



## ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

ΣΧΟΛΗ ΤΕΧΝΟΛΟΓΙΩΝ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΕΠΙΚΟΙΝΩΝΙΩΝ  
ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ

ΠΜΣ ΠΛΗΡΟΦΟΡΙΚΗ  
ΜΑΘΗΜΑ ΠΟΛΥΜΕΣΙΚΑ ΣΗΜΑΤΑ ΚΑΙ ΣΥΣΤΗΜΑΤΑ

## ΑΠΑΛΛΑΚΤΙΚΗ ΕΡΓΑΣΙΑ

Φοιτητής: Κωνσταντίνος Β. Ζιώβας

Υπεύθυνοι Καθηγητές: Καθηγητής: Χ. Δουληγέρης

Καθηγητής: Γ. Τσιχριντζής

ΑΘΗΝΑ

ΦΕΒΡΟΥΑΡΙΟΣ 2019

## Θέμα 1. Ανάκτηση εικόνας με βάση το περιεχόμενο (5/10)

### **1.1 Εισαγωγή**

Η εποχή μας χαρακτηρίζεται από έναν καταιγισμό πληροφορίας, που δεν έχει προηγούμενο στην ανθρώπινη ιστορία. Τεράστιος όγκος πληροφοριών παράγεται, αποθηκεύεται και διακινείται εξαιτίας της αλματώδης προόδου που έχει επιτευχθεί στους τομείς της πληροφορικής και των τηλεπικοινωνιών που εκφράστηκε κυρίως με την επανάσταση που προκάλεσε το διαδίκτυο με τον Παγκόσμιο Ιστό. Μέσα στο περιβάλλον αυτό της ανεπανάληπτης διαθεσιμότητας της πληροφορίας, το ζήτημα της πρόσβασης στην πληροφορία αποτελεί προτεραιότητα και αντικείμενο μεγάλης και εντεινόμενης ερευνητικής δραστηριότητας. Αντικειμενικός στόχος αποτελεί η αποτελεσματικότερη ανάκτηση της πληροφορίας από τον ενδιαφερόμενο σε όποια μορφή κι αν βρίσκεται.

Ένα πολύ σημαντικό φορέα πληροφορίας και μέσο επικοινωνίας για τον άνθρωπο αποτελούν οι εικόνες. Η ανάπτυξη των εικόνων, αναφορικά με τον αριθμό, τη διαθεσιμότητα και τη σημασία τους για τον άνθρωπο, υπήρξε αλματώδης. Ενδεικτικό της τεράστιας αύξησης των εικόνων είναι το γεγονός ότι κατά μέσο όρο κάθε ιστοσελίδα στον Παγκόσμιο Ιστό, που διαρκώς γιγαντώνεται, έχει υπολογιστεί ότι περιέχει 14,38 εικόνες. Η εκρηκτική ανάπτυξη των συλλογών με εικόνες, που υπήρξε αποτέλεσμα της προόδου της τεχνολογίας στον τομέα της δημιουργίας και αποθήκευσης της ψηφιακής εικόνας, έκανε επιτακτική την ανάγκη εύρεσης και εξέλιξης νέων τεχνικών για την διαχείριση και αποτελεσματική ανάκτηση των εικόνων. Μια από τις τεχνικές αυτές που στοχεύουν στην καλύτερη διαχείριση των ψηφιακών εικόνων είναι η ανάκτηση εικόνας με βάση το περιεχόμενο.

Σκοπός της παρούσης εργασίας αποτελεί η παρουσίαση της ανάκτησης των εικόνων με βάση το περιεχόμενο και η σημασία αυτής της τεχνικής. Η εργασία αυτή εστιάζει κυρίως στις σταθερές - ακίνητες εικόνες. Παρουσιάζονται επίσης τα πλεονεκτήματα και τα μειονεκτήματα της τεχνικής αυτής σε σχέση με τις παραδοσιακές μεθόδους ανάκτησης της εικόνας με βάση τις λέξεις κλειδιά καθώς και οι τάσεις και οι προοπτικές που διακρίνονται σ' αυτό τον τομέα.

## 1.2 Ορισμοί: εικόνα - ψηφιακή εικόνα

Γενικά εικόνα ονομάζεται η αναπαράσταση της ομοιότητας κάποιου θέματος, φυσικού αντικειμένου ή προσώπου και δημιουργείται από μια οπτική συσκευή - όπως φωτογραφική μηχανή, καθρέφτης, φακός, τηλεσκόπιο, μικροσκόπιο - ή από φυσικά αντικείμενα και φαινόμενα – όπως το ανθρώπινο μάτι ή η επιφάνεια του νερού. Οι εικόνες μπορεί να είναι είτε δύο διαστάσεων (π.χ. μια φωτογραφία) είτε τριών διαστάσεων (π.χ. ένα άγαλμα). Η λέξη εικόνα χρησιμοποιείται ως ευρύτερη έννοια για να περιγράψει κάθε αποτύπωση δύο διαστάσεων όπως ένας χάρτης, ένα διάγραμμα, ένας αφηρημένος πίνακας ζωγραφικής. Με αυτή την ευρύτερη έννοια μπορεί να δημιουργείται με το χέρι (ζωγραφική, χαρακτική), με τη χρήση της τεχνολογίας των γραφικών με υπολογιστή ή και με το συνδυασμό των δύο προηγουμένων.

Ψηφιακή εικόνα είναι η αναπαράσταση μιας δισδιάστατης εικόνας ως ένα πεπερασμένο σύνολο ψηφιακών αξιών που ονομάζονται εικονοστοιχεία (pixel). Τα εικονοστοιχεία αποθηκεύονται ως μια δισδιάστατη διάταξη μικρών ακέραιων αριθμών. Συχνά μεταδίδονται ή αποθηκεύονται σε μια συμπιεσμένη μορφή. Οι ψηφιακές εικόνες μπορούν να δημιουργηθούν από διάφορες συσκευές εισόδου ή τεχνικές όπως ψηφιακές φωτογραφικές μηχανές, σαρωτές, αερομεταφερόμενα ραντάρ, σεισμογράφους. Μπορεί να είναι είτε πρωτογενώς ψηφιακές δηλαδή να δημιουργούνται απευθείας σε ψηφιακή μορφή είτε δευτερογενώς ψηφιακές δηλαδή να είναι αποτέλεσμα ψηφιοποίησης ήδη υπαρχόντων εικόνων ή φωτογραφιών σε συμβατική μορφή. Επίσης μπορούν να δημιουργηθούν από μη εικονικά δεδομένα ή τρισδιάστατα γεωμετρικά μοντέλα· το τελευταίο αποτελεί ένα κύριο υποπεδίο των γραφικών με υπολογιστή. Οι ψηφιακές εικόνες αποθηκεύονται σε διάφορους μορφότυπους (format) όπως οι εξής: gif, jpeg, png, svg.

## 1.3 Σημασία της εικόνας

Η εικόνα αποτέλεσε μέσο επικοινωνίας για τον άνθρωπο ήδη από τα προϊστορικά χρόνια όπως αποδεικνύεται από τα ευρήματα τοιχογραφιών σε σπήλαια. Η χρήση χαρτών και σχεδίων κτιρίων για την παροχή πληροφόρησης έχει ξεκινήσει από την αρχαιότητα. Σήμερα, περισσότερο από ποτέ, η εικόνα είναι ένας σημαντικός φορέας πληροφορίας. Αποτελεί βασικό μέσο επικοινωνίας για την κοινωνία μας. Η εικόνα

μπορεί να αποδώσει με μεγάλη εκφραστικότητα νοήματα και ιδέες ξεπερνώντας γλωσσικούς περιορισμούς. Μάλιστα πολλοί τομείς της ανθρώπινης δραστηριότητας, όπως η δημοσιογραφία, οι εκδόσεις, οι διαφημίσεις, βασίζονται στην εικόνα και τις δυνατότητες που παρέχει. Η ανεπανάληπτη ανάπτυξη όσον αφορά τον αριθμό, τη διαθεσιμότητα και τη σημασία της εικόνας για την ανθρώπινη δραστηριότητα επήλθε κατά τη διάρκεια του τελευταίου αιώνα. Η τεχνολογική πρόοδος, που εκφράστηκε με εφευρέσεις και πρόοδο σε τομείς όπως η φωτογραφία και η τηλεόραση, συντέλεσε σε μεγάλο βαθμό στη διευκόλυνση της αποτύπωσης, διάδοσης και επικοινωνίας των δεδομένων της εικόνας. Τη μεγαλύτερη ώθηση όμως στην αλματώδη αυτή ανάπτυξη έδωσε ο υπολογιστής, που πρόσφερε δυνατότητες και τεχνικές για ψηφιακή αποτύπωση, επεξεργασία, αποθήκευση και μετάδοση της εικόνας. Η πρόοδος στην επιστήμη της πληροφορικής καθώς επίσης και η δημιουργία του Παγκόσμιου Ιστού (World Wide Web) με τις δυνατότητες που παρέχει στην πρόσβαση πληροφορίας σε ποικιλία μέσων από οπουδήποτε έδωσε ένα μαζικό ερέθισμα για την εκμετάλλευση της ψηφιακής εικόνας.

#### 1.4 Ανάκτηση εικόνας με βάση το κείμενο

Οι βιβλιοθήκες φωτογραφιών, τα μουσεία και οι εκθέσεις έργων τέχνης έδειξαν ιδιαίτερο ενδιαφέρον στην ψηφιοποίηση των συλλογών τους. Πρέπει όμως να τονιστεί ότι η διαδικασία της ψηφιοποίησης εικόνων από μόνη της δεν κάνει ευκολότερη τη διαχείριση των συλλογών. 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 φυλλομέτρηση (browsing), εξακολουθεί να είναι απαραίτητος κάποιος τρόπος για τη διαχείριση και ανάκτηση εικόνων από μεγάλες ψηφιακές βιβλιοθήκες εικόνων. Πιο αποτελεσματικές τεχνικές από την απλή φυλλομέτρηση είναι αναγκαίες για συλλογές που περιέχουν μεγάλο αριθμό εικόνων - χιλιάδες ή και πολλές δεκάδες χιλιάδες μερικές φορές. Για την οργάνωση των συλλογών των εικόνων χρησιμοποιήθηκαν λέξεις που αποδίδονταν στις εικόνες με τη μορφή μεταδεδομένων σύμφωνα με ορισμένα πρότυπα (π.χ. MARC, Dublin Core) 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 (text-based image retrieval ή metadata based image retrieval) και αποτελεί την παραδοσιακή τεχνική που χρησιμοποιήθηκε και παλιότερα στις αναλογικές συλλογές εικόνων.

Αναφορικά με την απόδοση λέξεων στις εικόνες, δηλαδή την καταλογογράφησή 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 - θησαυρούς, θεματικές επικεφαλίδες, ταξινομικά συστήματα - που ελέγχουν το βασικό λεξιλόγιο μιας γλώσσας θεματικής ευρετηρίασης και αποδίδουν σημασιολογικές σχέσεις στους όρους ή και ελεύθερο κείμενο με λεπτομερείς περιγραφές.

#### **1.4.1 Συστήματα ευρετηρίασης και ταξινόμησης εικόνων**

Παρουσιάζονται παρακάτω συνοπτικά ορισμένα συστήματα ταξινόμησης - ευρετηρίασης για εικόνες:

Ίσως το πιο γνωστό σύστημα ευρετηρίασης είναι ο Θησαυρός των Τεχνών και της Αρχιτεκτονικής (Art and Architecture Thesaurus) που δημιουργήθηκε στο Πολυτεχνείο του Rensselaer (Rensselaer Polytechnic Institute) το 1979 και σήμερα χρησιμοποιείται από βιβλιοθήκες τέχνης σε όλο τον κόσμο. Ο Θησαυρός υποστηρίζεται από το Ινστιτούτο Πληροφόρησης Getty (Getty Information Institute) 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 (Union List of Artist Names) και ο Θησαυρός Γεωγραφικών Ονομάτων του Getty (Getty Thesaurus of Geographic Names ή TGN).

