

Συστήματα Πολυμέσων και Εικονική Πραγματικότητα Εργασίες

A. Ντελόπουλος

Τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών
Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης

1 Εισαγωγικές Παρατηρήσεις

Οι παρακάτω εργασία αποτελεί *προαιρετικό* μέρος του μαθήματος 987H-Συστήματα Πολυμέσων και Εικονική Πραγματικότητα και η εκτέλεσή της συνεισφέρει 1 έως 4 επιπλέον μονάδες στην τελική βαθμολογία.

Οι εργασία θα πρέπει να εκτελεστεί σε αυστηρά ατομική βάση.

Η εργασία αποτελείται από 4 ενότητες. Η υλοποίηση της εργασίας μπορεί κατεπιλογή να περιλάβει την πρώτη ενότητα, την πρώτη και τη δεύτερη, κ.ο.κ. συνεισφέροντας 1 μονάδα για κάθε ενότητα. Δεν μπορεί όμως να εκτελεστεί μία ενότητα χωρίς να έχει ορθά εκτελεστεί η προηγούμενή της.

Κάθε ενότητα της εργασίας αφορά στην υλοποίηση μίας βασικής λειτουργίας ενός κωδικοποιητή/αποκωδικοποιητή MPEG-1.

2 Απλουστευμένη Κωδικοποίηση MPEG-1

Η προτεινόμενη εργασία στοχεύει στην υλοποίηση ενός απλουστευμένου κωδικοποιητή / αποκωδικοποιητή βίντεο κατά το πρότυπο MPEG-1 που περιλαμβάνει τα ακόλουθα στάδια :

1. Εκτίμηση και Αντιστάθμιση Κίνησης.
2. Block DCT.
3. Κβαντισμό (και αποκβαντισμό) των συντελεστών DCT.
4. Zig-zag scanning.
5. Κωδικοποίηση (και αποκωδικοποίηση) μήκους διαδρομής (Run length encoding).
6. Κωδικοποιητή (και αποκωδικοποιητή) Huffman.

Για απλούστευση η υλοποίηση θα υποθέσει ότι η προς κωδικοποίηση ακολουθία εικόνων είναι grayscale με 8 bits/pixel.

Για τον έλεγχο του κωδικοποιητή / αποκωδικοποιητή που θα κατασκευαστεί θα χρησιμοποιηθούν 13 frames της ακολουθίας house.seq100.png-house.seq112.png.

Στη συνέχεια περιγράφονται οι ενότητες της προτεινόμενης εργασίας.

2.1 Εκτίμηση και Αντιστάθμιση Κίνησης (1 μονάδα)

Να κατασκευαστεί συνάρτηση πρόβλεψης ενός frame από ένα frame αναφοράς,

$[comp_frame, mot_vec] = motestP(pred_frame, ref_frame, block_size, search_region, cost_fun)$

με τις παρακάτω εισόδους / εξόδους:

pred_frame: πίνακας, διάστασης $M \times N$, με τις τιμές των pixels του frame που πρόκειται να προβλεφθεί.

ref_frame: πίνακας, διάστασης $M \times N$, με τις τιμές των pixels του frame αναφοράς.

block_size: η διάσταση των blocks στα οποία υποδιαιρείται το *pred_frame* για τις ανάγκες της πρόβλεψης κίνησης (συνήθως 8 ή 16 για blocks διάστασης 8×8 ή 16×16 αντίστοιχα).

search_region: ακέραιος 1×4 πίνακας της μορφής $[v_{x,min}, v_{x,max}, v_{y,min}, v_{y,max}]$ που καθορίζει την περιοχή αναζήτησης του πλέον ταιριαστού block στο *ref_frame* σε σχέση με τη θέση του υπό πρόβλεψη block του *pred_frame* (Ισοδύναμα τις ελάχιστες / μέγιστες τιμές των συνιστωσών των διανυσμάτων κίνησης).

cost_fun: η μεταβλητή αυτή προσδιορίζει το είδος της απόστασης που θα χρησιμοποιηθεί για το ταίριασμα: 1 για τη νόρμα L_1 (άθροισμα απολυτών διαφορών), 2 για τη νόρμα L_2 (τετραγωνική ρίζα του αθροίσματος των τετραγώνων των διαφορών).

comp_frame: πίνακας, διάστασης $M \times N$, με τις τιμές των pixels του frame που αντικαθιστά το *pred_frame* και περιέχει τα σφάλματα πρόβλεψης.

mot_vec: πίνακας διάστασης $K \times L \times 2$ (όπου $K = \frac{M}{block_size}$, $L = \frac{N}{block_size}$) που περιέχει τα υπολογισθέντα διανύσματα κίνησης.

