

Τμήμα Μηχανικών Η/Υ και Πληροφορικής

Εργαστήριο Επεξεργασίας Σημάτων και Τηλεπικοινωνιών



Κινητά Δίκτυα Επικοινωνιών

Μέρος Α: Τηλεπικοινωνιακά Θέματα:

*Τεχνικές Κωδικοποίησης Πηγής
Η Περίπτωση της Φωνής*



Τεχνικές Κωδικοποίησης Φωνής



- Κωδικοποίηση Πηγής: Η πληροφορία μιας διακριτής πηγής κωδικοποιείται με σκοπό την οικονομική δυαδική αναπαράσταση. Μέσω της διαδικασίας αυτής απορρίπτεται η πλεονάζουσα «πληροφορία».
- Η κωδικοποίηση χωρίς απώλειες (lossless) υπόκειται στο θεμελιώδη περιορισμό ότι “μια πηγή με εντροπία H μπορεί να κωδικοποιηθεί με μηδενική πιθανότητα σφάλματος εφόσον ο ρυθμός που προκύπτει είναι $R > H$, και αντίστροφα, εάν $R < H$ τότε, ανεξάρτητα από την πολυπλοκότητα του κώδικα θα προκύψει μη μηδενική πιθανότητα σφάλματος”, (1ο Θεώρημα του Shannon).
- Κλασικές τεχνικές lossless, γενικού σκοπού:
 - Αλγόριθμος του Shannon
 - Αλγόριθμος των Lempel – Ziv
 - Αλγόριθμος του Huffman
- Κωδικοποίηση αναλογικών πηγών : Έχει αναπόφευκτα απώλειες (lossy) που οφείλονται στη διακριτοποίηση των τιμών (κβαντισμός). Ο στόχος στην περίπτωση αυτή είναι η ελαχιστοποίηση αυτών των απωλειών και εν συνεχεία η συμπίεση με ελεγχόμενες και πάλι απώλειες (Rate-Distortion Theory)



