



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΤΟΜΕΑΣ ΕΠΙΚΟΙΝΩΝΙΩΝ, ΗΛΕΚΤΡΟΝΙΚΗΣ ΚΑΙ ΣΥΣΤΗΜΑΤΩΝ ΠΛΗΡΟΦΟΡΙΚΗΣ

Αναζήτηση και Ανάκτηση Πολυτροπικού, Οπτικοακουστικού Περιεχομένου

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

του

ΠΑΠΠΑ Θ. ΧΡΗΣΤΟΥ

Επιβλέπων: Συμεών Παπαβασιλείου
Αναπληρωτής Καθηγητής

Αθήνα, Ιούλιος 2012



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΤΟΜΕΑΣ ΕΠΙΚΟΙΝΩΝΙΩΝ, ΗΛΕΚΤΡΟΝΙΚΗΣ
ΚΑΙ ΣΥΣΤΗΜΑΤΩΝ ΠΛΗΡΟΦΟΡΙΚΗΣ

Αναζήτηση και Ανάκτηση Πολυτροπικού, Οπτικοακουστικού Περιεχομένου

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ
του
ΠΑΠΠΑ Θ. ΧΡΗΣΤΟΥ

Επιβλέπων: Συμεών Παπαβασιλείου
Αναπληρωτής Καθηγητής

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την

.....
Συμεών Παπαβασιλείου
Αναπληρωτής Καθηγητής Ε.Μ.Π.

.....
Βασίλειος Μάγκλαρης
Καθηγητής Ε.Μ.Π.

.....
Νικόλαος Μήτρου
Καθηγητής Ε.Μ.Π.

Αθήνα, Ιούλιος 2012

.....

ΠΑΠΠΑΣ Θ. ΧΡΗΣΤΟΣ

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

© 2012 – Με επιφύλαξη παντός δικαιώματος. All rights reserved

Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας, εξολοκλήρου 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

Σκοπός της παρούσας διπλωματικής εργασίας είναι η μελέτη κατάλληλων και αποδοτικών μηχανισμών ευρετηρίου, αναζήτησης και ανάκτησης πολυτροπικού οπτικοακουστικού περιεχομένου μέσα από μια βάση δεδομένων.

Στα πλαίσια της εργασίας χρησιμοποιήθηκαν υπάρχουσες τεχνολογίες βάσεων δεδομένων για την αποθήκευση των μεταδεδομένων που συνοδεύουν το οπτικοακουστικό υλικό. Η χρήση σχεσιακών Συστημάτων Διαχείρισης Βάσεων Δεδομένων επιλέχθηκε μέσα από ένα σύνολο εναλλακτικών επιλογών, λόγω της ευρείας χρήσης και υποστήριξής τους. Μελετήθηκε η δημιουργία ενός σχεσιακού σχήματος που επιτρέπει τη μετέπειτα αποδοτική αναζήτηση και ανάκτηση περιεχομένου μέσω κατάλληλων ευρετηρίων. Η διαμόρφωση ενός τέτοιου σχήματος απαιτεί την κατανόηση και μελέτη του προτύπου MPEG-7, με βάση το οποίο έχουν κωδικοποιηθεί οι περιγραφές των μεταδεδομένων.

Το βασικότερο στάδιο της αναζήτησης και ανάκτησης αποτελεί ο αλγόριθμος, ο οποίος δεδομένης της αποθήκευσης της υπάρχουσας πληροφορίας σε ένα κατάλληλο σχήμα, καλείται να εξάγει αποδοτικά τις ζητούμενες πληροφορίες. Ο αλγόριθμος απαιτεί την ύπαρξη κατάλληλων ευρετηρίων που επιταχύνουν την αναζήτηση περιεχομένου χρησιμοποιώντας σημασιολογικά χαρακτηριστικά του περιεχομένου σε συνδυασμό με ορισμένα χαμηλού επιπέδου χαρακτηριστικά. Η αναζήτηση με βάση υψηλού επιπέδου χαρακτηριστικά επικεντρώνεται στις κινήσεις και αντιδράσεις ανθρώπινων οντοτήτων στο διαθέσιμο υλικό. Η δυνατότητα αναζήτησης εμπλουτίζεται επιπλέον με ένα μηχανισμό ανατροφοδότησης σχετικότητας των αποτελεσμάτων, ώστε να παρουσιάζονται στο χρήστη πρώτα τα σχετικά αποτελέσματα.

Εν κατακλείδι, δημιουργήθηκε μια γραφική διεπαφή χρήστη μέσω της οποίας ο χρήστης μπορεί εύκολα να θέσει τα ερωτήματα του στη βάση δεδομένων με σκοπό την ανάκτηση ενός video. Η διεπαφή καθοδηγεί το χρήστη στην κατάλληλη διαμόρφωση των ερωτημάτων με βάση υψηλού επιπέδου χαρακτηριστικά.

Λέξεις κλειδιά: «πολυτροπικό οπτικοακουστικό περιεχόμενο, ευρετήρια πολυμεσικού υλικού, πολυμεσικές βάσεις δεδομένων, αναζήτηση οπτικοακουστικού περιεχομένου»

Abstract

The objective of this diploma thesis is the efficient search and retrieval of multimodal audiovisual content through a database which contains audiovisual content.

Existing database technologies were studied and adopted in order to store metadata that accompanied the multimedia material. Relational database systems were devised among a series of alternative options due to their broad use and support. Furthermore, the main aim of the research includes the creation of a relational schema which enables the subsequent efficient search and retrieval of data through appropriate indexes. The formulation of such a schema requires an understanding of the MPEG-7 standard, which is used for encoding metadata descriptions.

The most important phase of the search and retrieval process is the algorithm, which given the stored information in an appropriate schema, is expected to efficiently extract the desired three-dimensional audiovisual material. The algorithm requires the existence of proper indexes that accelerate the search process by using the content's visual and audio high-level semantic features combined with certain low level features. Moreover, the search based on high level features focuses on the movements and reactions of humans in the available three-dimensional audiovisual content. The whole process is enriched with a Relevance Feedback mechanism to supply users only with the most relevant results.

Finally, a user interface was studied and created. User's interface enables a user to easily query the multimodal media database in order to retrieve audiovisual content. The interface guides the user to form queries mostly based on high level features.

Keywords: multimodal audiovisual content, multimedia indexing, multimedia database, audiovisual content search and retrieval

ΕΥΧΑΡΙΣΤΙΕΣ

Θέλω να ευχαριστήσω τον επιβλέποντα καθηγητή της συγκεκριμένης διπλωματικής εργασίας, κ. Συμεών Παπαβασιλείου, ο οποίος μου εμπιστεύθηκε ένα ιδιαίτερα ενδιαφέρον και απαιτητικό επιστημονικό έργο. Το ενδιαφέρον που έδειξε και η επιστημονική καθοδήγησή του συνέβαλλαν σημαντικά στην περάτωση της παρούσας εργασίας.

Επιπλέον, θα ήθελα να ευχαριστήσω ιδιαίτερα τους υποψήφιους διδάκτορες κ. Ειρήνη-Ελένη Τσιροπούλου και κ. Βασίλη Μερεκούλια του Εργαστηρίου Διαχείρισης και Βέλτιστου Σχεδιασμού Δικτύων (Netmode), καθώς υπήρξαν πολύτιμοι συνεργάτες καθ' όλη τη διάρκεια εκπόνησης της εργασίας. Ο χρόνος που αφιέρωσαν, οι συμβουλές και οι γνώσεις τους υπήρξαν καθοριστικοί παράγοντες για την ολοκλήρωση του έργου.

Αφιερώνεται στους γονείς μου
Θεόδωρο και Ute

ΠΕΡΙΕΧΟΜΕΝΑ

ΠΕΡΙΕΧΟΜΕΝΑ	15
ΕΥΡΕΤΗΡΙΟ ΣΧΗΜΑΤΩΝ	17
ΕΥΡΕΤΗΡΙΟ ΠΙΝΑΚΩΝ	19
ΠΙΝΑΚΑΣ ΑΚΡΩΝΥΜΙΩΝ	21
ΚΕΦΑΛΑΙΟ 1 ^ο	24
ΕΙΣΑΓΩΓΗ	25
1.1 Πρόλογος	26
1.2 Αντικείμενο της διπλωματικής εργασίας	27
1.3 Δομή της διπλωματικής εργασίας	29
ΚΕΦΑΛΑΙΟ 2 ^ο	30
ΤΟ ΠΡΟΤΥΠΟ MPEG-7	31
2.1 Εισαγωγή	32
2.2 Λόγοι υιοθέτησης του προτύπου MPEG-7	33
2.3 Περιγραφή προτύπου MPEG-7	34
2.3.1 Περιγραφείς (Descriptors)	34
2.3.2 Σχήματα περιγραφής (Description Schemes)	35
2.3.3 Γλώσσα ορισμού περιγραφής (Description Definition Language)	35
2.3.4 Εργαλεία συστήματος (System Tools)	35
2.4 Οργάνωση των πολυμεσικών σχημάτων περιγραφής	35
ΚΕΦΑΛΑΙΟ 3 ^ο	38
ΣΧΕΔΙΑΣΜΟΣ ΤΗΣ ΒΑΣΗΣ ΔΕΔΟΜΕΝΩΝ	39
3.1 Εισαγωγή	40
3.2 Απαιτήσεις σχεδιασμού της βάσης δεδομένων	41
3.2.1 Λειτουργικές απαιτήσεις	41
3.2.2 Μη λειτουργικές απαιτήσεις	42
3.3 Σχεδίαση της βάσης δεδομένων	43
3.3.1 Στρατηγική σχεδίασης	43
3.3.2 Σχεσιακό διάγραμμα	43
3.3.3 Περιγραφή πινάκων	44
3.4 Καταχώρηση πληροφορίας στη βάση δεδομένων	47
3.5 Επιλεχθέν Σύστημα Διαχείρισης Βάσεων Δεδομένων	48
ΚΕΦΑΛΑΙΟ 4 ^ο	50
ΜΗΧΑΝΙΣΜΟΣ ΑΝΑΖΗΤΗΣΗΣ	51
4.1 Εισαγωγή	52
4.2 Μηχανισμός αναζήτησης	53
4.3 Μηχανισμός ανατροφοδότησης σχετικότητας	59
4.4 Τεχνικές λεπτομέρειες υλοποίησης	61

ΚΕΦΑΛΑΙΟ 5°	64
ΣΧΕΔΙΑΣΜΟΣ ΓΡΑΦΙΚΗΣ ΔΙΕΠΑΦΗΣ ΧΡΗΣΤΗ	65
5.1 Εισαγωγή	66
5.2 Καθορισμός λειτουργικών απαιτήσεων	67
5.3 Καθορισμός μη λειτουργικών απαιτήσεων	72
5.4 Τεχνικές λεπτομέρειες υλοποίησης	74
ΚΕΦΑΛΑΙΟ 6°	78
ΣΥΜΠΕΡΑΣΜΑΤΑ & ΑΠΟΤΕΛΕΣΜΑΤΑ	79
6.1 Ανάλυση σεναρίων χρήσης	80
6.1.1 Σενάριο αναζήτησης εμφάνισης ενός ηθοποιού	80
6.1.2 Σενάριο αξιολόγησης αποτελέσματος και μετέπειτα αναζητήσεων	82
6.2 Εκτέλεση αλγορίθμου αναζήτησης	83
6.2.1 Σενάριο αναζήτησης εμφάνισης ενός ηθοποιού	85
6.2.2 Σενάριο αναζήτησης ταυτόχρονης εμφάνισης δύο ηθοποιών	87
ΚΕΦΑΛΑΙΟ 7°	90
ΕΠΙΛΟΓΟΣ & ΜΕΛΛΟΝΤΙΚΗ ΕΡΓΑΣΙΑ	91
7.1 Επίλογος	92
7.2 Μελλοντική εργασία	93
ΒΙΒΛΙΟΓΡΑΦΙΑ	95

ΕΥΡΕΤΗΡΙΟ ΣΧΗΜΑΤΩΝ

Σχήμα 1: Αρχιτεκτονική του συστήματος αναζήτησης	28
Σχήμα 2: Βασικά εργαλεία του MPEG-7	34
Σχήμα 3: Επισκόπηση των πολυμεσικών σχημάτων περιγραφής του MPEG-7	36
Σχήμα 4: Το υποσύστημα της βάσης δεδομένων	40
Σχήμα 5: Το σχεσιακό σχήμα της βάσης δεδομένων	44
Σχήμα 6: Το υποσύστημα του μηχανισμού αναζήτησης	52
Σχήμα 7: Σύγκριση πολυπλοκοτήτων αλγορίθμων	57
Σχήμα 8: Ο μηχανισμός ανατροφοδότησης σχετικότητας	60
Σχήμα 9: Το υποσύστημα της γραφικής διεπαφής χρήστη	66
Σχήμα 10: Λειτουργική απαίτηση προβολής οπτικοακουστικού αρχείου	67
Σχήμα 11: Λειτουργική απαίτηση καθορισμού κριτηρίων αναζήτησης	68
Σχήμα 12: Λειτουργική απαίτηση προβολής αποτελεσμάτων αναζήτησης	69
Σχήμα 13: Λειτουργική απαίτηση προβολής ενός αποτελέσματος αναζήτησης	70
Σχήμα 14: Λειτουργική απαίτηση καταχώρησης πολυμεσικού αρχείου	71
Σχήμα 15: Λειτουργική απαίτηση επεξεργασίας υπάρχουσας πληροφορίας	72
Σχήμα 16: Καθορισμός κριτηρίου αναζήτησης για εμφάνιση ενός ηθοποιού	80
Σχήμα 17: Προβολή των αποτελεσμάτων της αναζήτησης	81
Σχήμα 18: Προβολή του χώριου του αποτελέσματος αναζήτησης	82
Σχήμα 19: Αναδιάταξη των αποτελεσμάτων αναζήτησης μετά από αξιολόγηση	83

ΕΥΡΕΤΗΡΙΟ ΠΙΝΑΚΩΝ

Πίνακας 1: Εγγραφές της σχέσης ActorAppearances	84
Πίνακας 2: Εγγραφές της σχέσης Actors	84
Πίνακας 3: Εγγραφές της σχέσης RelevanceFeedback	84
Πίνακας 4: Αποτέλεσμα της πράξης ActorAppearances x Actors	86
Πίνακας 5: Αποτελέσματα εμφωλευμένου ερωτήματος	88
Πίνακας 6: Αποτελέσματα συνολικού ερωτήματος	88
Πίνακας 7: Τιμές μεταβλητών για τις επαναλήψεις του αλγορίθμου αναζήτησης	89

ΠΙΝΑΚΑΣ ΑΚΡΩΝΥΜΙΩΝ

κ.α.	και άλλα
κ.τ.λ.	και τα λοιπά
π.χ.	παραδείγματος χάριν
ΣΔΒΔ	Σύστημα Διαχείρισης Βάσεων Δεδομένων
ANSI	American National Standards Institute
API	Application Programming Interface
CLI	Command Line Interface
CSS	Cascading Style Sheets
DBMS	DataBase Management System
DDL	Data Definition Language
D	Descriptor
DS	Descriptor Scheme
GUI	Graphical User Interface
HTML	HyperText Markup Language
HTTP	HyperText Transmission Protocol
IEC	International Electrotechnical Commission
IETF	Internet Engineering Task Force
IIS	Internet Information Services
ISO	International Organization for Standardization
JSON	JavaScript Object Notation
MDS	Multimedia Description Scheme
MPEG	Moving Picture Experts Group
NTSC	National Television System Committee
PAL	Phase Alternating Line
PHP	PHP: Hypertext Preprocessor

RFC	Request For Comments
SMPTE	Society of Motion Picture and Television Engineers
SQL	Structured Query Language
TV	TeleVision
URI	Uniform Resource Identifier
URL	Uniform Resource Locator
URN	Uniform Resource Name
W3C	World Wide Web Consortium
XML	eXtended Markup Language
3D	Three Dimensional

ΚΕΦΑΛΑΙΟ 1ο

ΕΙΣΑΓΩΓΗ

1.1 Πρόλογος

1.2 Αντικείμενο της διπλωματικής εργασίας

1.3 Δομή της διπλωματικής εργασίας

1.1 Πρόλογος

Είναι γεγονός πως η εποχή που διανύουμε χαρακτηρίζεται από ραγδαίες τεχνολογικές εξελίξεις τόσο στον τομέα του λογισμικού (software) όσο και του υλικού (hardware). Ένας από τους κλάδους που γνωρίζει μεγάλη άνθηση, λόγω των παραπάνω εξελίξεων, είναι αυτός των πολυμέσων (multimedia). Το χαμηλό κόστος κατασκευής υλικού μειώνει παράλληλα το κόστος των συσκευών λήψης πολυμεσικού περιεχομένου (multimedia content), δίνοντας τη δυνατότητα στον τελικό χρήστη να δημιουργήσει εύκολα τη δική του συλλογή οπτικοακουστικού υλικού. Παράλληλα, η εξάπλωση των ευρυζωνικών συνδέσεων με τις υψηλές ταχύτητες που προσφέρουν, καθιστούν την ανταλλαγή πολυμεσικού περιεχομένου πολύ εύκολη, γρήγορη και οικονομική, γεγονός που φαίνεται και από νέες παρεχόμενες υπηρεσίες όπως το triple play [1] και το video on demand [2]. Επίσης, η ανταλλαγή περιεχομένου μεταξύ δύο ή περισσότερων χρηστών καθίσταται ακόμα πιο εύκολη μέσα από ψηφιακές πλατφόρμες όπως οι Youtube, Vimeo, Dailymotion και άλλες, στις οποίες συγκεντρώνεται πλέον τεράστιος όγκος περιεχομένου.

Η μετάβαση από το Web 1.0 [3] στο Web 2.0 [3], [4] προκάλεσε ουσιαστικά επανάσταση στην έως τότε μορφή του Παγκόσμιου Ιστού, αφού οι χρήστες από παθητικοί αναγνώστες απέκτησαν ενεργητικό ρόλο δημιουργώντας περιεχόμενο. Η μεταβολή αυτή όμως δημιούργησε και νέες απαιτήσεις. Ο τεράστιος όγκος περιεχομένου, που πλέον είναι δυναμικός, καθίσταται άχρηστος αν δεν μπορούν οι χρήστες να εντοπίσουν αυτό που επιθυμούν, συνεπώς έγινε προφανής η ανάγκη για αποδοτική και εύστοχη αναζήτηση. Η ανάγκη αυτή έχει πλέον ικανοποιηθεί ως ένα βαθμό (με αρκετά προβλήματα και συμβιβασμούς) από διάφορους αλγόριθμους και μηχανισμούς αναζήτησης, οι οποίοι βασίζονται κυρίως στο κείμενο. Η παρούσα κατάσταση που διαμορφώνεται στον κλάδο των πολυμέσων συνδέεται και έχει πολλές ομοιότητες με την εξέλιξη του Παγκόσμιου Ιστού που περιγράφηκε: η δυνατότητα της μαζικής παραγωγής πολυμεσικού περιεχομένου συνοδεύεται από την ανάγκη της αποδοτικής και εύστοχης αναζήτησης. Οι μέχρι τώρα δυνατότητες αναζήτησης σε τέτοιο περιεχόμενο βασίζονται κυρίως στο κείμενο (χρήση λέξεων-κλειδιών), προσαρμόζοντας ουσιαστικά τους υπάρχοντες αλγόριθμους και μηχανισμούς με αποτέλεσμα τη χαμηλή απόδοση και ευστοχία. Η ερευνητική προσπάθεια στην κατεύθυνση της αναζήτησης επικεντρώνεται στην εκμετάλλευση πληροφορίας, η οποία είναι εγγενής στο περιεχόμενο. Πιο συγκεκριμένα η αναζήτηση κινείται σε δύο άξονες: αναζήτηση με βάση χαμηλού επιπέδου χαρακτηριστικά, όπως είναι πληροφορίες χρώματος και σημασιολογικά χαρακτηριστικά υψηλού επιπέδου, όπως πρόσωπα, αντικείμενα και γεγονότα.

Ένας σημαντικός κλάδος του τομέα των πολυμέσων, που έχει επανέλθει δυναμικά τα τελευταία χρόνια στο προσκήνιο, είναι αυτός του τρισδιάστατου οπτικοακουστικού υλικού (three dimensional audiovisual content). Η τρισδιάστατη απεικόνιση περιεχομένου δεν είναι καινούργια πρακτική, με τις πρώτες προσπάθειες να χρονολογούνται στο 1915,

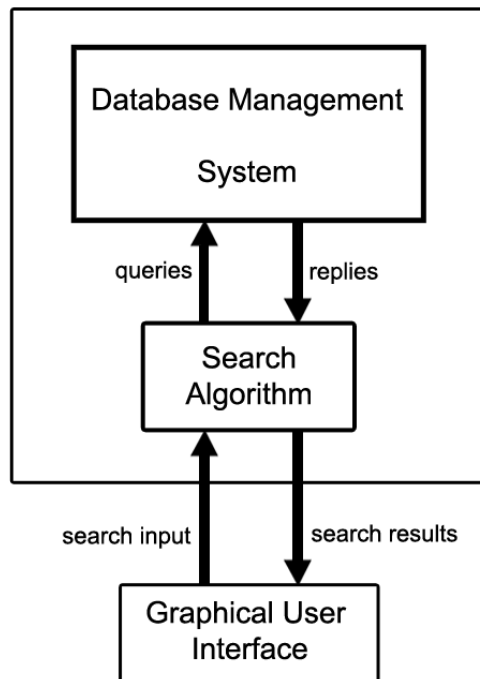
ωστόσο το ενδιαφέρον περιορίστηκε κυρίως λόγω του κόστους του υλικού (hardware) και των απαιτούμενων διαδικασιών. Τις δεκαετίες του '80 και του '90 το ενδιαφέρον αναζωπυρώθηκε και κορυφώθηκε στα τέλη του 2009 με την επιτυχία της τρισδιάστατης ταινίας Avatar [5]. Το κόστος δεν αποτελεί πλέον ανασταλτικό παράγοντα και ήδη υπάρχει μεγάλος όγκος τέτοιου περιεχομένου. Επιπρόσθετα, υπάρχουν πλέον στην αγορά συσκευές αναπαραγωγής και λήψης τέτοιου περιεχομένου που απευθύνονται σε απλούς χρήστες, οπότε αναμένεται έκρηξη της ποσότητας τρισδιάστατου οπτικοακουστικού υλικού, το οποίο θα αναρτάται σε κοινωνικά δίκτυα ή πλατφόρμες όπως το Youtube. Γίνεται σαφές, λοιπόν, ότι η αναζήτηση σε πολυμεσικό περιεχόμενο θα πρέπει να επεκταθεί σε έναν νέο τομέα, αυτόν του τρισδιάστατου οπτικοακουστικού περιεχομένου.

Η αναζήτηση απαιτεί ένα σύνολο πληροφορίας, το οποίο αντλείται από κάθε διαθέσιμο πολυμεσικό αρχείο και χρησιμοποιείται για την εύστοχη επιλογή περιεχομένου με βάση τις απαιτήσεις του χρήστη. Η πληροφορία αυτή προκύπτει μέσα από την ανάλυση του περιεχομένου και γίνεται αυτοματοποιημένα μέσω αλγορίθμων για χαμηλού επιπέδου χαρακτηριστικά και με ανθρώπινη παρέμβαση για υψηλού επιπέδου χαρακτηριστικά. Στο δισδιάστατο οπτικοακουστικό αρχείο υπάρχει ένα κανάλι οπτικής πληροφορίας και ένα ή περισσότερα κανάλια ήχου. Η οπτική πληροφορία προκύπτει από την ανάλυση του οπτικού καναλιού και η ηχητική από τα αντίστοιχα ακουστικά κανάλια. Στην περίπτωση του τρισδιάστατου οπτικοακουστικού περιεχομένου η ανάλυση είναι πιο περίπλοκη επειδή υπάρχει επιπλέον πληροφορία και είναι αυτή της τρίτης διάστασης, δηλαδή του βάθους. Η βασικότερη τεχνική για τη δημιουργία της αίσθησης τους βάθους είναι η στερεοσκοπία (stereoscopy) [6]: χρησιμοποιούνται δύο οπτικά κανάλια, τα οποία προέρχονται από δύο κάμερες, οι οποίες βιντεοσκοπούν την ίδια σκηνή από διαφορετικές θέσεις. Στη συνέχεια με ειδικό υλικό προβολής (special projection hardware) ή με ειδικά γυαλιά δημιουργείται η ψευδαίσθηση του βάθους. Επομένως, το τρισδιάστατο υλικό απαιτεί δύο κανάλια οπτικής πληροφορίας, γεγονός που επηρεάζει τόσο την ανάλυση όσο και την αναζήτηση, αφού υπάρχει επιπλέον πληροφορία που πρέπει να αξιοποιηθεί.

1.2 Αντικείμενο της διπλωματικής εργασίας

Η αναζήτηση πολυμεσικού περιεχομένου με τεχνικές βασισμένες στο κείμενο έχουν κριθεί ανεπαρκείς και ανεπιτυχείς. Στην παρούσα διπλωματική εργασία θα επιχειρήσουμε να προσεγγίσουμε την αναζήτηση πολυτροπικού οπτικοακουστικού υλικού με μεθόδους που είναι εγγενείς προς το περιεχόμενο αυτό. Πιο συγκεκριμένα, θα εστιάσουμε στην αναζήτηση περιεχομένου με βάση υψηλού επιπέδου, σημασιολογικά χαρακτηριστικά, όπως είναι οι εμφανίσεις ηθοποιών, ενέργειες προσώπων και γεγονότα. Επιπρόσθετα, η αναζήτηση θα εμπλουτιστεί με ένα μηχανισμό ανατροφοδότησης σχετικότητας (Relevance Feedback) [7] εξασφαλίζοντας για το χρήστη τα καλύτερα δυνατά αποτελέσματα. Πρέπει

να διασαφηνιστεί σε αυτό το σημείο ότι δεν επικεντρωνόμαστε στην εξαγωγή πληροφορίας από το πολυμεσικό αρχείο, αλλά θεωρούμε δεδομένη την πληροφορία αυτή, την οποία θα κληθούμε να διαχειριστούμε κατάλληλα. Στο παρακάτω σχήμα φαίνεται η αρχιτεκτονική ολόκληρου του συστήματος που απαιτείται για να υποστηρίξει πλήρως την αναζήτηση και θα ακολουθήσει συνοπτικά η ανάλυσή του, που είναι και ουσιαστικά το αντικείμενο της διπλωματικής εργασίας.



