

Κβαντοποίηση Διανυσμάτων σε Κωδικοποιητές Βίντεο

Καλός Πέτρος

Τμήμα Μηχανικών Η/Υ Τηλεπικοινωνιών και Δικτύων
Πανεπιστήμιο Θεσσαλίας
Ιούνιος 2013

- Περιγραφή Προβλήματος
- Ψηφιακό Βίντεο
- Θεωρία Πληροφοριών
- K-means
- VQ Training
- Τροποποίηση JM H.264
- Αποτελέσματα VQ H.264
- VQ H.264 vs. JM H.264
- Συμπεράσματα

Περιγραφή Προβλήματος

- Βίντεο συνεπάγεται τεράστιος όγκος δεδομένων
- Τεχνικές Συμπίεσης
 - Με απώλειες
 - Χωρίς απώλειες
- Η ευαισθησία του ανθρώπινου ματιού είναι μικρότερη των 38dB
- Μεγάλη πολυπλοκότητα σημερινών τεχνικών συμπίεσης

Ασυμπίεστη Εικόνα



43.07dB



38.19dB



32.76dB

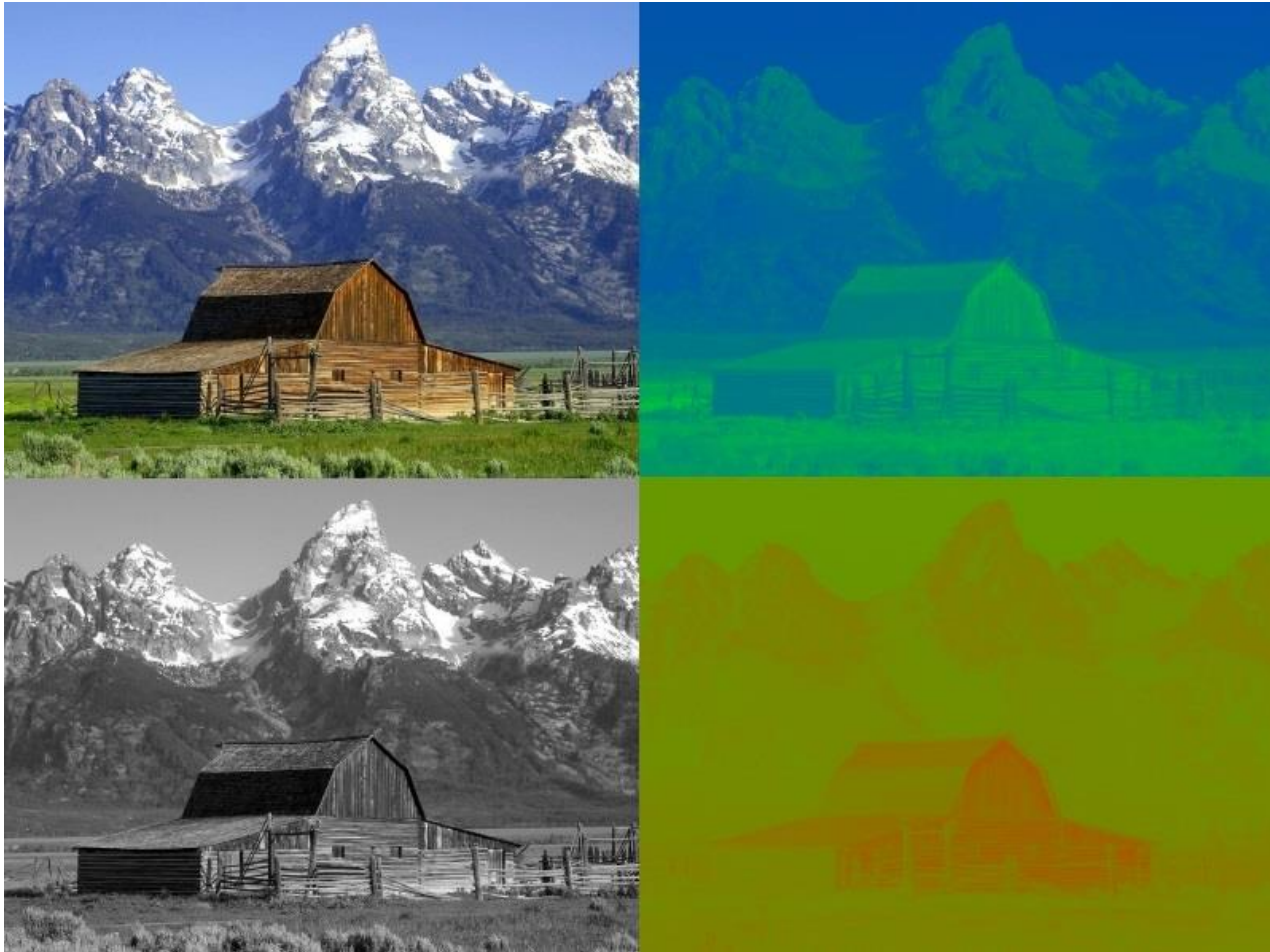


- Περιγραφή Προβλήματος
- Ψηφιακό Βίντεο
- Θεωρία Πληροφοριών
- K-means
- VQ Training
- Τροποποίηση JM H.264
- Αποτελέσματα VQ H.264
- VQ H.264 vs. JM H.264
- Συμπεράσματα

Ψηφιακό Βίντεο

- Αποτελείται από μία σειρά καρέ που αναπαράγονται με σταθερό ρυθμό (25 ή 30Hz)
- Καρέ είναι μια σειρά από pixels τοποθετημένα στον δυσδιάστατο χώρο. Οι διαστάσεις του καθορίζουν την ανάλυση του βίντεο
- Κάθε pixel έχει ένα βάθος (8 – 14bits)
- Το κάθε καρέ απεικονίζεται σε ένα χώρο χρωμάτων που ονομάζεται YUV, όπου το Y είναι η φωτεινότητα και το U,V η χρωματικότητα

Συνιστώσες ΥΥΥ



Τοποθέτηση των pixels

- Για να δημιουργήσουν ένα καρέ τα τοποθετούμε με διάφορους τρόπους (πχ YUV420,YUV444)

Single Frame YUV420:

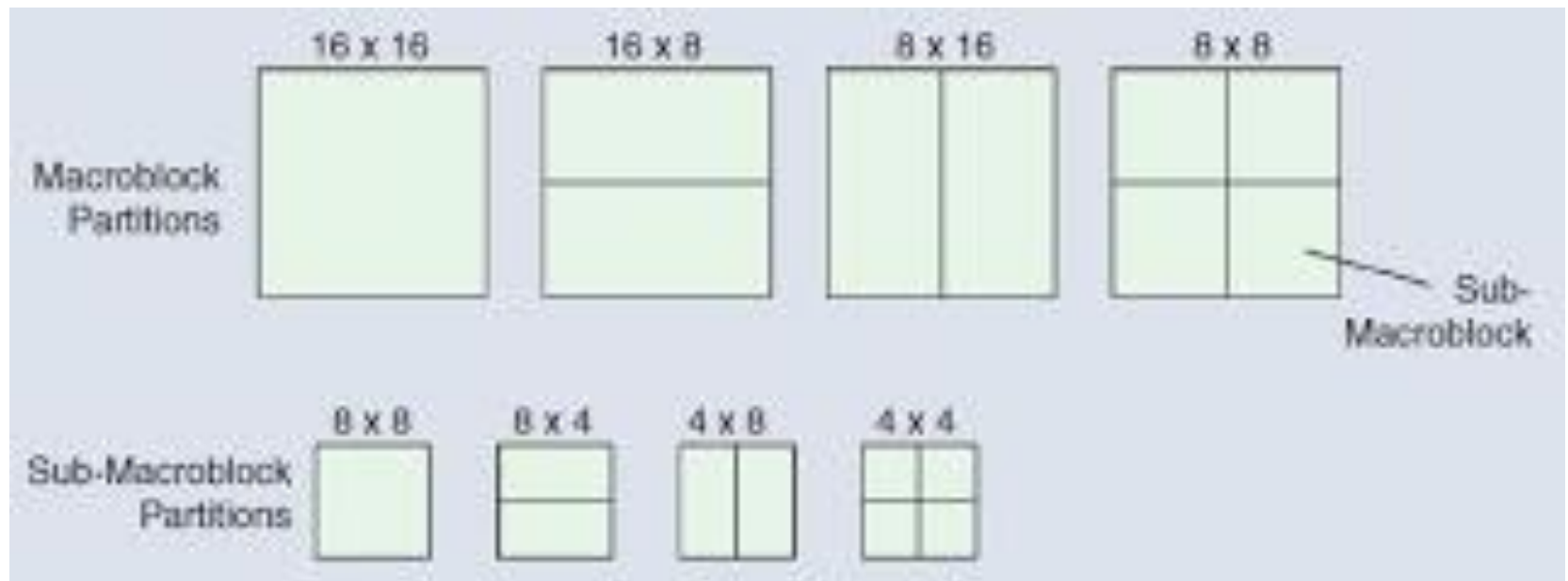


Position in byte stream:



Οργάνωση των pixels

- Οργάνωση σε macroblocks, blocks, subblocks

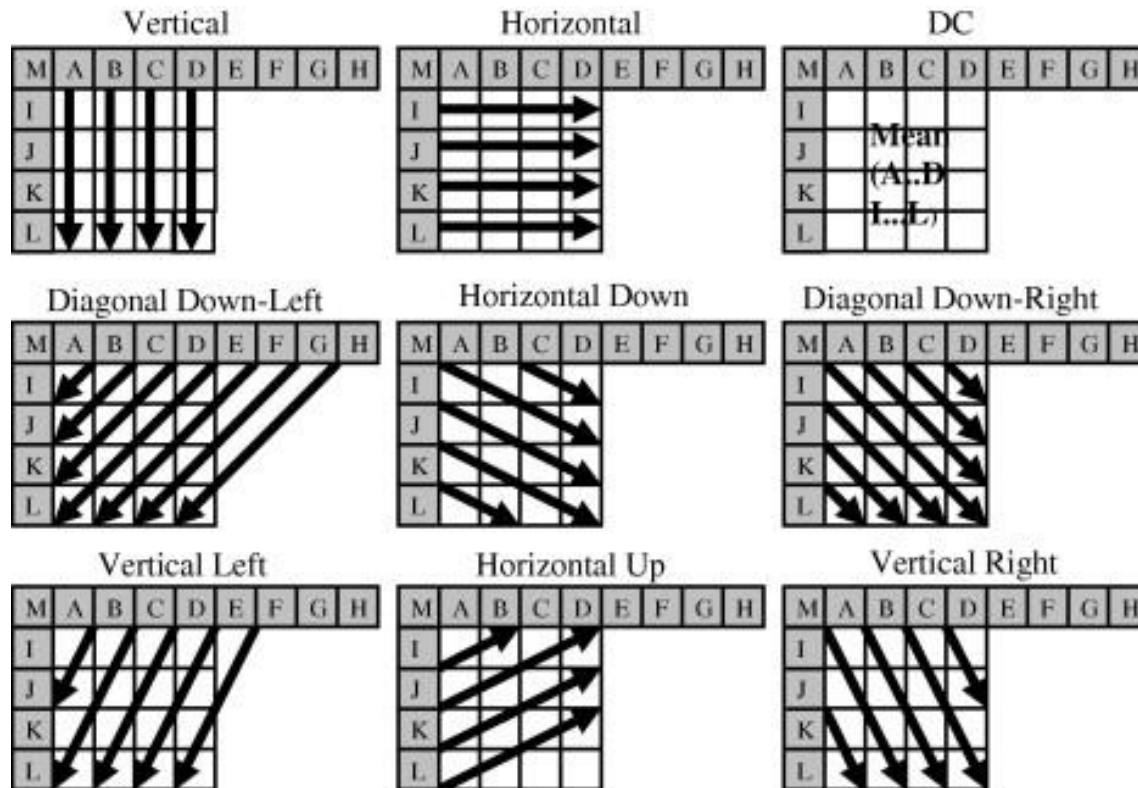


Οργάνωση των καρτέ

- Intra (Temporal)
 - I frames
- Inter (Special)
 - P,B frames
- Στόχος η δημιουργία διαφορών pixel (residuals)

Intra frames

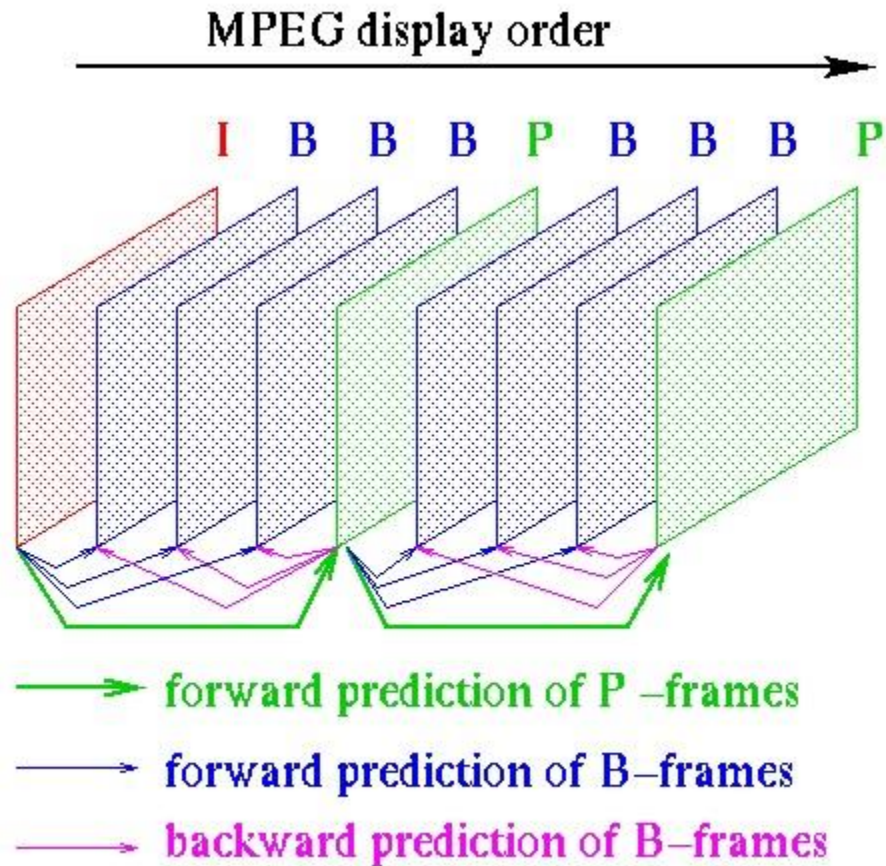
- Χρήση πληροφορίας μόνο εντός καρτέ
- Intra prediction modes



Inter frames

- *P (predictive) frames*
 - Δημιουργία διαφορών παίρνοντας ως pixels αναφοράς pixels από ένα συγκεκριμένο προηγούμενο καρέ
- *B (bidirectional) frames*
 - Δημιουργία διαφορών παίρνοντας ως pixels αναφοράς τον μέσο όρο των pixels από προηγούμενα ή επόμενα καρέ
- Motion Vectors