Ένα άλλο σύστημα που αναπτύχθηκε από τις Εικόνες Getty (Getty Images) 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. Έχει αναπτυχθεί λογισμικό που επιτρέπει στους χρήστες να κάνουν αναζήτηση σε διάφορα επίπεδα, από το ευρύ (π.χ. ελευθερία) ως το συγκεκριμένο (π.χ. ένα παιδί που κάνει κούνια).

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 (Library of Congress Thesaurus for Graphic Materials) που προήλθε από τις Θεματικές Επικεφαλίδες της Βιβλιοθήκης του Κογκρέσου (Library of Congress Subject Headings). Ακόμη το 1983 δημοσιεύτηκε ο Εικονογραφικός Θησαυρός (Thesaurus Iconographique: System Descriptif de Representations) 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.

Επιπλέον, έχει αναπτυχθεί ένας αριθμός από λιγότερο γνωστά συστήματα για την ταξινόμηση εικόνων για εξειδικευμένη χρήση. Παραδείγματα τέτοιων συστημάτων είναι η ταξινόμηση της Βιέννης για εικόνες εμπορικών σημάτων που δημιουργήθηκε το 1998 από τον Παγκόσμιο Οργανισμό Πνευματικής Ιδιοκτησίας (World Intellectual Property Organization) για να καταστήσει δυνατή την αναγνώριση εμπορικών σημάτων παγκοσμίως και το σύστημα κωδικοποίησης για μηχανικά μέρη του Opitz (1969) που χρησιμεύει στον εντοπισμό ομάδων από παρόμοια μέρη που μπορούν να τοποθετηθούν μαζί σε μηχανές.

Τελειώνοντας την παράθεση των συστημάτων αυτών, πρέπει να επισημανθεί ότι παρά την ύπαρξη ειδικών συστημάτων για τις εικόνες, γενικά ταξινομικά συστήματα όπως η Δεκαδική Ταξινόμηση του Dewey (Dewey Decimal Classification), η ταξινόμηση της Βιβλιοθήκης του Κογκρέσου (Library of Congress Classification) και η Παγκόσμια Δεκαδική Ταξινόμηση (Universal Decimal Classification) χρησιμοποιούνται σε ευρεία κλίμακα σε βιβλιοθήκες εικόνων και βίντεο.

#### **1.4.2 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 (ή/ και σύμβολα στην περίπτωση των ταξινομικών συστημάτων) τα χαρακτηριστικά τους ώστε να είναι δυνατή η ανάκτησή τους από το χρήστη της ψηφιακής βιβλιοθήκης.

Οι τεχνικές ευρετηρίασης έχουν μεγάλη εκφραστική δύναμη, 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. Υπάρχει άλλωστε διαθέσιμο ένα ευρύ φάσμα λογισμικού για την ανάκτηση με βάση το κείμενο που αυτοματοποιεί τις διαδικασίες της αναζήτησης και ανάκτησης.

Όμως η διαδικασία της ευρετηρίασης από τον άνθρωπο, είτε με λέξεις κλειδιά γίνεται είτε με ταξινομικά σύμβολα, έχει ορισμένα βασικά μειονεκτήματα. Κατά πρώτο λόγο είναι κοπιαστική εργασία και απαιτεί μεγάλη κατανάλωση πόρων σε χρόνο και χρήμα. Έχει υπολογιστεί ότι ο χρόνος ευρετηρίασης ποικίλει από περίπου 7 λεπτά ανά εικόνα στο Getty Images 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 (με άλλο τρόπο και σε άλλο βαθμό κατανόησης) από κάθε άνθρωπο. Η οπτική του ευρετηριαστή μπορεί να διαφέρει από αυτή του χρήστη. Άλλωστε οι λέξεις δε 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 ορθογραφημένες λέξεις (και κατά την ευρετηρίαση και κατά την αναζήτηση), 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.

Τα τελευταία χρόνια η συνεχής αύξηση των δυνατοτήτων των υπολογιστών, η δραστική μείωση του κόστους αποθήκευσης, η αλματώδης επέκταση του διαδικτύου και η ολοένα και μεγαλύτερη πρόσβαση σ' αυτό οδήγησε στην τεράστια αύξηση της ψηφιακής πρόσκτησης των πληροφοριών. Ένα μεγάλο μέρος αυτών των πληροφοριών αποτελούν οι εικόνες που η παραγωγή και η διάθεση τους είναι τεράστια και διαρκώς αυξανόμενη. Εξαιτίας της όξυνσης των προβλημάτων, που αναφέρθηκαν παραπάνω, στη διαχείριση και ανάκτηση των εικόνων - υποκειμενική αντίληψη του περιεχομένου της εικόνας, μεγάλη κατανάλωση πόρων, περιορισμοί στην ανάκτηση - σε συνδυασμό με την τεράστια παραγωγή εικόνων υπήρξε από τις αρχές της δεκαετίας του 1990 μια έντονη ερευνητική δραστηριότητα για την εύρεση

και εφαρμογή πιο αποτελεσματικών μεθόδων και τεχνικών ανάκτησης εικόνας από την παραδοσιακή, δηλαδή με βάση τις λέξεις κλειδιά. Η ερευνητική αυτή προσπάθεια έχει ως στόχο την ανάκτηση εικόνας με βάση το περιεχόμενο.

## 1.5 Ανάκτηση εικόνας με βάση το περιεχόμενο

Ανάκτηση εικόνας με βάση το περιεχόμενο (content-based image retrieval) ονομάζεται η τεχνική για την ανάκτηση εικόνας από μια ψηφιακή βιβλιοθήκη ή τον παγκόσμιο ιστό με τη χρήση χαρακτηριστικών στοιχείων (μαθηματικών μετρήσεων των χαρακτηριστικών) που εξάγονται αυτόματα με βάση το οπτικό της περιεχόμενο. Ονομάζεται επίσης και ανάκτηση οπτικής πληροφορίας με βάση το περιεχόμενο (content-based visual information retrieval) και ερώτημα στο περιεχόμενο της εικόνας (query by image content) και αποτελεί εφαρμογή του τομέα της όρασης των υπολογιστών (computer vision) 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 - γνωρισμάτων από την ίδια την εικόνα δηλαδή οι εικόνες ευρετηριάζονται αυτόματα (image indexing) 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 (π.χ. εξαγωγή χαρακτηριστικών σχετικά με το χρώμα για ένα σχήμα ή ένα τμήμα της εικόνας).

### **1.5.1 Χαρακτηριστικά στοιχεία που χρησιμοποιούνται για την ανάκτηση με βάση το περιεχόμενο**

Οι κυριότεροι τύποι χαρακτηριστικών στοιχείων – γνωρισμάτων που εξάγονται για την ανάκτηση της εικόνας περιγράφονται παρακάτω:

#### **1.5.1.1 Χρώμα**

Το χρώμα (colour) αποτελεί το πιο ευρέως χρησιμοποιούμενο οπτικό χαρακτηριστικό για την ανίχνευση και ανάκτηση εικόνας. Για την απεικόνιση των χρωμάτων χρησιμοποιούνται κάποια χρωματικά μοντέλα. Ένα χρωματικό μοντέλο (π.χ. RGB, CMY) είναι ένας καθορισμένος τρόπος να αναπαρασταθεί κάθε χρώμα με μαθηματικούς όρους. Τα περισσότερα μοντέλα χρησιμοποιούν ένα τρισδιάστατο σύστημα συντεταγμένων. Κάθε σημείο μέσα στον υποχώρο του συστήματος αναπαριστά και ένα μοναδικό χρώμα. Σε χρήση σήμερα υπάρχει ένας μεγάλος αριθμός από χρωματικά μοντέλα. Κάθε χρωματικό σύστημα έχει το δικό του σύνολο από χρωματικά μοντέλα που αποτελούν τις παραμέτρους του χρωματικού συστήματος. Για παράδειγμα στο RGB (Red, Green, Blue) σύστημα τα χρώματα εμφανίζονται με βάση τις πρωταρχικές φασματικές συνιστώσες του κόκκινου, πράσινου και μπλε (τα τρία αυτά χρώματα αποτελούν τις παραμέτρους του συστήματος αυτού).

Υπάρχουν πολλές μέθοδοι για την ανάκτηση εικόνας με βάση το χρώμα αλλά οι περισσότερες είναι παραλλαγές στην ίδια βασική ιδέα. Κάθε εικόνα που προστίθεται στη συλλογή αναλύεται ώστε να δημιουργηθεί ένα ιστόγραμμα χρώματος δηλαδή μια γραφική σύνοψη δεδομένων που αναπαριστά την κατανομή - την αναλογία των χρωμάτων μέσα στην εικόνα, δηλαδή σε τι ποσοστό υπάρχει σε κάθε εικονοστοιχείο κάθε αξία του χρωματικού μοντέλου που χρησιμοποιείται, στην περίπτωση του RGB τα χρώματα κόκκινο, πράσινο και μπλε. Το ιστόγραμμα χρώματος της κάθε εικόνας

αποθηκεύεται στη βάση δεδομένων. Κατά την αναζήτηση ο χρήστης έχει τη δυνατότητα είτε να δηλώσει την αναλογία των χρωμάτων που επιθυμεί (π.χ. 75% κόκκινο και 25% μπλε) είτε να δώσει μια εικόνα στο σύστημα για την οποία δημιουργείται ιστόγραμμα είτε να επιλέξει κάποια εικόνα από αυτές που παρέχει το σύστημα και για την οποία υπάρχει ήδη αποθηκευμένο το ιστόγραμμα. Μετά μετριέται ο βαθμός ομοιότητας του ιστογράμματος του ερωτήματος με τα ιστογράμματα των εικόνων της βάσης και γίνεται η ανάκτηση αυτών των εικόνων που τα ιστογράμματα τους μοιάζουν περισσότερο με το ερώτημα, που έχουν δηλαδή σε μεγαλύτερο βαθμό ίδια χρώματα.

### **1.5.1.2 Σχήμα**

Το σχήμα (shape) αποτελεί ένα άλλο βασικό γνώρισμα για την ανάκτηση εικόνας με βάση το περιεχόμενο. Τα φυσικά αντικείμενα αναγνωρίζονται πέρα από το χρώμα τους κι 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 (boundary-based) και σ' αυτές που βασίζονται στην περιοχή (region-based). Οι πρώτες χρησιμοποιούν μόνο τα εξωτερικά όρια του σχήματος, το περίγραμμά του ενώ οι δεύτερες χρησιμοποιούν τη συνολική περιοχή του σχήματος. Οι ερώτησεις 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.