Σχήμα 1: Αρχιτεκτονική του συστήματος αναζήτησης

Η έρευνά μας απαρτίζεται από περισσότερα σκέλη: αρχικά πρέπει να οριστεί ένας εύελκτος τρόπος περιγραφής της πληροφορίας που συνοδεύει κάθε οπτικοακουστικό αρχείο. Η περιγραφή της πληροφορίας δεν υπόκειται σε περιορισμούς, αρκεί η μορφή της να είναι προσυμφωνημένη και άρα αναγνωρίσιμη. Γι' αυτό το λόγο θα χρησιμοποιηθεί το πρότυπο MPEG-7 [8], ο σκοπός του οποίου είναι να προτυποποιήσει την περιγραφή πολυμεσικού περιεχομένου, προσφέροντας μεγάλο αριθμό έτοιμων εργαλείων.

Επόμενο στάδιο της έρευνάς μας είναι η μελέτη ενός σχήματος βάσης δεδομένων, η οποία θα φιλοξενήσει – με εύελκτο προς την αναζήτηση τρόπο – την πληροφορία από τα διάφορα αρχεία περιγραφής πολυμεσικής πληροφορίας. Πιο συγκεκριμένα, ερευνάται μια σειρά από διαθέσιμες τεχνολογίες βάσεων δεδομένων, ώστε να επιλεγθεί εκείνη που αντιστοιχεί καλύτερα στη φύση του προβλήματος. Εν συνεχεία μελετάται με λεπτομέρεια η σχεδίαση της βάσης δεδομένων, ορίζοντας με σαφήνεια τον τρόπο που αυτή θα φιλοξενήσει τα δεδομένα.

Ακολουθεί ο αλγόριθμος αναζήτησης, ο οποίος χρησιμοποιώντας το υπάρχον σύνολο πληροφορίας και με βάση την είσοδο του χρήστη, καλείται να του επιστρέψει ένα

υποσύνολο του συνολικά διαθέσιμου οπτικοακουστικού υλικού. Μελετώνται οι απαιτήσεις του χρήστη, οι οποίες θα καθορίσουν την έξοδο, αλλά και τη σχεδίαση του αλγορίθμου. Επίσης, ένα σημαντικό τμήμα του μηχανισμού αναζήτησης είναι η ανατροφοδότηση σχετικότητας (Relevance Feedback) [7] των αποτελεσμάτων, ώστε ύστερα από αξιολόγηση, οι μελλοντικές αναζητήσεις να επιστρέφουν πιο σχετικά αποτελέσματα.

Τέλος, σχεδιάζουμε και υλοποιούμε μια γραφική διεπαφή χρήστη (Graphical User Interface), η οποία πρέπει να του παρέχει ένα βολικό τρόπο πρόσβασης στην υποδομή αναζήτησης. Επιπρόσθετα, πρέπει να υπάρχει η δυνατότητα της παραστατικής παρουσίασης των αποτελεσμάτων αναζήτησης με επαρκή λεπτομέρεια. Μελετώνται, λοιπόν, οι απαιτήσεις που υπάρχουν από μια τέτοια γραφική διεπαφή και αφού γίνει αυτό παρουσιάζονται οι σχεδιαστικές μας αποφάσεις που θα καθορίσουν την τελική μορφή της διεπαφής.

1.3 Δομή της διπλωματικής εργασίας

Στην ενότητα αυτή γίνεται σύντομη αναφορά στα κεφάλαια της παρούσας διπλωματικής εργασίας. Η δομή της εργασίας ακολουθεί και τη σειρά της έρευνας, όπως αυτή αναλύθηκε στην προηγούμενη ενότητα. Το δεύτερο κεφάλαιο αναφέρεται στο πρότυπο MPEG-7, το οποίο θα χρησιμοποιηθεί για την περιγραφή της πληροφορίας που έχει αντληθεί από τα αρχεία πολυμέσων. Στο κεφάλαιο 3, αναλύεται η σχεδίαση της βάσης δεδομένων, η οποία θα φιλοξενήσει την απαραίτητη πληροφορία. Πιο συγκεκριμένα αναλύονται οι λόγοι για την επιλογή ενός Συστήματος Διαχείρισης Βάσεων Δεδομένων και στη συνέχεια εξετάζεται λεπτομερώς το σχήμα της βάσης που θα απαιτηθεί για την κάλυψη των αναγκών της εφαρμογής. Στο τέταρτο κεφάλαιο αναφέρονται οι απαιτήσεις που υπάρχουν από το μηχανισμό αναζήτησης και οι σχεδιαστικές λεπτομέρειες που υιοθετήθηκαν, ενώ αναλύεται και ο μηχανισμός ανατροφοδότησης σχετικότητας των αποτελεσμάτων. Το πέμπτο κεφάλαιο περιγράφει τις λειτουργικές και μη λειτουργικές απαιτήσεις της γραφικής διεπαφής χρήστη και τις σχεδιαστικές αποφάσεις που αφορούν την υλοποίησή της, καθώς αποτελεί σημαντικό τμήμα του συνολικού συστήματος, όντας το σημείο πρόσβασης του χρήστη στην αναζήτηση και την υπάρχουσα πληροφορία. Το κεφάλαιο 6 παρουσιάζει τα συμπεράσματα που προέκυψαν από την έρευνά μας, σχολιάζονται τα αποτελέσματα του μηχανισμού αναζήτησης, περιγράφονται κάποιες περιπτώσεις χρήσης και τέλος παρατίθενται κάποιες οθόνες (screenshots) της γραφικής διεπαφής χρήστη. Τέλος, στο έβδομο κεφάλαιο ακολουθεί ο επίλογος και γίνεται αναφορά σε μελλοντικές επεκτάσεις της παρούσας ερευνητικής εργασίας.

ΚΕΦΑΛΑΙΟ 2^ο

ΤΟ ΠΡΟΤΥΠΟ MPEG-7

2.1 Εισαγωγή

2.2 Λόγοι υιοθέτησης του προτύπου MPEG-7

2.3 Περιγραφή προτύπου MPEG-7

2.4 Οργάνωση πολυμεσικών σχημάτων περιγραφής

2.1 Εισαγωγή

Η συνεχής τεχνολογική εξέλιξη των συσκευών παραγωγής πολυμεσικού υλικού, η μείωση του κόστους τους και η διάδοση της ευρυζωνικότητας έχει οδηγήσει στην παραγωγή τεράστιου όγκου πολυμεσικού υλικού, με τον ρυθμό παραγωγής του να αυξάνεται συνεχώς και αναμένεται να αυξηθεί ακόμα περισσότερο μελλοντικά. Ωστόσο, αυτός ο όγκος πληροφορίας καθίσταται άχρηστος αν δεν μπορεί να υπάρξει αποδοτική πρόσβαση σε αυτόν εκ μέρους του χρήστη, ο οποίος πρέπει να έχει τη δυνατότητα να αναζητεί αυτό που επιθυμεί γρήγορα και με ικανοποιητική ευστοχία.

Μέχρι τώρα η αναζήτηση περιεχομένου γινόταν με βάση το κείμενο, μέσω περιγραφικών λέξεων κλειδιών (keywords) και υπομνημάτων (annotations), προσφέροντας μέτρια απόδοση, κυρίως λόγω της ανάμιξης του ανθρώπινου παράγοντα. Ο παράγοντας αυτός εισάγει μεγάλο βαθμό υποκειμενικότητας και οι περιγραφές που προκύπτουν είναι προσανατολισμένες στο είδος της εφαρμογής για την οποία δημιουργήθηκαν. Επιπλέον, η διαδικασία αυτή είναι ανεπιθύμητη γιατί είναι χρονικά δαπανηρή και το κόστος της αυξάνεται αναλογικά με τον όγκο της πληροφορίας. Συνεπώς, υπάρχει η ανάγκη της αυτοματοποιημένης και αντικειμενικής περιγραφής, της αρχειοθέτησης και του σχολιασμού του πολυμεσικού υλικού, χρησιμοποιώντας εργαλεία τα οποία εξάγουν πληροφορίες και χαρακτηριστικά από το περιεχόμενο και συμπληρώνουν ή αντικαθιστούν τις υπάρχουσες πληροφορίες που βασίζονται κατά κύριο λόγο στο κείμενο. Η αυτοματοποιημένη εξαγωγή πληροφορίας προφανώς υπερτερεί των ανθρωπίνων περιγραφών γιατί γίνεται χωρίς ανθρώπινη παρέμβαση, επειδή είναι πιο αντικειμενική και ανεξάρτητη του πεδίου εφαρμογής και τέλος επειδή μπορεί να είναι εγγενής στο οπτικοακουστικό περιεχόμενο.

Η αναγνώριση του παραπάνω προβλήματος οδήγησε την ομάδα εργασίας του MPEG (Moving Picture Experts Group) στην ανάπτυξη ενός νέου προτύπου, του MPEG-7 [8] ή αλλιώς Multimedia Content Description Interface (ISO/IEC 15938), το οποίο στοχεύει στον καθορισμό προτυποποιημένων τρόπων περιγραφής για διάφορα είδη πολυμεσικού υλικού. Οι περιγραφές αυτές μπορεί να μην περιλαμβάνουν κείμενο, όταν αναφέρονται για παράδειγμα σε πληροφορία όπως στατιστικά χαρακτηριστικά και παραμέτρους της κάμερας ή και να περιλαμβάνουν, όπως σε περιπτώσεις σχολίων/υπομνημάτων ή ονομάτων ηθοποιών κ.α. Το MPEG-7 [8], όπως και τα υπόλοιπα πρότυπα του MPEG, ορίζει μια προτυποποιημένη αναπαράσταση της πληροφορίας που θέλουμε να περιγράψουμε, ικανοποιώντας ένα σύνολο σαφώς ορισμένων απαιτήσεων. Ωστόσο, διαφέρει από τα άλλα πρότυπα του MPEG γιατί αναπαριστά πληροφορία που αφορά το περιεχόμενο (μεταπληροφορία), ενώ τα υπόλοιπα πρότυπα αναπαριστούν το ίδιο το περιεχόμενο.

2.2 Λόγοι υιοθέτησης του προτύπου MPEG-7

Στα πλαίσια της παρούσας διπλωματικής, στοχεύοντας στην έρευνα/εφαρμογή της αναζήτησης και ανάκτησης πολυτροπικού οπτικοακουστικού περιεχομένου θα κληθούμε να αντιμετωπίσουμε το πρόβλημα της περιγραφής του υπάρχοντος πολυμεσικού υλικού, προκειμένου να μπορεί να ακολουθήσει αποδοτική και εύστοχη αναζήτηση. Για την περιγραφή, λοιπόν, του διαθέσιμου υλικού θα υιοθετήσουμε το MPEG-7, καθώς προσφέρει μια σειρά σημαντικών πλεονεκτημάτων που θα αναλυθούν στη συνέχεια. Η περιγραφή πολυμεσικού υλικού προφανώς είναι δυνατή και χωρίς το MPEG-7, με τη δημιουργία συστημάτων που χρησιμοποιούν τη δική τους αναπαράσταση πληροφορίας, όμως το σημαντικότερο πλεονέκτημα που προσφέρεται από τη χρήση του προτύπου είναι αυτό της διαλειτουργικότητας.

Το MPEG-7 χρησιμοποιεί (και είναι φιλικό προς) τεχνολογίες και πρότυπα τα οποία έχουν καθιερωθεί με μεγάλη επιτυχία και είναι ευρέως αποδεκτά τόσο στις παραδοσιακές πολυμεσικές εφαρμογές όσο και στις νεότερες. Τα σημαντικότερα από αυτά είναι τα ακόλουθα [9]:

- XML [10] και XML Schema [11] (από World Wide Web Consortium, W3C)
- URI, URN και URL (από Internet Engineering Task Force, IETF)
- Dublin Core [12]
- ISO/ANSI Thesaurus Guidelines
- SMPTE Metadata Dictionary [13]
- TV-Anytime [14]

Αυτό ερμηνεύει τη διαλειτουργικότητα που προσφέρει το πρότυπο MPEG-7, διευκολύνοντας την ενσωμάτωση και χρησιμοποίηση από υπάρχοντα συστήματα και εφαρμογές.

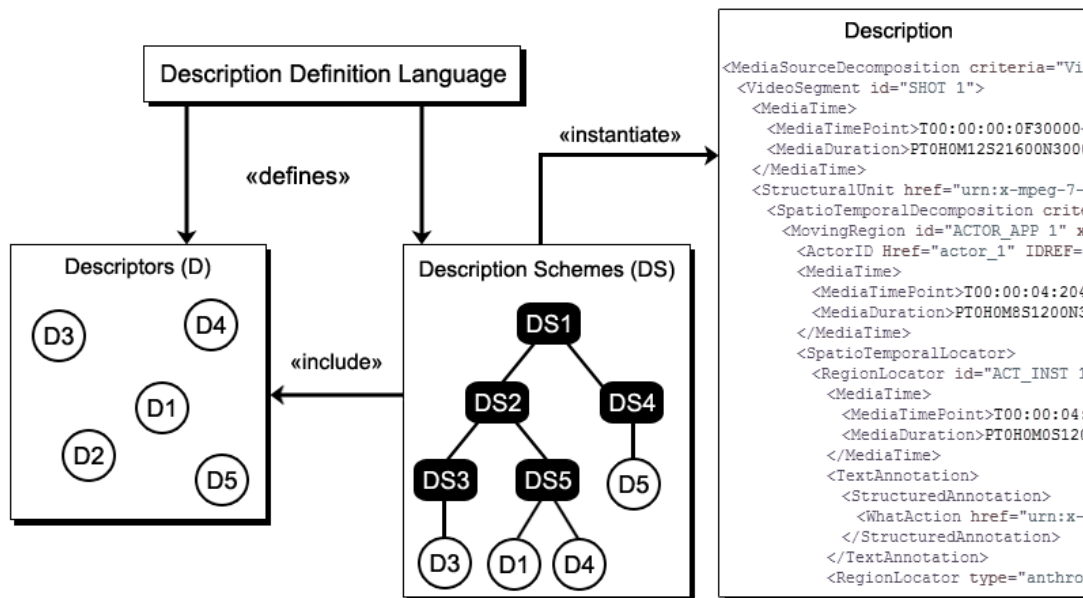
Σημαντικό πλεονέκτημα είναι και η δυνατότητα επεκτασιμότητας. Συγκεκριμένα, επιτρέπει την επέκταση των βασικών εργαλείων περιγραφής (description tools) με προτυποποιημένο τρόπο, έτσι ώστε να μπορούν να σχεδιαστούν δομές για την αντιμετώπιση νέων αναγκών μελλοντικά. Επίσης, υπάρχουν διαφορετικά επίπεδα αφαίρεσης, με περιγραφές χαρακτηριστικών από χαμηλού (π.χ. στατιστικά στοιχεία χρωμάτων) μέχρι υψηλού επιπέδου (χαρακτηριστικά με σημασιολογική έννοια, π.χ. πρόσωπα). Άλλο σημαντικό στοιχείο είναι η αντικειμενοστρέφεια: η περιγραφή του περιεχομένου γίνεται με αντικειμενοστραφή τρόπο και το περιεχόμενο μπορεί να θεωρηθεί ως μια σύνθεση πολυμεσικών αντικειμένων με δυνατή την ανεξάρτητη πρόσβαση των περιγραφών συγκεκριμένων αντικειμένων του περιεχομένου. Ακόμα, υπάρχει ανεξαρτησία της περιγραφής από την αναπαράσταση του περιεχομένου, είτε αυτό βρίσκεται σε αναλογική είτε σε ψηφιακή μορφή. Επιπρόσθετα, μπορεί να αναπαρίσταται σε μορφή PAL

(Phase Alternate Line), NTSC (National Television Standards Committee), MPEG-1, MPEG-2 ή MPEG-4 χωρίς να επηρεάζεται η περιγραφή.

Ένας μη τεχνικός παράγοντας που συντελεί στην επιλογή του MPEG-7, είναι η μέχρι τώρα επιτυχία των υπόλοιπων προτύπων που έχουν προταθεί από το MPEG, όντας ένα είδος εγγύησης. Τα μέχρι τώρα πρότυπά του γνωρίζουν μεγάλη επιτυχία στην αγορά κυρίως λόγω της ακολουθούμενης διαδικασίας, να διαμοιράζεται τεχνολογία αιχμής με παράλληλη μέριμνα για την προστασία της πνευματικής ιδιοκτησίας. Τέλος, όπως και με τα άλλα πρότυπα, η διαδικασία ανάπτυξης είναι ιδιαίτερα σχολαστική με τη συμμετοχή εμπειρογνομόνων σε όλα τα στάδια της ανάπτυξης, γεγονός που εγγυάται ένα ολοκληρωμένο πρότυπο ικανό να ανταπεξέλθει σε πολλά διαφορετικά και απαιτητικά είδη εφαρμογών.

2.3 Περιγραφή προτύπου MPEG-7

Οι στόχοι του προτύπου επιτυγχάνονται με τον ορισμό ενός συνόλου εργαλείων. Τα εργαλεία αυτά είναι οι περιγραφείς (Descriptors), τα σχήματα περιγραφής (Description Schemes), η γλώσσα ορισμού περιγραφής (Description Definition Language, DDL) και κάποια εργαλεία συστήματος (System Tools). Στο σχήμα που παρατίθεται απεικονίζεται η σχέση μεταξύ των εργαλείων αυτών.



Σχήμα 2: Βασικά εργαλεία του MPEG-7 [16]

2.3.1 Περιγραφείς (Descriptors)

Ένας περιγραφέας είναι η αναπαράσταση μιας ιδιότητας, όπου ως ιδιότητα ορίζεται ένα διακριτό χαρακτηριστικό του οπτικοακουστικού υλικού. Ο περιγραφέας ορίζει την σύνταξη και τη σημασιολογία της αναπαράστασης ενός χαρακτηριστικού και επιτρέπει την

ανάθεση μιας τιμής στο χαρακτηριστικό αυτό, μέσω της τιμής του περιγραφέα. Έχουν σχεδιαστεί για να περιγράφουν χαμηλού επιπέδου οπτικοακουστικά χαρακτηριστικά, όπως χρώμα, υφή, αλλά και υψηλού επιπέδου χαρακτηριστικά, όπως αντικείμενα και γεγονότα [17],[18]. Κατά κύριο λόγο τα χαμηλού επιπέδου χαρακτηριστικά εξάγονται με αυτοματοποιημένο τρόπο, ενώ για τα υψηλού επιπέδου απαιτείται η παρέμβαση ανθρώπινου παράγοντα.

2.3.2 Σχήματα περιγραφής (Description Schemes)

Τα σχήματα περιγραφής επεκτείνουν τους MPEG-7 περιγραφείς συνδυάζοντας μεμονωμένους περιγραφείς και άλλα σχήματα περιγραφής σε πιο πολύπλοκες δομές. Ένα σχήμα περιγραφής ορίζει τη δομή και σημασιολογία των σχέσεων μεταξύ των συστατικών του, τα οποία μπορεί να είναι περιγραφείς και σχήματα περιγραφής [17]. Είναι ένας τρόπος μοντελοποίησης και περιγραφής πολυμεσικού περιεχομένου από πλευράς δομής και σημασιολογίας. Ένα απλό παράδειγμα είναι μια ταινία που μπορεί να διαιρεθεί χρονικά σε σκηνές (scenes) και λήψεις (shots).

2.3.3 Γλώσσα ορισμού περιγραφής (Description Definition Language)

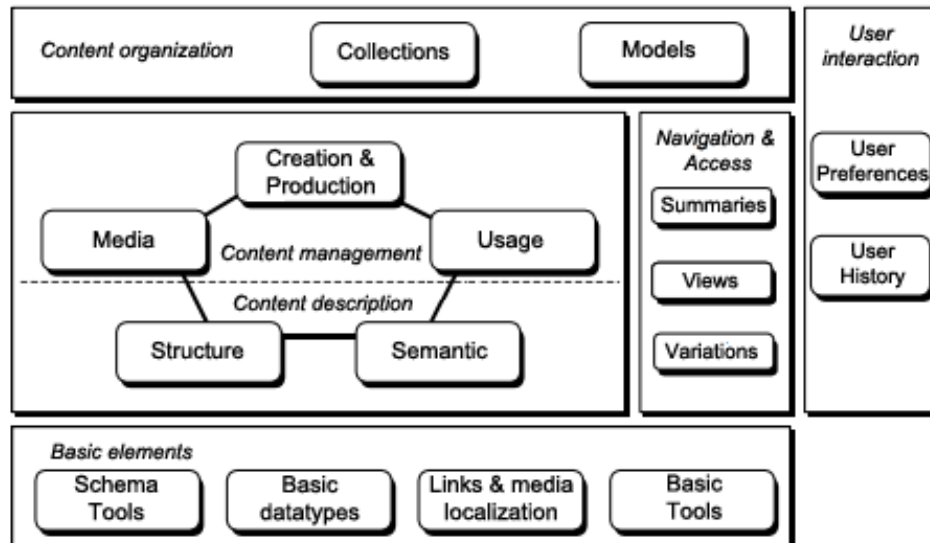
Η γλώσσα ορισμού περιγραφής επιτρέπει τη δημιουργία νέων σχημάτων περιγραφής και περιγραφών, καθώς και την επέκταση και τροποποίηση των υπάρχοντων σχημάτων περιγραφής. Η γλώσσα αυτή είναι μια επέκταση του XML Schema, που έχει οριστεί από το World Wide Web Consortium (W3C).

2.3.4 Εργαλεία συστήματος (System Tools)

Τα εργαλεία συστήματος σχετίζονται με το συγχρονισμό, τη μεταφορά και αποθήκευση των περιγραφών και τη διαχείριση και προστασία της πνευματικής ιδιοκτησίας.

2.4 Οργάνωση των πολυμεσικών σχημάτων περιγραφής

Τα πολυμεσικά σχήματα περιγραφής (Multimedia Description Schemes, MDS) του MPEG-7 (ISO/IEC 15938-5) ορίζουν τα διαφορετικά εργαλεία περιγραφής, που δεν αναφέρονται αποκλειστικά στον οπτικό ή ακουστικό τομέα, αφορούν το περιεχόμενο γενικά και μπορούν να κατηγοριοποιηθούν με βάση τη λειτουργικότητα που προσφέρουν στις εξής περιοχές: Βασικά στοιχεία (Basic elements), Περιγραφή περιεχομένου (Content description), Διαχείριση περιεχομένου (Content management), Οργάνωση περιεχομένου (Content organization), Πλοήγηση και Πρόσβαση (Navigation and Access) και στην Αλληλεπίδραση χρήστη (User interaction).



Σχήμα 3: Επισκόπηση των πολυμεσικών σχημάτων περιγραφής του MPEG-7 [9]

Ακολουθεί μια συνοπτική παρουσίαση των παραπάνω περιγραφικών εργαλείων [17].

- **Βασικά στοιχεία (Basic elements):** είναι γενικότερες οντότητες που χρησιμοποιούνται ως δομικά στοιχεία από άλλα εργαλεία περιγραφών και περιλαμβάνουν βασικούς τύπους δεδομένων (όπως αριθμούς, πίνακες, διανύσματα), συνδέσμους και εργαλεία περιγραφών για τοποθεσίες, άτομα, σχόλια κ.τ.λ.
- **Εργαλεία περιγραφής περιεχομένου (Content description tools):** χρησιμοποιούνται για την αναπαράσταση πληροφορίας και σε αυτά περιλαμβάνονται εργαλεία για την περιγραφή δομικής (Structure description tools) και σημασιολογικής πληροφορίας (Semantic description tools). Τα δομικά περιγραφικά εργαλεία επιτρέπουν το διαχωρισμό του περιεχομένου σε χωροχρονικά χωρία, οργανωμένα ιεραρχικά, ενώ τα σημασιολογικά περιγραφικά εργαλεία περιγράφουν το περιεχόμενο με υψηλού επιπέδου έννοιες, όπως αντικείμενα και γεγονότα.
- **Εργαλεία διαχείρισης περιεχομένου (Content management description tools):** είναι εργαλεία σχετικά με την αρχειοθέτηση του περιεχομένου, προσφέροντας πληροφορίες παραγωγής και χρήσης, κωδικοποίησης, ποιότητας, ταξινόμησης κ.τ.λ.
- **Εργαλεία οργάνωσης περιεχομένου (Content organization description tools):** επιτρέπουν τη δημιουργία και μοντελοποίηση συλλογών από πολυμεσικό περιεχόμενο και περιγραφές.

- **Εργαλεία πλοήγησης και πρόσβασης (Navigation and Access description tools):** χρησιμοποιούνται για τον ορισμό περιλήψεων, τμημάτων και αποσυνθέσεων του περιεχομένου διευκολύνοντας στην περιήγηση και ανάκτησή του.
- **Εργαλεία αλληλεπίδρασης χρήστη (User interaction description tools):** επιτρέπουν την περιγραφή των επιλογών και προτιμήσεων του χρήστη, καθώς και τη διατήρηση κάποιου ιστορικού σχετικού με τη χρήση του περιεχομένου, προκειμένου να διευκολύνουν τυχόν αναζητήσεις.

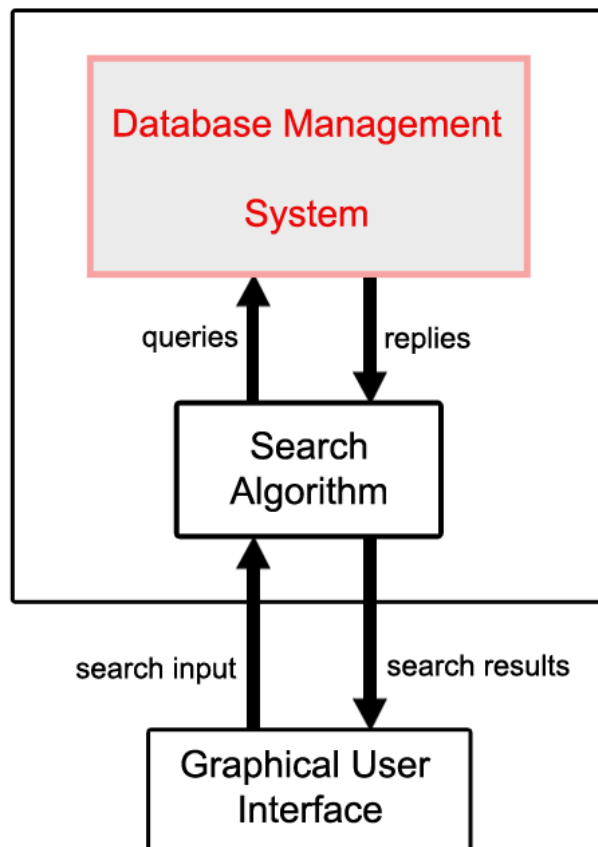
ΚΕΦΑΛΑΙΟ 3^ο

ΣΧΕΔΙΑΣΜΟΣ ΤΗΣ ΒΑΣΗΣ ΔΕΔΟΜΕΝΩΝ

-
- 3.1 Εισαγωγή**
 - 3.2 Απαιτήσεις σχεδιασμού της βάσης δεδομένων**
 - 3.3 Σχεδίαση της βάσης δεδομένων**
 - 3.4 Καταχώρηση πληροφορίας στη βάση δεδομένων**
 - 3.5 Επιλεγθέν Σύστημα Διαχείρισης Βάσεων Δεδομένων**
-

3.1 Εισαγωγή

Το σύστημα που θα αναπτυχθεί στα πλαίσια της παρούσας διπλωματικής εργασίας θα περιλαμβάνει μια συλλογή πολυτροπικού οπτικοακουστικού περιεχομένου, μέσα από την οποία θα έχουμε τη δυνατότητα αναζήτησης συγκεκριμένων υποσυνόλων που πληρούν κάποια σαφώς καθορισμένα κριτήρια. Πρωτού υλοποιηθούν οι μηχανισμοί αναζήτησης, το διαθέσιμο σύνολο πληροφορίας πρέπει να αποθηκευθεί με κάποιον αποδοτικό και ευέλικτο τρόπο για τη μετέπειτα αξιοποίησή του. Στο ρόλο αυτόν καλείται να ανταποκριθεί κάποιο Σύστημα Διαχείρισης Βάσεων Δεδομένων (DataBase Management System) [19] που είναι και το αντικείμενο έρευνας αυτού του κεφαλαίου, όπως φαίνεται στο ακόλουθο σχήμα.



