

Συστήματα Πολυμέσων:
Απαλλακτική Εργασία 2020
Μέρος Α΄



ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

UNIVERSITY OF PIRAEUS

Αγγελική Θεοδώρου: Π17032

Δανάη – Ιωάννα Μπαντάνα: Π17081

Εισαγωγή

Η εργασία αυτή είναι απαλλακτική για το μάθημα «Συστήματα Πολυμέσων», με υπεύθυνο καθηγητή τον Άγγελο Πικράκη και αναφέρεται στην διαδικασία συμπίεσης των εικόνων, στην αντιστάθμιση κίνησης και στην επικάλυψη κινούμενων σωμάτων σε βίντεο.

Συγκεκριμένα, στην πρώτη άσκηση (8.17) γίνεται υλοποίηση ενός μέρους της διαδικασίας της συμπίεσης των βίντεο, το οποίο είναι η πρόβλεψη των επόμενων πλαισίων με βάση το τρέχον πλαίσιο. Αυτή η μέθοδος πρόβλεψης των επόμενων πλαισίων χρησιμοποιεί την αντιστάθμιση κίνησης και το αποτέλεσμα που δίνει είναι πολύ κοντά στο αρχικό βίντεο.

Στην δεύτερη άσκηση (8.18) γίνεται ανίχνευση των σωμάτων που κινούνται σε κάθε πλαίσιο και στη συνέχεια γίνεται προσπάθεια αντικατάστασης αυτών με το παρασκήνιο.

Στην τρίτη άσκηση (μέρος β') εκτελείται ένα μέρος της διαδικασίας της συμπίεσης εικόνων και ζητείται η εφαρμογή της διαδικασίας αυτής σε εικόνα που κατασκευάζεται μέσω του κώδικα. Ακόμη ζητείται ο μέσος λόγος συμπίεσης της κβαντισμένης εικόνας που προκύπτει.

Μέθοδοι λύσης

- Άσκηση πρώτη

Για την υλοποίηση του πρώτου ερωτήματος χρησιμοποιείται η εξής μέθοδος:

1. Αποθηκεύεται το πρώτο πλαίσιο του βίντεο καθώς θεωρείται I-frame και μένει ίδιο στο βίντεο που θα παραχθεί.
2. Υπολογίζεται η διαφορά κάθε τρέχοντος και επόμενου πλαισίου.
3. Στο πρώτο πλαίσιο προσθέτω τη διαφορά των δύο πλαισίων που υπολόγισα πριν και έτσι ανακατασκευάζω το επόμενο frame.
4. Μαζεύοντας όλα τα ανακατασκευασμένα πλαίσια συνθέτω το ανακατασκευασμένο βίντεο.

Για την υλοποίηση του δεύτερου ερωτήματος χρησιμοποιείται η μετρική Sum of Absolute Differences (SAD), η οποία υπολογίζει τη διαφορά μεταξύ δύο πλαισίων με βάση τα blocks τους. Επομένως:

1. Διαιρεί κάθε πλαίσιο του βίντεο σε $k \times k$ macroblocks. Για την συγκεκριμένη εργασία για τη σταθερά k χρησιμοποιήθηκε ο αριθμός 16 και άρα κάθε πλαίσιο διαιρείται σε 16×16 macroblocks.

2. Για κάθε τρέχον macroblock επιλέγονται τα 4 (ή 3 αν πρόκειται για τα macroblocks που βρίσκονται στις ακριανές θέσεις του πλαισίου) γειτονικά macroblocks, συμπεριλαμβανομένου και του macroblock που έχει την ίδια θέση με το τρέχον στο επόμενο πλαίσιο.
3. Τα macroblocks αυτά συγκρίνονται μέσω του τύπου SAD:

$$SAD(i, j) = \sum_{p=1}^n \sum_{q=1}^m | frame_{n+1}[p, q] - frame_n[p + i, q + j] |$$

Όπου στη συγκεκριμένη περίπτωση $[m, n] = [16, 16]$ καθώς συμβολίζουν τις διαστάσεις των macroblocks και $[i, j] = \{ [0, 0], [0, 16], [16, 0], [0 - 16], [-16, 0] \}$ καθώς με βάση αυτές τις συντεταγμένες βρίσκονται οι γείτονες του εκάστοτε macroblock.

4. Από τις συγκρίσεις αυτές παράγονται το πολύ 5 macroblocks από τα οποία επιλέγεται το ελάχιστο.
5. Κρατώντας όλα τα προβλεπόμενα macroblocks συνθέτουμε το προβλεπόμενο επόμενο πλαίσιο.
6. Μαζεύοντας όλα τα προβλεπόμενα πλαίσια συνθέτουμε το παραγόμενο βίντεο που υπολογίσαμε με αντιστάθμιση κίνησης.

- Άσκηση δεύτερη

Για την υλοποίηση του πρώτου ερωτήματος χρησιμοποιείται το Computer Vision Toolbox της Matlab για να χωριστεί το κάθε πλαίσιο σε 17×17 macroblocks και να υπολογιστούν τα διανύσματα κίνησης. Για τον υπολογισμό των διανυσμάτων κίνησης χρησιμοποιείται η Mean Absolute Differences (MAD) μετρική. Τα διανύσματα αυτά αναπαρίστανται με βέλη και χρωματιστά blocks.

Για την υλοποίηση του δεύτερου ερωτήματος χρησιμοποιούνται τα ίδια διανύσματα κίνησης που υπολογίστηκαν στο προηγούμενο ερώτημα και ελέγχεται ποια από αυτά είναι διάφορα του 0. Έχοντας τις θέσεις των block του παρασκηνίου, σε όποιο block παρατηρείται κίνηση, εκείνο το block αντικαθίσταται με το αντίστοιχο block παρασκηνίου. Έπειτα, η εικόνα ανασυντίθεται από τα νέα blocks.

Σημειώνεται ότι στο τελευταίο πλαίσιο στο οποίο παρατηρήθηκε ότι δεν υπάρχει σώμα που κινείται παρά μόνο παρασκήνιο και από εκεί έχουμε τις θέσεις των μπλοκ του παρασκηνίου.

- Άσκηση τρίτη

Για την υλοποίηση του μέρους της συμπίεσης της εικόνας ακολουθείται η παρακάτω διαδικασία:

1. Εφαρμογή του μετασχηματισμού Discrete Cosine Transformation (DCT) σε κάθε block της εικόνας.
2. Κβάντιση κάθε block. Ως πίνακας κβάντισης χρησιμοποιήθηκε ο πίνακας που δίνεται από το βιβλίο στο οποίο βασίζεται η άσκηση.
3. Αντίστροφη κβάντιση κάθε block.
4. Εφαρμογή του αντίστροφου μετασχηματισμού Inverse Discrete Cosine Transformation (iDCT).

Στη συνέχεια, για την συμπίεση της κβαντισμένης εικόνας χρησιμοποιείται κωδικοποίηση Huffman και με βάση την κωδικοποίηση αυτή υπολογίζεται ο λόγος συμπίεσης της αρχικής εικόνας προς την συμπιεσμένη.