### **1.5.1.3 Υφή ή δομή**

Η υφή ή δομή ή σύσταση (texture) της εικόνας είναι ένα ακόμη σημαντικό χαρακτηριστικό. Η έννοια της δομής της εικόνας αναφέρεται στην παρουσία χωρικών οπτικών προτύπων που έχουν ιδιότητες ομοιογένειας. Η ομοιογένεια δεν προκύπτει από την παρουσία ενός μόνο χρώματος σε μια περιοχή. Η υφή είναι μια

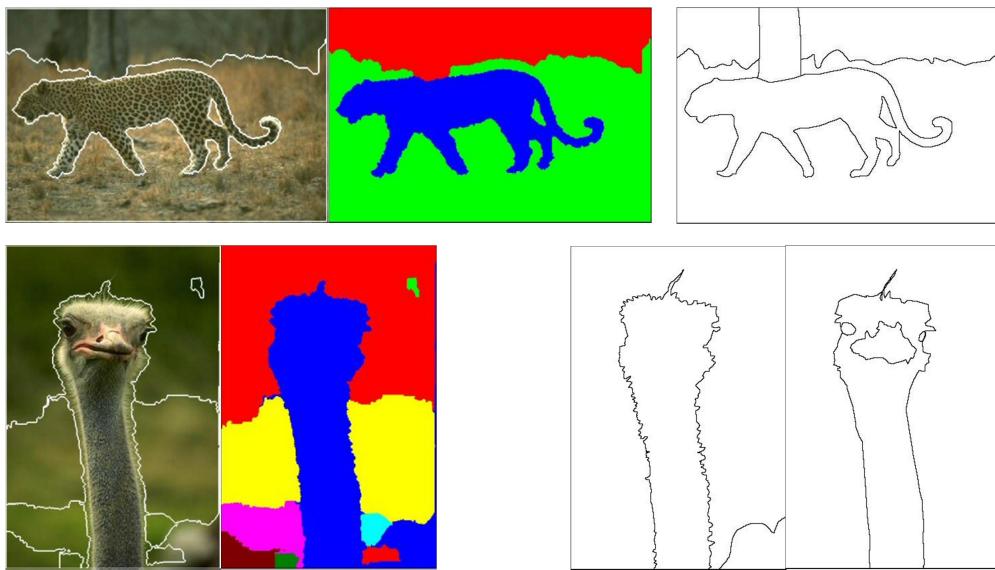
έμφυτη ιδιότητα των επιφανειών που συμπεριλαμβάνει μεταξύ άλλων για παράδειγμα τα σύννεφα, τα δέντρα, το ξύλο, τα υφάσματα, τα μαλλιά. Περιέχει σημαντική πληροφορία για τη δομική διάταξη των επιφανειών και τις σχέσεις τους με το περιβάλλον στο οποίο εντάσσονται. Η δομή της εικόνας που δίνεται στο ερώτημα υπολογίζεται και ανακτώνται οι εικόνες των οποίων η δομή ταιριάζει περισσότερο με αυτή του ερωτήματος. Για να γίνει κατανοητή η έννοια της υφής παρατίθεται το παρακάτω παράδειγμα. Ας υποθέσουμε ότι αναζητούμε εικόνες που να περιέχουν μια ζέβρα. Το ερώτημά μας μπορεί να αποτελείται από ένα σχεδιάγραμμα με το σχήμα του ζώου και τα χρώματα μαύρο και άσπρο. Σ' αυτή την περίπτωση το αποτέλεσμα μπορεί να είναι εικόνες που περιέχουν αγελάδες, όσο παράξενο κι αν φαίνεται γιατί τα σχήματα των δυο ζώων είναι πανομοιότυπα και τα χρώματα τους είναι και στις δύο περιπτώσεις το άσπρο και το μαύρο. Στην περίπτωση της αγελάδας το πρότυπο (υφή) είναι οι μαύρες κηλίδες ενώ στην περίπτωση της ζέβρας οι ρίγες.

#### **1.5.1.4 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 (location ή position) π.χ. το αντικείμενο 1 βρίσκεται στο δεξί πάνω μέρος της εικόνας καθώς και η χωρική του σχέση (spatial relation) με κάποιο ή κάποια άλλα αντικείμενα της εικόνας π.χ. το αντικείμενο 1 είναι τοποθετημένο πάνω κι αριστερά ή δίπλα ή εφάπτεται στο αντικείμενο 2. Εικόνες που έχουν όμοια αντικείμενα στην ίδια θέση ή συνδυασμό δύο ή περισσοτέρων όμοιων αντικειμένων με τη διάταξη (π.χ. το ένα δεξιά από το άλλο) που ανταποκρίνονται στο ερώτημα ανακτώνται.

#### **1.5.1.5 Κατάτμηση**

Η κατάτμηση (segmentation) της εικόνας αποτελεί μια σημαντική τεχνική για την ανάκτηση με βάση το περιεχόμενο. Η εικόνα χωρίζεται σε τμήματα ανάλογα με κάποιες παραμέτρους (π.χ. χρώματα, περιγράμματα) και στα τμήματα αυτά μπορούν να εφαρμοστούν κάποιες μέθοδοι εξαγωγής χαρακτηριστικών. Η εικόνα μπορεί να χωρίζεται σε μέρη άσχετα με το περιεχόμενό της π.χ. σε 4 ή 8 ίσα μέρη.



Εικόνα 1: Παράδειγμα κατάτμησης εικόνων

Για την εξαγωγή κάθε χαρακτηριστικού χρησιμοποιούνται διάφορες τεχνικές που διαφέρουν ως προς τη μεθοδολογία και την εφαρμογή. Τα συστήματα ανάκτησης εικόνας με βάση το περιεχόμενο (content-based image retrieval systems), δηλαδή οι εφαρμογές λογισμικού που υλοποιούν τις τεχνικές αναζήτησης αυτές συνήθως κάνουν συνδυασμό δύο ή περισσότερων χαρακτηριστικών για την εξαγωγή γνωρισμάτων ώστε να προκύπτει αποτελεσματικότερη ανάκτηση. Για παράδειγμα κάνουν εξαγωγή γνωρισμάτων για το χρώμα σε σχέση με μια περιοχή ή χρησιμοποιώντας τα χαρακτηριστικά του σχήματος και της θέσης για το σύνολο της εικόνας.

### 1.5.2 Σημασιολογικό χάσμα

Όλα τα παραπάνω χαρακτηριστικά αφορούν τα χαμηλού επιπέδου ή πρωτόγονα (low level ή primitive) στην ανίχνευση των οποίων έχει πραγματοποιηθεί μεγάλη πρόοδος όμως η δυσκολία είναι μεγάλη αναφορικά με τα υψηλού επιπέδου ή σημασιολογικά (high-level ή semantic) χαρακτηριστικά (όπως τα αντικείμενα, συναισθήματα που σχετίζονται με την εικόνα, περιγραφή θέματος). Το σημασιολογικό κενό (χάσμα) είναι το διάστημα που υπάρχει από την εξαγωγή γνωρισμάτων από τα οπτικά δεδομένα και την ερμηνεία που δίνουν οι χρήστες γι'

αυτά. Οι αλγόριθμοι που χρησιμοποιούνται δε μπορούν να αναγνωρίσουν αυτόματα έννοιες από τα οπτικά χαρακτηριστικά της εικόνας. Η οπτική ομοιότητα της εικόνας με μια άλλη εικόνα δεν είναι και σημασιολογική ομοιότητα. Ο πρωταρχικός στόχος ενός πρακτικού συστήματος ανάκτησης με βάση το περιεχόμενο είναι ουσιαστικά η ανάκτηση εικόνων που ανταποκρίνονται σε κάποια έννοια ή θέμα σε αντικατάσταση των μεταδεδομένων. Είναι στην πραγματικότητα εξαιρετικά δύσκολο για το χρήστη να αποδώσει μια περιγραφή της εικόνας που αναζητάει με βάση τα γνωρίσματα χαμηλού επιπέδου.

## 1.6 Αυτόματος σχολιασμός εικόνων

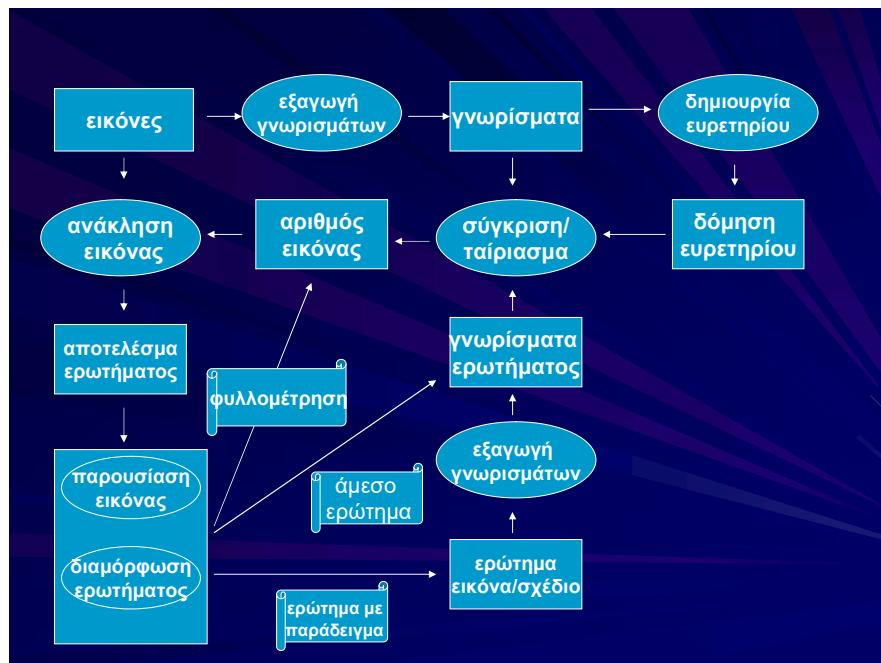
Τελευταία γίνονται προσπάθειες για τον αυτόματο σχολιασμό του περιεχομένου των εικόνων (μετατροπή του περιεχομένου της εικόνας σε γλωσσικούς όρους) δηλαδή αυτόματη εύρεση μεταδεδομένων ως λέξεις κλειδιά για την εικόνα από το περιεχόμενό της. Η τεχνική για την ανάκτηση εικόνας με βάση αυτούς τους όρους ονομάζεται ανάκτηση εικόνας με βάση το σχολιασμό (Annotation-Based Image Retrieval) και στοχεύει να ταξινομήσει τις εικόνες σε σημασιολογικές τάξεις για τη γεφύρωση του σημασιολογικού κενού που υπάρχει ανάμεσα στα χαμηλού επιπέδου χαρακτηριστικά και τις έννοιες εκμεταλλευόμενη τις δυνατότητες της φυσικής γλώσσας. Πραγματοποιείται ο αυτόματος σχολιασμός με την εκπαίδευση του συστήματος με ένα σύνολο αντιπροσωπευτικών εικόνων για κάθε έννοια στις οποίες έχουν προηγουμένως αποδοθεί λέξεις από τον άνθρωπο. Εντοπίζονται με αρκετά μεγάλη ακρίβεια κάποιες έννοιες όπως τοπίο, πόλη, ηλιοβασίλεμα, δάσος. Μπορεί να συνδυάσει τη δύναμη της φυσικής γλώσσας και το σημασιολογικό περιεχόμενο της εικόνας. Ένα τέτοιο πρόγραμμα είναι το ALIP (Automatic Linguistic Indexing of Pictures) που ξεκίνησε στο Πανεπιστήμιο του Στάνφορντ και συνεχίστηκε στο Πανεπιστήμιο Παρκ (The Pennsylvania State University, University Park). Προσδίδει αυτόματα όρους στις εικόνες κατατάσσοντας τις σε σημασιολογικές κατηγορίες. Για κάθε εικόνα που δίνεται το σύστημα συγκρίνει την εικόνα με τα μοντέλα στο εννοιολογικό λεξικό και εξάγει τους πιο σχετικούς όρους για να περιγράψει την εικόνα. Χρησιμοποιήθηκε ένα σύνολο λέξεων κλειδιών που δόθηκαν μαζί με κάποια σύνολο ομογενών εικόνων για την εκπαίδευση του συστήματος. Κατά μέσο όρο χρησιμοποιήθηκαν 3,6 λέξεις για την περιγραφή των εικόνων των 600 θεματικών κατηγοριών. Οι περιγραφές των εικόνων ποικίλουν από πολύ απλές ή χαμηλού

επιπέδου όπως λουλούδια ή μανιτάρια ως σύνθετες ή υψηλού επιπέδου όπως Αγγλία, τοπίο, μάχη, ιστορικό κτίριο.



