



ΤΕΧΝΟΛΟΓΙΚΟ ΕΚΠΑΙΔΕΥΤΙΚΟ ΙΔΡΥΜΑ
(ΤΕΙ) ΠΕΛΟΠΟΝΝΗΣΟΥ

ΣΧΟΛΗ ΤΕΧΝΟΛΟΓΙΚΩΝ ΕΦΑΡΜΟΓΩΝ (ΣΠΑΡΤΗ)
ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΠΛΗΡΟΦΟΡΙΚΗΣ Τ.Ε.

ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

“ΚΩΔΙΚΟΠΟΙΗΣΗ ΠΗΓΗΣ ΠΟΛΥΦΑΣΜΑΤΙΚΩΝ ΕΙΚΟΝΩΝ ΜΕ
ΤΗΝ ΧΡΗΣΗ ΤΩΝ ΑΛΓΟΡΙΘΜΩΝ JPEG”



ΦΟΙΤΗΤΡΙΑ: ΚΟΚΚΙΝΑΚΗ ΑΙΚΑΤΕΡΙΝΗ-ΕΙΡΗΝΗ Α.Μ.: 2010091

ΕΠΟΠΤΗΣ ΚΑΘΗΓΗΤΗΣ: ΔΡ. ΒΑΣΙΛΕΙΟΣ ΜΠΟΖΑΝΤΖΗΣ

ΣΠΑΡΤΗ, ΝΟΕΜΒΡΙΟΣ 2015

ΠΕΡΙΛΗΨΗ

Η εργασία αποτελεί μια εκτεταμένη επισκόπηση της μεθόδου συμπίεσης εικόνων με τον αλγόριθμο JPEG. Στο πρώτο μέρος της εργασίας, περιγράφονται οι μέθοδοι προεπεξεργασίας και ανάλυσης εικόνων που χρησιμοποιούνται στη Γεωπληροφορική.

Στη συνέχεια, περιγράφονται οι τεχνικές αναγωγής και ταυτοποίησης εικόνων. Η σωστή αναγωγή και ταυτοποίηση αποτελούν βασική προϋπόθεση για την επιτυχή ανάλυση των τηλεσκοπικών δεδομένων, η οποία ακολουθεί την προεπεξεργασία.

Ωστόσο, περιγράφεται αναλυτικά ο αλγόριθμος JPEG, όπως τα μαθηματικά και η λογική που τον διέπουν. Επίσης, υπάρχουν πειραματικά αποτελέσματα εικόνων που προέρχονται από τον κώδικα που έγραψα για να δοκιμάσω τον αλγόριθμο JPEG.

Οι εικόνες που παρήχθησαν από τα πειράματα, είναι από τη χρήση του Matlab.

ΔΗΛΩΣΗ ΜΗ ΛΟΓΟΚΛΟΠΗΣ ΚΑΙ ΑΝΑΛΗΨΗΣ ΠΡΟΣΩΠΙΚΗΣ ΕΥΘΥΝΗΣ

Με πλήρη επίγνωση των συνεπειών του νόμου περί πνευματικών δικαιωμάτων, δηλώνω ενυπογράφως ότι είμαι αποκλειστικός συγγραφέας της παρούσας Πτυχιακής Εργασίας, για την ολοκλήρωση της οποίας κάθε βοήθεια είναι πλήρως αναγνωρισμένη και αναφέρεται λεπτομερώς στην εργασία αυτή.

Έχω αναφέρει πλήρως και με σαφείς αναφορές, όλες τις πηγές χρήσης δεδομένων, απόψεων, θέσεων και προτάσεων, ιδεών και λεκτικών αναφορών, είτε κατά κυριολεξία είτε βάση επιστημονικής παράφρασης.

Αναλαμβάνω την προσωπική και ατομική ευθύνη ότι σε περίπτωση αποτυχίας στην υλοποίηση των ανωτέρω δηλωθέντων στοιχείων, είμαι υπόλογος έναντι λογοκλοπής, γεγονός που σημαίνει αποτυχία στην Πτυχιακή μου Εργασία και κατά συνέπεια αποτυχία απόκτησης του Τίτλου Σπουδών, πέραν των λοιπών συνεπειών του νόμου περί πνευματικών δικαιωμάτων.

Δηλώνω, συνεπώς, ότι αυτή η Πτυχιακή Εργασία προετοιμάστηκε και ολοκληρώθηκε από εμένα προσωπικά και αποκλειστικά και ότι, αναλαμβάνω πλήρως όλες τις συνέπειες του νόμου μου ανήκει διότι είναι στην περίπτωση κατά την οποία αποδειχθεί, διαχρονικά, ότι η εργασία αυτή ή τμήμα της δε προϊόν λογοκλοπής άλλης πνευματικής ιδιοκτησίας.

Όνομα και Επώνυμο Συγγραφέα (Με Κεφαλαία):.....

Υπογραφή (Ολογράφως, χωρίς μονογραφή):

Ημερομηνία (Ημέρα – Μήνας – Έτος):

ΕΥΧΑΡΙΣΤΙΕΣ

Σε αυτό το σημείο, θα ήθελα να ευχαριστήσω για ακόμη μια φορά τον επιβλέποντα καθηγητή μου κ. Βασίλειο Μποζαντζή για την εμπιστοσύνη, τον σημαντικό χρόνο που αφιέρωσε, καθώς και τις πολύτιμες πληροφορίες που μου μετέδωσε καθ' όλη τη διάρκεια εκπόνησης της εργασίας.

Επιπλέον, θα ήθελα να ευχαριστήσω την οικογένεια μου για όλα όσα μου έχει προσφέρει στην διάρκεια των μαθητικών και φοιτητικών μου χρόνων και την αμέριστη υποστήριξη της σε κάθε μου επιλογή.

Τέλος, θα ήθελα να ευχαριστήσω όλους τους φίλους μου, που ο καθένας ξεχωριστά με τον δικό του τρόπο έβαλε το λιθαράκι του για την περάτωση της εργασίας αυτής.

ΠΕΡΙΕΧΟΜΕΝΑ

ΠΕΡΙΛΗΨΗ 2

ΔΗΛΩΣΗ ΜΗ ΛΟΓΟΚΛΟΠΗΣ 3

ΕΚΦΡΑΣΗ ΕΥΧΑΡΙΣΤΙΩΝ 4

ΚΕΦΑΛΑΙΟ 1^ο

Εισαγωγή στις Τεχνικές Επεξεργασίας και Ανάλυσης Εικόνων με Εφαρμογή στη Γεωπληροφορική..... 7

ΚΕΦΑΛΑΙΟ 2^ο

Αναγωγή και Ταυτοποίηση Εικόνων 11

ΚΕΦΑΛΑΙΟ 3^ο

Ποιοτική Ανάλυση Εικόνων..... 14

3.1 Κατάταξη ή Ταξινόμηση Ψηφιακών Δεδομένων Εικόνων 14

3.2 Τμηματοποίηση Εικόνων Και Video 18

ΚΕΦΑΛΑΙΟ 4^ο

JPEG Συμπίεση 21

4.1 Αλγόριθμος JPEG..... 25

4.2 Διακριτός Μετασχηματισμός Συνημίτονου (DCT)..... 26

4.3 Ακύρωση Υψηλών Συχνοτήτων..... 34

4.4 JPEG Βελτίωση..... 35

4.5 Εξοικονόμηση Μνήμης..... 39

4.6 Γιατί 8 x 8 blocks 39

ΚΕΦΑΛΑΙΟ 5^ο

Πειράματα..... 41

ΚΕΦΑΛΑΙΟ 6^ο

Συμπεράσματα Και Μελλοντική Εργασία..... 45

ΚΕΦΑΛΑΙΟ 7^ο

Βιβλιογραφία 46

ΚΕΦΑΛΑΙΟ 1^ο

Εισαγωγή στις Τεχνικές Επεξεργασίας και Ανάλυσης Εικόνων με Εφαρμογή στη Γεωπληροφορική

Η τηλεπισκόπηση ή τηλεανίχνευση (remote sensing) είναι η επιστήμη παρατήρησης φαινομένων και χαρακτηριστικών από απόσταση, με μεγάλο εύρος εφαρμογών, εκτός των άλλων και στη Γεωπληροφορική.

Σήμερα, με τον όρο τηλεπισκόπηση εννοούμε την επιστήμη και τεχνολογία παρατήρησης και μελέτης των χαρακτηριστικών της γήινης επιφάνειας από απόσταση, βάσει της αλληλεπίδρασης των υλικών που βρίσκονται επάνω σε αυτή με την ηλεκτρομαγνητική ακτινοβολία.

Τα επιτεύγματα της τηλεπισκόπησης χρησιμοποιούνται σε εφαρμογές της Γεωπληροφορικής όπως το Κτηματολόγιο, το σχεδιασμό χρήσεων γής και τη δημιουργία ψηφιακών μοντέλων γής.

Η παρατήρηση της επιφάνειας της γης είναι δυνατή με τη χρήση τηλεπισκοπικών ανιχνευτών που ανιχνεύουν την αντανάκλαση της ηλεκτρομαγνητικής ακτινοβολίας της γήινης επιφάνειας και την αναπαριστούν ως ψηφιακή εικόνα. Εάν χρησιμοποιείται το ορατό τμήμα της ηλεκτρομαγνητικής ακτινοβολίας για την αναπαράσταση, προκύπτει μια έγχρωμη εικόνα, ισοδύναμη με αυτές που καταγράφουν οι ψηφιακές μηχανές φωτογράφισης που χρησιμοποιούμε καθημερινά.

Στην πράξη, οι ανιχνευτές «βλέπουν» πολύ πέρα από το ορατό φάσμα με αποτέλεσμα να λαμβάνουμε μια ποικιλία εικόνων. Έτσι, το αποτέλεσμα ποικίλει ανάλογα με τη φασματική ζώνη στην οποία λειτουργεί ο αισθητήρας του ανιχνευτή.

Η σωστή αξιοποίηση των τηλεπισκοπικών δεδομένων προϋποθέτει την προεπεξεργασία τους ώστε να απαλειφθούν μια σειρά από σφάλματα, αλλοιώσεις και μεταβολές που εάν παραμείνουν θα υποβαθμίσουν την ποιότητα της τελικής πληροφορίας και θα οδηγήσουν σε εσφαλμένα ή αλλοιωμένα συμπεράσματα.

Οι διορθώσεις στην εικόνα που λαμβάνεται από τον ανιχνευτή μπορεί να είναι:

- Γεωμετρικές διορθώσεις που έχουν να κάνουν με την αναγωγή της καμπύλης γεωμετρίας της εικόνας σε επίπεδη γεωμετρία απαλείφοντας τα σφάλματα που προέρχονται από την επιφάνεια του γήινου ελλειψοειδούς αλλά και της κίνησης της γης καθώς και του τεχνητού δορυφόρου πάνω στον οποίο βρίσκεται ο ανιχνευτής. Επίσης, γεωμετρικές διορθώσεις επιβάλλονται ώστε να διορθωθούν σφάλματα που οφείλονται στο έντονο ανάγλυφο μιας περιοχής, διαδικασία που είναι γνωστή ως ορθοδιόρθωση ή ορθοαναγωγή. Τέλος η εικόνα θα πρέπει να αναχθεί σε ένα σύστημα γεωγραφικών ή προβολικών συντεταγμένων ώστε να μπορεί να συνδυασθεί με άλλα γεωγραφικά δεδομένα.
- Ραδιομετρικές διορθώσεις, που αναφέρονται στην απαλοιφή σφαλμάτων τα οποία προέρχονται από τα τεχνικά στοιχεία της διαδικασίας καταγραφής και μετάδοσης.
- Ατμοσφαιρικές διορθώσεις, που αφορούν στην απαλοιφή σφαλμάτων τα οποία προέρχονται από την επίδραση των συστατικών της ατμόσφαιρας. Καθώς η ηλεκτρομαγνητική ακτινοβολία που αντανακλάστηκε από την γήινη επιφάνεια επιστρέφει στο διάστημα, διέρχεται μέσα από την ατμόσφαιρα, τα συστατικά της οποίας (νερό σε αέρια φάση και αερολύματα) αλλοιώνουν την ποσότητα της εισερχόμενης στον αισθητήρα ακτινοβολίας. Οι ατμοσφαιρικές διορθώσεις διακρίνονται σε μεθόδους σταθερής τιμής και σε μεθόδους μετάδοσης ακτινοβολίας.

