, για παράδειγμα 0 για ρεπορτάζ και 1 για πάνελ. Όπως είναι ευνόητο, η μεταφορά από το ένα τμήμα στο επόμενο έχει ως αποτέλεσμα την εναλλαγή των ετικετών τους. Κατά τη σύνοψη του δελτίου ειδήσεων εφιστούμε την προσοχή μας στην ετικέτα του κάθε τμήματος, οπότε υπάρχουν δύο περιπτώσεις:

Στην πρώτη περίπτωση, εάν η ετικέτα δηλώνει την ύπαρξη ρεπορτάζ τότε αναζητείται μια εικόνα που να περιέχει κάποιο τίτλο ειδήσεων τύπου ρεπορτάζ. Για το λόγο αυτό συγκρίνεται το χρωματικό ιστόγραμμα της εικόνας 6.2 με το χρωματικό ιστόγραμμα όμοια κομμένων εικόνων του τμήματος. Εάν βρεθεί κάποιος τίτλος ειδήσεων, δηλαδή χρωματική ομοιότητα, τότε αποθηκεύεται ολόκληρη η εικόνα.

Στη συνέχεια ελέγχεται ο αριθμός των διαδοχικών εικόνων που υπολείπονται για την ολοκλήρωση του ρεπορτάζ. Εάν υπολείπεται ένας σεβαστός αριθμός από εικόνες, για παράδειγμα πάνω από 7500 (frames) που αντιστοιχούν σε 5 λεπτά του βίντεο, τότε το πρόγραμμα προχωράει 5 λεπτά και η έρευνα για κάποια εικόνα που φέρει κάποιο τίτλο ειδήσεων συνεχίζεται. Εάν βρεθεί ξανά κάποιος τίτλος ειδήσεων τότε αποθηκεύεται επίσης ολόκληρη η εικόνα. Επαναλαμβάνεται το ίδιο έως ότου λήξει το ρεπορτάζ.

Εάν ψάχνοντας σε όλο το ρεπορτάζ δε βρεθεί κάποια εικόνα που να φέρει κάποιο τίτλο ειδήσεων, αποθηκεύεται ως εικόνα-κλειδί η τελευταία εικόνα του τμήματος.

Στη δεύτερη περίπτωση, εάν η ετικέτα ενός τμήματος δηλώνει την ύπαρξη πάνελ τότε η εύρεση των εικόνων-κλειδιών ακολουθεί παρόμοια διαδικασία.

Όπως είναι ευνόητο, σε ένα τμήμα με ετικέτα πάνελ όλες οι εικόνες του τμήματος φέρουν κάποιο τίτλο ειδήσεων, οπότε οι εικόνες-κλειδιά βρίσκονται άμεσα. Στην περίπτωση των ρεπορτάζ απαιτείται μεγαλύτερης χρονικής διάρκειας αναζήτηση ώστε να βρεθούν εικόνες που να εμφανίζουν κάποιο τίτλο ειδήσεων. Τέλος, ο αριθμός των εικόνων-κλειδιών που εξάγονται από ένα τμήμα αποτελεί ταυτόχρονα ένα δείκτη του μεγέθους του.

Ο παρακάτω αλγόριθμος συνοψίζει τα βήματα που χρειάστηκαν ώστε να επιτευχθεί τελικά μια ποιοτική σύνοψη των δελτίων ειδήσεων του Mega. Επισημάνεται ξανά ότι τα 7 βήματα του αλγορίθμου κατάτμησης που παρουσιάστηκε στο κεφάλαιο 5.2 θεωρούνται ότι έχουν πραγματοποιηθεί.

---

### *Αλγόριθμος 3: Σύνοψη δελτίων ειδήσεων*

---

1. Έλεγξε την ετικέτα του κάθε τμήματος.
  2. Εάν είναι ρεπορτάζ, ψάξε για κάποιο τίτλο ειδήσεων τύπου ρεπορτάζ. Εάν είναι πάνελ, ψάξε για κάποιο τίτλο ειδήσεων τύπου πάνελ.
  3. Εάν δε βρεθεί κάποια εικόνα που να φέρει κάποιο τίτλο ειδήσεων σε όλο το τμήμα, αποθήκευσε ως εικόνα-κλειδί την τελευταία εικόνα του τμήματος.
  4. Εάν βρεθεί μια τέτοια εικόνα αποθήκευσέ την και έλεγξε σε πόσα λεπτά τελειώνει το τμήμα.
  5. Εάν τελειώνει σε λιγότερο από 5 λεπτά τότε προχώρα στο επόμενο τμήμα.
  6. Στην αντίθετη περίπτωση προχώρα 5 λεπτά και ψάξε ξανά για κάποιο τίτλο ειδήσεων. Εάν βρεθεί μια τέτοια εικόνα πήγαινε πίσω στο βήμα 4.
  7. Πήγαινε στο βήμα 1 έως ότου τελειώσουν όλα τα τμήματα του δελτίου ειδήσεων.
-

## Κεφάλαιο 7: Εξαγωγή των τίτλων ειδήσεων

Στόχος του κεφαλαίου αυτού είναι η αυτόματη εξαγωγή όλων των τίτλων ειδήσεων που εμφανίζονται στα πάνελ αλλά και στα ρεπορτάζ ενός δελτίου ειδήσεων του Mega. Μια τέτοια διαδικασία είναι ιδιαίτερα χρήσιμη καθώς η παρατήρηση των τίτλων ειδήσεων ενός δελτίου μπορεί να οδηγήσει στην κατανόηση του περιεχομένου του.

### 7.1 Εξαγωγή των τίτλων ειδήσεων από τα πάνελ του Mega

Μέχρι στιγμής έχουμε καταφέρει να βρούμε ποιες εικόνες ενός δελτίου περιέχουν κάποιο τίτλο ειδήσεων τύπου πάνελ. Ωστόσο, όπως είναι φυσιολογικό, ένας τίτλος ειδήσεων εμφανίζεται σε πολλές διαδοχικές εικόνες του βίντεο. Πρέπει να βρεθεί ένας τρόπος έτσι ώστε το πρόγραμμά μας να μπορεί να αποθηκεύει έναν τίτλο όταν αυτός πρωτοεμφανίζεται και να αντιλαμβάνεται τότε οι επόμενες εικόνες φέρουν τον ίδιο τίτλο. Για το λόγο αυτό χρησιμοποιείται μια άλλη μετρική, αυτή της ευκλείδειας απόστασης δύο εικόνων. Ουσιαστικά, μετρώντας την ευκλείδεια απόσταση μεταξύ δύο εικόνων ίδιων διαστάσεων ανά εικονοστοιχείο (pixel by pixel) μπορούμε να αποφανθούμε για την ομοιότητά τους.