Εικόνα 2: Παράδειγμα αυτόματου σχολιασμού εικόνων από το σύστημα ALIP

## 1.7 Συστήματα ανάκτησης εικόνας με βάση το περιεχόμενο



Εικόνα 3: Σχεδιάγραμμα λειτουργίας τυπικού συστήματος ανάκτησης εικόνας με βάση το περιεχόμενο

Έχουν αναπτυχθεί αρκετά συστήματα που υλοποιούν την ανάκτηση εικόνας με βάση το περιεχόμενό της. Τα περισσότερα συστήματα δημιουργήθηκαν για ερευνητικούς σκοπούς αλλά υπάρχουν και κάποια που είναι διαθέσιμα ως εμπορικά πακέτα. Σε γενικές γραμμές τα περισσότερα συστήματα ανάκτησης εικόνας με βάση το περιεχόμενο λειτουργούν με τον παρακάτω τρόπο: για τις εικόνες που βρίσκονται αποθηκευμένες στη βάση δεδομένων του συστήματος εξάγονται χαρακτηριστικά στοιχεία τα οποία αποθηκεύονται, τα χαρακτηριστικά στοιχεία για κάθε ερώτημα που τίθεται κατά την αναζήτηση με κάποιον από τους τρόπους που αναφέρθηκαν συγκρίνονται με τα χαρακτηριστικά στοιχεία των εικόνων που βρίσκονται στη βάση και ανακαλούνται οι εικόνες (π.χ. 5 ή 10) που μοιάζουν περισσότερο με το ερώτημα, δηλαδή που τα χαρακτηριστικά τους στοιχεία έχουν τη μικρότερη απόκλιση σε σχέση με το ερώτημα.

### 1.7.1 Διεπαφή χρήστη

Η διεπαφή του συστήματος με το χρήστη αποτελεί μια βασική παράμετρο της ανάκτησης. Ο χρήστης έχει τη δυνατότητα αναζήτησης με τους εξής τρόπους: α) με

φυλλομέτρηση εικόνων (πλοϊγηση στη συλλογή των εικόνων), β) με άμεσο ερώτημα, δηλαδή καθορίζοντας τα χαρακτηριστικά που επιθυμεί να έχει η εικόνα που αναζητά (π.χ. κόκκινο και άσπρο χρώμα), γ) με ερώτηση με παράδειγμα, δίνοντας δηλαδή κάποια εικόνα στο σύστημα με τη χρήση σαρωτή ή επιλέγοντας κάποια από τις εικόνες που παρουσιάζονται. Η εικόνα αυτή χρησιμεύει ως παράδειγμα για την ανίχνευση ομοιότητας με τις εικόνες που υπάρχουν αποθηκευμένες στο σύστημα και ανάκληση παρόμοιων, δ) με ένα σχέδιο ή ζωγραφία που χρησιμεύει ως υπόδειγμα για την ανίχνευση παρόμοιων εικόνων. Επιπλέον σε ορισμένα συστήματα παρέχεται η δυνατότητα στο χρήστη να καθορίσει τη βαρύτητα των χαρακτηριστικών που εξάγονται από την εικόνα, δηλαδή να δηλώσει ποια γνωρίσματα και σε ποιο βαθμό είναι σημαντικά. Η εύχρηστη και φιλική διεπαφή του συστήματος με το χρήστη (user's interface) είναι πολύ σημαντική για την αποτελεσματική ανάκτηση.

### **1.7.2 Ανατροφοδότηση σχετικότητας**

Ιδιαίτερη παράμετρος κατά την ανάκτηση αποτελεί η ανατροφοδότηση σχετικότητας (relevance feedback) από το χρήστη προς το σύστημα. Είναι τεχνική που έχει χρησιμοποιηθεί με επιτυχία και στην ανάκτηση πληροφορίας. Η ανατροφοδότηση σχετικότητας είναι μια τεχνική τροποποίησης του ερωτήματος που χρησιμοποιείται στην ανάκτηση εικόνας και αποτελεί προσπάθεια να γίνουν αντιληπτές οι συγκεκριμένες ανάγκες του χρήστη μέσα από τη διαδικασία ανατροφοδότησης του συστήματος και επαναπροσδιορισμού του ερωτήματος. Η τεχνική αυτή είναι πολύ σημαντική, όσον αφορά τη βελτίωση της ανάκτησης της εικόνας, εξαιτίας της έλλειψης σημασιολογικού χαρακτηρισμού των εικόνων και της ανθρώπινης υποκειμενικότητας της αντίληψης. Ο χρήστης επαναπροσδιορίζει το ερώτημά του ύστερα από την ανάκτηση των εικόνων επιλέγοντας ποια ή ποιες από τις εικόνες που ανακτήθηκαν είναι σχετικές με το ερώτημά του και ποιες όχι για να επαναληφθεί το ερώτημα τροποποιημένο. Με αυτό τον τρόπο ο χρήστης δε χρειάζεται να καθορίσει τη βαρύτητα των χαρακτηριστικών στοιχείων της εικόνας γιατί αυτή διαμορφώνεται από το σύστημα. Τα συστήματα που εφαρμόζουν την τεχνική αυτή προσπαθούν να μάθουν την οπτική ανταπόκριση που υπάρχει ανάμεσα στις υψηλού επιπέδου έννοιες που χρησιμοποιεί ο άνθρωπος και τα χαμηλού επιπέδου γνωρίσματα που εξάγονται από τις εικόνες. Με την προσέγγιση

της ανατροφοδότησης σχετικότητας επιτυγχάνονται καλύτερα αποτελέσματα κατά την ανάκτηση.

### 1.7.3 Παρουσίαση συστημάτων

Στη συνέχεια παρουσιάζονται κατά αλφαριθμητική σειρά ορισμένα από τα συστήματα ανάκτησης 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.

#### 1.7.3.1 ALISA

To 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 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](http://www.seas.gwu.edu/~pbock/ALISA_Summary.html).

#### 1.7.3.2 Blobworld

Το Blobworld δημιουργήθηκε από το τμήμα Πληροφορικής του Πανεπιστημίου Μπέρκλευ της Καλιφόρνια (University of California, Berkeley). Εξάγει χαρακτηριστικά υπολογίζοντας το χρώμα και την υφή, τη θέση και τα σχήματα των περιοχών (blobs) 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>.

#### 1.7.3.3 CANDID

Το CANDID (Comparison Algorithm for Navigating Digital Image Database) 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>.

#### 1.7.3.4 CHROMA

Το σύστημα CHROMA (Colour Hierarchical Representation Oriented Management Architecture) αναπτύχθηκε από το Πανεπιστήμιο του Σάντερλαντ (School of Computing, Engineering and Technology, University of Sunderland, UK) στη Βρετανία. Τα γνωρίσματα που εξάγει βασίζονται στο χρώμα. Ο χρήστης μπορεί να πλοηγηθεί σε μια ιεραρχία χρωμάτων και να επιλέξει ομάδες χρωμάτων για το ερώτημά του ή να κάνει ερώτημα με ένα σχέδιο ή εικόνα ως παράδειγμα. Τα αποτελέσματα που ανακτώνται παρουσιάζονται σε σειρά ανάλογα με το βαθμό ομοιότητας. Πληροφορίες για το σύστημα διατίθενται στον παγκόσμιο ιστό στο <http://osiris.sund.ac.uk/~cs0sla>.

### 1.7.3.5 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/>

### 1.7.3.6 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 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>.

### 1.7.3.7 *FIDS*

Το FIDS (Flexible Image Database System) 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/imagedatabase>.

### 1.7.3.8 *FIR*

Το σύστημα FIR (Formula Image Retrieval) 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](http://www.igd.fhg.de/igd-a7/projects/formula/formula_e.html).

### 1.7.3.9 *ImageRETRO*

Το ImageRETRO (Image RETrieval by Reduction and Overview) 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 φίλτραρισμα 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/>.

### 1.7.3.10 *ImageRover*

Το σύστημα αυτό 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/ljm/imagescape.html>.

### 1.7.3.11 *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 εξυπηρετητή του συστήματος και επιστρέφεται η αντίστοιχη κατατμημένη. Ο χρήστης τότε έχει τη δυνατότητα επιλογής μιας ή περισσοτέρων περιοχών της εικόνας που τον ενδιαφέρουν για την ανάκτηση. Περισσότερα για το iPure στο: <http://www.research.ibm.com/irl/projects/mediamining/index.html>.

### 1.7.3.12 *KIWI*

Το σύστημα KIWI (Key-points Indexing Web Interface) αναπτύχθηκε από την εταιρεία INSA στη Λυών 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/>.

#### **1.7.3.13 MARS**

Το MARS (Multimedia Analysis and Retrieval System) 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>.

#### **1.7.3.14 Metaseek**

Το σύστημα Metaseek 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>.

#### **1.7.3.15 Photobook**

To Photobook 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./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>.

#### **1.7.3.16 PicSOM**

To 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 ζεκινά 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/>.

#### **1.7.3.17 QBIC**

To QBIC (Query By Image Content) 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://www.qbic.almaden.ibm.com>.

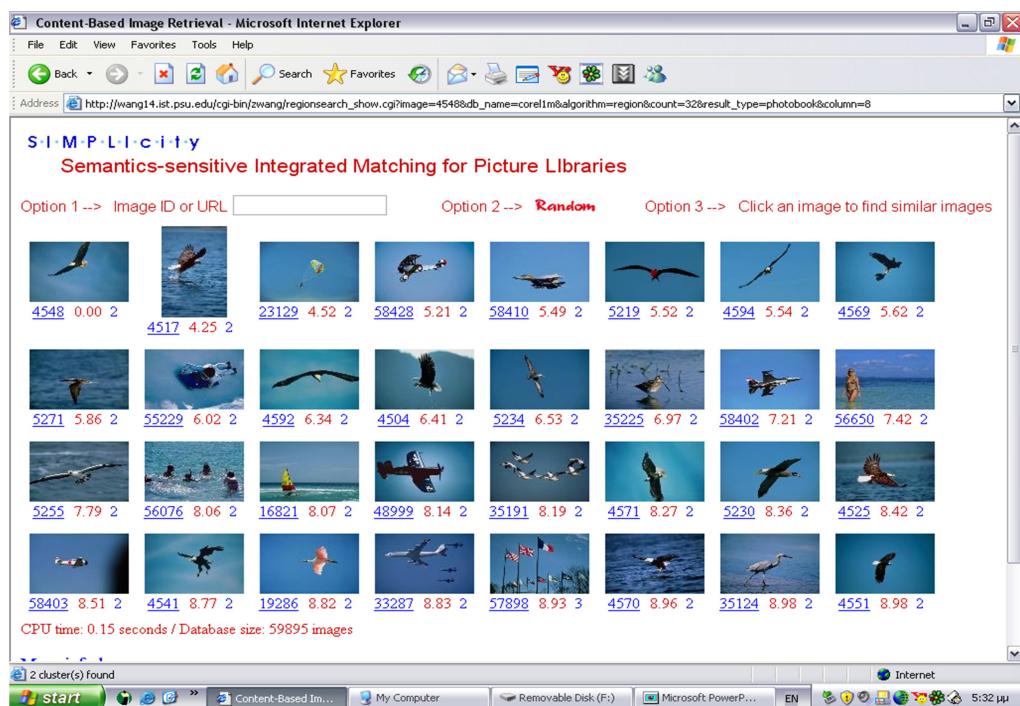
#### **1.7.3.18 SMURF**

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>.