Να κατασκευαστεί συνάρτηση πρόβλεψης ενός frame από δύο frames αναφοράς (bi-directional prediction),

$$[comp_frame, mot_vec, used_ref] = \text{motestB}(pred_frame, ref_frame_1, ref_frame_2, block_size, search_region, cost_fun)$$

με εισόδους / εξόδους αντίστοιχες της συνάρτησης *motestP()* με μόνες διαφορές ότι τα frames αναφοράς είναι δύο και ότι στην έξοδο έχει περιληφθεί ο πίνακας *used_ref* διαστάσεων $K \times L$ που δηλώνει ποιό από τα frames αναφοράς (1 για το *ref_frame_1* και 2 για το *ref_frame_2*) χρησιμοποιήθηκε στην πρόβλεψη του κάθε block.

Να κατασκευαστεί συνάρτηση

$$[comp_seq, mot_vec_seq] = \text{motest}(inp_seq, block_size, search_region, cost_fun)$$

που χρησιμοποιεί τις *motestP* και *motestB* και υλοποιεί την εκτίμηση/αντιστάθμιση κίνησης σε μια ακολουθία frames σύμφωνα με τη δομή GOP IBPB. Η συνάρτηση έχει τις ακόλουθες εισόδους / εξόδους:

inp_seq: πίνακας διάστασης $M \times N \times F$ που περιέχει τις τιμές των pixels των F frames της αρχικής ακολουθίας.

comp_seq: πίνακας διάστασης $M \times N \times F$ που περιέχει τις τιμές των pixels των F frames της ακολουθίας μετά την αντιστάθμιση κίνησης.

mot_vec_seq: : πίνακας διάστασης $K \times L \times 2 \times F$ που περιέχει τις τιμές των διανυσμάτων κίνησης για κάθε ένα από τα F frames της ακολουθίας. Ειδικά για τα I frames οι τιμές των διανυσμάτων θα είναι $[0, 0]$.

Να κατασκευαστεί συνάρτηση

$$[recon_seq] = \text{motestinv}(comp_seq, mot_vec_seq, block_size)$$

που θα αντιστρέφει τη διαδικασία της *motest()*.

2.2 Κβαντισμός / Αποκβαντισμός Συντελεστών DCT (1 μονάδα)

Να κατασκευαστεί συνάρτηση κβαντισμού της μορφής

$$symb_index = \text{quantizer}(x, dec_levels)$$

που δέχεται στην είσοδο μία βαθμωτή ποσότητα x και ένα διάνυσμα *dec_levels* με τις στάθμες απόφασης του κβαντιστή και παράγει στην έξοδο τον ακέραιο *symb_index* που δεικτοδοτεί το παραγόμενο σύμβολο (π.χ. 0 για το σύμβολο που αντιστοιχεί στη ζώνη που περιέχει την τιμή 0, -1 για την αμέσως αριστερότερη ζώνη, 1 για την αμέσως δεξιότερη, κ.λπ.).

Να κατασκευαστεί συνάρτηση αποκβαντισμού της μορφής

$$y = \text{dequantizer}(symb_index, quant_levels)$$

που δέχεται στην είσοδο τον ακέραιο *symb_index* που δεικτοδοτεί το σύμβολο εισόδου (βλ. παραπάνω) και το διάνυσμα *quant_levels* με τα επίπεδα και παράγει την αποκβαντισμένη τιμή y .

