



A R I S T O T L E  
U N I V E R S I T Y  
O F T H E S S A L O N I K I

Τμήμα Ηλεκτρολόγων Μηχανικών  
και Μηχανικών Υπολογιστών

Απλοποιημένος κωδικοποιητής/αποκωδικοποιητής AAC  
Συστήματα Πολυμέσων - Ομαδική Εργασία

Βαπόρης Δημήτριος	AEM 10625
Ντελόπουλος Εμμανουήλ	AEM 10693
Εργαστηριακός Υπεύθυνος	Αλέτρας Δημήτριος
Υπεύθυνος Καθηγητής	Ντελόπουλος Αναστάσιος

Χειμερινό Εξάμηνο 2025-2026

## Περιεχόμενα

<b>Περιεχόμενα</b>	<b>1</b>
<b>1 Εισαγωγή</b>	<b>2</b>
<b>2 Επίπεδο 1</b>	<b>2</b>
<b>3 Επίπεδο 2</b>	<b>2</b>
<b>4 Επίπεδο 3</b>	<b>2</b>
<b>5 Σύνοψη</b>	<b>7</b>
<b>Κατάλογος σχημάτων</b>	<b>8</b>

## 1 Εισαγωγή

Το παρόν αποτελεί την ηλεκτρονική αναφορά της εργασίας του μαθήματος Συστήματα Πολυμέσων με θέμα την υλοποίηση μίας απλοποιημένης εκδοχής κωδικοποιητή/αποκωδικοποιητή ήχου σύμφωνα με το πρότυπο Advanced Audio Coding (AAC). Η υλοποίηση χωρίζεται σε 3 διαφορετικά επίπεδα, διαμορφώνοντας τις 3 παρακάτω ενότητες στις οποίες παρατίθενται οι απαραίτητοι σχολιασμοί για κάθε επίπεδο. Στο 3ο επίπεδο, το οποίο αποτελεί την ολοκληρωμένη έκδοση του κωδικοποιητή και του αποκωδικοποιητή, επιδεικνύονται και μερικά ενδεικτικά αποτελέσματα από τα δοκιμαστικά προγράμματα που χρησιμοποιήθηκαν.

## 2 Επίπεδο 1

Το πρώτο επίπεδο περιλαμβάνει τις βαθμίδες Sequence Segmentation Control (SSC) και Filterbank και την αντίστροφή της. Περιλαμβάνονται ο ολοκληρωμένος κωδικοποιητής και αποκωδικοποιητής του επιπέδου αυτού, όπως και ένα δοκιμαστικό αρχείο `demo_aac_1.py`, που επιδεικνύει την κωδικοποίηση και την αποκωδικοποίηση του 1ου επιπέδου. Μέχρι το σημείο αυτό δεν υπάρχει συμπίεση και άρα απώλεια πληροφορίας, με αποτέλεσμα ο θόρυβος του αποκωδικοποιημένου σήματος να είναι μηδέν και άρα το SNR (signal to noise ratio) που επιστρέφεται να τείνει στο άπειρο (λόγω μικρών αριθμητικών αποκλίσεων το SNR που επιστρέφεται είναι στην πραγματικότητα 253.99dB). Εντός του φακέλου `level_1`, τρέχοντας το `demo_aac_1.py` γίνεται η δοκιμαστική κωδικοποίηση και αποκωδικοποίηση του ενδεικτικού αρχείου `LicorDeCalandraca.wav`, παράγεται και αποθηκεύεται η έξοδος `output_1.wav` και εκτυπώνεται το SNR.

