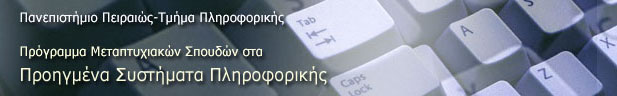
Πανεπιστήμιο Πειραιώς – Τμήμα Πληροφορικής

Πρόγραμμα Μεταπτυχιακών Σπουδών

«Προηγμένα Συστήματα Πληροφορικής»

Μεταπτυχιακή Διατριβή

|  |  |
| --- | --- |
| Τίτλος Διατριβής | Content Based Image Retrieval: A survey and an implementation |
| Ονοματεπώνυμο Φοιτητή | Xαραλάμπους Χαράλαμπος του Σταύρου |
| Αριθμός Μητρώου | ΜΠΣΠ-07026 |
| Κατεύθυνση | Συστήματα Υποστήριξης Αποφάσεων |
| Επιβλέπων | Τσιχριντζής Γεώργιος, Καθηγητής |



|  |  |
| --- | --- |
| Ημερομηνία Παράδοσης | Σεπτέμβριος 2009 |

|  |
| --- |
| Τριμελής Εξεταστική Επιτροπή |

|  |  |  |
| --- | --- | --- |
| (υπογραφή) | (υπογραφή) | (υπογραφή) |
| Όνομα Επώνυμο Βαθμίδα | Όνομα Επώνυμο Βαθμίδα | Όνομα Επώνυμο Βαθμίδα |

**Περίληψη**

Με τις μεγάλες βάσεις δεδομένων εικόνας να είναι πλέον πραγματικότητα τόσο στον επιστημονικό οσο και στον εμπορικό τομέα, οι μέθοδοι για την οργάνωση τέτοιων βάσεων δεδομένων και για την αποτελεσματική ανάκτηση των δεδομένων έχουν αναπτυχθεί σημαντικά. Έχουμε εργαστεί σε αυτό το πρόβλημα παρέχοντας μια μέθοδο για την επίλυση του που βρίσκει εικόνες ανάλογα με την ομοιότητα τους με την εικόνα ερώτημα. Η τεχνική Content Based Image Retrieval (θα την αναφέρουμε CBIR για συντομία από εδώ και στο εξής) είναι μια τεχνική για την ανάκτηση εικόνων με βάση τα αυτόματα εξαγώμενα χαρακτηριστικά όπως το χρώμα, η υφή και το σχήμα. Η CBIR διαφέρει από την κλασική ανάκτηση πληροφοριών σε βάσεις δεδομένων γιατί οι βάσεις δεδομένων εικόνων είναι κατ' ουσίαν αδόμητες, καθώς οι ψηφιακές εικόνες αποτελούνται από συστοιχίες τιμών εντάσεων των pixels που δεν είναι έχουν εγγενές νόημα. Η τεχνική CBIR ανακτά αποθηκευμένες εικόνες από μια τέτοια συλλογή από τη σύγκριση των χαρακτηριστικών που εξάγονται από τις ίδιες τις εικόνες. Τρέχουσες πρακτικές δημιουργίας ευρετηρίου για τις εικόνες βασίζονται σε μεγάλο βαθμό σε περιγραφικά κείμενα ή σε κώδικες ταξινόμησης, υποστηριζόμενες σε ορισμένες περιπτώσεις από τεχνικές ανάκτησης κείμενου που έχουν σχεδιαστεί ή προσαρμοστεί ειδικά για να χειρίζονται εικόνες. Το σύστημά μας είναι ένα αυτόματο σύστημα ανάκτησης εικόνων που βασίζεται σε ένα συνδυασμό των χαρακτηριστικών, συγκεκριμένα το χρώμα, την υφή και το σχήμα. Έτσι, χρησιμοποιώντας αλγόριθμους αντιστοίχισης και σύγκρισης, το χρώμα, η υφή και το σχήμα της εικόνας συγκρίνονται με τα αντίστοιχα χαρακτηριστικά μιας άλλης εικόνας. Αυτή η σύγκριση γίνεται με μετρήσεις απόστασης (distance metrics). Στο τέλος, οι μετρήσεις αυτές πραγματοποιούνται η μία μετά την άλλη, έτσι ώστε να ανακτηθούν εικόνες από τη βάση δεδομένων που είναι παρόμοιες με το ερώτημα. Η ομοιότητα μεταξύ των χαρακτηριστικών υπολογίζεται χρησιμοποιώντας αλγόριθμους που χρησιμοποιούνται από γνωστά συστήματα CBIR όπως το Qbic της IBM. Για κάθε συγκεκριμένο χαρακτηριστικό, υπάρχει ένας ειδικό αλγόριθμος για την ανάκτηση και ένας για την αντιστοίχιση.

1. Εισαγωγή

Η τεχνική CBIR λειτουργεί σε μία εντελώς διαφορετική αρχή από την αναζήτηση κειμένου. Τα πρωτογενή (primitives) χαρακτηριστικά που χαρακτηρίζουν το περιεχόμενο της εικόνας, όπως το χρώμα, η υφή και το σχήμα, υπολογίζονται τόσο για τις αποθηκευμένες όσο και για τις εικόνες προς αναζήτηση, και χρησιμοποιούνται για τον προσδιορισμό (π.χ.) 20 αποθηκευμένων εικόνων που ταιριάζουν καλύτερα με το ερώτημα. Τα σημασιολογικό χαρακτηριστικά, όπως το είδος του αντικειμένου που υπάρχει στην εικόνα είναι πιο δύσκολο να εξαχθούν, αν και αυτό παραμένει ενεργό θέμα έρευνας.

**Τρία εμπορικά συστήματα CBIR** είναι τώρα διαθέσιμα – το Qbic της IBM, η VIR μηχανή της Virage, και το Excalibur του RetrievalWare. Επιπλέον, δοκιμαστικές εκδόσεις των πολυάριθμων πειραματικών συστημάτων μπορεί κάποιος να βρει στο διαδίκτυο, συμπεριλαμβανομένων των Photobook του MIT, το WebSEEk του Πανεπιστημίου Κολούμπια, το Informedia του Carnegie-Mellon University, το Blobworld από το Πανεπιστήμιο της California, Berkeley, το SIMPLIcity [3]  και το ISTORAMA από το Ινστιτούτο Πληροφορικής και Τηλεματικής στη Θεσσαλονίκη. Τα CBIR συστήματα έχουν αρχίσει να βρίσκουν μια θέση στην αγορά, με τα κύρια πεδία εφαρμογής να περιλαμβάνουν την πρόληψη της εγκληματικότητας (αναγνώριση δακτυλικών αποτυπωμάτων και την αναγνώριση προσώπου), την πνευματική ιδιοκτησία (καταχώρισης εμπορικού σήματος), τη δημοσιογραφία και τη διαφήμιση (βίντεο διαχείριση περιουσιακών στοιχείων) και την αναζήτηση στο Web.

Η **αποτελεσματικότητα** του συνόλου των σημερινών συστημάτων CBIR είναι εγγενώς περιορισμένη από το γεγονός ότι μπορούν να λειτουργήσουν μόνο σε πρωτογενές (primitive) επίπεδο χαρακτηριστικών. Κανένα από αυτά δεν μπορεί να αναζητήσει αποτελεσματικά την, ας πούμε, φωτογραφία ενός σκύλου - αν και ορισμένες σημασιολογικές ερωτήματα θα μπορούν να διεκπεραιώνονται διευκρινίζοντας τα από την άποψη των πρωτογενών χαρακτηριστικών (primitives). Μια σκηνή σε μια παραλία για παράδειγμα μπορεί να εξορυχθεί με το να καθορίσουμε μεγάλες περιοχές μπλε χρώματος στο πάνω μέρος της εικόνας και κίτρινου στο κάτω. Υπάρχουν στοιχεία ότι συνδιάζοντας πρωτογενή χαρακτηριστικά με λέξεις-κλειδιά ή υπερσυνδέσμους μπορούμε να ξεπεράσουμε τέτοια προβλήματα αν και δεν ξέρουμε πως να τα συνδυάσουμε για ανάκτηση. Μια σκηνή παραλία, για παράδειγμα, μπορούν να ανακτηθούν με την ένδειξη μεγάλες περιοχές του μπλε στο πάνω μέρος της εικόνας, και το κίτρινο στο κάτω μέρος. Υπάρχουν αποδείξεις ότι ο συνδυασμός πρωτόγονα χαρακτηριστικά εικόνας με λέξεις-κλειδιά κειμένου ή υπερ-συνδέσεις μπορούν να ξεπεράσουν ορισμένα από αυτά τα προβλήματα, αν και λίγα είναι γνωστά για το πώς τα παραπάνω χαρακτηριστικά, μπορεί καλύτερα να συνδυάζονται για ανάκτηση εικόνων.

Η τεχνική **CBIR** αντλεί πολλές από **τις μεθόδους** της από τον τομέα της επεξεργασίας εικόνας και τεχνητής όρασης, και θεωρείται από ορισμένους ως υποσύνολο του εν λόγω τομέα. Διαφέρει από αυτά τα πεδία, κυρίως μέσω της έμφασης που δίνει για την ανάκτηση των εικόνων με τα επιθυμητά χαρακτηριστικά από μια συλλογή σημαντικού μεγέθους. Η επεξεργασία εικόνας καλύπτει ένα πολύ ευρύτερο πεδίο, συμπεριλαμβανομένων τη βελτίωση εικόνας, συμπίεση, μετάδοση, και την ερμηνεία. Ενώ υπάρχουν γκρίζες ζώνες (όπως η αναγνώριση αντικειμένων από την ανάλυση χαρακτηριστικών) η διάκριση μεταξύ mainstream ανάλυσης εικόνας και CBIR είναι συνήθως αρκετά σαφής. Ένα παράδειγμα μπορεί να το κάνει αυτό σαφές. Πολλές από τις αστυνομικές δυνάμεις χρησιμοποιούν σήμερα συστήματα αυτόματης αναγνώρισης προσώπου. Αυτά τα συστήματα μπορούν να χρησιμοποιηθούν με δύο τρόπους. Πρώτον, η εικόνα μπροστά από την κάμερα μπορεί να συγκριθεί με τη βάση δεδομένων που καταγράφει ένα μεμονωμένο άτομο για να εξακριβώσει την ταυτότητά του. Στην περίπτωση αυτή, μόνο δύο εικόνες συμφωνούν, μια διαδικασία που μόνο λίγοι παρατηρητές θα αποκαλούσαν CBIR. Δεύτερον, μπορεί να γίνει αναζήτηση σε ολόκληρη τη βάση δεδομένων για να βρεθούν οι πιο ταιριαστές τις εικόνες. Αυτό είναι ένα κλασικό παράδειγμα της τεχνικής CBIR.

1. **Υπάρχοντα συστήματα**

Εδώ παρουσιάζουμε μια σύντομη ανασκόπηση υπαρχόντων συστημάτων για τη δημιουργία ευρετηρίων εικόνων και την ανάκτηση αυτών. Αν και η τεχνολογία για την οργάνωση και την αναζήτηση εικόνων με βάση το περιεχόμενό τους είναι ακόμη στα σπάργανα, παρουσιάζει τεράστιες δυνατότητες.

**IBM's QBIC**

Το **Almaden Research Center** **της IBM** ανέπτυξε ένα σύστημα, το **Qbic (Query By Image Content)**. Παρέχει τις μεθόδους για την αναζήτησησε μια μεγάλη on-line βάση δεδομένων εικόνων χρησιμοποιώντας χαρακτηριστικά εικόνας ως βάση για τις ερωτήσεις. Τα χαρακτηριστικά που χρησιμοποιούν περιλαμβάνουν το χρώμα, την υφή και το σχήμα των αντικειμένων και των περιοχών. Οι αναζητήσεις του Qbic είναι κατά προσέγγιση. Αυτές οι αναζητήσεις χρησιμεύουν ως «φίλτρο πληροφοριών» και είναι διαδραστικές, ώστε ο χρήστης να μπορεί να χρησιμοποιήσει οπτική αναζήτηση, αξιολόγηση και οπτική βελτίωση προκειμένου να αποφασισεί τι να απορρίψει και τι να κρατήσει. Το σύστημα Qbic έχει τρεις λειτουργίες: το «γέμισμα» της βάσης, τον υπολογισμό χαρακτηριστικών και την αναζήτηση εικόνας. Δεδομένου ότι η περιγραφή των αντικειμένων της εικόνας δεν είναι αυτόματη, η ανάκτηση αντικειμένων στο Qbic γίνεται χειροκίνητα ή ημι-αυτόματα (μέθοδος ανάλυσης με τη βοήθεια ανθρώπου). Το έργο του υπολογισμού των χαρακτηριστικών γίνεται από το σύστημα. Τα χαρακτηριστικά γνωρίσματα περιλαμβάνουν το χρώμα, την υφή και το σχήμα. Το Qbic υπολογίζει το ιστόγραμμα ως χαρακτηριστικό για το χρώμα και το χαρακτηριστικό της υφής βασίζεται σε ένα τρισδιάστατο χώρο: coarseness, αντίθεση και κατευθυντικότητα. Επιλέγουν αυτό το μοντέλο χαρακτηριστικών κυρίως για χαμηλή πολυπλοκότητα υπολογισμού του. Για το σχήμα, το Qbic χρησιμοποιεί ένα συνδυασμό της κυκλικότητας, εκκεντρικότητας, προσανατολισμό του μεγάλου άξονα και μια σειρά από αλγεβρικές αναλλοίωτες ροπές. Το Qbic υποστηρίζει τόσο ερωτήματα "πλήρης σκηνής" (“full scene” query) όσο και ερωτήματα από παρόμοιες περιοχές. Οι συναρτήσεις που χρησιμοποιεί για την ομοιότητα είναι μετρήσεις, όπως Cityblock Ευκλείδεια ή σταθμισμένη Ευκλείδεια απόσταση. Υπολογίζει το histogram quadrature distance για το χρώμα. Η υφή υπολογίζεται ως σταθμισμένη Ευκλείδεια απόσταση στον τρισδιάστατο χώρο. Το Qbic επίσης υποστηρίζει "το ερώτημα από σκίτσο», το οποίο χρησιμοποιεί τα άκρα που σχεδιάζει ένας χρήστης για να εξάγει αυτόματα τα άκρα από τις εικόνες στη βάση δεδομένων. Τα αποτελέσματα της αναζήτησης βασίζονται στη μέθοδο του πλησιέστερου γείτονα.

**Photobook του MIT**

Στο **Media Lab του MIT** έχει αναπτύχθει ένα εργαλείο για ερωτήσεις σε εικόνα αλλά και σε βίντεο κλιπ, το **Photobook.** Όπως το Qbic, λειτουργεί συγκρίνοντας χαρακτηριστικά που εξάγονται από τις εικόνες. Τα χαρακτηριστικά αυτά είναι παραμετρικές τιμές των συγκεκριμένων μοντέλων που τοποθετούνται σε κάθε εικόνα. Τα μοντέλα είναι το χρώμα, η υφή και το σχήμα. Τα κριτήρια αντιστοίχισης περιλαμβάνουν Euclidean, Mahalanobis, απόσταση, ιστόγραμμα, Fourier peaks και wavelet tree distances, ή οποιοδήποτε γραμμικό συνδυασμό αυτών. Το Photobook έχει τρεις τύπους περιγραφών εικόνας, με κάθε έναν από αυτούς για τον χειρισμό ειδικού περιεχομένου της εικόνας. Περιγραφές εικόνας (“Appearance Photobook”) εφαρμόζονται σε βάσεις δεδομένων προσώπου και πλαισίων κλειδιών, περιγραφές υφής (“Texture Photobook”) εφαρμόζονται σε βάσεις δεδομένων υφής Brodatz και πλαισίων κλειδιών, και περιγραφές σχημάτων (“Shape Photobook”) εφαρμόζονται σε βάσεις δεδομένων εργαλείων χειρός και ψαριών. Το κύριο χαρακτηριστικό του Photobook είναι ότι περιλαμβάνει το FourEyes, ένα διαδραστικό παράγοντα μάθησης που επιλέγει και συνδυάζει τα μοντέλα με βάση το παράδειγμα από τον χρήστη. Στην παρούσα φάση, το FourEyes είναι ένα εργαλείο για την τμηματοποίσηση και την κατηγοριοποίηση των εικόνων. Αν και οι αλγόριθμοι τεχνητής όρασης δεν είναι αρκετά robust για να κάνουν αυτόματα τo annotation σε εικόνες γενικού θέματος, μπορούν να εφαρμοστούν σε αρκετές περιπτώσεις. Το πρόβλημα είναι πώς να επιλέξει κανείς το μοντέλο ή τον αλγόριθμο. Για παράδειγμα, στην περίπτωση της υφής, που ορίζεται μια ομάδα των μοντέλων, συμπεριλαμβανομένων συνεμφάνισης, τυχαίου τομέα, fractals, αντίδραση-διάχυσης, eigenpatterns, μορφολογία, Fourier bins, wavelets, κατευθυνόμενες πυραμίδες (steerable pyramids), autoregressive moving average, γραμματική, cluster-based probability, wold, συστήματα σωματιδίων, φίλτρα Gabor, κλπ. Δεδομένου ότι δεν υπάρχει ένα μοντέλο που θα είναι καλύτερο, ένα ημι-αυτόματο εργαλείο, το FourEyes, έχει συσταθεί για τον προσδιορισμό των μέτρων ομοιότητας που είναι κατάλληλα για την εργασία εκμάθησης από παραδείγματα που δίνονται από το χρήστη. Όταν ο χρήστης επιλέξει ορισμένες περιοχές και τους δώσει μια ετικέτα, το FourEyes μπορεί να προεκτείνει την ετικέτα και σε άλλες περιοχές της εικόνας και στη βάση δεδομένων. Λειτουργεί με το συνδυασμό κατάλληλων μοντέλων από την ομάδα των μοντέλων της, σύμφωνα με τα παραδείγματα που δίνονται από τον χρήστη. Το FourEyes δημιούργησε το Photobook, ένα ευέλικτο περιβάλλον που μπορεί να στηρίξει την έρευνα χρησιμοποιώντας διάφορα χαρακτηριστικά αλλά δεν προσφέρει βοήθεια στην επιλογή του σωστού χαρακηριστικού για ένα συγκεκριμένο σκοπό.

