ΣΥΣΤΗΜΑΤΑ ΠΟΛΥΜΕΣΩΝ ΚΑΙ ΕΙΚΟΝΙΚΗ ΠΡΑΓΜΑΤΙΚΟΤΗΤΑ

ΚΩΔΙΚΟΠΟΙΗΣΗ

ΑΠΟΚΩΔΙΚΟΠΟΙΗΣΗ JPEG



ΟΝΟΜΑ: ΚΩΝΣΤΑΝΤΙΝΟΣ

ΕΠΙΘΕΤΟ: ΛΕΤΡΟΣ

ΣΧΟΛΗ: ΑΡΙΣΤΟΤΕΛΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΟΝΙΚΗΣ

ΤΜΗΜΑ: ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧ. ΚΑΙ ΜΗΧ. ΥΠΟΛΟΓΙΣΤΩΝ

ΑΕΜ: 8851

ΕΤΟΣ: 2019-2020

**Περιεχόμενα**

[1. Εισαγωγή 3](#_Toc27670165)

[2. Τμήμα Α – JPEG Library 3](#_Toc27670166)

# 1. Εισαγωγή

Σκοπός της εργασίας αυτής είναι η δημιουργία ενός κωδικοποιητή-αποκωδικοποιητή ακίνητων εικόνων με απώλειες (Lossy Compression) με βάση το πρότυπο JPEG (ISO/IEC 109181:1994). Συγκεκριμένα, η υλοποίηση που ακολουθεί επιδιώκει να προσεγγίσει την εκδοχή baseline sequential DCTbased.



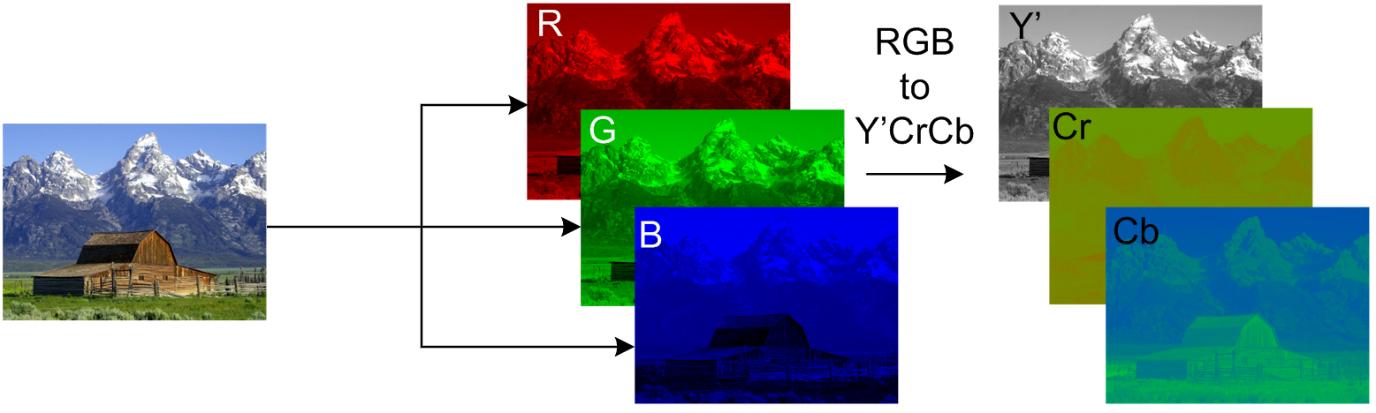
Η παραπάνω διαδικασία διακρίνεται σε 3 τμήματα, εκ των οποίων στην παρούσα εργασία υλοποιούνται τα δύο πρώτα. Παραλείπεται το τμήμα που αφορά την τελική κατασκευή του bitstream, με τη συμμετοχή των διάφορων headers της εικόνας.

# 2. Τμήμα Α – JPEG Library

Σκοπός του τμήματος αυτού είναι η επεξεργασία της προς συμπίεσης εικόνας προκειμένου να επιτευχθεί η κωδικοποίηση-αποκωδικοποίηση. Συγκεκριμένα, στο στάδιο αυτό γίνεται η δημιουργία όλων των συναρτήσεων που θα συμμετάσχουν στην υλοποίηση του προτύπου. Για κάθε συνάρτηση που αφορά τον κωδικοποιητή, υλοποιείται και η αντίστροφή της, η οποία θα βρίσκεται στο αντίστοιχο τμήμα του αποκωδικοποιητή.

