



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΕΠΙΚΟΙΝΩΝΙΩΝ, ΗΛΕΚΤΡΟΝΙΚΗΣ ΚΑΙ ΣΥΣΤΗΜΑΤΩΝ
ΠΛΗΡΟΦΟΡΙΚΗΣ

Επεξεργασία τηλεοπτικού σήματος για ταυτοποίηση διαφημιστικών μηνυμάτων

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Βασίλειος Ι. Βραχνάκης

Επιβλέπων : Βασίλειος Λούμος

Καθηγητής Ε.Μ.Π.

Αθήνα, Νοέμβριος 2010



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΕΠΙΚΟΙΝΩΝΙΩΝ, ΗΛΕΚΤΡΟΝΙΚΗΣ ΚΑΙ ΣΥΣΤΗΜΑΤΩΝ
ΠΛΗΡΟΦΟΡΙΚΗΣ

Επεξεργασία τηλεοπτικού σήματος για ταυτοποίηση διαφημιστικών μηνυμάτων

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Βασίλειος Ι. Βραχνάκης

Επιβλέπων : Βασίλειος Λούμος

Καθηγητής Ε.Μ.Π.

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 10^η Νοεμβρίου 2010.

.....

Β. Λούμος

Καθηγητής Ε.Μ.Π.

.....

Ε. Καγιάφας

Καθηγητής Ε.Μ.Π.

.....

Χ. – Ν. Αναγνωστόπουλος

Επίκουρος Καθηγητής Παν. Αιγαίου

Αθήνα, Νοέμβριος 2010

.....
Βασίλειος Ι. Βραχνάκης

Διπλωματούχος Ηλεκτρολόγος Μηχανικός και Μηχανικός Υπολογιστών Ε.Μ.Π.

Copyright © Βασίλειος Ι. Βραχνάκης, 2010

Με επιφύλαξη παντός δικαιώματος. All rights reserved.

Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας, εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα. Ερωτήματα που αφορούν τη χρήση της εργασίας για κερδοσκοπικό σκοπό πρέπει να απευθύνονται προς τον συγγραφέα.

Οι απόψεις και τα συμπεράσματα που περιέχονται σε αυτό το έγγραφο εκφράζουν τον συγγραφέα και δεν πρέπει να ερμηνευθεί ότι αντιπροσωπεύουν τις επίσημες θέσεις του Εθνικού Μετσόβιου Πολυτεχνείου.

Περίληψη

Στόχος της παρούσας εργασίας είναι η ανάλυση και υλοποίηση ενός αυτοματοποιημένου συστήματος για ταύτιση διαφημιστικών εικονοσειρών τηλεοπτικού σήματος με τη χρήση ευφυών αλγορίθμων κατάτμησης εικόνων. Οι διαφημιστικές εταιρίες πληρώνουν μεγάλα χρηματικά ποσά για την προβολή των διαφημιστικών μηνυμάτων τους στην τηλεόραση, ένα μέσο μετάδοσης με μεγάλη επίδραση στο κοινό. Όμως, ο έλεγχος για την ομαλή λειτουργία χειρωνακτικά είναι μια δαπανηρή, χρονοβόρα και με μεγάλη πιθανότητα σφάλματος διαδικασία. Επιτακτική ανάγκη, λοιπόν, είναι η δημιουργία ενός αυτόματου τρόπου ταύτισης των εικονοσειρών με τη βοήθεια υπολογιστή. Στη διπλωματική εργασία αυτή αρχικά διερευνάται διεξοδικά η σχετική βιβλιογραφία στο αντικείμενο και οι διάφορες προσεγγίσεις του προβλήματος. Έπειτα, παρουσιάζονται τα βασικά σημεία του αναγκαίου θεωρητικού υποβάθρου επάνω στην επεξεργασία εικόνας καθώς και των αλγορίθμων που χρησιμοποιήθηκαν. Ο αλγόριθμος Επαγωγικής Σάρωσης καθώς και ο αλγόριθμος Ταύτισης Εικονοσειρών μελετώνται διεξοδικά. Στη συνέχεια αναλύεται το σύστημα που υλοποιήθηκε για την εφαρμογή των ανωτέρω. Στα επόμενα κεφάλαια, παρουσιάζονται παραδείγματα λειτουργίας καθώς και τα πειραματικά αποτελέσματα για μεγάλο εύρος μετρήσεων σε σημαντικό όγκο δεδομένων. Τα αποτελέσματα αυτά επιβεβαίωσαν την αποδοτικότητα των αλγορίθμων. Τέλος, γίνεται μια επισκόπηση της εργασίας και των αποτελεσμάτων που καθιστούν το σύστημα ταχύ και αποτελεσματικό, ενώ παρατίθενται και οι κατευθύνσεις για μελλοντικές επεκτάσεις.

Λέξεις Κλειδιά : << επεξεργασία εικόνας, αλγόριθμος Επαγωγικής Σάρωσης, ταύτιση εικονοσειρών διαφημίσεων, τηλεοπτικό σήμα, ψηφιακή επεξεργασία εικονοσειράς, σύγκριση εικόνων, εύρεση ακμών >>

Abstract

The purpose of this thesis is the analysis and implementation of an automatic system for video matching of television advertisement spots using intelligent algorithms for dividing an image. Advertising companies pay large sums of money to show their ads on television, a media with large impact on the public. However, checking normal function manually is a costly, time-consuming and with high error probability procedure. Thus there is a compelling need to create an automated system for matching video streams using computers. In this thesis, initially there is a thorough research on related work on the subject and on different approaches on the problem. Afterwards, the key points of the needed theoretical background on image processing are presented as well as the algorithms used. The Operator Context Scanning algorithm and the Video Matching algorithm are studied in detail. Next, the system created for applying these algorithms is examined. On the following chapters, examples of use are presented as well as the experimental results for a wide range of tests on a big volume of data. These results confirmed the efficiency of the algorithms used. Finally, there is a summary of the thesis and of the results that make the system fast and effective. The possibilities for future work are also presented.

Keywords : << digital image processing, video processing, operator context scanning algorithm, video matching algorithm, television signal, compare images, edge detection >>

Ευχαριστίες

Για την εκπόνηση αυτής της διπλωματικής ήταν σημαντική η συμβολή του υποψήφιου διδάκτορα Ιωάννη Γιαννούκου τον οποίο ευχαριστώ θερμά για όλη την ενασχόληση και βοήθεια που μου προσέφερε. Επίσης θα ήθελα να ευχαριστήσω την οικογένεια μου και τους φίλους μου και όλους όσοι με βοήθησαν και με στήριξαν στην προσπάθεια αυτή.

“I fade out again...”

Thom Yorke

ΠΕΡΙΕΧΟΜΕΝΑ

Ευχαριστίες	7
1. Εισαγωγή	11
2. Σχετική Βιβλιογραφία.....	15
3. Μέθοδος.....	19
3.1. Αλγόριθμος Επαγωγικής Σάρωσης κατά μπλοκ	19
3.2. Αλγόριθμος Ταύτισης Εικονοσειρών	26
4. Σύστημα.....	31
4.1. Τεχνικά Χαρακτηριστικά Συστήματος	31
4.2. Βάση Δεδομένων	32
4.3. Επεξεργασία Εικόνας	35
4.4. Σύστημα ταυτοποίησης εικονοσειρών τηλεοπτικού σήματος	37
4.5. Περιγραφή κώδικα του προγράμματος.....	39
5. Παράδειγμα Λειτουργίας Συστήματος	45
5.1. Παράδειγμα τυπικής λειτουργίας	45
5.2 Παράδειγμα παθολογικής λειτουργίας	50
6. Πειραματικά αποτελέσματα	57
7. Επίλογος	67
Αναφορές	68

1. Εισαγωγή

Σε αυτή τη διπλωματική εργασία εξετάζεται η ταύτιση εικονοσειρών τηλεοπτικού σήματος με τη χρήση ευφώνων αλγορίθμων κατάτμησης εικόνων. Τα τελευταία χρόνια, με την πρόοδο της τεχνολογίας έχουν γίνει μεγάλα βήματα ανάπτυξης στον τομέα ανάλυσης και μετάδοσης της οπτικοακουστικής πληροφορίας. Ο διαρκώς αυξανόμενος όγκος οπτικοακουστικού υλικού, απαιτεί συνεχώς εργαλεία τα οποία χρησιμοποιούν νέες μεθόδους και τεχνικές, για γρήγορη και αποτελεσματική επεξεργασία του. Παράλληλα, αντίστοιχη εξέλιξη ακολουθεί και ο κλάδος της τηλεόρασης με το ψηφιακό ή υψηλής ευκρίνειας σήμα (high definition television - HDTV), να κερδίζει ολοένα και περισσότερο έδαφος έναντι παλαιότερων προδιαγραφών. Τα διαφημιστικά μηνύματα είναι αναπόσπαστο κομμάτι της τηλεόρασης, όπως τη γνωρίζουμε, και οι νέες εξελίξεις αυτές δε θα μπορούσαν να το αφήσουν ανεπηρέαστο. Στην τομή των δύο περιοχών αυτών, βρίσκεται το πρόβλημα το οποίο αναφέρεται στο υψηλότατο κόστος των μηνυμάτων αυτών για κάθε δευτερόλεπτο προβολής τους. Μάλιστα, σε συγκεκριμένες περιπτώσεις υψηλής τηλεθέασης αυτό μπορεί να αγγίξει τις τάξεις των εκατοντάδων χιλιάδων ευρώ. Η επαλήθευση για την τήρηση των αντίστοιχων συμβολαίων όμως είναι αρκετά χρονοβόρα και δύσκολη διαδικασία. Η παρούσα εργασία πραγματεύεται μια αυτοματοποιημένη λύση του προβλήματος αυτού με τη βοήθεια των υπολογιστών, σε αντίθεση με τις μέχρι τώρα λύσεις που βασίζονταν στον ανθρώπινο παράγοντα. Σκοπός της είναι η μελέτη τεχνικών που θα μπορούσαν να δώσουν μια ταχεία και αποτελεσματική λύση.

Η δομή της διπλωματικής εργασίας είναι η ακόλουθη. Στο Κεφάλαιο 2 γίνεται μια γενική θεωρητική ανάλυση του προβλήματος. Στο Κεφάλαιο 3 παρουσιάζεται η σχετική βιβλιογραφία που έχει δημοσιευτεί γύρω από το θέμα αυτό. Στο Κεφάλαιο 4 εξηγείται το θεωρητικό κομμάτι των αλγορίθμων που χρησιμοποιήθηκαν στην εργασία αυτή, ενώ στο Κεφάλαιο 5 περιγράφεται το σύστημα πάνω στο οποίο υλοποιήθηκαν οι προαναφερθείς αλγόριθμοι. Στο Κεφάλαιο 6 παρουσιάζονται παραδείγματα λειτουργίας και στο Κεφάλαιο 7 γίνεται η έκθεση των πειραματικών αποτελεσμάτων. Τέλος, το Κεφάλαιο 8 αποτελεί τον επίλογο της διπλωματικής εργασίας με τη σύνοψη και τα συμπεράσματά της.

Η βιομηχανία ανάπτυξης και προώθησης των καταναλωτικών προϊόντων, έχει ωθήσει τον τομέα των διαφημίσεων σε μεγάλη ανάπτυξη. Εξέχουσα θέση στον τομέα αυτό ως μέσο μετάδοσης έχει η τηλεόραση. Η τηλεόραση αποτελεί ένα μέσο που είναι εξαιρετικά διαδεδομένο σε όλο τον κόσμο με εύκολη, άμεση και βαθειά επίδραση στο τηλεοπτικό κοινό. Οι διαφημιστικές εταιρίες εκμεταλλεύονται αυτό το γεγονός πληρώνοντας αδρά ποσά για μερικά δευτερόλεπτα προβολής. Με την πρόοδο της τεχνολογίας όμως, έχουν βρεθεί διάφορες τεχνικές και συστήματα για την αποκοπή των διαφημιστικών μηνυμάτων από την κανονική ροή του προγράμματος προς διευκόλυνση των τηλεθεατών. Αυτό όπως είναι φυσικό, προκαλεί μεγάλο πρόβλημα στις εταιρίες, οι οποίες έχουν πληρώσει ακριβά τον τηλεοπτικό χρόνο προβολής, χωρίς το επιθυμητό αποτέλεσμα. Ως εκ τούτου, η επίδραση της τεχνολογικής προόδου στο συγκεκριμένο ζήτημα μπορεί να συνοψιστεί σε 2 κύριες κατηγορίες όπου εφαρμόζονται ευφυείς τεχνικές επεξεργασίας σήματος. Τους αλγορίθμους εύρεσης και αφαίρεσης των διαφημίσεων (commercial killers) και τους αλγορίθμους παρακολούθησης διαφημίσεων (commercial trackers) [Satterwhite και Marques, 2004].

Στην πρώτη κατηγορία εντάσσονται συστήματα που εντοπίζουν τις περιοχές μιας εικονοσειράς (video) όπου υπάρχουν διαφημίσεις και τις απομακρύνουν κατά τη ζωντανή μετάδοση ή κατά την καταγραφή του σήματος σε προσωπικό σύστημα εγγραφής βίντεο (Personal video recorder – PVR), διευκολύνοντας τους θεατές. Στη δεύτερη κατηγορία υπάγονται οι τεχνικές ανάλυσης και ταυτοποίησης συγκεκριμένων διαφημίσεων στο τηλεοπτικό σήμα για την επαλήθευση της τήρησης των αντίστοιχων συμβολαίων από τις εταιρίες τηλεοπτικών καναλιών. Μέχρι πρότινος αυτή η επαλήθευση γινόταν είτε από το ανθρώπινο δυναμικό των ιδίων των εταιριών είτε προσλαμβάνοντας κάποια άλλη εταιρία για το σκοπό αυτό. Σε κάθε περίπτωση, η χειρωνακτική καταγραφή και επαλήθευση των διαφημιστικών μηνυμάτων είναι μια χρονοβόρα διαδικασία, με μεγάλη πιθανότητα σφάλματος, καθώς και αρκετά δαπανηρή.

Υπάρχει λοιπόν η ανάγκη για ένα αυτοματοποιημένο σύστημα το οποίο μέσω υπολογιστή, θα διεκπεραιώνει αυτή τη διαδικασία ανάλυσης και ταυτοποίησης της εικονοσειράς διαφημιστικών μηνυμάτων. Το σύστημα αυτό θα μπορούσε να παρακολουθεί τη ζωντανή τηλεοπτική ροή πολλών καναλιών ταυτόχρονα. Παράλληλα, θα επεξεργαζόταν το σήμα των καναλιών και θα μπορούσε να κάνει την ταύτιση των εικονοσειρών με τη βοήθεια μιας βάσης δεδομένων. Διάφορα στατιστικά στοιχεία όπως χρόνος και διάρκεια εκπομπής των μηνυμάτων, καθώς και το πλήθος εκπομπών τους ανά κανάλι ή συνολικά,

θα ενημερώνονταν σχεδόν σε πραγματικό χρόνο, άμεσα και αυτόματα. Μια τέτοια υλοποίηση επιτυγχάνει μια αποτελεσματική, οικονομική αλλά και αποδοτική λύση του προβλήματος. Μάλιστα, δεν περιορίζεται εκεί, αλλά μπορεί να επεκταθεί μέσω της χρήσης του από διάφορα εργαλεία για ανάλυση αγοράς (market analysis), τα οποία προϋποθέτουν αυτόματα και γρήγορη επίβλεψη διαφορετικών μέσων μετάδοσης.

Το σύστημα που υλοποιήθηκε σε αυτή τη διπλωματική κινείται σε αυτή τη κατεύθυνση. Στοχεύει στη δυνατότητα δημιουργίας εύκολα ενός τέτοιου συστήματος, το οποίο να χρησιμοποιεί χαμηλού κόστους υλικό αλλά να διατηρεί την υψηλή ταχύτητα και αποτελεσματικότητα. Επίσης, βασίζεται σε αλγορίθμους για την ανάλυση της εικόνας και την εξαγωγή πληροφορίας από αυτή, καθώς και σε αλγορίθμους επεξεργασίας των καρτέ των εικονοσειρών που θα εξασφαλίζουν την ορθή λειτουργία του ακόμα και σε παθολογικές περιπτώσεις. Αυτό οφείλεται σε διάφορες δυσκολίες που εισάγουν τα κανάλια κατά την προβολή των διαφημίσεων. Πιο συγκεκριμένα, διαφορετικά κανάλια παρουσιάζουν με διαφορετικό τρόπο τις ίδιες διαφημίσεις, εισάγοντας ποικίλα μεταβατικά εφέ (transition effects). Σε αυτά περιλαμβάνονται η σταδιακή εισαγωγή των διαφημιστικών καρτέ (fade in), η σταδιακή μείωση της έντασης της φωτεινότητας των διαφημιστικών (fade out), η περιστροφή (Rotation), μετακίνηση (Translation) και κλιμάκωση (Scaling) των διαφημιστικών, κ.α. Επίσης, τροποποιούν τα χρώματα των διαφημιστικών, αλλάζουν τη θέση του λογότυπου του καναλιού και εισάγονται μηδέν ή περισσότερα μαύρα πλαίσια στο τέλος ή την αρχή των διαφημιστικών. Πέρα από την τροποποίηση των διαφημιστικών για λόγους ομοιομορφίας, οι επεμβάσεις στις εικονοσειρές των διαφημίσεων γίνονται για να καταστήσουν τα συστήματα απομάκρυνσης διαφημιστικών (commercial killers) αναποτελεσματικά. Στα πλαίσια αυτών των δυσκολιών οι μέθοδοι επεξεργασίας που εφαρμόζονται οφείλουν να μένουν ανεπηρέαστοι.

Στις επόμενες ενότητες θα εξεταστούν αναλυτικά οι μέθοδοι και η υλοποίηση που εφαρμόστηκε στην εργασία αυτή, αφού συζητηθούν πρώτα και διάφορες άλλες εργασίες που αγγίζουν το θέμα αυτό από την ίδια ή άλλες σκοπιές.

2. Σχετική Βιβλιογραφία

Στον τομέα της αναγνώρισης εικονοσειρών (video), αλλά και ειδικότερα διαφημιστικών μηνυμάτων περιλαμβανομένων σε αυτές, υπάρχει αρκετή έρευνα και βιβλιογραφία. Έχουν αναπτυχθεί διάφορες τεχνικές που μπορούν να ταξινομηθούν ανάλογα με τον τρόπο και τα στοιχεία που χρησιμοποιούν για τη σύγκριση των καρτέ. Μια απλή τεχνική χρησιμοποιεί το γεγονός της ύπαρξης μαύρων καρτέ ή μεταβατικών εφέ από τη μία διαφήμιση στην επόμενη [Sadlier et al., 2002; Gauch και Shivadas, 2006]. Τέτοιες απότομες αλλαγές στο φωτισμό ή στο χρώμα, καθώς και μεταβατικά εφέ όπως η βαθμιαία εισαγωγή/απομάκρυνση εικόνας (fade in/ fade out), υποδηλώνουν τα όρια για τις εκάστοτε εικονοσειρές. Στην ίδια φιλοσοφία κινούνται και εργασίες που ερευνούν την ύπαρξη ή μη του λογότυπου (logo) στις εικονοσειρές [Schoffmann et al., 2009; Albiol et al., 2004]. Εκμεταλλεύονται το γεγονός πως το λογότυπο εμφανίζεται πάντα στην ίδια θέση της εικόνας κατά την κανονική ροή του προγράμματος και εξαφανίζεται κατά την προβολή των διαφημιστικών μηνυμάτων. Παρόλα αυτά, τα παραπάνω δεν αποτελούν ολοκληρωμένη λύση του προβλήματος, αφού μπορεί εύκολα κάποιο κανάλι να τροποποιήσει τον τρόπο που γίνονται αυτές οι μεταβάσεις στις διαφημίσεις. Έτσι, μπορούν να καταστήσουν τα συστήματα απομάκρυνσης διαφημιστικών σποτ αναποτελεσματικά.

Μια άλλη προσέγγιση περιλαμβάνει την αξιοποίηση οπτικής πληροφορίας όπως το χρώμα ή τη φωτεινότητα, που περιέχουν τα καρτέ μιας εικονοσειράς. Σε αυτή τη κατηγορία εντάσσονται οι εργασίες [Nagasaka και Miyatake, 2000; Ostermann et al., 2004], όπου η πληροφορία εξάγεται κάθε φορά από τα χρωματικά χαρακτηριστικά ή τις χρωματικές ψηφιακές υπογραφές [Lu et al., 2001]. Με την ίδια λογική στην [Cheung et al., 2003], η χρωματική πληροφορία είναι τα ιστογράμματα χρώματος και η σύγκριση γίνεται μέσω του υπολογισμού της απόστασης μεταξύ τους στο χρωματικό μοντέλο HSV. Η απόσταση για τα βασικά καρτέ (key frames) της έντασης της φωτεινότητας και των κενών περιθωρίων της εικόνας (letterboxing) είναι τα χαρακτηριστικά διαχωρισμού και εντοπισμού των διαφημίσεων στην [Agnihotri et al., 2003], η οποία χρησιμοποιεί γενετικούς αλγορίθμους (genetic algorithms). Εκτός όμως από τις εικόνες, μπορεί να γίνει αξιοποίηση της πληροφορίας που απορρέει από κείμενα (textual information). Η εξαγωγή αυτής της πληροφορίας γίνεται από τους υποτίτλους [Huang et al., 2007], ή από τα κανάλια (tracks) χαρακτήρων [Dimitrova et al. 2000; Meng et al., 2009], είτε εξάγοντας σημασιολογική

πληροφορία από τις εικόνες [Duan et al., 2008; Katsiouli et al., 2007; Colombo et al., 2001].

Εκτός όμως από την επεξεργασία ολόκληρης της εικόνας, υπάρχουν τεχνικές που βασίζονται στην κατάτμηση των καρέ σε μπλοκ έτσι ώστε να εξαχθεί χωρική πληροφορία. Στην [Nijmeijer, 2008] γίνεται με τη χρήση του αναλλοίωτου στην κλιμάκωση μετασχηματισμού χαρακτηριστικών (Scale Invariant Feature Transform-SIFT), ενώ στη [Duan et al., 2008] υπολογίζεται το ιστόγραμμα του κάθε μπλοκ. Επίσης, στην εργασία [Hamrapur et al., 2001] προτείνεται η δημιουργία ψηφιακής υπογραφής μιας εικονοσειράς, υπολογίζοντας το ενδιάμεσο σημείο των εντάσεων της παραγώγου της εικόνας μαζί με το επικρατές χρώμα (dominant color), ενώ στην [Oostveen et al., 2002] η υπογραφή προκύπτει από τις διαφορές στην ένταση των μπλοκ.