Μια σημαντική διαφορά όσον αφορά την πορεία της υλοποίησης που ακολουθήθηκε μέχρι τώρα είναι ότι η εικόνα βάσης πρέπει να ανανεώνεται συνεχώς. Αυτό σημαίνει ότι, αν και αρχικά χρησιμοποιείται ως εικόνα βάσης μια εικόνα όπως η 5.4 που πάρθηκε τυχαία, η πρώτη εξεταζόμενη εικόνα που θα βρεθεί να φέρει κάποιο τίτλο ειδήσεων θα γίνει η νέα εικόνα βάσης. Η ανανέωση αυτή προσφέρει ένα βασικό πλεονέκτημα. Εάν στην επόμενη εικόνα συγκριθεί η νέα εικόνα βάσης με την εξεταζόμενη εικόνα και μετρηθεί η ευκλείδεια απόστασή τους, αυτή πρέπει να είναι πολύ μικρή αφού ο τίτλος ειδήσεων είναι λογικά ο ίδιος. Οπότε, όταν περάσουν κάποιες εικόνες και αλλάξει ο τίτλος ειδήσεων, εάν συγκριθεί η νέα εικόνα βάσης με την εξεταζόμενη εικόνα η ευκλείδεια απόστασή τους θα είναι εμφανώς μεγαλύτερη από πριν. Στη συνέχεια αποθηκεύεται η εξεταζόμενη εικόνα, η οποία και αντικαταστεί ξανά την εικόνα βάσης. Με τον τρόπο αυτό είναι δυνατόν να ανιχνευτεί τότε αλλάζει ένας τίτλος ειδήσεων.

Όπως είναι κατανοητό, είναι απαραίτητη η χρήση ανάλογου κατωφλίου ώστε να καθοριστεί για το πρόγραμμα τι σημαίνει μεγάλη και τι μικρή ευκλείδεια απόσταση. Η μέθοδος αυτή, με την οποία μπορούν να αποθηκευτούν όλοι οι τίτλοι ειδήσεων που εναλλάσσονται στα πάνελ ενός δελτίου ειδήσεων του Mega, μπορεί να συνοψιστεί στον παρακάτω αλγόριθμο.

---

#### *Αλγόριθμος 4: Εξαγωγή των τίτλων ειδήσεων από τα πάνελ*

---

1. Κόψε μια εικόνα βάσης από ένα πάνελ του Mega.
  2. Σε κάθε επανάληψη, δηλαδή για κάθε εικόνα του βίντεο, δημιούργησε την εξεταζόμενη εικόνα.
  3. Σύγκρινε το χρωματικό ιστόγραμμα της εικόνας βάσης με αυτό της εξεταζόμενης εικόνας.
  4. Εάν τα ιστογράμματα δεν είναι όμοια προχώρησε στην επόμενη εικόνα. Αν είναι όμοια εξέτασε την ευκλείδεια απόστασή τους.
  5. Εάν η ευκλείδεια απόσταση είναι μικρή προχώρησε στην επόμενη εικόνα. Εάν είναι μεγάλη αποθήκευσε την εξεταζόμενη εικόνα και όρισέ την ως εικόνα βάσης.
  6. Πήγαινε στο βήμα 2 έως ότου τελειώσουν όλες οι εικόνες του βίντεο.
-

## 7.2 Εξαγωγή των τίτλων ειδήσεων από τα ρεπορτάζ του Mega

Ο τρόπος με τον οποίον εξάγουμε τους τίτλους ειδήσεων από τα ρεπορτάζ του Mega διαφέρει κάπως από τον τρόπο όπου εξάγαμε τους τίτλους ειδήσεων από τα πάνελ. Στο κεφάλαιο 7.1 η ύπαρξη μιας εικόνας βάσης αρκούσε και για τη χρωματική σύγκριση και για τη μέτρηση της ευκλείδειας απόστασης. Ωστόσο, στην περίπτωση των ρεπορτάζ υπάρχουν δύο εικόνες βάσης, οπότε από κάθε εικόνα (frame) προκύπτουν δύο εικόνες προς εξέταση. Αυτό συμβαίνει διότι, όπως αναφέρθηκε και στο κεφάλαιο 6, οι τίτλοι ειδήσεων στα ρεπορτάζ του Mega διαφέρουν αρκετά χρωματικά. Ουσιαστικά, η χρήση της πρώτης εικόνας βάσης σε συνδυασμό με τη χρωματική σύγκριση καταστεί σαφές ποιες εικόνες του βίντεο φέρουν κάποιο τίτλο ειδήσεων τύπου ρεπορτάζ, ενώ η χρήση της δεύτερης εικόνας βάσης σε συνδυασμό με τη σύγκριση της ευκλείδειας απόστασης καταστεί σαφές πότε αλλάζει ο τίτλος ειδήσεων.

Καταρχήν, θα χρησιμοποιήσουμε ως πρώτη εικόνα βάσης την εικόνα 6.2.

Οπότε με την ίδια λογική, αν συγκριθεί χρωματικά η πρώτη εικόνα βάσης με τις αντίστοιχες πρώτες εικόνες προς εξέταση, είναι δυνατόν να ανιχνευτούν ποιες εικόνες του βίντεο φέρουν κάποιο τίτλο ειδήσεων τύπου ρεπορτάζ.

Στη συνέχεια, ως δεύτερη εικόνα βάσης χρησιμοποιείται μια εικόνα όπως οι εικόνες 6.1. Η δεύτερη εικόνα βάσης συγκρίνεται με τις δεύτερες εικόνες προς εξέταση. Η σύγκριση πραγματοποιείται με τη μετρική της ευκλείδειας απόστασης, ώστε να γίνεται αντιληπτό πότε αλλάζει ένας τίτλος ειδήσεων. Η δεύτερη εικόνα βάσης πρέπει επίσης να ανανεώνεται όποτε ανιχνεύεται ένας νέος τίτλος ειδήσεων.

Ο παρακάτω αλγόριθμος μπορεί να συνοψίσει τη διαδικασία που ακολουθήσαμε ώστε να εξάγουμε τους τίτλους ειδήσεων από τα ρεπορτάζ του Mega:

---

### *Αλγόριθμος 5: Εξαγωγή των τίτλων ειδήσεων από τα ρεπορτάζ*

---

1. Χρησιμοποίησε ως πρώτη εικόνα βάσης την εικόνα 6.2. Χρησιμοποίησε ως δεύτερη εικόνα βάσης μια εικόνα όπως η 6.1.
  2. Σε κάθε επανάληψη, δηλαδή για κάθε εικόνα του βίντεο, δημιούργησε τις πρώτες και δεύτερες εικόνες προς εξέταση.
  3. Σύγκρινε το χρωματικό ιστόγραμμα της πρώτης εικόνας βάσης με αυτό της πρώτης εικόνας προς εξέταση.
  4. Εάν τα ιστογράμματα δεν είναι όμοια συνέχισε στην επόμενη εικόνα. Αν είναι όμοια υπολόγισε την ευκλείδεια απόσταση μεταξύ της δεύτερης εικόνας βάσης και της δεύτερης εικόνας προς εξέταση.
  5. Εάν η ευκλείδεια απόσταση είναι μικρή συνέχισε στην επόμενη εικόνα. Εάν είναι μεγάλη αποθήκευσε τη δεύτερη εξεταζόμενη εικόνα και όρισέ την ως δεύτερη εικόνα βάσης.
  6. Πήγαινε στο βήμα 2 έως ότου τελειώσουν όλες οι εικόνες του βίντεο.
-