Τεχνικές Κωδικοποίησης Φωνής

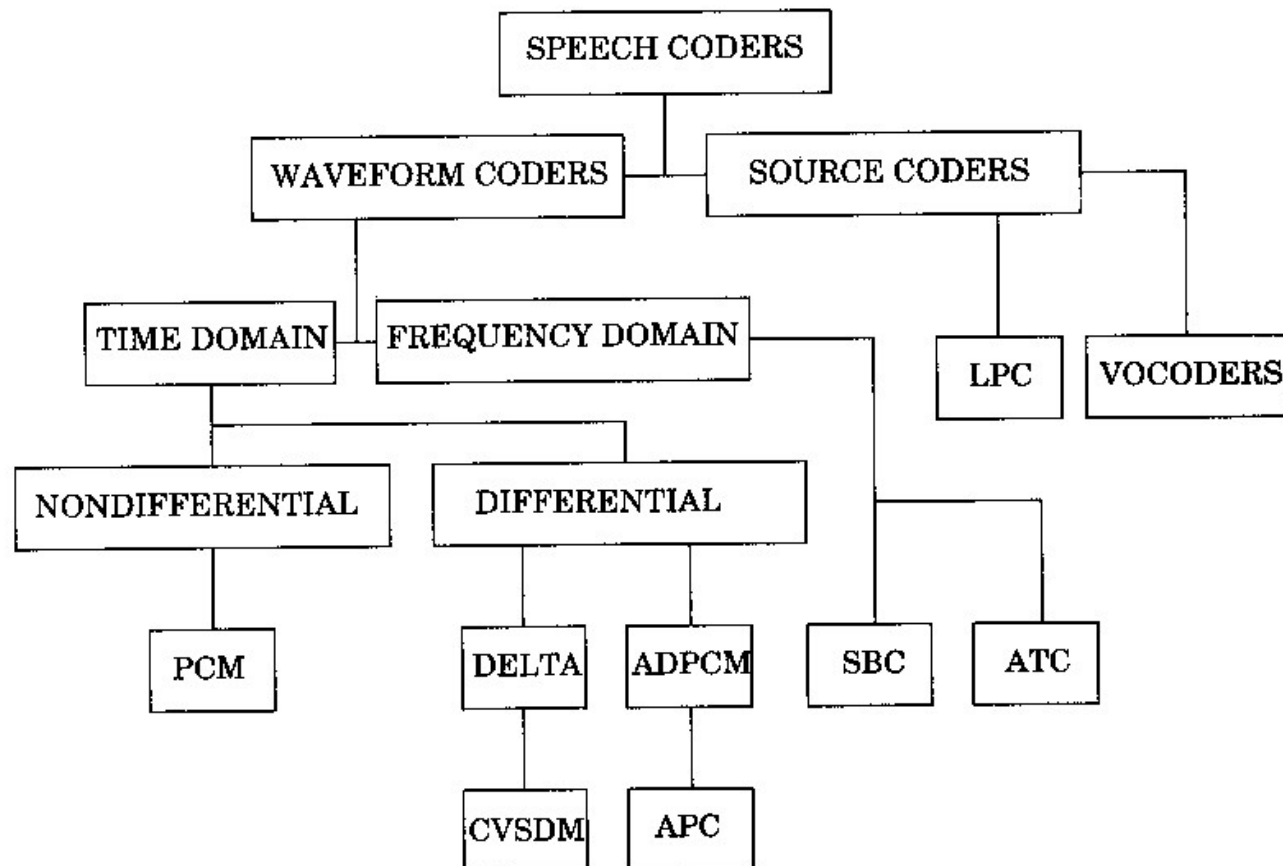


- Κωδικοποίηση Φωνής: Ιδιαίτερα χρήσιμη διαδικασία αν αναλογιστούμε ότι μεγάλο μέρος της διακινούμενης πληροφορίας είναι φωνή και ότι μέσω της συμπίεσης της μπορεί να επιτευχθεί σημαντική μείωση του απαιτούμενου εύρους ζώνης (μέχρι και 4-5 φορές, χωρίς αισθητή υποβάθμιση της ποιότητας).
- Γενική κατηγοριοποίηση των τεχνικών
 - 1) Waveform Coding
 - Temporal
 - Spectral (Transform)
 - 2) Model-based Coding
- Παρόμοια κατηγοριοποίηση ισχύει και σε άλλου τύπου αναλογικές πηγές, π.χ. εικόνα ή video (η φυσική σκηνή είναι σχεδόν πάντα αναλογική)



Τεχνικές Κωδικοποίησης Φωνής

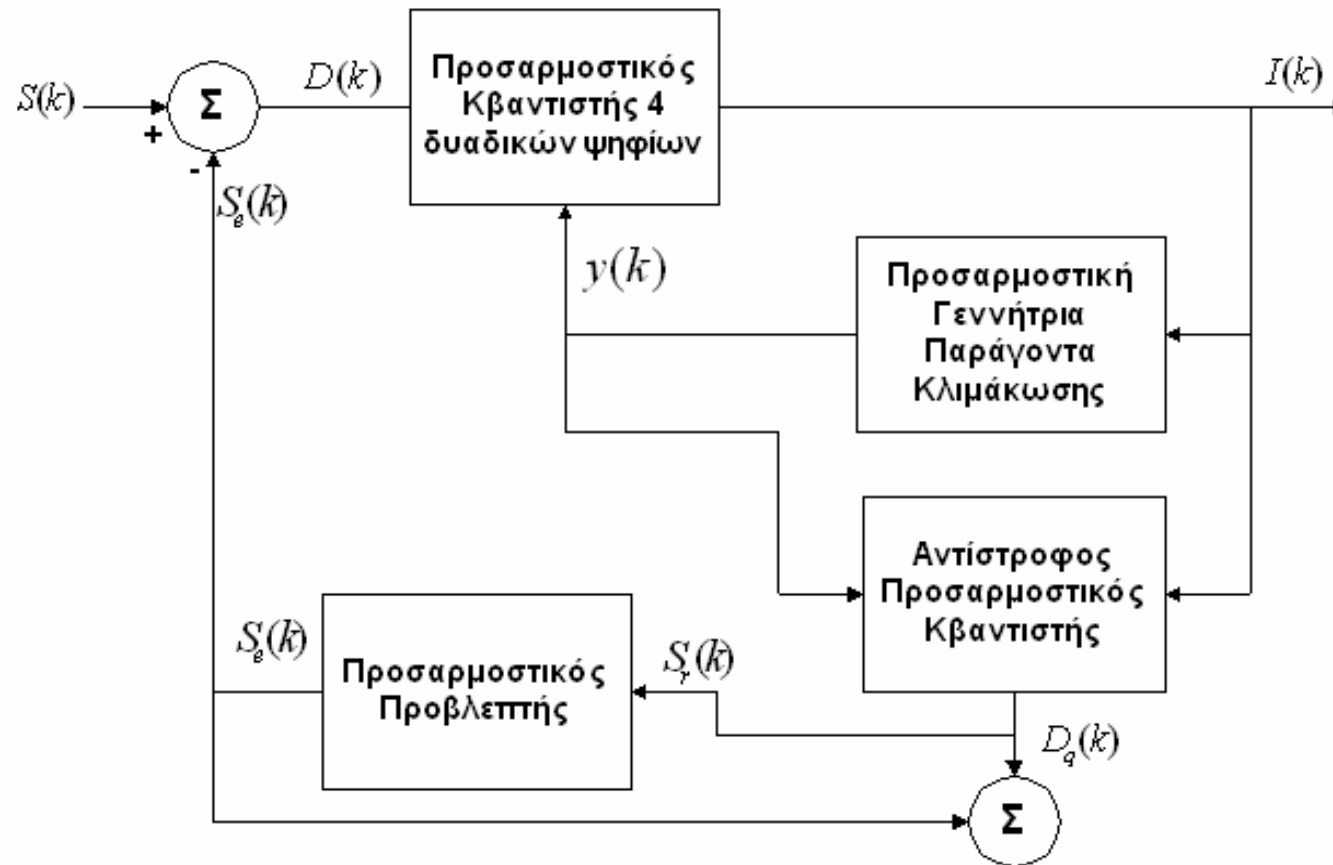
Ιεραρχική κατηγοριοποίηση των διαφόρων τεχνικών κωδικοποίησης





Τεχνικές Κωδικοποίησης Φωνής

Διάγραμμα βαθμίδων ενός κωδικοποιητή ADPCM (CT2 cordless telephone system)



Το σύστημα είναι προσαρμοστικό (προβλέπτης και κβαντιστής) και έτσι είναι σε θέση να παρακολουθεί τις στατιστικές αλλαγές του σήματος εισόδου $S(k)$ 5

