



**ΤΕΧΝΟΛΟΓΙΚΟ ΕΚΠΑΙΔΕΥΤΙΚΟ
ΙΔΡΥΜΑ ΠΑΤΡΑΣ**

**ΣΧΟΛΗ ΔΙΟΙΚΗΣΗΣ & ΟΙΚΟΝΟΜΙΑΣ
ΤΜΗΜΑ ΔΙΟΙΚΗΣΗΣ ΕΠΙΧΕΙΡΗΣΕΩΝ**

ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

**ΘΕΜΑ: ΜΕΛΕΤΗ ΚΑΙ ΣΥΓΚΡΙΣΗ ΣΥΣΤΗΜΑΤΩΝ ΓΙΑ
ΑΝΑΚΤΗΣΗ ΕΙΚΟΝΩΝ ΜΕ ΒΑΣΗ ΤΟ ΠΕΡΙΕΧΟΜΕΝΟ**

Εισηγήτρια» ΦΩΚΑ ΑΜΑΛΙΑ

Ονόματα Σπουδαστών»

**ΑΛΕΤΡΑ ΝΙΚΗ
ΑΡΑΝΚΙ ΙΕΗΑΔ
ΜΟΥΡΤΖΙΝΗΣ ΓΙΩΡΓΟΣ**

ΠΑΤΡΑ 2011

ΠΕΡΙΕΧΟΜΕΝΑ

ΕΥΡΕΤΗΡΙΟ ΕΙΚΟΝΩΝ.....	- 5 -
ΚΕΦΑΛΑΙΟ 1: ΕΙΣΑΓΩΓΗ.....	- 7 -
1.1 ΑΝΤΙΚΕΙΜΕΝΟ ΕΡΓΑΣΙΑΣ.....	- 7 -
1.2 ΔΙΑΦΘΩΡΩΣΗ ΕΡΓΑΣΙΑΣ.....	- 7 -
1.3 ΕΥΧΑΡΙΣΤΙΕΣ.....	- 8 -
ΚΕΦΑΛΑΙΟ 2: ΑΝΑΚΤΗΣΗ ΕΙΚΟΝΩΝ.....	- 9 -
2.1 ΕΙΣΑΓΩΓΗ – ΙΣΤΟΡΙΚΗ ΑΝΑΔΡΟΜΗ.....	- 9 -
2.2 ΝΕΕΣ ΕΞΕΛΙΞΕΙΣ ΣΤΟΝ ΕΙΚΟΣΤΟ ΠΡΩΤΟ ΑΙΩΝΑ	- 10 -
2.3 ΑΝΑΚΤΗΣΗ ΕΙΚΟΝΑΣ ΜΕ ΒΑΣΗ ΤΟ ΚΕΙΜΕΝΟ.....	- 11 -
2.3.1 Συστήματα ευρετηρίασης και ταξινόμησης εικόνων.....	- 11 -
2.3.2 Πλεονεκτήματα - μειονεκτήματα ανάκτησης εικόνας με βάση το κείμενο - ανάγκη για ανάκτηση με βάση το περιεχόμενο	- 13 -
2.4 ΑΝΑΚΤΗΣΗ ΕΙΚΟΝΑΣ ΜΕ ΒΑΣΗ ΤΟ ΠΕΡΙΕΧΟΜΕΝΟ (CBIR)	- 14 -
2.4.1 Περιγραφή CBIR συστήματος.....	- 14 -
2.4.2 Εξαγωγή χαρακτηριστικών.....	- 16 -
2.5 ΑΝΑΚΤΗΣΗ ΕΙΚΟΝΑΣ ΒΑΣΗ ΟΠΤΙΚΗΣ ΠΛΗΡΟΦΟΡΙΑΣ (VBIR)	- 19 -
ΚΕΦΑΛΑΙΟ 3: ΑΝΑΚΤΗΣΗ ΕΙΚΟΝΩΝ ΜΕ ΒΑΣΗ ΤΟ ΠΕΡΙΕΧΟΜΕΝΟ	- 20 -
3.1 ΕΙΣΑΓΩΓΗ.....	- 20 -
3.2 ΧΡΩΜΑ	- 21 -
3.3 ΥΦΗ.....	- 25 -
3.4 ΣΧΗΜΑ	- 25 -
3.5 ΘΕΣΗ	- 26 -
3.6 ΤΕΧΝΙΚΕΣ ΣΥΓΚΡΙΣΗΣ ΙΣΤΟΓΡΑΜΜΑΤΩΝ.....	- 27 -
3.7 ΤΕΧΝΙΚΕΣ ΤΑΙΡΙΑΣΜΑΤΟΣ ΕΙΚΟΝΩΝ	- 30 -
3.8 ΤΕΧΝΙΚΕΣ ΔΕΙΚΤΟΔΟΤΗΣΗΣ	- 31 -
3.9 ΤΕΧΝΙΚΕΣ ΜΕ ΧΡΗΣΗ ΟΠΤΙΚΟΥ ΛΕΞΙΚΟΥ.....	- 32 -
ΚΕΦΑΛΑΙΟ 4: ΤΟ ΠΡΟΤΥΠΟ MPEG7.....	- 33 -
4.1 ΕΙΣΑΓΩΓΗ.....	- 33 -
4.1.1 MPEG-1.....	- 33 -
4.1.2 MPEG-2.....	- 34 -
4.1.3 MPEG-4.....	- 34 -
4.1.4 MPEG-7.....	- 35 -
4.2 MPEG7 ΚΑΙ ΒΑΣΕΙΣ ΔΕΔΟΜΕΝΩΝ	- 38 -
4.3 ΠΕΡΙΓΡΑΦΗ ΤΟΥ MPEG7	- 41 -
4.4 ΠΕΡΙΓΡΑΦΕΙΣ	- 42 -
ΚΕΦΑΛΑΙΟ 5: ΣΥΣΤΗΜΑΤΑ ΑΝΑΚΤΗΣΗΣ ΕΙΚΟΝΑΣ ΜΕ ΒΑΣΗ ΤΟ ΠΕΡΙΕΧΟΜΕΝΟ.....	- 48 -

5.1 ADL.....	- 48 -
5.2 ALISA.....	- 49 -
5.3 AMORE.....	- 50 -
5.4 BLOBWORLD.....	- 51 -
5.5 CANDID.....	- 52 -
5.6 C-BIRD.....	- 53 -
5.7 CBVQ.....	- 54 -
5.8 CHROMA.....	- 55 -
5.9 COMPASS.....	- 56 -
5.10 DRAWSEARCH.....	- 57 -
5.11 EXCALIBUR VISUAL RETRIEVALWARE.....	- 58 -
5.12 FIDS.....	- 59 -
5.13 FIR.....	- 60 -
5.14 FIRE.....	- 60 -
5.15 FOCUS.....	- 61 -
5.16 IMAGEFINDER.....	- 62 -
5.17 IMAGERETRO.....	- 63 -
5.18 IMAGEROVER.....	- 64 -
5.19 IMAGESCAPE.....	- 65 -
5.20 IPURE.....	- 66 -
5.21 KIWI.....	- 67 -
5.22 LCPD.....	- 67 -
5.23 MARS.....	- 68 -
5.24 METASEEK.....	- 70 -
5.25 MIR.....	- 71 -
5.26 NETRA.....	- 72 -
5.27 PHOTOBOOK.....	- 73 -
5.28 PICSOM.....	- 74 -
5.29 PICTOSEEK.....	- 75 -
5.30 QBIC.....	- 76 -
5.31 QUICKLOOK.....	- 77 -
5.32 SIMBA.....	- 78 -
5.33 SMURF.....	- 79 -
5.34 SIMPLICITY.....	- 80 -
5.35 SQUID.....	- 81 -
5.36 SURFIMAGE.....	- 82 -
5.37 SYNAPSE.....	- 83 -
5.38 TODAI.....	- 84 -
5.39 VISUALSEEK.....	- 85 -
5.40 VP IMAGE RETRIEVAL SYSTEM.....	- 87 -
5.41 WEBSEEK.....	- 88 -
5.42 WISE.....	- 89 -
ΚΕΦΑΛΑΙΟ 6: ΣΥΜΠΕΡΑΣΜΑΤΑ.....	- 91 -

Ευρετήριο εικόνων

<i>Εικόνα 1: Γενική δομή ενός CBIR συστήματος.....</i>	<i>- 16 -</i>
<i>Εικόνα 2: Μοντέλο RGB.....</i>	<i>- 23 -</i>
<i>Εικόνα 3: Το μοντέλο CMYK.....</i>	<i>- 24 -</i>
<i>Εικόνα 4: Παρεμβολή δύο ιστογραμμάτων A και B</i>	<i>- 29 -</i>
<i>Εικόνα 5: Το σύστημα ADL</i>	<i>- 49 -</i>
<i>Εικόνα 6: Το σύστημα ALISA.....</i>	<i>- 50 -</i>
<i>Εικόνα 7: Το σύστημα Amore</i>	<i>- 51 -</i>
<i>Εικόνα 8: Αποτελέσματα ερωτήσεων με το Blobworld</i>	<i>- 52 -</i>
<i>Εικόνα 9: Το σύστημα C-bird</i>	<i>- 54 -</i>
<i>Εικόνα 10: Το σύστημα CBVQ.....</i>	<i>- 55 -</i>
<i>Εικόνα 11: Το σύστημα CHROMA</i>	<i>- 56 -</i>
<i>Εικόνα 12: Το σύστημα COMPASS</i>	<i>- 57 -</i>
<i>Εικόνα 13: Το σύστημα DrawSearch.....</i>	<i>- 58 -</i>
<i>Εικόνα 14: Το σύστημα Excalibur Visual RetrievalWare.....</i>	<i>- 59 -</i>
<i>Εικόνα 15: Το σύστημα FIDS</i>	<i>- 60 -</i>
<i>Εικόνα 16: Το σύστημα FIRE</i>	<i>- 61 -</i>
<i>Εικόνα 17: Το σύστημα FOCUS.....</i>	<i>- 62 -</i>
<i>Εικόνα 18: Το σύστημα ImageFinder.....</i>	<i>- 63 -</i>
<i>Εικόνα 19: Το σύστημα ImageRETRO.....</i>	<i>- 64 -</i>
<i>Εικόνα 20: Το σύστημα ImageRover</i>	<i>- 65 -</i>
<i>Εικόνα 21: Το σύστημα ImageScape</i>	<i>- 66 -</i>
<i>Εικόνα 22: Το σύστημα KIWI.....</i>	<i>- 67 -</i>
<i>Εικόνα 23: Το σύστημα LCPD.....</i>	<i>- 68 -</i>
<i>Εικόνα 24: Το σύστημα MARS.....</i>	<i>- 70 -</i>

<i>Εικόνα 25: Το σύστημα METASEEK.....</i>	<i>- 71 -</i>
<i>Εικόνα 26: Το σύστημα MIR.....</i>	<i>- 72 -</i>
<i>Εικόνα 27: Το σύστημα NETRA.....</i>	<i>- 73 -</i>
<i>Εικόνα 28: Το σύστημα Photobook.....</i>	<i>- 74 -</i>
<i>Εικόνα 29: Το σύστημα PicSOM</i>	<i>- 75 -</i>
<i>Εικόνα 30: Το σύστημα PicToSeek.....</i>	<i>- 76 -</i>
<i>Εικόνα 31: Το σύστημα QBIC.....</i>	<i>- 77 -</i>
<i>Εικόνα 32: Το σύστημα QuickLook</i>	<i>- 78 -</i>
<i>Εικόνα 33: Το σύστημα SIMBA</i>	<i>- 79 -</i>
<i>Εικόνα 34: Το σύστημα SMURF.....</i>	<i>- 80 -</i>
<i>Εικόνα 35: SIMPLIcity.....</i>	<i>- 81 -</i>
<i>Εικόνα 36: Το σύστημα SQUID.....</i>	<i>- 82 -</i>
<i>Εικόνα 37: Το σύστημα SurfImage</i>	<i>- 83 -</i>
<i>Εικόνα 38: Το σύστημα Synapse.....</i>	<i>- 84 -</i>
<i>Εικόνα 39: Το σύστημα Todai.....</i>	<i>- 85 -</i>
<i>Εικόνα 40: Το σύστημα VISUALSEEK.....</i>	<i>- 86 -</i>
<i>Εικόνα 41: Το σύστημα VP Image Retrieval System</i>	<i>- 88 -</i>
<i>Εικόνα 42: Το σύστημα WEBSEEK</i>	<i>- 89 -</i>
<i>Εικόνα 43: Το σύστημα WISE.....</i>	<i>- 90 -</i>

ΚΕΦΑΛΑΙΟ 1: Εισαγωγή

1.1 Αντικείμενο εργασίας

Η τεχνολογική ανάπτυξη που παρατηρείται τα τελευταία χρόνια έχει οδηγήσει στην δημιουργία, αποθήκευση και διακίνηση τεράστιου όγκου δεδομένων. Η επανάσταση του διαδικτύου έχει οξύνει το πρόβλημα της πρόσβασης και διαθεσιμότητας όλης αυτής της πληροφορίας. Οξύτερο είναι το πρόβλημα σήμερα της λεγόμενης πολυμεσικής πληροφορίας. Με τον όρο πολυμεσική πληροφορία αναφερόμαστε σε οποιοδήποτε πόρο περιλαμβάνει εικόνα, ήχο και κινούμενη εικόνα ή συνδυασμό όλων αυτών.

Ένα σημαντικό πρόβλημα που χρήζει αντιμετώπισης στην εποχή μας σχετικά με την πολυμεσική πληροφορία είναι σίγουρα η ανάκτηση της πολυμεσικής πληροφορίας από συστήματα βάσεων δεδομένων. Οι κατηγοριοποιήσεις που χρησιμοποιήθηκαν παλιότερα με τη χρήση απλών συστημάτων βάσεων δεδομένων ή απλά με την οπτική ή την αναγνώριση με βάση το κείμενο που χρησιμοποιούνταν έχουν πλέον απαρχαιωθεί και δεν προσφέρουν κανενός είδους διαχείρισης κι εύρεσης της κατάλληλης πληροφορίας λόγω του τεράστιου όγκου που διαχειρίζονται.

Στην παρούσα εργασία γίνεται μια σε βάθος ανάλυση των χαρακτηριστικών που απαρτίζουν μια εικόνα καθώς επίσης και πως μπορεί αυτή η πληροφορία να γίνει διαχειρίσιμη. Παράλληλα γίνεται μια παρουσίαση των συστημάτων ανάκτησης δεδομένων που υπάρχουν σήμερα καθώς και του προτύπου MPEG7.

1.2 Διάρθρωση εργασίας

Στο 2^ο κεφάλαιο γίνεται μία σύντομη αναδρομή στην ανάκτηση εικόνων και παρουσιάζονται οι διάφορες τεχνικές ανάκτησης εικόνων, όπως η ανάκτηση εικόνων με βάση το κείμενο, το περιεχόμενο, την οπτική πληροφορία και βιολογικά εμπνευσμένες μεθόδους. Στο 3^ο κεφάλαιο γίνεται μία εκτενής επισκόπηση της τεχνική

ανάκτησης εικόνων με βάση το περιεχόμενο. Στο 4^ο κεφάλαιο γίνεται μια σε βάθος ανάλυση του προτύπου MPEG7.

Στο 5^ο κεφάλαιο παρουσιάζονται μερικά από τα πιο ευρέως διαδεδομένα συστήματα ανάκτησης εικόνων με βάση το περιεχόμενο. Τέλος, στο 6^ο κεφάλαιο παρέχονται τα συμπεράσματα της παρούσας εργασίας.

1.3 Ευχαριστίες

Με την ολοκλήρωση της πτυχιακής εργασίας μας θα θέλαμε να ευχαριστήσουμε θερμά την καθηγήτρια μας κα. Φωκά Αμαλία για την βοήθεια και την συνεργασία που μας πρόσφερε.

Τα εύσημα πρέπει να αποδοθούν και στους υπόλοιπους καθηγητές και στην γραμματεία του τμήματος μας για την βοήθεια τους προς εμάς, των απαραίτητων γνώσεων κατά την διάρκεια της φοίτησης μας στο Α.Τ.Ε.Ι Πατρών καθώς και όσων συνέβαλαν, μεριμνούν και θα συνεχίζουν να συνδράμουν στις επιστήμες και στην έννοια της ηθικής στην θεωρία και στην πράξη.

ΚΕΦΑΛΑΙΟ 2: Ανάκτηση εικόνων

Στο κεφάλαιο αυτό θα παρουσιαστούν οι πιο ευρέως διαδεδομένες τεχνικές που χρησιμοποιούνται στην ανάκτηση εικόνας.

2.1 Εισαγωγή – Ιστορική Αναδρομή

Η έρευνα για ψηφιακή γνώση ξεκίνησε πριν από πολλές δεκαετίες όταν υπήρξε η ιδέα για ψηφιοποίηση σε όλους τους τομείς της επιστήμης για αρχή και αργότερα και σε όλους τους τομείς της κοινωνίας. Πριν από αυτήν την εποχή η ανάκτηση εικόνας υπήρχε μόνο σε θεωρητικές συζητήσεις μεταξύ των επιστημόνων. Από τη θεωρητική πλευρά, περιοχές όπως η τεχνητή νοημοσύνη, θεωρίες βελτιστοποίησης, οπτικοποίηση των υπολογιστών και αναγνώριση υποδειγμάτων βοήθησαν σημαντικά στην δημιουργία του μαθηματικού υποβάθρου στον τομέα της ανάκτησης εικόνας. Η ψυχολογία και σχετικές με αυτήν περιοχές όπως η αισθητική και η εργονομία παρέχουν τις βασικές αρχές για την αλληλεπίδραση μεταξύ των χρηστών. Επιπρόσθετα, εφαρμογές για αναζήτηση εικόνων σε μια βάση με εικόνες σε χαρτί προϋπήρχαν σε απλοϊκές μορφές για την αναγνώριση προσώπων και χαρακτήρων γραφής.

Στα πρώιμα χρόνια της Ανάκτησης Εικόνας έχουν γραφεί διάφορα βιβλία για την επιτυχία τέτοιων συστημάτων με την χρήση ηλεκτρονικού υπολογιστή. Οι εργασίες των Ballard και Brown (1982), Levine (1985), και Haralick et al (1993), εστιάζουν στους αλγόριθμους που χρησιμοποιούνται για την ανάκτηση εικόνων με βάση τις ομοιότητες μεταξύ εικόνων.

Υλοποιήσεις των παραπάνω αλγορίθμων μπορούν να θεωρηθούν συστήματα όπως το QBIC, που προτάθηκε από τους Flickner et al. (1995) και το Virage, που προτάθηκε από τους Bach, et al. (1996) τα οποία δημιουργήθηκαν στα μέσα τις δεκαετίες του 1990. Μέσα σε λίγα χρόνια η βασική ιδέα της ομοιότητας μεταξύ εικόνων μεταφέρθηκε σε διάφορους δικτυακούς τόπους μηχανών αναζήτησης εικόνων. Έτσι έχουμε τέτοιες υλοποιήσεις όπως το Webseek, που δημιουργήθηκε από τους Smith et al (1997) και το Webseer, που δημιουργήθηκε από τους Frankel, et al. (1996). Μεγάλη προσπάθεια έγινε επίσης και στον τομέα των επιχειρήσεων όπου η αναζήτηση με βάση τις ομοιότητες εικόνων ήταν η βάση για την δημιουργία των βάσεων δεδομένων της

Informix, μία επέκταση της ευρέως γνωστής βάσης δεδομένων *IBM DB2* καθώς επίσης και του συστήματος διαχείρισης βάσεων δεδομένων *Oracle*, που δημιουργήθηκε από τους *Bliujute, et al. (1999)* και *Egas, et al. (1999)* έτσι ώστε να μπορεί να γίνει ευκολότερη η χρήση των εικόνων στον τομέα των επιχειρήσεων.

2.2 Νέες εξελίξεις στον εικοστό πρώτο αιώνα

Ξεκινώντας από τις αρχές του εικοστού πρώτου αιώνα οι ερευνητές παρατήρησαν ότι η χρήση αλγορίθμων εύρεσης ομοιοτήτων είχε περισσότερο χαρακτήρα διαίσθησης παρά ήταν φιλική προς τον χρήστη όπως περίμεναν. Αυτά τα συστήματα τα οποία είχαν δημιουργηθεί από επιστήμονες μέσω ερευνητικών προγραμμάτων μπορούσαν να χρησιμοποιηθούν αποτελεσματικά μόνο από τους ίδιους τους δημιουργούς τους. Η καινούργια φιλοσοφία ήταν να δημιουργηθούν συστήματα φιλικά προς τον χρήστη που θα μπορούσαν να παρέχουν την πληροφορία από βάσεις δεδομένων και συλλογές σε όλων τον κόσμο, έτσι ώστε να είναι πιο εύκολα διαχειρίσιμα από τους χρήστες. Για την υλοποίηση όμως τέτοιων συστημάτων τα συστήματα επόμενης γενιάς θα έπρεπε να καταλάβουν τη σημασιολογία των ερωτημάτων όχι με την έννοια της σε χαμηλό επίπεδο λειτουργίας του υπολογιστή.

Από την πλευρά της αναγνώρισης υποδειγμάτων αυτό σήμαινε την μετάφραση των χαμηλού επιπέδου οπτικών περιεχομένων μιας εικόνας σε υψηλού επιπέδου όρους που θα γίνονταν εύκολα διαχειρίσιμοι από τον χρήστη. Μερικές από τις κυριότερες εργασίες που ασχολήθηκαν με το συγκεκριμένο τομέα υπάρχουν μόνο για την αναγνώριση προσώπων από τους *Rowley et al (1996)* και από τους *Lew et al (1996)*. Το πιο πρόσφατο σύστημα *Ανάκτησης Εικόνας με βάση το περιεχόμενο* που αναπτύχθηκε και προσπάθησε να καλύψει το κενό στον τομέα της αναζήτησης με ερωτήματα, της ευρετηριοποίησης και ταξινόμησης ήταν η μηχανή αναζήτησης *ImageScare*, που προτάθηκε από τον *Lew (2000)*.

Στην μηχανή αναζήτησης εικόνων *ImageScare*, ο χρήστης μπορούσε να δημιουργήσει απευθείας ερωτήματα με πληθώρα οπτικών αντικειμένων, χρησιμοποιώντας εικονίδια τα οποία βρίσκονταν στον δικτυακό τόπο ως βάση της αναζήτησής του και ξεπερνούσαν σε αριθμό τα 10 εκατομμύρια. Το σύστημα

χρησιμοποιούσε αλγορίθμους πληροφοριών επέστρεφε τις εικόνες εκείνες που χρησιμοποιούσαν σε μεγαλύτερο βαθμό τα χαρακτηριστικά που ζητούσε ο χρήστης.

2.3 Ανάκτηση εικόνας με βάση το κείμενο

Οι βιβλιοθήκες φωτογραφιών, τα μουσεία και οι εκθέσεις έργων τέχνης έδειξαν ιδιαίτερο ενδιαφέρον στην ψηφιοποίηση των συλλογών τους. Οι εικόνες οργανώνονται με απώτερο στόχο ο ενδιαφερόμενος να έχει τη δυνατότητα να εντοπίσει την επιθυμητή εικόνα. Ενώ είναι εφικτό να εντοπίσει κάποιος μια επιθυμητή εικόνα από μια μικρού μεγέθους συλλογή απλώς με ξεφύλλισμα, εξακολουθεί να είναι απαραίτητος κάποιος τρόπος για τη διαχείριση και ανάκτηση εικόνων από μεγάλες ψηφιακές βιβλιοθήκες εικόνων.

Πιο αποτελεσματικές τεχνικές από το απλό ξεφύλλισμα είναι αναγκαίες για συλλογές που περιέχουν μεγάλο αριθμό εικόνων. Για την οργάνωση των συλλογών των εικόνων χρησιμοποιήθηκαν λέξεις που αποδίδονταν στις εικόνες με τη μορφή μεταδεδομένων σύμφωνα με ορισμένα πρότυπα και αφορούν στοιχεία της εικόνας όπως ο δημιουργός, ο τόπος, ο χρόνος καθώς και το θέμα της. Η τεχνική αυτή που βασίζεται σε λέξεις για την ανάκτηση εικόνας ονομάζεται ανάκτηση εικόνας με βάση το κείμενο και αποτελεί την κλασική τεχνική που χρησιμοποιήθηκε και παλιότερα στις αναλογικές συλλογές εικόνων.

Αναφορικά με την απόδοση λέξεων στις εικόνες, δηλαδή την ευρετηριοποίησή τους, το θέμα της εικόνας έχει ιδιαίτερη σημασία γιατί αποτελεί το σημαντικότερο σημείο πρόσβασης από την πλευρά του χρήστη για την αναζήτηση και ανάκτηση εικόνων από μια συλλογή. Για την απόδοση των θεμάτων των εικόνων δηλαδή για την ευρετηρίαση και ταξινόμησή τους οι βιβλιοθήκες εικόνων χρησιμοποιούν διάφορα δομημένα συστήματα, που ελέγχουν το βασικό λεξιλόγιο μιας γλώσσας θεματικής ευρετηρίασης και αποδίδουν σημασιολογικές σχέσεις στους όρους ή και ελεύθερο κείμενο με λεπτομερείς περιγραφές.

2.3.1 Συστήματα ευρετηρίασης και ταξινόμησης εικόνων

Ένα από τα πιο γνωστά συστήματα ευρετηρίασης είναι ο Θησαυρός των Τεχνών και της Αρχιτεκτονικής που δημιουργήθηκε στο Πολυτεχνείο του Rensselaer το 1979 και σήμερα χρησιμοποιείται από βιβλιοθήκες τέχνης σε όλο τον κόσμο. Ο θησαυρός υποστηρίζεται από το Ινστιτούτο Πληροφόρησης Getty και αποτελείται από περίπου 120.000 όρους για την περιγραφή αντικειμένων, εικόνων και άλλων υλικών πολιτιστικής κληρονομιάς. Διαιρείται σε επτά κατηγορίες που υποδιαιρούνται σε 33 υποκατηγορίες για την περιγραφή εννοιών, περιόδων, δραστηριοτήτων, υλικών, τεχνοτροπιών και αντικειμένων. Άλλα εργαλεία για την περιγραφή εικόνων από το Ινστιτούτο Getty είναι ο Συλλογικός Κατάλογος Ονομάτων Καλλιτεχνών και ο Θησαυρός Γεωγραφικών Ονομάτων του Getty.

Ένα άλλο σύστημα που αναπτύχθηκε από το Ινστιτούτο Getty για την ευρετηρίαση των φωτογραφιών είναι ένας θησαυρός που ανταποκρίνεται στην ειδική φύση της συλλογής τους και αποτελείται από περίπου 10.000 όρους, κατανεμημένους σε εννέα σημασιολογικές ομάδες που περιλαμβάνουν τη γεωγραφία, τους ανθρώπους, τις δραστηριότητες και τις έννοιες. Οι θεματικοί όροι αποδίδονται στην εικόνα συνολικά, στα αντικείμενα που περιέχονται και στις ομάδες τους.

Στη συνέχεια, ένα άλλο σύστημα, που σχεδιάστηκε για την ευρετηρίαση και παροχή πρόσβασης σε συλλογές ιστορικών κυρίως εικόνων στο αυτοματοποιημένο περιβάλλον, είναι ο Θησαυρός για Γραφικά Υλικά της Βιβλιοθήκης του Κογκρέσου που προήλθε από τις Θεματικές Επικεφαλίδες της Βιβλιοθήκης του Κογκρέσου.

Ακόμη το 1983 δημοσιεύτηκε ο Εικονογραφικός Θησαυρός που περιέχει περίπου 5.500 όρους από το Υπουργείο Πολιτισμού της Γαλλίας για τη διαχείριση των μνημείων και των καλλιτεχνικών θησαυρών της Γαλλίας

Επίσης, υπάρχουν αρκετά συστήματα ευρετηρίασης που χρησιμοποιούν ταξινομικούς κωδικούς αντί για λέξεις κλειδιά ή θεματικούς περιγραφείς για να αποδώσουν το περιεχόμενο της εικόνας επειδή παρέχεται με αυτό τον τρόπο ανεξαρτησία από τη γλώσσα σε μεγαλύτερο βαθμό και οι εννοιολογικές ιεραρχίες απεικονίζονται πιο καθαρά. Τέτοια συστήματα είναι το ICONCLASS που αναπτύχθηκε από το Πανεπιστήμιο του Leiden και το TELCLASS που δημιουργήθηκε από το BBC. Το πρώτο σχεδιάστηκε για την ταξινόμηση έργων τέχνης ενώ το δεύτερο δίνει έμφαση στα προγράμματα της τηλεόρασης και τα βίντεο.

2.3.2 Πλεονεκτήματα - μειονεκτήματα ανάκτησης εικόνας με βάση το κείμενο - ανάγκη για ανάκτηση με βάση το περιεχόμενο

Στις περιπτώσεις αυτές η αναζήτηση και ανάκτηση των εικόνων γίνεται με βάση το κείμενο δηλαδή με τη χρήση των μεταδεδομένων που έχουν αποδοθεί στις εικόνες από τον ειδικό που τις επεξεργάστηκε κατά την ένταξή τους στη συλλογή και απέδωσε με λέξεις τα χαρακτηριστικά τους ώστε να είναι δυνατή η ανάκτησή τους από το χρήστη της ψηφιακής βιβλιοθήκης.

Οι τεχνικές ευρετηρίασης έχουν μεγάλη εκφραστική δύναμη και μπορούν να χρησιμοποιηθούν για να περιγράψουν σχεδόν κάθε πλευρά του περιεχομένου της εικόνας. Τα συστήματα ευρετηρίασης είναι επεκτάσιμα και ευέλικτα. Μπορούν εύκολα να προστεθούν σ' αυτά νέες έννοιες όπου απαιτούνται για τη λεκτική απόδοση του περιεχομένου των εικόνων. Επίσης παρέχεται η δυνατότητα περιγραφής του περιεχομένου των εικόνων σε ποικίλους βαθμούς πολυπλοκότητας. Ανάλογα με την πολιτική, τους στόχους της ψηφιακής βιβλιοθήκης και το κοινό στο οποίο απευθύνεται η περιγραφή μπορεί να είναι από πολύ γενική ως πολύ εξειδικευμένη και αναλυτική.

Όμως η διαδικασία της ευρετηρίασης από τον άνθρωπο, είτε με λέξεις κλειδιά είτε με ταξινομικά σύμβολα, έχει ορισμένα βασικά μειονεκτήματα. Κατά πρώτο λόγο είναι χρονοβόρα διαδικασία και απαιτεί μεγάλη κατανάλωση πόρων σε χρόνο και χρήμα. Έχει υπολογιστεί ότι ο χρόνος ευρετηρίασης ποικίλει από περίπου 7 λεπτά ανά εικόνα στο Ινστιτούτο Getty που χρησιμοποιείται το δικό τους σύστημα ως και περισσότερο από 40 λεπτά για κάθε εικόνα σε μια συλλογή διαφανειών στο Πολυτεχνείο του Rensselaer, που χρησιμοποιεί το Θησαυρό των Τεχνών και της Αρχιτεκτονικής. Όσον αφορά το βίντεο ο χρόνος ευρετηρίασης είναι μεγαλύτερος.

Επιπρόσθετα, το άλλο βασικότερο πρόβλημα πηγάζει από το πλούσιο περιεχόμενο των εικόνων και από την υποκειμενικότητα της ανθρώπινης αντίληψης. Η ευρετηρίαση, παρά το γεγονός ότι γίνεται από κάποιον ειδικό σχολιαστή -ευρετηριαστή, έχει το στοιχείο της υποκειμενικότητας γιατί το περιεχόμενο της ίδιας εικόνας γίνεται διαφορετικά αντιληπτό από κάθε άνθρωπο. Η οπτική του ευρετηριαστή μπορεί να διαφέρει από αυτή του χρήστη. Άλλωστε οι λέξεις δεν έχουν τη δυνατότητα να εκφράσουν το πλήρες νόημα της εικόνας. Αφηρημένες έννοιες και αισθήματα, όπως η ελευθερία και η χαρά, έχουν διαφορετικό νόημα για κάθε άτομο. Από έρευνες που

έχουν διεξαχθεί προέκυψαν μεγάλες διαφορές στις λέξεις κλειδιά που διαφορετικά άτομα προσέδωσαν στην ίδια εικόνα.