Σε διαφορετικό πλαίσιο κινούνται οι εργασίες που επικεντρώνονται στην ηχητική πληροφορία των διαφημιστικών μηνυμάτων. Εκμεταλλεύονται το γεγονός ότι κατά τη διάρκεια των διαφημιστικών μηνυμάτων υπάρχει συχνότερη αλλά και εντονότερη μουσική παρασκηνίου, στοιχείο που διαφοροποιεί τα φασματικά χαρακτηριστικά τους. Τέτοια ταξινόμηση με βάση το αν ο ήχος είναι αρμονικός ή όχι, δηλαδή αν περιέχει μουσική ή ομιλία αντίστοιχα, συναντάμε στην [Nwe και Li, 2005]. Αντίστοιχα, για την αναγνώριση του ομιλητή στις [Duan], [Liu2004], [Liu1998] χρησιμοποιούνται χαρακτηριστικά όπως συντελεστές σάφματος συχνοτήτων Μελ (Mel-Frequency Cepstral Coefficients- MFCCs), βραχύχρονη ενέργεια, τόνος, δομική συχνότητα (fundamental frequency), αρμονικές και εύρος ζώνης ήχου. Τέλος, το ηχητικό περιεχόμενο μπορεί να βοηθήσει και στην ορθή ταυτοποίηση των σιωπηλών μαύρων καρέ, εξετάζοντας το μέσο επίπεδο της έντασης του ήχου [Satterwhite και Marques, 2004]. Βέβαια είναι δυνατός και ο συνδυασμός των δύο τεχνικών. Έτσι, στην [Wang et al., 2008] γίνεται χρήση περιγραφών εικόνας (image descriptors) αλλά και ήχου, καθώς και στην [Duygulu et al. 2004] η αναγνώριση προκύπτει από την ανάλυση χρωματικών και ηχητικών χαρακτηριστικών.

Τέλος, η χρήση βάσης δεδομένων για τη σύγκριση εικονοσειρών είναι ευρέως διαδεδομένη. Οι Sánchez et al [2002] για τη σύγκριση χρησιμοποιούν τα βασικά καρέ της εικονοσειράς μαζί με μια βάση δεδομένων, ενώ οι Anusha et al [2008] υπολογίζουν την ευκλείδεια απόσταση των ακμών όπως έχουν αποθηκευτεί στη βάση. Εναλλακτικά, η μέθοδος που παρουσιάζεται στην [Roover et al., 2005] χρησιμοποιεί συναρτήσεις

κατακερματισμού (hash functions) για να συγκρίνει τις εικόνες της κανονικής ροής του προγράμματος μαζί με αυτές που είναι αποθηκευμένες στη βάση δεδομένων.

Γενικά, παρατηρείται από τα παραπάνω πως υπάρχει πλήθος τεχνικών για εύρεση διαφημιστικών μηνυμάτων σε μια εικονοσειρά τηλεοπτικού σήματος. Το πρόβλημα έγκειται στο ότι βασίζονται σε ευριστικά στοιχεία (heuristics) όπως η ύπαρξη μαύρων καρέ ή μορφής του ηχητικού περιεχομένου, τα οποία αν αλλάξουν ουσιαστικά αχρηστεύουν τις τεχνικές. Η μέθοδος της εξαντλητικής σάρωσης και ανάλυσης των καρέ από την άλλη, προσφέρει μεγαλύτερη αξιοπιστία με υψηλό όμως υπολογιστικό κόστος, ιδιαίτερα αν αναλογιστούμε την περίπτωση του σήματος υψηλής ευκρίνειας (High Definition) και το μεγάλο όγκο δεδομένων στη βάση. Η αντιμετώπιση αυτών των δυσκολιών σε αυτή τη διπλωματική εργασία γίνεται μειώνοντας το χώρο αναζήτησης από τη βάση, μειώνοντας παράλληλα και το χρόνο υπολογισμού της πληροφορίας της εικόνας, με τη βοήθεια του αλγορίθμου Επαγωγικής Σάρωσης [Giannoukos et al., 2010].

3. Μέθοδος

Σε αυτό το κεφάλαιο θα γίνει η περιγραφή της μεθόδου που χρησιμοποιήθηκε για την ανάλυση και ταυτοποίηση των εικονοσειρών. Για την κατανόηση του τρόπου με τον οποίο λειτουργεί ο αλγόριθμος είναι απαραίτητες μερικές θεμελιώδεις γνώσεις στο αντικείμενο της επεξεργασίας εικόνας. Ψηφιακή εικόνα θεωρείται μια δισδιάστατη απεικόνιση $f(x, y)$ όπου x, y οι χωρικές συντεταγμένες και το πλάτος της f ονομάζεται ένταση της εικόνας στο σημείο (x, y) . Η ψηφιακή εικόνα αποτελείται από πεπερασμένο αριθμό στοιχείων καθένα από τα οποία έχει συγκεκριμένη θέση και τιμή. Αυτά τα στοιχεία ονομάζονται εικονοστοιχεία (pixel) [Torre και Poggio, 1980]. Η ψηφιακή επεξεργασία εικόνας συνίσταται στην επεξεργασία της ψηφιακής εικόνας μέσω του υπολογιστή. Στη διπλωματική αυτή, ο περιγραφέας που επιλέχθηκε για να χαρακτηρίζει και ταυτοποιεί μια εικόνα είναι ο περιγραφέας ακμών (edge descriptor). Ανίχνευση ακμών ονομάζεται η διαδικασία που επιχειρεί να χαρακτηρίσει τις διαφοροποιήσεις στην ένταση των εικονοστοιχείων της εικόνας, σε σχέση με τα φυσικά αίτια που τις προκάλεσαν [Pitas, 2000; González και Woods, 2008]. Όσον αφορά στην ανίχνευση ακμών υπάρχει πληθώρα τεχνικών που βασίζεται σε διαφορετικό μαθηματικό υπόβαθρο.

Η γενική μέθοδος μπορεί να χωριστεί σε δύο κύριες υποκατηγορίες, την ανάλυση της εικόνας και εξαγωγή της πληροφορίας με τον αλγόριθμο Επαγωγικής Σάρωσης (Operator Context Scanning - OCS), και την σύγκριση, επεξεργασία και μετέπειτα ταυτοποίηση όλων των δεδομένων της εικονοσειράς.

3.1. Αλγόριθμος Επαγωγικής Σάρωσης κατά μπλοκ

Ο αλγόριθμος Επαγωγικής Σάρωσης (Operator Context Scanning) διαχωρίζει σε μια εικόνα εισόδου τις περιοχές ενδιαφέροντος (Regions of Interest), σύμφωνα με την πυκνότητα που παρουσιάζουν σε ενδιαφέροντα χαρακτηριστικά. Χρησιμοποιεί τελεστές εικονοστοιχείων και των γειτονιών τους (pixel neighborhoods) για τη διαχώριση του χώρου των εικονοστοιχείων της εικόνας, ο οποίος κυλίνεται στην εικόνα (sliding window operator).

Η ανάλυση εικόνας με κυλιόμενα παράθυρα είναι συνήθης τεχνική στην επεξεργασία εικόνας. Οι τελεστές κυλιόμενων παραθύρων στη διεθνή βιβλιογραφία

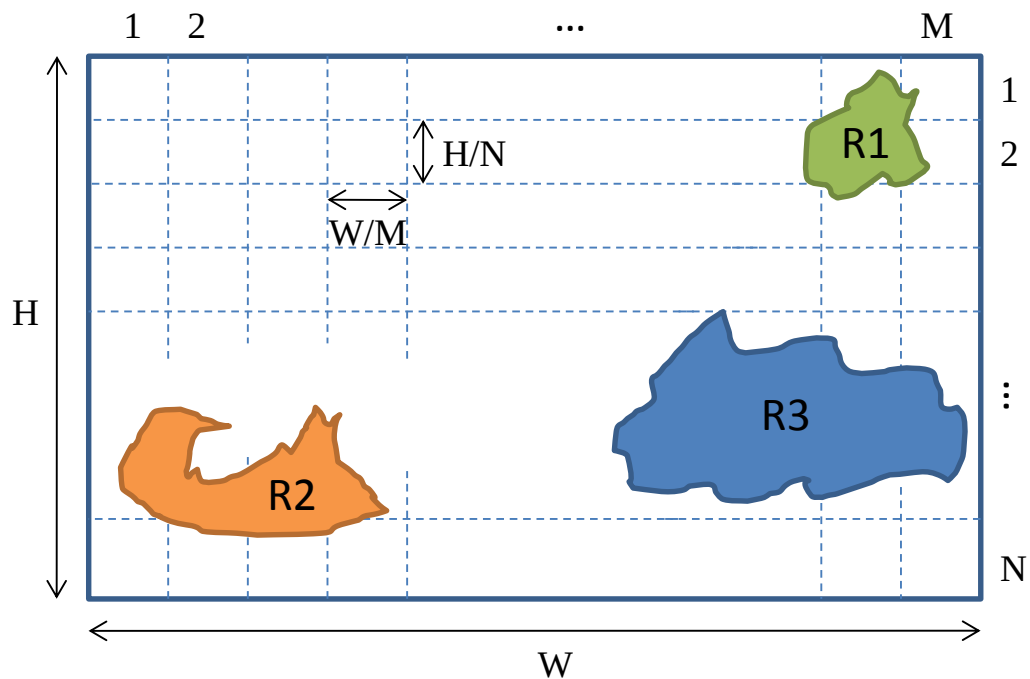
αναφέρονται και ως φίλτρα πυρήνων (kernel filters) και ουσιαστικά αποδίδουν μια τιμή σε ένα εικονοστοιχείο ή και τη γειτονιά του, ανάλογα με τη πιθανότητα να ανήκει η εν λόγω περιοχή σε ένα πρότυπο που αναζητείται στην εικόνα. Αυτές οι περιοχές αναζητούνται με τη μέθοδο της εξαντλητικής σάρωσης όλων των εικονοστοιχείων της εικόνας. Διάφοροι τομείς στους οποίους έχει γίνει χρήση αυτής της μεθόδου είναι ο εντοπισμός προσώπων ή ανθρώπων (face/human detection) [Huang και Shimizu, 2006; Meynet et al., 2007; Dalal και Triggs, 2005], η ανίχνευση ακμών [Shen και Zen, 2006], η κατάτμηση υφής [Sayadi et al., 2007; Kimesova, 2000] και η κατηγοριοποίηση εικονοσειρών [Lampert et al., 2008; Feng et al., 2005]. Το πρόβλημα με αυτή τη μέθοδο είναι το μεγάλο υπολογιστικό κόστος που παρουσιάζει, αφού σαρώνουν εξαντλητικά όλο το χώρο των εικονοστοιχείων της εικόνας. Εναλλακτικά, χρησιμοποιούνται στρατηγικές αδρομερούς-προς-το-λεπτομερές (coarse-to-fine strategies) οι οποίες εφαρμόζουν μερικά γρήγορα περάσματα της εικόνας, από τα οποία προκύπτουν υποψήφιες περιοχές πάνω στις οποίες εφαρμόζεται έπειτα η τεχνική κυλιόμενων παραθύρων, μειώνοντας έτσι το χώρο αναζήτησης [Lampert et al., 2008; Feng et al., 2005].

Ο αλγόριθμος Επαγωγικής Σάρωσης μειώνει το χώρο επεξεργασίας αφού δεν εφαρμόζει εξαντλητική σάρωση όλου του χώρου των εικονοστοιχείων. Επιπλέον, εφαρμόζει μόνο ένα αλγοριθμικό πέρασμα σαρώνοντας τις περιοχές ενδιαφέροντος και υποδειγματοληπτώντας τις υπόλοιπες.

Μπορούμε να θεωρήσουμε πως το αντικείμενο προς αναζήτηση σε μια εικόνα περιγράφεται από ένα σύνολο χαρακτηριστικών (feature set). Ο αλγόριθμος εστιάζει στις περιοχές που πληρούν το σύνολο των χαρακτηριστικών αυτών και μειώνει τη δειγματοληψία στις υπόλοιπες περιοχές, μειώνοντας έτσι το χώρο αναζήτησης. Επίσης, η επεξεργασία γίνεται κατά μπλοκ δηλαδή δεν βασίζεται στα εικονοστοιχεία αλλά σε ομάδες αυτών, δηλαδή σε υποπεριοχές της εικόνας.

Φορμαλιστικά η μέθοδος μπορεί να περιγραφεί ως εξής: Έστω μια εικόνα I η οποία αποτελείται από H (height) εικονοστοιχεία κατά ύψος και W (width) εικονοστοιχεία κατά πλάτος. Η εικόνα περιέχει το αντικείμενο προς κατάτμηση το οποίο περιγράφεται από ένα διάνυσμα n χαρακτηριστικών $A=[a_1, a_2, \dots, a_n]$. Ο Τελεστής Κυλιόμενων Παραθύρων ΤΚΠ (Sliding Windows Operator) εφαρμόζεται σε περιοχή R της εικόνας με κέντρο το εικονοστοιχείο p και εφαρμόζει μια συνάρτηση καταλληλότητας (fitness function) f στην περιοχή R , μέσω του διανύσματος A .

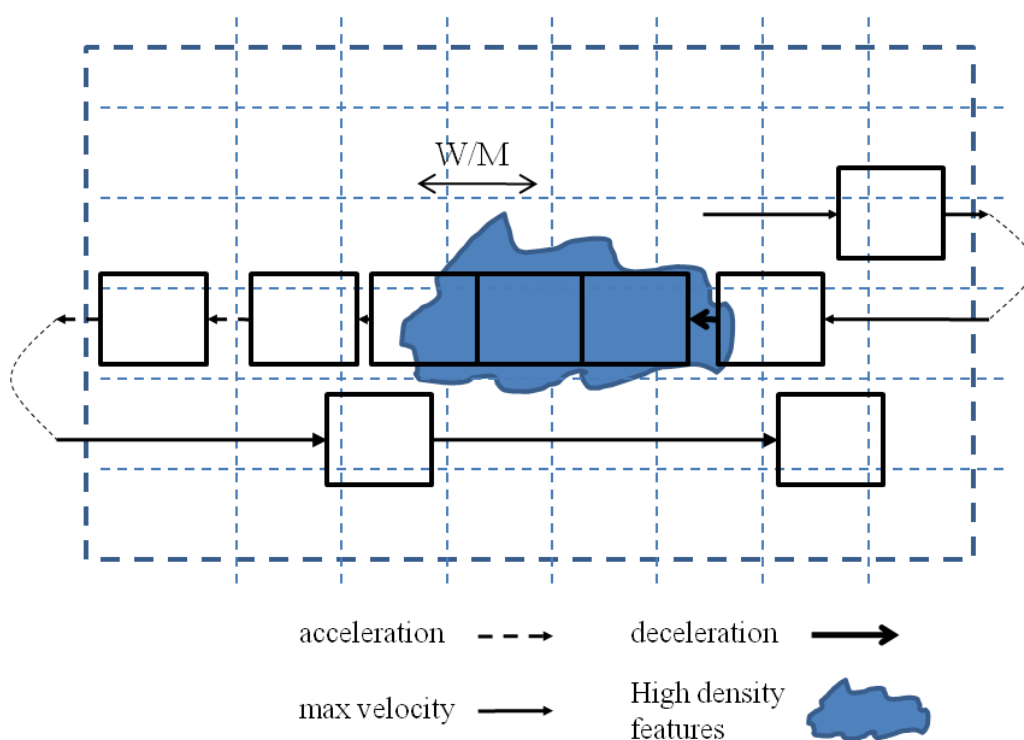
Στη γενική περίπτωση της επεξεργασίας κατά μπλοκ, η εικόνα αυτή χωρίζεται σε M περιοχές κατά πλάτος και N περιοχές κατά ύψος, δηλαδή $K=M \times N$ υποπεριοχές. Σε μια τυπική επεξεργασία κατά μπλοκ θα γινόταν εφαρμογή ενός τελεστή (πχ ανίχνευση ακμών) σε κάθε μία από τις K υποπεριοχές καλύπτοντας ολόκληρη την εικόνα ακόμα και αν μεγάλο μέρος δεν περιείχε χρήσιμη πληροφορία. Όπως φαίνεται στο σχήμα 4.1.1. οι υποπεριοχές $R1, R2$ και $R3$ έχουν υψηλή πυκνότητα χαρακτηριστικών οπότε κατά τον έλεγχο των υπολοίπων πρακτικά γίνεται σπατάλη των πόρων του συστήματος.



Σχήμα 3.1.1. Εικόνα χωρισμένη σε μπλοκ με τρεις περιοχές με υψηλή πυκνότητα χαρακτηριστικών

Αντιθέτως, στον αλγόριθμο Επαγωγικής Σάρωσης χρησιμοποιείται μία κυλιόμενη περιοχή, με διαστάσεις $W \times H$, όπως μια τυπική περιοχή της εικόνας. Η συνάρτηση f χρησιμοποιείται για τη διαπίστωση αν μια περιοχή είναι υποψήφια να περιέχει το ζητούμενο αντικείμενο. Αν το αποτέλεσμα της f ξεπερνά ένα προκαθορισμένο κατώφλι (threshold) τότε η περιοχή σημειώνεται με μονάδα '1' ενώ σε αντίθετη περίπτωση με μηδέν '0'. Τελικά παράγεται μια δυαδική μάσκα (binary mask) στην οποία τα θετικά εικονοστοιχεία αντιστοιχούν σε περιοχές ενδιαφέροντος.

Η συνάρτηση f δεν εφαρμόζεται σε όλα τα μπλοκ της εικόνας αλλά ακολουθείται ένας συγκεκριμένος αλγόριθμος σάρωσης. Ορίζουμε την έννοια της ταχύτητας κύλισης ως τον αριθμό εικονοστοιχείων που παραλείπονται μέχρι την εκ νέου εφαρμογή του τελεστή στην επόμενη περιοχή της εικόνας. Για να αποφύγουμε τον υπολογισμό του χαρακτηριστικού διανύσματος πολλαπλές φορές σε μία υποπεριοχή, εισάγεται η έννοια της ελάχιστης ταχύτητας κύλισης. Στην περίπτωση αυτή, ο αλγόριθμος δεν μπορεί να έχει μηδενική ταχύτητα, αλλά στην οριζόντια και κατακόρυφη διεύθυνση έχει το λιγότερο — και — ταχύτητα αντίστοιχα.



Σχήμα 3.1.2. Επαγωγική Σάρωση κατά μπλοκ

Έστω $p(x,y)$ το κεντρικό εικονοστοιχείο της περιοχής R στην οποία εφαρμόζουμε τη συνάρτηση καταλληλότητας f και $p'(x',y')$ το επόμενο κεντρικό εικονοστοιχείο προς εφαρμογή της συνάρτησης. Για λόγους απλότητας θεωρούμε ότι τα y και y' είναι μικρότερα του συνολικού πλάτους της εικόνας. Το επόμενο στοιχείο προς εξέταση θα είναι στην ίδια γραμμή ($x'=x$) και στο πλάτος που περιγράφεται από την εξίσωση

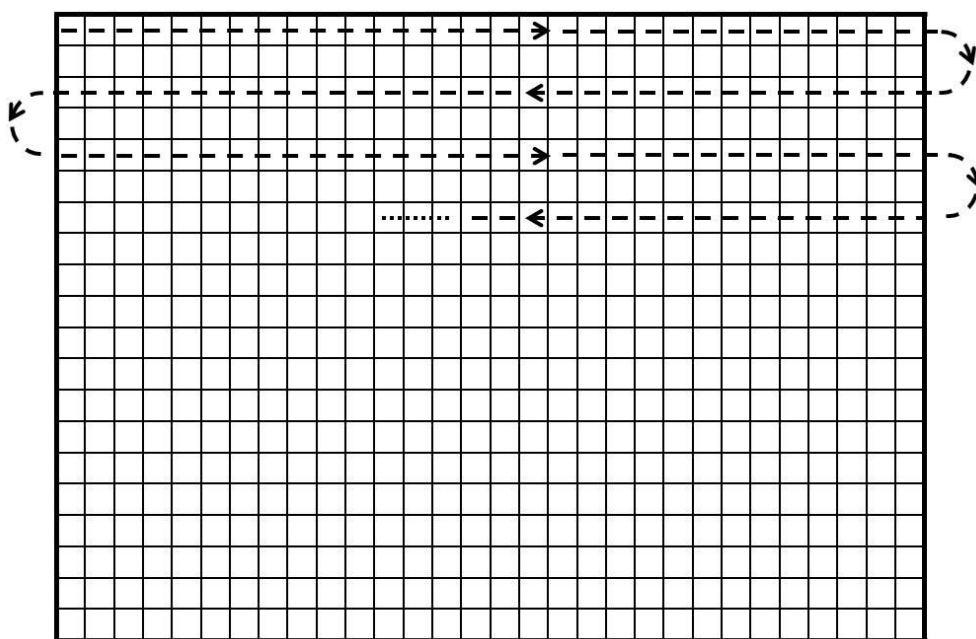
$$(1)$$

όπου είναι η παράμετρος της οριζόντιας ταχύτητας του αλγορίθμου. Πρακτικά δηλώνει πόσα εικονοστοιχεία παραλείπονται κατά πλάτος, μέχρι να φθάσουμε στο επόμενο προς εξέταση. Για την παράμετρο αυτή της ταχύτητας ισχύει όπου η μέγιστη τιμή που μπορεί να λάβει η παράμετρος και η ελάχιστη ταχύτητα κύλισης. Η παράμετρος που επηρεάζει την τιμή της αναφέρεται στο εικονοστοιχείο p που εξετάζεται, στοιχείο που υποδηλώνει το γεγονός ότι η ταχύτητα επηρεάζεται από τα εικονοστοιχεία που σαρώνονται. Ισχύει λοιπόν, όπου το πλήθος των εικονοστοιχείων της εικόνας. Η παράμετρος της ταχύτητας ακολουθεί την εξίσωση:

Όπως φαίνεται στην ανωτέρω εξίσωση εφαρμόζουμε το μοντέλο Προσθετικής Αύξης Πολλαπλασιαστικής Μείωσης ΠΑΠΜ (Additive Increase Multiplicative Decrease). Ο αλγόριθμος ΠΑΠΜ χρησιμοποιείται ευρέως στα δίκτυα υπολογιστών για να αποτραπεί η κυκλοφοριακή συμφόρηση των μεταδιδόμενων δεδομένων (Network congestion avoidance) [Chiu και Jain, 1989]. Στη συγκεκριμένη περίπτωση η κυκλοφοριακή συμφόρηση υποδηλώνει την περιοχή ενδιαφέροντος. Όταν η συνάρτηση καταλληλότητας f για το τρέχον εικονοστοιχείο p δεν ξεπεράσει το κατώφλι, και συνεπώς επιστρέψει '0', τότε η εξεταζόμενη περιοχή δεν περιέχει το αντικείμενο ενδιαφέροντος και αυξάνουμε προσθετικά την ταχύτητα με την σταθερά προσθετικής αύξης. Έτσι μειώνεται η δειγματοληψία στις περιοχές που δεν παρουσιάζουν χρήσιμα χαρακτηριστικά. Αντίθετα, όταν η συνάρτηση f επιστρέψει '1' σημαίνει ότι η περιοχή είναι υποψήφια να περιέχει τα επιθυμητά χαρακτηριστικά και μειώνεται πολλαπλασιαστικά η ταχύτητα με τη σταθερά, ώστε να μη χαθεί κομμάτι της ζητούμενης περιοχής. Βέβαια, σύμφωνα με τις αρχές του αλγορίθμου, η ταχύτητα εξαρτάται από το κατάλληλο μέγεθος παραθύρου που αντιστοιχεί στην περιοχή R που εξετάζεται.

Η παραπάνω διαδικασία εφαρμόζεται όχι μόνο στον οριζόντιο άξονα αλλά και στον κατακόρυφο. Αν κατά την εξέταση μιας γραμμής ο αλγόριθμος δεν εντοπίσει επαρκή αριθμό υποψήφιων εικονοστοιχείων, τότε παραλείπεται ένας αριθμός γραμμών κατά την

ίδια λογική με προηγούμενως. Τελικά η σάρωση όλης της εικόνας γίνεται όπως στο σχήμα που ακολουθεί.