**VisualSEEK Columbia**

Το **Center for Image Technology for New Media** στο **Columbia** έχει αναπτύξει ένα σύστημα αναζήτησης εικόνας που ονομάζεται **VisualSEEK.** Η τρέχουσα έκδοση VisualSEEK είναι κυρίως για την αναζήτηση εικόνων μέσω του Παγκόσμιου Ιστού και έχει ένα αρχείο από 12.000 εικόνες. Το VisualSEEK αποτελείται από τρία μέρη:

1. H εφαρμογή πελάτη, η οποία είναι μια σουίτα από Java applets που εκτελούνται απο το browser. Συλλέγει το ερώτημα από το χρήστη και το μετατρέπει στο ανάλογο αλφαριθμητικό.

2. H εφαρμογή δικτύου και επικοινωνίας, η οποία χειρίζεται την επικοινωνία.

3. Η εφαρμογή εξυπηρετητή, η οποία λαμβάνει το ερώτημα, εκτελεί το ερώτημα και επιστρέφει τα αποτελέσματα στο χρήστη. Η εφαρμογή εξυπηρετητή δημιουργεί HTML κώδικα που εμφανίζει τα αποτελέσματα του ερωτήματος του χρήστη. Το VisualSEEK υποστηρίζει σήμερα ολική ανάκτηση ομοιότητας εικόνας και ανάκτηση από τοπικά και περιφερειακά χαρακτηριστικά (χρώμα και υφή). Η αναζήτηση του είναι διαφορετική από το Qbic και το Photobook στο ότι ο χρήστης μπορεί να εκτελέσει το ερώτημα για τις εικόνες χρησιμοποιώντας τόσο τις οπτικές ιδιότητες όσο και τις χωρικές. Όσον αφορά τη ολική ανάκτηση εικόνας, το VisualSEEK έχει αρκετά μέτρα για την ομοιότητα, η οποία μπορεί να επιλεγεί από το χρήστη, μεταξύ των οποίων color histogram intersection, color histogram moments, color region intersection, color region Euclidean, color histogram Euclidean, texture set intersection, texture histogram intersection, texture and color intersection, texture and color Euclidean and color histogram quadratic (ταξινομημένα ανάλογα με την πολυπλοκότητα του αλγορίθμου). Για ερωτήματα με χαρακτηριστικά από την περιοχή και τις χωρικές σχέσεις, το σύστημα αυτόματα εξάγει τις κυριότερες περιοχές χρώματος και υφής από μια εικόνα και εκτελεί το ερώτημα με χαρακτηριστικά και με σχετικά σημεία. Η τεχνική ανάκτησης που τελικά χρησιμοποιείται από το VisualSEEK είναι το back-projection του δυαδικού χρώματος και των συνόλων υφής.

Στον **Computer Science Division** του **University of California at Berkeley** επίσης διεξάγεται έρευνα για τις ψηφιακές βιβλιοθήκες. Στόχος τους είναι να παρέχουν ένα γενικό πλαίσιο για την ανάκτηση εικόνας που βασίζεται στο περιεχόμενο και επιτρέπει την αναζήτηση από χαμηλού επιπέδου σε υψηλού επιπέδου. Καθορίζουν μια μικρή αναπαράσταση, η οποία είναι μια μετάβαση από τα pixel δεδομένων σε ένα μικρό σύνολο συνεκτικών περιοχών με ένα χρώμα και μία υφή του χώρου. Σε αυτό το επίπεδο, είναι αρκετά παρόμοια με τα VisualSEEK και Photobook. Η διαφορά είναι ότι προσπαθούν να ταξινομήσουν τις εικόνες σε οπτικές κατηγορίες, κάτι που χρειάζεται μια διαδικασία εκμάθησης να δώσει πιθανοτική ερμηνεία των blob περιοχών σε μια εικόνα. Το σύστημα περιλαμβάνει επίσης χωρική πληροφορία στην διαδικασία μάθησης του. Ο χώρος χρωμάτων που χρησιμοποιούν στην επεξεργασία τους είναι HSV διότι είναι πιο αποτελεσματικός και μπορεί να βοηθήσει στην ομαδοποίηση και αναγνώριση. Ενσωματώνουν μια λίστα με 13 χρώματα και να δημιουργήσουμε ένα πίνακα αναζήτησης να διαιρέσουν κάθε εικόνας σε αυτά τα κανάλια χρώματος. Η εξαγωγή της υφής ως χαρακτηριστικό βασίζεται σε πληροφορίες από το "windowed image second moment matrix", το οποίο μπορεί να κατατάξει 6 είδη τύπων υφή (μη-υφή, 2D υφή, 4 1D υφή σε διαφορετικές κατευθύνσεις). Μετά την απόκτηση των χαρακτηριστικών χρώματος και υφής, χρησιμοποιούν έναν Expectation Maximization αλγόριθμο για να πάρουν την κατάτμηση της εικόνας. Το πλεονέκτημα του αλγορίθμου Expectation Maximization είναι ότι μπορεί να αποφευχθεί ο κατακερματισμός των κύριων περιοχών στην κατάτμηση. Η μαθησιακή διαδικασία για τη δημιουργία κατηγοριοποίησης βασίζεται σε έναν ταξινομητής Bayes, το οποίο χρησιμοποιεί το χρώμα, την υφή και χωρική πληροφορία. Πείραματα έγιναν σε περίπου 1.200 φωτογραφίες που εμπίπτουν σε 12 κατηγορίες. Τα training και testing sets του ταξινομητή Bayes περιέιχαν τα 2/3 και 1/3 των εικόνων αντίστοιχα. Οι επιδόσεις είναι αρκετά συνεπείς με την οπτική αντίληψη, αν και υπάρχουν κάποιες εσφαλμένες κατατάξεις σε ορισμένες κατηγορίες.

**Virage**

H **Virage Inc.** έχει τις ρίζες της στην έρευνα που έγινε στο **University of California at San Diego.** Η τεχνολογία της ανάκτηση της εικόνας σε αυτην την εφαρμογή είναι χτισμένη γύρω από έναν πυρήνα που ονομάζεται Virage Engine. Το Virage Engine είναι μία ανοικτή, φορητή και επεκτάσιμη αρχιτεκτονική που μπορεί να συμπεριλάβει οποιαδήποτε θεματολογία εικόνων. Το Virage χρησιμοποιεί τρία επίπεδα άντλησης πληροφοριών για τις εικόνες: την εικόνα αυτή καθ’ αυτή (το επίπεδο εκπροσώπησης εικόνας), την επεξεργασία εικόνας (το επίπεδο αντικείμενου εικόνας) και ένα χαρακτηριστικό για το οποίο ενδιαφέρεται ο χρήστης. Τα υπολογίζομενα οπτικά χαρακτηριστικά ονομάστηκε από Virage ως «πρωτογενή» (primitives), τα οποία είναι είτε global είτε τοπικά. Σύμφωνα με Virage, ένα πρωτογενές χαρακτηριστικό θα πρέπει να είναι ουσιαστικό για την αντίληψη, συμπαγές για την αποθήκευση, αποτελεσματικό υπολογιστικά, ακριβές, και θα πρέπει να μπορεί να χρησιμοποιθεί για τη δημιουργία ευρετηρίου. Μερικά primitives του Virage είναι global χρώμα, local χρώμα, τη δομή και την υφή. Αυτά τα primitives είναι καθολικά, με την έννοια ότι είναι χρήσιμα στις περισσότερες domain-ανεξάρτητες εφαρμογές και μπορούν να υπολογίζονται αυτόματα από το Virage Engine. Ο χρήστης μπορεί να επιλέξει να συνδυάσει τα global primitives σε συνδυασμό με συγκεκριμένα για το πεδίο που τον ενδιαφέρει primitives ώστε να κατασκευάσει μια συγκεκριμένη εφαρμογή, όπως μια ιατρική ή multimedia εφαρμογή. Η επεκτάσιμη αρχιτεκτονική του δίνει τη δυνατότητα στο χρήστη να εφαρμόζει κάθε φορά καινούριους βελτιστοποιημένους αλγόριθμους για την επίλυση συγκεκριμένων προβλημάτων. Η Virage Μηχανή παραδίδεται σε μία δυναμική βιβλιοθήκη για διάφορες πλατφόρμες (Sun, SGI, Windows και MAC). Η βιβλιοθήκη παρέχει τη δυνατότητα χρήσης διαφόρων ειδών βάσεων δεδομένων και APIs. Βάσεις δεδομένων όπως Oracle, Sybase ή τα εργαλεία επεξεργασίας εικόνας όπως Photoshop, ή CorelDraw θα μπορούσαν να χρησιμοποιήσουν τη μηχανή Virage για την αναζήτηση εικόνων και τη διαχείριση. Μια μηχανή αναζήτησης στο Διαδίκτυο, όπως η WebCrawler ή Infoseek μπορεί να επεκτείνει τις ικανότητες της με την εύρεση εικόνας στο δίκτυο και έτσι να βοηθήσει την οικοδόμηση ενός αποθηκευτικού συστήματος με εικόνες που να επιτρέπει την αναζήτηση και να διανέμεται μέσω του διαδικτύου.

**KRDL's CORE**

Το **CORE, a COntent-based Retrieval Engine** για βάσεις δεδομένων πολυμέσων, αναπτύχθηκε από το RWC Lab of KRDL. Το CORE παρέχει λειτουργίες πολυμέσων στα συστήματα πληροφοριών για δημιουργία αντικείμενων πολυμέσων, ανάλυση, αποθήκευση και ανάκτηση. Μερικές από τις βασικές τεχνικές του CORE είναι: πολλαπλές μέθοδοι εξαγωγής χαρακτηριστικών, πολλαπλές μεθόδοι ανάκτησης με βάση το περιεχόμενο, μία νέα τεχνική που βασίζεται σε ευρετηρίαση (δημιουργία ευρετηρίου) σύνθετων μέτρων χαρακτηριστικών με τη χρήση των νευρωνικών δικτύων και μια νέα τεχνική για fuzzy ανάκτηση των πληροφοριών πολυμέσων. Το CORE έχει τρεις κύριες ενότητες:

1. Μονάδα ανάλυσης δεδομένων πολυμέσων (ανάλυση)

2. Query ενότητα (ανάκτηση και ευρετηρίαση)

3. Ενότητα Customization (Training)

Το CORE παρέχει στοιχειώδεις κλάσεις πολυμέσων ως δομικά στοιχεία για την ανάπτυξη μιας εφαρμογής. Η αναπαράσταση της εικόνας αποτελείται από τέσσερα επίπεδα: ασπρόμαυρη η έγχρωμη εικόνα, κατακερματισμένη εικόνα, περιγραφές και χαρακτηριστικά, και ερμηνεία. Το CORE προβλέπει τρία είδη των λειτουργιών του κατακερματισμού, συγκεκριμένα τον κατακερματισμό ανάλογα με το χρώμα, ανάλογα με τη μορφή, και ανάλογα με το περίγραμμα. Η λειτουργία εξαγωγής χαρακτηριστικών γίνεται για τον υπολογισμό των μέτρων που αναγράφονται στις κατακερματισμένες ανά αντικείμενα περιοχές. Αυτές οι λειτουργίες περιλαμβάνουν: principal component analysis, αναλλοίωτες ροπές, Fourier συντελεστές για καθορισμό σχήματος, προβολές, οι παράμετροι σχήματος, texture energy, χαρακτηριστικά χρώματος κλπ. Το ευρετήριο με βάση το περιεχόμενο βρίσκεται ανάμεσα στα χαρακτηριστικά και στο επίπεδο ερμηνείας. Χρησιμοποιεί την έννοια του ταξινομητή και του content-based indexing, και χρησιμοποιεί το LEP (Learning based on Experiences and Perspective) μοντέλο νευρωνικών δικτύων για την παροχή μιας στέρεης θεωρητική βάσης για να είναι σε θέση να σχηματίσει σύνθετα μέτρα χαρακτηριστικών. Η τεχνική αυτή έχει εφαρμοστεί στο CORE για την υποστήριξη αναζήτησης με βάση την ομοιότητα. Μπορεί επίσης να χρησιμοποιηθεί για την ευρετηρίαση fuzzy multi-variate fuzzy membership functions. Υπάρχουν τέσσερα είδη μεθόδων ανάκτησης στο CORE, δηλαδή οπτική περιήγηση, ανάκτηση με βάση την ομοιότητα, fuzzy ανάκτηση και ανάκτηση με βάση το κειμένο. Η ανάκτηση με βάση την ομοιότητα παρέχει πρόσβαση σε πολυμεσικά αντικείμενα, μέσω μετρήσεων των χαρακτηριστικών που διαθέτουν, ενώ η fuzzy ανάκτηση και ανάκτησημε βάση το κειμένο παρέχουν πρόσβαση μέσω της ερμηνείας αντικειμένων πολυμέσων. Η ανάδραση των χρηστών είναι συχνά αναγκαία για την αποτελεσματική ανάκτηση. Με τη χρήση λειτουργίων ανάδρασης, ο χρήστης μπορεί να επιλέξει ένα ή περισσότερα αντικείμενα από τα αποτελέσματα αναζήτησης τα οποία νομίζει ότι είναι πολύ κοντά στο επιθυμητό. Η λειτουργία της ανάδρασης (feedback) υπολογίζει στη συνέχεια τις πληροφορίες από τα επιλεγμένα αντικείμενα και βελτιώνει το ερώτημα. Η λειτουργία ανάδρασης στο CORE υποθέτει ότι τα αντικείμενα ανατροφοδότησης που θα περιληφθούν είναι σε τέτοια σειρά ώστε το πρώτο επιλεγμένο να είναι και το πιο επιθυμητό, και η επιλογή θα πρέπει να βασίζεται στα πλέον παρόμοια χαρακτηριστικά μεταξύ των επιλεγμένων αντικειμένων.

Το CORE έχει χρησιμοποιηθεί για την ανάπτυξη των εξής δύο εφαρμογών:

Computer- Aided Facial Image Inference and Retrieval (CAFIIR) και

System for Trademark Archival and Registration (STAR)

Επιπλέον, μια βάση δεδομένων ιατρικών εικόνων για διάγνωση και εγχείρηση με τη βοήθεια υπολογιστή αναπτύσεται με χρήση του CORE. Το CORE έχει πλούσιες και ολοκληρωμένες λειτουργίες. Ωστόσο, κάθε εφαρμογή έχει ειδικά προβλήματα ανάλογα με το τομέα της, και ειδικές γνώσεις από κάθε τομέα πρέπει να προστεθούν για να προσαρμοστεί η μονάδα ανάκτησης. Αυτό σημαίνει ότι η εκπαίδευση αποτελεί ένα κρίσιμο βήμα στην ανάπτυξη εφαρμογών.

Εφαρμογές Content-Based Image Retrieval

Στο [31], βλέπουμε τρεις γενικές κατηγορίες των στόχων των χρηστών κατά τη χρήση ενός συστήματος CBIR. Υπάρχει μια ευρεία κατηγορία των μεθόδων και συστημάτων που αποσκοπούν στην περιήγηση μέσα από ένα μεγάλο σύνολο εικόνων από απροσδιόριστες πηγές. Οι χρήστες που χρησιμοποιούν αναζήτηση μέσω σύνδεσης (search by association) **στην αρχή** δεν έχουν συγκεκριμένο στόχο, εκτός από το να βρουν ενδιαφέροντα πράγματα. Η αναζήτηση μέσω σύνδεσης συχνά συνεπάγεται την επαναληπτική βελτίωση της αναζήτησης με βάση την ομοιότητα ή τα παραδείγματα με τα οποία ξεκίνησε η αναζήτηση. Συστήματα σε αυτή την κατηγορία κατά κανόνα είναι εξαιρετικά διαδραστικά, όπου η προδιαγραφή μπορεί να δωθεί με σκίτσο [30] ή με παραδείγματα από εικόνες. Το παλαιότερο ρεαλιστικό παράδειγμα ενός τέτοιου συστήματος είναι πιθανότατα το [88]. Το αποτέλεσμα της αναζήτησης μπορεί να παραποιηθεί από διαδραστική ανατροφοδότηση [68]. Για να υποστηριχθεί η αναζήτηση σχετικών αποτελεσμάτων, άλλες πηγές εκτός από τις εικόνες χρησιμοποιούνται, για παράδειγμα το [168].