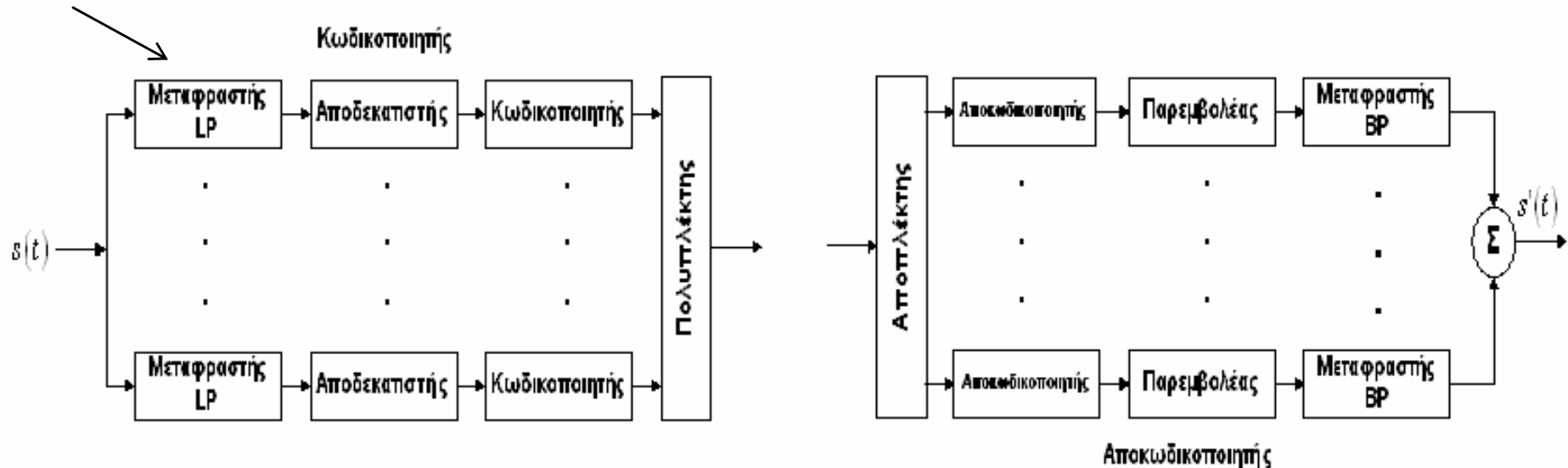


Τεχνικές Κωδικοποίησης Φωνής



Διαγράμματα βαθμίδων υπο-ζωνικού κωδικοποιητή και αποκωδικοποιητή

Μετατοπιστής



- Το σήμα φωνής χωρίζεται σε ζώνες συχνοτήτων, που κβαντίζονται και κωδικοποιούνται ξεχωριστά
- Στις ζώνες που είναι περισσότερο σημαντικές για την ανθρώπινη ακοή ανατίθενται περισσότερα bits
- Χρήση των filter banks QMF για ακύρωση της αναδίπλωσης φάσματος



Τεχνικές Κωδικοποίησης Φωνής

Τυπικός διαχωρισμός υπο-ζωνών

Αριθμός Υποζώνης	Ζώνη Συχνοτήτων (Hz)	Πλήθος δυναμικών ψηφίων
1	225-450	4
2	450-900	3
3	1000-1250	2
4	1800-2700	1

Adaptive Transform Coding (ADC)

- Κωδικοποίηση τμημάτων φωνής στο πεδίο συχνοτήτων
- Συνήθως χρησιμοποιείται ο Διακριτός Μετασχηματισμός Συνημιτόνου (DCT) του οποίου οι σχέσεις ανάλυσης και σύνθεσης φαίνονται παρακάτω ($g(0)=1$, $g(k)=2^{1/2}$) :