Επίσης υπάρχουν περιορισμοί στην ευρετηρίαση με λέξεις κλειδιά, ιδιαίτερα σε μεγάλες συλλογές, όπως οι λάθος συντακτικά λέξεις, λεξικά που αποτυγχάνουν να ξεπεράσουν τα αποτελέσματα των ανακριβών περιγραφών καθώς επίσης ότι υπάρχουν εικόνες με ελάχιστα ή και καθόλου μεταδεδομένα. Αυτοί οι περιορισμοί σημαίνουν ότι η ανάκτηση εικόνων ορισμένες φορές πρέπει να βασίζεται στη γνώση και την εμπειρία του προσωπικού. Το συμπέρασμα είναι ότι η υποκειμενική αντίληψη και ο ατελής ή ανακριβής σχολιασμός της εικόνας μπορούν να προκαλέσουν πολλά προβλήματα στις διαδικασίες ανάκτησης και να οδηγήσουν σε ελλιπή ή λανθασμένη ανάκτηση από το χρήστη.

2.4 Ανάκτηση εικόνας με βάση το περιεχόμενο (CBIR)

Ο όρος CBIR πρωτοεμφανίστηκε το 1992 από μία ομάδα ερευνητών, θέλοντας να περιγράψουν τα πειράματά ανάκτησης βάσει σχήματος και χρώματος τα οποία υλοποίησαν. Στην παρούσα ενότητα, επιχειρείται η περιγραφή των βασικών χαρακτηριστικών που εξάγονται σε CBIR συστήματα, η επισκόπηση των γνωστών μεθόδων καθώς και οι προκλήσεις στην ανάκτηση εικόνας βάσει περιεχομένου.

2.4.1 Περιγραφή CBIR συστήματος

Παρότι υπάρχουν πολλά διαφορετικά CBIR συστήματα, τα περισσότερα από αυτά ακολουθούν περίπου την ίδια διαδικασία. Η μεθοδολογία των CBIR συστημάτων βασίζεται στην εξαγωγή χαρακτηριστικών. Η διαδικασία αυτή υλοποιείται στα ακόλουθα βήματα:

- **Εξαγωγή χαρακτηριστικών**

Αφού επιλεχθούν τα χαρακτηριστικά σύμφωνα με τα οποία θα γίνει η σύγκριση, εξάγονται οι πίνακες μέσω μεθόδων επεξεργασίας εικόνας για κάθε εικόνα της βάσης.

- **Αίτηση**

Σε αυτό το στάδιο γίνεται η περιγραφή των επιθυμητών ανακτώμενων εικόνων από το σύστημα. Η αίτηση μπορεί να γίνει με πολλούς τρόπους. Ο πιο συνηθισμένος είναι η αίτηση μέσω παραδείγματος, κατά την οποία εισάγεται μία εικόνα στο σύστημα ζητώντας στην έξοδο όμοιες προς αυτή.

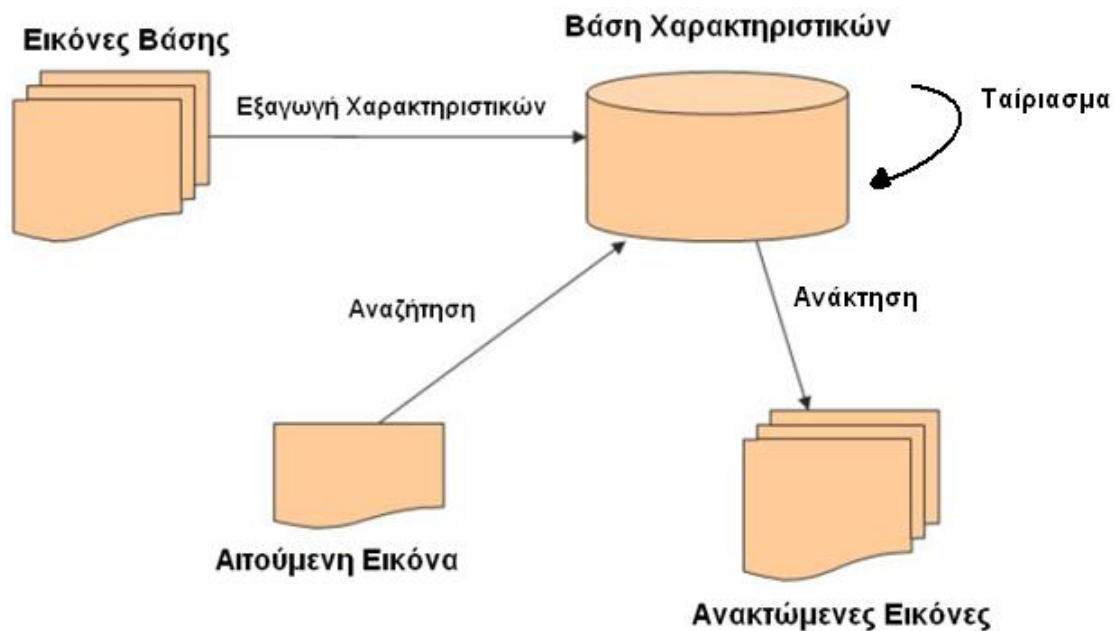
- **Ταίριασμα**

Σε αυτό το στάδιο συγκρίνονται τα διανύσματα των εικόνων βάσης με τα επιθυμητά χαρακτηριστικά του διανύσματος. Στην περίπτωση όπου το αίτημα είναι μία εικόνα, η σύγκριση γίνεται μέσω του πίνακα χαρακτηριστικών της. Ο τρόπος με τον οποίο γίνεται η σύγκριση, εξαρτάται από το μέτρο ομοιότητας που χρησιμοποιείται.

- **Ανάκτηση**

Κατά τη διάρκεια του ταιριάσματος δημιουργείται ένας πίνακας αποστάσεων μεταξύ όλων των χαρακτηριστικών διανυσμάτων, από τον οποίο μπορεί να γίνει άμεση κατάταξη των εικόνων που μοιάζουν περισσότερο με τη ζητούμενη. Ο τρόπος απεικόνισης, μπορεί να γίνει σε μία διάσταση υπό μορφή λίστας, ή σε δύο έως τρεις διαστάσεις σε ένα χάρτη (Cox et al, 2001).

Η σχηματική απεικόνιση της παραπάνω διαδικασίας παρατίθεται στην παρακάτω εικόνα.



Εικόνα 1: Γενική δομή ενός CBIR συστήματος

Η αποτελεσματικότητα ενός CBIR συστήματος κρίνεται συνήθως με βάση το πλήθος των ανακτώμενων εικόνων που σχετίζονται με τη ζητούμενη. Τα πιο συνηθισμένα μέτρα αξιολόγησης των αποτελεσμάτων, είναι οι ποσότητες ακρίβειας και ανάκλησης. Οι δύο αυτές μεταβλητές μπορούν να χρησιμοποιηθούν ταυτόχρονα σε μία γραφική παράσταση ανάκλησης. Ένας άλλος τρόπος αξιολόγησης προϋποθέτει τη σύγκριση των αποτελεσμάτων με αυτά που προκύπτουν από προηγούμενα επιτυχημένα συστήματα. Το πρόβλημα σε αυτή την περίπτωση είναι η εύρεση των κατάλληλων εικόνων που θα μπορέσουν να καλύψουν ολόκληρη τη βάση. Μια άλλη προσέγγιση προϋποθέτει τη χρήση μιας μικρότερης ενδεικτικής βάσης εικόνων με μικρό αριθμό κατηγοριών. Σε αυτή την υλοποίηση, επιλέγονται αντιπροσωπευτικές εικόνες με τις οποίες θα μπορεί να γίνει όσο το δυνατόν πιο αντικειμενική η προσομοίωση του συστήματος με την αυθεντική βάση.

2.4.2 Εξαγωγή χαρακτηριστικών

Τα χαρακτηριστικά τα οποία εξάγονται από τις εικόνες μπορούν να χωριστούν σε βασικά και σημασιολογικά. Τα πρώτα, είναι χαμηλού επιπέδου χαρακτηριστικά, όπως η περίμετρος αντικειμένων και το χρωματικό ιστόγραμμα. Τα σημασιολογικά χαρακτηριστικά αναφέρονται σε εννοιολογικής φύσεως πληροφορία, όπως η

αναγνώριση ενός προσώπου σε μία μεγάλη βάση. Στις υποενότητες που ακολουθούν, γίνεται μια σύντομη περιγραφή των πιο συνηθισμένων χαρακτηριστικών.

Χαρακτηριστικά χρώματος

Τα χαρακτηριστικά χρώματος είναι ευρέως διαδεδομένα στην ανάκτηση εικόνας, δεδομένου ότι το χρώμα είναι άμεσα αντιληπτό από τον άνθρωπο κατά την παρατήρηση μίας εικόνας. Είναι ανεξάρτητα από τον θόρυβο, το μέγεθος και τη γωνία της εικόνας, αλλά μπορεί να επηρεαστεί από την υφή, τη φωτεινότητα και τα φαινόμενα σκίασης. Η αντιμετώπιση αυτών των χρωματικών μεταβολών, είναι εξαιρετικά δύσκολη, καθώς δεν εξάγεται πληροφορία για το περιεχόμενο της εικόνας.

Τα συστήματα ανάκτησης βάσει χρώματος διαφέρουν ως προς τις ιδιότητές τους, καθιστώντας κάθε ένα από αυτά κατάλληλο για συγκεκριμένες εφαρμογές. Είναι πολύ σημαντικό ότι οι αποστάσεις των χρωμάτων στο χρωματικό σύστημα να αντιστοιχούν σε αυτές του συστήματος όρασης. Πολλά επιτυχημένα CBIR συστήματα βασίζονται σε έγχρωμα ιστογράμματα. Γενικά, το έγχρωμο ιστόγραμμα παρέχει χρήσιμα στοιχεία για τη μετέπειτα έκφραση της ομοιότητας μεταξύ δύο εικόνων, εξαιτίας της βέλτιστης απόδοσής του σε περίπλοκα φόντα των εικόνων και τις παραμορφώσεις αντικειμένων. Η αναπαράστασή τους προκύπτει από την καταμέτρηση του πλήθους των εικονοστοιχείων κάθε χρώματος. Τα συστήματα τα οποία χρησιμοποιούν ιστογράμματα, επιδεικνύουν σημαντικά γρήγορη ταχύτητα ανάκτησης η οποία κάνει τις υλοποιήσεις πραγματικού χρόνου ευκολότερες. Το ολικό ιστόγραμμα (Stricker et al, 1995) υπολογίζει την κοινή πιθανότητα των εντάσεων των χρωματικών καναλιών, περιγράφοντας την ολική κατανομή του χρώματος. Συνήθως, προτιμώνται για χρωματικές κατανομές υψηλών διαστάσεων.

Χαρακτηριστικά υφής

Η υφή είναι μια γνωστή έννοια χωρίς όμως σαφή ορισμό. Τα χαρακτηριστικά τα οποία μπορούν να την περιγράψουν είναι τρία:

- Η τοπική επαναληπτικότητα, η οποία εμφανίζεται σε μία μεγάλη περιοχή της εικόνας.
- Η μη τυχειότητα της επαναληπτικής δομής.
- Τα δομικά στοιχεία τα οποία χαρακτηρίζουν τις περιοχές με μία υφή, είναι μονάδες με ομοιόμορφη κατανομή που έχουν τις ίδιες διαστάσεις.

Τα δομικά αυτά στοιχεία πήραν την ονομασία *textons* το 1981. Οι υφές περιγράφονται κατά κύριο λόγο από τη χωρική τους συχνότητα και από τη διεύθυνσή τους. Επίσης, έχουν τη δυνατότητα να διαχωρίζουν περιοχές με το ίδιο χρώμα. Σε συνδυασμό με χρωματικά χαρακτηριστικά, η υφή μπορεί να αποτελέσει μέτρο ανίχνευσης αντικειμένων. Η περιγραφή τους μπορεί να γίνει μέσω στατιστικών, δομικών, στοχαστικών και συχνοτικών προσεγγίσεων. Οι στατιστικές προσεγγίσεις, χρησιμοποιούν μια ομάδα από χαρακτηριστικά όπως η αντίθεση και η εντροπία, για να αντιπροσωπεύσουν εικόνες υφής. Οι δομικές, χρησιμοποιούν δισδιάστατα πρότυπα, όπως παράλληλες γραμμές, κύκλοι και περιοδικά σημεία. Οι στοχαστικές προσεγγίσεις, βασίζονται σε στοχαστικές διαδικασίες οι οποίες μπορούν να χρησιμοποιηθούν για την εξαγωγή χαρακτηριστικών σε τμηματοποίηση και κατηγοριοποίηση υφής. Τέλος, οι συχνοτικές προσεγγίσεις, στηρίζονται στην ανάλυση συναρτήσεων φασματικής κατανομής στο πεδίο των συχνοτήτων. Συνήθως, χρησιμοποιούνται για την ανίχνευση περιοδικότητας σε μία εικόνα εντοπίζοντας στενές περιοχές με υψηλή ενέργεια στο φάσμα.

Χαρακτηριστικά σχήματος

Σύμφωνα με τους *Shomaker et al (1999)*, ο άνθρωπος εξάγει χαρακτηριστικά σχήματος, παρά χρώματος ή υφής. Ο ανθρώπινος εγκέφαλος έχει τη δυνατότητα αναγνώρισης ενός αντικειμένου μέσα από ασαφή σχήματα τα οποία μπορούν να το περιγράψουν. Αντιθέτως, η ανάκτηση εικόνας βάσει σχήματος παραμένει ακόμα πρόκληση για τα *CBIR* συστήματα. Γενικά, οι σχηματικές αναπαραστάσεις μπορούν να ταξινομηθούν σε τρεις κύριες κατηγορίες, σύμφωνα με τους *Mehtre et al (1997)*. Η επιλογή της κατηγορίας εξαρτάται από τις ανάγκες της εφαρμογής, όπως για παράδειγμα η αντιμετώπιση του θορύβου. Στην πραγματικότητα, υπάρχουν πολλές

περισσότερες μέθοδοι στην υπολογιστική όραση για τον υπολογισμό των χαρακτηριστικών του σχήματος, με αρκετές από αυτές να παρατίθενται στην εργασία (Veltkamp, 2001).

2.5 Ανάκτηση εικόνας βάση οπτικής πληροφορίας (VBIR)

Η χρήση *Eye Tracker* σε ένα *CBIR* σύστημα, δίνει τη δυνατότητα εισαγωγής των αναγκών του χρήστη από την καταγραφή των οφθαλμικών κινήσεων. Μέσω της οπτικής ανάδρασης, το σύστημα εκπαιδεύεται και προσαρμόζεται στις διαφορετικές ανάγκες του εκάστοτε ατόμου. Μία από τις πρώτες ολοκληρωμένες εργασίες ανάκτησης εικόνας βάσει οπτικής πληροφορίας δημοσιεύτηκε από τον Essing (2007).

Εκτελώντας ένα πείραμα συγκριτικής οπτικής αναζήτησης, ο χρήστης καλείται να ανακαλύψει διαφορές ανάμεσα στις εικόνες βάσης σε μια ζητούμενη εικόνα. Τα χαρακτηριστικά τα οποία εξάγονται από περιοχές που παρουσιάζουν πολλά σημεία εστίασης, έχουν μεγαλύτερη βαρύτητα σε σχέση με αυτά τα οποία δεν παρατηρούνται. Ο *Eye Tracker* εκμεταλλεύεται την ικανότητα του ανθρώπινου εγκεφάλου να λαμβάνει στιγμιαίες αποφάσεις για την αναγνώριση και κατηγοριοποίηση αντικειμένων από το οπτικό ερέθισμα το οποίο δέχεται. Επιπλέον, τα σημεία εστίασης πάνω σε μία εικόνα αποκαλύπτουν τις σημασιολογικά σημαντικές περιοχές από τις οποίες μπορεί να γίνει πιο αξιόπιστο το σύστημα ανάκτησης.

Η αλγοριθμική εύρεση περιοχών ενδιαφέροντος έχει προσελκύσει το ερευνητικό ενδιαφέρον τις τελευταίες δεκαετίες. Στην εργασία των Privitera et al (1998), παρουσιάζεται μια συγκριτική μελέτη των μέχρι τότε τεχνικών. Τα τελευταία χρόνια, έχουν αναπτυχθεί κι άλλοι αλγόριθμοι εύρεσης με κυριότερους τον SIFT, από τον Lowe (1999) και αυτόν που προτάθηκε από τον Stentiford (2007).

ΚΕΦΑΛΑΙΟ 3: Ανάκτηση εικόνων με βάση το περιεχόμενο

Στο κεφάλαιο αυτό θα αναλυθούν σε βάθος οι τεχνικές ανάκτησης εικόνων με βάση το περιεχόμενο.

3.1 Εισαγωγή

Ανάκτηση εικόνας με βάση το περιεχόμενο ονομάζεται η τεχνική για την ανάκτηση εικόνας από μια ψηφιακή βιβλιοθήκη ή τον παγκόσμιο ιστό με τη χρήση χαρακτηριστικών στοιχείων που εξάγονται αυτόματα με βάση το οπτικό της περιεχόμενο. Ονομάζεται επίσης και ανάκτηση οπτικής πληροφορίας με βάση το περιεχόμενο και ερώτημα στο περιεχόμενο της εικόνας και αποτελεί εφαρμογή του τομέα της όρασης των υπολογιστών στο πρόβλημα της ανάκτησης της εικόνας. Είναι σημαντική σ' αυτό το σημείο η επισήμανση ότι η ανάκτηση με βάση λέξεις κλειδιά ή θεματικούς όρους δεν αποτελεί ανάκτηση βασισμένη στο περιεχόμενο ακόμη κι όταν οι όροι αυτοί αφορούν το θέμα, δηλαδή το περιεχόμενο της εικόνας.

Η παρακάτω βασική ιδέα αποτελεί τη βάση της ανάκτησης εικόνας με βάση το περιεχόμενο: γίνεται εξαγωγή και ανάλυση χαρακτηριστικών από την ίδια την εικόνα δηλαδή οι εικόνες ευρετηριάζονται αυτόματα με τη χρήση αλγόριθμων επεξεργασίας εικόνας και μετατρέπονται σε μαθηματικά δεδομένα και στατιστικά στοιχεία ώστε να είναι μετρήσιμα και συγκρίσιμα από τον υπολογιστή για την ανίχνευση ομοιότητας μεταξύ των εικόνων. Χρήσιμα χαρακτηριστικά για την ανάκτηση με βάση το περιεχόμενο θεωρούνται αυτά που μιμούνται τα χαρακτηριστικά που βλέπει ο άνθρωπος, αυτά που γίνονται δηλαδή αντιληπτά από την ανθρώπινη όραση. Η χρήση αυτών των οπτικών γνωρισμάτων, που αντανακλούν κάποια πλευρά της ομοιότητας των εικόνων όπως γίνεται αντιληπτή από τον άνθρωπο ακόμη κι αν αυτός έχει μεγάλη δυσκολία να τα περιγράψει, αυξάνει την πιθανότητα ότι το σύστημα ανακαλεί τις εικόνες που είναι παρόμοιες ή μοιάζουν σύμφωνα με την ανθρώπινη αντίληψη.

Τα χαρακτηριστικά που χρησιμοποιούνται για τη βασισμένη στο περιεχόμενο ανάκτηση εικόνας κατηγοριοποιούνται στα παγκόσμια χαρακτηριστικά και τα τοπικά

χαρακτηριστικά. Τα μεν είναι αυτά που αφορούν και υπολογίζονται για το σύνολο της εικόνας, τα δε είναι αυτά που εξάγονται από κάποιο μέρος ή τμήμα της εικόνας.

3.2 Χρώμα

Το χρώμα είναι ένα από τα πιο διαδεδομένα χαρακτηριστικά μιας εικόνας. Η έρευνα για τα χρώματα ξεκίνησε ουσιαστικά από τα πρώτα βήματα των συστημάτων ανάκτησης εικόνων Wang *et al.* (1997). Μεγάλη έμφαση δόθηκε τότε στην εξερεύνηση διαφόρων χρωματικών χώρων που φαίνεται να προσεγγίζουν καλύτερα τον τρόπο με τον οποίο το ανθρώπινο μάτι αντιλαμβάνεται το χρώμα σε σχέση με τον κλασικό RGB χώρο ο οποίος ήταν ευρέως διαδεδομένος μέχρι εκείνη την στιγμή. Μελετήθηκε επίσης το πώς μεταβάλλεται το χρώμα σε συνάρτηση με τις αλλαγές στον φωτισμό, την θέση της κάμερας και την κλίση των επιφανειών και έχουν γίνει προσπάθειες να αναπτυχθούν μοντέλα που να λαμβάνουν υπόψη τις μεταβολές αυτές Sebe *et al* (2001).

Το μοντέλο της αντίθετης αναπαράστασης χρωμάτων, που προτάθηκε από τους Swain *et al* (1991), κατάφερε να λάβει υπόψη τη μεταβολή της φωτεινότητας σε έναν άξονα, με αποτέλεσμα να υπάρχει η δυνατότητα να απαλειφθεί αυτή η συνιστώσα. Το ιστόγραμμα μιας εικόνας ή των χρωματικών της αποδόσεων είναι μια πάγια τεχνική που χρησιμοποιήθηκε σε πολλά συστήματα ανάκτησης της προηγούμενης δεκαετίας, όπως για παράδειγμα το σύστημα QBIC, που προτάθηκε από τους Flickner *et al.* (1995) της IBM. Στατιστικά το ιστόγραμμα χρωμάτων υποδηλώνει την από κοινού πιθανότητα των εντάσεων των τριών χρωματικών καναλιών. Μια εναλλακτική του ιστογράμματος αυτού είναι η χρήση των χρωματικών στιγμιότυπων, που προτάθηκε από τον Stricker *et al* (1995).

Με το πέρασμα των χρόνων η έρευνα εστιάστηκε στην εξαγωγή μιας όσο το δυνατόν καλύτερης περίληψης των χρωμάτων της εικόνας, λαμβάνοντας υπόψη και την χωρική τους κατανομή. Το πρότυπο MPEG-7 περιλαμβάνει σύγχρονους χρωματικούς περιγραφείς οι οποίοι θα αναλυθούν σε επόμενο κεφάλαιο.

Οι χρωματικοί χώροι ουσιαστικά χρησιμοποιούνται για να περιγράψουν τον τρόπο κωδικοποίησης της χρωματικής πληροφορίας σε μια εικόνα. Το κάθε χρωματικό μοντέλο ορίζει ένα σύστημα συντεταγμένων και έναν υποχώρο μέσα σε αυτό. Το κάθε χρώμα είναι μοναδικό και αναπαρίσταται σε ένα μοναδικό σημείο.

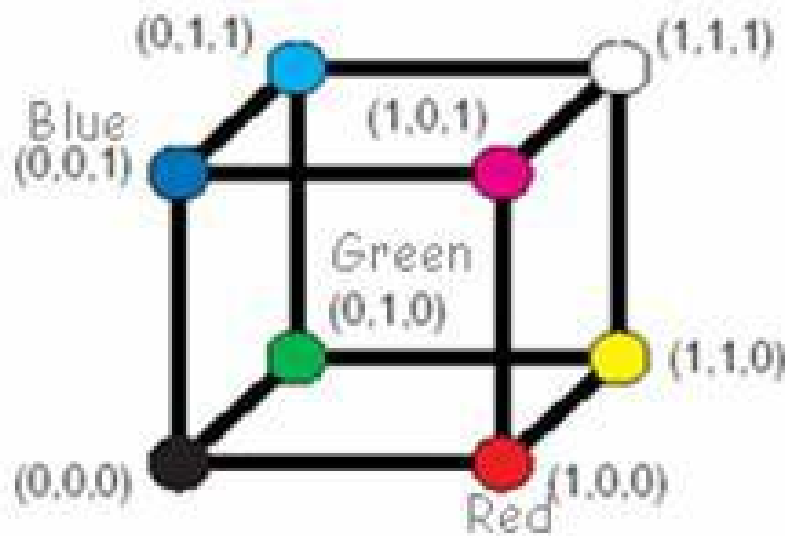
RGB χρωματικό σύστημα

Το πρότυπο χρώματος RGB είναι ένα προσθετικό πρότυπο στο οποίο τα χρώματα κόκκινο, πράσινο και μπλε (χρώματα που χρησιμοποιούνται συχνά σε προσθετικά χρωματικά πρότυπα) συνδυάζονται με διάφορους τρόπους για να αναπαραχθούν άλλα χρώματα. Το όνομα του προτύπου και η σύντηξη RGB προέρχονται από τα τρία βασικά χρώματα, το κόκκινο, πράσινο, και το μπλε. Αυτά τα τρία χρώματα δεν πρέπει να συγχέονται με τα τρία ανακλαστικά χρώματα κόκκινο, μπλε, και κίτρινο, τα οποία αναφέρονται στον χώρο των τεχνών ως βασικά χρώματα.

Ένα χρώμα στο πρότυπο χρώματος RGB μπορεί να περιγραφεί με το προσδιορισμό του πόσο κάθε ένα από το κόκκινο, πράσινο και μπλε χρώματα συμπεριλαμβάνεται. Κάθε ένα μπορεί να ποικίλει μεταξύ του ελάχιστου (καθόλου χρώμα) και του μεγίστου (πλήρης ένταση). Εάν όλα τα χρώματα είναι στο ελάχιστο το αποτέλεσμα είναι μαύρο. Εάν όλα τα χρώματα είναι στο μέγιστο, το αποτέλεσμα είναι το άσπρο. Τα χρώματα μπορούν να περιγραφούν ποσοτικά με διάφορους τρόπους:

- Οι επιστήμονες του χρώματος συχνά τοποθετούν τα χρώματα στην κλίμακα 0 (ελάχιστο) έως 1 (μέγιστο). Πολλοί μαθηματικοί τύποι που σχετίζονται με το χρώμα χρησιμοποιούν αυτές τις τιμές. Π.χ. το μέγιστο κόκκινο είναι 1,0,0 για Κόκκινο, Πράσινο, Μπλε.
- Οι τιμές χρώματος μπορούν να γραφτούν επίσης ως ποσοστά, από 0% (ελάχιστο) ως 100% (μέγιστο). Το μέγιστο κόκκινο είναι 100%, 0%, 0%.
- Οι τιμές χρώματος μπορούν να γραφτούν ως αριθμοί στην κλίμακα 0 έως 255, απλά με τον πολλαπλασιασμό της κλίμακας 0.0 έως 1.0 με 255. Αυτό το μοντέλο απαντάται συνήθως στην πληροφορική, όπου οι προγραμματιστές προτιμούν να αποθηκεύουν κάθε αξία χρώματος σε ένα byte (8 bit). Αυτή η σύμβαση έχει γίνει τόσο διαδεδομένη ώστε πολλοί συγγραφείς την θεωρούν αυτονόητη και δεν παρέχουν το σωστό υπόβαθρο αναφοράς. Το μέγιστο κόκκινο είναι το 255,0,0.

- Η ίδια σειρά, 0 έως 255, γράφεται μερικές φορές σε δεκαεξαδικό, και ίσως με ένα πρόθεμα (π.χ. #). Επειδή οι δεκαεξαδικοί αριθμοί σε αυτήν την κλίμακα μπορούν να γραφτούν με ένα σταθερό σχήμα δύο ψηφίων, το μέγιστο κόκκινο #FF, #00, #00 μπορεί να γραφτεί και σαν #ff0000. Αυτή η σύμβαση χρησιμοποιείται στα χρώματα στο διαδίκτυο και θεωρείται επίσης από μερικούς συγγραφείς αυτονόητη.



Εικόνα 2: Μοντέλο RGB

CMYK

Το χρωματικό μοντέλο CMYK είναι ένα αφαιρετικό χρωματικό μοντέλο. Χρησιμοποιεί τέσσερα βασικά χρώματα, το κυανό, το ματζέντα, το κίτρινο και το μαύρο. Χρησιμοποιείται επίσης για την εξαγωγή χρωμάτων και ανοιχτόχρωμα υπόβαθρα κυρίως τα λευκά. Το μοντέλο αυτό είναι αφαιρετικό γιατί το χρώμα αφαιρεί την λάμψη από το λευκό υπόβαθρο κι έτσι δημιουργείται η αίσθηση ενός διαφορετικού χρώματος.

Η διαφορά των δύο μοντέλων RGB και CMYK είναι ότι ενώ στο RGB με την τριάδα (0,0,0) αναπαριστούμε το μαύρο δηλαδή την πλήρη απορρόφηση του φωτός και η (1,1,1) το λευκό δηλαδή την πλήρη αντανάκλαση του φωτός, στο CMYK μοντέλο ισχύει ακριβώς το αντίστροφο. Δηλαδή εδώ αντί να προσθέτουμε φως στο Μαύρο,

μετράμε πόσο χρώμα αφαιρούμε από το Λευκό. Άρα στο CMYK με το 1,1,1 εννοούμε το μαύρο (απορροφάται όλο το χρώμα).

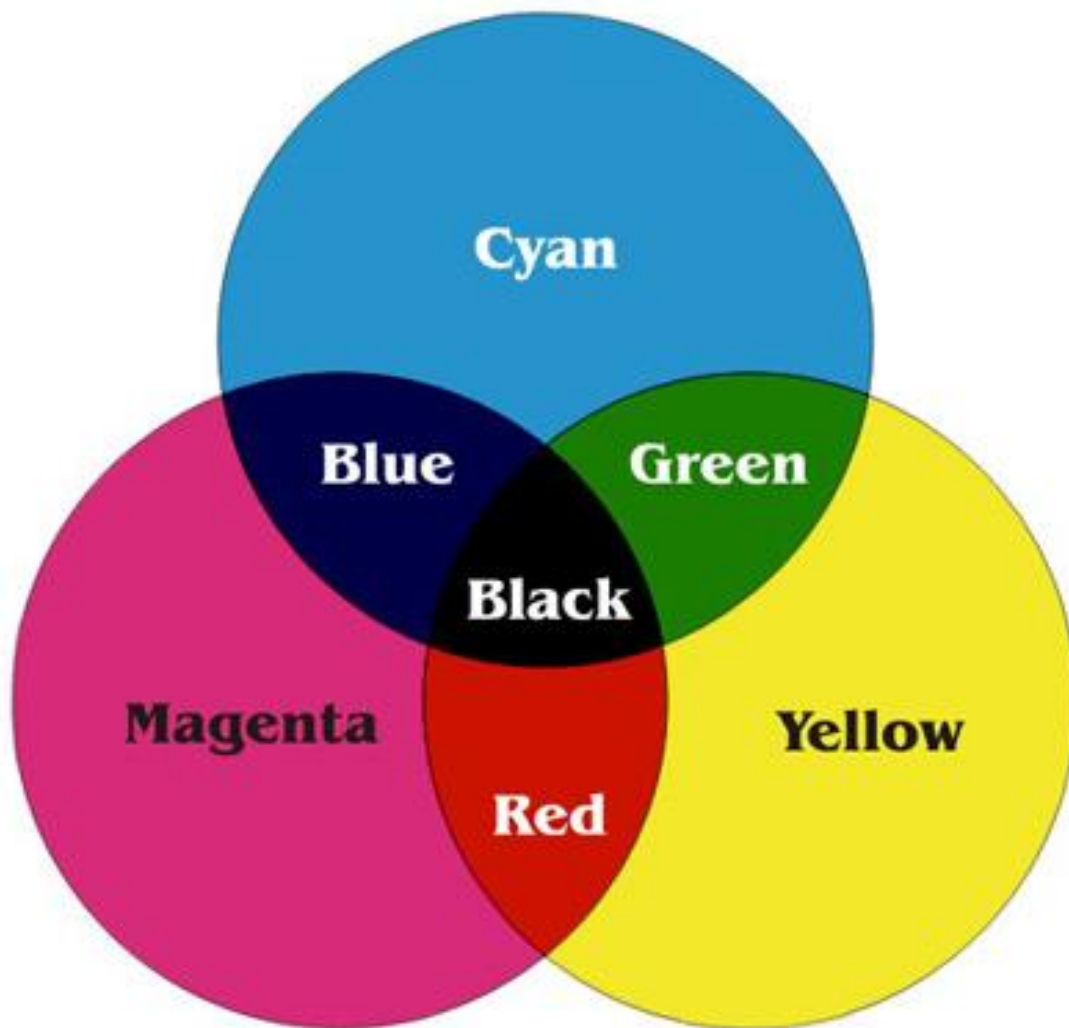
Προσεγγιστικά, για μετατροπές από το ένα μοντέλο στο άλλο έχουμε τις εξής απλές εξισώσεις:

$$R = 1 - C (1)$$

$$G = 1 - M (2)$$

$$B = 1 - Y (3)$$

Πρακτικά το CMYK μοντέλο χρησιμοποιείται περισσότερο στους εκτυπωτές, οι οποίοι κατασκευάζουν το κάθε χρώμα σαν μείξη των Κυανό, Πορφυρό, Κίτρινο.