Σχήμα 4: Το υποσύστημα της βάσης δεδομένων

Η πληροφορία που πρόκειται να αποθηκευτεί στη βάση δεδομένων βρίσκεται κατά κύριο λόγο στα αρχεία MPEG-7 XML [10] και αφορούν την περιγραφή του περιεχομένου των οπτικοακουστικών αρχείων. Επιπρόσθετα, θα χρειαστούν κάποιες επιπλέον πληροφορίες για την ορθή και αποδοτική λειτουργία του συστήματος, όπως είναι τεχνικές πληροφορίες για τα ίδια τα οπτικοακουστικά αρχεία. Οι πληροφορίες αυτές είναι δευτερεύουσας σημασίας, ωστόσο είναι απαραίτητες.

3.2 Απαιτήσεις σχεδιασμού της βάσης δεδομένων

Με σκοπό την επιλογή ενός Συστήματος Διαχείρισης Βάσεων Δεδομένων [19], αλλά και το σχεδιασμό ενός κατάλληλου σχήματος βάσης πρέπει πρώτα να καθοριστούν οι απαιτήσεις που υπάρχουν και αφορούν την αποθήκευση της διαθέσιμης πληροφορίας. Οι απαιτήσεις μπορούν να καταταγούν σε δύο κατηγορίες: τις λειτουργικές (functional requirements) και τις μη λειτουργικές απαιτήσεις (non-functional requirements), οι οποίες θα εξεταστούν στη συνέχεια. Ένα σημαντικό ζήτημα είναι η βασική αρχιτεκτονική του Συστήματος Διαχείρισης Βάσεων Δεδομένων, όπου οι επιλογές είναι περιορισμένες. Οι δύο πιο ταιριαστές λύσεις για το είδος της εφαρμογής μας είναι τα παραδοσιακά Σχεσιακά Συστήματα Διαχείρισης Βάσεων Δεδομένων (Relational DataBase Management Systems) και τα «Καθαρά XML» Συστήματα Διαχείρισης Βάσεων Δεδομένων (Native XML DataBase Management Systems) [20]. Η πρώτη επιλογή είναι ευρέως διαδεδομένη και βασίζεται στη χρήση σχεσιακών σχημάτων, όπου τα δεδομένα αποθηκεύονται σε πίνακες (σχέσεις), ενώ στη δεύτερη επιλογή τα δεδομένα βρίσκονται σε μορφή XML [10] και το βασικό μοντέλο αποθήκευσης της πληροφορίας είναι η συλλογή (collection). Στη δική μας περίπτωση θα χρησιμοποιήσουμε ένα Σχεσιακό Σύστημα Διαχείρισης Βάσεων Δεδομένων, καθώς αποτελεί μια δοκιμασμένη λύση, προσφέροντας σημαντικό βαθμό αξιοπιστίας. Επίσης, υπάρχουν πιο πολλές εναλλακτικές για την τελική επιλογή του συστήματος που θα χρησιμοποιήσουμε.

3.2.1 Λειτουργικές απαιτήσεις

Οι λειτουργικές απαιτήσεις (functional requirements) που έχουμε από το Σύστημα Διαχείρισης Βάσεων Δεδομένων δεν απαιτούν ιδιαίτερη ανάλυση, καθώς είναι οι στοιχειώδεις λειτουργίες που προσφέρει οποιοδήποτε τέτοιο πλήρες σύστημα. Πιο συγκεκριμένα, η βασικότερη απαίτηση είναι η δυνατότητα καταχώρισης/αποθήκευσης και ανάκτησης πληροφοριών με την ευελιξία που προσφέρουν εξειδικευμένες γλώσσες γι' αυτόν το σκοπό, όπως για παράδειγμα η SQL (Structured Query Language) [21]. Η SQL είναι μια από τις πιο διαδεδομένες γλώσσες ερωτημάτων σε βάσεις δεδομένων και η υποστήριξή της είναι η σημαντικότερη απαίτηση.

Δεδομένου ότι τα αρχεία περιγραφής του οπτικοακουστικού περιεχομένου είναι σε XML, θα θέλαμε το Σύστημα Διαχείρισης Βάσεων Δεδομένων να παρέχει υποστήριξη για XML δεδομένα. Τα περιεχόμενα ενός XML εγγράφου θα μπορούσαν να αποθηκευτούν σαν κείμενο σε έναν πίνακα, ωστόσο αυτό έχει μικρή αξία λόγω της απώλειας της δομής του XML δέντρου, με το οποίο αναπαρίσταται κάθε έγγραφο. Υπάρχουν Συστήματα Διαχείρισης Βάσεων Δεδομένων που προσφέρουν δυνατότητες αναζήτησης εγγενείς στο XML, λαμβάνοντας υπόψη τα ιδιαίτερα χαρακτηριστικά ενός τέτοιου εγγράφου όπως είναι η ιεραρχία. Συνήθως τα συστήματα αυτά εμπλουτίζουν τη γλώσσα SQL με χαρακτηριστικά της γλώσσας XPath. Η XPath [22] είναι μια γλώσσα ερωτημάτων για την επιλογή κόμβων

μέσα σε XML έγγραφα. Η απαίτηση αυτή δεν είναι καθοριστική για την επιλογή μας, ωστόσο θα θέλαμε το σύστημά μας να υποστηρίζει XPath [22] ερωτήματα, κυρίως με την προοπτική της μελλοντικής επέκτασής του και τον εμπλουτισμό με περισσότερες και πιο σύνθετες δυνατότητες.

Τέλος, μια άλλη δυνατότητα που θέλουμε να παρέχεται είναι αυτή της απομακρυσμένης πρόσβασης στο Σύστημα Διαχείρισης Βάσεων Δεδομένων. Αυτή η απαίτηση δε σχετίζεται άμεσα με τη λειτουργία του συστήματος, αλλά αφορά την ευκολία διαχείρισης και τη μελλοντική συντήρηση.

3.2.2 Μη λειτουργικές απαιτήσεις

Στην κατηγορία των μη λειτουργικών απαιτήσεων (non-functional requirements) του Συστήματος Διαχείρισης Βάσεων Δεδομένων κατατάσσονται κυρίως ποιοτικά μεγέθη που θέλουμε να το χαρακτηρίζουν.

Θέλουμε το σύστημα να είναι συμβατό με το τετράπτυχο απαιτήσεων: ατομικότητα, συνέπεια, απομόνωση και μονιμότητα (atomicity, consistency, isolation, durability) [23]. Αναλυτικότερα, πρέπει οι συναλλαγές με το σύστημα της βάσης να είναι ατομικές, δηλαδή να υπάρχουν εκείνοι οι μηχανισμοί που να εξασφαλίζουν ότι δεν υπάρχουν ημιτελείς συναλλαγές. Σε περίπτωση απόρριψης ή αποτυχίας μιας συναλλαγής όλες οι αλλαγές αναιρούνται. Η απομόνωση εξασφαλίζει ότι η εναλλαγή στην εκτέλεση των λειτουργιών των διάφορων συναλλαγών είναι ισοδύναμη με τη σειριακή εκτέλεση των συναλλαγών. Έτσι έχουμε βελτίωση της απόδοσης με το ισοδύναμο τελικό αποτέλεσμα. Η μονιμότητα εξασφαλίζει ότι η επίδραση των ολοκληρωμένων συναλλαγών θα παραμείνει ακόμα και σε περίπτωση κατάρρευσης του συστήματος και τέλος η συνέπεια αφορά τη συνεπή κατάσταση της βάσης μετά το πέρας εκτέλεσης μιας συναλλαγής και εξασφαλίζεται σε ένα βαθμό μέσω των προηγούμενων απαιτήσεων. Όλα τα ολοκληρωμένα Συστήματα Διαχείρισης Βάσεων Δεδομένων ανταποκρίνονται στις παραπάνω απαιτήσεις, αφού αυτές είναι ιδιαίτερης σημασίας για τη στοιχειώδη λειτουργία.

Μια άλλη σημαντική απαίτηση είναι αυτή της επεκτασιμότητας (scalability) [24]. Αυτή αφορά τη μετάβαση από ένα σύστημα μικρού όγκου δεδομένων και ενεργών χρηστών σε ένα σύστημα σημαντικά μεγαλύτερου όγκου πληροφορίας και αριθμού χρηστών, όπου ο φόρτος εργασίας θα είναι μεγαλύτερος. Έτσι, απαιτείται μέριμνα για την εξασφάλιση αποδεκτού χρόνου απόκρισης στα ερωτήματα, μέσω διάφορων μηχανισμών, όπως για παράδειγμα η δυνατότητα δημιουργίας ευρετηρίων (indexes) [25]. Επιπλέον, η μετάβαση από ένα συγκεντρωτικό Σύστημα Διαχείρισης Βάσεων Δεδομένων σε ένα κατανεμημένο είναι μια πολύ αποτελεσματική λύση στο πρόβλημα της επεκτασιμότητας. Θα πρέπει προφανώς να υπάρχει η υποστήριξη για κάτι τέτοιο από το επιλεγθέν σύστημα.

3.3 Σχεδίαση της βάσης δεδομένων

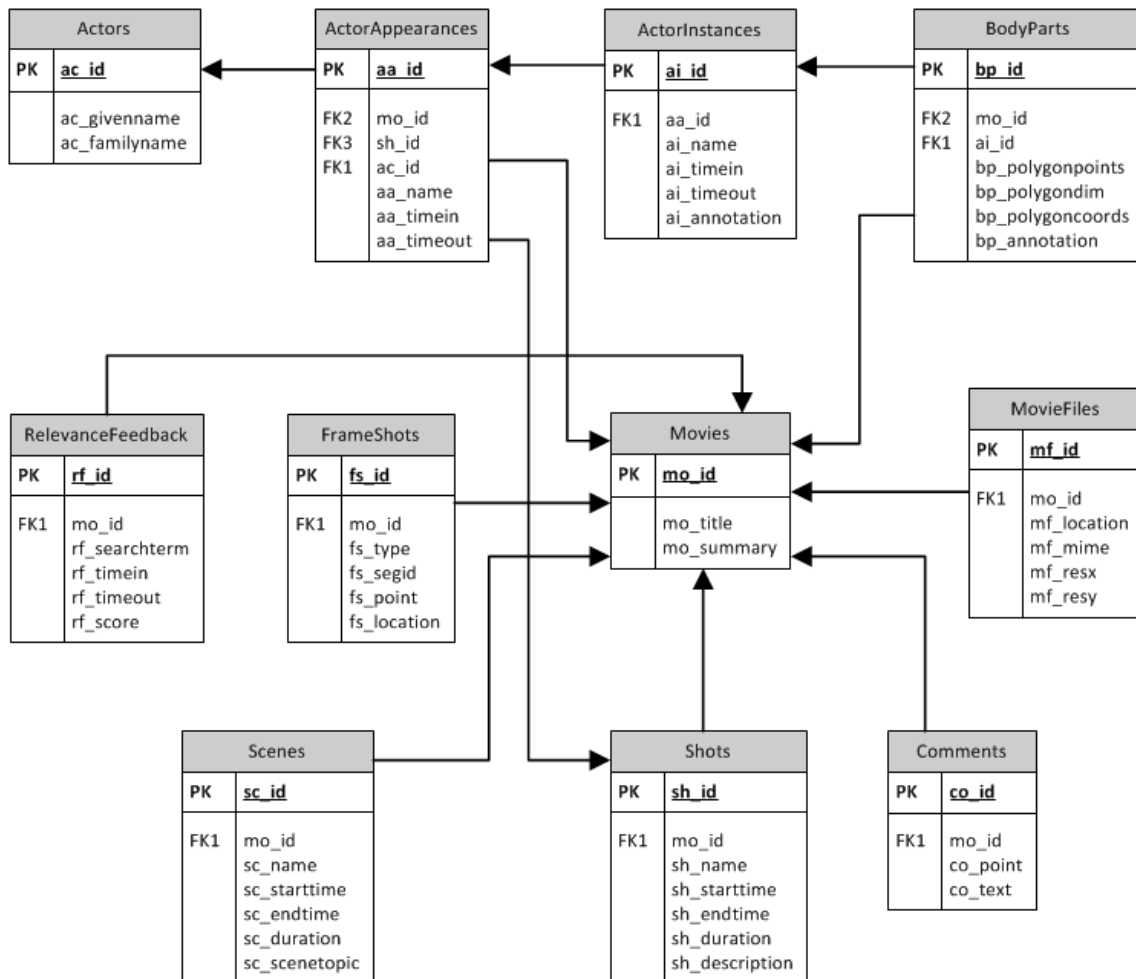
3.3.1 Στρατηγική σχεδίασης

Η σχεδίαση του σχεσιακού διαγράμματος πρέπει να γίνει με τέτοιον τρόπο, ώστε να μην καταστρατηγείται η δομή της πληροφορίας από τα MPEG-7 XML αρχεία. Τα MPEG-7 [8] αρχεία δομούνται ιεραρχικά, αφού πρόκειται ουσιαστικά για XML [10] έγγραφα. Για παράδειγμα ένα οπτικοακουστικό αρχείο αποτελείται από λήψεις (αντικείμενα Shots) και σε κάθε λήψη ανήκουν αντικείμενα εμφάνισης ηθοποιών (αντικείμενα Actor Appearance). Ο ορισμός του κάθε αντικειμένου (object) προκύπτει από το πρότυπο του MPEG-7 μέσω της χρήσης τύπων (type), δηλαδή το εκάστοτε αντικείμενο δημιουργείται με βάση κάποιον τύπο. Πιο συγκεκριμένα, το πρότυπο MPEG-7 καθορίζει ένα μεγάλο αριθμό τύπων, αλλά προσφέρει και τη δυνατότητα ορισμού νέων ή επέκταση των ήδη υπαρχόντων. Η προσέγγιση που επιλέξαμε είναι ότι για κάθε τύπο ορίζουμε και μια σχέση με πεδία τις ιδιότητες του τύπου. Μια ιδιότητα τύπου μπορεί να είναι σύνθετη, δηλαδή να είναι μια συλλογή αντικειμένων κάποιου άλλου τύπου. Η αναπαράσταση αυτής της ιδιότητας γίνεται χρησιμοποιώντας τον απαραίτητο αριθμό σχέσεων και συνδέοντας αυτές κατάλληλα μέσω ξένων κλειδιών (foreign keys). Το MPEG-7 περιλαμβάνει ένα μεγάλο σύνολο προκαθορισμένων τύπων (της τάξης των μερικών εκατοντάδων), οπότε είναι πρακτικά αδύνατο να ορίσουμε σχέσεις για κάθε τέτοιον τύπο. Συνεπώς, θα κάνουμε τη διαδικασία αυτή για ένα μικρό υποσύνολο τύπων, αυτών που χρησιμοποιούνται από τα MPEG-7 αρχεία που έχουμε στη διάθεσή μας. Η προσέγγιση αυτή λύνει πολύ ικανοποιητικά το πρόβλημά μας, καθώς η συντριπτική πλειοψηφία των αρχείων MPEG-7 χρησιμοποιεί ένα μικρό σύνολο τύπων για την αναπαράσταση της απαραίτητης πληροφορίας. Στη συνέχεια ακολουθεί το σχεσιακό διάγραμμα που σχεδιάσαμε με βάση την παραπάνω στρατηγική, κάνοντας κάποιες τροποποιήσεις, όπου ήταν απαραίτητο.

3.3.2 Σχεσιακό διάγραμμα

Όπως αναφέραμε, η αποθήκευση των διαθέσιμων δεδομένων θα γίνει σε ένα σχεσιακό Σύστημα Διαχείρισης Βάσεων Δεδομένων, επομένως πρέπει να καταστρώσουμε ένα κατάλληλο σχεσιακό σχήμα που θα φιλοξενήσει τις απαραίτητες πληροφορίες. Ένα τέτοιο σχήμα πρέπει να χαρακτηρίζεται από δύο σημαντικούς κανόνες σχεδίασης βάσεων δεδομένων: κανονικοποίηση (database normalization) [26] και αποφυγή απώλειας πληροφοριών (zero information loss). Η κανονικοποίηση είναι εκείνο το στοιχείο που εξασφαλίζει ότι κάθε σχέση του σχήματος δεν περιλαμβάνει πλεονάζουσα πληροφορία, αλλά μόνο την ελάχιστη δυνατή, έτσι ώστε να αποφεύγεται όμως η απώλεια πληροφορίας. Απώλεια πληροφορίας έχουμε όταν δεν είναι δυνατή η ανάκτηση μιας ολοκληρωμένης και ορθής «εικόνας» της πραγματικής πληροφορίας που όφειλε να αποθηκευτεί. Στη δική μας σχεδίαση θα καταστρατηγήσουμε το στοιχείο της κανονικοποίησης, δηλαδή θα υπάρχει πλεονάζουσα πληροφορία στους πίνακες που θα σχεδιάσουμε. Αυτή είναι μια συνήθης

τακτική στο σχεδιασμό σχεσιακών σχημάτων, γνωστή ως αποκανονικοποίηση (database denormalization) [27] και χρησιμοποιείται για λόγους βελτιστοποίησης των ερωτημάτων που γίνονται. Το μειονέκτημα της αποθήκευσης πλεονάζουσας πληροφορίας μετριάζεται από τα οφέλη που προκύπτουν στην ταχύτητα εκτέλεσης πολύπλοκων ερωτημάτων, τα οποία απαιτούν συνενώσεις (joins) [28],[29] περισσότερων πινάκων. Στην παρακάτω εικόνα φαίνεται το τελικό σχεσιακό σχήμα που θα χρησιμοποιηθεί, αφού έχει γίνει η απαιτούμενη αποκανονικοποίηση. Κάθε ορθογώνιο αναπαριστά έναν πίνακα με τα πεδία του, ενώ τα βέλη τις σχέσεις μεταξύ τους, χρησιμοποιώντας κατάλληλα ξένα κλειδιά (foreign keys).



Σχήμα 5: Το σχεσιακό σχήμα της βάσης δεδομένων

3.3.3 Περιγραφή πινάκων

Στο τμήμα αυτό θα παρουσιαστούν με μεγαλύτερη λεπτομέρεια οι παραπάνω πίνακες, συνοδευόμενοι από μια σύντομη περιγραφή για το κάθε πεδίο.

Movies: Περιέχει εγγραφές για όλες τις περιγραφές αρχείων οπτικοακουστικού υλικού που υπάρχουν στη βάση δεδομένων, όπου υπάρχει μια καταχώρηση ανά περιγραφή

mo_id: μοναδικό αναγνωριστικό για την καταχώρηση της περιγραφής οπτικοακουστικού αρχείου (πρωτεύον κλειδί)

mo_title: τίτλος του αρχείου οπτικοακουστικού περιεχομένου

mo_summary: σύντομη περίληψη/περιγραφή του αρχείου οπτικοακουστικού περιεχομένου

MovieFiles: Πίνακας με τεχνικές πληροφορίες για τα πολυμεσικά αρχεία που υπάρχουν, είναι βοηθητικός πίνακας για την ορθή λειτουργία του συστήματος

mf_id: μοναδικό αναγνωριστικό για το πολυμεσικό αρχείο (πρωτεύον κλειδί)

mo_id: δείκτης σε καταχώρηση του πίνακα Movies προς το αρχείο περιγραφής το οποίο αντιστοιχεί στο συγκεκριμένο πολυμεσικό αρχείο

mf_location: απόλυτη διαδρομή (absolute path) του αρχείου

mf_mime: τύπος (mime type) του πολυμεσικού αρχείου

mf_resx: η οριζόντια ανάλυση του πολυμεσικού αρχείου

mf_resy: η κάθετη ανάλυση του πολυμεσικού αρχείου

Scenes: Καταχωρήσεις σκηνών του οπτικοακουστικού περιεχομένου

sc_id: μοναδικό αναγνωριστικό σκηνής (πρωτεύον κλειδί)

mo_id: δείκτης σε καταχώρηση περιγραφής του πίνακα Movies, στην οποία ανήκει η σκηνή

sc_name: όνομα σκηνής

sc_starttime: χρονικό σημείο έναρξης της σκηνής

sc_endtime: χρονικό σημείο λήξης της σκηνής

sc_duration: χρονική διάρκεια της σκηνής

sc_scenetopic: περιγραφή της σκηνής

Shots: Καταχωρήσεις λήψεων του οπτικοακουστικού περιεχομένου

sh_id: μοναδικό αναγνωριστικό λήψης (πρωτεύον κλειδί)

mo_id: δείκτης σε καταχώρηση περιγραφής του πίνακα Movies, στην οποία ανήκει η λήψη

sh_name: όνομα λήψης

sh_starttime: χρονικό σημείο έναρξης της λήψης

sh_endtime: χρονικό σημείο λήξης της λήψης

sh_duration: χρονική διάρκεια της λήψης

sh_description: περιγραφή της λήψης

FrameShots: Πίνακας με πληροφορίες για εικόνες στιγμιότυπων (frame snapshots) που έχουν ληφθεί από τα οπτικοακουστικά αρχεία για λόγους παρουσίασης, είναι βοηθητικός πίνακας

fs_id: μοναδικό αναγνωριστικό στιγμιότυπου (πρωτεύον κλειδί)

mo_id: δείκτης σε καταχώρηση του πίνακα Movies, στην οποία ανήκει το στιγμιότυπο

fs_type: τύπος του στιγμιότυπου, καθορίζει αν ανήκει σε στιγμιότυπο από σκηνή, λήψη ή εμφάνιση ηθοποιού

fs_segid: δείκτης που σε συνδυασμό με το προηγούμενο πεδίο καθορίζουν το συγκεκριμένο αντικείμενο στο οποίο ανήκει το στιγμιότυπο

fs_point: χρονικό σημείο για το οποίο έχει ληφθεί το στιγμιότυπο

fs_location: απόλυτη διαδρομή (absolute path) της εικόνας του στιγμιότυπου

Comments: Πίνακας σχολίων που αφορούν το κάθε οπτικοακουστικό περιεχόμενο

co_id: μοναδικό αναγνωριστικό σχολίου (πρωτεύον κλειδί)

mo_id: δείκτης σε καταχώρηση του πίνακα Movies, στην οποία ανήκει το σχόλιο

co_point: χρονικό σημείο στο οποίο αναφέρεται το σχόλιο

co_text: το κείμενο του σχολίου

Actors: Πίνακας με πληροφορίες ηθοποιών που εμφανίζονται στο αποθηκευμένο περιεχόμενο

ac_id: μοναδικό αναγνωριστικό ηθοποιού (πρωτεύον κλειδί)

ac_givenname: όνομα του ηθοποιού

ac_familyname: επίθετο του ηθοποιού

ActorAppearances: Πίνακας με πληροφορίες για τις εμφανίσεις ηθοποιών

aa_id: μοναδικό αναγνωριστικό εμφάνισης ηθοποιού (πρωτεύον κλειδί)

mo_id: δείκτης σε καταχώρηση του πίνακα Movies, στην οποία αντιστοιχεί η εμφάνιση του ηθοποιού

sh_id: δείκτης σε καταχώρηση του πίνακα Shots, στην οποία αντιστοιχεί η εμφάνιση του ηθοποιού

ac_id: δείκτης σε καταχώρηση του πίνακα Actors, συνδέοντας την εμφάνιση ενός ηθοποιού με τον ηθοποιό που εμφανίζεται από τον πίνακα Actors

aa_name: ονομασία για την εμφάνιση του ηθοποιού

aa_timein: χρονικό σημείο έναρξης της εμφάνισης του ηθοποιού

aa_timeout: χρονικό σημείο λήξης της εμφάνισης του ηθοποιού

ActorInstances: Πίνακας με πληροφορίες για στιγμιότυπα εμφανίσεων των ηθοποιών, κάθε εμφάνιση ηθοποιού (Actor Appearance) αποτελείται από πολλά στιγμιότυπα

ai_id: μοναδικό αναγνωριστικό του στιγμιότυπου εμφάνισης ηθοποιού (πρωτεύον κλειδί)

aa_id: δείκτης στην εμφάνιση ηθοποιού του πίνακα ActorAppearances, στην οποία ανήκει το στιγμιότυπο

ai_name: όνομα του στιγμιότυπου της εμφάνισης του ηθοποιού

ai_timein: χρονικό σημείο έναρξης του στιγμιότυπου

ai_timeout: χρονικό σημείο λήξης του στιγμιότυπου

ai_annotation: σχολιασμός του στιγμιότυπου

BodyParts: Πίνακας με τα τμήματα σωμάτων των ηθοποιών που έχουν αναγνωριστεί, για κάθε στιγμιότυπο εμφάνισης ηθοποιού (Actor Instance)

br_id: μοναδικό αναγνωριστικό του τμήματος σώματος του ηθοποιού (πρωτεύον κλειδί)

mo_id: δείκτης σε καταχώρηση του πίνακα Movies, στην οποία αντιστοιχεί το τμήμα του ηθοποιού που έχει αναγνωριστεί

ai_id: δείκτης σε καταχώρηση του πίνακα ActorInstances, στην οποία ανήκει το τμήμα ηθοποιού που έχει αναγνωριστεί

br_polygonpoints: αριθμός σημείων για την οπτική σημείωση του τμήματος του σώματος του ηθοποιού, μέσω πολυγώνου

br_polygondim: αριθμός διαστάσεων για την οπτική σημείωση μέσω πολυγώνου

br_polygoncoords: συντεταγμένες των σημείων του πολυγώνου για την οπτική σημείωση

br_annotation: περιγραφή που σχετίζεται με το τμήμα σώματος του ηθοποιού που έχει αναγνωριστεί

3.4 Καταχώρηση πληροφορίας στη βάση δεδομένων

Από τη στιγμή που έχει σχεδιαστεί ένα κατάλληλο σχήμα για τη βάση μας, πρέπει να ακολουθήσει η καταχώρηση των δεδομένων σε αυτή. Το ενδιάμεσο αυτό στάδιο απαιτεί την κατάλληλη επεξεργασία και διάβασμα των δοθέντων XML MPEG-7 αρχείων. Οι πληροφορίες θα διαβάζονται από τα αρχεία αυτά και στη συνέχεια θα εισάγονται στους κατάλληλους πίνακες του σχεσιακού μας σχήματος, όμως πρέπει να σχεδιαστεί ένας μηχανισμός που να επιτελεί τη λειτουργία αυτή (parser), ο οποίος είναι και το αντικείμενο ανάλυσης αυτού του χωρίου.