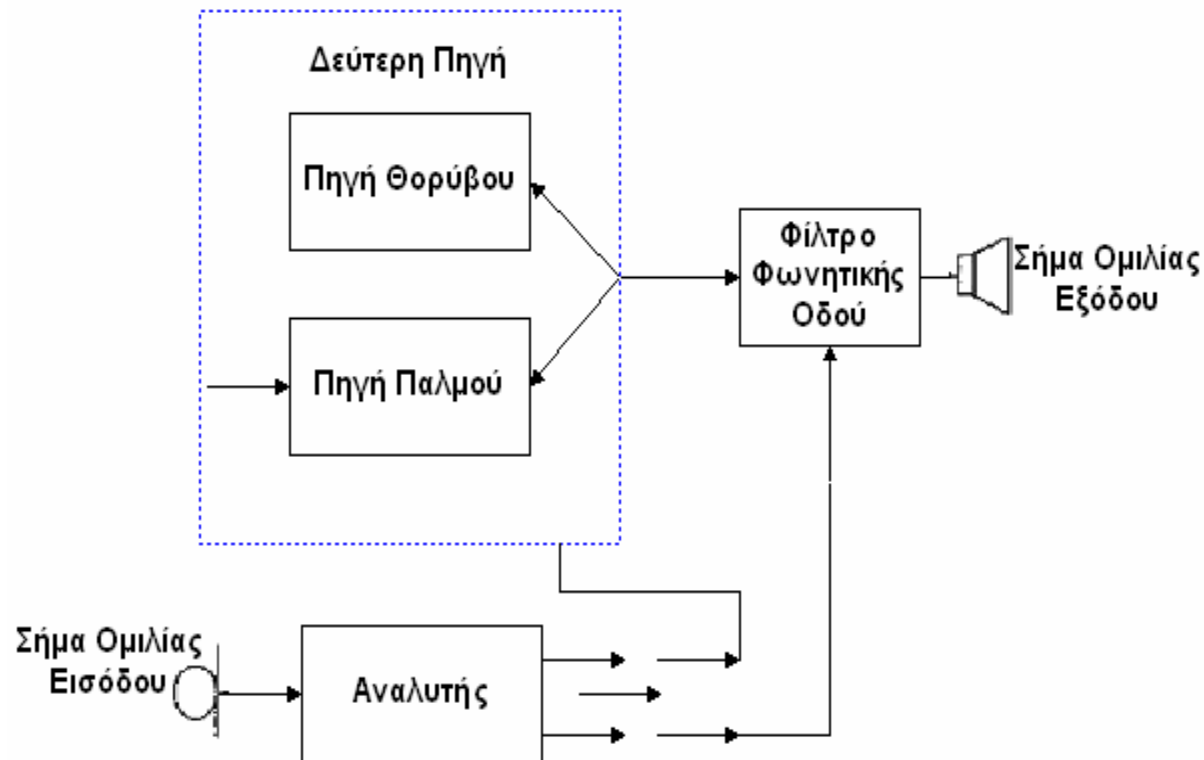
$$X_c(k) = \sum_{n=0}^{N-1} x(n) g(k) \cos\left[\frac{(2n+1)k\pi}{2N}\right] \quad k = 0, 1, 2, \dots, N-1$$

$$x(n) = \frac{1}{N} \sum_{k=0}^{N-1} X_c(k) g(k) \cos\left[\frac{(2n+1)k\pi}{2N}\right] \quad n = 0, 1, 2, \dots, N-1$$



Τεχνικές Κωδικοποίησης Φωνής

Μοντέλο παραγωγής φωνής



Με δεδομένη τη βασική φωνητική οδό, απαιτείται επιπλέον η εκτίμηση των εξής παραμέτρων:

- Αν το φώνημα είναι εύφωνο (voiced) ή άφωνο (unvoiced) -
- Θεμελιώδης συχνότητα (pitch)
- Μέση ισχύς του φωνήματος



Κωδικοποιητές φωνής (Vocoders) τύπου LPC (Linear Predictive Coding)

Ένας LPC Vocoder μοντελοποιεί τη φωνητική οδό ως ένα γραμμικό σύστημα τύπου all-pole με συνάρτηση μεταφοράς:

$$H(z) = \frac{G}{1 + \sum_{k=1}^M b_k z^{-k}}$$

Προσδιορισμός των συντελεστών του LPC Vocoder :

Το τρέχον δείγμα γράφεται ως
γραμμικός συνδυασμός προηγούμενων
δειγμάτων :

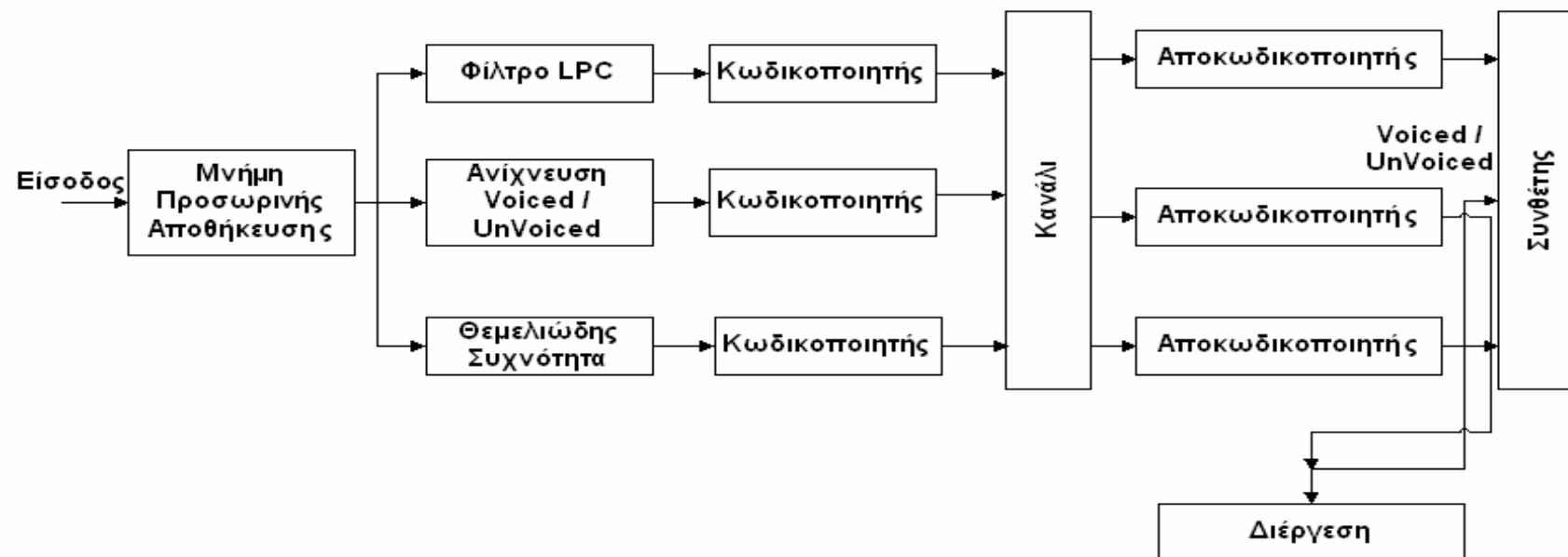
$$s_n = \sum_{k=1}^p a_k s_{n-k} + e_n$$