**Μια άλλη κατηγορία χρηστών** έχει ως στόχο την αναζήτηση μιας συγκεκριμένης εικόνας. Η αναζήτηση μπορεί να γίνεται για ένα ακριβές αντίγραφο της εικόνας που έχει ο χρήστης κατά νου, όπως η αναζήτηση σε καταλόγους τέχνης, π.χ. [48]. Στόχος της αναζήτησης μπορεί επίσης να είναι μια άλλη εικόνα από το ίδιο αντικείμενο του οποίου ο χρήστης έχει ήδη μια εικόνα. Αυτή είναι η αναζήτηση στόχου με βάση το παράδειγμα. Αυτού του είδους η αναζήτηση μπορεί επίσης να εφαρμόζεται όταν ο χρήστης έχει μια συγκεκριμένη εικόνα στο μυαλό του και ο στόχος του αλληλεπιδραστικά προσδιορίζεται ως παρόμοιος με μια ομάδα εικόνων, όπως για παράδειγμα στο [31]. Τα συστήματα αυτά ανταποκρίνονται σε γενικές γραμμές στην αναζήτηση για γραμματόσημα, τέχνη, βιομηχανικά εξαρτήματα, καθώς και σε καταλόγους.

**Η τρίτη κατηγορία** των εφαρμογών αναζήτησης έχει ως στόχο την ανάκτηση αυθαίρετης αντιπροσωπευτικής εικόνας μιας συγκεκριμένης κατηγορίας. Μπορεί να είναι η περίπτωση που ο χρήστης έχει ένα παράδειγμα και η αναζήτηση γίνεται για άλλες εικόνες της ίδιας κατηγορίας. Οι κατηγορίες μπορεί να προέρχονται από τις λεζάντες των εικόνων ή προκύπτουν από τη βάση δεδομένων [170]. Σε αναζήτηση με βάση την κατηγορία, ο χρήστης μπορεί να έχουν στη διάθεσή του μια ομάδα εικόνων και η αναζήτηση είναι για πρόσθετες εικόνες της ίδιας κατηγορίας [29]. Μια τυπική εφαρμογή της αναζήτησης με βάση την κατηγορία είναι σε καταλόγους ποικιλιών (catalogues of varieties). Στο [74], [79], τα συστήματα είναι σχεδιασμένα για την ταξινόμηση των εμπορικών σημάτων. Συστημάτα σε αυτή την κατηγορία είναι συνήθως διαδραστικά και αναφέρονται σε συγκεκριμένη θεματική περιοχή που ορίζει και το είδος τη ομοιότητας.

Συναφείς εργασίες σε Content-Based Image Retrieval

Η τεχνική CBIR για βάσεις δεδομένων εικόνας γενικού σκοπού είναι ένα εξαιρετικά δύσκολο πρόβλημα, λόγω του μεγάλου μεγέθους της βάσης δεδομένων, της δυσκολίας κατανόησης των εικόνων τόσο από τους ανθρώπους όσο και από τους υπολογιστές, της δυσκολία διαμόρφωσης ενός ερωτήματος, και της αξιολόγησης των αποτελεσμάτων τους. Ένας μεγάλος αριθμός μηχανών αναζήτησης σε τέτοιες βάσεις δεδομένων έχει αναπτυχθεί. Το κοινό έδαφος για τα CBIR συστήματα είναι να εξαγάγουν μια υπογραφή για κάθε εικόνα με βάση τις τιμές της σε pixel και να ορίσουν έναν κανόνα για τη σύγκριση των εικόνων. Η υπογραφή χρησιμεύει ως μια αναπαράσταση της εικόνας στο σύστημα. Οι συνιστώσες της υπογραφής ονομάζονται χαρακτηριστικά. Ένα πλεονέκτημα της υπογραφής πάνω από τις αρχικές τιμές των pixel είναι η σημαντική συμπίεση στην αναπαράσταση της εικόνας. Στην πραγματικότητα, το κύριο έργο του σχεδιασμού μιας υπογραφής είναι να γεφυρωθεί το χάσμα μεταξύ της σημασιολογίας της εικόνας και της pixel εκπροσώπησης της, δηλαδή, να δημιουργηθεί μια καλύτερη συσχέτιση με τη σημασιολογία της εικόνας. Τα υπάρχοντα CBIR γενικού σκοπού εμπίπτουν περίπου σε τρεις κατηγορίες ανάλογα με την προσέγγιση για την εξαγωγή των υπογραφών: ιστόγραμμα, χρώμα, και αναζήτηση με βάση την περιοχή.

Αναζήτηση με βάση το Ιστόγραμμα

Οι αλγόριθμοι αναζήτησης με βάση το ιστόγραμμα [4], [18] χαρακτηρίζουν μια εικόνα από τη κατανομή του χρώματος ή από το ιστόγραμμα της. Έχουν χρησιμοποιηθεί πολλές αποστάσεις για να καθορίστει η ομοιότητα μεταξύ δύο αναπαραστάσων ιστογραμμάτων. Η Ευκλείδεια απόσταση και οι παραλλαγές της είναι η πιο συχνά χρησιμοποιούμενη [4] μέθοδος. Ο Rubner από το Πανεπιστήμιο του Stanford πρότεινε την απόσταση του πρωτοπόρου γης (earth mover's distance EMD) [18] που χρησιμοποιεί γραμμικό προγραμματισμό για την αντιστοίχιση ιστογραμμάτων.

Το μειονέκτημα της ολικής αναπαράστασης της εικόνας από το ιστόγραμμα είναι ότι οι πληροφορίες σχετικά με τοποθεσία αντικειμένου, το σχήμα και την υφή [10] χάνονται. Επίσης η αναζήτηση με βάση το ιστόγραμμα είναι ευαίσθητη στις διακυμάνσεις της έντασης και στις στρεβλώσεις του χρώματος.

Αναζήτηση με βάση τη διάταξη χρώματος

Η προσέγγιση αυτή προσπαθεί να ξεπεράσει το μειονέκτημα της αναζήτησης με βάση το ιστόγραμμα. Στην απλή ευρετηρίαση με βάση τη διάταξη χρώματος [4], οι εικόνες είναι χωρισμένες σε τμήματα και αποθηκεύεται το μέσο χρώμα του κάθε μπλοκ. Έτσι, η διάταξη με βάση το χρώμα είναι ουσιαστικά μια μικρή εκπροσώπηση χαμηλότερης ανάλυσης της αρχικής εικόνας. Ένα σχετικά πρόσφατο σύστημα, το WBIIS [28], χρησιμοποιεί συντελεστές Daubechies αντί του μέσου όρου. Με την αναπροσαρμογή των μεγεθών κατά κατηγορίες ή τα επίπεδα των μετασχηματισμών wavelet, η ταχύτητα στην παράσταση διάταξης χρώματος μπορεί να είναι συντονισμένη. Η καλύτερη διάταξη χρώματος χρησιμοποιώντας ένα μόνο pixel μπλοκ είναι η αρχική αναπαράσταση pixel. Ως εκ τούτου, μπορούμε να δούμε μια αναπαράσταση διάταξης με βάση το χρώμα ως το αντίθετο άκρο του ιστογράμματος. Σε καλή ανάλυση, η εκπροσώπηση διάταξης χρώματος διατηρεί το σχήμα, τη θέση και την υφή.

Ωστόσο, όπως με την αναπαράσταση με pixel, αν και οι πληροφορίες όπως το σχήμα διατηρούνται στην αναπαράσταση με διάταξη χρώματος, το σύστημα ανάκτησης μπορεί να αντιληφθεί ότι δεν μπορεί να το διατηρήσει άμεσα. Η αναζήτηση με βάση τη διάταξη χρώματος είναι ευαίσθητη στις μεταβαλλόμενες, cropping, scaling και περιστροφή και επειδή οι εικόνες περιγράφονται από ένα σύνολο τοπικών ιδιότητων [28]. Η προσέγγιση που υιοθετείται από το πρόσφατο σύστημα Walrus [14] για τη μείωση της ευαισθησία στην μετατόπισης και την κλιμάκωσης της αναζήτηση με βάση τη διάταξη χρώματος, είναι να αναπαραγάγει εξαντλητικά πολλές υποεικόνες (subimage) βασιζώμενο σε μια πρωτότυπη εικόνα. Οι υποεικόνες σχηματίζονται με συρόμενα παράθυρα διαφόρων μεγεθών και μία υπογραφή διάταξης χρώματος υπολογίζεται για κάθε υποεικόνα. Η ομοιότητα μεταξύ των εικόνων καθορίζεται στη συνέχεια, από τη σύγκριση των υπογραφών των subimages. Ένα προφανές μειονέκτημα του συστήματος είναι η ραγδαία αύξηση της υπολογιστικής πολυπλοκότητας και η αύξηση του μεγέθους του χώρου αναζήτησης που οφείλεται στην πλήρη παραγωγή του subimages. Furthermore, texture and shape information is discarded in the signatures because every subimage is partitioned into four blocks and only average colors of the blocks are used as features. This system is also limited to intensity-level image representations. Επιπλέον, οι πληροφορίες για την υφή και το σχήμα απορρίπτοναι κατά την υπογραφή, επειδή κάθε subimage είναι χωρισμένο σε τέσσερις ενότητες και μόνο το μέσο χρώμα χρησιμοποείται ως χαρακτηριστικό.

**Region-Based Αναζήτηση**

Τα συστήματα ανάκτησης με βάση την περιοχή κάνουν μία προσπάθεια να ξεπεραστούν οι ελλείψεις της αναζήτησης με διάταξη χρώματος από εικόνες που αντιπροσωπεύουν το αντικείμενο. Σε τέτοια συστήματα αρχικά γίνεται ο κατακερματισμός της εικόνας [20], [27] για να αποσυντεθούν οι εικόνες σε περιοχές, οι οποίες αντιστοιχούν στα αντικείμενα εάν η αποσύνθεση είναι ιδανική. Ωστόσο, ο κατακερματισμός της εικόνας είναι σχεδόν τόσο δύσκολη υπόθεση όσο η κατανόηση εικόνας, επειδή οι εικόνες είναι 2D προβολές των τριδιάστατων αντικειμένων και οι υπολογιστές δεν είναι εκπαιδευμένοι στον τρισδιάστατο κόσμο με τον τον τρόπο που είναι οι άνθρωποι. Δεδομένου ότι το σύστημα ανάκτησης έχει προσδιορίσει ποια είναι τα αντικείμενα της εικόνας, είναι ευκολότερο για το σύστημα να αναγνωρίζει παρόμοια αντικείμενα σε διαφορετικές θέσεις και με διαφορετικές κατευθύνσεις και μεγέθη. Συστήματα ανάκτησης Region-based περιλαμβάνουν το σύστημα Netra [11], το σύστημα Blobworld [2], και το σύστημα ερωτημάτων με πρότυπο το χρώμα της περιοχής [22].

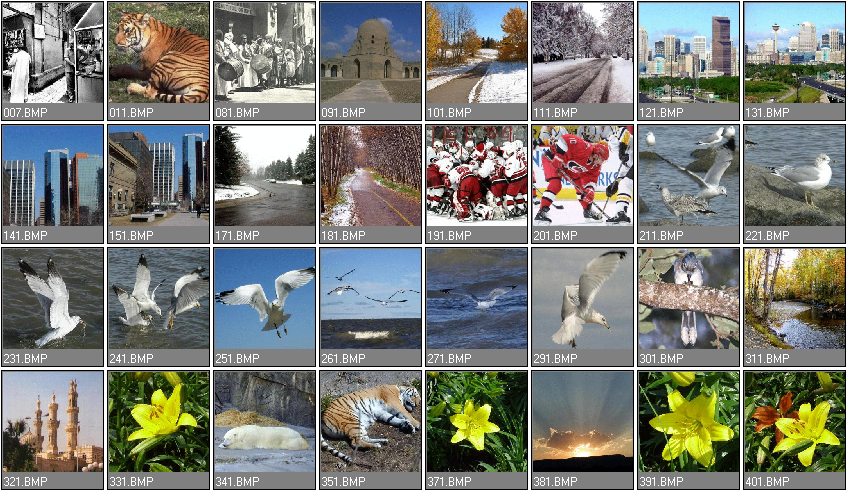
Το Netra και το σύστημα Blobworld συγκρίνουν τις εικόνες βασιζόμενοι σε μεμονωμένες περιοχές. Παρόλο που ερώτημα που βασίζεται σε περιορισμένο αριθμό περιοχών επιτρέπεται, το ερώτημα γίνεται με τη συγχώνευση των αποτελεσμάτων αναζήτησης μίας μόνο περιοχής. Το κίνητρο είναι να στραφεί μέρος της έργου της σύγκρισης στους χρήστες. Για να γίνει ένα ερώτημα εικόνας, ένας χρήστης είναι εφοδιασμένο με την κατά τμηματοποιημένες περιοχές της εικόνας και καλείται να επιλέξει τις περιοχές ενώ θα πρέπει να συμπληρώνεται και χαρακτηριστικά, για παράδειγμα, το χρώμα και την υφή, των περιοχών που πρέπει να χρησιμοποιηθούν για την αξιολόγηση της ομοιότητας. Τέτοια συστήματα αναζήτησης παρέχουν περισσότερο έλεγχο στο χρήστη. Ωστόσο, η σημασιολογική κατανόηση του χρήστη για μία εικόνα βρίσκεται σε υψηλότερο επίπεδο από ό, τι η αναπαράσταση με βάση την περιοχή. Για αντικείμενα που δεν έχουν οξυδερκή χαρακτηριστικά, όπως η ειδική υφή, δεν είναι προφανές για τον χρήστη πώς να επιλέξει ένα ερώτημα από μία μεγάλη ποικιλία επιλογών. Έτσι, μια τέτοια αναζήτηση μπορεί να προσθέτει βάρη για χρήστες που δεν έχουν σημαντική ανταμοιβή. Από την άλλη πλευρά, λόγω της μεγάλης δυσκολίας για την επίτευξη ακριβούς κατάτμησης, συστήματα όπως [11],η το [2] συχνά χωρίζουν ενα αντικειμένο σε διάφορες περιοχές με καμία από αυτές να είναι αντιπροσωπευτική για το αντικείμενο, ιδίως για τις εικόνες δεν διαθέτουν διακριτά αντικείμενα και σκηνές.

Δεν έχει δοθεί μεγάλη προσοχή στην ανάπτυξη μέτρων ομοιότητας που συνδυάζουν στοιχεία από όλες τις περιοχές. Μία προσπάθεια προς αυτή την κατεύθυνση είναι η αναζήτηση σε σύστημα που αναπτύχθηκε από τους Smith και Li [22]. Το σύστημα τους χωρίζει μια εικόνα σε περιοχές με προκαθορισμένους χαρακτηρισμούς από έναν πεπερασμένο μοτίβο. Με κάθε πρότυπο χαρακτηρισμένο από ένα σύμβολο, οι εικόνες εκπροσωπούνται στη συνέχεια από region strings. Τα region strings μετατρέπονται σε σύνθετους πίνακες (CRT descriptor matrices) που παρέχουν τη σχετική διάταξη των συμβόλων. Η ομοιότητα μεταξύ των εικόνων μετράται από την εγγύτητα μεταξύ των πινάκων CRT. Το μέτρο αυτό είναι ευαίσθητο στη μετατόπιση αντικείμενου αφού ένας πίνακας CRT καθορίζεται αποκλειστικά από την διάταξη των συμβόλων. Το μέτρο αυτό υστερεί όσο αφορά την περιστροφή και την αλλαγή κλίμακας. Επειδή ο ορισμός του πίνακα CRT επικαλείται τη βιβλιοθήκη πρότυπο, η απόδοση του συστήματος εξαρτάται καθοριστικά από τη βιβλιοθήκη. Οι επιδόσεις μειώνονται αν τα είδη των περιοχών σε μια εικόνα δεν εκπροσωπούνται από τα πρότυπα της βιβλιοθήκης. Το σύστημα χρησιμοποιεί μία CRT βιβλιοθήκη με τα πρότυπα που περιγράφονται μόνο από το χρώμα. Αν η υφή και το σχήμα ως χαρακτηριστικό γνωρίσμα χρησιμοποιηθούν επίσης για τη διάκριση των πρότυπων, ο αριθμός των προτύπων στην βιβλιοθήκη θα αυξηθεί δραματικά, περίπου εκθετικά του αριθμού των χαρακτηριστικών, εφόσον τα πρότυπα έχουν παραχθεί από ομοιόμορφα χαρακτηριστικά.

Οργάνωση μία συλλογής εικόνων

Ενώ αυτή η έρευνα επικεντρώνεται κυρίως στις τεχνικές για την αποθήκευση και ανάκτηση των ηλεκτρονικών εικόνων, είναι χρήσιμο να προβληματιστούμε σχετικά με τις παραδοσιακές πρακτικές της εικόνας και άλλων συλλογών εικόνων και βίντεο που υπάρχουν και έχουν παραχθεί με μη ηλεκτρονικό τρόπο. Υπάρχουν έτσι συλλογές εικόνας διαφόρων τύπων που συντηρούνται από ένα ευρύ φάσμα οργανισμών, όλων των μεγεθών και σε ποικίλους τομείς.