GOP



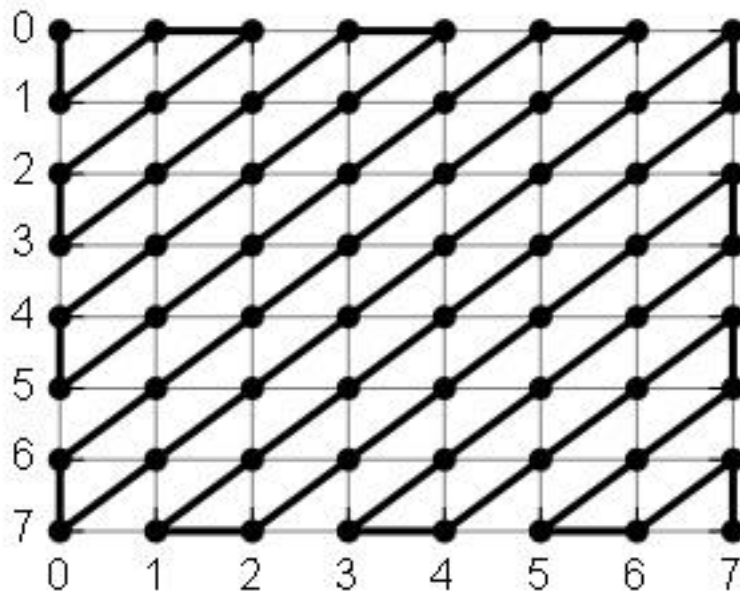
Encoding

- Μετασχηματισμός DCT 4x4, 8x8, 16x16
- ✓ Κβαντοποίηση
- Zigzag Scan
- Run Length Encoding
- Entropy encoding

Κβαντοποίηση

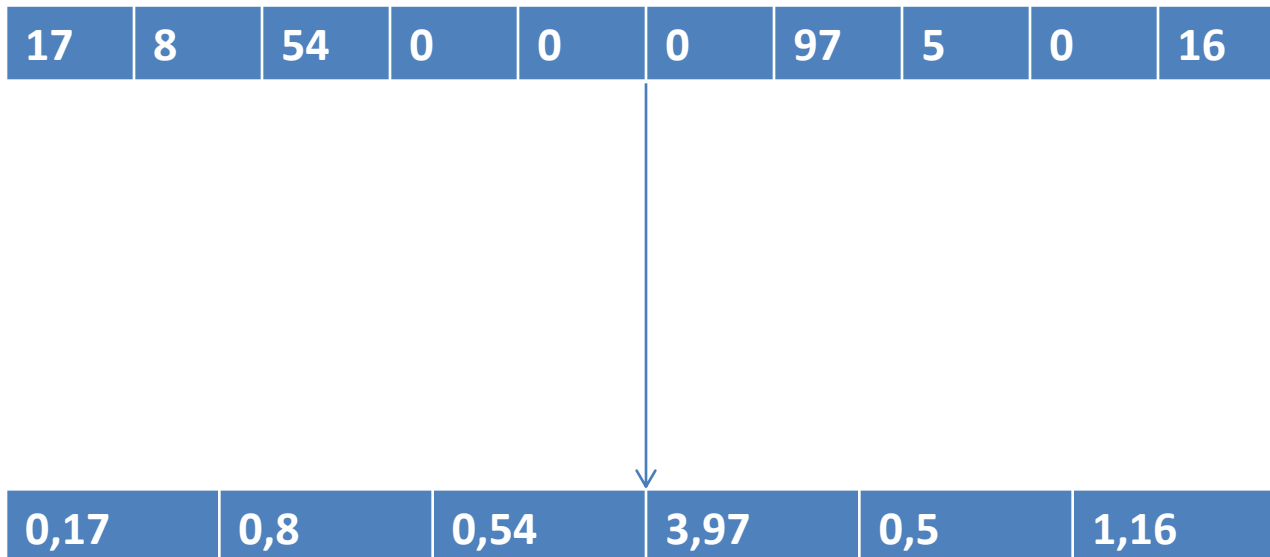
- Εισαγωγή σφάλματος
- Ακέραια διαίρεση συντελεστών DCT με κάποια ακέραια τιμή, πιθανόν διαφορετική για κάθε συντελεστή
- Quantization Parameter (QP) καθορίζει την ποιότητα

ZigZag Scan



Run Length Encoding

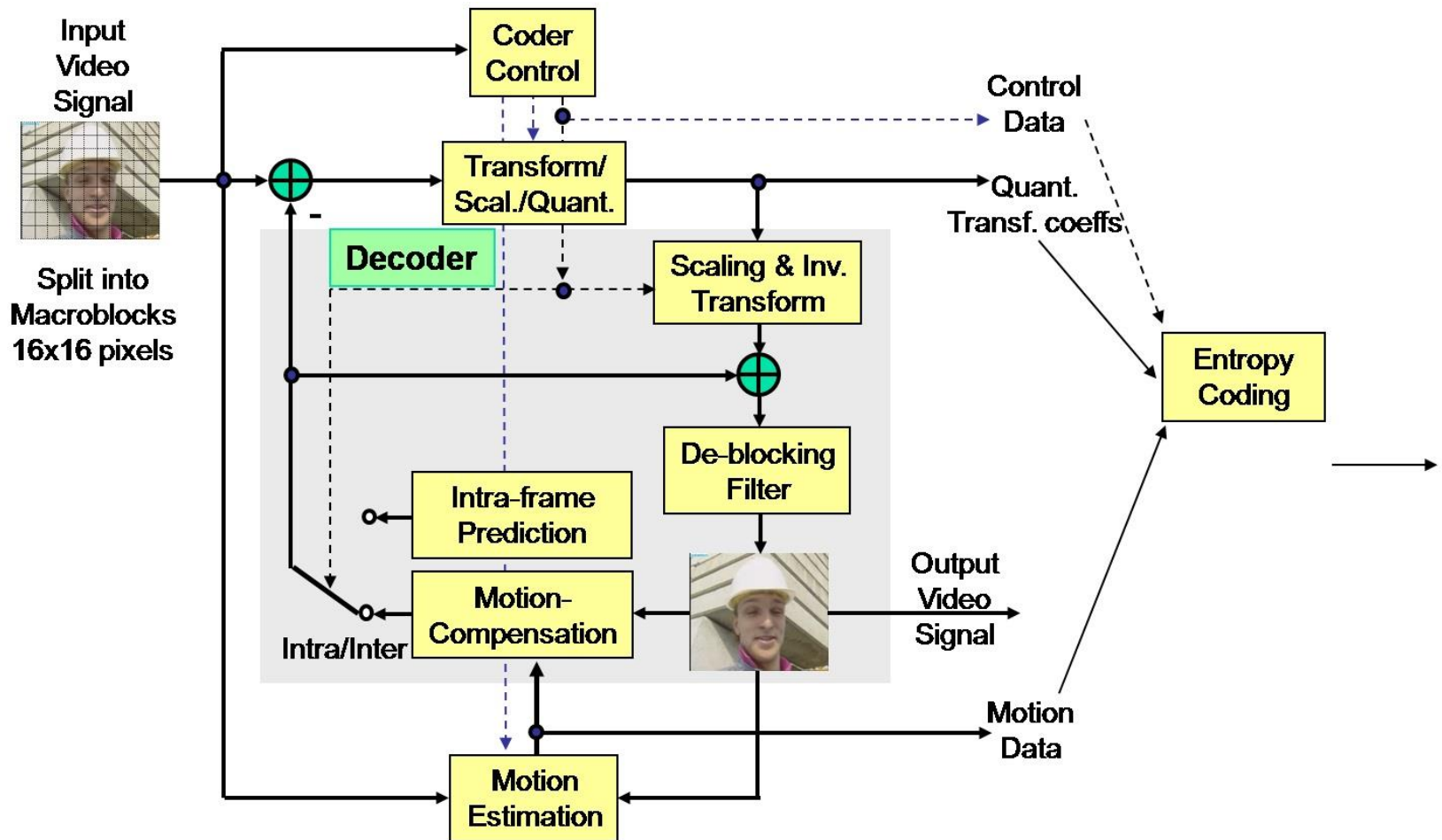
- Μείωση αριθμών προς κωδικοποίηση



Ποιότητα Βίντεο

- $PSNR = 10 \times \log_{10} \left(\frac{MAX_i^2}{MSE} \right)$
 - $MSE = \frac{\sum_{i=0}^{X*Y} (Source_i - Reconstructed_i)^2}{X*Y}$
 - $MAX_i^2 = bitdepth^2 - 1$
- Υπολογίζεται για κάθε συνιστώσα YUV ξεχωριστά αλλά ως μετρική λαμβάνεται το PSNR του Y