Εικόνα 3: Το μοντέλο CMYK

3.3 Υφή

Η υφή είναι από τα βασικά χαρακτηριστικά της εικόνας, η οποία όμως δεν μπορεί να οριστεί με τρόπο κατάλληλο όπως η χρωματική απόδοση, Mezaris et al (2003), Stanchev (2003). Αρκετά από τα συστήματα ανάκτησης δεν λαμβάνουν υπόψη τους την υφή. Παρ' όλ' αυτά η υφή παρέχει σημαντικές πληροφορίες για την κατηγοριοποίηση της εικόνας καθώς περιγράφει το περιεχόμενο πολλών πραγματικών εικόνων. Τέτοια παραδείγματα είναι η φλούδα των φρούτων, τα σύννεφα, τα δέντρα, τα τούβλα και τα υφάσματα. Οπότε η υφή είναι ένα πολύ σημαντικό χαρακτηριστικό για τον ορισμό της υψηλού επιπέδου σημασιολογίας μιας εικόνας για την καλύτερη ανάκτησή της. Ο διαχωρισμός της υφής γίνεται σε δύο κατηγορίες, στην φυσική υφή και στην οπτική υφή. Με τον όρο φυσική υφή εννοούμε την απτές μεταβλητές ενός αντικειμένου.

Δηλαδή την αίσθηση που μας προκαλείται με το άγγιγμα της επιφάνειας ενός αντικειμένου. Παραδείγματα υπάρχουν πολλά, όπως η τραχύτητα της επιφάνειας ενός ξύλου και η αναγνώριση των στρώσεων του κορμού, η αντίληψη των κόκκων της άμμου κτλ. Με τον όρο οπτική υφή εννοούμε την παραισθητική αντίληψη της φυσικής υφής. Από μικρό παιδί ο άνθρωπος μαθαίνει να αναγνωρίζει αντικείμενα μόλις τα δει ανακτώντας ουσιαστικά την πληροφορία της υφής από τον εγκέφαλό του και αναγνωρίζοντάς την μόνο αφού την έχει αντιληφθεί έστω και μία φορά στην ζωή του. Κάθε υλικό και κάθε επιφάνεια έχει την δική της υφή.

3.4 Σχήμα

Για να πετύχουμε σε ένα σύστημα ανάκτηση εικόνας με βάση το σχήμα αυτό συνεπάγεται ότι έχουμε πρόσβαση σε σχήματα που έχουν εξαχθεί από πραγματικές εικόνες. Οπότε ουσιαστικά πρέπει πρώτα να αναλύσουμε τις εικόνες αφού τα πρωτογενή δεδομένα που είναι διαθέσιμα είναι τα ίδια τα πολυμέσα που θέλουμε να ανακτήσουμε. Τα σχήματα συνήθως είναι τα περιγράμματα των αντικειμένων που εμφανίζονται στις εικόνες. Εδώ πρέπει να προσέξουμε, γιατί θέλουμε να υπάρχει ανεξαρτησία της μεθόδου ανάκτησης από τα ίδια τα δεδομένα ανάκτησης.

Το σχήμα αποτελεί ένα άλλο βασικό γνώρισμα για την ανάκτηση εικόνας με βάση το περιεχόμενο. Τα φυσικά αντικείμενα αναγνωρίζονται πέρα από το χρώμα τους και από το σχήμα τους. Ένας αριθμός χαρακτηριστικών γνωρισμάτων των σχημάτων (ανεξάρτητα από το μέγεθος ή τον προσανατολισμό) των αντικειμένων της εικόνας εξάγονται για κάθε αποθηκευμένη εικόνα από το σύστημα. Όταν διεξάγεται μια αναζήτηση το σύστημα ανακαλεί αυτές τις εικόνες που τα χαρακτηριστικά τους ταιριάζουν περισσότερο με αυτά του ερωτήματος υπολογίζοντας τα χαρακτηριστικά της εικόνας που δίνεται στο ερώτημα. Γενικά οι σχηματικές αναπαραστάσεις διαιρούνται σε δύο κατηγορίες, σ' αυτές που βασίζονται στα όρια του σχήματος και σ' αυτές που βασίζονται στην περιοχή. Οι πρώτες χρησιμοποιούν μόνο τα εξωτερικά όρια του σχήματος, το περίγραμμά του ενώ οι δεύτερες χρησιμοποιούν τη συνολική περιοχή του σχήματος. Οι ερωτήσεις κατά την αναζήτηση γίνονται είτε με τη δήλωση μιας εικόνας (ή κάποιου σχήματος αυτής) ως υπόδειγμα είτε με τη σχεδίαση ενός σχήματος από το χρήστη.

Παρακάτω παρουσιάζονται μερικά από τα πιο γνωστά συστήματα Ανάκτησης Εικόνας και Βίντεο που εκμεταλλεύονται Τεχνικές Σύγκρισης Σχημάτων:

1. Το σύστημα *Query by Image Content* της IBM παρέχει δυνατότητες ανάκτησης εικόνας μέσω ενός συνόλου κριτηρίων επιλογής που βασίζονται στο χρώμα, την υφή, τα σχετική θέση και το σχήμα αντικειμένων. Τα πολυμέσα ανακτώνται με βάση σχήματα που είτε έχουν εξαχθεί από άλλη εικόνα είτε σχεδιάζονται από το χρήστη του συστήματος.
2. Το σύστημα *VIR*, της *Virage*, χρησιμοποιεί πληροφορίες σχήματος για την ανάκτηση εικόνων.
3. Το σύστημα *Excalibur Visual Retrieval Ware SDK* (παρέχει δυνατότητες ανάκτησης εικόνας με βάση διάφορα χαρακτηριστικά, ανάμεσα στα οποία και πληροφορίες σχήματος).

3.5 Θέση

Ένα ακόμη χαρακτηριστικό για την ανάκτηση βασισμένη στο περιεχόμενο αποτελεί η τοποθεσία ή η θέση που βρίσκεται το αντικείμενο στην εικόνα. Η θέση του αντικειμένου στο χώρο της εικόνας καθώς και η χωρική του σχέση με κάποιο ή κάποια άλλα αντικείμενα της εικόνας. Εικόνες που έχουν όμοια αντικείμενα στην ίδια θέση ή συνδυασμό δύο ή περισσότερων όμοιων αντικειμένων με τη διάταξη (π.χ. το ένα δεξιά από το άλλο) που ανταποκρίνονται στο ερώτημα ανακτώνται.

3.6 Τεχνικές σύγκρισης ιστογραμμάτων

Αντί του ακριβούς ταιριάσματος, η ανάκτηση εικόνας βασισμένη στο περιεχόμενο (CBIR), υπολογίζει οπτικές ομοιότητες μεταξύ της εικόνας ερώτησης και των εικόνων σε μια βάση δεδομένων. Συνεπώς, το αποτέλεσμα ανάκτησης είναι όχι μια ενιαία εικόνα αλλά ένας κατάλογος εικόνων που ταξινομούνται σύμφωνα με το βαθμό ομοιότητάς τους με την εικόνα ερώτησης. Πολλά μέτρα ομοιότητας έχουν αναπτυχθεί για την ανάκτηση εικόνας τα τελευταία χρόνια, βασισμένα σε εμπειρικές εκτιμήσεις της κατανομής των χαρακτηριστικών γνωρισμάτων της εικόνας.

Διαφορετικά μέτρα ομοιότητας/απόστασης έχουν επιπτώσεις στην απόδοση τέτοιων συστημάτων. Σε αυτό το τμήμα, θα εισάγουμε τα πιο ευρέως χρησιμοποιούμενα μέτρα ομοιότητας.

Για τη μέτρηση της ομοιότητας των χρωματικών ιστογραμμάτων, δυο αποστάσεις $L1$ και $L2$ η αλλιώς Τετραγωνική και Ευκλείδεια απόσταση αντίστοιχα είναι οι πιο συχνά χρησιμοποιούμενες. Οι αποστάσεις αυτές μεταξύ του ιστογράμματος της εικόνας ερώτησης $HQ(I)$ και αυτού της εικόνας στόχου $HC(I)$ είναι:

$$D_H(I_Q, I_C) = \sum_{j=1}^n |H_Q(j) - H_C(j)|$$

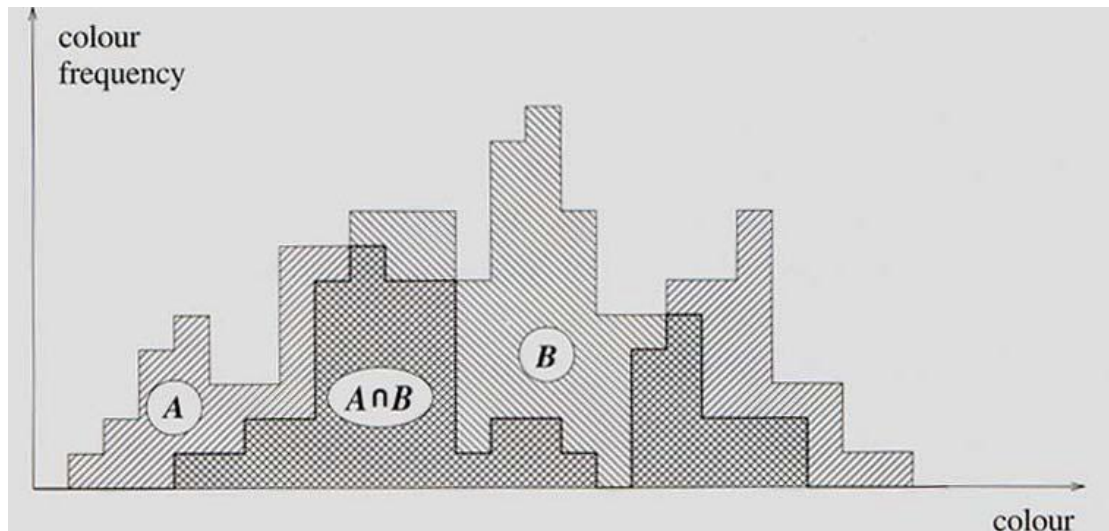
$$D_H(I_Q, I_C) = \left(\sum_{j=1}^n (H_Q(j) - H_C(j))^2 \right)^{\frac{1}{2}}$$

όπου j είναι ο δείκτης του bin στο ιστόγραμμα. Όμοιες εικόνες θεωρούνται αυτές που ελαχιστοποιούν τις αποστάσεις αυτές. Η αποτελεσματικότητα των δυο αυτών μετρικών δεν είναι και τόσο ικανοποιητική.

Ο Swain και ο Ballard πρότειναν τη μέθοδο παρεμβολής ιστογραμμάτων για μεγάλες βάσεις δεδομένων. Η τεχνική αυτή είναι πολύ ισχυρή όσον αφορά αλλαγές στην ανάλυση, στο μέγεθος ιστογραμμάτων, στο βάθος και στον ενδιάμεσο άξονα- medial axis της εικόνας. Η αναλογία ομοιότητας που ανήκει στο διάστημα [0 έως 1] συγκρίνεται με ένα δεδομένο κατώτατο όριο. Μπορεί να εκφραστεί μέσω της ακόλουθης εξίσωσης:

$$H(H_Q, H_C) = \frac{\sum_i \min(H_Q(i), H_C(i))}{\min\left(\sum_i H_Q(i), \sum_i H_C(i)\right)}$$

Χρώματα τα οποία δεν είναι παρόντα στην εικόνα ερώτηση δε συνεισφέρουν στον υπολογισμό της απόστασης μέσω της παρεμβολής, όπως φαίνεται και στην επόμενη εικόνα.



Εικόνα 4: Παρεμβολή δύο ιστογραμμάτων A και B

Παρά το γεγονός ότι η μέθοδος της παρεμβολής ιστογραμμάτων προτάθηκε το 1991, παραμένει ακόμα και σήμερα μία από τις καλύτερες τεχνικές για σύγκριση ιστογραμμάτων σε εφαρμογές ανάκτησης εικόνας. Σε αυτή την τεχνική έχουν βασισθεί οι περισσότερες μεταγενέστερες. Παραδείγματος χάριν, ασαφείς μέθοδοι έχουν προταθεί για την αναπαράσταση της απόστασης των χρωμάτων στο ανθρώπινο μάτι. Οι ασαφείς μέθοδοι παρουσιάζουν πολύ μεγάλη ακρίβεια, αλλά η επεξεργασία του ιστογράμματος παρουσιάζει πολυπλοκότητα $O(n^2)$, με n τα χρώματα που εμφανίζονται. Η απαιτούμενη υπολογιστική ισχύς τα καθιστά μη πρακτικά. Επίσης συζευγμένα ιστογράμματα εμπεριέχουν και κάποια χωρική πληροφορία, σε στατιστική πάλι μορφή.

Η τεχνική *Divergence Factor* μετρά πόσο συμπαγής είναι η κατανομή χρώματος του ιστογράμματος ερώτησης, σε σύγκριση με αυτά των εικόνων στη βάση δεδομένων.

$$D(H_Q, H_C) = \sum_i \left[(H_Q(i) - H_C(i)) \ln \frac{H_Q(i)}{H_C(i)} \right]$$

Η απόσταση *Matusita* είναι ένα μέτρο διαχωρισμού που παρέχει αξιόπιστα αποτελέσματα. Εκφράζεται από την παρακάτω εξίσωση:

$$M(H_Q, H_C) = \sqrt{\sum_i (\sqrt{H_Q(i)} - \sqrt{H_C(i)})^2}$$

Η απόσταση *Bhattacharya* μετρά τη στατιστική διαχωριστικότητα των φασματικών κατηγοριών δίνοντας μια εκτίμηση της πιθανότητας σωστής ταξινόμησης. Αυτή η απόσταση δε λαμβάνει υπόψη τμήματα ιστογραμμάτων με μηδενικές τιμές. Για τα ιδιαίτερα δομημένα ιστογράμματα (δηλ. εκείνα, τα οποία δεν είναι ομοιόμορφα καταναμημένα), αυτό μπορεί να οδηγήσει στην επιλογή των αντιστοιχιών στην οποία υπάρχει μια ισχυρή συνταύτιση μεταξύ της δομής του ιστογράμματος της εικόνας ερώτησης και του ιστογράμματος της εικόνας της βάσης δεδομένων.

$$B(H_Q, H_C) = -\ln \sum_i \sqrt{H_Q(i) \times H_C(i)}$$

3.7 Τεχνικές ταιριάσματος εικόνων

Είναι πολύ σημαντικό να έχουμε τόσο διακριτούς περιγραφείς των οπτικών χαρακτηριστικών της εικόνας, όπως επίσης να υπάρχει και κάποια μέθοδος για τη μέτρηση της ομοιότητας μεταξύ δύο εικόνων. Μια εκτενής περιγραφή των μεθόδων για τον υπολογισμό της ομοιότητας των χαρακτηριστικών των εικόνων μπορεί να βρεθεί στην εργασία (Jolion, 2001).

Όταν ο χώρος των χαρακτηριστικών είναι διανυσματικός, τότε το πιο αποτελεσματικό κριτήριο για την ομοιότητα των εικόνων είναι η Ευκλείδεια απόσταση. Βέβαια, δε μπορεί να θεωρηθεί ότι η Ευκλείδεια απόσταση είναι η αποτελεσματικότερη σε κάθε περίπτωση, αλλά όπως έχει αποδειχθεί από μελέτες το μέτρο απόσταση πρέπει να γίνεται πάντα σε σχέση με τη συλλογή εικόνων που θέλουμε να ερευνήσουμε και με τη φύση των διανυσμάτων των περιγραφέων που συγκρίνουμε (Sebe et al, 2000).

Για τον υπολογισμό της απόστασης ιστογραμμάτων υπάρχουν διάφορες μέθοδοι, η πιο ευρέως διαδεδομένη από τις οποίες είναι η απόσταση Minkowski που συγκρίνει την απόσταση των σχετικών κορυφών των ιστογραμμάτων. Ο πιο συνηθισμένος, όμως, τρόπος για το χειρισμό των ιστογραμμάτων είναι η μετατροπή τους σε διανύσματα. Το πρόβλημα της μεθόδου αυτής είναι ότι χάνεται η πληροφορία της θέσης από την οποία εξήχθησαν οι κορυφές του ιστογράμματος (Rubner et al, 1998). Μια βελτίωση αυτής της τεχνικής που λύνει αυτό το πρόβλημα έχει δοθεί από την εργασία (Rubner et al, 2000). Μια άλλη πολύ ευρέως διαδεδομένη απόσταση είναι η απόσταση Mahalanobis, που έχει χρησιμοποιηθεί κυρίως σε συστήματα ανάκτησης κατά περιοχές (Carson et al, 2002).

3.8 Τεχνικές δεικτοδότησης

Στα πρώιμα συστήματα ανάκτησης εικόνων, οι πληροφορίες που εξάγονταν από τις εικόνες αποθηκεύονταν είτε σε κάποιο αρχείο είτε σε κάποια βάση δεδομένων. Καθώς, όμως, το μέγεθος της προς αποθήκευση πληροφορίας μεγάλωνε, η απόδοση των δύο αυτών μεθόδων έπεφτε από τη σκοπιά της υπολογιστικής αποτελεσματικότητας. Ένα ακόμα πρόβλημα ήρθε να προστεθεί στο ήδη υπάρχων και είναι το γεγονός ότι η διάσταση των διανυσμάτων χαρακτηριστικών άρχισε να αυξάνει δραματικά. Οπότε, ήταν επιτακτική πλέον η ανάγκη για χρήση κάποιας πιο έξυπνης μορφής δεικτοδότησης ή ομαδοποίησης της πληροφορίας.

Μερικές από τις μεθόδους που έχουν χρησιμοποιηθεί κατά καιρούς είναι οι εξής:

- Η εφαρμογή κάποιας τεχνικής ελάττωσης των διανυσμάτων, όπως για παράδειγμα με την τεχνική *Principal Component Analysis* (Ng et al, 1996).
- Η χρησιμοποίηση κάποιου ευρετηρίου για την επιτάχυνση της αναζήτησης.
- Κάποια δενδρική δομή, όπως τα *k-d δένδρα* ή τα *R-δένδρα* (Scott et al, 2003).
- Ο κβαντισμός των διανυσμάτων (Ye et al, 2003).

- Η χρησιμοποίηση κάποιας τεχνικής κατακερματισμού για την απεικόνιση των διανυσμάτων σε άλλους διανυσματικούς χώρους μικρότερης διάστασης (Greene et al, 1994).

3.9 Τεχνικές με χρήση οπτικού λεξικού

Όταν μας ενδιαφέρει η εξαγωγή χαρακτηριστικών από κάποια περιοχή των εικόνων, τότε ο όγκος της εξαγόμενης πληροφορίας αυξάνει δραματικά. Οπότε ακόμα και οι πιο αποδοτικοί τρόποι δεικτοδότησης θα είναι ιδιαίτερα χρονοβόροι από υπολογιστικής απόψεως. Για τη λύση αυτού του προβλήματος έχει προταθεί η συσταδοποίηση των διανυσμάτων σε συστάδες. Κάθε συστάδα έχει ως αντιπροσωπευτικό στοιχείο το κέντρο της.

Μεταξύ άλλων, η πιο ευρέως διαδεδομένη μέθοδος συσταδοποίησης είναι η *k-means*. Το χαρακτηριστικό της μεθόδου αυτής είναι ότι ο αριθμός των συστάδων πρέπει να είναι γνωστός εκ των προτέρων. Κάποιες άλλες μέθοδοι που έχουν χρησιμοποιηθεί είναι οι:

- *Kernel mapping*
- *Hierarchical mapping*
- *Metric learning*

Όταν τα διανύσματα των χαρακτηριστικών πρόκειται να προέλθουν από περιοχές γύρω από σημεία ενδιαφέροντος της εικόνας, τότε αν εφαρμοστεί κάποια μέθοδος συσταδοποίησης, τότε θα παραχθεί ένα οπτικό λεξικό μικρών περιοχών. Κάθε μία συστάδα θα εκφράζει μια περιοχή ως προς τις κατευθύνσεις των φωτεινότητων μέσα σε αυτή.

ΚΕΦΑΛΑΙΟ 4: Το πρότυπο MPEG7

Στο κεφάλαιο αυτό παρουσιάζεται αναλυτικά το πρότυπο MPEG7.

4.1 Εισαγωγή

Το MPEG-7 (*Multimedia Content Description Interface*) αποτελεί ένα επίσημο πρότυπο από τον Φεβρουάριο του 2002 και είναι το πρότυπο που προβλέπεται από το MPEG (*Moving Pictures Expert Group*) για τη διαχείριση και την περιγραφή της πολυμεσικής πληροφορίας.

Το πρότυπο αυτό περιλαμβάνει πολλές μορφές πολυμέσων, όπως κείμενο, σταθερές εικόνες, γραφικά, 3D μοντέλα, ήχο, ομιλία, video και συνδυασμούς όλων των προηγούμενων μορφότυπων. Οι γενικοί στόχοι που καλείται να εξυπηρετήσει το πρότυπο είναι δύο:

1. Να γίνει ευκολότερος ο στόχος της ανάκτησης της πολυμεσικής πληροφορίας και
2. Η επίτευξη της διαλειτουργικότητας μεταξύ των εφαρμογών και του λογισμικού που χρησιμοποιούν οι εφαρμογές, και όλα αυτά να πραγματοποιηθούν με την χρήση των μεταδεδομένων.

Το MPEG είναι μία ομάδα του Διεθνούς Οργανισμού Προτυποποίησης (*ISO, International Organization for Standardization*), που σκοπός της είναι η ανάπτυξη προτύπων για συμπίεση, επεξεργασία, κωδικοποίηση και διαχείριση εικόνων, ήχου και των συνδυασμών τους, που καθιερώνονται διεθνώς. Πολύ γνωστά πρότυπα που έχουν δημιουργεί είναι τα MPEG-1, MPEG-2, MPEG-3 (το γνωστό σε όλους ως mp3), MPEG-4 (αντίστοιχα το γνωστό ως mp4), MPEG-7 και MPEG-21.

4.1.1 MPEG-1

Το MPEG-1 αφορά την συμπίεση ήχου και εικόνας. Μέρος του προτύπου είναι και το MPEG-1 Audio που αφορά αποκλειστικά τον ήχο. Επίσης, το MPEG-1 είναι το πρώτο διεθνές πρότυπο για την ψηφιακή συμπίεση ήχου υψηλής πιστότητας. Στο πρότυπο αυτό ανήκει και το παγκοσμίως γνωστό MP3 (MPEG - Layer 3), που είναι το αποτέλεσμα της αναζήτησης για μεγαλύτερη συμπίεση και ταυτόχρονα καλή ποιότητα ήχου.

4.1.2 MPEG-2

Το MPEG-2 είναι το πρότυπο για την ψηφιακή τηλεόραση και αποτελεί την συνέχεια του MPEG-1 ως προς τις επιπλέον δυνατότητες που παρέχει. Σαν μία συμβατική συνέχεια λοιπόν του MPEG-1 υποστηρίζει μορφές video και έναν αριθμό άλλων προηγμένων χαρακτηριστικών, όπως τέτοιων που να μπορούν να υποστηρίξουν τον τύπο HDTV (High Definition TV). Επίσης, το πρότυπο αυτό αποτέλεσε και τη βάση για το DVD Video (Digital Versatile Disc). Το πρότυπο MPEG-2 παρέχει επίσης ένα σύστημα με ορισμούς για το πώς το video, ο ήχος και άλλα δεδομένα συνδυάζονται σε ένα ή πολλαπλά ρεύματα (stream) που να είναι κατάλληλα για την αποθήκευση και μετάδοσή τους.

4.1.3 MPEG-4

Το πρότυπο MPEG-4 αφορά την παραγωγή, κατανομή και πρόσβαση του πολυμεσικού περιεχομένου. Παρόλο που τα MPEG-1 και MPEG-2 κάλυπταν ένα αρκετά μεγάλο εύρος εφαρμογών, όπως η ψηφιακή τηλεόραση και η διαδραστική κινούμενη εικόνα, σύντομα έγινε αντιληπτό ότι οι εφαρμογές που έχουν σχέση με τα πολυμέσα έχουν περισσότερες απαιτήσεις. Αναπτύχθηκε λοιπόν το πρότυπο MPEG-4 με σκοπό να παρέχει ένα προτυποποιημένο σύνολο τεχνολογιών, που θα καταστήσει ικανή την ενοποίηση της παραγωγής, κατανομής, τα παραδείγματα για την πρόσβαση στο περιεχόμενο και τα διαδραστικά πολυμέσα για την ανάπτυξη της χρήσης των εφαρμογών πολυμέσων στο Internet.

Θα πρέπει επίσης να σημειώσουμε ότι το πρότυπο MPEG-4 επιτρέπει την επισύναψη μετά-δεδομένων σχετικά με το περιεχόμενο επάνω στο αντικείμενο. Επομένως ο χρήστης μπορεί να χρησιμοποιήσει αυτή τη ροή δεδομένων με την πληροφορία του περιεχομένου του αντικειμένου OCI (Object Content Information) για να στείλει τα μεταδεδομένα ταυτόχρονα με το περιεχόμενο του MPEG-4. Ωστόσο στο συγκεκριμένο πρότυπο δεν υπάρχει κάποια τυποποιημένη δομή και καθορισμένη μορφή που να αφορά τα μεταδεδομένα.

4.1.4 MPEG-7

Ενώ τα προηγούμενα πρότυπα του MPEG είχαν ως στόχο την καλύτερη παρουσίαση και αναπαράσταση της πληροφορίας, ο στόχος του προτύπου MPEG-7 είναι να παρέχει ένα προτυποποιημένο περιβάλλον για την περιγραφή της πληροφορίας αυτής. Στόχος δηλαδή είναι η πλήρης περιγραφή της πολυμεσικής πληροφορίας, τα μετα-δεδομένα και όχι το περιεχόμενό της. Τα μετα-δεδομένα είναι απαραίτητα γιατί διευκολύνουν την ανταλλαγή, την αναζήτηση, την ανάκτηση και το φιλτράρισμα των πληροφοριών που υπάρχουν στα πολυμεσικά τεκμήρια.

Το κίνητρο για την ανάπτυξη του MPEG-7 δόθηκε από τις ξεκάθαρες πλέον τάσεις στην τεχνολογία, στην αγορά και στους χρήστες. Με την εξάπλωση των εργαλείων για την δημιουργία οπτικού και ακουστικού περιεχομένου και την ψηφιοποίησή του και με την ευρεία χρήση των πολυμέσων σε ολοένα και περισσότερους τομείς δημιουργήθηκε ένα περιβάλλον, όπου παράγεται όλο και περισσότερη πληροφορία οπτικοακουστικού περιεχομένου. Η ίδια η φύση αυτής της πληροφορίας υπαγορεύει νέους τρόπους ως προς την επεξεργασία, την οργάνωση και την αξιοποίηση της.

Όπως είναι, λοιπόν, φυσιολογικό αναπτύσσονται νέα συστήματα και εργαλεία που να καλύπτουν τις ανάγκες για συλλογή, διαχείριση, ευρετηρίαση και αναζήτηση πολυμεσικού περιεχομένου. Εκτός από τις ανάγκες των χρηστών, τέτοιες λύσεις πρέπει να δοθούν και για τους διάφορους παραγωγούς των υπηρεσιών και για τους δημιουργούς του περιεχομένου. Γι' αυτό λοιπόν και το πρότυπο MPEG-7 προκειμένου να επιτύχει τη μέγιστη δια-λειτουργικότητα και να επιτρέψει την δημιουργία καινοτόμων

εργαλείων και εφαρμογών, παρέχει ένα διαλειτουργικό περιβάλλον, στο οποίο ορίζεται η δομή και η σημασιολογία ποικίλων εργαλείων περιγραφής.

Κάθε ένα από αυτά τα εργαλεία περιγραφής μπορεί να έχει σχεδιαστεί για κάποια εξειδικευμένη ή γενική μορφή πληροφορίας (π.χ. για ήχο, για εικόνα ή για πολυμέσα), πλευράς (π.χ. δομική, σημασιολογική) και εφαρμογής (π.χ. για μηχανή αναζήτησης ή για πλοήγηση). Ως απάντηση στις ανάγκες αυτές έχουν πραγματοποιηθεί από διάφορες ομάδες και οργανισμούς αρκετές δραστηριότητες για τον καθορισμό διαλειτουργικών πλαισίων και για τον καθορισμό της αναπαράστασης της περιγραφής των μετα-δεδομένων.

Μέσα σε αυτό το πλαίσιο λοιπόν αναπτύχθηκε το πρότυπο MPEG-7. Το MPEG-7 δε στοχεύει σε κάποια συγκεκριμένη εφαρμογή, αλλά θα μπορούσε να διατυπωθεί ότι τα στοιχεία που προτυποποιεί υποστηρίζουν μία ευρεία σειρά εφαρμογών. Αυτή είναι επίσης και η διαφορά του σε σχέση με κάποια άλλα πρότυπα που αφορούν τα μετα-δεδομένα, ότι δηλαδή στοχεύει στο να είναι ένα γενικό πρότυπο και όχι στο να χρησιμοποιείται σε μία και μόνο εφαρμογή ή σε ένα μόνο πεδίο δραστηριοτήτων.

Όπως προαναφέρθηκε στόχος του προτύπου MPEG-7 είναι να προάγει την διαλειτουργικότητα ανάμεσα στα συστήματα και στις εφαρμογές που χρησιμοποιούνται για την δημιουργία, διαχείριση, κατανομή και χρήση των περιγραφών του οπτικοακουστικού περιεχομένου. Αυτές οι περιγραφές έχουν ως στόχο να βοηθήσουν το χρήστη και τις εφαρμογές στον προσδιορισμό της πληροφορίας, στην ανάκτηση και στο να φιλτράρουν αυτήν την πληροφορία. Το πρότυπο MPEG-7 μπορεί να χρησιμοποιηθεί τόσο σε επιστημονικές όσο και σε απλές εφαρμογές που αφορούν έναν καταναλωτή, όπως:

- Στην εκπαίδευση
- Στην δημοσιογραφία (στην αναζήτηση κάποιας ομιλίας ενός πολιτικού, χρησιμοποιώντας το όνομά του ή την φωνή του)
- Σε πολιτισμικές υπηρεσίες (όπως σε ένα μουσείο, σε μία αίθουσα τέχνης)
- Στην ψυχαγωγία (για διάφορα παιχνίδια)
- Σε συστήματα Γεωγραφικής Πληροφόρησης

- Σε βιοϊατρικές εφαρμογές
- Σε ερευνητικές υπηρεσίες (για την αναγνώριση ανθρώπινων χαρακτηριστικών)
- Στην αρχιτεκτονική και στην διακόσμηση εσωτερικών χώρων
- Για πληροφορίες σε σχέση με τον τουρισμό
- Σε κοινωνικές εφαρμογές και
- Σε αρχεία video, μιας ταινίας ή ενός ραδιοφωνικού σταθμού.

Παραδείγματα εφαρμογών αποτελούν οι ψηφιακές βιβλιοθήκες (ένας κατάλογος χρηστών με βάση την εικόνα τους), υπηρεσίες του χρυσού οδηγού με την χρήση πολυμέσων, επεξεργασία με την χρήση πολυμέσων (για παράδειγμα υπηρεσίες για εξατομίκευση των ειδήσεων σε ηλεκτρονική μορφή). Δεδομένης όλης αυτής της ποικιλίας των εφαρμογών, οι περιγραφές του MPEG-7 δίνουν την δυνατότητα στους χρήστες ή στις εφαρμογές να εκτελέσουν τις παρακάτω εργασίες που αναφέρονται παρακάτω.

Στην κατηγορία των πολυμέσων βοηθούν στην δημιουργία ενός εξειδικευμένου προγράμματος ενός οδηγού ή μιας περίληψης από κάποιο τηλεοπτικό οπτικό ή/ και ακουστικό περιεχόμενο σύμφωνα με τις προτιμήσεις του εκάστοτε χρήστη. Ακόμη βοηθούν στην δημιουργία μιας βάσης δεδομένων εικόνων όπου μπορεί να γίνει η ανάκτηση της πολυμεσικής πληροφορίας.

Στα αρχεία δημιουργούν περιγραφές από κάποια μεμονωμένα στοιχεία μιας συλλογής ή ακόμα και από ολόκληρη την συλλογή από οπτικοακουστικό περιεχόμενο και γίνεται ανταλλαγή των περιγραφών αυτών με τους ιδιοκτήτες του περιεχομένου, τους οργανισμούς ή τους καταναλωτές.

Στην προσαρμοστικότητα, φιλτράρουν και μετατρέπουν τις ροές των πολυμέσων σε περιβάλλοντα περιορισμένων πόρων ταιριάζοντας τις προτιμήσεις των χρηστών, τους διαθέσιμους πόρους και τις περιγραφές του περιεχομένου.

Στον ήχο, δίνουν την δυνατότητα να παίζουμε μερικές νότες και να μας επιστραφεί μία λίστα από μουσικά κομμάτια που να περιλαμβάνουν αυτόν τον ρυθμό, συνδυάζοντας κάπως τις νότες.