Η βάση δεδομένων που θα χρησιμοποιηθηκε στην έργασία μας περιέχει 1000 8-bit ασυμπίεστα **BMP** που έχουν επιλεγεί από τη συλλογή της Corel. Η παρακάτω εικόνα απεικονίζει ένα δείγμα των εικόνων στη βάση δεδομένων:



**Εικόνα: Βάση δεδομένων**

Προτεινόμενη λύση

Η προτεινόμενη λύση αρχικά ήταν να εξαχθούν τα πρωτόγονα χαρακτηριστικά γνωρίσματα του ερωτήματος και να συγκριθούν με εκείνα των εικόνων της βάσης δεδομένων. Τα χαρακτηριστικά γνωρίσματα της εικόνας υπό εξέταση είναι το χρώμα, υφή και η μορφή. Κατά συνέπεια, χρησιμοποιώντας τους αλγορίθμους ταιριάσματος και σύγκρισης, τα χαρακτηριστικά γνωρίσματα χρώματος, υφής και μορφής μιας εικόνας συγκρίνονται και αντιστοιχούνται με τα αντίστοιχα χαρακτηριστικά γνωρίσματα μιας άλλης εικόνας. Αυτή η σύγκριση εκτελείται χρησιμοποιώντας τις μετρικές απόστασης χρώματος, υφής και μορφής. Στο τέλος, αυτές οι μετρικές διενεργούνται η μια μετά από άλλη, ώστε να ανακτηθούν οι εικόνες της βάσεων δεδομένων που είναι παρόμοιες με την ερώτηση. Η ομοιότητα μεταξύ των χαρακτηριστικών γνωρισμάτων επρόκειτο να υπολογιστεί χρησιμοποιώντας τους αλγορίθμους που χρησιμοποιήθηκαν από τα γνωστά συστήματα CBIR όπως QBIC της ΙΒΜ. Για κάθε συγκεκριμένο χαρακτηριστικό γνώρισμα υπήρξε ένας συγκεκριμένος αλγόριθμος για την εξαγωγή και άλλος για το ταίριασμα.

# Color Space

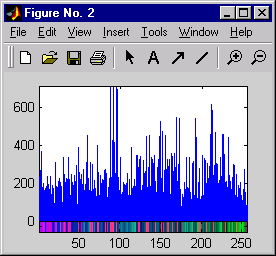
Κάθε pixel της εικόνας μπορεί να αντιπροσωπευθεί ως σημείο σε ένα τρισδιάστατο χώρο χρώματος. Το συνήθως χρησιμοποιημένο διάστημα χρώματος για την ανάκτηση εικόνας περιλαμβάνει RGB, Munsell, *CIE L\*a\*b\**, *CIE L\*u\*v\**,*HSV* (ή HSL, HSB. Δεν υπάρχει καμία συμφωνία στην πιο είναι το καλύτερο. Εντούτοις, ένα από τα επιθυμητά χαρακτηριστικά ενός κατάλληλου διαστήματος χρώματος για την ανάκτηση εικόνας είναι η ομοιομορφία του [ 65 ]. Η ομοιομορφία (**Uniformity**) σημαίνει ότι δύο ζευγάρια χρώματος που είναι ίσα στην απόσταση ομοιότητας σε περιοχή χρώματος γίνονται αντιληπτά σαν ίσα από ένα παρατηρητή. Με άλλα λόγια, η μετρημένη εγγύτητα μεταξύ των χρωμάτων πρέπει να αφορά άμεσα την ψυχολογική ομοιότητα μεταξύ τους. Το RGB διάστημα είναι ένα ευρέως χρησιμοποιημένο διάστημα χρώματος για την παρουσίαση εικόνας. Αποτελείται από τρεις περιοχές χρώματος κόκκινο, πράσινο, και το μπλε. Αυτά τα συστατικά καλούνται " *additive primaries* " δεδομένου ότι ένα χρώμα στο RGB διάστημα παράγεται με το να τα προσθέσει από κοινού. Αντίθετα, το διάστημα CMY είναι ένα διάστημα χρώματος που χρησιμοποιείται κυρίως για την εκτύπωση. Τα τρία τμήματα που συνθέτουν ένα χρώμα είναι κυανό, μωβ (magenta), και κίτρινo. Αυτά τα τρία συστατικά καλούνται "subtractive *primaries* " δεδομένου ότι ένα χρώμα στο διάστημα CMY παράγεται μέσω της ελαφριάς απορρόφησης. Και το RGB και το διάστημα CMY είναι εξαρτώμενα από τη συσκευή και ανομοιόμορφα στην αντίληψη.

Το CIE L\*a\*b\* και CIE L\*u\*v\* διαστήματα είναι ανεξάρτητα συσκευών και είναι ομοιόμορφο στην αντίληψη. Αποτελούνται από ένα τμήμα φωτεινότητας ή ελαφρότητας (L) και δύο χρωματικά συστατικά α και β ή u και V. CIE L\*a\*b \* σχεδιάζονται για να εξετάσουν τα αφαιρετικά (subtractive) μίγματα χρωμάτων, ενώ CIE L\*u\*v \* έχει ως σκοπό να εξετάσει τα προσθετικά (additive) μίγματα χρωμάτων. Το HSV (ή HSL, ή HSB) διάστημα χρησιμοποιείται ευρέως στην ηλεκτρονική γραφιστική και είναι ένας πιό διαισθητικός τρόπος για να περιγραφεί το χρώμα. Τα τρία τμήματα χρώματος είναι η απόχρωση, ο κορεσμός (ελαφρότητα) και η φωτεινότητα . Η απόχρωση είναι αμετάβλητη στις αλλαγές στο φωτισμό και την κατεύθυνση των φωτογραφικών μηχανών και ως εκ τούτου πιο κατάλληλη στην ανάκτηση αντικειμένου. Οι RGB συντεταγμένες μπορούν να μεταφραστούν εύκολα στις συντεταγμένες HSV (ή HLS, ή HSB) από έναν απλό τύπο [ 27 ].

Το αντίπαλο διάστημα χρώματος χρησιμοποιεί τους αντίπαλους άξονες χρώματος (ρ-γ, 2ψ-ρ-γ, R+G+B). Αυτή η αντιπροσώπευση έχει το πλεονέκτημα τις απομόνωσης της πληροφορίας φωτεινότητας στον τρίτο άξονα. Με αυτήν την λύση, οι πρώτοι δύο άξονες χρωματικότητας, που είναι αμετάβλητοι στις αλλαγές της έντασης και των σκιων φωτισμού, μπορούν να κάτω-επιλεχτούν δεδομένου ότι οι άνθρωποι είναι πιό ευαίσθητοι στη φωτεινότητα από είναι στις χρωματικές πληροφορίες.

## Μέθοδος αναπαράστασης

Η βασική μέθοδος αναπαράστασης της πληροφορία χρώματος των εικόνων στα συστήματα CBIR είναι μέσω του ιστογράμματος χρώματος. Ένα ιστόγραμμα χρώματος είναι ένας τύπος γραφικής παράστασης ράβδων, όπου κάθε ράβδος αντιπροσωπεύει ένα ιδιαίτερο χρώμα του διαστήματος χρώματος που χρησιμοποιείται. Σε MatLab παραδείγματος χάριν μπορείτε να πάρετε ένα ιστόγραμμα χρώματος μιας εικόνας στο RGB ή στο HSV διάστημα. Οι ράβδοι σε ένα ιστόγραμμα χρώματος αναφέρονται ως δοχεία (bins) και αντιπροσωπεύουν τον Χ-άξονα. Ο αριθμός των bins εξαρτάται από τον αριθμό χρωμάτων που υπάρχουν σε μια εικόνα. Ο Υ-άξονας δείχνει τον αριθμό pixels που υπάρχουν κάθε bin. Με άλλα λόγια πόσα pixels σε μια εικόνα είναι ενός ιδιαίτερου χρώματος. Ένα παράδειγμα ενός ιστογράμματος χρώματος στο διάστημα HSV μπορεί να δει με την ακόλουθη εικόνα:

Για να δει κάποιος ένα ιστόγραμμα αριθμητικά πρέπει να εξετάσει το χάρτη χρώματος ή την αριθμητική αντιπροσώπευση κάθε bin.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ***Colour Map***  ***(x-axis)*** | | |  | ***Number of Pixels per Bin***  ***(y-axis)*** |
| H | ***S*** | ***V*** |  |
| 0.9922 | 0.9882 | 0.9961 |  | 106 |
| 0.9569 | 0.9569 | 0.9882 |  | 242 |
| 0.9725 | 0.9647 | 0.9765 |  | 273 |
| 0.9176 | 0.9137 | 0.9569 |  | 372 |
| 0.9098 | 0.8980 | 0.9176 |  | 185 |
| 0.9569 | 0.9255 | 0.9412 |  | 204 |
| 0.9020 | 0.8627 | 0.8980 |  | 135 |
| 0.9020 | 0.8431 | 0.8510 |  | 166 |
| 0.9098 | 0.8196 | 0.8078 |  | 179 |
| 0.8549 | 0.8510 | 0.8941 |  | 188 |
| 0.8235 | 0.8235 | 0.8941 |  | 241 |
| 0.8471 | 0.8353 | 0.8549 |  | 104 |
| 0.8353 | 0.7961 | 0.8392 |  | 198 |
| . | . | . |  | . |
| . | . | . |  | . |
| . | . | . |  | . |

Όπως φαινεταιι από το χάρτη χρώματος κάθε σειρά αντιπροσωπεύει το χρώμα ενός bin. Η σειρά αποτελείται από τις τρεις συντεταγμένες του διαστήματος χρώματος. Η πρώτη συντεταγμένη αντιπροσωπεύει την απόχρωση, η δεύτερη κορεσμό (saturation), και η τρίτη, αξία (value). Τα ποσοστά από κάθε μια από αυτές τις συντεταγμένες είναι από τι τελικά αποτελεί το χρώμα ενός bin. Επίσης κάποιος μπορεί να δει τους αντίστοιχους αριθμούς pixels για κάθε bin, οι οποίοι φαινονται από τις μπλε γραμμές στο ιστόγραμμα.

## Quadratic Distance Metric

**Η εξίσωση που χρησιμοποιήσαμε στην παραγωγή της απόστασης μεταξύ δύο ιστογράμματων είναι η δευτέρου βαθμού μετρική απόσταση:**



Η εξίσωση αποτελείται από τρεις όρους. Το πως προκύπτει κάθε ένας από αυτούς τους όρους θα εξηγηθεί στη συνέχεια. Ο πρώτος όρος αποτελείται από τη διαφορά μεταξύ δύο ιστογράμματων ή ακριβέστερα τη διαφορά στον αριθμό pixels σε κάθε bin. Αυτός ο όρος είναι προφανώς ένα διάνυσμα δεδομένου ότι αποτελείται από μια σειρά. Ο αριθμός στηλών σε αυτό το διάνυσμα είναι ο αριθμός bins σε ένα ιστόγραμμα. Ο τρίτος όρος είναι που μεταθέτει εκείνο το διάνυσμα. Ο μέσος όρος είναι ο πίνακας ομοιότητας. Το τελικό αποτέλεσμα d αντιπροσωπεύει την απόσταση χρώματος μεταξύ δύο εικόνων. Όσο πιό κοντινή είναι η απόσταση στο μηδέν που τόσο πιό κοντά είναι οι εικόνες στην ομοιότητα χρώματος. Ο περαιτέρω η απόσταση από μηδέν λιγότερο ο παρόμοιος οι εικόνες είναι στην ομοιότητα χρώματος.

**Histograms**

Η κβαντοποίηση από την άποψη των ιστογράμμάτων χρώματος αναφέρεται στη διαδικασία της μείωσης του αριθμού bins με τη λήψη των χρωμάτων που είναι πολύ παρόμοια το ένα με το άλλο και βάζοντας τους στο ίδιο bin. Εξ ορισμού ο μέγιστος αριθμός δοχείων που κάποιο μπορεί να λάβει με τη χρησιμοποίηση της λειτουργίας ιστογράμμάτων σε MatLab είναι 256. Με σκοπό την εξοικονόμηση χρόνου κατά προσπάθεια για να συγκρίνει τα ιστογράμματα ς, το μπορεούμε να κβαντοποιήσουμε τον αριθμό των bins. Προφανώς η κβαντοποίηση μειώνει τις πληροφορίες σχετικά με το περιεχόμενο των εικόνων αλλά όπως αναφέρθηκε αυτό είναι η ανταλλαγή που πρέπει να γίνει όταν θέλουμε να μειώσουμε το χρόνο επεξεργασίας. Υπάρχουν δύο τύποι ιστογραμμάτων χρώματος, σφαιρικά ιστογράμματα χρώματος (GCHs) και τοπικά ιστογράμμων χρώματος (LCHs). Ένα GCH αντιπροσωπεύει μια ολόκληρη εικόνα με ένα ενιαίο ιστόγραμμα χρώματος. Ένα LCH διαιρεί μια εικόνα σε σταθερές ομάδες δεδομένων και παίρνει το ιστόγραμμα χρώματος για κάθε μια από εκείνες τις ομάδες δεδομένων [ 7 ]. Τα LCHs περιέχουν περισσότερες πληροφορίες για μια εικόνα αλλά είναι υπολογιστικά ακριβά στη σύγκριση των εικόνων. Το GCH είναι η παραδοσιακή μέθοδος για την ανάκτηση εικόνας βασισμένη στο χρώμα. Εντούτοις, δεν περιλαμβάνει τις πληροφορίες σχετικά με τη διανομή χρώματος των περιοχών [ 7 ] μιας εικόνας. Κατά συνέπεια κατά σύγκριση GCHs τον κάποιος μπορεί να μην πάρει πάντα ένα κατάλληλο αποτέλεσμα από την άποψη της ομοιότητας των εικόνων.

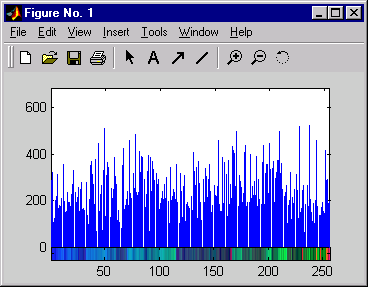
Χρησιμοποιήσαμε GCHs στην εξαγωγή των χαρακτηριστικών γνωρισμάτων χρώματος των εικόνων. Στην ανάλυση των ιστογραμμάατων υπήρξαν μερικά ζητήματα που έπρεπε να εξεταστούν. Πρώτα υπήρξε το ζήτημα πόσο θα κβαντοποιούσαμε τον αριθμό δοχείων σε ένα ιστόγραμμα. Εξ ορισμού ο αριθμός δοχείων που αντιπροσωπεύονται στο ιστόγραμμα χρώματος μιας εικόνας που χρησιμοποιεί τη λειτουργία imhist () σε MatLab είναι 256.

Αυτό σημαίνει ότι στους υπολογισμούς του πίνακα ομοιότητας και διαφοράς ιστογραμμάτων, η επεξεργασία θα ήταν υπολογιστικά ακριβή. Αρχικά αποφασίσαμε να κβαντοποιήσουμε τον αριθμό δοχείων σε 20. Αυτό σημαίνει ότι τα χρώματα που είναι ευδιάκριτα όμως παρόμοια ανατίθενται στο ίδιο δοχείο που μειώνει τον αριθμό δοχείων από 256 έως 20. Αυτό μειώνει προφανώς το περιεχόμενο πληροφοριών των εικόνων, αλλά μειώνει το χρόνο στον υπολογισμό της απόστασης χρώματος μεταξύ δύο ιστογραμμάτων. Αφ' ετέρου το να διατηρήσουμε των αριθμού των bins σε 256 δίνει ένα ακριβέστερο αποτέλεσμα από την άποψη της απόστασης χρώματος. Αργότερα επιστρέψαμε σε 256 bins λόγω μερικών ασυνεπειών που λήφθηκαν στις αποστάσεις χρώματος μεταξύ των εικόνων. Αυτό δεν είχε καμία σχέση με την κβαντοποιήση της εικόνας αλλά μάλλον με τους τύπους εικόνων χρησιμοποιούσαμε που θα αναλύσουμε περαιτέρω στο τμήμα αποτελεσμάτων.

Το δεύτερο ζήτημα ήταν σε ποιό διάστημα χρώματος θα παρουσιάζαμε το χάρτη χρώματός μας. Θα έπρεπε αυτό να είναι RGB ή HSV; Αυτό λύθηκε αμέσως όταν διαπιστώσαμε ότι η εξίσωση που παράγει τον πίνακα ομοιότητας QBIC χρησιμοποιούσε το διάστημα χρώματος HSV στους υπολογισμούς της. Δεν έχουν υπάρξει οποιαδήποτε στοιχεία που να λένε ποιο διάστημα χρώματος παράγει τα καλύτερα αποτελέσματα ανάκτησης, κατά συνέπεια η χρήση αυτού του διαστήματος χρώματος δεν μας περιόρισε.