Το κυρίως ζήτημα είναι το διάβασμα ενός XML [10] αρχείου και υπάρχουν δύο διαθέσιμες επιλογές προς υιοθέτηση. Η πρώτη επιλογή είναι η χρήση της διαπροσωπείας (interface) του XML DOM [30], κατά την οποία δημιουργείται το δέντρο της δομής του XML στη μνήμη. Κάθε στιγμή όλο το XML δέντρο βρίσκεται στη μνήμη, προσφέροντας ευέλικτη και υψηλού επιπέδου δυνατότητα επεξεργασίας και προσπέλασης. Το βασικότερο

μειονέκτημα αυτής της μεθόδου είναι οι υψηλές απαιτήσεις μνήμης που υπάρχουν επειδή παραμένει όλο το έγγραφο στην κύρια μνήμη. Ο δεύτερος εναλλακτικός μηχανισμός είναι γνωστός ως SAX Parser [31], ο οποίος διαβάζει σειριακά το αρχείο και ενεργοποιείται σε συγκεκριμένα συμβάντα (π.χ. όταν συναντάει συγκεκριμένες XML ετικέτες) που έχει ορίσει ο χρήστης. Δεν προσφέρει την ευελιξία μιας XML DOM [30] διαπροσωπείας, αλλά δεν έχει και υψηλές απαιτήσεις μνήμης αφού μπορεί κάθε στιγμή να βρίσκεται ένα μόνο μέρος του εγγράφου στη μνήμη και επίσης υπερτερεί στον τομέα της ταχύτητας. Παρ' όλο που το XML DOM υστερεί από άποψη απόδοσης, υιοθετήσαμε αυτή την προσέγγιση, λόγω της ευελιξίας που προσφέρει η διατήρηση της δομής του εγγράφου και της μεγαλύτερης ευκολίας ανάπτυξης. Τα αρχεία που πρόκειται να διαβαστούν είναι μεγάλου μεγέθους, ωστόσο σε καμία περίπτωση δεν δημιουργούν ζητήματα απόδοσης και διαθέσιμη μνήμη για να προτιμηθεί ο SAX Parser [31], μια προσέγγιση κατώτερου επιπέδου.

Μεταξύ του διαβάσματος και της εισαγωγής των δεδομένων μεσολαβεί μια αντικειμενοστραφής βιβλιοθήκη που δημιουργήσαμε και αποθηκεύει προσωρινά τα δεδομένα. Η βιβλιοθήκη αυτή χρησιμοποιήθηκε κυρίως για τη διατήρηση της δομικής πληροφορίας του MPEG-7 και για να προσφέρει σε άλλους προγραμματιστές του συστήματος εύκολη πρόσβαση σε όλη τη διαθέσιμη πληροφορία, χωρίς να πρέπει να ασχοληθούν με το διάβασμα του XML αρχείου. Μια τέτοια αρχιτεκτονική σχεδίαση προσφέρει το πλεονέκτημα της ανεξαρτησίας μεταξύ των διάφορων υποσυστημάτων, γιατί μπορεί να αλλάξει η υλοποίησή τους, χωρίς να αλλάξουν οι υπηρεσίες που προσφέρουν. Θα μπορούσε έτσι να αλλάξει η υλοποίηση του διαβάσματος του XML εγγράφου και να χρησιμοποιηθεί ο SAX Parser, χωρίς να αλλάξει ο τρόπος πρόσβασης στη διαθέσιμη πληροφορία. Το τελικό στάδιο που ακολουθεί είναι το διάβασμα της πληροφορίας, μέσω της διεπαφής που προσφέρει η βιβλιοθήκη μας και η καταχώρηση στους πίνακες της βάσης δεδομένων μας. Λόγω της βιβλιοθήκης μας, η λειτουργία αυτή θεωρείται προγραμματιστικά τετριμμένη.

3.5 Επιλεχθέν Σύστημα Διαχείρισης Βάσεων Δεδομένων

Έχοντας αναλύσει διεξοδικά τις απαιτήσεις και το σχήμα που απαιτείται για να καλυφθούν αυτές είμαστε σε θέση να δικαιολογήσουμε την επιλογή που κάναμε για το Σύστημα Διαχείρισης Βάσεων Δεδομένων. Θα χρησιμοποιήσουμε το Σχεσιακό Σχήμα Διαχείρισης Βάσεων Δεδομένων της MySQL [32]. Παρακάτω παραθέτουμε συνοπτικά τους σημαντικότερους λόγους που οδήγησαν σε αυτή την επιλογή.

Είναι ένα ευρέως χρησιμοποιούμενο Σύστημα Διαχείρισης Βάσεων Δεδομένων [19] που βρίσκει εφαρμογή σε πλήθος ερευνητικών και εμπορικών συστημάτων. Υπάρχει έτσι το προτέρημα των συνεχών αναβαθμίσεων και μιας ενεργής κοινότητας προγραμματιστών

που μπορεί να προσφέρει τεχνική υποστήριξη. Άλλος ένας λόγος είναι η εύκολη επικοινωνία με τη γλώσσα που επιλέξαμε για την επικοινωνία με τη βάση δεδομένων μας. Ο συνδυασμός τους βρίσκεται σε χρήση από πολλά συστήματα και αποτελεί έναν αξιόπιστο συνδυασμό βάσης δεδομένων και γλώσσας προγραμματισμού υποδοχής (host programming language). Ένα ακόμη σημαντικό στοιχείο είναι αυτό του ανοιχτού κώδικα και του ελεύθερου λογισμικού, αφού η MySQL [32] είναι τέτοιο προϊόν και μπορούμε να το χρησιμοποιήσουμε χωρίς περιορισμούς. Ακόμα, προσφέρει σε κάποιον βαθμό δυνατότητες εργασίας με XML δεδομένα, στοιχείο που όπως αναφέραμε είναι επιθυμητό, κυρίως για πιθανές μελλοντικές επεκτάσεις. Τέλος, επιτρέπει τη δημιουργία κατανεμημένων συστάδων (clusters) σε περίπτωση που απαιτηθεί επέκταση του συστήματός μας.

ΚΕΦΑΛΑΙΟ 4^ο

ΜΗΧΑΝΙΣΜΟΣ ΑΝΑΖΗΤΗΣΗΣ

4.1 Εισαγωγή

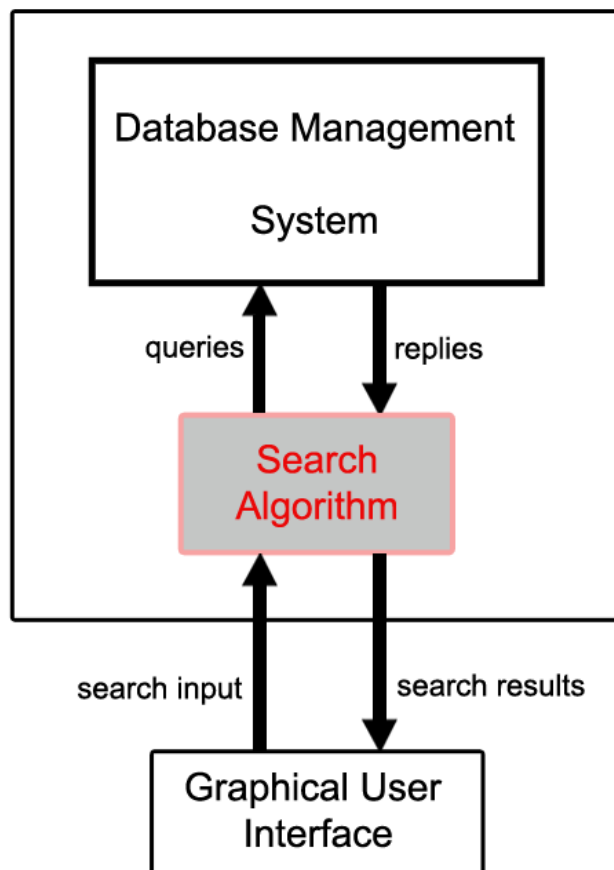
4.2 Μηχανισμός αναζήτησης

4.3 Μηχανισμός ανατροφοδότησης σχετικότητας

4.4 Τεχνικές λεπτομέρειες υλοποίησης

4.1 Εισαγωγή

Το επόμενο στάδιο της έρευνάς μας αφορά τη σχεδίαση αλγορίθμων αναζήτησης, οι οποίοι έχοντας μια βάση πληροφορίας, καλούνται να εξάγουν εκείνο το υποσύνολο οπτικοακουστικού περιεχομένου, το οποίο πληροί κάποια σαφώς καθορισμένα κριτήρια αναζήτησης. Αρχικά θα πρέπει να καθοριστούν οι απαιτήσεις που υπάρχουν από την αναζήτηση, δηλαδή τα ερωτήματα του χρήστη στα οποία θα κληθεί να απαντήσει ο μηχανισμός. Εν συνεχεία, αφού αναλυθούν οι αλγόριθμοι που θα απαντήσουν στα ερωτήματα, θα εξετάσουμε την υλοποίηση ενός μηχανισμού ανατροφοδότησης σχετικότητας (Relevance Feedback) [7]. Ο μηχανισμός αναζήτησης είναι το κυριότερο τμήμα της συνολικής μας έρευνας, αφού από αυτόν εξαρτάται η αποδοτική ή μη αξιοποίηση της υπάρχουσας βάσης πληροφορίας. Στο ακόλουθο σχήμα φαίνεται το τμήμα που θα μελετήσουμε στο παρόν κεφάλαιο.



Σχήμα 6: Το υποσύστημα του μηχανισμού αναζήτησης

4.2 Μηχανισμός αναζήτησης

Θα καθορίσουμε τα σενάρια αναζήτησης που υπάρχουν και για κάθε απαίτηση θα παρουσιάσουμε την προσέγγιση που ακολουθήσαμε. Σε όλα τα ερωτήματα που γίνονται η απάντηση αποτελείται από ένα σύνολο διαστημάτων κάποιων οπτικοακουστικών αρχείων. Τα χωρία αυτά μπορεί να συμπίπτουν με μια σκηνή, μια λήψη, μια εμφάνιση ηθοποιού (δηλαδή να συμπίπτουν οι χρονικές στιγμές έναρξης και λήξης τους) ή με τίποτα από τα παραπάνω και να είναι απλώς ένα κατάλληλα επιλεγμένο χρονικό διάστημα. Σε κάποιες περιπτώσεις οι δυνατότητες που μας δίνει η γλώσσα της SQL [21], μέσω των ερωτημάτων της, επαρκούν για να απαντηθούν τα ερωτήματα του χρήστη. Στις περιπτώσεις που οι υπολογιστικές δυνατότητες της SQL δεν επαρκούν, θα ακολουθεί περαιτέρω επεξεργασία μέσω κατάλληλων αλγορίθμων σε μια γλώσσα προγραμματισμού της επιλογής μας. Στη συνέχεια παρατίθενται τα σενάρια αναζήτησης που έχουμε υλοποιήσει. Οι περιγραφές των υλοποιήσεων των σεναρίων που ακολουθούν περιλαμβάνουν ερωτήματα της SQL και σε περίπτωση που χρειάζεται περαιτέρω επεξεργασία παρατίθεται ο απαραίτητος ψευδοκώδικας του αλγορίθμου.

1. Αναζήτηση εμφανίσεων ενός ηθοποιού με μια ελάχιστη χρονική διάρκεια εμφάνισης

Ο χρήστης καθορίζει μια χρονική διάρκεια και τον ηθοποιό που επιθυμεί, μέσω του ονόματός του. Θα του επιστραφούν καταχωρήσεις του πίνακα εμφανίσεων ηθοποιών (ActorAppearances), οι οποίες αφορούν τον ηθοποιό που καθορίστηκε και ξεπερνούν τη δοσμένη χρονική διάρκεια.

Οι μεταβλητές της αναζήτησης είναι το όνομα (Name) του ηθοποιού, το επίθετό (Surname) του και η χρονική διάρκεια σε δευτερόλεπτα (Duration). Το ακόλουθο ερώτημα επιστρέφει τις εμφανίσεις εκείνες που πληρούν τα κριτήρια.

```
SELECT
    *
FROM
    ActorAppearances aa, Actors ac
WHERE
    aa.ac_id = ac.ac_id
AND
    ac.ac_givenname=Name
AND
    ac.ac_familyname=Surname
AND
    aa.aa_timeout - aa.aa_timein > Duration
```

Το ερώτημα εκτελείται με τη χρήση μιας συνένωσης (join) [28],[29] δύο πινάκων, καθώς ο πίνακας των εμφανίσεων των ηθοποιών (ActorAppearances) δεν περιλαμβάνει την πληροφορία του ονόματος ενός ηθοποιού. Η συνένωση δύο πινάκων είναι η πράξη του καρτεσιανού γινομένου, στο οποίο εφαρμόζονται κάποιες συνθήκες προκειμένου να εξαλειφθούν κάποιες εγγραφές, οι οποίες δεν αντιστοιχούν σε πραγματική πληροφορία. Η συνένωση περισσότερων πινάκων ορίζεται επαγωγικά, ωστόσο ο υπολογισμός της (δηλαδή ενός καρτεσιανού γινομένου) γίνεται από το Σύστημα Διαχείρισης Βάσεων Δεδομένων με περίπλοκους αλγόριθμους και δεν αποτελεί αντικείμενο της έρευνάς μας.

2. Αναζήτηση ταυτόχρονης εμφάνισης αυθαίρετου αριθμού συγκεκριμένων ηθοποιών

Η είσοδος του αλγορίθμου αποτελείται από μια λίστα ονοματεπωνύμων ηθοποιών (Name1 Surname1, Name2 Surname2, ..., NameX SurnameX) βάσει των οποίων θα αναζητηθούν χωρία οπτικοακουστικού υλικού, στα οποία εμφανίζονται όλοι οι δοθέντες ηθοποιοί. Σε αυτή την περίπτωση η εξαγωγή απάντησης είναι πιο περίπλοκη διαδικασία και δεν μπορεί να απαντηθεί με ένα ερώτημα.

Αρχικά επιλέγονται εκείνες οι καταχωρήσεις του πίνακα των εμφανίσεων ηθοποιών (ActorAppearances), οι οποίες αντιστοιχούν σε κάποιο εκ των ονομάτων που έχουν δοθεί σαν είσοδος. Επιστρέφονται όλες εκείνες οι εμφανίσεις ηθοποιών στις οποίες μετέχουν οι δοθέντες ηθοποιοί και εν συνεχεία θα εξεταστούν για το αν πληρούν το κριτήριο της ταυτόχρονης εμφάνισης. Το ερώτημα είναι το εξής:

```
SELECT
    *
FROM
    ActorAppearances
WHERE
    ac_id IN (
        SELECT
            ac_id
        FROM
            Actors
        WHERE
            CONCAT(ac_givename, ac_familyname) IN (
                Name1+Surname1,
                Name2+Surname2,
                ...,
                NameX+SurnameX
            )
    )
ORDER BY
    aa_timein
ASC
```

Οι εμφανίσεις που επιστρέφονται θα εξεταστούν από τον αλγόριθμο αν είναι ταυτόχρονες. Πρέπει να σημειωθεί εδώ ότι οι εγγραφές θα επιστραφούν σε αύξουσα σειρά έναρξης εμφάνισης (aa_timein), στοιχείο που απαιτείται από τον αλγόριθμο και θα εξηγηθεί παρακάτω. Ο αλγόριθμος που επιστρέφει τα κατάλληλα χωρία από τα οπτικοακουστικά αρχεία, αφορά την ταυτόχρονη εμφάνιση ηθοποιών σε ένα οπτικοακουστικό αρχείο, οπότε θα πρέπει να γίνει η κατάλληλη επανάληψη για όλα τα οπτικοακουστικά αρχεία, στα οποία ανήκουν οι εμφανίσεις που επέστρεψε το προηγούμενο ερώτημα. Ακολουθεί ο αλγόριθμος για ένα οπτικοακουστικό αρχείο:

/*

overlaps: variable for counting the number of overlapping actor appearances in each iteration

previousEndtime: variable containing the timeout value of the previous iteration over the actor appearances, if overlap found it contains the value where the overlap ends in order to check for another overlap in the next iteration

queryResult: contains the result of the previous query, grouped by each movie, as overlapping check has no meaning between intervals in different movies

actorNumber: the number of actors supplied by the user to check for simultaneous appearances

aa: iteration variable containing every time one of the fetched rows that belong to the result

endResult: array containing objects of the desired actor appearances of all the specified actors, contains the final result for the user

*/

overlaps ← 1

previousEndtime ← 0

endResult ← {}

foreach queryResult as aa {

 if aa.timein < previousEndtime {

 overlaps ← overlaps + 1

 previousEndtime ← min(aa.timeout, previousEndtime)

 }

 else {

 overlaps ← 1

 previousEndtime ← aa.timeout

 }

```

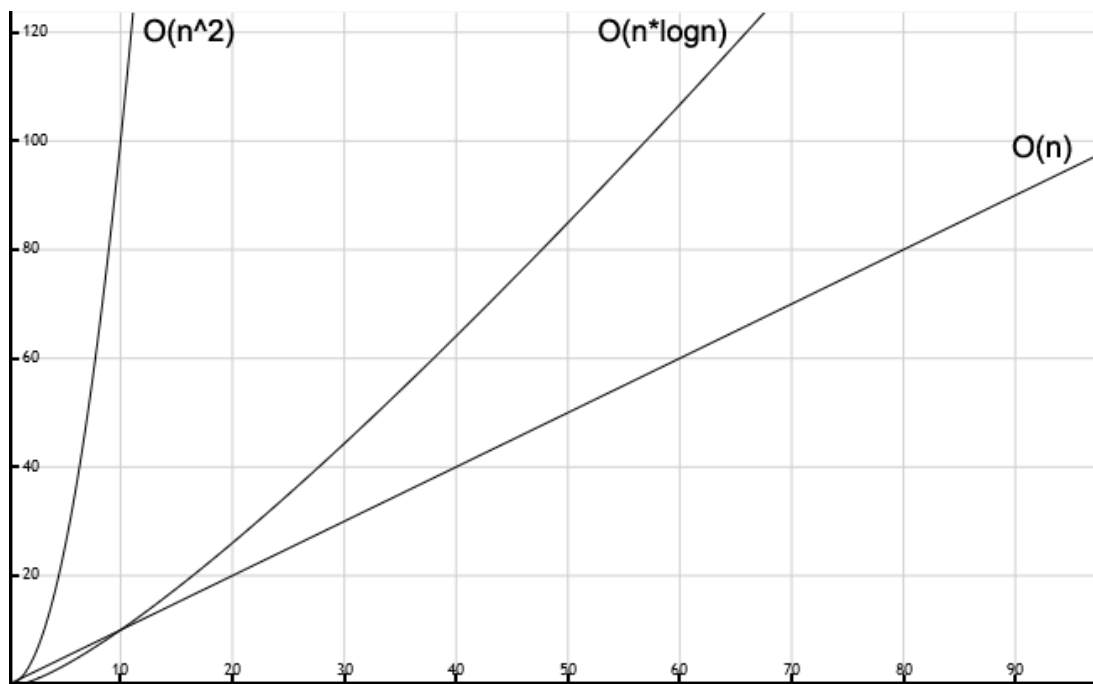
if overlaps = actorNumber {
    //found: a dummy variable for creating a result object and pushing it
    found.mo_id ← aa.mo_id
    found.timein ← aa.timein
    found.timeout ← min(previousEndtime, aa.timeout)
    //push the found overlap to the final results table
    endResult.push(found)
    overlaps ← 1
    previousEndtime ← aa.timeout
}
}

```

Η προσέγγισή μας ανήκει στην κατηγορία των άπληστων αλγορίθμων (greedy algorithms) [33], που σημαίνει ότι σε κάθε βήμα γίνεται μια αμετάκλητη επιλογή για τη λύση, η άπληστη επιλογή. Η άπληστη επιλογή είναι εκείνη η επιλογή, η οποία φαίνεται να είναι η βέλτιστη με βάση την τρέχουσα κατάσταση και κάποιο απλό κριτήριο απόφασης. Οι αλγόριθμοι αυτοί χρησιμοποιούνται για προβλήματα βελτιστοποίησης και ως τέτοιο μπορούμε να θεωρήσουμε και το δικό μας, προσαρμόζοντας κατάλληλα τη διατύπωσή του: δοθέντων των διαστημάτων εμφάνισης των ηθοποιών, θέλουμε να τα διατάξουμε με τέτοιο τρόπο, ώστε να μεγιστοποιείται ο αριθμός των επικαλύψεων. Ο αριθμός αυτός στην περίπτωση μας φράσσεται από τον αριθμό των ηθοποιών που μας έδωσε ο χρήστης σαν είσοδο, οπότε κάθε φορά που ανακαλύπτουμε επικάλυψη διαστημάτων που είναι σε πλήθος ίση με τον αριθμό των ηθοποιών έχουμε βρει κοινή εμφάνιση όλων των ηθοποιών. Να σημειώσουμε ότι δεν μπορεί να υπάρξει επικάλυψη των διαστημάτων δύο εμφανίσεων του ίδιου ηθοποιού, τα διαστήματα αυτά θα είναι απαραίτητα ξένα μεταξύ τους.

Η αμετάκλητη επιλογή που αναφέρθηκε είναι εκείνη που υποδεικνύει την ταξινόμηση κατά αύξουσα σειρά χρόνου έναρξης των εμφανίσεων. Με τον τρόπο αυτόν επιτυγχάνεται η εύρεση λύσης σε χρόνο γραμμικό ως προς τον αριθμό των εμφανίσεων όλων των δοθέντων ηθοποιών, δηλαδή σε $O(n)$. Σε αυτό πρέπει να προστεθεί η πολυπλοκότητα της ταξινόμησης, η οποία όμως γίνεται από το Σύστημα Διαχείρισης Βάσεων Δεδομένων, οπότε δεν μπορούμε να αποφανθούμε με μεγάλη ακρίβεια για την πολυπλοκότητα. Μπορούμε όμως να καθορίσουμε ένα θεωρητικό άνω όριο καθώς μπορούμε σίγουρα να κάνουμε την ταξινόμηση σε χρόνο $O(n \log n)$, οπότε συνολικά έχουμε: $O(n) + O(n \log n) = O(n \log n)$. Ο αλγόριθμός είναι ιδιαίτερα αποδοτικός, γεγονός που σημαίνει ότι μπορεί να υιοθετηθεί χωρίς πρόβλημα για μεγάλες εισόδους. Στο ακόλουθο σχήμα φαίνεται η πολυπλοκότητά του σε σχέση με αυτήν ενός γραμμικού αλγορίθμου και ενός με πολυπλοκότητα $O(n^2)$. Στον οριζόντιο άξονα έχει σημειωθεί το

μέγεθος της εισόδου του αλγορίθμου και στον κατακόρυφο οι χρονικές μονάδες που διαρκεί η εκτέλεση για το αντίστοιχο μέγεθος εισόδου.



Σχήμα 7: Σύγκριση πολυπλοκότητων αλγορίθμων

3. Αναζήτηση εμφάνισης ενός ηθοποιού αμέσως μετά την εμφάνιση κάποιου άλλου καθορισμένου ηθοποιού

Σε αυτή τη μορφή αναζήτησης ο χρήστης δίνει σαν είσοδο δύο ονόματα ηθοποιών και ο αλγόριθμος αναζητά για εμφανίσεις του δεύτερου ηθοποιού που ακολουθούν εμφανίσεις του πρώτου. Ένα σημείο που απαιτεί διευκρίνιση εκ μέρους μας είναι η έννοια του «ακολουθεί» και υπάρχουν δύο δυνατές προσεγγίσεις: η εμφάνιση του δεύτερου ηθοποιού ακολουθεί αυτήν του πρώτου εντός ενός μικρού χρονικού διαστήματος, ώστε να θεωρούνται διαδοχικές ή η εμφάνιση του δεύτερου ηθοποιού είναι η αμέσως επόμενη εμφάνιση ηθοποιού μετά την εμφάνιση του πρώτου. Η διαφορά είναι λεπτή και έγκειται στο γεγονός ότι μεταξύ των δύο εμφανίσεων μπορεί να παρεμβάλλεται κάποια σκηνή μεγάλης διάρκειας χωρίς κάποια εμφάνιση ηθοποιού. Επιλέξαμε την πρώτη προσέγγιση, καθώς η συνήθης απαίτηση αναζήτησης είναι οι δύο εμφανίσεις να σχετίζονται μεταξύ τους, γεγονός που δεν ισχύει απαραίτητα στη δεύτερη προσέγγιση. Η αναζήτηση αυτή μπορεί να απαντηθεί μέσω ενός κατάλληλου ερωτήματος, του ακόλουθου:

```

SELECT
*
FROM
ActorAppearances a1, ActorAppearances a2, Actors ac1, Actors ac2
WHERE
a2.aa_timein – a1.aa_timeout <1
AND
a2.aa_timein – a1.aa_timeout > 0
AND
a1.mo_id = a2.mo_id
AND
a1.ac_id = ac1.ac_id
AND
a2.ac_id = ac2.ac_id
AND
CONCAT(ac1.ac_givename, ac1.ac_familyname) = Name1+Surname1
AND
CONCAT(ac2.ac_givename, ac2.ac_familyname) = Name2+Surname2

```

Στο ερώτημα αυτό υπολογίζουμε το καρτεσιανό γινόμενο του πίνακα εμφανίσεων ηθοποιών (ActorAppearances) με τον εαυτό του και φιλτράρουμε εκείνες τις εγγραφές που αφορούν τις διαδοχικές εμφανίσεις των ζητούμενων ηθοποιών, μέσω ενός ενδιάμεσου χρονικού διαστήματος μέγιστης διάρκειας ενός δευτερολέπτου. Οι συνενώσεις με τον πίνακα των ηθοποιών (Actors) γίνονται για τη δυνατότητα αναζήτησης με βάση τα ονόματα των ηθοποιών και όχι των μοναδικών αναγνωριστικών τους, τα οποία είναι άγνωστα στο χρήστη.

