



Εργασία 2

(Προθεσμία: -----)

Κωδικοποιητής CELP

α) Εξαγωγή Συντελεστών LPC

Φορτώστε το αρχείο ήχου “voice.wav” και αναλύστε το με την `frame_wind` με `frame = 256` δείγματα και `ovrlp = 0.5`. Για κάθε frame t , χρησιμοποιείτε τη ρουτίνα “`lpc_new.m`” που υπάρχει στο `eclass` για να βρείτε $p = 40$ συντελεστές γραμμικής πρόβλεψης $a_{i,t}$ και το αντίστοιχο κέρδος G_t που περιγράφουν το κάθε παράθυρο.

β) Εξαγωγή Σφάλματος Γραμμικής Πρόβλεψης

Υπολογίστε το λάθος $e[n,t]$ του γραμμικού προβλεπτή για κάθε frame t του αρχείου “voice.wav”, σύμφωνα με τη σχέση:

$$e[n,t] = (x[n,t] - \sum_{i=1}^p a_{i,t} x[n-i,t]) / G_t = x[n,t] * (1 - \sum_{i=1}^p a_i \delta[n-i]) / G_t$$

γ) Υπολογισμός Λεξικού για τον κωδικοποιητή CELP

Ο κωδικοποιητής CELP αναμεταδίδει τους συντελεστές LPC και τα κέρδη G_t . Αντί να στέλνει όλα τα λάθη $e[n,t]$ του γραμμικού προβλεπτή, χτίζει ένα λεξικό από N τέτοια λάθη και η όλη κωδικοποίηση γίνεται με βάση τα λάθη που υπάρχουν στο λεξικό.

Πρέπει να ορίσουμε ένα κριτήριο με το οποίο θα επιλέξουμε τα καλύτερα N λάθη για να μπουν στο λεξικό. Ένα τέτοιο κριτήριο, θα μπορούσε να είναι ένα στατικό μέτρο απόστασης από την Γκαουσιανή Κατανομή: η κύρτωση (kurtosis). Όσο πιο μεγάλη είναι η κύρτωση, τόσο πιο μη-Γκαουσιανό είναι το σφάλμα μας. Γράψτε μια ρουτίνα, η οποία θα υπολογίζει την κύρτωση των $e[n,t]$ και να επιλέξετε τα N σφάλματα με την μεγαλύτερη κύρτωση. Ενδεικτικά, επιλέξτε $N=50$ λάθη για το λεξικό.

γ) Κωδικοποίηση/Αποκωδικοποίηση Σήματος με τον κωδικοποιητή CELP

Τώρα που έχουμε το λεξικό με τα σφάλματα, μπορούμε να κωδικοποιήσουμε όλα τα σφάλματα με βάση το λεξικό. Για να γίνει αυτό, πρέπει να βρούμε ποιο σφάλμα από το λεξικό ταιριάζει καλύτερα με κάθε $e[n,t]$ και να το αντικαταστήσουμε το $e[n,t]$ με το αντίστοιχο σφάλμα. Υπολογίζουμε την ανακατασκευή που σχηματίζει κάθε σφάλμα του λεξικού με τις αντίστοιχες παραμέτρους γραμμικής πρόβλεψης ($a_{i,t}$ και G_t) για το συγκεκριμένο frame t . Η ανακατασκευή

γίνεται με φιλτράρισμα του σφάλματος με το φίλτρο $G_t/A_t(z)$ (όπου το φίλτρο του παρανομαστή έχει τους συντελεστές γραμμικής πρόβλεψης). Υπολογίζουμε το μέσο τετραγωνικό σφάλμα κάθε ανακατασκευής με το συγκεκριμένο $X(:,t)$. Το σφάλμα του λεξικού που οδηγεί σε μικρότερο μέσο τετραγωνικό σφάλμα ανακατασκευής είναι αυτό που θα αντικαταστήσει το αρχικό σφάλμα $e[n,t]$, δίνοντας το νέο $e_{\text{new}}[n,t]$.

Υπολογίστε όλες τις ανακατασκευές με όλα τα νέα $e_{\text{new}}[n,t]$, δημιουργώντας το Y και με τη `frame_recon.m` ανακατασκευάστε την y . Ακούστε το αποτέλεσμα και υπολογίστε το SNR με τη βοήθεια της σχέσης:

$$SNR_{dB} = 10 \log_{10} \frac{\sum_{n=0}^{N_0-1} x(n)^2}{\sum_{n=0}^{N_0-1} [x(n) - y(n)]^2}$$

Δοκιμάστε διάφορες τιμές του $N=[100, 200, 300, 350]$ και δείτε την επίδραση τους στο SNR.

Κρατήστε το λεξικό για $N=100$ από το αρχείο `'voice.wav'` και κωδικοποιήστε κατά CELP τις κυματομορφές `'guit1.wav'` και `'guit2.wav'`. Τι παρατηρείτε ? Έχετε καλή ανακατασκευή σε σχέση με το προηγούμενο σας πείραμα ?

Ξάνθη, 7/4/2020