© Technion - Israel Institute of Technology, Elyachar Central Library

הטכניון – מכון טכנולוגי לישראל
Technion – Israel Institute of Technology

ספריות הטכניון
The Technion Libraries

בית הספר ללימודי מוסמכים ע"ש ארווין וג'ואן ג'ייקובס
Irwin and Joan Jacobs Graduate School

©
All rights reserved

This work, in whole or in part, may not be copied (in any media), printed, translated, stored in a retrieval system, transmitted via the internet or other electronic means, except for "fair use" of brief quotations for academic instruction, criticism, or research purposes only.
Commercial use of this material is completely prohibited.

©
כל הזכויות שמורות

אין להעתיק (במדיה כלשהי), להדפיס, לתרגם, לאחסן במאגר מידע, להפיץ באינטרנט, חיבור זה או כל חלק ממנו, למעט "שימוש הוגן" בקטעים קצרים מן החיבור למטרות לימוד, הוראה, ביקורת או מחקר. שימוש מסחרי בחומר הכלול בחיבור זה אסור בהחלט.

קידוד סקלרי ווקטורי של אותות דיבור
בפסי תדר נפרדים

חיבור על מחקר
לשם מילוי חלקי של הדרישות לקבלת התואר
מגיסטר למדעים
בהנדסת חשמל

מאת
רן ערד

הוגש לסנט הטכניון - מכון טכנולוגי לישראל
תשרי תשמ"ז חיפה אוקטובר 1986

טכניון
פקולטה להנדסת חשמל
פקולטה למדעי המחשב
896 מספר הס...

2033449

000001011142

2.01.89

© Technion - Israel Institute of Technology, Elyachar Central Library

המחקר נעשה בהנחייתו של פרופסור דוד מלאך
בפקולטה להנדסת חשמל.

תודתי העמוקה נתונה לפרופסור דוד מלאך, על הצעת הנושא, ועל הדרכתו, מעורבותו ותרומתו למהלך העבודה.

תודתי נתונה גם לצוות המעבדה לעבוד אותות, ובעיקר ליורם אור-חן, ציפי פורטנוי, זיוה אבני וברוך כוכבי, שתמכו בי בשעות קשות.

© Technion - Israel Institute of Technology, Elyachar Central Library

תוכן הענינים

עמוד

קציר

שימת סמלים וקיצורים

רק 1 - מבוא

1. קידוד דיבור בתחום התדר.....1

1. קידוד דיבור בפסי תדר נפרדים (SBC).....3

1.2.1 הסינון ב-SBC.....4

1.2.2 הקוואנטיזציה ב-SBC.....7

1. מטרת המחקר.....9

1. מבנה העבודה.....10

רק 2 - מערכת SBC עם קוואנטיזציה סקלרית

2. קוואנטיזציה סקלרית.....11

2. בעית הקצאת הסיביות.....14

2.2.1 רקע עיוני.....14

2.2.2 שיקולים מעשיים.....16

2. מערכת עם הקצאה סטטית של סיביות.....19

2. מערכת עם הקצאה דינמית של סיביות.....24

2.4.1 אדפטציה אחורית.....26

2.4.2 אדפטציה קדמית.....27

רק 3 - קוואנטיזציה וקטורית

3. בעית הקוואנטיזציה הוקטורית.....30

3. אלגוריתם LBG לתכנון קוואנטיזר וקטורי אופטימלי.....32

3.2.1 תנאי Lloyd.....32

3.2.2 אלגוריתם LBG.....33

3.2.3 תנאי התחלה לאלגוריתם.....34

3.2.4 בעית התא הריק.....35

© Technion - Israel Institute of Technology, Elyachar Central Library

תוכן הענינים (המשך)

עמוד

3.3 קוואנטייזרים וקטוריים תת-אופטימליים.....37
3.3.1 קודי מכפלה.....37
3.3.2 קוואנטייזר במבנה עץ.....38
3.3.3 קוואנטייזר רב דרגות.....41
3.4 מערכות קדוד דבור עם קוואנטיזציה וקטורית.....43
3.4.1 קוואנטיזציה וקטורית במקודדי מקור.....43
3.4.2 מערכות עם חזוי.....44
3.4.3 מערכות עם בקרת הגבר.....47
3.4.4 קדוד וקטורי בתחום התדר.....49

פרק 4 - מערכת SBC עם קוואנטיזציה וקטורית

4.1 מערכת עם קוואנטיזציה "אנכית".....51
4.1.1 הוספת בקרת הגבר.....53
4.1.2 הוספת שקלול בתחום התדר.....55
4.2 מערכת עם קוואנטיזציה "אפקית".....59
4.2.1 הוספת בקרת הגבר.....62
4.2.2 הקצאה דינמית של מילונים.....64
4.2.3 מערכת עם מילונים במבנה עץ.....65

פרק 5 - מאפייני המערכות המומלצות

5.1 המערכת המומלצת עם קוואנטייזר סקלרי.....67
5.1.1 סבוכיות המימוש.....68
5.1.2 עמידות לרעשים.....71
5.1.3 Tandeming.....75
5.1.4 טיפול באותות נתונים (data).....76

© Technion - Israel Institute of Technology, Elyachar Central Library

תוכן הענינים (המשך)

עמוד

5.2 המערכת המומלצת עם קוואנטייזר וקטורי.....80
5.2.1 סבוכיות המימוש.....81
5.2.2 עמידות לרעשים.....83
5.2.3 Tandeming.....86
5.2.4 טפול באותות נתונים.....86
5.3 השוואה בין קוואנטיזציה סקלרית ווקטורית במערכת SBC.....88

פרק 6 - סיכום

6.1 סיכום העבודה.....90
6.2 כוונים למחקר המשך.....92

נספח א' - נוסחאות ה-QMF והוכחתן.....94

נספח ב' - נוסחאות ההקצאה הדינמית של סיביות והוכחתן.....96

מקורות

תקציר (באנגלית)

© Technion - Israel Institute of Technology, Elyachar Central Library

תקציר

עבודה זו עוסקת במערכות קידוד דיבור בפסי תדר נפרדים - SBC (Subband Coding), ומתמקדת בעקר במנגנון הקוואנטיזציה במערכות אלה.

קידוד דיבור בפסי תדר נפרדים הינה שיטה מקובלת בקידוד ספרתי של דיבור בתחום התדר, ויעילה במיוחד בתחום קצבי הביניים (8-16Kbps). בשיטה זו מופרד אות הדיבור ע"י מערך מסננים למספר פסי תדר, וכל פס תדר מקודד בנפרד לפי קריטריונים אנליטיים (כגון - יחס אות לרעש) או סובייקטיביים (כגון - תכונות האוזן האנושית). האותות המקודדים משודרים בערוץ תקשורת ספרתי, ובמקלט הם מפוענחים, ומורכבים מחדש ליצירת האות המשוחזר.

מקובל בדרך-כלל להפריד את תכנון המערכת לשניים: תכנון מערכי המסננים שבמשדר ובמקלט (הקרויים גם "מערכת אנליזה-סינתיזה"), ותכנון הקוואנטייזרים, המשמשים בקידוד ובפענוח של האותות בתת-הפסים. עבודה זו מתרכזת, כאמור, בבעיה השניה.

במערכות הפועלות כיום מקובל להשתמש בקוואנטייזרים כגון ADPCM, שניתן לכנותם "קוואנטייזרים סקלריים", שכן הם פועלים על האות המקודד דגם-דגם. עם זאת, גובר בשנים האחרונות השימוש בקוואנטיזציה וקטורית - VQ (Vector Quantization) המבוססת על קידוד בבת אחת של מספר דגמים של האות. קוואנטייזרים וקטוריים ידועים זה מכבר בספרות, אולם הפופולריות שלהם גדלה לאחרונה, לאחר פרסומו של אלגוריתם נוח לתכנונם, ובעקבות ההתקדמות הטכנולוגית שאפשרה מימוש שלהם בזמן אמת.

קוואנטיזרים וקטוריים משמשים כיום בסכמות קידוד דיבור רבות, ובין היתר הועלו הצעות לשלבם במערכת SBC. עם זאת, לא פורסמה עדיין עבודה מקיפה הבוחנת את כל האספקטים של מערכת כזו, ומכילה הצעות קונקרטיות למערכות מעשיות, המבוססות על השוואה של כמה שיטות. השוואות והצעות כאלו מופיעות בעבודה זו.

© Technion - Israel Institute of Technology, Elyachar Central Library

במסגרת העבודה נבחנו מחד המערכות המוכרות, עם קוואנטייזרים סקלריים, ומאידך מספר אלגוריתמים לשלוב של מקודדים וקטוריים במערכת. המערכות המומלצות נבדקו כדי לוודא התאמתן לשימושים מעשיים, ואף בדיקות אלה מופיעות בעבודה. המסקנה העיקרית מהעבודה לגבי שלוב VQ במערכת SBC היא, כי גם אם בקצבי ביניים גבוהים (כמו 16Kbps) לא מצדיק השיפור באיכות, את הגידול הניכר בסיבוכיות, הרי בקצבי ביניים נמוכים (9.6Kbps) אין למקודד הוקטורי בן תחרות סקלרי.

בסיכום העבודה מועלות גם הצעות להמשך המחקר במערכות SBC-VQ, במטרה לשפר את איכותן, וכן לרדת בסיבוכיותן מעבר למפורט בעבודה.

© Technion - Israel Institute of Technology, Elyachar Central Library

© Technion - Israel Institute of Technology, Elyachar Central Library

רשימת הסמלים והקיצורים

$a(n)$ - סדרה (אות או מסנן)

$a_i(n)$ - $a(n)$ עבור פס התדר ה-i-י

$A(e^{jw})$ - התמרת פוריה של $a(n)$

$\hat{a}(n)$ - $a(n)$ לאחר קידוד ושחזור

$\tilde{a}(n)$ - חזוי של $\hat{a}(n)$

$cent(S)$ - צנטרואיד הקבוצה S

$d(x,\hat{x})$ - עוות בין מקור ושחזור

D - תוחלת העוות

$d(n)$ - שארית החזוי בחוג סגור

$d^*(n)$ - שארית החזוי בחוג פתוח

$E\{\cdot\}$ - אופרטור התוחלת

$f(n)$ - מסנן סינתיזה

f - תדר דגימה (KHz)

$h(n)$ - מסנן אנליזה

I - קצב מעבר אינפורמציה ספרתית (Kbps)

I_i - אינמרוול ההחלטה ה-i-י של קוואנטייזר סקלרי

K - גודל בלוק-קבוע - וריאנס

k - מימד הוקטור המקודד

L - מספר מקדמים של מסנן

M - מספר פסי התדר

N - מספר המילים במילון קידוד

P - חזאי ליניארי

$p(x)$ - צפיפות הפילוג של x

Q - קוואנטייזר

R - מספר סיביות (ממוצע) לדגם/וקטור

R_{max} - מספר סיביות מקסימלי להקצאה

R^k - המרחב האוקלידי ה-k ממדי

S_i - קבוצת ההחלטה ה-i-ית של קוואנטייזר וקטורי

W - רוחב הסרט של אות הדיבור

w - גורמי שיקלול

$x(n)$ - אות הדיבור

Y - מילון קידוד

y_i - מילת הקוד ה-i-ית במילון

α - פרמטר החלון האקספוננציאלי הדועך

τ - פרמטר שיקלול בתחום התדר

Δ - גודל הצעד של קוואנטייזר סקלרי יוניפורמי

ε - קבוע הקוואנטיזציה

β - מקדם החיזוי (הקורלציה) הראשון

σ^2 - וריאנס

σ_e^2 - תוחלת העוות הריבועי (וריאנס השגיאה)

δ_i - נקודת קצה של I_i

$\downarrow j$ - הורדת קצב (דצימציה) פי j

$\uparrow j$ - העלאת קצב (אינטרפולציה) פי j

$||\cdot||$ - הנורמה האוקלידית

ADPCM - Adaptive Differential Pulse Code Modulation

BER - Bit Error Rate

BLER - Block Error Rate

CCITT - Comité Consultatif International Téléphonique et Télégraphique

Kbps - Kilobit per second

LBG - Linde, Buzo and Gray (Algorithm)

QMF - Quadrature Mirror Filters

QPSK - Quadrature Phase Shift Keying

SBC - Subband Coding

SNR - Signal to Noise Ratio

SNRSEG - Segmented SNR (seg-size 256)

VBD - Voice Band Data

VQ - Vector Quantization

© Technion - Israel Institute of Technology, Elyachar Central Library

פרק 1 - מבוא

1.1 קידוד דיבור בתחום התדר

וכיוון שמערכות קידוד הדיבור, בהן עוסקת עבודה זו, שייכות למשפחה של
וקודדי דיבור בתחום התדר, מן הראוי לפתוח בסקירה קצרה על שיטות קידוד אלו
:כלל, ועל מקומם של המקודדים, בהם עוסקת העבודה, בתוך משפחה זו בפרט.

קידוד דיבור, בכללותו, הנו שם כולל לתחום העוסק בהעברת אותות דיבור
:ערוצי תקשורת ספרתיים. השם "קידוד" (או - "דחיסה") מרמז על כך
שמעוניינים לצמצם את הקצב בו משודר האות מבלי לפגוע באיכות השדור, במטרה
יעל ולהוזיל ככל האפשר את התקשורת. מערכות קידוד דיבור באיכות טובה
המקובלות כיום, הם מקודדי צורת גל (בתחום הזמן) הפועלים בקצב של 64Kbps,
רק לאחרונה התקבל סטנדרט של CCITT לתקשורת ב-32Kbps. אחד מכווני המחקר
העקריים בתחום זה כיום, הוא פיתוחן של מערכות שיאפשרו תקשורת בקצבים
:מוכים יותר, ובעקר בתחום המכונה "קצבי ביניים" (8-16Kbps), באיכות
:בוהה (Toll Quality) ובשמושים מסוימים מסתפקים אף באיכות תקשורת
(Communication Quality).

וקובל לחלק את שיטות קידוד הדיבור לשתי מחלקות עקריות: קידוד מקורות
(Source Coding), בו מתאימים מודל לאות הדיבור ומשדרים את פרמטרי המודל,
קידוד צורת-גל (Waveform Coding), בו מקודדים את דגמי האות ממש. המערכות
והסוג הראשון פועלות בקצבים נמוכים יותר, אך גם באיכות פחותה, מאשר אלה
והסוג השני.

:ם את המחלקה של מקודדי צורת-גל מקובל לחלק לשניים: מקודדים בתחום הזמן,
הפועלים על האות עצמו (כגון - ADPCM, DM וכדומה), ומקודדים בתחום התדר
(או תחום הטרנספורם), הפועלים על ורסיה מותמרת של האות. המקודדים מהסוג
השני, שבהם נתמקד בהמשך סעיף זה, הינם בדרך-כלל מסובכים יותר לעומת אלה
והסוג הראשון, אך מאפשרים ירידה לתחום קצבים נמוך יותר, תוך שמירה על
איכות טובה.

© Technion Israel Institute of Technology Elyachar Central Library

המעבר לתחום התדר (או תחום התמרה כלשהו), נועד למלא אחר שתי מטרות עקריות: הראשונה, הוצאת הקורלציה מהמידע, כך שאפשר יהיה לישם שיטות קידוד יעילות הידועות לאותות חסרי-קורלציה (חסרי-זכרון); השניה, להעביר את המידע לתחום בו אפשר יהיה לנתחו, בתלות ביצירתו הפיזיקלית ובמודלים של התפישה האנושית. בשתי מטרות אלה נעוצה האטרקטיביות של שיטות קידוד דיבור בתחום התדר.

ישנן שתי קבוצות עקריות של מקודדים השייכות למשפחת המקודדים בתחום התדר: מקודדי התמרה (TC - Transform Coders), ומקודדים בפסי תדר נפרדים (SBC - Subband Coders); עבודה זו מתמקדת במקודדים מטפוס SBC.

ניתן לסכם את ההבדלים העקריים ואת התכונות המשותפות בין שתי הקבוצות באופן הבא: במקודדים מטפוס SBC, מחולק ספקטרום הדיבור למספר קטן יחסית של תת-פסים (2-16) בעזרת מערך של מסננים; במקודדים מטפוס TC, הספקטרום מחולק למספר גדול יחסית של רכיבי תדר (512 - 64) ע"י התמרה בבלוקים (למשל DCT, DFT); בשני המקרים מבצעים לאחר מכן קוואנטיזציה (קידוד) של המידע המתקבל בטכניקות אדפטיביות המוכרות מתחום הזמן (כגון ADPCM). גם ב-SBC וגם ב-TC משתדלים להתאים את המקודד של כל רכיב תדר לאופיו של אותו רכיב, בהתאם לשתי המטרות שהוזכרו לעיל - יצירת אי תלות בין הרכיבים, והתאמת הקידוד לתכונות האוזן האנושית.

כפי שניתן לראות, שתי השיטות מנסות לבצע אנליזה ספקטרלית לזמן-קצר של אות הדיבור, אולם ברזולוציה ספקטרלית שונה. שיטת ה-TC היא שיטה בעלת סיבוכיות גבוהה, ו-SBC בעלת סיבוכיות בינונית-עד-גבוהה (בתלות, בין היתר, בצורה שבה מתבצעת הקוואנטיזציה).

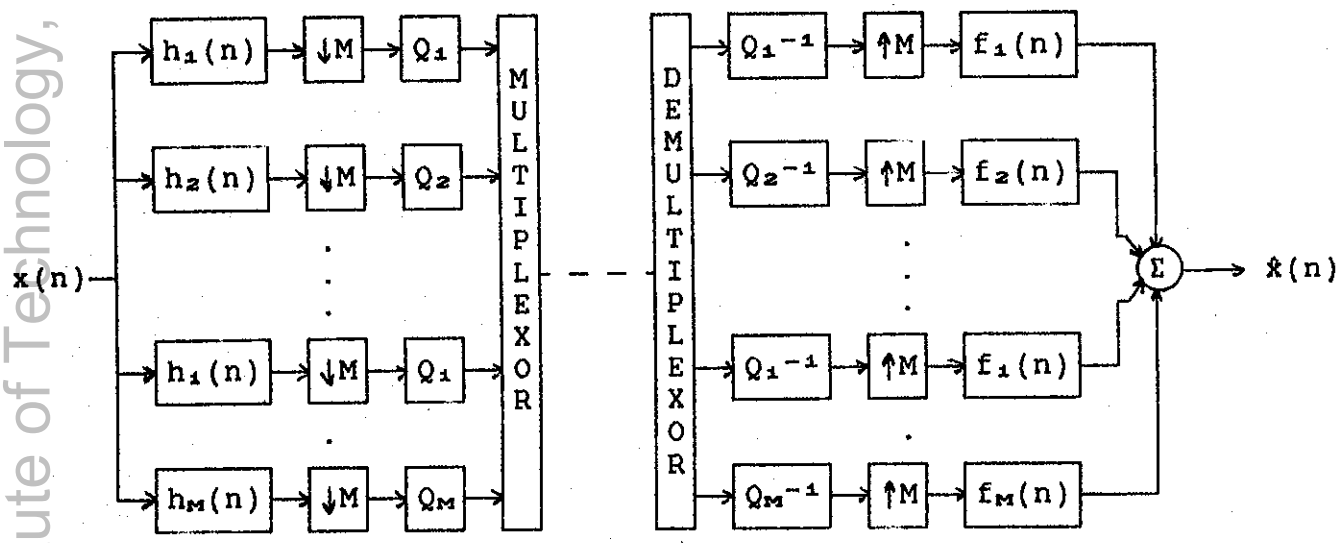
סקירה ממצה של קידוד דיבור ניתן למצוא ב-[4]; לסקירה של קידוד דיבור בתחום התדר מומלץ לעיין ב-[5].

© Technion - Israel Institute of Technology, Elizachar Central Library

1.2 קידוד דיבור בפסי תדר נפרדים (Subband Coding - SBC)

כפי שצוין, נעשית האנליזה הספקטרלית בשיטת SBC ע"י מערך של מסננים, המחלק את ספקטרום האות למספר קטן יחסית של תת-פסים. כל תת-פס מקודד בנפרד, והאותות המקודדים משודרים בערוץ התקשורת. במקלט מתבצע פענוח האותות, ולאחר מכן מתבצעת סינתיזה ספקטרלית בעזרת מערך מסננים הופכי, כדי לקבל את הדיבור המשוחזר.

תאור סכמטי של מערכת SBC מופיע בציור 1.1.



ציור 1.1 - תאור סכמטי של מערכת SBC.

Figure 1.1 - Schematic description of a SBC system.