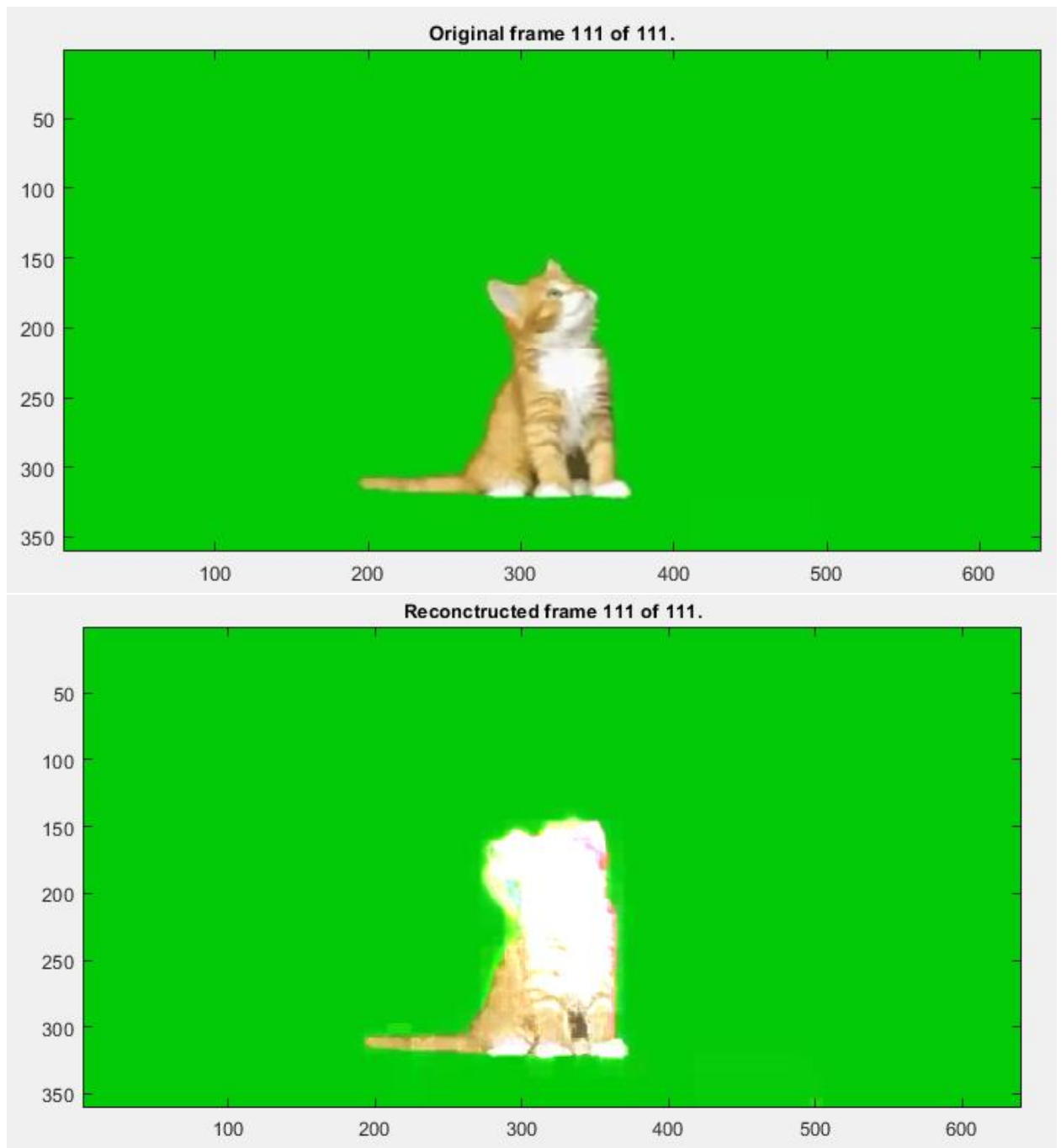
Σημειώνεται ότι για την συμπίεση με Huffman χρησιμοποιήθηκε το Communications Toolbox της Matlab.

Αποτελέσματα

- Άσκηση πρώτη

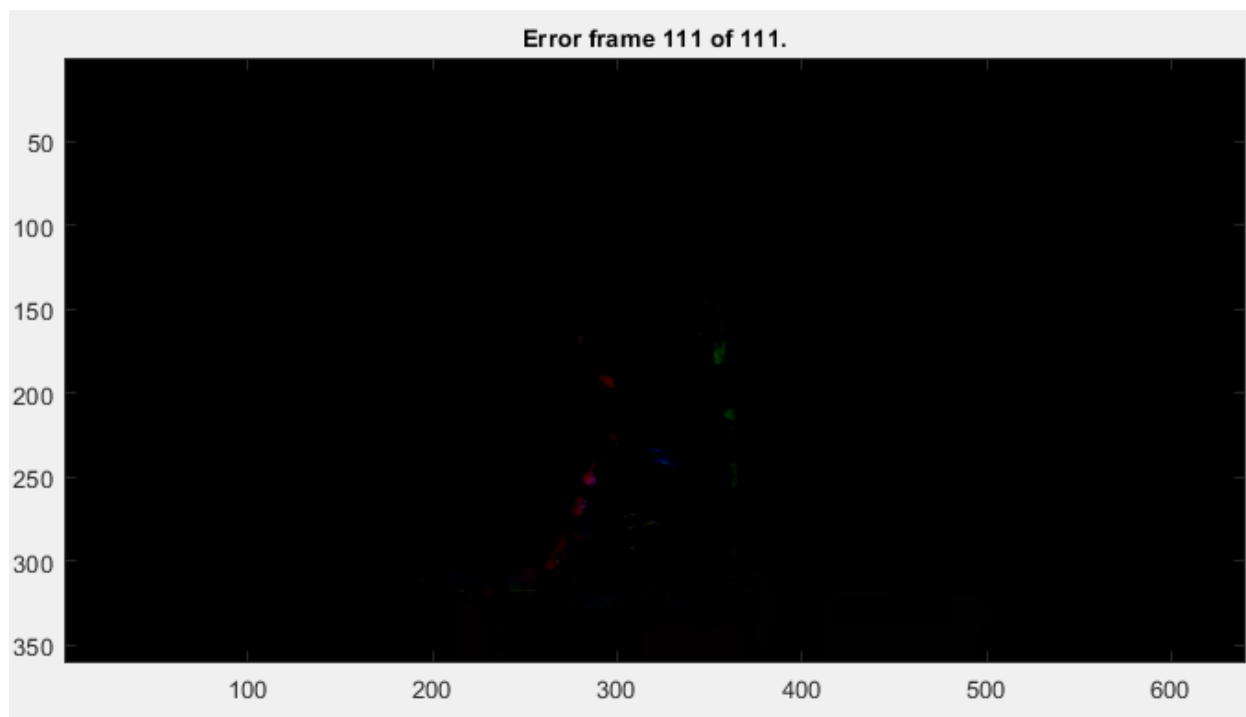
Ερώτημα α

Το αποτέλεσμα που προκύπτει για την πρόβλεψη κίνησης δεν είναι ικανοποιητικό. Παρακάτω δίνονται κάποια ενδεικτικά στιγμιότυπα που δείχνουν το αρχικό πλαίσιο και το ανακατασκευασμένο πλαίσιο.