#### **1.7.3.19 SIMPLIcity**

To SIMPLIcity (Semantic-sensitive Integrated Matching for Picture Libraries) αναπτύχθηκε από ερευνητική ομάδα στο Πανεπιστήμιο του Στάνφορντ (Stanford

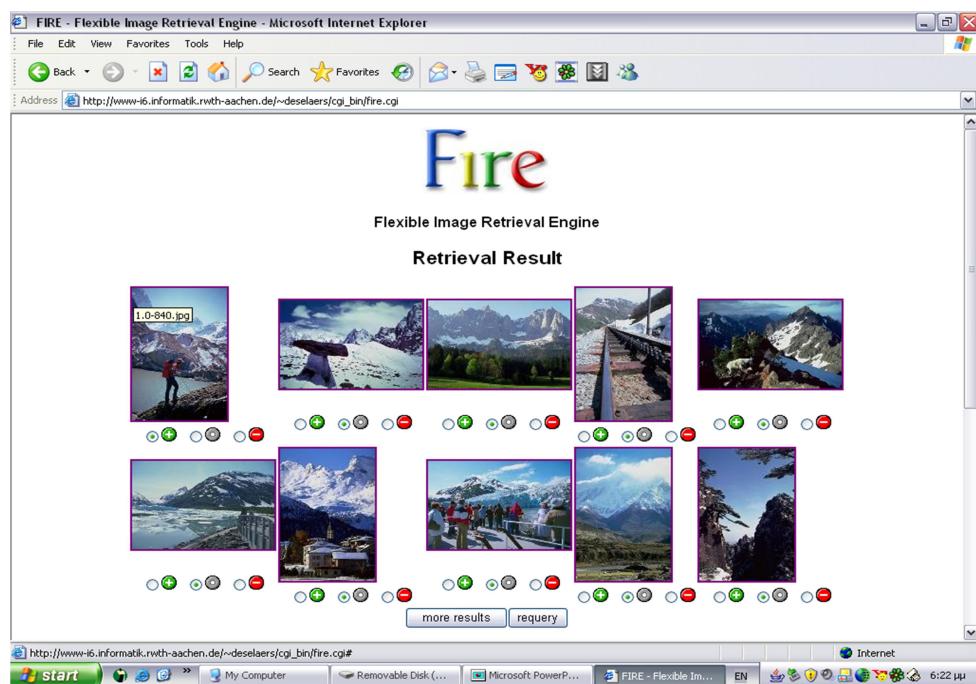
University). Η εικόνα στο σύστημα αυτό αναπαριστάται ως ένα σύνολο από περιοχές που αντιστοιχούν γενικά σε αντικείμενα και χαρακτηρίζονται από το χρώμα, τη δομή, το σχήμα και τη θέση τους. Ταυτίζοντας τα αντικείμενα στην εικόνα, το σύστημα μπορεί ευκολότερα να αναγνωρίσει παρόμοια αντικείμενα σε διαφορετικές 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 αναπτύξμες μεθόδους έρευνας. Ο χρήστης απλά κάνει κλικ σε μια εικόνα ή δηλώνει έναν αριθμό που έχει η εικόνα στη βάση δεδομένων ή ένα URL και ανακτά παρόμοιες φωτογραφίες. Περισσότερες πληροφορίες καθώς και παρουσίαση του συστήματος αυτού στο: <http://www-db.stanford.edu/IMAGE/>.



Εικόνα 4: Παράδειγμα ανάκτησης εικόνων από το σύστημα SIMPLIcity

### 1.7.3.20 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 εικόνωνέχει 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 5: Παράδειγμα ανάκτησης εικόνων από το σύστημα FIRE.

## 1.8 Συμπεράσματα - επίλογος

Η ανάκτηση εικόνας με βάση το περιεχόμενο αποτελεί το ιερό δισκοπότηρο για τον τομέα της ανάκτησης πληροφοριών, όπως εύστοχα χαρακτηρίστηκε, κατά πρώτο λόγο για να καταδείξει τη σπουδαιότητα της τεχνικής αυτής και κατά δεύτερο γιατί υπάρχει έντονη ερευνητική δραστηριότητα στον τομέα αυτό. Παρόλη την πρόοδο που έχει πραγματοποιηθεί στον τομέα αυτό η ανάκτηση εικόνας με βάση το

περιεχόμενο βρίσκεται ακόμη σε ανώριμο στάδιο και δε μπορεί να ανταποκριθεί επαρκώς στις ανθρώπινες ανάγκες. Αν και λειτουργεί με ικανοποιητικά αποτελέσματα αναφορικά με τα χαρακτηριστικά στοιχεία χαμηλού επιπέδου, υστερεί στα υψηλού επιπέδου ή σημασιολογικά χαρακτηριστικά που αποτελούν και τα σημαντικότερα για τον άνθρωπο. Εξαιτίας των αδυναμιών και τον περιορισμών της που αφορούν το σημασιολογικό κενό έχει προς το παρόν μικρή εφαρμογή. Ορισμένα συστήματα όπως το QBIC αποτελούν εμπορικές εφαρμογές είτε από μόνα τους είτε σε συνδυασμό με ανάκτηση εικόνας με βάση το κείμενο. Σε αρκετά από τα νέα συστήματα διαχείρισης ψηφιακών βιβλιοθηκών συμπεριλαμβάνεται και υποσύστημα για την ανάκτηση με βάση το περιεχόμενο. Μάλιστα ορισμένα συστήματα διατίθενται ως ελεύθερο λογισμικό (open source software) για τη χρήση και την παραπέρα ανάπτυξή τους από ψηφιακές βιβλιοθήκες (π.χ. FIRE). Είναι κατάλληλη για ορισμένες χρήσεις όπως για ιατρικές διαγνώσεις που βασίζονται στη σύγκριση ακτινογραφιών και εντοπισμό προσώπων για την αντιμετώπιση του εγκλήματος.

Περαιτέρω ανάπτυξη των τεχνικών αυτών και έρευνα για τη γεφύρωση του σημασιολογικού κενού θα οδηγούσε σε πολλές και εκτεταμένες πρακτικές εφαρμογές σε διάφορους τομείς όπως: η αντιμετώπιση του εγκλήματος, η άμυνα, η πνευματική ιδιοκτησία, η αρχιτεκτονική, ο σχεδιασμός συστημάτων, η μόδα, τα συστήματα γεωγραφικής πληροφόρησης, η δημοσιογραφία και η διαφήμιση, η ιατρική, η πολιτιστική κληρονομιά, η εκπαίδευση, η διασκέδαση, οι ψηφιακές βιβλιοθήκες φωτογραφιών και η αναζήτηση εικόνων στον Παγκόσμιο Ιστό. Η παραπέρα βελτίωση της ανάκτησης της εικόνας με βάση το περιεχόμενο θα μπορούσε να προσφέρει πολλά οφέλη στην κοινωνία μας. Ιδιαίτερα ο συνδυασμός της με την αναζήτηση με βάση το κείμενο μπορεί να αποδώσει καλύτερα αποτελέσματα όσον αφορά την ακρίβεια και την ανάκληση κατά την ανάκτηση εικόνας και οι μελλοντικές τάσεις δείχνουν συνδυασμό των δύο αυτών τεχνικών ανάκτησης για τη βελτίωση της ανάκτησης.

## Θέμα 2. Υπολογιστικές Ερωτήσεις (6/10)

### Ερώτημα 1

#### Υποερώτημα 1

Αρχικά πρέπει να βρούμε την μέγιστη συχνότητα του σήματος. Το σήμα αποτελείται από το άρθροισμα 2 συνημιτονοιδών σημάτων:

$$x(t) = 10 \cos(1000\pi t + \pi/3) + 20 \cos(2000\pi t + \pi/6)$$

Οι συχνότητες των δυο συνιστοσών είναι  $f_1=1000/2=500\text{ Hz}$  και  $f_2=2000/2=1000\text{ Hz}$ . Το θεώρημα Δειγματοληψίας Nyquist μας λέει ότι για να έχουμε πλήρη αναπαράσταση του σήματος η συχνότητα Δειγματοληψίας πρέπει να είναι διπλάσια της μέγιστης συχνότητας του σήματος. Η μέγιστη συχνότητα που μόλις υπολογίσαμε είναι  $f_2=1000\text{Hz}$  επομένως η συχνότητα Δειγματοληψίας είναι:  $f_s=2000\text{Hz}$ . Η μέγιστη τιμή του διαστήματος μεταξύ δύο δειγμάτων είναι επομένως  $t=1/f_s=0.0005\text{ sec}=0.5\text{ msec}$ .

#### Υποερώτημα 2

Με βάση τον παραπάνω υπολογισμό για μια ώρα σήματος  $1\text{ h} = 3600000\text{ ms}$  χρειαζόμαστε  $N=3600000/0.5=7200000$  δείγματα.

#### Υποερώτημα 3

Στη μετάδοση με 256 επίπεδα χρησιμοποιούμε 8-bit για να αποθηκευσουμε την τιμή κάθε δείγματος. Αν πολλαπλασιάσουμε τον αριθμό των δειγμάτων με το μέγεθος κάθε δείγματος σε bit έχουμε:  $S=N*8=7200000*8=57,600,000\text{ bit}=7.2\text{ MB}$ .

### Ερώτημα 2

Ένα σύστημα που βρίσκεται σε αρχική ηρεμία ονομάζεται γραμμικό, όταν και μόνον όταν δοθέντων δύο σημάτων εισόδου  $x_1(t)$  και  $x_2(t)$ , ισχύει η σχέση:  $S a_1 x_1 t + a_2 x_2 t = a_1 S[x_1 t] + a_2 S[x_2 t]$ . Δηλαδή, η απόκριση του συστήματος σε είσοδο που είναι ο γραμμικός συνδυασμός δύο σημάτων, ισούται με τον αντίστοιχο γραμμικό

συνδυασμό των αποκρίσεων του συστήματος σε καθένα από τα σήματα αυτά. Με βάση αυτόν τον ορισμό είναι προφανές από βασικά μαθηματικά ότι το **σήμα (1) δεν είναι γραμμικό** καθώς περιέχει πολλαπλασιασμό μεταξύ  $x(n-2)*x(n-3)$ . Ομοίως και το **σήμα (3)** είναι προφανές ότι **δεν είναι γραμμικό** λογώ του εκθετικού όρου. Το **σήμα (2) είναι γραμμικό** γιατί περιέχει απλος μια πρόσθεση των 2 τιμών εισόδου.

Ένα σύστημα είναι **αιτιατό (causal)** όταν η παρούσα τιμή της εξόδου του δεν εξαρτάται από μελλοντικές τιμές της εισόδου. Δηλαδή, για κάθε σήμα εισόδου  $x(t)$  η αντίστοιχη έξοδος  $y(t)$  εξαρτάται μόνο από την παρούσα ή/και τις προηγούμενες τιμές της εισόδου. Αντίστροφα, αν η έξοδος  $y(t_0)$  τη χρονική στιγμή  $t_0$  εξαρτάται από μεταγενέστερες τιμές του σήματος εισόδου  $x(t)$ , δηλ. για  $t \geq t_0$ , τότε το σύστημα είναι μη-αιτιατό. Με βάση αυτό τον ορισμό βλέπουμε ότι και **τα σήματα (1) και (3) είναι αιτιατά** καθώς σε καμία από τις 2 εξισώσεις δεν υπάρχει κάποια εξαρτηση από μελλοντική τιμή του  $n$ , παρα μόνο από παροντές και παρελθοντικές τιμές. Αντιθέτως το **σήμα(2)  $y(n)=x(n)+x(-n)$  δεν είναι αιτιατό** καθώς για αρνητικές τιμές του  $n$  ο δεύτερος όρος πέρνει θετικές τιμές δηλαδή τιμές μεγαλύτερες ( και επομένως μελλοντικές)σε σχεσή με την τρεχουσα τιμή  $n$  του  $y(n)$ .

Τέλος ένα σύστημα λέγεται χρονικά αμετάβλητο ή χρονικά αναλοίωτο, όταν και μόνον όταν χρονικές ολισθήσεις του σήματος εισόδου μεταφράζονται σε αντίστοιχες χρονικές ολισθήσεις στο σήμα εξόδου. Σε ένα χρονικά αμετάβλητο σύστημα αν ένα σήμα εισόδου  $x(t)$  προκαλεί έξοδο  $y(t)$ , τότε ένα σήμα εισόδου  $x(t - t_0)$  παράγει σήμα εξόδου  $y(t - t_0)$  Αναλύοντας τα 3 σήματα βλέπουμε ότι το σήμα (1) και (3) δεν εξαρτάται απευθείας από το χρόνο. Για παράδειγμα ας αναλύσουμε το πρώτο σήμα:

$$y(n)=x(n-1)+x(n-2)*x(n-3)$$

Αντικαθιστώντας το  $x(n)$  με  $x(n-\kappa)$ :

$$y(n-\kappa)=x(n-1-\kappa)+x(n-2-\kappa)*x(n-3-\kappa)$$

Τώρα αν συγκρίνουμε αυτό το σήμα με το σήμα  $y(n-\kappa)$  πρέπει να είναι όμοια για να είναι χρονικά αναλοίωτα:

$$y(n-\kappa)=x(n-1-\kappa)+x(n-2-\kappa)*x(n-3-\kappa)$$

Όντως οι 2 εξισώσεις είναι ίδιες άρα το σήμα είναι χρονικά αναλοίωτο. Με όμοιο τρόπο αποδυκνείσουμε ότι και σήμα (3) είναι χρονικά αναλοίωτα. Επομένως το σήμα **(1) και (3) ειναι χρονικά αναλοίωτα**.