$\{h_i(n)\}_{i=1}^{M}$ הינו מערך מסננים, כלומר קבוצת מסננים ספרתיים המחלקת את ספקטרום האות לפסי-תדר זרים ומשלימים. $\{f_i(n)\}_{i=1}^{M}$ הנם המסננים ההפכים להם, כלומר - המערכת $\{(h_i,f_i)\}_{i=1}^{M}$ היא מערכת יחידה (עד כדי השהיה קבועה). הבלוקים המסומנים ב-$\downarrow M$ ו- $\uparrow M$ מבצעים הורדה והעלאה בהתאמה של קצב הדגימה פי M, מכיוון שבכל פס-תדר תחום התדרים של האות קטן פי M מתחום התדרים המקורי. הבלוקים המסומנים ב Q_i ו-Q_i^{-1} מציינים סכמת קוואנטיזציה כלשהי (למשל ADPCM) ופענוחה בהתאמה.

© Technion - Israel Institute of Technology, Elyachar Central Library

כפי שניתן לראות, רוחב תת-הפסים הוא גדול יחסית למבנה הספקטרום של אות הדיבור, מה שמאפיין מערכת זו כמערכת בעלת רזולציה ספקטרלית נמוכה.

תכנון החלוקה לפסים מתחשב בדרך כלל במבנה הספקטרום של אות הדיבור. ספקטרום זה איננו שטוח, אלא מאופין ע"י מספר שיאים הקרויים פורמאנטים (תדירויות התהודה של המעבר הקולי). אנו מעוניינים בדרך כלל לבודד את הפורמאנטים ע"י החלוקה בתחום התדר, וע"י כך להשיג אחידות יחסית בכל פס. מכיוון שהפורמאנטים הם יותר צרים בתדרים נמוכים, מקובל לבחור פסי תדר הולכים ומתרחבים עם התדר, למרות שהדבר אינו הכרחי.

גם מנגנון הקידוד מבוסס בדרך-כלל על מבנה הספקטרום של הדיבור והקשר שלו לתכונות האוזן האנושית. גם משקולים אנליטיים (יחס אות לרעש) וגם משקולים תפישתיים (תכונות האוזן), מסתבר שבד"כ עדיף לקודד את פסי התדר הנמוכים ביותר סיביות מאשר את הגבוהים. בכל מקרה, בשל האופי הלא סטאציונארי של הדיבור, משתמשים במקודדים אדפטיביים.

בדרך כלל מקובל לחלק את בעיית תכנון המערכת לשני מרכיביה העקריים, וכך גם ננהג בהמשך סעיף זה: תת הסעיף הבא עוסק בתכנון מערכי המסננים עבור המשדר והמקלט, כלומר בבצוע האנליזה והסינתיזה הספקטרליות לזמן קצר; תת הסעיף שלאחריו עוסק בתכנון הקוואנטייזרים (קידוד ופענוח) במערכת.

1.2.1 הסינון ב-SBC

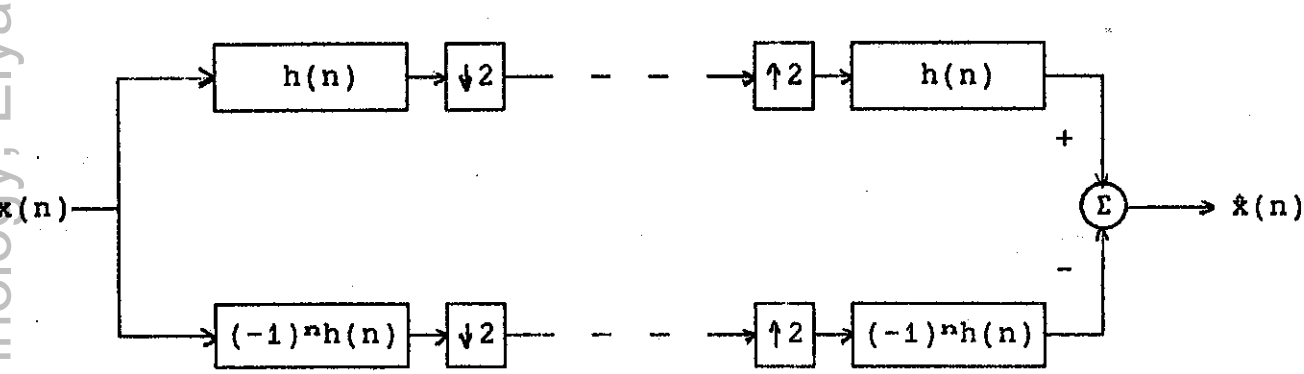
בתכנון מערכת הסנון ב-SBC, הכוונה לתכנון שני מערכי מסננים, $\{h_i(n)\}_{i=1}^{M}$ ו-$\{f_i(n)\}_{i=1}^{M}$, כאשר המערך h מחלק את תחום התדרים לפי פרוט נתון, והמערך f משלים אותו למערכת יחידה.

שיטה יעילה היא שמוש באותו מסנן, h(n), בכל הפסים, כאשר בכל פס מבצעים הזזה לתחום התדרים המתאים ע"י מודולציה. המסנן h(n) נקרא אז מסנן אבטיפוס (Prototype). בשיטה זו נוח להשתמש ב-DFT לשם ממוש המערכת. עם זאת, השיטה היא בעייתית, שכן בשל חוסר האידיאליות של המסננים ובשל השמוש ב-DFT, גוברות ההשפעות של זליגה בין פסי-תדר; בנוסף לכך קשה לתכנן את $\{h_i(n)\}_{i=1}^{M}$ ו- $\{f_i(n)\}_{i=1}^{M}$ כך שייצרו מערכת יחידה.

© Technion Israel Institute of Technology, Elizabeth Cental Library

ופיכך השיטה המקובלת ביותר כיום מבוססת על סוג של מסננים הקרויים QMF (Quadrature Mirror Filters). מסננים אלה מתוכננים כך שיקיימו את דרישת וערכת היחידה, וכן שיצמצמו למינימום את השפעת הזליגה בין פסי התדר.

יטת QMF מבוססת על המערכת הבסיסית המתוארת בציור 1.2, והמבצעת חלוקה לשני תת-פסים שווים.



ציור 1.2 - מערכת QMF בסיסית.

Figure 1.2 - Basic QMF system.

h(n) הינו מסנן אבטיפוס, מעביר נמוכים, סימטרי ומסדר זוגי. ההכפלה ב-$(-1)^n$ הנה למעשה מודולציה ספרתית. בנוסף לכך דורשים שיתקיים הקשר -

$$|H(e^{jw})|^2 + |H(e^{j(w+\pi)})|^2 = 1 \qquad (1.1)$$

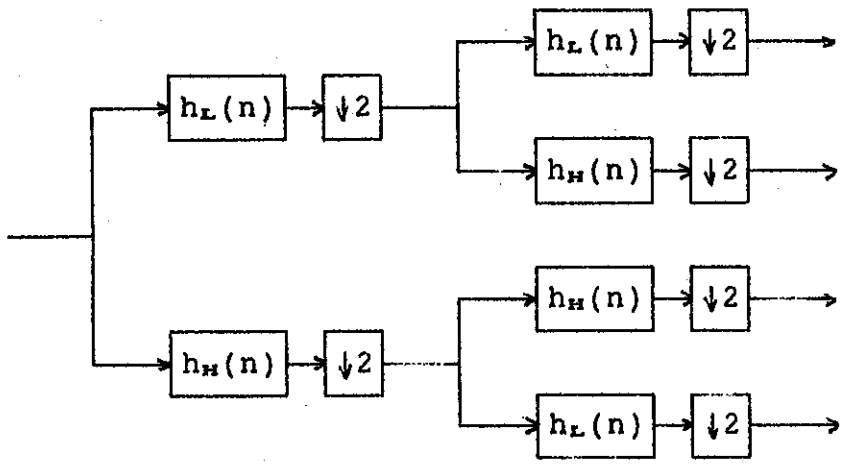
קשרים אלו גורמים לכך שיהיו למסננים שתי התכונות הרצויות, של מערכת יחידה ושל בטול הזליגה (ההתחזות) בין הפסים. בטול ההתחזות מתקבל מתוך תכונות המסננים ללא צורך בתכנונם כמסננים בעלי מעבר חד, דבר ההופך את שיטת QMF לאטרקטיבית מאד. הבעיה העקרית בתכנון היא בעובדה שתנאי (1.1) לא ניתן למלוי במדויק, פרט למקרים פרטיים לא מעשיים; עם זאת ניתן להתקרב אליו כרצוננו בעזרת שיטות אופטימיזציה סטנדרטיות, וכך מחושבים מקדמי ה-QMF המשתמשים בהם במערכות מעשיות (ר. למשל [2] עמ' 386-382).

בעיה נוספת שיש לציינה היא, כי במערכות המכילות קוואנטיזציה, בטול ההתחזות הוא עד כדי רמת רעש הקוואנטיזציה.

© Technion - Israel Institute of Technology, Elyachar Central Library

פתוח מפורט של נוסחאות ה- QMF מופיע בנספח א'. המאמר המקורי בו הוצעו מסננים אלה הוא [6], למרות שכיום ישנם כבר פרסומים רבים המציגים טכניקות חדשניות לשפור האיכות והסיבוכיות של מסננים מטיפוס QMF.

הרחבת הסכמה הבסיסית ליותר משני תת- פסים נעשת ע"י שמוש חוזר בסכמה של ציור 1.2, ומתקבל מערך מסננים במבנה עץ. לדוגמא, עבור מערכת עם ארבעה פסי תדר נקבל עבור מערך האנליזה עץ בן שתי רמות, כמתואר בציור 1.3.



ציור 1.3 - עץ QMF בן שתי רמות (אנליזה).
Figure 1.3 - A two-level QMF tree (analysis).

המסננים $h_L(n)$ ו- $h_H(n)$ הם מסנני ה-QMF כפי שהוצגו מקודם. יש לשים לב להיפוך התפקידים בין h_H ו-h_L בענף התחתון. זו תופעה של "הפוך ספקטרלי", הנובעת כתוצאה מהורדת קצב הדגימה.

מובן שלא חייבים לבחור אותו מסנן, או אף אותו סדר-מסנן בכל רמה של העץ, ואכן מקובל לבחור במסננים מסדרים יורדים והולכים ככל שמתקדמים בעץ, כדי לקבל תחומי- מעבר שווים עבור הסנון בכל שלב.

ניתן להמיר את מבנה העץ למבנה מקבילי רגיל, ע"י מציאת המסננים השקולים לכל ענף, תוך שמוש בתכונות פשוטות של פעולות הדצימציה (הורדת קצב הדגימה) והאינטרפולציה (העלאת קצב הדגימה). המסננים המתקבלים אינם בנויים על

© Technion - Israel Institute of Technology, Elyachar Central Library

אבטיפוס משותף, ולפיכך הם פחות אטרקטיביים למימוש. עם זאת יתכן וניתן דווקא בעזרתם לחסוך עומס חישובי, שכן יותר קל לקרבם ע"י מסננים מסדר נמוך יותר, וזאת ע"י בצוע קצוץ זהיר של "זנב" המסנן.

סיבוכיות השמוש במערך QMF מהווה חלק בעל משקל כבד בממוש המערכת, ולפיכך יתכן ויש מקום לנסות ולמצוא מסננים אלטרנטיביים להם, או שיטות אחרות להורדת העומס החשובי, בעיות שלא נדונו בעבודה זו. פרוט של חשובי הסיבוכיות מופיע בסעיף 5.1.1 .

מקדמי ה-QMF בהם נעשה שמוש בעבודה זו, לקוחים מנספח 7.1 (עמ' 401-404 ב-[2]). לא נוסו כל המסננים המפורטים שם: השמוש שנעשה בכל המערכות הוא במערך עם שלוש רמות, במבנה עץ, אשר מסנן 64E מהנספח הנ"ל משמש בכל רמותיו (אלא אם כן מצוין אחרת).

1.2.2 הקוואנטיזציה ב- SBC

שאלת התכנון השניה במערכת SBC היא, כאמור, תכנון המקודדים בפסי התדר השונים.

מקודד (או - קוואנטייזר), הנה מערכת המקבלת אות כניסה ומקטינה את הדיוק שבו מיוצגים ערכיו. הכניסה למערכת קוואנטיזציה יכולה להיות אנלוגית (דיוק אינסופי) או ספרתית (בדיוק גבוה), ויציאתה היא ספרתית, בדיוק נמוך יחסית.

התאמת מקודד לכל פס מורכבת בדרך כלל משתי הפעולות הבאות:

1. התאמת דיוק הקוואנטייזר (מס' סיביות ביצוג אות המוצא), לפי האופן שבו האות בפס מתייחס לאותות בפסים אחרים. בעיה זו היא בעיית הקצאת הסיביות שתידון בהרחבה בסעיף 2.2.

2. אדפטציה של התחום הדינמי של הקוואנטייזר בהתאם לאופי האות בפס. בעיה זו היא בעית תכנון סטנדרטית ומתבצעת בנפרד לכל פס. ישנה גם אפשרות לשלב בקוואנטייזר פעולת חיזוי, להורדת היתירות באות הכניסה, וגם החזאי יכול להיות אדפטיבי. אספקטים אלו ידונו בסעיף 2.1.

© Technion Israel Institute of Technology, Elyachar Central Library

החלטה נוספת צריכה להיות לגבי מימד הקוואנטייזר - האם להשתמש בקוואנטייזר סקלרי סטנדרטי או בקוואנטייזר וקטורי. בעיה זו היא למעשה הנושא המרכזי בו דנה העבודה, ולכן לא נפרט אותה כאן, והיא תפורט בהרחבה בהמשך. יש רק לציין כי במערכות המקובלות כיום הקידוד הוא סקלרי, ואילו שיטות קידוד וקטוריות, שהן חדשות יותר, עדיין פחות מקובלות. לפיכך מתרכזת עבודה זו באלגוריתמים שונים לשלוב של קידוד וקטורי במערכת SBC, כאשר המערכת עם קידוד סקלרי משמשת בעקר כאמת מידה, לצרכי השוואה.

© Technion - Israel Institute of Technology, Elyachar Central Library

1.3 מטרות המחקר

כפי שצוין, מערכת SBC היא בהחלט מועמדת מעשית לתקשורת דיבור בקצבי ביניים. מצד שני, שיטות הקוואנטיזציה הוקטורית נראות כיום בין השיטות המבטיחות ביותר להורדת קצב שדור ושפור איכות השמיעה במערכות תקשורת דיבור. לפיכך נראה ששילוב של שתי הטכניקות יכול להביא למערכת אטרקטיבית למטרות אלו.

ניתן לומר, אם כן, שמטרת המחקר היא לשלב את טכניקת הקוואנטיזציה הוקטורית במקום הקידוד הסקלרי המקובל במערכת SBC, תוך שימת דגש על שני הכוונים הבאים:

1. נסיון למצוא מערכת SBC-VQ עבור קצבי ביניים גבוהים (16Kbps), באיכות טובה כמו זו של המערכת עם קידוד סקלרי (שאיכותה בהחלט טובה), וזאת ללא תשלום נוסף בסיבוכיות.

2. נסיון למצוא מערכת SBC-VQ עבור קצבי ביניים נמוכים (9.6Kbps, 8Kbps) שם אין למערכת כזו תחרות מצד המערכות המקובלות (עם קידוד סקלרי); במקרה זה התשלום בסיבוכיות יכול להיות גבוה יחסית, משום שקשה למצוא פתרון אלטרנטיבי.

במשך המחקר נעשו, לפיכך, סימולציות של מערכות SBC שונות והוסקו לגביהן המסקנות הרלוונטיות: פותחה סימולציה של מערכת מקובלת, עם קידוד סקלרי, ולעומתה הושוו מספר סוגים של מערכות בקידוד וקטורי, באספקטים של איכות וסיבוכיות. לגבי המערכות המומלצות נעשו גם בדיקות נוספות הנדרשות ממערכות קידוד דיבור, כגון עמידות לרעשים וכדומה.

© Technion - Israel Institute of Technology, Elyachar Central Library

1.4 מבנה העבודה

הפרק השני בעבודה מוקדש למערכת SBC המקובלת, בה משתמשים בקוואנטיזציה סקלרית לקידוד האותות בתת-הפסים. לאחר הצגת השיטות המקובלות בקידוד סקלרי, נדונה בהרחבה בעית הקצאת הסיביות שהוזכרה כבר, מאספקטים עיוניים ומעשיים כאחד. מוצגות המערכות שנעשו עליהן סימולציות, המתאימות להקצאה סטטית ודינמית של סיביות בהתאמה, ובמקרה השני עם שני סוגי אדפטציה - קדמית ואחורית.

הפרק השלישי מוקדש כולו לסקירה בנושא הקוואנטיזציה הוקטורית: בתור נושא שהחל לעורר ענין יחסית לאחרונה, הוא דורש הצגה מפורטת וממצה. לאחר הגדרת הבעיה מוצג האלגוריתם המשמש לפתרונה, מבחינה עיונית ומעשית כאחד. מוצגות שיטות מקובלות לחסכון בסיבוכיות (שהיא ה"אגוז הקשה לפצוח" בקידוד הוקטורי), ונסקרות מערכות תקשורת דיבור שונות בהן משתמשים בקוואנטיזציה וקטורית.

בפרק הרביעי מוצגים אלגוריתמים שונים לשלוב הקידוד הוקטורי במערכת SBC. בנוסף לבעיות הסטנדרטיות של תכנון מקודד, שהוצגו בסעיף 1.2, יש צורך גם להחליט מהיכן ילקחו הוקטורים לקידוד. מוצגות שתי סכמות, האחת עם קידוד "אנכי" (הוקטור מכיל דגם מכל פס תדר) והאחרת עם קידוד "אפקי" (הוקטור מכיל מספר דגמים מפס מסוים, וכל פס מקודד בנפרד); במערכת ה"אפקית" (שנמצאה נותנת בצועים טובים יותר) מוצגים גם שפורים אפשריים, במטרה להקטין את הסיבוכיות ולשפר את האיכות.

בפרק החמישי נסקרים אפיוניהן של המערכות המומלצות - סקלרית ווקטורית; בכל מערכת מוצגת הסיבוכיות, וכן עמידות לרעשים ול-tandeming, וטיפול באותות נתונים. כמו כן נעשית השוואה מסכמת בין הקידוד הסקלרי והוקטורי במערכת SBC.

בפרק השישי, החותם את העבודה, מסוכם המחקר, ומוצגים מספר כוונים אפשריים למחקר המשך, במטרה לשפר עוד את המערכות, מבחינת הסיבוכיות והאיכות.

© Technion - Israel Institute of Technology Elyachar Central Library

פרק 2 - מערכת SBC עם קוואנטיזציה סקלרית

2.1 קוואנטיזציה סקלרית

קוואנטיזר סקלרי Q מוגדר פורמלית ע"י מיפוי, $Q:R \dashrightarrow Y$, כאשר $Y=\{y_i\}_{i=1}^{N}$ היא קבוצה של N ערכים ממשיים, הקרויים "ערכי המוצא של הקוואנטיזר". הגדרת המיפוי היא:

$$Q(x) = y_i \quad \text{if} \quad x \in I_i \tag{2.1}$$

כשהקבוצה $\{I_i\}_{i=1}^{N}$ היא קבוצת קטעים, המהווה חלוקה של R. הקטע I_i מוגדר ע"י נקודות הקצה שלו, δ_i ו-δ_{i+1}, וה-δ_i -ים קרויים גם "ערכי ההחלטה של הקוואנטיזר".

הערך $R=Log_2N$ הוא הדיוק של הקוואנטיזר בסיביות לדגם: זהו מספר הסיביות הדרוש ליצוג ערך מוצא. בשמוש בקוואנטיזרים במערכות תקשורת מקובל לשדר עבור הערך y_i את האינדקס שלו, i, או קידוד כלשהו של אינדקס זה, C_i. במשך הדיון נשתמש לכן במינוח "קוואנטיזר" או "מקודד" לחלופין, והכוונה תהיה ברורה מן ההקשר.

הדרך המקובלת למדידת איכותו של קוואנטיזר, היא ע"י הגדרת פונקצית שגיאה (או עוות), $d(x,Q(x))$ (מקובל לסמן $Q(x)=\hat{x}$). אם מתייחסים ל-x כאל משתנה אקראי עם פלוג $p(x)$, אזי מקובל להגדיר כמדד לבצועי הקוואנטיזר את העוות הממוצע

$$D = E\{d(x,\hat{x})\}. \tag{2.2}$$

בהנתן הפלוג של x, $p(x)$, ופונקצית העוות, $d(x,\hat{x})$, ידועים שני תנאים הכרחיים לאופטימליות הקוואנטיזר (במובן של עוות ממוצע מינימלי). תנאים אלה, הידועים כתנאי Lloyd (או Max-Lloyd), יפורטו בתת-סעיף 3.2.1, בהמשך. במאמר קלסי שלו, שלא פורסם, והודפס מחדש כעבור 25 שנה ([7]), נותן Lloyd גם שיטה קונסטרוקטיבית למציאת הקוואנטיזר האופטימלי. עם זאת, עבור אותות כמו אות הדיבור למשל, לא ידוע הפלוג $p(x)$. מסיבה זו, וכן גם ממניעי סיבוכיות, מקובל להשתמש בקוואנטיזר פשוט יותר, הקוואנטיזר היוניפורמי.