## Similarity Matrix

Όπως μπορεί να δει κανείς από τα ιστογράμματα χρώματος δύο εικόνων q και i στα παρακάτω σχήματα, τα πρότυπα χρώματος που παρατηρούνται στη ράβδο χρώματος είναι συνολικά διαφορετικά. Αυτό επιβεβαιώνεται περαιτέρω όταν βλέπει το ένα τους αντίστοιχους χάρτες χρώματος που ακολουθούν:



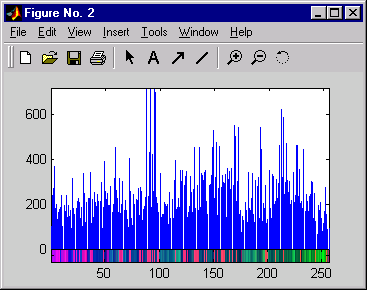
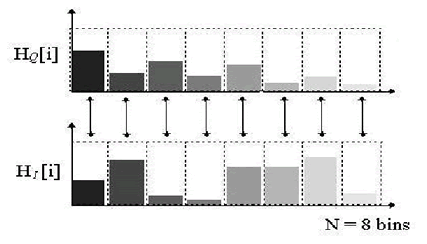


Figure: Ιστογράμματα δύο εικόνων.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| ***Color Map of image Q*** | | |  | ***Color Map of image I*** | | |
| 0.9608 | 0.8980 | 0.7843 |  | 0.9922 | 0.9882 | 0.9961 |
| 0.9373 | 0.9059 | 0.8235 |  | 0.9569 | 0.9569 | 0.9882 |
| 0.9098 | 0.8510 | 0.7765 |  | 0.9725 | 0.9647 | 0.9765 |
| 0.9255 | 0.8588 | 0.8039 |  | 0.9176 | 0.9137 | 0.9569 |
| 0.8627 | 0.8275 | 0.7961 |  | 0.9098 | 0.8980 | 0.9176 |
| 0.9098 | 0.8431 | 0.7216 |  | 0.9569 | 0.9255 | 0.9412 |
| 0.9137 | 0.8392 | 0.6627 |  | 0.9020 | 0.8627 | 0.8980 |
| 0.9059 | 0.7882 | 0.6510 |  | 0.9020 | 0.8431 | 0.8510 |
| 0.9451 | 0.8275 | 0.6824 |  | 0.9098 | 0.8196 | 0.8078 |
| 0.9569 | 0.7882 | 0.5922 |  | 0.8549 | 0.8510 | 0.8941 |
| 0.9137 | 0.7765 | 0.5961 |  | 0.8235 | 0.8235 | 0.8941 |
| 0.9412 | 0.7961 | 0.5569 |  | 0.8471 | 0.8353 | 0.8549 |
| 0.8471 | 0.7843 | 0.7176 |  | 0.8353 | 0.7961 | 0.8392 |
| 0.8275 | 0.7843 | 0.6745 |  | 0.8431 | 0.7804 | 0.7843 |
| 0.9020 | 0.8392 | 0.6667 |  | 0.7961 | 0.7804 | 0.8353 |
| 0.8980 | 0.7333 | 0.5843 |  | 0.7882 | 0.7725 | 0.7882 |
| 0.9020 | 0.7216 | 0.5333 |  | 0.8235 | 0.8314 | 0.8118 |
| . | . | . |  | . | . | . |
| . | . | . |  | . | . | . |
| . | . | . |  | . | . | . |

###### Πίνακας: Χάρτης χρώματος των δύο εικόνων.

Μια απλή μετρική απόστασης περιλαμβάνει την αφαίρεση του αριθμού pixels στο 1ο bin ενός ιστογράμματος από το 1ο bin ενός άλλου αλλά δεν είναι και τόσο επαρκής. Αυτή η μετρική αναφέρεται ως **απόσταση *Minkowski***, και εμφανιζεται πιο κάτω, και η οποία συγκρίνει μόνο τα bins τα οποία είναι ίδια μεταξύ των ιστογραμμάτων χρώματος [ 3 ].



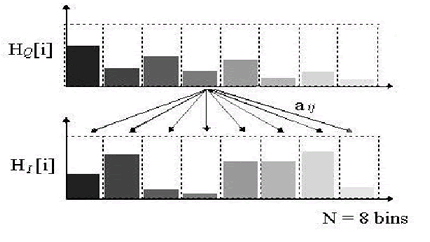
Σχήμα: Minkowski απόσταση

Η **απόσταση Minkowski** μεταχειρίζεται όλα τα bins του ιστογράμματος εξ ολοκλήρου ανεξάρτητα και δεν λαμβάνει υπ’όψη το γεγονός ότι ορισμένα ζευγάρια των bins αντιστοιχούν στα χαρακτηριστικά γνωρίσματα που είναι σχετικά πιό παρόμοια από τα άλλα ζευγάρια.

Αυτό είναι ο βασικός λόγος για τη μετρική απόστασης δευτέρου βαθμού. Ακριβέστερα είναι ο μέσος όρος της εξίσωσης ή του πίνακα Α ομοιότητας που μας βοηθά να υπερνικήσουμε το πρόβλημα των διαφορετικών χαρτών χρώματος. Ο πίνακας ομοιότητας λαμβάνεται μέσω ενός σύνθετου αλγορίθμου:

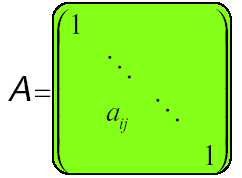


ο οποίος συγκρίνει βασικά ένα bin χρώματος του Hq με όλους εκείνους Ηi για να προσπαθήσει και να ανακαλύψει πιο bin είναι πίο όμοιο, όπως εμφανίζεται κατωτέρω:



Σχήμα: Quadratic απόσταση

Αυτό συνεχίζεται έως ότου έχουμε συγκρίνει όλα τα δοχεία χρώματος του Hq. Με αυτό τον τρόπο παίρνουμε ΝxΜ πίνακα, με το Ν να αντιπροσωπεύει τον αριθμό των bins. Αυτό που δείχνει εάν τα πρότυπα χρώματος δύο ιστογραμμάτων είναι παρόμοια είναι η διαγώνιος του πίνακα, που εμφανίζεται κατωτέρω. Εάν η διαγώνιος αποτελείται εξ ολοκλήρου από άσσους έπεται ότι τα πρότυπα χρώματος είναι ίδια. Όσο πιο μακριά είναι οι αριθμοί στη διαγώνιο από το ένα, τόσο λιγότερο τα πρότυπα χρώματος είναι όμοια. Κατά συνέπεια το πρόβλημα της συγρισης ανεξάρτητων bins λύνεται.



**Σχήμα: Πίνακας ομοιότητας A, με διαγώνιο από άσσους [3]**

Έχει αποδειχθεί ότι η δευτέρου βαθμού απόσταση μορφής μπορεί να οδηγήσει στα πιό επιθυμητά αποτελέσματα από την euclidean μέθοδο και μέθοδο ιστογραμμάτων δεδομένου ότι εξετάζει τη διαγώνια ομοιότητα μεταξύ των χρωμάτων.

**Texture**

Στον τομέα της τεχνητής όρασης και της επεξεργασίας εικόνας δεν υπάρχει κανένας ευδιάκριτος ορισμός της υφής. Αυτό είναι επειδή οι διαθέσιμοι ορισμοί υφής είναι βασισμένοι στις μεθόδους ανάλυσης της υφής και των χαρακτηριστικών γνωρισμάτων που εξάγονται από την εικόνα. Εντούτοις η υφή μπορεί να θεωρηθεί ως επαναλαμβανόμενα πρότυπα των pixels πέρα από μια χωρική δικτυακή γειτονιά, της οποίας η προσθήκη του θορύβου στα πρότυπα και οι επαναληπτικές συχνότητές τους, οδηγούν τις υφές που είναι τυχαίες και μη δομημένες. Οι ιδιότητες της υφής είναι τα οπτικά πρότυπα σε μια εικόνα που έχουν τις ιδιότητες της ομοιογένειας και που δεν προκύπτουν από την παρουσία μόνο ενός ενιαίου χρώματος ή μιας έντασης. Οι διαφορετικές ιδιότητες υφής όπως γίνονται αντιληπτές από το ανθρώπινο μάτι είναι τακτικότητα, κατευθυντικότητα, η ομαλότητα και οξύτητα (coarseness). Οι υφές των εικόνων έχουν χρήσιμες εφαρμογές στην επεξεργασία εικόνας και στην τεχνιτή όραση. Περιλαμβάνουν

* Την αναγνώριση των περιοχών εικόνας χρησιμοποιώντας τις ιδιότητες υφής, διαφορετικά γνωστή ως ταξινόμηση υφής.
* Αναγνώριση των ορίων υφής χρησιμοποιώντας τις ιδιότητες της υφής, διαφορετικά γνωστό και ως κατάτμηση υφής.
* Σύνθεση υφής, παραγωγή των εικόνων υφής από τα γνωστά μοντέλα υφής.
* Εξαγωγή των σχημάτων της εικόνας που χρησιμοποιεί τις ιδιότητες υφής.

**Σχήμα : Παραδείγματα υφής**

## Texture feature extraction

Δεδομένου ότι δεν υπάρχει κανένας αποδεκτός μαθηματικός καθορισμός για την υφή, πολλές διαφορετικές μέθοδοι για τα χαρακτηριστικά γνωρίσματα της έχουν προταθεί μέχρι σήμερα. Δυστυχώς, δεν υπάρχει ακόμα ούτε μια μέθοδος που λειτουργεί καλύτερα με όλους τους τύπους υφής. Οι συνήθως χρησιμοποιημένες μέθοδοι για την περιγραφή χαρακτηριστικών γνωρισμάτων σύστασης είναι στατιστικές και βασισμένες σε μετατροπή μέθοδοι [ 8 ],[ 9 ].

#### Statistical method

Οι στατιστικές μέθοδοι αναλύουν τη χωρική διανομή των γκρίζων τιμών με τον υπολογισμό των τοπικών χαρακτηριστικών γνωρισμάτων σε κάθε σημείο στην εικόνα, και την άντληση ενός συνόλου στατιστικών από τη διανομή των τοπικών χαρακτηριστικών. Για βέλτιστους λόγους ταξινόμησης, αυτό που μας ενδιαφέρει είναι οι στατιστικές τεχνικές χαρακτηρισμού. Αυτό είναι επειδή είναι αυτές οι τεχνικές που οδηγούν στον υπολογισμό των ιδιοτήτων υφής. Οι δημοφιλέστερες στατιστικές αντιπροσωπεύσης της υφής είναι:

* Co-occurrence Matrix
* Tamura Texture
* Wavelet Transform

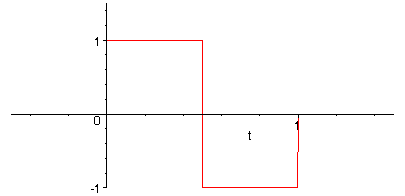
**Pyramid-Structured Wavelet Transform**

Οι υφές μπορούν να διαμορφωθούν ως περιοδικά πρότυπα με τη χωρική αντιπροσώπευση συχνότητας. Ο μετασχηματισμός wavelet μετασχηματίζει την εικόνα σε μια αντιπροσώπευση multi-scale και με χωρικά και με χαρακτηριστικά συχνότητας. Αυτό επιτρέπει την αποτελεσματική ανάλυση μιας multi-scale εικόνας είναι το χαμηλότερο υπολογιστικό κόστος [ 10 ]. Σύμφωνα με αυτόν τον μετασχηματισμό, μια λειτουργία, που μπορούν να αντιπροσωπεύεται σε μια εικόνα, με μια καμπύλη, η με ένα σήμα κ.λπ., μπορούν να περιγραφούν από την άποψη μιας χονδροειδούς περιγραφής επιπέδων με τις λεπτομέρειες να κυμαίνονται από ευρείες σε στενές κλίμακες [ 12 ].

Αντίθετα από τη χρήση ημιτονικών συναρτήσεων που χρησιμοποιούνται σε μετασχηματισμούς Fourier, στο μετασχηματισμό wavelet, χρησιμοποιούμε τις λειτουργίες γνωστές ως wavelets. Τα wavelets είναι πεπερασμένα στο χρόνο, όμως η μέση τιμή ενός wavelet είναι μηδέν [ 10 ]. Από μία άποψη, ένα wavelet είναι μία κυματομορφή που είναι φραγμένη τόσο στη συχνότητα όσο και στο χρόνο. Ενώ ο μετασχηματισμός κατά Fourier μετατρέπει ένα σήμα σε μια συνεχή σειρά κυμάτων ημιτόνου, κάθε μια από την οποίες είναι σταθερής συχνότητας και του εύρους και άπειρης διάρκειας, τα περισσότερα πραγματικά σήματα (όπως η μουσική ή οι εικόνες) έχουν μια πεπερασμένη "διάρκεια" και απότομες αλλαγές στη συχνότητα. Για αυτό και οι μετασχηματισμοί με wavelets είναι αποδοτικοί. Αυτό είναι επειδή οι μετασχηματισμοί με wavelets μετατρέπουν ένα σήμα σε μια σειρά wavelets, η οποία μπορεί να καταχωρηθεί αποτελεσματικότερα λόγω του πεπερασμένου χρόνου, και μπορεί να κατασκευαστεί με τραχιές άκρες, με αυτόν τον τρόπο προσεγγίζοντας καλύτερα τα πραγματικά σήματα.

**Haar Wavelet**

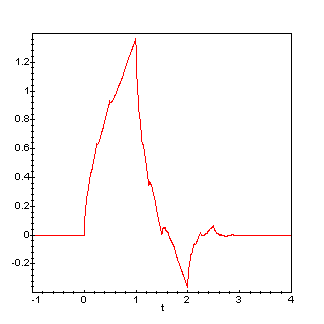
Η οικογένεια Haar wavelet ορίζεται ως [2]:



**Σχήμα: Παράδειγμα Haar Wavelet**

***Daubechies Wavelet***

Η οικογένεια Daubechies wavelet ορίζεται ως εξύς [2]:



Σχήμα: Παράδειγμα Daubechies Wavelet

Οι μετασχηματισμοί wavelets [ 21 ..62 ] παρέχει μια προσέγγιση multi-resolution στην ανάλυση και την ταξινόμηση υφής [ 19 ..54 ]. Οι μετασχηματισμοί αυτοί αποσυνθέτουν ένα σήμα με ένα σύνολο βασικών συναρτήσεων:

που έχουν αποκτηθεί μέσω της μετάφρασης και της διαστολής ενός αρχικού wavelet ψ(x), δηλ.,

όπου το m και το n είναι παράμετροι διαστολή και μεταφράσης. Μία συνάρτηση f(x) μπορεί να αναπαρασταθεί όπως:

Ο υπολογισμός των wavelets ενός δισδιάστατου σήματος περιλαμβάνει το επαναλαμβανόμενο φιλτράρισμα και την υπο--δειγματοληψία. Σε κάθε επίπεδο το σήμα αποσυντίθεται στις υποζώνες LL, LH, HL, HH τεσσάρων συχνοτήτων όπου το L δείχνει τη χαμηλή συχνότητα και το H δείχνει την υψηλή συχνότητα. Δύο σημαντικοί τύποι μετατροπής wavelet που χρησιμοποιούνται για την ανάλυση υφής είναι μετατροπή waveletς δομών πυραμίδων (pyramid wavelet trasform ή PWT) και μετατροπή wavelets δομών δέντρων (tree wavelet trasform ή TWT). Μετά από την αποσύνθεση, τα διανύσματα των χαρακτηριστικών γνωρισμάτων μπορούν να κατασκευαστούν χρησιμοποιώντας τη μέση και σταθερή απόκλιση της ενεργειακής διανομής κάθε υποζώνης σε κάθε επίπεδο. Για την τριών επιπέδων αποσύνθεση, η PWT οδηγεί σε ένα διάνυσμα χαρακτηριστικών μεγέθους 3x4x2. Για TWT, το χαρακτηριστικό γνώρισμα θα εξαρτηθεί από τον τρόπο με τον οποίο οι υποζώνες σε κάθε επίπεδο αποσυντίθενται. Ένα σταθερό δέντρο αποσύνθεσης μπορεί να ληφθεί με διαδοχικές αποσυνθέσεςι των LL, το LH, και της ζώνης HL, και έτσι τα αποτελέσματα να μπούν σε ένα διάνυσμα χαρακτηριστικών με 52x2 στοιχεία. Σημειώστε ότι για παράδειγμα, το χαρακτηριστικό γνώρισμα που λαμβάνεται από τον PWT μπορεί να θεωρηθεί ως υποσύνολο του χαρακτηριστικού γνωρίσματος που λαμβάνεται από τον TWT.

Στην εργασία αυτή χρησιμοποιήσαμε τη μετατροπή wavelets δομών πυραμίδων για την ταξινόμηση υφής. Το όνομά της προέρχεται από το γεγονός ότι αποσυνθέτει κατ' επανάληψη τα σήματα σε κανάλια χαμηλής συχνότητας. Αυτό είναι συνήθως σημαντικό για υφές με κυρίαρχα κανάλια συχνότητας. Για αυτόν τον λόγο, είναι συνήθως κατάλληλο για τα σήματα που αποτελούνται από στοιχεία με τις πληροφορίες να συγκεντρώνονται στα κανάλια χαμηλότερης συχνότητας [ 10 ]. Λόγω του ότι οι εικόνες διαθέτουν έμφυτες ιδιότητες που επιτρέπουν τις περισσότερες πληροφορίες να υπάρχουν στις χαμηλότερες υποζώνες, η μετατροπή waveletς δομών πυραμίδων είναι ιδιαίτερα ικανοποιητική. Χρησιμοποιώντας την μετατροπή wavelets δομών πυραμίδων, η εικόνα αποσυντίθεται σε τέσσερις υπο- εικόνες, στο low-low σημείο, στη low-high, high-low και high-high υποζώνη. Σε αυτό το σημείο, το ενεργειακό επίπεδο κάθε υποζώνης υπολογίζεται (δείτε τον αλγόριθμο ενεργειακών επιπέδων στο επόμενο τμήμα). Αυτό είναι η πρώτη αποσύνθεση επιπέδων. Χρησιμοποιούμε τη low-low υποζώνη για περαιτέρω υπολογισμούς.

## Ενεργειακό επίπεδο

Αλγόριθμος ενεργειακών επιπέδων:

1. Αποσυνθέτουμε την εικόνα σε τέσσερις υπο-εικόνες

2. Υπολογίζουμε την ενέργεια όλων των αποσυντεθειμένων εικόνων στην ίδια κλίμακα, με τη χρησιμοποίηση του παρακάτω τύπου:



όπου το Μ και το Ν είναι οι διαστάσεις της εικόνας, και το Χ είναι η ένταση του pixel που βρίσκεται στη σειρά ι και τη στήλη j στο χάρτη εικόνας.