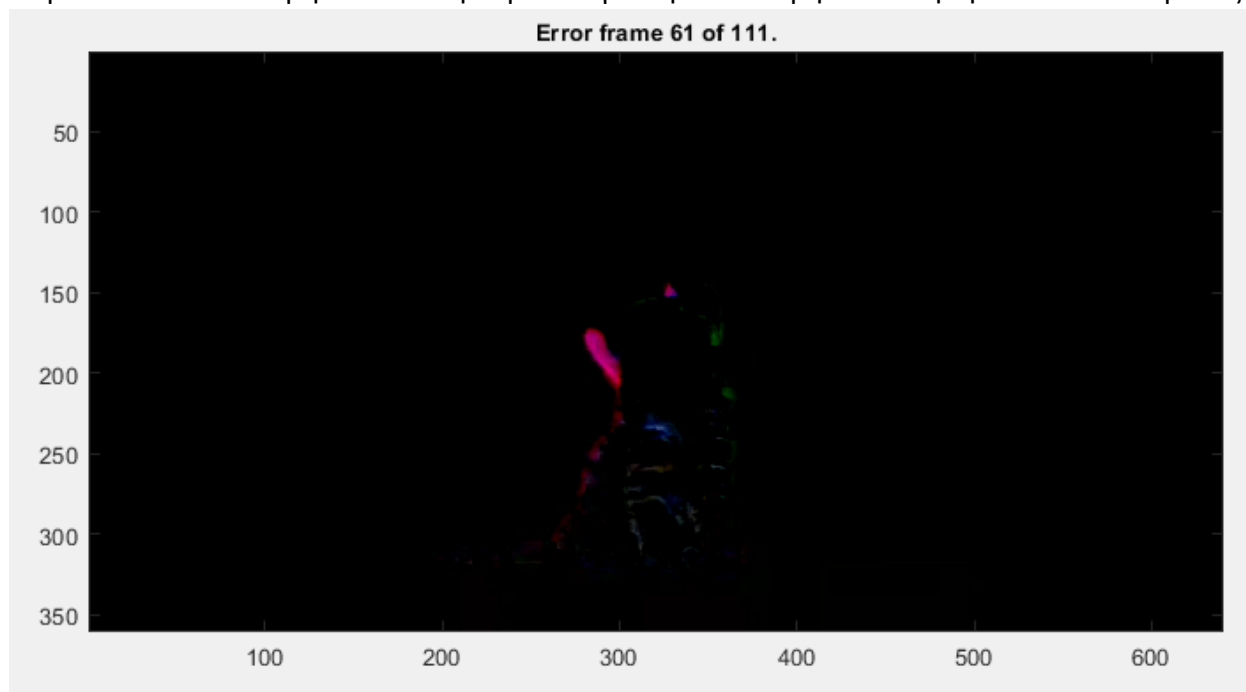


Είναι φανερό πως η πρόβλεψη της κίνησης δεν έχει δώσει κάποιο επιθυμητό αποτέλεσμα καθώς στο ανακατασκευασμένο πλαίσιο δεν διακρίνεται καν η μορφή του σώματος που κινείται.

Αντίστοιχα, υπολογίζοντας τα πλαίσια σφαλμάτων γίνεται αντιληπτό πως οι διαφορές ανάμεσα στο κανονικό και το ανακατασκευασμένο πλαίσιο είναι μεγάλες. Για το πλαίσιο 111.

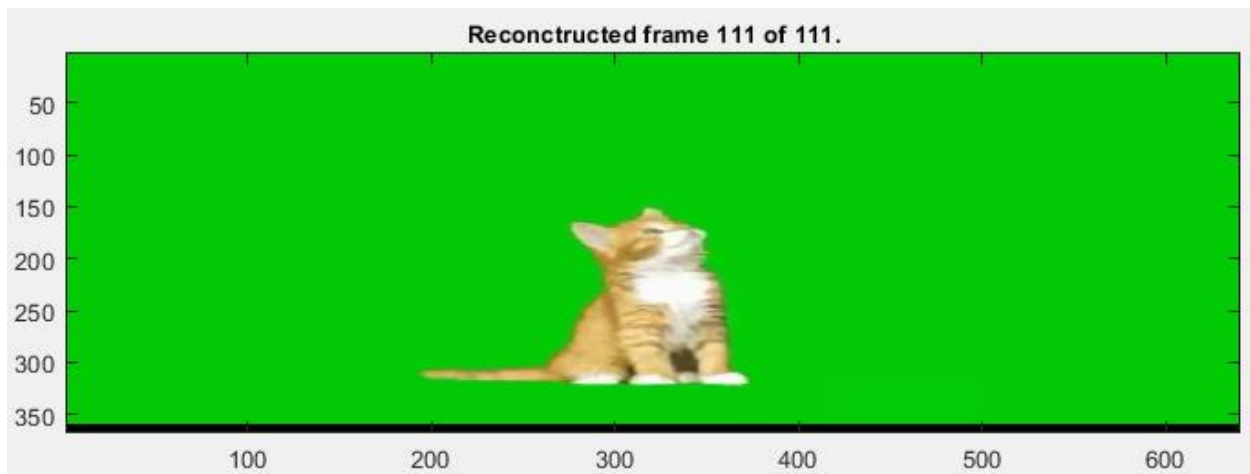
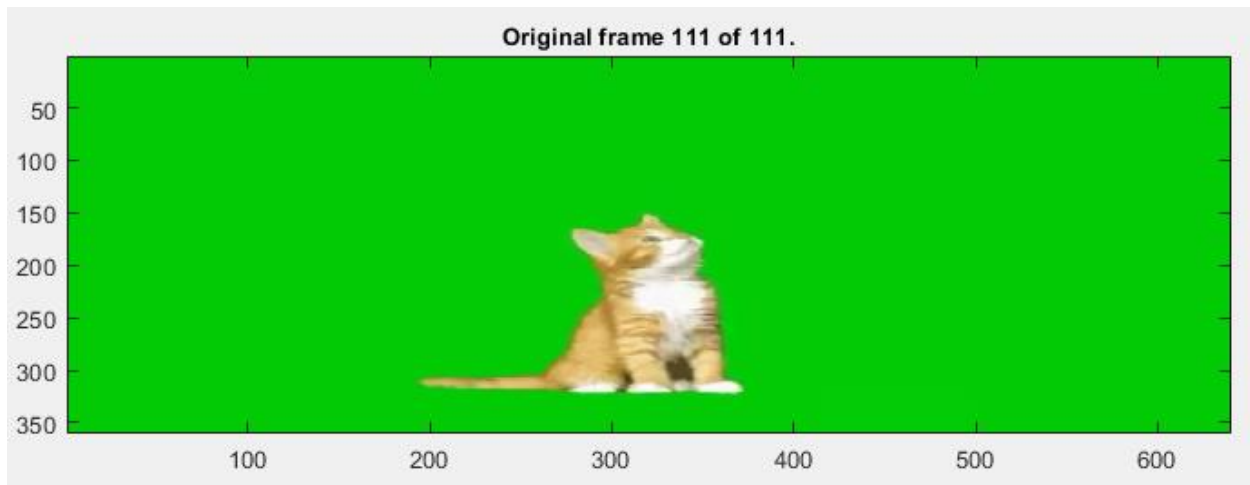


Παρατίθεται ένα ακόμη πλαίσιο σφαλμάτων για την καλύτερη κατανόηση του αποτελέσματος.



Ερώτημα β

Το αποτέλεσμα που προκύπτει σημειώνει μεγάλη βελτίωση σε σχέση με το προηγούμενο καθώς κρίνεται δύσκολο από το ανθρώπινο μάτι να παρατηρήσει διαφορές ανάμεσα στα προβλεπόμενα πλαίσια και τα κανονικά.