Σχήμα 3.1.3. Βουστροφηδόν σάρωση της εικόνας

Όπως φαίνεται στο Σχήμα 3.1.3. η σάρωση της εικόνας γίνεται ‘βουστροφηδόν’. Δηλαδή όταν η σάρωση φθάσει στο τέλος του πλάτους της εικόνας αλλάζει η γραμμή αλλά και η φορά της σάρωσης. Έτσι διατηρείται μια ενιαία μορφή κίνησης και κάθε περιοχή που εξετάζεται είναι γειτονική με την προηγούμενη. Αυτό εισάγει ένα χαρακτήρα μνήμης και σημαίνει πως δε χρειάζεται κάποιος μηδενισμός της ταχύτητας κάθε φορά που αλλάζει η γραμμή.

Όσον αφορά στις παραμέτρους του αλγορίθμου αυτές είναι η παράμετρος της μέγιστης ταχύτητας, της επιτάχυνσης καθώς και της επιβράδυνσης για οριζόντιο και κάθετο άξονα αντίστοιχα και το μέγεθος και σχήμα του τελεστή ΤΚΠ. Εξαρτώνται από δύο παράγοντες: τις διαστάσεις των αντικειμένων στην εικόνα και την ικανότητα του τελεστή ΤΚΠ να διακρίνει τις περιοχές ενδιαφέροντος από τις περιοχές όπου δεν υπάρχει η επιθυμητή πληροφορία.

Πρώτον, οι παράμετροι της μέγιστης ταχύτητας του αλγορίθμου στον οριζόντιο και κάθετο άξονα, εξαρτώνται από το ελάχιστο πλάτος και ύψος των αντικειμένων στην εικόνα αντίστοιχα. Αν η μέγιστη ταχύτητα υπερβεί το μέγεθος ενός αντικειμένου τότε υπάρχει ο

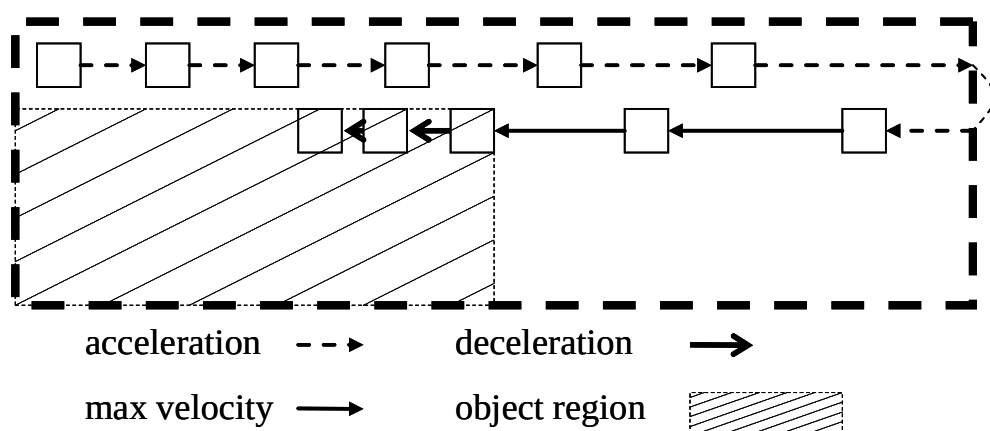
κίνδυνος ο αλγόριθμος να ξεπεράσει εντελώς το αντικείμενο και λανθασμένα να χαρακτηρίσει μια περιοχή ως μη ενδιαφέρουσα.

Δεύτερον, η παράμετρος της επιτάχυνσης λειτουργεί κατά το πρότυπο ΠΑΠΜ αθροιστικά. Αυτό βοηθά στην περίπτωση που ο τελεστής ΤΚΠ αποτύχει να εξάγει μερικά υποψήφια χαρακτηριστικά μέσα στη σωστή περιοχή ενδιαφέροντος να μην υπάρχει μεγάλη αύξηση της ταχύτητας και συνεπώς να μη χαθεί μεγάλο επιπλέον κομμάτι της ζητούμενης περιοχής.

Τρίτον, αν ο τελεστής ΤΚΠ δίνει σαν αποτέλεσμα ένα μεγάλο αριθμό από κίβδηλα εικονοστοιχεία (spurious pixels), τότε είναι αναποτελεσματικό να μηδενίζεται τελείως σε κάθε τέτοιο εικονοστοιχείο η ταχύτητα του αλγορίθμου. Για αυτό το λόγο η παράμετρος της επιβράδυνσης λειτουργεί πολλαπλασιαστικά, μειώνοντας γρήγορα αλλά όχι αμέσως την ταχύτητα, ελαχιστοποιώντας έτσι το υπολογιστικό κόστος. Η τιμή της παραμέτρου εξαρτάται από το μέγεθος τόσο των κίβδηλων όσο και των υποψηφίων περιοχών στην εικόνα.

Οι τελευταίες παράμετροι, το μέγεθος και το σχήμα του ΤΚΠ, εξαρτώνται από τους εξής παράγοντες: τις διαστάσεις του αντικειμένου στην εικόνα και την ελάχιστη περιοχή που πρέπει να επεξεργαστεί ώστε να εξάγει αξιόπιστα συμπεράσματα σύμφωνα με το διάνυμα χαρακτηριστικών A.

Οι αρχές λειτουργίας του αλγορίθμου Επαγωγικής Σάρωσης συνοψίζονται στο Σχήμα 3.1.4.



Σχήμα 3.1.4. Οι αρχές λειτουργίας του αλγορίθμου Επαγωγικής Σάρωσης

3.2. Αλγόριθμος Ταύτισης Εικονοσειρών

Ο αλγόριθμος ταύτισης των εικονοσειρών χρησιμοποιεί μια βάση δεδομένων για την ταυτοποίηση μιας δεδομένης εικονοσειράς με κάποια από τις καταχωρημένες στη βάση εικονοσειρές. Το αποτέλεσμα του αλγορίθμου Επαγωγικής Σάρωσης κατά μπλοκ είναι μια δυαδική μάσκα με πληροφορία για τα εικονοστοιχεία ενδιαφέροντος της εικόνας. Η πληροφορία αυτή αποθηκεύεται σε μια βάση δεδομένων με σκοπό αργότερα να γίνεται γρήγορα η ανάκληση, σύγκριση και ενδεχομένως ταυτοποίηση των εικόνων της τηλεοπτικής εκπομπής που εξετάζεται. Το σύστημα που θα εφαρμόζει τα παραπάνω όμως πρέπει να αποκρίνεται σε σχεδόν πραγματικό χρόνο, στοιχείο που σημαίνει πως πρέπει να είναι ταχύ και αποτελεσματικό. Εξαιτίας αυτού, η επιλογή της εξαντλητικής σάρωσης και ταύτισης όλων των καρέ της εικονοσειράς δεν αποτελεί αποδεκτή λύση.

Σημαντικό λοιπόν σημείο του αλγορίθμου είναι η υποδειγματοληψία που γίνεται στα καρέ. Παραλείποντας έναν αριθμό από καρέ μειώνεται το υπολογιστικό φορτίο ενώ παράλληλα το κόστος σε αξιοπιστία είναι ασήμαντο, αφού γειτονικά καρέ έχουν πάρα πολλές ομοιότητες μεταξύ τους.

Για τη σύγκριση των εικόνων εισόδου με αυτών της βάσης δεδομένων του συστήματος, υπολογίζεται το μέσο τετραγωνικό σφάλμα μεταξύ των ιστογραμμάτων των μπλοκ. Έστω κάθε εικόνα χωρίζεται σε K μπλοκ και επομένως το ιστόγραμμα αυτής θεωρούμε πως έχει K καλάθια (bins). Επίσης, έστω H_{in} το ιστόγραμμα εικόνας εισόδου και H_i , $0 \leq i \leq M$ τα ιστογράμματα των αποθηκευμένων εικονοσειρών στη βάση δεδομένων. Το μέσο τετραγωνικό σφάλμα υπολογίζεται με βάση τον ακόλουθο τύπο:

—

Ο δείκτης αναφέρεται στο j μπλοκ του ιστογράμματος H_i . Κατά τη διάρκεια των πειραμάτων φάνηκε πως η εικονοσειρά που περιέχει τα καρέ με το μικρότερο μέσο τετραγωνικό σφάλμα δεν είναι πάντοτε η σωστή. Αυτό μπορεί να οφείλεται σε μικρές αποκλίσεις των καρέ, λόγω της δειγματοληψίας, και στο γεγονός ότι μπορεί να υπάρχει κάποιο καρέ άλλης διαφήμισης με παρόμοιες περιοχές ενδιαφέροντος. Έτσι, ορίστηκε ένα κατώφλι (threshold) με σκοπό το διαχωρισμό των καρέ που παρουσιάζουν ομοιότητες με την εικόνα εισόδου. Για να θεωρηθεί μια διαφήμιση υποψήφια, πρέπει το καρέ της με το

χαμηλότερο μέσο τετραγωνικό σφάλμα να μην υπερβαίνει το κατώφλι αυτό. Όμως, κατά την επεξεργασία μπορεί να βρεθεί κάποιο τμήμα εικονοσειράς το οποίο να είναι κίβδηλο (spurious candidate) αλλά παρ' όλα αυτά μη ξεπερνά το κατώφλι. Αντίστοιχα μπορεί να συμβεί και το ανάποδο δηλαδή για τη σωστή διαφήμιση να μη βρίσκουμε κάποιο καρέ που να είναι όμοιο με αυτό προς αναζήτηση. Ο λόγος για την πρώτη περίπτωση του κίβδηλου καρέ είναι πιθανή ομοιότητα στη σκηνή ή στις ακμές της εικόνας. Στη δεύτερη εξαιτίας της δειγματοληψίας μπορεί να μην υπάρχει το συγκεκριμένο καρέ στη βάση δεδομένων και τα κοντινά του καρέ λόγω γρήγορης κίνησης στην εικόνα να διαφέρουν αρκετά. Όμως αλλαγή της τιμής για το κατώφλι αυτό δε λύνει το πρόβλημα και έτσι ορίζονται χρονικά παράθυρα επεξεργασίας και εφαρμόζεται η ακόλουθη μέθοδος σύγκρισης όχι ενός αλλά μιας ομάδας από καρέ στη σειρά (Group of Frames).

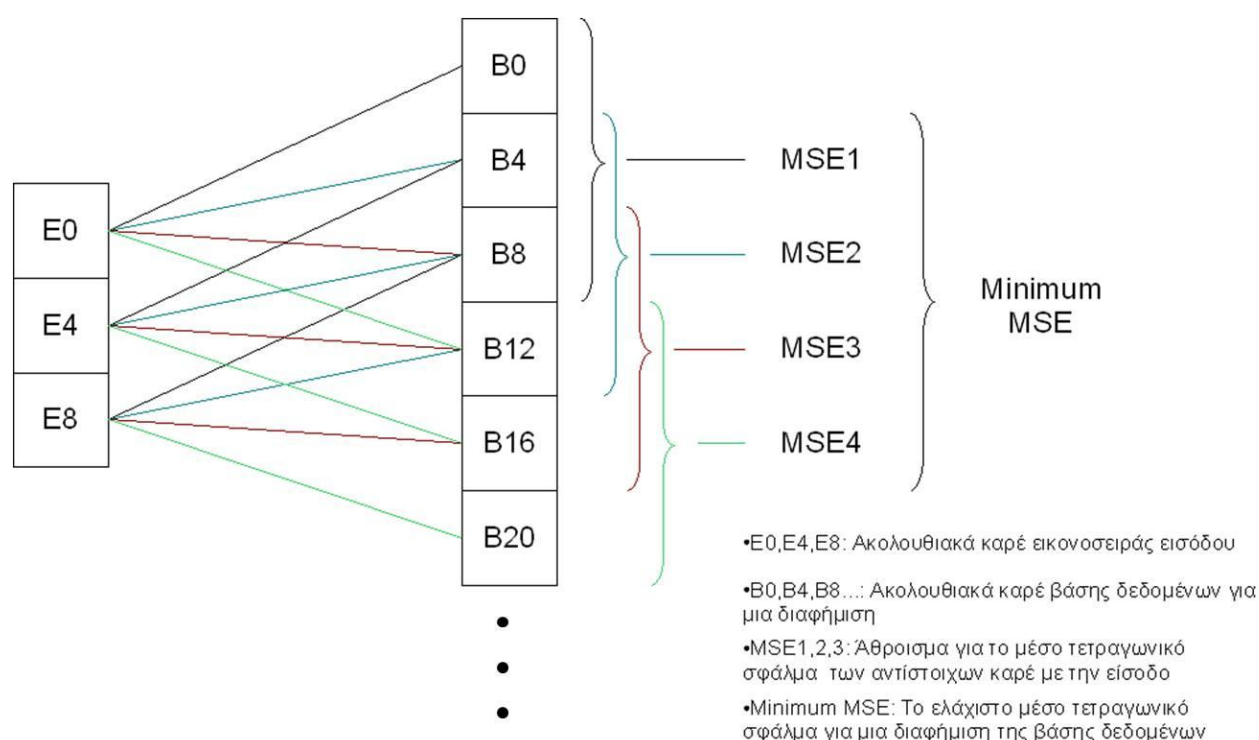
Αποθηκεύονται τα δεδομένα των ιστογραμμάτων των μπλοκ για ένα σύνολο έστω T διαδοχικών καρέ σύμφωνα με παράθυρο κύλισης στη διάσταση του χρόνου. Έπειτα, γίνεται έλεγχος στη βάση και ελέγχεται το μέσο τετραγωνικό σφάλμα για τα T διαδοχικά καρέ εξασφαλίζοντας ότι ακόμα και αν κάποιο καρέ της εσφαλμένης διαφήμισης μοιάζει με αυτό της εισόδου της εικονοσειράς, η πιθανότητα να μοιάζουν και τα υπόλοιπα μειώνεται αρκετά. Δηλαδή, στην πραγματικότητα το μέσο τετραγωνικό σφάλμα για ένα καρέ της βάσης αναφέρεται στο άθροισμα των σφαλμάτων για αυτό το καρέ και των $T-1$ επόμενων από την ίδια διαφήμιση, σε σχέση πάντα με την ακολουθία των καρέ από την εικονοσειρά εισόδου.

όπου H_{in} το ιστόγραμμα εικόνας εισόδου και H_j το ιστόγραμμα του καρέ j της διαφήμισης i της βάσης. Έτσι ο έλεγχος που γίνεται στην ουσία είναι έλεγχος ομοιότητας όχι ενός μόνο καρέ εισόδου αλλά μιας περιοχής συνεχόμενων καρέ. Σύμφωνα με αυτή την τεχνική, τη χρονική στιγμή t αντιστοιχεί σε παράθυρο μεγέθους T ως εξής.

Το χρονικό αυτό παράθυρο εφαρμόζεται τόσο στη ροή εισόδου όσο και στις διαφημίσεις της βάσης. Όσον αφορά στη ροή εισόδου το χρονικό αυτό παράθυρο κυλίνεται γραμμικά σύμφωνα με την κανονική ροή του προγράμματος, ενώ για να γίνει ταύτιση με τις διαφημίσεις της βάσης υπολογίζεται το αν ακολουθείται η ίδια συμπεριφορά.

Τέλος, για κάθε διαφήμιση της βάσης αναζητείται το ελάχιστο μέσο τετραγωνικό σφάλμα (MSE_{min}) δηλαδή

όπου F το σύνολο των καρτέ για μια διαφήμιση. Αυτό απεικονίζεται και στο σχήμα 3.2.1. που ακολουθεί.

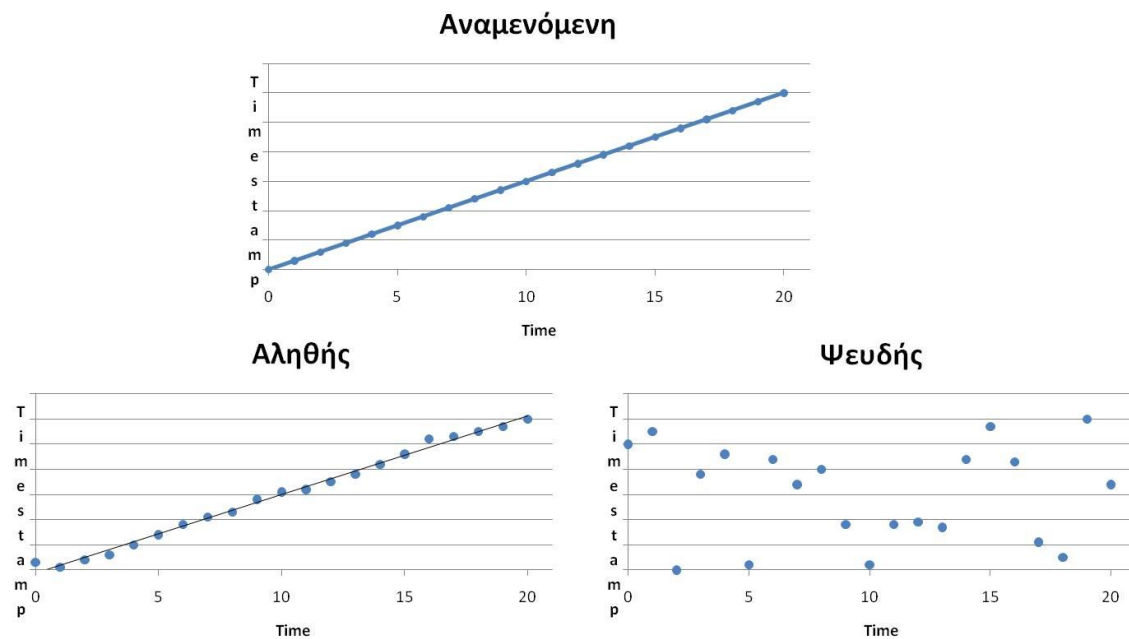


Σχήμα 3.2.1. Εύρεση ελαχίστου μέσου τετραγωνικού σφάλματος για μια διαφήμιση της βάσης δεδομένων με χρήση παραθύρων κύλισης στη διάσταση του χρόνου

Η διαφήμιση είναι υποψηφία για ταύτιση μόνο εάν

Παράλληλα εφαρμόζεται και μια δεύτερη μέθοδος σύγκρισης βασισμένη στις χρονοσφραγίδες (timestamps) του κάθε καρτέ. Η χρονοσφραγίδα δείχνει σε ποια χρονική θέση βρίσκεται ένα καρτέ στην εικονοσειρά κάποιας διαφήμισης. Κάνοντας έλεγχο στις χρονοσφραγίδες των καρτέ για κάθε υποψήφια διαφήμιση εξάγονται βοηθητικά στοιχεία για το αν είναι η σωστή διαφήμιση ή όχι. Αυτό γίνεται γιατί αναμένεται η εύρεση μιας ακολουθίας χρονοσφραγίδων η οποία να είναι γραμμική και αύξουσα, με την κάθε

χρονοσφραγίδα να μην απέχει πολύ από την επόμενη. Συνεπώς, μόνο σε περίπτωση που δεν αποκλίνει σημαντικά από αυτό το μοντέλο θεωρείται μια διαφήμιση υποψήφια. Έτσι, απορρίπτονται οι διαφημίσεις που αν και μπορεί να έχουν χαμηλό μέσο τετραγωνικό σφάλμα που να μην υπερβαίνει το κατώφλι, έχουν ταυτιστεί καρέ από διάσπαρτα τμήματα της χωρίς λογική ακολουθία. Για τη σύγκριση αυτή εφαρμόζεται η μέθοδος ελαχίστων τετραγώνων σε συνδυασμό με ευριστικές τεχνικές οι οποίες ελέγχουν την τιμή των χρονοσφραγίδων σε σχέση με την αναμενόμενη.



Σχήμα 3.2.2. Γραφήματα για τις χρονοσφραγίδες μια αληθούς και μιας κίβδηλης υποψήφιας διαφήμισης σε σχέση με την κανονική αναμενόμενη μορφή

Όλα τα παραπάνω λοιπόν, βοηθούν στο να βρίσκονται οι ομοιότητες μεταξύ των καρέ. Η διαφήμιση που τελικά θα αντιστοιχηθεί σε κάποιο τμήμα της ροής εισόδου θα είναι αυτή που έχει ταυτιστεί με τα περισσότερα καρέ της ροής. Αυτό υπολογίζεται ορίζοντας δύο σύνολα, το αληθές T και το ψευδές F . Αρχικά όλες οι διαφημίσεις της βάσης δεδομένων είναι εν δυνάμει υποψήφιες, οπότε συμπεριλαμβάνονται στο σύνολο T . Με την έναρξη των συγκρίσεων και την κύλιση των παραθύρων, υπολογίζονται τα μέσα τετραγωνικά σφάλματα όλων των διαφημίσεων του συνόλου T . Οι διαφημίσεις που υπερβαίνουν το κατώφλι MSE_{thres} μεταφέρονται στο σύνολο F . Αυτό συνεχίζεται για τα επόμενα καρέ που εξετάζονται με τα χρονικά παράθυρα, μέχρι η ροή ή κάποια ταυτοποιημένη διαφήμιση να φθάσει στο τέλος της είτε να μην υπάρχουν άλλες

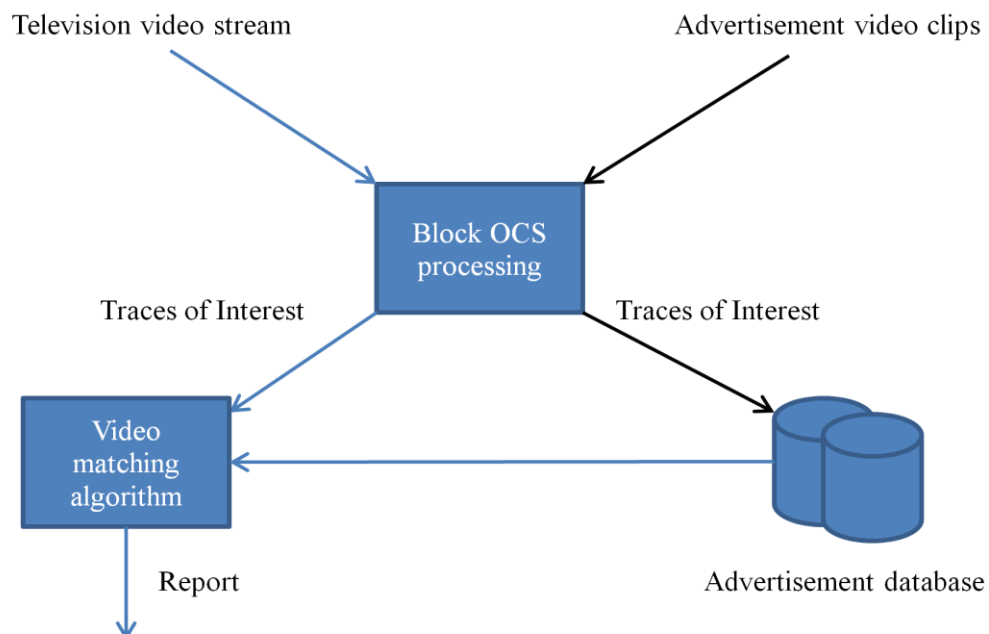
διαφημίσεις στο σύνολο T . Τότε, γίνεται έλεγχος για τις διαφημίσεις που έχουν απομείνει στο T ώστε να διαπιστωθεί αν κάποια από αυτές μπορεί να αντιστοιχηθεί στο κομμάτι της ροής που εξετάστηκε. Μια διαφήμιση θεωρείται ότι έχει ταυτιστεί σωστά αν η διάρκεια που έμεινε στο σύνολο T συμπίπτει με τη διάρκεια που έχει στη βάση δεδομένων και αν η ακολουθία στις χρονοσφραγίδες της έχει την επιθυμητή σειρά. Αλλιώς, είτε θεωρείται ότι το μέρος της ροής δεν υπάρχει στη βάση (σημειώνεται ως «άγνωστη» ροή) είτε εμφανίζεται προειδοποίηση ότι μια καταχωρημένη διαφήμιση στη βάση εμφανίστηκε συμπτυγμένη.