## 3 Επίπεδο 2

Στο δεύτερο επίπεδο υλοποιούνται οι βαθμίδες Temporal Noise Shaping (TNS) και η αντίστροφή της, σύμφωνα με τις προδιαγραφές της εκφώνησης. Οι βαθμίδες αυτές σε συνδυασμό με αυτές του επιπέδου 1 δημιουργούν το ζευγάρι κωδικοποιητή-αποκωδικοποιητή του επιπέδου 2, η λειτουργία του οποίου επιδεικνύεται με τρόπο πλήρως ανάλογο με αυτόν του επιπέδου 1, με το αρχείο `demo_aac_2.py`. Εξακολουθεί να μην υπάρχει συμπίεση και απώλεια πληροφορίας, οδηγώντας πάλι σε ιδιαίτερα υψηλές τιμές SNR (και πάλι 253.99dB). Εντός του φακέλου `level_2`, τρέχοντας το `demo_aac_2.py` γίνεται η δοκιμαστική κωδικοποίηση και αποκωδικοποίηση του ενδεικτικού αρχείου `LicorDeCalandraca.wav`, παράγεται και αποθηκεύεται η έξοδος `output_2.wav` και εκτυπώνεται το SNR.

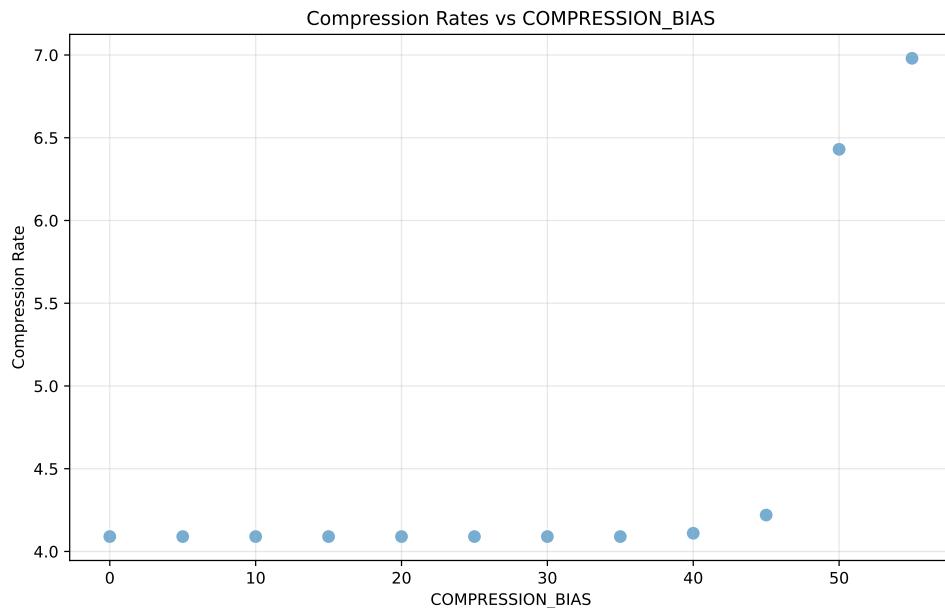
## 4 Επίπεδο 3

Στο 3ο και τελευταίο επίπεδο, προστίθενται οι βαθμίδες του ψυχοακουστικού μοντέλου, του κβαντιστή και του αποκβαντιστή, ενώ χρησιμοποιούνται και οι συναρτήσεις κωδικοποίησης και αποκωδικοποίησης Huffman που δίνονται στο αρχείο `huff_utils.py` έπειτα από τις