4. Αναζήτηση λήψης με βάση την περιγραφή της και κάποιων λέξεων κλειδιών

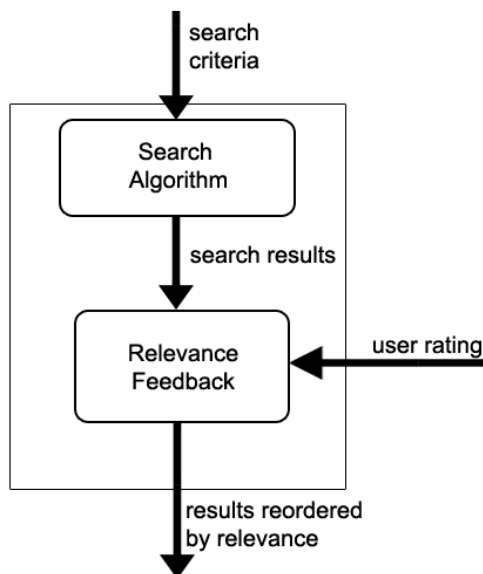
Σε αυτήν την περίπτωση ο χρήστης μπορεί να αναζητήσει λήψεις μέσα από το σύνολο των οπτικοακουστικών αρχείων με βάση την περιγραφή της λήψης. Η περιγραφή της λήψης αποτελείται από κείμενο και η είσοδος του χρήστη είναι επίσης κείμενο. Είναι η κλασσική δυνατότητα αναζήτησης που προσφέρεται από όλες τις μηχανές αναζήτησης και την υιοθετήσαμε, ώστε να προσφέρει το σύστημά μας μια παραπάνω δυνατότητα αναζήτησης, της πιο δημοφιλούς σε αντίστοιχα εμπορικά συστήματα. Το είδος της αναζήτησης δεν είναι εγγενές στο περιεχόμενο, όπως στις προηγούμενες περιπτώσεις, αλλά είναι πιο οικείο για το χρήστη. Το SQL ερώτημα είναι το εξής, όπου η μεταβλητή εισόδου είναι η SearchTerm:

```
SELECT
  *
FROM
  Shots
WHERE
  MATCH(sh_description) AGAINST (SearchTerm)
```

Η δυνατότητα αναζήτησης κειμένου είναι αυτή που μας προσφέρει το Σύστημα Διαχείρισης Βάσεων Δεδομένων και χρησιμοποιεί ξεχωριστά τις λέξεις του όρου αναζήτησης προκειμένου να ψάξει μέσα από τις διαθέσιμες περιγραφές. Κάθε περιγραφή ελέγχεται για το αν περιέχει όρους από το κείμενο προς αναζήτηση και αν περιέχει αυξάνεται μια εσωτερική βαθμολογία που της έχει αποδοθεί. Σε περίπτωση που περιέχει περισσότερους όρους η βαθμολογία αυξάνεται αντίστοιχα, ενώ αν ο όρος αναζήτησης που έχει εντοπιστεί είναι σπάνιος η βαθμολογία που αποδίδεται είναι αυξημένη. Τέλος, δεν λαμβάνονται υπόψη όροι αναζήτησης μικρού μήκους (κάτω των τεσσάρων χαρακτήρων), αφού στην πλειοψηφία των περιπτώσεων δεν προσφέρουν πληροφορία και είναι συνήθως άρθρα ή συνδυαστικές λέξεις. Με τον τρόπο αυτόν πετυχαίνεται η εύρεση λήψεων που ταιριάζουν με το κείμενο εισόδου προς αναζήτηση και επιστρέφονται στο χρήστη.

4.3 Μηχανισμός ανατροφοδότησης σχετικότητας

Ο μηχανισμός ανατροφοδότησης σχετικότητας (Relevance Feedback Mechanism) [7] είναι ένα σημαντικό τμήμα της έρευνάς μας, που ανήκει στον ευρύτερο τομέα της αναζήτησης, χωρίς όμως να είναι απαραίτητος για τη σωστή λειτουργία της. Είναι ένα πρόσθετο σύστημα που στοχεύει στη βελτιστοποίηση των αποτελεσμάτων της αναζήτησης, λαμβάνοντας υπόψη τη γνώμη του χρήστη. Ο μηχανισμός αυτός είναι μια ξεχωριστή μονάδα που δέχεται σαν είσοδο τα αποτελέσματα της αναζήτησης και επιστρέφει την ίδια λίστα αναδιοργανωμένη με τα πιο σχετικά αποτελέσματα πρώτα. Στο παρακάτω σχήμα φαίνεται ο ρόλος αυτού του μηχανισμού.



Σχήμα 8: Ο μηχανισμός ανατροφοδότησης σχετικότητας

Όπως είδαμε, ο μηχανισμός αναζήτησης επιστρέφει ένα σύνολο αποτελεσμάτων που ικανοποιούν τα κριτήρια του χρήστη. Το σύνολο αυτό έχει επιλεγεί με απόλυτο τρόπο, που σημαίνει ότι με απόλυτη σιγουριά δεν υπάρχουν περισσότερα πιθανά αποτελέσματα που να ικανοποιούν τα κριτήρια αναζήτησης, αλλά και τα αποτελέσματα που βρέθηκαν καλύπτουν τα κριτήρια αναζήτησης. Ωστόσο, αυτό δεν εξασφαλίζει και τη σχετικότητα των αποτελεσμάτων ως προς το χρήστη, ο οποίος προφανώς έχει διαφορετική αντίληψη από την απολυτότητα των αποτελεσμάτων του μηχανισμού αναζήτησης. Αυτό το χάσμα αντίληψης που υπάρχει μεταξύ ενός ανθρώπου και ενός ντετερμινιστικού μηχανισμού αναζήτησης καλείται να γεφυρώσει ο μηχανισμός ανατροφοδότησης σχετικότητας, μέσω της παρέμβασης του ίδιου του χρήστη.

Στο χρήστη επιστρέφεται ένα σύνολο αποτελεσμάτων σε φθίνουσα σειρά σχετικότητας και δίνεται η δυνατότητα της αξιολόγησης του κάθε αποτελέσματος. Η αξιολόγηση αποτελείται από μια θετική ή αρνητική ψήφο που υποδηλώνει σχετικό ή μη αποτέλεσμα αντίστοιχα. Σε μετέπειτα αναζητήσεις με το ίδιο κριτήριο τα αποτελέσματα θα ταξινομηθούν εκ νέου σε φθίνουσα σειρά σχετικότητας, αντανακλώντας τις απόψεις των χρηστών. Μετά από μεγάλο αριθμό αναζητήσεων και αξιολογήσεων από τους χρήστες, η συνολική εικόνα της σχετικότητας των αποτελεσμάτων θα συγκλίνει στην πραγματική αντίληψη του μέσου χρήστη. Σε αυτό συμβάλλει το ίδιο βάρος που έχει κάθε αξιολόγηση, οπότε ένας μικρός αριθμός «λανθασμένων» αξιολογήσεων δεν επηρεάζει την τελική εικόνα του αποτελέσματος, το οποίο θα συγκλίνει στο αν κάτι είναι σχετικό ή όχι.

Σε αυτό το σημείο θα αναλύσουμε με μεγαλύτερη λεπτομέρεια τη λειτουργία του μηχανισμού. Κάθε αίτημα αναζήτησης επιστρέφει κάποια αποτελέσματα, τα οποία είναι χωρία από τα διαθέσιμα οπτικοακουστικά αρχεία, επομένως κάθε αποτέλεσμα

ταυτοποιείται από το χρονικό σημείο έναρξης, λήξης και το μοναδικό αναγνωριστικό του αρχείου οπτικοακουστικού περιεχομένου. Αυτό όμως δεν αρκεί για τη λειτουργία του μηχανισμού μας, γιατί το κάθε αποτέλεσμα πρέπει να συνδέεται και με ένα συγκεκριμένο ερώτημα αναζήτησης. Το ίδιο αποτέλεσμα μπορεί να εξάγεται για διαφορετικά ερωτήματα, όμως η σχετικότητα του να διαφέρει για κάθε ερώτημα. Έτσι, όταν αναζητούμε τη σχετικότητα (αξιολόγηση, score) ενός αποτελέσματος, αυτή θα εντοπίζεται από την τετράδα: κριτήρια αναζήτησης, αναγνωριστικό οπτικοακουστικού αρχείου, χρονικό σημείο έναρξης και χρονικό σημείο λήξης:

(searchterm, movie_id, time_in, time_out) → score

Για όλα τα αποτελέσματα που επέστρεψε ο μηχανισμός αναζήτησης, εντοπίζουμε τις αξιολογήσεις τους, αν αυτές υπάρχουν και ταξινομούμε τα αποτελέσματα με βάση αυτές. Δεν υπάρχει απαραίτητα αξιολόγηση για κάθε αποτέλεσμα, καθώς μπορεί να μην έχει ξαναγίνει η συγκεκριμένη αναζήτηση ή σε περίπτωση που έχει γίνει να μην έχει προηγηθεί αξιολόγηση από κάποιον χρήστη. Σε αυτή την περίπτωση η τιμή του score είναι μηδέν.

Γίνεται κατανοητό πως ο μηχανισμός ανατροφοδότησης σχετικότητας λειτουργεί σαν ένα ξεχωριστό, ανεξάρτητο υποσύστημα, δεχόμενο είσοδο από το μηχανισμό αναζήτησης. Η ανεξαρτησία των ψηφίδων (system components) είναι ιδιαίτερα επιθυμητό στοιχείο σε ένα σύστημα λογισμικού γιατί μπορούν να γίνουν εύκολα τροποποιήσεις σε κάθε ψηφίδα, χωρίς να επηρεαστεί η λειτουργία του συνολικού συστήματος και σε αυτή την αρχή βασίσαμε την αρχιτεκτονική μας σχεδίαση.

4.4 Τεχνικές λεπτομέρειες υλοποίησης

Μια πρώτη τεχνική λεπτομέρεια είναι ο τρόπος με τον οποίον κωδικοποιούμε τα κριτήρια αναζήτησης του μηχανισμού ανατροφοδότησης σχετικότητας, τα οποία σε κάθε είδος αναζήτησης είναι διαφορετικά. Πρέπει να βρεθεί ένας τρόπος που να μπορεί να χρησιμοποιηθεί από όλα τα είδη αναζήτησης και να υποστηρίξει άλλους πιθανούς μελλοντικούς τρόπους. Ένα αίτημα αναζήτησης ταυτοποιείται από το είδος της αναζήτησης (για παράδειγμα αναζήτηση εμφάνισης ενός ηθοποιού ή αναζήτηση με βάση την περιγραφή μιας λήψης) και από τα κριτήρια που καθορίζει ο χρήστης. Αυτές οι πληροφορίες πρέπει με κάποιον τρόπο να συνδυαστούν για να προκύψει ένα μοναδικό αναγνωριστικό για την αναζήτηση. Για να το πετύχουμε αυτό, χρησιμοποιούμε για κάθε είδος αναζήτησης ένα αναγνωριστικό αλφαριθμητικό, το οποίο το συνενώνουμε με την είσοδο του χρήστη. Στη συνέχεια, περνώντας το προκύπτον αλφαριθμητικό από μια συνάρτηση κατακερματισμού (hash function) [34] προκύπτει ένα νέο αλφαριθμητικό που ταυτοποιεί κάθε αναζήτηση με τα κριτήριά της. Το τελικό αλφαριθμητικό που προέκυψε έχει πάντα σταθερό μήκος και εξαρτάται από τη συνάρτηση κατακερματισμού. Εμείς χρησιμοποιήσαμε τη συνάρτηση MD5 [35], που παράγει μια τιμή κατακερματισμού 128-

bit και αναπαρίσταται συνήθως σαν δεκαεξαδικός αριθμός 32 ψηφίων. Επειδή μια συνάρτηση κατακερματισμού απεικονίζει ένα άπειρο σύνολο αλφαριθμητικών σε ένα πεπερασμένο σύνολο τιμών είναι προφανές ότι υπάρχει η πιθανότητα σύγκρουσης, δηλαδή δύο διαφορετικά αλφαριθμητικά να απεικονίζονται στην ίδια τιμή, γεγονός που θα δημιουργούσε πρόβλημα. Ωστόσο η πιθανότητα ενός τέτοιου γεγονότος είναι αμελητέα ($2^{-128} * 2^{-128} = 2^{-256} = 8.636*10^{-78}$) και μπορούμε να χρησιμοποιήσουμε την παραπάνω προσέγγιση χωρίς πρόβλημα.

Μια δεύτερη, τεχνικής φύσεως επιλογή είναι η γλώσσα προγραμματισμού για την υλοποίηση όσων περιγράφηκαν στο κεφάλαιο αυτό. Η γλώσσα που επιλέξαμε είναι η PHP [36], κυρίως λόγω κάποιων πλεονεκτημάτων που προσφέρει και ταιριάζουν στη φύση της εφαρμογής μας. Η PHP [36] προσφέρει ένα πλήρες σύνολο βιβλιοθηκών και ευκολιών, επιταχύνοντας πολύ την ανάπτυξη εφαρμογών και επιτρέποντας στον προγραμματιστή να εστιάσει στην ουσία της εφαρμογής και να μην αναλώνεται στον προγραμματισμό απολύτως βασικών λειτουργιών, που θα χρησιμοποιηθούν ως βάση μετέπειτα. Ένα άλλο πλεονέκτημα είναι η υποστήριξη αντικειμενοστρέφειας [37], μιας σημαντικής δυνατότητας κυρίως για την ανάπτυξη μεγάλων σε όγκο συστημάτων και ταιριάζει στο είδος της δικής μας εφαρμογής. Ακόμα, οι εφαρμογές της γλώσσας αυτής μπορούν να επεκταθούν και να χρησιμοποιηθούν από μεγάλο αριθμό χρηστών χωρίς καμία τροποποίηση, πέραν των αναγκαίων βελτιστοποιήσεων που θα προκύψουν λόγω αυξημένης χρήσης. Επιπρόσθετα, προσφέρεται μεγάλος βαθμός ευελιξίας, αφού οι εφαρμογές σε PHP μπορούν να λειτουργήσουν σε συστήματα Linux, Mac και Windows, ενώ υπάρχει και πολύ καλή υποστήριξη με εκσυγχρονισμένες διεπαφές προγραμματισμού (Application Programming Interface) για ένα πλήθος Συστημάτων Διαχείρισης Βάσεων Δεδομένων, μεταξύ των οποίων και το σύστημα που επιλέξαμε εμείς (MySQL [32]). Τέλος, η PHP είναι μια γλώσσα ανοιχτού κώδικα (open source) και η χρήση της σε εφαρμογές γίνεται ελεύθερα, προσφέροντας πολλά επιπλέον πλεονεκτήματα, όπως πληθώρα διαδικτυακά διαθέσιμου υλικού και έναν μεγάλο αριθμό ομάδων υποστήριξης (support groups) σε περίπτωση εμφάνισης τεχνικών προβλημάτων.

ΚΕΦΑΛΑΙΟ 5^ο

ΣΧΕΔΙΑΣΜΟΣ ΓΡΑΦΙΚΗΣ ΔΙΕΠΑΦΗΣ ΧΡΗΣΤΗ

5.1 Εισαγωγή

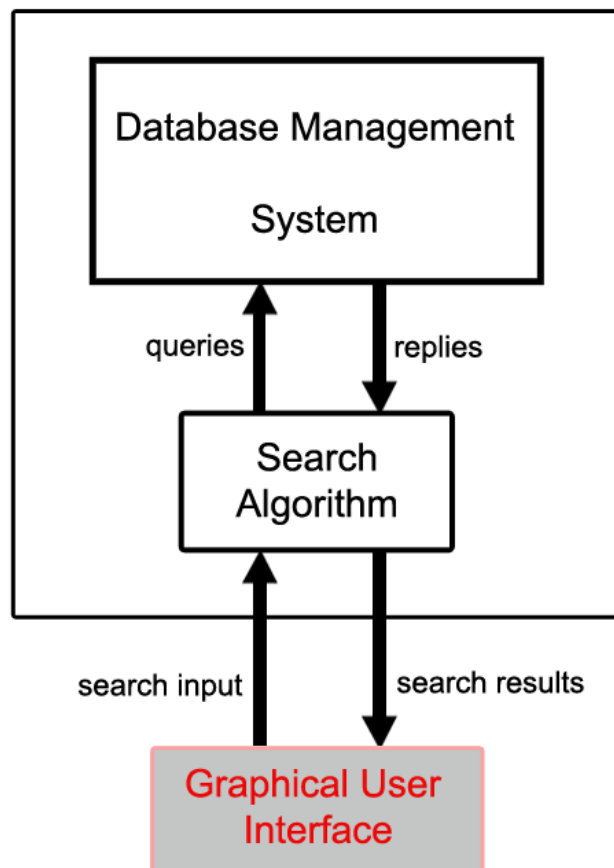
5.2 Καθορισμός λειτουργικών απαιτήσεων

5.3 Καθορισμός μη λειτουργικών απαιτήσεων

5.4 Τεχνικές λεπτομέρειες υλοποίησης

5.1 Εισαγωγή

Η γραφική διεπαφή χρήστη (Graphical User Interface) είναι ένα σημαντικό τμήμα του συστήματός μας και γι' αυτό ο σχεδιασμός του απαιτεί ιδιαίτερη προσοχή. Η σημασία της διεπαφής έγκειται στο γεγονός ότι είναι το σημείο πρόσβασης του τελικού χρήστη στην υπάρχουσα βάση πληροφορίας. Μέσω αυτής επιτυγχάνεται η αναπαράσταση και η αλληλεπίδραση με το σύνολο της αποθηκευμένης πληροφορίας, η οποία καθίσταται άχρηστη αν δεν μπορεί να προσπελαστεί εύκολα. Ο χρήστης δεν χρειάζεται να έχει γνώσεις χρήσης εξειδικευμένων συστημάτων και γλωσσών, ενώ η εμπειρία που έχει από άλλες υπάρχουσες διεπαφές τον διευκολύνει. Επίσης, το γεγονός ότι η εφαρμογή μας σχετίζεται με οπτικά δεδομένα (πολυτροπικό οπτικοακουστικό υλικό) ουσιαστικά επιβάλλει τη χρήση γραφικής διεπαφής σε σχέση με άλλες εναλλακτικές επιλογές, όπως θα μπορούσε να είναι μια γραμμή εντολών (Command Line Interface). Στο παρακάτω σχήμα φαίνεται το τμήμα του συστήματος που θα αναλυθεί σε αυτό το κεφάλαιο της διπλωματικής εργασίας.



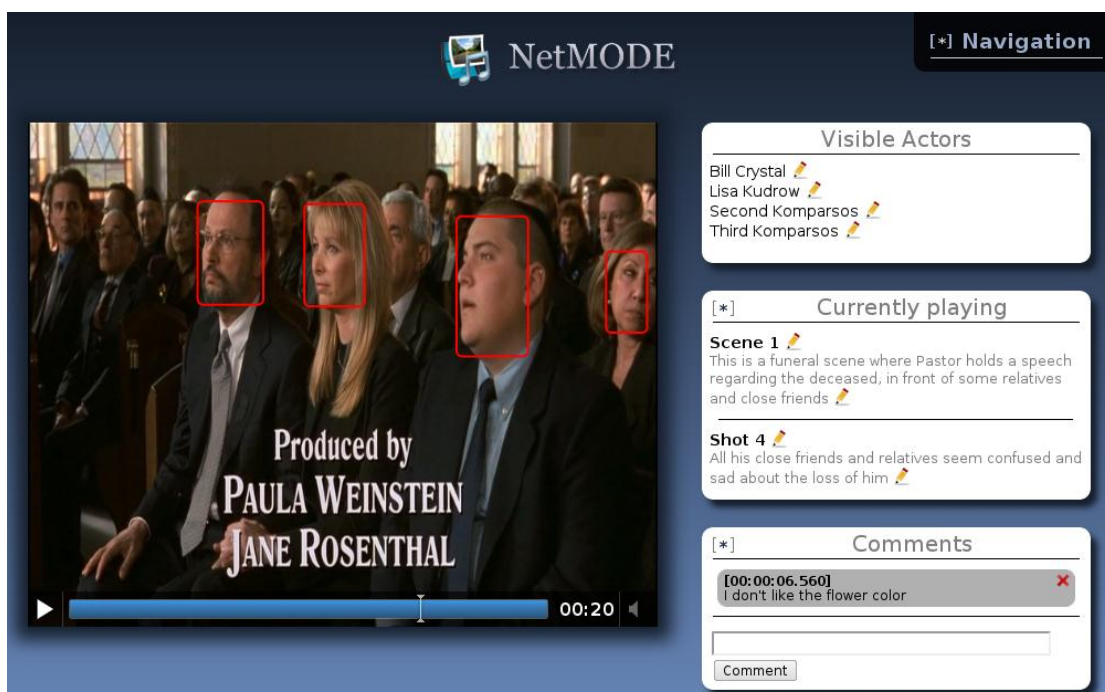
Σχήμα 9: Το υποσύστημα της γραφικής διεπαφής χρήστη

5.2 Καθορισμός λειτουργικών απαιτήσεων

Το πρώτο στάδιο της διαδικασίας σχεδιασμού μιας γραφικής διεπαφής χρήστη είναι η λεπτομερής καταγραφή των λειτουργικών απαιτήσεων (functional requirements). Οι λειτουργικές απαιτήσεις ουσιαστικά καθορίζονται από τις δυνατότητες που θέλουμε να έχει ο χρήστης, οι οποίες προκύπτουν από το σκοπό της παρούσας διπλωματικής εργασίας και επίσης βρίσκονται σε αντιστοιχία με τις δυνατότητες που προσφέρει ο μηχανισμός αναζήτησης. Στη συνέχεια θα αναλυθεί κάθε λειτουργική απαίτηση, συνοδευόμενη από αντίστοιχες οθόνες (screenshots) της διεπαφής μας, προκειμένου να γίνει κατανοητό πως αντιμετωπίσαμε την κάθε απαίτηση.

Τίτλος: Αναπαραγωγή ενός οπτικοακουστικού αρχείου

Περιγραφή: Ο χρήστης πρέπει να έχει τη δυνατότητα της αναπαραγωγής ενός οπτικοακουστικού αρχείου από τη συλλογή πολυμεσικού περιεχομένου, μαζί με όλες τις συνοδευόμενες πληροφορίες. Οι πληροφορίες αυτές περιλαμβάνουν τους τίτλους και τις περιγραφές των σκηνών (scenes) και των λήψεων (shots) του περιεχομένου αντίστοιχα, οπτική σημείωση των ηθοποιών που εμφανίζονται πάνω στο πλαίσιο αναπαραγωγής, μαζί με τα ονόματά τους και τέλος τα σχόλια που έχουν γίνει και αφορούν το περιεχόμενο. Επιπρόσθετα, οι παραπάνω πληροφορίες πρέπει να εμφανίζονται στα κατάλληλα χρονικά σημεία σε συγχρονισμό πάντα με την αναπαραγωγή του οπτικοακουστικού αρχείου.

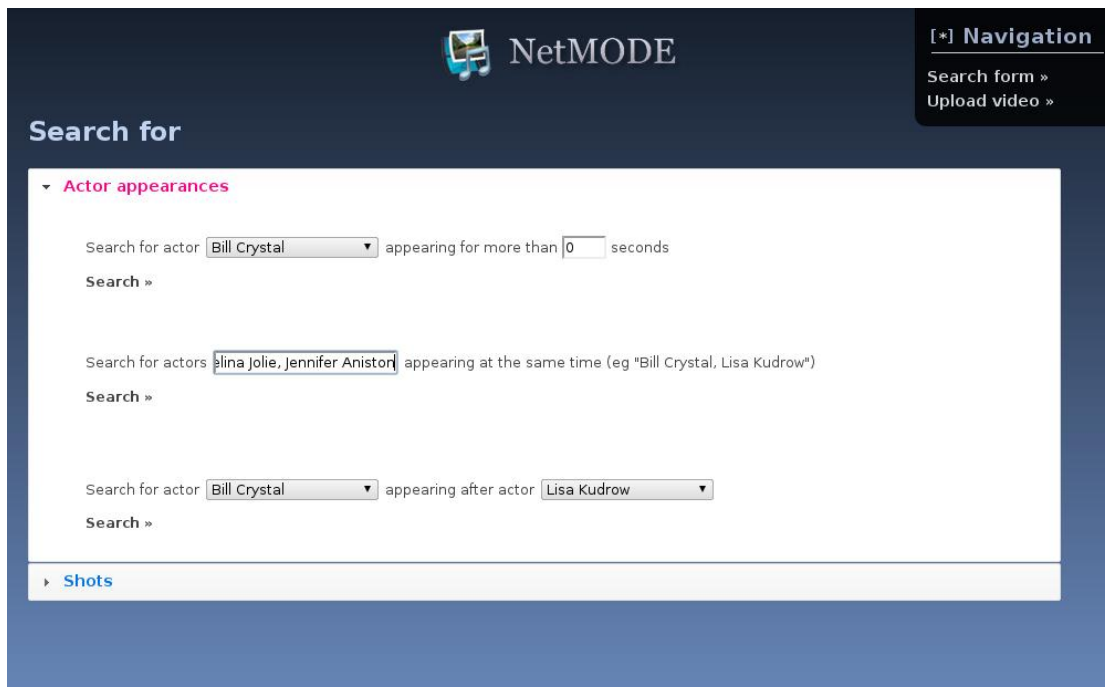


Σχήμα 10: Λειτουργική απαίτηση προβολής οπτικοακουστικού αρχείου

Τίτλος: Καθορισμός κριτηρίων αναζήτησης οπτικοακουστικού περιεχομένου

Περιγραφή: Ο χρήστης έχει τη δυνατότητα αναζήτησης οπτικοακουστικού περιεχομένου μέσα από μια φόρμα, η οποία του επιτρέπει να καθορίσει τα κριτήρια βάση των οποίων θα πραγματοποιηθεί η αναζήτηση. Οι δυνατότητες αναζήτησης που προσφέρει η διεπαφή πρέπει να βρίσκονται σε ένα προς ένα («1-1») αντιστοιχία με τις δυνατότητες αναζήτησης που έχουν υλοποιηθεί από το μηχανισμό αναζήτησης. Πιο συγκεκριμένα έχουμε:

- Αναζήτηση για την εμφάνιση ενός ηθοποιού, η οποία υπερβαίνει μια συγκεκριμένη χρονική διάρκεια. Ο χρήστης επιλέγει από μια λίστα το όνομα του ηθοποιού και πληκτρολογεί μια χρονική διάρκεια σε δευτερόλεπτα (προεπιλεγμένη χρονική διάρκεια 0).
- Αναζήτηση για εμφάνιση περισσότερων ηθοποιών ταυτόχρονα. Ο χρήστης μπορεί να καταχωρήσει αυθαίρετο πλήθος ηθοποιών, μέσω των ονομάτων τους σε ένα πεδίο κειμένου.
- Αναζήτηση για την εμφάνιση ενός ηθοποιού, αμέσως μετά την εμφάνιση κάποιου άλλου συγκεκριμένου ηθοποιού. Ο χρήστης μπορεί να επιλέξει μέσα από δύο λίστες το όνομα του κάθε ηθοποιού αντίστοιχα.
- Αναζήτηση μιας λήψης (shot) με βάση την περιγραφή της. Ο χρήστης έχει τη δυνατότητα να καταχωρήσει κείμενο με βάση το οποίο θα γίνει αναζήτηση στις περιγραφές των λήψεων όλων των διαθέσιμων οπτικοακουστικών αρχείων.