4. Σύστημα

Σε αυτό το κεφάλαιο της διπλωματικής περιγράφεται το σύστημα που χρησιμοποιήθηκε για την υλοποίηση των προαναφερθέντων μεθόδων για την ταύτιση των εικονοσειρών.

4.1. Τεχνικά Χαρακτηριστικά Συστήματος

Η γλώσσα προγραμματισμού που χρησιμοποιήθηκε ήταν η γλώσσα “C” και έγινε χρήση της πλατφόρμας “Microsoft Visual Studio 2008 Professional Edition” [VS2008-Pro]. Επίσης, χρησιμοποιήθηκε τη βιβλιοθήκη “Open CV” (**O**pen **S**ource **C**omputer **V**ision Library [OpenCV, 2010]), η οποία περιέχει πλήθος συναρτήσεων για την επεξεργασία εικόνων. Όσον αφορά στη βάση δεδομένων, έγινε χρήση του συστήματος διαχείρισης βάσεων δεδομένων “MySQL” [MySQL, 2010], ενώ η σύνδεση με το πρόγραμμα έγινε με τη βοήθεια των βιβλιοθηκών “mysql-connector-5.1.48” [MySQL connectors, 2010]. Παράλληλα, χρησιμοποιήθηκε το εργαλείο “PHPMYAdmin” [PHPMYAdmin, 2010] για τη διαχείριση και επισκόπηση της βάσης με διακομιστή (server) τον “Apache” [Apache, 2010].



Σχήμα 4.1.1. Διάγραμμα επισκόπησης γενικής λειτουργίας του συστήματος

4.2. Βάση Δεδομένων

Η βάση δεδομένων που δημιουργήθηκε έχει σκοπό την αποθήκευση των δεδομένων και στατιστικών στοιχείων για τις διαφημίσεις που έχουν καταχωρηθεί. Όταν δίνεται μια εικονοσειρά προς ταυτοποίηση, ελέγχεται η βάση για την εύρεση κάποιας καταχωρημένης διαφήμισης στην εικονοσειρά εισόδου. Για να επιτευχθούν τα παραπάνω δημιουργήθηκαν 2 πίνακες, οι “ads” και “frames”. Ο πίνακας ads έχει τα εξής πεδία:

- **ID**

Το πεδίο αυτό αποτελεί και το πρωτεύον κλειδί του πίνακα μιας και είναι μια ξεχωριστή για κάθε διαφήμιση τιμή, που δίνουμε όταν πρώτο-δημιουργούμε την εγγραφή μιας διαφήμισης.

- **Name**

Στο πεδίο αυτό σημειώνεται το όνομα της διαφήμισης ή της εταιρίας στην οποία αναφέρεται.

- **Channel**

Το κανάλι από το οποίο προβλήθηκε. Παρόλο που ο αλγόριθμος ταυτοποίησης δεν εξετάζει το κανάλι προβολής και μπορεί να μην υπάρχει, είναι μια πληροφορία χρήσιμη κυρίως για τον έλεγχο περιπτώσεων όπου προβάλλεται η διαφήμιση από διαφορετικό κανάλι από αυτό που υπάρχει στη βάση.

- **Seconds**

Η διάρκεια της διαφήμισης σε δευτερόλεπτα. Ουσιαστικά αποθηκεύει το αποτέλεσμα της διαίρεσης του συνολικού αριθμού των καρέ της διαφήμισης με το 25 που είναι ο αριθμός των καρέ ανά δευτερόλεπτο του PAL σήματος.

- **Frames**

Ο αριθμός των συνολικών καρέ για μια διαφήμιση.

- **PricePerSec**

Η τιμή χρέωσης ανά δευτερόλεπτο προβολής.

- **FirstFrame**

Το αναγνωριστικό id του πίνακα frames της βάσης για το πρώτο καρέ που αντιστοιχεί στη συγκεκριμένη διαφήμιση. Χρησιμοποιεί στην υλοποίηση του αλγορίθμου και αποτελεί εξωτερικό κλειδί (foreign key) από τον πίνακα frames.

- **LastFrame**

Το αναγνωριστικό id του πίνακα frames της βάσης για το τελευταίο καρέ που αντιστοιχεί στη συγκεκριμένη διαφήμιση. Επίσης, αποτελεί εξωτερικό κλειδί από τον πίνακα frames.

Ο πίνακας frames έχει τα πεδία:

- **ID**

Το αναγνωριστικό του κάθε καρέ. Αποτελεί και το πρωτεύον κλειδί του πίνακα.

- **Ads_id**

Το αναγνωριστικό (id) της διαφήμισης η οποία περιέχει το καρέ. Είναι εξωτερικό κλειδί από τον πίνακα ads.

- **Name**

Περιέχει το όνομα του αρχείου εικόνας που αναφέρεται στο καρέ.

- **Stats**

Το πεδίο αυτό έχει τα δεδομένα μετά την εφαρμογή του αλγορίθμου Επαγωγικής Σάρωσης. Είναι το πιο σημαντικό πεδίο μιας και έχει όλα τα δεδομένα πάνω στα οποία γίνεται η σύγκριση δύο καρέ μέσω του υπολογισμού του μέσου τετραγωνικού σφάλματος. Η μορφή των δεδομένων είναι ένας αριθμός για κάθε μπλοκ, που δηλώνει τον αριθμό των ακμών της μάσκας στην αντίστοιχη περιοχή.

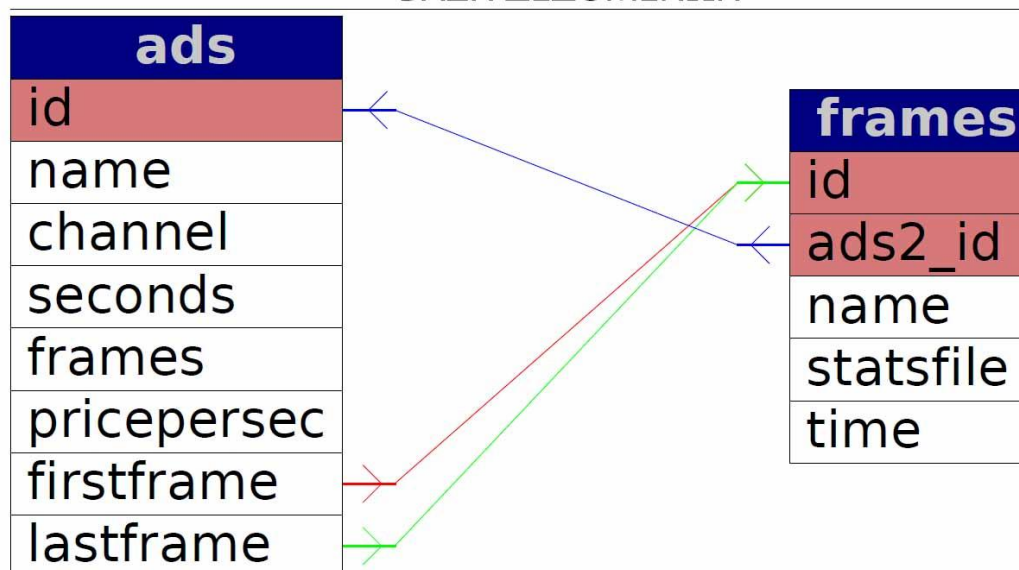
- **Time**

Η χρονοσφραγίδα (timestamp) του κάθε καρέ. Ο αριθμός αυτός χρησιμοποιείται για τον προσδιορισμό του χρονικού σημείου στο οποίο βρίσκεται κάποιο καρέ σε σχέση με τα υπόλοιπα.

Βάσει των παραπάνω λοιπόν, ο πίνακας ads περιέχει εγγραφές για τις διαφημίσεις ενώ ο πίνακας frames έχει τα καρέ όλων των διαφημίσεων. Με τα εξωτερικά κλειδιά (ads_id) μπορεί να γίνει η αντιστοίχιση του κάθε καρέ με τη διαφήμιση στην οποία ανήκει. Να σημειωθεί ακόμα, πως η χρήση των εξωτερικών κλειδιών βοηθά και στο να είναι πιο σωστά δομημένη η βάση και όταν για παράδειγμα διαγραφεί μια διαφήμιση από τον πίνακα ads, τότε να σβήνονται και οι εγγραφές του πίνακα frames των καρέ που αντιστοιχούν στην διαφήμιση αυτή.

[illegible]

ΒΑΣΗ ΔΕΔΟΜΕΝΩΝ



Εικόνα 4.2.3. Το σχεσιακό διάγραμμα (schema) της βάσης δεδομένων

4.3. Επεξεργασία Εικόνας

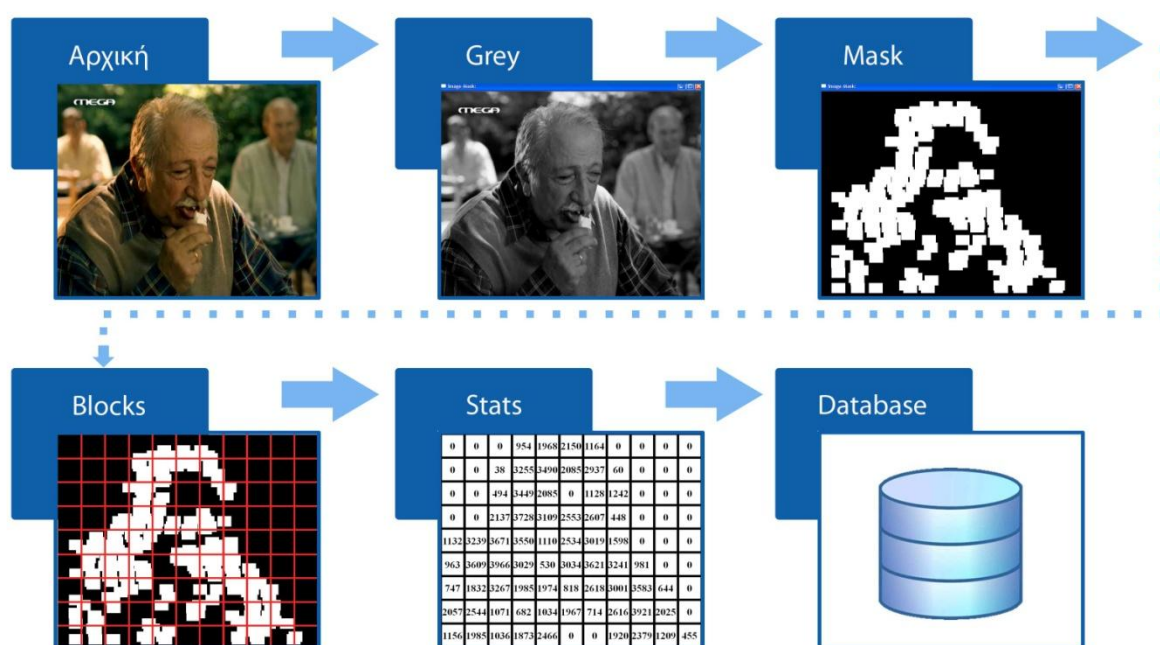
Σε αυτό το σημείο θα παρουσιαστούν οι βασικές πληροφορίες σχετικά με την επεξεργασία των εικόνων και τον τρόπο εξαγωγής των περιοχών ενδιαφέροντος (Regions of Interest) βάσει περιγραφέα ακμών (edge descriptor).

Αρχικά, η εικόνα φορτώνεται όχι σε έγχρωμη αλλά σε ασπρόμαυρη μορφή, με αποχρώσεις του γκριζου βάθους 8-bit. Αυτό γίνεται γιατί η μετέπειτα επεξεργασία της εικόνας αρκεί να γίνει σε ένα κανάλι και όχι και στα 3 κανάλια RGB (Red Green Blue) όπως είναι αρχικά.

Επόμενο βήμα είναι η εύρεση των ακμών της εικόνας, αφού αυτό θα είναι το κριτήριο σύμφωνα με το οποίο θα εξετάζεται αν υπάρχει ταύτιση των καρέ και κατ' επέκταση των διαφημίσεων. Σε αυτό το σημείο λοιπόν γίνεται χρήση του αλγορίθμου Επαγωγικής Σάρωσης ο οποίος επιλέγεται αντί της χρήσης κάποιου έτοιμου αλγορίθμου λόγω της ικανότητας του να βρίσκει γρήγορα της ακμές μιας εικόνας παραλείποντας τα κομμάτια που δεν περιέχουν πληροφορία. Τα βασικά σημεία της υλοποίησής του περιγράφονται στη συνέχεια.

Το παράθυρο κύλισης ορίζεται να είναι ίσο με 37x28 δηλαδή 37 εικονοστοιχεία (pixels) πλάτος και 28 εικονοστοιχεία ύψος. Παράλληλα έχουν οριστεί και οι απαραίτητες παράμετροι για την Επαγωγική Σάρωση, δηλαδή η μέγιστη ταχύτητα ίση με 20 εικονοστοιχεία και η αρχική ταχύτητα ίση με 37 εικονοστοιχεία, όσο και το πλάτος του παραθύρου κύλισης. Στο παράθυρο που ορίζεται κάθε φορά εφαρμόζεται η μέθοδος “Canny” [Canny, 2010] για ανίχνευση ακμών. Αν το αποτέλεσμα περιέχει πολλές ακμές τότε το παράθυρο που ορίζεται από το κεντρικό του εικονοστοιχείο παίρνει άσπρο χρώμα στην δυαδική μάσκα, αλλιώς παραμένει μαύρο.

Έπειτα, το αποτέλεσμα του αλγορίθμου επαγωγικής σάρωσης, δηλαδή η δυαδική μάσκα με την πληροφορία για τις ακμές, χωρίζεται σε 99 μπλοκ. Χωρίζοντας την εικόνα σε $99=11 \times 9$ τμήματα επιτυγχάνεται τα τελικά δεδομένα να είναι ένας πίνακας ακεραίων 11×9 όπου κάθε αριθμός υποδηλώνει τον αριθμό των ακμών ανά καρέ, ή αλλιώς πόσα εικονοστοιχεία στο αντίστοιχο τμήμα της εικόνας είναι λευκά αντί για μαύρα. Τέλος, να σημειωθεί πως επειδή ο αλγόριθμος πρέπει να είναι ανεξάρτητος και ανεπηρέαστος από το κανάλι προβολής, παραλείπεται ένα μέρος από το πάνω αριστερά κομμάτι της εικόνας μιας και σε αυτό το σημείο εμφανίζεται το σήμα του καναλιού το οποίο προβάλλει τη διαφήμιση.

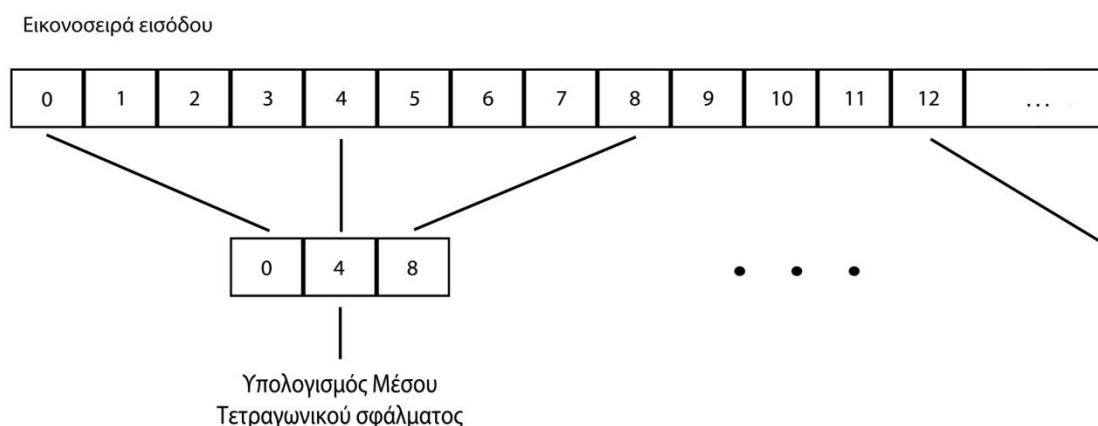


Σχήμα 4.3.1. Τα βήματα του αλγορίθμου για την αποθήκευση της πληροφορίας μιας εικόνας εισόδου

4.4. Σύστημα ταυτοποίησης εικονοσειρών τηλεοπτικού σήματος

Είναι επίσης σημαντική και η βασική δομή του υπόλοιπου προγράμματος καθώς και ο τρόπος υλοποίησης των αλγορίθμων που παρουσιάστηκαν στα προηγούμενα κεφάλαια.

Αρχικά, κατά την ανάλυση της εικονοσειράς εισόδου δε διαβάζονται όλα τα καρέ που περιέχει. Σύμφωνα με τα τρέχοντα τεχνολογικά δεδομένα του σήματος PAL στην τηλεόραση, η προβολή γίνεται με ρυθμό 25 καρέ το δευτερόλεπτο [PAL, 2010]. Αυτό σημαίνει πως σε μια ακολουθία εικόνων οι αλλαγές από το ένα καρέ στο επόμενο είναι πάρα πολύ μικρές. Έτσι, εφαρμόζεται δειγματοληψία με σταθερό ρυθμό 6 καρέ ανά δευτερόλεπτο στην επεξεργασία των εικόνων. Επιπλέον, το χρονικό παράθυρο κύλισης ορίζεται ίσο με 3 καρέ. Τα παραπάνω οδηγούν τη δειγματοληψία να έχει τη μορφή που απεικονίζεται στο ακόλουθο σχήμα :

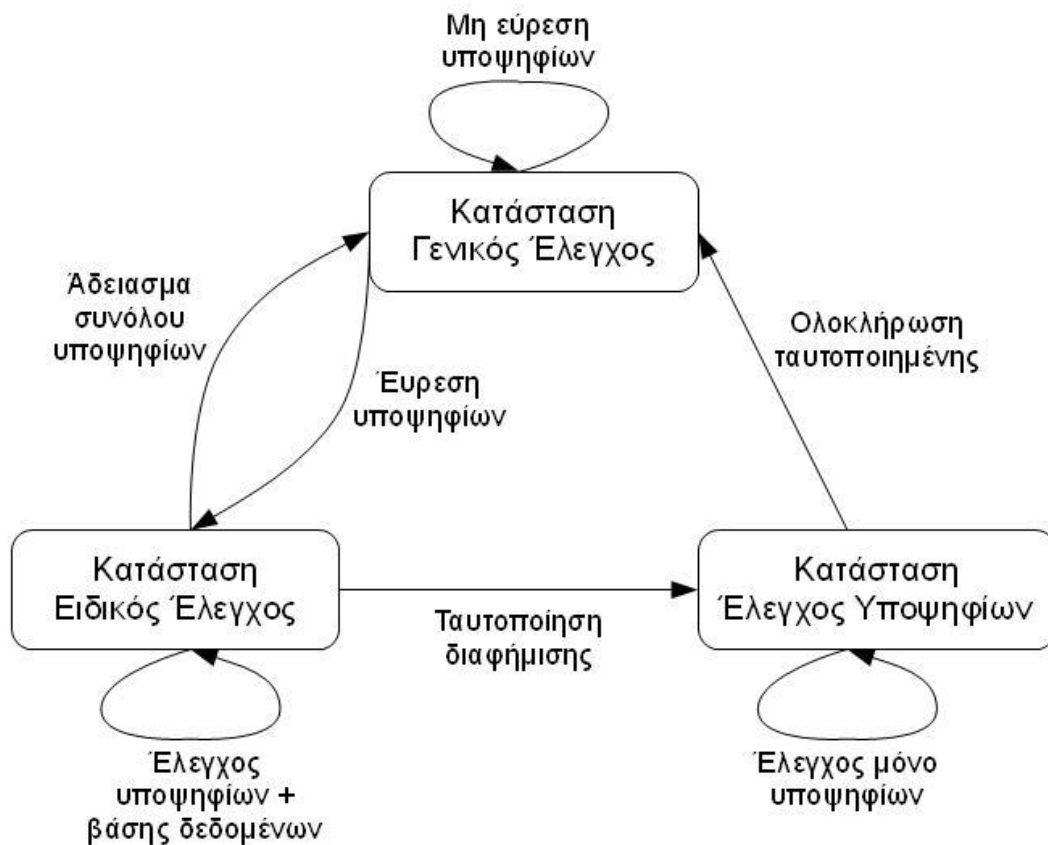


Σχήμα 4.4.1. Δειγματοληψία για τα καρέ εισόδου

Ο τρόπος με τον οποίο ταυτοποιείται μια διαφήμιση γίνεται με τα σύνολα αληθών και ψευδών υποψηφίων όπως παρουσιάστηκαν στο Κεφάλαιο 4. Να σημειωθεί μόνο πως για να αποσυρθεί μια διαφήμιση από το σύνολο των ψευδών υποψηφίων δεν αρκεί να αποτύχει στον έλεγχο του μέσου τετραγωνικού σφάλματος μια φορά αλλά συνεχόμενες φορές. Ακόμα, μια διαφήμιση ταυτοποιείται αν έχουμε βρει τουλάχιστον τα μισά καρέ της στην ακολουθία.

Όσον αφορά στις διαφημίσεις με τις οποίες ελέγχεται και συγκρίνεται η εικονοσειρά εισόδου, ο έλεγχος με όλα τα πιθανά καρέ είναι αρκετά χρονοβόρος αλλά και άσκοπος. Στην αρχή και στη γενική περίπτωση που οι λίστες με τις υποψηφίες διαφημίσεις είναι άδειες, διεξάγεται έλεγχος σε όλα τα καρέ που είναι στα πρώτα δευτερόλεπτα μιας διαφήμισης από τη βάση. Αν υπάρχουν μερικές υποψηφίες διαφημίσεις αλλά καμία εξ' αυτών δεν είναι ταυτοποιημένη, τότε γίνεται έλεγχος στα καρέ τους που περιέχονται σε ένα χρονικό παράθυρο γύρω από τη χρονοσφραγίδα του τελευταίου ταυτοποιημένου καρέ. Παράλληλα, σε αυτή τη περίπτωση, συνεχίζεται ο έλεγχος με τις υποψηφίες διαφημίσεις από τη βάση αλλά με μικρότερο αντίστοιχο χρονικό παράθυρο. Ο λόγος που γίνεται αυτό είναι για να αντιμετωπιστεί το ενδεχόμενο μετά από ένα άγνωστο κομμάτι της εικονοσειράς να ακολουθεί μια καταχωρημένη διαφήμιση, η οποία χωρίς αυτό τον έλεγχο δε θα μπορούσε να εισέλθει στο σύνολο των πιθανών υποψηφίων και τελικώς να ταυτοποιηθεί. Τέλος, στην περίπτωση που ο αλγόριθμος έχει βρει πάνω από τα μισά καρέ μιας διαφήμισης, τότε τη θεωρεί σωστά ταυτοποιημένη και συνεχίζει μέχρι αυτή να τελειώσει, ελέγχοντας πλέον μόνο τις διαφημίσεις που είναι στα σύνολα με τις υποψηφίες. Συνοπτικά το πρόγραμμα λειτουργεί στις εξής 3 καταστάσεις :

- *Κατάσταση Γενικός Έλεγχος Βάσης* : Αρχικοποίηση παραθύρων κύλισης στο χρόνο και έλεγχος όλων των αρχικών καρέ της βάσης εντός των παραθύρων αυτών.
- *Κατάσταση Ειδικός Έλεγχος Υποψηφίων και Βάσης* : Αρχικοποίηση αυστηρότερων παραθύρων κύλισης και έλεγχος όλων των αρχικών καρέ της βάσης καθώς και των υποψηφίων σύμφωνα με τα δεδομένα για τη χρονική στιγμή.
- *Κατάσταση Έλεγχος Υποψηφίων* : Έλεγχος μόνο των καρέ των υποψηφίων διαφημίσεων εντός των παραθύρων κύλισης.