1. Επαναλάμβάνουμε το βήμα 1 για τη low-low υποζώνη, έως ότου το ευρετήριο IND είναι ίσο με 5

Χρησιμοποιώντας τον παραπάνω αλγόριθμο, υπολογίστηκαν τα ενεργειακά επίπεδα των υποζωνών. Αυτό επαναλαμβάνεται πέντε φορές, για να φθάσει στην πέμπτη αποσύνθεση επιπέδων. Αυτές οι τιμές ενεργειακών επιπέδων καταχωρούνται για να χρησιμοποιηθούν στο euclidean αλγόριθμο απόστασης.

Euclidean Distance αλγόριθμος:

1. Αποσυνθέτουμε την εικόνα ερώτησης.

2. Πάρνουμε τις ενέργειες των πρώτων κυρίαρχων καναλιών Κ.

3. Για την εικόνα i στη βάση δεδομένων λάμβάνουμε τις ενέργειες Κ.

4. Υπολογίζουμε τη euclidean απόσταση μεταξύ των δύο συνόλων ενεργειών, με τον τύπο:



1. Αυξάνουμε το i. Επανάληψη από το βήμα 3.

Χρησιμοποιώντας τον παραπάνω αλγόριθμο, η εικόνα ερώτησης αναζητάται στη βάση δεδομένων εικόνας. Η euclidean απόσταση υπολογίζεται μεταξύ της εικόνας ερώτησης και κάθε εικόνας στη βάση δεδομένων. Αυτή η διαδικασία επαναλαμβάνεται έως ότου έχουν συγκριθεί όλες οι εικόνες στη βάση δεδομένων με την εικόνα ερώτησης. Με την ολοκλήρωση του euclidean αλγορίθμου απόστασης, έχουμε μια σειρά euclidean αποστάσεων, η οποία ταξινομείται στη συνέχεια. Οι πέντε κορυφαίες εικόνες παρουσιάζονται έπειτα ως αποτέλεσμα της αναζήτησης υφής.

**Ορισμός σχήματος**

Η μορφή ενός αντικειμένου είναι η χαρακτηριστική διαμόρφωση της επιφάνειας όπως αντιπροσωπεύεται από το περίγραμμα ή το περίγραμμα. Η αναγνώριση μορφής/σχήματος είναι ένας από τους τρόπους μέσω των οποίων η ανθρώπινη αντίληψη για το περιβάλλον εκτελείται. Είναι σημαντικό στην τεχνική CBIR επειδή καθορίζει περιοχές ενδιαφερόντων στις εικόνες. Στην επεξεργασία εικόνας, το σχήμα είναι μία δυαδική εικόνα που αποτελείται από το περίγραμμα ή το περίγραμμα των αντικειμένων, που λαμβάνεται μετά από την κατάτμηση. Στο σύστημα CBIR που σχεδιάζεται για τη συγκεκριμένη χώρο όπως τα εμπορικά σήματα και οι σκιαγραφίες των εργαλείων, η κατάτμηση σχήματος μπορεί να είναι αυτόματη και αποτελεσματική. Εντούτοις αυτό δεν είναι ο κανόνας για ένα σύστημα CBIR που έχει μια ετερογενή βάση δεδομένων. Σε αυτήν την περίπτωση η κατάτμηση σχήματος μπορεί να είναι δύσκολη ή μερικές φορές αδύνατη.

Τα χαρακτηριστικά γνωρίσματα των σχημάτων είναι ταξινομημένα σύμφωνα με τις χρησιμοποιούμενες τεχνικές. Είναι βασισμένα στα όρια (boundary based) και βασισμένα στην περιοχή (region based). Η boundary based τεχνική περιγράφει την περιοχή ενός σχήματος με τη χρησιμοποίηση των εξωτερικών χαρακτηριστικών της, παραδείγματος χάριν pixel κατά μήκος του ορίου ενός αντικειμένου, ενώ η τεχνική περιγράφει την περιοχή ενός σχήματος με τη χρησιμοποίηση των εσωτερικών χαρακτηριστικών της, παραδείγματος χάριν το pixel που περιλαμβάνεται στην περιοχή. Οι απλοί region based boundary based περιγραφείς (descriptors) σχήματος περιλαμβάνουν την περιοχή, την περίμετρο, την πυκνότητα, την εκκεντρικότητα, την επιμήκυνση, και τον προσανατολισμό. Οι σύνθετοι region based περιγραφείς περιλαμβάνουν τους περιγραφείς Fourier, τους περιγραφείς πλέγματος, τους κώδικες αλυσίδων και τις στατιστικές ροπές [ 23 ].

**Περιοχή**

Η περιοχή είναι ο αριθμός pixels στην περιοχή που περιγράφεται από το σχήμα. Ο πραγματικός τομέας κάθε pixel μπορεί να ληφθεί υπόψη για να πάρουμε το πραγματικό μέγεθος μιας περιοχής. Στο σχήμα που ακολουθεί κάθε pixel έχει τον τομέα μιας τετραγωνικής μονάδας. Η συνολική περιοχή που αντιπροσωπεύεται από το σχήμα είναι 28 τετραγωνικές μονάδες επειδή ο συνολικός αριθμός των pixels μέσα στην περιοχή του σχήματος είναι 28.

**Περίμετρος**

Η περίμετρος είναι ο αριθμός των pixel στο όριο του σχήματος. Στο σχήμα που ακολουθεί ο συνολικός αριθμός pixel στο όριο του σχήματος είναι 32.

**Πυκνότητα**

Η πυκνότητα είναι ένα μέτρο για το πόσο πολύ “packed” είναι ένα σχήμα. Η συμπαγέστερη μορφή είναι ένας κύκλος ενώ όλες οι άλλες μορφές έχουν την πυκνότητα μεγαλύτερη από αυτή ενός κύκλου. Το παράδειγμα της compact και non-compact μορφής εμφανίζεται στη συνέχεια 

**Επίδειξη της περιοχής ως περιγραφέα σχήματος**



**Επίδειξη της περιμέτρου ως περιγραφέα σχήματος** 

**Παράδειγμα compact and non-compact σχήματος**

**Εκκεντρικότητα (Eccentricity)**

Η εκκεντρικότητα είναι η αναλογία της μακρύτερης χορδής ενός διαμορφωμένου αντικειμένου στην πιό μακριά χορδή κάθετο σε αυτό. Η εκκεντρικότητα είναι ένα μέτρο για το πόσο κυκλικό είναι ένα σχήμα. Για μια τέλεια κυκλική μορφή η εκκεντρικότητα είναι μηδέν. Οι ελλειπτικές τροχιές έχουν τα eccentricities μεταξύ μηδενός και του ένα. Παρακάτω βλέπουμε αντικείμενα με τα διαφορετικά eccentricities .



**Αντικείμενα με διαφορετικά eccentricities**

**Ροπές**

Οι κεντρικές και αναλοίωτες ροπές είναι συντελεστές που μετρούν τη διανομή της μάζας σχετικά με τους άξονες μέσω του κέντρου βάρους. Είναι βασισμένοι στην έννοια φυσικής της στιγμής της αδράνειας. Η κλασσική αντιπροσώπευση χρησιμοποιεί ένα σύνολο ροπών στιγμής. Εάν το αντικείμενο R αντιπροσωπεύεται ως δυαδική εικόνα, κατόπιν οι κεντρικές ροπές τάξης p+q για τη μορφή του αντικειμένου R ορίζονται ως:



όπου τα αθροίσματα λαμβάνονται σε όλα τα σημεία (Χ, Y) μέσα στην περιοχή. Το κέντρο βάρους της περιοχής είναι

και



το n, ο συνολικός αριθμός σημείων που περιλαμβάνονται στην περιοχή. Αυτή η κεντρική ροπή μπορεί να κανονικοποιηθεί για να είναι σταθερής κλίμακας:



Από τις ροπές , μπορούν να ληφθούν πολλές χρήσιμες πληροφορίες για το αντικείμενο. Παραδείγματος χάριν για μια δυαδική εικόνα:



Η περιοχή, δίνεται με τη μηδενική ροπή:



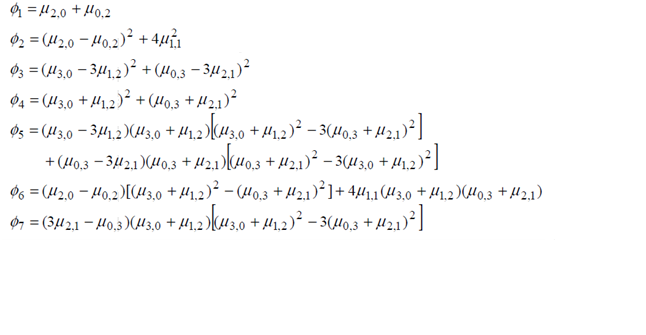
Το κέντρο της μάζας είναι η πρώτη ροπή:



Επτά ροπές προέρχονται από τις κανονικοποιημένες δεύτερες και τρίτες κεντρικές ροπές. Ονομάζονται:

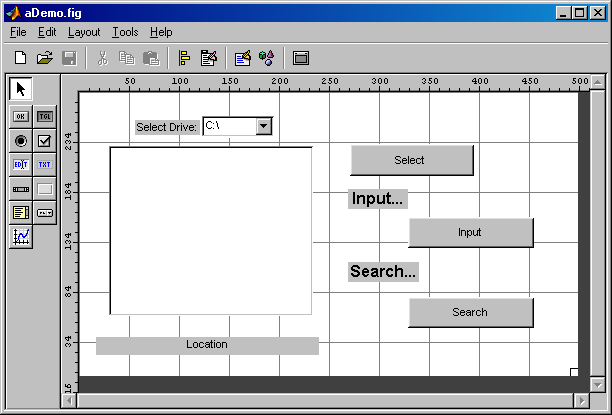
**

και είναι αμετάβλητες στις αλλαγές της θέσης, της κλίμακας και του προσανατολισμού [ 24 ]. Οι τιμές κάθε μια από επτά αμετάβλητες ροπές παρέχουν τις πληροφορίες για τη μορφή του αντικειμένου [ 25 ]. Τα φ1, φ2 είναι πάντα θετικα. Υψηλότερο φ2 σημαίνει ότι ένα σχήμα είναι πιο μεγάλο σε πλάτος από είναι σε ύψος, ενώ φ2 θα είναι μεγαλύτερο εάν το σχήμα είναι πιο ψηλό. Τα φ4, φ5 είναι μέτρα της συνδιακύμανσης. Αυτό σημαίνει είναι ότι αν τα σχήματα είναι έντονα διαγώνια, ή διαστρεφόμενα, θα δώσουν τις υψηλότερες τιμές. Τα φ6, φ7 είναι μέτρα της ασυμμετρίας. Εάν φ6 είναι θετικό, σημαίνει ότι η μορφή είναι ογκωδέστερη στο αριστερό και περισσότερο «τεντωμένη» δεξιά του κέντρου βάρους της. Εάν το φ7 είναι αρνητικό η μορφή είναι περισσότερος τεντωμένη προς τα πάνω και εάν είναι θετικό προς τα κάτω. Το σύνολο σταθερών ροπών αμετάβλητων την περιστροφή, και την αλλαγή κλίμακας είναι:



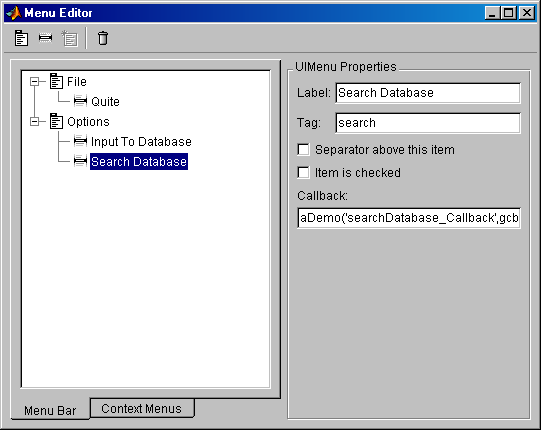
### GUI

Χρησιμοποιώντας τα εργαλεία που παρέχει το Matlab σχεδιάσαμε ένα γραφικό περιβάλλον για την εφαρμογή.



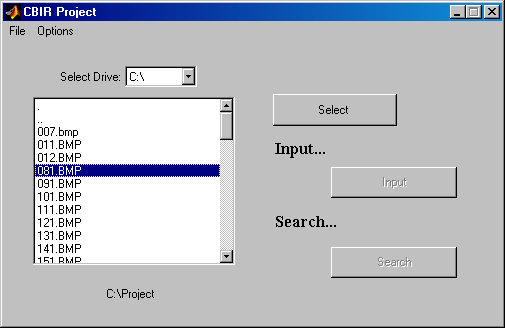
Σχήμα: GUI Design

Εκτός από το περιβάλλον που φαίνεται παραπάνω, σχεδιάσαμε επίσης μια απλή δομή καταλόγων επιλογής, χρησιμοποιώντας τον ***Menu Editor***, όπως εμφανίζεται παρακάτω:



**Σχήμα: Menu Editor επιλογές για τη δημιουργία menu**

Έτσι όταν τρέχουμε το πρόγραμμά μας βλέπουμε την παρακάτω οθόνη:



**Σχήμα: Παράθυρο εφαρμογής στο runtime**

1. **Αποτελέσματα - Παραδείγματα**

**Χρώμα**

Μετά από τη λήψη όλων των τους απαραίτητων όρων, του πίνακα ομοιότητας, και τις διαφορές ιστογράμμων χρώματος, για έναν αριθμό εικόνων στη βάση δεδομένων μας, εφαρμόσαμε τα αποτελέσματα στην τελική εξίσωση, που υπολογίζει τη δευτέρου βαθμού μετρική απόσταση. Προς μεφάλη έκπληξη διάφορες ασυνέπειες συνέχισαν από την άποψη των αποστάσεων χρώματος μεταξύ ορισμένων εικόνων. Εικόνες όπου συνολικά δεν ήταν σχετιζόμενες είχαν αποστάσεις χρώματος μικρότερες από εκείνες που ήταν πολύ παρόμοιες. Ένα παράδειγμα αυτού μπορούμε να δούμε με τις ακόλουθες τρεις εικόνες: ένα μουσουλμανικό τέμενος, ένα παιχνίδι χόκεϋ, και μια άλλη εικόνα του ίδιου παιχνιδιού χόκεϋ, όπως βλέπουμε πιο κάτω:

***(a) faceoff3 (b) faceoff4 (c) mosque***

Σχήμα: Εικόνες προς δοκιμά…

Όπως βλέπουμε από τον ακόλουθο πίνακα, τα αποτελέσματα δεν είναι σύμφωνα με το πώς οι εικόνες φαίνονται στο ανθρώπινο μάτι.

|  |  |
| --- | --- |
| ***Images*** | ***Colour distance between image histograms*** |
| faceoff3 vs faceoff4 | 10.77 |
| faceoff3 vs mosque | 9.99 |

Πίνακας: Αποστάσεις χρώματος για συμπιεσμένες εικόνες

Αυτό έγινε επανειλημμένως με διάφορες εικόνες, και οδήγησε στις ίδιες ασυνέπειες. Αυτό που αποδείχθηκε η αιτία όλου αυτού, ήταν ο τύπος εικόνων που χρησιμοποιούσαμε. Πρώτα σκεφτήκαμε ότι το μόνο πράγμα που θα μπορούσε να δώσει τα ασυμβίβαστα αποτελέσματα είναι σύγκριση εικόνων διαφορετικών μεγεθών, αλλά είχαμε επαναταξινομήσει όλες τις εικόνες στη βάση δεδομένων μας σε 256x256 πρίν εξετάσουμε τον αλγόριθμό μας. Οι εικόνες που είχαμε στη βάση δεδομένων μας ήταν όλες 24-bits JPEGs. Το πρόβλημα με τις εικόνες JPEG είναι ότι συμπιέζονται και ο αλγόριθμος συμπίεσης φαίνεται να έχει επιπτώσεις στον τρόπο που παράγονται τα ιστόγραμμα. Αυτό το ανακαλύψαμε αυτό με τη μετατροπή μερικών από τις εικόνες στη βάση δεδομένων μας σε 8-bit ασυμπίεστα bitmaps. Οι ίδιες εικόνες που εξετάζονταν με τη μορφή JPEG τώρα εξετάζονται πάλι ως BMPs. Το αποτέλεσμα ήταν σύμφωνο με το πώς οι εικόνες έμοιαζαν στο ανθρώπινο μάτι. Οι εικόνες που φάνηκαν παρόμοιες έδωσαν μικρές αποστάσεις χρώματος έναντι σε εκείνες που φαίνονταν πολύ διαφορετικές. Αυτό μπορεί να δει στον ακόλουθο πίνακα, ο οποίος παρουσιάζει τις ίδιες εικόνες με εκείνους στον προηγούμενο πίνακα αλλά με bmp τη μορφή.

|  |  |
| --- | --- |
| ***Images*** | ***Colour distance between image histograms*** |
| faceoff3 vs faceoff4 | 4.39 |
| faceoff3 vs mosque | 6.10 |

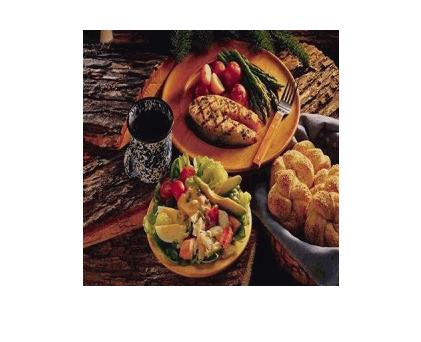
###### Πίνακας: αποστάσεις χρώματος για ασυμπίεστες εικόνες

Όταν συνειδητοποίησαμε ότι το σφάλμα μας οφειλόταν στον τύπο της εικόνας, αποφασίσαμε να μετατρέψουμε όλες τις εικόνες μας σε ασυμπίεστο BMP. Αυτό δεν είναι προφανώς σύμφωνο με τα πλήρη συστήματα CBIR που είναι διαθέσιμα στην αγορά, τα οποία παίρνουν οποιοδήποτε τύπο εικόνας ως ερώτηση με οποιαδήποτε μορφή, αλλά για τους σκοπούς αυτής της εργασίας δεν θελήσαμε να ερευνήσουμε στα ζητήματα συμπίεσης. Στη μετατροπή όλων εικόνων μας σε οκτάμπιτο ασυμπίεστο BMPs υπήρξε μια μικρή αλλαγή στον τρόπο που εξετάσαμε τους αντίστοιχους χάρτες χρώματός τους. Αυτο κάναμε προηγουμένως ήταν να δημιουργούμε ένα ευρετήριο για μια εικόνα πρίν χρησιμοποιηθεί η λειτουργία imhist. Το indexing αυτό που κάνει είναι να κβαντοποιεί το χάρτη χρώματος με να αφήνει το χρήστη να προσδιορίσει τον αριθμό των bins. Αυτό μειώνει προφανώς το χρόνο επεξεργασίας από την άποψη του υπολογισμού των αποστάσεων χρώματος δεδομένου ότι δεν έχουμε 256 bins για να συγκρίνουμε. Κατά φόρτωση των οκτάμπιτων ασυμπίεστων εικόνων σε μια μεταβλητή, το MatLab δεν μας αφήνει να κβαντοποιήσουμε τους χάρτες χρώματός τους. Μας δίνει ένα σφάλμα όταν προσπαθούμε να τους εντάξουμε σε index. Κατά συνέπεια αναγκαστήκαμε να χρησιμοποιήσουμε την προκαθορισμένη τιμή των 256 bins για όλα τα ιστόγραμμα χρώματός μας.