## Κεφάλαιο 8: Πειραματικά αποτελέσματα

### 8.1 Περιγραφή της βάσης δεδομένων

Η βάση δεδομένων που βρίσκεται στην κατοχή μας περιέχει δελτία ειδήσεων των καναλιών Mega, ΣΚΑΙ, NET και ET1. Ωστόσο, εκτός από τα βίντεο, για κάθε δελτίο ειδήσεων υπάρχουν ξεχωριστά δύο ειδών αρχεία. Στο πρώτο αρχείο έχει απομονωθεί ο ήχος. Στο δεύτερο αρχείο, το οποίο μπορεί να επεξεργαστεί από το εργαλείο transcriber [20], υπάρχει αποθηκευμένη χρήσιμη μετά-πληροφορία όσον αφορά τον ήχο. Ειδικότερα, στα αρχεία αυτά επισημαίνονται οι ομιλητές, τα λόγια των ομιλητών και τα σημεία του δελτίου όπου δεν υπάρχει ήχος. Ωστόσο, η εργασία αυτή ενδιαφέρεται μόνο για τα βίντεο και συγκεκριμένα τα βίντεο του καναλιού Mega, καθώς το ενδιαφέρον αυτό επικεντρώνεται στα οπτικά χαρακτηριστικά των δελτίων. Συγκεκριμένα:

- Η υπάρχον βάση δεδομένων περιέχει 14 δελτία ειδήσεων του Mega διάρκειας 47-51 λεπτών που προβλήθηκαν από τις 23 Μαρτίου έως τις 17 Απριλίου του 2007.
- Τα δώδεκα από αυτά λήγουν πριν την έναρξη των αθλητικών ειδήσεων.
- Η ταχύτητα με την οποία εναλλάσσονται οι διαδοχικές εικόνες των βίντεο είναι 25 frames/sec.
- Το μέγεθος της κάθε εικόνας είναι 352 x 288 εικονοστοιχεία (pixels).

## 8.2 Το εργαλείο Anvil

Για την εξαγωγή των τελικών αποτελεσμάτων χρησιμοποιήθηκαν 7 βίντεο του καναλιού Mega από τη βάση δεδομένων. Όσον αφορά την εξαγωγή των τίτλων ειδήσεων από τα πάνελ και την κατάτμηση των βίντεο, χρησιμοποιήθηκε το εργαλείο Anvil [21]. Ουσιαστικά το εργαλείο αυτό είναι ένα video annotation tool, δηλαδή μας βοηθάει να απεικονίσουμε και να αποθηκεύσουμε μετά-πληροφορία για τα δελτία ειδήσεων.

Για τα 7 αυτά δελτία δημιουργήσαμε 7 αρχεία Anvil όπου καταγράψαμε τα τμήματα (shots) των βίντεο και τις λεζάντες των τίτλων ειδήσεων που περνάνε από τα πάνελ του Mega.



**Εικόνα 8.1: Εικόνα από το εργαλείο Anvil**

Η εικόνα 8.5 είναι διαφωτιστική. Αριστερά βλέπουμε την ονομασία των δύο μπαρών που χρησιμοποιούμε. Στην πρώτη μπάρα με ονομασία λεζάντες καταγράφονται όλες οι λεζάντες με τους τίτλους ειδήσεων που περνάνε από τα πάνελ. Για παράδειγμα, στο 45<sup>ο</sup> λεπτό και 28<sup>ο</sup> δευτερόλεπτο εμφανίζεται η 22<sup>η</sup> λεζάντα με κάποιο τίτλο ειδήσεων. Επίσης, την ίδια χρονική στιγμή τελειώνει ένα ρεπορτάζ και αρχίζει ένα τμήμα με ετικέτα πάνελ, όπως μας πληροφορεί η δεύτερη μπάρα. Στο 45<sup>ο</sup> λεπτό και 45<sup>ο</sup> δευτερόλεπτο τελειώνει το τμήμα με ετικέτα πάνελ και αρχίζει ένα ρεπορτάζ. Επίσης ταυτόχρονα εξαφανίζεται η λεζάντα νούμερο 22. Όπως είναι εύκολο να διαπιστωθεί, τα τμήματα με ετικέτα πάνελ παρουσιάζονται με κόκκινο χρώμα ενώ τα ρεπορτάζ με πράσινο χρώμα.

Με τον τρόπο αυτό μπορέσαμε να συγκρίνουμε με ακρίβεια τα αποτελέσματα που εξάγαμε σε σύγκριση με την πραγματικότητα, να κατανοήσουμε δηλαδή αν η ανίχνευση που πραγματοποιήσαμε ανταποκρίνεται στα πραγματικά δεδομένα και σε ποια σημεία υπήρξε σφάλμα.

### 8.3 Αποτελέσματα κατάτμησης, σύνοψης και εξαγωγής λεζαντών

Στα 7 αυτά βίντεο περνάνε συνολικά 195 τίτλοι ειδήσεων από τα πάνελ του Mega. Από αυτούς καταφέραμε να ανιχνεύσουμε τους 184, με ποσοστό επιτυχίας 94.4%. Ωστόσο, ανιχνεύτηκαν επίσης 12 φωτογραφίες που δεν εμφανίζουν κάποιο τίτλο ειδήσεων, οπότε παρουσιάστηκε σφάλμα της τάξης του 6.1%

Παρόμοια έρευνα δεν είχε νόημα να πραγματοποιηθεί για τους τίτλους ειδήσεων των ρεπορτάζ, καθώς, όπως αναφέρθηκε και στο κεφάλαιο 6, εκτός από τους τίτλους ειδήσεων ανιχνεύονται ταυτόχρονα οι λεζάντες με τους ρεπόρτερ και τους κεντρικούς ομιλητές.

Η κατάτμηση των δελτίων ειδήσεων χωρίζει τα 7 αυτά βίντεο σε 93 τμήματα-shots, δηλαδή κατά μέσο όρο λαμβάνουμε 13.3 τμήματα ανά δελτίο ειδήσεων. Δεν υπάρχει αποτυχία, καθώς εξάγονται όλα τα υπάρχοντα πάνελ και ρεπορτάζ.