$\min\{\sum e_n^2\} \text{ w.r.t. } a_k \rightarrow \mathbf{R} \mathbf{a} = \mathbf{r}$ (\mathbf{R} = Toeplitz) (Levinson-Durbin Algorithm)



Τεχνικές Κωδικοποίησης Φωνής

Διάγραμμα βαθμίδων ενός κωδικοποιητή LPC



Απαιτούμενη πληροφορία στον δέκτη:

- Συντελεστές LPC (κβαντισμένοι)
- Voiced/unvoiced decision
- Pitch period
- Gain

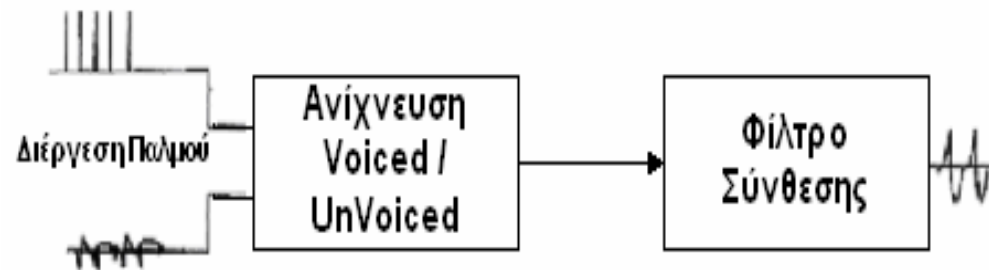
> Τεχνικές εκτίμησης/ανίχνευσης για τις παραπάνω παραμέτρους

Η διαδικασία LPC θυμίζει το ADPCM με τη βασική διαφορά ότι αντί να αποστέλλεται το κβαντισμένο σφάλμα αποστέλλονται κάποια χαρακτηριστικά του



Τεχνικές Κωδικοποίησης Φωνής

Διάφορες εναλλακτικές μέθοδοι δημιουργίας του σήματος διέγερσης στον αποκωδικοποιητή LPC

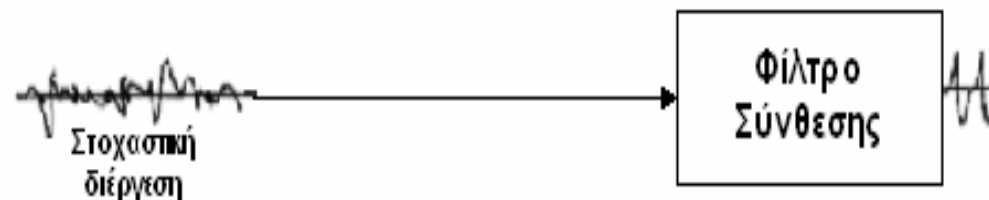


LPC Vocoder



MPE-LPC (Multipulse Excitation)

Χρήση συγκεκριμένου αριθμού παλμών ανά περίοδο με μεταβαλλόμενα πλάτη και θέσεις (επιλέγεται από ένα codebook η ακολουθία που ελαχιστοποιεί κατάλληλη συνάρτηση κόστους)



CELP

(Code Excited LP)

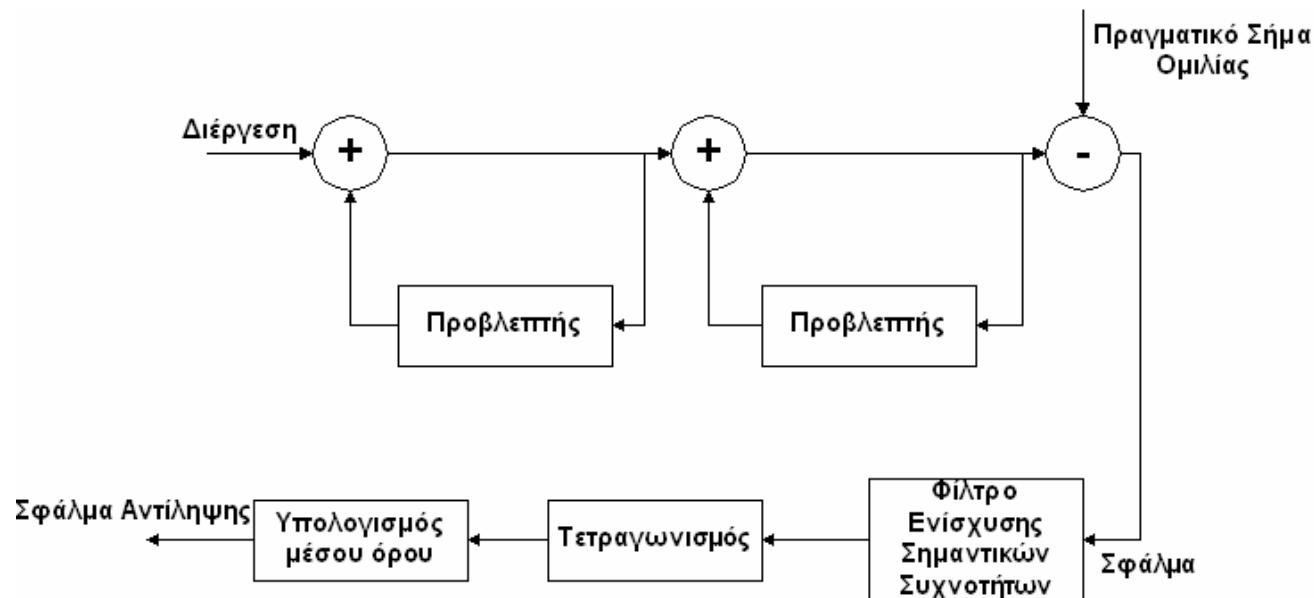
Χρήση pitch predictor + διέγερση (η οποία επιλέγεται από ένα προκαθορισμένο codebook με σήματα διέγερσης) (βλ. επόμενο slide)