## 2.1 Προεπεξεργασία

Συγκεκριμένα, η διαδικασία κωδικοποίησης ξεκινάει με τη μετατροπή της εικόνας από το σύστημα χρωματικών συνιστωσών RGB (Red – Green – Blue) στο σύστημα YCbCr (Luminance – Chrominance).



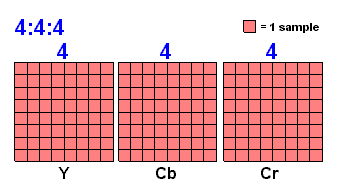
Η εικόνα RGB έχει τη μορφή τρισδιάστατου πίνακα μεγέθους , όπου R, C η οριζόντια και κατακόρυφη διάσταση ενώ η τρίτη διάσταση λαμβάνει τις τιμές 1 (Red), 2 (Green), 3 (Blue) καθώς περιέχει τις τρεις χρωματικές συνιστώσες της εικόνας.

Για να επιτευχθεί η παραπάνω μετατροπή, για κάθε εικονοστοιχείο, της αρχικής εικόνας RGB πραγματοποιείται η παρακάτω πράξη:

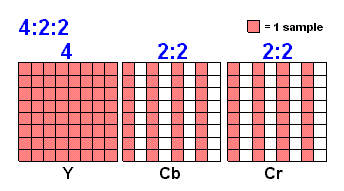
όπου οι δείκτες υποδεικνύουν τη θέση του εκάστοτε pixel της εικόνας. Έτσι τελικά κατασκευάζεται ένας τρισδιάστατος πίνακας με τις τρεις συνιστώσες Y, Cb και Cr.

Στη συνέχεια ακολουθείται ένα στάδιο υποδειγματοληψίας χρώματος προκειμένου να μειωθεί ο όγκος της εικόνας. Συγκεκριμένα, η διαδικασία αυτή διακρίνεται σε 3 κατηγορίες.

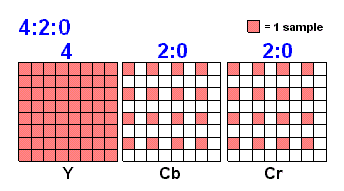
Η πρώτη κατηγορία ονομάζεται υποδειγματοληψία χρώματος 4:4:4, κατά την οποία λαμβάνεται ολόκληρη η εικόνα που σχηματίστηκε χωρίς κάποια δειγματοληψία.

**

Η δεύτερη κατηγορία ονομάζεται υποδειγματοληψία χρώματος 4:2:2. Στην κατηγορία αυτή διαγράφονται οι στήλες των χρωματικών συνιστωσών Cb, Cr που βρίσκονται σε άρτιες θέσεις στον πίνακα YCbCr. Είναι εμφανές ότι η κατηγορία αυτή εισάγει απώλειες καθώς δεν υπάρχει τρόπος πλήρους ανακατασκευής της εικόνας μετά από τη διαδικασία αυτή. Το αποτέλεσμα που προκύπτει είναι τρεις πίνακες διαστάσεων , , , για τις συνιστώσες Y, Cb, Cr αντίστοιχα.

**

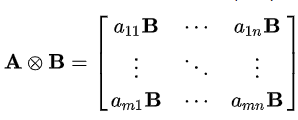
Τέλος, η τρίτη κατηγορία που υλοποιείται ονομάζεται υποδειγματοληψία χρώματος 4:2:0. Στην κατηγορία αυτή διαγράφονται τόσο οι στήλες όσο και οι γραμμές των χρωματικών συνιστωσών Cb, Cr που βρίσκονται σε άρτιες θέσεις στον πίνακα YCbCr. Είναι εμφανές ότι η κατηγορία αυτή εισάγει ακόμα μεγαλύτερες απώλειες καθώς και πάλι δεν υπάρχει τρόπος πλήρους ανακατασκευής της εικόνας μετά από τη διαδικασία αυτή. Το αποτέλεσμα που προκύπτει είναι τρεις πίνακες διαστάσεων , , , για τις συνιστώσες Y, Cb, Cr αντίστοιχα.