Αντιθέτως το σήμα (2) είναι χρονικά μεταβλητό γιατί αν ακολουθήσουμε την παραπάντας διαδικασία:

$$y(n,\kappa) = x(n-\kappa) + x(-n-\kappa)$$

Τώρα αν συγκρίνουμε αυτό το σήμα με το σήμα  $y(n-\kappa)$  παρατηρούμε ότι οι εξισώσεις είναι διαφορετικές:

$$y(n-\kappa) = x(n-\kappa) + x(-n+\kappa)$$

Συγκεκριμένα βλέπουμε ότι το πρόσημο του κ στο δεύτερο όρο αναστρέφεται. Επομένως το **σήμα(2) είναι χρονικά μεταβαλώμενο.**

### Ερώτημα 3

Για τους μαθηματικούς υπολογισμούς και τη δημιουργία των γραφημάτων χρησιμοποιήθηκε η γλώσσα Python μαζί με τις βιβλιοθήκες numpy και matplotlib. Οι υπολογισμοί πραγματοποιήθηκαν με βάση τους μαθηματικούς τύπους που περιγράφουν κάθε τύπο διαμόρφωσης όπως αυτος αναφέρεται στις σημειώσεις του μαθήματος (Χ.Δουληγέρης, Γ.Α.Τσιχριντζής).

#### (a) DSB-SC

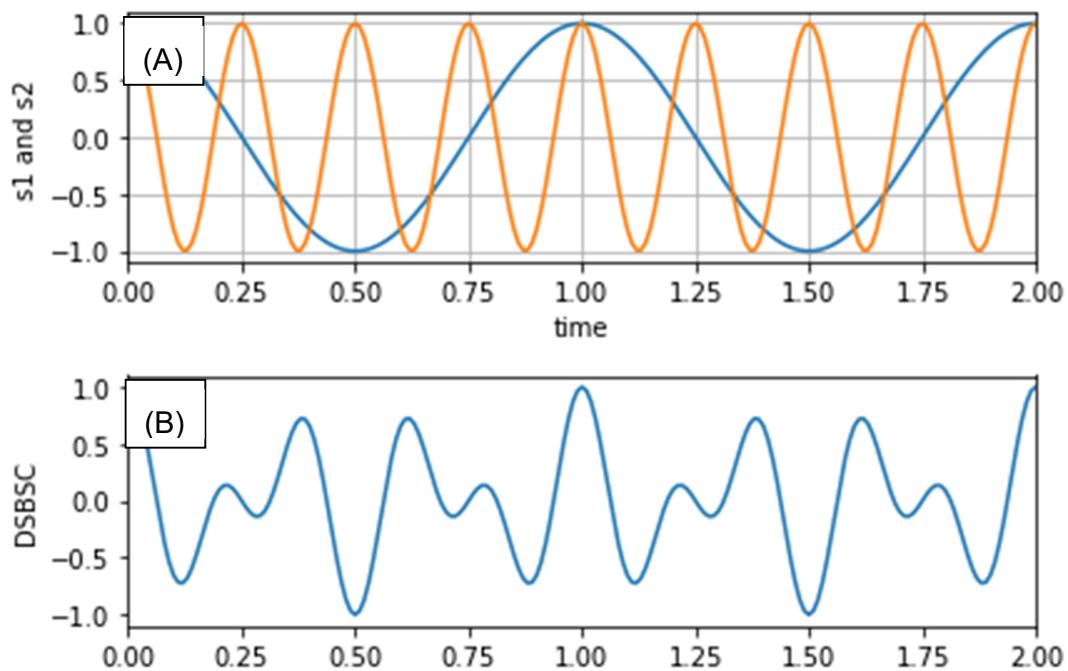


Figure 1. (A) Τα 2 αρχικά Στο γράφημα (B) Το τελικό DSB-SC κύμα.

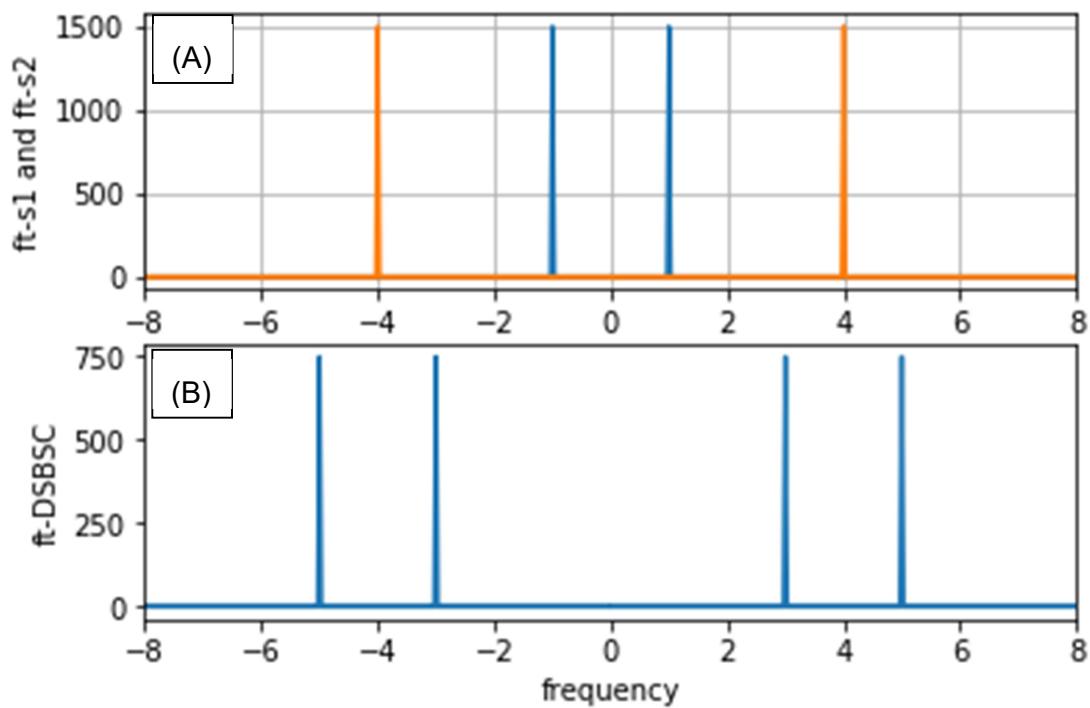


Figure 2. (A) Οι συχνότητες των 2 αρχικών σημάτων, (B) Η συχνότητες του τελικού DSB-SC σήματος.

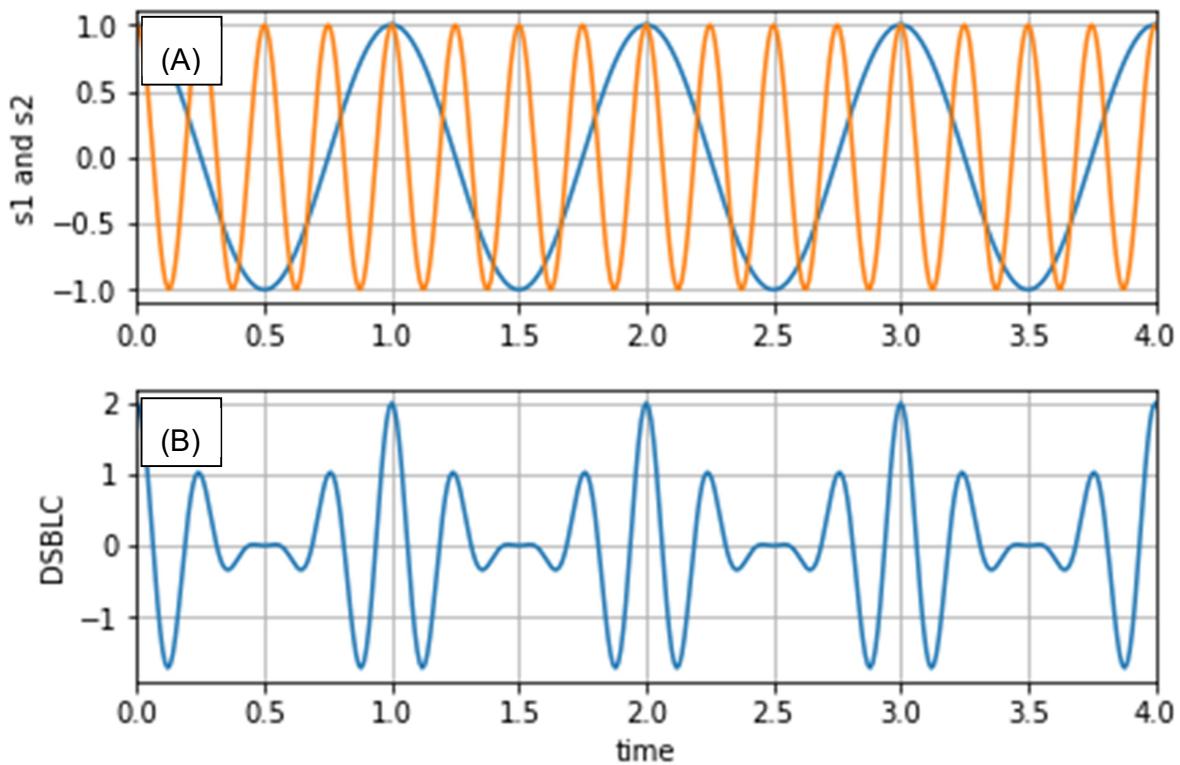
Όπως φένεται στο σχήμα 2 το τελικό DSB-SC σήμα έχει 2 συχνότητες που ισούνται με το αθροισμα και τη διαφορά των συχνοτήτων των αρχικών σημάτων κάτι που συμφωνεί απόλυτα με την θεωρία.

### (b) DSB-LC m=1

Για αυτή την διαμόρφωση το κύμα περιγράφεται από την εξίσωση:

$$\varphi AM(t) = [A + f(t)] \cos \omega_c t$$

Όπως φένεται παρακάτω στο σχήμα 4 στο πεδίο των συχνοτήτων η διαμόρφωση DSB-LC μοιάζει με την DSB-SC μόνο που στο φάσμα των συχνοτήτων προστίθεται τώρα και η συχνότητα του φέροντος κύματος (4Hz) όπως ακριβώς προβλέπεται και από τη θεωρία.