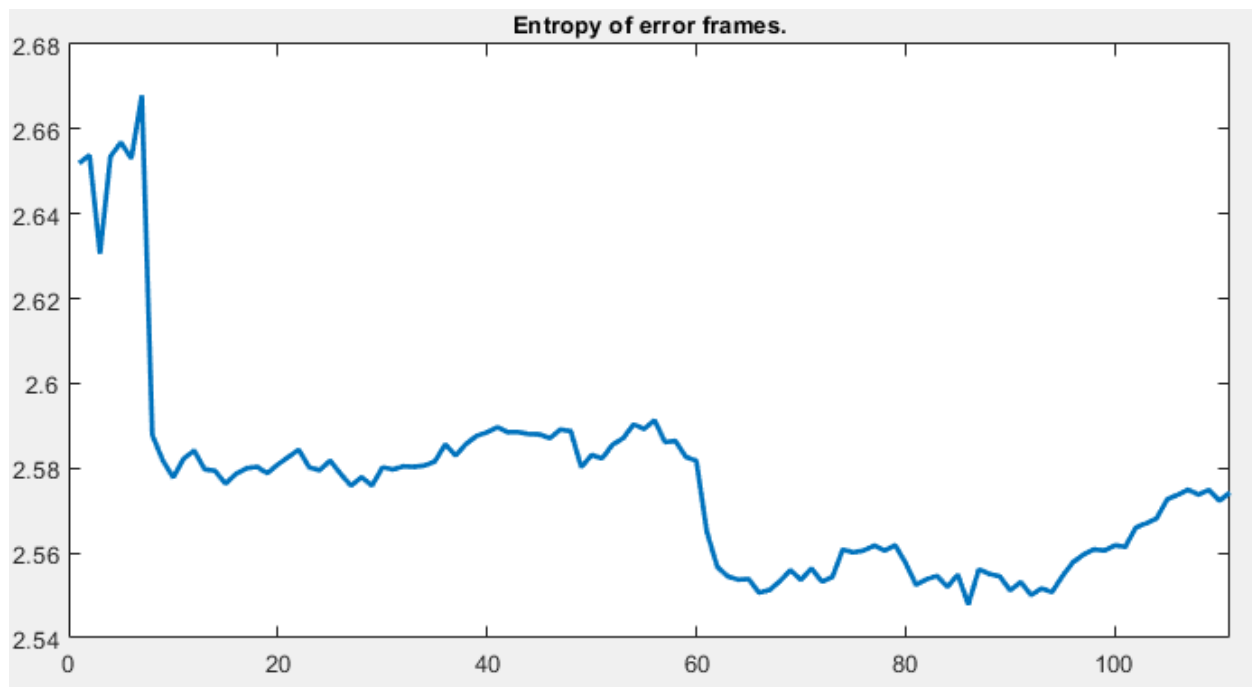
Σημειώνεται πως η μαύρη γραμμή στο τέλος του πλαισίου έχει προστεθεί έτσι ώστε να μπορεί να γίνει η διαίρεση του πλαισίου σε 16x16 macroblocks.

Με αυτή τη μέθοδο δεν φαίνεται να υπάρχουν διαφορές μεταξύ των δύο πλαισίων. Το αντίστοιχο πλαίσιο σφαλμάτων:

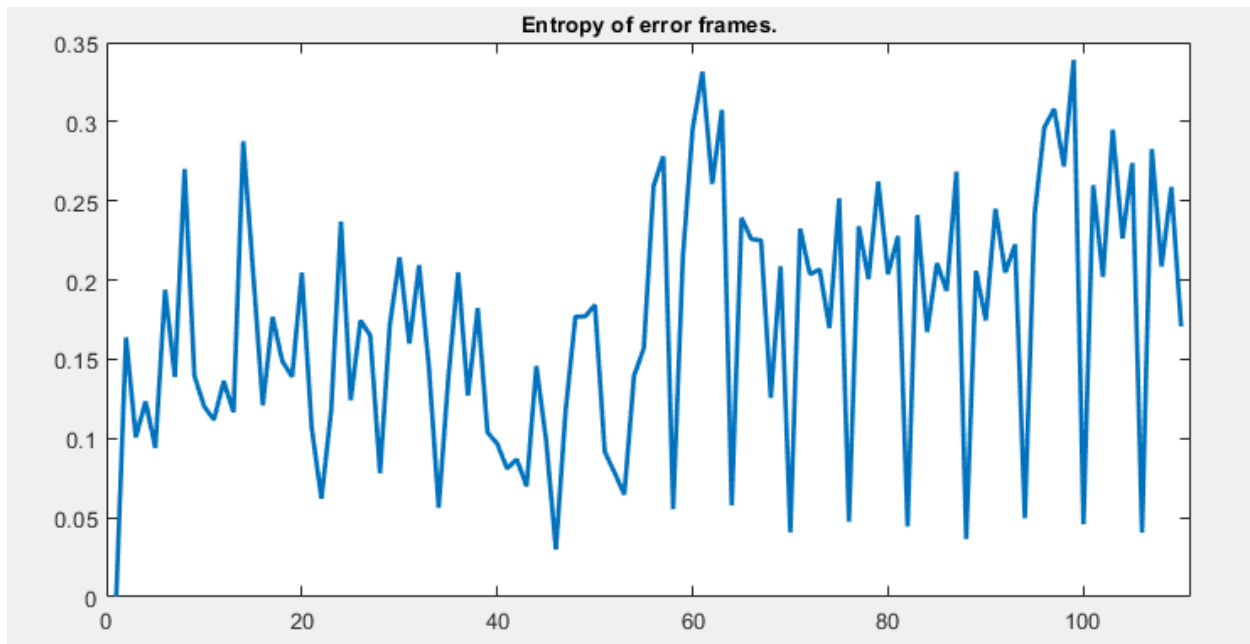


Για να γίνει πιο αντιληπτή η διαφορά ανάμεσα στους δύο τρόπους πρόβλεψης του επόμενου πλαισίου παραθέτουμε τα διαγράμματα εντροπίας των πλαισίων σφαλμάτων σε κάθε περίπτωση.

1. Για το ερώτημα α.



2. Για το ερώτημα β.



Παρατηρείται ότι η εντροπία των πλαισίων σφαλμάτων που προβλήθηκαν με την μετρική SAD είναι πολύ χαμηλότερη από την εντροπία των πλαισίων σφαλμάτων που προβλήθηκαν χωρίς αντιστάθμιση κίνησης. Στην πρώτη περίπτωση οι τιμές της εντροπίας κυμαίνονται στο διάστημα $[2.54, 2.67]$, ενώ στη δεύτερη κυμαίνονται $[0, 0.35]$.

Τέλος, παρατηρείται ότι η μέθοδος του δεύτερου ερωτήματος έχει μεγαλύτερο χρόνο εκτέλεσης από την πρώτη:

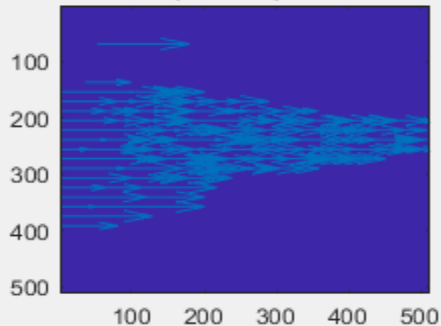
```
>> multimedia8_17a
Elapsed time is 43.308764 seconds.
>> multimedia8_17b
Elapsed time is 51.450843 seconds.
```

Ωστόσο, η διαφορά είναι τάξεως δευτερολέπτων για ένα βίντεο των 111 πλαισίων.

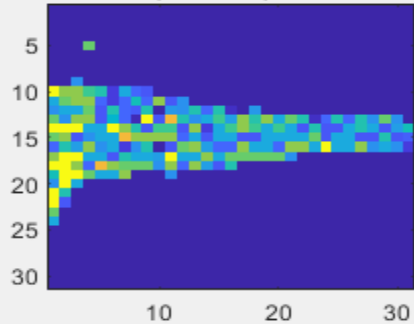
- Άσκηση δεύτερη
Ερώτημα α

Ο υπολογισμός των διανυσμάτων κίνησης δίνεται παρακάτω.

Original video. Frame 13 of 51. Motion Vectors (Arrows) for frame 14 of 51.



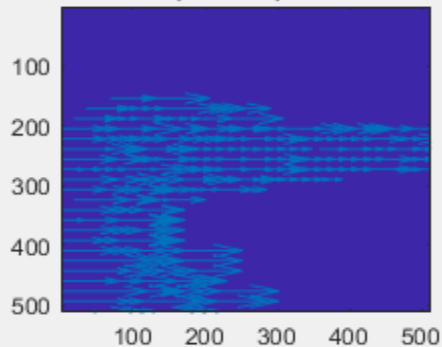
Motion Vectors (Colours) for frame 14 of 51.