Δομή H.264



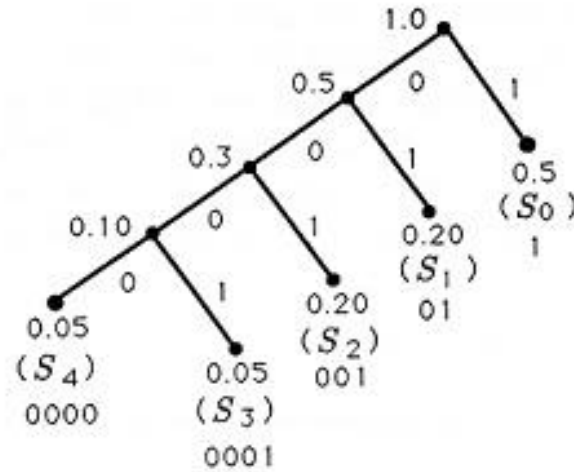
- Περιγραφή Προβλήματος
- Ψηφιακό Βίντεο
- Θεωρία Πληροφοριών
- K-means
- VQ Training
- Τροποποίηση JM H.264
- Αποτελέσματα VQ H.264
- VQ H.264 vs. JM H.264
- Συμπεράσματα

Θεωρία Πληροφοριών

- Εντροπία
 - $H(X) = -\sum_{i=1}^n (p(x_i) \times \log_b p(x_i))$, $p(x_i)$ η πιθανότητα του ενδεχομένου x_i
 - Το απόλυτο κάτω όριο που η πληροφορία μίας πηγής μπορεί να συμπιεστεί

Κωδικοποιητές Εντροπίας

- Μέθοδος Huffman
 - Μικρή πολυπλοκότητα
 - $H(X) \leq L_c \leq H(X) + 1bit$

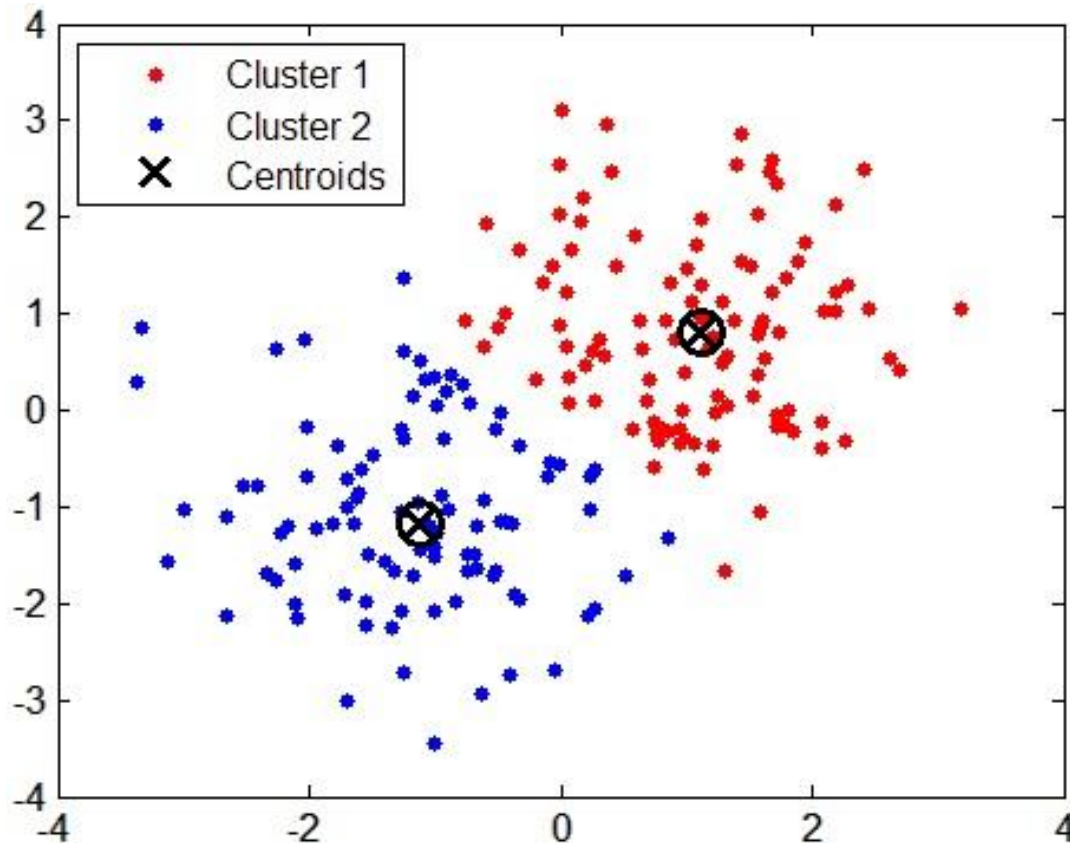


- Αριθμητική Κωδικοποίηση
 - Context Adaptive Binary Arithmetic Encoding (CABAC)
 - Μεγάλη πολυπλοκότητα
 - Πλησιάζει “κοντά” στο όριο εντροπίας

- Περιγραφή Προβλήματος
- Ψηφιακό Βίντεο
- Θεωρία Πληροφοριών
- **K-means**
- VQ Training
- Τροποποίηση JM H.264
- Αποτελέσματα VQ H.264
- VQ H.264 vs. JM H.264
- Συμπεράσματα

K-means

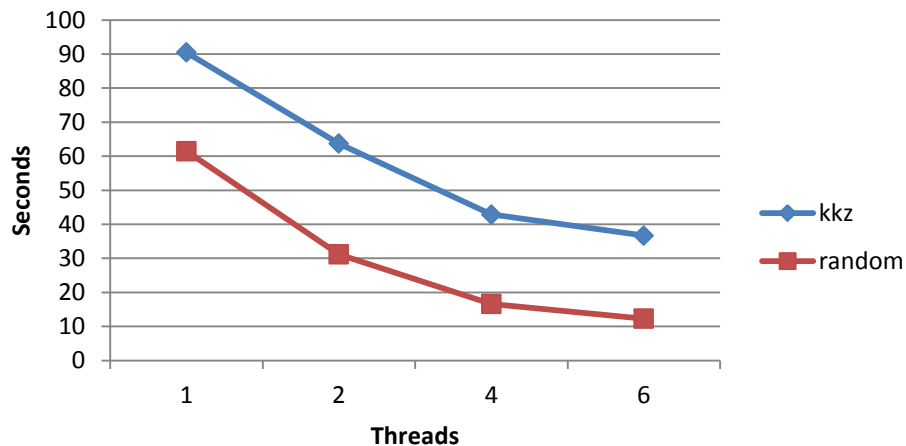
- Επαναληπτικός αλγόριθμος που χωρίζει με το ελάχιστο σφάλμα n σημεία σε διάσταση χώρου R^d σε k περιοχές $k \leq d$