Στα γραφικά μπορούμε να σχεδιάσουμε μερικές γραμμές στην οθόνη και να μας επιστρέψει ένα σύνολο από εικόνες, που να περιέχουν παρόμοια σχέδια ή εικόνες, δηλαδή ανάκτηση μέσω κάποιου σκίτσου.

Στην κίνηση, έχοντας ένα σύνολο από αντικείμενα video, μπορούμε να περιγράψουμε κινήσεις και σχέσεις ανάμεσα στα αντικείμενα και να μας επιστραφεί μία λίστα από video clips που να εκπληρώνουν τις χωρικές και χρονικές σχέσεις.

Στο σενάριο μπορούμε σε ένα οπτικοακουστικό περιεχόμενο που μας δίνεται να περιγράψουμε τις ενέργειες και να λάβουμε μία λίστα από σενάρια, στα οποία θα περιλαμβάνονται παρόμοιες πράξεις.

Οι εφαρμογές του προτύπου MPEG-7 που προβλέπονται από το ISO, καθώς και οι προδιαγραφές για την εκπλήρωσή τους είναι ποικίλες και καλύπτουν πολλούς τομείς. Για μία καλύτερη κατανόηση του τι δυνατότητες παρέχει το πρότυπο MPEG-7 και ποιες είναι οι λειτουργίες του, θα παρατεθούν κάποιες εφαρμογές που έχουν σχέση με την εικόνα. Αυτό δεν δηλώνει κάποια προτεραιότητα ως προς τις εφαρμογές, απλά εκπληρώνει τους στόχους αυτής της διπλωματικής εργασίας.

4.2 MPEG7 και βάσεις δεδομένων

Σε ότι αφορά λοιπόν το οπτικοακουστικό περιεχόμενο, το πρότυπο MPEG-7 ξεκίνησε σαν ένα πρότυπο που θα μπορούσε να καταστήσει το περιεχόμενο αυτό εύκολο στην αναζήτηση, όπως συμβαίνει και στην απλή πληροφορία που περιέχει κείμενο. Πράγματι μπορεί να γίνει εύκολα αντιληπτό, ότι η δομή και η αυστηρότητα που απαιτείται για την περιγραφή ακόμα και της ελάχιστης πολυμεσικής πληροφορίας, είναι αρκετά πιο δύσκολη σε σχέση με ένα απλό, γραπτό κείμενο.

Οι μέθοδοι ανάκτησης επομένως πρέπει να περιλαμβάνουν βάσεις δεδομένων, οπτικοακουστικά αρχεία και τις υπηρεσίες του Internet, που βασίζονται στον Παγκόσμιο Ιστό (παράδειγμα, αναζήτηση μιας εικόνας από μία μηχανή αναζήτησης). Τυπικό παράδειγμα αυτής της κατηγορίας αποτελούν τα αρχεία από ταινίες και

διάφορες τηλεοπτικές εκπομπές, που αποθηκεύουν τεράστιες ποσότητες πολυμεσικού περιεχομένου σε διάφορα μέσα (CD-ROM, DVD κ.τ.λ.) μαζί με ακριβή περιγραφική πληροφορία (μεταδεδομένα). Αυτά τα μετα-δεδομένα βρίσκονται αποθηκευμένα σε βάσεις δεδομένων, όπου υπάρχει ένα αρκετά μεγάλο ενδιαφέρον για την αποθήκευση και ανταλλαγή των περιγραφών που θα μπορούσαν να διασφαλίσουν:

1. Την διαλειτουργικότητα ανάμεσα σε χειριστές αρχείων video
2. Την ουσιαστική σημασία των μετα-δεδομένων και
3. Μία ευρύτερη εξάπλωση των δεδομένων τόσο στο επιστημονικό όσο και στο ευρύ κοινό.

Οι προδιαγραφές του προτύπου του MPEG-7 για τέτοιου είδους εφαρμογές περιλαμβάνουν:

- Περιγραφές πλήρους κειμένου και δομημένα πεδία (περιγραφές της βάσης δεδομένων).
- Έναν μηχανισμό με τον οποίο διαφορετικές περιγραφές του MPEG-7 θα μπορούν να υποστηρίξουν την ικανότητα για δια-λειτουργικότητα ανάμεσα σε διαφορετικές σημασιολογίες
- Έναν αυτόνομο μηχανισμό
- Μία δομή για να χειρίζεται πολλαπλές εκδόσεις του ίδιου τεκμηρίου κατά τα διάφορα στάδια της παραγωγής του και περιγραφές που να μπορούν να εφαρμοστούν σε πολλαπλά αντίγραφα του ίδιου υλικού.

Μία παρόμοια κατάσταση συναντάμε και στις βάσεις δεδομένων που αποθηκεύουν ήχο. Η μουσική βιομηχανία προσπαθεί να προσεγγίσει όσο το δυνατό περισσότερους πελάτες, από τους οποίους ο καθένας μπορεί να έχει ποικίλες μουσικές προτιμήσεις. Οι ερευνητές-καταναλωτές μπορεί να αναζητούν ένα τραγούδι βασισμένοι σε μία σχετική απόδοση του τραγουδιού αυτού από ένα κατάστημα ή από το σπίτι τους.

Εναλλακτικά θα μπορούσαν να ζητήσουν ένα μουσικό κομμάτι με χαρακτηριστικά που ήδη γνωρίζουν (όπως για παράδειγμα το όνομα του τραγουδιστή)

και από εκείνο το σημείο μπορούν να ακούσουν το κατάλληλο δείγμα (και ίσως να δουν σε αυτό συνδεδεμένη πληροφορία, όπως τους στίχους ή το video clip) και να το αγοράσουν.

Οι εξειδικευμένες απαιτήσεις για τέτοιου είδους εφαρμογές, που έχουν σχέση με τον ήχο είναι οι εξής:

- Ένας μηχανισμός που να μπορεί να υποστηρίξει μελωδία και άλλα μουσικά χαρακτηριστικά, που θα επιτρέπουν το περιθώριο λάθους, για κάποιον που θα κάνει αναζήτηση ενός μουσικού τραγουδιού με βάση ένα απλό μουρμουρητό.
- Έναν μηχανισμό που θα υποστηρίξει τους περιγραφείς που βασίζονται στην πληροφορία που είναι συνδεδεμένη με τα δεδομένα
- Θα πρέπει να υποστηρίζονται σχήματα περιγραφής που να περιέχουν περιγραφείς οπτικών, ηχητικών και άλλων χαρακτηριστικών, καθώς και να παρέχει συνδέσμους ανάμεσα στα διάφορα μέσα.
- Εκτός από την μουσική βιομηχανία, ενδιαφέρουσες εφαρμογές που έχουν σχέση με τον ήχο είναι βάσεις δεδομένων που περιέχουν ομιλίες ιστορικών προσώπων.

Οι απαιτήσεις που προβλέπονται από το MPEG-7 και που απορρέουν από τις παραπάνω εφαρμογές είναι οι εξής:

- Μηχανισμούς με τους οποίους οι περιγραφείς μπορούν να συνδέονται με εξωτερική πληροφορία, όπως οι υπηρεσίες του World Wide Web και η HTML (Hypertext Markup Language).
- Υποστήριξη για δια-λειτουργικότητα μεταξύ των περιγραφέων.
- Η ικανότητα να επιτρέπονται λειτουργίες σε πραγματικό χρόνο σε συνδυασμό με τις βάσεις δεδομένων.

- Δείκτες που να αναφέρονται απευθείας στα δεδομένα, ώστε να δίνεται η δυνατότητα στους χρήστες να επεξεργάζονται τα πολυμέσα.

4.3 Περιγραφή του MPEG7

Οι στόχοι που καλύπτει και προβλέπεται να καλύψει το πρότυπο MPEG-7 είναι τρεις:

- Την περιγραφή του πολυμεσικού περιεχομένου.
- Την ευέλικτη διαχείριση των δεδομένων.
- Την δια-λειτουργικότητα των συστημάτων

Ο πιο σημαντικός στόχος του προτύπου MPEG-7 (ο ουσιαστικός λόγος για την ανάπτυξή του) είναι να παρέχει ένα σύνολο με προτυποποιημένες μεθόδους και εργαλεία για τις διάφορες κατηγορίες του πολυμεσικού περιεχομένου. Εκτός από την διάκριση του υλικού σε κατηγορίες οπτικού και ακουστικού ή συνδυασμό αυτών, γίνεται και διάκριση στις κλάσεις περιγραφών.

Όταν αναφερόμαστε σε κλάσεις περιγραφών, στην πραγματικότητα εννοούμε τις διάφορες πιθανές πλευρές που θα μπορούσε να καλύψει μία περιγραφή οπτικοακουστικού περιεχομένου. Ένα βασικό στοιχείο είναι ότι υπάρχουν πολλοί τρόποι για να περιγράψουμε μία οντότητα, ανάλογα με το πως αυτή θα χρησιμοποιηθεί. Για αυτόν τον λόγο το πρότυπο MPEG-7 πρέπει να συμβιβάσει αυτές τις περιγραφές και να τις καταστήσει συμπληρωματικές την μία ως προς την άλλη, παρά να χρησιμοποιηθεί αποκλειστικά μόνο μία.

Υπάρχουν τέσσερις θεμελιώδεις κλάσεις για την περιγραφή του περιεχομένου, που σχετίζονται περισσότερο με τα δεδομένα, δηλαδή με το υλικό που πρόκειται να περιγραφεί, παρά μεταξύ τους. Οι τέσσερις αυτές κατηγορίες είναι:

- Η εγγραφή,
- Η φυσική
- Η νοητική και
- Η περιγραφή που βασίζεται στο μέσο.

Στην κορυφή αυτών των σχημάτων βρίσκεται η περιγραφή της αρχιτεκτονικής τους, που σχεδιάζει τις σχέσεις ανάμεσα σε μεγάλες ενότητες δεδομένων καθώς και τις σχέσεις μεταξύ των περιγραφών και μέσα στις ίδιες τις περιγραφές. Η περιγραφή της σήμανσης, όπως ονομάζεται, βρίσκεται στην κορυφή όλων των κατηγοριών και αγγίζει την κάθε μία ξεχωριστά.

Το πιο πιθανό πάντως είναι ότι στην πραγματικότητα κάθε περιγραφή που θα χρησιμοποιείται στο MPEG-7 θα περιλαμβάνει μόνο μία ή δύο από αυτές τις κλάσεις. Στη συνέχεια θα αναλυθούν οι διαφορετικοί πιθανοί τύποι περιγραφών που υπάρχουν με περισσότερη λεπτομέρεια.

4.4 Περιγραφείς

Οι περιγραφείς του προτύπου MPEG-7 θα μπορούσαν να ομαδοποιηθούν ως εξής:

- Ο «περιγραφέας που βασίζεται στο μέσο» είναι εκείνος που περιγράφει το μέσο με το οποίο εκφράζονται τα δεδομένα. Τέτοιος περιγραφέας θα μπορούσε να είναι ο ρυθμός δειγματοληψίας ενός ψηφιακού αρχείου, εάν είναι αναλογική ή ψηφιακή η πηγή, σε ποιο σημείο κόβονται οι σκηνές, σε ποιο σημείο εστιάζει κάθε φορά η κάμερα, και θα μπορούσαν να παρατεθούν και πολλά άλλα στοιχεία. Ένας περιγραφέας που βασίζεται στο μέσο μπορεί να πραγματεύεται τέτοια είδη στοιχείων χαμηλού επιπέδου, επιφανειακά χαρακτηριστικά που περιγράφουν την εγγραφή ή την επαναπροβολή του ίδιου του μέσου. Υπάρχουν αρκετές τεχνικές για να αποκτηθεί αυτό του είδους η περιγραφή, μέσω της ανάλυσης του ήχου ή της εικόνας. Η κωδικοποίηση αυτών των περιγραφών θα μπορούσε να γίνει πολύ πιο εύκολα κατά την επεξεργασία της δημιουργίας του υλικού.
- Ο «φυσικός περιγραφέας» είναι ο περιγραφέας που ενδεχομένως θα μπορούσε να συνδυαστεί και με την

υποκειμενική πλευρά και αυτή η προσέγγιση μπορεί να καλύψει όλες εκείνες τις πλευρές που έχουν σχέση με τα υπολογιστικά χαρακτηριστικά, που δεν ανταποκρίνονται στην ανθρώπινη αντίληψη. Πρακτικά μιλώντας θα μπορούσαμε να εξάγουμε αυτά τα χαρακτηριστικά εύκολα από τα ακατέργαστα πολυμεσικά δεδομένα. Αυτά τα δεδομένα έχουν αδιαμφισβήτητες τιμές και καλά εδραιωμένους αλγόριθμους για την εξαγωγή των δεδομένων.

- Ο «περιγραφέας που βασίζεται στην υποκειμενική αντίληψη»: Η υποκειμενική αντίληψη τεμαχίζει τα μέσα σε αντικείμενα, δηλαδή αναφέρεται σε περιγραφές τέτοιων χαρακτηριστικών όπως το χρώμα, την υφή και το σχήμα. Τα αντικείμενα περιγράφουν την εικόνα ή τον ίδιο τον ήχο και δεν πραγματεύονται γενικά την δομή της καταγραφόμενης (ή δημιουργούμενης) σκηνής μέσα στα δεδομένα.
- Ο «περιγραφέας της εγγραφής ή η κλάση της εγγραφής»: τυπικά αναπαριστά μία επαναδόμηση της δομής του κόσμου, όπως αυτή συλλαμβάνεται από τα δεδομένα.. Στην μουσική αυτή η κλάση εξυπηρετεί σαν την πραγματική αντιγραφή της μουσικής, δηλαδή οι νότες που παίζονται. Φυσικά το να καταφέρνουμε τον εμπλουτισμό της δομής από ένα απλό video ή ήχο (χωρίς κάποια άλλη πληροφορία) προάγει την σημερινή τεχνολογία. Υπάρχουν περισσότεροι από ένας τρόποι για να αποκτήσουμε μία τέτοια περιγραφή.
- Ο «περιγραφέας της αρχιτεκτονικής» βρίσκεται πάνω από τις τρεις προαναφερθείσες. Αυτός ο περιγραφέας περιέχει τα δομικά στοιχεία των άλλων κλάσεων και επομένως τα δεδομένα που αυτές περιγράφουν. Αυτή η κλάση περιγράφει την περιοχή των τεκμηρίων, στην οποία ένας χρήστης ενδιαφέρεται να μάθει μόνο τις

σχέσεις μεταξύ των τμημάτων και όχι το τι αυτά περιέχουν. Αυτή είναι επίσης και η κλάση που καθορίζει την συντακτική δομή, που πρέπει να στηρίζεται απαραίτητα σε χαμηλότερα σημασιολογικά επίπεδα. Αυτό το είδος της περιγραφής μπορεί να αποκτηθεί με πολλούς τρόπους. Οι περιγραφές μπορούν να παραχθούν με το χέρι, μπορούν να είναι υποπροϊόντα μίας αυτοματοποιημένης ανάλυσης ή να υποδηλώνονται σε ένα χειρόγραφο αντίγραφο και απλά να εξυπηρετούν στο να κάνουν αυτές τις σχέσεις πιο ξεκάθαρες.

- Ο «περιγραφέας της σήμανσης βρίσκεται τοποθετημένος στη κορυφή όλων των άλλων κλάσεων, όπως επίσης στην κορυφή βρίσκονται και τα ίδια τα δεδομένα. Είναι ο τομέας της σήμανσης και της ανάλυσης άλλων ήδη υπάρχοντων μετα-δεδομένων που πραγματοποιούνται από τον ίδιο τον άνθρωπο. Στην πραγματική της χρήση, η περιγραφή αυτή, στη μουσική, υπήρξε ο χώρος για τους μουσικολόγους, στον οποίο μπορούσαν να πραγματοποιήσουν τα σχόλιά τους για διάφορα χαρακτηριστικά ενός μουσικού κομματιού, όπως γενικά ζητήματα μουσικολογικής ανάλυσης.

Συνηθισμένα ζητήματα που αφορούν την μουσικολογία περιλαμβάνουν μουσικές φόρμες (που αναφέρονται στην αρχιτεκτονική του μουσικού κομματιού), τις ίδιες τις νότες (που συνδέονται με το αντίγραφο του μουσικού κομματιού), τα σχόλια σχετικά με το συναισθηματικό περιεχόμενο της μουσικής (που αναφέρονται στα ίδια τα δεδομένα) και τις σχέσεις που υπάρχουν με άλλα μουσικά κομμάτια (που συνδέονται με άλλες περιγραφές).

Το πρότυπο MPEG-7 παρέχει ένα πλαίσιο όπου επιτρέπεται να γίνονται αναφορές στα τμήματα ενός τεκμηρίου, σε ολόκληρο το τεκμήριο, ή σε μία σειρά από τεκμήρια. Με την ευελιξία στην διαχείριση των δεδομένων στηρίζεται και η πολλαπλή μορφή που μπορεί να έχει ένα τεκμήριο. Αυτό σημαίνει ότι θα μπορούσε να είναι πιθανή η περιγραφή πολυμεσικού περιεχομένου με τέτοιον τρόπο που να επιτρέπει να

θέτουμε ερωτήματα βασισμένα σε ηχητικά δεδομένα και να ανακτούμε οπτικά δεδομένα, καθώς και το αντίστροφο.

Επιπρόσθετα, η ευελιξία αυτή βοηθά στο να διασφαλιστεί η μακροβιότητα του προτύπου, που σημαίνει ότι τα σχήματα των περιγραφών που χρησιμοποιούνται για μία συγκεκριμένη εργασία, μπορούν να τροποποιηθούν για μία διαφορετική, αλλά σχετική εφαρμογή. Το πρότυπο MPEG-7 πραγματεύεται εφαρμογές που μπορούν να αποθηκευτούν ή να είναι με την μορφή της ροής (όπως για παράδειγμα συμβαίνει στην αναμετάδοση) και μπορούν να λειτουργήσουν ταυτόχρονα σε περιβάλλοντα πραγματικού χρόνου και σε περιβάλλοντα μη πραγματικού χρόνου.

Οι περιγραφείς του προτύπου MPEG-7 μπορούν να εντοπισθούν μαζί με το συνδεδεμένο με αυτές οπτικοακουστικό υλικό στην ίδια ροή δεδομένων ή στο ίδιο αποθηκευτικό σύστημα όμως οι περιγραφείς θα μπορούσαν να υπάρξουν και αλλού. Όταν το περιεχόμενο και οι περιγραφείς του δεν βρίσκονται μαζί, θα μπορούσαν να αποδειχθούν χρήσιμοι μηχανισμοί που συνδέουν το οπτικό ακουστικό υλικό μαζί με τους περιγραφείς τους.

Αυτοί οι σύνδεσμοι μπορούν να χρησιμεύουν και στις δύο κατευθύνσεις. Ο συνδυασμός της ευελιξίας και της καθολικότητας επιτρέπει τόσο στους ανθρώπους όσο και στις μηχανές να ανταλλάσσουν, να ανακτούν και να επαναχρησιμοποιούν το σχετικό υλικό.

Το βασικό είναι ότι το πρότυπο MPEG-7 υπάρχει για να εξυπηρετεί αυτόν τον στόχο. Η προτυποποίηση κυρίως αναζητεί τρόπους για να φτάσει πέρα από κάθε απλή, αυθαίρετη λύση και να παρέχει όχι μόνο ένα πλαίσιο, αλλά ένα στέρεο μέσο με το οποίο οι λύσεις που προτείνουν οι βιομηχανίες θα μπορούσαν να συνεργαστούν. Εάν το MPEG-7 γίνει ακόμη πιο εξειδικευμένο ή αρκετά γενικό ως προς τις εφαρμογές του, τότε αυτός ο στόχος, που είναι και κύριος, θα αποτύχει.

Η παρακάτω εικόνα απεικονίζει την διαδικασία της επεξεργασίας του προτύπου MPEG-7 με στόχο να εξηγήσει την λειτουργία του προτύπου με έναν απλό τρόπο. Στη συνέχεια ακολουθεί η επόμενη ενότητα, που έχει σχέση με την παρουσίαση και ανάλυση των ρυθμιστικών στοιχείων του MPEG-7.

Ως περιγραφέας ορίζεται η τιμή ενός χαρακτηριστικού. Ένας περιγραφέας καθορίζει την σύνταξη και την σημασιολογία της αναπαράστασης του χαρακτηριστικού.

Για να λειτουργήσει ένας περιγραφέας στο MPEG-7, πρέπει να προσδιορίσει με ακρίβεια την σημασιολογία του χαρακτηριστικού, τον συνδεόμενο με αυτόν τύπο δεδομένων, τις επιτρεπτές τιμές και μία ερμηνεία των τιμών των περιγραφέων. Ένα παράδειγμα θα μπορούσε να είναι το εξής. Το χρώμα για παράδειγμα είναι ένα αλφαριθμητικό. Ο τύπος δεδομένων μπορεί να είναι σύνθετος, το οποίο σημαίνει ότι μπορεί να διαμορφωθεί με το να συνδέσουμε αλυσιδωτά πολλαπλές τιμές ενός τύπου δεδομένων, π.χ. *RGB color: {integer, integer, integer}*.

Το MPEG-7 πραγματεύεται διάφορα επίπεδα αφαίρεσης. Στο χαμηλότερο επίπεδο αφαίρεσης, οι περιγραφείς μπορεί να περιλαμβάνουν χαρακτηριστικά όπως η μορφή, το χρώμα, το μέγεθος για τις εικόνες και το video, η αρμονία και το ηχόχρωμα για την μουσική.

Στο υψηλό επίπεδο μπορεί να βρίσκονται τα γεγονότα που συμβαίνουν, οι σχέσεις μεταξύ των ατόμων, κάποιες αφηρημένες έννοιες. Οι περιγραφείς ήχου και εικόνας αναπαριστούν εξειδικευμένα χαρακτηριστικά που σχετίζονται με ακουστικό και οπτικό περιεχόμενο αντίστοιχα. Υπάρχουν επίσης και οι γενικοί περιγραφείς που περιγράφουν τα γενικά χαρακτηριστικά.

Είναι πιθανό να έχουμε ποικίλους περιγραφείς που να αναπαριστούν ένα μοναδικό χαρακτηριστικό και αυτό σημαίνει ότι πραγματεύεται διαφορετικές σχετικές απαιτήσεις. Παραδείγματα πολλαπλών περιγραφέων για ένα χαρακτηριστικό αποτελούν οι αριθμητικές λίστες και τα ιστογράμματα με χρώματα.

Σε ένα σύνολο δεδομένων (ή και σε ένα υποσύνολό τους) μία τιμή ενός περιγραφέα (*description value*) είναι η απόδοση της τιμής (*instantiation*) που δίδεται σε αυτόν. Οι τιμές συνδυάζονται μέσω ενός Σχήματος Περιγραφέων ώστε να δημιουργήσουν μία περιγραφή.

Ένα σχήμα περιγραφέων προδιαγράφει την δομή και την σημασιολογία των σχέσεων ανάμεσα στα συστατικά στοιχεία του, τα οποία μπορεί να είναι ταυτόχρονα περιγραφείς και σχήματα περιγραφής.

Η διάκριση ανάμεσα στα σχήματα περιγραφέων και στους περιγραφείς είναι η εξής. Ένας περιγραφέας πραγματεύεται την αναπαράσταση και την παρουσίαση ενός χαρακτηριστικού. Το σχήμα περιγραφέων από την άλλη ασχολείται με την δομή μιας περιγραφής. Γίνεται αντιληπτό λοιπόν ότι οι περιγραφείς και τα σχήματα περιγραφής

αναφέρονται σε διαφορετικές πλευρές και για αυτό αποτελούν δύο πολύ διαφορετικές έννοιες.

Τα σχήματα περιγραφέων επιτρέπουν την δημιουργία πολύπλοκων περιγραφών με το να προσδιορίζουν την δομή και την σημασιολογία των σχέσεων ανάμεσα στους περιγραφείς και στα σχήματά τους. Για παράδειγμα, ένα σχήμα περιγραφής για ένα τμήμα μιας εικόνας μπορεί να προσδιορίζει την σύνταξη και την σημασιολογία των συστατικών στοιχείων του, όπως την βαθύτερη αποσύνθεση των τμημάτων, τα χαρακτηριστικά κάθε τμήματος ξεχωριστά και τις σχέσεις ανάμεσα στα συστατικά στοιχεία.

Όπως συμβαίνει και με τους περιγραφείς, έτσι και τα σχήματα περιγραφέων μπορούν να κατηγοριοποιηθούν σε σχήματα για τον ήχο, την εικόνα, ή να είναι γενικά. Τα γενικά Σχήματα συνήθως αναπαριστούν μία γενική μεταπληροφορία που σχετίζεται με όλα τα είδη των μέσων (ήχο, εικόνα, κείμενο, γραφικά, κ.τ.λ.).

Σε μία περιγραφή εκτός από τους Περιγραφείς και τα Σχήματα Περιγραφής που προέρχονται εσωτερικά από το κείμενο, το πρότυπο MPEG-7 περιλαμβάνει και Περιγραφείς και Σχήματα Περιγραφής που έχουν σχέση με την δημιουργία, την παραγωγή και την διαχείριση του οπτικοακουστικού υλικού. Αυτά τα μετα-δεδομένα μπορεί να περιέχουν πληροφορίες για τις προϋποθέσεις της πρόσβασης στο υλικό (για παράδειγμα, πληροφορία για τα δικαιώματα της πνευματικής ιδιοκτησίας του υλικού), για την ταξινόμηση (να περιλαμβάνει ταξινόμηση του υλικού σε έναν αριθμό προκαθορισμένων κατηγοριών) και συνδέσμους με άλλα σχετικά στοιχεία (όπου και η πληροφορία αυτή πιθανό να κάνει πιο σύντομη την έρευνα του χρήστη).

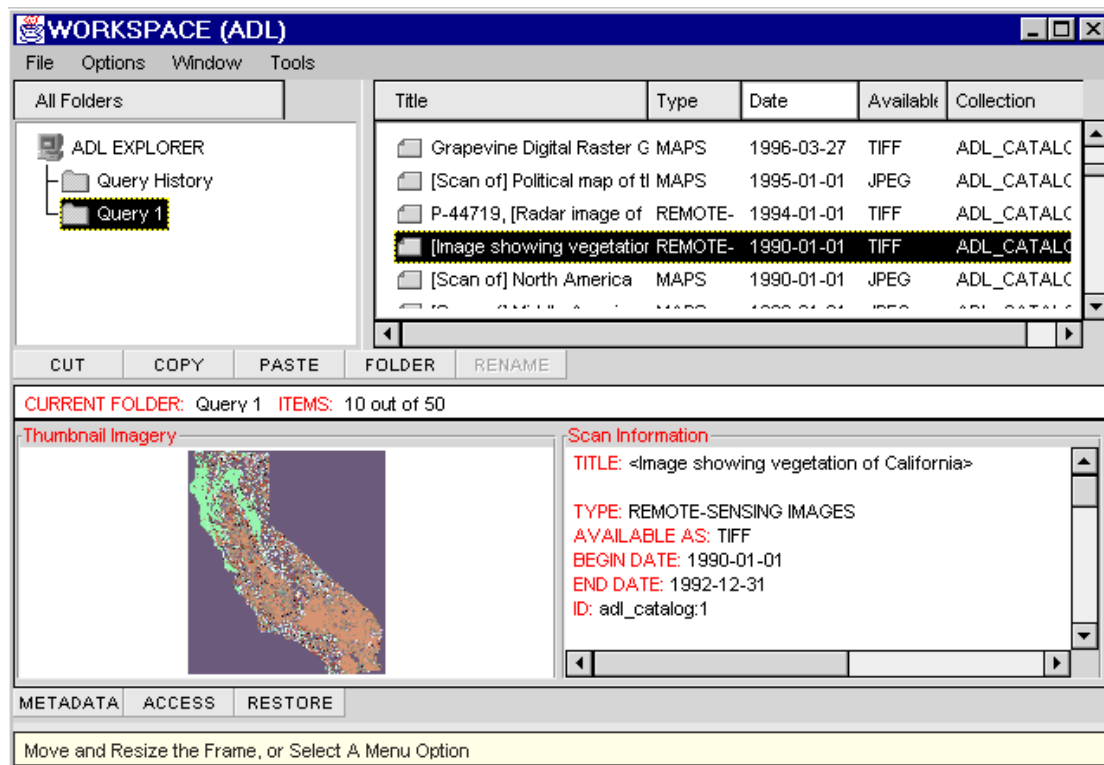
ΚΕΦΑΛΑΙΟ 5: Συστήματα ανάκτησης εικόνας με βάση το περιεχόμενο

Στο κεφάλαιο αυτό θα περιγραφούν μερικά από τα πιο ευρέως διαδεδομένα συστήματα ανάκτησης εικόνας με βάση το περιεχόμενο. Η παρουσίαση τους γίνεται προκειμένου να καταδειχθεί η δυναμική της ερευνητικής δραστηριότητας που επιτελείται στον τομέα αυτό καθώς και η ποικιλία των συστημάτων που υπάρχουν με τις διαφορετικές προσεγγίσεις του θέματος της ανάκτησης των εικόνων.

5.1 ADL

Το σύστημα ADL προτάθηκε από τον Manjunath (1995). Χρησιμοποιώντας ένα εύχρηστο περιβάλλον, ο χρήστης μπορεί να επιλέξει μια περιοχή σε ένα δισδιάστατο παγκόσμιο χάρτη που τον ενδιαφέρει και να αναζητήσει τις εικόνες με βάση το περιεχόμενο που επιθυμεί. Στη συνέχεια, ο χρήστης μπορεί να επιλέξει ένα κατάλογο εικόνων και να θέσει όποιες παραμέτρους επιθυμεί.

Όταν το ερώτημα του χρήστη αποστέλλεται στις βάσεις δεδομένων, τότε η περιοχή ερωταποκρίσεων αναζητεί τις εικόνες και βρίσκει το πιθανό σύνολο εικόνων που ενδιαφέρει το χρήστη. Οι εικόνες εμφανίζονται ως μικρογραφίες και ένα σημάδι εμφανίζεται στο χάρτη για καθεμία εικόνα. Πληροφορίες για το σύστημα υπάρχουν στο διαδίκτυο στο <http://www.alexandria.ucsb.edu/adl.html>.



Εικόνα 5: Το σύστημα ADL

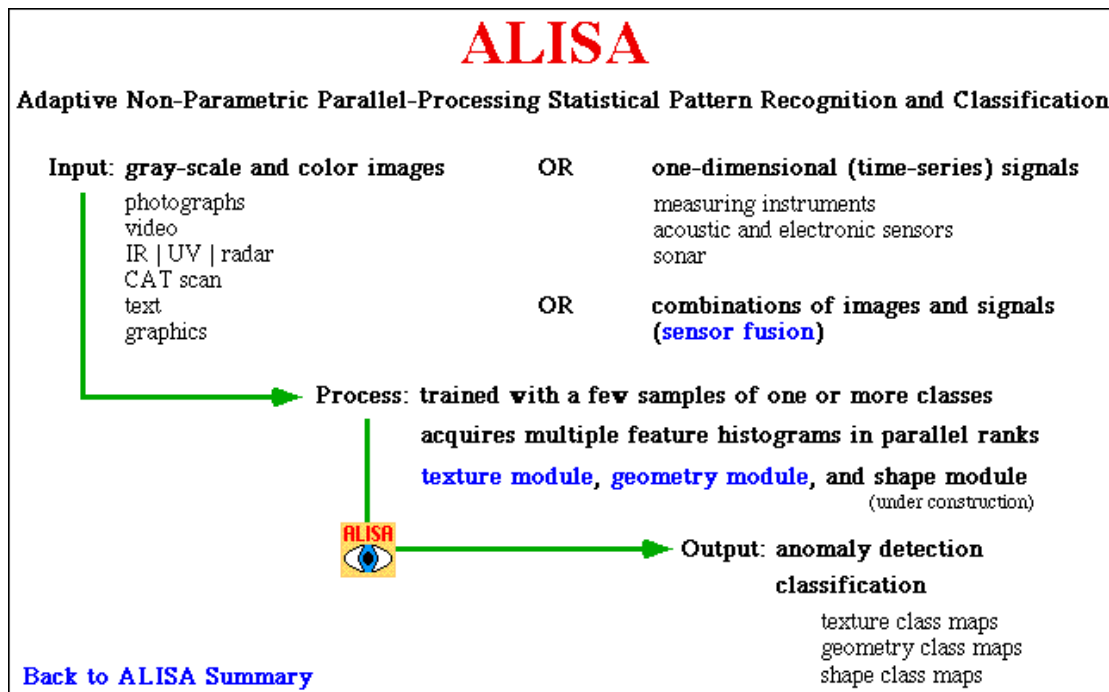
5.2 ALISA

Το ALISA (*Adaptive Learning Image and Signal Analysis*) αναπτύχθηκε από την ALIAS Corporation σε συνεργασία με το Πανεπιστήμιο George Washington (*George Washington University*) και το Ερευνητικό Ινστιτούτο για την Επεξεργασία της Εφαρμοσμένης Γνώσης (*Research Institute for Applied Knowledge Processing*) στη Γερμανία. Το σύστημα αυτό χρησιμοποιεί έναν αλγόριθμο που ονομάζεται *Collective Learning* για τη στατιστική αναπαράσταση μιας εικόνας.