Τα 93 αυτά τμήματα δίνουν 118 εικόνες-κλειδιά, δηλαδή κατά μέσο όρο λαμβάνουμε 16.9 εικόνες-κλειδιά ανά δελτίο ειδήσεων. Από τις 118 εικόνες, 3 εικόνες που προέρχονται από τμήματα με ετικέτα ρεπορτάζ απέτυχαν να βρουν κάποιο τίτλο ειδήσεων και αποθήκευσαν ως εικόνα-κλειδί την τελευταία εικόνα του τμήματος.

Στη συνέχεια παραθέτονται ορισμένες ενδεικτικές εικόνες, οι οποίες καταστούν σαφή τα αποτελέσματα των κεφαλαίων 6 και 7:



Εικόνες 8.2: Οι εικόνες-κλειδιά του δελτίου που προβλήθηκε στις 27/03/2007:

ΒΡΕΘΗΚΑΝ ΤΑ ΛΕΦΤΑ ΑΠΟ ΤΙΣ ΒΑΛΙΤΣΕΣ

ΕΦΟΔΟΣ ΤΟΥ ΣΔΟΕ ΜΕ ΛΕΙΑ 4 ΕΚΑΤ. ΕΥΡΩ

ΒΡΕΘΗΚΑΝ 4 ΕΚΑΤ. ΕΥΡΩ ΜΕΤΡΗΤΑ ΣΕ ΧΡΗΜΑΤΟΚΙΒΩΤΙΟ

"ΡΑΜΠΟ" ΤΟΥ ΣΔΟΕ ΒΡΗΚΑΝ ΤΗΝ ΠΡΟΜΗΘΕΙΑ ΓΙΑ ΤΟ ΟΜΟΛΟΓΟ

ΕΦΟΔΟΣ ΤΟΥ ΣΔΟΕ ΜΕ ΛΕΙΑ 4 ΕΚΑΤ. ΕΥΡΩ

ΤΟ ΣΧΕΔΙΟ ΤΗΣ ΜΕΓΑΛΗΣ "ΑΡΠΑΧΤΗΣ"

ΚΑΝΟΥΝ "ΦΥΛΛΟ ΚΑΙ ΦΤΕΡΟ" ΒΕΝΖΙΝΑΔΙΚΑ ΣΤΗΝ ΚΡΗΤΗ

ΣΤΑ ΧΕΡΙΑ ΤΟΥ ΕΙΣΑΓΓΕΛΕΑ ΖΟΡΜΠΑ ΠΟΛΥΤΙΜΑ ΣΤΟΙΧΕΙΑ

ΤΙ ΒΡΕΘΗΚΕ ΣΤΗΝ ΕΤΑΙΡΕΙΑ - ΜΥΣΤΗΡΙΟ ΤΗΣ ΝΕΑΣ ΙΩΝΙΑΣ

ΠΑΡΑΠΤΗΣΗ - ΜΥΣΤΗΡΙΟ ΣΤΗΝ ΕΤΑΙΡΕΙΑ - ΓΡΙΦΟ

ΚΑΝΟΥΝ "ΦΥΛΛΟ ΚΑΙ ΦΤΕΡΟ" ΒΕΝΖΙΝΑΔΙΚΑ ΣΤΗΝ ΚΡΗΤΗ

ΠΑΡΑΠΤΗΣΗ - ΜΥΣΤΗΡΙΟ ΣΤΗΝ ΕΤΑΙΡΕΙΑ - ΓΡΙΦΟ

ΤΙ ΒΡΕΘΗΚΕ ΣΤΗΝ ΕΤΑΙΡΕΙΑ - ΜΥΣΤΗΡΙΟ ΤΗΣ ΝΕΑΣ ΙΩΝΙΑΣ

ΤΙ ΒΡΗΚΑΝ ΟΙ "ΡΑΜΠΟ" ΤΟΥ ΣΔΟΕ ΣΤΗΝ ΕΤΑΙΡΕΙΑ - ΓΡΙΦΟ

ΣΑΛΟΣ ΑΠΟ ΤΙΣ ΑΠΟΚΑΛΥΨΕΙΣ ΓΙΑ ΤΑ ΟΜΟΛΟΓΑ

ΠΟΥ ΠΗΓΑΝ ΟΙ ΒΑΛΙΤΣΕΣ ΜΕ ΤΙΣ ΜΙΖΕΣ ;

"ΑΝ ΥΠΑΡΧΕΙ ΠΟΙΝΙΚΗ ΕΥΘΥΝΗ - ΥΠΑΡΧΕΙ ΚΑΙ ΠΟΛΙΤΙΚΗ"

ΠΟΥ ΠΗΓΑΝ ΟΙ ΒΑΛΙΤΣΕΣ ΜΕ ΤΙΣ ΜΙΖΕΣ ;

"ΓΑΛΑΖΙΟ ΦΙΔΑΚΙ" ΜΕ ΤΑ ΟΜΟΛΟΓΑ ΤΩΝ ΤΑΜΕΙΩΝ

ΗΠΑ: Ο "ΑΗ - ΣΤΡΑΤΗΣ" ΕΙΝΑΙ ΕΛΛΗΝΙΚΟΣ

ΗΠΑ: Ο "ΑΗ - ΣΤΡΑΤΗΣ" ΕΙΝΑΙ ΕΛΛΗΝΙΚΟΣ

ΜΥΡΙΖΕΙ ΜΠΑΡΟΥΤΙ Ο ΠΕΡΣΙΚΟΣ ΚΟΛΠΟΣ



Εικόνες 8.3: Η εναλλαγή των τίτλων ειδήσεων στα πάνελ του Mega για το δελτίο που προβλήθηκε στις 27/03/2007:



**ΤΑΠΙΛΤΑΤ-ΚΑΤΑΓΓΕΛΙΕΣ**  
**ΠΟΥΛΗΣΑΝ ΤΑ ΟΜΟΛΟΓΑ ΦΘΗΝΟΤΕΡΑ**  
**ΑΠΟ ΟΤΙ ΤΑ ΑΓΟΡΑΣΑΝ**

**ΦΩΤΗΣ ΔΗΜΗΤΡΟΥΛΑΣ**  
 αντιπρόεδρος ΤΑΠΙΛΤΑΠ

**ΣΥΝΔΙΚΑΛΙΣΤΕΣ-ΤΑΠΙΛΤΑΠ**  
**ΓΙΑ ΠΕΡΙΕΡΓΗ ΑΔΙΑΦΟΡΙΑ**  
**ΚΑΤΑΓΓΕΛΛΟΥΝ ΤΟΝ Σ. ΤΣΙΤΟΥΡΙΔΗ**

**ΒΙΒΛΙΟ ΙΣΤΟΡΙΑΣ**  
 ΛΑΘΗ ΚΑΙ ΠΑΡΑΛΕΙΨΕΙΣ  
 ΕΝΤΟΠΙΖΕΙ Η ΑΚΑΔΗΜΙΑ ΑΘΗΝΩΝ

**ΓΡΗΓΟΡΗΣ ΣΚΑΛΚΕΑΣ**  
 Ακαδημαϊκός

**ΒΙΒΛΙΟ ΙΣΤΟΡΙΑΣ**  
 ΛΑΘΗ ΚΑΙ ΠΑΡΑΛΕΙΨΕΙΣ  
 ΕΝΤΟΠΙΖΕΙ Η ΑΚΑΔΗΜΙΑ ΑΘΗΝΩΝ

**ΒΙΒΛΙΟ ΙΣΤΟΡΙΑΣ**  
 ΕΝΣΤΑΣΕΙΣ ΤΗΣ ΑΚΑΔΗΜΙΑΣ ΓΙΑ  
 ΜΙΚΡΑΣΙΑΤΙΚΗ ΚΑΤΑΣΤΡΟΦΗ ΚΑΙ ΤΟΥΡΚΟΚΡΑΤΙΑ

**ΑΓΟΡΑ - ΠΑΣΧΑ**  
**ΚΕΡΔΟΣΚΟΠΙΚΑ ΠΑΙΧΝΙΔΙΑ ΜΕ ΤΟ ΑΡΝΙ**  
**ΚΑΤΑΓΓΕΛΛΟΥΝ ΚΤΗΝΟΤΡΟΦΟΙ-ΚΑΤΑΝΑΛΩΤΕΣ**

ρεπορτάζ  
**ΓΙΩΡΓΟΣ ΣΑΡΑΝΤΑΚΟΣ**

**ΑΓΟΡΑ - ΠΑΣΧΑ**  
**ΚΕΡΔΟΣΚΟΠΙΚΑ ΠΑΙΧΝΙΔΙΑ ΜΕ ΤΟ ΑΡΝΙ**  
**ΚΑΤΑΓΓΕΛΛΟΥΝ ΚΤΗΝΟΤΡΟΦΟΙ-ΚΑΤΑΝΑΛΩΤΕΣ**

ρεπορτάζ  
**ΓΙΩΡΓΟΣ ΣΑΡΑΝΤΑΚΟΣ**

**ΑΓΟΡΑ - ΠΑΣΧΑ**  
**ΣΤΑ 8 ΕΥΡΩ ΠΩΛΕΙΤΑΙ ΣΗΜΕΡΑ**  
**ΤΟ ΑΡΝΙ ΣΤΗ ΧΟΝΔΡΙΚΗ**

**ΠΕΡΣΙΚΟΣ ΚΟΛΠΟΣ**  
**ΚΡΙΣΗ ΜΕ ΤΗΝ ΤΕΧΕΡΑΝΗ-15 ΒΡΕΤΑΝΟΙ**  
**ΚΟΜΑΝΤΟΣ ΣΤΑ ΧΕΡΙΑ ΤΩΝ ΙΡΑΝΩΝ**

**ΓΕΡΜΑΝΙΑ**  
**ΔΕΚΑΔΕΣ ΟΡΕΙΝΑ ΧΩΡΙΑ**  
**ΑΠΟΚΛΕΙΣΜΕΝΑ ΑΠΟ ΤΟ ΧΙΟΝΙ**



Εικόνες 8.4: Η εναλλαγή των τίτλων ειδήσεων στα ρεπορτάζ του Mega για το δελτίο που προβλήθηκε στις 27/03/2007:



## 8.4 Αποτελέσματα ταχείας ανίχνευσης προσώπων

Κατ' αρχήν, όλα τα βίντεο της βάσης δεδομένων παίζουν με ταχύτητα 25 εικόνες ανά δευτερόλεπτο (25 frames per second). Αυτό σημαίνει ότι σε ένα δευτερόλεπτο περνάνε 25 διαδοχικές εικόνες. Αρχικός στόχος είναι να παίξει το δελτίο ειδήσεων με την κανονική του ταχύτητα. Για το λόγο αυτό γίνεται χρήση της συνάρτησης `CvWaitKey()` της `OpenCV`. Η `CvWaitKey` λαμβάνει ως παράμετρο τον αριθμό των `milliseconds` που πρέπει να περιμένει το πρόγραμμα ώστε να παιχτεί η επόμενη εικόνα (`frame`). Για παράδειγμα, καθώς ξεκινάει ένα βίντεο μέσω της `OpenCV`, η ύπαρξη της εντολής `CvWaitKey(40)` έχει ως αποτέλεσμα το πρόγραμμα να περιμένει 40 `milliseconds` ώστε να ληφθεί η επόμενη εικόνα. Επειδή  $40 \text{ milliseconds} \times 25 = 1 \text{ second}$ , σε ένα δευτερόλεπτο θα έχουν περάσει 25 εικόνες, οπότε το βίντεο παίζει με την κανονική του ταχύτητα.

Στη συνέχεια πραγματοποιείται ανίχνευση προσώπων σε κάθε εικόνα του βίντεο. Δίνεται αρχικά στη `CvWaitKey` η τιμή ένα (`CvWaitKey(1)`) έτσι ώστε το βίντεο να παίζει με τη μέγιστη δυνατή ταχύτητα. Έπειτα, πραγματοποιείται όσο δυνατόν πιο ακριβής ανίχνευση προσπαθώντας να διατηρηθεί η ταχύτητα του βίντεο κανονική, δηλαδή 25 frames/sec ή 40 `milliseconds` επεξεργασίας ανά εικόνα. Παρατηρούμε ότι ενώ ο ανιχνευτής καταφέρνει να βρει μεγάλα πρόσωπα, για παράδειγμα έναν ομιλητή ή 2 ομιλητές σε 2 παράθυρα, αποτυγχάνει να βρει μικρά πρόσωπα όπως 6 ομιλητές σε 6 παράθυρα. Στην περίπτωση που πραγματοποιηθεί πιο αυστηρή ανίχνευση μειώνοντας την παράμετρο που αφορά το μέγεθος του μικρότερου προσώπου που αναζητείται, παρατηρούμε ότι το βίντεο αρχίζει να “κολλάει”, δηλαδή να παίζει με ταχύτητα μικρότερη από την κανονική.

Παρακάτω εμφανίζονται τα αποτελέσματα της προσπάθειας να αυξηθεί η ταχύτητας της ανίχνευσης προσώπων στα δελτία ειδήσεων του `Mega` με τη βοήθεια της ανίχνευσης γραμμών, όπως αυτή παρουσιάστηκε στο τρίτο κεφάλαιο.