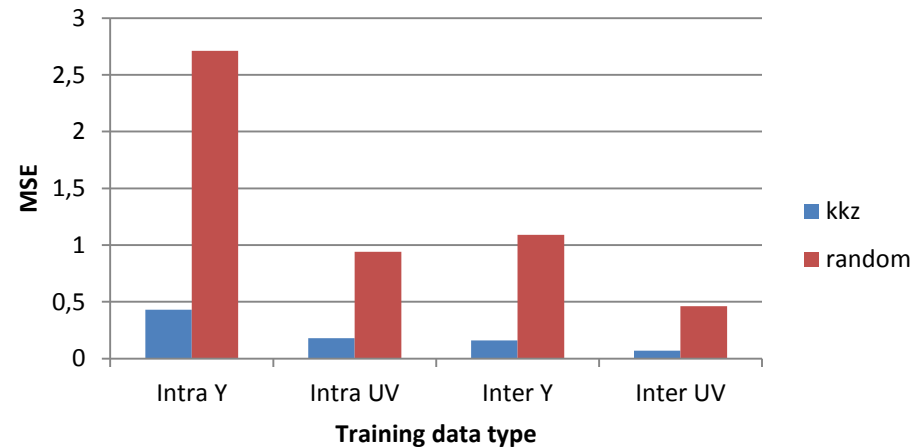
Αρχικοποίηση K-means

- Random η με κάποια στρατηγική
 - $n=100000, k=65536, d=16$

KKZ vs. Random Duration

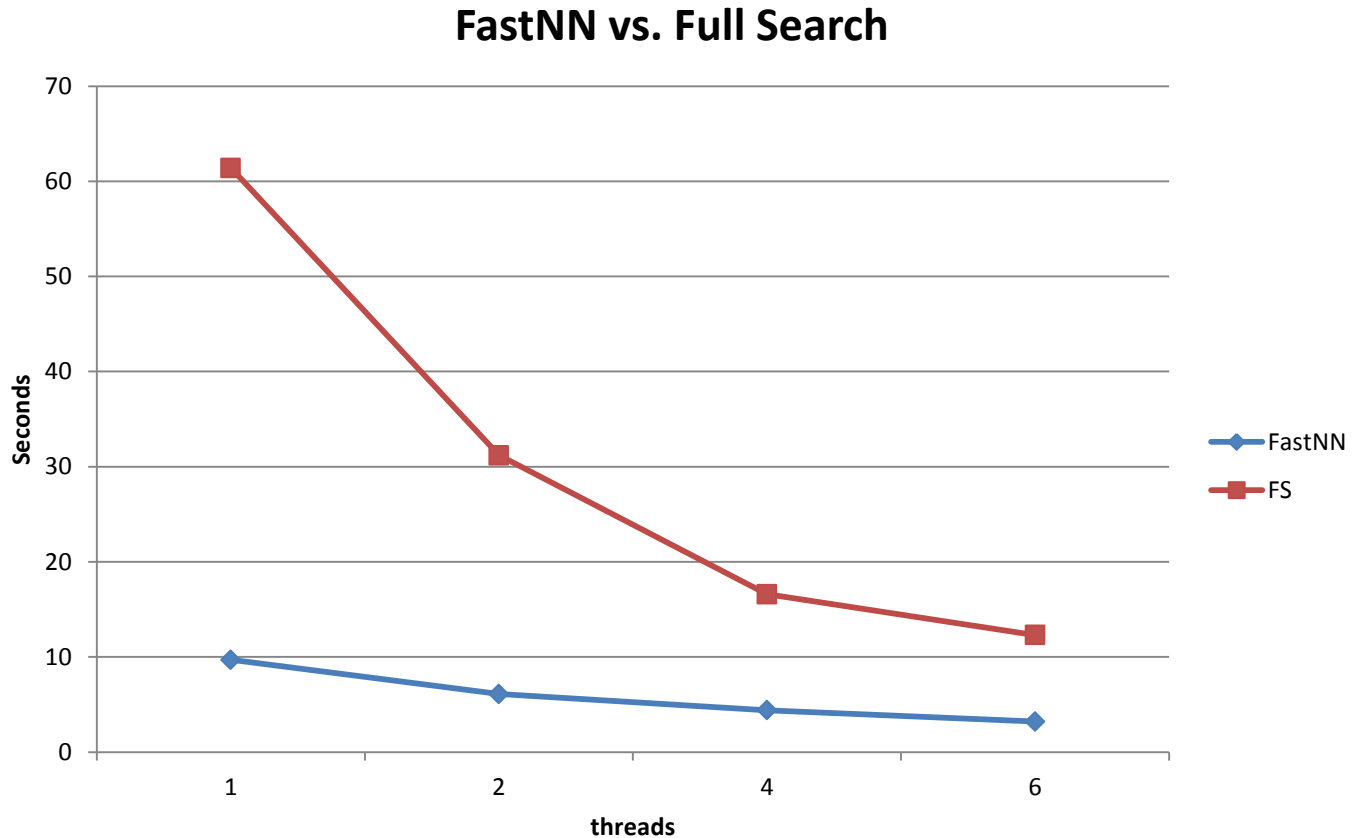


KKZ vs. Random MSE (5 iterations)



Αναζήτηση κοντινότερου cluster

- Με Full Search η με FastNN



- Περιγραφή Προβλήματος
- Ψηφιακό Βίντεο
- Θεωρία Πληροφοριών
- K-means
- **VQ Training**
- Τροποποίηση JM H.264
- Αποτελέσματα VQ H.264
- VQ H.264 vs. JM H.264
- Συμπεράσματα

K-means Training

- Επιλογή των residuals ως training set
- Υπάρχουν τόσα residuals όσα και pixels
- Διαίρεση των καρέ σε $m \times m$ κομμάτια με $m=4$
- Χρησιμοποιήθηκαν 2600 καρέ από 10 βίντεο με διαφορετικό περιεχόμενο

Εξαγωγή του training set από τον H.264

- Τροποποίηση Decoder
- Βήματα για την εξαγωγή
 - Encoding σε lossless mode με δύο διαφορετικά GOP
 1. I-I-I-....
 2. I-P-P-B-P-P-B-...
 - Decoding
 - KeepI = 1, 1^ο GOP
 - KeepP = 1, 2^ο GOP
 - ~~• KeepB = 1, 2^ο GOP~~

Codebooks

- Codebooks για IntraY, UV και InterY, UV

Τύπος	d	n	k	Εντροπία	PSNR(dB)	Επαναλήψεις	Διάρκεια (minutes)
IntraY	16,000	56160000,000	65536	0,712229	33,6	3249	12154
IntraUV	16,000	28080000,000	65536	0,743071	42,1	2697	3119
InterY	16,000	42117616,000	65536	0,692577	40,5	3270	9120
InterUV	16,000	21058808,000	65536	0,707785	48,1	4221	8509

Εντροπία υπό συνθήκη

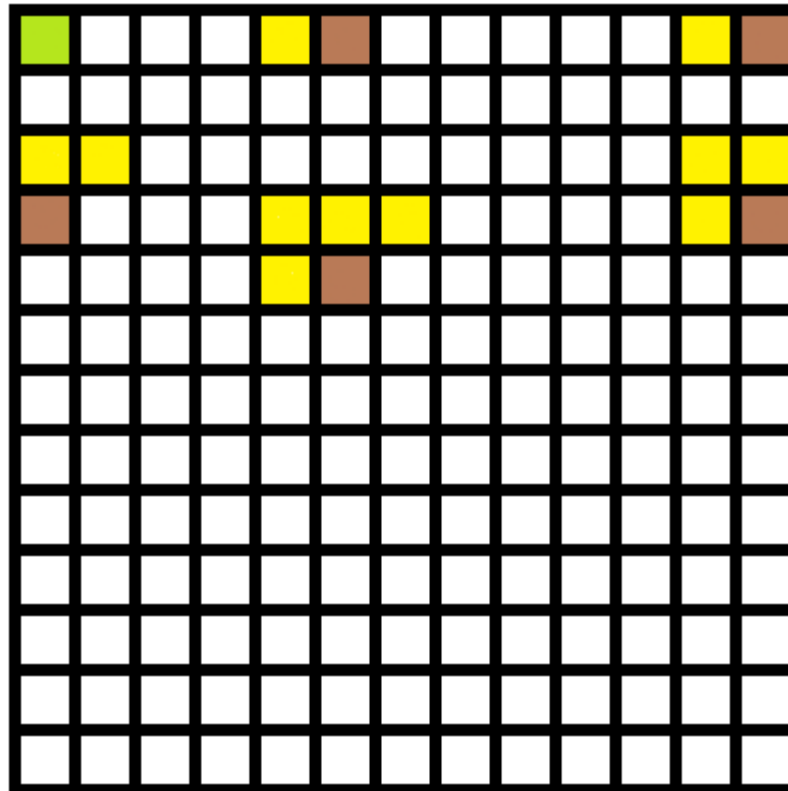
- Η πληροφορία για μια τυχαία μεταβλητή Y μπορεί μόνο να μας μειώσει την εντροπία της μεταβλητής X

$$H(X|Y) \leq H(X)$$

- Δημιουργία 8 ισοπίθανων περιοχών με βάση την ενέργεια των codewords
- Παραγωγή στατιστικών των 8 contexts με βάση την ενέργεια των γειτόνων