Το ALISA έχει τη δυνατότητα να εξάγει χαρακτηριστικά για ολόκληρη την εικόνα ή για τμήματά της. Ο αλγόριθμος αυτός εκπαιδεύεται για να αναγνωρίζει μια νέα δομή ή σχήμα με λίγες μόνο εικόνες. Εξάγει χαρακτηριστικά στοιχεία σχετικά με το χρώμα, την υφή και το σχήμα των εικόνων.

Επιτυχημένες εφαρμογές του συστήματος περιλαμβάνουν την ανίχνευση εισβολέων σε περιορισμένο περιβάλλον, την ανίχνευση οχημάτων σε περιβάλλον ερήμου, την ανίχνευση ανευρυσμάτων στην καρδιά, ταξινόμηση μπαταριών ανά μάρκα, νομισμάτων, φρούτων, γραμματοσήμων κ.ά. Χρησιμοποιείται εδώ και χρόνια σε

βιομηχανικές εφαρμογές. Πληροφορίες για το σύστημα υπάρχουν στο διαδίκτυο στο http://www.seas.gwu.edu/~pbock/ALISA_Summary.html.



Εικόνα 6: Το σύστημα ALISA

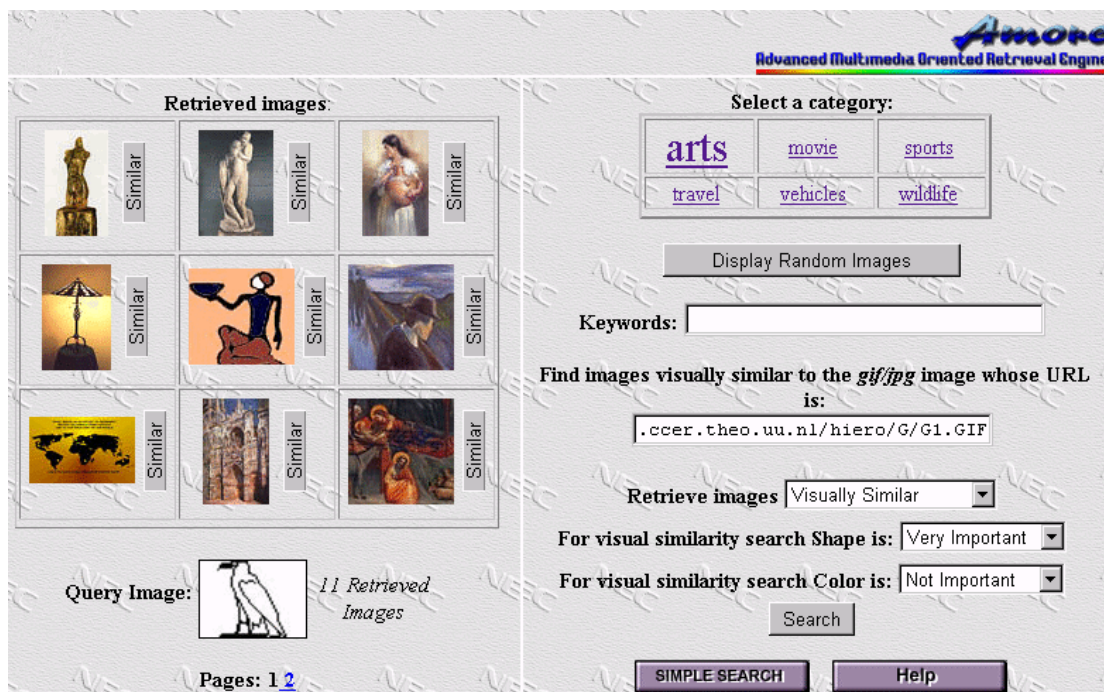
5.3 Amore

Το σύστημα Amore (Mukherjea et al, 1997) έχει κατασκευαστεί από την εταιρεία C & C Research Laboratories NEC USA, Inc. Η εικόνα διαχειρίζεται το πολύ σε οχτώ περιοχές κάποιου χρώματος και το μέγεθος τους αλλάζει σε διαστάσεις 24 x 24 pixels. Οι περιοχές αυτές χρησιμοποιούνται απευθείας για την εύρεση παρόμοιων εικόνων.

Ο χρήστης πρώτα επιλέγει μία κατηγορία εικόνων και εμφανίζεται ένα αρχικό σύνολο εικόνων που μπορούν να επιλεγθούν είτε στην τύχη είτε με κάποια λέξη κλειδί. Από αυτές τις εικόνες, το σύστημα ανακτά παρόμοιες εικόνες. Επίσης, το ερώτημα μπορεί να οριστεί επίσης και από τη διεύθυνση URL που έχει.

Στην αναζήτηση των παρόμοιων εικόνων, αναζητείται αρχικά μία αντιστοιχία μεταξύ των περιοχών του ερωτήματος και των υπόλοιπων εικόνων. Η ομοιότητα μεταξύ δύο περιοχών βασίζεται στον αριθμό των pixel που επικαλύπτονται και στις δύο

περιοχές. Πληροφορίες για το σύστημα υπάρχουν στο διαδίκτυο στο <http://www.crl.com/amore/>.



Εικόνα 7: Το σύστημα Amore

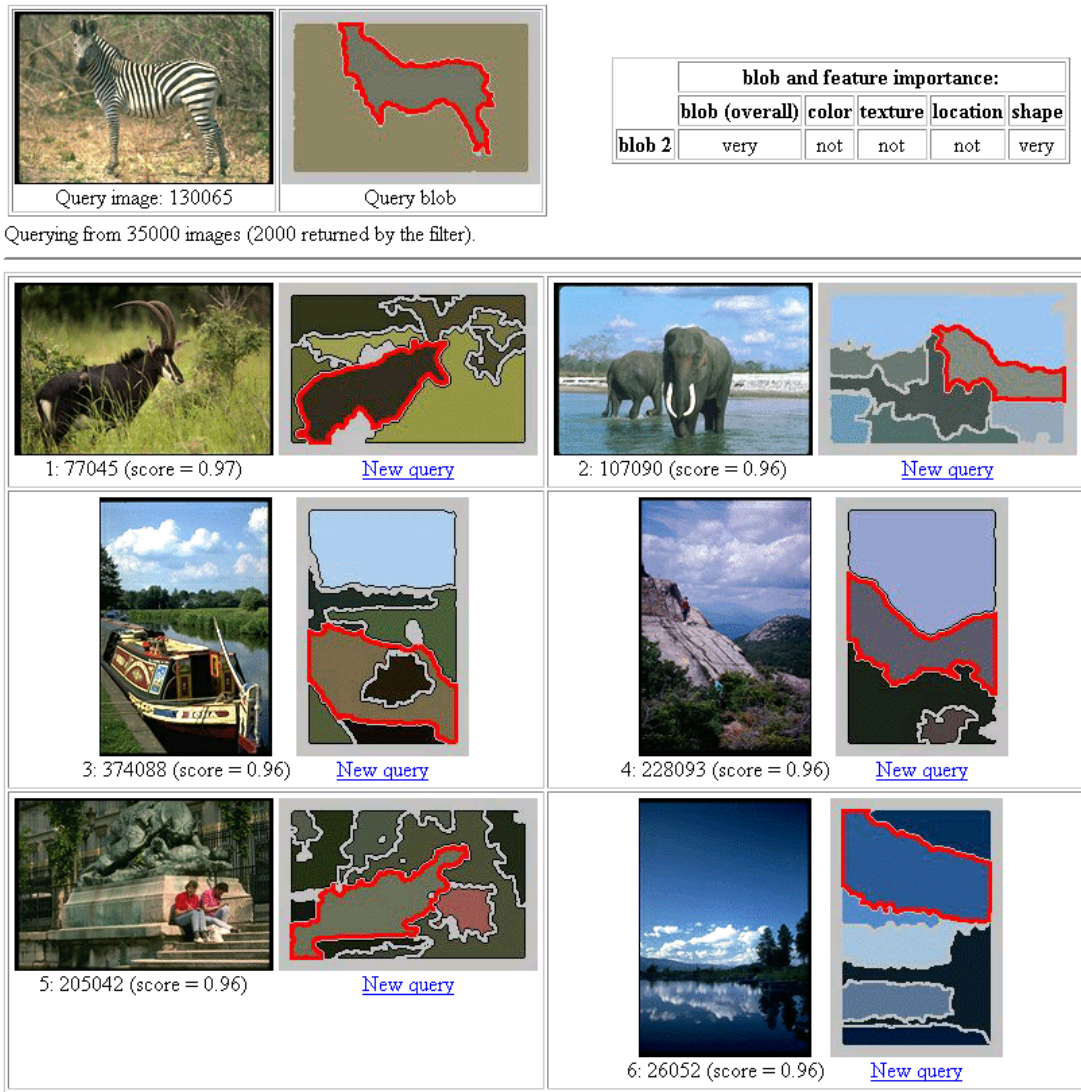
5.4 Blobworld

Το Blobworld (Carson et al, 1999) δημιουργήθηκε από το τμήμα Πληροφορικής του Πανεπιστημίου Μπέρκλεϋ της Καλιφόρνια (University of California, Berkeley). Εξάγει χαρακτηριστικά υπολογίζοντας το χρώμα και την υφή, τη θέση και τα σχήματα των περιοχών των εικόνων.

Ο χρήστης πρώτα επιλέγει μια κατηγορία που περιορίζει το χώρο έρευνας του συστήματος. Έπειτα ο χρήστης σε μια αρχική εικόνα επιλέγει μια περιοχή και δηλώνει την σημασία της αναφορικά με το χρώμα, την υφή, τη θέση και το σχήμα (επιλέγοντας για καθένα απ' αυτά τα χαρακτηριστικά μια από τις λέξεις: not, somewhat, very που υπάρχουν).

Μπορούν να χρησιμοποιηθούν περισσότερες από μία περιοχές για τον προσδιορισμό ερωτήματος. Τα αποτελέσματα παρουσιάζονται σε κατάταξη, ανά ζεύγος η εικόνα και η εκδοχή της μετά την κατάτμηση. Πληροφορίες για το σύστημα υπάρχουν

διαθέσιμες στο διαδίκτυο στο <http://elib.cs.berkeley.edu/photos/blobworld/>. Η παρουσίαση του είναι διαθέσιμη στο <http://elib.cs.berkeley.edu/photos/blobworld/start/html>.



Εικόνα 8: Αποτελέσματα ερωτήσεων με το Blobworld

5.5 CANDID

Το CANDID (*Comparison Algorithm for Navigating Digital Image Database*) (Kelly et al, 1995) δημιουργήθηκε από το Εθνικό Εργαστήριο του Λος Άλαμος των ΗΠΑ (Los Alamos National Laboratory, USA). Υπολογίζει γνωρίσματα για το χρώμα και/ ή για τη δομή των εικόνων. Όσον αφορά τη διεπαφή του με το χρήστη, λειτουργεί

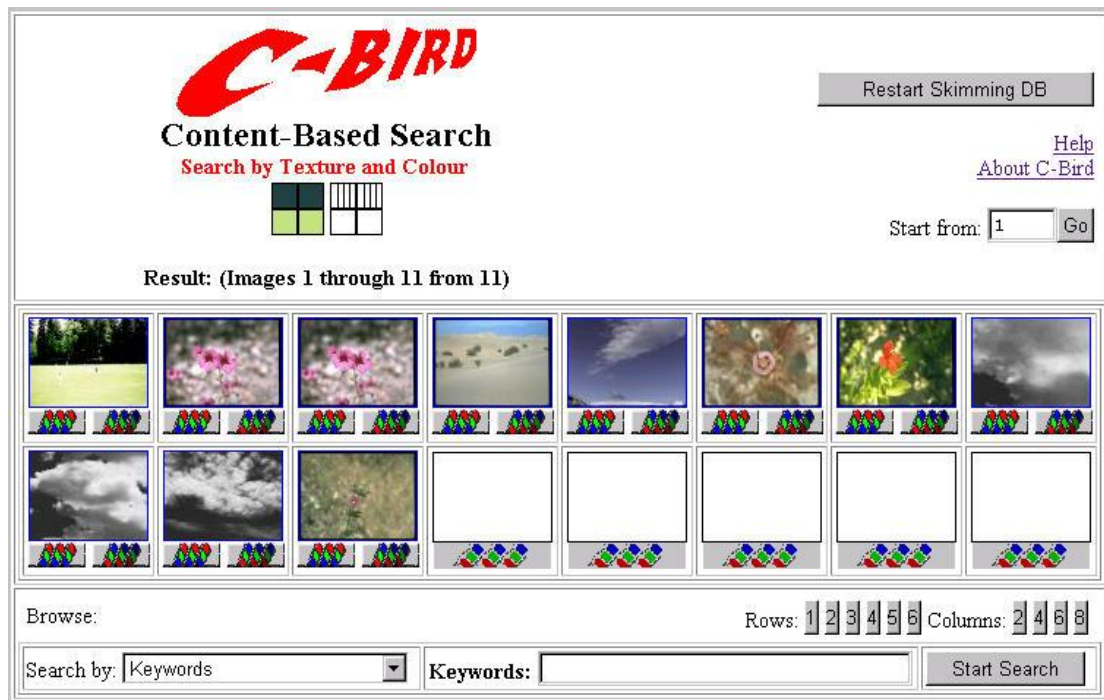
απλά με ένα κλικ στην εικόνα που χρησιμεύει ως παράδειγμα και ανακτά παρόμοιες εικόνες. Στο διαδίκτυο μπορεί κάποιος να βρει πληροφορίες για το CANDID στο <http://public.lanl.gov/kelly/CANDID/index.shtml>.

5.6 C-BIRD

Το σύστημα C-bird (Li et al, 1998) έχει αναπτυχθεί από στο τμήμα School of Computing Science του πανεπιστημίου Fraser του Καναδά. Για κάθε εικόνα που υπάρχει στη βάση δεδομένων του συστήματος, δημιουργείται ένας περιγραφέας για τη δομή και την εμφάνισή της.

Αρχικά, εμφανίζεται στο χρήστη μία συλλογή από τυχαίες εικόνες, όπου επιλέγοντας κάποια μπορεί να κάνει μία αναζήτηση με όμοιες εικόνες ή να κάνει μία αναζήτηση εικόνων με βάση το χρώμα. Ο χρήστης μπορεί να επιλέξει το πλήθος των γραμμών και των στηλών του πλαισίου, στο οποίο θα εμφανίζονται οι εικόνες. Επιλέγοντας μία μικρογραφία, βλέπει μερικά χαρακτηριστικά για το χρώμα και την υφή της εικόνας.

Στο διαδίκτυο μπορεί κάποιος να βρει πληροφορίες για το C-bird στη σελίδα <http://jupiter.cs.sfu.ca/cbird/>.

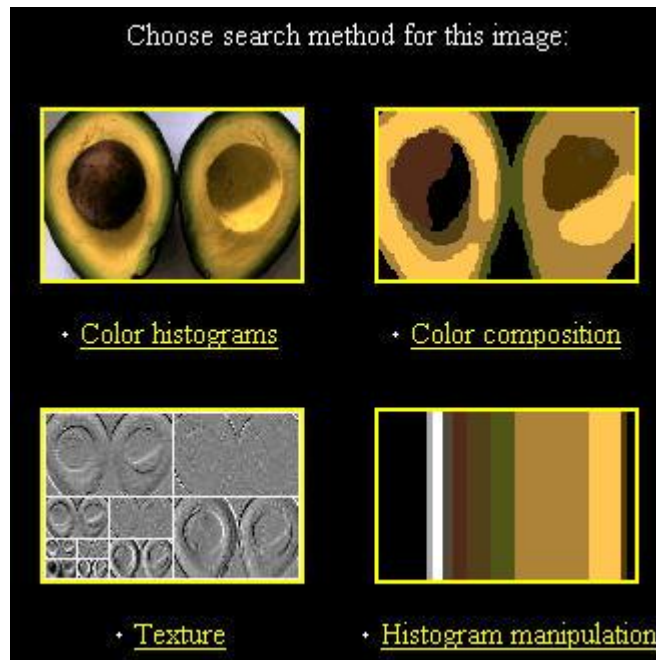


Εικόνα 9: Το σύστημα C-bird

5.7 CBVQ

Το σύστημα CBVQ (Smith et al, 1995) έχει αναπτυχθεί από την ομάδα Image and Advanced Lab στο Columbia University. Το σύστημα επιτρέπει να γίνονται ερωτήματα με παραδείγματα και εκτελεί τα ερωτήματα δημιουργώντας ένα χρωματικό ιστόγραμμα. Ο χρήστης μπορεί να επιλέξει κάποια από τις τυχαίες εικόνες που προβάλλονται από το λογισμικό ή να δώσει τη διεύθυνση URL οποιασδήποτε εικόνας. Μετά την επιλογή της εικόνας, ο χρήστης επιλέγει ένα από τους διαθέσιμους τρόπους αναζήτησης: χρωματικό ιστόγραμμα, σύνθεση χρώματος, υφή ή ιστόγραμμα. Τα 20 καλύτερα αποτελέσματα παρουσιάζονται στο χρήστη σε φθίνουσα διάταξη, αρχίζοντας από την εικόνα με το μεγαλύτερο σκορ.

Στο διαδίκτυο μπορεί κάποιος να βρει πληροφορίες για το CBVQ στη σελίδα <http://maya.ctr.columbia.edu:8088/cbvq/>.



Εικόνα 10: Το σύστημα CBVQ

5.8 CHROMA

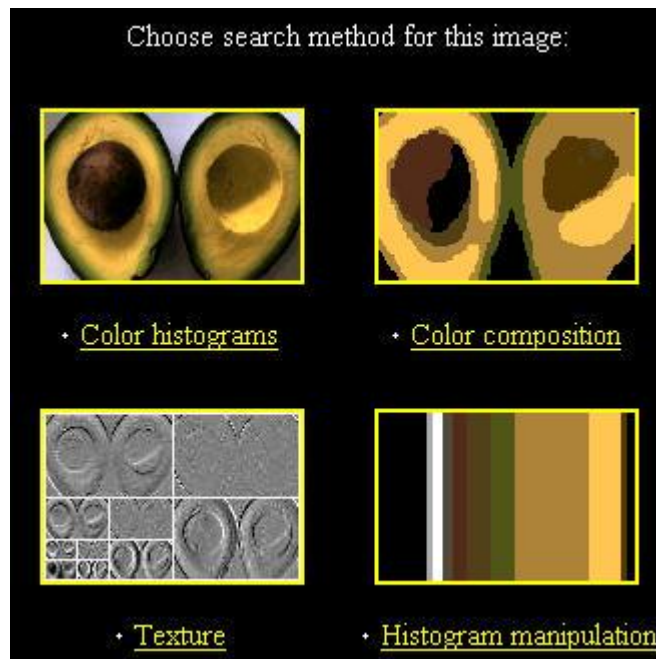
Το σύστημα CHROMA έχει αναπτυχθεί από την ομάδα Image and Advanced Lab στο Columbia University. Το λογισμικό αυτό επιτρέπει ερωταποκρίσεις με βάση το περιεχόμενο. Για κάθε pixel στην εικόνα δημιουργείται ένα διάνυσμα 9 διαστάσεων, που περιέχει τα μεγέθη των αποτελεσμάτων ενός φίλτρου Haar. Στη συνέχεια, με το αποτέλεσμα κάθε φίλτρου δημιουργούνται 9 δυαδικές εικόνες. Στη συνέχεια, οι εικόνες αυτές συγχωνεύονται σε περιοχές με παρόμοιο περιεχόμενο χρησιμοποιώντας ένα μη γραμμικό φιλτράρισμα σε κάθε δυαδική εικόνα για τη μείωση του θορύβου στις εικόνες. Για κάθε περιοχή που ανιχνεύεται, εξάγεται χωρική πληροφορία για αυτή. Για κάθε εικόνα της βάσης δεδομένων δημιουργείται ένα ιστόγραμμα και μία κατάτμηση χρωμάτων.

Το σύστημα επιτρέπει ερωτήματα με παράδειγμα (QBE) και κατευθύνει τα ερωτήματα δημιουργώντας ένα ιστόγραμμα χρώματος. Μετά την επιλογή της εικόνας προς επερώτηση, ο χρήστης επιλέγει μία μέθοδο από αυτές που είναι διαθέσιμες:

- Ιστόγραμμα χρώματος
- Σύνθεση χρώματος

- Σύσταση και χειρισμός ιστογραμμάτων

Στο διαδίκτυο μπορεί κάποιος να βρει πληροφορίες για το CHROMA στη σελίδα <http://maya.ctr.columbia.edu:8088/cbvq/>.



Εικόνα 11: Το σύστημα CHROMA

5.9 COMPASS

Το COMPASS (Computer Aided Search System) αναπτύχθηκε από το Κέντρο Επιστημονικής και Τεχνολογικής Έρευνας στο Τρέντο της Ιταλίας. Εξάγει χαρακτηριστικά από το χρώμα και την υφή των εικόνων. Η αναζήτηση πραγματοποιείται με εικόνα που χρησιμεύει ως παράδειγμα. Επίσης παρέχεται η δυνατότητα στο χρήστη να φυλλομετρήσει εικόνες στη βάση δεδομένων του συστήματος.

Ο χρήστης έχει τη δυνατότητα να δηλώσει ποιες εικόνες από αυτές που ανακτήθηκαν είναι σχετικές με το ερώτημα του και ποιες όχι ώστε οι σχετικές με το ερώτημα εικόνες να ενσωματωθούν στο ερώτημα και να αποτελέσουν νέο ερώτημα για καλύτερα αποτελέσματα κατά την ανάκτηση (ανατροφοδότηση σχετικότητας). Οι εικόνες κατά την ανάκτηση παρουσιάζονται με φθίνουσα κατάταξη ανάλογα με την

ομοιότητα. Πληροφορίες για το σύστημα αυτό υπάρχουν στο διαδίκτυο στο <http://compass.itc.it/>

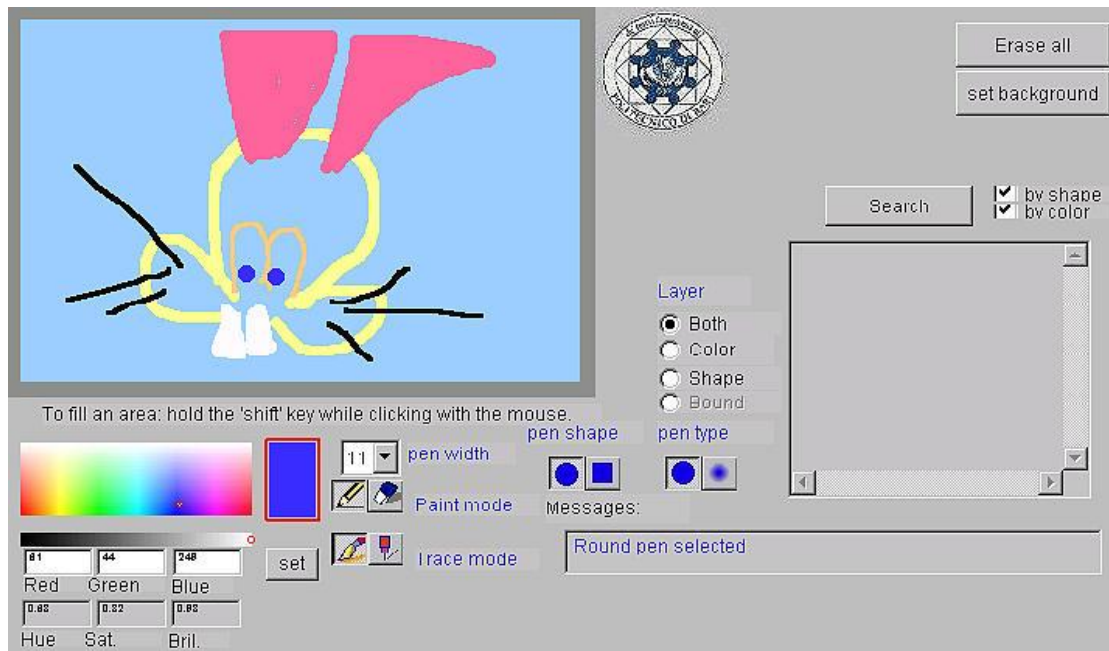


Εικόνα 12: Το σύστημα COMPASS

5.10 DrawSearch

Το DrawSearch (Sciascio et al, 1999) δημιουργήθηκε από το τμήμα Ηλεκτρολόγων και Ηλεκτρονικών Μηχανικών του πανεπιστημίου του Μπάρι στην Ιταλία. Το λογισμικό έχει δύο διεπαφές για το χρήστη: η μία επιτρέπει ερωτήματα με βάση το χρώμα και το σχήμα και η άλλη επιτρέπει ερωτήματα με βάση την υφή. Στο πρώτο υποσύστημα ο χρήστης σχεδιάζει ένα προσχέδιο στον καμβά, επιλέγει τις χρωματικές ιδιότητες για τις γραμμές και τις κλειστές περιοχές. Στο δεύτερο υποσύστημα, ένα ερώτημα επιτυγχάνεται επιλέγοντας μία περιοχή υφής από μία βάση δεδομένων με εικόνες. Τα αποτελέσματα εμφανίζονται σε φθίνουσα σειρά ομοιότητας.

Σχετικές πληροφορίες για το σύστημα αυτό στην ιστοσελίδα: <http://deecom03.poliba.it/DrawSearch/DrawSearch.html>.

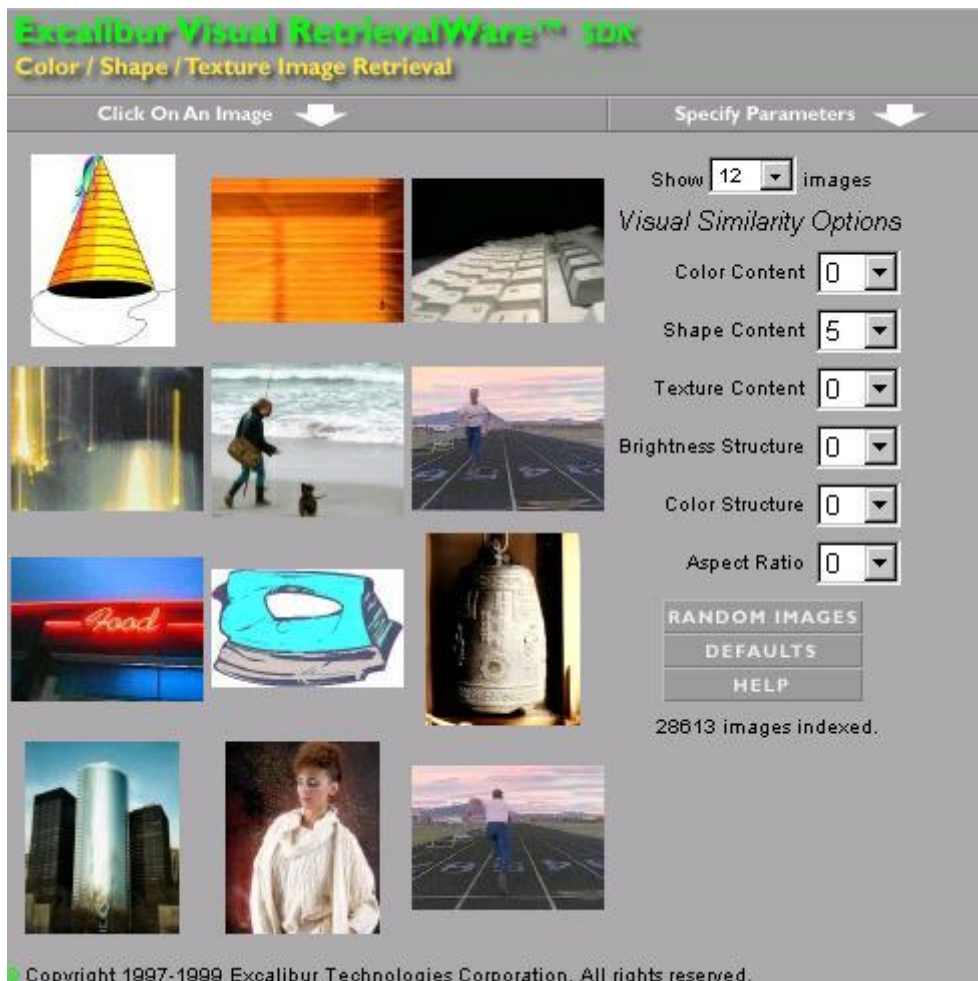


Εικόνα 13: Το σύστημα DrawSearch

5.11 Excalibur Visual RetrievalWare

Το Visual RetrievalWare δημιουργήθηκε από την Excalibur Technologies. Προσφέρει δυνατότητες για την επεξεργασία, την ευρετηρίαση και την ανάκτηση εικόνων με βάση το περιεχόμενο. Ουσιαστικά αποτελεί ένα εργαλείο εφαρμογών λογισμικού, που μπορούν να χρησιμοποιηθούν ανεξάρτητα για την δημιουργία άλλων συστημάτων, παρά ένα αυτοτελές σύστημα.

Χρησιμοποιεί για την ανίχνευση χαρακτηριστικών το χρώμα, το σχήμα και τη δομή. Υποστηρίζεται επίσης η εξαγωγή ειδικών γνωρισμάτων που μπορούν να χρησιμοποιηθούν για τη σύγκριση και το ταίριασμα δακτυλικών αποτυπωμάτων ή προσώπων. Ο χρήστης πρώτα ορίζει τις παραμέτρους της επιθυμητής οπτικής ομοιότητας, δηλαδή σε ποια χαρακτηριστικά θα βασιστεί η ομοιότητα και επιλέγει μια από τις εικόνες που εκθέτονται. Οι εικόνες παρουσιάζονται χωρίς κατάταξη σχετικότητας. Το λογισμικό του χρησιμοποιείται από το Yahoo Image Surfer. Περισσότερα για το Visual RetrievalWare μπορεί κάποιος να πληροφορηθεί στο <http://www.aa-lab.cs.uu.nl/cbirsurvey/cbir-survey/node15.html>.

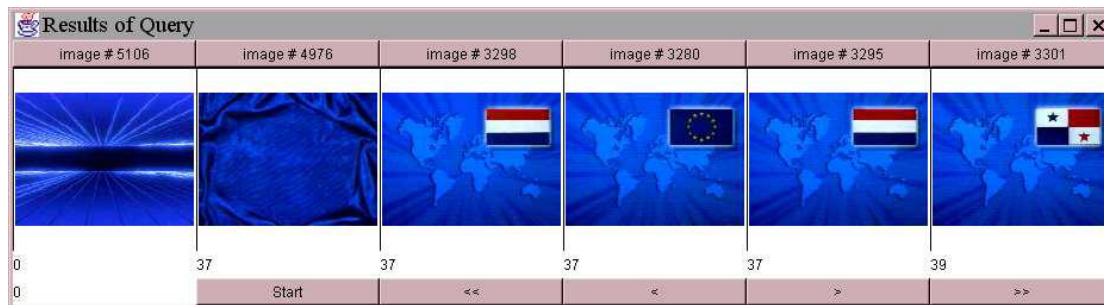


Εικόνα 14: Το σύστημα Excalibur Visual RetrievalWare

5.12 FIDS

Το *FIDS* (*Flexible Image Database System*) (Berman et al, 1999) δημιουργήθηκε από το Πανεπιστήμιο της Ουάσινγκτον (*Department of Computer Science and Engineering, University of Washington, Seattle*) στις Ηνωμένες Πολιτείες. Χρησιμοποιεί διάφορες τεχνικές για την ανίχνευση γνωρισμάτων αναφορικά με το χρώμα.

Ο χρήστης επιλέγει μια εικόνα ως ερώτημα. Κάποια εικόνα από τα αποτελέσματα μπορεί να χρησιμοποιηθεί για αναζήτηση. Η ιστοσελίδα του συστήματος αυτού είναι η παρακάτω: <http://www.cs.washington.edu/research/imagetdatabase>.