$$e(n) = \beta e(n - T) + c(n)$$



Τεχνικές Κωδικοποίησης Φωνής

Διάγραμμα βαθμίδων της διαδικασίας code book search του συστήματος κωδικοποίησης CELP (Code Excited Linear Predictive)



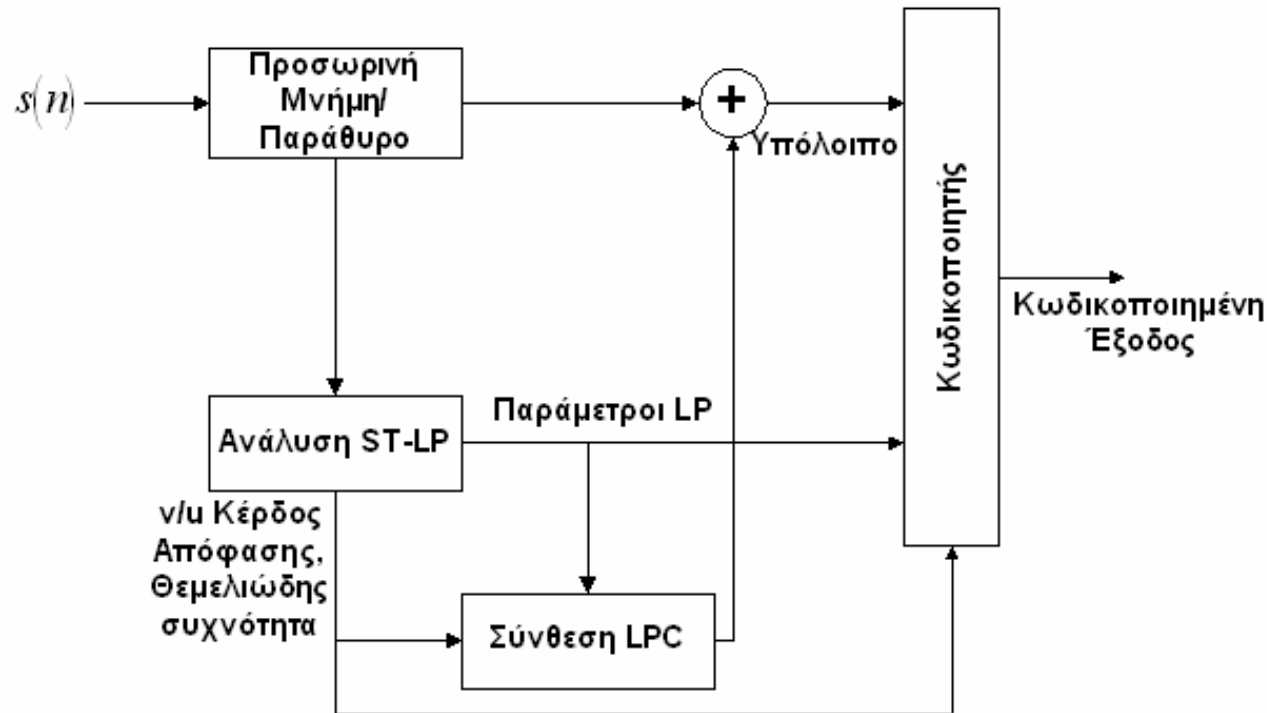
Οι βασικές διαδικασίες στον κωδικοποιητή CELP:

- Υπολογισμός των δύο προβλεπτών (LTP, STP) από το εκάστοτε τμήμα φωνής (Long Term & Short Term Prediction)
- Εύρεση της βέλτιστης διέγερσης (από το codebook). Το αντικειμενικό σφάλμα ανακατασκευής (error) υφίσταται περαιτέρω μετασχηματισμό σύμφωνα με τον αντιληπτικό μηχανισμό μας (perceptual masking)



Τεχνικές Κωδικοποίησης Φωνής

Διάγραμμα βαθμίδων του συστήματος κωδικοποίησης RELP (Residual Excited Linear Predictive)