© Technion Israel Institute of Technology, Elyachar Central Library

בקוואנטיזר יוניפורמי, כל האינטרוולים I_i הם שווים בגדלם, והערכים y_i הם מרכזיהם. הפרמטר היחיד של הקוואנטיזר הוא גודל האינטרוול, Δ , וכאשר מדברים על קוואנטיזר יוניפורמי אופטימלי, מתכוונים לבחירה אופטימלית של Δ. הקוואנטיזר היוניפורמי האופטימלי מזדהה לקוואנטיזר הכללי האופטימלי רק עבור פלוג כניסה יוניפורמי, אולם הוא נותן תוצאות טובות גם במקרים אחרים.

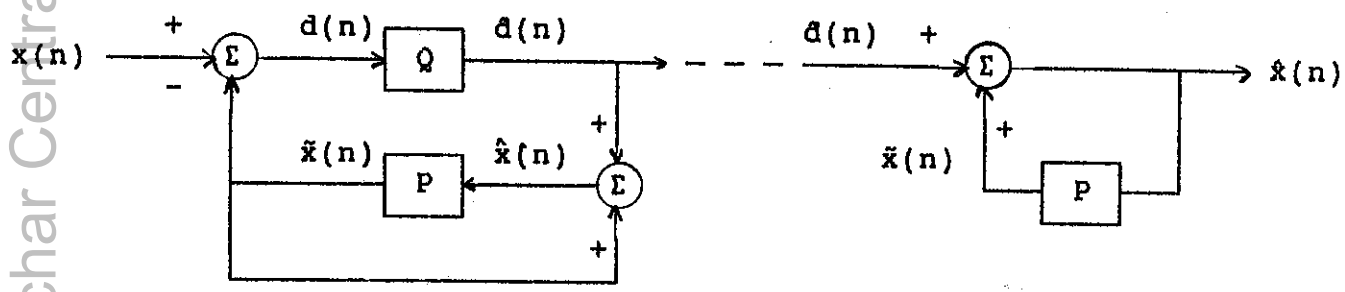
עבור המקרה של אות כניסה לא סטאציונארי, (כמו למשל אות הדיבור), מקובל לתכנן קוואנטיזר אדפטיבי, שיתאים את עצמו לתכונות האות. הדרך הפשוטה והמקובלת ביותר לעשות זאת היא לעבוד עם קוואנטייזר יוניפורמי, כאשר האדפטציה היחידה הנדרשת היא של הגודל Δ. כמו בתחומים רבים אחרים, מקובלות גם כאן שתי שיטות לאדפטציה: קדמית (Forward), המתבצעת ע"ס ערכי אות הכניסה (שאינם ידועים למקלט, ולכן יש לשלוח אינפורמצית צד), ואחורית (Backward), המתבצעת ע"ס ערכים של האות המשוחזר (ואינה מחייבת משלוח אינפורמצית צד).

אדפטציה אחורית היא פשוטה יותר, אולם נותנת תוצאות גרועות יותר בד"כ (וכן רגישה לשגיאות ערוץ). שני סוגי האדפטציה יכולים להיות רגעיים (Instantanous) או מתמשכים (Long-Term).

באותות שהם בעלי יתירות גבוהה, מקובל לשלב במערכת הקוואנטיזציה חזאי, המקטין את יתירות הכניסה לקוואנטייזר. החזאי הוא בד"כ ליניארי, כאשר מספר המקדמים שלו הנו פרמטר בתכנון המערכת. גם החזאי יכול להיות אדפטיבי או קבוע.

מערכת קוואנטיזציה סקלרית כללית (עם אדפטציה אחורית וחיזוי), מוצגת בציור 2.1.

© Technion - Israel Institute of Technology, Elyachar Central Library



ציור 2.1 - מערכת קוואנטיזציה סקלרית כללית.
figure 2.1 - A general scalar quantization system.

הבלוק Q מייצג את הקוואנטיזר ואילו הבלוק P - את החזאי. כאשר הקוואנטיזר הוא יוניפורמי ואדפטיבי והחזאי קבוע, מקובל לקרוא למערכת זו ADPCM (Adaptive Differential Pulse Code Modulation). סכמה כזו, עם מספר וריאציות, שמשה לקידוד האותות בתת-הפסים במערכת SBC עם קוואנטיזציה סקלרית, שתוצג בהמשך פרק זה.

סקירה ממצה של נושא הקוואנטיזציה הסקלרית ניתן למצוא בפרק 4 של [1].

© Technion - Israel Institute of Technology, Elyachar Central Library

2.2 בעית הקצאת הסיביות

בעית הקצאת הסיביות הנה בעיה כללית מתחום הקידוד, השמושית במיוחד במערכות SBC. לפני שנגש להצגת מערכת SBC ספציפית, נדון בבעיה זו באופן כללי.

2.2.1 רקע עיוני (ר' גם [1], פרקים 11-12).

נסתכל על M הדגמים המופיעים יחד במוצא ערוצי מערך מסננים, בעל פסי-תדר שווי רוחב, כעל וקטור אקראי - $\underline{x}=(x_1 \ldots x_M)$. לכל רכיב בוקטור יש ממוצע אפס ווריאנס $\sigma_i^2=E\{x_i^2\}$.

עלינו לקודד כל רכיב x_i, בדיוק של R סיביות לדגם בממוצע, כלומר - ברשותנו MR סיביות לקידוד $\underline{x}$. נשאלת השאלה כיצד ניתן לחלק את MR הסיביות בין רכיבי $\underline{x}$ באופן אופטימלי, במובן של מינימום עוות (שגיאה).

נסמן את מספר הסיביות המוקצה לקידוד הרכיב ה-i בתור R_i. אם פונקצית העוות היא רבוע השגיאה, אנו רושמים את העוות כ-

$$D = 1/M \sum_{i=1}^{M} E\{(x_i-\hat{x}_i)^2\} = 1/M \sum_{i=1}^{M} D_i \qquad (2.3)$$

כאשר $\hat{x}_i = Q_i(x_i)$ היא הורסיה של x_i לאחר קוואנטיזציה, ו-D_i הנו העוות (וריאנס השגיאה) ברכיב ה-i-י.

תוצאה ידועה מתורת הקוואנטיזציה היוניפורמית היא הקשר הבא בין העוות והדיוק של הקוואנטייזר (בהנחת קוואנטיזציה עדינה):

$$D_i = \varepsilon\, 2^{-2R_i}\, \sigma_i^2 \qquad (2.4)$$

כאשר ε הוא (בערך) קבוע של הקוואנטייזר.

© Technion Israel Institute of Technology, Elyachar Central Library

הצבה של (2.4) ב-(2.3) תתן לנו לפיכך-

$$D=\varepsilon/M \sum_{i=1}^{M} 2^{-2R_i} \sigma_i^2 \tag{2.5}$$

(אם מניחים קוואנטייזרים זהים בתת-הפסים השונים). בעית האופטימיזציה העומדת לפנינו היא, לכן, הבעיה הבאה:

$$\min \quad D(R_1 \ldots R_M) \tag{2.6}$$

$$\text{s.t.} \quad \sum_{i=1}^{M} R_i = MR$$

פתרון של בעיה זו בעזרת כופלי לגרנז' (ר' נספח ב'), נותן את נוסחת ההקצאה האופטימלית של סיביות לפסים השונים:

$$R_i = R + 1/2 \log_2 \frac{\sigma_i^2}{\left(\prod_{j=1}^{M} \sigma_j^2\right)^{1/M}} \qquad i=1,2,\ldots,M \tag{2.7}$$

במקרה בו פסי התדר אינם שווי רוחב, ניתן להרחיב בקלות את (2.7) לקבלת התוצאה הבאה:

$$R_i = I/2W + 1/2 \log_2 \frac{\sigma_i^2 / f_i}{\prod_{j=1}^{M} \left(\sigma_j^2 / f_j\right)^{f_j/2W}} \qquad i = 1,\ldots,M \tag{2.8}$$

כאשר I הוא קצב שדור האינפורמציה הכולל במערכת (ב-Kbps), W רוחב הסרט של אות הכניסה (ב-KHz) ו-f_i תדר הדגימה של הפס ה-i-י (ב-KHz). גם ההוכחה לנוסחא זו נמצאת בנספח ב'.

© Technion Israel Institute of Technology, Elyachar Central Library

2.2.2 שיקולים מעשיים

לכאורה, נראה כי נוסחאות (2.7) ו-(2.8) נותנות פתרון מלא לבעיות הקצאת הסיביות במערכות SBC. עם זאת, ישנן שתי בעיות מעשיות שנוסחאות אלו אינן נותנות להן תשובה.

הבעיה הראשונה היא בעיית שערוך הוראיינס. הגדלים σ_i^2 המופיעים בנוסחאות אינם ידועים בד"כ א-פריורי, ובמקרים רבים (כמו למשל באותות דיבור) הם אף משתנים בזמן. לפיכך יש צורך באלגוריתם טוב כדי לשערך גדלים אלה. ניתן לשלב את בעית שערוך הוראיינס בבעית האדפטציה של הקוואנטיזרים בכל פס (שכן, מכיוון שהקוואנטיזר הנו יוניפורמי, האדפטציה של Δ קשורה ישירות לשנוי הוראיינס בפס). גישה זו ננקטה אכן, מאוחר יותר, במערכת עם קוואנטיזציה וקטורית. עם זאת, במקרה של קוואנטיזציה סקלרית, טופלו הבעיות בנפרד.

אלגוריתם נוח לשערוך הוראיינס הינו חישוב של וראיינס הדגם. משערך זה מחושב ע"ס בלוק בן K דגמים, ומתקבלת הסדרה הבאה:

$$\hat{\sigma}_i^2(l) = 1/K \sum_{n=0}^{K-1} x_i^2(lK+n) \tag{2.9}$$

משערך זה מפעיל למעשה "חלון" מלבני על הסדרה לפני סיכום הריבועים, אולם ניתן באותה מידה לבצע שערוך תוך הפעלת חלון אקספוננציאלי דועך:

$$\left[\begin{array}{l} \hat{\sigma}_i^2(n) = (1-\alpha) \sum_{j=0}^{\infty} \alpha^j x_i^2(n-j) \\ \underline{\text{או}} \\ \hat{\sigma}_i^2(n) = \alpha\hat{\sigma}_i^2(n-1) + (1-\alpha)x_i^2(n) \quad . \end{array}\right. \tag{2.10}$$

השערוך בעזרת חלון ריבועי נותן ערך של הוראיינס עבור כל בלוק בן K דגמים; גודל בלוק-קבוע-וראיינס , K, הנו פרמטר של האלגוריתם. לעומתו, שערוך

© Technion - Israel Institute of Technology, Elia Schar Central Library

בעזרת חלון אקספוננציאלי נותן ערך של הואריאנס עבור כל דגם, כאשר הפרמטר α מבקר את מהירות ההשתנות של הואריאנס המשוערך (עם זאת, השמוש בערכי המשערך יעשה כמובן במרווחי זמן גדולים יותר).

אלגוריתמים נוספים לשערוך הואריאנס מתבטסים על נורמות אחרות - שערוך ע"ס סכום ערכים מוחלטים, שערוך ע"ס ערך שיא וכו'. בכל מקרה, השערוך יכול להיות קדמי או אחורי, בהתאם לאינפורמציה עליה מסתמכים.

הבעיה השניה שנוסחאות ההקצאה הדינמית אינן מתחשבות בה, היא העובדה ש-R_i (הקצאת הסיביות לפס ה-i-י) איננו גודל רציף בלתי מוגבל: ראשית, מעוניינים בדרך כלל ש-R_i יהיה גודל שלם (למרות שלמעשה מספיק לדרוש ש-2^{R_i} יהיה גודל שלם, אולם דרישה זו פחות מקובלת). בנוסף לכך R_i הוא אי-שלילי, ונהוג גם להגבילו מלמעלה ע"י גודל R_{max}, כדי למנוע "בזבוז" בהקצאת משאבים.

כדי להתגבר על הגבלות אלו, ניתן לגשת לפתרון בעית הקצאת הסיביות כאל בעית אופטימיזציה בשלמים עם מספר אילוצים. גישה מסוג זה ניתן למצוא ב-[8]. עם זאת מקובלת הרבה יותר היא הגישה האלגוריתמית, המתבססת בדרך כלל על הקצאה ראשונית לפי (2.7) או (2.8), ותיקונה לפי שיטה כלשהי. בהמשך מפורטים מספר אלגוריתמים הפועלים בצורה זו.

אלגוריתם פשוט הוא אלגוריתם חלוקת עודפים בשיטת ה"פוקר". באלגוריתם זה מתבצע ראשית חישוב לפי (2.7) או (2.8), ולאחר מכן קיצוץ התוצאה לערך שלם והגבלתה לתחום $[0, R_{max}]$. לאחר הקצאה ראשונית זו מחושב עודף הסיביות, ומוסיפים סיבית לכל פס-תדר לפי סדר (מתדר נמוך לגבוה או להיפך), עד שהעודף ממוצה.

אלגוריתם אפשרי אחר הנו ורסיה של אלגוריתם המשמש להקצאת סיביות ב-Transform Coder (ר' [9]). בשיטה זו מבצעים ראשית חישוב לפי (2.7) או (2.8), ואז, אם S מוגדרת כקבוצת כל הערוצים שקבלו הקצאה אי-שלילית, מחשבים את -

$$\theta = \frac{\sum_{i \in S} R_i f_i - I}{\sum_{i \in S} f_i} \qquad (2.11)$$

© Technion Israel Institute of Technology Elyacar Central Library

מחברים את θ להקצאות שהתקבלו בחשוב הראשוני עבור אברי S ומבצעים קיצוץ לערך שלם והגבלה ל- $[0,R_{max}]$ (וההקצאות לאברים שאינם ב-S הן אפס). במידה ונשארים עודפים, מקצים בשיטת ה"פוקר".

אלגוריתם שלישי להקצאה, אף הוא מקורו ב-Transform Coding (ר' [5]). שיטה זו מבצעת תיקונים איטרטיביים של ההקצאה, בהסתמך על הרעיון הבא: מכיוון שההקצאה לפי (2.7) או (2.8) מתבצעת ע"פ הלוגריתם של אנרגית פס-התדר, אפשר לנסות להוסיף ללוגריתמים אלה ערך קבוע אופטימלי, כך שלאחר הקצוץ וההגבלה ישאר עודף קטן ככל האפשר. מכיוון שלכל הלוגריתמים מוסיפים אותו קבוע, היחס ביניהם נשמר ונוסחאות (2.7) ו-(2.8) לא משתנות. האלגוריתם מציע, לפיכך, חיפוש של ערכו האופטימלי של הקבוע האדיטיבי, למשל בשיטת החציה ע"י נסוי וטעיה. את העודפים מחלקים כרגיל בשיטת ה"פוקר".

אלגוריתם אחרון מוצע ב-[10] ויתרונו הרב בפשטותו. בשיטה זו מקצים ראשית סיבית אחת לפס-התדר בעל הואריאנס המקסימלי, ומקטינים ואריאנס זה פי r (שהוא פרמטר) כדי "להוציאו מהמשחק"; מבצעים שוב חיפוש ואריאנס מקסימלי והקצאת סיבית אחת, ושוב מקטינים פי r, וחוזר חלילה, עד שחולקו כל הסיביות. r הנו גודל קבוע הנקבע שרירותית, וע"פ נסיונות המופיעים במאמר, הערך r=2 הוא מספק.

לסיום הדיון באספקטים המעשיים של הקצאת הסיביות, יש לציין דבר נוסף - נוסחה (2.4), המשמשת בפיתוח נוסחאות ההקצאה, מבוססת כאמור על הנחת קוואנטייזר יוניפורמי עם קוואנטיזציה עדינה. הנחות אלה לא תמיד מתקיימות (ובעיקר ההנחה השניה), אולם למרות זאת משתמשים בהן. הדבר נובע מהסיבה הפשוטה (והמקובלת), שנוסחאות קצב-עוות מעין אלה אינן ידועות עבור הקוואנטייזרים המעשיים בהם משתמשים.

© Technion - Israel Institute of Technology, Elyachar Central Library

2.3 מערכת עם הקצאה סטטית של סיביות

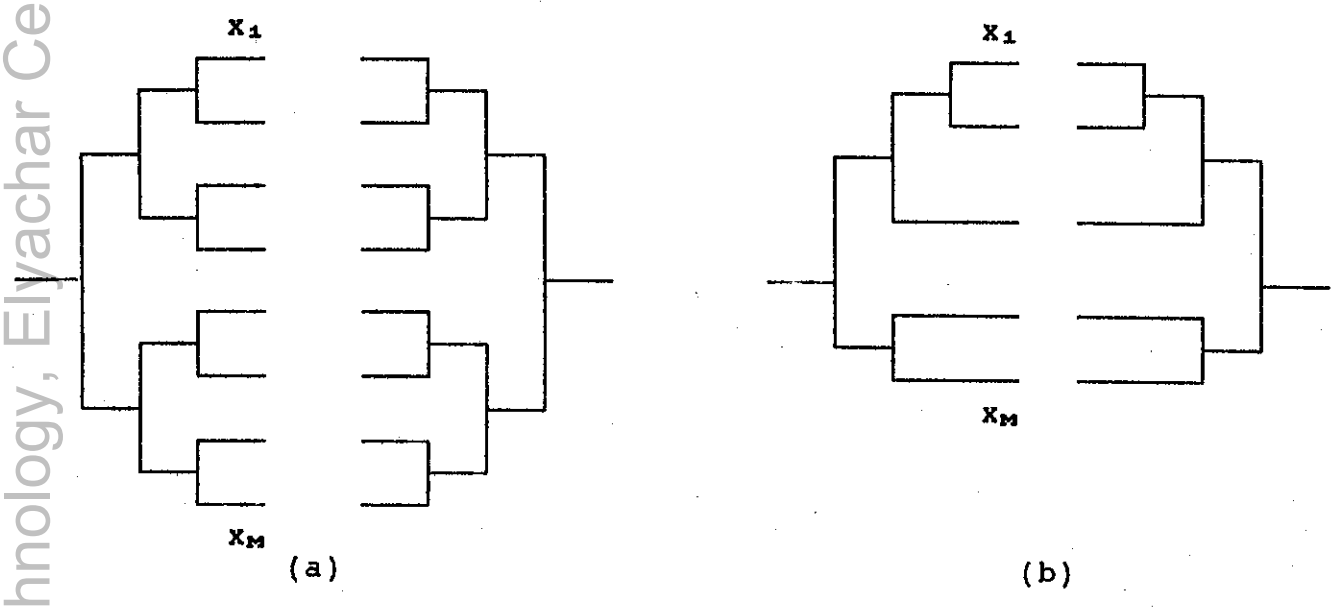
בסעיף זה מתוארת מערכת SBC מעשית המשתמשת בהקצאה סטטית של סיביות. כן מפורטים הנסויים שהביאו לקביעת הפרמטרים שלה, והמסקנות שהופקו מנסויים אלה.

מערכי המסננים שנבדקו היו, כמצוין בסעיף 1.2.1, מערכי QMF במבנה עץ בן שלוש רמות. מכיוון שעבודה זו התרכזה בעיקרה בנושא הקידוד במערכת, נבחרו כל המסננים במערך להיות בני 64 מקדמים. מספר רב כזה של מקדמים מבטיח דיוק גבוה בסינון, כך שאפשר לראות את שגיאת השחזור של המערכת כנובעת מהקידוד בלבד. במדידות שנעשו על מערכת שהכילה רק מסננים, התקבל SNR של מעל 50dB בין האות המקורי והמשוחזר, וזו כמובן איכות מעולה.

במהלך יצירת המערכת הועמדו זה מול זה שני מערכי מסננים. האחד, מערך אחיד, במבנה עץ מלא, המחלק את תחום התדרים של אות הדיבור (0-4KHz) לשמונה פסים שווים בני 0.5KHz כל אחד. השני מבוסס על מערך אחיד במבנה עץ בן שתי רמות (ארבעה פסי-תדר), שפס התדר הנמוך שלו פוצל פעם נוספת, לקבלת חמשה פסים בסה"כ (ר' [1] עמ' 502). המערך המלא הוא קל יותר לממוש בשל אחידותו. המערך החלקי הנו, אמנם, מסובך יותר לממוש, אולם ככל הנראה מותאם יותר לתכונות סוביקטיביות של האוזן, בעזרת חלוקה השמה דגש על רזולוציה גבוהה יותר בפסים נמוכים.