Μεγάλο μέρος των διορθώσεων αυτών γίνονται με τεχνικές επεξεργασίας εικόνων. Η επεξεργασία ψηφιακών εικόνων (digital image processing) αποτελεί τομέα συνεχιζόμενης έρευνας για περισσότερα από 60 χρόνια στην επιστήμη υπολογιστών, με σημαντικά αποτελέσματα και πληθώρα εφαρμογών, από την τηλεόραση έως και την τομογραφία.

Θεωρώντας την ψηφιακή εικόνα ως διδιάστατο σήμα, η επεξεργασία εικόνων αποτελεί επέκταση της επεξεργασίας σημάτων (signal processing) στις 2 διαστάσεις. Υπό αυτή τη θεώρηση, αρκετές τεχνικές επεξεργασίας σημάτων εφαρμόζονται και στην επεξεργασία εικόνων. Η έξοδος της επεξεργασίας μπορεί να είναι μια νέα βελτιωμένη εικόνα ή κάποιες παράμετροι ενδιαφέροντος που εξάγονται από την εικόνα εισόδου.

Μετά το στάδιο προεπεξεργασίας, λαμβάνει χώρα η ανάλυση των επεξεργασμένων εικόνων της τηλεπισκόπησης. Η ανάλυση των τηλεσκοπικών δεδομένων μπορεί να διακριθεί σε τρεις γενικές κατηγορίες: ποσοτική, ποιοτική και οπτική ανάλυση ή φωτοερμηνεία.

Στην πρώτη περίπτωση ζητούμενο είναι η μέτρηση μιας ιδιότητας-μεταβλητής, όπως για παράδειγμα η θερμοκρασία στην επιφάνεια της θάλασσας, γνωστή ως Sea Surface Temperature (SST).

Στην περίπτωση της ποιοτικής ανάλυσης, ζητούμενο είναι η αποτύπωση χαρακτηριστικών όπως οι χρήσεις γης ή ο εντοπισμός και η αναγνώριση συγκεκριμένων υλικών.

Τέλος, όταν ο αναλυτής εικόνας ερμηνεύει τα δεδομένα με οπτικό τρόπο, δηλαδή μεταφράζει το μοτίβο, το χρώμα και το σχήμα της διάταξης των pixels ώστε να εξάγει πληροφορίες, αναφερόμαστε σε οπτική ανάλυση. Κάθε κατηγορία απαιτεί διαφορετική αντιμετώπιση και διαφορετικές μεθόδους και εργαλεία ανάλυσης.

Επίσης, ακόμα και σε κάθε μία από τις παραπάνω κατηγορίες απαιτείται διαφορετική αντιμετώπιση ανάλογα με τα χαρακτηριστικά των διαθέσιμων δεδομένων (χωρική, φασματική και ραδιομετρική ανάλυση).

Στο επόμενο Κεφάλαιο παρουσιάζονται τεχνικές γεωμετρικών διορθώσεων και ειδικότερα αναγωγής/ταυτοποίησης εικόνων που λαμβάνονται από τους αισθητήρες σε γνωστούς χάρτες αναφοράς. Στο Κεφάλαιο 3 περιγράφονται αναλυτικά οι διεργασίες ποιοτικής ανάλυσης πολυφασματικών δεδομένων τηλεπισκόπησης που βασίζονται σε γνωστές τεχνικές κατάταξης και τμηματοποίησης οπτικών δεδομένων από τον τομέα ανάλυσης εικόνων της επιστήμης υπολογιστών.

ΚΕΦΑΛΑΙΟ 2^ο

Αναγωγή και Ταυτοποίηση Εικόνων

Τα ψηφιακά δεδομένα τα οποία λαμβάνονται από τους αισθητήρες των δορυφόρων υφίστανται προεπεξεργασία ώστε να αναχθούν γεωμετρικά σε ένα χάρτη αναφοράς (base map), βάσει συγκεκριμένου συντελεστή κλίμακας και πίνακα προβολής (projection map).

Εναλλακτικά, η γεωμετρική προεπεξεργασία μπορεί να αναφέρεται στην αντιπαραβολή ή ταυτοποίηση (registration) της ληφθείσας εικόνας σε σχέση με μια δοσμένη σειρά εικόνων που έχουν ληφθεί προγενέστερα, με σκοπό για παράδειγμα, την παρακολούθηση αλλαγών που έχουν συντελεστεί στο μεσοδιάστημα.

Γενικά, η ταυτοποίηση είναι μια αυτόματη διαδικασία κατά την οποία η θέση του δορυφόρου σε σχέση με τη γη κατά τη στιγμή λήψης της εικόνας προσδιορίζεται με βάση:

- το λεγόμενο μοντέλο ephemeris των παραμέτρων τροχιάς του δορυφόρου,
- τις προδιαγραφές λήψης του αισθητήρα και
- ενός σήματος χρόνου που στέλνονται στην επίγεια βάση.

Αν και οι αλγόριθμοι υπολογισμού θέσης βασίστηκαν αρχικά στην υπόθεση ότι οι δορυφόροι βρίσκονται συνεχώς στη σωστή στάση (attitude) και ακολουθούν ακριβώς την προδιαγεγραμμένη τροχιά τους, κάτι τέτοιο δεν ισχύει στην πράξη. Λόγω των παρεκλίσεων από τις σχεδιαστικές τιμές στάσης και τροχιάς καθώς και της διασποράς αυτών των παρεκλίσεων, για κάποιες από τις λαμβανόμενες εικόνες δεν υπάρχει κατάλληλος χάρτης αναφοράς στον οποίο θα μπορούσαν να αναχθούν με ακρίβεια.

Οι δορυφορικές εικόνες πάσχουν επίσης από γεωμετρικές παραμορφώσεις λόγω παραγόντων όπως η πανοραμική παραμόρφωση, η καμπυλότητα της γης, η διάθλαση λόγω της ατμόσφαιρας, μετατοπίσεις λόγω του αναγλύφου και η μή γραμμικότητα του πεδίου θέασης των αισθητήρων.

Τα λάθη αυτά μπορεί να είναι συστηματικά ή τυχαία. Η γεωμετρική διόρθωση των συστηματικών λαθών περιλαμβάνει συνήθως τη μοντελοποίηση των πηγών λάθους με τη βοήθεια μαθηματικών μοντέλων και στη συνέχεια την εφαρμογή αυτών των μοντέλων στα δεδομένα της εικόνας υπό διόρθωση.

Η διόρθωση των τυχαίων παραμορφώσεων βασίζεται στη μέτρηση της μετατόπισης γνωστών σημείων ελέγχου (ground control points - GCP), δηλαδή διακριτικών γεωγραφικών χαρακτηριστικών με γνωστή θέση πάνω στην εικόνα αναφοράς. Η θέση των σημείων αυτών προσδιορίζεται από την τομή ποταμών, αυτοκινητοδρόμων, αεροδρομίων, χειμάρων ή από μικρές λίμνες, μικρά νησιά κλπ. Στη συνέχεια, η εικόνα του ανιχνευτή δειγματοληπτείται ξανά βάσει αυτής της μετατόπισης, οδηγώντας στη νέα, διορθωμένη εικόνα.

Ειδικότερα, με τον όρο «μετατόπιση» εννοούμε τη σχέση f μεταξύ των συντεταγμένων (X, Y) της εικόνας του αισθητήρα και των δεδομένων ελέγχου με γνωστό γεωγραφικό πλάτος και μήκος (x, y) , αντίστοιχα. Η σχέση αυτή προσδιορίζεται συνήθως με την εφαρμογή τεχνικών στρέβλωσης (warping), με βάση τις εξισώσεις:

$$\begin{aligned}x' &= f_1(X, Y) \\y' &= f_2(X, Y)\end{aligned}$$

Η έξοδος των μεθόδων είναι το γεωμετρικά σωστό πλέγμα (grid) συντεταγμένων για τη ληφθείσα εικόνα. Η αντιστοίχιση της πληροφορίας των pixel της ληφθείσας εικόνας στο νέο πλέγμα γίνεται με την απλή μέθοδο παρεμβολής (interpolation) του πλησιέστερου γειτονικού pixel (nearest neighboring pixel) ή με τις πιο «λεπτομερείς» μεθόδους παρεμβολής βάσει διγραμμικών (bilinear) ή κυβικών (cubic) splines.

Σε κάθε περίπτωση, οι μέθοδοι αυτές μπορεί να μεταβάλλουν τη ραδιομετρική πιστότητα και πρέπει να εφαρμόζονται με ιδιαίτερη προσοχή. Για μεγαλύτερη ακρίβεια, τα πιο πρόσφατα συστήματα χρησιμοποιούν ταυτόχρονα το μοντέλο ephemeris και τη διαδικασία αναγωγής βάσει γνωστών σημείων ελέγχου.

Επίσης, είναι γνωστό το πρόβλημα των δεδομένων που προκύπτουν από αισθητήρες οι οποίοι μεταφέρονται από δορυφόρους που διαγράφουν πολική τροχιά.

Στην περίπτωση αυτή τα δεδομένα είναι σειρές από «λωρίδες» (strips) οι οποίες αφού πρώτα ταυτοποιηθούν μεταξύ τους και διορθωθούν γεωμετρικά, συνδυάζονται στη συνέχεια σε διαδοχικές εικόνες.

Αξίζει να σημειωθεί ότι η απώλεια χωρικής ανάλυσης (spatial resolution) που μπορεί να εμφανισθεί κατά την ταυτοποίηση ή την αναγωγή των εικόνων, δεν αποτελεί πάντα πρόβλημα στις εφαρμογές της τηλεπισκόπησης. Αντίθετα, σε μερικές περιπτώσεις αρκεί μια «χονδρική» χωρική περιγραφή των δεδομένων, σε ανάλυση πολύ μικρότερη από αυτή του pixel.

Τέλος, σε μερικές περιπτώσεις, η τελική ταυτοποίηση επιτυγχάνεται με την επικάλυψη (overlay) της επεξεργασμένης εικόνας με κάθε έναν από τους χάρτες αναφοράς.

ΚΕΦΑΛΑΙΟ 3^ο

Ποιοτική Ανάλυση Εικόνων

Η ταχεία ανάπτυξη στη συλλογή τηλεσκοπικών δεδομένων τις δύο τελευταίες δεκαετίες, έχει οδηγήσει στη ύπαρξη πολλαπλών πηγών δεδομένων εικόνων, με ολοένα αυξανόμενη ανάλυση και αριθμό διαστάσεων χωρικής, χρονικής, φασματικής και ακτινοβόλας πληροφορίας.

Για την ποιοτική ανάλυση ψηφιακών εικόνων τηλεπισκόπησης, χρησιμοποιούνται κυρίως τεχνικές κατάταξης (classification) και τμηματοποίησης (segmentation) από τον τομέα της αναγνώρισης προτύπων (pattern recognition), της επιστήμης υπολογιστών.

Η ανάλυση βασίζεται τόσο σε τεχνικές που έχουν αναπτυχθεί και εφαρμοσθεί με επιτυχία από τις αρχές της δεκαετίας του 1970, όσο και σε νεότερες καινοτόμες ιδέες, οι οποίες βρίσκονται σε πολλές περιπτώσεις στο στάδιο έρευνας ακόμη, με ενθαρρυντικά αποτελέσματα στους τομείς που έχουν εφαρμοσθεί.

3.1 Κατάταξη ή Ταξινόμηση Ψηφιακών Δεδομένων Εικόνων

Με τον όρο «κατάταξη» εννοούμε την ταξινόμηση των pixel της εικόνας σε κλάσεις ως προς κάποια χαρακτηριστικά όπως τιμές φάσματος, το χρώμα ή η υφή (texture).