**Figure 3. (A)** Τα 2 αρχικά Στο γράφημα **(B)** Το τελικό DSB-LC κύμα.

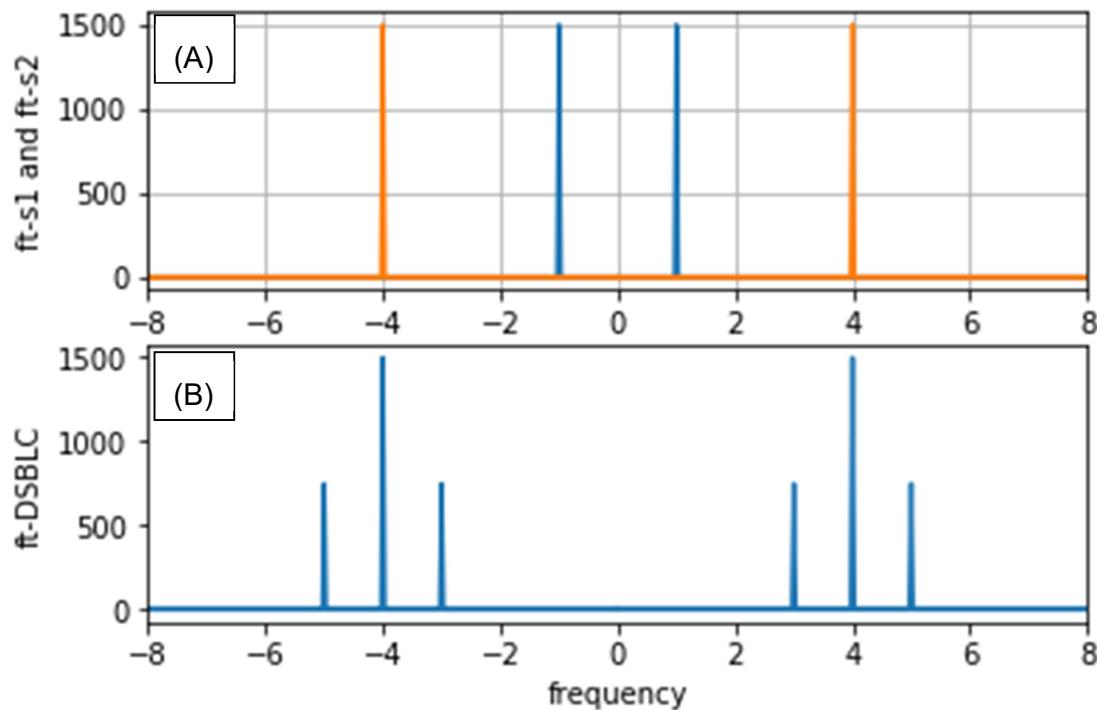


Figure 4. (A) Οι συχνότητες των 2 αρχικών σημάτων, (B) Η συχνότητες του τελικού DSB-LC σήματος.

### (c) SSB-SC+

Για την διαμόρφωση του άνω μισού της πλήρους AM διαμόρφωσης το κύμα περιγράφεται από την εξίσωση:

$$s(t)=AmAc2\cos[2\pi(fc+fm)t]$$

Στο σχήμα 6 παρατηρούμε ότι το διαμορφωμένο σήμα αποτελείται μόνο από το άνω μέρος των 2 συχνοτήτων του DSB-SC σήματος από το ερώτημα (a). Δηλαδή όπως προκύπτει και από την ανωτέρω εξίσωση η συχνότητα του διαμορφωμένου άνω SSB-SC+ είναι το άρθροισμα των συχνοτήτων του φέροντος κύματος και του κύματος διαμόρφωσης (5Hz).

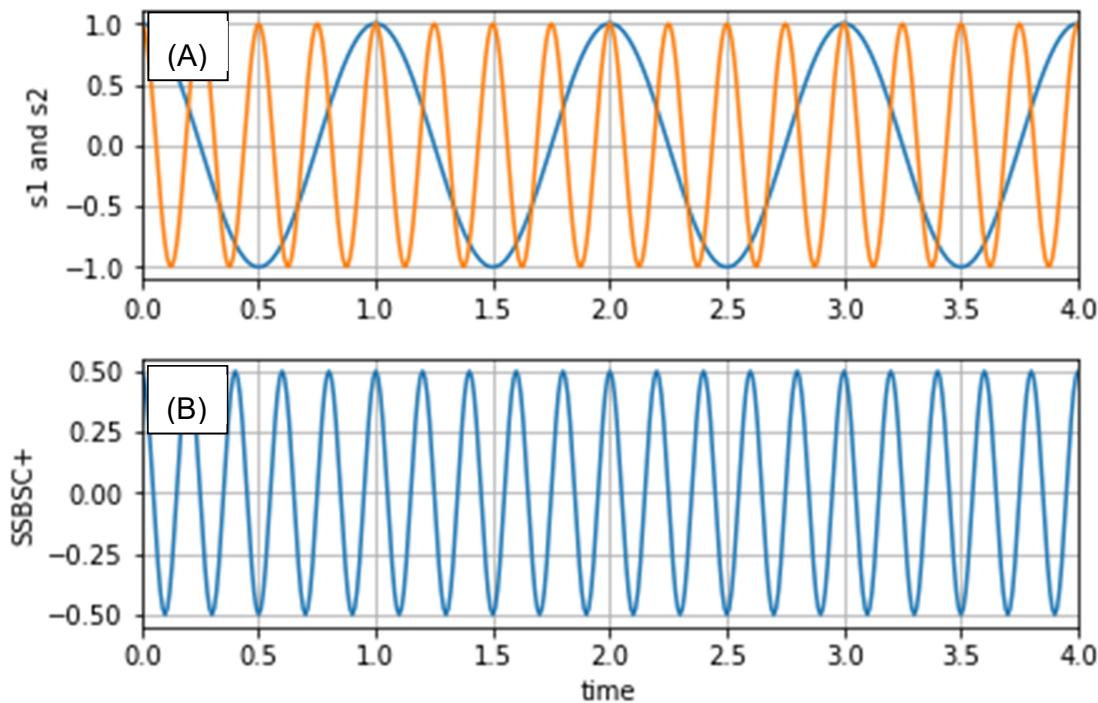


Figure 5. (A) Τα 2 αρχικά Στο γράφημα (B) Το τελικό SSB-SC+ κύμα.

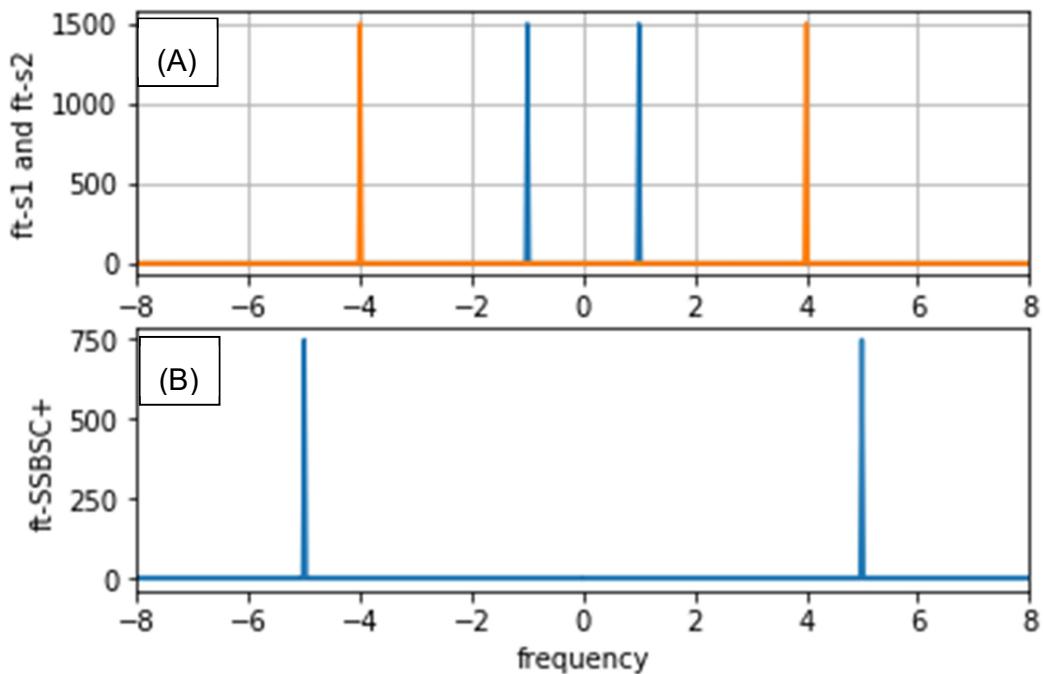


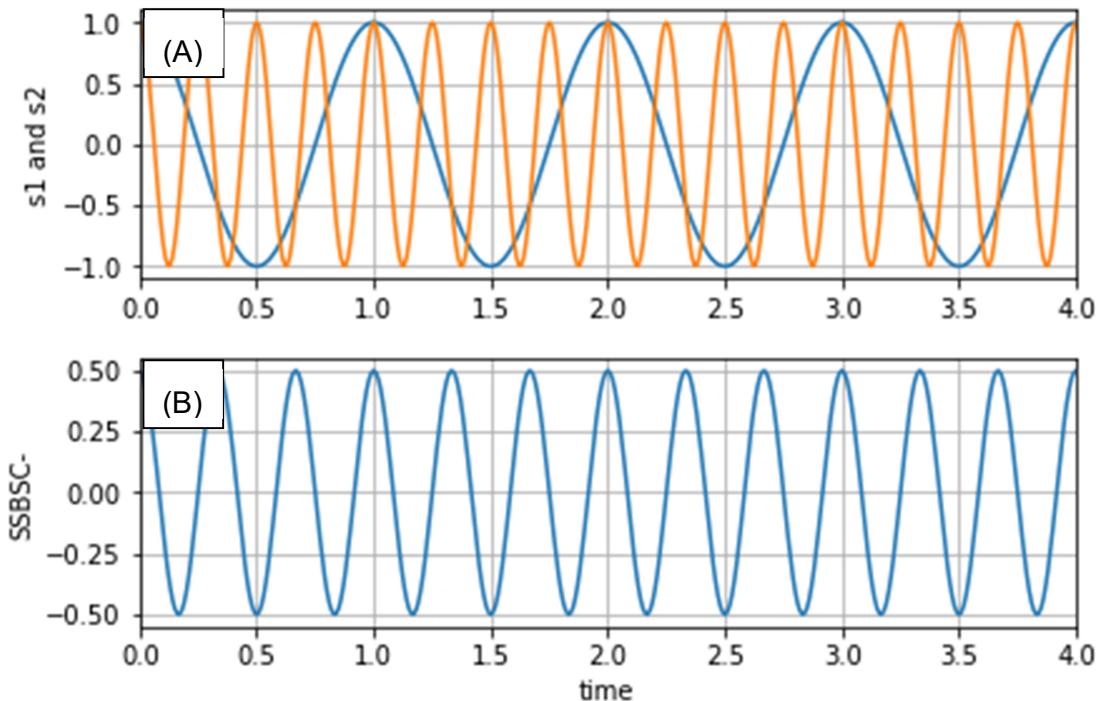
Figure 6. (A) Οι συχνότητες των 2 αρχικών σημάτων, (B) Η συχνότητες του τελικού SSB-SC+ σήματος.

**(d) SSB-SC-**

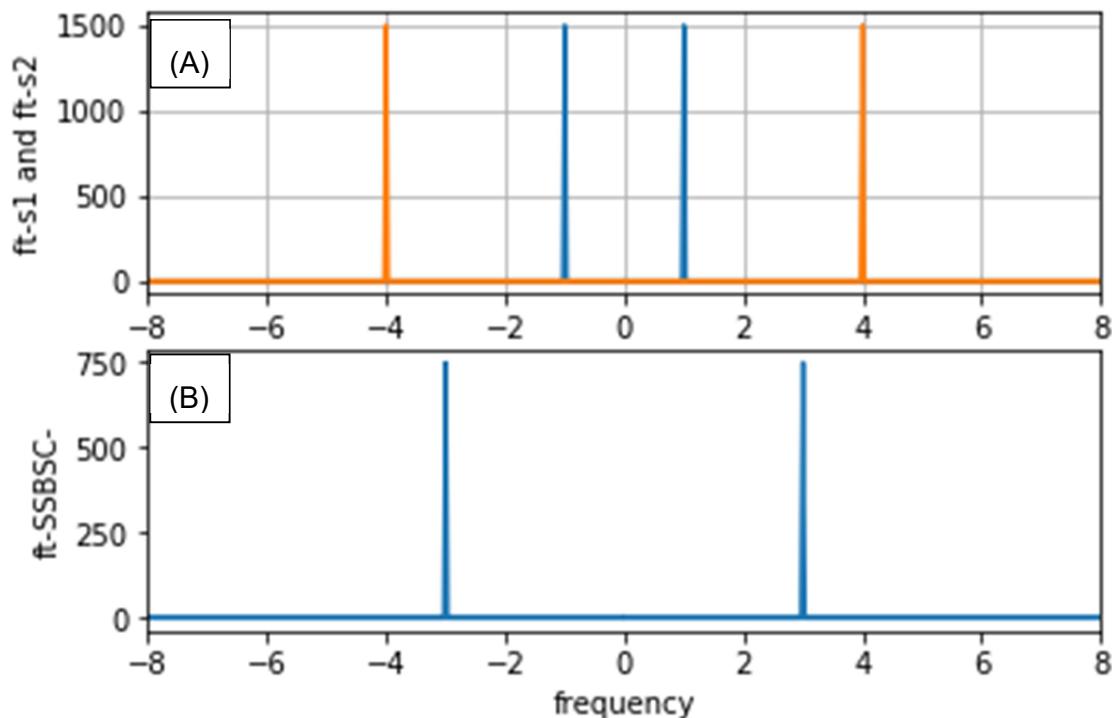
- (e) Για την διαμόρφωση του κάτω μισού της πλήρους AM διαμόρφωσης το κύμα περιγράφεται από την εξίσωση:

$$(f) s(t)=AmAc2\cos[2\pi(fc-fm)t]$$

Όπως και στο προηγουμένο ερώτημα παρατηρούμε στα σχήματα 7 και 8 ότι το διαμορφωμένο σήμα συμφωνεί πλήρως με το αναμενόμενο από την θεωρία σήμα. Συγκεκριμένα παρατηρούμε ότι το διαμορφωμένο σήμα αποτελείται μόνο από το κάτω μέρος των 2 συχνοτήτων του DSB-SC σήματος από το ερώτημα (a). Δηλαδή όπως προκύπτει και από την ανωτέρω εξίσωση η συχνότητα του διαμορφομένου άνω SSB-SC- είναι η διαφορά των συχνοτήτων του φέροντος κύματος και του κύματος διαμόρφωσης (3Hz).