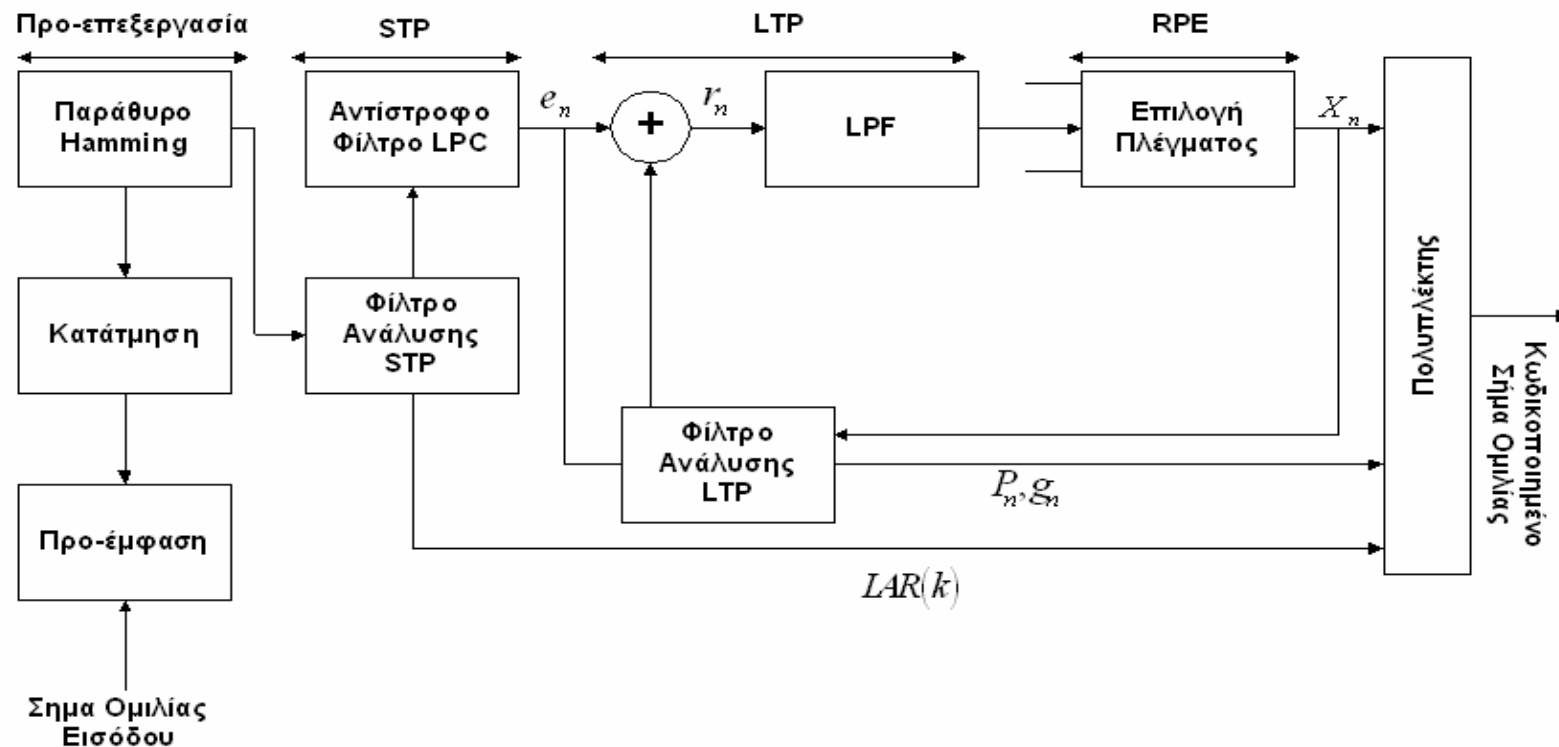


Το σύστημα κωδικοποίησης RELP βασίζεται ουσιαστικά στην ίδια λογική με αυτή του συστήματος ADPCM με επιπλέον στοιχεία την ενσωμάτωση χαρακτηριστικών του τρέχοντος δείγματος (v/u, gain, pitch) . Σχηματίζεται το σήμα σφάλματος (residual) το οποίο κβαντίζεται και μεταδίδεται.



Τεχνικές Κωδικοποίησης Φωνής

Διάγραμμα βαθμίδων του κωδικοποιητή φωνής του συστήματος GSM (Regular Pulse Excited - Long Term Prediction – RPE-LTP)