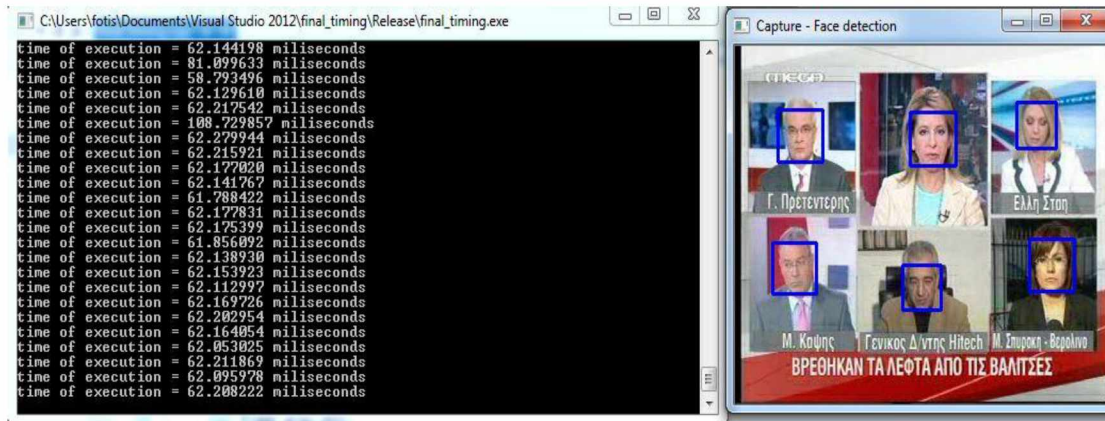
Για την εξαγωγή των αποτελεσμάτων “κόψαμε” με την βοήθεια του `ffmpeg` ορισμένα, συνήθως δίλεπτα, κομμάτια των βίντεο όπου χρησιμοποιείται μόνο ένα πάνελ και υπολογίσαμε τη μέση ταχύτητα και την ακρίβεια της ανίχνευσης.

	Ταχύτητα	Ακρίβεια
Πάνελ 2 ομιλητών	47 ms	99%
Πάνελ 5 ομιλητών	52 ms	94%
Πάνελ 6 ομιλητών	63 ms	98%

**Πίνακας 8.1: Αποτελέσματα επεξεργασίας ανά εικόνα με ιδανικές τιμές 40 milliseconds ταχύτητα, 100% ακρίβεια σε Intel Core i3 2.53 GHz 2.53 GHz:**

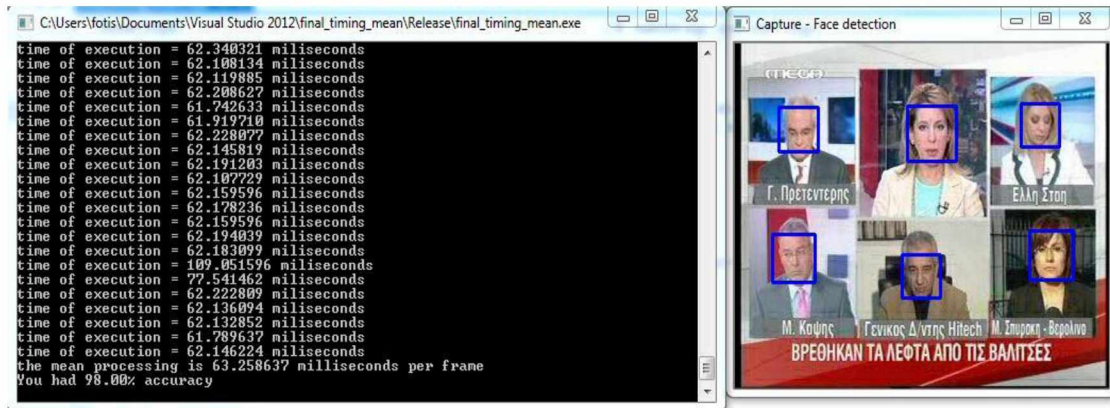
Στην πρώτη και δεύτερη στήλη δίνεται η μέση ταχύτητα και η ακρίβεια που επιτυγχάνεται χρησιμοποιώντας τον αλγόριθμο του κεφαλαίου 3.2. Γενικά, στην περίπτωση όπου πραγματοποιούνταν ανίχνευση προσώπων και όχι γραμμών σε κάθε εικόνα, για να επιτευχθεί τόσο μεγάλη ακρίβεια για μικρά πρόσωπα, δηλαδή για πάνελ με 5 ή 6 ομιλητές, χρειαζόμαστε περίπου **100 ms** ανά εικόνα. Η διαφορά είναι σαφώς τεράστια, ωστόσο και πάλι το βίντεο δεν παίζει σε πραγματικό χρόνο.

Η ανίχνευση γραμμών γενικά δεν αποτυγχάνει. Η πιθανότητα να υπάρχει πάνελ και να μη βρεθεί κάποια από τις χαρακτηριστικές γραμμές είναι 4% στο πάνελ με 5 ομιλητές, ενώ στα άλλα δύο πάνελ είναι λιγότερο από 1%. Το υπόλοιπο σφάλμα στην ακρίβεια οφείλεται στον ανιχνευτή προσώπων.

