

5/12/2024

סמסטר חורף תשפ"ה

אוניברסיטת חיפה, החוג למדעי המחשב

ضغط الصوت والصورة

דחיסת קול ותמונה

تمرين بيت 1

תרגיל בית מספר 1

- ההגשה (במודל!) בזוגות, אחרת- רק באישור המרצה. יש לציין בבירור בראש העבודה ובשם קובץ ה ZIP שמות המגישים ומספר ת.ז.
- הגשה עד 24/12/24 בחצות
- כל יום איחור ממועד ההגשה הסופי ללא אישור יגרור קנס של 5 נקודות
- העתקות אסורות ויטופלו בחומרה

שאלה 1: קורלציה וכימוי (10%)

נתונה פונקצית האוטוקורלציה:

$$r_k = \sum_n^{N-K-1} X_n * X_{n+k} \quad k = 0, \dots, N-1$$

1. יש להוכיח כי $r(k)$ סימטרית, כלומר, $r_k = r_{-k}$.
2. ניתן להוכיח (בקלות – נסו!) כי המקסימום הגלובלי של הפונקציה מתקבל ב- $k = 0$, יש להסביר למה נקודה זו לא רלוונטית לחישוב ה- $Pitch$.
3. יש להסביר (או להראות בצורה גרפית, אין חובה להוכחה פורמלית), מדוע אם האות $X(n)$ מחזורי במחזור P , מתקבל מקסימום מקומי ב- $0, \pm P, \pm 2P, \pm 3P, \dots$ בפונקצית האוטוקורלציה.

2. יחס קצב עיוות (15%)

לפניכם עקומות (גרפים) שמתארות התנהגות של שני קוונטיזרים (מכמתים: A,B), מבחינת קצב

הסיביות במוצא הקוונטיזר לעומת העיוות שנגרם לאות המכומת

בתליך הכימוי. (גרפים אלו נקראים Rate-Distortion או בקיצור R-D)

יש לענות לשאלות ולנמק היטב!

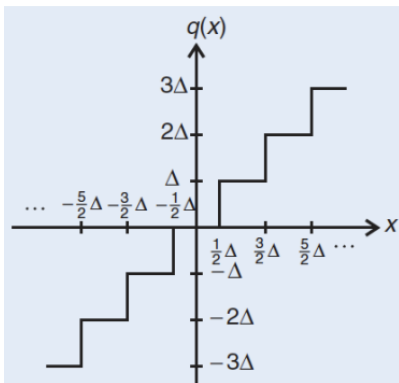
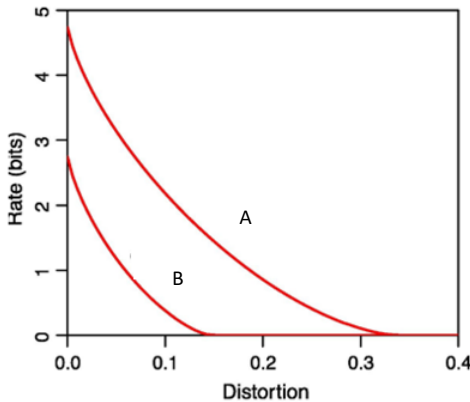
א. איזו עקומה מתארת לדעתך קוונטיזר שנותן ביצועים טובים יותר?

ב. כידוע, הדרך הטריויאלית לשנות את קצב הסיביות היא שינוי הקוונטיזר. לפניך עקומת כימוי של קוונטיזר יוניפורמי. צעד

הכימוי (step size) בו הוא Δ . נניח ש Δ יכול לקבל את הערך 1 או את הערך 2.

האם ניתן לדעת איזו עקומת R-D מהסעיף הקודם מתארת את הבחירה

בכל אחת משתי האפשרויות של Δ ?



שאלה 3: שאלות כלליות בקידוד דיבור (25%)

1. קיבלת משימה לדגום אות פיזיולוגי – לתכנן מכשיר המודד את הדופק. דופק אופייני למבוגרים הוא 60-100 פעימות לדקה.
א. מהו קצב הדגימה הנדרש (ב Hz)?
ב. כיצד יראה מסנן ה LPF (Anti-Aliasing) בכניסת מערכת הדגימה (צורת המסנן ותדר ההחסימה הנדרש)
2. עליך להחליט באיזה שיטת קידוד להשתמש לקידוד רעש לבן (Random Noise).
באיזה מקודד כדאי יותר להשתמש – DPCM, DM?, PCM? יש לנמק!
3. א. יש להשוות (לפי קריטריון MSE) את ביצועי חזאי DPCM פשוט לשני חזאים המבצעים: חיזוי לפי הדגימה הקודמת, ומיצוע על הדגימה הקודמת והדגימה הבאה, בהתאמה.
יש להשוות על אות הכניסה הבא:
$$X(n) = [98, 100, 98, 102, 105, 106, 105, 102]$$

(ניתן להניח שהדגימות הקודמות והבאות כולן בעלות הערך: 100)
- מה היתרונות והחסרונות של שתי השיטות בסעיף הקודם?