Να κατασκευαστεί συνάρτηση κβαντισμού ενός blockτης μορφής

$$block_symb_index = \text{blockquantizer}(block_x, dec_levels, resolution)$$

που δέχεται στην είσοδο έναν πίνακα πραγματικών αριθμών, $block_x$, το διάνυσμα dec_levels με τις στάθμες απόφασης ενός πρότυπου κβαντιστή και τον πίνακα $resolution$ διάστασης ίδιας με το $block_x$ και παράγει στην έξοδο τον ακέραιο πίνακα $block_symb_index$ που δεικτοδοτεί τα παραγόμενα σύμβολα για κάθε θέση του block εισόδου. Για κάθε θέση (i,j) του block εισόδου ο κβαντιστής θα χρησιμοποιεί τις στάθμες απόφασης που προσδιορίζονται από το διάνυσμα $\frac{dec_levels}{resolution(i,j)}$.

Να κατασκευαστεί συνάρτηση αποκβαντισμού ενός block της μορφής

$$block_y = \text{blockdequantizer}(block_symb_index, quant_levels, resolution)$$

Να κατασκευαστεί συνάρτηση υπολογισμού των συντελεστών DCT και κβαντισμού αυτών, της μορφής:

$$quant_DCT = \text{QDCT}(frame, DCT_block_size, dec_levels, resolution)$$

με τις παρακάτω εισόδους / εξόδους:

$frame$: Πίνακας $M \times N$ με τις τιμές του frame εισόδου.

DCT_block_size : Συνήθως 8 για 8×8 blocks.

dec_levels και $resolution$: όπως παραπάνω.

$quant_DCT$: Πίνακας $M \times N$ με τα σύμβολα κβαντισμού για τους συντελεστές DCT που προέκυψαν.

Να κατασκευαστεί συνάρτηση αποκβαντισμού των συντελεστών DCT και αντιστροφής του μετασχηματισμού, της μορφής:

$$recon_frame = \text{IQDCT}(quant_DCT, DCT_block_size, quant_levels, resolution)$$

με έξοδο το ανακατασκευασμένο frame, $recon_frame$ υπό τη μορφή $M \times N$ πίνακα πραγματικών αριθμών.

2.3 Zig-zag scanning και Κωδικοποίηση / Αποκωδικοποίηση Μήκους Διαδρομής (1 μονάδα)

Να κατασκευαστεί συνάρτηση υπολογισμού των συμβόλων μήκους διαδρομής για τους πίνακες των κβαντισμένων συντελεστών DCT. Τα σύμβολα μήκους διαδρομής θα είναι διάδες της μορφής $(quant_DCT_coeff, following_zeros)$. Τα μήκη διαδρομής θα υπολογίζονται χωριστά για κάθε block συντελεστών. Η συνάρτηση θα είναι της μορφής:

$$run_symbols = \text{RLE}(quant_DCT, DCT_block_size,)$$

όπου ο πίνακας $run_symbols$ θα είναι διάστασης $R \times 2$ όπου R τα μήκη διαδρομής που εντοπίστηκαν για όλα τα blocks του frame $quant_DCT$.

Να κατασκευαστεί η συνάρτηση

$$quant_DCT = \text{IRLE}(run_symbols, DCT_block_size)$$

που αντιστρέφει την προηγούμενη

2.4 Κωδικοποίηση Huffman (1 μονάδα)

Να κατασκευαστεί συνάρτηση κωδικοποίησης Huffman για τα μήκη διαδρομής της προηγούμενης ενότητας:

$$[frame_stream, frame_symbol_prob] = \text{huff}(run_symbols)$$

όπου η έξοδος $frame_stream$ είναι μία συμβολοσειρά από 0 και 1 και ο πίνακας $frame_symbol_prob$ περιέχει δύο στήλες με τα σύμβολα μήκους διαδρομής που ανιχνεύτηκαν και μία τρίτη με τις αντίστοιχες εκτιμώμενες πιθανότητες εμφάνισής τους.

Να κατασκευαστεί η αντίστροφη συνάρτηση :

$$run_symbols = ihuff (frame_stream, frame_symbol_prob)$$

Να συναρμολογηθούν τα παραπάνω στις συναρτήσεις *MPEGcod()* και *MPEGdecod()* και να υπολογιστεί ο συνολικός βαθμός συμπίεσης.