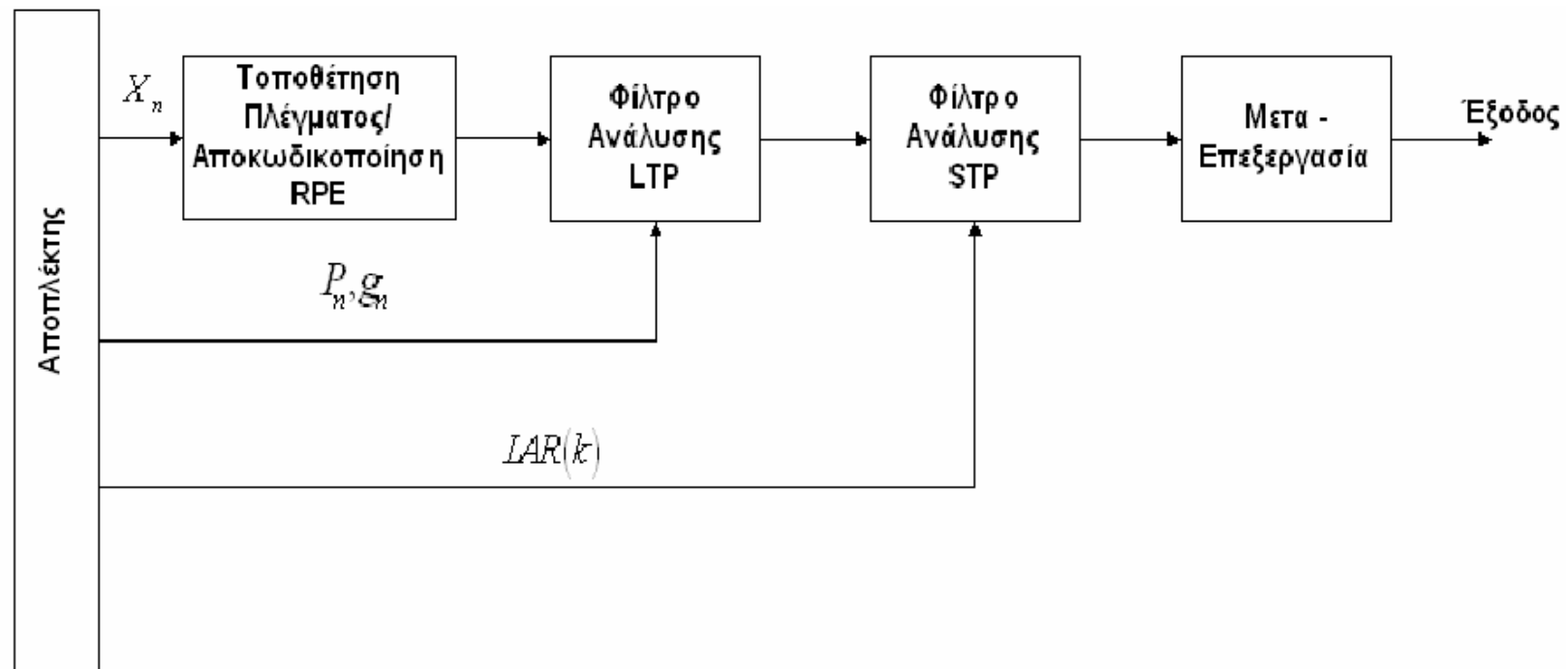


Το σύστημα RPE-LTP συνδυάζει τα πλεονεκτήματα του RELP και του MPE-LTP. Ουσιαστικά είναι σύστημα τύπου RELP που έχει ενσωματώσει τη διαδικασία LTP.



Τεχνικές Κωδικοποίησης Φωνής

Διάγραμμα βαθμίδων του αποκωδικοποιητή φωνής του συστήματος GSM



- Η ανάλυση STP (στον κωδικοποιητή) δίνει τους λεγόμενους reflection coefficients και όχι τους ίδιους τους συντελεστές πρόβλεψης. Είναι μαθηματικά ισοδύναμη παραμετροποίηση.
- Στη συνέχεια από τους r.c. υπολογίζονται οι συντελεστές L.A.R. (logarithmic area ratios) μέσω ενός απλού μη-γραμμικού μετασχηματισμού. Οι συντελεστές που προκύπτουν έχουν καλύτερες ιδιότητες κβάντισης.



Τεχνικές Κωδικοποίησης Φωνής

Τεχνικές κωδικοποίησης που χρησιμοποιούνται σε διάφορα συστήματα
κινητών επικοινωνιών

Πρότυπο	Τύπος Υπηρεσίας	Κωδικοποιητής Φωνής	Bit Rate (kbps)
GSM	Cellular	RPE-LTP	13
CD-900	Cellular	SBC	16
USDC (IS-54)	Cellular	VSELP	8
IS-95	Cellular	CELP	1.2, 2.4, 4.8, 9.6
IS-95 PCS	PCS	CELP	14.4
PDC	Cellular	VSELP	4.5, 6.7, 11.2
CT2	Cordless	ADPCM	32
DECT	Cordless	ADPCM	32
PHS	Cordless	ADPCM	32
DCS-1800	PCS	RPE-LTP	13
PACS	PCS	ADPCM	32

- Συστήματα 3ης γενιάς: Adaptive Multirate (AMR) **speech** codec

Το σύστημα επιτρέπει συμπίεση σε διαφορετικούς ρυθμούς ανάλογα με τις συνθήκες. Στον πυρήνα του συστήματος είναι η τεχνική ACELP (Algebraic CELP) που είναι κατά βάση η CELP με διαφορετικό codebook design and search.



Τεχνικές Κωδικοποίησης Φωνής

Αξιολόγηση της απόδοσης διαφόρων τεχνικών κωδικοποίησης
με βάση τον δείκτη MOS (Mean Opinion Score)

Κωδικοποιητής	MOS
64 kbps PCM	4.3
14.4 kbps QCELP13	4.2
32 kbps ADPCM	4.1
8 kbps ITU-CELP	3.9
8 kbps CELP	3.7
13 kbps GSM Codec	3.54
9.6 kbps QCELP	3.45
4.8 kbps CELP	3.0
2.4 kbps LPC	2.5

Διαβάθμιση Ποιότητας	Βαθμολογία	Προσπάθεια ακρόασης
Εξαιρετική	5	Σαφήνεια χωρίς καμία προσπάθεια
Καλή	4	Σαφήνεια χωρίς ιδιαίτερη προσπάθεια
Μέτρια	3	Σαφήνεια με σχετική προσπάθεια
Ανεπαρκής	2	Σαφήνεια με σημαντική προσπάθεια
Κακή	1	Πολύ περιορισμένη σαφήνεια παρά την σημαντική προσπάθεια