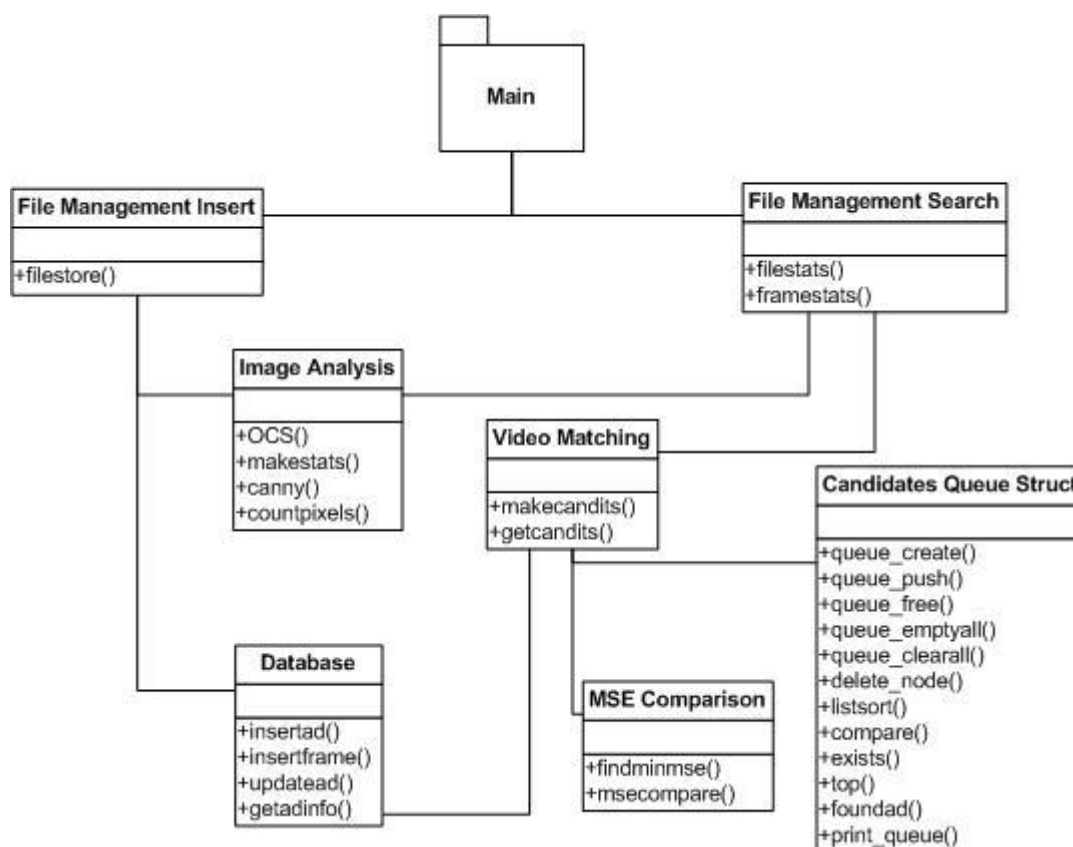
Σχήμα 4.4.2. Διάγραμμα καταστάσεων του προγράμματος

Τελικά, όταν ολοκληρωθεί η ταυτοποιημένη διαφήμιση ελέγχουμε ένα μετρητή που έχει οριστεί για να υπολογίσει τη διάρκεια που έπαιξε σε σχέση με τη συνολική διάρκεια της διαφήμισης όπως αυτή είναι καταχωρημένη στη βάση. Σε περίπτωση που η διαφορά της διάρκειας είναι μεγαλύτερη από 1,5 δευτερόλεπτα, τότε εμφανίζεται κατάλληλο προειδοποιητικό μήνυμα για πιθανή αλλοίωση της διαφήμισης. Όσον αφορά στα κομμάτια που δεν έχουν ταυτοποιηθεί, αν η διάρκεια τους υπερβαίνει τα 2,5 δευτερόλεπτα τότε χαρακτηρίζονται ως άγνωστο κομμάτι της εικονοσειράς πάλι με την εμφάνιση κατάλληλου μηνύματος.

4.5. Περιγραφή κώδικα του προγράμματος

Ο κώδικας του προγράμματος μπορεί να χωριστεί σε δύο λειτουργίες εκτέλεσης. Την επεξεργασία εικονοσειρών για την αποθήκευσή τους στη βάση δεδομένων και την

επεξεργασία εικονοσειρών με σκοπό την ταύτιση τους με τις αποθηκευμένες εικονοσειρές. Ακολουθεί το UML διάγραμμα κλάσεων του κώδικα που εκτελείται.



Σχήμα 4.5.1. Διάγραμμα κλάσεων του συστήματος

Όπως φαίνεται στο Σχήμα 4.5.1. οι συναρτήσεις που χρησιμοποιούνται είναι χωρισμένες ανάλογα με τη λειτουργία που επιτελούν. Αναλυτικότερα, οι συναρτήσεις μπορούν να ταξινομηθούν σύμφωνα με το παραπάνω σχήμα ως εξής :

- **Επεξεργασία Εικόνας (Image Analysis)**

Σε αυτή τη κατηγορία εντάσσονται οι συναρτήσεις οι οποίες έχουν να κάνουν με την επεξεργασία των καρτέ και των χαρακτηριστικών τους.

- **OCS ()**

Η συνάρτηση υλοποίησης του αλγορίθμου Επαγωγικής Σάρωσης. Δέχεται ως είσοδο μια εικόνα και επιστρέφει μια δυαδική μάσκα με τη πληροφορία για τις περιοχές ενδιαφέροντος.

- **makestats()**

Δημιουργεί τον πίνακα των δεδομένων για τις ακμές βασισμένο στην πληροφορία της δυαδικής μάσκας.

- **canny()**
Υλοποιεί τον αλγόριθμο Canny για εύρεση ακμών σε μια εικόνα. Χρησιμοποιείται από την συνάρτηση OCS ως περιγραφέας ακμών.
- **countpixels()**
Βοηθητική συνάρτηση η οποία μετρά τον αριθμό των λευκών εικονοστοιχείων σε μια περιοχή της εικόνας.
- **Βάση δεδομένων (Database)**
Οι συναρτήσεις της κατηγορίας αυτής βοηθούν στην επικοινωνία με τη βάση δεδομένων με σκοπό την εισαγωγή δεδομένων ή άντληση πληροφοριών. Για τη σύνδεση με τη βάση χρησιμοποιούνται επιπλέον και οι συναρτήσεις της βιβλιοθήκης της MySQL.
 - **insertad ()**
Η συνάρτηση αυτή εισάγει μια νέα εγγραφή στον πίνακα ads της βάσης που αντιστοιχεί στην εικονοσειρά μιας νέας διαφήμισης.
 - **insertframe()**
Η συνάρτηση αυτή εισάγει μια νέα εγγραφή στον πίνακα frames της βάσης που αντιστοιχεί σε ένα νέο καρέ μιας διαφήμισης.
 - **updatead ()**
Μετά την ολοκλήρωση της εισαγωγής όλων των καρέ για μια διαφήμιση αυτή η συνάρτηση καλείται για να ενημερώσει τα υπόλοιπα πεδία στατιστικών του πίνακα ads.
 - **getadinfo**
Βοηθητική συνάρτηση που επιστρέφει τον αριθμό των συνολικών καρέ για μια διαφήμιση. Χρησιμοποιείται από τη makecandits() στη ταύτιση εικονοσειρών.
- **Ταύτιση εικονοσειρών (video matching)**
Για την ταύτιση των εικονοσειρών εφαρμόζονται οι αλγόριθμοι που περιγράφηκαν στο Κεφάλαιο 4.4. Αυτοί περιλαμβάνουν τους ελέγχους καταστάσεων του προγράμματος, την εύρεση υποψηφίων διαφημίσεων και τους υπολογισμούς για το μέσο τετραγωνικό σφάλμα καθώς και για την ακολουθία των χρονοσφραγίδων.
 - **makecandits ()**
Με τη συνάρτηση αυτή δημιουργούνται οι υποψήφιες διαφημίσεις και αρχικοποιούνται κατάλληλα όλα τα πεδία της αντίστοιχης δομής δεδομένων που έχει οριστεί.
 - **getcandits ()**
Η getcandits διαχειρίζεται τις λίστες Αληθών και Ψευδών Υποψηφίων και είναι υπεύθυνη για τον έλεγχο του αποτελέσματος του μέσου τετραγωνικού σφάλματος. Εδώ γίνεται η ταύτιση των εικονοσειρών.
- **Δομή Δεδομένων διαφημίσεων (candidates queue struct)**
Για το χειρισμό των διαφημίσεων και όλων των δεδομένων που τις αφορούν ήταν αναγκαία η δημιουργία μιας νέας δομής που να εξυπηρετεί τις ανάγκες αυτές. Βασικό στοιχείο είναι η δομή queueNode η οποία αντιστοιχεί σε μια οντότητα διαφημιστικής εικονοσειράς και έχει τα εξής πεδία :

- **int id;**
Το αναγνωριστικό id της διαφήμισης στον πίνακα ads της βάσης δεδομένων.
- **bool found;**
Αληθές αν έχει ταυτιστεί η διαφήμιση και ψευδές διαφορετικά.
- **int start;**
Αποθηκεύει το σημείο έναρξης της διαφήμισης.
- **int counter;**
Μετρητής της διάρκειας για την οποία η διαφήμιση ήταν στο σύνολο Αληθών υποψηφίων.
- **int falsecounter;**
Μετρητής της διάρκειας για την οποία η διαφήμιση ήταν στο σύνολο Ψευδών υποψηφίων. Σε συνδυασμό με τον μετρητή counter μπορεί να υπολογιστεί η συνολική διάρκεια προβολής της διαφήμισης.
- **int totalfrm;**
Ο συνολικός αριθμός των καρέ.
- **double mse;**
Το άθροισμα των τιμών για το μέσο τετραγωνικό σφάλμα καθενός καρέ της διαφήμισης.
- **int timestamp;**
Η χρονοσφραγίδα του πιο πρόσφατα ταυτισμένου καρέ.
- **int state;**
Ανάλογα με τη τιμή αυτής της μεταβλητής τα κατώφλια για το μέσο τετραγωνικό σφάλμα και τις χρονοσφραγίδες ορίζονται κατάλληλα.
- **int * times;**
Πίνακας που αποθηκεύει τις χρονοσφραγίδες των καρέ που έχουν ήδη ταυτιστεί ώστε να μην γίνεται εκ νέου έλεγχος για αυτά τα καρέ.
- **int toptms;**
Βοηθητικός μετρητής των φορών που η διαφήμιση βρισκόταν στην κορυφή της λίστας των υποψηφίων διαφημίσεων.
- **struct queueitem *next;**
Δείκτης στο κόμβο της επόμενης διαφήμισης που ακολουθεί στο αντίστοιχο σύνολο των υποψηφίων.

Για τη διευκόλυνση στον χειρισμό των ανωτέρω δεδομένων έχουν οριστεί οι ακόλουθες συναρτήσεις :

- **queue_create ()**
Συνάρτηση για τη δημιουργία των συνόλων Αληθών κι Ψευδών υποψηφίων.
- **queue_push ()**
Εισάγει τον κόμβο μιας διαφήμισης σε μία λίστα υποψηφίων.
- **queue_emptyall ()**
Διαγράφει όλους τους κόμβους από μια λίστα.
- **queue_clearall ()**
Διαγράφει όλους τους κόμβους που δεν πληρούν τις κατάλληλες προϋποθέσεις ως προς της διάρκεια.

- **delete_node ()**
Διαγράφει το κόμβο ο οποίος έχει αναγνωριστικό id ίσο με την είσοδο της συνάρτησης.
- **listsort ()**
Συνάρτηση ταξινόμησης των κόμβων.
- **compare ()**
Το κριτήριο ταξινόμησης για την παραπάνω συνάρτηση.
- **exists ()**
Η αναζητά μια διαφήμιση με ένα δοσμένο αναγνωριστικό σε ένα σύνολο υποψηφίων διαφημίσεων.
- **top ()**
Η αναζητά μια διαφήμιση με κριτήριο την ύπαρξη στη κορυφή της λίστας των υποψηφίων μετά την ταξινόμηση.
- **print_queue ()**
Βοηθητική συνάρτηση για την τύπωση στην οθόνη όλων των διαφημίσεων σε κάποιο σύνολο υποψηφίων και των στοιχείων τους.
- **Υπολογισμός μέσου τετραγωνικού σφάλματος (mse comparison)**
Σε αυτήν την κατηγορία εντάσσονται οι συναρτήσεις που συγκρίνουν την εικονοσειρά εισόδου με τις αποθηκευμένες στη βάση για τον υπολογισμό του μέσου τετραγωνικού σφάλματος.
 - **findminmse ()**
Η συνάρτηση αυτή είναι υπεύθυνη για την εύρεση του καρέ με το χαμηλότερο μέσο τετραγωνικό σφάλμα για μια διαφήμιση. Εδώ εφαρμόζεται η τεχνική με τα χρονικά παράθυρα στο σύνολο από καρέ (group of frames).
 - **msecompare ()**
Ο υπολογισμός του μέσου τετραγωνικού σφάλματος για σύμφωνα με τον τύπο (3) του Κεφαλαίου 3.2.
- **Διαχείριση αρχείων (File Management)**
Η κατηγορία αυτή περιλαμβάνει τις συναρτήσεις που διαχειρίζονται τους φακέλους με περιεχόμενο την ζητούμενη εικονοσειρά. Η ανάγνωση των εικόνων γίνεται ομαλά και διαδοχικά μεταξύ των διαφορετικών φακέλων. Εδώ πραγματοποιείται η κλήση των συναρτήσεων για την αποθήκευση ή ταύτιση της εικονοσειράς.
 - **filestore ()**
Η κατάλληλη συνάρτηση για την αποθήκευση μιας εικονοσειράς.
 - **filestats ()**
Η διαχείριση των φακέλων κατά τη λειτουργία ταύτισης της εικονοσειράς εκτελείται από αυτή τη συνάρτηση.
 - **framestats ()**
Η συνάρτηση αυτή καλείται από τη συνάρτηση filestats και είναι υπεύθυνη για το διάβασμα των καρέ από το πρόγραμμα. Επίσης, μετά το διάβασμα των καρέ καλεί τις αντίστοιχες συναρτήσεις των κατηγοριών για την επεξεργασία εικόνας και την ταύτιση της εικονοσειράς.

- **Κυρίως Πρόγραμμα (main)**

Στο κυρίως πρόγραμμα ο χρήστης εισάγει 2 εισόδους. Η πρώτη αφορά στη διεύθυνση στον υπολογιστή που περιέχει τα αρχεία της εικονοσειράς εισόδου. Η δεύτερη αναφέρεται στην επιλογή της λειτουργίας που θα εκτελέσει για την εικονοσειρά που δόθηκε. Οι διαθέσιμες λειτουργίες είναι αποθήκευση στη βάση δεδομένων και προσπάθεια ταύτισης της εικονοσειράς.

5. Παράδειγμα Λειτουργίας Συστήματος

Βλέποντας ένα παράδειγμα λειτουργίας του προγράμματος, μπορεί να γίνει πιο εύκολα κατανοητός ο τρόπος με τον οποίο λειτουργεί ο αλγόριθμος Επαγωγικής Σάρωσης, αλλά και ο τρόπος εύρεσης και ταυτοποίησης τελικά της σωστής διαφήμισης.

5.1. Παράδειγμα τυπικής λειτουργίας

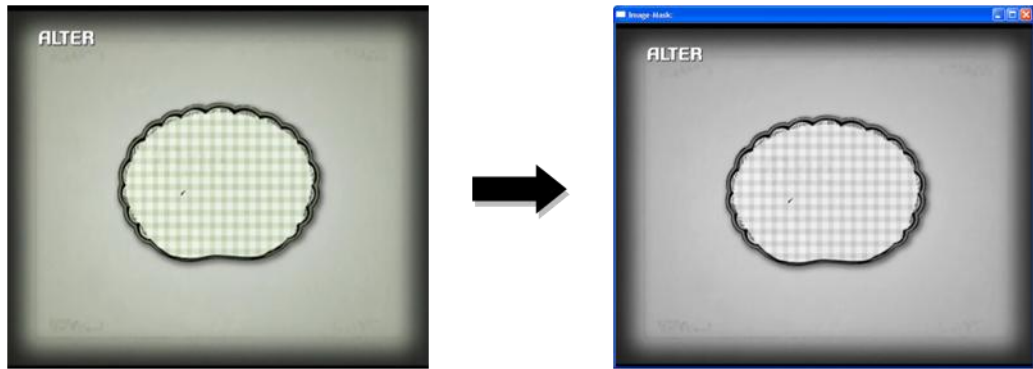
Στο παράδειγμα που εκτελέσαμε, ζητούμενη εικονοσειρά είναι η διαφήμιση για ένα προϊόν στη κατηγορία τροφίμων, της οποίας μια απλή ακολουθία εικόνων είναι:



Εικόνα 5.1.1. Θεματική ακολουθία εικόνων (storyboard) της αρχικής διαφήμισης

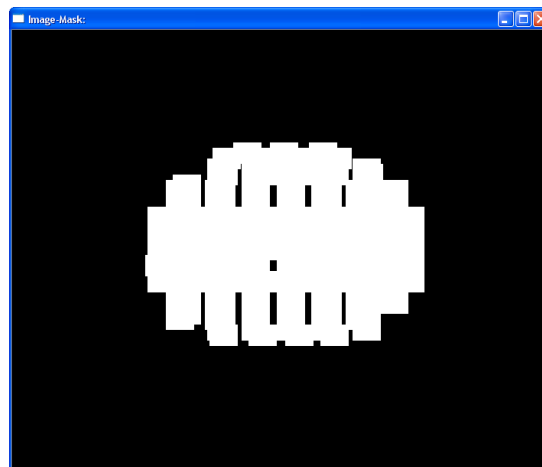
Συνολικά η διαφήμιση έχει 461 καρέ και δεδομένου ότι ένα δευτερόλεπτο αντιστοιχεί σε 25 καρέ, διαρκεί 18,4 δευτερόλεπτα. Στη βάση δεδομένων είναι αποθηκευμένα δεδομένα για 116 καρέ αυτής, αφού όπως έχει αναφερθεί η δειγματοληψία γίνεται κάθε 4 καρέ ή αλλιώς 6 καρέ το δευτερόλεπτο (συνεπώς — καρέ).

Ξεκινώντας ο αλγόριθμος μετατρέπει κάθε εικόνα που διαβάζει από την ακολουθία σε ασπρόμαυρη με αποχρώσεις γκριζου βάθους 8-bit, για να την επεξεργαστεί.



Εικόνα 5.1.2. Αποθήκευση εικόνας εισόδου από έγχρωμη σε ασπρόμαυρη για την επεξεργασία

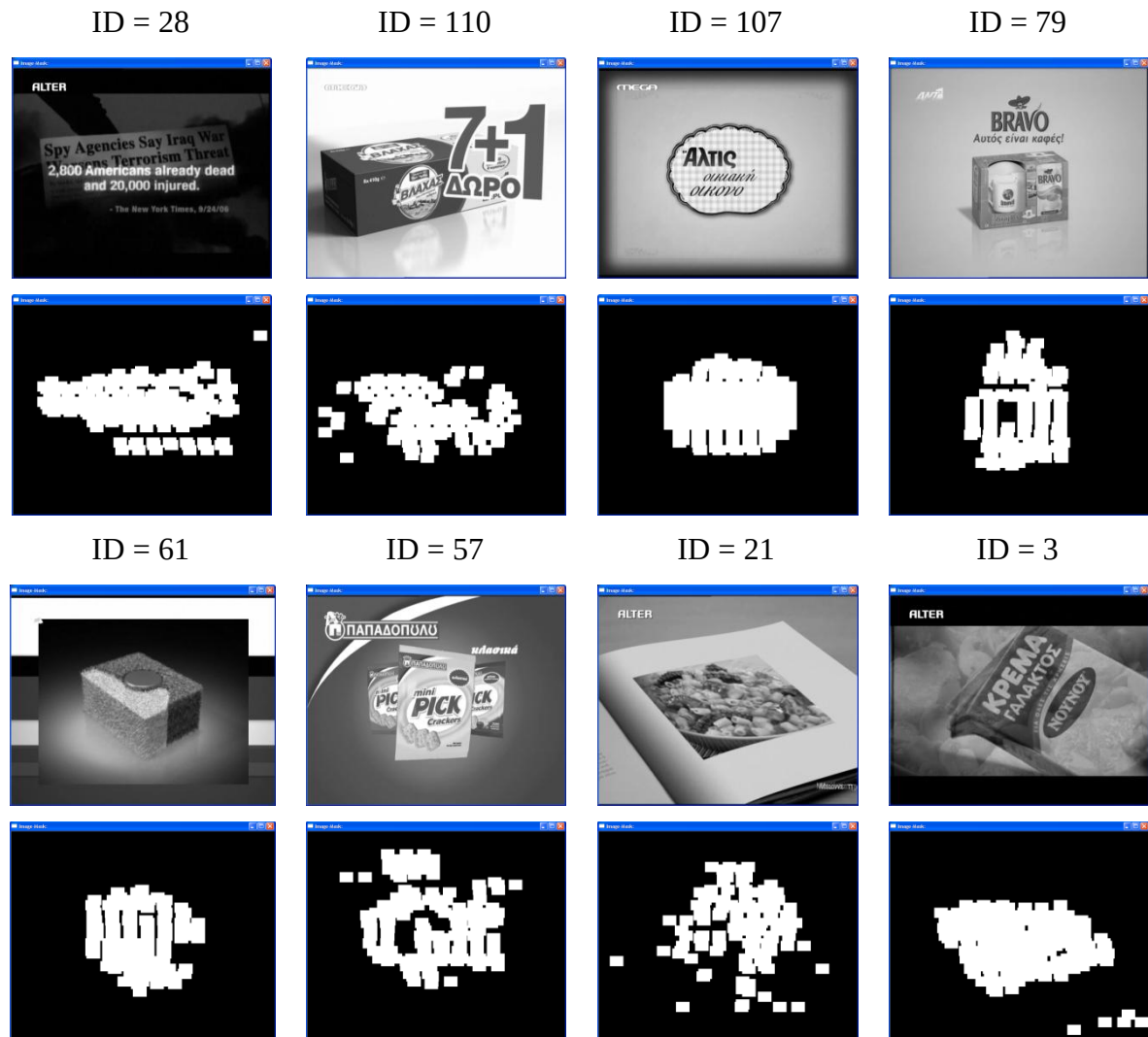
Έπειτα, εκτελείται ο αλγόριθμος επαγωγικής σάρωσης ώστε να βρεθούν οι περιοχές που περιέχουν πληροφορία, δηλαδή οι περιοχές με μεγάλο πλήθος ακμών. Μετά την εύρεση αυτών των περιοχών, σχηματίζεται η δυαδική μάσκα (binary mask) της αρχικής εικόνας, η οποία έχει την πληροφορία που επέστρεψε ο αλγόριθμος επαγωγικής σάρωσης. Η μάσκα για την παραπάνω εικόνα ακολουθεί.