Η πολυφασματική κατάταξη (multispectral classification) είναι μια από τις ευρέως χρησιμοποιούμενες τεχνικές για την εξαγωγή χρήσιμης πληροφορίας. Θεωρώντας την οργάνωση των δεδομένων για μια γεωγραφική περιοχή σε πολλαπλές ζώνες (bands) ή επίπεδα (layers) με καλή γεωμετρική ταυτοποίηση, κάθε pixel χαρακτηρίζεται

από ένα διάνυσμα τιμών, μία τιμή για κάθε ζώνη ή επίπεδο δεδομένων. Υπό αυτή την έννοια, κάθε pixel αντιστοιχίζεται μέσω του διανύσματος χαρακτηριστικών του σε ένα σημείο του πολυδιάστατου χώρου χαρακτηριστικών (attribute space). Οι άξονες του χώρου αυτού αντιστοιχούν σε κάθε μια από τις ζώνες της πολυεπίπεδης οργάνωσης των δεδομένων.

Κάθε ομάδα από pixel, με αναφορά στον πολυδιάστατο χώρο χαρακτηριστικών, ονομάζεται συλλογή (cluster). Δύο pixel ανήκουν στην ίδια συλλογή όταν τα διανύσματά τους ή κάποιες από τις τιμές των διανυσμάτων τους είναι «παρόμοια». Κάθε κλάση αντιστοιχεί σε μια ομάδα από pixel. Στην

περίπτωση πολυφασματικών εικόνων τηλεπισκόπησης οι κλάσεις περιγράφουν δάσος, νερό, εκτάσεις καλλιέργειας δημητριακών κλπ.

Οι προσεγγίσεις κατάταξης διακρίνονται σε εποπτευόμενες (supervised) και μη εποπτευόμενες (unsupervised). Κατά την εποπτευόμενη κατάταξη, οι κλάσεις ή ετικέτες (labels) είναι γνωστές εκ των προτέρων.

Επιπλέον είναι γνωστές κάποιες αντιπροσωπευτικές θέσεις της μελετούμενης περίπτωσης (case study) για κάθε κλάση. Οι θέσεις αυτές αποτελούν τα δεδομένα εκπαίδευσης (training data) και πρέπει να επιλέγονται από τον αναλυτή ώστε και να περιγράφουν ικανοποιητικά τα χαρακτηριστικά κάθε κλάσης και να διακρίνουν την κάθε κλάση από τις υπόλοιπες. Ο σκοπός σε αυτή την περίπτωση είναι να ανατεθεί σεκάθε μονάδα χώρου μια από τις ετικέτες με βάση κάποιον αλγόριθμο κατάταξης. Γνωστοί αλγόριθμοι εποπτευόμενης κατάταξης είναι :

- ο αλγόριθμος κατάταξης ελάχιστης απόστασης (minimum distance),
- ο αλγόριθμος κατάταξης του εγγύτερου γείτονα (nearest neighbor),
- αλγόριθμοι βασισμένοι σε νευρωνικά δίκτυα και
- ο αλγόριθμος μέγιστης πιθανοφάνειας (maximum likelihood).

Ο τελευταίος ανήκει στους παραμετρικούς αλγορίθμους κατάταξης, αφού, σε σχέση με τους άλλους αλγορίθμους, προϋποθέτει ότι είναι γνωστή η στατιστική κατανομή των

δεδομένων κάθε κλάσης.

Κατά την μη εποπτευόμενη κατάταξη, δεν έχουμε καμία γνώση εκ των προτέρων. Στην συγκεκριμένη περίπτωση, πρέπει ο αναλυτής να ταξινομήσει τα στοιχεία του με οποιοδήποτε φυσικό χαρακτήρα επιθυμεί, όπως για παράδειγμα με βάση το χρώμα, το μέγεθος, το σχήμα, κλπ. Γνωστοί αλγόριθμοι μη-εποπτευόμενης κατάταξης είναι :

- Αλγόριθμοι Ιεραρχικής Συσταδοποίησης (Hierarchical Clustering Algorithm)
- Αλγόριθμοι Διαιρετικής Συσταδοποίησης (Partitional Algorithm)

Αποτελεί επίσης μια αποτελεσματική προσέγγιση διαμέρισης δεδομένων του πολυφασματικού χώρου χαρακτηριστικών.

Σε αντίθεση με την εποπτευόμενη κατάταξη, απαιτείται ελάχιστη καθοδήγηση από τον αναλυτή της μελετούμενης περίπτωσης διότι δεν είναι αναγκαία η συλλογή και επεξεργασία δεδομένων εκπαίδευσης.

Η ομαδοποίηση των πολυφασματικών δεδομένων σε συλλογές γίνεται από γνωστούς υπολογιστικούς αλγορίθμους, βάσει στατιστικών κριτηρίων.

Ειδικότερα, οι αλγόριθμοι αυτοί βασίζονται σε επαναληπτικές, αριθμητικές μεθόδους για τη διερεύνηση της φυσικής ομαδοποίησης των χαρακτηριστικών των pixel.

Μετά την ομαδοποίηση ακολουθεί ο μετασχηματισμός των φασματικών ομάδων στις θεματικές κλάσεις πληροφορίας της μελέτης (δάσος, γεωργικές καλλιέργειες κλπ.).

Πάντως, κοινό γνώρισμα και των δύο προσεγγίσεων είναι ότι ακολουθείται η λογική της ακριβούς ή σαφούς κατάταξης (hard classification), ώστε να προκύψει ένας χάρτης αντιστοίχισης της κάθε μονάδας χώρου σε ακριβώς μία, διακριτή κατηγορία χαρακτηριστικών.

Στον αντίποδα, έχουν αναπτυχθεί μέθοδοι μη ακριβούς ή ασαφούς λογικής (fuzzy classification) οι

οποίες λαμβάνουν υπόψη την ανομοιογένεια και ασάφεια που παρατηρείται στον πραγματικό κόσμο.

Για την κατάταξη έχουν επίσης χρησιμοποιηθεί δένδρα απόφασης (decision trees) που βασίζονται σε κανόνες (rule based classification).

Για τη διατύπωση των κανόνων κατάταξης χρησιμοποιούνται τεχνικές εξόρυξης γνώσης από δεδομένα (data mining).

Στόχος των τεχνικών αυτών είναι η περιγραφή της εξειδικευμένης γνώσης, που προϋπάρχει για την υπό μελέτη περίπτωση, από λειτουργικούς κανόνες. Η προσέγγιση αυτή εφαρμόζεται στην κατάταξη τόσο δεδομένων που παίρνουν τιμές σε κάποιο διάστημα των πραγματικών αριθμών, όσο και σε προεπεξεργασμένα δεδομένα με ονομαστικές ή διακριτές ακέραιες τιμές.

Μια τελευταία, σημαντική διάκριση των μεθόδων κατάταξης προκύπτει από το αν αυτές αναφέρονται στην ανάθεση ετικέτας σε επίπεδο pixel ή αντικειμένων (objects). Η συνήθης πρακτική στην κατάταξη ψηφιακών εικόνων είναι η ταξινόμηση να βασίζεται στην επεξεργασία όλης της εικόνας, pixel προς pixel.

Αντίθετα, οι τεχνικές κατάταξης που βασίζονται σε αντικείμενα, επιτρέπουν στον αναλυτή να αποσυνθέσει την εικόνα σε πολλά, σχετικά ομοιογενή αντικείμενα, τα οποία αναφέρονται στη βιβλιογραφία ως τμήματα (segments) ή τεμάχια (patches), χρησιμοποιώντας κάποια από τις διαδικασίες τμηματοποίησης της εικόνας οι οποίες περιγράφονται στην επόμενη υποενότητα.

Σε αυτή την περίπτωση, αντικείμενο των παραδοσιακών μεθόδων σαφούς ή ασαφούς κατάταξης είναι η ταξινόμηση των αντικειμένων, αντί των pixel, με βάση τα στατιστικά χαρακτηριστικά τους.

Η κατάταξη βάσει αντικειμένων εφαρμόζεται κυρίως σε δεδομένα μεσαίας και μεγάλης χωρικής ανάλυσης, τα οποία είναι γνωστό ότι χαρακτηρίζονται από χωρικό πλεονασμό (spatial redundancy), λόγω της τοπικής ομοιογένειας τμημάτων της εικόνας.

Η χρήση αυτής της νέας προσέγγισης ενισχύθηκε από τη διαπίστωση ότι η απόδοση των μεθόδων κατάταξης των pixel δεν αυξάνεται σε ικανοποιητικό βαθμό καθώς αυξάνεται η χωρική ανάλυση.

Επιπλέον, είναι γνωστό ότι το επίπεδο της απαιτούμενης λεπτομέρειας κατά την αντιστοίχιση των pixel σε κλάσεις, υπαγορεύει και το επίπεδο χωρικής ανάλυσης που πρέπει να χρησιμοποιηθεί.

Με άλλα λόγια, η χωρική ανάλυση των τηλεπισκοπικών εικόνων, καθορίζει σε σημαντικό βαθμό το επίπεδο λεπτομέρειας που μπορεί να επιτευχθεί από τους αλγορίθμους κατάταξης.

Τέλος, ένας ακόμη σημαντικός παράγοντας επιτυχούς κατάταξης είναι η σωστή επιλογή ενός αριθμού από τις διαθέσιμες ζώνες των φασματικών δεδομένων, έχοντας ως κριτήριο τη σχετικότητα τους με τη μελετούμενη περίπτωση.

3.2 Τμηματοποίηση Εικόνων και Video

Πέρα από τη χρήση της στην κατάταξη αντικειμένων, η τμηματοποίηση ή κατάτμηση (segmentation) εικόνων αποτελεί αυτόνομη διεργασία ανάλυσης πολυμεσικών δεδομένων εικόνων και video με πληθώρα εφαρμογών στην επεξεργασία και μετάδοση των δεδομένων αυτών.

Η χρησιμότητά της έγκειται κυρίως στο γεγονός ότι μειώνει τον χωροχρονικό πλεονασμό (spatio-temporal redundancy) που διαπιστωμένα χαρακτηρίζει τα οπτικά δεδομένα.

Η μείωση του πλεονασμού οδηγεί σε μείωση του υπολογιστικού χρόνου επεξεργασίας και, με κατάλληλη κωδικοποίηση, του χρόνου μετάδοσης εικόνων/video.

Με τον όρο «τμηματοποίηση» εννοούμε την αποσύνθεση της εικόνας σε αντικείμενα ενδιαφέροντος.

Τυπικά, ως αντικείμενα ορίζονται συνεκτικές περιοχές pixel στο χώρο συντεταγμένων της εικόνας οι οποίες χαρακτηρίζονται από ομοιογένεια ως προς κάποια γνωρίσματα. Η καταλληλότητα των γνωρισμάτων κρίνεται από το κατά πόσο είναι ικανά να

διακρίνουν τα αντικείμενα από την υπόλοιπη εικόνα.

Οπτικά γνωρίσματα (visual cues) που έχουν χρησιμοποιηθεί με επιτυχία στην τμηματοποίηση video είναι για παράδειγμα η κίνηση αντικειμένων και η ανίχνευση αλλαγών (change detection).

Για την τμηματοποίηση στατικών (still) εικόνων έχουν χρησιμοποιηθεί χαρακτηριστικά όπως οι ακμές (edges) και η ομοιογένεια περιοχών ως προς την ένταση ή φωτεινότητα (intensity), το χρώμα και την υφή (texture).

Σε πολλές περιπτώσεις τα χαρακτηριστικά αυτά συνδυάζονται για καλύτερα αποτελέσματα.

Επίσης, αξίζει να σημειωθεί ότι η τμηματοποίηση μπορεί να εφαρμοσθεί σε οποιαδήποτε χωρικά κατανεμημένη μεταβλητή, έστω κι αν η κατανομή της στο χώρο δεν έχει τη μορφή του πλέγματος (grid) των εικόνων.

Οι κυριότεροι αλγόριθμοι τμηματοποίησης τηλεπισκοπικών εικόνων έχουν τη βάση τους στον τομέα μηχανικής όρασης και ρομποτικής της επιστήμης υπολογιστών.