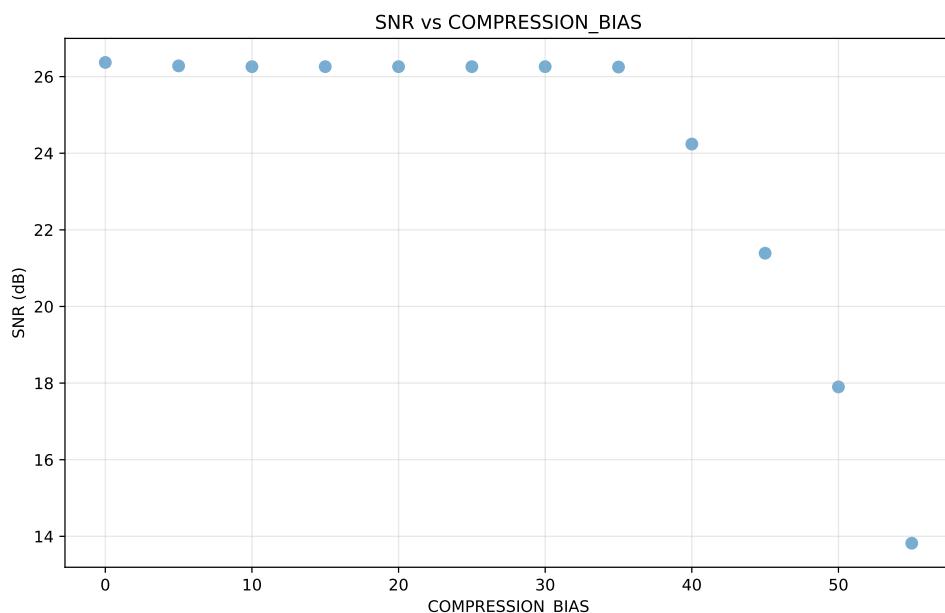
διορθώσεις που ακολούθησαν τη συζήτηση στο forum του μαθήματος. Οι βαθμίδες αυτές σε συνδυασμό με αυτές του επιπέδου 2 οδηγούν στον τελικό κωδικοποιητή και αποκωδικοποιητή του 3ου επιπέδου. Το αρχείο demo\_aac\_3.py επιδεικνύει τη λειτουργία του ζεύγους κωδικοποιητή-αποκωδικοποιητή του 3ου επιπέδου, με SNR μεταβαλλόμενο από την προαιρετική μεταβλητή COMPRESSION\_BIAS (που συζητείται παρακάτω). Για default τιμή COMPRESSION\_BIAS = 0, επιτεύχθη SNR ίσο με 26.37dB, συμπίεση κατά 4.09 και αποτέλεσμα χωρίς ιδιαίτερα προφανή θόρυβο, ενώ για τιμή ίση με 57 υπάρχει SNR 12.47dB, συμπίεση κατά 7.15 αλλά αισθητός θόρυβος στο αποτέλεσμα. Πέρα από την έξοδο output\_3.wav, το αποσυμπιεσμένο αρχείο ήχου που είναι αποτέλεσμα της αποκωδικοποίησης, περιλαμβάνεται στις εξόδους του δοκιμαστικού αρχείου το aac\_coded\_3.mat, που περιλαμβάνει την κωδικοποιημένη έξοδο της συνάρτησης aac\_coder\_3, η οποία καλείται στο demo.

Προστέθηκε μία περαιτέρω προαιρετική float μεταβλητή COMPRESSION\_BIAS με default τιμή 0 στη συνάρτηση aac\_quantizer (και σε όσες την καλούν, δηλαδή την demo\_aac\_3). Η λειτουργία της παραπάνω μεταβλητής είναι η εξής: προστίθεται στην αρχική προσέγγιση του Scalefactor Gain της κάθε μπάντας  $\hat{\alpha}(b)$ , ουσιαστικά μειώνοντας τα βήματα της επαναληπτικής αύξησης της προσέγγισης όσο αυξάνεται η τιμή της, άρα υπό περιπτώσεις αυξάνοντας την ισχύ του σφάλματος κραντισμού του τελικού scalefactor gain που επιλέγεται. Αυτό έχει ως αποτέλεσμα να αυξάνεται ο θόρυβος (να μειώνεται το SNR) του τελικού αποκωδικοποιημένου σήματος, αλλά να αυξάνεται και η τελική συμπίεση που μπορεί να επιτευχθεί, δημιουργώντας ένα trade-off μεταξύ συμπίεσης και αισθητικής του αποτελέσματος και μία ακόμα πηγή πειραματισμού, όπως φαίνεται παρακάτω.