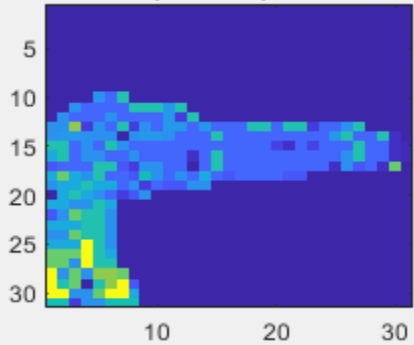
Motion Vectors (Numeric)

	1	2	
1	0	0	^
2	0	0	
3	0	0	
4	0	0	
5	0	0	
6	0	0	
7	0	0	v
< >			

Original video. Frame 38 of 51. Motion Vectors (Arrows) for frame 39 of 51.



Motion Vectors (Colours) for frame 39 of 51.

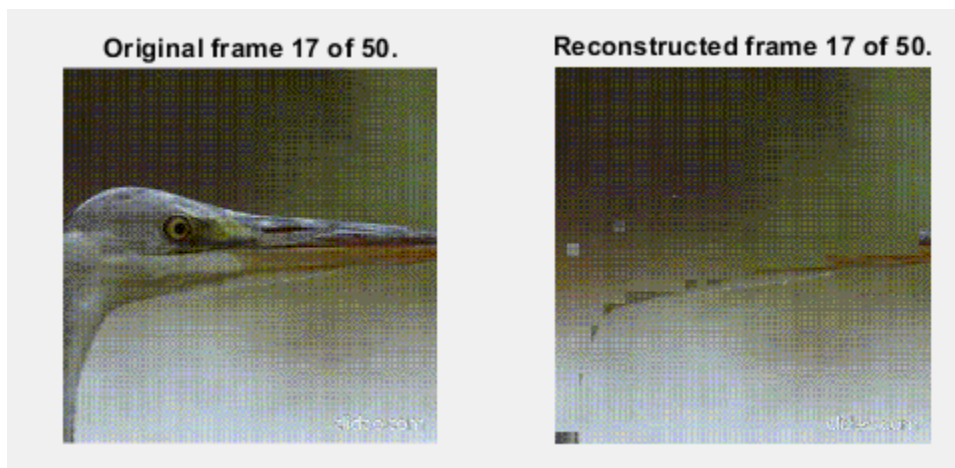
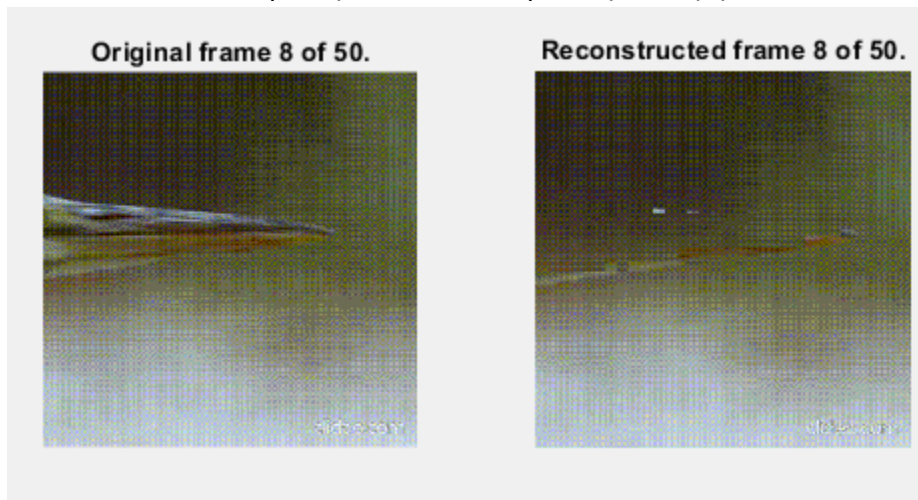


Motion Vectors (Numeric)

	1	2	
1	0	0	^
2	0	0	
3	0	0	
4	0	0	
5	0	0	
6	0	0	
7	0	0	v
< >			

Ερώτημα β

Κατά την υλοποίηση το σώμα που κινείται καλύπτεται με επιτυχία αλλά όχι πλήρως, καθώς στα όρια ακόμη εμφανίζεται. Ωστόσο, παρατηρήθηκε ότι για μέγεθος macroblock 16x16 έγινε η πιο αποτελεσματική κάλυψη.



μέγεθος macroblock = 16x16.

Original frame 6 of 50.



Reconstructed frame 5 of 50.



Original frame 12 of 50.



Reconstructed frame 12 of 50.



μέγεθος macroblock = 8x8

Original frame 19 of 50.

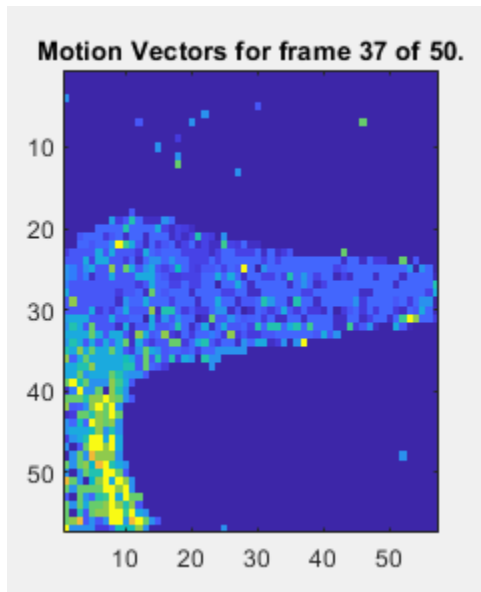


Reconstructed frame 19 of 50.

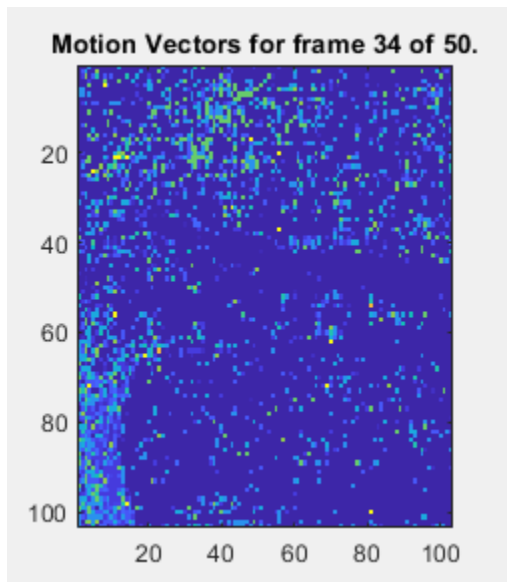


μέγεθος macroblock = 4x4.

Είναι φανερό πως για μέγεθος macroblock = 16x16 γίνεται η καλύτερη επικάλυψη, εφόσον για μέγεθος 8x8 παρατηρείται μεγάλο μέρος των ορίων του κινούμενου σώματος και για μέγεθος 4x4 δεν γίνεται σχεδόν καμία επικάλυψη. Αυτό συμβαίνει διότι η ακρίβεια των διανυσμάτων κίνησης μειώνεται όσο μειώνεται και το μέγεθος των macroblock. Στα macroblock των 8x8 τα διανύσματα ανιχνεύουν αρκετά καλά τις κινήσεις, με διάφορα σφάλματα, ωστόσο. Ειδικότερα, στα 4x4 macroblocks τα διανύσματα κίνησης μετά βίας ανιχνεύουν σωστά τις κινήσεις των macroblock.



Διανύσματα κίνησης με macroblock των 8x8.



