



ΕΛΛΗΝΙΚΗ ΔΗΜΟΚΡΑΤΙΑ
ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ

Ψηφιακή Επεξεργασία Φωνής

Άσκηση 7η

Στυλιανού Ιωάννης

Τμήμα Επιστήμης Υπολογιστών

ΗΥ578: 7^η Σειρά Ασκήσεων

Απορίες: yannis@csd.uoc.gr

1. Θεωρήστε το σήμα:

$$x[n] = A \cos(\omega_0 n)$$

και $y[n] = x[n]w[n]$, όπου $w[n]$ ένα παράθυρο ανάλυσης.

(α') Ποιο είναι το φάσμα $X(\omega)$ του $x[n]$.

(β') Ποιο είναι το φάσμα $Y(\omega)$ του $y[n]$.

(γ') Ποιες συνθήκες πρέπει να πληρεί το $w[n]$ ώστε:

$$Y(\pm\omega_0) = A/2$$

2. Αντί να χρησιμοποιήσουμε ένα πολυώνυμο 3ου βαθμού για την παρεμβολή της φάσης μπορούμε εναλλακτικά να θεωρήσουμε 2 πολυώνυμα 2ου βαθμού: ένα για $0 \leq t \leq T/2$ και ένα δεύτερο για $T/2 \leq t \leq T$. Αυτό σημαίνει ότι η συχνότητα σε κάθε ένα από τα διαστήματα αυτά θα είναι γραμμική συνάρτηση του χρόνου. Θεωρώντας ότι

$$\underline{t = 0}$$

$$\theta(0) = \theta^l$$

$$\omega(0) = \Omega^l$$

$$\underline{t = T/2}$$

$$\theta(T/2) = \theta(T/2)$$

$$\omega(T/2) = \tilde{\Omega}$$

$$\underline{t = T}$$

$$\theta(T) = \theta^{l+1} + 2\pi M$$

$$\omega(T) = \Omega^{l+1}$$

να εκφράσετε το $\theta(t)$ ως συνάρτηση του $\tilde{\Omega}$ (άγνωστη μεταβλητή) και των άλλων γνωστών μεγεθών (το M να θεωρηθεί γνωστό).

Βοήθεια: Η παραπάνω άσκηση αναφέρεται σε σημεία της άσκησης 9.4 (Quatieri) όπου δείτε και το σχήμα 9.28 σελ. 485.

3. Είτε επειδή το επιλέγουμε είτε επειδή μας αναγκάζουν, δεν μπορούμε να στείλουμε τις παραμέτρους ενός ημιτονοειδούς μοντέλου σε κάθε παράθυρο. Για παράδειγμα στην πρώτη περίπτωση θα μπορούσε να είναι αυτό που αναφέρεται ως half-rate representation ενώ η 2η περίπτωση θα μπορούσε να είναι σε ένα σενάριο Voice Over IP όπου κάποια πακέτα/frames χάνονται. Βασισμένοι στα αποτελέσματα της προηγούμενης άσκησης προτείνεται μία μέθοδο εκτίμησης των παραμέτρων του ημιτονοειδούς μοντέλου που δεν έχουν σταλεί/φτάσει από τις παραμέτρους που είναι διαθέσιμες σε half-rate συνθήκες.

Τι προβλήματα θα συναντούσε η παραπάνω τεχνική δεδομένου της μη στασιμότητας του σήματος της φωνής;

4. Θεωρούμε ένα στάσιμο έμφωνο μέρος ενός σήματος φωνής που έστω ότι αποτελείται από N συνιστώσες σε αρμονικές συχνότητες:

$$s(t) = \sum_{k=0}^{N-1} A_k \cos(\theta_k(t))$$

όπου $\theta_k(t) = a_k + b_k t + g_k t^2 + c_k t^3$ και $\dot{\theta}_k(0) = k\omega_0$. Τοποθετούμε στο σήμα ένα παράθυρο μήκους $2P+1$, όπου P είναι η περίοδος του σήματος ($P = 2\pi/\omega_0$) και θέλουμε να υπολογίσουμε τα πλάτη A_k του σήματος (π.χ. με peak-picking).

(α') Πως πρέπει να κανονικοποιηθεί το παράθυρο ανάλυσης ώστε πράγματι το peak-picking να μας δώσει τα σωστά πλάτη του σήματος;

(β') Σχεδιάστε το φάσμα πλάτους θεωρώντας μόνο τον κεντρικό λοβό του φάσματος πλάτους του μετ.Fourier του παράθυρου.