**Figure 7. (A)** Τα 2 αρχικά Στο γράφημα **(B)** Το τελικό SSB-SC- κύμα.



**Figure 8. (A)** Οι συχνότητες των 2 αρχικών σημάτων, **(B)** Η συχνότητες του τελικού SSB-SC- σήματος.

Τέλος στο ερώτημα για το λόγο των ισχύων των σημάτων σε σχέση με την ισχύ του σήματος (a) μπορούμε να συμβολευτούμε τις μαθηματικές εξισώσεις ισχύος των αντίστοιχων σημάτων. Έτσι η διαμόρφωση DSB-SC έχει ισχύ:

$$P_{DSBSC} = \frac{A_m^2 A_c^2}{4R}$$

Η διαμόρφωση DSB-LC έχει συνολική ισχύς, το άθροισμα ισχύος του φέροντος σήματος και των πλευρικών:

$$P_{DSBLC} = \frac{A_c^2}{2R} \left( 1 + \frac{m^2}{2} \right)$$

Επίσης το κλάσμα της συνολικής ισχύος που περιέχεται στις πλευρικές συχνότητες (πέριξ της φέροντος) ως προς την συνολική ισχή δίδεται από:

$$\mu = \frac{P_S}{P_t} = \frac{P_{DSBSC}}{P_{DSBLC}} = \frac{m^2}{2 + m^2}$$

Στην άσκηση μας το  $m=1$  επομένως, το κλάσμα ισχύος είναι 33.3%. Επομένως η διαμόρφωση DSB-LC απαιτεί 3 φορές περισότερη ισχύ.

Τέλος κάθε μια (η άνω και η κάτω) διαμόρφωση SSB-SC έχει συνολική ισχύς:

$$P_{SSBSC} = \frac{A_m^2 A_c^2}{8R} = \frac{P_{DSBSC}}{2}$$

Όπου με απλή σύγκριση με την DSB-SC βλέπουμε ότι η SSB-SC έχει την μισή ισχύ.

## Παράρτημα-Κώδικας Python

```
# -*- coding: utf-8 -*-

"""
Created on Tue Sep 24 19:52:35 2019

@author: Kostas Ziovas
Signal Processing Project 2019
"""

import numpy as np
import matplotlib.pyplot as plt

"""(a) DSB-SC MODULATION"""
# Fixing random state for reproducibility
np.random.seed(19680801)

dt = 0.01
t = np.arange(0, 30, dt)
#nse1 = np.random.randn(len(t))                                # white noise 1
#nse2 = np.random.randn(len(t))                                # white noise 2

# Two signals and their DSB-SC MODULATION
s1 = np.cos(2 * np.pi * t)
s2 = np.cos(2 * np.pi * 4 * t)
DSBSC=s1*s2

fig, axs = plt.subplots(2, 1)
axs[0].plot(t, s1, t, s2)
axs[0].set_xlim(0, 2)
axs[0].set_xlabel('time')
axs[0].set_ylabel('s1 and s2')
axs[0].grid(True)

axs[1].plot(t,DSBSC)
axs[1].set_xlim(0, 2)
axs[1].set_ylabel('DSBSC')

#cxy, f = axs[1].cohere(s1, s2, 256, 1. / dt)
#axs[1].set_ylabel('coherence')

fig.tight_layout()
plt.show()

#The fft of the same signals
fts1=np.fft.fft(s1)
fts2=np.fft.fft(s2)
ftDSB=np.fft.fft(DSBSC)
freq = np.fft.fftfreq(len(s1),dt)

fig, axs = plt.subplots(2, 1)
axs[0].plot(freq, fts1, freq, fts2)
axs[0].set_xlim(-8, 8)
axs[0].set_xlabel('frequency')
axs[0].set_ylabel('ft-s1 and ft-s2')
axs[0].grid(True)
```

```

axs[1].plot(freq, ftDSB)
axs[1].set_xlim(-8, 8)
axs[1].set_xlabel('frequency')
axs[1].set_ylabel('ft-DSBSC')

"""(b) DSB-LC m=1 MODULATION"""
#XXXXXXXXXXXXXXXXXXXXXXXXXXXX Define AM modulation Index XXXXXXXXXXXXXXXXXXXXXXX
#example: m=1 means 100% modulation
#m=input(' Enter the value of modulation index (m) = ');
m=1; # for 100% modulation

#XXXXXXXXXXXXXXXXXXXX modulating signal generation XXXXXXXXXXXXXXXXXXXXXXX
Am=1 # Amplitude of modulating signal
fa=1 # Frequency of modulating signal
#Ta=1/fa; % Time period of modulating signal
#t=0:Ta/999:6*Ta; # Total time for simulation
dt = 0.01
t = np.arange(0, 30, dt)

ym=Am*np.cos(2*np.pi*fa*t)# Equation of modulating signal

fig, axs = plt.subplots(2, 1)
fig.tight_layout()
#axs[0].plot(t, ym)
#axs[0].set_xlim(0, 4)
#axs[0].set_xlabel('time')
#axs[0].set_ylabel('s1')
#axs[0].grid(True)

#XXXXXXXXXXXXXXXXXXXX carrier signal generation XXXXXXXXXXXXXXXXXXXXXXX
Ac=Am/m# Amplitude of carrier signal [ where, modulation Index (m)=Am/Ac ]
fc=fa*4# Frequency of carrier signal
#Tc=1/fc;# Time period of carrier signal

yc=Ac*np.cos(2*np.pi*fc*t)#Equation of carrier signal

axs[0].plot(t,ym,t, yc)
axs[0].set_xlim(0, 4)
#axs[1].set_xlabel('time')
axs[0].set_ylabel('s1 and s2')
axs[0].grid(True)

#XXXXXXXXXXXXXXXXXXXXXXXXXXXX AM Modulation XXXXXXXXXXXXXXXXXXXXXXX

yDSBLC=Ac*(1+m*np.cos(2*np.pi*fa*t))*np.cos(2*np.pi*fc*t) # Equation of
DSBLC

axs[1].plot(t, yDSBLC)
axs[1].set_xlim(0, 4)
axs[1].set_xlabel('time')
axs[1].set_ylabel('DSBLC')
axs[1].grid(True)

#The fft of the same signals
fts1=np.fft.fft(ym)
fts2=np.fft.fft(yc)

```

```

ftDSB=np.fft.fft(yDSBLC)
freq = np.fft.fftfreq(len(s1),dt)

fig, axs = plt.subplots(2, 1)
axs[0].plot(freq, fts1, freq, fts2)
axs[0].set_xlim(-8, 8)
axs[0].set_xlabel('frequency')
axs[0].set_ylabel('ft-s1 and ft-s2')
axs[0].grid(True)

axs[1].plot(freq,ftDSB)
axs[1].set_xlim(-8, 8)
axs[1].set_xlabel('frequency')
axs[1].set_ylabel('ft-DSBLC')

"""(c) SSB-SC+ MODULATION"""

# Message Signal
dt = 0.01
t = np.arange(0, 30, dt)
Am=1 # Amplitude of modulating signal
fa=1 # Frequency of modulating signal
ym=Am*np.cos(2*np.pi*fa*t)# Equation of modulating signal

Ac=1# Amplitude of carrier signal
fc=4# Frequency of carrier signal
yc=Ac*np.cos(2*np.pi*fc*t)#Equation of carrier signal

ySSBSC=np.cos(2*np.pi*(fc+fa)*t)*(Am*Ac)/2 # Equation of SSBSC

fig, axs = plt.subplots(2, 1)
fig.tight_layout()

axs[0].plot(t,ym,t, yc)
axs[0].set_xlim(0, 4)
#axs[1].set_xlabel('time')
axs[0].set_ylabel('s1 and s2')
axs[0].grid(True)

axs[1].plot(t, ySSBSC)
axs[1].set_xlim(0, 4)
axs[1].set_xlabel('time')
axs[1].set_ylabel('SSBSC+')
axs[1].grid(True)

#The fft of the same signals
fts1=np.fft.fft(ym)
fts2=np.fft.fft(yc)
ftDSB=np.fft.fft(ySSBSC)
freq = np.fft.fftfreq(len(s1),dt)

fig, axs = plt.subplots(2, 1)
axs[0].plot(freq, fts1, freq, fts2)
axs[0].set_xlim(-8, 8)
axs[0].set_xlabel('frequency')
axs[0].set_ylabel('ft-s1 and ft-s2')
axs[0].grid(True)

```

```

axs[1].plot(freq, ftDSB)
axs[1].set_xlim(-8, 8)
axs[1].set_xlabel('frequency')
axs[1].set_ylabel('ft-SSBSC+')

"""(c) SSB-SC- MODULATION"""

# Message Signal
dt = 0.01
t = np.arange(0, 30, dt)
Am=1 # Amplitude of modulating signal
fa=1 # Frequency of modulating signal
ym=Am*np.cos(2*np.pi*fa*t)# Equation of modulating signal

Ac=1# Amplitude of carrier signal
fc=4# Frequency of carrier signal
yc=Ac*np.cos(2*np.pi*fc*t)#Equation of carrier signal

ySSBSC=np.cos(2*np.pi*(fc-fa)*t)*(Am*Ac)/2 # Equation of SSBSC

fig, axs = plt.subplots(2, 1)
fig.tight_layout()

axs[0].plot(t,ym,t, yc)
axs[0].set_xlim(0, 4)
#axs[1].set_xlabel('time')
axs[0].set_ylabel('s1 and s2')
axs[0].grid(True)

axs[1].plot(t, ySSBSC)
axs[1].set_xlim(0, 4)
axs[1].set_xlabel('time')
axs[1].set_ylabel('SSBSC-')
axs[1].grid(True)

#The fft of the same signals
fts1=np.fft.fft(ym)
fts2=np.fft.fft(yc)
ftDSB=np.fft.fft(ySSBSC)
freq = np.fft.fftfreq(len(s1),dt)

fig, axs = plt.subplots(2, 1)
axs[0].plot(freq, fts1, freq, fts2)
axs[0].set_xlim(-8, 8)
axs[0].set_xlabel('frequency')
axs[0].set_ylabel('ft-s1 and ft-s2')
axs[0].grid(True)

axs[1].plot(freq,ftDSB)
axs[1].set_xlim(-8, 8)
axs[1].set_xlabel('frequency')
axs[1].set_ylabel('ft-SSBSC-')

```