© Technion - Israel Institute of Technology, Elizachar Central Library

המערך המלא והחלקי מוצגים באופן סכמטי בציורים 2.2a ו-2.2b בהתאמה.



ציור 2.2 - מערך QMF במבנה עץ בין שלוש רמות
(a) מערך מלא, M=8 (b) מערך חלקי, M=5.

figure 2.2 - A 3-level, tree-structured, QMF bank
(a) full bank, M=8 (b) partial bank, M=5.

לקידוד האותות בתת-הפסים, נוסו מספר ורסיות של הסכמה הכללית המוצגת בציור 2.1. בדיון נתייחס בנפרד למקודדים בני סבית אחת ולמקודדים בני $2+R_{max}$ סיביות, מכיוון שמקובלים אלגוריתמים שונים עבור שני מקרים אלה. כמו כן נתייחס בנפרד למקודדים פרדיקטיביים ולמקודדים לא פרדקטיביים.

האלגוריתם לקידוד בתחום ערכי הסיביות $2+R_{max}$ קרוי בשם ADPCM (למרות ששם זה יאה למעשה לסכמה הכללית המוצגת בציור 2.1). הסכמה היא בעלת חזאי קבוע במקדם אחד, כלומר -

$$P(Z) = \beta Z^{-1} \tag{2.12}$$

כאשר הבחירה האופטימלית עבור המקדם β היא בחירת מקדם הקורלציה של האות. הקוואנטייזר Q הוא יוניפורמי ואדפטיבי, כאשר האדפטציה, המבוצעת על הפרמטר

© Technion - Israel Institute of Technology, Elyachar Central Library

Δ, היא לפי אלגוריתם ה-One-Word-Memory של Jayant (ר' למשל [1] עמ' 199-210).

אדפטצית One-Word-Memory היא שיטת אדפטציה אחורית, המעדכנת את Δ לפי הנוסחה הבאה:

$$\Delta(n) = M(H(n-1))\cdot\Delta(n-1) \qquad (2.13)$$

כאשר $M(\cdot)$ הוא מערך מקדמים ("כופלים") לבצוע האדפטציה, ו-$H(\cdot)$ הוא מספרו הסדורי של אינטרוול הקוואנטיזציה. העקרון עליו מתבססת שיטת אדפטציה זו הנו הבא: אם האות מקודד בעזרת אינטרוולים בעלי מספר סידורי גבוה, יתכן שהקוואנטייזר אינו מכסה את התחום הדינמי של האות, ועל כן יש להגדיל את Δ; לעומת זאת, קידוד בעזרת אינטרוולים עם מספר סדורי נמוך, מצביע כנראה על רזולוציה גסה מדי של התחום הדינמי הנוכחי, ולכן כדאי להקטין את Δ. בהתאם לכך מתוכנן מערך הכופלים.

הפיכת המקודד ללא-פרדיקטיבי נעשתה במקרה זה בפשטות ע"י אפוס ערכו של מקדם החיזוי β. באופן זה הושגה סכמה פשוטה (שניתן לכנותה APCM), הכוללת קוואנטיזציה (אדפטיבית) בלבד.

האלגוריתם ששמש לקידוד בסיבית אחת, ידוע בשם CVSD (Continuous Variable Slope Delta-modulation). גם כאן הסכמה היא כמוצג בציור 2.1 עם חזאי כבנוסחה (2.12), אולם אלגוריתם האדפטציה של הקוואנטייזר הוא שונה (ר' למשל [3], עמ' 223). נוסחת העדכון היא:

$$\Delta(n) = \Delta(n-1) + \begin{cases} D_1 & \text{if } \hat{d}(n-1)=\hat{d}(n-2)=\hat{d}(n-3) \\ D_2 & \text{otherwise} \end{cases} \qquad (2.14)$$

כאשר בד"כ בוחרים $D_2=0$ (וזהו המקרה של שנויים תכופים ביציאת הקוואנטייזר, כלומר מיצוי התחום הדינמי) ו-D_1 בעל ערך קטן יחסית, המאפשר הגדלה של Δ במצב בו מוצא הקוואנטייזר איננו משתנה תכופות (שהוא כנראה אינדיקציה לחוסר-כיסוי של התחום הדינמי). גם את אלגוריתם CVSD קל להפוך ללא-פרדיקטיבי ע"י אפוס β, כמו באלגוריתם ADPCM.

© Technion - Israel Institute of Technology Ilya Schar Central Library

באשר להקצאת הסיביות במערכת - זו התבצעה, כאמור, באופן סטטי, כלומר נקבעה מראש ולא שונתה. ערכי ההקצאה עבור העץ החלקי נלקחו מ-[1], עמ' 502, והתברר בדיעבד שהם מתאימים לממוצעים לזמן ארוך של הקצאת הסיביות הדינמית. עבור העץ המלא נקבעו ערכי ההקצאה רק לפי סטטיסטיקה לזמן ארוך של הקצאה דינמית. מספרי הסיביות ($R_1 \ldots R_M$) עבור קצבי פעולה של 32Kbps ו- 16Kbps ועבור העץ המלא והחלקי, מפורטים בטבלה 2.1:

	32 Kbps	16 Kbps
עץ מלא	(5,5,5,4,4,4,3,2)	(4,4,2,2,2,1,1,0)
עץ חלקי	(5,5,4,4,3)	(4,4,2,2,0)

טבלה 2.1: הקצאה סטטית של סיביות במערכת SBC.
Table 2.1: Static bit-allocation in a SBC system.

המסקנה החשובה ביותר לגבי המערכת עם הקצאת סיביות סטטית, נוגעת למבנה מערך המסננים. בבדיקות שנערכו במערכת זו, הסתבר שישנו הבדל משמעותי בין השמוש במערך מסננים מלא וחלקי. מערך המסננים החלקי נותן תוצאות טובות הרבה יותר מהמלא, גם אוביקטיבית וגם סוביקטיבית.

בטבלה 2.2 נתונה השוואת ביצועים בין שתי המערכות (שפרמטריהן האחרים מפורטים בהמשך), עבור קצב של 16Kbps. הקצב השני בו נערכו הנסויים, 32Kbps מעניין פחות, שכן המערכת לא משמשת לעבודה בקצב כה גבוה; ב-32Kbps מתברר- ר' גם [1] עמ' 504 - כי מערכות קידוד בתחום הזמן עדיפות על מערכת קידוד בתחום התדר.

© Technion - Israel Institute of Technology, Elyachar Central Library

	SNR(dB)	SNRSEG(dB)
עץ מלא	13.06	11.99
עץ חלקי	15.51	15.83

טבלה 2.2: בצועי מערכות SBC ב-16Kbps.

Table 2.2: Performance of SBC systems at 16 Kbps.

כפי שצוין, ההבדלים בין העצים הינם לא רק אוביקטיביים כי אם גם סוביקטיביים. בשתי המערכות מאופיין האות המשוחזר ע"י צפצופי רקע (תופעה זו אופיינית למערכות SBC עם קוואנטיזציה סקלרית בכלל). צפצופים אלה נובעים, ככל הנראה, מההפרדה הלא-מדויקת בין פסי התדר (במערכות עם קוואנטיזציה), בשילוב עם הקוואנטייזרים הסקלריים המוצעים. תופעת הצפצופים בולטת ומפריעה הרבה יותר במערכת עם עץ מסננים מלא.

הסיבה ליתרונו של המערך החלקי על המלא, היא ככל הנראה, כפי שכבר צוין, בעובדה שהמערך החלקי מתאים יותר לתכונות האוזן: מערך זה נותן רזולוציה טובה יותר בתדרים נמוכים מאשר בגבוהים.

באשר לתכונות המקודדים במערכות בעלות הקצאת הסיביות הסטטית, המסקנה העיקרית היתה כי אין צורך בחיזוי. אולם, מכיוון שמסקנה זו נמצאה גם לגבי המערכת עם הקצאה דינמית של סיביות, ומכיוון שזו נבחרה לבסוף כמערכת ה"סקלרית" המומלצת, נשאיר את הדיון בנושא זה לסעיף הבא, הדן במערכת ה"דינמית".

© Technion - Israel Institute of Technology, Elyachar Central Library

2.4 מערכת עם הקצאה דינמית של סיביות

גם במערכת SBC עם הקצאה דינמית של סיביות, נוסו אותן סכמות סנון (עץ מלא וחלקי) וקידוד (פרדיקטיבי / לא פרדיקטיבי) שפורטו בהרחבה עבור המערכת ה"סטטית" בסעיף 2.3. אולם להבדיל מהמערכת שתוארה שם, נוסו עבור המערכת המתוארת בסעיף זה סכמות שונות של הקצאת סיביות דינמית, שיפורטו בהמשך. כפי שצוין בסעיף 2.2.2, יש לטפל במערכת עם הקצאת סיביות דינמית באלגוריתמים לשערוך הואריאנס ובאלגוריתמים להקצאת סיביות תת-אופטימלית. בנוסף לכך, הועמדו זו מול זו גם סכמות אדפטציה אחורית וקדמית עבור הקצאת הסיביות, וגם הן יתוארו בהמשך.

באשר למבנה מערך המסננים, הסתבר שההבדל בין מערך מלא וחלקי איננו בולט כמו במערכת ה"סטטית": ככל הנראה יתרונה הגדול של ההקצאה הדינמית, הוא חוסר תלות במבנה מערך המסננים, והיא מצליחה לחפות על חסרונותיו של המערך המלא ולהביאו לאותה רמה כשל המערך החלקי. חוסר ההבדל בין שני המערכים הנו בעקר סוביקטיבי, למרות שהוא מתבטא גם במדדים האוביקטיביים, שערכיהם עבור שני העצים קרובים זה לזה יותר מאשר במערכת ה"סטטית".

לגבי המקודדים, הסתבר (כפי שצוין גם עבור המערכת ה"סטטית") כי הוספת החזאי (הקבוע, במקדם אחד) אינה חיונית. אמנם ע"פ המדדים האוביקטיביים יש ירידה בבצועים, אולם הירידה היא קטנה (כ-0.1dB) ובאוזן לא ניתן להבחין בה. תוצאה זו היא הגיונית למדי שכן ההפרדה לפסי-התדר השונים והטיפול בכל פס-תדר בנפרד מקטינים את היתירות של כל אות מקודד (שכן הקורלציה בין דגמי האות בפס התדר קטנה). השפעה זו דומה להשפעתו של חיזוי, ולפיכך הוספת החיזוי איננה משמעותית (למרות שניתן להניח שחזאי עם כמה מקדמים יכול לשפר משמעותית את הבצועים, כמו גם אדפטציה של החיזוי).

מסקנה נוספת הקשורה בקידוד האותות בתת-הפסים, מתייחסת לאדפטציה של המקודד בסיבית אחת. מקודד זה הוא בעייתי בשל הקוואנטיזציה הגסה שהוא מבצע (לרוב בתדרים גבוהים) - קוואנטיזציה גסה כזו מורגשת בד"כ היטב ע"י השומע. לפיכך נוסו שתי סכמות קידוד שונות בסיבית אחת (ללא חיזוי).

השיטה הראשונה היתה שיטת CVSD (עם איפוס של β) כפי שהוצגה בסעיף 2.3. לעומתה נוסתה שיטה שניה, שבה האדפטציה של Δ היא אטית (פעם בבלוק) ומתבצעת

© Technion Israel Institute of Technology, Elyachar Central Library

לפי הנוסחה:

$$\Delta(l) = (1/3\ \hat{\sigma}_1^2(l-1))^{1/2} \tag{2.15}$$

כאשר כאן

$$\hat{\sigma}_1^2(l) = 1/K \sum_{n=0}^{K-1} \hat{x}_1^2(lK+n) \tag{2.16}$$

(נוסחה (2.16) זהה לנוסחה (2.9) פרט לעובדה שהסכימה נעשית על דגמים שעברו קוואנטיזציה).

האדפטציה לפי נוסחה (2.15) מניחה שהאותות בתת-הפסים מפולגים פלוג ז (ר' [1] עמ' 129 טבלה 4.2). הנחה זו, המקובלת עבור אותות דיבור, נבדקה אמפירית, ונמצאה בקירוב טוב נכונה.

בהשוואת שתי שיטות האדפטציה הסתבר שהשיטה ה"איטית" עדיפה יותר. ככל הנראה, בקידוד כה גס, שנויים תכופים מדי בקוואנטייזר רק מפריעים לאוזן. שיטת האדפטציה לפי נוסחה (2.15) הצליחה להביא לביטול חלק נכבד מהרעש במערכת, למרות שלא שנתה באופן משמעותי את התוצאות האוביקטיביות.

בדיקות נוספות שנערכו בסכמת הקידוד הסקלרי הדינמי, כללו שיטות שונות לבקרת המיתוג בין מקודדים בכל פס, במטרה לאלץ הקצאה יותר "חלקה" על פני תקופות זמן ארוכות. מכיוון שנסויים אלה לא העלו שום שנוי משמעותי, לא נפרטם כאן.

המסקנות המרכזיות במערכת עם הקצאה דינמית של סיביות, מתייחסות כמובן להקצאה הדינמית עצמה. נפרט כעת את תוצאות הנסויים על אלגוריתמי שערוך הואריאנס ואלגוריתמי ההקצאה; ההבדלים בין סכמות האדפטציה האחורית והקדמית ידונו בתת-סעיפים נפרדים.

אלגוריתמי שערוך הואריאנס שנוסו במערכת ה"דינמית" היו, כמפורט בסעיף 2.2.2, מבוססים על חשוב ואריאנס הדגם, עם חלונות אקספוננציאלי ומלבני. לאחר שהתברר כי השערוך עם חלון מלבני נותן תוצאות טובות יותר, נעשו

© Technion Israel Institute of Technology, Elyachar Central Library

בדיקות לקביעת פרמטר השערוך K, המציין את גודל הבלוק עליו מתבצע השערוך (ר' נוסחה (2.9)). בבחירת הפרמטר יש Trade-Off ברור בין הצורך בהגדלת הבלוק, כך שהמתוג בין מקודדים שונים לא יהיה מהיר מדי ביחס לכושר הסתגלותם, לבין הצורך בהקטנתו, כדי לעמוד בדרישות הדינאמיות. לאחר שנעשו מספר נסויים נבחר כאופטימלי גודל בלוק של 32 דגמים (32msec בקצב דגימה של 1KHz, לאחר דצימציה).

אלגוריתמי הקצאת הסיביות השונים שנוסו מפורטים בסעיף 2.2.2. המסקנה המענינת שעלתה מההשוואה בין האלגוריתמים השונים היא, כי אין הבדל משמעותי בתוצאות שהם מניבים. מסקנה זו מפתיעה במקצת, מכיוון שהאלגוריתימים השונים מושתתים על תפישות שונות. עם זאת, ניתן להסבירה, לאור העובדה שכל האלגוריתמים קרובים לאופטימליות, ומכיוון שמספר דרגות החופש בהקצאה הוא קטן. לאור זאת הוחלט לבחור באלגוריתם חלוקת העודפים בשיטת ה"פוקר", מתדר נמוך לגבוה, שתואר בסעיף 2.2.2, בשל פשטותו היחסית. הגודל R_{max} (ר' סעיף 2.2.2) נבחר, מבין הערכים 5 ו-6, להיות 5.

הנושא החשוב ביותר עבור המערכת ה"דינמית" הנו האופן בו מתבצעת האדפטציה של הקצאת הסיביות. בסעיפים הבאים יטקרו שתי שיטות האדפטציה שנוסו ותוצאותיהן.

2.4.1 אדפטציה אחורית

כמו בכל מערכת אדפטיבית, גם בהקצאה הדינמית של סיביות ניתן לבחור באדפטציה אחורית (backward) או קדמית (forward).

שיטת האדפטציה האחורית מאופיינת בעיקר ע"י העובדה שאיננה מחייבת שדור אינפורמצית-צד מהמשדר למקלט לצורך בצוע האדפטציה. כדי להמנע משדור אינפורמצית הצד, מתבצעים השנויים במקודדי המשדר והמקלט ע"פ אינפורמציה הנמצאת בשניהם. אינפורמציה זו מתאימה לפיכך לתכונות של קטע דיבור קודם ולא של הנוכחי; בנוסף לכך, זו איננה אינפורמציה מקורית, אלא אינפורמציה שעברה קוואנטיזציה. טכמות אדפטציה אחוריות סובלות גם מרגישות רבה לשגיאות ערוץ, בשל העובדה שכל אדפטציה תלויה בקודמותיה.

© Technion - Israel Institute of Technology, Elyachar Central Library

לפיכך לא מפתיעה העובדה שסכמת האדפטציה האחורית הניבה תוצאות גרועות למדי, גרועות אפילו מאלה של המערכת ה"סטטית". טבלה 2.3 נותנת את האיכות האוביקטיבית של המערכת ה"דינמית" עם אדפטציה אחורית, בקצב של 16Kbps.

	SNR (dB)	SNRSEG (dB)
עץ מלא	7.88	15.15
עץ חלקי	9.28	15.63

טבלה 2.3: בצועי מערכות SBC עם הקצאה דינמית אחורית של סיביות ב-16Kbps.

Table 2.3: Performance of SBC Systems with backward dynamic bit allocation at 16Kbps.

גם האיכות הסוביקטיבית של מערכות אלו היתה גרועה, והורגשו במיוחד אפקטים של סנון-גבוהים ושל התבהרויות והתעמעמויות פתאומיות בדיבור המשוחזר. כפי שצוין קודם, לא הסתמן הבדל משמעותי בתוצאות בין שני מערכי המסננים, המלא והחלקי.

לאור תוצאות אלו הוחלט לנסות סכמת אדפטציה קדמית שבה יש אמנם צורך בשדור אינפורמצית-צד, אולם אינפורמציה זו היא עדכנית ולא-מקוונטטת.

2.4.2 אדפטציה קדמית

בשיטת הקצאת הסיביות עם אדפטציה קדמית, יש ליחד חלק מכמות הסיביות העוברת בערוץ לשדור אינפורמצית צד על האדפטציה, כדי לתאם בין האדפטציות במשדר ובמקלט. ניתן היה לבחור בשדור הואריאנסים המשוערכים, שלפיהם מתבצעת ההקצאה, אולם שדור ההקצאה עצמה הוא פחות בזבזני.

כמות הסיביות הדרושה לשדור ההקצאות מחושבת באופן הבא: אורך בלוק-קבוע-ואריאנס הוא K דגמים; מכאן שבכל שניה יש f_s/K הקצאות בכל פס, כלומר, בסה"כ Mf_s/K הקצאות בשניה. כל הקצאה דורשת 3 סיביות (מספר שלם

© Technion Israel Institute of Technology, Elyachar Central Library

בתחום 0÷5) ולכן כמות הסיביות הכוללת היא $3Mf_1/K$. גודל זה יש כמובן לעגל כלפי מעלה.

עבור הפרמטרים שנבחרו (f_1=1KHz, K=32, M=5-ו M=8 או) התקבל שיש צורך בהעברת כ-500 או כ-750 סיביות בשניה במערכות עם מערך מסננים חלקי ומלא בהתאמה, עבור אינפורמצית הצד. מכיוון שהאינפורמציה העיקרית חייבת להיות משודרת בכל פס בכפולות שלמות של f_1, נקבל שהמערכות עם העץ החלקי והמלא עובדות בקצבים של 15.5Kbps ו-15.75Kbps בהתאמה. "בזבוז" הסיביות השיוריות עד 16Kbps הנו למעשה יתרון, שכן הסיביות ה"עודפות" ינוצלו להגנה על סיביות ההקצאה מפני שגיאות ערוץ (סיביות ההקצאה חשובות ביותר לשחזור תקין של האות).

הביצועים האוביקטיביים של המערכת ה"דינמית" עם אדפטציה קדמית נתונים בטבלה 2.4, עבור קצב של 16Kbps:

	SNR (dB)	SNRSEG (dB)
עץ מלא	18.26	19.38
עץ חלקי	17.06	18.36

טבלה 2.4: בצועי מערכות SBC עם הקצאה דינמית קדמית של סיביות בכ-16Kbps.

table 2.4: Performance of SBC systems with forward dynamic bit allocation at about 16Kbps.

מבחינה סוביקטיבית ניתן לומר כי הדיבור המשוחזר הוא באיכות טובה מאד, עם מעט מאד רעש. גם כאן אין הבדל בין מערכי המסננים השונים, ולכן, בשל ה-SNR הגבוה יותר ונוחיות הממוש, הוחלט לבחור בעץ מלא.