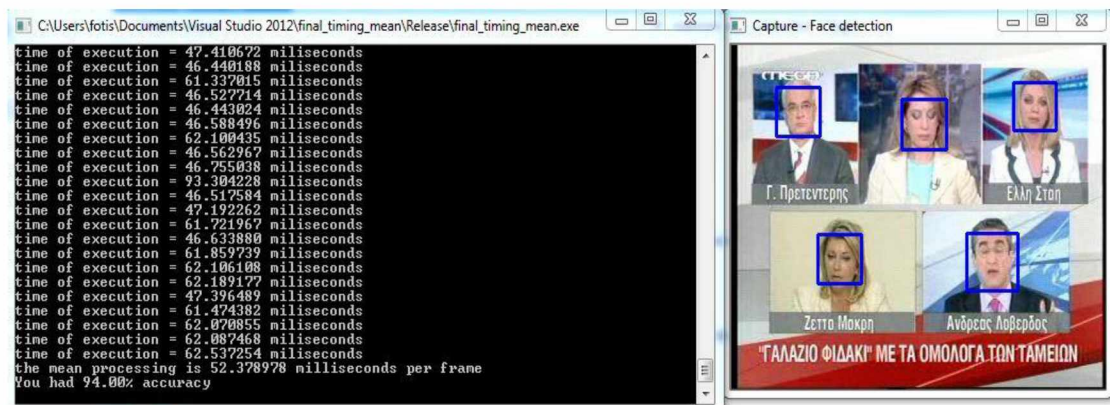


Εικόνα 8.5: Στιγμιότυπο της ταχείας ανίχνευσης προσώπων

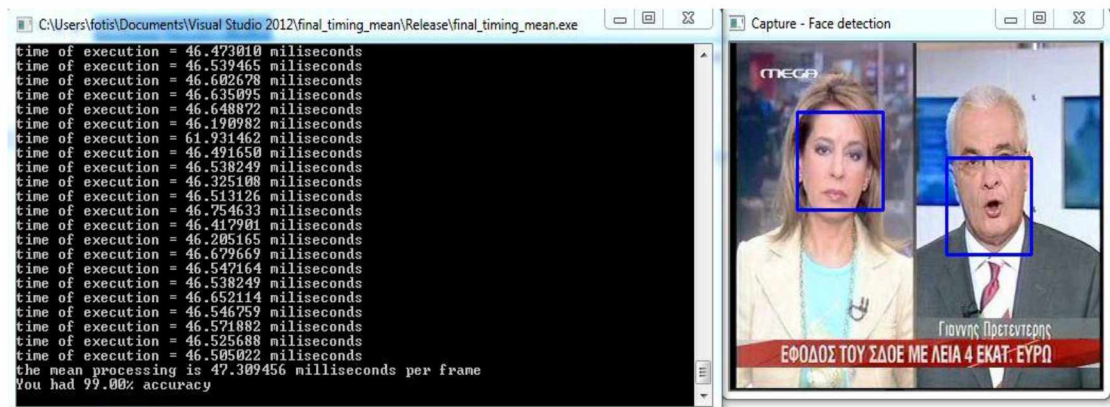
Στην εικόνα 8.1 εμφανίζεται ένα στιγμιότυπο της ταχείας ανίχνευσης προσώπων. Συγκεκριμένα, παρουσιάζεται ο χρόνος επεξεργασίας για 24 συνεχόμενες εικόνες. Επειδή τα πρόσωπα είναι μικρά, δηλαδή ο αριθμός των ομιλητών που εμφανίζονται είναι μεγάλος, πραγματοποιείται ανίχνευση γραμμών σε κάθε εικόνα και ανίχνευση προσώπων ανά 50 εικόνες. Η επεξεργασία της κάθε εικόνας απαιτεί περίπου 62 milliseconds για 21 από τις 24 εικόνες και 58 και 81 milliseconds αντίστοιχα για 2 εικόνες. Το γεγονός αυτό δείχνει ότι υπάρχει μια μικρή αστάθεια όσον αφορά την ταχύτητα της ανίχνευσης γραμμών. Είναι εύκολα αντιληπτό ότι η εικόνα για την οποία ο χρόνος επεξεργασίας είναι 108 milliseconds είναι εικόνα στην οποία έχει γίνει αυστηρή ανίχνευση προσώπων. Τα αποτελέσματα αυτά φυσικά δεν είναι πάγια για όλους τους υπολογιστές, αφού πιο γρήγοροι υπολογιστές μπορούν να δώσουν καλύτερες ταχύτητες. Τέλος, παρακάτω εμφανίζονται τα πειραματικά αποτελέσματα σε μορφή εικόνων, όπου υπολογίστηκε η μέση ταχύτητα και η ακρίβεια του αλγορίθμου του κεφαλαίου 3.2 για κάθε πάνελ.



Εικόνα 8.6: Η μέση ταχύτητα και ακρίβεια για 6 ομιλητές



Εικόνα 8.7: Η μέση ταχύτητα και ακρίβεια για 5 ομιλητές



Εικόνα 8.8: Η μέση ταχύτητα και ακρίβεια για 2 ομιλητές. Από τις 22 εικόνες υπάρχει μία και μοναδική εικόνα που ο χρόνος επεξεργασίας της δεν είναι 46 ms αλλά 62 ms. Στην εικόνα αυτή πραγματοποιείται χαλαρή ανίχνευση προσώπων.

## Κεφάλαιο 9: Συμπεράσματα

### 9.1 Συνεισφορά της διπλωματικής εργασίας

Τα αποτελέσματα της έρευνάς μας είναι ιδιαίτερα ικανοποιητικά διότι βασίστηκαν στην βαθιά παρακολούθηση και σαφή κατανόηση της δομής των δελτίων ειδήσεων του Mega, ενώ παράλληλα είναι σαφή η απουσία περίπλοκων μεθόδων. Ειδικότερα, η έρευνα αυτή μπορεί να επεκταθεί για οποιαδήποτε δελτία ειδήσεων που χρησιμοποιούν στατικής μορφής τηλεπαράθυρα ή εμφανίζουν τίτλους ειδήσεων σε συγκεκριμένα σημεία της οθόνης. Πιο αναλυτικά τα συμπεράσματα της διπλωματικής αυτής μπορούν να συνοψιστούν στα παρακάτω σημεία:

- Η ταχύτητα με την οποία ανιχνεύονται τα μικρά πρόσωπα στα πάνελ του Mega μπορεί να αυξηθεί 2 φορές περισσότερο, με μηδαμινή επίπτωση στην ακρίβεια της ανίχνευσης. Κύρια χαρακτηριστικά της μεθόδου που προτάθηκε είναι η ανίχνευση γραμμών ανά εικόνα και η λιγότερο συχνή ανίχνευση προσώπων, καθώς οι ομιλητές στα τηλεπαράθυρα δε μετακινούνται ιδιαίτερα.
- Μέσω της ανίχνευσης γραμμών, δηλαδή τηλεπαραθύρων, σε συνδυασμό με την ανίχνευση προσώπων είναι δυνατή η κατάτμηση των ειδήσεων του Mega σε πάνελ και ρεπορτάζ.
- Μέσω της σύγκρισης χρωματικών ιστογραμμάτων και την έξυπνη χρήση του γεγονότος ότι οι τίτλοι ειδήσεων στα δελτία του Mega εμφανίζονται σε συγκεκριμένη θέση στην οθόνη, τα δελτία ειδήσεων μπορούν να κατατμηθούν σε πάνελ και ρεπορτάζ εύκολα, γρήγορα και αποτελεσματικά.
- Με παρόμοια λογική τα δελτία ειδήσεων συνοψίζονται σε εικόνες-κλειδιά.
- Κάνοντας χρήση χρωματικών ιστογραμμάτων και χρησιμοποιώντας τη μετρική της ευκλείδειας απόστασης δύο εικόνων είναι δυνατή η εξαγωγή όλων των τίτλων ειδήσεων που εναλλάσσονται στα τηλεπαράθυρα και τα ρεπορτάζ του Mega.