Γνωστοί αλγόριθμοι τμηματοποίησης που έχουν χρησιμοποιηθεί σε τηλεπισκοπικά δεδομένα είναι:

- αλγόριθμοι που βασίζονται σε κατώφλια,
- αλγόριθμοι επέκτασης περιοχών (region growing),
- αλγόριθμοι πλημμυρίδας (flooding),
- τεχνικές που βασίζονται στα ισοσταθμικά σύνολα (level sets),
- αλγόριθμοι ενεργών περιγραμμάτων (active contours).

Κύριο χαρακτηριστικό των αλγορίθμων τμηματοποίησης είναι ότι συνδυάζουν τη φασματική με τη χωρική πληροφορία.

Στις εικόνες τηλεπισκόπησης, το αποτέλεσμα της τμηματοποίησης είναι διακριτές περιοχές που χαρακτηρίζονται από ομοιογένεια ως προς το φάσμα και το σχήμα τους και αποτελούν τμήματα ή τεμάχια του απεικονιζόμενου τοπίου.

Σε πολλές περιπτώσεις, η εξαγωγή αντικειμένων οδηγεί σε μεγαλύτερο αριθμό χαρακτηριστικών που περιγράφουν καλύτερα την υπό μελέτη διεργασία και οδηγούν σε καλύτερα αποτελέσματα κατάταξης των τμημάτων σε κλάσεις.

Από την άλλη πλευρά, έχουν αναπτυχθεί μέθοδοι όπου η κατάταξη των pixels σε κλάσεις προηγείται ή συνδυάζεται με το αποτέλεσμα της τμηματοποίησης βάσει κυρίως της στατιστικής μοντελοποίησης κατά Bayes, με πολύ καλά αποτελέσματα.

ΚΕΦΑΛΑΙΟ 4^ο

JPEG Συμπίεση

Οι μέθοδοι συμπίεσης δεδομένων με μηδενική απώλεια πληροφοριών έχουν χρησιμοποιηθεί στα δεδομένα εικόνας για κάποιο χρονικό διάστημα. Στην πραγματικότητα, η δημοφιλής GIF μορφή χρησιμοποιεί ένα σύστημα LZW (η βασική μέθοδος που χρησιμοποιείται στο UNIX) για να συμπίεσει 256-χρωμάτων εικόνες. Η PNG μορφή είναι πιο εξελιγμένη και ικανή, χρησιμοποιώντας ένα προγνωστικό (ή φίλτρο) προετοιμάζει τα δεδομένα για ένα συμπίεστη τύπου gzip. Ωστόσο, οι εφαρμογές χρησιμοποιώντας υψηλής ανάλυσης εικόνες με χιλιάδες χρώματα μπορεί να χρειαστούν περισσότερη συμπίεση από ότι μπορεί να επιτευχθεί με αυτές τις μεθόδους χωρίς απώλειες.

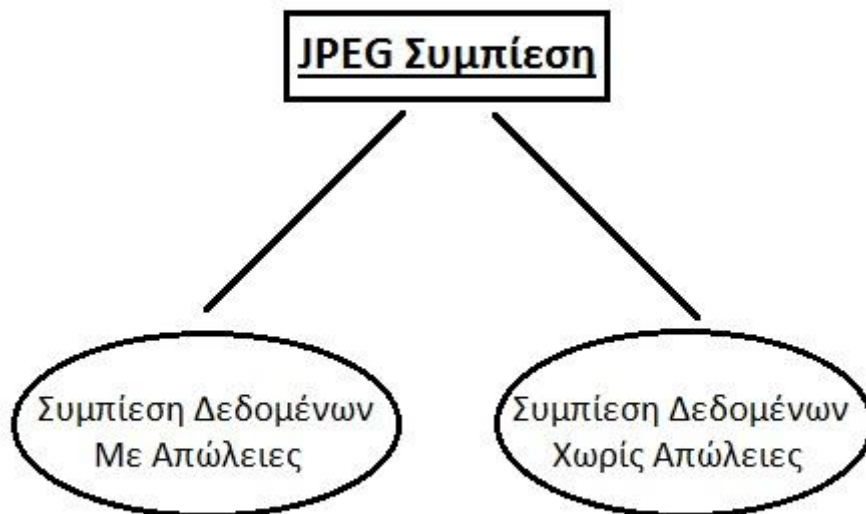
Τα συστήματα με απώλειες απορρίπτουν ορισμένα από τα δεδομένα προκειμένου να επιτύχουν καλύτερη συμπίεση. Το πρόβλημα, βέβαια, είναι να αποφασιστεί ακριβώς ποιές πληροφορίες πρέπει να τεθούν σε κίνδυνο. Απώλειες πληροφοριών σε συμπίεση κειμένου είναι τυπικά απαράδεκτο, αν και απλά συστήματα όπως η εξάλειψη κάθε φωνήεν από αγγλικό κείμενο μπορεί να τύχει κάπου. Η κατάσταση είναι διαφορετική με τις εικόνες και τον ήχο, σε αυτές τις περιπτώσεις, κάποιες απώλειες δεδομένων μπορεί να είναι αρκετά αποδεκτές, ακόμη και ανεπαίσθητες.

Στη δεκαετία του 1980, η Joint Photographic Experts Group (JPEG) ιδρύθηκε για να αναπτύξει πρότυπα για τη συμπίεση ακίνητης εικόνας. Ο ορισμός περιλαμβάνει τόσο

χωρίς απώλειες όσο και με απώλειες τρόπους, αν και ο τελευταίος είναι ίσως ο πιο ενδιαφέρον (και συνήθως είναι αυτός που εννοείται όταν λέμε jpeg συμπίεση).

Η μέθοδος jpeg με απώλειες εξαρτάται από ένα σημαντικό μαθηματικό και φυσικό θέμα για συμπίεση, την τοπική προσέγγιση. Η jpeg ομάδα πήρε αυτή την ιδέα και τελειοποίησε με αποτελέσματα που προέκυψαν από τις μελέτες για το ανθρώπινο οπτικό σύστημα. Το προκύπτον σύστημα βρίσκεται σε ευρεία χρήση, εν μέρει επειδή είναι ένα ανοιχτό πρότυπο αλλά κυρίως γιατί λειτουργεί καλά σε μία μεγάλη κατηγορία εικόνων, με αρκετά μέτριες απαιτήσεις πόρων. Τα jpeg και wavelet συστήματα εμπίπτουν στην γενική κατηγορία των μεθόδων μετασχηματισμού. Η ανάπτυξη των wavelet τεχνικών έχει λάβει χώρα πιο πρόσφατα από την κλασική μέθοδο σε jpeg, και είναι μία συνέπεια της ατέρμονης αναζήτησης για “καλύτερες” βασικές εικόνες.

Σε γενικές γραμμές, το πρώτο βήμα σε συστήματα συμπίεσης με απώλειες, όπως jpeg και wavelets είναι να διαχωρίσουν μια εικόνα σε μία σταθμισμένη ακολουθία απλούστερων, πιο βασικών εικόνων. Σε αυτό το στάδιο, η εικόνα μπορεί να ανακατασκευαστεί ακριβώς από τη γνώση των βασικών εικόνων και των αντίστοιχων φορτίων τους. Η αποτελεσματικότητα της μεθόδου εξαρτάται σε μεγάλο βαθμό από την επιλογή των βασικών εικόνων. Μόλις ένα σύνολο βασικών εικόνων, ή βάση έχει επιλεγεί, αυθαίρετες εικόνες μπορούν να αντικατασταθούν από ισοδύναμες συλλογές φορτίων. Μία βασική εικόνα που έχει ένα αντίστοιχα μεγάλο φορτίο αποτελεί ένδειξη της χαρακτηριστικής σημασίας του στη συνολική εικόνα. (Η παραδοχή εδώ είναι ότι οι βασικές εικόνες έχουν εξομαλυνθεί, έτσι ώστε να έχουν το ίδιο μαθηματικό μέγεθος.)



Τα μαθηματικά πίσω από αυτή τη διαδικασία εκφράζονται στη γλώσσα της γραμμικής άλγεβρας. Υπάρχει αρκετή μαθηματική ελευθερία στην επιλογή των βασικών εικόνων, ωστόσο, στην πράξη είναι συνήθως επιλεγμένο να εμφανίζει χαρακτηριστικά εγγενή στην κατηγορία των εικόνων του ενδιαφέροντος. Για παράδειγμα, το jpeg επιλέγει βασικές εικόνες κατά παραγγελία για να αντανakλούν ορισμένες κλασσικές χωρικές συχνότητες.

Η διαδικασία της χρήσης μιας βάσης για την επίλυση μιας εικόνας σε μία συλλογή φορτίων, ονομάζεται μετασχηματισμός. Για να απλοποιήσουμε τα πράγματα, θα εξετάσουμε τις εικόνες σε κλίμακα του γκρι, οι οποίες μπορούν να παρασταθούν ως $m \times n$ πίνακες ακεραίων. Το εύρος των τιμών δεν είναι σημαντικό για την κατανόηση της μαθηματικής ιδέας, αν και είναι σύνηθες να περιορίσεις τις τιμές στο διάστημα $[0,255]$, δίνοντας συνολικά 256 επίπεδα του γκρι.

Από μαθηματική άποψη, οποιαδήποτε βάση για το χώρο των $m \times n$ εικόνων σε κλίμακα του γκρι πρέπει να περιέχει ακριβώς $m \times n$ εικόνες - ο αριθμός των pixels σε μία $m \times n$ εικόνα. Κατά συνέπεια, ο μετασχηματισμός μιας $m \times n$ εικόνας θα έχει mn φορτίο. Τα

φορτία μπορούν να ταξινομηθούν εύκολα σε μία $m \times n$ σειρά που ονομάζεται μετασχηματισμένη εικόνα, παρόλο που δεν είναι καθόλου πραγματική εικόνα.

Η διαδικασία μετασχηματισμού, από μόνη της, σε καμιά περίπτωση δεν είναι μία τεχνική συμπίεσης (δεδομένου ότι η μετασχηματισμένη εικόνα είναι το ίδιο μέγεθος με την πρωτότυπη), αλλά μπορεί να οδηγήσει σε μία. Ας υποθέσουμε ότι η βάση των εικόνων μπορεί να επιλέγει έτσι ώστε, για μία ευρεία κατηγορία εικόνων, πολλά από τα φορτία να καταλήξουν να είναι μικρά: για μία συγκεκριμένη εικόνα, ορίστε αυτά τα μικρά φορτία ως 0 και χρησιμοποιήστε την προκύπτουσα σειρά από τροποποιημένα φορτία για να την εκπροσωπήσει. Δεδομένου ότι ο μετασχηματισμός της εικόνας έχει τροποποιηθεί, μπορεί να χρησιμοποιηθεί μόνο για την προσέγγιση της αρχικής.



Πόσο καλή είναι η προσέγγιση; Αυτό εξαρτάται από το πόσο καλό είναι το σύστημα για την απόρριψη των μη-μηδενικών φορτίων, δηλαδή, για την καταλληλότητα των στοιχείων βάσης και τον αριθμό των φορτίων που μπορεί να απορριφθεί.

Οι jpeg και wavelet μέθοδοι, χρησιμοποιούν αυτό τον τύπο διαδικασίας και προσφέρουν σημαντικά οφέλη συμπίεσης, συχνά με ελάχιστη επίπτωση στην ποιότητα του αντίγραφου.

