



Εργασία 2

(Προθεσμία: Τετάρτη 21 Δεκεμβρίου)

Μοντέλα Γραμμικής Πρόβλεψης (Linear Prediction Coefficients - LPC)

(Θα χρησιμοποιήσουμε τις ρουτίνες `frame_wind` και `frame_recon` που είχατε αναπτύξει στην Εργασία 1, οπότε φροντίστε να είναι στο `working directory`)

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 χρησιμοποιείτε την εντολή “lpc” του MATLAB για να βρείτε $p = 20$ συντελεστές γραμμικής πρόβλεψης a_i που περιγράφουν το κάθε παράθυρο. Για κάθε frame υπολογίστε την αυτοσυσχέτιση $R_{xx}(k)$ χρησιμοποιώντας την εντολή “xcorr”. Υπολογίστε και κρατήστε το κέρδος G της περιβάλλουσας, σύμφωνα με τη σχέση:

$$G^2 = R_{xx}(0) - \sum_{i=1}^p a_i R_{xx}(i)$$

(Προσοχή: αν το $x[n]$ έχει μήκος `frame`, τότε η συνάρτηση αυτοσυσχέτισης $R_{xx}(k)$ παίρνει τιμές για k από $[-\text{frame}+1, \text{frame}-1]$, άρα στο MATLAB θα έχει μήκος $2*\text{frame}-1$)

β) Συνθέτης Φωνής

Χρησιμοποιείτε τον κώδικα που φτιάξατε στην Εργασία 1 και χαρακτηρίστε το κάθε frame ως `voiced` και `unvoiced`. Αντικαταστήστε κάθε `unvoiced` frame με Γκαουσιανό θόρυβο με $\sigma^2 = 0.01$, που θα πάρετε από την εντολή `randn`. Αντικαταστήστε κάθε `voiced` frame με μια σειρά από δέλτα συχνότητας 110 Hz στην οποία θα προσθέσετε Γκαουσιανό θόρυβο με $\sigma^2 = 0.005$. Μετά φιλτράρετε κάθε frame με την εντολή `filter` και το φίλτρο $H(z) = G/A(z)$, ανακατασκευάστε με τη ρουτίνα `frame_recon` και ακούστε το αποτέλεσμα. Ηχογραφείστε τη δική σας φωνή και επαναλάβετε τα βήματα (για γυναικείες φωνές χρησιμοποιείτε δέλτα συχνότητας 200 Hz).

γ) Συμπίεστης Φωνής

Υπολογίστε το λάθος $e(n)$ του γραμμικού προβλεπτή για κάθε frame του αρχείου “voice.wav”.

$$e(n) = s(n) - \sum_{i=1}^p a_i s(n-i)$$

Ανακατασκευάστε με τη ρουτίνα `frame_recon` και ακούστε το σφάλμα. Μετά φιλτράρετε κάθε frame του σφάλματος με την εντολή `filter` και το φίλτρο $H(z) = G/A(z)$, ανακατασκευάστε με τη ρουτίνα `frame_recon` και ακούστε το αποτέλεσμα. Τί παρατηρείτε; Ηχογραφείστε τη δική σας φωνή και επαναλάβετε τα παραπάνω βήματα.

δ) *Voice Morphing*

Φορτώστε τα αρχεία ήχου "voice1.wav" και "voice2.wav" με συχνότητα δειγματοληψίας 8 KHz. Περιέχουν την ίδια φράση από δύο διαφορετικά άτομα όμως με διαφορετικό χρονισμό. Φορτώστε το αρχείο ήχου "voice2align.wav", όπου έχει γίνει μια διόρθωση ώστε η φράση του δεύτερου ομιλητή να ταιριάζει χρονικά με του δεύτερου. Επιβεβαιώστε το με μια γραφική παράσταση και των 2 κυματομορφών στο χρόνο ή με 2 φασματογραφήματα.

Κάντε ανάλυση LPC στις 2 κυματομορφές και υπολογίστε τους συντελεστές A_1 και G_1 για τον πρώτο ομιλητή και τους συντελεστές A_2 και G_2 για το δεύτερο ομιλητή. Κατόπιν, υπολογίστε τα σφάλματα e_1 και e_2 της ανάλυσης LPC για τους 2 ομιλητές. Μετά φιλτράρετε κάθε frame του σφάλματος e_1 με την εντολή `filter` και το φίλτρο $H_2(z) = G_2/A_2(z)$, ανακατασκευάστε με τη ρουτίνα `frame_recon` και ακούστε το αποτέλεσμα. Τί παρατηρείτε;

Κάντε και το αντίστροφο, δηλαδή φιλτράρετε κάθε frame του σφάλματος e_2 με την εντολή `filter` και το φίλτρο $H_1(z) = G_1/A_1(z)$, ανακατασκευάστε με τη ρουτίνα `frame_recon`. Τί παρατηρείτε; Ποιο είναι το συμπέρασμα για την πληροφορία που περιέχουν οι συντελεστές LPC;

Ξάνθη 30/11/2016