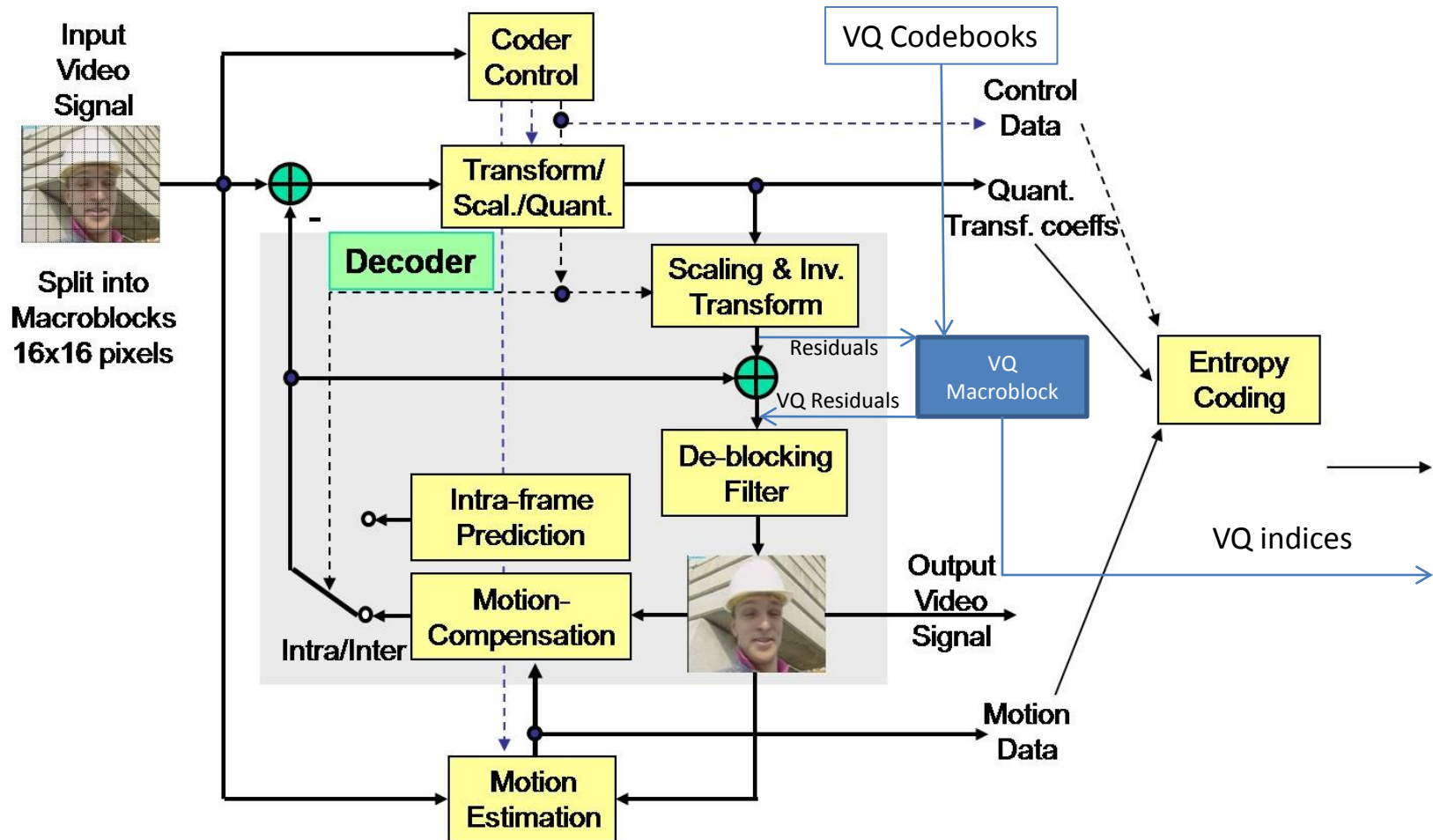
Contexts

I Y entropy	I UV Entropy	P Y Entropy	P UV Entropy
0,5400	0,6434	0,5443	0,6314

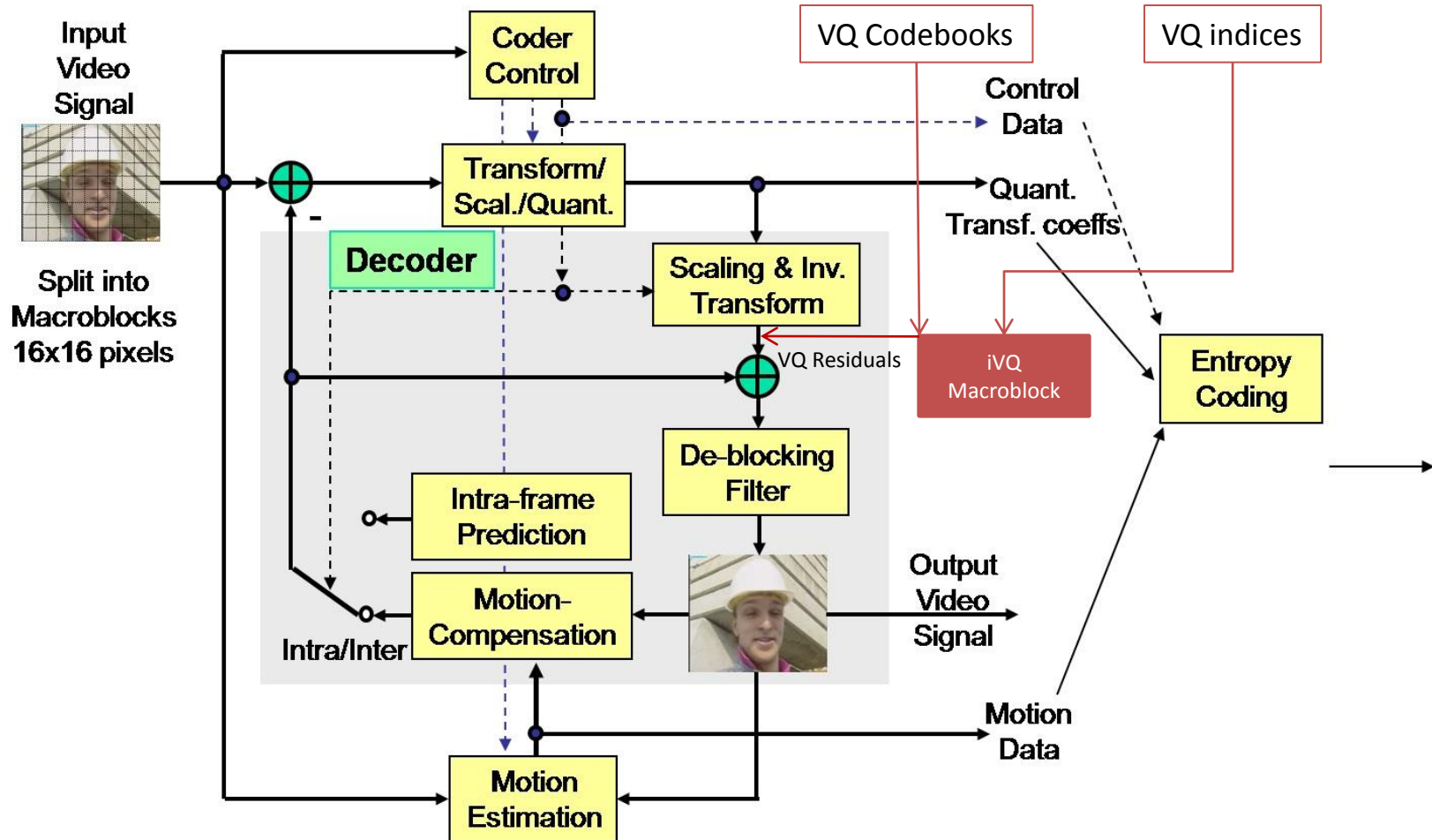


- Περιγραφή Προβλήματος
- Ψηφιακό Βίντεο
- Θεωρία Πληροφοριών
- K-means
- VQ Training
- Τροποποίηση JM H.264
- Αποτελέσματα VQ H.264
- VQ H.264 vs. JM H.264
- Συμπεράσματα

VQ H.264 Encoder

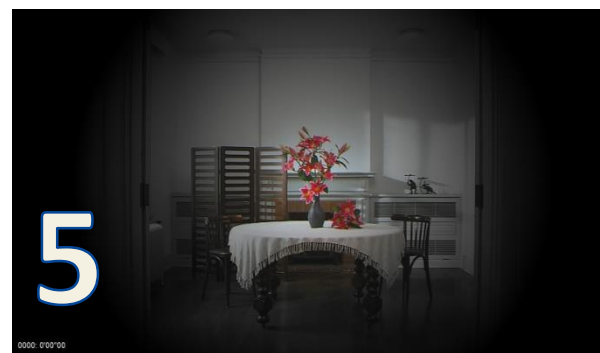


VQ H.264 Decoder



- Περιγραφή Προβλήματος
- Ψηφιακό Βίντεο
- Θεωρία Πληροφοριών
- K-means
- VQ Training
- Τροποποίηση JM H.264
- Αποτελέσματα VQ H.264
- VQ H.264 vs. JM H.264
- Συμπεράσματα