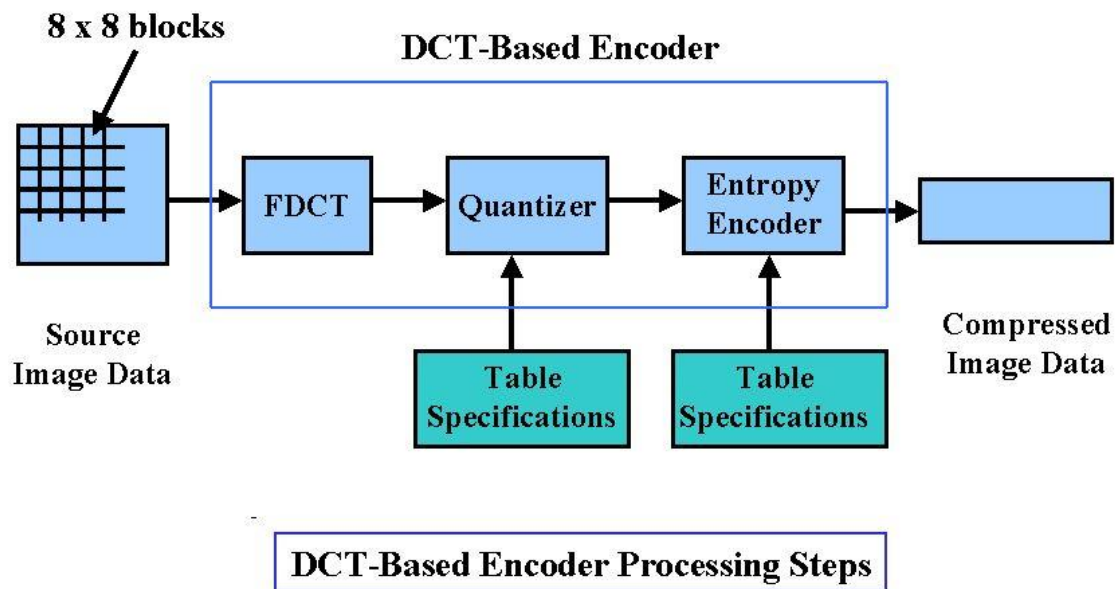
Διαφέρουν στην επιλογή των βάσεων εικόνων, δηλαδή, στο μετασχηματισμό που χρησιμοποιείται και στη συνέχεια στη μέθοδο που χρησιμοποιείται για να απορρίψει τα μικρά φορτία. Ωστόσο, και οι δύο μοιράζονται την ιδέα του να παίρνουν μία βάση που μπορεί να αντιπροσωπεύει αποτελεσματικά μία εικόνα, χρησιμοποιώντας συχνά μόνο ένα μικρό αριθμό των βασικών εικόνων της.

4.1 Αλγόριθμος JPEG

Ο αλγόριθμος του jpeg είναι σχετικά απλός και μπορεί να εξηγηθεί μέσα από τα ακόλουθα βήματα:

- 1) Παίρνουμε μία εικόνα και τη χωρίζουμε σε 8×8 blocks. Εάν η εικόνα δεν μπορεί να διαιρεθεί σε 8×8 blocks, τότε μπορούμε να προσθέσουμε άδεια pixels στα άκρα, ουσιαστικά γεμίζουμε μηδενικά την εικόνα.
- 2) Για κάθε 8×8 block, παίρνουμε τα δεδομένα της εικόνας έτσι ώστε να έχουμε τιμές να αντιπροσωπεύουν το χρώμα σε κάθε pixel.
- 3) Παίρνουμε το διακριτό μετασχηματισμό συνημιτόνου (DCT) για κάθε 8×8 block.
- 4) Μετά τη λήψη του DCT ενός block, πολλαπλασιάζουμε το block με μία μάσκα που θα μηδενίσει ορισμένες τιμές από τον DCT πίνακα.

- 5) Τέλος, για να πάρουμε τα δεδομένα για την συμπιεσμένη εικόνα, παίρνουμε το αντίστροφο DCT του κάθε block. Όλα αυτά τα blocks συνδυάζονται πίσω σε μία εικόνα του ίδιου μεγέθους με την πρωτότυπη.

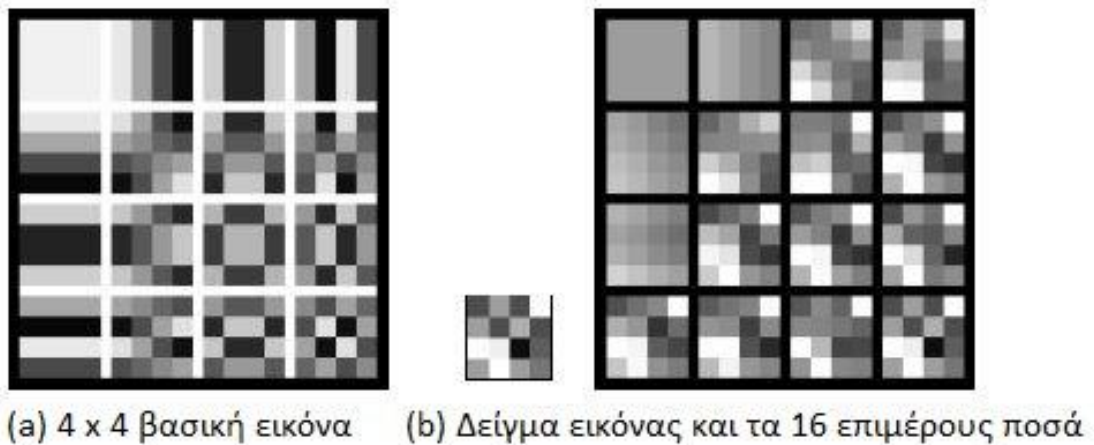


Δεδομένου ότι ενδέχεται να είναι ασαφές γιατί αυτά τα βήματα έχουν σαν αποτέλεσμα μία συμπιεσμένη εικόνα, θα εξηγήσω τώρα τα μαθηματικά και τη λογική πίσω από τον αλγόριθμο.

4.2 Διακριτός Μετασχηματισμός Συνημίτονου (DCT)

Σε αυτή την ενότητα, παρουσιάζονται διάφορα παραδείγματα με το μετασχηματισμό συνημίτονου. Αυτός ο μετασχηματισμός χρησιμοποιείται από την jpeg, εφαρμόζεται σε 8 x 8 τμήματα μιας εικόνας. Υπάρχει ένας μετασχηματισμός συνημίτονου $N \times N$ για κάθε N , το οποίο ανταλλάσσει χωρική πληροφορία για πληροφορία συχνότητας. Για

την περίπτωση $N = 4$, ένα καθορισμένο 4×4 τμήμα μιας εικόνας μπορεί να γραφτεί ως γραμμικός συνδυασμός των εικόνων με βάση 16, όπως φαίνεται στο Σχήμα 1(a).



Σχήμα 1: Στοιχεία εικόνας για το 2D μετασχηματισμό συνημίτονου ($N = 4$), δείγμα εικόνας, και τα 16 επιμέρους ποσά.

Ο μετασχηματισμός παρέχει τους συντελεστές στον γραμμικό συνδυασμό, επιτρέποντας προσεγγίσεις ή προσαρμογές στην αρχική εικόνα με βάση το περιεχόμενο της συχνότητας. Μία πιθανότητα είναι απλά να εξαλείψει ορισμένες συχνότητες, αποκτώντας ένα είδος επιμέρους ποσού της προσέγγισης. Η εγγενής υπόθεση σε μορφή jpeg, για παράδειγμα, είναι ότι οι πληροφορίες υψηλότερης συχνότητας σε μία εικόνα τείνουν να είναι μικρότερης σημασίας για το μάτι.

Θα χρησιμοποιήσουμε “>” για να υποδηλώσει την προτροπή που εμφανίζεται από το Matlab, αλλά αυτό θα διαφέρει ανάλογα με την πλατφόρμα.

Ορίστε την εικόνα της δοκιμής:

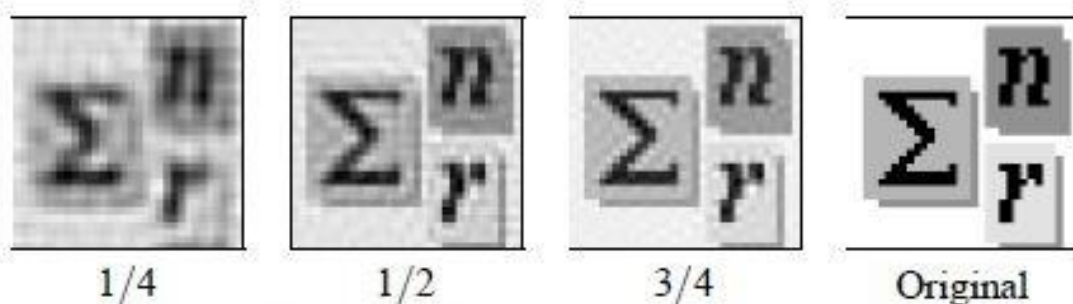
```
> x = round(rand(4)*50) % 4 x 4 τυχαίος πίνακας  
% είσοδος ακεραίου [0,50]
```

Αυτό θα εμφανίσει κάποιο (τυχαίο) πίνακα, ίσως

$$x = \begin{bmatrix} 10 & 20 & 10 & 41 \\ 20 & 10 & 22 & 10 \\ 40 & 30 & 2 & 12 \\ 20 & 35 & 20 & 15 \end{bmatrix}$$

και μπορείτε να δείτε αυτή την “εικόνα” με τις οδηγίες:

> `imagesc(x);`



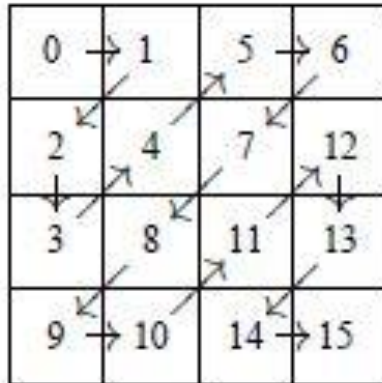
Σχήμα 2: Τα επιμέρους ποσά που συσσωρεύονται στην αρχική εικόνα.

Κάτι παρόμοιο θα επιδειχθεί με την μικρότερη εικόνα στο Σχήμα 1(b). Τώρα ζητήστε τον πίνακα του επιμέρους ποσού (η μεγαλύτερη εικόνα στο Σχήμα 1(b)):

> `imagesc(psumgrid(x)); % προβολή των 16 επιμέρους ποσών`

Τα επιμέρους ποσά έχουν δημιουργηθεί από τα στοιχεία βάσης με τη σειρά που εμφανίζονται στην zigzag ακολουθία. Αυτή η διαδρομή μέσα από το Σχήμα 1(a) βασίζεται στην αυξανόμενη συχνότητα των στοιχείων βάσης. Σε γενικές γραμμές, η

τεχνητή εικόνα στο Σχήμα 1(b) είναι η χειρότερη όσον αφορά την jpeg συμπίεση. Δεδομένου ότι είναι τυχαίο, ενδέχεται να έχει σημαντικά υψηλής συχνότητας όρους.



Αυτό μπορούμε να το δούμε εκτελώντας το διακριτό μετασχηματισμό συνημίτονου (DCT):

> $Tx = \text{dct}(x, 4)$ % 4 x 4 διακριτός μετασχηματισμός συνημίτονου του x

Για το παραπάνω παράδειγμα, μας δίνεται ο πίνακας

$$Tx = \begin{bmatrix} 79.25 & 9.47 & 4.75 & -11.77 \\ -5.92 & -22.36 & 10.42 & 1.52 \\ 6.25 & -19.69 & -4.25 & -11.60 \\ 5.97 & 8.02 & 12.73 & -15.64 \end{bmatrix}$$

των συντελεστών που χρησιμοποιούνται για την κατασκευή των επιμέρους ποσών στο Σχήμα 1 από τα στοιχεία βάσης. Η επάνω αριστερά καταχώρηση παίρνει ειδική αναγνώριση ως ο DC συντελεστής, που αντιπροσωπεύει το μέσο γκρι επίπεδο, οι άλλοι είναι οι AC συντελεστές $AC_{0,1}$ έως $AC_{0,3}$.

Οι όροι στην κάτω δεξιά γωνία του Tx αντιστοιχούν στο υψηλής συχνότητας τμήμα της εικόνας. Σημειώστε ότι ακόμη και σε αυτή την “χειρότερη περίπτωση”, το Σχήμα 1 υποδεικνύει ότι μία αρκετά καλή εικόνα μπορεί να επιτευχθεί χωρίς τη χρήση όλων των 16 όρων.