Δημιουργήθηκε ένας ακόμη φάκελος με τίτλο plotting, που περιλαμβάνει 2 ακόμη δοκιμαστικά αρχεία. Το πρώτο από αυτά είναι το SNRandCRvsBIAS.py, η εκτέλεση του οποίου εντός του φακέλου folders τρέχει ως υποδιαδικασία το demo\_aac\_3.py για διάφορες τιμές της προαιρετικής μεταβλητής COMPRESSION\_BIAS και δημιουργεί τα παρακάτω διαγράμματα που δείχνουν τη σχέση μεταξύ Compression Rate, SNR και COMPRESSION\_BIAS. Παρατηρούμε ότι μέχρι την τιμή περίπου 40, δεν παρουσιάζεται σημαντική διαφορά ούτε στο SNR ούτε στη συμπίεση που πετυχαίνει η κωδικοποίηση. Ωστόσο, για μεγαλύτερες τιμές (κοντά στο 50), υπάρχει αύξηση τόσο στη συμπίεση όσο και στον θόρυβο του τελικού σήματος. Δε δοκιμάστηκαν τιμές άνω του 60.



Σχήμα 1: Σχέση μεταξύ COMPRESSION\_BIAS και Compression Rate



Σχήμα 2: Σχέση μεταξύ COMPRESSION\_BIAS και SNR

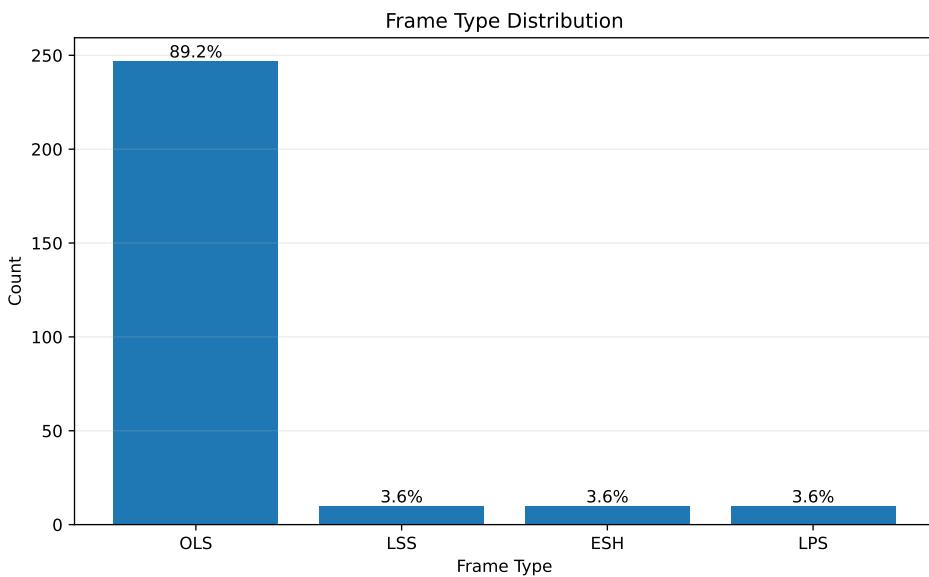
Το δεύτερο δοκιμαστικό αρχείο του φακέλου plotting είναι το generate\_report\_figures.py. Το αρχείο αυτό προϋποθέτει να υπάρχουν στον εκάστοτε φάκελο οι έξοδοι output\_1.wav, output\_2.wav και output\_3.wav που παράγονται από τα αρχεία demo. Η εκτέλεση ακολουθεί την εξής διαδικασία: εκτελείται σε κάθε έναν από τους φακέλους level\_1, level\_2 και level\_3 το αντίστοιχο demo (στην περίπτωση του level\_3, υπάρχει η δυνατότητα να μεταβληθεί το COMPRESSION\_BIAS χειροκίνητα στην αρχή του αρχείου demo\_aac\_3.py από την τιμή 57 που δίνεται). Έπειτα, στον φάκελο plotting εκτελείται το αρχείο generate\_report\_figures.py, το οποίο παράγει τα παρακάτω αποτελέσματα.