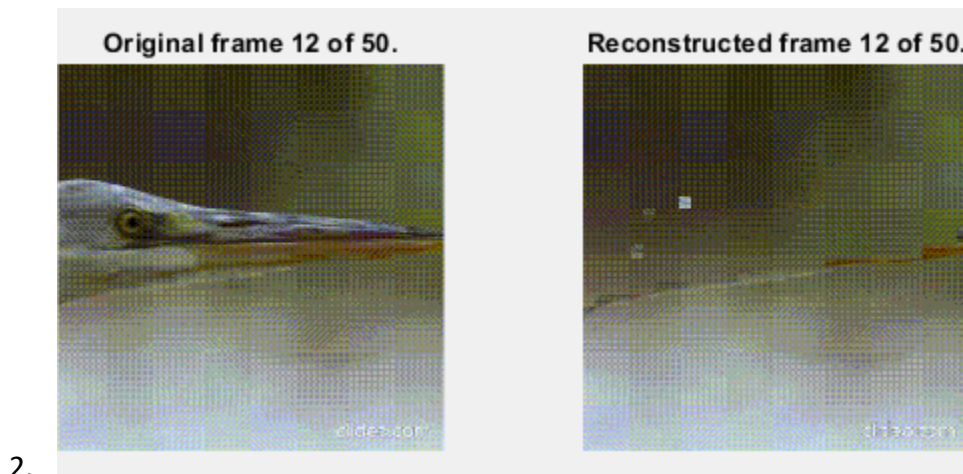
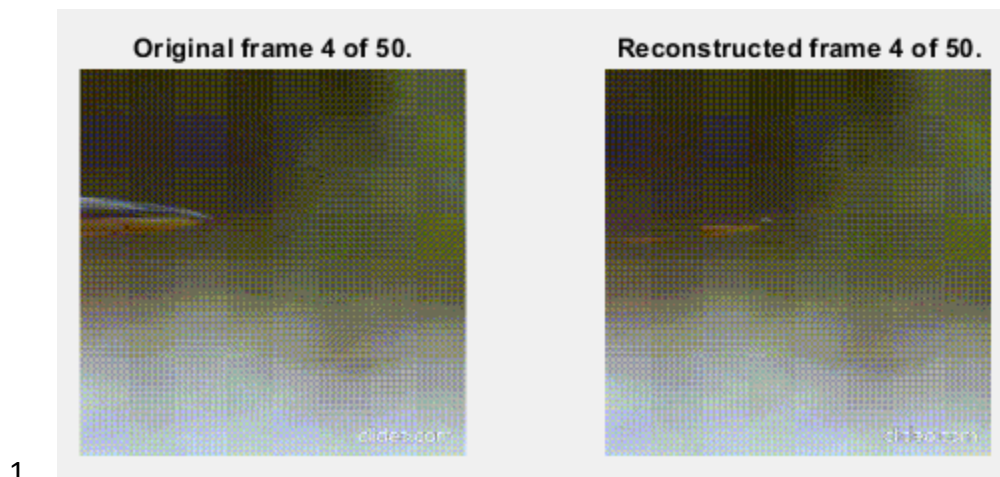
Διανύσματα κίνησης με macroblock των 4x4.

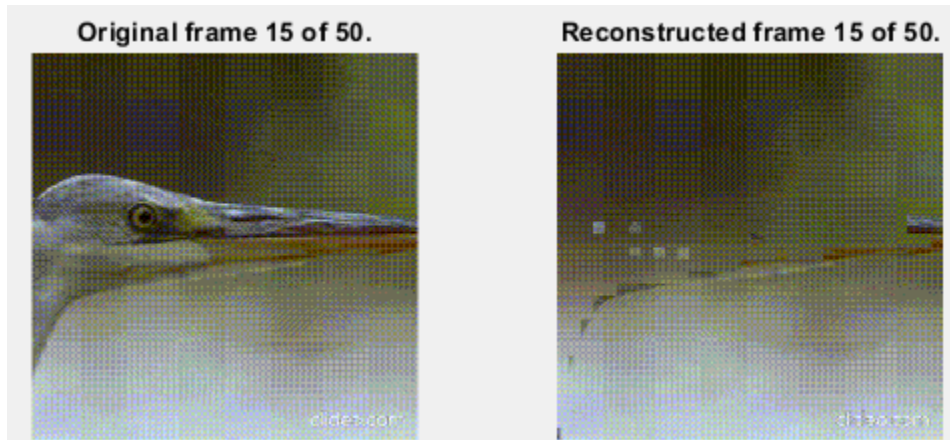
Ερώτημα γ

Κατά την υλοποίηση του προγράμματος που «καλύπτει» με μπλοκ παρασκήνιου τις κινήσεις των αντικειμένων σε βίντεο, παρατηρούνται ασυνέχειες στα όρια των μπλοκ που αντικαθίστανται. Είναι δηλαδή αντιληπτό το ότι υπάρχει σώμα στο πλαίσιο το οποίο έχει καλυφθεί, καθώς τα όρια του σώματος δεν είναι πλήρως καλυμμένα.

Τέλος, είναι σημαντικό να σημειωθεί ότι ακόμη και το κινούμενο σώμα κατά τη διάρκεια του βίντεο μπορεί να ακινητοποιηθεί στιγμιαία. Τότε δεν γίνεται ανίχνευση κάποιας κίνησης και έτσι δεν μπορεί να γίνει «κάλυψη» του σώματος με μπλοκ παρασκήνιου. Επομένως, κατά τη διάρκεια του ανακατασκευασμένου βίντεο υπάρχουν πλαίσια στα οποία το σώμα είναι πλήρως καλυμμένο από το παρασκήνιο και πλαίσια στα οποία στιγμιαία το σώμα εμφανίζεται κανονικά, καθώς μένει ακίνητο.

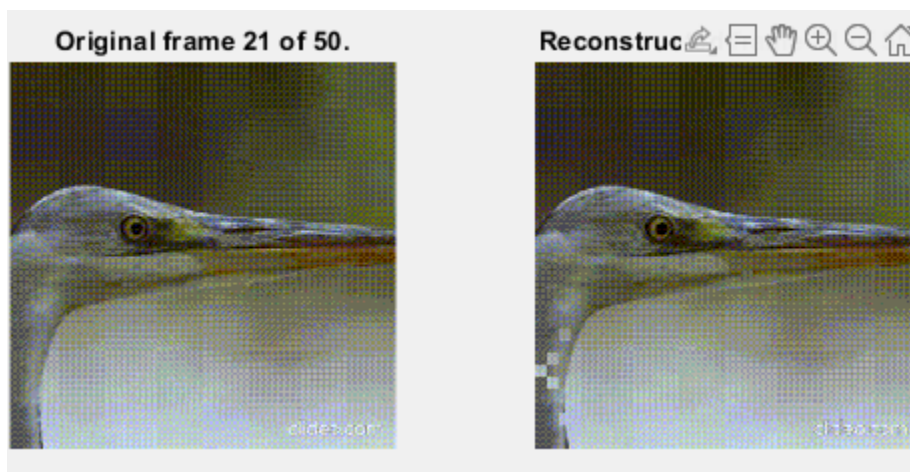
Παρατίθενται αναφορικά στιγμιότυπα από το πρόγραμμα και τις παραπάνω ασυνέχειες:



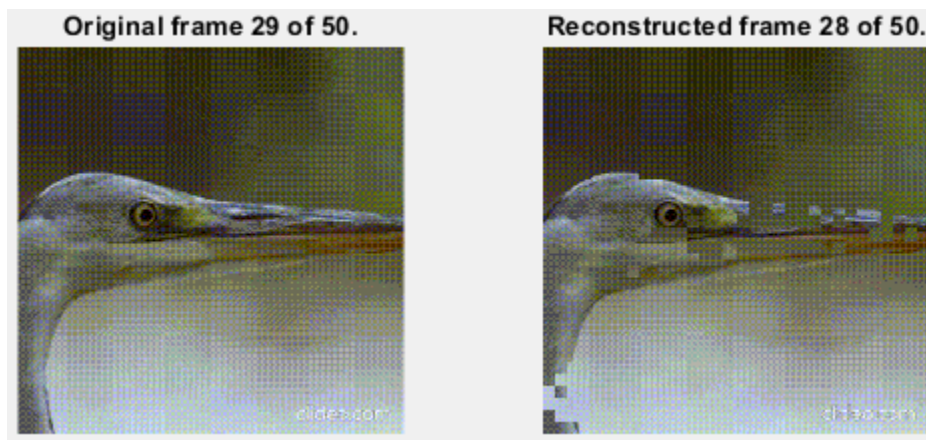


3.