שאלה 4: מקודד LPC (15%)

1. במקודד LPC10 סטנדרטי (ראו שקפים במצגות הרלוונטיות):
 - קצב הדגימה 8kHz
 - אורך כל מסגרת אנליזה 180 דגימות
 - כל מסגרת מקודדת ב 54 סיביות, וביציאת המקודד קצב כולל של 2400bit/Sec
 - במסגרת מסוג Voiced (קולית) מוקדשות 4 סיביות לקידוד כל אחד מעשרת מקדמי המסנן
- אם נתכנן מקודד דומה שייקרא "LPC12" ובו נקודד 12 מקדמים (במקום 10) תוך שמירה על מספר הסיביות הכולל ביציאה.
 - א. כמה סיביות יוקצו עכשיו לכל מקדם?
 - ב. האם וכיצד תשתנה לדעתך איכות הקול במוצא המקלט?
 - ג. מה יקרה לקצב הסיביות ולאיכות הקול במוצא המקודד אם נקטין את מסגרת האנליזה ל 120 דגימות תוך שמירה על כל שאר הפרמטרים (המקוריים לפי התקן)?

שאלה 5: SPDemo (35%)

1. יש לטעון את הקובץ lcyWind.wav. בעזרת העכבר (לחצן ימני). בכמה סיביות לדגימה מיוצג הקובץ?
2. בצעו כימוי PCM ל-15 סיביות ומטה עד שהבדל נשמע לאוזן. בכמה סיביות לכל הפחות נצטרך לייצג את האות אם נרצה לקבל אות ללא רעש?
3. מה הקשר בין איכות האות לרמת ה-SNR המתקבלת?
4. בצעו קידוד בשיטת PCM ו-ADPCM לאות ב-5 סיביות. האם ניתן להסיק ששיטה אחת יעילה מהשנייה? נמק.
5. השתמשו בשני קטעי הדיבור: speech0.wav ו speech8.wav הנתונים לצורך אנליזת LPC תחת: SPDemo. האזינו לקבצים!
6. יש לבחור קטע קולי וא-קולי בכל אחד מהקבצים ולבצע אנליזת LPC (לסמן: Use Autocorrelation method).
 - כמה פורמנטים ניתן לזהות בבירור בכל אחד?
 - האם ניתן לאפיין הבדלים בין קול גברי לנשי בעזרת בדיקה מסוג זה (רמז: נסו לעקוב אחרי השתנות הפורמנטים וה-Pitch לאורך מילה שלמה)?
7. הקליטו את עצמכם אומרים מילה כלשהי הכוללת אזור קולי ואזור א-קולי. הדגימו על הקלטה זו: (בדגימה בקצב של 16KHz)
 - אילו ערכי Pitch אתם מגלים?
 - כמה פורמנטים ואיפה? (יש להראות את האות, לסמן עליו את האזורים הקוליים והא-קוליים). בהגשה יש לצרף את הקטעים שהקלטתם בקבצי wav !
8. העבירו את הקבצים Depeche Mode.wav & Vega.wav ב-Phone line simulator תחת לשונית: PHONE כל אחד בתורו (מצב זה מדמה מעבר בקו טלפון אופייני).
 - יש להאזין לאות המקורי ולאות שעבר בקו הטלפון. השוו ביניהם והסבירו ממה לדעתכם נובעים ההבדלים בין האות המקורי והאות החדש.
 - ציינו איפה מורגש ההבדל במיוחד בקטע של Depeche Mode ולמה?

שאלה 6: שאלת בונוס פייתון/ מטלב (10 נקודות)

יש להשתמש בפונקצית `xcorr` (מובנית במטלב) לחישוב קרוס-קורלציה בין שני אותות.

א. עבור אות באורך n ואות באורך m . באיזה אורך תתקבל התוצאה ?

ב. היכן יתקבל הערך המירבי ?

יש לכתוב את הפונקציה: `function s = snr(in_vec, out_vec)` המחשבת את snr האות לרעש בין שני וקטורים.

התוכנית הבאה מבצעת קריאת קובץ שמע דגום ספרתי ומבצעת לו קוואנטיזציה ל- n סיביות:

```
n=16;
[x, Fs, bps] = wavread('vega.wav');

xq=floor((x+1)*2^(n-1));
xq=xq/(2^(n-1));
xq=xq-(2^(n)-1)/2^(n);

xe=x-xq;
```

יש לבצע כימוי (קוואנטיזציה) ל-15 סיביות ומטה עד לסיבית 1. ולחשב את SNR . בכל פעם. (בעזרת הפונקציה שכתבתם בסעיף הקודם) יש להציג את התוצאות על גבי גרף אחד.

במטלב ניתן לבצע תצוגה גרפית של מידע באמצעות פונקציות `plot` או `bar`.

(רמז: הפשוט ביותר להשתמש בלולאת `For` ולשנות את n בכל איטרציה של הלולאה).