Εικόνα 15: Το σύστημα FIDS

5.13 FIR

Το σύστημα *FIR* (*Formula Image Retrieval*) (Volmer, 1997) αναπτύχθηκε ως μέρος του προγράμματος *FORMULA* από το Ινστιτούτο Υπολογιστικών Γραφικών της Γερμανίας (*Fraunhofer Institute for Computer Graphics, Darmstadt, Germany*) σε συνεργασία με εταιρείες από την Ιταλία (*Txt Ignegneria Informatica SPA, Giunti Multimedia Srl*) και από την Ελλάδα (*Epsilon Software, Kino TV & Movie Productions S.A.*

Τα χαρακτηριστικά ανιχνεύονται με βάση το σχήμα και το χρώμα των εικόνων. Το ερώτημα στο σύστημα πραγματοποιείται με μια εικόνα ως υπόδειγμα για την ανάκτηση. Πληροφορίες γι' αυτό το σύστημα υπάρχουν στο διαδίκτυο στο http://www.igd.fhg.de/igd-a7/projects/formula/formula_e.html.

5.14 FIRE

Το σύστημα *FIRE* (*Flexible Image Retrieval Engine*) δημιουργήθηκε στο Πανεπιστήμιο του Άαχεν (*Aachen University*) της Γερμανίας. Εξάγει χαρακτηριστικά για το χρώμα, το σχήμα και τη δομή των εικόνων. Ο χρήστης ξεκινά την αναζήτηση επιλέγοντας μια εικόνα από ένα σύνολο εικόνων έχει τη δυνατότητα ανατροφοδότησης σχετικότητας του ερωτήματος του σημειώνοντας κάτω από κάθε εικόνα αν η εικόνα είναι σχετική, άσχετη ή ουδέτερη με το ερώτημά του.

Το λογισμικό που χρησιμοποιεί το σύστημα αυτό διατίθεται ελεύθερα. Περισσότερες πληροφορίες καθώς και παρουσίαση του συστήματος FIRE υπάρχουν στο: <http://www-i6.informatik.rwth-aachen.de/~deselaers/fire.html>.



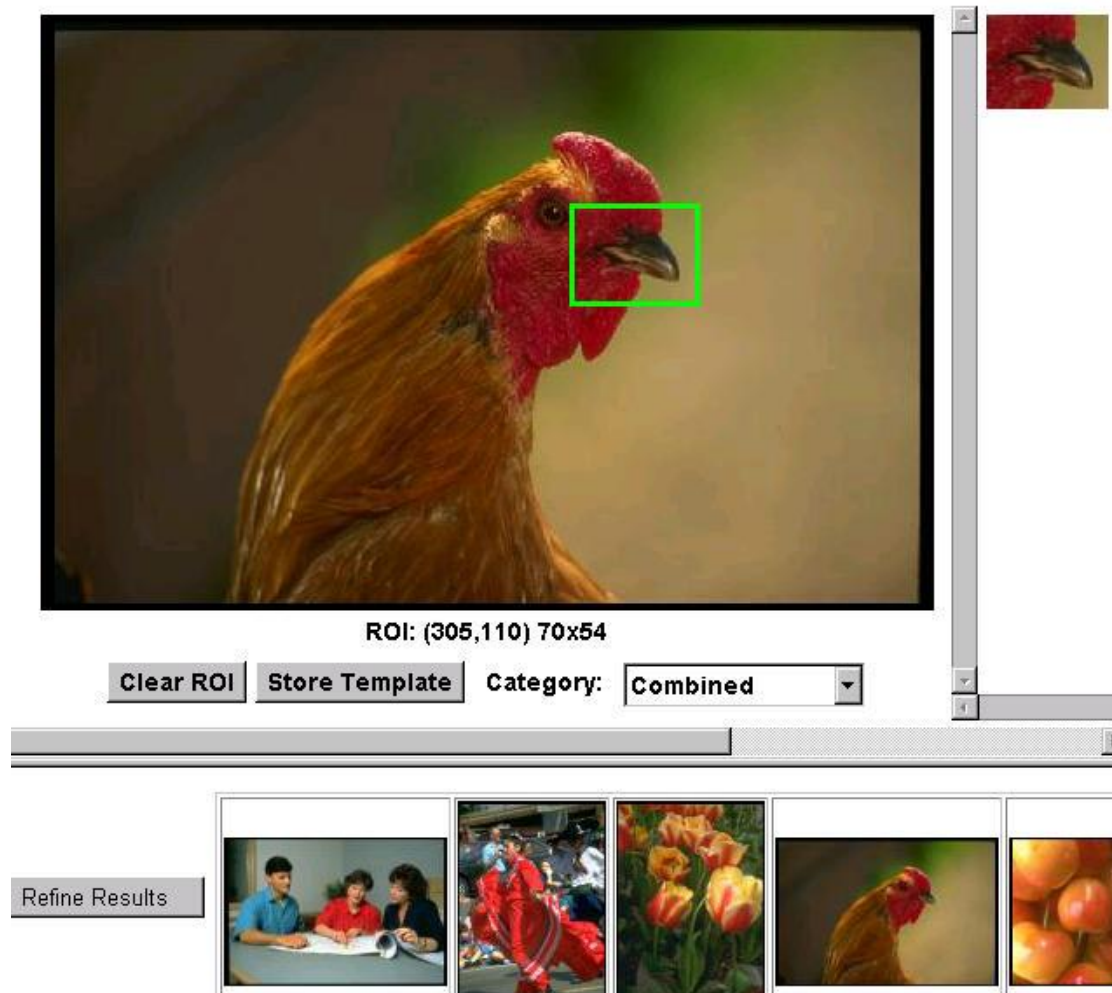
Εικόνα 16: Το σύστημα FIRE

5.15 Focus

Το σύστημα Focus (Das et al, 1997) δημιουργήθηκε από το τμήμα Επιστήμης Υπολογιστών του University of Massachusetts. Κάθε εικόνα τεμαχίζεται σε κελιά που το καθένα περιέχει 100 x 100 pixels. Για κάθε κελί δημιουργείται ένα ιστόγραμμα χρώματος στο χώρο HSV. Οι ακμές από κάθε τοπικό ιστόγραμμα καθορίζονται και συγχωνεύονται σε μία λίστα από μοναδικές κορυφές ολόκληρης της εικόνας. Επίσης, δημιουργείται ένας πίνακας συχνοτήτων για κάθε χρώμα στο χώρο HSV που δίνει τον αριθμό των εικόνων που έχουν μία ακμή του κάθε χρώματος. Η χωρική σχέση μεταξύ των χρωματικών περιοχών αναπαρίσταται με τη χρήση ενός γραφήματος χωρικής εγγύτητας (SPG graph).

Πληροφορίες γι' αυτό το σύστημα υπάρχουν στο διαδίκτυο στη σελίδα http://wagga.cs.umass.edu/~mdas/color_proj.html. Μία δοκιμαστική έκδοση του

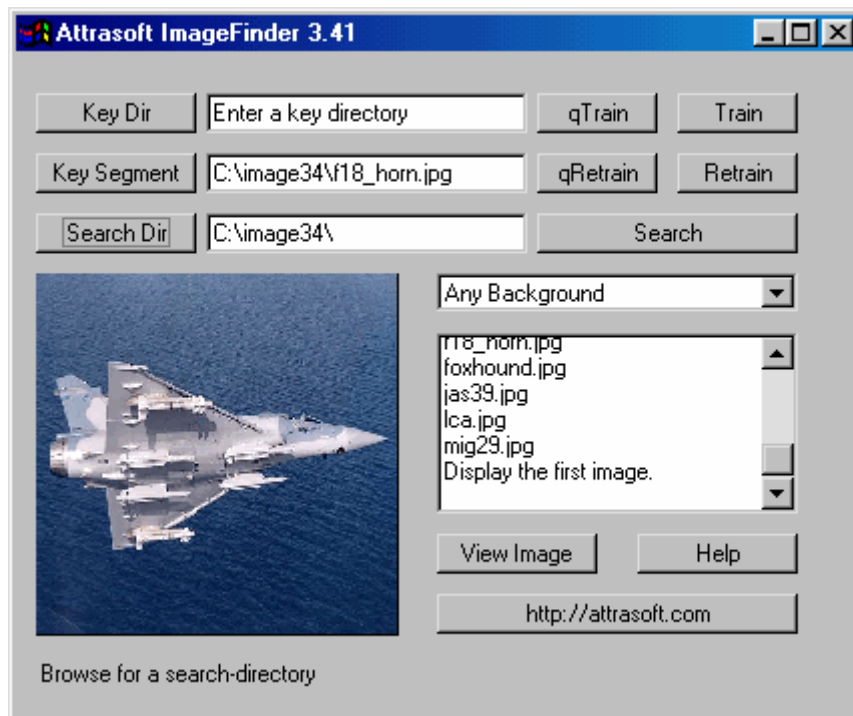
<http://cowarie.cs.umass.edu/~colordemo/mdas/demo1/phase0.html>.



Εικόνα 17: Το σύστημα FOCUS

5.16 ImageFinder

Το ImageFinder αναπτύχθηκε από την εταιρία Attrasoftware. Τα ερωτήματα γίνονται δίνοντας μία εικόνα και τα αποτελέσματα επιστρέφονται ως μία ιστοσελίδα με μία λίστα υπερσύνδεσμων προς τις εικόνες που βρέθηκαν. Σχετικές πληροφορίες για το σύστημα αυτό στην ιστοσελίδα: http://attrasoftware.com/abm3_4.html.



Εικόνα 18: Το σύστημα ImageFinder

5.17 ImageRETRO

To ImageRETRO (Image RETrieval by Reduction and Overview) (Vendrig et al, 1999) αναπτύχθηκε στην Ολλανδία από το Πανεπιστήμιο του Άμστερνταμ (Department of Computer Science, University of Amsterdam). Χρησιμοποιούνται γνωρίσματα σχετικά με το χρώμα των εικόνων. Εκτίθεται στο χρήστη ένα σύνολο εικόνων κάθε μια από τις οποίες είναι αντιπροσωπευτική για ένα τμήμα της συλλογής. Επιλέγοντας, στη συνέχεια, μια από αυτές ο χρήστης, επιλέγει ένα μέρος της βάσης δεδομένων όπου βρίσκονται ομαδοποιημένες και εκτίθεται στο χρήστη ένα άλλο σύνολο αντιπροσωπευτικών εικόνων. Με αυτό το φιλτράρισμα ο χρήστης ανακτά ένα μικρό σύνολο εικόνων το οποίο μπορεί να φυλλομετρήσει. Σχετικές πληροφορίες για το σύστημα αυτό στην ιστοσελίδα: <http://carol.wins.uva.nl/~vendrig/imageretro/>.



Εικόνα 19: Το σύστημα ImageRETRO

5.18 ImageRover

Το σύστημα ImageRover (Sclaroff et al, 1997) δημιουργήθηκε στο Πανεπιστήμιο της Βοστώνης στις Ηνωμένες Πολιτείες (Department of Computer Science, Boston University). Τα γνωρίσματα που χρησιμοποιούνται είναι σχετικά με το χρώμα και την υφή της εικόνας. Ο χρήστης ξεκινά ένα ερώτημα δηλώνοντας ένα σύνολο από λέξεις κλειδιά που σχετίζονται με τις επιθυμητές εικόνες. Από το σύνολο των εικόνων που παρουσιάζονται, ο χρήστης σημειώνει (μαρκάρει) μία ή περισσότερες που είναι παρόμοιες με αυτή που αναζητάει.

Εκτός από την επιλογή σχετικών εικόνων για την ανατροφοδότηση του ερωτήματος του ο χρήστης μπορεί επιπλέον να αφαιρέσει μια εικόνα (ή περισσότερες) από το τελευταίο ερώτημα. Οι εικόνες που μοιάζουν με το ερώτημα παρουσιάζονται σε φθίνουσα σειρά κατά την ανάκτηση. Το σύστημα αυτό αποτελεί μια μηχανή αναζήτησης για τον Παγκόσμιο Ιστό. Πληροφορίες γι' αυτό το σύστημα μπορεί να βρει ο ενδιαφερόμενος στον παρακάτω ιστότοπο:
<http://www.wi.leidenuniv.nl/home/lim/image.scape.html>.

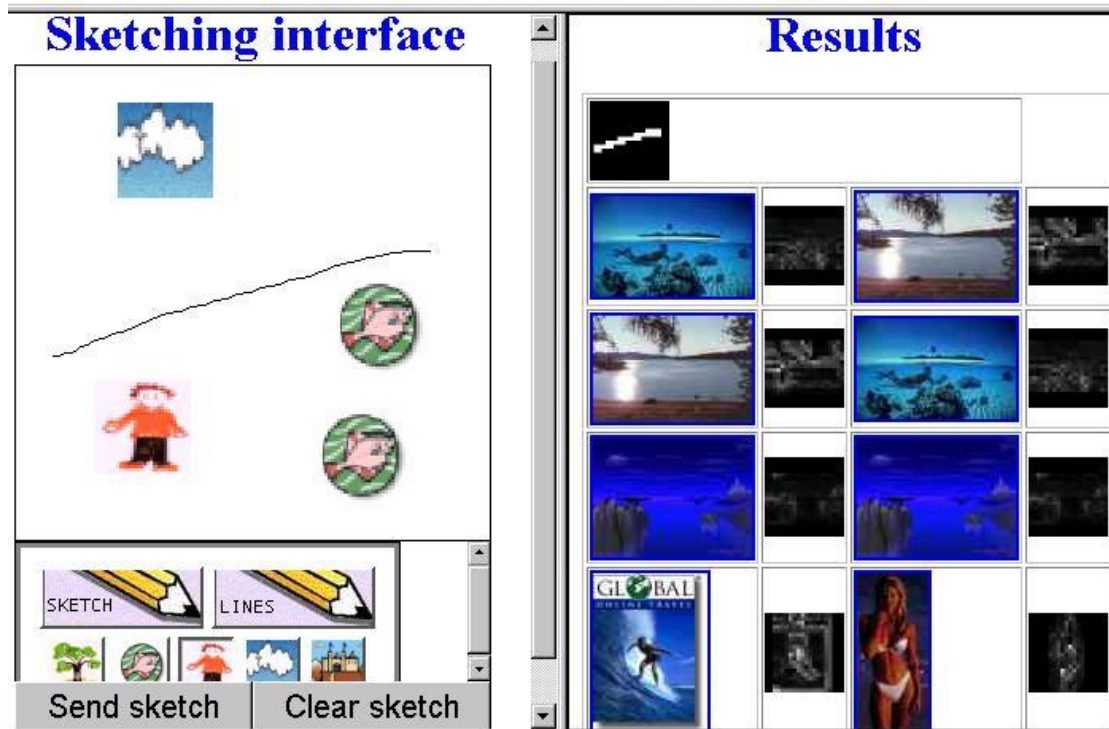


Εικόνα 20: Το σύστημα ImageRover

5.19 ImageScape

Το ImageScape (Lew et al, 1997) δημιουργήθηκε από το τμήμα Πληροφορικής του πανεπιστημίου Leiden στην Ολλανδία. Χρησιμοποιώντας τη διεπαφή, ο χρήστης μπορεί να σχεδιάσει ένα περίγραμμα της επιθυμητής εικόνας. Επίσης, ο χρήστης μπορεί να εισάγει εικόνες στον καμβά, οι οποίες αναπαριστούν αντικείμενα που αναζητά. Μερικές κατηγορίες που μπορεί να αναζητήσει ο χρήστης είναι: ανθρώπινα πρόσωπα, άμμο, νερό, ουρανό, δένδρα, γραμμές, σημεία κ.τ.λ. Περισσότερα για το ImageScape στη σελίδα <http://www.wi.leidenuniv.nl/home/lim/image.scape.html>.

ImageScape



Εικόνα 21: Το σύστημα ImageScape

5.20 iPure

Το iPure (*Perceptual and User-friendly Retrieval of Images*) δημιουργήθηκε από το Ερευνητικό Εργαστήριο της IBM στην Ινδία (*IBM India Research Lab, New Delhi, India*). Γίνεται κατάτμηση των εικόνων και για κάθε περιοχή εξάγονται χαρακτηριστικά για το χρώμα, την υφή, το σχήμα και τη θέση. Ο χρήστης ξεκινά την αναζήτηση με μια εικόνα που χρησιμεύει ως πρότυπο για την αναζήτηση που στέλνεται στον εξηρητητή του συστήματος και επιστρέφεται η αντίστοιχη κατατμημένη. Ο χρήστης τότε έχει τη δυνατότητα επιλογής μιας ή περισσότερων περιοχών της εικόνας που τον ενδιαφέρουν για την ανάκτηση. Περισσότερα για το iPure στο: <http://www.research.ibm.com/irl/projects/mediamining/index.html>.

5.21 KIWI

Το σύστημα KIWI (*Key-points Indexing Web Interface*) αναπτύχθηκε από την εταιρεία INSA στη Λυόν της Γαλλίας. Χρησιμοποιεί σημεία κλειδιά που εντοπίζονται στην εικόνα και γι' αυτά εξάγονται γνωρίσματα σχετικά με το χρώμα, τη θέση τους σε σχέση με άλλα σημεία και τις περιοχές γύρω από αυτά. Ο χρήστης μπορεί να δηλώσει μια εικόνα ως υπόδειγμα είτε επιλέγοντας μια εικόνα από τη βάση είτε δίνοντας το URL μιας εικόνας. Οι εικόνες κατά την ανάκτηση παρουσιάζονται σε διάταξη σύμφωνα με τη μεγαλύτερη ομοιότητα ως προς την εικόνα που χρησιμοποιήθηκε ως ερώτημα. Πληροφορίες για το σύστημα αυτό μπορεί να βρει κανείς στην παρακάτω ιστοσελίδα <http://telesun.insa-lyon.fr/kiwi/>.



Εικόνα 22: Το σύστημα KIWI

5.22 LCPD

Το σύστημα LCPD (*Lew et al, 1996*) αναπτύχθηκε από το τμήμα Επιστήμης Υπολογιστών στο Πανεπιστήμιο του Leiden. Ο χρήστης έχει διάφορες επιλογές, όπως να επιλέξει το χαρακτηριστικό διάνυσμα, μια περιοχή τιμών των *pixel* και το επίπεδο ανάλυσης. Ένα χαρακτηριστική διάνυσμα είναι μια οριζόντια ή κάθετη προβολή του διανύσματος. Για κάθε εικόνα με $m \times n$ *pixels*, το χαρακτηριστικό διάνυσμα έχει $m + n$ στοιχεία που υπολογίζονται ως ο μέσος όρος των τιμών των *pixel* των γραμμών και των στηλών για την επιλεγμένη περιοχή.

Ο χρήστης κάνει τις ερωτήσεις του μέσω παραδειγμάτων. Επιλέγει αρχικά τη μέθοδο αναζήτησης από την επιλεγμένη λίστα. Πληροφορίες για το σύστημα αυτό μπορεί να βρει κανείς στην παρακάτω ιστοσελίδα <http://ind156b.wi.leidenuniv.nl:2000/>. Μία δοκιμαστική έκδοση του προγράμματος μπορεί να βρεθεί στην ιστοσελίδα http://ind156b.wi.leidenuniv.nl:2000/cgi-bin/image_demo6.0.pl.



Εικόνα 23: Το σύστημα LCPD

5.23 MARS

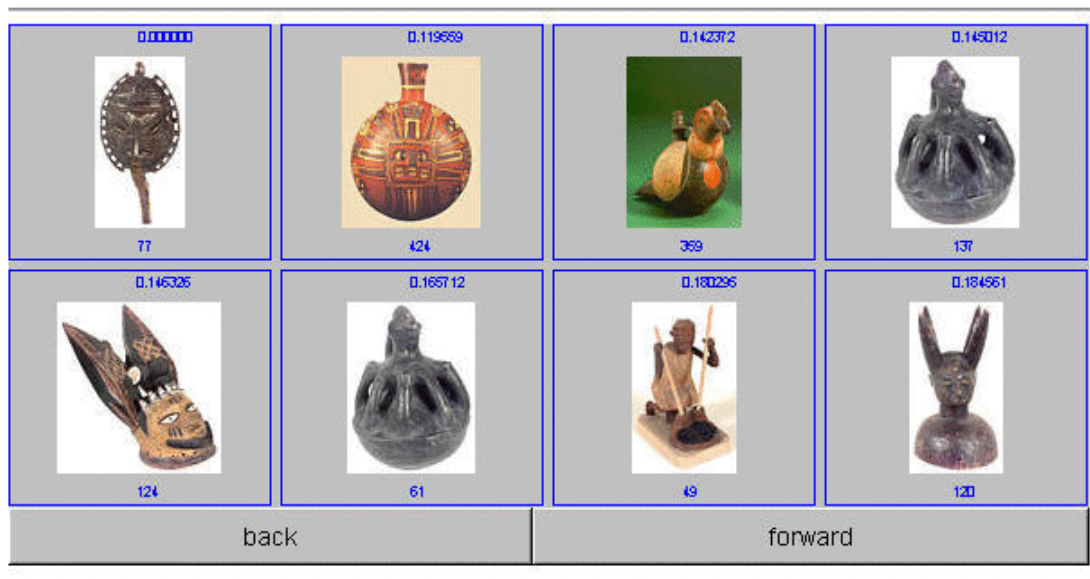
Το MARS (Multimedia Analysis and Retrieval System) (Ortega et al, 1997) δημιουργήθηκε στο Πανεπιστήμιο του Ιλινόις (Department of Computer Science, University of Illinois at Urbana-Champaign) και αναπτύχθηκε περαιτέρω στο Πανεπιστήμιο της Καλιφόρνια (Department of Information and Computer Science,

University of California at Irvine). Αποτελεί μια ερευνητική προσπάθεια που συνδυάζει πολλές ερευνητικές κοινότητες: όραση των υπολογιστών, συστήματα διαχείρισης βάσεων δεδομένων και ανάκτηση πληροφορίας.

Το σύστημα αυτό υποστηρίζει ερωτήματα σε συνδυασμούς για χαρακτηριστικά χαμηλού επιπέδου και περιγραφές κειμένων. Τα χαρακτηριστικά εξάγονται είτε από το σύνολο της εικόνας είτε από τμήματα της εικόνας (η εικόνα χωρίζεται σε 20 (5 x 5) τμήματα. Μπορούν να διεξαχθούν σύνθετα ερωτήματα με τη χρήση των τελέστων Boolean. Τα επιθυμητά χαρακτηριστικά μπορούν να δηλωθούν είτε με παράδειγμα (επιλέγοντας μια εικόνα της βάσης δεδομένων) ή άμεσα (π.χ. επιλέγοντας χρώματα από ένα χρωματολόγιο ή κάποια υφή από ένα σύνολο προτύπων).

Το σύστημα υλοποιεί την αρχιτεκτονική ανατροφοδότησης σχετικότητας από το χρήστη. Ο χρήστης καθορίζει τη βαρύτητα των χαρακτηριστικών και αναπροσαρμόζει το ερώτημα. Οι εικόνες παρουσιάζονται κατά την ανάκτηση με σειρά φθίνουσας ομοιότητας. Για περισσότερες πληροφορίες σχετικά με το MARS μπορεί κανείς να επισκεφτεί στο διαδίκτυο την παρακάτω ιστοσελίδα: <http://www-db.ics.uci.edu/pages/research/mars.shtml>.





Εικόνα 24: Το σύστημα MARS

5.24 Metaseek

Το σύστημα *Metaseek* (Benitez et al, 1998) αναπτύχθηκε από το Πανεπιστήμιο Κολούμπια (*Image and Advanced Television Lab, Columbia University*) στις Ηνωμένες Πολιτείες. Η ανάκτηση εικόνας διεξάγεται με βάση το χρώμα και τη δομή. Ο χρήστης έχει τη δυνατότητα να επιλέξει μια κατηγορία, να δώσει μια λέξη κλειδί, να παρέχει ένα URL μιας εικόνας ή να επιλέξει μια εικόνα που παρουσιάζεται. Ο χρήστης μπορεί να δηλώσει αν τον ικανοποιεί ή όχι για καθεμιά από τις ανακτημένες εικόνες και να ανατροφοδοτήσει την αναζήτηση. Πληροφορίες για το *Metaseek* μπορεί να δει κανείς στην ιστοσελίδα <http://www.ctr.columbia.edu/metaseek>.



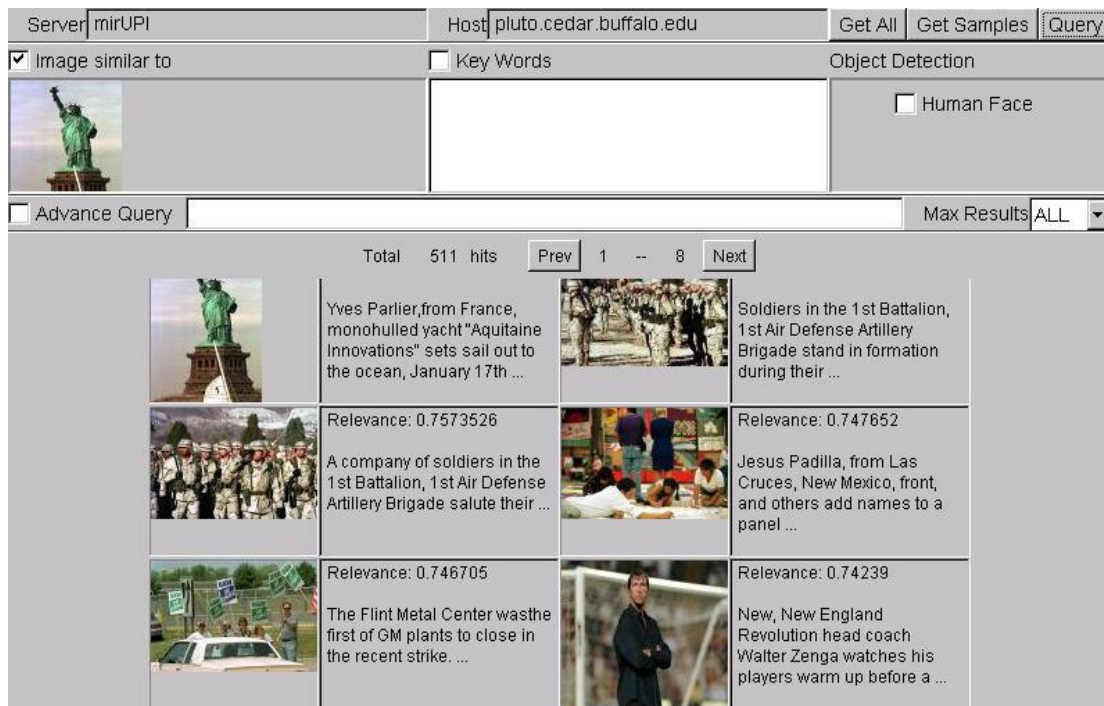
Εικόνα 25: Το σύστημα METASEEK

5.25 MIR

Το σύστημα MIR (Srihari et al, 1995) αναπτύχθηκε από το Center of Excellence for Document Analysis and Recognition στο πανεπιστήμιο του Buffalo των Ηνωμένων Πολιτειών. Το MIR σύστημα περιέχει διάφορες τεχνικές για την επεξεργασία του κειμένου και της εικόνας σε μια προσπάθεια να εξάγει σημασιολογικές περιγραφές των εικόνων. Με τη χρησιμοποίηση τεχνικών επεξεργασίας φυσικής γλώσσας (NLP techniques) στην ανάλυση εικόνων, μπορούν να εξαχθούν πληροφορίες για το περιεχόμενο της εικόνας. Αυτές οι πληροφορίες μπορούν να είναι οι εξής:

- Ύπαρξη ανθρώπων στις εικόνες
- Τα ονόματα των εικόνων
- Περιοχή και χρόνος λήψης της φωτογραφίας
- Χωρικές σχέσεις των ανθρώπων και
- Άλλα οπτικά χαρακτηριστικά των ανθρώπων

Πληροφορίες για το MIR μπορεί να δει κανείς στην ιστοσελίδα <http://www.cedar.buffalo.edu/MMIR/>. Μια δοκιμαστική έκδοση του συστήματος υπάρχει στην ιστοσελίδα <http://www.cedar.Buffalo.EDU/MMIR/demo/index.html>.



Εικόνα 26: Το σύστημα MIR

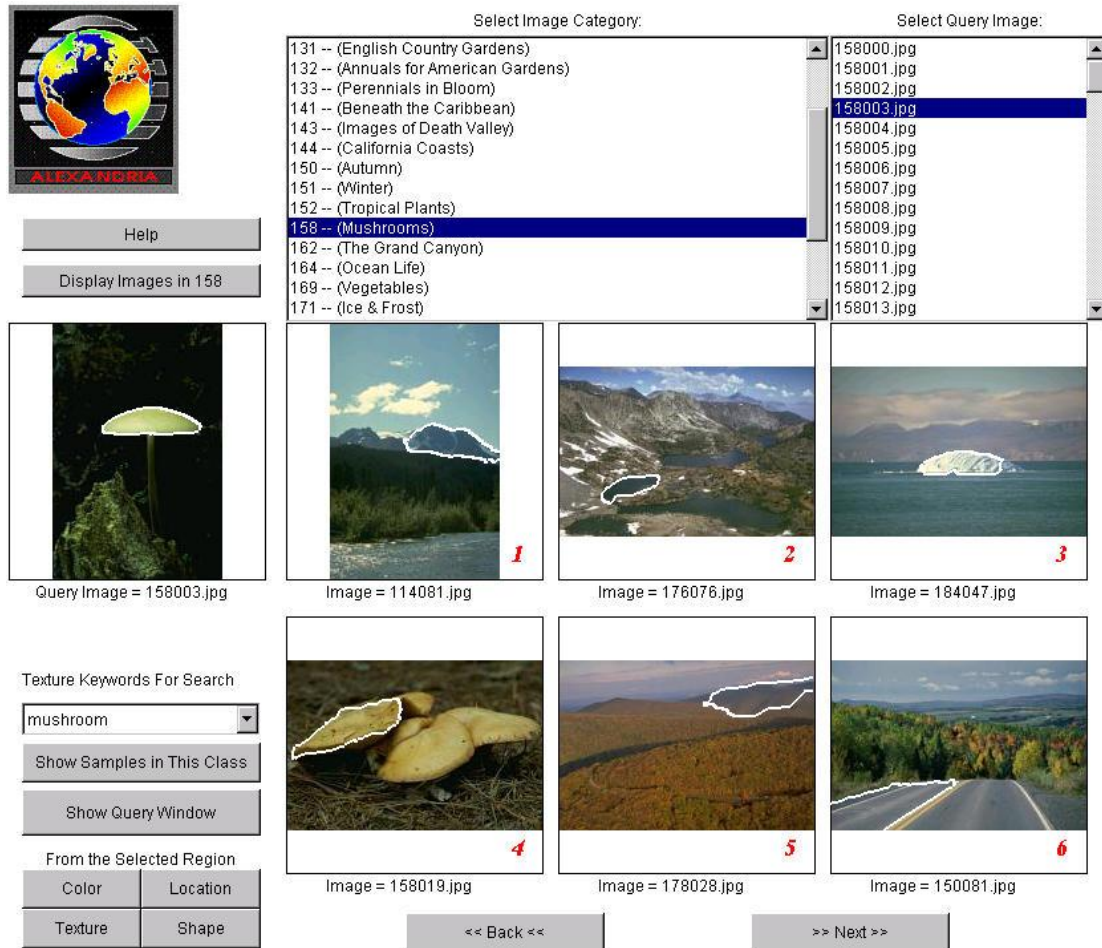
5.26 Netra

Το σύστημα NETRA (Ma, 1997) αναπτύχθηκε από το τμήμα Ηλεκτρολόγων και Μηχανικών Υπολογιστών του πανεπιστημίου της California των Ηνωμένων Πολιτειών. Οι εικόνες στη βάση δεδομένων είναι χωρισμένες σε περιοχές ομογενών χρωμάτων. Από αυτές τις περιοχές εξάγονται τα εξής χαρακτηριστικά:

- Χρώμα
- Υφή
- Σχήμα και
- Χωρική τοποθεσία

Στη βάση δεδομένων υπάρχουν 2500 εικόνες από τη συλλογή φωτογραφιών Corel και είναι οργανωμένες σε 25 κατηγορίες των 100 εικόνων η καθεμία. Ο χρήστης μπορεί να επιλέξει μία από αυτές τις εικόνες ως την εικόνα για ερώτηση. Επίσης, ο χρήστης μπορεί να επιλέξει μία περιοχή μιας εικόνας και να επιλέξει μία από τις τέσσερις ιδιότητες, όπως αναφέρθηκε παραπάνω.

Περισσότερες πληροφορίες για αυτό το σύστημα διατίθενται στον Παγκόσμιο Ιστό στην ιστοσελίδα <http://maya.ece.ucsb.edu/Netra/>. Επιπρόσθετα, παρουσίαση των εφαρμογών του συστήματος αυτού πραγματοποιείται στον ιστότοπο <http://maya.ece.ucsb.edu/Netra/netra.html>.