Εικόνα 5.1.3. Δυαδική μάσκα της εικόνας εισόδου, μετά την εφαρμογή του αλγορίθμου επαγωγικής σάρωσης

Παρατηρεί κανείς πως ο αλγόριθμος σωστά απομόνωσε το κεντρικό κομμάτι της εικόνας, αφού σε εκείνο το σημείο έχουμε μεγάλο πλήθος ακμών. Επίσης, να σημειωθεί πως το κομμάτι όπου εμφανίζεται το κανάλι που προβάλλει τη διαφήμιση σωστά έχει παραλειφθεί από την μάσκα.

Έπειτα ο αλγόριθμός ανατρέχει στη βάση δεδομένων για να βρει ποιες διαφημίσεις είναι υποψήφιες για να έχουν δώσει την παραπάνω μάσκα. Ο αλγόριθμος βρίσκει 8 υποψήφιες διαφημίσεις με τα εξής αναγνωριστικά :



Εικόνα 5.1.4. Χαρακτηριστικά καρέ υποψήφιων διαφημίσεων

Η επιλογή αυτών των υποψηφίων έγινε γιατί όπως φαίνεται στις φωτογραφίες, είναι διαφημίσεις που η μάσκα τους έχει αρκετές ομοιότητες με αυτή της διαφήμισης που αναζητούμε, δηλαδή συγκεντρωμένη πληροφορία στο κεντρικό μέρος της εικόνας.

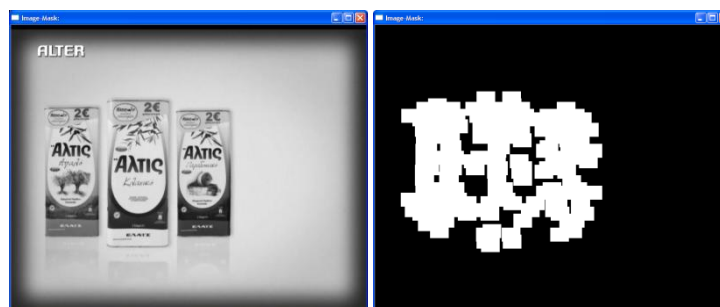
Πιο συγκεκριμένα, υπολογίστηκε η τιμή του μέσου τετραγωνικού σφάλματος της κάθε φωτογραφίας σε σχέση με την αρχική, παίρνοντας τα ακόλουθα αποτελέσματα.

Πίνακας 5.1.1. Τιμές για το μέσο τετραγωνικό σφάλμα των υποψήφιων ως προς την αρχική διαφήμιση

Ad - ID	28	110	107	79	61	57	21	3
MSE	259.892	233.050	15.618	232.940	163.048	243.635	247.376	252.423

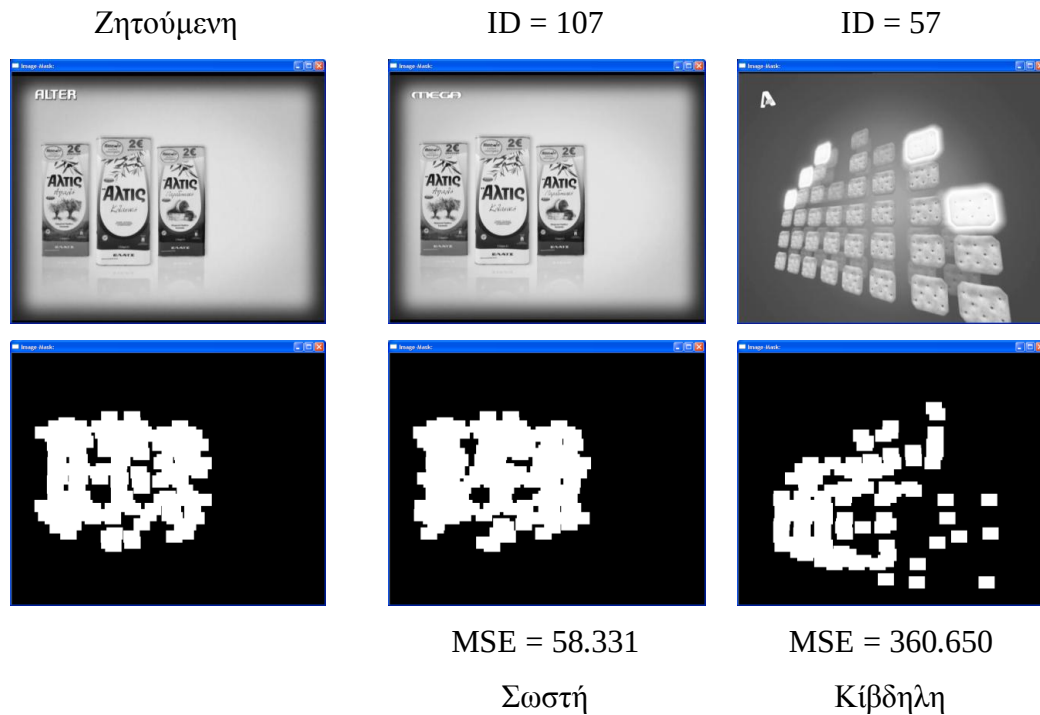
Προφανώς η σωστή διαφήμιση είναι αυτή με αναγνωριστικό από τη βάση δεδομένων id = 107, στοιχείο που φαίνεται και από την πολύ μικρή τιμή στο μέσο τετραγωνικό σφάλμα.

Ο αλγόριθμος στη συνέχεια εφαρμόζει την ίδια διαδικασία στα επόμενα καρέ της διαφήμισης ελέγχοντας αν οι υποψήφιες ως τώρα διαφημίσεις συνεχίζουν να μην αποκλίνουν από την αρχική. Τελικά, μετά από 145 καρέ και ενώ μέχρι τότε δεν υπήρχε σημαντική διαφοροποίηση των εικόνων μεταξύ τους, στην αρχική μας διαφήμιση αλλάζει το πλάνο και γίνεται:



Εικόνα 5.1.5. Καρέ αλλαγής πλάνου στην αρχική διαφήμιση και η αντίστοιχη μάσκα

Τότε όλες οι υποψήφιες, εκτός της σωστής με id=107, εμφανίζουν μάσκα που αποκλίνει από την ζητούμενη τόσο ώστε να ξεπερνάει το κατώφλι (threshold) που έχουμε θέσει, με αποτέλεσμα να μετακινούνται στο σύνολο Ψευδών Υποψηφίων (False), όπου μετά την παραμονή τους εκεί για 3 συνεχόμενες φορές, διαγράφονται εντελώς από την λίστα με τις υποψήφιες.



Εικόνα 5.1.6. Χαρακτηριστικές εικόνες σωστής και λανθασμένης διαφήμισης από τις υποψήφιες, μαζί με τα αντίστοιχα σφάλματα

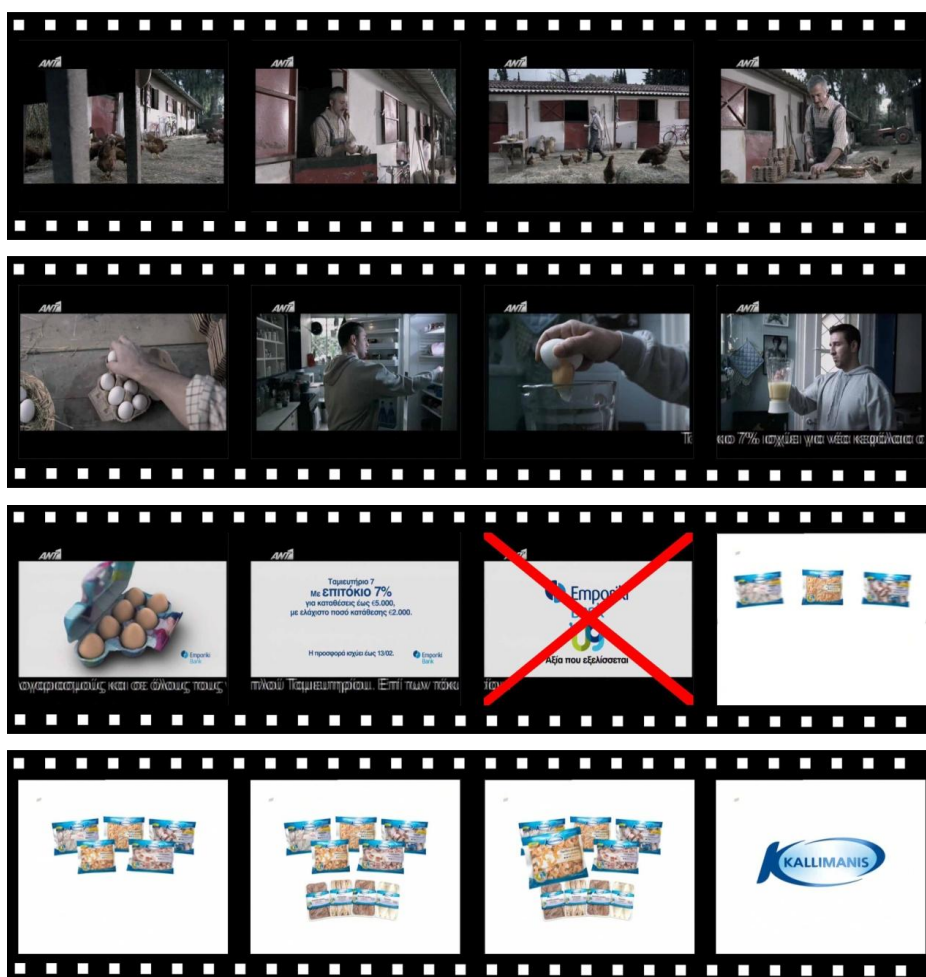
Σε αυτό το σημείο λοιπόν, μοναδική υποψήφια απομένει η διαφήμιση με αναγνωριστικό id=107, η οποία είναι και η σωστή. Έτσι, ο αλγόριθμός φθάνει τελικά μετά από 228 καρέ να ταυτοποιήσει πάνω από τα μισά καρέ αυτής της διαφήμισης. Τότε μετά από έλεγχο για το μέσο όρο του μέσου τετραγωνικού σφάλματος αλλά και της ακολουθίας των καρέ αν είναι χρονικά γραμμική, θεωρεί ότι η αρχική ζητούμενη εικονοσειρά πράγματι είναι η διαφήμιση με αναγνωριστικό id=107 της βάσης δεδομένων. Στη συνέχεια και μέχρι το τέλος δεν παρουσιάζεται κάποιο πρόβλημα, με αποτέλεσμα το πρόγραμμα να επιστρέφει το αποτέλεσμα

«Found Ad - ID = 107»

Τελικά λοιπόν, σωστά έγινε ταυτοποίηση της αρχικής δοσμένης εικονοσειράς με την καταχωρημένη στη βάση δεδομένων διαφήμιση με αναγνωριστικό id=107. Μάλιστα, η μέτρηση της διάρκειας επιστρέφει τιμή για τον αντίστοιχο μετρητή ίση με 114 που αντιστοιχεί σε $4 \times 114 = 456$ καρέ αντί για 461 που είναι καταχωρημένα στην βάση. Αυτή η διαφορά των 5 καρέ είναι πολύ μικρή και θεωρείται ότι η διαφήμιση έπαιξε ολόκληρη χωρίς πρόβλημα.

5.2 Παράδειγμα παθολογικής λειτουργίας

Σε αυτό το σημείο θα παρουσιαστεί ακόμα ένα παράδειγμα λειτουργίας του συστήματος σε μια πιο σύνθετη περίπτωση αυτή τη φορά. Η εικονοσειρά προς ταυτοποίηση είναι μια διαφήμιση τραπέζης ακολουθούμενη από μια διαφήμιση ενός προϊόντος τροφίμων. Στη διαφήμιση της τράπεζας έχουν αφαιρεθεί 100 καρτέ από το τέλος, δηλαδή είναι κομμένη κατά 4 δευτερόλεπτα. Τα βασικά πλάνα της ακολουθίας φαίνονται στην ακόλουθη εικόνα.



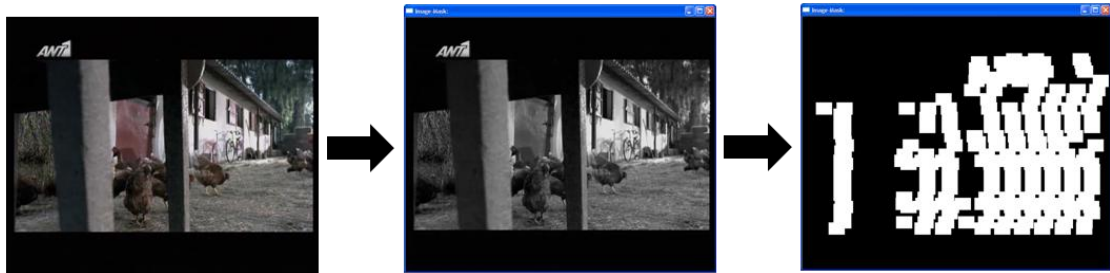
Εικόνα 5.2.1. Θεματική ακολουθία διαφημιστικού (storyboard)

Το κόκκινο σημάδι “X” υποδηλώνει ότι τα καρτέ που αναφέρονται στο συγκεκριμένο πλάνο είναι τα καρτέ που έχουν αφαιρεθεί.

Στην εικονοσειρά η πρώτη διαφήμιση περιέχει 659 καρτέ, ενώ στην ολοκληρωμένη μορφή της έχει 759 καρτέ (30,36 sec) 190 από τα οποία (αφού η δειγματοληψία γίνεται ανά 4 καρτέ) είναι αποθηκευμένα στη βάση δεδομένων με το αναγνωριστικό id=119. Η δεύτερη

διαφήμιση έχει 347 καρέ συνολικά (87 αποθηκευμένα στη βάση με αναγνωριστικό id=72) και εμφανίζεται ολόκληρη.

Όπως έχει αναφερθεί στο πρώτο παράδειγμα η επεξεργασία των εικόνων δεν γίνεται σε έγχρωμη μορφή, αλλά σε ασπρόμαυρη βάθους 8-bit και στην συνέχεια με την εφαρμογή του αλγορίθμου επαγωγικής σάρωσης γίνεται εξαγωγή της δυαδικής μάσκας βασισμένη στις ακμές που περιέχει η εικόνα.



Εικόνα 5.2.2. Παράδειγμα μετατροπής εικόνας εισόδου σε ασπρόμαυρη και εύρεση της αντίστοιχης δυαδικής μάσκας



Εικόνα 5.2.3. Χαρακτηριστικό στιγμιότυπο μετά από 36 καρέ των υποψηφίων διαφημίσεων από την βάση δεδομένων μαζί με τις μάσκες τους και το μέσο τετραγωνικό σφάλμα

Ξεκινώντας ο αλγόριθμος ελέγχει ακολουθιακά τα καρέ της εικονοσειράς και βρίσκει τις υποψήφιες διαφημίσεις που έχουν μάσκα που δεν απέχει πολύ από την δοσμένη μάσκα της εισόδου. Ο έλεγχος γίνεται με την μέθοδο του μέσου τετραγωνικού σφάλματος και η τιμή του για τις υποψήφιες διαφημίσεις δεν πρέπει να ξεπερνά το κατώφλι που έχει οριστεί. Έτσι, στο παράδειγμα, βρίσκονται δύο υποψήφιες διαφημίσεις από την βάση δεδομένων με αναγνωριστικά id=119 και id=83.

Εδώ πρέπει να σημειωθεί πως άλλη μια ιδιαιτερότητα της εικονοσειράς αυτής, είναι το γεγονός πως οι υποψήφιες διαφημίσεις με αναγνωριστικά id=119 και id=83 έχουν την ίδια αρχή και το ίδιο λογότυπο στο τέλος και διαφοροποιούνται μόνο στο μέσον τους, δηλαδή αποτελούν διαφορετικές παραλλαγές μιας διαφήμισης της ίδιας εταιρίας. Έτσι, μια επιπλέον δυσκολία του παραδείγματος αυτού έγκειται στο ότι πρέπει ο αλγόριθμος να μην μπερδέψει τις δυο διαφημίσεις, παρόλο που σε αρκετά σημεία αποτελούνται από όμοια καρέ. Για αυτόν τον λόγο λοιπόν το μέσο τετραγωνικό σφάλμα έχει περίπου την ίδια τιμή και για τις δύο διαφημίσεις με την μικρή απόκλιση να οφείλεται σε μια μικροδιαφορά στον αριθμό των καρέ των δύο διαφημίσεων



Εικόνα 5.2.4. Ομοιότητες και διαφορές στα καρέ για δύο διαφορετικές παραλλαγές διαφήμισης ίδιας εταιρίας

Στη συνέχεια της εκτέλεσης παραμένουν όπως αναμένεται και οι δύο διαφημίσεις αυτές στο σύνολο με τις υποψήφιες, μαζί με κάποιες άλλες που μπορεί να εισέρχονται και αυτές για λίγο στο σύνολο των υποψηφίων λόγω κάποιας μικρής ομοιότητας σε κάποιο καρέ. Στο σημείο όμως της διαφοροποίησης των δύο παραλλαγών, και συγκεκριμένα μετά από 176 καρέ η λανθασμένη υποψήφια με αναγνωριστικό id=83 διαγράφεται σωστά από το σύνολο των υποψηφίων και τελικά αναγνωρίζεται από τον αλγόριθμο η διαφήμιση id=119.



Εικόνα 5.2.5. Ενδεικτικά οι υποψήφιες διαφημίσεις από την βάση μαζί με τις αντίστοιχες μάσκες σε σύγκριση με το καρέ εισόδου μετά από 288 καρέ

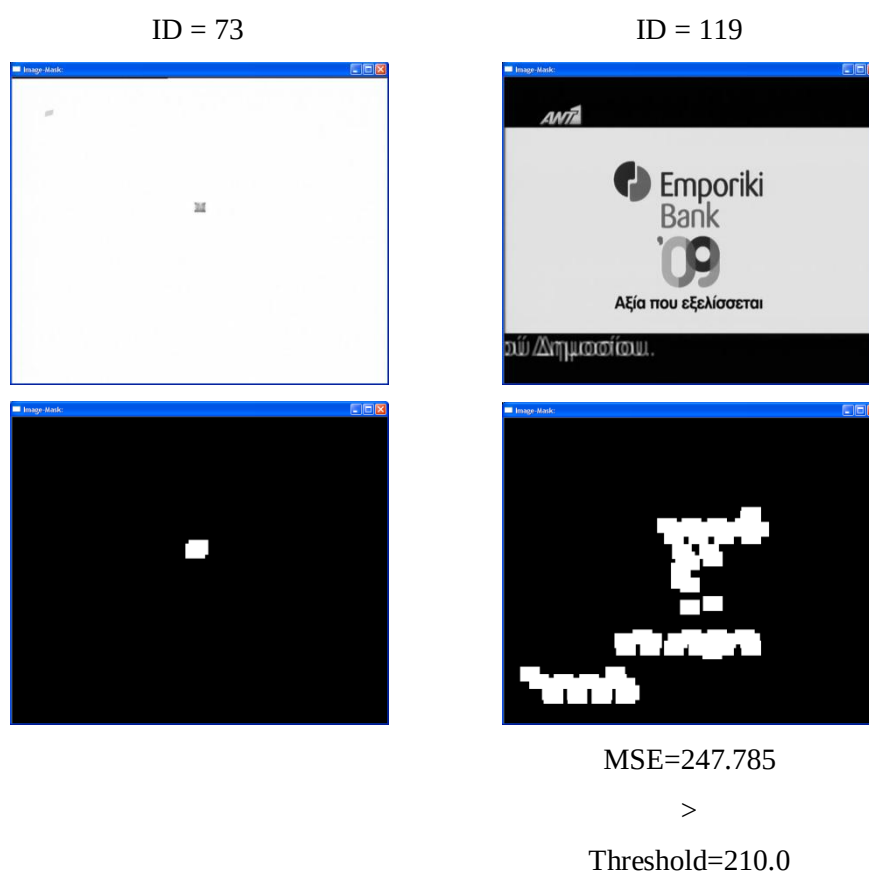
Πίνακας 5.2.1. Τιμές για τη μέση τιμή του μέσου τετραγωνικού σφάλματος των υποψήφιων διαφημίσεων

Ad - ID	119	41	51	4	28
MSE	125.649	203.377	222.835	223.132	236.363

Το σημείο αυξημένου ενδιαφέροντος στη συνέχεια είναι το σημείο που κόβεται πρόωρα η διαφήμιση που έχει ταυτοποιηθεί με αυτή της βάσης με αναγνωριστικό id=119

και ξεκινά να προβάλλεται στην εικονοσειρά η δεύτερη διαφήμιση, η οποία αντιστοιχεί στη διαφήμιση με αναγνωριστικό id=72 της βάσης δεδομένων.

Αυτό που συμβαίνει είναι πως με το που γίνει η αλλαγή της διαφήμισης η ταυτοποιημένη διαφήμιση id=119 μετακινείται στο σύνολο των Λανθασμένων υποψηφίων (F) αφού δεν βρέθηκε κάποιο καρέ που το μέσο τετραγωνικό σφάλμα της μάσκας του να μην υπερβαίνει το κατώφλι(threshold). Όταν διαπιστωθεί ότι μια ταυτοποιημένη διαφήμιση πηγαίνει στο σύνολο των Λανθασμένων υποψηφίων όμως, ο αλγόριθμος αρχίζει να προσπαθεί να αντιστοιχίσει τα επόμενα καρέ αναζητώντας στη βάση, κάποια νέα διαφήμιση που πιθανώς ξεκίνησε.



Εικόνα 5.2.6. Το πρώτο καρέ κατά το οποίο εισέρχεται στην εικονοσειρά η δεύτερη διαφήμιση με αποτέλεσμα την αποτυχία στον έλεγχο του μέσου τετραγωνικού σφάλματος της ταυτοποιημένης

Πράγματι λοιπόν βρίσκονται 49 νέες υποψήφιες διαφημίσεις οι οποίες εισάγονται στην αντίστοιχη λίστα. Έπειτα, αφού η ταυτοποιημένη διαφήμιση παραμένει για 3

συνεχόμενες φορές στη λίστα με τις υποψήφιες μετά διαγράφεται και θεωρείται πλέον ότι έχει ολοκληρωθεί η προβολή της. Τότε, γίνεται ο έλεγχος της διάρκειας της διαφήμισης όπου διαπιστώνεται πως μετρήθηκαν 660 καρέ της διαφήμισης αντί 759 που είναι στο σύνολο, διαφορά που είναι σημαντική. Έτσι στο αποτέλεσμα της διαφήμισης περιέχεται προειδοποιητικό μήνυμα για αυτά τα $759-660=99$ καρέ ή 3.96 δευτερόλεπτα που χάθηκαν.