(γ') Κατασκευάζουμε μια νέα φάση $\tilde{\theta}_k(t) = B\theta_k(t)$ και συνθέτουμε το σήμα με την νέα αυτή φάση. Ποια είναι η θεμελιώδης συχνότητα του νέου σήματος; Σχεδιάστε το φάσμα πλάτους του σήματος για $B = 1/2$. Ποιο μήκος τώρα θα πρέπει να έχει το παράθυρο ανάλυσης για να εκτιμήσουμε σωστά τα πλάτη και τις φάσεις;

5. Έχουμε το σήμα:

$$x[n] = A \cos(\omega n) w[n]$$

όπου $w[n]$ είναι ένα τετραγωνικό παράθυρο. Η συχνότητα ω του σήματος δεν είναι ακέραιο πολλαπλάσιο της συχνότητας $2\pi/NFFT$ όπου $NFFT$ είναι το μέγεθος του διακριτού μετ.Fourier.

- (α') Σχεδιάστε το φάσμα πλάτους του $x[n]$ ως διακριτό φάσμα ($X[k]$) θεωρώντας μόνο τις θετικές συχνότητες και πάνω σε αυτό τοποθετήστε το συνεχές φάσμα πλάτους του σήματος ($X(\omega)$)
- (β') Εστω ότι έχετε τον διακριτό μετ. Fourier ενός σήματος και θέλετε να κάνετε peak-picking με στόχο τον υπολογισμό μεταξύ άλλων των συχνοτήτων του σήματος. Βασισμένοι στην απάντησή σας στο προηγούμενο ερώτημα προτείνετε μια τεχνική για μία βελτιωμένη εκτίμηση της συχνότητας μέσω της τεχνικής του peak-picking (δηλ. προτείνετε ένα βελτιωμένο peak-picking αλγόριθμο βασισμένοι στις γνώσεις σας στην επεξεργασία σήματος).

(

ό

$A \quad \omega \quad n \quad)$

ι

Σημειώματα

Σημείωμα αναφοράς

Copyright Πανεπιστήμιο Κρήτης, **Στυλιανού Ιωάννης**. «Ψηφιακή Επεξεργασία Φωνής. Άσκηση 7η». Έκδοση: 1.0. **Ηράκλειο/Ρέθυμνο** 2015. Διαθέσιμο από τη δικτυακή διεύθυνση: <http://www.csd.uoc.gr/~hy578>.

Σημείωμα Αδειοδότησης

Το παρόν υλικό διατίθεται με τους όρους της άδειας χρήσης Creative Commons Αναφορά, Μη Εμπορική Χρήση, Όχι Παράγωγο Έργο 4.0 [1] ή μεταγενέστερη, Διεθνής Έκδοση. Εξαιρούνται τα αυτοτελή έργα τρίτων π.χ. φωτογραφίες, διαγράμματα κ.λ.π., τα οποία εμπεριέχονται σε αυτό και τα οποία αναφέρονται μαζί με τους όρους χρήσης τους στο «Σημείωμα Χρήσης Έργων Τρίτων».



[1] <http://creativecommons.org/licenses/by-nc-nd/4.0/>

Ως **Μη Εμπορική** ορίζεται η χρήση:

- που δεν περιλαμβάνει άμεσο ή έμμεσο οικονομικό όφελος από την χρήση του έργου, για το διανομέα του έργου και αδειοδόχο
- που δεν περιλαμβάνει οικονομική συναλλαγή ως προϋπόθεση για τη χρήση ή πρόσβαση στο έργο
- που δεν προσπορίζει στο διανομέα του έργου και αδειοδόχο έμμεσο οικονομικό όφελος (π.χ. διαφημίσεις) από την προβολή του έργου σε διαδικτυακό τόπο

Ο δικαιούχος μπορεί να παρέχει στον αδειοδόχο ξεχωριστή άδεια να χρησιμοποιεί το έργο για εμπορική χρήση, εφόσον αυτό του ζητηθεί.

Διατήρηση Σημειωμάτων

Οποιαδήποτε αναπαραγωγή ή διασκευή του υλικού θα πρέπει να συμπεριλαμβάνει:

- το Σημείωμα Αναφοράς
- το Σημείωμα Αδειοδότησης
- τη δήλωση Διατήρησης Σημειωμάτων
- το Σημείωμα Χρήσης Έργων Τρίτων (εφόσον υπάρχει)

μαζί με τους συνοδευόμενους υπερσυνδέσμους.

Χρηματοδότηση

- Το παρόν εκπαιδευτικό υλικό έχει αναπτυχθεί στα πλαίσια του εκπαιδευτικού έργου του διδάσκοντα.
- Το έργο «Ανοικτά Ακαδημαϊκά Μαθήματα στο Πανεπιστήμιο Κρήτης» έχει χρηματοδοτήσει μόνο τη αναδιαμόρφωση του εκπαιδευτικού υλικού.
- Το έργο υλοποιείται στο πλαίσιο του Επιχειρησιακού Προγράμματος «Εκπαίδευση και Δια Βίου Μάθηση» και συγχρηματοδοτείται από την Ευρωπαϊκή Ένωση (Ευρωπαϊκό Κοινωνικό Ταμείο) και από εθνικούς πόρους.