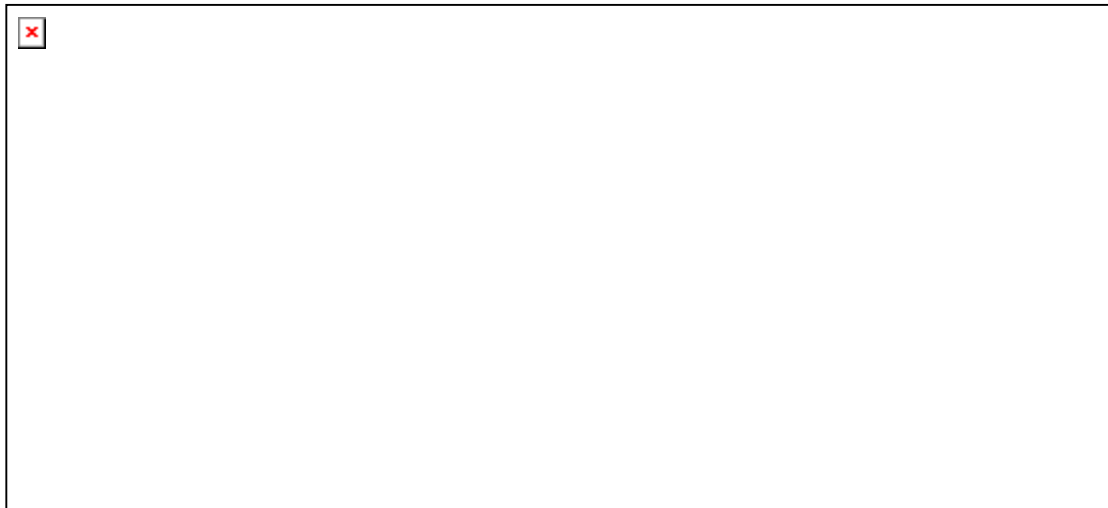
Εικόνα 27: Το σύστημα NETRA

5.27 Photobook

Το Photobook (Pentland et al, 1996) υλοποιήθηκε στο Πανεπιστήμιο MIT (Vision and Modeling Group, Media Laboratory, Massachusetts Institute of Technology). Εξάγει χαρακτηριστικά με βάση το σχήμα και την υφή. Για την εκτέλεση ενός ερωτήματος ο χρήστης επιλέγει κάποιες εικόνες από αυτές που παρουσιάζονται, επιπλέον μπορεί να ανατροφοδοτήσει την ερώτηση επιλέγοντας κάποια άλλη από τις εικόνες.

Το σύστημα *FourEyes* αποτελεί μια πιο πρόσφατη αναβάθμιση του *Photobook* που συνδυάζει τα χαρακτηριστικά. Αντί να επιτρέπει στο χρήστη να καθορίσει τη βαρύτητα στα χαρακτηριστικά, στρέφεται στη διαδραστικότητα του χρήστη με τα δεδομένα και μαθαίνει τα καλύτερα γνωρίσματα αφήνοντας το χρήστη να δίνει θετικά και αρνητικά παραδείγματα εικόνων.

Παρόλο που το *Photobook* δεν έγινε εμπορικό προϊόν από μόνο του, η τεχνολογία του για την αναγνώριση προσώπων εφαρμόζεται σε πολλά αστυνομικά τμήματα στις ΗΠΑ. Περισσότερες πληροφορίες για αυτό το σύστημα διατίθενται στον Παγκόσμιο Ιστό στην παρακάτω ιστοσελίδα <http://vismod.www.media.mit.edu/vismod/demos/photobook/>. Επιπρόσθετα, παρουσίαση των εφαρμογών του συστήματος αυτού πραγματοποιείται στον ιστότοπο <http://vismod.www.media.mit.edu/cgi-bin/tpminka/query?vistex...10>.



Εικόνα 28: Το σύστημα Photobook

5.28 PicSOM

Το *PicSOM* αναπτύχθηκε στη Φιλανδία από το Πολυτεχνείο του Ελσίνκι (*Laboratory of Computer and Information Sciences, Helsinki University of Technology*). Το σύστημα εξάγει γνωρίσματα σχετικά με το χρώμα, το σχήμα και την υφή χωρίζοντας την εικόνα σε πέντε περιοχές (μια κεντρική περιοχή και τέσσερις περιοχές στις άκρες).

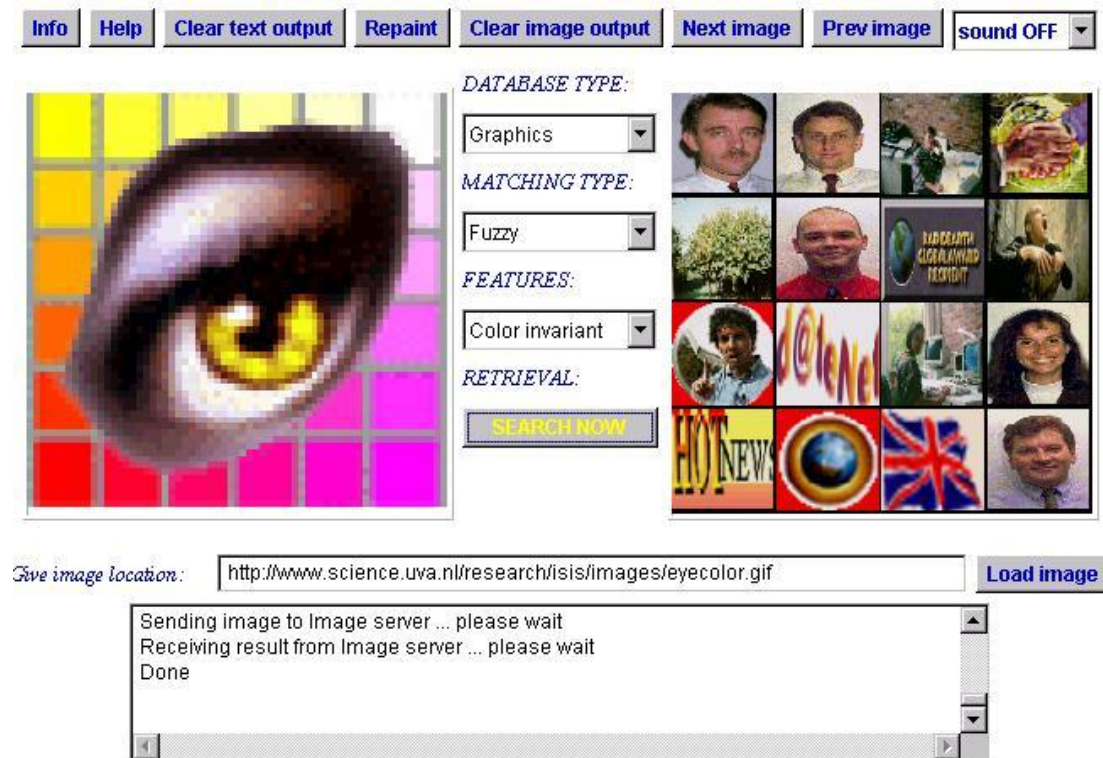
Το ερώτημα ξεκινά από μια σταθερή επιλογή ανάμεσα σε αντιπροσωπευτικές εικόνες. Η βελτίωση του ερωτήματος επιτυγχάνεται με ανατροφοδότηση. Ο χρήστης μπορεί να αποδώσει θετική ή αρνητική αξία στις εικόνες ανάλογα με το αν τις βρίσκει παρόμοιες ή όχι. Στο διαδίκτυο πληροφορίες καθώς και παρουσίαση του συστήματος αυτού υπάρχουν στο παρακάτω URL: <http://www.cis.hut.fi/picsom/>.



Εικόνα 29: Το σύστημα PicSOM

5.29 PicToSeek

Το PicToSeek (Gevers et al, 2000) αναπτύχθηκε από το τμήμα Πληροφορικής του πανεπιστημίου του Άμστερνταμ στην Ολλανδία. Η εικόνα προς επερώτηση μπορεί να επιλεγεί από μία βάση δεδομένων ή εισάγοντας τη URL διεύθυνση της εικόνας. Πριν την αποστολή του ερωτήματος, ο χρήστης πρέπει να δώσει μερικές επιλογές, που αφορούν τον τρόπο με τον οποίο το σύστημα θα αναζητήσει όμοιες εικόνες. Επίσης, επιτρέπονται ερωτήματα με παραδείγματα. Οι εικόνες που επιστράφηκαν εμφανίζονται σε αύξουσα σειρά ανάλογα με το βαθμό ομοιότητάς τους. Περισσότερες πληροφορίες για το PicToSeek υπάρχουν στην ιστοσελίδα: <http://www.science.uva.nl/research/isis/pictoseek/>.



Εικόνα 30: Το σύστημα PicToSeek

5.30 QBIC

Το QBIC (*Query By Image Content*) (Niblack et al, 1993) αναπτύχθηκε από την IBM (IBM Almaden Research Center, San Jose). Είναι διαθέσιμο για εμπορική χρήση είτε από μόνο του είτε ως τμήμα άλλων προϊόντων της IBM όπως το DB2 Digital Library. Προσφέρει ανάκτηση με οποιοδήποτε συνδυασμό χαρακτηριστικών χρώματος, σχήματος και υφής. Τα ερωτήματα μπορούν να διαμορφωθούν με επιλογή από ένα δειγματολόγιο χρωμάτων ή προτύπων υφής ή με ορισμό εικόνας ως παράδειγμα ή με τη σχεδίαση ενός σχήματος στην οθόνη.

Κάθε εικόνα που έχει ανακτηθεί μπορεί να χρησιμεύσει ως παράδειγμα για ανατροφοδότηση. Τα αποτελέσματα εκτίθενται με φθίνουσα σειρά ομοιότητας και – προαιρετικά – τη βαθμολογία τους (που δείχνει σε τι ποσοστό ταιριάζουν με τα χαρακτηριστικά που δόθηκαν κατά την ερώτηση). Υπάρχει επίσης η δυνατότητα χρήσης λέξεων κλειδιών και ο συνδυασμός με τα άλλα χαρακτηριστικά. Περισσότερες πληροφορίες για το QBIC υπάρχουν στην ιστοσελίδα: <http://wwwqbic.almaden.ibm.com>.



Εικόνα 31: Το σύστημα QBIC

5.31 QuickLook

Το σύστημα *QuickLook* (Ciocca et al, 2000) αναπτύχθηκε από το ινστιτούτο *CNR Institute of Multimedia Information Technologies* στο Μιλάνο της Ιταλίας. Ο χρήστης, αρχικά, επιλέγει την εικόνα με την οποία θα γίνει η ερώτηση. Τα χαρακτηριστικά της εικόνας εξάγονται από την πρώτη αναζήτηση. Ο χρήστης μπορεί να υποδείξει ποιες εικόνες είναι όμοιες. Από τις όμοιες αυτές εικόνες, μπορεί να γίνει ένα νέο ερώτημα. Τα αποτελέσματα εμφανίζονται σε φθίνουσα σειρά ομοιότητας. Επιπρόσθετα, παρουσίαση των εφαρμογών του συστήματος αυτού πραγματοποιείται στον ιστότοπο <http://www.itim.mi.cnr.it/Linee/Linea1/Sottolinea3/relfeme.htm>.



Εικόνα 32: Το σύστημα QuickLook

5.32 SIMBA

Το σύστημα SIMBA (Siggelkow et al, 2001) αναπτύχθηκε από το ινστιτούτο *Pattern Recognition and Image Processing* στο πανεπιστήμιο του Freiburg της Γερμανίας. Αρχικά, παρουσιάζεται στο χρήστη μία συλλογή από 25 τυχαίες εικόνες από τη βάση δεδομένων. Ο χρήστης επιλέγει μία από αυτές και στέλνει στο σύστημα ένα ερώτημα για παρουσίαση παρόμοιων εικόνων. Εναλλακτικά, ο χρήστης μπορεί να εισάγει μία διεύθυνση URL με μία δική του εικόνα, της οποίας το μέγεθος πρέπει να είναι συγκρίσιμο με το μέγεθος των εικόνων της βάσης δεδομένων (384 x 256).

Περισσότερες πληροφορίες για αυτό το σύστημα διατίθενται στον Παγκόσμιο Ιστό στην ιστοσελίδα <http://lmb.informatik.uni-freiburg.de/SIMBA>. Επιπρόσθετα, παρουσίαση των εφαρμογών του συστήματος αυτού πραγματοποιείται στον ιστότοπο <http://obelix.informatik.uni-freiburg.de/cgi-bin/SIMBA/SIMBA.cgi>.

SIMBA - Search IMages By Appearance

on the MPEG-7 test set of 2343 images

Query Interface:

Please select number of rows and columns and subsampling rate for presenting results.

Then provide the url of your favorite JPEG image (note: the image will be compared unscaled, the database images have size 384x256)

URL

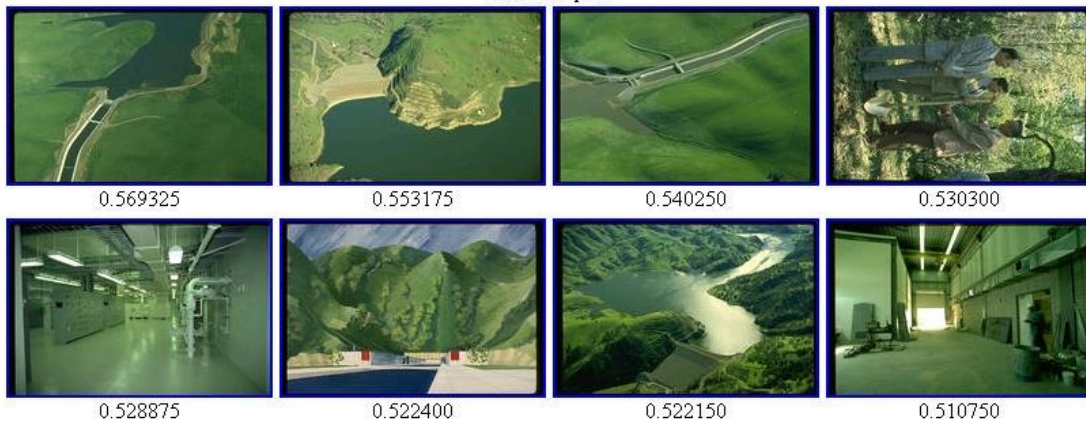
or use to get random images from the database

or select an image from the results' section to look for similar ones.

Results:



Search template

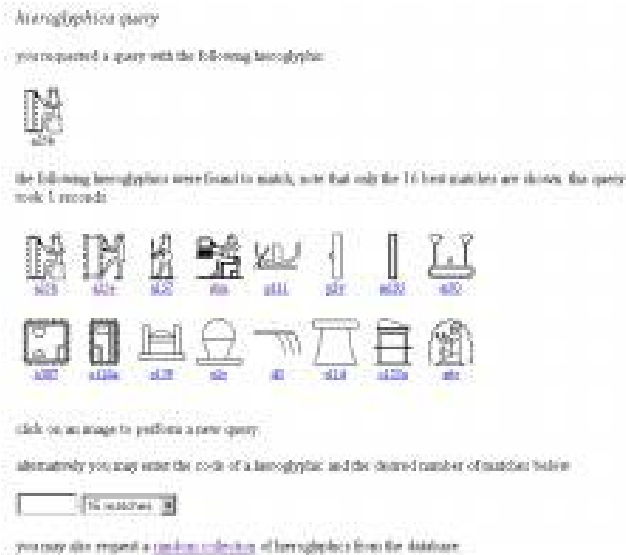


Εικόνα 33: Το σύστημα SIMBA

5.33 SMURF

To SMURF (Similarity-based Multimedia Retrieval Framework) δημιουργήθηκε από το Πανεπιστήμιο της Ουτρέχτης στην Ολλανδία (Center for Geometry, Imaging, and Virtual Environments, Institute of Information and Computing Sciences, Utrecht University). Ανιχνεύει γνωρίσματα αναφορικά με το χρώμα, το σχήμα και τη δομή των εικόνων.

Η αναζήτηση αρχίζει με την επιλογή μιας εικόνας από μια τυχαία επιλογή. Ο χρήστης επιλέγει μια νέα εικόνα από τα αποτελέσματα για να φτάσει σταδιακά στην επιθυμητή εικόνα. Τα αποτελέσματα παρουσιάζονται σε φθίνουσα διάταξη ανάλογα με το βαθμό ομοιότητας. Στην παρακάτω ιστοσελίδα υπάρχουν περισσότερες πληροφορίες για το SMURF <http://give-lab.cs.uu.nl>.



Εικόνα 34: Το σύστημα SMURF

5.34 SIMPLIcity

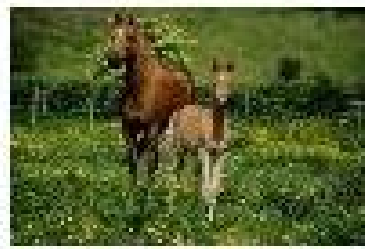
Το SIMPLIcity (*Semantic-sensitive Integrated Matching for Picture Libraries*) αναπτύχθηκε από ερευνητική ομάδα στο Πανεπιστήμιο του Στάνφορντ (*Stanford University*). Η εικόνα στο σύστημα αυτό αναπαριστάται ως ένα σύνολο από περιοχές που αντιστοιχούν γενικά σε αντικείμενα και χαρακτηρίζονται από το χρώμα, τη δομή, το σχήμα και τη θέση τους. Ταυτίζοντας τα αντικείμενα στην εικόνα, το σύστημα μπορεί ευκολότερα να αναγνωρίσει παρόμοια αντικείμενα σε διαφορετικές θέσεις και με διαφορετικό προσανατολισμό και μέγεθος.

Βασιζόμενο στις περιοχές της κατατμημένης εικόνας το σύστημα ταξινομεί τις εικόνες σε κάποιες σημασιολογικές κατηγορίες με σκοπό να γίνουν διακριτές διαφορές με σημασιολογικό νόημα, για παράδειγμα εσωτερικός – εξωτερικός χώρος, γραφικό – μη γραφικό, δομημένο – μη δομημένο μικραίνοντας το πεδίο έρευνας στη βάση δεδομένων επιτρέποντας σημασιολογικά αναπτύξιμες μεθόδους έρευνας.

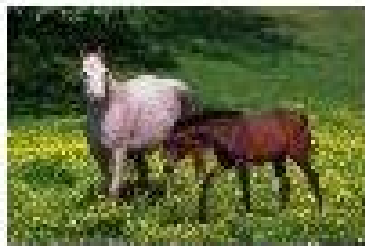
Ο χρήστης απλά κάνει κλικ σε μια εικόνα ή δηλώνει έναν αριθμό που έχει η εικόνα στη βάση δεδομένων ή μία διεύθυνση URL και ανακτά παρόμοιες φωτογραφίες. Περισσότερες πληροφορίες καθώς και παρουσίαση του συστήματος αυτού στην ιστοσελίδα <http://www-db.stanford.edu/IMAGE/>.



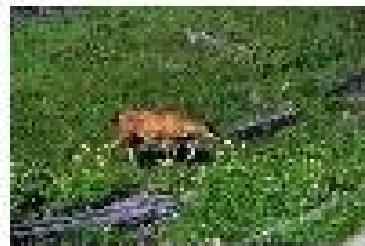
6258 12.59 4



36644 12.68 4



6229 13.31 5

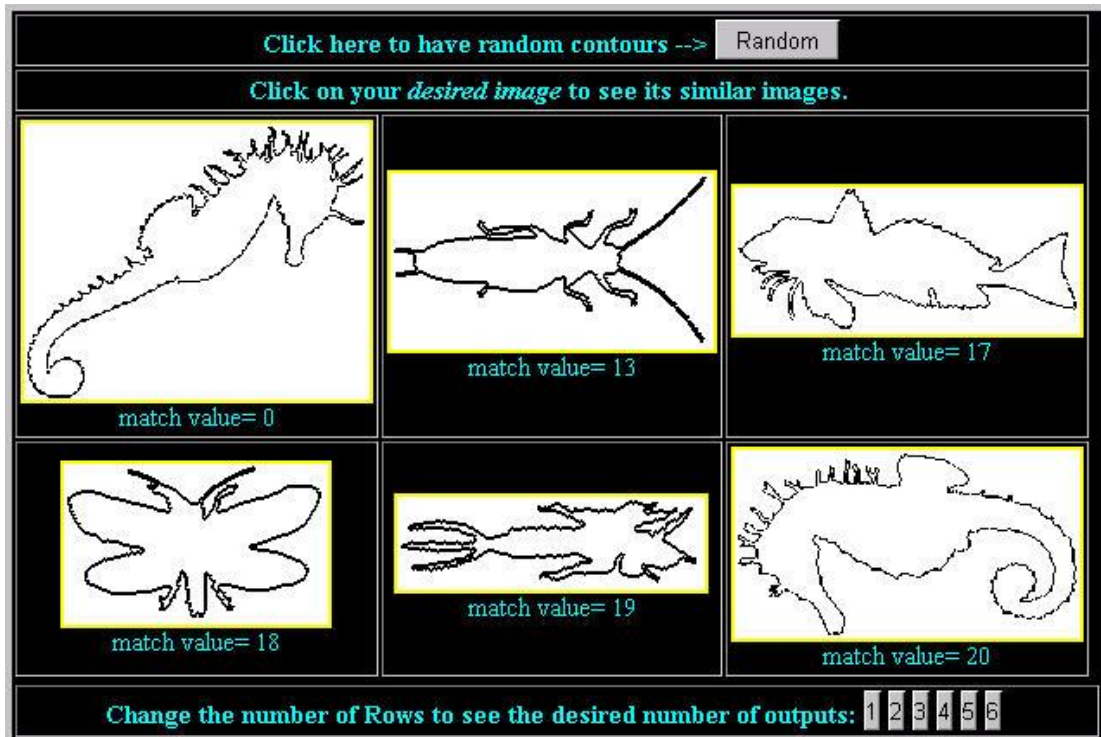


33769 13.45 3

Εικόνα 35: SIMPLIcity

5.35 SQUID

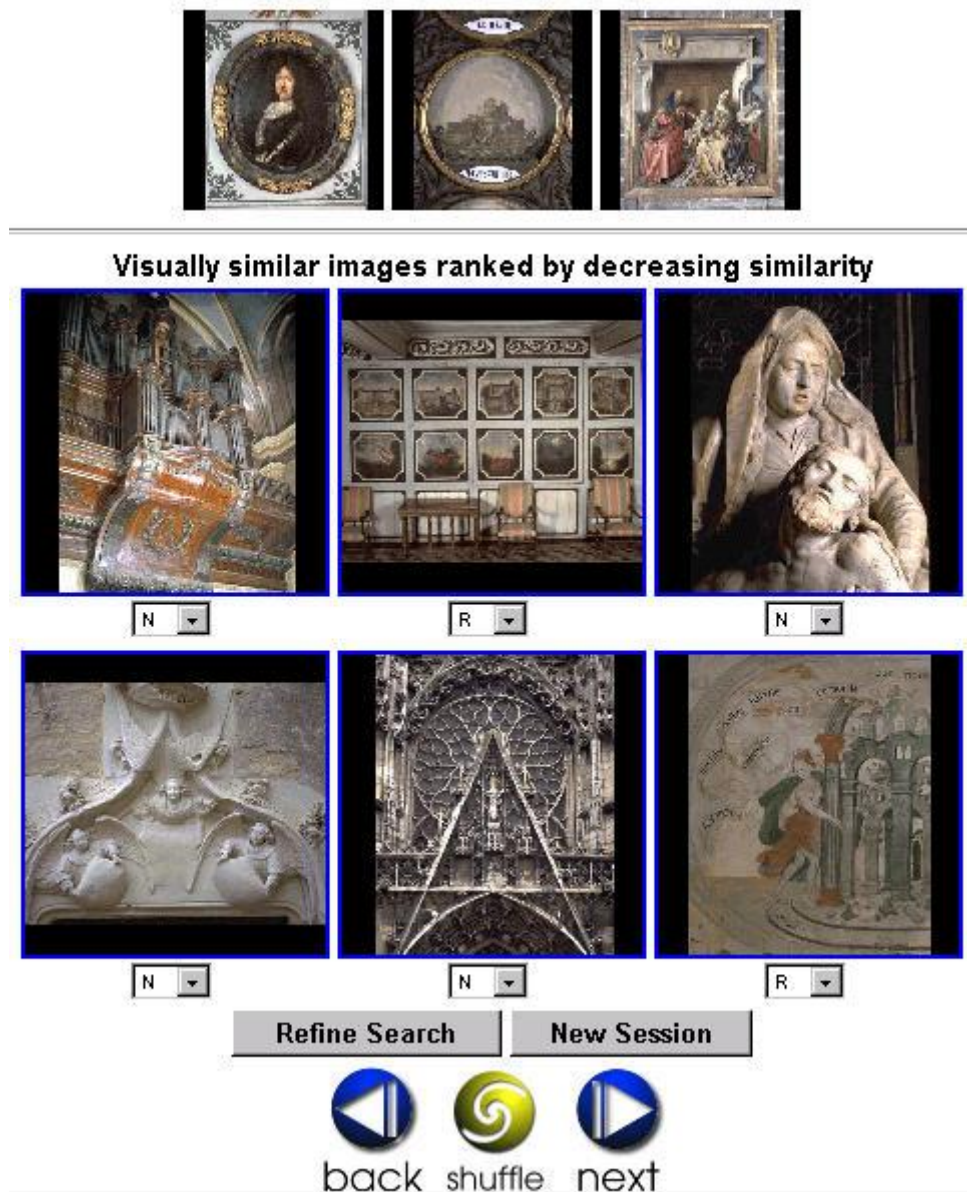
Το σύστημα SQUID (Mokhtarian et al, 1996) αναπτύχθηκε από το Centre for Vision, Speech, and Signal Processing στο πανεπιστήμιο του Surrey του Ηνωμένου Βασιλείου. Ο χρήστης μπορεί να επιλέξει μία εικόνα από τις εικόνες αντικειμένων που εμφανίζονται και να δημιουργήσει ένα ερώτημα για παρόμοιες εικόνες. Το σύστημα εμφανίζει τις πρώτες εικόνες που μοιάζουν με την εικόνα προς επερώτηση. Η βάση δεδομένων του λογισμικού αποτελείται από 1100 εικόνες. Περισσότερες πληροφορίες καθώς και παρουσίαση του συστήματος αυτού στην ιστοσελίδα <http://www.ee.surrey.ac.uk/Research/VSSP/imagedb/demo.html>.



Εικόνα 36: Το σύστημα SQUID

5.36 SURFIMAGE

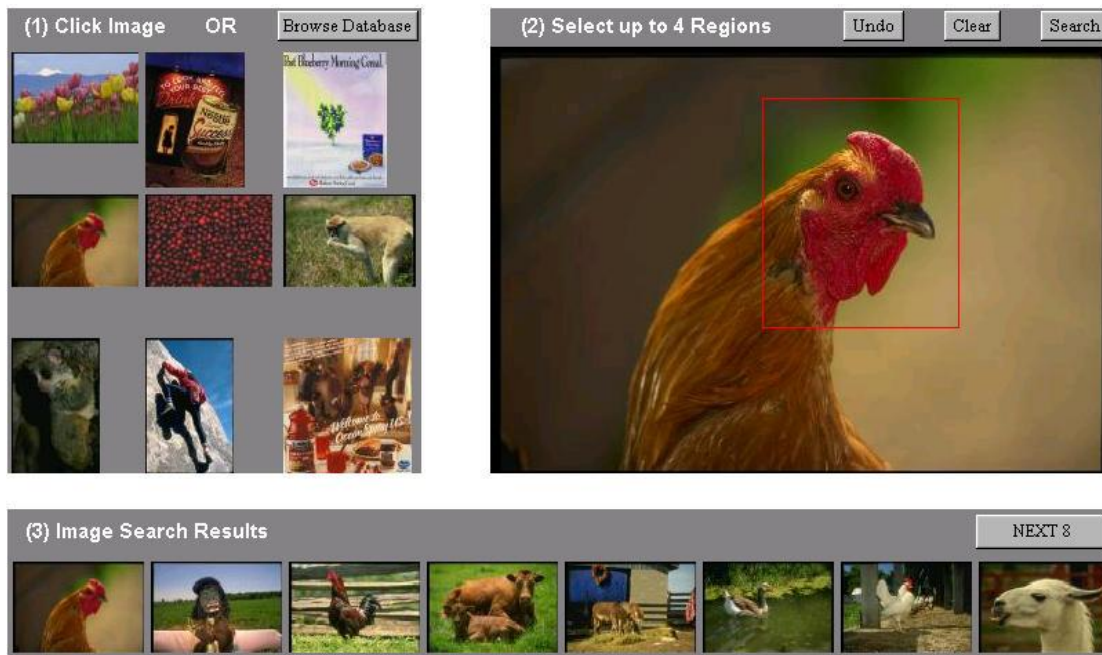
Το σύστημα *SurfImage* (Nastar et al, 1998) αναπτύχθηκε από το ινστιτούτο INRIA της Γαλλίας. Ο χρήστης μπορεί να επιλέγει κάποιο χαρακτηριστικό μιας εικόνας και κατάλληλους δείκτες ομοιότητας από τις λίστες που απεικονίζονται, καθορίζοντας τη σημασία που θα έχει κάθε χαρακτηριστικό στην ανάκτηση των όμοιων εικόνων. Το λογισμικό διαθέτει πολλές βάσεις δεδομένων, όπως η βάση MIT που περιέχει πρόσωπα, η βάση δεδομένων BTphoto που διαθέτει εικόνες από πόλεις και η βάση δεδομένων MIT Vistex που διαθέτει εικόνες σχετικές με την υφή των αντικειμένων. Περισσότερες πληροφορίες καθώς και παρουσίαση του συστήματος αυτού στην ιστοσελίδα http://www-rocq.inria.fr/imedia/index_UK.html.



Εικόνα 37: Το σύστημα SurfImage

5.37 SYNAPSE

Το σύστημα SYNAPSE (Manmatha et al, 1997) αναπτύχθηκε από το Center for Intelligent Information Retrieval του University of Massachusetts. Το λογισμικά επιτρέπει στους χρήστες να εκτελούν ερωτήματα με παραδείγματα και τα αποτελέσματα εμφανίζονται σε φθίνουσα σειρά σχετικά με το σκορ ομοιότητας. Περισσότερες πληροφορίες καθώς και παρουσίαση του συστήματος αυτού στην ιστοσελίδα <http://hobart.cs.umass.edu/~mmedia/>.



Εικόνα 38: Το σύστημα Synapse

5.38 TODAI

Το σύστημα TODAI (Michel et al, 1996) αναπτύχθηκε από το τμήμα Ηλεκτρολόγων Μηχανικών του πανεπιστημίου EPFL της Ελβετίας και στη συνέχεια από το εργαστήριο Intelligent Systems Laboratory του πανεπιστημίου του Halmstad της Σουηδίας. Ο χρήστης μπορεί να επιλέξει μία τοποθεσία από την εικόνα προς επερώτηση. Πρόσθετα κριτήρια, όπως το μέγεθος, μπορούν να καθοριστούν από το χρήστη. Η ομοιότητα δύο εικόνων υπολογίζεται από το άθροισμα της ευκλείδειας απόστασης των έξι χαρακτηριστικών διανυσμάτων. Περισσότερες πληροφορίες καθώς και παρουσίαση του συστήματος αυτού στην ιστοσελίδα http://www.unil.ch/BCU/docs/collecti/res_prec/en/wtres/todai_doc.html.