### Παράδειγμα

Για να δείξουμε πως δουλέυει η εφαρμογή, εφαρμόσαμε το ακόλουθο παράδειγμα:

* Ξεκινήσαμε την εφαρμογή με τη δακτυλογράφηση aDemo στο παράθυρο εντολής Matlab. Το παράθυρο εφαρμογής αρχίζει.
* Στο παράθυρο εφαρμογής, επιλέξαμε το menu *Options*, και επιλέξαμε *Search Database*. Από εκεί επιλέξαμε ένα bmp αρχείο. Σημείωση: Μόνο οκτάμπιτα ασυμπίεστα BMPs είναι κατάλληλα για αυτήν την εφαρμογή. Σε αυτό το παράδειγμα, επιλέξαμε ακόλουθο bmp.
* Το τονισμένο bmp επιλέγεται
* Πατάμε το *Search*  και ξεκινά η αναζήτηση.



Σχήμα: Η εικόνα ερώτημα: 935.bmp

* 1. Colour Extraction & Matching

Χρησιμοποιώντας τον αλγόριθμο εξαγωγής χαρακτηριστικών γνωρισμάτων χρώματος που περιγράφηκε, όπου τα ιστόγραμμα της εικόνας ερώτησης και των εικόνων στη βάση δεδομένων συγκρίνονται χρησιμοποιώντας τη δευτέρου βαθμού μετρική απόσταση, πήραμε τα ακόλουθα 10 αποτελέσματα:

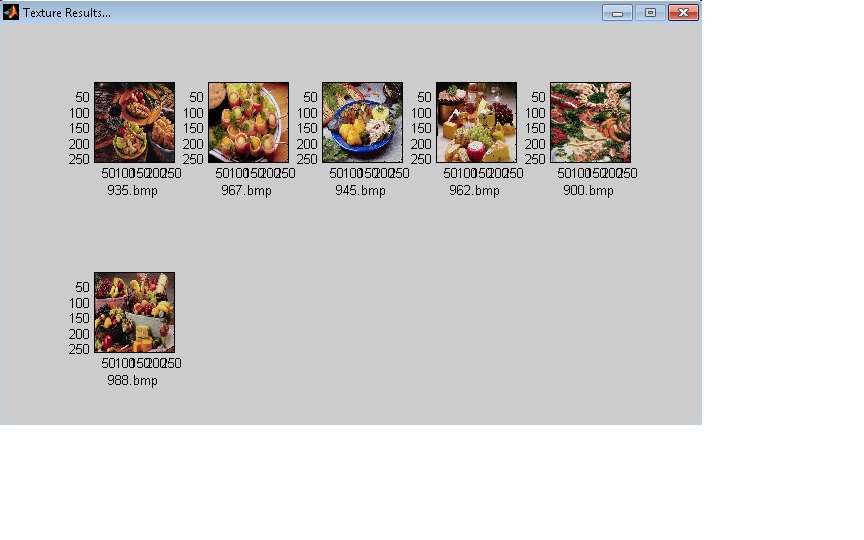


Σχήμα: Αποτελέσματα αναζήτησης με βάση το χρώμα για το 935.bmp

Τα ανωτέρω αποτελέσματα έιναι ταξινομημένα σύμφωνα με τη δευτέρου βαθμού απόσταση.

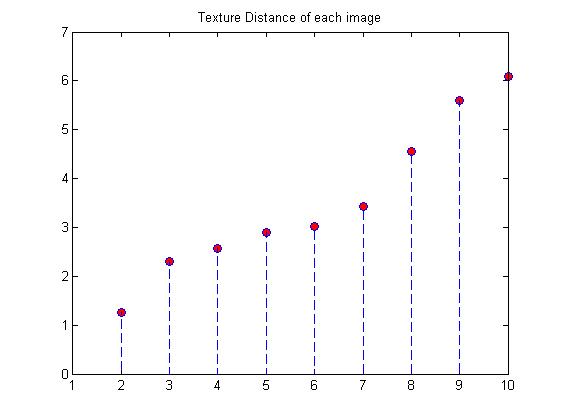
* 1. Texture Extraction & Matching

Χρησιμοποιώντας τον αλγόριθμο εξαγωγής χαρακτηριστικών γνωρισμάτων υφής που περιγράφηκε, όπου οι ενέργειες της εικόνας ερώτησης και οι υποζώνες των εικόνων που πήραμε σαν αποτέλεσμα από την εξαγωγή χρώματος συγκρίνονται, χρησιμοποιώντας τη ευκλείδια απόσταση, επιτύχαμε τα ακόλουθα 5 αποτελέσματα:



Σχήμα: Αποτελέσματα υφής για το 935.bmp

Τα παραπάνω αποτελέσματα έχουν ταξινομηθεί σύμφωνα με τη ευκλείδια απόσταση. Αυτές οι αποστάσεις εμφανίζονται στον παρακάτω πίνακα:

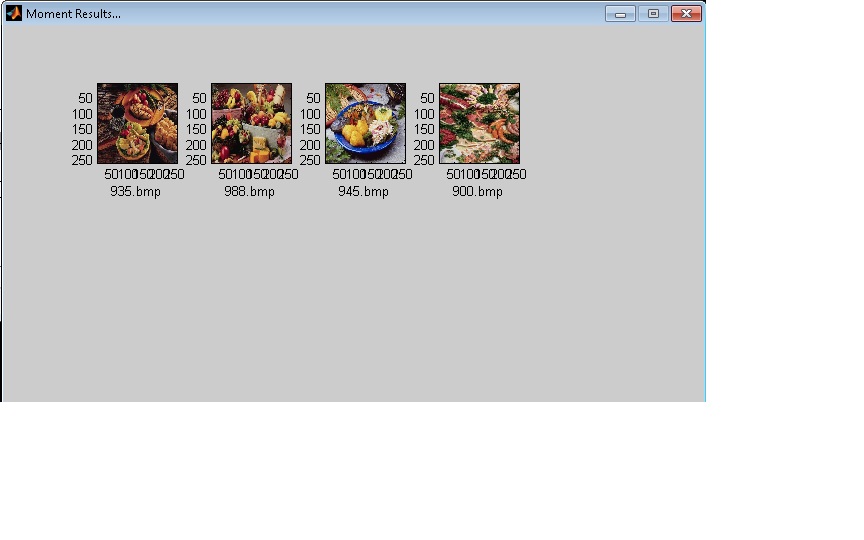


###### **Σχήμα: Ευκλείδια απόσταση μεταξύ query και αποτελεσμάτων…**

Παρατηρώντας των εικόνων στη βάση δεδομένων μας, μπορούμε πραγματικά να πούμε ότι τα παραπάνω αποτελέσματα αντιπροσωπεύουν τις πιό στενές αντιστοιχίες με την επιλεγμένη εικόνα ερώτηση.

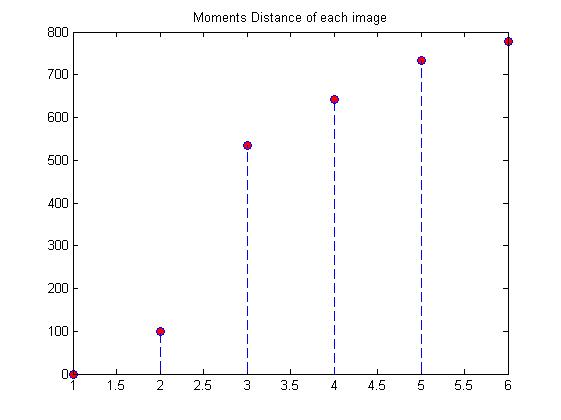
* 1. **Moments Extraction & Matching**

Χρησιμοποιώντας τις τις κεντρικές και αναλοίωτες ροπές με τον αλγόριθμο εξαγωγής που περιγράφεται στο προηγούμενο κεφάλαιο, όπου οι κεντρικές ροπές της εικόνας ερώτησης και των εικόνων που πήραμε σαν αποτέλεσμα από την εξαγωγή της υφής, συγκρίνονται χρησιμοποιώντας τη ευκλείδια απόσταση, και πετύχαμε τα ακόλουθα 4 αποτελέσματα:

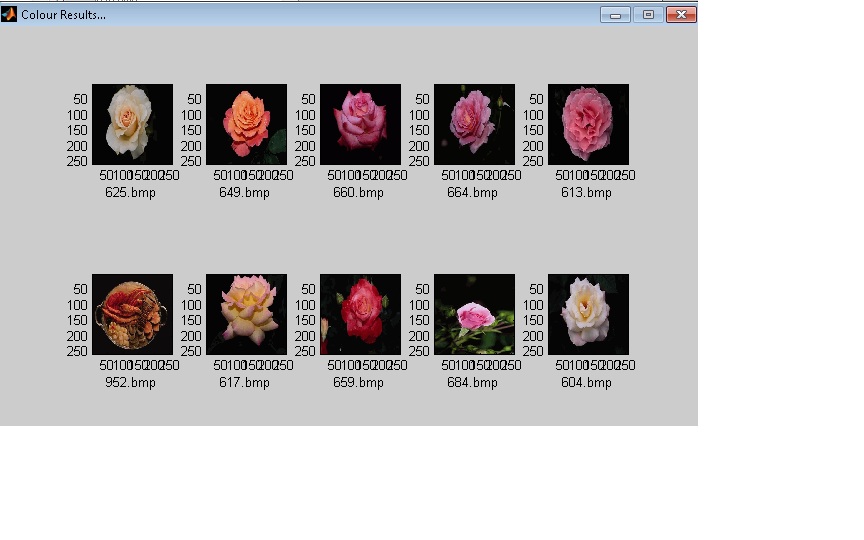


Σχήμα: Αποτελέσματα μετά την εφαρμογή των ροπών για το 935.bmp

Τα ανωτέρω αποτελέσματα ταξινομούνται σύμφωνα με τη ευκλείδια απόσταση όπως εμφανίζεται παρακάτω:



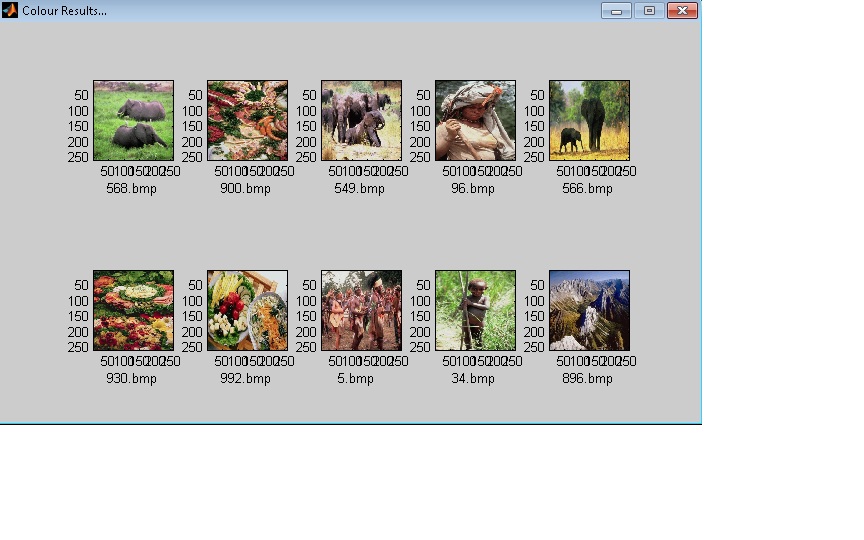
Στη συνέχεια παραθέτουμε μερικά άλλα πειραματικά αποτελέσματα για να επιδείξουμε πώς η προσέγγισή μας αποδίδει στις διάφορες εικόνες. Η ακόλουθη εικόνα ερώτησης ενός λουλουδιού απαντήθηκε αρκετά καλά, αφού δίνει 9 από τις 10 παρόμοιες εικόνες λουλουδιών αντίστοιχα.



Αντιθέτως αν δοκιμάσουμε την αναζήτηση με μια διαφορετική εικόνα λουλουδιών με ένα όχι και τόσο καθαρό υπόβαθρο ο αλγόριθμος έδωσε πιο φτωχά αποτελέσματα.



Η παρακάτω εικόνας δίνει μερικές αντιστοιχίες, αλλά ακόμα ο αλγόριθμος χρώματος που είναι το πρώτο βήμα στην προσέγγισή μας δεν αποδίδει καλά στο να μας δόσει περισσότερες παρόμοιες εικόνες στα κορυφαία δέκα αποτελέσματα, δεδομένου ότι η βάση δεδομένων αποτελείται από πολλές άλλες εικόνες με τα παρόμοια πρότυπα χρώματος από την αυτό της ερώτησης.



1. **Συμπεράσματα**

Η δραματική άνοδος στα μεγέθη των βάσεων δεδομένων εικόνων έχει δώσει ώθηση την ανάπτυξη αποτελεσματικών και αποδοτικών συστημάτων ανάκτησης. Η ανάπτυξη αυτών των συστημάτων άρχισε με την ανάκτηση των εικόνων χρησιμοποιώντας το κείμενο αλλά στη συνέχεια εισήχθεί η ανάκτηση εικόνας βασισμένη στο περιεχόμενο. Αυτό έγινε φνωστό ως CBIR ή ανάκτηση εικόνας με βάση το περιεχόμενο. Τα συστήματα που χρησιμοποιούν CBIR ανακτούν τις εικόνες βασισμένες στα οπτικά χαρακτηριστικά γνωρίσματα όπως το χρώμα, η υγή και η μορφή, σε αντιδιαστολή με ανάλογα συστήματα που χρησιμοποιούν τις περιγραφές εικόνας ή την κειμενική ευρετηρίαση. Σε αυτό το πρόγραμμα, έχουμε ερευνήσει τους διάφορους τρόπους φια την αναπαράσταση και εξαγωγή των ιδιοτήτων της εικόνας όπως του χρώματος, της υφής και της μορφής.

Η εφαρμογή εκτελεί μια απλή βασισμένη στο χρώμα αναζήτηση σε μια βάση δεδομένων μιας εικόνας ερώτησης εισόδου, χρησιμοποιώντας τα ιστόγραμμα χρώματος. Συγκρίνει έπειτα τα ιστόγραμμα χρώματος των διαφορετικών εικόνων χρησιμοποιώντας τη δευτέρου βαθμού εξίσωση απόστασης. Περαιτέρω ενισχύοντας την αναζήτηση, η εφαρμογή εκτελεί μια βασισμένη στην υφή αναζήτηση στα αποτελέσματα χρώματος, χρησιμοποιώντας την αποσύνθεση wavelets και τον υπολογισμό ενεργειακών επιπέδων. Συγκρίνει έπειτα τα χαρακτηριστικά γνωρίσματα σύστασης αποκτηθέντα χρησιμοποιώντας τη ευκλείδια εξίσωση απόστασης. Ένα πιο λεπτομερές βήμα ενίσχυει περαιτέρω αυτά τα αποτελέσματα υφής, χρησιμοποιώντας την αναζήτηση βασισμένη στη μορφή η στο σχήμα, με τις κεντρικές και αναλοίωτες ροπές.