The screenshot shows the NetMODE search interface. At the top right, there is a navigation menu with the text "[*] Navigation" and two links: "Search form »" and "Upload video »". The main content area is titled "Search for" and contains a section for "Actor appearances" which is currently expanded. This section has three search options:

- Search for actor appearing for more than seconds. Below this is a "Search »" button.
- Search for actors appearing at the same time (eg "Bill Crystal, Lisa Kudrow"). Below this is a "Search »" button.
- Search for actor appearing after actor . Below this is a "Search »" button.

At the bottom of the search area, there is a link for "Shots" with a right-pointing arrow.

Σχήμα 11: Λειτουργική απαίτηση καθορισμού κριτηρίων αναζήτησης

Τίτλος: Προβολή των αποτελεσμάτων αναζήτησης πολυμεσικού περιεχομένου

Περιγραφή: Σαν περίπτωση χρήσης (use case) της αναζήτησης, η προβολή των αποτελεσμάτων είναι αυτή που ακολουθεί τον καθορισμό των κριτηρίων αναζήτησης. Πρέπει να υπάρχει η δυνατότητα προβολής μιας λίστας με τα αποτελέσματα που επέστρεψε η αναζήτηση ή σε περίπτωση αποτυχίας να υπάρχει η κατάλληλη ενημέρωση. Το αποτέλεσμα της αναζήτησης είναι ένα σύνολο τμημάτων κάποιων εκ των πολυμεσικών αρχείων της βάσης, σχετικών με τα κριτήρια της αναζήτησης. Έτσι, για κάθε αποτέλεσμα που επιστρέφεται υπάρχει η πληροφορία του τίτλου για το συγκεκριμένο αποτέλεσμα καθώς και η πληροφορία των χρονικών σημείων έναρξης και λήξης. Επίσης, υπάρχει μια προεπισκόπηση (preview) του εκάστοτε επιστραφέντος αποτελέσματος, μέσω ενός αριθμού πλαισίων (frames) του αντίστοιχου κάθε φορά χωρίου. Τέλος, δίνεται στο χρήστη η δυνατότητα της αξιολόγησης του κάθε αποτελέσματος, το οποίο είναι το σημείο επαφής των χρηστών με το μηχανισμό ανατροφοδότησης σχετικότητας, επηρεάζοντας έτσι μελλοντικές αναζητήσεις. Για λόγους παρουσίασης του μηχανισμού ανατροφοδότησης σχετικότητας εμφανίζεται και η αξιολόγηση του κάθε αποτελέσματος.

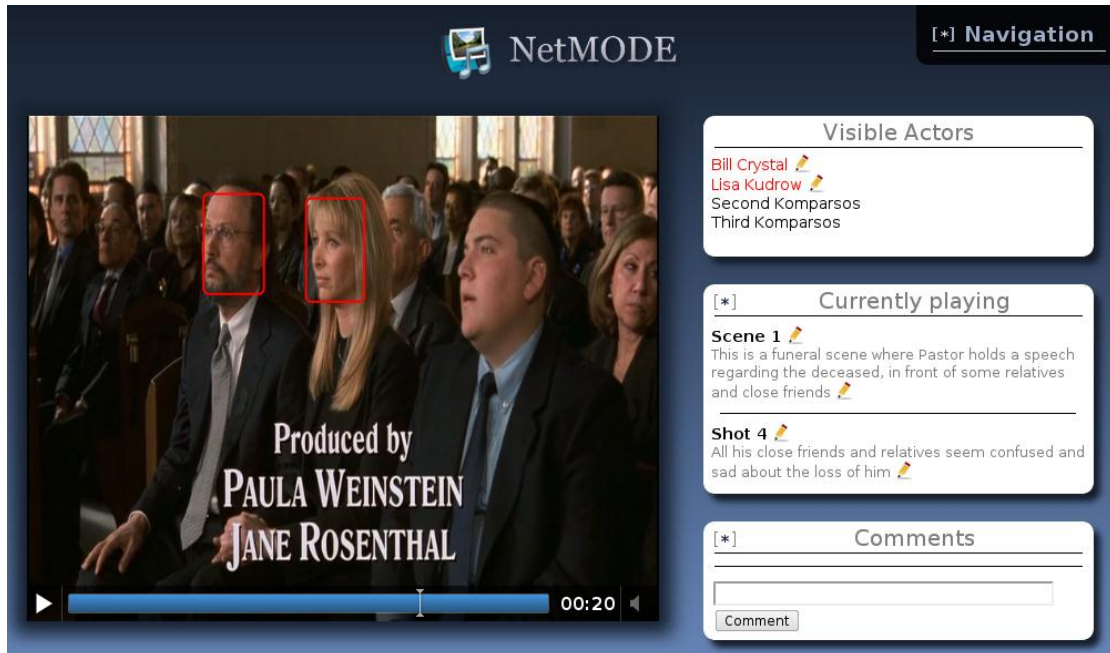


Σχήμα 12: Λειτουργική απαίτηση προβολής αποτελεσμάτων αναζήτησης

Τίτλος: Προβολή ενός αποτελέσματος αναζήτησης

Περιγραφή: Το επόμενο βήμα της αναζήτησης είναι να μεταβεί ο χρήστης στην προβολή του αποτελέσματος, δηλαδή ενός κατάλληλα επιλεγμένου χωρίου από τον αλγόριθμο αναζήτησης. Ο χρήστης αφού επιλέξει να δει κάποιο αποτέλεσμα από αυτά που

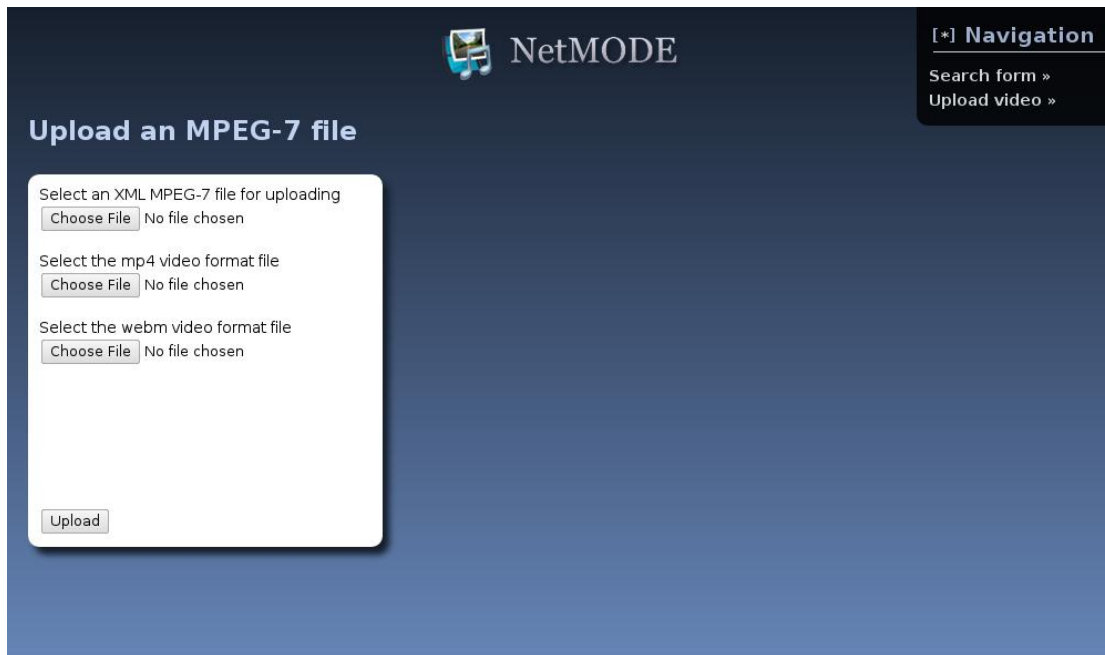
παρουσιάστηκαν στη σελίδα αποτελεσμάτων, μεταβαίνει στην αναπαραγωγή του κατάλληλου χωρίου του οπτικοακουστικού αρχείου, με παράλληλη προβολή των αντίστοιχων μόνο πληροφοριών. Σημειώνεται οπτικά μόνο ο ηθοποιός ή οι ηθοποιοί για τους οποίους έγινε η αναζήτηση.



Σχήμα 13: Λειτουργική απαίτηση προβολής ενός αποτελέσματος αναζήτησης

Τίτλος: Καταχώρηση πολυμεσικού αρχείου με τις συνοδευόμενες πληροφορίες

Περιγραφή: Απαραίτητη προϋπόθεση προκειμένου ο χρήστης να χρησιμοποιήσει όλες τις προηγούμενες δυνατότητες είναι να μπορεί να καταχωρήσει ένα πολυμεσικό αρχείο με τις πληροφορίες που το συνοδεύουν. Αυτό γίνεται μέσω μιας κατάλληλα σχεδιασμένης φόρμας, όπου καταχωρούνται τα πολυμεσικά αρχεία και τα αντίστοιχα αρχεία περιγραφής MPEG-7 [8]. Αφού επικυρωθεί η επιλογή του χρήστη μεταβαίνουν προς κατάλληλη επεξεργασία και αποθήκευση στη βάση δεδομένων.



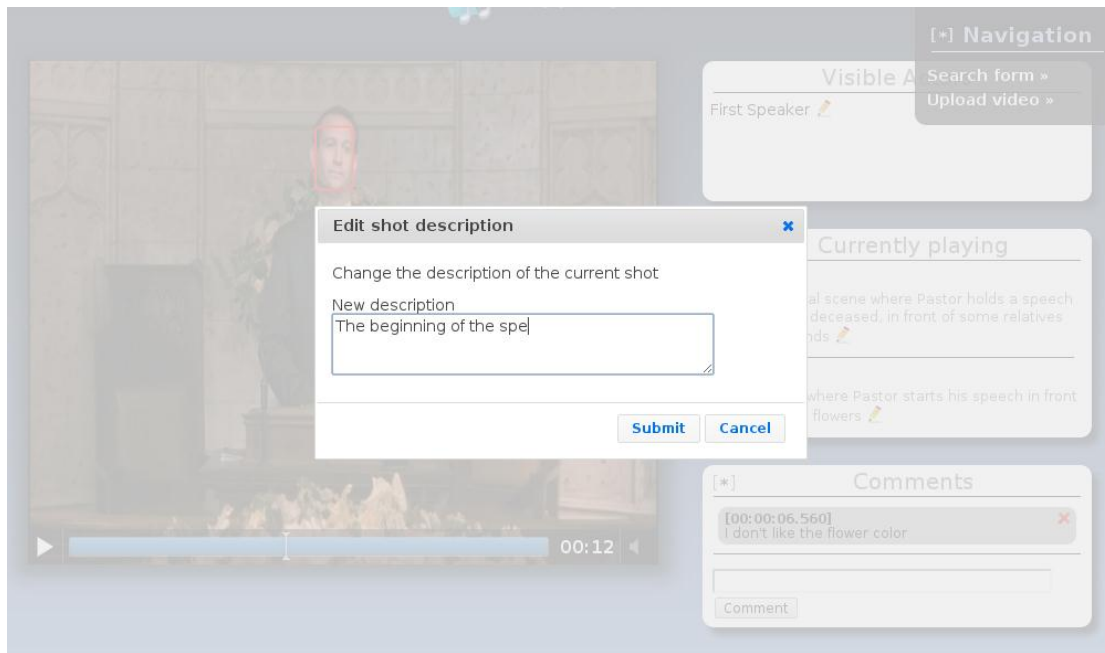
Σχήμα 14: Λειτουργική απαίτηση καταχώρησης πολυμεσικού αρχείου

Τίτλος: Επεξεργασία της υπάρχουσας πληροφορίας ενός πολυμεσικού αρχείου

Περιγραφή: Ο χρήστης πρέπει να έχει τη δυνατότητα αλληλεπίδρασης με το σύστημα πιο δυναμικά, μεταβάλλοντας τις υπάρχουσες πληροφορίες. Αυτό γίνεται κατά την αναπαραγωγή ενός πολυμεσικού αρχείου, όπου κάθε πληροφορία μπορεί να επεξεργαστεί πατώντας το αντίστοιχο εικονίδιο επεξεργασίας. Η απαίτηση διασπάται περαιτέρω σε μικρότερες:

- Επεξεργασία του ονόματος ενός ηθοποιού σε μια εμφάνισή του
- Επεξεργασία του τίτλου μιας σκηνής (scene name)
- Επεξεργασία της περιγραφής μιας σκηνής (scene description)
- Επεξεργασία του τίτλου μιας λήψης (shot name)
- Επεξεργασία της περιγραφής μιας λήψης (shot description)

Όλες οι παραπάνω απαιτήσεις ικανοποιούνται με τον ίδιο τρόπο, μέσω ενός διαλογικού περιβάλλοντος, στο οποίο ο χρήστης δίνει τις νέες τιμές για τις προαναφερθείσες ιδιότητες. Παρουσιάζεται ενδεικτικά πως γίνεται η επεξεργασία της περιγραφής μιας σκηνής.



Σχήμα 15: Λειτουργική απαίτηση επεξεργασίας υπάρχουσας πληροφορίας

5.3 Καθορισμός μη λειτουργικών απαιτήσεων

Εκτός από τις λειτουργικές απαιτήσεις, οι οποίες περιγράφουν τις παρεχόμενες δυνατότητες, πρέπει να καθοριστούν και οι μη λειτουργικές απαιτήσεις (non-functional requirements). Αυτές ουσιαστικά αποτελούν τα ποιοτικά στοιχεία που θέλουμε να χαρακτηρίζουν τη διεπαφή μας. Θα παρουσιάσουμε μια λίστα με τις σημαντικότερες μη λειτουργικές απαιτήσεις που θα προσπαθήσουμε να ικανοποιήσουμε, συνοδευόμενες από μια σύντομη περιγραφή.

Χρηστικότητα (usability): Η χρηστικότητα αναφέρεται στην ευκολία χρήσης και εκμάθησης της διεπαφής. Επιτυγχάνεται μέσω σωστού σχεδιασμού παρέχοντας εύκολη πρόσβαση στις προσφερόμενες δυνατότητες, παρουσίαση των πληροφοριών με σαφή και συνοπτικό τρόπο και ελαχιστοποίηση των ασαφειών για τα αποτελέσματα που θα επιφέρουν οι ενέργειες του χρήστη. Η κάλυψη αυτής της απαίτησης είναι ιδιαίζουσας σημασίας για κάθε γραφική διεπαφή, γιατί καθορίζει το βαθμό στον οποίον ο χρήστης θα μπορέσει να εκμεταλλευτεί τις δυνατότητες που του προσφέρει το σύστημα για το οποίο σχεδιάστηκε η διεπαφή.

Σταθερότητα (stability): Αφορά τη σταθερότητα λειτουργίας ενός συστήματος, δηλαδή την ικανότητα παροχής υπηρεσιών αδιάλειπτα. Εκτός από την εξάλειψη προγραμματιστικών λαθών (bugs), επιτυγχάνεται και με τη χρήση δοκιμασμένων τεχνολογιών, καθώς αυτές αποτελούν ένα είδος εγγύησης προς αυτή την κατεύθυνση. Στην περίπτωση μας πρέπει προφανώς να φροντίσουμε για την εξάλειψη προγραμματιστικών

λαθών και να υιοθετήσουμε συγκεκριμένους κανόνες σχεδίασης διεπαφών για την αποφυγή απρόσμενων καταστάσεων για το χρήστη.

Ταχύτητα/χρόνος απόκρισης (speed/response time): Η απαίτηση γι' αυτό το ποιοτικό χαρακτηριστικό καθορίζεται από το είδος της εφαρμογής. Στην περίπτωσή μας πρόκειται για ένα σύστημα πραγματικού χρόνου, όπου ο χρήστης αλληλεπιδρά με το σύστημα και περιμένει απαντήσεις από αυτό ακαριαία. Ο ανεκτός χρόνος απόκρισης για τη δική μας περίπτωση δεν πρέπει να ξεπερνάει τα λίγα δευτερόλεπτα (2-3) και ιδανικά θα έπρεπε να είναι της τάξης των μερικών milliseconds. Ο παράγοντας αυτός επηρεάζεται από πολλές παραμέτρους (όπως για παράδειγμα ταχύτητα σκληρού δίσκου, φόρτος δικτύου κ.τ.λ.) και για το λόγο αυτό θα πρέπει να ληφθεί υπόψη στις σχεδιαστικές επιλογές που θα γίνουν.

Scalability (επεκτασιμότητα): Αυτό είναι ένα σημαντικό στοιχείο και αφορά τη δυνατότητα που προσφέρει ο υπάρχων σχεδιασμός για μελλοντική επέκταση. Ακόμα και αν μια τέτοια επέκταση δεν φαντάζει πιθανή, η δυνατότητα επέκτασης είναι ένα σημαντικό επιθυμητό στοιχείο. Ιδανικά, λοιπόν, θέλουμε η διεπαφή μας να μπορεί να χρησιμοποιηθεί χωρίς μεγάλες τροποποιήσεις από μεγάλο αριθμό χρηστών, παρόλο που αρχικά απευθύνεται σε μια μικρή κοινότητα ατόμων.

Ασφάλεια (security): Η σημασία αυτής της απαίτησης συνδέεται στενά με τον αριθμό των χρηστών στους οποίους αναφέρεται η εφαρμογή. Μεγάλος αριθμός χρηστών συνεπάγεται μεγαλύτερους κινδύνους, αφού αυξάνει και ο αριθμός των πιθανών κακόβουλων χρηστών. Στην περίπτωσή μας, απευθυνόμαστε σε ένα μικρό αριθμό χρηστών, οπότε το ποιοτικό χαρακτηριστικό της ασφάλειας είναι δευτερεύουσας σημασίας, ωστόσο κάνουμε κάποιους στοιχειώδεις ελέγχους στην είσοδο του χρήστη προκειμένου να αποφευχθούν ανεπιθύμητες παρενέργειες στη λειτουργία του συστήματος (SQL injections, Cross-site scripting).

Συντηρησιμότητα (maintainability): Η συντηρησιμότητα αφορά την ευκολία συντήρησης ενός προγράμματος, τόσο από τον ίδιο το δημιουργό όσο και από άλλους προγραμματιστές. Είναι ένα ιδιαίτερα επιθυμητό χαρακτηριστικό προκειμένου να καθίσταται εύκολη η διόρθωση λαθών και συντήρηση από τον ίδιο το δημιουργό ή από άλλους μελλοντικούς προγραμματιστές, κατά τον κύκλο ζωής ενός πακέτου λογισμικού. Προς αυτή την κατεύθυνση βοηθάει μια άρτια αρχιτεκτονική σχεδίαση και η σωστή οργάνωση του κώδικα.

Ανοιχτά πρότυπα (open source standards): Ένα ακόμα στοιχείο που θα θέλαμε για την υλοποίηση της εφαρμογής μας είναι η χρήση ανοιχτών προτύπων. Παρ' όλο που αυτό δεν εγγυάται απαραίτητα κάτι για την ποιότητα μιας εφαρμογής, τα ανοιχτά πρότυπα συνήθως προσφέρουν μια εγγύηση λόγω της υιοθέτησής τους από μεγάλο πλήθος εφαρμογών.

Οι μη λειτουργικές απαιτήσεις θα επηρεάσουν σε μεγάλο βαθμό τις αποφάσεις που αφορούν την υλοποίηση και την επιλογή των κατάλληλων εργαλείων, καθώς η επίτευξή τους διευκολύνεται σε σημαντικό βαθμό από τη χρήση των καταλληλότερων – και πιο ταιριαστών για το είδος της εφαρμογής - γλωσσών και συστημάτων.

5.4 Τεχνικές λεπτομέρειες υλοποίησης

Σε αυτό το χωρίο θα ασχοληθούμε με κάποιες σχεδιαστικές αποφάσεις τεχνικής φύσεως και τους λόγους που οδήγησαν σε αυτές, χωρίς υπερβολική τεχνική ανάλυση. Πιο συγκεκριμένα θα αναλυθούν οι τεχνολογίες και οι γλώσσες που επιλέχθηκαν για την υλοποίηση της γραφικής διεπαφής χρήστη και οι βασικότεροι λόγοι που οδήγησαν στις επιλογές αυτές. Αποφασίσαμε να υλοποιήσουμε τη γραφική διεπαφή σαν διαδικτυακή εφαρμογή ιστού (web application) χρησιμοποιώντας τις γλώσσες HTML, CSS και Javascript. Συνεπώς, ο χρήστης θα προσπελάει την διεπαφή μέσω ενός προγράμματος φυλλομετρητή ιστού (web browser). Οι φυλλομετρητές ιστού αποτελούν ένα αξιόπιστο και δοκιμασμένο περιβάλλον εκτέλεσης του κώδικά μας, προσφέροντας σταθερότητα, ενώ υπάρχει μια μεγάλη και ενεργή κοινότητα υποστήριξης (community support) σε περίπτωση εμφάνισης προβλημάτων.

Η HTML5 [38] – η πιο πρόσφατη έκδοση της HTML – προσφέρει πλέον πολλές δυνατότητες για την ανάπτυξη πολυμεσικών διαδικτυακών εφαρμογών [39], παρέχοντας ένα πολύ ισχυρό πλαίσιο ανάπτυξης στους προγραμματιστές (development framework). Ουσιαστικά η HTML5 [38] βρίσκεται ακόμη υπό ανάπτυξη και έχει υλοποιηθεί ένα υποσύνολο μόνο των τελικών της δυνατοτήτων, το οποίο ωστόσο χρησιμοποιείται ευρέως τα τελευταία χρόνια σε διαδικτυακές εφαρμογές. Η γλώσσα αυτή ενδείκνυται πλέον για την ενσωμάτωση οπτικοακουστικού περιεχομένου σε μια σελίδα, δίνοντας στους προγραμματιστές τη δυνατότητα να ασχοληθούν με υψηλότερου επιπέδου προγραμματισμό, χωρίς να εστιάζουν σε τεχνικές λεπτομέρειες χαμηλού επιπέδου, όπως είναι οι βιβλιοθήκες αναπαραγωγής ενός πολυμεσικού αρχείου (video codecs), αφήνοντας τις αρμοδιότητες αυτές στη δικαιοδοσία του εκάστοτε φυλλομετρητή.

Η Javascript [40] είναι η γλώσσα που προσφέρει το μεγαλύτερο βαθμό αλληλεπίδρασης και επικοινωνίας με το χρήστη στη διεπαφή μας, επιτρέποντας τον προγραμματισμό των αντικειμένων που εμφανίζονται μέσω HTML στο χρήστη. Η δημοτικότητά της αυξάνεται συνεχώς, γεγονός που οδηγεί σε νέες εκδόσεις που προσφέρουν περισσότερες δυνατότητες και αυξημένη ταχύτητα. Ακόμα, υπάρχει μια πολυπληθής ενεργή κοινότητα προγραμματιστών ικανή να προσφέρει υποστήριξη, στοιχείο ιδιαίτερα ενθαρρυντικό για την υιοθέτηση μιας γλώσσας σε κάποια εφαρμογή. Η CSS [41] (Cascading Style Sheets) είναι μια γλώσσα μορφοποίησης της εμφάνισης μιας σελίδας και στην περίπτωσή μας χρησιμοποιείται για την επίτευξη της επιθυμητής χρηστικότητας

(usability), δηλαδή για τη δημιουργία μιας ξεκάθαρης και περιεκτικής διεπαφής, εύκολης στη χρήση.

Ένα από τα μεγαλύτερα πλεονεκτήματα που προσφέρει η επιλογή μας είναι ότι δε χρειάζεται εξειδικευμένο λογισμικό από την πλευρά του χρήστη, αποφεύγοντας πολύπλοκες και χρονοβόρες εγκαταστάσεις. Η μόνη απαίτηση είναι ένας σύγχρονος φυλλομετρητής ιστού, εφαρμογή που βρίσκεται εγκατεστημένη στη συντριπτική πλειοψηφία των προσωπικών υπολογιστών. Επιπλέον, η χρήση φυλλομετρητών αποτελεί εγγύηση για τη σταθερότητα (stability) της γραφικής διεπαφής ως ένα βαθμό, αφού προσφέρουν την απαιτούμενη αξιοπιστία, όντας από τις πιο απαιτητικές εφαρμογές σε κυκλοφορία. Τέλος, ένα ακόμη θετικό στοιχείο είναι η χρήση ανοιχτών προτύπων, έχοντας έτσι αυξημένη υποστήριξη από ενεργές κοινότητες προγραμματιστών, αλλά και προσαρμοστικότητας σε περίπτωση που αυτή απαιτηθεί για την επέκταση ή τροποποίηση της γραφικής διεπαφής.