Η διαδικασία προσέγγισης από τα επιμέρους ποσά εφαρμόζεται σε μία “πραγματική” εικόνα στο Σχήμα 2, όπου το $\frac{1}{4}$, το $\frac{1}{2}$ και τα $\frac{3}{4}$ του όρου 1024 για μία 32 x 32 εικόνα εμφανίζονται. Αυτά μπορούν να παραχθούν με τις εντολές της μορφής:

- > `x = getpgm ('math4.pgm');` % Πάρτε μια graymap εικόνα

- > `n = length (x);` % Το n είναι ο αριθμός των γραμμών
% στην τετράγωνη εικόνα

- > `y = psum (x, n*n / 2);` % Το y είναι το επιμέρους ποσό που χρησιμοποιεί
% το $\frac{1}{2}$ των όρων

- > `imagesc (y);` % Εμφανίζει το αποτέλεσμα

Οι προσεγγίσεις μας διατηρούν όλες τις πληροφορίες συχνότητας που αντιστοιχούν στους όρους από τη zigzag ακολουθία κάτω από μερικές επιλεγμένες οριακές τιμές, οι υπόλοιπες υψηλής συχνότητας πληροφορίες απορρίπτονται. Αν και αυτό μπορεί να θεωρηθεί ως ειδική περίπτωση της jpeg – σαν σύστημα, η jpeg επιτρέπει την πιο εξελιγμένη χρήση της πληροφορίας συχνότητας.

Η jpeg εκμεταλλεύεται την ιδέα της τοπικής προσέγγισης για τη συμπίεση του: 8 x 8 τμήμα της πλήρης εικόνας μετατρέπονται χρησιμοποιώντας τον μετασχηματισμό συνημίτονου, τότε κάθε block κβαντίζεται με μία μέθοδο η οποία τείνει να καταστείλει τα υψηλής συχνότητας στοιχεία και να μειώσει τον αριθμό των bits που απαιτείται για κάθε όρο. Για να “ανακτηθεί” η εικόνα, χρησιμοποιείται ένα βήμα αποκβάντισης ακολουθούμενο από έναν αντίστροφο μετασχηματισμό. (Αγνοήσαμε το τμήμα του jpeg που κάνει συμπίεση χωρίς απώλειες στην “έξοδο” της κβάντισης, αλλά αυτό δεν επηρεάζει την ποιότητα της εικόνας.)

Οι πράξεις του πίνακα μπορούν να απεικονιστούν ως εξής:

$$x \xrightarrow{\text{transform}} Tx \xrightarrow{\text{quantize}} QTx \xrightarrow{\text{dequantize}} Ty \xrightarrow{\text{invert}} y,$$

Στο Matlab, τα επιμέρους βήματα μπορούν να γραφτούν:

- > `x = getpgm ('bird.pgm');` % Πάρτε μια graymap εικόνα
- > `Tx = dct (x);` % Κάνετε τον 8 x 8 μετασχηματισμό συνημίτονου
- > `QTx = quant (Tx);` % Κβαντοποιήστε, χρησιμοποιώντας 8 x 8
% φωτεινότητα
- > `Ty = dequant (QTx);` % Αποκβαντοποιήστε
- > `y = invdct (Ty);` % Ανακτήστε την εικόνα
- > `imagesc (y);` % Εμφανίστε την εικόνα

Για να είμαστε ακριβείς, μια διαδικασία στρογγυλοποίησης πρέπει να γίνει στον πίνακα y . Επιπλέον, έχουμε αγνοήσει τη μηδενική μετατόπιση που καθορίζεται στο πρότυπο, η οποία επηρεάζει τους κβαντισμένους συντελεστές DC.

Θα πρέπει να τονιστεί ότι δεν μπορούμε να ανακτήσουμε την εικόνα πλήρως - υπήρξε απώλεια πληροφοριών κατά το στάδιο κβαντισμού.

Στη σχολή μελετήσαμε τον διακριτό μετασχηματισμό Fourier, ο οποίος είναι ένας τρόπος να αναθέσουμε μία συνάρτηση στη συχνότητα πεδίου ορισμού με δειγματοληψία μιας συνάρτησης σε περιοδικό διάστημα. Το DCT είναι παρόμοιο, εκτός από το ότι χρησιμοποιεί μόνο τη συνάρτηση συνημίτονου, επομένως δεν αλληλεπιδρά με μιγαδικούς αριθμούς. Υπάρχουν μερικές παραλλαγές του DCT, ανάλογα με το αν το όριο του διαστήματος είναι άρτιο ή περιττό, και σε ποιο σημείο η συνάρτηση είναι άρτια.

Το DCT που χρησιμοποιείται συχνότερα, και το οποίο χρησιμοποιούμε εδώ, ονομάζεται DCT-II, το οποίο είναι μέσα από το απομένον όριο σε ένα ενδιάμεσο σημείο μεταξύ του πρώτου σημείου και του σημείου πριν από αυτό. Στα σωστά όρια, έχουμε ακόμη μία επέκταση της συνάρτησης, η οποία δημιουργεί μια συνεχής συνάρτηση σε όλο το όριο. Αυτό προκαλεί στους DCT συντελεστές να εξαλειφθούν ακόμα πιο γρήγορα, το οποίο είναι καλό για την συμπίεση της εικόνας, αφού μας επιτρέπει να προσεγγίσουμε τα δεδομένα της εικόνας με όσο το δυνατόν λιγότερους όρους. Το DCT δεν θα είχε αυτή τη δυνατότητα, δεδομένου ότι είναι ασυνεχής πέρα από τα όρια και επομένως έχει βραδύτερη εξάλειψη των συντελεστών.

Από την εξαγωγή του DCT, θα δείτε ότι οι τιμές του DCT είναι οι μισές από αυτές του DFT (διακριτός μετασχηματισμός Fourier), έτσι ώστε να αναπαριστά τη σχέση ως:

$$f_m = 2 \sum_{j=0}^{N-1} f_j \cos\left(\frac{m\pi}{2}\left(j + \frac{1}{2}\right)\right)$$

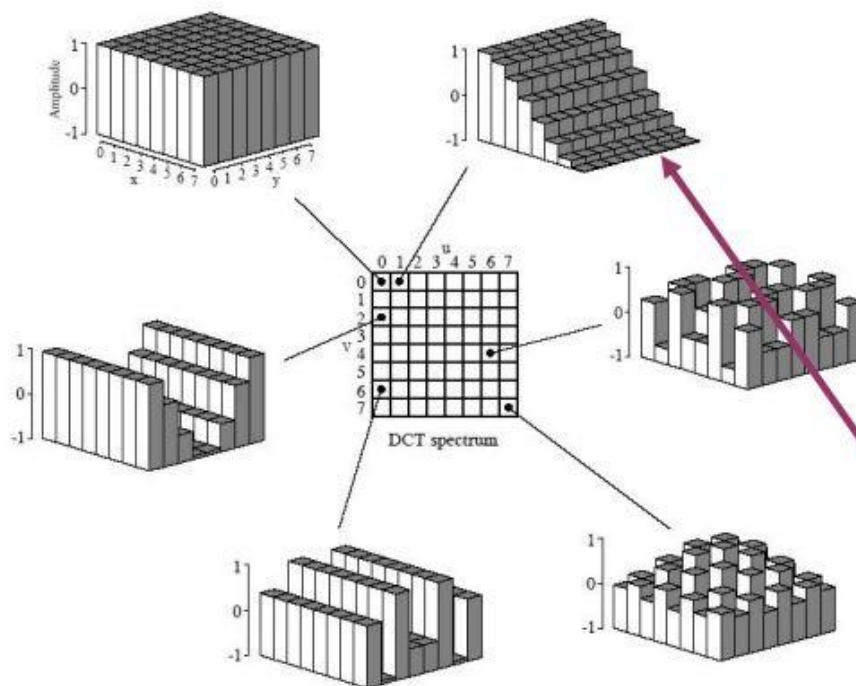
όπου f_m είναι ο DFT συντελεστής και το άθροισμα στα δεξιά είναι οι τιμές του DCT.

Ως εκ τούτου, οι τιμές του DCT είναι απλά:

$$g_m = \sum_{j=0}^{N-1} f_j \cos\left(\frac{m\pi}{2}\left(j + \frac{1}{2}\right)\right)$$

Όταν παίρνουμε το DCT, οι τιμές με την χαμηλότερη συχνότητα θα συγκεντρωθούν γύρω από την επάνω αριστερή γωνία του DCT πίνακα. Για να το καταλάβουμε αυτό, μπορούμε να δούμε τον DFT πίνακα. Στην επάνω αριστερή γωνία, η τιμή είναι ω^0 . Κάτω δεξιά, θα είχαμε ω^{-N^2} , εάν ο πίνακας ήταν $N \times N$.

Δεδομένου ότι με το DCT χρησιμοποιούμε μόνο συνημίτονο, οι συντελεστές στο DCT πίνακα θα έχουν υψηλότερη συχνότητα όταν η είσοδος του όρου συνημίτονου της είναι μεγαλύτερο. Αυτό θα μπει στο παιχνίδι καθώς εργαζόμαστε για να ακυρώσουμε τις υψηλές συχνότητες στην εικόνα.



Στο σχήμα φαίνονται μερικά παραδείγματα συναρτήσεων βάσης για μετασχηματισμό DCT σε μια εικόνα διαστάσεων 8×8 :

- ◊ Ο συντελεστής DCT στη συχνότητα $(u,v)=(1,0)$ μιας εικόνας 8×8 προκύπτει ως το άθροισμα των γινομένων (στοιχείο προς στοιχείο) της συνάρτησης βάσης $(1,0)$ με τα pixel της εικόνας

4.3 Ακύρωση Υψηλών Συχνοτήτων

Όπως πρωτίστως αναφέρθηκε, οι άνθρωποι αδυνατούν να δουν πτυχές μιας εικόνας που είναι σε πολύ υψηλές συχνότητες. Από την λήψη του DCT μας επιτρέπεται να απομονώσουμε το που βρίσκονται αυτές οι υψηλές συχνότητες, μπορούμε να επωφεληθούμε από αυτό επιλέγοντας ποιες τιμές να διατηρήσουμε. Πολλαπλασιάζοντας τον DCT πίνακα με κάποια μάσκα, μπορούμε να μηδενίσουμε στοιχεία του πίνακα, απελευθερώνοντας έτσι την μνήμη που εκπροσωπούσε αυτές τις τιμές.

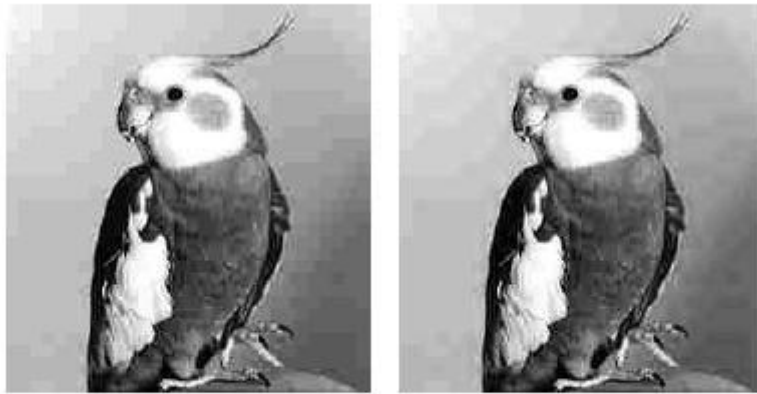
Όταν δουλεύουμε με 8 x 8 block, για παράδειγμα η μάσκα θα μπορούσε να μοιάζει έτσι:

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Αυτός ο πίνακας θα διατηρήσει μόνο εκείνες τις τιμές στις χαμηλότερες συχνότητες μέχρι ένα ορισμένο σημείο, το οποίο είναι ακριβώς αυτό που θέλουμε.