«Found Ad - ID = 107. Warning! 3.96 seconds less than original. »

Τέλος, όπως αναφέρθηκε αρχικά, στην πραγματικότητα είχαν κοπεί 4 δευτερόλεπτα από τη διαφήμιση αυτή, όμως αυτή η διαφορά του ενός καρέ είναι πολύ μικρή και μέσα στα όρια σφάλματος που έχουν θεωρηθεί.

Κατόπιν, δε συναντάται κάποιο πρόβλημα και η εκτέλεση του προγράμματος συνεχίζεται ομαλά. Μοναδική υποψήφια μετά από 74 καρέ μένει η διαφήμιση της βάσης με αναγνωριστικό id=72 η οποία είναι και η διαφήμιση που προβάλλεται στην εικονοσειρά. Τελικά μετά την ταυτοποίηση και ολοκλήρωση της διαφήμισης αυτής, ο έλεγχος για την διάρκεια επιστρέφει σαν αποτέλεσμα ότι έπαιξε για 336 καρέ αντί 347 που αναμέναμε. Αυτή η διαφορά 11 καρέ ή 0.44 δευτερολέπτων οφείλεται στο γεγονός πως ο αλγόριθμος δεν ανέμενε την απότομη αλλαγή διαφήμισης χωρίς να έχει φθάσει στο τέλος της η πρώτη διαφήμιση. Σε κάθε περίπτωση όμως μια τέτοια διαφορά της τάξης του μισού δευτερολέπτου, θεωρείται πολύ μικρή και για αυτό το πρόγραμμα επιστρέφει :

«Found Ad - ID = 72.»

6. Πειραματικά αποτελέσματα

Το σύνολο δεδομένων που χρησιμοποιήθηκε για τη δοκιμή του συστήματος απαρτίζεται από εικονοσειρές τηλεοπτικού σήματος ελληνικών καναλιών.

Το σήμα λήφθηκε μέσω της ψηφιακής πλατφόρμας της NOVA [NOVA, 2010] καθώς θεωρήθηκε ότι το σήμα θα ήταν καλής ποιότητας, λόγω της ψηφιακής κωδικοποίησής του. Ωστόσο, ήδη από το 2010 μεταδίδεται το τηλεοπτικό σήμα δημοφιλών καναλιών, δωρεάν, μέσω της επίγειας ψηφιακής πλατφόρμας βασισμένης στο πρωτόκολλο DVB-T [Digea, 2010]. Έτσι, κρίνεται ότι το σύστημα δε θα απαιτεί κατά την πλήρη λειτουργία του, υψηλού κόστους υλικό βάσης.

Συνολικά λήφθηκαν περίπου 26 ώρες βίντεο από 6 δημοφιλή ελληνικά τηλεοπτικά κανάλια. Η κατανομή σε διάρκεια του τηλεοπτικού σήματος που λήφθηκε ανά κανάλι φαίνεται στον πίνακα 6.1. Το δείγμα εικονοσειρών λήφθηκε στο διάστημα 25/2/2009 με 3/3/2009 ώστε να είναι πιθανό μία διαφήμιση που έχει αποθηκευτεί από ένα κανάλι να έχει παρουσιαστεί στα διαφημιστικά διαλλείματα κάποιου άλλου καναλιού.

Πίνακας 6.1. Δείγμα αποθηκευμένης τηλεοπτικής ροής

Κανάλι	Διάρκεια αποθηκευμένης ροής
ALPHA	4:22:06
ALTER	2:52:14
ANT1	2:52:57
MEGA	5:02:14
NET	9:05:00
STAR	1:56:22
Σύνολο:	26:10:53

Στη συνέχεια, το τηλεοπτικό δείγμα που συλλέχθηκε κατατμήθηκε στις εικονοσειρές διαφημίσεων που το απαρτίζει. Κατατμήθηκαν συνολικά 138 διαφημιστικά, τα οποία περιλαμβάνουν και εικονοσειρές που προβλήθηκαν σε πολλαπλά κανάλια. 113

από αυτά, δηλαδή περίπου το 80% του συνόλου δεδομένων, αποθηκεύτηκαν στη βάση δεδομένων της μεθόδου.

Έπειτα, σχεδιαστήκαν τρία διαφορετικά σύνολα δοκιμής, με κλιμακούμενη δυσκολία, όπως φαίνεται και στον Πίνακα 6.2. Το πρώτο σύνολο δοκιμής απαρτίζεται από κανονική ροή προγράμματος των καναλιών και διαφημίσεις που υπάρχουν στη βάση. Με το σύνολο αυτό επιχειρείται να ελεγχθεί το κατά πόσο το σύστημα μπορεί να ταυτίσει τις διαφημίσεις που έχει αποθηκεύσει στη βάση δεδομένων με την τηλεοπτική ροή.

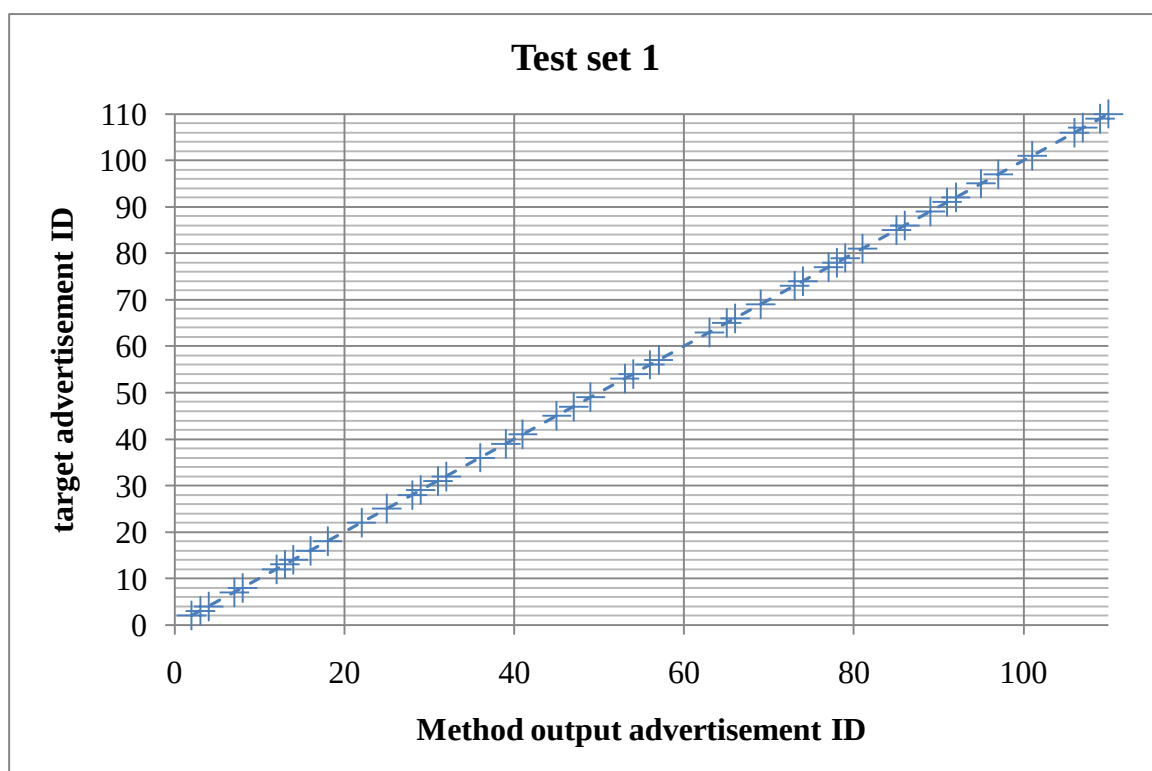
Το επόμενο σύνολο δοκιμής εκτός από πρόγραμμα ροής του καναλιού και γνωστές διαφημίσεις στο σύστημα, επιπλέον εμπεριέχει άγνωστες διαφημίσεις που δεν είναι καταχωρημένες στη βάση. Επίσης, σε αυτή τη κατηγορία εντάσσονται και μερικές γνωστές διαφημίσεις με διαφορετικό κανάλι προβολής απ' ότι στη βάση δεδομένων. Έτσι, γίνεται δυνατή η εξαγωγή ασφαλών συμπερασμάτων σχετικά με το πόσο αποτελεσματικά λειτουργεί το σύστημα υπό κανονικές συνθήκες.

Το τρίτο και τελευταίο σύνολο δοκιμής περιλαμβάνει ένα σύνολο παθολογικών περιπτώσεων ώστε να ελεγχθούν τα όρια αποτελεσματικότητας του συστήματος. Στο σύνολο δοκιμής υπάρχουν όπως στα δύο προηγούμενα γνωστές και άγνωστες διαφημίσεις καθώς πρόγραμμα ροής. Οι παθολογικές περιπτώσεις περιλαμβάνουν διαφημίσεις που έχουν παραποιηθεί, εσκεμμένα ή μη, για την εξοικονόμηση τηλεοπτικού χρόνου. Σε αυτή την περίπτωση το σύστημα πρέπει είτε να μη σημειώσει ότι έχει προβληθεί μία διαφήμιση είτε να εκδώσει μία προειδοποίηση.

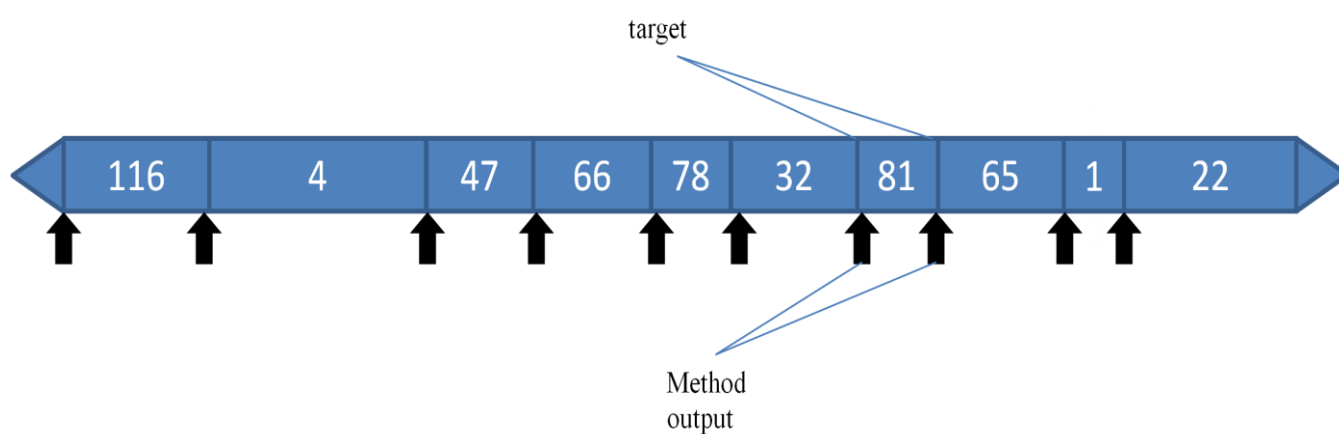
Πίνακας 6.2. Σύνολα δοκιμής του συστήματος ταύτισης εικονοσειρών

Σύνολο δεδομένων	Βασικό τηλεοπτικό πρόγραμμα	Διαφημίσεις στη βάση δεδομένων	Άγνωστες διαφημίσεις	Αλλοιωμένες διαφημίσεις
1	~20 λεπτά	50	0	0
2	~20 λεπτά	20	20	0
3	~20 λεπτά	13	13	9

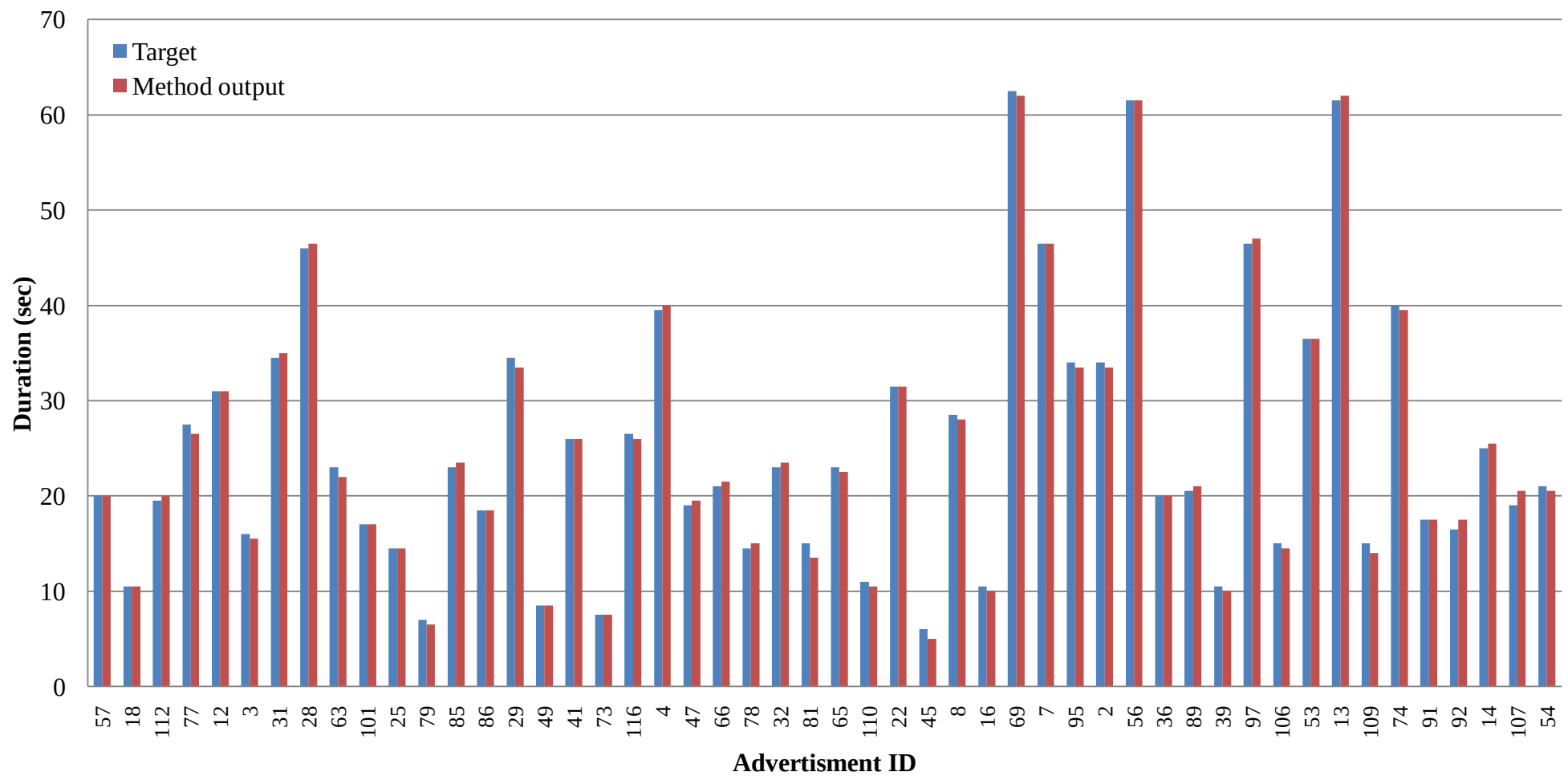
Όσον αφορά στο πρώτο σύνολο δοκιμής, η αποτελεσματικότητα του συστήματος είναι 100%, όπως φαίνεται και στο Σχήμα 7.1. Επίσης, δεν αναγνώρισε καμία διαφήμιση που υπάρχει στη βάση στο πρόγραμμα κανονικής ροής των 20' που εμπεριέχεται στο σύνολο δοκιμής. Στο Σχήμα 7.2, φαίνεται το πώς αποκρίθηκε το σύστημα σε ένα κομμάτι του συνόλου δοκιμής. Όπως φαίνεται στο συγκεκριμένο σχήμα το σύστημα αναγνώρισε σωστά τις διαφημίσεις και τα χρονικά τους όρια στην εικονοσειρά.



Σχήμα 6.1. Αποτελεσματικότητα μεθόδου για το πρώτο σύνολο δοκιμής



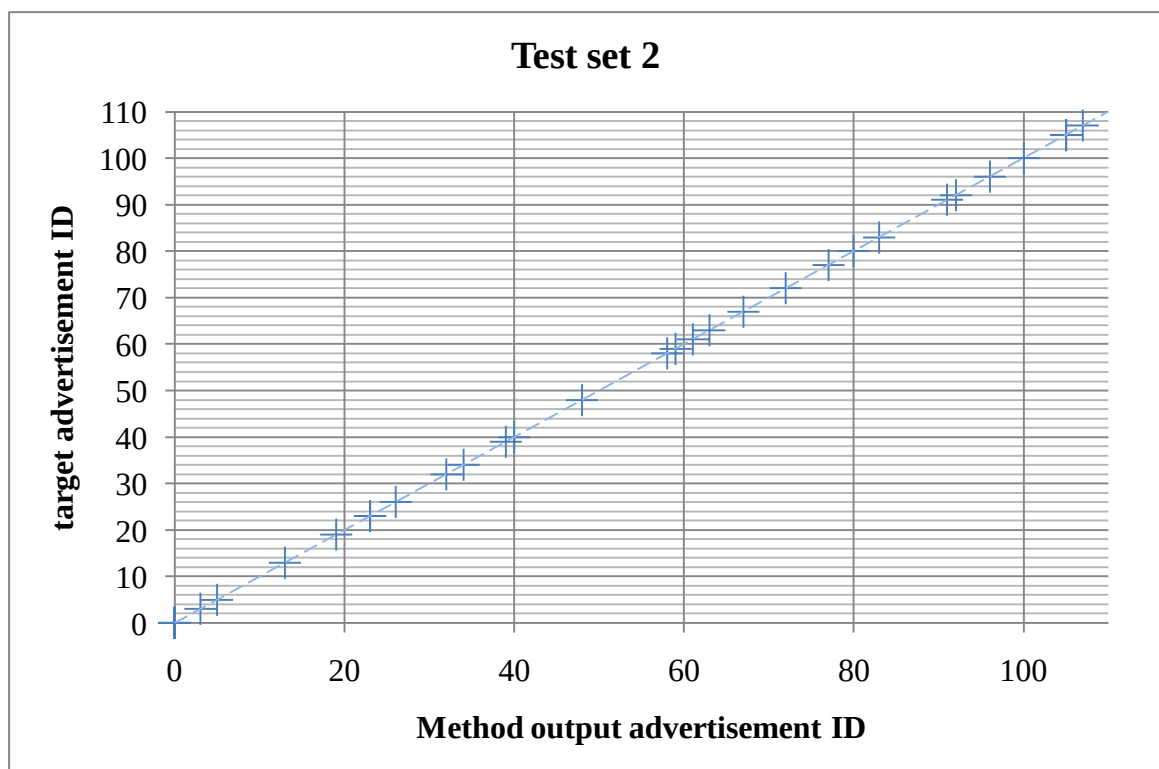
Σχήμα 6.2. Παράδειγμα εξόδου του συστήματος για το πρώτο σύνολο δοκιμής



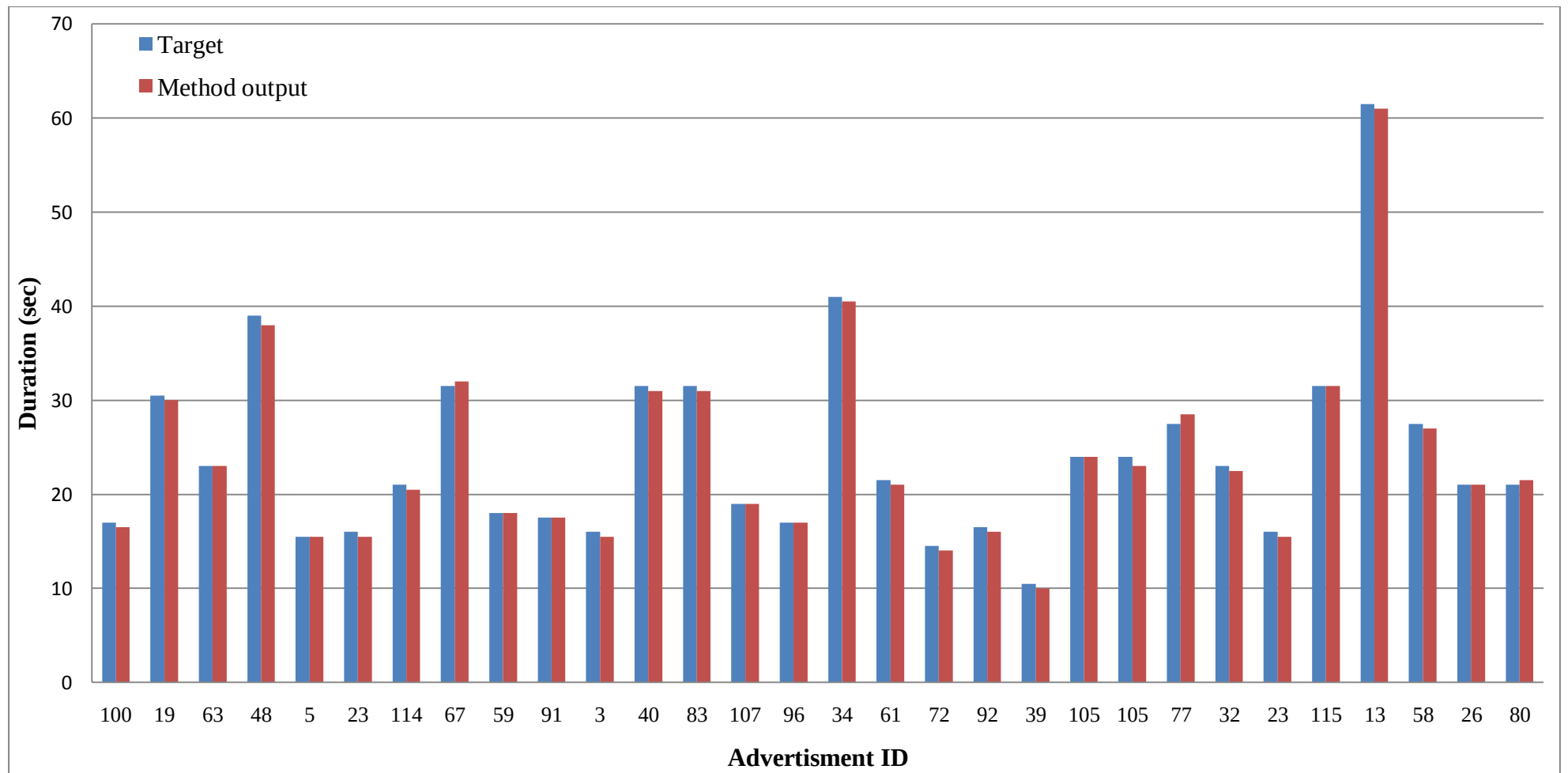
Σχήμα 6.3. Αποτελέσματα χρονικής διάρκειας του συνόλου δοκιμής 1 για τις επιθυμητές (target) και πραγματικές (method output) εξόδους

Το Σχήμα 6.3. δείχνει την αποτελεσματικότητα του συστήματος στο να προσδιορίζει τη σωστή χρονική διάρκεια των αναγνωρισμένων διαφημίσεων. Όπως φαίνεται και στο προαναφερθέν Σχήμα, ο αλγόριθμος πετυχαίνει να προσδιορίζει ακριβώς τη διάρκεια των περισσότερων διαφημιστικών, εκτός από κάποιες περιπτώσεις όπου η έξοδος διαφέρει από το πραγματικό κατά 0,5 έως 1,5 δευτερόλεπτα. Η διαφορά αυτή εξαρτάται από το πλήθος των καρέ (μαύρων ή μη) που υπάρχουν ως συμπλήρωμα στις ροές των τηλεοπτικών καναλιών (στην αρχή ή το τέλος των διαφημιστικών) και στις διαφημίσεις που υπάρχουν στη βάση. Επίσης, πολλές διαφημίσεις έχουν κοινά χαρακτηριστικά ως προς άλλες σε ό,τι αφορά το τελικό τους λογότυπο, το οποίο μπορεί να είναι το σήμα της εταιρίας που παράγει το προϊόν ή της εταιρείας που δημιούργησε τη διαφήμιση. Σε κάθε περίπτωση ο αλγόριθμος φαίνεται να λειτουργεί πολύ καλά σε οτιδήποτε έχει να κάνει με το χαμηλής δυσκολίας σύνολο δοκιμής.