Αρχικά, παρατίθενται οι κυματομορφές του αρχικού σήματος και των παραγόμενων εξόδων. Η διαφορά μεταξύ της ανακατασκευής του επιπέδου 3 και των υπόλοιπων κυματομορφών γίνεται εμφανής μόνο με λεπτομερή ανάλυση, ειδικά των αρκετά υψηλών συχνοτήτων.



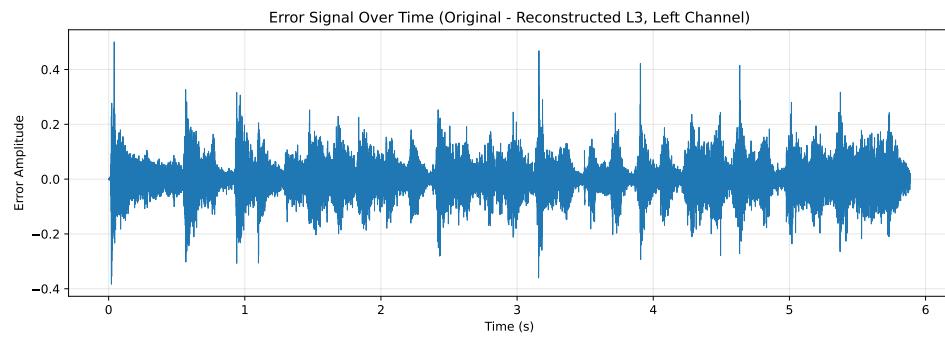
Σχήμα 3: Σύγκριση κυματομορφών (COMPRESSION\_BIAS=57)

Έπειτα, το παρακάτω bar plot περιγράφει την κατανομή των τύπων πλαισίων για το συγκεκριμένο αρχικό αρχείο ήχου. Παρατηρούμε ότι η συντριπτική πλειοψηφία των frames ανήκει στην κατηγορία OLS (Only Long Sequence).



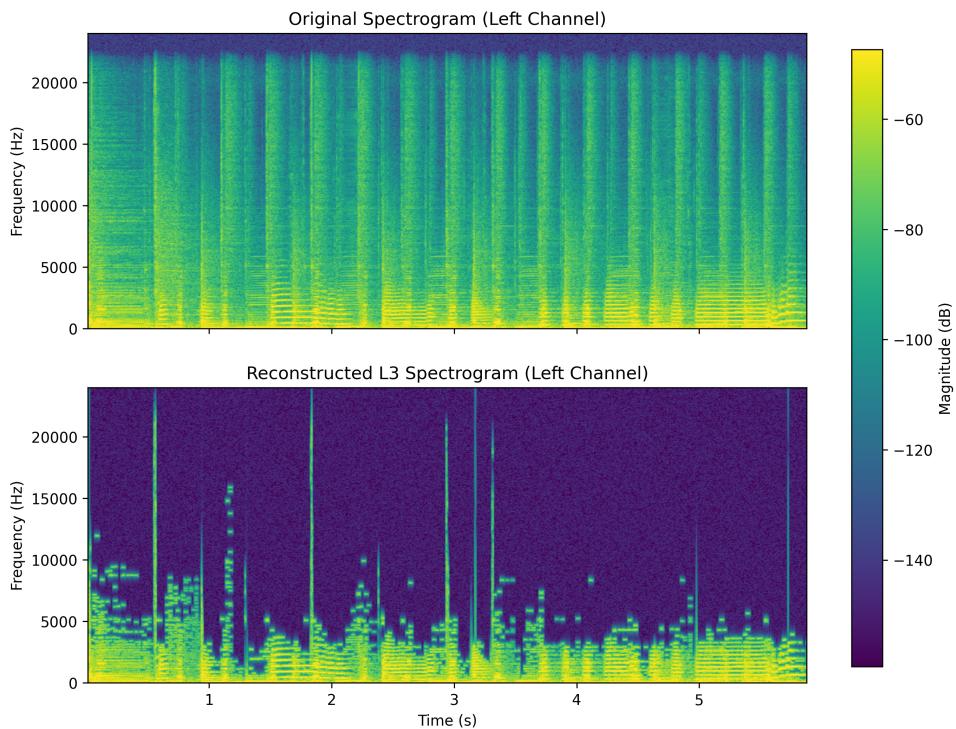
Σχήμα 4: Κατανομή τύπων πλαισίων (COMPRESSION\_BIAS=57)