Στο σημείο αυτό θα αναλύσουμε τα αρνητικά σημεία και μειονεκτήματα της επιλογής μας: οι επιλογές που προσφέρει η HTML5 [38] στον πολυμεσικό τομέα είναι ακόμα σχετικά περιορισμένες, καθώς το όλο πρότυπο βρίσκεται ακόμα υπό ανάπτυξη και έχει υλοποιηθεί ένα μικρό μόνο μέρος από αυτά που πρόκειται να γίνουν διαθέσιμα κάποια στιγμή μελλοντικά. Επιπλέον, οι δυνατότητες που παρέχονται στον προγραμματιστή περιορίζονται από το πρόγραμμα του φυλλομετρητή ιστού, στον οποίον εκτελείται ο κώδικας που έχει συνταχθεί, δεν υπάρχει δηλαδή η ελευθερία που θα προσέφερε η ανάπτυξη μιας ανεξάρτητης, αυτόνομης εφαρμογής. Το τελευταίο μειονέκτημα έχει να κάνει με τα προγράμματα φυλλομετρητή, τα οποία σε αρκετές περιπτώσεις δε σέβονται τα καθορισμένα πρότυπα με αποτέλεσμα ο κώδικας να εκτελείται με κάποιες παρενέργειες. Η εφαρμογή οφείλει να τρέχει σε διαφορετικούς φυλλομετρητές ιστού για να είναι όσο το δυνατόν πιο προσβάσιμη από μεγάλο αριθμό χρηστών. Το πρόβλημα αυτό απαιτεί ειδική μέριμνα από την πλευρά του προγραμματιστή, ώστε να αντιμετωπίσει τυχόν ασυμβατότητες μεταξύ των διάφορων φυλλομετρητών.

Ένα ζήτημα τεχνικής φύσεως είναι η επικοινωνία της γραφικής διεπαφής χρήστη με το μηχανισμό αναζήτησης και την αποθηκευμένη πληροφορία της βάσης, προκειμένου να προβάλλει στο χρήστη αυτά που εκείνος ζητάει. Η σχεδιαστική επιλογή για υλοποίηση με τη μορφή διαδικτυακής εφαρμογής ιστού απαιτεί σε πρώτο στάδιο την ύπαρξη ενός εξυπηρετητή ιστού (web server), ο οποίος θα μεσολαβήσει για την αποστολή του περιεχομένου στη διεπαφή. Ο εξυπηρετητής πρέπει να είναι συμβατός με τη γλώσσα προγραμματισμού του μηχανισμού αναζήτησης, απαίτηση που όμως καλύπτεται, αφού η PHP [36] είναι προσανατολισμένη προς αυτή την κατεύθυνση. Έχουμε τη δυνατότητα επιλογής μέσα από ένα περιορισμένο πλήθος λογισμικού εξυπηρετητών (Apache, Lighttpd, Internet Information Services, nginx κ.τ.λ.) και επιλέξαμε τον Apache, καθώς είναι ο δημοφιλέστερος τέτοιος εξυπηρετητής. Επιπλέον, είναι λογισμικό ανοιχτού κώδικα, με

συνεχείς αναβαθμίσεις, στοιχείο απαραίτητο για την εξάλειψη προβλημάτων ασφάλειας (security issues), τα οποία είναι από τα σημαντικότερα τρωτά σημεία αυτών των εφαρμογών. Θα μπορούσαμε όμως να χρησιμοποιήσουμε και κάποιον άλλον εξυπηρετητή, αφού οι απαιτήσεις μας είναι περιορισμένες και δε χρησιμοποιούμε εξεζητημένες δυνατότητες.

Επόμενο τεχνικό ζήτημα είναι η μορφή της πληροφορίας που δέχεται η διεπαφή μας. Επιλέξαμε ο εξυπηρετητής να αποστέλλει πληροφορίες σε μορφή JSON [42] γιατί είναι άμεσα προσπελάσιμη και συμβατή με τη γλώσσα Javascript, η οποία αναλαμβάνει την αξιοποίηση των πληροφοριών που σχετίζονται με το οπτικοακουστικό υλικό. Η μορφή JSON [42] χρησιμοποιείται για τη σειριοποίηση (serialization) και τη μεταφορά δεδομένων μέσω δικτύου και παρ' όλο που συνδέεται στενά με την Javascript [40], μπορεί να χρησιμοποιηθεί ανεξάρτητα από αυτήν. Εναλλακτική επιλογή θα μπορούσε να είναι η αποστολή της πληροφορίας σε μορφή XML [10], με την οποία είμαστε ίδιοι οικείοι, όμως η ευκολία χρήσης της πληροφορίας από τη Javascript μας οδήγησαν σε αυτή την επιλογή. Επιπλέον, η κωδικοποίηση της πληροφορίας σε μορφή JSON [42] από την πλευρά του εξυπηρετητή γίνεται άμεσα, με την κλήση μιας μόνο συνάρτησης, οπότε η ευκολία και η ευελιξία που μας προσφέρεται την καθιστούν την πιο λογική επιλογή.

Αφού αναλύσαμε τη δική μας τεχνική προσέγγιση στο ζήτημα της διεπαφής του χρήστη, θα ασχοληθούμε με τις εναλλακτικές επιλογές που είχαμε, σχολιάζοντάς τις συνοπτικά. Για το ζήτημα της ανάπτυξης της γραφικής διεπαφής είχαμε μια πληθώρα διαθέσιμων επιλογών και θα μπορούσαμε να την αναπτύξουμε σαν μια εφαρμογή για κάποιο λειτουργικό σύστημα και όχι απαραίτητα σαν διαδικτυακή εφαρμογή. Μια πρώτη εναλλακτική επιλογή θα ήταν η ανάπτυξη διεπαφής σε μια γλώσσα όπως η C++, όπου θα είχαμε τον απόλυτο έλεγχο σε κάθε στοιχείο της διεπαφής μας, σε αντίθεση με τη δική μας επιλογή. Στη δική μας περίπτωση οι δυνατότητές μας περιορίζονται από την ελευθερία που μας παρέχει ο φυλλομετρητής ιστού, ωστόσο αυτοί οι περιορισμοί δεν μας αποτρέπουν από την ανάπτυξη των δυνατοτήτων που θέλουμε και επομένως δεν αποτελούν ανασταλτικό παράγοντα. Η χρήση C++ θα προσέφερε μια ελαφρώς αισθητά καλύτερη επίδοση, αφού βρίσκεται πιο κοντά στο υλικό του υπολογιστή (hardware). Το στοιχείο αυτό όμως είναι παράλληλα και το μεγαλύτερο μειονέκτημα χρήσης αυτής της γλώσσας, αφού η ανάπτυξη θα ήταν μια πολύ χρονοβόρα διαδικασία και θα αναλώναμε πολύ χρόνο στην ανάπτυξη λειτουργιών που σε άλλες γλώσσες υπάρχουν έτοιμες. Αυτό το πρόβλημα μετριάζεται σε σημαντικό βαθμό αν αναπτύσσαμε τη διεπαφή μας σε Java. Η Java προσφέρει στον προγραμματιστή όλες εκείνες τις ευκολίες μέσω βιβλιοθηκών, προκειμένου να μειωθεί κατά πολύ ο χρόνος ανάπτυξης. Επιπλέον, υπάρχει το πλεονέκτημα της λειτουργίας της διεπαφής ανεξαρτήτως συστήματος, αφού ο κώδικας μιας Java εφαρμογής τρέχει φαινομενικά σε μια εικονική μηχανή και όχι απευθείας στο λειτουργικό σύστημα. Από τεχνική άποψη, η Java θα μπορούσε κάλλιστα να αποτελέσει την επιλογή μας, ωστόσο μια

απόφαση βασίζεται συχνά και στις επιθυμίες, εμπειρίες και γνώσεις του εκάστοτε προγραμματιστή. Επιπρόσθετα, με την επιλογή μας χρησιμοποιήσαμε κάποιες από τις πιο σύγχρονες τεχνολογίες, με σπουδαίες προοπτικές για το μέλλον, στοιχείο που αποτελεί οπωσδήποτε μια τεχνολογική πρόκληση. Από τη στιγμή που επιλέξαμε να αναπτύξουμε τη γραφική διεπαφή χρήστη σε μορφή διαδικτυακής εφαρμογής, υπήρχε άλλη μια δυνατή εναλλακτική επιλογή, η ανάπτυξη μιας εφαρμογής Flash . Οι εφαρμογές Flash είναι ευρέως διαδεδομένες στον παγκόσμιο ιστό, ωστόσο υπάρχει πληθώρα λόγων για τους οποίους αποφύγαμε μια τέτοια επιλογή. Αν και προσφέρει πολλές δυνατότητες διαδραστικότητας και επικοινωνίας με το χρήστη και ίσως υπερτερεί στον τομέα αυτόν έναντι της Javascript, έχει χαμηλό χρόνο απόκρισης και η απόδοση εκτέλεσης είναι χαμηλή. Επιπρόσθετα, απαιτείται η εγκατάσταση ειδικού λογισμικού στο φυλλομετρητή (Flash plugin), γεγονός προφανώς μη επιθυμητό από τους χρήστες. Επίσης, δεν υπάρχει καλή υποστήριξη για όλα τα λειτουργικά συστήματα και δεν βασίζεται σε ανοιχτά πρότυπα, όπως η δική μας επιλογή. Οι υπόλοιποι αποτρεπτικοί λόγοι αφορούν κυρίως τη μελλοντική επέκταση του συστήματός μας: σε περίπτωση που το σύστημα χρησιμοποιηθεί ευρέως από μεγάλο αριθμό χρηστών, η χρήση Flash αποτρέπει τις μηχανές αναζήτησης από το να αρχειοθετήσουν τις πληροφορίες των σελίδων μας. Η φιλικότητα προς τις μηχανές αναζήτησης είναι ένα πολύ σημαντικό στοιχείο εφαρμογών που χρησιμοποιούνται ευρέως από χρήστες.

ΚΕΦΑΛΑΙΟ 6^ο

ΣΥΜΠΕΡΑΣΜΑΤΑ & ΑΠΟΤΕΛΕΣΜΑΤΑ

6.1 Ανάλυση σεναρίων χρήσης

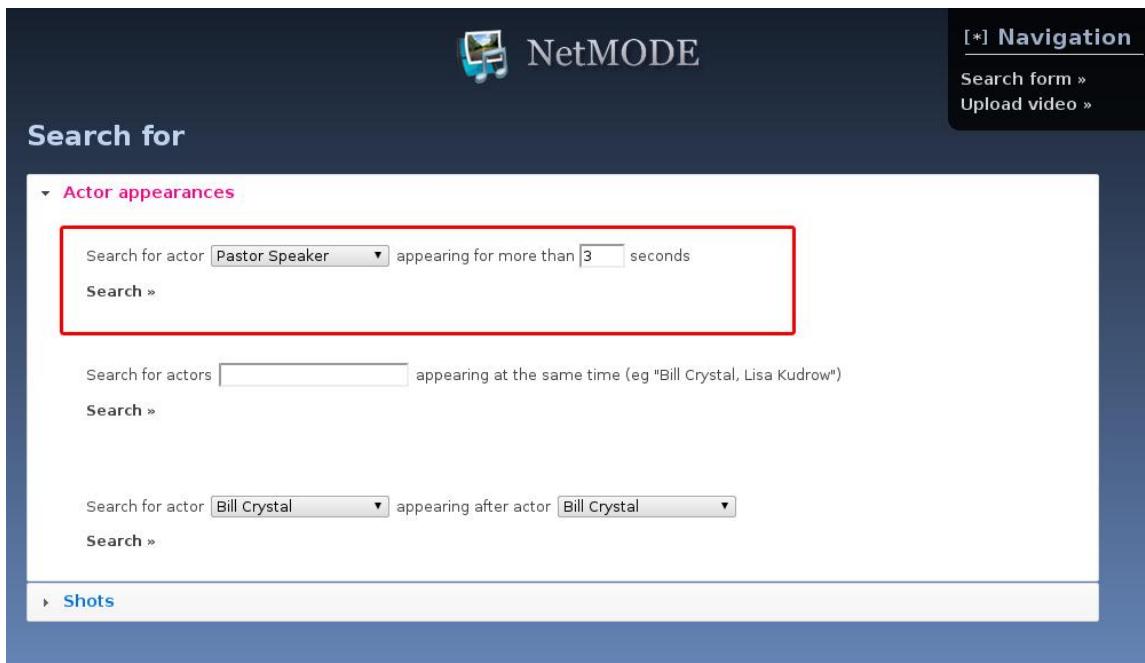
6.2 Ανάλυση εκτέλεσης αλγορίθμου αναζήτησης

6.1 Ανάλυση σεναρίων χρήσης

Η εξαγωγή μιας περιγραφής από ένα οπτικοακουστικό αρχείο δεν είναι μια τετριμμένη διαδικασία και απαιτεί εξειδικευμένους αλγόριθμους για την εξόρυξη της πληροφορίας. Αυτό έχει ως αποτέλεσμα να μην έχουμε διαθέσιμο μεγάλο πλήθος οπτικοακουστικών αρχείων συνοδευόμενων από τις περιγραφές τους και να μη μπορούμε να κάνουμε έτσι ανάλυση επίδοσης των αλγορίθμων που έχουμε δημιουργήσει, αφού δεν μπορούν να εξαχθούν χρήσιμα συμπεράσματα. Για το λόγο αυτό θα εξετάσουμε δύο σενάρια χρήσης του όλου συστήματος από κάποιο χρήστη, παραθέτοντας τις αντίστοιχες οθόνες (screenshots) της διεπαφής, όπου κρίνεται απαραίτητο.

6.1.1 Σενάριο αναζήτησης εμφάνισης ενός ηθοποιού

Ο χρήστης θέλει να αναζητήσει μέσα από μια βάση οπτικοακουστικών αρχείων τα χωρία εκείνα στα οποία εμφανίζεται ένας ηθοποιός της αρεσκείας του (π.χ. Pastor Speaker) για μια ελάχιστη χρονική διάρκεια (π.χ. 3 δευτερόλεπτα). Το πρώτο στάδιο της αναζήτησης είναι η μετάβαση στη φόρμα αναζήτησης, όπου μπορεί ο χρήστης να καθορίσει τα κριτήρια που επιθυμεί. Από τις διαθέσιμες δυνατότητες αναζήτησης εκείνη που ανταποκρίνεται στην επιθυμία μας είναι η πρώτη, όπου πρέπει να επιλεγθεί το όνομα του ηθοποιού μέσα από μια λίστα και να πληκτρολογηθεί μια χρονική διάρκεια που αντιστοιχεί στην ελάχιστη διάρκεια των εμφανίσεων προς αναζήτηση. Αφού συμπληρωθούν τα παραπάνω, μπορεί να πατήσει στον αντίστοιχο σύνδεσμο και να μεταβεί στα αποτελέσματα της αναζήτησης. Στην παρακάτω εικόνα φαίνεται ο καθορισμός των απαιτήσεων μέσω της φόρμας.



The screenshot shows the NetMODE search interface. At the top, there is a navigation menu with 'Search form »' and 'Upload video »'. The main search area is titled 'Search for' and contains three search options under the 'Actor appearances' section. The first option is highlighted with a red box and shows a search for 'Pastor Speaker' appearing for more than 3 seconds. The second option is for searching for multiple actors at the same time. The third option is for searching for an actor appearing after another actor. Below the search options, there is a 'Shots' section.

Σχήμα 16: Καθορισμός κριτηρίου αναζήτησης για εμφάνιση ενός ηθοποιού

Αφού ο χρήστης έχει καθορίσει τα κριτήρια αναζήτησης και πατήσει στο σύνδεσμο μεταβαίνει στη σελίδα των αποτελεσμάτων. Σε περίπτωση που δεν υπάρχουν αποτελέσματα ενημερώνεται κατάλληλα, αλλιώς του εμφανίζεται μια λίστα με αυτά. Στην περίπτωσή μας υπάρχουν δύο εμφανίσεις που ανταποκρίνονται στα κριτήριά μας, όπως φαίνεται και στην εικόνα παρακάτω. Ο χρήστης μπορεί να προΐδεαστεί από τα περιεχόμενα των χωρίων που επιλέχθηκαν, αφήνοντας το ποντίκι πάνω από την αντίστοιχη εικόνα και να δει κάποια στιγμιότυπα του χωρίου.

NetMODE

[*] Navigation

Search form »

Upload video »

Search results

for appearances of **Pastor Speaker** for more than 3 seconds

Appearance: ACTOR_APP 10
Start: 00:00:24.160
End: 00:00:27.840
Score: 0

Appearance: ACTOR_APP 1
Start: 00:00:04.680
End: 00:00:12.720
Score: 0

Σχήμα 17: Προβολή των αποτελεσμάτων της αναζήτησης

Το επόμενο βήμα είναι η μετάβαση σε κάποιο από τα προηγούμενα αποτελέσματα. Έστω ότι ο χρήστης θέλει να δει το δεύτερο αποτέλεσμα, πρέπει να το επιλέξει πατώντας στην εικόνα που εμφανίζεται. Αφού το κάνει αυτό μεταβαίνει σε μια σελίδα αναπαραγωγής, όπου το κομμάτι προς αναπαραγωγή αντιστοιχεί στην αναζήτηση που κάναμε και οι πληροφορίες που εμφανίζονται αφορούν μόνο τα άτομα που αναζητήσαμε και το αντίστοιχο χρονικό χωρίο του αποτελέσματος, όπως βλέπουμε και στην επόμενη εικόνα.



Σχήμα 18: Προβολή του χωρίου του αποτελέσματος αναζήτησης

6.1.2 Σενάριο αξιολόγησης αποτελέσματος και μετέπειτα αναζητήσεων

Το σενάριο αυτό αποτελεί μια συνέχεια του προηγούμενου σεναρίου. Ο χρήστης στη σελίδα αποτελεσμάτων αποφασίζει να βαθμολογήσει τα αποτελέσματα που του επιστράφηκαν πατώντας στα αντίστοιχα εικονίδια. Επιλέγει να αξιολογήσει ως σχετικό αποτέλεσμα το δεύτερο και ως μη σχετικό το πρώτο. Η αξιολόγηση αυτή επηρεάζει τη σειρά των αποτελεσμάτων σε μελλοντικές αναζητήσεις, σε συνάρτηση πάντα με τις αξιολογήσεις που έχουν προηγηθεί από άλλους χρήστες. Για λόγους επίδειξης έχουμε φροντίσει να μην υπάρχουν άλλες αξιολογήσεις, οπότε κάνοντας ξανά την ίδια αναζήτηση παρατηρούμε μια αναδιάταξη των αποτελεσμάτων. Η αναδιάταξη αυτή αντανακλά τη γνώμη του χρήστη ως προς τη σχετικότητα των αποτελεσμάτων αναζήτησης. Στην παρακάτω εικόνα φαίνεται η αναδιάταξη που περιγράψαμε.

The screenshot shows the NetMODE search results page. At the top, there is a navigation menu with a search form and an upload video button. The main content area displays search results for 'Pastor Speaker' appearances. Two results are shown:

- Appearance: ACTOR_APP 1**: Start: 00:00:04.680, End: 00:00:12.720, Score: 1. The thumbnail shows a man in a suit standing at a podium with flowers.
- Appearance: ACTOR_APP 10**: Start: 00:00:24.160, End: 00:00:27.840, Score: -1. The thumbnail shows a man in a suit standing at a podium with flowers, viewed from a different angle.

Σχήμα 19: Αναδιάταξη των αποτελεσμάτων αναζήτησης μετά από αξιολόγηση

Προφανώς, η σημασία μιας τέτοιας δυνατότητας δεν μπορεί να φανεί μέσα από ένα τόσο μικρό πλήθος οπτικοακουστικών αρχείων και χρηστών, αφού και τα αποτελέσματα και οι αξιολογήσεις είναι πολύ λίγες. Προκειμένου να φανεί η αξία του μηχανισμού μας πρέπει να επιστρέφονται πολλά αποτελέσματα και να υπάρχουν πολλές αξιολογήσεις χρηστών, ώστε να υπάρχει μια πιο σαφής εικόνα της σχετικότητας του κάθε αποτελέσματος. Ωστόσο, ο μηχανισμός έχει σχεδιαστεί με τέτοιο τρόπο, ώστε η αύξηση του όγκου των αρχείων και των χρηστών να μην απαιτεί περαιτέρω αλλαγές.

6.2 Ανάλυση εκτέλεσης αλγορίθμου αναζήτησης

Στη συνέχεια θα εξετάσουμε διεξοδικά την εσωτερική λειτουργία του μηχανισμού αναζήτησης για δύο διαφορετικά σενάρια. Πρώτα, μέσω των κατάλληλων ερωτημάτων θα επιλέξουμε ένα υποσύνολο της διαθέσιμης πληροφορίας και στη συνέχεια θα το διαχειριστούμε περαιτέρω αν αυτό είναι απαραίτητο. Θα δούμε ποιο είναι το σύνολο της υπάρχουσας πληροφορίας, αλλά και ποιο υποσύνολό της επιλέγεται κάθε φορά. Τέλος, θα παρατηρήσουμε τις μεταβολές στο επιλεγμένο υποσύνολο πληροφορίας, ανάλογα με τις ενέργειες του αλγορίθμου. Η πληροφορία που μας ενδιαφέρει για τα σενάρια που θα εξετάσουμε βρίσκεται σε τρεις πίνακες: τον πίνακα εμφανίσεων των ηθοποιών (ActorAppearances), τον πίνακα με τα στοιχεία των ηθοποιών (Actors) και τον πίνακα με τις αξιολογήσεις του μηχανισμού ανατροφοδότησης σχετικότητας (RelevanceFeedback). Παρακάτω βλέπουμε τη συνολική πληροφορία που περιέχουν οι πίνακες αυτοί.

ActorAppearances

aa_id	mo_id	sh_id	ac_id	aa_name	aa_timein	aa_timeout
1	1	1	10	ACTOR_APP 1	4.68	12.72
2	1	2	2	ACTOR_APP 2	12.76	16.08
3	1	2	3	ACTOR_APP 3	12.76	16.08
4	1	3	4	ACTOR_APP 4	16.16	19.32
5	1	3	5	ACTOR_APP 5	16.16	19.32
6	1	4	2	ACTOR_APP 6	19.40	23.96
7	1	4	9	ACTOR_APP 7	19.40	24.08
8	1	4	7	ACTOR_APP 8	19.40	24.08
9	1	4	8	ACTOR_APP 9	19.40	24.08
10	1	5	10	ACTOR_APP 10	24.16	27.84

Πίνακας 1: Εγγραφές της σχέσης ActorAppearances

Actors

ac_id	ac_givename	ac_familyname
2	Bill	Crystal
4	First	GrandMama
3	First	Komparsos
1	First	Speaker
6	Janthia	Friends
9	Lisa	Kudrow
10	Pastor	Speaker
5	Second	GrandMama
7	Second	Komparsos
8	Third	Komparsos

Πίνακας 2: Εγγραφές της σχέσης Actors

RelevanceFeedback

rf_id	rf_searchterm	mo_id	rf_timein	rf_timeout	rf_score
1	8c6aa27c7bda1822bf1045f2a5f2b15	1	24.16	27.84	-1
2	8c6aa27c7bda1822bf1045f2a5f2b16	1	4.68	12.72	1
3	0c082c7997a9fc66dd5ffde9a04232f0	1	16.16	19.32	1
4	c44d1f7978dd5195ebc12592e3419861	1	19.40	24.08	-1

Πίνακας 3: Εγγραφές της σχέσης RelevanceFeedback

6.2.1 Σενάριο αναζήτησης εμφάνισης ενός ηθοποιού

Το πρώτο σενάριο που θα εξετάσουμε είναι αυτό που αναλύσαμε και στην προηγούμενη ενότητα, μόνο που τώρα θα εστιάσουμε στους μηχανισμούς που οδήγησαν στην εμφάνιση των αποτελεσμάτων εκείνων. Η είσοδος της αναζήτησης είναι το όνομα ενός ηθοποιού (στην περίπτωση μας Pastor Speaker) και μια χρονική διάρκεια (3 δευτερόλεπτα). Αρχικά θα εκτελεστεί το ακόλουθο ερώτημα:

```
SELECT
    *
FROM
    ActorAppearances aa, Actors ac
WHERE
    aa.ac_id = ac.ac_id
AND
    ac.ac_givenname='Pastor'
AND
    ac.ac_familyname='Speaker'
AND
    aa.aa_timeout - aa.aa_timein > '3'
```

Το ερώτημα εκτελείται με τη χρήση ενός καρτεσιανού γινομένου υπό συνθήκες με αποτέλεσμα να επιλέγονται μόνο μερικές από τις εγγραφές του καρτεσιανού γινομένου, αυτές που έχουν σημειωθεί στον παρακάτω πίνακα. Οι συνθήκες έχουν επιλεγθεί προκειμένου να απορριφθούν αρχικά συνδυασμοί ηθοποιών και εμφανίσεων που δεν αντιστοιχούν στην πραγματικότητα (καθώς από τον υπολογισμό του καρτεσιανού γινομένου προκύπτουν τέτοιοι συνδυασμοί) και στη συνέχεια από το εναπομείναν σύνολο επιλέγονται εκείνες οι εμφανίσεις που πληρούν το κριτήριο της ελάχιστης διάρκειας εμφάνισης. Στον επόμενο πίνακα φαίνονται παραστατικά οι τελικές εγγραφές που επιλέγονται από το καρτεσιανό γινόμενο. Για πρακτικούς λόγους δεν μπορεί να αναρτηθεί ολόκληρο το αποτέλεσμα του καρτεσιανού γινομένου και παρουσιάζουμε ένα υποσύνολό του.