VQ H.264 vs. JM H.264

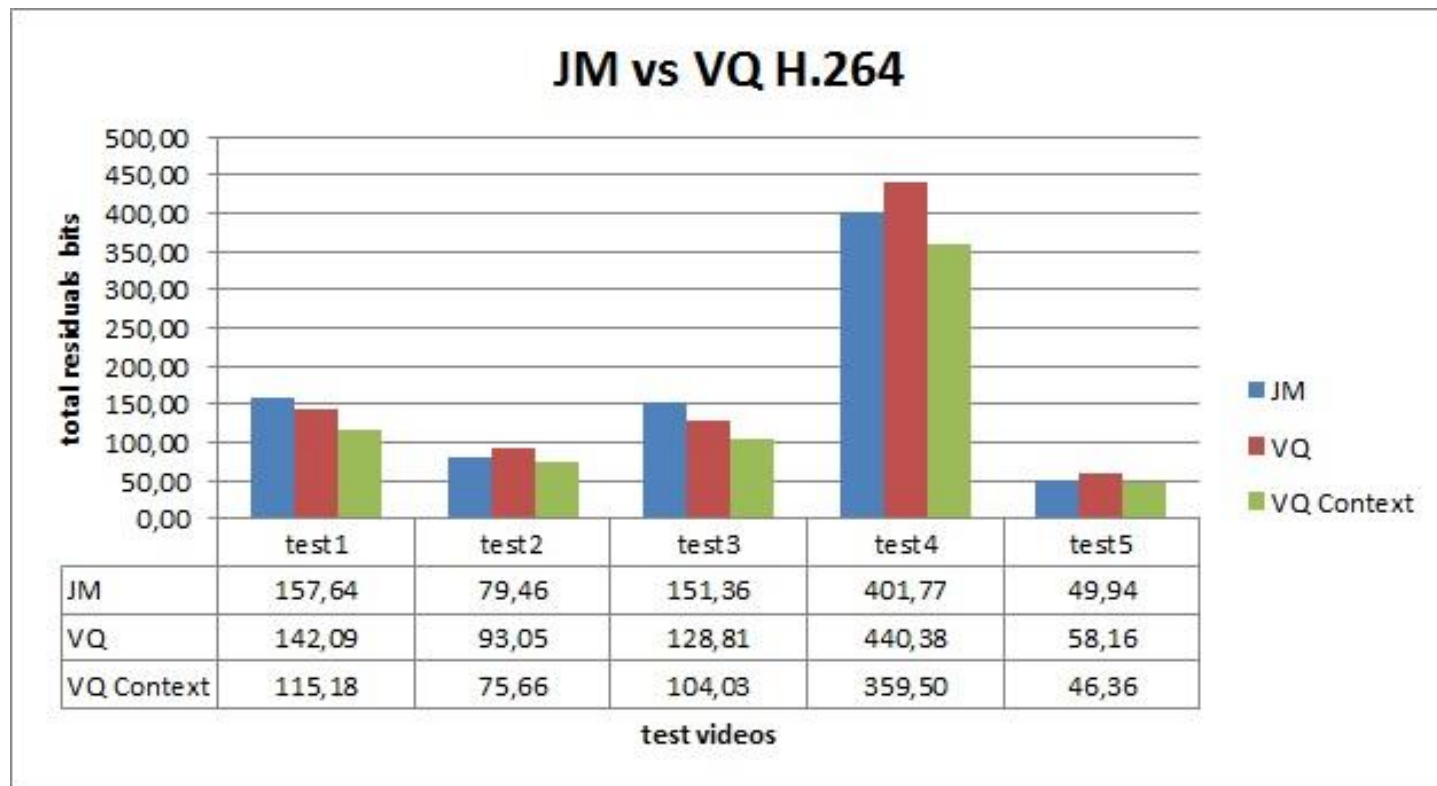


Αποτελέσματα VQ H.264

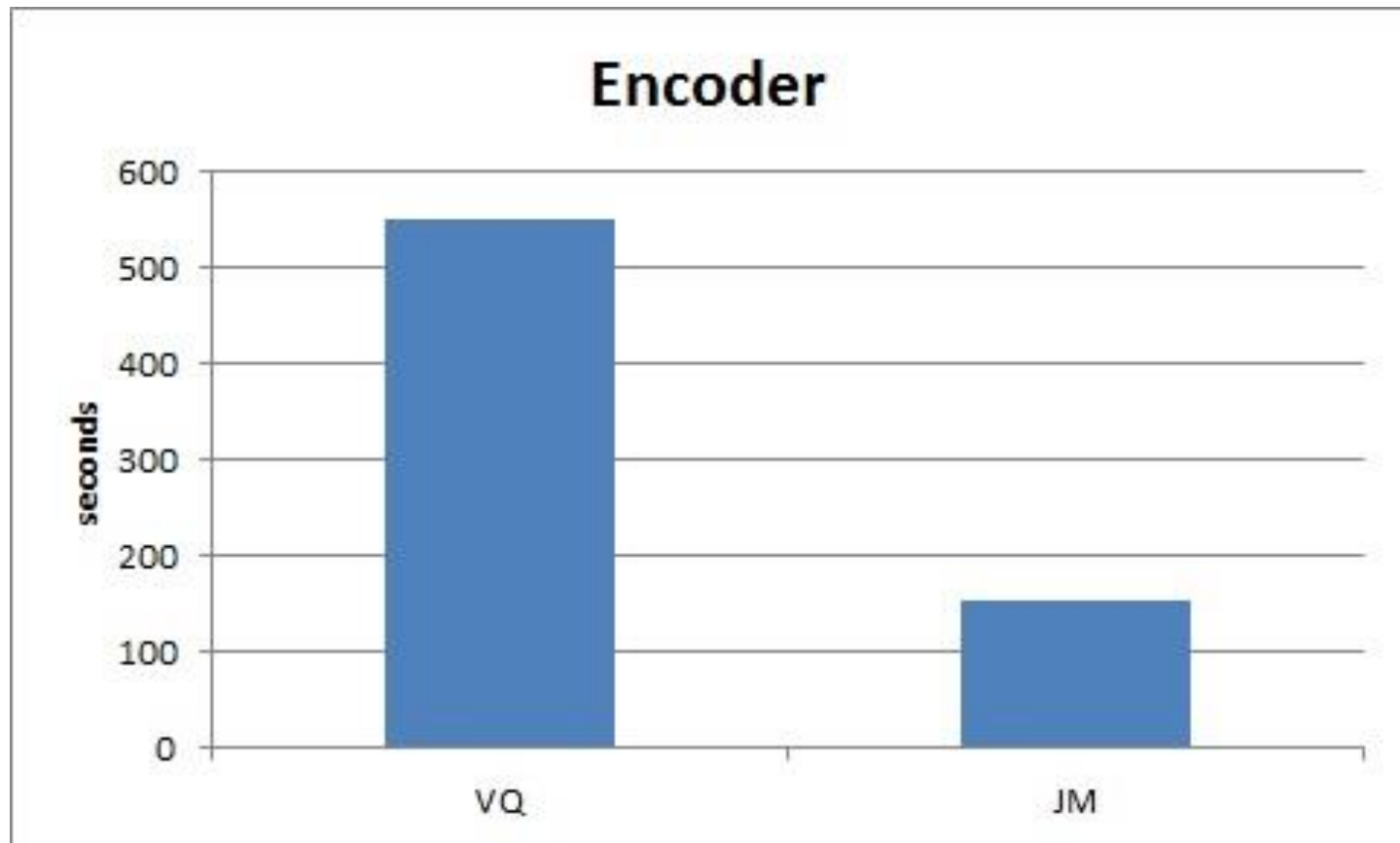
	YI	YU	YV	PI	PU	PV	BI	BU	BV
test1	35,2	39,72	39,67	41,2	43,91	43,87	43,00	45,1	45,22
test2	35,5	41,1	42,83	38,9	41,53	43,28	40,38	43	44,71
test3	30,7	39,86	41	36,7	42,64	43,76	38,00	43,7	44,90
test4	44,1	47,36	47,91	46	47,2	47,75	46,55	47,4	48,05
test5	37	50,12	48,7	44	49,9	49,65	44,78	50,5	50,29

- Περιγραφή Προβλήματος
- Ψηφιακό Βίντεο
- Θεωρία Πληροφοριών
- K-means
- VQ Training
- Τροποποίηση JM H.264
- Αποτελέσματα VQ H.264
- VQ H.264 vs. JM H.264