Το παρακάτω διάγραμμα δείχνει το σφάλμα του ανακατασκευασμένου σήματος σε σχέση με τον χρόνο. Παρατηρούμε ότι οι τιμές κυμαίνονται στην περιοχή [-0.2, 0.2] κατά κύριο λόγο, με μερικές ακραίες τιμές που ξεπερνούν το 0.4 κατά απόλυτη τιμή.



Σχήμα 5: Σφάλμα σήματος σε σχέση με τον χρόνο (COMPRESSION\_BIAS=57)

Τέλος, παρατίθεται και μία σύγκριση των φασματογραμμάτων (spectrograms) του αρχικού σήματος και του αποκωδικοποιημένου από το level 3 σήματος. Και πάλι επιβεβαιώνεται η παρατήρηση ότι η σημαντικότερη διαφορά μεταξύ των δύο είναι στις μεγάλες συχνότητες.



Σχήμα 6: Σύγκριση φασματογράμματος (COMPRESSION\_BIAS=57)

## 5 Σύνοψη

Συνοπτικά, η εργασία αυτή περιλαμβάνει την υλοποίηση και τη δοκιμή της κωδικοποίησης και αποκωδικοποίησης σύμφωνα με το απλοποιημένο πρότυπο AAC που περιγράφεται στην εκφώνηση. Συμπεριλαμβάνονται 3 διαφορετικά επίπεδα της υλοποίησης και ένας φάκελος με περισσότερα δοκιμαστικά αρχεία που αναδεικνύουν τη λειτουργία της τελικής εκδοχής του κωδικοποιητή-αποκωδικοποιητή.

Κατά την εργασία αυτή, ο Βαπόρης Δημήτριος ανέλαβε την ανάπτυξη του κώδικα για τα πρώτα 2 επίπεδα, τον έλεγχο του κώδικα του 3ου επιπέδου και των δοκιμαστικών αρχείων και τη συγγραφή της συγκεκριμένης αναφοράς. Ο Ντελόπουλος Εμμανουήλ ανέλαβε τον έλεγχο του κώδικα των πρώτων 2 επιπέδων, την υλοποίηση του κώδικα του 3ου επιπέδου, τη συγγραφή των δοκιμαστικών αρχείων του φακέλου plotting και την παραγωγή των σχημάτων που παρουσιάστηκαν.

## Κατάλογος σχημάτων

1	Σχέση μεταξύ COMPRESSION_BIAS και Compression Rate . . . . .	4
2	Σχέση μεταξύ COMPRESSION_BIAS και SNR . . . . .	4
3	Σύγκριση κυματομορφών (COMPRESSION_BIAS=57) . . . . .	5
4	Κατανομή τύπων πλαισίων (COMPRESSION_BIAS=57) . . . . .	5
5	Σφάλμα σήματος σε σχέση με τον χρόνο (COMPRESSION_BIAS=57) . . . . .	6
6	Σύγκριση φασματογράμματος (COMPRESSION_BIAS=57) . . . . .	6