Παρατηρούνται ασυνέχειες στα όρια του σώματος. Είναι εύκολο στο ανθρώπινο μάτι να αντιληφθεί ότι έχει γίνει επεξεργασία του βίντεο και κάλυψη του κινούμενου σώματος.



4.



5.

Εδώ παρατηρείται το γεγονός ότι εφόσον το σώμα μένει ακίνητο, δεν γίνεται ανίχνευση κίνησης και το σώμα εμφανίζεται κανονικά στο ανακατασκευασμένο βίντεο. Αντίστοιχα στην εικόνα 5 υπάρχουν ασυνέχειες στο καλυμμένο σώμα επειδή υπάρχουν μπλοκ που έμειναν ακίνητα και δεν άλλαξαν με μπλοκ παρασκηνίου.

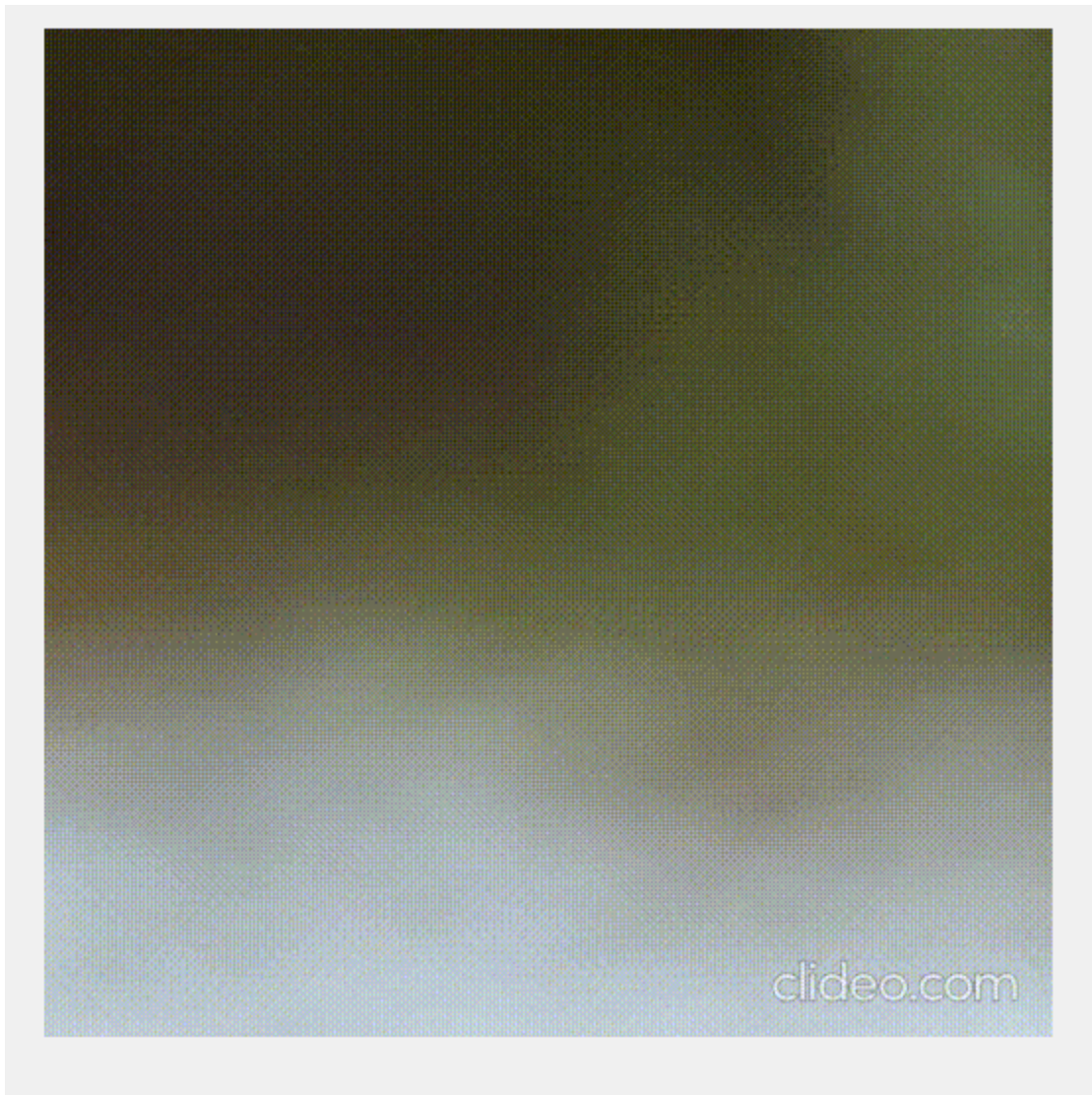
Αυτά τα φαινόμενα μπορούν να ελαχιστοποιηθούν εάν μπορέσουμε με κάποιον τρόπο να ανιχνεύσουμε όλες τις θέσεις του παρασκηνίου και μετά τις αντιστοιχίσουμε στις θέσεις των μπλοκ που κινήθηκαν.

Ερώτημα δ

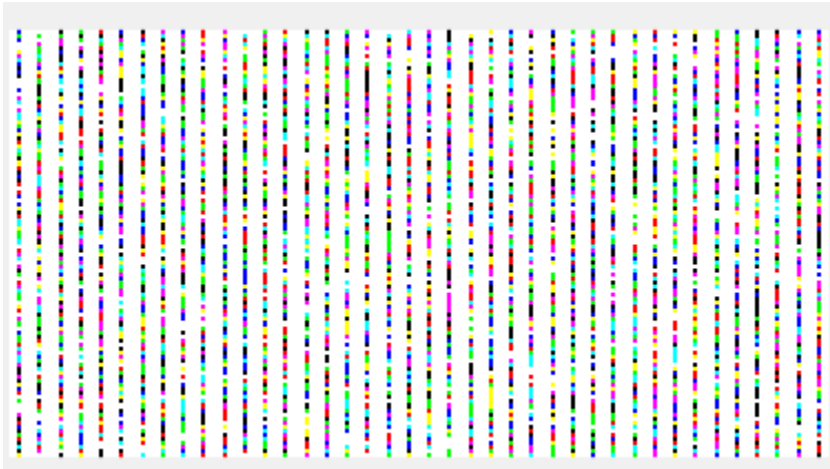
«Αλλά ας υποθέσουμε ότι μπορείτε να εντοπίσετε όλα τα μακρομπλόκ ενός πλαισίου που αντιστοιχούν στο παρασκήνιο. Πώς μπορείτε να αξιοποιήσετε αυτό το γεγονός, πλέον της χρήσης αντιστάθμισης κίνησης σε επίπεδο μακρομπλόκ;»

Εφόσον έχουμε ανιχνεύσει ποια μπλοκ κινούνται σε κάθε πλαίσιο μπορούμε με αντιστάθμιση κίνησης να προβλέψουμε την κίνησή τους στα επόμενα πλαίσια και έχοντας τις θέσεις κάθε μπλοκ παρασκηνίου μπορούμε να αντικαταστήσουμε τα μπλοκ που θα κινηθούν με τα μπλοκ παρασκηνίου.