4.4 JPEG Βελτίωση

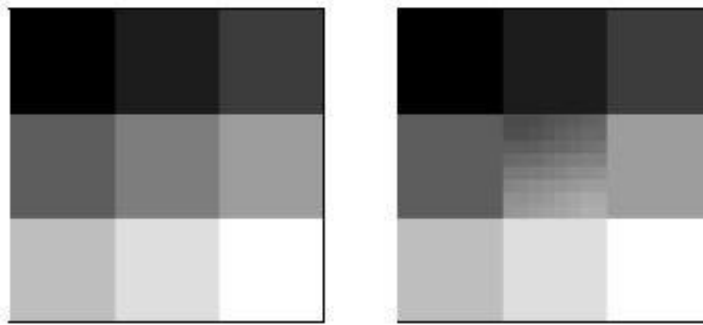
Μία ενοχλητική πτυχή της jpeg, όπως τα συστήματα είναι η εμφάνιση του “κλειδωμένα αντικείμενα”, οι ενδεικτικές ασυνέχειες μεταξύ των blocks οι οποίες συχνά ακολουθούν επιθετικό κβαντισμό. Η εικόνα στο Σχήμα 3 παρήχθη χρησιμοποιώντας ένα κλιμακωτό πολλαπλασιασμό της προτεινόμενης φωτεινότητας κβαντιστή. Ευδιάκριτα blocks φαίνονται, ειδικά στις “ομαλότερες” περιοχές της εικόνας.



Σχήμα 3: " Πτηνό" με επιθετικό κβαντισμό, που στη συνέχεια εξομαλύνεται.

Η jpeg λειτουργεί σε μεμονωμένα 8 x 8 block στην εικόνα και τα επεξεργάζεται ανεξάρτητα. Μπορεί να υπάρξει σημαντική απώλεια της λεπτομερείς πληροφορίας εντός των μεμονωμένων τμημάτων, εάν ο κβαντισμός είναι επιθετικός. Ο μετασχηματισμός συνημίτονου χρησιμοποιείται σε μορφή jpeg έχοντας ιδιότητες οι οποίες μπορούν (έμμεσα) να συμβάλλουν στην ομαλή μετάβαση μεταξύ των γειτονικών blocks, ωστόσο, τα κομμάτια της επεξεργασίας block με block μπορεί να είναι εμφανής όταν τα blocks επανασυναρμολογηθούν και αποκατασταθεί η εικόνα. Σε αυτή την περίπτωση, θα ήταν επιθυμητό να εφαρμοστεί ένα σύστημα εξομάλυνσης ως μέρος της διαδικασίας αποκατάστασης. Αυτή η ενότητα εξετάζει την διαδικασία εξομάλυνσης back-end.

Ο jpeg αποσυμπιεστής μπορεί να έχει μόνο πρόχειρες εκτιμήσεις σχετικά με ένα μεγάλο μέρος των αρχικών πληροφοριών συχνότητας, αλλά τυπικά αρκετά καλές εκτιμήσεις του μέσου επιπέδου του γκρι σε κάθε αρχικό 8 x 8 block (λόγω του τρόπου που επιλέγονται οι κβαντιστές). Η ιδέα είναι να χρησιμοποιήσουμε τη σύνηθες πληροφορία του γκρι (DC-συντελεστής) από τους πλησιέστερους “γείτονές” της για να προσαρμόσει μία συγκεκριμένη πληροφορία συχνότητας του block (AC-συντελεστής).



Σχήμα 4: Το πρωτότυπο και το "εξομαλυμένο" superblock.

Το Σχήμα 4 απεικονίζει τη διαδικασία με ένα μόνο “superblock”, αποτελούμενο από μία κεντρική 8 x 8 εικόνα και τους πλησιέστερους γείτονες της. Το κεντρικό block στην εικόνα στα δεξιά έχει “εξομαλυνθεί” από την επιρροή των πλησιέστερων γειτόνων της (τα γύρω 8 x 8 blocks).

Η διαδικασία σε μία πιο περίπλοκη εικόνα απεικονίζεται στο Σχήμα 5. Εδώ, η εικόνα καταγράφεται ως μία επιφάνεια όπου, σε κάθε pixel (y,x) το ύψος της επιφάνειας του αντιπροσωπεύει την τιμή του γκρι. Για ένα δεδομένο 8 x 8 block, το 3 x 3 “superblock” αποτελείται από τους πλησιέστερους γείτονες περιέχοντας $3^2 \cdot 8^2$ συνολικές καταχωρήσεις. Το πολυώνυμο

$$p(y,x) = a_1 x^2 y^2 + a_2 x^2 y + a_3 x y^2 + a_4 x^2 + a_5 x y + a_6 y^2 + a_7 x + a_8 y + a_9$$

ταιριάζει με την απαίτηση ότι η μέση τιμή πάνω από κάθε υπο-μπλόκ ταιριάζει με τη μέση εκτίμηση του γκρι (αυτό δίνει 9 εξισώσεις για τους αγνώστους $\alpha_1, \dots, \alpha_9$).

Το πολυώνυμο ορίζει μια επιφάνεια πάνω από το κεντρικό block, το οποίο προσεγγίζει το αντίστοιχο τμήμα της αρχικής επιφάνειας. Το Σχήμα 5 δείχνει μία επιφάνεια στο (a) και την πολυωνυμική προσέγγισή της στο (b).

Ο αποσυμπιεστής jpeg μπορεί να εκτελέσει τη διαδικασία μετασχηματισμού σε μία πολυωνυμική προσέγγιση, λαμβάνοντας ένα σύνολο από δείκτες πρόβλεψης για την πληροφορία συχνότητας της αρχικής εικόνας.

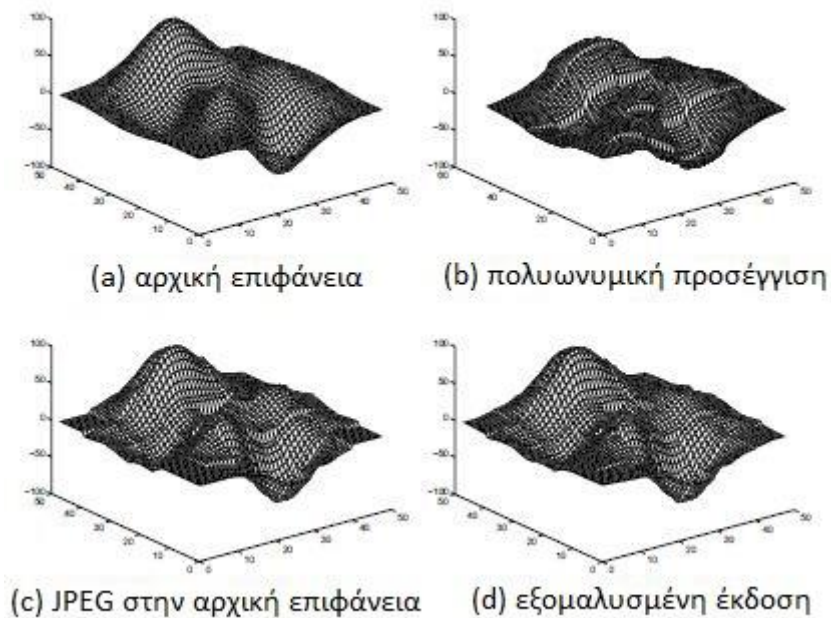
Οι πρωτότυπες εκτιμήσεις που πέρασαν από τον συμπιεστή μπορούν να ρυθμιστούν χρησιμοποιώντας αυτούς τους προγνωστικούς παράγοντες με την ελπίδα της μείωσης του προβλήματος μπλοκαρίσματος.

Στο Σχήμα 5, οι χαμηλότερες πέντε συχνότητες λήφθηκαν για προσαρμογή από τους προγνωστικούς παράγοντες, οι μηδενικές τιμές πέρασαν από το συμπιεστή όπου αντικαταστάθηκαν από τις προβλεπόμενες τιμές (υπόκεινται σε μία ορισμένη σύσφιξη). Η διαδικασία που εφαρμόστηκε σε μία επιθετικά κβαντισμένη εικόνα εμφανίζεται στο Σχήμα 3. Η εντολή `deblock.m` εκτελεί την εξομάλυνση.

Ο ακόλουθος κώδικας χρησιμοποιήθηκε για να δημιουργήσει την δεξιά εικόνα:

- > `x = getpgm ('bird.pgm');` % Πάρτε μια `graymap` εικόνα
- > `Tx = dct (x);` % Κάνετε τον 8 x 8 μετασχηματισμό συνημίτονου
- > `QTx = quant (Tx, 4*stdQ);` % Κβαντοποιήστε, χρησιμοποιώντας 4*φωτεινότητα
- > `Ty = dequant (QTx);` % Αποκβαντοποιήστε
- > `Tz = deblock (Ty);` % Εξομάλυνση
- > `z = invdct (Tz);` % Ανακτήστε την εικόνα
- > `imagesc (z);` % Εμφανίστε την εικόνα

Αυτού του τύπου σύστημα εξομάλυνσης είναι ελκυστικό, εν μέρει της απλότητας του και του γεγονότος ότι μπορεί να χρησιμοποιηθεί ως μία back-end διαδικασία σε jpeg (ανεξάρτητα από το αν το αρχικό αρχείο συμπιέστηκε με αυτό κατά νου). Ωστόσο, η jpeg επιτυγχάνει την αρκετά εντυπωσιακή συμπίεση με την απόρριψη πληροφορίας. Η διαδικασία εξομάλυνσης μερικές φορές κάνει καλές εικασίες σχετικά με τα δεδομένα που λείπουν, αλλά δεν μπορεί να ανακτήσει την αρχική πληροφορία.



Σχήμα 5: Η διαδικασία εξομάλυνσης.

4.5 Εξοικονόμηση Μνήμης

Είναι σημαντικό να καταλάβουμε ακριβώς πώς αυτή η διαδικασία εξοικονομεί μνήμη. Η ιδέα αποτελεί τη βάση από το γεγονός ότι η εκπροσώπηση των αριθμών απαιτεί μνήμη (για την ακρίβεια 1 byte ανά αριθμό).

Όταν κάθε pixel σε μία εικόνα έχει μία τιμή να το εκπροσωπεί, αυτό απαιτεί μνήμη. Μηδενίζοντας τα στοιχεία στο DCT, ωστόσο, απελευθερώνουμε μνήμη. Έτσι, το ποσό της μνήμης που χρησιμοποιείται στην εικόνα μπορεί να προσδιοριστεί με τον υπολογισμό του αριθμού των μη μηδενικών όρων στον DCT πίνακα. Αυτό θα σας δώσει τον αριθμό των bytes που απαιτούνται για την αποθήκευση μιας εικόνας. Για να μετατρέψετε σε kilobites, θα διαιρέσετε την τιμή με το 1024, τον αριθμό των bytes σε kilobyte.