**

Η διαδικασία αποκωδικοποίησης προκύπτει με αντίστροφη σειρά των παραπάνω βημάτων.

Αρχικά γίνεται μετατροπή από τους τρεις πίνακες Y,Cb, Cr με μειωμένες διαστάσεις σε ένα τρισδιάστατο πίνακα διάστασης . Για να επιτευχθεί αυτό γίνεται παρεμβολή προκειμένου να συμπληρωθούν οι θέσεις των δειγμάτων που χάθηκαν κατά την υποδειγματοληψία. Για τη διαδικασία της παρεμβολής στις διάφορες κατηγορίες υποδειγματοληψίας γίνεται χρήση του γινομένου Kronecker.

Σύμφωνα με τον ορισμό, αν o **Α** είναι ένας πίνακας και ο **Β** είναι ένας πίνακας, τότε το γινόμενο Kronecker **A** ⊗ **B** είναι ο μπλοκ πίνακας :



Αρχικά, για την κατηγορία υποδειγματοληψίας χρώματος 4:4:4 δεν υπάρχει ανάγκη για παρεμβολή. Για την κατηγορία 4:2:2 εφαρμόζεται το γινόμενο Kronecker ανάμεσα στους χρωματικούς πίνακες Cb, Cr με το διάνυσμα ενώ για την κατηγορία 4:2:0 εφαρμόζεται το γινόμενο Kronecker ανάμεσα στους χρωματικούς πίνακες Cb, Cr με τον πίνακα .

Τέλος, η μετατροπή από εικόνα YCbCr σε RGB, γίνεται αντίστροφα και συγκεκριμένα μέσω του παρακάτω μετασχηματισμού:

***\* Σημείωση****: Σε περίπτωση που το πλήθος γραμμών και στηλών του τρισδιάστατου πίνακα RGB δεν είναι ακέραιο πολλαπλάσιο του 16, αφαιρείται από το τέλος των εικόνων το απαραίτητο πλήθος γραμμών και στηλών ώστε να πληρείται η παραπάνω συνθήκη.*

## 2.2 Μετασχηματισμός DCT

Στη συνέχεια οι πίνακες Y, Cb και Cr που προέκυψαν από την προηγούμενη διαδικασία διαχωρίζονται σε μπλοκ μεγέθους στα οποία εφαρμόζεται στη συνέχεια ο μετασχηματισμός DCT2 (Discrete Cosine Transform 2D) μέσω μιας συνάρτησης FDCT (Fast DCT). Έτσι προκύπτει ένα σύνολο από μπλοκ που περιέχουν τους συντελεστές DCT. Το πρώτο στοιχείο του κάθε μπλοκ ονομάζεται DC όρος ενώ τα υπόλοιπα στοιχεία AC όροι. Ο μετασχηματισμός DCT2 φαίνεται παρακάτω:

και ο αντίστροφός του είναι:

## 2.3 Κβαντισμός

Στο σημείο αυτό γίνεται κβαντισμός των στοιχείων κάθε μπλοκ μέσω του παρακάτω τύπου:

όπου οι δείκτες υποδεικνύουν τη θέση του εκάστοτε δείγματος στο συγκεκριμένο block, qScale σταθερά και qTable ένας πίνακας μεγέθους που ονομάζεται πίνακας κβαντισμού και διαφέρει για τα μπλοκ της χρωματικής συνιστώσας Y από τις υπόλοιπες. Οι διάφοροι πίνακες qTable που χρησιμοποιούνται στον παρών κωδικοποιητή-αποκωδικοποιητή είναι οι προτεινόμενοι πίνακες όπως ορίζονται από το πρότυπο JPEG.

Η διαδικασία αποκβαντισμού των στοιχείων κάθε μπλοκ στον αποκωδικοποιητή γίνεται αντίστροφα,

Είναι φανερό πως η διαδικασία αυτή, όπως και εκείνη της υποδειγματοληψίας χρώματος, εισάγει απώλειες στη ποιότητα της εικόνας.