## 9.2 Μελλοντικές ερευνητικές κατευθύνσεις

Καθώς παρουσιάζονται οι ειδήσεις των δελτίων ειδήσεων του καναλιού Mega παρατηρείται ότι υπάρχει μια εναλλαγή μεταξύ τμημάτων όπου είτε πραγματοποιείται συζήτηση σε τηλεπαράθυρα είτε παρουσιάζεται κάποιο ρεπορτάζ. Η παρούσα διπλωματική κατάφερε να πετύχει μια τέτοιου είδους κατάτμηση. Ωστόσο, τα τμήματα που χαρακτηρίζονται ως πάνελ περιορίζονται σε ένα θέμα συζήτησης μέχρι την έναρξη κάποιου ρεπορτάζ. Αντίθετα, στα τμήματα που χαρακτηρίζονται ως ρεπορτάζ αποτελεί συχνό φαινόμενο η παρουσίαση πολλών τίτλων ειδήσεων. Συγκεκριμένα, μετά το τέλος μιας συζήτησης στα τηλεπαράθυρα η κεντρική παρουσιάστρια εισάγει το θεατή στο ρεπορτάζ που θα παρουσιαστεί στη συνέχεια. Πολλές φορές ωστόσο τα θέματα που καλύπτονται μπορεί να είναι περισσότερα από ένα. Ορισμένες φορές, μετά το τέλος ενός ρεπορτάζ, είναι πιθανόν να μην ξεκινήσει ένα πάνελ αλλά η κεντρική παρουσιάστρια να εισάγει τον θεατή σε ένα καινούργιο ρεπορτάζ. Αυτό συμβαίνει συνήθως όταν παρουσιάζονται πολλά γεγονότα χωρίς μεγάλο όμως ειδησεογραφικό ενδιαφέρον. Για το λόγο αυτό μπορεί να πραγματοποιηθεί περαιτέρω κατάτμηση των ρεπορτάζ με τη χρήση κάποιας anchor detection τεχνικής.

Τέλος, η εργασία αυτή βασίστηκε κυρίως σε οπτικά χαρακτηριστικά (χρώματα, λεζάντες ειδήσεων). Ο συνδυασμός οπτικών χαρακτηριστικών με διάφορα άλλα χαρακτηριστικά, όπως για παράδειγμα χαρακτηριστικά ήχου, παρουσιάζει ιδιαίτερο ερευνητικό ενδιαφέρον και μπορεί να εφαρμοστεί με σκοπό τη λύση πολλών προβλημάτων κατά την επεξεργασία δελτίων ειδήσεων (κατάτμηση, αυτόματη αναγνώριση ομιλίας κ.α.).



## Βιβλιογραφία

- [1] D. Dimitriadis, A. Metallinou, I. Konstantinou, G. Goumas, P. Maragos and N. Koziris, "GRIDNEWS: A Distributed Automatic Greek Broadcast Transcriptions System", *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, 2009
- [2] M. De Santo, P. Foggia, G. Percannella, C. Sansone and M. Vento, "An Un-supervised Algorithm for Anchor Shot Detection", *18<sup>th</sup> International Conference on Pattern Recognition*, Volume 2, Pages 1238-1241, 2006.
- [3] G. J. Poulisse and M. F. Moens, "Multimodal News Story Segmentation", *Proceedings of the First International Conference on Intelligent Human Computer Interaction*, Pages 95-101, 2009.
- [4] L. Chaisorn, T. S. Chua, C. K. Koh, Y. Zhao, H. Xu, H. Feng and Q. Tian, "A Two-Level Multi-Modal Approach for Story Segmentation of Large News Video Corpus", *TRECVID Workshop*, 2003.
- [5] W. Zhu, C. Toklu and S. P. Liou, "Automatic News Video Segmentation and Categorization Based on Closed-Captioned Text", *Proceedings of the 2001 IEEE International Conference on Multimedia and Expo*, Pages 829-832, 2001.
- [6] Y. Zhai, A. Yilmaz and M. Shah, "Story Segmentation in News Videos Using Visual and Text Cues", *Proceedings of the 4<sup>th</sup> International Conference on Image and Video Retrieval*, Pages 92-102, 2005.
- [7] A. Goyal, P. Punitha, F. Hopfgartner and J. M. Jose, "Split and Merge Based Story Segmentation in News Videos", *Proceedings of the 31<sup>th</sup> European Conference on IR Research on Advances in Information Retrieval*, Pages 766-770, 2009.
- [8] H. Misra, F. Hopfgartner, A. Goyal, P. Punitha, and J. M. Jose, "TV News Story Segmentation based on Semantic Coherence and Content Similarity", *Proceedings of the 16<sup>th</sup> International Conference on Advances in Multimedia Modeling*, Pages 347-357, 2010.



[9] H. Jiang, T. Lin and H. Zhang, "Video Segmentation with the Support of Audio Segmentation and Classification", *Proceedings of the 2000 IEEE International Conference on Multimedia and Expo*, 2000.

[10] C. Wang, Y. Wang, H. Y. Liu and Y. X. He, "Automatic Story Segmentation of News Video Based on Audio-Visual Features and Text Information", *International Conference of Machine Learning and Cybernetics*, Volume 5, Pages 3008-3011, 2003.

[11] X. Wang, L. Xie, M. Lu, B. Ma, E. S. Chng and H. Li, "Broadcast News Story Segmentation Using Conditional Random Fields and Multimodal Features", *Special Section on Recent Advances in Multimedia Signal Processing Techniques and Applications*, 2012.

[12] B. Han, X. Gao and H. Ji, "A Shot Boundary Detection Method for News Video Based on Rough-Fuzzy Sets", *Proceedings of Image Analysis and Recognition, Second International Conference*, Pages 231-238, 2005.

[13] Βιβλιοθήκη OpenCV - Ανίχνευση γραμμών. [Online]. Available: [http://docs.opencv.org/doc/tutorials/imgproc/imgtrans/hough\\_lines/hough\\_lines.html](http://docs.opencv.org/doc/tutorials/imgproc/imgtrans/hough_lines/hough_lines.html)

[14] Βιβλιοθήκη OpenCV - Ανίχνευση ακμών. [Online]. Available: [http://docs.opencv.org/doc/tutorials/imgproc/imgtrans/canny\\_detector/canny\\_detector.html](http://docs.opencv.org/doc/tutorials/imgproc/imgtrans/canny_detector/canny_detector.html)

[15] Ανίχνευση γραμμών - Αλγόριθμος LSWMS. [Online]. Available: <http://marcosnietoblog.wordpress.com/2012/04/28/line-segment-detection-opencv-c-source-code>

[16] Ανίχνευση γραμμών - Αλγόριθμος LSWMS. [Online]. Available: <http://marcosnietoblog.wordpress.com/2013/12/25/faster-line-segment-detection-opencv-c-source-code>

[17] Βιβλιοθήκη OpenCV - Ανίχνευση προσώπων. [Online]. Available: [http://docs.opencv.org/doc/tutorials/objdetect/cascade\\_classifier/cascade\\_classifier.html#cascade-classifier](http://docs.opencv.org/doc/tutorials/objdetect/cascade_classifier/cascade_classifier.html#cascade-classifier)

[18] Ταχύτητα ανίχνευσης προσώπων και γραμμών. [Online]. Available: <http://www.shervinemami.info/timingTests.html>

[19] J. S. Boreczky and L. A. Rowe, "Comparison of video shot boundary detection techniques", *Journal of Electronic Imaging* 5, Issue 2, Pages 122-128, 1996.

[20] Transcriber. [Online]. Available: <http://trans.sourceforge.net/en/presentation.php>

[21] Anvil. [Online]. Available: <http://www.anvil-software.org>