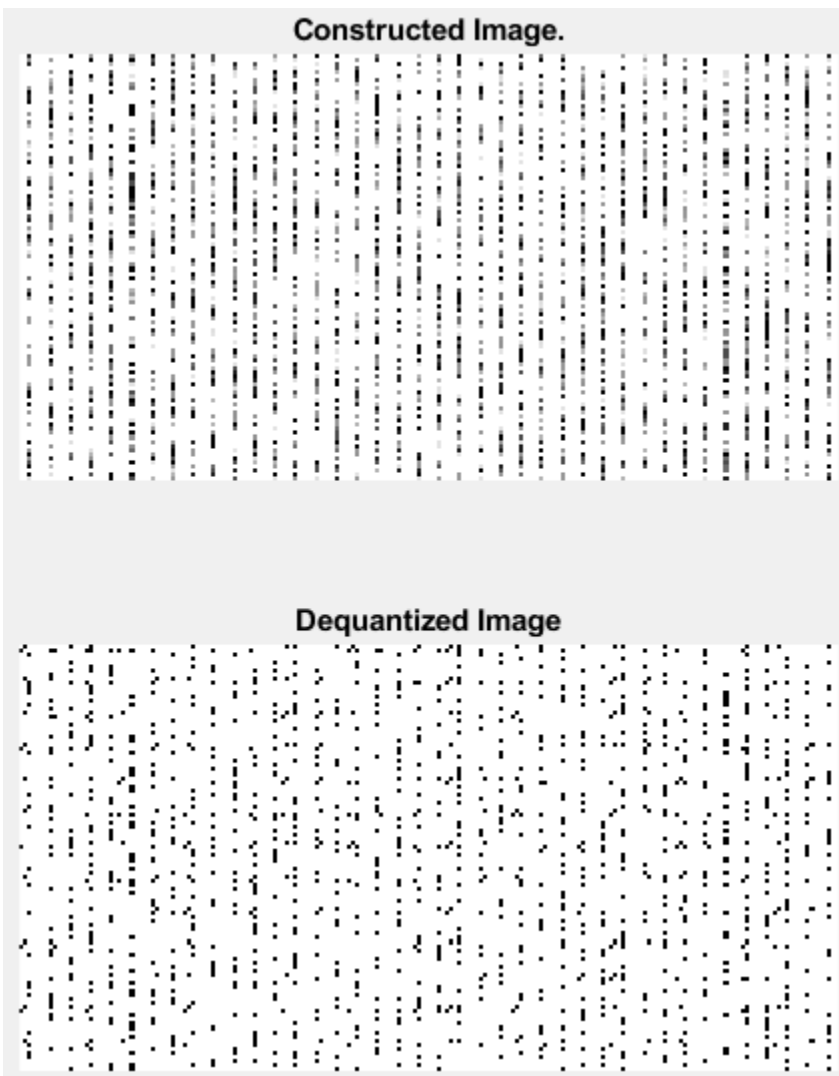
Υ.Γ.: Στο συγκεκριμένο βίντεο που επιλέξαμε έτυχε το τελευταίο πλαίσιο δείχνει μόνο το παρασκήνιο και έτσι καταφέραμε να ανιχνεύσουμε ποιο μπλοκ κίνησης αντιστοιχεί σε ποιο μπλοκ παρασκηνίου. Το τελευταίο πλαίσιο:



- Άσκηση Τρίτη
Η εικόνα που κατασκευάσαμε είναι η εξής (σε κάθε επανάληψη του κώδικα αλλάζει καθώς βασίζεται σε τυχαία δεδομένα):



Η «από-κβαντισμένη» εικόνα που προκύπτει είναι η εξής:



Είναι φανερό πως υπάρχει έλλειψη πληροφορίας παρόλο που εφαρμόζεται μια διαδικασία και στη συνέχεια η αντίστροφή της, και άρα θα έπρεπε η εικόνα που προκύπτει να είναι ίδια με την αρχική.

Έπειτα αφού εφαρμόζεται κωδικοποίηση Huffman, προκύπτει ο λόγος συμπίεσης:

```
>> constructedIMG
The compression ratio of the quantized DCT image is:

compressionRatio =

    0.9395
```

Ο λόγος που προκύπτει είναι αρκετά κοντά στο 1 γεγονός που δείχνει ότι πιθανώς η συμπίεση να μην ήταν πολύ αποτελεσματική, ειδικά η κατασκευασμένη εικόνα είχε μεγάλη εντροπία και συμπίεστηκε με δυσκολία.

Βιβλιογραφία

1. Συστήματα Πολυμέσων: Αλγόριθμοι, Πρότυπα και Εφαρμογές Parag Havaladar, Gerard Medioni.
2. Computer Vision Toolbox: Matlab Documentation
<https://www.mathworks.com/help/vision/>
3. Communications Toolbox: Matlab Documentation
<https://www.mathworks.com/help/comm/>

Παραρτήματα

Στην άσκηση 8.17β και στο μέρος β' χρησιμοποιούνται κάποιες συναρτήσεις τις οποίες υλοποιήσαμε για τα απαιτούμενα της κάθε άσκησης.

8.17β

1. kCoordinates(row,col)

Η συνάρτηση αυτή εξυπηρετεί το σκοπό του να προσδιορισθεί με ποια μπλοκ μπορεί να συγκριθεί το εκάστοτε μπλοκ που εξετάζεται.

Παραδείγματος χάριν, αν το μπλοκ που εξετάζεται είναι το πρώτο πάνω αριστερά στο πλαίσιο (frame_block(1,1)) τότε αυτό μπορεί να συγκριθεί μόνο με το επόμενο μπλοκ στη σειρά (frame_block(1,2)) και το επόμενο μπλοκ στη στήλη (frame_block(2,1)).

Αντίστοιχα αν το μπλοκ που εξετάζεται είναι στη μέση του πλαισίου (frame_block(r,c))

Τότε αυτό το μπλοκ μπορεί να συγκριθεί και με τα 4 γειτονικά του μπλοκ. Η συνάρτηση αυτή επιστρέφει μια λίστα με τις συντεταγμένες των μπλοκ που μπορούν να εξετασθούν.

2. `minimumSAD(sad)`

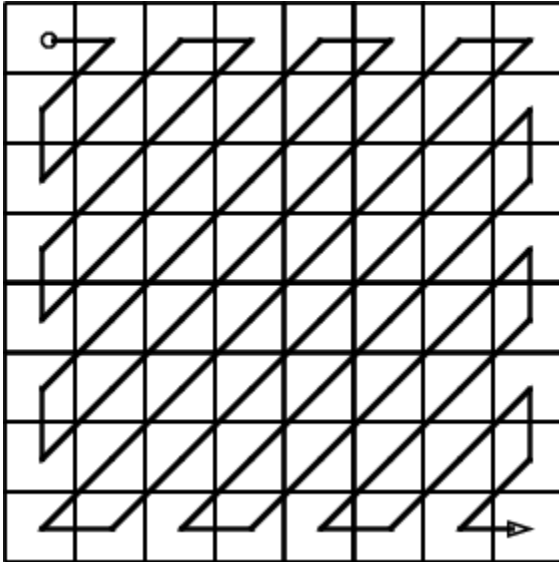
Η συνάρτηση αυτή παίρνει ως όρισμα τον πίνακα κελιών στον οποίο έχουν αποθηκευτεί οι πίνακες SAD - μπλοκ που συγκρίθηκαν με το εκάστοτε μπλοκ. Η συνάρτηση αυτή επιστρέφει τον μικρότερο πίνακα εξ αυτών.

Μέρος β΄

3. `zigzag(img)`

Η συνάρτηση αυτή μετατρέπει την εικόνα που δίνεται ως όρισμα σε διάνυσμα.

Διαβάζει και επιστρέφει τα στοιχεία του πίνακα που παριστάνει την εικόνα με τον εξής τρόπο:



Άρα ο Huffman κωδικοποιεί ένα διάνυσμα αντί για έναν πίνακα, διευκολύνοντας έτσι τη διαδικασία.