Στο μέσης δυσκολίας σύνολο δεδομένων, το σύστημα παρουσιάζει επίσης υψηλή αποτελεσματικότητα. Όπως φαίνεται στο Σχήμα 6.4. η μέθοδος ταύτισε τις αποθηκευμένες διαφημίσεις που υπήρχαν στην ροή του συνόλου. Επίσης, δεν παρουσιάστηκε κάποιο σφάλμα κατηγοριοποίησης μίας άγνωστης εικονοσειράς με μία της βάσης (false positive) είτε το αντίθετο (false negative).



Σχήμα 6.4. Αποτελεσματικότητα μεθόδου για το δεύτερο σύνολο δοκιμής



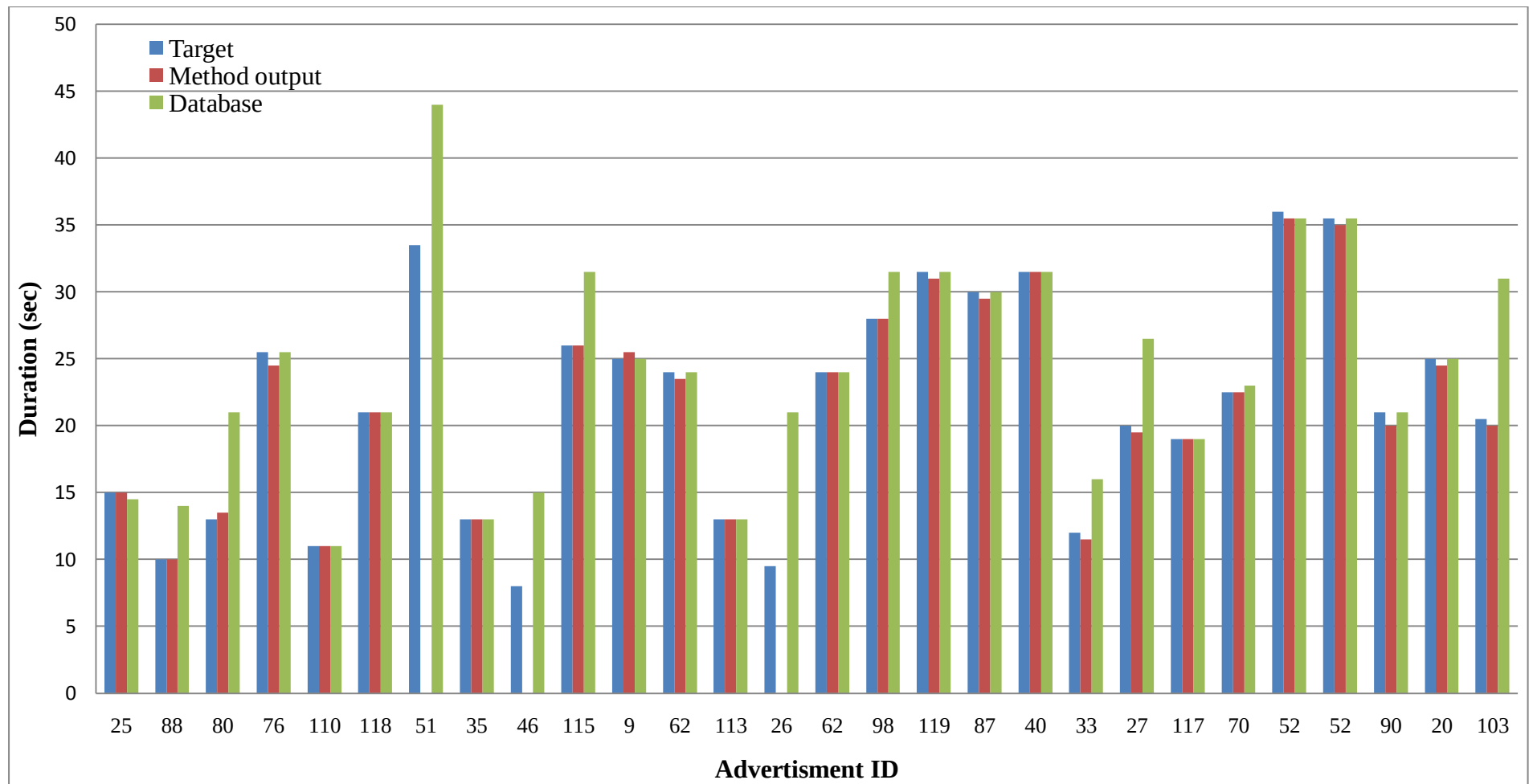
Σχήμα 6.5. Αποτελέσματα χρονικής διάρκειας του συνόλου δοκιμής 2 για τις επιθυμητές (target) και πραγματικές (method output) εξόδους

Στο Σχήμα 6.5 φαίνεται ότι όπως και κατά την επεξεργασία του πρώτου συνόλου δεδομένων το σύστημα πέτυχε να αναγνωρίσει σωστά τα χρονικά όρια κάθε «γνωστής» διαφήμισης.

Τα αποτελέσματα του συστήματος για το σύνολο δοκιμής 3 φαίνονται στον πίνακα 6.3 και στο Σχήμα 6.6. Ο πίνακας αποτελεί το μητρώο σύγχυσης (confusion matrix) συστήματος για τις πραγματικές και επιθυμητές εξόδους σχετικές με τις γνωστές, άγνωστες και αλλοιωμένες διαφημίσεις. Όπως φαίνεται στον πίνακα, το σύστημα επιτυγχάνει απόλυτα να διακρίνει τις γνωστές και τις άγνωστες διαφημίσεις, όπως και στα προηγούμενα δύο σύνολα δοκιμής. Σε ό,τι αφορά στις αλλοιωμένες διαφημίσεις, ανάλογα με το βαθμό που αυτές έχουν παραποιηθεί ταξινομούνται είτε ως άγνωστες, σε να μην έχουν δηλαδή εκπεμφθεί καθόλου, είτε αναγνωρίζεται ότι έχουν προβληθεί αλλά εκδίδεται προειδοποίηση αλλοιωμένης διαφήμισης. Στο Σχήμα 6.6 φαίνεται ότι αναγνωρίζονται σωστά οι διαφημίσεις που η διάρκεια στη βάση δεδομένων και στη ροή συμπίπτει (γνωστές διαφημίσεις). Οι διαφημίσεις που η διάρκειά τους στη ροή του συνόλου δεδομένων είναι διαφορετική είτε ταυτίζονται ως αλλοιωμένες, όπως για παράδειγμα οι διαφημίσεις με αναγνωριστικό (ID) 88 και 33, είτε δεν αναγνωρίζονται καθόλου (διαφήμιση 51, 46, κ.α.).

Πίνακας 6.3. Αποτελεσματικότητα συστήματος για το σύνολο δοκιμής 3

	Πραγματική έξοδος			
		Γνωστές διαφημίσεις	Άγνωστες διαφημίσεις	Κομμένες διαφημίσεις
Έξοδος συστήματος	Γνωστές διαφημίσεις	18	0	0
	Άγνωστες διαφημίσεις	0	9	0
	Κομμένες διαφημίσεις	0	4	7



Σχήμα 6.6. Αποτελέσματα χρονικής διάρκειας του συνόλου δοκιμής 3 συμπεριλαμβανομένων αλλοιωμένων διαφημίσεων

Επιπλέον, τα αποτελέσματα έδειξαν ότι το πρόγραμμα ικανοποιεί τις απαιτήσεις για γρήγορη εκτέλεση εκτός από αποδοτικότητα. Όπως έχει αναφερθεί, οι προδιαγραφές του συστήματος αναδείκνυν την ανάγκη για εκτέλεση σε σχεδόν πραγματικό χρόνο. Οι μετρήσεις στη διάρκεια εκτέλεσης έδειξαν ότι επετεύχθη αυτός ο σκοπός αφού η επεξεργασία γινόταν με ρυθμό 9-11 καρέ το δευτερόλεπτο. Δεδομένου ότι υπάρχει δειγματοληψία 6 καρέ το δευτερόλεπτο ουσιαστικά το πρόγραμμα μπορεί να επεξεργαστεί σε ένα δευτερόλεπτο 36-44 καρέ από τη ροή εισόδου του τηλεοπτικού σήματος, τη στιγμή που η προβολή τους στη τηλεόραση γίνεται με ρυθμό 25 καρέ ανά δευτερόλεπτο.

Παράλληλα, έγιναν συμπληρωματικές μετρήσεις διαφοροποιώντας τις παραμέτρους του αλγορίθμου Επαγωγικής Σάρωσης. Έτσι, ορίζοντας την ταχύτητα ίση με το μέγεθος του κυλιομένου παραθύρου και την επιτάχυνση ίση με μηδέν η επεξεργασία απλοποιείται στην κατά μπλοκ ανάλυση όλων των περιοχών της εικόνας. Αναμένεται σε αυτή τη περίπτωση η παρατήρηση μιας χρονικής επιβάρυνσης, αφού γίνεται επεξεργασία ακόμα και στα κομμάτια που δεν περιέχουν πληροφορία. Επίσης, πραγματοποιήθηκε δοκιμή αυξάνοντας τη μέγιστη ταχύτητα αλλά και επιτάχυνση του αλγορίθμου. Σε αυτή τη περίπτωση παρατηρήθηκε μείωση της διάρκειας εκτέλεσης, στοιχείο λογικό αφού παραλείπεται μεγαλύτερο ποσοστό από κάθε εικόνα. Τα παραπάνω αποτελέσματα φαίνονται στον Πίνακα 6.4. που ακολουθεί. Φαίνεται πως οι αποκλίσεις αυτές που παρατηρούνται δεν είναι πολύ μεγάλες. Αυτό οδηγεί στο συμπέρασμα πως περισσότερη βαρύτητα στο πρόγραμμα έχει τελικά ο αλγόριθμος ταύτισης των εικονοσειρών καθώς και στο ότι μπορεί δευτερογενείς παράγοντες όπως η αλληλεπίδραση με τη βάση δεδομένων ή την εσωτερική μνήμη του συστήματος, να έχουν εξίσου μεγάλη βαρύτητα στο χρόνο εκτέλεσης.

Πίνακας 6.4. Χρονική Διάρκεια για διαφορετικές παραμέτρους στα 3 σύνολα δοκιμής

	Μέγιστη Ταχύτητα (pixels)			
		0	20	75
	easy	633,8s	591,5s	594,7s
	medium	502,4s	497,1s	435,1s
	hard	994,7s	960,8s	900,6s

7. Επίλογος

Συνοψίζοντας, στα παραπάνω κεφάλαια αναλύθηκε το πρόβλημα της ταύτισης εικονοσειρών τηλεοπτικού σήματος για διαφημίσεις. Μετά από τη μελέτη της τρέχουσας τεχνολογικής στάθμης και ερευνητικής δραστηριότητας, ανέκυψε η ανάγκη υλοποίησης ενός γρήγορου συστήματος που να μην επηρεάζεται από διακυμάνσεις στα χαρακτηριστικά προβολής των διαφημιστικών μηνυμάτων. Η μέθοδος που χρησιμοποιήθηκε εφαρμόζει τον αλγόριθμο Επαγωγικής Σάρωσης για την επεξεργασία του κάθε καρέ και την εύρεση της πληροφορίας, μειώνοντας το χώρο αναζήτησης στις υπόλοιπες αδιάφορες περιοχές. Σύμφωνα με αυτά τα στοιχεία γίνεται η αποθήκευση των δεδομένων σε μια βάση δεδομένων, δημιουργημένη για αυτό ακριβώς το σκοπό. Επίσης, ο αλγόριθμος για την ταύτιση των εικονοσειρών φροντίζει για το γρήγορο και «έξυπνο» τρόπο ταυτοποίησης του σήματος με τις καταχωρημένες διαφημίσεις. Όπως φάνηκε, τα πειραματικά αποτελέσματα επιβεβαίωσαν την αποδοτικότητα των ανωτέρω τεχνικών εφαρμόζοντας ελέγχους σε μεγάλο δείγμα δεδομένων με σκοπό την αύξηση της αξιοπιστίας.

Στο μέλλον υπάρχουν αρκετά περιθώρια επέκτασης της παρούσας εργασίας. Μπορούν να χρησιμοποιηθούν άλλοι περιγραφείς εικόνων αντί ενός ανιχνευτή ακμών κυρίως σε περιγραφείς χρωματικών συνιστωσών των εικόνων. Εναλλακτικά, για ακόμη μεγαλύτερη αποτελεσματικότητα, μπορεί να γίνει συνδυασμός των περιγραφέων εικόνων με άλλους, όπως περιγραφείς για την ηχητική πληροφορία του σήματος. Επιπλέον, η ταύτιση των εικονοσειρών μπορεί να γίνεται με τη βοήθεια στατιστικών ή στοχαστικών μεθόδων. Μία τέτοια είναι τα κρυμμένα μοντέλα Markov (Hidden Markov Models), τα οποία έχουν χρησιμοποιηθεί και παρουσιάζει θετικά αποτελέσματα σε παρόμοιες εφαρμογές. Τέλος, ολόκληρη η εφαρμογή μπορεί να γενικευθεί για ταύτιση όχι μόνο διαφημιστικών εικονοσειρών, αλλά και εικονοσειρών λοιπού τηλεοπτικού προγράμματος ή εικονοσειρών στο διαδίκτυο.

Αναφορές

- Agnihotri, L., N. Dimitrova, T. Mcgee, S. Jeannin, D. Schaffer and J. Nesvadba (2003). Envolvable visual commercial detector. IEEE Conference on Computer Vision and Pattern Recognition (CVPR '03). Los Alamitos, CA, USA. 2: 79.
- Albiol, A., M. J. Fulla, A. Albiol and L. Torres (2004). Commercials detection using HMMs. Proc. of the International Workshop on Image Analysis for Multimedia Interactive Services, Lisboa, Portugal.
- Anusha, V., M. Kalaiselvigetha and S. Palanivel (2008). Classifying Video into Predefined Genres. Proc. of 4th International Conference on Networked Computing and Advanced Information Management: 66-70.
- Apache 2010. “Διακομιστής (Server) Apache έκδοσης 2.2.15” Retrieved 22/9/2010, from <http://www.apache.org/>
- Canny, J., A Computational Approach To Edge Detection, IEEE Trans. Pattern Analysis and Machine Intelligence, 8(6):679–698, 1986.
- Cheung, S. C. and A. Zakhor (2003). "Efficient video similarity measurement with video signature." IEEE Trans. Circuits and Systems for Video Technology 13(1): 59-74.
- Chiu, D. M. and R. Jain (1989). "Analysis of the increase and decrease algorithms for congestion avoidance in computer networks." Computer Networks and ISDN Systems 17(1): 1–14.
- Colombo, C., A. D. Bimbo and P. Pala (2001). "Retrieval of Commercials by Semantic Content: The Semiotic Perspective." Multimedia Tools and Applications 13(1): 93—118.
- Dalal, N. and B. Triggs (2005). Histograms of Oriented Gradients for Human Detection. Proceedings of the 2005 Conference on Computer Vision and Pattern Recognition (CVPR 2005): 886-893.

- Digea (2010). "Digea DVB-T." Retrieved 22/9/2010, from <http://www.digea.gr/>.
- Dimitrova, N., L. Agnihotri and G. Wei (2000). Video classification based on HMM using text and faces. European Signal Processing Conference. Tampere, Finland.
- Duan, L.-Y., Y.-T. Zheng, J. Wang, H. Lu and J. S. Jin (2008). "Digesting Commercial Clips from TV Streams." IEEE Multimedia 15(1): 28-41.
- Duygulu, P., M.-Y. Chen and A. Hauptmann (2004). Comparison and Combination of Two Novel Commercial Detection Methods. Proc.of 2004 IEEE International Conference on Multimedia and Expo, Taipei: 1267- 1270.
- Feng, H., W. Fang, S. Liu and Y. Fang (2005). A new general framework for shot boundary detection and key-frame extraction. Proceedings of the 7th ACM SIGMM international workshop on Multimedia information retrieval: 121-126.
- Gauch, J. M. and A. Shivadas (2006). "Finding and identifying unknown commercials using repeated video sequence detection." Computer Vision and Image Understanding 103: 80–88.
- Giannoukos, I., C.-N. Anagnostopoulos, V. Loumos and E. Kayafas (2010). "Operator context scanning to support high segmentation rates for real time license plate recognition." Pattern Recognition 43(11): 3866-3878.
- González Rafael C., Richard Eugene Woods. (2008). Digital image processing, Prentice Hall, ISBN013168728X, 9780131687288
- Hampapur, A. and R. M. Bolle (2001). VideoGREP: Video Copy Detection using Inverted File Indices, IBM Research.
- Huang, L.-L. and A. Shimizu (2006). "A multi-expert approach for robust face detection." Pattern Recognition 39 (9): 1695-1703.
- Huang, Y.-P., L.-W. Hsu and F.-E. Sandnes (2007). "An Intelligent Subtitle Detection Model for Locating Television Commercials." IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics 37(2): 485 - 492.

- Katsioulis, P., V. Tsetsos and S. Hadjiefthymiades (2007). Semantic video classification based on subtitles and domain terminologies. Proc. of 2nd International Conference on Semantic and Digital media Technologies (SAMT'07), Workshop on Knowledge Acquisition from Multimedia Content (KAMC), Genova, Italy.
- Klimesova, D. (2000). "Sliding window for relations mapping." International Archives of Photogrammetry and Remote Sensing 33(B7): 688-692.
- Lampert, C. H., M. B. Blaschko and T. Hofmann (2008). Beyond Sliding Windows: Object Localization by Efficient Subwindow Search. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2008): 1-8.
- Liu, T.-Y., T. Qin and H.-J. Zhang (2004). Time-constraint boost for TV commercials detection. Proc. of International Conference on Image Processing: 1617 - 1620.
- Liu, Z. and Y. Wang (1998). "Audio Feature Extraction and Analysis for Scene Segmentation and Classification." Journal of VLSI Signal Processing 20(1-2): 61-79.
- Lu, L., H. Jiang and H. J. Zhang (2001). A Robust Audio Classification and Segmentation Method. Proc. of the 9th ACM international conference on Multimedia: 203-211.
- Meng, L., Y. Cai, M. Wang and Y. Li (2009). TV Commercial Detection Based on Shot Change and Text Extraction. 2nd International Congress on Image and Signal Processing: 1-5.
- Meynet, J., V. Popovici and J.-P. Thiran (2007). "Face detection with boosted Gaussian features." Pattern Recognition 40(8): 2283-2291.
- MySQL (2010). "MySQL στην έκδοση 5.1.45." Retrieved 22/9/2010, from <http://www.mysql.com/>
- MySQL connectors 2010. "mysql-connector libraries στην έκδοση 5.1.48." Retrieved 22/9/2010, from <http://www.mysql.com/products/connector/>

- Nagasaka, A. and T. Miyatake (2000). "A Quick Scene Classification Method Based on Compact Encoding of Video Feature Sequence." *Systems and Computers in Japan* 31(1): 102-108.
- Nijmeijer, T. (2008). *Automatic Classification of Television Commercials*. Human Media Interaction, Electrical Engineering, Mathematics and Computer Science Dpt. Twente, University of Twente, Netherlands. Master.
- NOVA (2010). "Nova DVB-S." Retrieved 22/9/2010, from <http://www.nova.gr>.
- Nwe, T. L. and H. Li (2005). Broadcast news segmentation by audio type analysis. *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing*: 1065 - 1068.
- Oostveen, J., T. Kalker and J. Haitsma (2002). Feature Extraction and a Database Strategy for Video Fingerprinting. *Proc. International Conference on Recent Advances in Visual Information Systems*: 117-128.
- OpenCV (2010). "Open Source Computer Vision Library έκδοσης 2.1." Retrieved 22/9/2010, from <https://sourceforge.net/projects/opencvlibrary/files/opencv-win/2.1/>
- Ostermann, J., J. Bormans, P. List, D. Marpe, M. Narroschke, F. Pereira, T. Stockhammer and T. Wedi (2004). "Video coding with H. 264/AVC: tools, performance, and complexity." *IEEE Circuits and Systems Magazine* 4(1): 7-28.
- PAL 2010. "NTSC, PAL, and SECAM Overview". Retrieved 22/9/2010, from <http://www.deetc.isel.ipl.pt/>
- PHPMyAdmin 2010. "PHPMyAdmin στην έκδοση 3.3.2" Retrieved 22/9/2010, from <http://www.phpmyadmin.net/>
- Pitas I. (2000). *Digital Image Processing Algorithms and Applications*, Wiley-Interscience; ISBN-10: 0471377392
- Roover, C. D., C. D. Vleeschouwer, F. Lefebvre and B. Macq (2005). "Robust video hashing based on radial projections of key frames." *IEEE Trans. Signal Processing* 53(10): 4020-4037.

- Sadlier, D. A., S. Marlow, N. O'Connor and N. Murphy (2002). "Automatic TV advertisement detection from MPEG bitstream." *Pattern Recognition* 35(12): 2719-2726.
- Sanchez, J. M., X. Binefa and J. Vitria (2002). "Shot partitioning based recognition of TV commercials." *Multimedia Tools and Applications* 18(3): 233–247.
- Satterwhite, B. and O. Marques (2004). "Automatic detection of TV commercials." *IEEE Potentials* 23(2): 9-12.
- Sayadi, M., LTlig and F. Fnaiech (2007). "A New Texture Segmentation Method Based on the Fuzzy C-Mean Algorithm and Statistical Features." *Applied Mathematical Sciences* 1(60): 2999-3007.
- Schoffmann, K., M. Lux and L. Boszormenyi (2009). Novel Approach for Fast and Accurate Commercial Detection in H.264/AVC Bit Streams Based on Logo Identification. *Proc. of the 15th International Multimedia Modeling Conference on Advances in Multimedia Modeling*: 119 – 127.
- Shen, J.-J. and C.-M. Zen (2006). Using Sliding Window Technique to Explore the Variations of Image Pixels for Edge Detection. *Proceedings of the International Computer Symposium (ICS 2006)*: 1109-1114.
- Torre, V. and T.A. Poggio. (1980). On edge detection. *IEEE Trans. Pattern Anal. Machine Intell.* , 8 : 2 pp. 147–163
- VS2008-Pro “Visual Studio 2008 Professional Edition” Retrieved 22/9/2010, from <http://www.microsoft.com/>
- Wang, X. and Z. Guo (2008). A Novel Real-Time Commercial Detection Scheme. *Proc. of 3rd International Conference on Innovative Computing Information and Control*: 536 – 536.