המערכת ה"סקלרית" הסופית, עליה ניתן להמליץ, היא לפיכך בעלת האפיונים הבאים: מערך QMF מלא במבנה עץ בן 3 רמות (כלומר 8 פסים); מקודדים אדפטיביים ללא חזאי (APCM) עם קוואנטיזציה יוניפורמית (2÷5 סיביות - אדפטציה לפי נוסחה (2.13), סיבית אחת - לפי נוסחה (2.15)); הקצאה דינמית

© Technion Israel Institute of Technology Elyacar Central Library

של סיביות עם אלגוריתם הקצאה תת-אופטימלי המבצע שערוך ואריאנס לפי נוסחה (2.9) עם K=32, חשוב לפי נוסחה (2.7), קיצוץ וחלוקת עודפים בשיטת ה"פוקר", מתדר נמוך לגבוה; אדפטצית ההקצאה - קדמית, עם שדור ערכי ההקצאה כאינפורמצית-צד. בקצב של 15.75Kbps נותנת המערכת בצועים אוביקטיביים כמפורט בטבלה 2.4 בשורה הראשונה, וביצועים סוביקטיביים טובים מאד.

עם זאת יש למערכת המוצעת חסרון משמעותי: כאשר מנסים להוריד את קצב השדור לערכים של 8Kbps או 9.6Kbps, האיכות מתדרדרת מאד. ערכי SNR אופיניים לקצבים אלה הם כ-11-13dB, והאיכות הסוביקטיבית נמוכה (מורגשים צפצופים ואפקטים של סנון-גבוהים).

לפיכך הצעד הבא בשפור המערכת הוא שלוב של קוואנטייזרים וקטוריים בקידוד האותות בתת-הפסים, ע"מ לאפשר ירידה בקצב ללא פגיעה משמעותית באיכות. הפרקים הבאים מוקדשים למערכות ה"וקטוריות", כאשר בפרק הבא נסקרת באופן מפורט שיטת הקוואנטיזציה הוקטורית, וכן נסקרים שמושיה במערכות תקשורת דיבור בכלל ובמערכות SBC בפרט.

© Technion - Israel Institute of Technology, Elyachar Central Library

פרק 3 - קוואנטיזציה וקטורית

3.1 בעית הקוואנטיזציה הוקטורית

דומה לאופן שבו הוגדר הקוואנטייזר הסקלרי בסעיף 2.1, ניתן להגדיר את הקוואנטייזר הוקטורי כהכללה שלו. ההגדרה הפורמלית הינה ע"י מיפוי $Q:R^k \to Y$, כאשר $Y=\{\underline{y}_i\}_{i=1}^{N}$ היא קבוצה של N וקטורים ממשיים במימד k, הקרויה "מילון" (הוקטורים במילון קרויים "מילות קוד"). הגדרת המיפוי היא:

$$Q(\underline{x}) = \underline{y}_i \quad \text{if} \quad \underline{x} \in S_i \tag{3.1}$$

כאשר $\{S_i\}_{i=1}^{N}$ היא קבוצת איזורים k-ממדיים המהווה חלוקה של R^k.

ל מילת קוד במוצא הקוואנטייזר מיוצגת ע"י $R=\log_2 N$ סיביות, ולפיכך מיוצג ל דגם מוצא ע"י R/k סיביות (מענין לראות, שבקוואנטיזציה וקטורית ניתן הגיע לרזולוציה קטנה כרצוננו ביצוג, ע"י הגדלת ערך k). אם נכפיל את R/k בקצב הדגימה של האות המקוונטט, f_s, נקבל את קצב העבודה של הקוואנטיזר הוקטורי:

$$I=f_s/k \log_2 N \ (\text{Kbps}) \tag{3.2}$$

ם בקוואנטיזציה וקטורית, כבסקלרית, מקובל למדוד את ביצועי הקוואנטייזר הוקטורי ע"י תוחלת של פונקצית עוות מסוימת:

$$D = E\{d(\underline{x},\hat{\underline{x}})\} \quad . \tag{3.3}$$

פונקצית העוות המקובלת ביותר היא העוות הרבועי, המוגדר ע"י

$$d(\underline{x},\hat{\underline{x}}) = ||\underline{x}-\hat{\underline{x}}||^2 \quad . \tag{3.4}$$

עבור מקור סטאציונארי וארגודי, מתקיים הקשר הבא, המקל על חשוב העוות הממוצע:

$$D = \lim_{n\to\infty} 1/n \sum_{i=0}^{n-1} d(\underline{x}_i,\hat{\underline{x}}_i) \tag{3.5}$$

© Technion - Israel Institute of Technology, Elyachar Central Library

כאשר $\{\underline{x}_i\}_{i=0}^{I}$ הם הוקטורים המיוצרים ע"י המקור.

המודל של קוואנטיזציה וקטורית מופיע בספרות כבר בכתביו של Shannon, ב-1948. עם זאת, תקופה ארוכה לא יושם, שכן לא היתה ידועה דרך נוחה לתכנון ולמימוש של קוואנטיזרים כאלה. לאחר פרסומו של אלגוריתם LBG לתכנון קוואנטיזר וקטורי אופטימלי ב-1980, החלה התעניינות מחודשת בתחום זה. (אלגוריתם LBG ידון בהרחבה בסעיף 3.2).

כיום משמשת הקוואנטיזציה הוקטורית במערכות קידוד רבות, כמו גם בתחומים נוספים (כגון זיהוי דיבור, נקוי מרעש וכד'). בפרק זה ננסה לסקור את נושא הקוואנטיזציה הוקטורית, מצדיו התאורטיים והמעשיים כאחד. מאמרי סקירה מומלצים בנושא זה הם [11] ו-[12].

© Technion - Israel Institute of Technology, Eyachad Central Library

3.2 אלגוריתם LBG לתכנון קוואנטייזר וקטורי אופטימלי

3.2.1 תנאי Lloyd

כפי שצויין בסעיף 2.1, ידועים שני תנאים הכרחיים לאופטימליות של קוואנטייזר, במובן של עוות ממוצע מינימלי. תנאים אלה מפורטים ב-[7] עבור קוואנטייזר סקלרי, אולם ניתן בקלות להכלילם למקרה הוקטורי. בצורתם המוכללת התנאים הם:

(L1): הקבוצה S_i מכילה את כל הוקטורים במרחב ה"קרובים" ל-y_i יותר מאשר לכל מילת קוד אחרת:

$$S_i = \{ \underline{x} \mid d(\underline{x},\underline{y_i}) < d(\underline{x},\underline{y_j}) \ \forall \ j \neq i \}$$

(וקטורים גבוליים משויכים בדרך כלל לפי כלל שרירותי כלשהו).
תנאי זה קרוי גם "כלל השכן הקרוב" (nearest neighbour rule).

(L2): מילת הקוד $\underline{y_i}$ צריכה להיות הצנטרואיד של הקבוצה S_i, כלומר הוקטור בקבוצה שממוצע "מרחקיו" מכל אבריה הוא מינימלי:

$$\underline{y_i} = \text{cent}\ (S_i) = \underset{\underline{u} \in S_i}{\text{argmin}}\ E\{d(\underline{x},\underline{u}) \mid \underline{x} \in S_i\}$$

עבור עוות ריבועי (וגם עוות ריבועי משוקלל), הצנטרואיד של S_i הינו פשוט הממוצע של הקבוצה.

הוכחת תנאי Lloyd (למקרה הסקלרי) נמצאת ב-[7], וניתן להגיע אליה בפשטות תוך הפתרון האנליטי של בעית האופטימיזציה.

נוסף לתנאים עצמם ולהוכחתם, מוצעות ב-[7] גם שתי שיטות נומריות לתכנון קוואנטייזר הסקלרי האופטימלי. אחת משתי השיטות (שיטה II) שמשה באופן סורתי כשיטת התכנון המקובלת לצורך זה. עם זאת, אלגוריתם LBG לתכנון קוואנטייזר וקטורי אופטימלי, מבוסס דווקא על השיטה האחרת (שיטה I), והכללתה למקרה הוקטורי פשוטה יותר.

© Technion - Israel Institute of Technology, Elyachar Central Library

3.2.2 אלגוריתם LBG

אלגוריתם LBG לתכנון קוואנטייזר וקטורי אופטימלי הוצג לראשונה ב-[13], וקרוי לפי ראשי התיבות של מחבריו. במאמר מורחבת שיטה I של Lloyd לתכנון הקוואנטייזר האופטימלי לא רק מבחינת המימד, אלא גם עבור משפחה גדולה של פונקציות עיוות. המחברים מעידים, כי אלגוריתמים דומים כבר הופיעו לפני כן בספרות (ברוב המקרים בהקשר של זיהוי צורות, ותחת הכותרת של "אלגוריתמי קיבוץ" - "Clustering algorithms") אולם לא באופן כה ממצה כבמאמר שלהם.

יש להדגיש את העובדה, כי האלגוריתם משיג מינימום לוקאלי בלבד, מתוך כך שתנאי Lloyd הם תנאים הכרחיים ולא מספיקים. בנוסף לכך, נדרש שהמקור יהיה סטאציונארי-בממוצע-אסימפטוטית (Asymptotically Mean Stationary) (ר. [14] לניתוח אנליטי של האלגוריתם). במידה ולא ידוע מודל סטטיסטי למקור (דבר Lloyd-ו דרש), משתמש אלגוריתם LBG בסטטיסטיקה הנלמדת מסדרת לימוד ארוכה, ושואף להתכנס לקוואנטייזר המתאים לפילוג, ככל שסדרת הלימוד גדלה באורכה. האלגוריתם שימושי בעיקר בוורסיה זו, ולכן זו הוורסיה שתפורט בהמשך. פרוט האלגוריתם הוא:

(0) **אתחול:** נתונים - N, מספר הוקטורים במילון.
$\varepsilon \geq 0$, סף עוות סביר (טולרנס).
Y_0, מילון התחלתי בגודל N (ר. סעיף 3.2.3).
$\{\underline{x}_j\}_{j=0}^{n-1}$, סדרת לימוד של וקטורי מקור.

מאתחלים - $m=0$ (צעד איטרציה).
$D_{-1}=\infty$ (עוות).

(1) בהנתן $Y_m=\{\underline{y}_i^m\}_{i=1}^N$, מצא חלוקה ע"פ עוות מינימלי של סדרת הלימוד: $\underline{x}_j \in S_i^m$, אם הוא קרוב יותר ל-$\underline{y}_i^m$ מאשר לכל $\underline{y}_k^m$ אחר. חשב את העוות הממוצע בצעד זה באופן

$$D_m = 1/n \sum_{j=0}^{n-1} \min_{\underline{y} \in Y_m} d(\underline{x}_j, \underline{y})$$

© Technion - Israel Institute of Technology, Elyacha Central Library

(זהו העוות הממוצע בקידוד עזרת הלימוד בעזרת המילון Y_m).

(2) אם $(D_{m-1}-D_m)/D_m \leq \varepsilon$ עצור: Y_m הוא המילון הסופי.

(3) לקבלת Y_{m+1} מ-Y_m, החלף כל איבר $\underline{y_i^m}$ ב-$\underline{y_i^{m+1}} = cent(S_i^m)$ (כפי שהוגדר בתנאי Lloyd השני).
$m+1 \dashrightarrow m$ וחזור ל-(1).

כפי שניתן לראות, הפעולה של האלגוריתם היא בפשטות הפעלה של תנאי Lloyd, (L1) ו-(L2), לסרוגין על המילון בכל שלב, עד שההתכנסות משביעת רצון. באופן אינטואיטיבי, הצעד הנוקט בכל שלב הוא אופטימלי: עבור מילון נתון, קבוצות ההחלטה $\{S_i\}_{i=1}^{N}$ הטובות ביותר יתקבלו ע"י חלוקת "שכן קרוב" של סדרת הלימוד (למשל - בשלילה), ועבור קבוצות החלטה נתונות, מילות הקוד הטובות ביותר הן הצנטרואידים שלהן (מתוך הגדרת הצנטרואיד בתנאי Lloyd השני).

3.2.3 תנאי התחלה לאלגוריתם

כפי שצויין בסעיף הקודם, אלגוריתם LBG מבטיח התכנסות למינימום לוקאלי בלבד, מתוך העובדה שתנאי Lloyd הם תנאים הכרחיים ולא מספיקים. התוצאה המתקבלת תלויה לכן, בצורה חזקה, בתנאי ההתחלה שנבחרים, כלומר במילון ההתחלתי Y_0. יש לפיכך לבחור את Y_0 בצורה הטובה ביותר האפשרית.

שיטות שונות שהועלו לבחירת מילון התחלתי הן הבאות (ר. [11]):

- מילון התחלתי אקראי: בחירת N וקטורים מסדרת הלימוד כמילון התחלתי (למשל N הראשונים, או N המרוחקים ביותר זה מזה), או בחירה שרירותית.

- מילון מכפלה: המילון ההתחלתי הוא מכפלה קרטזית של מילונים ממימדים נמוכים יותר (למשל - k מילונים סקלריים).

© Technion - Israel Institute of Technology, Elyachar Central Library

- מילון פיצול (Splitting): שיטה זו, שהוצגה ב-[13], נותנת אמפירית את התוצאות הטובות ביותר, והיא שימשה ליצור המילונים ההתחלתיים בניסויים שנערכו בעבודה זו. לפיכך תפורט להלן:

(0) אתחול: Y_0^1 הוא הצנטרואיד של כל סדרת הלימוד.
M=1 (גודל המילון).

(1) בהנתן $Y_0^M=\{\underline{y_i}\}_{i=1}^M$, פצל כל מילת קוד $\underline{y_i}$ לשתי המילים $\underline{y_i}\pm\underline{\delta}$, כאשר $\underline{\delta}$ וקטור פרטורבציה קבוע. התקבל Y_0^{2M}. M --> 2M.

(2) אם M=N, $Y_0=Y_0^M$ ועצור. אחרת $Y_0^M=LBG(Y_0^M)$ וחזור ל-(1).

כפי שניתן לראות, אלגוריתם הפיצול בונה מילון התחלתי בגודל N ע"י פרטורבציה של מילון אופטימלי בגודל N/2 (מניחים כאן תמיד ש-N הוא חזקה שלמה של 2, אולם זו בלאו הכי דרישה מעשית מקובלת). עבור וקטור פרטורבציה $\underline{\delta}$ קטן וסדרת לימוד טיפוסית, צעד הפרטורבציה יקטין את העוות הממוצע, (או במקרה הגרוע ישאירו ללא שינוי), וניתן לכן להסתכל על כל תהליך יצור המילון, ממילון התחלתי בגודל 1 ועד מילון סופי בגודל N, כעל תהליך אופטימיזציה.

נקודה שכדאי להדגישה כעת (כהכנה לדיון במילונים במבנה עץ, שידונו בסעיף 3.3.2) היא, שהרצת אלגוריתם LBG בשלב (2) של סכמת הפיצול מתבצעת עם סדרת הלימוד כולה, ולכן מבנה המילון ההתחלתי איננו נשמר, ונעלם אחרי ביצוע איטרציות ה"תקון" של אלגוריתם LBG.

3.2.4 בעית התא הריק

כמו פתרונות תאורטיים רבים (ראה למשל בעית הקצאת הסיביות שנדונה בסעיף 2.2), גם אלגוריתם LBG סובל ממספר בעיות מעשיות כשבאים ליישמו במציאות. החשובה בהן, שעליה נרחיב כעת את הדיון, היא בעית התא הריק (ראה למשל [15], שהוא מאמר המשך של [13]).

תא ריק הינה קבוצת החלטה S_i שאיננה מכילה אף וקטור למוד, ולכן חסרת משמעות במסגרת האלגוריתם. הווצרותו של תא ריק היא כמעט בלתי נמנעת, כאשר סדרת הלימוד דלה מדי (כלומר "יחס הלימוד", או היחס בין גודל הסדרה, ת,

© Technion - Israel Institute of Technology, Elyachar Central Library

וגודל המילון, N , הוא קטן מדי - על פי רוב פחות מ-10). עם זאת, תא ריק יכול גם לנבוע כתוצאה מבעיות בסכמת האתחול של האלגוריתם. למשל, אם המלון ההתחלתי נבחר באקראי, יכול להיות תא התחלתי מרוחק מאוד, שלא יצטרפו אליו כלל וקטורי לימוד.

בשיטת הפיצול תווצר בעיה של תא ריק כאשר ננסה לפצל תא סינגולרי (כלומר תא המכיל וקטור לימוד בודד, בין אם בהעתק אחד או במספר העותקים; תא כזה מאופין ע"י העובדה, שהעוות הכולל בו הוא אפס). פיצול תא סינגולרי יתן תא אחד שיכיל את התוכן המקורי, ותא שני ריק.

ון הראוי להעיר, כי בעית התא הריק לא תווצר בשיטת הפיצול עבור תא לא סינגולרי, שכן הפיצול מתבצע על צנטרואיד התא, הנמצא במרכז אבריו. אם נזניח בעיות נומריות ובעיות של בחירת וקטור הפרטוברציה $\underline{\delta}$, סביר להניח, שבערך מחצית מאברי התא יפלו בכל אחד משני תתי-התאים שנוצרו.

פתרון מיידי לבעית התא הריק הוא הבא: לא לפצל את התא הבעייתי, ובמקום זאת לבצע פיצול נוסף לתא אחר. השאלה הנשאלת היא - איזה תא עדיף לבחור.

בחירה הגיונית היא בחירת התא בעל העוות המקסימלי. הבעיה העקרית היא, שתא זה יכול להכיל מעט איברים (לא טפוסיים) ולכן פיצול כפול שלו עלול להפוך את האזור, בו הוא נמצא, למוקד לבעית תא-ריק נוספת. בחירה אפשרית אחרת, היא של התא המאוכלס ביותר. בחירה זו נראית אומנם כאילו היא מונעת את הבעיה, אולם למעשה תא בעל מספר איברים גדול יכול להכיל עותקים רבים של אותו וקטור, ואנו חוזרים לבעית תא-ריק נוספת.

נראה, אם כן, שהשיטה הטובה ביותר היא לבחור תא, שיהיה בעל עוות ואוכלוסיה גדולים כאחד. הפתרון שאיתו עבדנו, ושנתן תוצאות טובות, מסתכל כמדד על הכפלת האוכלוסיה והעוות בכל תא, ובוחר לפיצול נוסף את התא בעל המדד המקסימלי. יש לציין, עם זאת, כי בסדרות לימוד דלות מאוד, גם פתרון זה נכשל, ואין מנוס מהגדלת הסדרה.

© Technion - Israel Institute of Technology, Elyachar Central Library

3.3 קוואנטייזרים וקטוריים תת-אופטימליים

הבעיה העיקרית בשימוש מעשי בקוואנטיזציה וקטורית היא הסיבוכיות הגבוהה שיש להתמודד עמה, גם מבחינת זמן (לקידוד ולפענוח) וגם מבחינת מקום (אחזקת המילון). קידוד וקטור בעזרת המילון שתוכנן ע"י אלגוריתם LBG, מתבצע ע"י חיפוש מלא במילון, למציאת מילת הקוד הקרובה ביותר לוקטור המקודד. המילון עצמו הוא סט של N וקטורים k-ממדיים, שחייבים להיות מוחזקים בשלמותם גם במקלט וגם במשדר. לפיכך הוצעו בספרות אלגוריתמים תת-אופטימליים רבים, שנועדו להקטין את סיבוכיות המימוש ולאפשר את השימוש בקוואנטיזציה הוקטורית באופן מעשי. בפרק זה נסקור את החשובים בהם, הנוגעים ישירות לעבודה זו.

3.3.1 קודי מכפלה

הרעיון בקוד מכפלה הוא, להגדיר, על הוקטור שיש לקודד, מספר תכונות, שכולן יחד מהוות תאור מלא שלו, ולקודד כל אחת מהן בנפרד. אם נגדיר A תכונות, ונקודד כל תכונה ב-b_i סיביות, נקבל A מילונים שכל אחד מהם מגודל 2^{b_i}, כלומר הגודל הכולל יהיה

$$\sum_{i=1}^{A} 2^{b_i}$$

וזאת במקום מילון אחד שגודלו -

$$2^{\sum_{i=1}^{A} b_i}$$

הדבר מביא כמובן לחסכון משמעותי במקום ובזמן קידוד.

הבעיה המרכזית היא הגדרה טובה של התכונות. אפשרות טריוויאלית היא להגדיר כל תכונה בתור תת-וקטור של הוקטור לקידוד, כלומר - וקטור מסדר קטן יותר. במקרה הקיצוני נחזור ל-k מקודדים סקלריים שכל אחד מהם פועל על אחד מרכיבי הוקטור. עם זאת, ישנן הגדרות עדיפות המותאמות לתכונות האותות המקודדים.