ActorAppearances x Actors

aa_id	mo_id	ac_id	aa_timein	aa_timeout	ac_id	ac_givename	ac_familyname
6	1	2	19.40	23.96	9	Lisa	Kudrow
7	1	9	19.40	24.08	9	Lisa	Kudrow
8	1	8	19.40	24.08	9	Lisa	Kudrow
9	1	7	19.40	24.08	9	Lisa	Kudrow
10	1	10	24.16	27.84	9	Lisa	Kudrow
1	1	10	4.68	12.72	10	First	Pastor
2	1	2	12.76	16.08	10	First	Pastor
3	1	3	12.76	16.08	10	First	Pastor
4	1	4	16.16	19.32	10	First	Pastor
5	1	5	16.16	19.32	10	First	Pastor
6	1	2	19.40	23.96	10	First	Pastor
7	1	9	19.40	24.08	10	First	Pastor
8	1	7	19.40	24.08	10	First	Pastor
9	1	8	19.40	24.08	10	First	Pastor
10	1	10	24.16	27.84	10	First	Pastor

Πίνακας 4: Αποτέλεσμα της πράξης ActorAppearances x Actors

Για κάθε εγγραφή που επιλέχθηκε τελικώς πρέπει να γίνει ένα ερώτημα στη βάση δεδομένων για να βρεθεί η αξιολόγηση της απάντησης σε σχέση πάντα με το ερώτημα που έγινε. Το πρώτο βήμα είναι να καθοριστεί το μοναδικό αναγνωριστικό του ερωτήματος αναζήτησης του χρήστη και όπως έχουμε αναφέρει αυτό προκύπτει από το είδος της αναζήτησης (αναζήτηση εμφάνισης ενός ηθοποιού, που έχουμε ορίσει να έχει μοναδικό αναγνωριστικό αναζήτησης “aa1”) και τα κριτήρια που έχει καθορίσει ο χρήστης. Στην περίπτωσή μας η είσοδος του χρήστη είναι το όνομα Pastor Speaker, ενώ η παράμετρος της χρονικής διάρκειας δεν λαμβάνεται υπόψη στην αξιολόγηση, καθώς είναι ένα μέγεθος που λαμβάνει τιμές από ένα συνεχές σύνολο τιμών και η αξιολόγηση με βάση ένα άπειρο σύνολο πιθανών τιμών δεν έχει πρακτική αξία. Η τιμή που ταυτοποιεί το ερώτημα αναζήτησης προκύπτει από την τιμή της συνάρτησης md5 στη συνένωση των παραπάνω μεγεθών, δηλαδή η τιμή αυτή είναι η md5(“aa1PastorSpeaker”) = “8c6aa27c7bda1822bf1045f2a5f2fb15”. Το ερώτημα που θα επιστρέψει την αξιολόγηση για το πρώτο από τα παραπάνω αποτελέσματα - αν η αξιολόγηση υπάρχει - είναι το εξής:

```

SELECT
    *
FROM
    RelevanceFeedback rf
WHERE
    rf_searchterm ='8c6aa27c7bda1822bf1045f2a5f2fb15'
AND
    rf_timein='4.68'
AND
    rf_timeout='12.72'
AND
    rf.mo_id='1'
LIMIT 1

```

Με βάση τα περιεχόμενα της σχέσης RelevanceFeedback, βλέπουμε ότι η αξιολόγηση που θα επιστραφεί έχει τιμή 1. Αντίστοιχα εκτελείται και το ερώτημα για το δεύτερο αποτέλεσμα, με την αξιολόγηση που επιστρέφεται να έχει τιμή -1. Αν για ένα αποτέλεσμα δεν υπάρχει αξιολόγηση αυτή θεωρείται μηδενική, οπότε έχοντας μια τιμή αξιολόγησης για κάθε αποτέλεσμα ακολουθεί η φθίνουσα ταξινόμησή τους με βάση την τιμή αυτή, ώστε να προκύψει το τελικό αποτέλεσμα που παρουσιάζεται στο χρήστη μέσω της γραφικής διεπαφής, όπως είδαμε παραπάνω.

6.2.2 Σενάριο αναζήτησης ταυτόχρονης εμφάνισης δύο ηθοποιών

Θα εξετάσουμε ένα ακόμα σενάριο αναζήτησης, όπου η τελική απάντηση δεν μπορεί να δοθεί μόνο μέσα από ένα ερώτημα. Στο σενάριο αυτό ο χρήστης αναζητάει χωρία όπου εμφανίζονται ταυτόχρονα δύο συγκεκριμένοι ηθοποιοί, οι Bill Crystal και Lisa Kudrow. Αρχικά επιλέγονται όλες οι εμφανίσεις ηθοποιών που αντιστοιχούν είτε στον έναν είτε στον άλλον ηθοποιό. Αυτό γίνεται μέσω του ακόλουθου ερωτήματος:

```

SELECT
    *
FROM
    ActorAppearances
WHERE
    ac_id IN (
        SELECT
            ac_id
        FROM
            Actors
        WHERE
            CONCAT(ac_givename, ac_familyname) IN
            ('BillCrystal', 'LisaKudrow')
    )
ORDER BY

```

aa_timein
ASC

Το συνολικό ερώτημα αποτελείται από ένα ερώτημα εμφωλευμένο σε ένα άλλο εξωτερικό, όπου το εμφωλευμένο επιλέγει τα μοναδικά αναγνωριστικά των ηθοποιών με βάση τα ονόματα που έδωσε ο χρήστης και στη συνέχεια με βάση τα αναγνωριστικά αυτά επιλέγονται οι αντίστοιχες εμφανίσεις που αφορούν τους χρήστες αυτούς. Στον επόμενο πίνακα φαίνονται τα αναγνωριστικά των ηθοποιών που επιλέγονται από το εμφωλευμένο ερώτημα.

ac_id
2
9

Πίνακας 5: Αποτελέσματα του εμφωλευμένου ερωτήματος

Το όλο ερώτημα, λοιπόν, είναι ισοδύναμο με το εξής:

```
SELECT
    *
FROM
    ActorAppearances
WHERE
    ac_id IN ('2', '9')
ORDER BY
    aa_timein
ASC
```

Ακολουθεί το αποτέλεσμα του παραπάνω τελικού ερωτήματος:

aa_id	mo_id	sh_id	ac_id	aa_name	aa_timein	aa_timeout
2	1	2	2	ACTOR_APP 2	12.76	16.08
6	1	4	2	ACTOR_APP 6	19.40	23.96
7	1	4	9	ACTOR_APP 7	19.40	24.08

Πίνακας 6: Αποτελέσματα του συνολικού ερωτήματος

Οι επιλεγθείσες εμφανίσεις θα αποτελέσουν την είσοδο στον αλγόριθμο που έχουμε εξετάσει, ο οποίος θα αποφανθεί για την ταυτόχρονη ή μη εμφάνιση των ηθοποιών, ώστε να επιστρέψει το ζητούμενο. Πρέπει να σημειωθεί ότι οι εμφανίσεις είναι ταξινομημένες σε αύξουσα σειρά χρόνου έναρξης, ενώ πρέπει να κατηγοριοποιηθούν και με βάση το οπτικοακουστικό αρχείο στο οποίο ανήκουν, που στην περίπτωσή μας όμως έχει γίνει ήδη, αφού το σύνολο εμφανίσεων που επιστράφηκε ανήκει στο ίδιο οπτικοακουστικό αρχείο. Θα παρακολουθήσουμε τώρα την εξέλιξη της εκτέλεσης του αλγορίθμου που παρουσιάστηκε στο 4^ο κεφάλαιο. Η μεταβλητή actorNumber είναι 2 γιατί ψάχνουμε για

ταυτόχρονη εμφάνιση 2 ηθοποιών και η queryResult έχει 3 αντικείμενα, αυτά που είδαμε στον προηγούμενο πίνακα.

Iteration	previousEndtime	overlaps	endResult
0	0	1	{}
1	16.08	1	{}
2	23.96	1	{}
3	23.96	2	{{mo_id:1, timein:19.40, timeout:23.96}}

Πίνακας 7: Τιμές μεταβλητών για τις επαναλήψεις του αλγορίθμου αναζήτησης

Στον παραπάνω πίνακα βλέπουμε τις τιμές των μεταβλητών για κάθε επανάληψη της εκτέλεσης, ώστε να γίνει κατανοητό πως προκύπτει η έξοδος του αλγορίθμου. Αφού βρεθεί μια ζητούμενη ταυτόχρονη εμφάνιση ο αλγόριθμος αρχικοποιεί κατάλληλα ξανά τις μεταβλητές και συνεχίζει να ψάχνει για επιπλέον ταυτόχρονες εμφανίσεις. Για το στάδιο της αναζήτησης αξιολογήσεων θα γίνει αντίστοιχο ερώτημα με την προηγούμενη περίπτωση με χρόνο έναρξης (rf_timein) 19.40 και χρόνο λήξης (rf_timeout) 23.96.

ΚΕΦΑΛΑΙΟ 7^ο

ΕΠΙΛΟΓΟΣ & ΜΕΛΛΟΝΤΙΚΗ ΕΡΓΑΣΙΑ

7.1 Επίλογος

7.2 Μελλοντική εργασία

7.1 Επίλογος

Στην παρούσα διπλωματική εργασία μελετήθηκε το πρόβλημα της δημιουργίας ευρετηρίου, αναζήτησης και ανάκτησης πολυτροπικού, οπτικοακουστικού περιεχομένου. Αντιμετωπίσαμε το πρόβλημα αναπτύσσοντας ένα κατάλληλα σχεδιασμένο σύστημα και η έρευνα για την ανάπτυξη κινήθηκε σε τρεις άξονες, οι οποίοι βρίσκονται σε αντιστοιχία με τα επιμέρους προβλήματα στα οποία διασπάσαμε το αρχικό.

Στο πρώτο στάδιο μελετήθηκαν αποδοτικοί και ευέλικτοι τρόποι για την αποθήκευση περιγραφών οπτικοακουστικού περιεχομένου σε μια βάση δεδομένων. Το στάδιο αυτό διαδραματίζει καθοριστικό ρόλο για την ανάπτυξη κατάλληλων αλγορίθμων αναζήτησης με βάση την αποθηκευμένη πληροφορία. Οι αλγόριθμοι αναζήτησης είναι το δεύτερο στάδιο της έρευνας και αξιοποιούν τη διαθέσιμη πληροφορία για να απαντήσουν στα ερωτήματα των χρηστών. Για την απάντηση στις απαιτήσεις αναζήτησης χρησιμοποιούνται κατάλληλα σχεδιασμένα ερωτήματα προς τη βάση δεδομένων και για τις περιπτώσεις όπου απαιτείται περαιτέρω επεξεργασία έχουμε αναπτύξει τους αντίστοιχους αλγόριθμους προκειμένου να δώσουν τη ζητούμενη απάντηση. Στο σύστημα της αναζήτησης ανήκει και ο μηχανισμός ανατροφοδότησης σχετικότητας που έχουμε αναπτύξει, αλλά σαν μια ξεχωριστή ψηφίδα (component) που δεν επηρεάζει τη λειτουργία της αναζήτησης. Ο μηχανισμός αυτός δέχεται σαν είσοδο τα αποτελέσματα της αναζήτησης και τα αναδιατάσσει με βάση τη σχετικότητα. Το μέτρο της σχετικότητας κάθε αποτελέσματος προκύπτει από τις αξιολογήσεις των χρηστών, βελτιώνοντας έτσι τη λειτουργία όλου του συστήματος αναζήτησης. Το τελευταίο στάδιο της έρευνάς μας είναι ανεξάρτητο από τα δύο προηγούμενα και αφορά τη σχεδίαση μιας γραφικής διεπαφής χρήστη (Graphical User Interface), που αποσκοπεί στην οπτική παρουσίαση της υπάρχουσας πληροφορίας και την αλληλεπίδραση με το χρήστη. Αποτελεί το σημείο αλληλεπίδρασης του χρήστη με το μηχανισμό αναζήτησης και το σύνολο της αποθηκευμένης πληροφορίας.

Η έρευνα που πραγματοποιήθηκε αποτελεί ένα μικρό μέρος ενός ευρύτερου, επιστημονικά δραστήριου τομέα, ο οποίος αφορά τις πολυμεσικές υπηρεσίες. Παρουσιάστηκαν κάποιες βασικές δυνατότητες αναζήτησης οπτικοακουστικού περιεχομένου, οι οποίες είναι εγγενείς στο περιεχόμενο και αποτελούν ένα μικρό μόνο δείγμα των δυνατοτήτων που θα μπορούσαν να έχουν οι χρήστες μελλοντικά στην αναζήτηση τέτοιου περιεχομένου.

7.2 Μελλοντική εργασία

Όπως αναφέραμε, το σύστημά μας αποτελεί ένα μικρό μόνο τμήμα ενός ευρύτερου τομέα με σημαντικές μελλοντικές προοπτικές, επομένως υπάρχει πληθώρα επεκτάσεων και τροποποιήσεων που μπορούν να βρουν εφαρμογή σε αυτό που υλοποιήσαμε. Μπορούμε να κατηγοριοποιήσουμε τη μελλοντική εργασία σε τρία στάδια, τα οποία αντιστοιχούν σε αυτά που έχουμε χωρίσει τη δική μας έρευνα.

Στον τομέα της βάσης δεδομένων υπάρχει ένα πολύ σημαντικό στοιχείο που αξίζει λεπτομερούς έρευνας και ανάλυσης και είναι η υιοθέτηση «Καθαρά» XML Συστημάτων Διαχείρισης Βάσεων Δεδομένων (Native XML DataBase Management Systems). Τα συστήματα αυτά προσφέρουν ένα τελείως διαφορετικό μοντέλο αποθήκευσης της πληροφορίας από τα Σχεσιακά Συστήματα Διαχείρισης Βάσεων Δεδομένων (Relational DataBase Management Systems) και βασίζονται στην έννοια της συλλογής XML εγγράφων. Για κάθε οπτικοακουστικό αρχείο έχουμε και μια αντίστοιχη περιγραφή σε MPEG-7 XML και η βάση πληροφορίας μας αποτελείται από πολλά τέτοια οπτικοακουστικά αρχεία και αρχεία περιγραφών, οπότε το μοντέλο που μας προσφέρει ένα τέτοιο Σύστημα Διαχείρισης Βάσεων Δεδομένων φαντάζει ταιριαστό. Η αναζήτηση μπορεί να πραγματοποιηθεί με μεθόδους εγγενείς στο XML περιεχόμενο στοιχείο επίσης επιθυμητό. Παρ' όλα αυτά μια τέτοια απόφαση για μετάβαση σε μια «καθαρή» XML βάση δεδομένων δεν είναι μονόδρομος, αφού υπάρχουν αρκετά σημεία που χρίζουν προσοχής. Τα συστήματα αυτά δεν προσφέρουν την αξιοπιστία των σχεσιακών συστημάτων και η εμπειρία και τεχνογνωσία που υπάρχει πάνω σε αυτά υπολείπεται από εκείνη των σχεσιακών, τα οποία βρίσκονται στο προσκήνιο πολλές δεκαετίες. Επίσης, οι διαθέσιμες επιλογές για τέτοια συστήματα είναι ιδιαίτερα περιορισμένες. Θα μπορούσε ακόμη να εξεταστεί η χρήση ενός υβριδικού συστήματος, δηλαδή η χρήση Native XML Συστήματος Διαχείρισης Βάσεων Δεδομένων μαζί με ένα σχεσιακό σύστημα με σκοπό να αξιοποιηθούν τα θετικά στοιχεία του καθενός, εξαλείφοντας τα αρνητικά και τις πιθανές τεχνικές δυσκολίες. Τέλος, στον κλάδο αυτό θα είχε ενδιαφέρον η σχεδίαση μιας γενικής αρχιτεκτονικής για μια πολυμεσική βάση δεδομένων. Έχουν γίνει δειλά βήματα προς αυτή την κατεύθυνση και σίγουρα υπάρχουν πολλές δυνατότητες εξέλιξης. Ο απώτερος στόχος πρέπει να είναι η ανάπτυξη μιας αρχιτεκτονικής για τη βάση, η οποία να είναι σε θέση να φιλοξενήσει περιγραφές MPEG-7 με τη μέγιστη δυνατή γενικότητα, χωρίς να είναι προσανατολισμένη σε συγκεκριμένες εφαρμογές ή επεκτάσεις του MPEG-7.

Και στο δεύτερο στάδιο της έρευνάς μας υπάρχει πληθώρα πιθανών μελλοντικών επεκτάσεων που θα αυξήσουν τη λειτουργικότητα του συστήματός μας. Καταρχήν μπορούν να εμπλουτιστούν οι δυνατότητες αναζήτησης, προσφέροντας νέα ερωτήματα στο χρήστη. Μερικά από τα πιο χρήσιμα είναι η αναζήτηση με βάση δραστηριότητες ηθοποιών (π.χ. ηθοποιοί που τρέχουν ή ηθοποιοί που κινούνται προς τα δεξιά), αναζήτηση με βάση εκφράσεις του προσώπου των ηθοποιών, αλλά και δυνατότητες αναζήτησης περιεχομένου

με βάση χαμηλού επιπέδου χαρακτηριστικά του περιεχομένου. Επίσης, μπορούν να προστεθούν δυνατότητες αναζήτησης με βάση υψηλού και χαμηλού επιπέδου χαρακτηριστικά της ηχητικής πληροφορίας (π.χ. σκηνές που ακούγεται κάποιος πυροβολισμός). Μεγάλο ερευνητικό ενδιαφέρον παρουσιάζει η εξέλιξη του μηχανισμού ανατροφοδότησης σχετικότητας (Relevance Feedback), καθώς έχουμε υλοποιήσει τις βασικές μόνο λειτουργίες. Το κυριότερο στοιχείο που πρέπει να προστεθεί στο μηχανισμό είναι η εξατομίκευση. Αυτό προϋποθέτει την ύπαρξη διαφορετικών χρηστών, οπότε πρέπει να τροποποιηθεί κατάλληλα η βάση δεδομένων για την υποστήριξη περισσότερων χρηστών, με τις προσωπικές πληροφορίες και ρυθμίσεις του καθενός. Στη συνέχεια ο μηχανισμός πρέπει να επεκταθεί προκειμένου να λαμβάνει υπόψη τα ιδιαίτερα χαρακτηριστικά και τις προτιμήσεις του κάθε χρήστη. Πιο συγκεκριμένα θα είχε ενδιαφέρον η μελέτη ενός μηχανισμού με τη χρήση βαρών, όπου κάθε παράμετρος που επηρεάζει το τελικό αποτέλεσμα να έχει ένα βάρος ανάλογο με τη σημασία που έχει στον καθορισμό της σχετικότητας των αποτελεσμάτων. Τέτοιοι παράγοντες είναι η ηλικία του χρήστη, το είδος (genre) των ταινιών που του αρέσουν, οι αγαπημένοι ηθοποιοί, παραγωγοί και σκηνοθέτες, αλλά και πιο δυναμικοί παράγοντες, όπως τα χαρακτηριστικά (είδος, δημοτικότητα, διάρκεια) των οπτικοακουστικών αρχείων που παρακολουθεί ο χρήστης, η συχνότητα χρήσης του μηχανισμού ανατροφοδότησης σχετικότητας και η εμπειρία χρήσης του 3DTV συστήματος.

Το τελευταίο στάδιο της έρευνας ήταν η ανάπτυξη μιας γραφικής διεπαφής χρήστη (Graphical User Interface). Σε αυτό το τμήμα προφανώς πρέπει να προστεθούν δυνατότητες αντίστοιχες με αυτές της αναζήτησης που θα υλοποιηθούν στο προηγούμενο στάδιο, ωστόσο αυτό δεν αποτελεί κάποια σημαντική μελλοντική πρόκληση. Αυτό που έχει μεγαλύτερο ενδιαφέρον είναι οι δυνατότητες αναπαραγωγής τρισδιάστατου, πολυτροπικού οπτικοακουστικού περιεχομένου μέσα από την HTML5, όπου έχουν γίνει ήδη αξιολογικές προσπάθειες και αναμένονται εντυπωσιακές δυνατότητες μελλοντικά. Το ενδιαφέρον επικεντρώνεται στην αξιοποίηση των προγραμματιστικών δυνατοτήτων που θα δίνονται στον προγραμματιστή, όπως η πρόσβαση στα μεμονωμένα κανάλια οπτικής πληροφορίας. Ιδιαίτερο ενδιαφέρον έχει και η προσαρμογή της οπτικής σημείωσης για χρήση σε τρισδιάστατο οπτικοακουστικό υλικό, ώστε να δίνεται η αίσθηση του βάθους. Τέλος, ένα ακόμη σημείο ενδιαφέροντος είναι η επέκταση του μηχανισμού της οπτικής σημείωσης σε συμβάντα που συμβαίνουν μεταξύ ενός ηθοποιού και κάποιου άλλου εικονιζόμενου προσώπου ή αντικειμένου, έτσι ώστε να αποδίδεται με παραστατικό τρόπο η ενέργεια που γίνεται και τα πρόσωπα και αντικείμενα που συμμετέχουν.

ΒΙΒΛΙΟΓΡΑΦΙΑ

- [1] C.Hellberg, D.Greene, T.Boyes, "Next Generation Triple-Play Services," in *Broadband Network Architectures: Designing and Deploying Triple-Play Services*, pp. 35-36, Prentice Hall, May 2007.
- [2] C.Hellberg, D.Greene, T.Boyes, "Video on Demand," in *Broadband Network Architectures: Designing and Deploying Triple-Play Services*, pp. 45-51, Prentice Hall, May 2007.
- [3] C. Fuchs, W. Hofkirchner, M. Schafranek et. al, "*Theoretical Foundations of the Web: Cognition, Communication and Co-operation. Towards an Understanding of Web 1.0, 2.0, 3.0*", 2010.
- [4] T. O'Reilly, "*What is Web 2.0: Design Patterns and Business Models for the Next Generation of Software*", *Communications and Strategies*, no. 65, March 2007.
- [5] Wikipedia, *Avatar (2009 film)* [Online]. Available: [http://en.wikipedia.org/wiki/Avatar_\(2009_film\)](http://en.wikipedia.org/wiki/Avatar_(2009_film)).
- [6] B. Mendiburu, "Stereoscopic Vision and 3D Cinematography," in *3D Movie Making: Stereoscopic Digital Cinema from Script to Screen*, pp. 11-34, Focal Press, 2009.
- [7] E. Onasoglou, P. Daras, "Semantic force relevance feedback, content-free 3D object retrieval and annotation propagation: bridging the gap and beyond," *Multimedia Tools Appl.*, 39(2), pp. 217-241, Thessaloniki 2008.
- [8] B.S. Manjunath, P. Salembier and T. Sikora, "Context, Goals and Procedures," in *Introduction to MPEG-7: Multimedia Content Description Interface*, pp. 7-29, Wiley, England 2002.
- [9] S. Chang, T. Sikora and A.Puri, "Overview of the MPEG-7 Standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 11, no. 6, June 2001.
- [10] E. T. Ray, "What is XML?," in *Learning XML*, pp. 2-11, O' Reilly Media, USA 2003.
- [11] E. T. Ray, "W3C XML Schema," in *Learning XML*, pp. 132-134, O' Reilly Media, USA 2003.
- [12] H. Chu, "Dublin Core (DC)," in *Information Representation and Retrieval in the Digital Age*, pp. 38-39, American Society for Information Science & Technology, 2007.
- [13] A. Lugmayr, S. Niiranen, S. Kalli, "SMPTE Metadata Definitions," in *Digital Interactive TV and Metadata: Future Broadcast Multimedia*, pp. 69-71, Springer, 2004.
- [14] A. Lugmayr, S. Niiranen, S. Kalli, "TV-Anytime," in *Digital Interactive TV and Metadata: Future Broadcast Multimedia*, pp. 63-69, Springer, 2004.
- [15] S. Chang, T. Sikora and A.Puri, "Overview of the MPEG-7 Standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 11, no. 6, June 2001.
- [16] "MPEG-7 White Paper", Sonera MediaLab, Finland, Oct. 13 2003.
- [17] J.M. Martvnez, *MPEG-7 Overview (10th ed.)* [Online]. Available: <http://mpeg.chiariglione.org/standards/mpeg-7/mpeg-7.htm>, Oct. 2004.
- [18] B.S. Manjunath, P. Salembier and T. Sikora, "Overview of Multimedia Description Schemes and Schema Tools," in *Introduction to MPEG-7: Multimedia Content Description Interface*, pp. 83-88, Wiley, England 2002.

- [19] J. L. Harrington, "Database Software: DBMSs," in *Relational Database Design and Implementation: Clearly Explained*, pp. 8-10, Morgan Kaufmann, 2009.
- [20] R. Bourret, *XML and Databases* [Online]. Available: <http://www.rpbouret.com/xml/XMLAndDatabases.htm>, Sep. 2005.
- [21] A. Silberschatz, H. F. Korth and S. Sudarshan, "SQL," in *Database System Concepts*, pp. 141-178, McGraw Hill, July 2001.
- [22] M. Kay, "What Is XPath?" in *XPath 2.0 Programmer's Reference*, pp. 1-15, Wiley, Canada 2004.
- [23] P. A. Bernstein, E. Newcomer, "Atomicity, Consistency, Isolation and Durability," in *Principles of Transaction Processing*, pp. 9-15, Morgan Kaufmann, 2009.
- [24] H. H. Liu, "Scalability Testing," in *Software Performance and Scalability: A Quantitative Approach*, pp. 75-82, Wiley, 2009.
- [25] P. Domanski, D. Irvine, "Creating Indexes," in *A Practical Guide to Relational Database Design*, pp. 117-120, Diaxon, 2000.
- [26] J. L. Harrington, "Unnecessary Duplicated Data and Data Consistency," in *Relational Database Design and Implementation: Clearly Explained*, pp 47-48, Morgan Kaufmann, 2009.
- [27] R. Stephens, R. Plew, A. D. Jones, "Denormalizing a Database," in *Teach Yourself SQL in 24 Hours*, pp. 69, Sams, 2008.
- [28] A. Silberschatz, H. F. Korth and S. Sudarshan, "The Natural-Join Operation," in *Database System Concepts*, pp. 107-109, McGraw Hill, July 2001.
- [29] A. Silberschatz, H. F. Korth and S. Sudarshan, "The Select Clause," in *Database System Concepts*, pp. 144-147, McGraw Hill, July 2001.
- [30] D. Hunter, J. Rafter, J. Fawcett et. al, "Purpose of the XML DOM," in *Beginning XML*, pp. 443-446, Wiley, 2007.
- [31] V. Vaswani, "PHP and SAX," in *XML and PHP*, pp. 21-34, New Riders, 2002.
- [32] V. Vaswani, "History and Evolution," in *MySQL: The Complete Reference*, pp. 6-12, McGraw Hill, 2004.
- [33] T. H. Cormen, C. E. Leiserson, R. L. Rivest et. al, "Greedy Algorithms," in *Introduction to Algorithms*, pp. 371-383, MIT Press, 1991.
- [34] S. Friedl, *An Illustrated Guide to Cryptographic Hashes* [Online]. Available: <http://unixwiz.net/techtips/iguide-crypto-hashes.html>
- [35] P. Yuen, "The 128-bit MD5 algorithm," in *Practical Cryptography and Web Security*, pp. 178-182, Addison Wesley, 2006.
- [36] D. Sklar, "Orientation and First Steps," in *Learning PHP 5*, pp. 1-6, O'Reilly Media, USA 2004.
- [37] M. Zandstra, "PHP and Objects," in *PHP Objects, Patterns, and Practice*, pp. 11-14, Apress, 2010.
- [38] M. MacDonald, "The Story of HTML5," in *HTML5: The Missing Manual*, pp. 13-18, O'Reilly Media, USA 2011.
- [39] M. Pilgrim, "Detecting HTML5 Features," in *HTML5: Up and Running*, pp. 18-21, O'Reilly Media, USA 2010.
- [40] D. Crockford, *Javascript: The World's Most Misunderstood Programming Language* [Online]. Available: <http://www.crockford.com/javascript/javascript.html>, 2001.
- [41] Mozilla Developer Network, *CSS* [Online]. Available: <https://developer.mozilla.org/en/CSS/>
- [42] D. Crockford, *RFC4627* [Online]. Available: <http://tools.ietf.org/html/rfc4627>, July 2006.

- [43] P. McIntire, "Review of Web Technology," in *Visual Design for the Modern Web*, pp. 4-17, New Riders, 2008.