Η τεχνική CBIR είναι ακόμα ένας αναπτυσσόμενος ερευνητικός τομέας. Δεδομένου ότι η συμπίεση εικόνας, η ψηφιακή επεξεργασία εικόνας, και οι τεχνικές εξαγωγής χαρακτηριστικών γνωρισμάτων εικόνας αναπτυσονται, η CBIR διατηρεί έναν σταθερό ρυθμό ανάπτυξης στο ερευνητικό πεδίο. Επιπλέον, η ανάπτυξη ισχυρής επεξεργαστικής δύναμης και οι γρηγορότερες και φτηνότερες μνήμες συμβάλλουν πολύ στην ανάπτυξη της CBIR. Αυτή η ανάπτυξη υπόσχεται μια απέραντη σειρά μελλοντικών εφαρμογών χρησιμοποιώντας CBIR Στο τέλος αυτής της παρουσίασης, θα επιθυμούσαμε να παρουσιάσουμε την άποψη μας σχετικά με μερικές τάσεις:

1. **Η κατευθυντήρια δύναμη**

Στην εργασία μας, οι περισσότερες από τις αναφορές περιοδικών είναι από τα τελευταία πέντε έτη. Γνωρίζουμε το γεγονός ότι ένα μεγάλο μέρος αυτού που έχουμε πει εδώ θα είναι ξεπερασμένο σύντομα. Η ώθηση πίσω από μία ικανοποιητηκή ανάκτηση εικόνας δίνεται από την ευρεία διαθεσιμότητα ψηφιακών αισθητήρων, του Διαδικτύου, και της μειωμένης τιμής των συσκευών αποθήκευσης. Λαμβάνοντας υπόψη το μέγεθος αυτών των κατευθυντήριων δυνάμεων, θεωρείται ότι η ικανοποίηση στην ανάκτηση θα συνεχίσει να αυξάνεται σε κάθε κατεύθυνση στα επόμενα έτη.

Αυτό που απαιτείται για το μέλλον είναι ακριβέστερα θεμέλια. Για μερικά από τα έγγραφα στις αναφορές, δεν ήταν σαφές ποιο πρόβλημα προσπαθούσαν να λύσουν ή εάν η προτεινόμενη μέθοδος θα εκτελείτο καλύτερα από μια εναλλακτική λύση. Μια ταξινόμηση των τύπων, των στόχων, και των σκοπών θα ήταν πολύ χρήσιμη εδώ, συμπεριλαμβανομένων των κριτηρίων για τη διάκριση μεταξύ των περιοχών εφαρμογής. Παρά τις δυσκολίες, είναι πεποίθησή μας ότι η ανάκτηση βασισμένη στην εικόνα στο τέλος δεν θα είναι μέρος του πεδίου της όρασης υπολογιστών. Η σχεδίαση man-machine interfaces, η γνώση του πεδίου εφαρμογής και οι τεχνολογίες βάσεων δεδομένων θα ασκήσουν επίδρασή τους στο προϊόν.

1. **Η κληρονομιά της όρασης υπολογιστών**

Ένα σημαντικό εμπόδιο που πρέπει να υπερνικηθεί πριν να μπορέσει να απογειωθεί η ανάκτηση εικόνας ήταν να συνειδητοποιήσουμε ότι η ανάκτηση εικόνας δεν συνεπάγεται το γενικό πρόβλημα κατανόησης εικόνας. Μπορεί να είναι ικανοποιητικό ότι ένα σύστημα ανάκτησης δίνει παρόμοιες εικόνες, παρόμοιες υπό κάποια καθορισμένη από το χρήστη έννοια. Η ισχυρή κατάτμηση της σκηνής και οι πλήρεις περιγραφές χαρακτηριστικών γνωρισμάτων μπορούν να μην είναι καν απαραίτητα για να επιτύχουμε την ταξινόμηση (ranking) της ομοιότητας. Φυσικά, όσο πιο βαθύτερα πηγαίνουμε στη σημασιολογία των εικόνων, τόσο βαθύτερη η κατανόηση της εικόνας θα πρέπει να είναι, αλλά αυτό μπορεί τελικά να είναι αποτέλεσμα της ταξινόμησης των εικόνων παρά μιας ακριβούς κατανόησης.

1. **Η επιρροή στην τεχνητή όραση**

Αντιθέτως, η ανάκτηση εικόνας βασισμένη στο περιεχόμενο προσφέρει ένα διαφορετικό βλέμμα στα παραδοσιακά προβλήματα όρασης υπολογιστών. Αρχικά, η ανάκτηση έχει φέρει μεγάλα σύνολα στοιχείων. Ο αριθμός εικόνων δοκιμής σε μία δημοσιευμένη εργασία ήταν κατά μέσο όρο κάτω από εκατό μέχρι πολύ πρόσφατα, μια πολύ καλή δημοσίευση ανάκτησης πειραματίζεται σε χιλιάδες εικόνες. Φυσικά, ο σκοπός είναι διαφορετικός για την όραση υπολογιστών και την ανάκτηση βασισμένη στο περιεχόμενο. Είναι πολύ ευκολότερο να συντεθεί ένα γενικό σύνολο στοιχείων αυθαίρετων εικόνων παρά να συλλεχθούν οι συγκεκριμένες που απαιτούνται σε μια εφαρμογή όρασης υπολογιστών, αλλά το στάδιο αυτό έχει τεθεί για περισσότερη ευρωστία. Άλλωστε, για την επεξεργασία χιλίων εικόνων απαιτείται ισχυρή υποδομή σε λογισμικό και επεξεργαστική ισχύ.

Δεύτερον, η ανάκτηση εικόνας βασισμένη στο περιεχόμενο έχει τρέξει στην απουσία μιας γενικής μεθόδου για την ισχυρή κατάτμηση. Τα συστήματα ανάκτησης έχουν εξετάσει τη δυσχέρεια κατάτμησης με μερικούς δημιουργικούς τρόπους.

Τρίτον, η ανάκτηση εικόνας βασισμένη στο περιεχόμενο έχει αναζωογονήσει το ενδιαφέρον για την επεξεργασία εικόνας. Αυτό οφείλεται στον προσδιορισμό των εντάσεων στον προσδιορισμό ενός αντικειμένου, καθώς επίσης και στη σημασία του χρώματος στην αντίληψη για τις εικόνες. Δεδομένου ότι η ανάκτηση με βάση το περιεχόμενο είναι προσανατολισμένη προς το χρήστη, το χρώμα δεν μπορεί να αφεθεί έξω. Ο σκοπός της περισσότερης επεξεργασίας του χρώματος της εικόνας είναι εδώ να μειωθεί η επιρροή των τυχαίων όρων της σκηνής και της αντίληψης (δηλ. το αισθητήριο χάσμα) με τον υπολογισμό των αμετάβλητων αντιπροσωπεύσεων αντίληψης και σκηνής.

Πρόοδος έχει σημειωθεί στην αντιπροσώπευση χρώματος για τις καλά περιγραμμένες κλάσεις των διάφορων όρων. Επίσης, η εφαρμογή των τοπικών γεωμετρικών περιγραφών που προέρχονται από τη θεωρία κλίμακας θα αποκαλύψει σύνολα άποψης και ανεξάρτητα εμφανή σημεία σκηνής, ανοίγοντας κατά συνέπεια το δρόμο στην ομοιότητα των εικόνων σε έναν μικρό αριθμό περισσότερων πληροφοριακών περιοχών ή σημείων. Τέλος, η προσοχή για τη σταθερότητα έχει αναζωογονηθεί επίσης με πολλά νέα χαρακτηριστικά γνωρίσματα και μέτρα ομοιότητας. Για την ανάκτηση βασισμένη στο περιεχόμενο, η σταθερότητα είναι μόνο μια πλευρά του νομίσματος, η διακριτική δύναμη είναι άλλη. Λίγη εργασία έχει αναφερθεί μέχρι τώρα για να καθιερώσει την υπόλοιπη διακριτική δύναμη των ιδιοτήτων. Αυτό είναι ουσιαστικό ως ισορροπία μεταξύ της σταθερότητας ενάντια στις παραλλαγές και η διατηρημένη διακριτική δύναμη καθορίζει την αποτελεσματικότητα μιας ιδιοκτησίας.

1. **Η ομοιότητα**

Η ομοιότητα είναι μια ερμηνεία της εικόνας βασισμένη στη διαφορά μεταξύ δύο στοιχείων ή ομάδων στοιχείων. Για κάθε έναν από τους τύπους χαρακτηριστικών γνωρισμάτων, ένα διαφορετικό μέτρο ομοιότητας απαιτείται. Για την ομοιότητα μεταξύ των συνόλων χαρακτηριστικών γνωρισμάτων, η ιδιαίτερη προσοχή έχει πάει στην καθιέρωση της ομοιότητας μεταξύ των ιστογράμμάτων λόγω τις υπολογιστικής αποδοτικότητας και την αποτελεσματικότητας ανάκτησής τους. Μέχρι τώρα η περισσότερη προσοχή έχει πάει στα ιστογράμματα χρώματος, αλλά αναμένεται ότι τα ιστόγραμμα τοπικών γεωμετρικών ιδιοτήτων και η υφή θα ακολουθήσουν.

1. **Η ανάγκη για τις βάσεις δεδομένων**

Όταν αυξάνονται τα data sets στο μέγεθος και όταν τα μεγαλύτερα σύνολα καθορίζουν πιό ενδιαφέροντα προβλήματα, και επιστημονικά και για το κοινό, οι υπολογιστικές πτυχές δεν μπορούν πλέον να αγνοηθούν. Η σύνδεση μεταξύ της ανάκτησης εικόνας βασισμένη στο περιεχόμενο και της έρευνας βάσεων δεδομένων είναι πιθανό να αυξηθεί στο μέλλον. Ήδη, οι πιό ελπιδοφόρες προσπάθειες είναι διεπιστημονικές, αλλά μέχρι τώρα, τα προβλήματα όπως ο καθορισμό των κατάλληλων γλωσσών διατύπωσης ερωτήσεων, η αποδοτική αναζήτηση στο υψηλό διαστατικό διάστημα χαρακτηριστικών, η αναζήτηση παρουσία των μεταβαλλόμενων μέτρων ομοιότητας είναι κατά ένα μεγάλο μέρος άλυτα.

**References**

[2] C. Carson, M. Thomas, S. Belongie, J.M. Hellerstein, and J. Malik,

ªBlobworld: A System for Region-Based Image Indexing and

Retrieval,º Proc. Visual Information Systems, pp. 509-516, June 1999

[3] James Z. Wang, Jia Li and Gio Wiederhold, ``SIMPLIcity: Semantics-Sensitive Integrated Matching for Picture Libraries,'' **IEEE Transactions on Pattern Analysis and Machine Intelligence,** vol. 23, no. 9, pp. 947-963, 2001.

[4] M. Flickner, H. Sawhney, W. Niblack, J. Ashley, Q. Huang, B. Dom

et al. ªQuery by Image and Video Content: The QBIC System,º

IEEE Computer, vol. 28, no. 9, 1995.

[5]. John P. Eakins and Margaret E. Graham, Content-based image retrieval, a report to the JISC technology applications programme**,** *Institute for image database research, University of Northumbria at Newcastle, U.K,* January 1999.

[6]. Hideyuki Tamura and Naokazu Yokoya. Image Database Systems: A Survey**,** *PatternRecognition*, 17(1):29–49, 1984

[7] Shengjiu Wang, “A Robust CBIR Approach Using Local Color Histograms,” Department of Computer Science, University of Alberta, Edmonton, Alberta, Canada, Tech. Rep. TR 01-13, October 2001, Found at:

<http://citeseer.nj.nec.com/wang01robust.html>

[8]. Zijun Yang and Jay Kuo, Survey on content-based analysis, indexing and retrieval techniques and status report of MPEG-7,*Tamkang journal of science and engineering, vol.2, No.3 pp.101-118, 1999*

[9]. T. Ojala and M. Pietikäinen, Texture Classfication*, Machine vision and Media signal Processing Unit*, University of Oulu, Finland

[10] Sharmin Siddique, “A Wavelet Based Technique for Analysis and Classification of Texture Images,” Carleton University, Ottawa, Canada, Proj. Rep. 70.593, April 2002.

[11] W.Y. Ma and B. Manjunath, ªNaTra: A Toolbox for Navigating

Large Image Databases,º Proc. IEEE Int'l Conf. Image Processing,

pp. 568-571, 1997.

[12] G. D. Magoulas, S. A. Karkanis, D. A. Karras and M. N. Vrahatis, “Comparison Study of Textural Descriptors for Training Neural Network Classifiers”, in Proceedings of the 3rd IEEE-IMACS World Multi-conference on Circuits, Systems, Communications and Computers, vol. 1, 6221-6226, Athens, Greece, July 1999, Found at:

<http://www.brunel.ac.uk/~csstgdm/622.pdf>

[18] Y. Rubner, L.J. Guibas, and C. Tomasi, ªThe Earth Mover's

Distance, Multi-Dimensional Scaling, and Color-Based Image

Retrieval,º Proc. DARPA Image Understanding Workshop, pp. 661-

668, May 1997.

[20] J. Shi and J. Malik, ªNormalized Cuts and Image Segmentation,º

Proc. Computer Vision and Pattern Recognition, pp. 731-737, June

1997.

[21] J.R. Smith and S.-F. Chang, ªVisualSEEk: A Fully Automated

Content-Based Image Query System,º Proc. ACM Multimedia,

pp. 87-98, Nov. 1996.

[22] J.R. Smith and C.S. Li, ªImage Classification and Querying Using

Composite Region Templates,º Int'l J. Computer Vision and Image

Understanding, vol. 75, nos. 1-2, pp. 165-174, 1999.

[23] Dengsheng Zhang, and Guojun Lu, Content-based shape retrieval using different shape descriptors *Grippsland school of computing and information technology, Monash university*, Churchill, Victoria, Australia.

[24]. Rafael C. Gonzales and Richaerds E. Woods, *Digital image processing* second edition, 2002

[25]. Jean-Marc Pelletier, Computer vision documentation,www.iamas.ac.jp/~jovan02/cv/jit\_cv\_doc.pdf

[27] J.Z. Wang, J. Li, R.M. Gray, and G. Wiederhold, ªUnsupervised

Multiresolution Segmentation for Images with Low Depth of

Field,º IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 23,

no. 1, pp. 85-91, Jan. 2001.

[28] J.Z. Wang, G. Wiederhold, O. Firschein, and X.W. Sha, ªContent-

Based Image Indexing and Searching Using Daubechies' Wavelets,

º Int'l J. Digital Libraries, vol. 1, no. 4, pp. 311-328, 1998.

[29] G. Ciocca and R. Schettini, ªUsing a Relevance Feedback

Mechanism to Improve Content-Based Image Retrieval,º Proc.

Visual '99: Information and Information Systems, pp. 107-114, 1999.

[88] T. Kato, T. Kurita, N. Otsu, and K. Hirata, ªA Sketch Retrieval

Method for Full Color Image DatabaseÐQuery by Visual

Example,º Proc. ICPR, Computer Vision and Applications, pp. 530-

533, 1992.

[68] A. Hiroike, Y. Musha, A. Sugimoto, and Y. Mori, ªVisualization of

Information Spaces to Retrieve and Browse Image Data,º Proc.

Visual '99: Information and Information Systems, pp. 155-162, 1999.

[168] M.J. Swain, ªSearching for Multimedia on the World Wide Web,º

Proc. Int'l Conf. Multimedia Computing and Systems, pp. 33-37, 1999.

[31] I.J. Cox, M.L. Miller, T.P. Minka, and T.V. Papathomas, ªThe

Bayesian Image Retrieval System, PicHunter: Theory, Implementation,

and Pychophysical Experiments,º IEEE Trans. Image

Processing, vol. 9, no. 1, pp. 20-37, 2000.

[48] M. Flickner, H. Sawhney, W. Niblack, J. Ashley, Q. Huang, B. Dom,

M. Gorkani, J. Hafner, D. Lee, D. Petkovic, D. Steele, and P. Yanker,

ªQuery by Image and Video Content: The QBIC System,º IEEE

Computer, 1995.

[170] D.J. Swets and J. Weng, ªHierarchical Discriminant Analysis for

Image Retrieval,º IEEE Trans. Pattern Analysis and Machine

Intelligence, vol. 21, no. 5, pp. 386-401, May 1999.

[79] A.K. Jain and A. Vailaya, ªShape-Based Retrieval: A Case Study

with Trademark Image Databases,º Pattern Recognition, vol. 31,

no. 9, pp. 1,369-1,390, 1998.

[74] B. Huet and E.R. Hancock, ªLine Pattern Retrieval Using

Relational Histograms,º IEEE Trans. Pattern Analysis and Machine

Intelligence, vol. 21, no. 12, pp. 1,363-1,371, Dec. 1999.

**Περιεχόμενα**

Περίληψη…………………………………………………………..σελ. 3

Εισαγωγή…………………………………………………………..σελ. 4

Υπάρχοντα Συστήματα…………………………………………..σελ. 5

Εφαρμογές Content Based Image Retrieval…………………...σελ. 9

Συναφείς εργασίες σε Content Based Image Retrieval……σελ. 10

Οργάνωση μια συλλογής εικόνων……………………………..σελ. 12

Προτεινόμενη λύση…………………...……………………...σελ. 13

GUI………………………………………………………………..σελ. 30

Αποτελέσματα-Παραδείγματα………………………………….σελ. 32

Συμπεράσματα…………………………………………………..σελ. 39

References……………………………………………………….σελ. 42

Περιεχόμενα……………………………………………………σελ. 45