מקודד מכפלה קלאסי הוא מקודד מטפוס "הגבר/צורה". במקודד זה מפרידים את הוקטור המקודד לפי שתי תכונות - ההגבר (נורמה) שלו, והצורה (הוקטור המנורמל). את ההגבר מקודדים סקלרית, בדרך כלל לוגריתמית, כדי להתחשב בתחום הדינמי הרחב של הכניסה. הצורה מקודדת וקטורית. אם מקדישים R_g

© Technion - Israel Institute of Technology, Eyal Gal Central Library

ז יביות לקידוד ההגבר ו-R_s סיביות לקידוד הצורה, נקבל שני מילונים בגודל 2^{R_g} ו-2^{R_s} בהתאמה, במקום מילון אחד בגודל $2^{R_g+R_s}$. מקודדים כאלה הוצגו ראשונה ב-[15], ולאחר מכן שוכללו במאמרים רבים נוספים. ואריאנט של מקודד הגבר/צורה הוא מקודד מפריד-ממוצע, המקודד לחוד את ממוצע הוקטור ואת הוקטור הממורכז (השיטה הראשונה מקובלת בעקר בדיבור, והשניה בעיקר בתמונות - ר' [11]).

זוג אחר של קודי מכפלה עולה בקידוד וקטורים מרוכבים (למשל וקטורי התמרה). במקרה זה בחירות הגיוניות להפרדת הוקטור לתכונותיו הן למשל - הפרדת החלק הממשי מהחלק המדומה, או הפרדת האמפליטודה והפאזה.

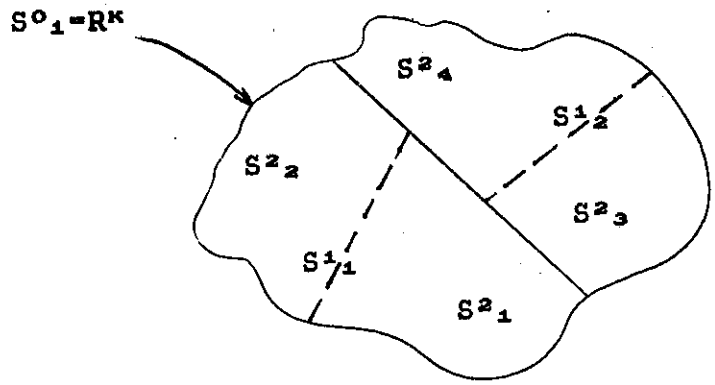
3.3.2 קוואנטיזר במבנה עץ

קוואנטייזר וקטורי תת-אופטימלי המשפר באופן ניכר את סיבוכיות זמן הקידוד על חשבון סיבוכיות המקום, הינו קוואנטייזר במבנה עץ. קוואנטייזר זה הוצג ראשונה ב-[15].

קוואנטייזר במבנה עץ מסתמך על כך שבמשדר ובמקלט מוחזק לא רק קוואנטייזר וקטורי מגודל N, אלא גם קוואנטייזרים מגודל 1,2,...,N/4,N/2 (N צריך להיות חזקה שלמה של 2). סדרת הקוואנטייזרים הזו מקיימת את התכונה הבאה: קבוצות ההחלטה של הקוואנטייזר מגודל 2^i הן עידון של קבוצות ההחלטה של

© Technion - Israel Institute of Technology, Elyachar Central Library

הקוואנטייזר מגודל 2^{i-1}, באופן המוצג בציור 3.1:



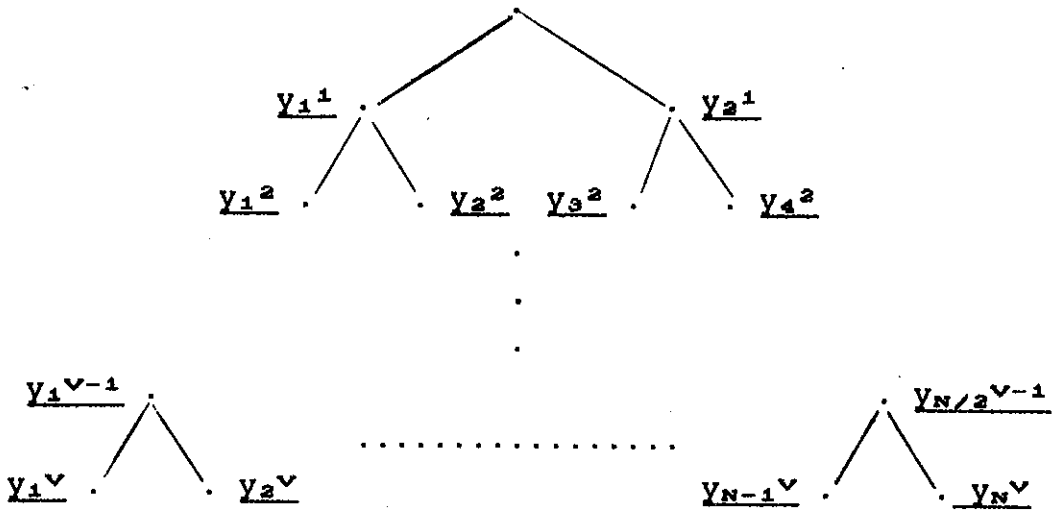
ציור 3.1 - קבוצות ההחלטה עבור קוואנטייזר-עץ מגודל 4.

Figure 2.1 - Decision regions for a size-4 tree-structured vector quantizer.

כפי שניתן לראות, קבוצות ההחלטה של הקוואנטייזר מסדר 4 נוצרות בציור 3.1 ע"י פיצול אופטימלי של קבוצות ההחלטה של הקוואנטייזר מסדר 2. כדי לייצר קוואנטייזר העומד בדרישה זו, יש להפעיל את שיטת הפיצול (ר' סעיף 3.2.3) על סדרת הלימוד, אולם להקפיד להפעיל את אל'גוריתם LBG לאחר הפיצול בכל קבוצה, רק על סדרת הלימוד המתאימה לקבוצה זו, ולא על כל סדרת הלימוד. באופן זה, קבוצות ההחלטה שנוצרות לאחר הפיצול והאופטימיזציה, מוכלות בקבוצת ההחלטה המקורית, דבר שלא מתקיים בשיטת יצור הקוואנטייזר האופטימלי, אך הכרחי ביצור קוואנטייזר במבנה עץ.

© Technion - Israel Institute of Technology, Elyachar Central Library

קבוצת המילונים הנוצרת מוחזקת במבנה עץ, כפי שניתן להבין מהשם. מבנה כזה מתואר בציור 3.2 (האינדקס העליון מייצג כאן את הרמה בעץ).



ציור 3.2 - קוואנטייזר וקטורי במבנה עץ. $V=\log_2 N$.

Figure 3.2 - a tree-structured vector quantizer. $V=\log_2 N$.

הקידוד בעזרת מילון העץ מתבצע באופן הבא: הוקטור לקידוד משווה ראשית ל-y_1^1 ו-y_2^1. אם הוא קרוב יותר ל-y_1^1, הוא יושווה בהמשך ל-y_1^2 ו-y_2^2, ואם הוא קרוב יותר ל-y_2^1, הוא יושווה בהמשך ל-y_3^2 ו-y_4^2. כך ממשיכים, עד למציאת הוקטור המתאים ברמה האחרונה, שהוא מילת הקוד המבוקשת.

כפי שניתן לראות, סיבוכיות הקידוד קטנה מ-$O(N)$ ל-$O(\log_2 N)$ (מספר ההשוואות הוקטוריות לביצוע), ואילו האינפורמציה שיש לאכסן כעת היא $2N-1$ וקטורים במקום N (שכן יש לשמור את כל "מילוני העזר" בעזרתם מתבצע הקידוד, בנוסף למילון בגודל N).

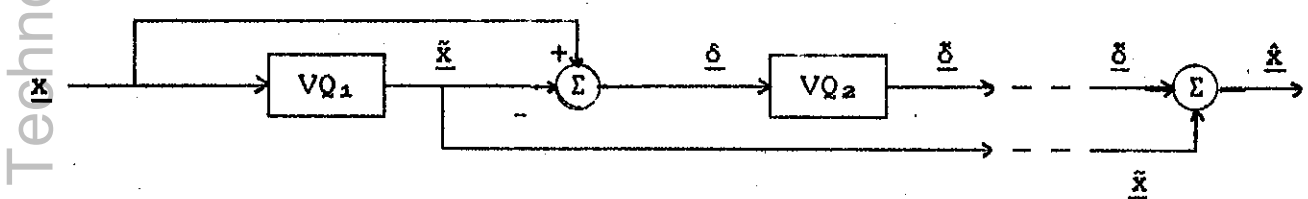
ובן, שמבנה של עץ בינארי אינו הכרחי, למרות שהוא הנוח והמקובל ביותר למימוש. ניתן בפרוש לבנות מילונים במבנה עץ m-ארי, כאשר גם על המילון האופטימלי ניתן להסתכל כעל מילון-עץ N-ארי, בעומק 1.

© Technion - Israel Institute of Technology, Elyachar Central Library

בהשוואה למילונים אופטימליים, הדגרדציה המתקבלת במילוני-עץ היא קטנה יחסית, ואילו החסכון בזמן הוא כאמר משמעותי מאוד, ומכאן האטרקטיביות הגדולה של מילונים אלה.

3.3.3 קוואנטייזר רב-דרגות

קוואנטייזר רב-דרגות (Multistage) הוצג לראשונה ב-[16]. המדובר במערכת תת-אופטימלית, שבה מקודדים תחילה את וקטור הכניסה באמצעות קוואנטייזר וקטורי גס יחסית, ולאחר מכן יוצרים את אות ההפרש בין הכניסה והיציאה, ומקודדים אותו בעזרת קוואנטייזר נוסף. על תהליך זה ניתן לחזור מספר פעמים כרצוננו. הורסיה המקוונטטת הסופית, מורכבת כמובן מסיכום מוצאי הרמות. קוואנטייזר רב דרגות בעל שתי דרגות מוצג בציור 3.3.



ציור 3.3 - קוואנטייזר רב-דרגות בעל שתי דרגות
Figure 3.3 - A two stage multistage vector quantizer

בכדי לעבוד באותו קצב כמו קוואנטייזר אופטימלי בעל $N=2^V$ מילות קוד, הקוואנטייזרים בכל דרגה צריכים להיות בעלי $N_i=2^{V_i}$ מילות קוד, כך ש-

$$\sum_{i=1}^{M} V_i = V$$

(ו-M מספר הדרגות). למשל, במערכת בת שתי דרגות, שוות קצב, $V_1=V_2=V/2$, כלומר למילונים בכל דרגה יש $N^{1/2}$ מילות קוד. מכאן עולה מיד החסכון הנכבד במקום ובזמן קידוד.

המגרעת העיקרית של קוואנטייזר רב-דרגות היא בהנחה, שהמקור $\underline{\delta}$ (ר' ציור 3.3) מכיל וקטורים שווי-פילוג. למעשה וקטורי ההפרש הנובעים משתי קבוצות

© Technion - Israel Institute of Technology, Elyachar Central Library

חלטה שונות של המילון בדרגה הראשונה, הם בעלי פילוג שונה. לפיכך קיבוץ
ל וקטורי ההפרש לשם קידודם ביחד בדרגה השניה גורם לדגרדציה חזקה
בצועים, ביחס לקוואנטייזר אופטימלי. עם זאת, החסכון הנכבד בסיבוכיות
עניק בכל זאת לקוואנטייזר רב-דרגות אטרקטיבות רבה.

© Technion - Israel Institute of Technology, Elyachar Central Library

3.4 מערכות קידוד דיבור עם קוואנטיזאציה וקטורית

כפי שצויין בסעיף 3.1, הרי בעקבות פיתוחו של אלגוריתם LBG, החלה הקוואנטיזציה הוקטורית לתפוש מקום גדל והולך במערכות קידוד דיבור. בסעיף זה יסקרו השמושים החשובים של קוואנטיזציה וקטורית במערכות כאלה, וכן יפורטו מספר ניסויים שנערכו על ידינו באותן מערכות.

3.4.1 קוואנטיזציה וקטורית במקודדי מקור

קידוד מקור (Source Coding), להבדיל מקידוד צורת גל (Waveform Coding), הינה שיטת קידוד דיבור בה מושם הדגש על קידוד פרמטרי המודל של יצור הדיבור, ולא על שימור של צורת הגל הספציפית שנוצרה. באופן טיפוסי, משמשים מקודדי מקור בתחום קצבים נמוך בהרבה מזה של מקודדי צורת גל, אולם איכותם נמוכה יחסית, והם לוקים בבעיות עמידות קשות, לרעש רקע, לריבוי דוברים וכדומה.

מרבית מקודדי המקור מסתמכים על שיטת קידוד הידועה בשם LPC (Linear Predictive Coding). בשיטת קידוד זו מיוצג אות הדיבור כתגובה של מערכת ליניארית משתנה בזמן, לערור, הקרוי גם "אות-שארית" (residual). זהו למעשה מודל של מערכת יצור הדיבור, כמורכבת מפעולת המעבר הקולי על הערור המיוצר במיתרי הקול. במערכות כאלה יש לשדר את פרמטרי המערכת הליניארית (פרמטרי LPC) וכן קידוד כלשהו של אות הערור. כל אחד משני אלה יכול להתבצע על ידי קוואנטיזר וקטורי.

המערכות הראשונות שישמו קוואנטיזציה וקטורית במקודדי דיבור, השתמשו בה לשם קידוד פרמטרי ה-LPC. פרמטרים אלה ניתנים בדרך כלל לקיבוץ, ליצירת וקטורים במימד קטן יחסית, ובאופן זה התקבלה ירידה משמעותית בקצב העבודה של המקודדים. מערכת כזו מוצגת אפילו במאמר המקורי על אלגוריתם LBG ([13]).

לאחרונה הוחל לבצע גם קידוד של אות השארית בטכניקות של קוואנטיזציה וקטורית. אות זה הוא קשה לקידוד, משום שמעוניינים להעבירו בקצב נמוך מאוד, ועם זאת לנצל את מירב האינפורמציה שבו. מבנה אות השארית גם איננו מותאם מיידית לקידוד וקטורי, כפי שמותאמים פרמטרי ה-LPC. עם זאת, תולים

© Technion - Israel Institute of Technology, Elyachar Central Library

:יום תקווה גדולה במערכות המכונות CELP (Code Excited Linear Prediction)
ו גם VXC (Vector Excited Coding), בהן מקודד בד"כ אות השארית בעזרת
וילון של וקטורים אקראיים במימד גבוה. המאמר המקורי בנושא זה הוא [17],
מאז הופיעו מאמרים רבים על מערכות כאלה. מערכות CELP מאפשרות תקשורת
יבור באיכות טובה אף ב-4.8 Kbps. עם זאת, מגרעתן הגדולה היא בסיבוכיות
העצומה שלהן (כ- 5x10⁹ mads/sec), שאי אפשר כיום להתמודד אתה בזמן אמת,
מסגרת הטכנולוגיה העכשווית.

3.4.2 מערכות עם חיזוי

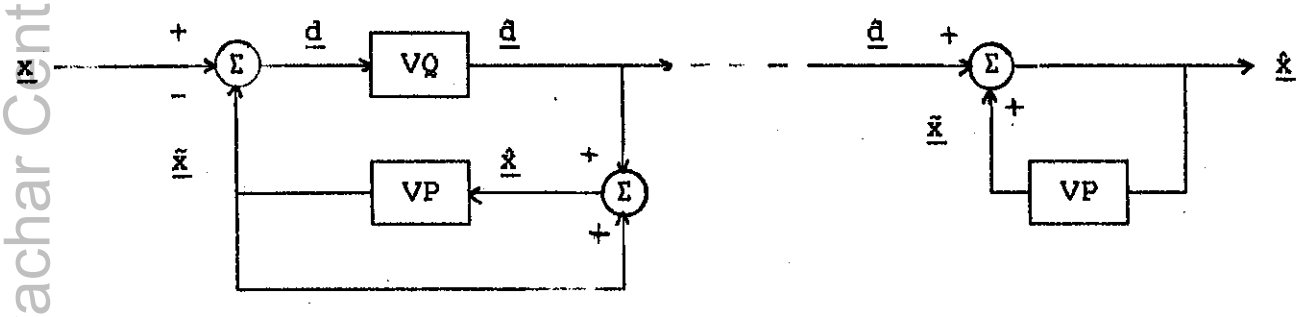
:קידוד וקטורי של אותות דיבור, ניצול היתירות שבאות הוא גדול יותר מאשר
:קידוד סקלרי, בשל עצם העובדה שמקודדים קבוצה של מספר דגמים (תלויים)
:מקום דגם יחיד. עם זאת, יש עדיין מקום לניצול טוב יותר של יתירות זו בעת
הקידוד: באותו אופן שמנצלים עקרונות של חיזוי בקידוד סקלרי, אפשר לנצל
נקרונות אלה גם בקידוד וקטורי.

אחת השיטות הראשונות שהוצעו היא שיטה של "מילונים מותנים". הסתבר כי,
:הנתן תוצאת הקידוד ברגע מסוים, הוקטור הבא במוצא הקוואנטייזר נבחר מתוך
ושפחה מצומצמת של וקטורים, קטנה בהרבה מאשר המילון כולו. אבחנה זו קוראת
לקוואנטיזציה בעזרת מכונת מצבים (ר' למשל [18], אך רעיון זה פותח ע"י
זוקרים נוספים, במקביל). בשיטה זו מגדיר הקידוד בכל רגע מצב של המכונה,
וע"פ תוצאת הקידוד והמצב הנוכחי, עוברת המכונה למצב שונה, לפי פונקצית
זעברים נתונה. מובן שיש להגדיר אלגוריתם לתכנון פונקצית המעברים, וכן
לתכנון סט המילונים המתאימים למצבים השונים, ואכן מוצעות שיטות כאלה
:ספרות.

עם זאת, שיטה אטרקטיבית הרבה יותר היא הכללה פשוטה של החיזוי, המתבצע
במערכות DPCM סקלריות, למקרה הוקטורי. הכללה כזו הוצעה לראשונה ב-[19],
וקרויה שם VPC (Vector Predictive Coding). ניתן להסתכל עליה כעל מכונת
מצבים עם מספר אינסופי של מצבים (בבדיקות שנעשו שם התברר כי, ככל שמספר
המצבים גדל, מתקרבת מכונת המצבים הסופית למערכת VPC). מכיוון ש-VPC נוחה
יותר לתכנון ומימוש, ברור שהיא העדיפה מבין השתיים. סכמה של שיטת VPC

© Technion - Israel Institute of Technology, Elyachar Central Library

מוצגת בציור 3.4 (למעשה זו בדיוק ורסיה וקטורית של ציור 2.1).



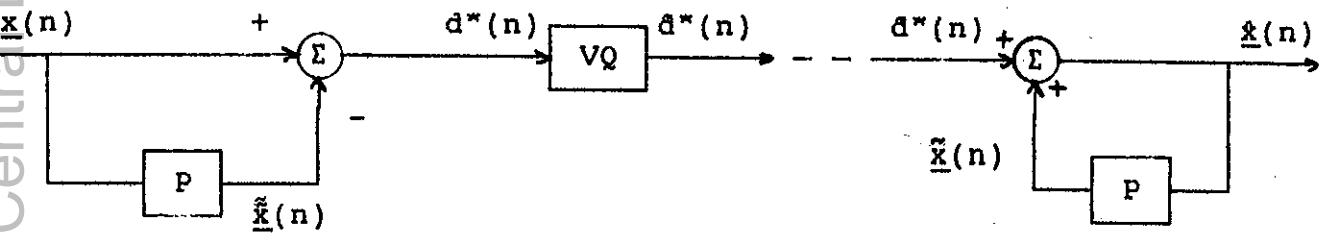
ציור 3.4 - מערכת VPC.

Figure 3.4 - A VPC system.

הבעיה העיקרית של VPC, היא בעובדה שהחיזוי חייב להיות וקטורי, מתוך כך שרוצים לעבוד במערכת עם חוג-סגור (על מנת שבמשדר ובמקלט יתבצע החיזוי על אותו האות). לשם כך מקדמי החזאי צריכות להיות מטריצות. במידה ורוצים לעבוד עם חזאי אדפטיבי, בעל מספר מטריצות חיזוי (כמקובל במערכות סקלריות), הדבר מגדיל את סיבוכיות המערכת במידה ניכרת. בנוסף לכך, חיזוי סקלרי עדיף על חיזוי וקטורי מבחינת הורדת היתירות, כאשר מדובר על תהליכים סקלריים שדגמיהם קובצו מלאכותית ליצירת וקטורים. לפיכך רצוי למצוא שיטה המפאשרת שלוב של חיזוי סקלרי עם קוואנטיזציה וקטורית.

השיטה הפשוטה ביותר היא כמובן לעבוד בחוג פתוח. במסגרת הכנת התשתית לשימוש בקוואנטיזציה וקטורית בנסויים עם מערכת SBC, נוסתה על ידינו גם ורסיה כזו של קידוד וקטורי פרדיקטיבי. סכמה של המערכת מוצגת בציור 3.5.