## 2.4 Zig Zag Scanning και Run Length Encoding

Στο επόμενο στάδιο κάθε μπλοκ αναδιατάσσεται στη μορφή διανύσματος, ώστε να γίνει η απαραίτητη προετοιμασία για κωδικοποίηση. Η παραπάνω αναδιάταξη γίνεται μέσω μιας σάρωσης zig-zag όπως φαίνεται στο παρακάτω σχήμα. Ο λόγος που συμβαίνει αυτό αφορά ένα βασικό χαρακτηριστικό των συντελεστών DCT. Οι συντελεστές DCT που προηγούνται, σε κάθε διάσταση (οριζόντια και κατακόρυφη), στο μετασχηματισμένο μπλοκ αντικατοπτρίζουν τις μικρότερες συχνότητες που εμφανίζονται στο αντίστοιχο μπλοκ εικόνας ενώ οι συντελεστές που έπονται αντικατοπτρίζουν τις μεγαλύτερες συχνότητες. Με τη σάρωση αυτή έχει παρατηρηθεί ότι το νέο διάνυσμα συχνά συμπληρώνεται με μεγάλες ακολουθίες μηδενικών στοιχείων οι οποίες θα κωδικοποιηθούν στη συνέχεια με κατάλληλο τρόπο για περεταίρω συμπίεση. Για τον ίδιο λόγο επίσης, αντί για την κωδικοποίηση του συντελεστή DC κάθε μπλοκ, γίνεται κωδικοποίηση της διαφοράς του εκάστοτε DC συντελεστή από το DC συντελεστή του προηγούμενου μπλοκ (DCpred). Ως αρχική τιμή του DCpred τίθεται η τιμή μηδέν για κάθε συνιστώσα Y,Cb και Cr.

Με το πέρας της παραπάνω διαδικασίας έχει δημιουργηθεί ένα διάνυσμα μήκους 64 στοιχείων, με πρώτο στοιχείο τη διαφορά κβαντίσμεων DC-DCpred και υπόλοιπα στοιχεία τους κβαντισμένους συντελεστές AC. Το διάνυσμα αυτό θα αναφέρεται στη συνέχεια ως διάνυσμα zig zag. Για την κωδικοποίηση των τελευταίων 63 στοιχείων ακολουθείται μια παραλλαγή της κωδικοποίησης Run Length Encoding (RLE). Συγκεκριμένα, η παραλλαγή αυτή κωδικοποιεί το παραπάνω διάνυσμα σε ένα πίνακα διάστασης , ο οποίος στη συνέχεια θα αναφέρεται ως πίνακας rle.

Η πρώτη γραμμή του πίνακα αυτού συμπληρώνεται με ένα μηδενικό στοιχείο ακολουθούμενο από τη DC διαφορά που αναφέρθηκε προηγουμένως.

Για τις υπόλοιπες γραμμές η διαδικασία συμπλήρωσης είναι η εξής:

Γίνεται καταμέτρηση των συνεχόμενων μηδενικών στοιχείων του διανύσματος zig zag μέχρι την εμφάνιση ενός μη μηδενικού στοιχείου. Το πλήθος που υπολογίστηκε εισάγεται στην πρώτη στήλη του πίνακα ενώ το μη μηδενικό στοιχείο εισάγεται στη δεύτερη στήλη. Η διαδικασία αυτή επαναλαμβάνεται μέχρι το τέλος του διανύσματος zig zag. Στην περίπτωση που η ακολουθία μηδενικών στοιχείων φτάνει σε πλήθος τον αριθμό 16, η γραμμή του πίνακα rle συμπληρώνεται από το διάνυσμα και η καταμέτρηση των μηδενικών μηδενίζεται. Η περίπτωση αυτή θα αναφέρεται στη συνέχεια ως ZRL (Zero Run Length). Τέλος, μετά την καταμέτρηση του τελευταίου μη μηδενικού συντελεστή DCT ο πίνακας rle σταματάει να συμπληρώνεται ενώ προστίθεται ως τελευταία γραμμή το διάνυσμα που υποδηλώνει το τέλος του μπλοκ EOB (End Of Block).

Για την αποκωδικοποίηση του πίνακα rle και τη μετατροπή του πίσω σε κβαντισμένα μπλοκ ακολουθείται η αντίστροφη διαδικασία μετατρέποντας αρχικά τον πίνακα rle σε ένα διάνυσμα 64 θέσεων, υπολογισμό του DC όρου με βάση τον προηγούμενο και τέλος, ανακατασκευή του κβαντισμένου μπλοκ με zig zag σειρά.