Submitting an ornament


Select the ornament *browser* and click on *Open picture*:

H:\art\match\cbir-survey\G1.gif Browse...

jpg Open picture
 tiff
 gif

Banque internationale d'ornements d'imprimerie
 International Bank of Printers' Ornaments
 International Bank von Buchdruck Ornamente
 Banca Internazionale d'ornamenti di stamperia
Only working with Netscape >=4.5

? % **Selected picture G1**



Find reset More criteria

More criteria

Ornament's use	ornament's nature
<input type="checkbox"/> fleuron/vignette	<input type="checkbox"/> woodcut/cliché
<input type="checkbox"/> headpiece and rule	<input type="checkbox"/> intaglio
<input type="checkbox"/> decorative initial and factotum	<input type="checkbox"/> typographical composition
<input type="checkbox"/> frame	<input type="checkbox"/> mixt
<input type="checkbox"/> ornamental pagination	

Site

width X height width
 [mm] [mm] height

X

devise

signature(s):

Seek by number(s)

ex: 1,3,4
 enter *rech* to look for the previous found

Εικόνα 39: Το σύστημα Today

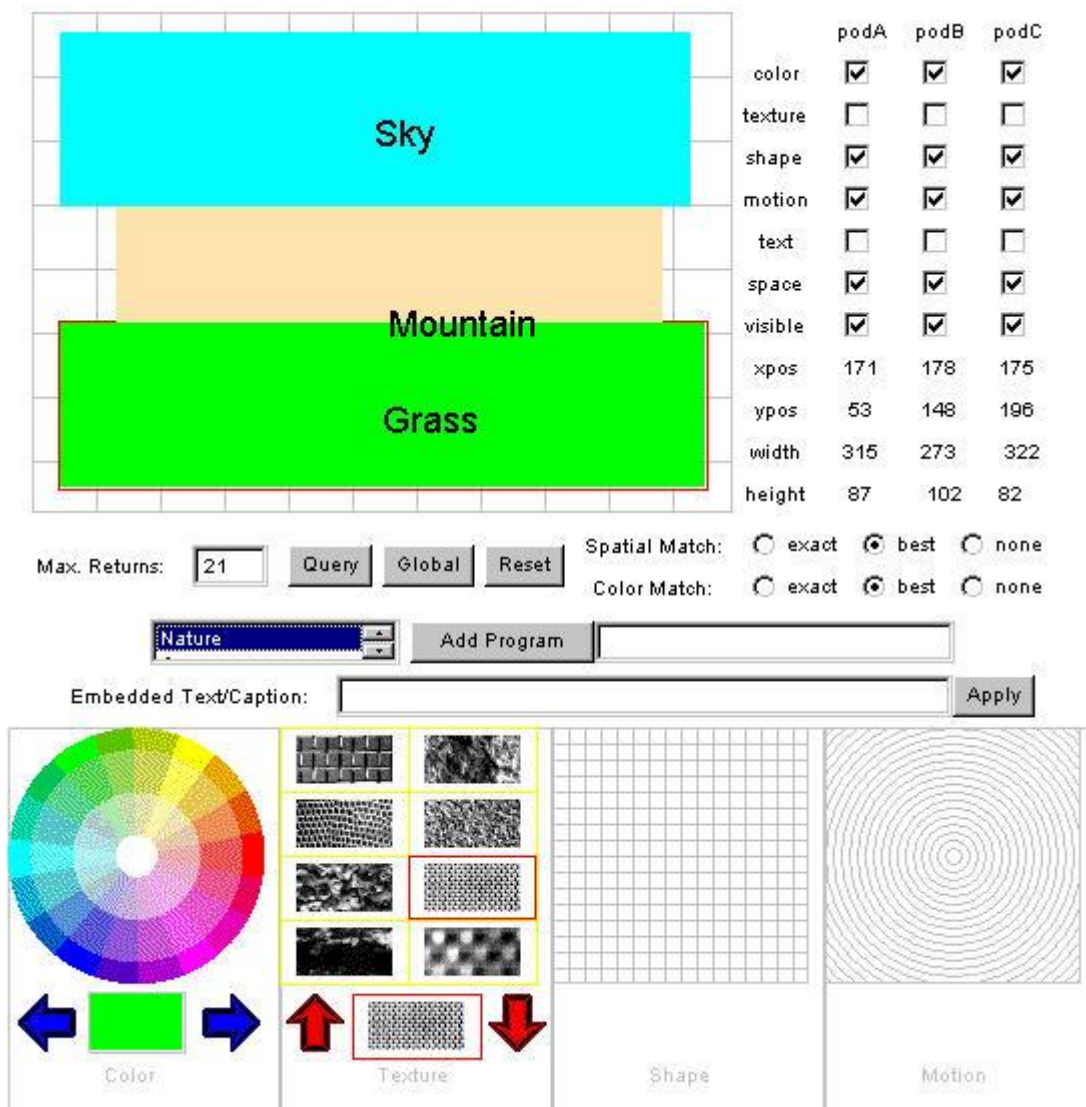
5.39 VISUALSEEK

Το σύστημα VISUALSEEK (Smith et al, 1997) αναπτύχθηκε από το εργαστήριο Image and Advanced Television Lab στο πανεπιστήμιο του Columbia των Ηνωμένων Πολιτειών. Κάθε εικόνα διαχωρίζεται σε περιοχές από εξίσου κυρίαρχα θέματα. Για κάθε περιοχή διατηρούνται χαρακτηριστικές και χωρικές πληροφορίες. Για κάθε ένα

ερώτημα ουσιαστικά εκτελείται μία αναζήτηση και εμφανίζονται οι εικόνες που περιέχουν τις περισσότερες αντιστοιχίες από παρόμοιες περιοχές.

Για τη δημιουργία ενός ερωτήματος, ο χρήστης επιλέγει ένα αριθμό περιοχών, θέσεων και διαστάσεων και επιλέγει ένα χρώμα για κάθε περιοχή. Επίσης, ο χρήστης καθορίζει τα όρια κάθε περιοχής και το μέγεθος ή/και τις χωρικές σχέσεις μεταξύ των περιοχών. Στη συνέχεια, το σύστημα επιστρέφει μικρογραφίες εικόνων από τις καλύτερες αντιστοιχίσεις και ο χρήστης μπορεί να αναζητήσει με παράδειγμα μία από αυτές τις εικόνες.

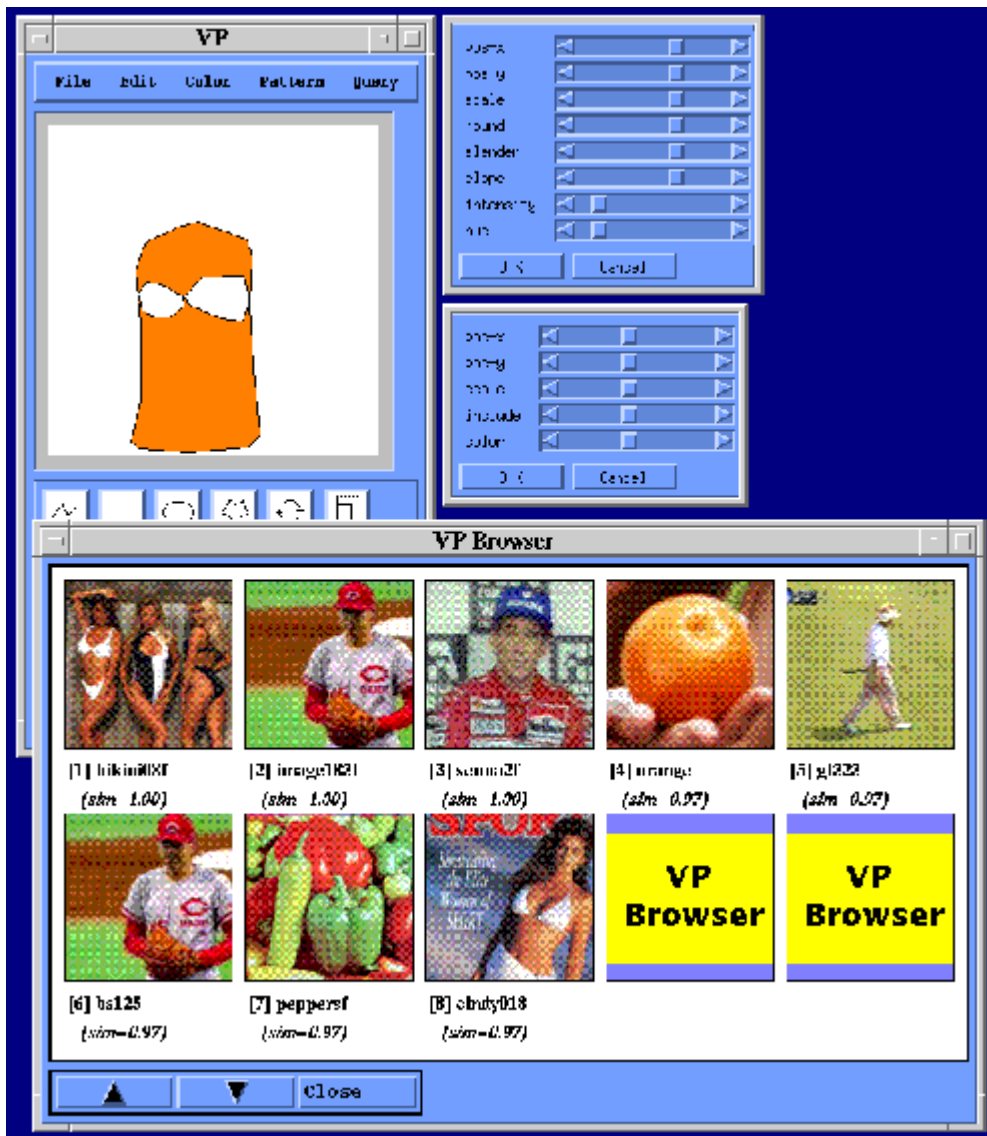
Περισσότερες πληροφορίες καθώς και παρουσίαση του συστήματος αυτού στην ιστοσελίδα <http://www.ctr.columbia.edu/VisualSEEK/>.



Εικόνα 40: Το σύστημα VISUALSEEK

5.40 VP Image Retrieval System

Το σύστημα VP Image Retrieval System (Kanehara, 1995) έχει αναπτυχθεί από το National Center for Science Information Systems του πανεπιστημίου του Τόκιο της Ιαπωνίας. Το λογισμικό απαιτεί από το χρήστη να επιτελέσει τρία βήματα για τη διεξαγωγή ενός ερωτήματος. Αρχικά, ο χρήστης σχεδιάζει ένα προσχέδιο στον καμβά, το οποίο αυτόματα καταταμείται σε τμήματα. Τελικά, ο χρήστης δίνει ένα ξεχωριστό βάρος σε καθένα από αυτά τα τμήματα. Τα βάρη αυτά καθορίζουν το βαθμό ομοιότητας για τη διαδικασία της επερώτησης από τις ιδιότητες και τις σχέσεις που έχουν τα τμήματα της επερώτησης. Περισσότερες πληροφορίες καθώς και παρουσίαση του συστήματος αυτού στην ιστοσελίδα <http://www.rd.nacsis.ac.jp/>.



Εικόνα 41: Το σύστημα VP Image Retrieval System

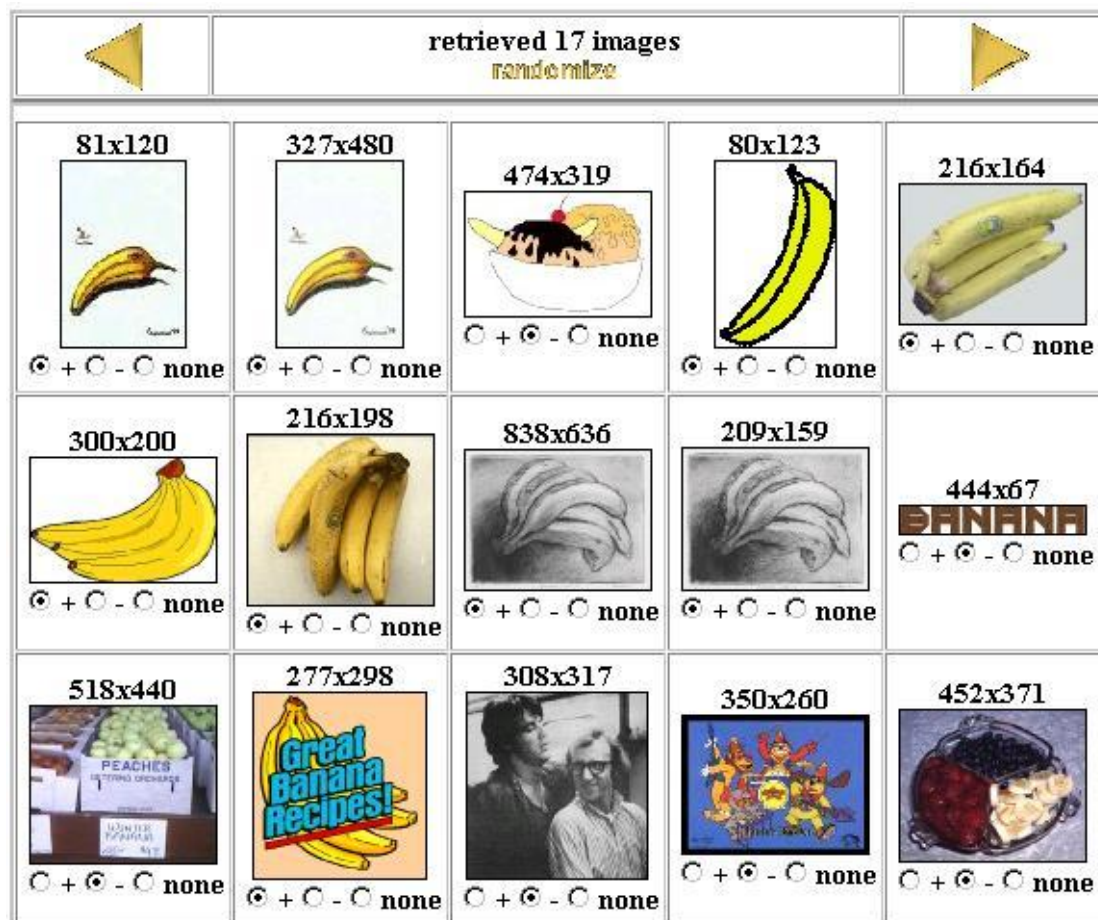
5.41 WEBSEEK

Το σύστημα Webseek (Smith, 1997) έχει αναπτυχθεί από το εργαστήριο Image and Advanced Television Lab του πανεπιστημίου της Columbia των Ηνωμένων Πολιτειών. Το σύστημα αυτό εκτελεί ερωτήματα με βάση ένα κείμενο ή μία εικόνα σε ένα κατάλογο από εικόνες και video που έχουν συλλεχθεί από το Διαδίκτυο. Το χρώμα αναπαρίσταται από ένα κανονικοποιημένο ιστόγραμμα στο χώρο HSV.

Ο χρήστης μπορεί να δημιουργήσει ένα ερώτημα επιλέγοντας ένα θέμα από τον διαθέσιμο κατάλογο ή εισάγει ένα θέμα. Τα αποτελέσματα του ερωτήματος μπορούν να χρησιμοποιηθούν για ένα ερώτημα με βάση το χρώμα σε ολόκληρο τον κατάλογο ή

μπορεί να ταξινομηθούν σε φθίνουσα σειρά σχετικά με τη χρωματική ομοιότητα του επιλεγμένου αντικειμένου. Επίσης, ο χρήστης έχει τη δυνατότητα να αλλάζει χειροκίνητα το ιστόγραμμα μίας εικόνας ή ενός video πριν επαναλάβει την αναζήτηση.

Περισσότερες πληροφορίες καθώς και παρουσίαση του συστήματος αυτού στην ιστοσελίδα <http://www.ctr.columbia.edu/WebSEEK/>.



Εικόνα 42: Το σύστημα WEBSEEK

5.42 WISE

Το σύστημα Wise (Wang et al, 1997) έχει αναπτυχθεί από το τμήμα Πληροφορικής του πανεπιστημίου του Stanford. Ένα ερώτημα στο λογισμικό γίνεται μέσω παραδείγματος, όπου η εικόνα προς επερώτηση μπορεί να είναι είτε μία εικόνα από τη βάση δεδομένων του λογισμικού είτε ένα σκίτσο που μπορεί να ζωγραφίσει ο χρήστης στον καμβά. Επίσης, εικόνες χαμηλής ανάλυσης ή τμήματα εικόνων μπορούν να χρησιμοποιηθούν ως εικόνες για τα ερωτήματα. Περισσότερες πληροφορίες καθώς

και παρουσίαση του συστήματος αυτού στην ιστοσελίδα <http://www-db.stanford.edu/~wangz/project/imsearch/>.

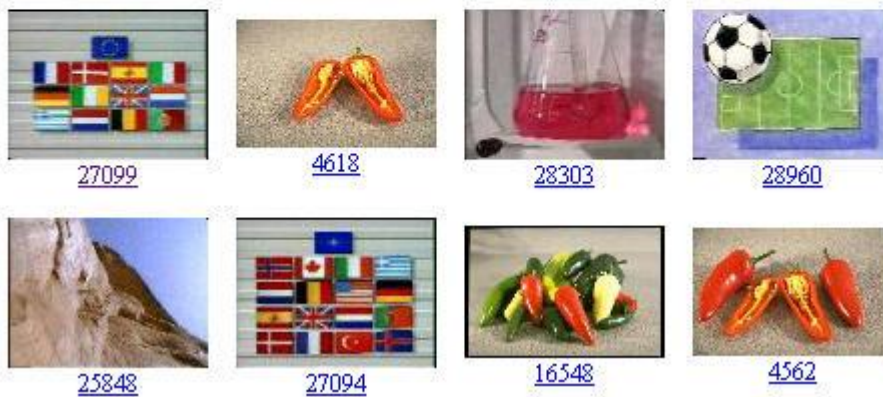
Searching Algorithm

Click one of the following images to search for similar images,

or click **Random** to get a random selection.

There are a total of 30,000 images in this database.

DONE! The search time is within 0.05 second CPU time on a PC. But it takes time to transmit the images.



Εικόνα 43: Το σύστημα WISE

ΚΕΦΑΛΑΙΟ 6: Συμπεράσματα

Ένα σημαντικό πρόβλημα που χρήζει αντιμετώπισης στην εποχή μας σχετικά με την πολυμεσική πληροφορία είναι σίγουρα η ανάκτηση της πολυμεσικής πληροφορίας από συστήματα βάσεων δεδομένων. Οι κατηγοριοποιήσεις που χρησιμοποιήθηκαν παλιότερα με τη χρήση απλών συστημάτων βάσεων δεδομένων ή απλά με την οπτική ή την αναγνώριση με βάση το κείμενο που χρησιμοποιούνταν έχουν πλέον απαρχαιωθεί και δεν προσφέρουν κανενός είδους διαχείρισης κι εύρεσης της κατάλληλης πληροφορίας λόγω του τεράστιου όγκου που διαχειρίζονται.

Στην παρούσα εργασία έγινε μια σε βάθος ανάλυση των χαρακτηριστικών που απαρτίζουν μια εικόνα καθώς επίσης και πως μπορεί αυτή η πληροφορία να γίνει διαχειρίσιμη. Παράλληλα έγινε μια παρουσίαση των συστημάτων ανάκτησης δεδομένων που υπάρχουν σήμερα καθώς και του προτύπου MPEG7.

Συμπερασματικά, ο τομέας της ανάκτησης πολυμεσικής πληροφορίας έχει πολύ δρόμο ακόμα για να μπορέσει να δώσει αποτελέσματα τα οποία θα είναι αδιαμφισβήτητα σωστά. Η αναζήτηση μιας πληροφορίας γενικότερα δεν είναι εύκολο πόσο μάλλον όταν πρόκειται για πολυμεσική πληροφορία.

Βήματα προόδου έχουν γίνει μετά την προτυποποίηση κάποιων κανόνων με το MPEG7, όμως, παρατηρούμαι ότι στο πεδίο της εικόνας λίγες είναι οι μηχανές αναζήτησης οι οποίες παράγουν αποτελέσματα σχετικά με την αναζήτησή μας. Έτσι, λοιπόν, υπάρχει ακόμη τεράστιο περιθώριο έρευνας στον τομέα αυτό. και σίγουρα οι λύσεις που θα μας προσφέρονται στο μέλλον θα είναι πολύ καλύτερες από τις σημερινές.

Βιβλιογραφία

Bach, J.R., Fuller, C., Gupta, A., Hampapur, A., Horowitz, B., Humphrey, R., Jain, R., And Shu, C.F. 1996. Virage image search engine: An open framework for image management. In Proceedings of the SPIE Storage and Retrieval for Still Image and Video Databases, California, USA, 76-87.

Ballard, D.H. And Brown, C.M. 1982. Computer Vision. Prentice Hall, New Jersey, USA. Levine, M. 1985. Vision in Man and Machine, Mcgraw Hill, Columbus.

Benitez, A. B., Beigi, M. and Chang, S. F. 1998. Using relevance feedback in content-based image metasearch. IEEE Internet Computing, 2(4):59-69.

Berman, A. and Shapiro, L. 1999. A flexible image database system for content-based retrieval. Computer Vision and Image Understanding, 75(1/2):175-195.

Bliujute, R., Saltenis, S., Slivinskas, G., And Jensen, C.S. 1999. Developing a DataBlade for a New Index. In Proceedings of IEEE International Conference on Data Engineering, IEEE, Sydney, March 1999, 314-323.

Carson, C., Belongie, S., Greenspan, H., & Malik, J. 2002. Blobworld: Image segmentation using expectation-maximization and its application to image querying. IEEE Trans. Pattern Anal. Mach. Intell., 24(8), 1026-1038.

Carson, C., Thomas, M., Belongie, S., Hellerstein, J. and Malik J. 1999. Blobworld: A system for region-based image indexing and retrieval. In Huijismans and Smeulders.

Ciocca, G., Gagliardi, I. and Schettini, R. 2000. *Quicklook2: An integrated multimedia system*, *International Journal of Visual Languages and Computing, Special Issue on Multimedia Databases and Image Communication*.

Cox T.F and Cox M.A.A. 2001. *Multidimensional Scaling*, Chapman and Hall, Second edition.

Das, M., Riseman, E. M. and Draper, B. 1997. *Focus: Searching for multi-colored objects in a diverse image database*. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition '97*, pages 756-761.

Egas, R., Huijismans, N., Lew, M.S., And Sebe, N. 1999. *Adapting kd Trees to Visual Retrieval*. In *Proceedings of the International Conference on Visual Information Systems, Amsterdam, June 1999*, A. SMEULDERS AND R. JAIN, Eds., 533-540.

Flickner, M. Sawhney, H. Niblack, W. Ashley, J. Qian Huang Dom, B. Gorkani, M. Hafner, J. Lee, D. Petkovic, D. Steele, D. Yanker, P. 1995. *Query by image and video content: the QBIC system*, *IEEE Computer*, September, 23-32.

Frankel, C., Swain, M.J., And Athitsos, V. 1996. *WebSeer: An Image Search Engine for the World Wide Web*. University of Chicago Technical Report 96-14, University of Chicago, USA.

Gevers, T. and Smeulders, A. 2000. *Pictoseek: Combining color and shape invariant features for image retrieval*. *IEEE Transactions on Image Processing*, vol. 9(1), pp. 102 – 119.

Greene, D., Parnas, M., & Yao, F. 1994. *Multi-index hashing for information retrieval. In Foundations of Computer Science, 1994 Proceedings., 35th Annual Symposium, pp. 722-731.*

Grossberg S. 2004. *Visual Boundaries and Surfaces, The Visual Neurosciences, Vol. 2, pp. 1624-1639, MIT Press.*

Haralick, R.M. And Shapiro, L.G. 1993. *Computer and Robot Vision. Addison-Wesley, New York, USA.*

Jolion, J.-M. 2001. *Feature similarity. pp. 121-143.*

Kai Essing, *Phd. Thesis (2007). Vision-Based Image Retrieval (VBIR)-A New Eye-Tracking Based Approach to Efficient and Intuitive Image Retrieval, University of Bielefeld.*

Kanehara, F., Satoh, S. and Hamada, T. 1995. *A flexible image retrieval using explicit visual instruction. In Proceedings of the Third International Conference on Document Analysis Recognition, Montreal, Canada, August '95, pp. 175 – 178.*

Kelly, P. M., Cannon, T. M. and Hush, D. R. 1995. *Query by image example: the CANDID approach. In SPIE Vol. 2420, Storage and Retrieval for Image and Video Databases III, pages 238-248.*

Laaksonen, J., Koskela, M., Oja, E. 2002. *PicSOM—Self-organizing image retrieval with MPEG-7 content descriptions. IEEE Transactions on Neural Networks, Special Issue on Intelligent Multimedia Processing.*

Levine, M. 1985. Vision in Man and Machine, Mcgraw Hill, Columbus.

Lew, M.S. 2000. Next Generation Web Searches for Visual Content. IEEE Computer, November, 46-53.

Lew, M.S. And Huijismans, N. 1996. Information Theory and Face Detection. In Proceedings of the International Conference on Pattern Recognition, Vienna, Austria, 601-605.

Lew, M. S., Huijismans, N. and Denteneer, D. 1996. Content based image retrieval: KLT, projections, or templates. In Smeulders and Jain, pages 27-34.

Lew, M., Lempinen, K. and Huijismans, N. 1997. Webcrawling using sketches. In Proceedings of the 2nd International Conference on Visual Information Systems, San Diego, December '97, pp. 77-84.

Li, Z. N., Zaïane, O. R, and Yan, B. 1998. C-bird: Content-based image retrieval from image repositories using chromaticity and recognition kernel. Technical Report CMPT98-03, School of Computing Science, Simon Fraser University, Canada.

Lowe, David G. 1999. Object recognition from local scale-invariant features. Proceedings of the International Conference on Computer Vision.

Ma, W. Y. 1997. NETRA: A Toolbox for Navigating Large Image Databases. PhD thesis, Dept. of Electrical and Computer Engineering, University of California at Santa Barbara.

Ma, W. Y. and Manjunath, B. S. 1999. *Netra: A toolbox for navigating large image databases. Multimedia Systems, vol. 7(3), pp. 184 – 198.*

Manmatha, R. and Ravela, S. 1997. *A syntactic characterization of appearance and its application to image retrieval. In Proceedings of the SPIE conference on Human Vision and Electronic Imaging II, Vol, 3016, San Jose, CA, Feb. '97.*

Manjunath, B. S. 1995. *Image browsing in the alexandria digital library project. D-Lib Magazine.*

Mehetre, B., Kankanhalli, M. & Lee, W. 1997. *Shape measures for content based image retrieval: A comparison. Information Processing and Management, 33, 319–337.*

Mezaris V., Kompatsiaris I., Strintzis M.G. 2003. *An ontology approach to object-based image retrieval, Proceedings of the ICIP, vol. II, pp. 511– 514.*

Michel, S., Karoubi, B., Bigun, J. and Corsini, S. 1996. *Orientation radiograms for indexing and identification in image databases. In Proceedings Eusipco-96, European conference on signal processing, pp. 1693 – 1696.*

Mokhtarian, F., Abbasi, S. and Kittler, J. 1996. *Efficient and robust retrieval by shape content through curvature scale space. In Smeulders and Jain, pp. 35 – 42.*

Mukherjea, S., Hirata, K. and Hara Y. 1997. *Towards a multimedia world wide web information retrieval engine. In Sixth International WWW Conference, 7-11 April '97, Santa Clara, CA, USA.*

Nastar, C., Mitschke, M., Meilhac, C. and Boujemaa, N. 1998. Surfimage: A flexible content-based image retrieval system. In Proceedings of the ACM International Multimedia Conference, 12-16 September '98, Bristol, England, pp. 339 – 344.

Ng, R. T. & Sedighian, A. 1996. Evaluating multidimensional indexing structures for images transformed by principal component analysis. volume 2670, pp. 50-61.

Niblack, W., Barber, R., Equitz, W., Flickner, M., Glasman, E., Petkovic, D., Yanker, P., Faloutsos, C. and Taubin, G. 1993. The qbic project: Querying images by content using color, texture, and shape. In Proceedings of the SPIE Conference on Storage and Retrieval for Image and Video Databases, 2-3 February '93, San Jose, CA, pages 173-187.

Ortega, M., Rui, Y., Chakrabarti, K., Mehrotra, S. and Huang, T. 1997. Supporting similarity queries in MARS. In Proceedings of the 5th ACM International Multimedia Conference, Seattle, Washington, 8-14 Nov. '97, pages 403-413.

Pentland, A., Picard, R. and Sclaroff, S. 1996. Photobook: Content-based manipulation of image databases. International Journal of Computer Vision, 18(3):233-254.

Porkaew K. and Mehrotra S. 1999. Query Refinement for Content Based Multimedia Retrieval in MARS. Multimedia Computing and Systems, pp. 747-751.

Pinna B., Brelstaff G. and Spillmann L. 2001. Surface color from boundaries: A new 'watercolor' illusion, Vision Research, Vol. 41, pp. 2669 – 2676.

Privitera C.M. & Stark L.W. 1998. Evaluating image processing algorithms that predict regions of interest. UC Berkeley, USA.

Rowley, H., Baluja, S., And Kanade, K. 1996. *Human Face Detection in Visual Scenes. Advances in Neural Information Processing Systems 8 (Proceedings of NIPS), Denver, USA, November, 875-881.*

Rubner, Y., Tomasi, C., & Guibas, L. J. 1998. *A metric for distributions with applications to image databases. iccv, 00, 59.*

Rubner, Y., Tomasi, C., & Guibas, L. J. 2000. *The earth mover's distance as a metric for image retrieval. Int. J. Comput. Vision, 40(2), 99-121.*

Schomaker, L., de Leau, E. & Vuurpijl, L. 1999. *Using pen-based outlines for object-based annotation and image-based queries. Lecture notes in Computer Science, 1614, 585–592.*

Di Sciascio, E., Mingolla, G. and Mongiello, M. 1999. *Content-based image retrieval over the web using query by sketch and relevance feedback. In Huijismans and Smeulders, pp 123 – 130.*

Sclaroff, S., Taycher, L. and La Cascia, M. 1997. *Imagerover: A content-based image browser for the world wide web. In Proceedings IEEE Workshop on Content-based Access of Image and Video Libraries, June '97.*

Scott, G. J. & Shyu, C.-R. 2003. *Ebs k-d tree: An entropy balanced statistical k-d tree for image databases with ground-truth labels. In CIVR, pp. 467-476.*

Sebe, N. & Lew, M. S. 2001. *Color-based retrieval. Pattern Recogn. Lett., 22(2), 223-230.*

Sebe, N., Lew, M. S., & Huijsmans, D. P. 2000. *Toward improved ranking metrics. IEEE Trans. Pattern Anal. Mach. Intell.*, 22(10), 1132-1143.

Shim S.O. and Choi T.S. 2002. *Edge Color Histogram for Image Retrieval*, In *IEEE International Conference on Image Processing, Rochester (NY)*, pp. 957-960.

Siggelkow, S. and Burkhardt, H. 2001. *Fast invariant feature extraction for image retrieval*. In Remco C. Veltkamp, Hans Burkhardt, and Hans-Peter Kriegel, editors, *State-of-the-Art in Content-Based Image and Video Retrieval*. Kluwer.

Smith, J. R. 1997. *Integrated Spatial and Feature Image Systems: Retrieval, Compression and Analysis*. PhD thesis, Graduate School of Arts and Sciences, Columbia University.

Smith, J. R. And Chang, S.F. 1995. *Automated image retrieval using color and texture*. Technical Report CU/CTR 408-95-14, CTR, Columbia University.

Smith, J. R. And Chang, S.F. 1997. *Visually Searching the Web for Content*. *IEEE Multimedia* 4(3), 12- 20.

Smith, J. R. And Chang, S.F. 1997. *Querying by color regions using the VisualSEEK content-based visual query system*. In M. T. Maybury, editor, *Intelligent Multimedia Information Retrieval*. AAAI Press.

Srihari, R. 1995. *Automatic indexing and content-based retrieval of captioned images*. *IEEE Computer*, 28(9):49-56.

Stanchev P. 2003. *Using image mining for image retrieval*, *IASTED Conference "Computer Science and Technology," Cancun, Mexico*, pp. 214–218.

Stentiford FWM, 2007. Attention based similarity, Pattern Recognition.

Stricker, M.A. & Orengo, M. 1995. Similarity of colour images. Proceedings Storage and Retrieval for Image and Video Databases III, 381–392.

Swain, M. J. & Ballard, D. H. 1991. Color indexing. Int. J. Comput. Vision, 7(1), 11-32.

Tamura H., Mori S. and Yamawaki T., 1978. Texture features corresponding to visual perception, IEEE Trans. on System, Man and Cybernetics.

Tanaka K., Saito H., Fukada Y. and Moriya M., Coding visual images of objects in the inferotemporal cortex of the macaque monkey, Journal of Neurophysiology, Vol. 66, pp. 170-180.

Veltkamp R.C. & Hagedoorn M. 2001. State-of-the-art in shape matching. In M.S. Lew (Ed.), Principles of Visual Information Retrieval (87–119). Berlin: Springer.

Vendrig, J., Worring, M. and Smeulders, A. W. M. 1999. Filter image browsing: Exploiting interaction in image retrieval. In Huijsmans and Smeulders, pages 147-154.

Volmer, S. 1997. Tracing images in large databases by comparison of wavelet fingerprints. In Proceedings of the 2nd International Conference on Visual Information Systems, San Diego, December '97, pages 163-172.

Wang J. Z. and Du Y. 2001. Scalable Integrated Region-based Image Retrieval using IRM and Statistical Clustering, Proc. of ACM/IEEE Joint Conference on Digital Libraries.

Wang, J. Z., Wiederhold, G., Firschein, O. and Wei, S. 1997. Wavelet-based image indexing techniques with partial sketch retrieval capability. In Proceedings of the Fourth Forum on Research and Technology Advances in Digital Libraries, Washington D.C., May '97, pp. 13 – 24.

Wang, J., Jann Yang, W., & Acharya, R. 1997. Color clustering techniques for color-content-based image retrieval from image databases. icmcs, 00, 442.

Ye, H. & Xu, G. 2003. Fast search in large-scale image database using vector quantization. In CIVR, pp. 477-487.