© Technion - Israel Institute of Technology, Elyachar Central Library



ציור 3.5 - מערכת עם קוואנטיזציה וקטורית וחיזוי סקלרי.
Figure 3.5 - A System with vector quantization and scalar prediction.

כדי להמחיש את כוחה של הקוואנטיזציה הוקטורית לעומת הסקלרית, הושוו על ידנו ביצועיה של מערכת כזו עם וקטורים במימד 4, לביצועי מערכת זהה לה, עם קוואנטיזציה סקלרית. למען הוגנות ההשוואה, נבנתה המערכת הסקלרית סביב קוואנטייזר לא אדפטיבי שתוכנן אף הוא ע"י אלגוריתם LBG. החזאי היה קבוע ובעל מקדם אחד, שערכו נקבע להיות מקדם הקורלציה של סדרת הלימוד. בקצב של 16Kbps השיגה המערכת ה"וקטורית" SNR של כ-13dB, לעומת SNR של כ-5dB במערכת ה"סקלרית", וביצועיה הסובקיטיביים היו טובים בהרבה. גם ב-8Kbps היה הפרש ניכר בין התוצאות האובייקטיביות בשתי המערכות, אולם השיפור הסוביקטיבי היה פחות בולט.

בהשוואה לקוואנטייזר הוקטורי האופטימלי נוסו במערכת זו גם קוואנטייזרים תת-אופטימליים במבנה עץ ובמבנה רב-דרגות (ר' סעיף 3.3). הניסויים נערכו בקצב של 16Kbps, והתוצאות היו כצפוי: הקוואנטייזר מטיפוס עץ האיץ מאוד את פעולת הקידוד, והדגרדציה באיכות שהפיק היתה קטנה למדי (סוביקטיבית, וגם אוביקטיבית - כ-1dB); קוואנטייזר במבנה רב דרגות, עם שתי דרגות זהות, הביא לחסכון משמעותי בסיבוכיות הזמן והמקום, אולם לדגרדציה רצינית בביצועים (כ-3dB).

לסיכום הדיון במערכות קוואנטיזציה וקטורית עם חיזוי, יש לציין כי לאחרונה יצא מאמר המציג טכניקה של שלוב חזאי סקלרי וקוואנטייזר וקטורי במערכת עם חוג סגור ([20]). השיטה המוצגת שם מבוססת על גישה של אנליזה-סינתיזה, ובנויה על העובדה כי מוצא הקוואנטייזר של המערכת עבור וקטור כניסה כלשהו, הוא תמיד אחד מ-N וקטורים קבועים. לפיכך ניתן לייצר מכל אחד מ-N הוקטורים (ומתנאי הסיום של הקידוד הקודם) וקטור הפרש בכניסת הקוואנטייזר (פשוט ע"י שמוש בחזאי הסקלרי), ולבחור את וקטור המוצא, שהעוות בינו לבין הכניסה

© Technion Israel Institute of Technology, Elyachar Central Library

המאולצת על ידו הוא מינימלי. באופן זה בצענו למעשה את הקוואנטיזציה, וסיבוכיות הקידוד היתה כשל חיפוש ממצה על המילון, דבר שבכל מקרה היינו נאלצים לעשות. מכיוון שמאמר זה התפרסם רק לאחרונה, ומכיוון שאינו נוגע ישירות למערכות SBC שסביבן נסבה העבודה, לא ניסינו שיטה זו, אולם היא בפרוש מועמדת מעניינת לקידוד דיבור בקצבי ביניים נמוכים.

3.4.3 מערכות עם בקרת הגבר

כמו בקוואנטיזציה סקלרית, גם קוואנטייזר וקטורי, המשמש לקידוד אותות לא סטאציונאריים, צריך להיות מלווה בסכמת אדפטיציה, שתעדכן אותו לאופיו המשתנה של אות הכניסה.

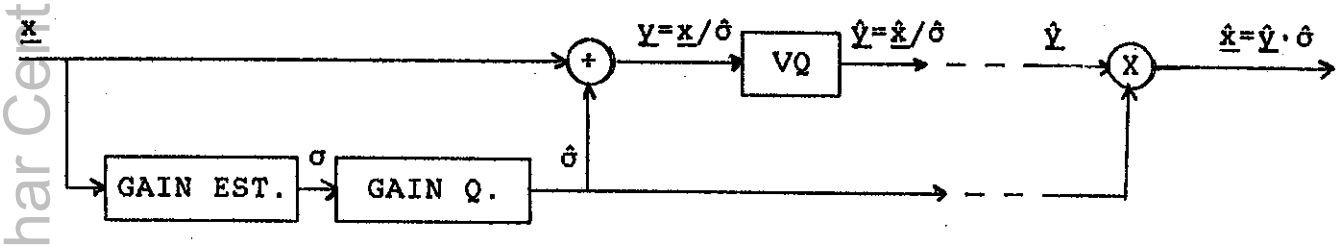
הבעיה המרכזית באדפטציה של קוואנטייזרים וקטוריים, היא כמובן הסיבוכיות העצומה של עדכון מילון שלם בזמן אמיתי. אלטרנטיבה נוחה היא, לפיכך, לבצע אדפטציה רק לגודל תאי הקוואנטיזציה (כפי שהדבר מקובל בקוואנטיזציה סקלרית יוניפורמית). בשיטה זו אנו מתעלמים מהאדפטציה שיש לבצע לצורתם של התאים, אולם מצד שני הסיבוכיות קטנה בהרבה.

הדרך המתבקשת לאדפטציה כזו היא בעזרת AGC (Adaptive Gain Controller), הפועל על וקטורי הכניסה. ע"י נרמול אות הכניסה באופן שוטף לפי הגברו נקבל במבוא הקוואנטייזר אות ברמת אנרגיה קבועה פחות או יותר, שנוח יותר לקודדו ע"י מילון בגודל נתון. סיכום של שיטות שונות לביצוע אדפטציה כזו ניתן למצוא ב-[21], שם מחולקות המערכות האדפטיביות, כמקובל, לפי אדפטציה אחורית וקדמית. בכל מקרה מוצגים גם מספר אלגוריתמים לשערוך ההגבר המשמש לנירמול וקטורי הכניסה.

מערכת עם אדפטצית הגבר קדמית מבצעת שערוך של הגבר וקטורי הכניסה לפי שיטה כלשהי, ולאחר מכן קוואנטיזציה של ההגבר, ונרמול הוקטורים לפי ההגבר המקוונטט. בנוסף לקוואנטיזצית הוקטורים, נשלחת כאינפורמצית צד גם קוואנטיזצית ההגבר. המקלט מבצע כמובן הכפלה של ההגבר בוקטור, לשחזור

© Technion-Israel Institute of Technology, Eyachar Central Library

המידע המקורי. מערכת כזו מוצגת בציור 3.6.



ציור 3.6 - מערכת קוואנטיזציה וקטורית עם בקרת הגבר קדמית.
Figure 3.6 - A vector quantization system with forward gain control.

לשערוך ההגבר, ניתן פשוט לחשב עבור כל וקטור את הנורמה שלו (למשל נורמה אוקלידית). אולם בשיטה זו, בשל התכיפות בשינוי ההגבר, הופכת אינפורמציית הצד למעמסה של ממש (אנו חוזרים למעשה לקוואנטייזר מטפוס הגבר/צורה שתואר בסעיף 3.3.1). לפיכך נוח לבצע אדפטציה בבלוקים בני n וקטורים, ולהגדיר את ההגבר כהגבר הממוצע של וקטורי הבלוק:

$$\hat{\sigma} = 1/n \sum_{i=1}^{n} ||x_i|| \tag{3.6}$$

עבור מערכת כזו, תכנון המילון לקידוד הוקטורים מתבצע לפי אלגוריתם LBG ובד"כ בשימוש בעוות רבועי. נוסחת הצנטרואידים כעת שונה במקצת, והיא:

$$\tilde{\underline{y}}_i = \mathrm{cent}(S_i) = \frac{E\{\hat{\sigma}^2 \underline{y} \mid \underline{y} \in S_i\}}{E\{\hat{\sigma}^2 \mid S_i\}} \tag{3.7}$$

הוכחת נוסחה זו היא פשוטה, אם נזכור כי המטרה היא לבצע מינימיזציה של העוות הכולל במערכת. בהנתן תאי הקוואנטיזציה, $\{S_i\}_{i=1}^{N}$, מלות הקוד האופטימליות יהיו אלה שישיגו, בכל תא S_i,

$$\min_{\tilde{\underline{y}}_i} E\{||\underline{x} - \hat{\sigma}\tilde{\underline{y}}_i||^2 \mid S_i\} \tag{3.8}$$

© Technion - Israel Institute of Technology, Elyachar Central Library

אם נזכור שמתקיים (לפי סימוני ציור 3.6), $\underline{x}=\hat{\sigma}\underline{y}$, ניתן לרשום את הבעיה כ-

$$\min_{\tilde{\underline{y}}_1} E\{\hat{\sigma}^2||\underline{y} - \tilde{\underline{y}}_1||^2 \mid S_1\} \qquad (3.9)$$

כתיבת הביטוי המפורש לנורמה, גזירה לפי אברי $\tilde{\underline{y}}_1$ והשוואה לאפס, נותנות מיד את (3.7).

לעומת מערכת עם אדפטציית הגבר קדמית, סובלת המערכת ה"אחורית" משתי בעיות מרכזיות. ראשית, יש לבצע כאן חיזוי של ההגבר, במקום שערוך שלו, ולכן יש לבחור אלגוריתם טוב, מבין הרבה אלגוריתמי חיזוי ידועים. שנית, תכנון המילון מסובך יותר, משום שהחזאי נחוץ לשם התכנון, אולם אנו צריכים לדעת את המילון בכדי לבצע חיזוי. מבחינת האיכות, שתי שיטות האדפטציה נותנות תוצאות קרובות זו לזו, כאשר מסתבר אמפירית כי השיטה הקדמית טובה יותר בקצבים נמוכים, והאחורית בגבוהים. מהשיקולים שפורטו נראה, כי השיטה הקדמית עדיפה על האחורית, ולפיכך נעשה בה שימוש לאדפטציית מערכת ה-SBC עם קוואנטיזציה וקטורית (ר' פרק 4).

3.4.4 קידוד וקטורי בתחום התדר

לסיכום הדיון בשימושיה של הקוואנטיזציה הוקטורית במערכות קידוד דיבור, יש לסקור כמובן מערכות קוואנטיזציה וקטורית בתחום התדר, ובפרט מערכות SBC עם קידוד וקטורי שכבר דווחו בספרות.

כאשר באים לשלב קוואנטיזציה וקטורית במערכות קידוד דיבור בתחום התדר, ישנה שאלת תכנון נוספת שאיננה קימת במערכות "סקלריות": כיצד לבחור את הוקטורים לקידוד. אפשרות אחת (שתקרא להלן "אנכית"), היא להגדיר וקטור בכולל את כל רכיבי התדר של האות בזמן מסוים. האפשרות השניה ("אפקית"), היא לקבץ לווקטור מספר דגמים עוקבים באותו תחום תדר. ברור, שקיימת גם האפשרות לקידוד מטריצי, בשני המישורים, שהיא כל הנראה היעילה ביותר, אולם היא בד"כ מסובכת מדי לממוש מעשי.

ניתן לאמר, כי המערכות ה"אנכיות" מנצלות את חוסר הקורלציה בין רכיביו של האות שעבר טרנספורמציה. לעומת זאת מערכות "אפקיות" יכולות בדרך כלל לנצל

© Technion - Israel Institute of Technology, Elyachar Central Library

יותר תכונות סוביקטיביות, הבאות לידי הדגשה בתחום ההתמרה, למשל ע"י הקצאה ותאימה של גדלי מילוני קידוד בפסים השונים. דיון מעמיק יותר במערכות יאנכיות" ו"אפקיות" יובא בפרק 4, שם יוצגו תוצאות נסויים שנערכו במערכות SBC משני הסוגים.

וספר המאמרים בספרות המוקדשים למערכות עם קידוד וקטורי ו-Transform Coding איננו רב. הפרסום היחיד עד כתיבת שורות אלה הדן במקודדי התמרה הינו [22]. על מערכות SBC עם קוואנטיזציה וקטורית נעשתה יותר עבודה, מקורות רלוונטיים הם [23], [24] ו-[25].

מאמר [23] מתרכזים הכותבים במערכת מהטפוס ה"אנכי" בלבד. המסקנה העקרית שלהם היא, כי עדיף לשלב קידוד וקטורי במערכת SBC מאשר לבצעו ישירות על גימות האות, אולם הם אינם מביאים השוואות בין קידוד סקלרי וקטורי במערכות SBC. במאמר המשך מציע אחד המחברים הוספת חזאי מטפוס LPC לפני ערך המסננים, לשפור איכות המערכת, (כלומר הקידוד הוא עתה של אות השארית לבד).

מאמר [24] מציג מערכת "אפקית" בעלת סיבוכיות גבוהה, המבצעת חלוקה לפי גבר/צורה גם במישור הזמן וגם במישור התדר, וכן כוללת הקצאה דינמית של מילונים (השקולה להקצאה דינמית של סיביות, במקרה הסקלרי). המסקנה העיקרית אליה מגיעים מחברי המאמר היא יתרון הקידוד הסקלרי על פני הוקטורי עבור ערך מסננים בגודל נתון, אולם עדיפות להגדלת מערך המסננים על פני הגדלת מימד הוקטור.

במאמר [25] מוצגת השוואה של שיטות קידוד סקלריות וקטוריות במערכות SBC, וכן פירוט של הפרמטרים במערכות אלה. למרבה הצער רב החסר על הנמצא בגוף המאמר, והפירוט איננו מושלם וממצה.

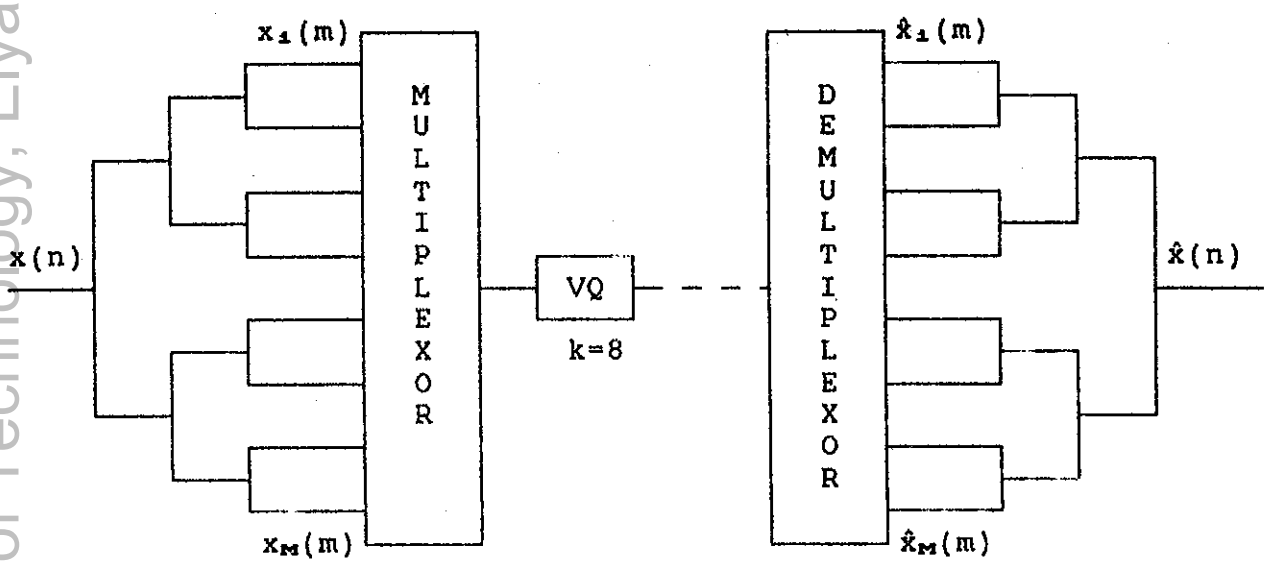
למעשה, כפי שכבר צוין, לא הופיעה עד היום סקירה ממצה של שיטות המשתמשות בקידוד וקטורי במערכות SBC, עם השוואה לקידוד סקלרי ופרוט נרחב של נסויים ומסקנות, וזהו בין השאר יעדה של עבודה זו. בפרק הבא יפורטו לכן בהרחבה הנסויים שנערכו במערכות SBC עם קידוד וקטורי, ואלו הפרק שאחריו יוקדש לתאור אפייני המערכות הסקלרית והווקטורית ולהשוואה ממצה שלהן.

© Technion - Israel Institute of Technology, Elyachar Central Library

פרק 4 - מערכת SBC עם קוואנטיזציה וקטורית

4.1 מערכת עם קוואנטיזציה "אנכית"

כפי שצוין, מערכת עם קוואנטיזציה "אנכית" מתאפיינת בכך שהוקטורים המקודדים בה מכילים רכיב אחד מכל פס-תדר של אות הכניסה. תאור סכמטי של מערכת כזו, עם שמונה פסי תדר, מופיע בציור 4.1:



ציור 4.1 - מערכת SBC עם קוואנטיזאציה וקטורית "אנכית".

Figure 4.1 - A SBC system with "vertical" vector quantization.

מערכת עם קוואנטיזציה "אנכית" מנצלת למעשה עקרון מקובל בקידוד מקורות בעלי זכרון: מכיוון שהידע על קידוד מקורות כאלה הוא מועט, ומכיוון שיש ידע רב לגבי קידוד מקורות חסרי זכרון, מבצעים בד"כ טרנספורמציה על המקור, לקבלת מקור חסר זכרון, ואז משתמשים בשיטות הקידוד, המקובלות עבור מקורות כאלה. מכיוון שהתמרה מורידת-תלות היא לרוב לא מעשית (פרט למקרה הגאוסי, בו הורדת תלות זהה להורדת קורלציה), מקובל להשתמש בהתמרות מורידות קורלציה: אופטימלית (KLT) או תת-אופטימליות (DCT,DFT).

© Technion - Israel Institute of Technology, Elyachar Central Library

הפרוק לרכיבי תדר, שהינו למעשה מעין התמרת פוריה גסה, משרת לפיכך גם הוא מטרה כזו: כל k דגמים עוקבים של אות הכניסה "מותמרים" ל-k דגמים, אחד בכל פס תדר של האות, המקודדים כוקטור. הקורלציה בין רכיבי וקטור זה קטנה בהרבה מהקורלציה בין דגמי מקור עוקבים. בחירה של הוקטור לקידוד כמורכב ממספר רכיבי תדר של אות הדיבור ברגע מסוים, היא גם בחירה "טבעית", שכן הוקטור לא נוצר ע"י קיבוץ של סקלרים, אלא מיצג ישות אחת כוללת (הספקטרום לזמן קצר).

החלוקה לשמונה פסי תדר, הכתיבה את מימד הוקטור המקודד, k, להיות שמונה (והמערכת היא לפיכך בדיוק כמוצג בציור 4.1). סדרת הלימוד לפיה תוכננו הקוואנטייזרים הוקטוריים הכילה כ-130,000 וקטורי לימוד, על בסיס של ששה משפטים אנגליים שנאמרו ע"י ששה דוברים שונים (שלושה גברים ושלוש נשים). הוכן מילון אופטימלי בן 256 וקטורים, המאפשר עבודה בקצב של 8Kbps, ולפיו נבנה מילון נוסף, באותו גודל, המהווה חוליה שניה במערכת שתי-דרגות (ר' סעיף 3.3.3), בעלת קצב של 16Kbps. הביצועים האוביקטיביים של המערכת בשני הקצבים נתונים בטבלה 4.1*.

SNRSEG (dB)	SNR (dB)	
15.30	14.91	8Kbps
21.18	21.34	16Kbps

טבלה 4.1: בצועי מערכת SBC עם קוואנטיזציה וקטורית "אנכית".
Table 4.1: Performance of a SBC system with "vertical" vector. quantization.

---

* - כאן ובהמשך, אלא אם צוין אחרת, נעשות המדידות על סדרת הלימוד שלפיה תוכננו המילונים.

© Technion - Israel Institute of Technology, Elyachar Central Library

כפי שניתן לראות, התוצאות האוביקטיביות הן בהחלט טובות: עבור קצב של 16Kbps, המערכת המוצגת משיגה תוצאות עדיפות על המערכת ה"סקלרית" המומלצת (ר' טבלה 2.4), וגם התוצאות ב-8Kbps הן טובות - בהשוואה למערכת ה"סקלרית" ולמערכות קידוד אחרות, המקובלות בקצב זה.