- Συμπεράσματα

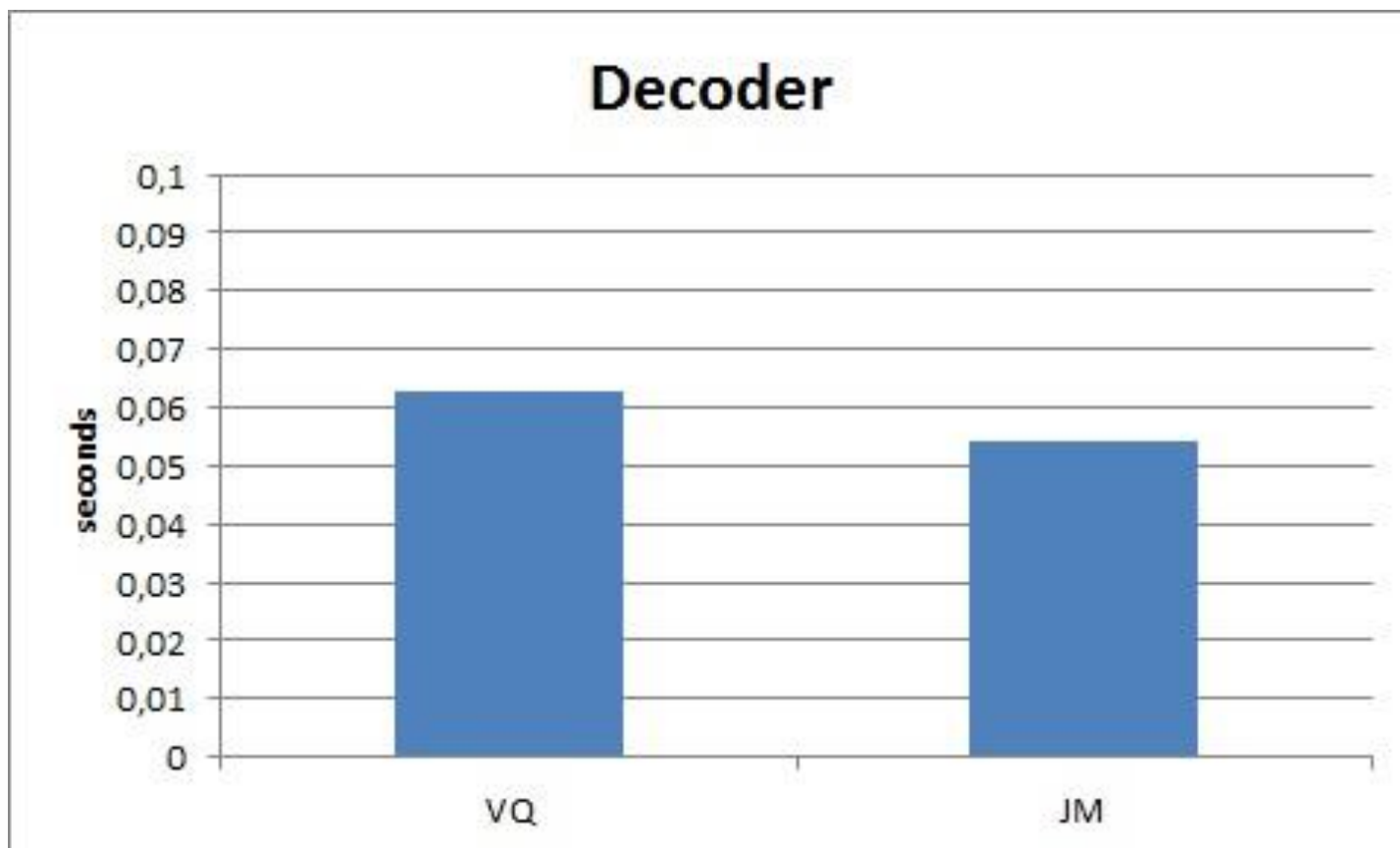
Σύγκριση Mbits



Σύγκριση Πολυπλοκότητας Encoding



Σύγκριση Πολυπλοκότητας Decoding



- Περιγραφή Προβλήματος
- Ψηφιακό Βίντεο
- Θεωρία Πληροφοριών
- K-means
- VQ Training
- Τροποποίηση JM H.264
- Αποτελέσματα VQ H.264
- VQ H.264 vs. JM H.264
- Συμπεράσματα

Συμπεράσματα

- ✓ Καλύτερη απόδοση του VQ και ιδιαίτερα στα δύσκολα βίντεο
- ✓ Εύκολη και γρήγορη βελτίωση των codebooks
- ✓ Μείωση πολυπλοκότητας του decoder
- Μεγάλη αύξηση της πολυπλοκότητας του encoder
- Μεγάλη απόκλιση του PSNR κάθε συνιστώσας