4.6 Γιατί 8 x 8 blocks

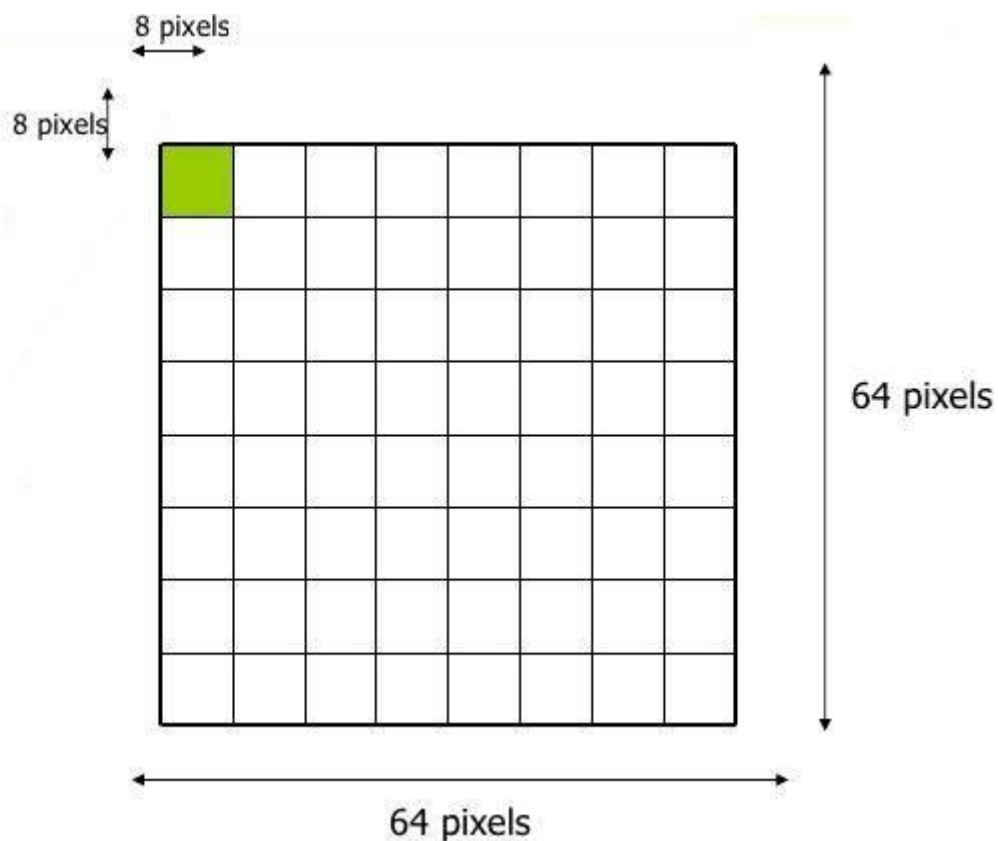
Μία άλλη βασική πτυχή στον jpeg αλγόριθμο είναι η απόφαση να εκτελέσεις αυτές τις λειτουργίες σε 8 x 8 blocks. Οι διαστάσεις αυτές μπορεί να φαίνονται κάπως αυθαίρετες, και αυτό είναι εν μέρει. Ωστόσο, υπάρχουν επίσης μερικοί λόγοι που υποστηρίζουν αυτή την απόφαση.

Πρώτον, αν τα μεγέθη patch ήταν μεγαλύτερα τότε είναι πιθανό ότι η εικόνα θα έχει μεγαλύτερες χρωματικές κλίσεις μεταξύ αυτών των blocks. Είναι χρήσιμο να σκεφτούμε το βήμα συγκάλυψης ως βασικό υπολογίζοντας τον κοινό μέσο όρο των τιμών στο block πρώτου να επιστρέψει πίσω στον θεατή. Αν υπάρχουν μικρές διαφορές σε μια εικόνα μεταξύ των μικρότερων blocks των pixel, αυτή η διαδικασία δεν θα αποδώσει αυτές τις διαφορές καλά, το οποίο θα καταστήσει τα αποτελέσματα της συμπίεσης σας χειρότερα.

Ωστόσο, αυτό φαίνεται να δείχνει ότι ένα ακόμη πιο μικρό μέγεθος block, όπως 4 x 4 ή 2 x 2 θα πρέπει να χρησιμοποιούνται. Αυτό δεν ισχύει επειδή αυτό θα κάνει την συμπίεση πιο περίπλοκη στο να εκτελεστεί. Για κάθε block, θα πρέπει να πάρετε το DCT, πολλαπλασιάστε με την μάσκα, και να πάρετε το αντίστροφο DCT. Με περισσότερα blocks, αυτή η διαδικασία θα διαρκέσει περισσότερο.

Όπως θα δούμε τα πειραματικά αποτελέσματα αργότερα, είναι επίσης δύσκολο να συμπίεσεις μια εικόνα με την ίδια ακρίβεια ενώ επιτυγχάνεις μικρότερα μεγέθη αρχείων όταν χρησιμοποιείς μικρότερα blocks.

Ως εκ τούτου, η προσέγγιση μεγέθους 8 x 8 block έχει αναδειχθεί ως ο κυρίαρχος τρόπος για να χωρίζεις την εικόνα.



ΚΕΦΑΛΑΙΟ 5^ο

Πειράματα

Μια εξέχων διάσταση του αλγορίθμου είναι η επιλογή του μεγέθους block 8 x 8. Επομένως, έτρεξα μερικά πειράματα για να δούμε τον αντίκτυπο αυτής της επιλογής.

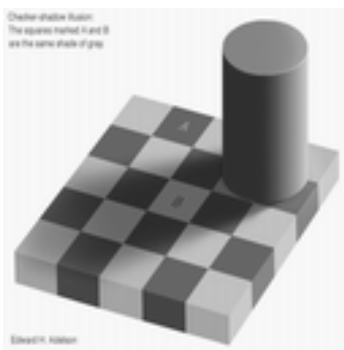
1. Αρχική εικόνα:



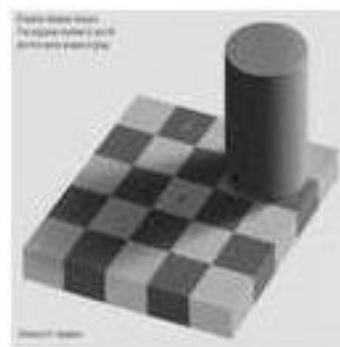
Συμπιεσμένη εικόνα



2. Αρχική εικόνα:



Συμπιεσμένη εικόνα:



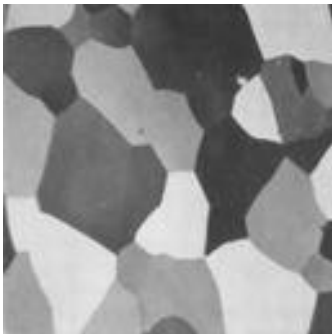
3. Αρχική εικόνα:



Συμπιεσμένη εικόνα:



4. Αρχική εικόνα:



Συμπιεσμένη εικόνα:



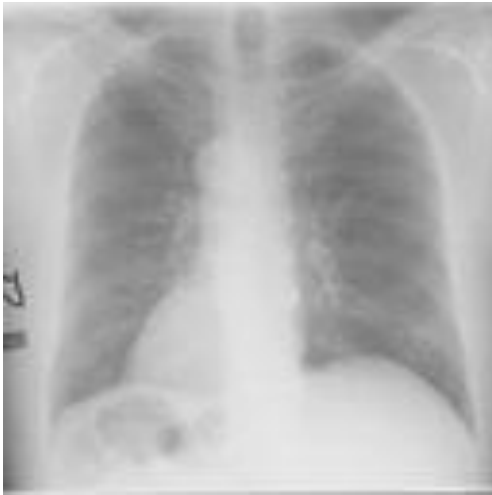
5. Αρχική εικόνα:



Συμπιεσμένη εικόνα:



6. Αρχική εικόνα:



Συμπιεσμένη εικόνα:



7. Αρχική εικόνα:



Συμπιεσμένη εικόνα:



8. Αρχική εικόνα:



Συμπιεσμένη εικόνα:



9. Αρχική εικόνα:



Συμπιεσμένη εικόνα:



ΚΕΦΑΛΑΙΟ 6^ο

Συμπεράσματα Και Μελλοντική Εργασία

Σε αυτή τη πτυχιακή εργασία καταγράψαμε τις τεχνικές επεξεργασίας και ανάλυσης εικόνων που έχουν εφαρμογή στη Γεωπληροφορική.

Η ανάλυση του jpeg επικεντρώθηκε στις ασπρόμαυρες εικόνες. Οι έγχρωμες εικόνες μπορούν να αναθέσουν κόκκινο, πράσινο και μπλέ (R,G,B) τριπλό σε κάθε pixel, μολονότι και άλλες επιλογές είναι δυνατές. Το χρώμα καθορίζεται από άποψη φωτεινότητας, η απόχρωση και ο κορεσμός, γνωστά ως φωτεινότητα - χρωματικές αναπαραστάσεις, μπορεί να είναι επιθυμητά από άποψη συμπίεσης, δεδομένου ότι το ανθρώπινο οπτικό σύστημα είναι πιο ευαίσθητο σε σφάλματα στη φωτεινότητα από ότι στα χρώματα. Λαμβάνοντας υπόψη μια χρωματική αναπαράσταση, το jpeg σύστημα μπορεί να εφαρμοστεί σε κάθε μία από τις τρεις περιπτώσεις.

Αυτή τη πτυχιακή εργασία εξηγεί πώς λειτουργεί ο αλγόριθμος συμπίεσης εικόνας jpeg. Από την εξέταση των πειραμάτων, φαίνεται ότι για κάποιον που γράφει κώδικα για συμπίεση εικόνας jpeg, υπάρχει μία αμφιταλάντευση από την άποψή του πως θα ρυθμίσει το μέγεθος του block και θα επιλέξει τη μάσκα που θα χρησιμοποιήσει. Αν και το τυπικό μέγεθος block, το οποίο είναι το 8 x 8, είναι συνήθως η πιο ακέραιη επιλογή, διαφορετικό μέγεθος block μπορεί να ταιριάζει καλύτερα σε μία "ιδιαίτερη" εικόνα. Μεταβάλλοντας τη μάσκα, η ποιότητα της εικόνας και η απόδοση του μεγέθους αρχείου μπορεί να αλλάξουν εξίσου.

Ο jpeg είναι κατηγορηματικά μία ενδιαφέρουσα επέκταση και ελπίζω ότι αυτή η πτυχιακή εργασία θα σας κεντρίσει το ενδιαφέρον για να μάθετε περισσότερα για τους αλγόριθμους συμπίεσης εικόνας. Καθώς επίσης, να αξιολογήσετε σε επόμενο στάδιο, αν και κατά πόσο οι εξελίξεις στον τομέα της Πληροφορικής που ασχολείται με την επεξεργασία εικόνων εισάγονται και αλληλεπιδρούν με την Γεωπληροφορική.

ΚΕΦΑΛΑΙΟ 7^ο

ΒΙΒΛΙΟΓΡΑΦΙΑ

- Wacker A.G. and Landgrebe D.A., “Minimum Distance Classification in Remote Sensing”, LARS Tech. Report No. 030772, 1972.
- “A Parallel Network of Modified 1-NN and KNN Classifiers: Application to Remote Sensing Image Classification”, Jozwik A., Serpico S. and Roli, F., Pattern Recognition Letters, 1998, pp. 57-62.
- Kanellopoulos I., Wilkinson G.G. and Megier J., “Integration of neural network and statistical image classification for land cover mapping”, Proc. IGARSS 93, 1993, pp. 511-513.
- “A fast maximum likelihood classifier”, Venkateswarlu N.B. and Singh R.P., International Journal of Remote Sensing, Vol. 16, 1995, pp 313-320.
- “Data clustering: A review”, Jain A.K., Murty M.N. and Flynn P.J., ACM Computing Surveys, vol. 31, no. 3, pp. 264-323, 1999.
- Jian Zheng, Zhanzhong Cui, Anfei Liu, Yu Jia, “A K-Means Remote Sensing Image Classification Method Based On AdaBoost”, 4th International Conference on Natural Computation, Vol. 4, 2008, pp. 27-32.

- Akgün A., Eronat A.H. and Turk N., “Comparing Different Satellite Image Classification Methods: An Application In Ayvalık District,Western Turkey”, XXthInternational Congress for Photogrammetry and Remote Sensing, Istanbul,Turkey, 2004.
- “Fuzzy sets in remote sensing classification”, Gomez D. and Montero J., Soft Computing - A Fusion of Foundations, Methodologies and Applications, Vol. 12, 2007, pp. 243-249.
- “Decision Tree Classification of Land Cover From Remotely Sensed Data”, Friedl M.A. and Brodley C.E., Remote Sensing of Environment, 61, 1997, pp. 399-409.
- “Expert System Based on Object-oriented Approach for Land Cover Mapping”, Zhang Lei, Zhou Yueming and Wu Bingfang, The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XXXVII, Part B7, Beijing, 2008.
- “Transform Methods and Image Compression”, Darrel Hankerson and Greg A. Harris
- “JPEG Image Compression”, Matt Marcus
- “Επισκόπηση των Τεχνικών Επεξεργασίας και Ανάλυσης Εικόνων με Εφαρμογή στη Γεωπληροφορική”, Η. Γκρίνιας και Δ. Κοτζίνος , 2010