עם זאת, מבחינה סוביקטיבית היו התוצאות מאכזבות למדי. בקצב של 16Kbps קשה לדבר על שיפור ביחס למערכת ה"סקלרית" המומלצת (יש לציין שההשוואה איננה פשוטה, בשל האופי השונה של הרעשים הנלווים לאות בשני המקרים). ב-8Kbps הושגה איכות העולה על זו של המערכת ה"סקלרית" (שהיא גרועה מאוד), אולם לא במידה מספקת - הרעשים המתלווים לאות המשוחזר הם עדין בולטים ומפריעים.

זכיוון שכך, הוחלט לנסות ולעבור למערכת עם בקרת הגבר, מהטיפוס שתואר בסעיף 3.4.3, במטרה לשפר את האיכות (בעיקר בשמיעה). מערכת זו מתוארת בסעיף הבא.

4.1.1 הוספת בקרת הגבר

הוספת בקרת ההגבר למערכת נעשתה בדיוק בטכניקה המתוארת בציור 3.6. תחילה מתבצע חישוב של גורם הגבר, המשוערך בתור הנורמה הממוצעת של בלוק בן ששה עשר וקטורים (128 דגמים), לאחר מכן מקוונטט הגבר זה ע"י קוואנטייזר לוגריתמי מטיפוס μ-law בן ארבע סיביות (ר' [1] ע"מ 145). הוקטורים לקידוד מנורמלים לפי הגבר זה, ולאחר מכן מקודדים בעזרת מילון, שתוכנן כמתואר בסעיף 3.4.3. על הערוץ משודרת אינפורמציה גם על קוואנטיזצית הווקטורים וגם על קוואנטיזצית ההגבר. מכיוון שלהגבר יש להקדיש 250bps (4 סיביות כל 16msec), ניתן לנצל את שאר קיבול הערוץ לשידור האינפורמציה העיקרית.

המערכות שנוסו היו בקצבים 7.25Kbps (מילון בן 128 וקטורים) ו-15.25Kbps (מילון בן שתי דרגות, בנות 256 ו-128 וקטורים בהתאמה). קצבים אלה אמורים להתאים לקצבי שדור מעשיים של 8Kbps ו-16Kbps. התוצאות האוביקטיביות שהושגו נתונות בטבלה 4.2.

© Technion - Israel Institute of Technology, Elyashar Central Library

SNRSEG (dB)	SNR (dB)	
14.82	13.76	7.25Kbps
20.67	20.37	15.25Kbps

טבלה 4.2: ביצועי מערכת SBC "אנכית", עם בקרת הגבר בבלוקים של 16msec.

Table 4.2: Performance of a "vertical" SBC system with gain adaptation in 16msec blocks.

כפי שניתן לראות (מהשוואה לטבלה 4.1), הבצועים האוביקטיביים גרועים מאלה של המערכת חסרת בקרת ההגבר. ניתן להסביר זאת, חלקית, ע"י העובדה שהקצב בו משדרים נמוך יותר. הבצועים הסובייקטיביים הם בצפוי טובים יותר מאלה של המערכת הקודמת, אולם לא באופן משמעותי. נסיון להגדיל את חלקו של ההגבר במערכת, ע"י הקטנת גודל בלוק השערוך ל-4 וקטורים (4ms), לא שפרה בהרבה את בצועי המערכת. אמנם כעת מנוצל כל קיבול הערוץ (שכן יש להקדיש 1Kbps לשידור אינפורמצית הצד), אולם השפור ב-SNR הוא בכ-0.1dB ביחס לנתונים בטבלה 4.2, והשפור הסוביקטיבי כמעט ולא מורגש.

בסך הכל, שילוב של בקרת הגבר במערכת איננו מביא לפריצת דרך משמעותית. ככל הנראה, ההשתנות באופיו של הספקטרום לזמן קצר אינה מתרכזת בעיקרה בואריאנסים אלא דווקא במומנטים אחרים. (דבר זה הגיוני למדי, לאור ההשתנות של התכולה הספקטרלית של אות הדיבור במשך הזמן).

לפיכך הוחלט לזנוח את בקרת ההגבר, ולנסות לתקוף את המערכת ה"אנכית" בנקודה אחרת, שלה יש בד"כ משקל נכבד מבחינה סוביקטיבית (גם אם לא אוביקטיבית) - המדובר בהכנסה של גורם שקלול בתדר לפונקצית העוות לפיה נמדדים בצועי הקוואנטייזר. באופן זה מושג אפקט של תצורת רעש (Noise Shaping), המקובל במערכות קידוד דיבור רבות.

© Technion - Israel Institute of Technology, Elyachar Central Library

4.1. הוספת שקלול בתחום התדר

צורת רעש (Noise Shaping) הנה טכניקה מקובלת במערכות קידוד דיבור, מיועדת לשפר את איכותו הסוביקטיבית של האות המשוחזר. בשיטה זו נמנעים מתכוון מיצירת שגיאה לבנה, שהיא לכאורה סוג הרעש ה"רצוי" במערכת. במקום את "צובעים" את אות השגיאה, בהסתמך על כך שהאוזן איננה רגישה לרעש באופן חיד על פני תחום התדרים. בדרך כלל משתדלים לשוות לשגיאה ספקטרום דומה זה של אות הדיבור. עם זאת משתדלים שהדמיון יהיה רב יותר בתדרים שמעותיים, ומודגש פחות בתדרים אחרים, בשל אפקט של מיסוך הרעש ע"י האות, אשר זה האחרון בעל עצמה תדרית גבוהה.

צורת רעש מתבצעת במערכות בתחום הזמן ע"י הכנסה מלאכותית של מסנן צובע מערכת (ר' למשל [1], פרק 7). היא מקובלת גם במערכות בתחום התדר, ע"י קצאת סיביות המתבצעת לפי ואריאנסים משוקללים במקום לפי הואריאנס האמיתי ר' למשל [1], עמ' 532). בכל מקרה, האפקט המושג הוא של עליה בואריאנס רעש, אולם של ירידה סוביקטיבית ברעש בו האוזן מבחינה. הדרך המקובלת ליצג את, היא ע"י הצגת המערכת כמביאה למינימום לא את הספק השגיאה, אלא הספק ל שגיאה משוקללת בתדר.

תמקד כעת בדרך המקובלת לתצורת רעש בתחום התדר. כזכור (ר' משוואה (2.3)), עוות שאנו מעוניינים להביא למינימום בשיטת הקצאת הסיביות נתון ע"י:

$$D = 1/M \sum_{i=1}^{M} E\{(x_i-\hat{x}_i)^2\} = 1/M \sum_{i=1}^{M} \sigma_{ei}^2 \qquad (4.1)$$

הנחה המקובלת של קוואנטיזציה עדינה ויוניפורמית אנחנו מקרבים את σ_{ei}^2 ע"י (משוואה (2.4)):

$$\sigma_{ei}^2 = \varepsilon\, 2^{-2R_i}\, \sigma_i^2 \qquad (4.2)$$

© Technion - Israel Institute of Technology, Elyachar Central Library

ם נציב בבטוי זה את ההקצאה האופטימלית (משוואה (2.7)) נקבל:

$$\sigma_{ei}^2{}_{opt} = \varepsilon\ 2^{-2R} \left(\prod_{i=1}^{M} \sigma_i^2\right)^{1/M} \tag{4.3}$$

לומר ואריאנס השגיאה שווה בכל פסי התדר (או, במלים אחרות, השגיאה בנה).

די להשיג שגיאה "צבועה", כפי שאנו רוצים, נגדיר את D בצורה אחרת מאשר משוואה (4.1), ע"י:

$$D = 1/M \sum_{i=1}^{M} w_i \sigma_{ei}^2 \tag{4.4}$$

אשר w_i הם מקדמים משקלים. הקצאת הסיביות המתקבלת היא בדומה לזו שבנוסחה 2.7), אלא שהבטוי $w_i\sigma_i^2$ מחליף בה בכל מקום את הבטוי σ_i^2. במקום לקבל, מקודם, שגיאה לבנה, נקבל כעת שגיאה הצבועה לפי מקדמי השקלול: משוואה 4.3) מוחלפת כעת ב:

$$\sigma_{ei}^2{}_{opt} = 1/w_i\ \varepsilon\ 2^{-2R} \left(\prod_{i=1}^{M} \sigma_i^2\right)^{1/M} \tag{4.5}$$

ל מנת להשיג באותות דיבור את השקלול הרצוי (דמוי ספקטרום הדיבור עצמו) גדירים פשוט:

$$w_i = (\sigma_i^2)^{\tau} \tag{4.6}$$

אשר $\tau=0$ ו $\tau=-1$ הם המקרים הקיצוניים של שגיאה לבנה ושגיאה שספקטרומה ספקטרום הדיבור, בהתאמה. ערך מקובל (אמפירית) עבור τ הוא 0.3-.

די למצוא את הטכניקה הדרושה לשם תצורת רעש כזו במערכת SBC עם קוואנטיזציה וקטורית "אנכית", עלינו להזכר כי מילון הקוואנטיזציה הינו אופטימלי במובן של מינימום הגודל:

© Technion - Israel Institute of Technology, Elyachar Central Library

$$(4.7)\qquad D = E\{||\underline{x}-\hat{\underline{x}}||^2\} = E\{\sum_{i=1}^{M}(x_i-\hat{x}_i)^2\} = E\{\sum_{i=1}^{M} e_i^2\} = \sum_{i=1}^{M}\sigma_{e_i}^2$$

טוי זה הוא בדיוק כבטוי במשוואה (4.1) (עד כדי הגורם 1/M). מכאן, שנשיג מערכת זו את האפקט של תצורת רעש בדיוק באותו אופן עצמו כבבעית הקצאת סיביות: נגדיר את העוות כמו במשוואה (4.4) ע"י:

$$(4.8)\qquad D = E\{||\underline{w}(\underline{x}-\hat{\underline{x}})||^2\}$$

-$\underline{w}$ הוא וקטור שרכיבו ה-i הוא $(\sigma_i^2)^{\tau}$. (את σ_i^2 יש כמובן צורך לשערך, אותן טכניקות שהוצגו לשערוכו בסעיף 2.2.2).

ם לשים לב לעובדה, שמכיוון שקריטריון העוות שוב איננו רבועי, כי אם בועי-משוקלל*, הרי נוסחת חשוב הצנטרואידים ביצור המילון האופטימלי שתנה, כפי שהדבר קרה במערכת עם בקרת הגבר (ר' סעיף 3.4.3). הנוסחה כעת יא:

$$(4.9)\qquad \tilde{y}_{ij} = cent(S_i)_j = \frac{E\{w_j^2 x_j \mid \underline{x}\in S_i\}}{E\{w_j^2 \mid S_i\}}$$

להבדיל מנוסחה (3.7), שיצגה את הגדרת הצנטרואיד במערכת עם בקרת הגבר, יש חשב כאן ביטוי שונה לכל רכיב $\tilde{y}_{ij}$ של הצנטרואיד $\underline{\tilde{y}}_i$).

נסויים שנערכו נבחר הגודל τ להיות 0.3- כמוצע בספרות (ר' למשל [1] עמ' 56). שערוך הואריאנס התבצע לפי משערך ואריאנס-הדגם עם חלון רבועי בן K=1 דגמים (ר' נוסחה (2.9)). התוצאות האוביקטיביות עבור מערכות זו זונות בטבלה 4.3.

- למעשה השוני איננו נובע מעצם השקלול, אלא מהעובדה שזהו שקלול <u>משתנה בזמן</u>.

© Technion - Israel Institute of Technology, Elyachar Central Library

SNRSEG (dB)	SNR (dB)	
15.00	14.00	8Kbps
20.39	19.82	16Kbps

טבלה 4.3: בצועי מערכת SBC "אנכית" עם תצורת רעש.
Table 4.3: Performance of a "vertical" SBC system with noise shaping.

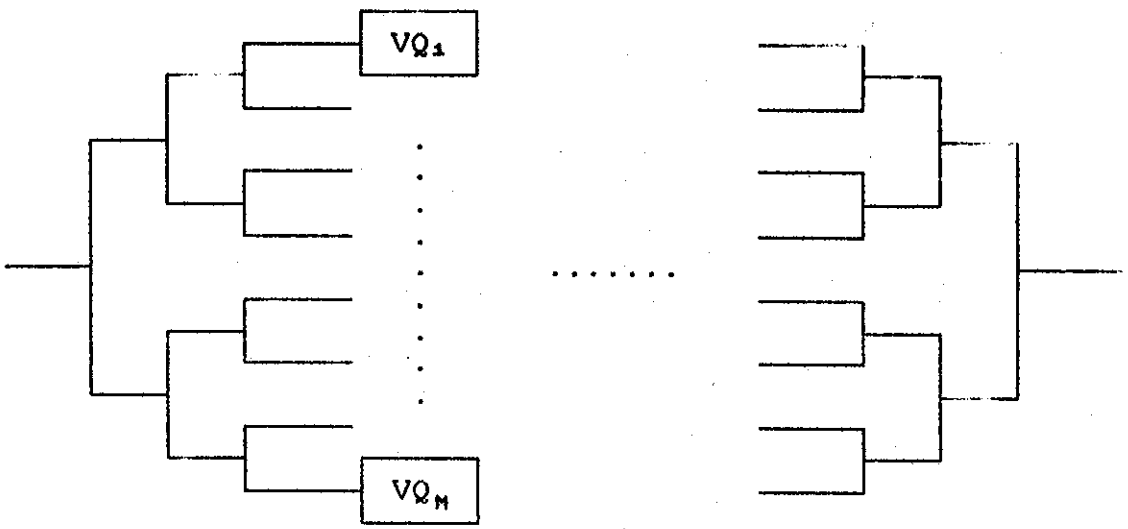
כפי שניתן לראות, התוצאות האוביקטיביות הן טובות למדי, למרות שיש לציין כי מידע זה הוא מטעה במקצת, שכן המטרה במערכת זו היא , כאמור, להביא למינימום לא את הספק הרעש אלא הספק-רעש-משוקלל. סוביקטיבית, ישנה התאמה לתוצאות האוביקטיביות: ב-16Kbps המערכת מניבה בצועים דומים לאלה של מערכת "אנכית" עם בקרת הגבר, ואילו ב-8Kbps הבצועים טובים יותר (אולם עדין יש אפקטים ברורים של עמעום תדרים גבוהים ושל צרידות). בסה"כ, למרות הצפיות הגבוהות, קשה לומר שתצורת הרעש תרמה לשפור משמעותי באיכות המערכת.

לסיכום ניתן לומר שהמערכת ה"אנכית" הניבה תוצאות מאכזבות למדי, למרות הצפיות (שהיו גם מבוססות מבחינה תאורטית). נפנה לכן לתאור של מערכת "אפקית", בה מקודד כל פס תדר בנפרד, ושם באמת הושגה "קפיצת דרך" - שפור ושמעותי על פני המערכת ה"סקלרית".

© Technion - Israel Institute of Technology, Elyachar Central Library

4.2 מערכת עם קוואנטיזציה "אפקית"

ערכת עם קוואנטיזציה "אפקית" הנה, כאמור, מערכת שבה מתבצע קידוד כל פס נדר בנפרד. בכל פס תדר יוצרים וקטורים לקידוד ע"י קיבוץ של מספר דגמים נוקבים, ומקודדים וקטורים אלה בעזרת מילון, המתוכנן במיוחד לפס התדר. נאור סכמטי של המערכת, במקרה של שמונה פסי תדר, מופיע בציור 4.2.



ציור 4.2: מערכת SBC עם קוואנטיזציה וקטורית "אפקית".
Figure 4.2: A SBC system with "horizontal" vector quantization.

ש לציין, כי קוואנטיזציה וקטורית "אפקית" מתאימה בדיוק למודל של מערכת SB כפי שהוצגה בסעיף 1.2: קידוד כל פס מתבצע בנפרד, והקשר בין הפסים ינו מפורש כי אם מובלע בתכנון המערכת (כמו למשל - הקצאת משאבים לפסים פי יחס האנרגיות ביניהם). תפישה זו שונה בברור מהקוואנטיזציה ה"אנכית", ם מושם דגש דווקא על קידוד במשותף של הפסים.

יתרון העקרי של שיטת הקוואנטיזציה ה"אפקית" היא בעובדה, שניתן לשלב בה ודיניות של הקצאת מילונים, השקולה לגישת הקצאת הסיביות במערכות ה"סקלריות". דבר זה לא היה קיים במערכת ה"אנכית", שם היה זה מתפקידו של מילון לדאוג לחלוקת משאבים אופטימלית בין הפסים. ע"י הכנסה של הקצאת מילונים, סטטית או דינמית, ניתן להשיג במערכת ה"אופקית" את כל היתרונות

© Technion - Israel Institute of Technology, Elyachar Central Library

ל המערכת ה"סקלרית", ובנוסף לכך גם את יתרונות הקוואנטיזציה הוקטורית.
סה"כ נראה כי, גם אם למערכת ה"אנכית" יש בסיס תאורטי חזק, אזי המערכת
"אפקית" מתאימה יותר לעקרונות הסוביקטיביים שביסוד מערכות SBC. אין
שכוח עם זאת, כי המערכת ה"אפקית" גם מסובכת יותר, בשל ריבוי המילונים.

כל המערכות ה"אפקיות" שנוסו, נבחר, מטעמי סיבוכיות, אורך וקטור של k=4.
כל המערכות נבחר גם עץ מסנני ה-QMF להיות מלא. במערכת הראשונית הוחלט על
קצאה סטטית של מילוני קידוד, וגדלי המילונים נבחרו (כמו במקרה הסקלרי)
פי ממוצעים לזמן ארוך של הקצאה (סקלרית) דינמית. במקרה הוקטורי, להבדיל
הסקלרי, ניתן לבחור גם הקצאות סיביות שבורות (בקפיצות של 1/k - ר' נוסחה
3.2)), ואכן כך נעשה הדבר. מספר הסיביות ($R_1,\ldots,R_8$) עבור קצבי פעולה של
16Kbp. ו-8Kbps מפורטים בטבלה 4.4. יחד אתם מפורטים גם גדלי המילונים
מתאימים (במקום שנדרש מילון עם יותר מ-256 וקטורים, הוא פוצל, מטעמי
יבוכיות, לשתי דרגות - אחת בת 256 וקטורים, והשניה משלימה לקצב הרצוי).

8Kbps			16Kbps			מספר הפס
גודל מילון 2	גודל מילון 1	R_1	גודל מילון 2	גודל מילון 1	R_1	
64	256	3.5	256	256	4.0	1
-	256	2.0	64	256	3.5	2
-	8	0.75	4	256	2.5	3
-	4	0.5	-	128	1.75	4
-	4	0.5	-	64	1.5	5
-	4	0.25	-	64	1.5	6
-	2	0.25	-	8	0.75	7
-	1	0	-	4	0.5	8

טבלה 4.4: פרמטרי מערכת SBC עם קוואנטיזציה וקטורית "אפקית".
Table 4.4: Parameters of a SBC system with "horizontal" vector quantization.

© Technion - Israel Institute of Technology, Elyachar Central Library

בצועים האוביקטיביים של המערכת נתונים בטבלה 4.5:

	SNR (dB)	SNRSEG (dB)
8Kbps	12.59	11.57
16Kbps	18.38	17.10

טבלה 4.5: בצועי מערכת SBC עם קוואנטיזציה וקטורית "אפקית".
Table 4.5: Performance of a SBC system with "horizontal" vector quantization.

פי שניתן לראות, התוצאות האוביקטיביות טובות פחות מאלה של המערכת "אנכית", וקרובות לאלה שהתקבלו במערכת ה"סקלרית". מענין לציין, כי כאן פעם הראשונה מתקבל SNRSEG נמוך מה-SNR. עם זאת, סוביקטיבית, מערכת זו ובה יותר מהמערכת ה"אנכית" על כל הואריאציות שלה (אם כי לא בצורה שמעותית). ביחס למערכת ה"סקלרית", הרי ב-16Kbps האיכות היא בערך כמו זו התקבלה בהקצאה סטטית של סיביות, אולם פחותה מזו של המערכת ה"סקלרית" המומלצת. שבה נעשה שימוש בהקצאת סיביות דינמית. ב-8Kbps מתקבל (כמו מערכת ה"אנכית") שפור על פני בצועי המערכת ה"סקלרית", אולם לא מתקבלת איכות דרך משמעותית.

ש לציין, כי למען השלמות, נבחנה גם מערכת "אופקית" ללא הקצאת סיביות כלל כלומר עם חלוקה שווה של המשאבים), כדי לראות עד כמה גורם זה הוא כבד שקל (כלומר - האם יכולה הקוואנטיזציה הוקטורית לפצות על העובדה, המשאבים מחולקים לכל הפסים בצורה שווה, ולא לפי יחסי האנרגיות). הסתבר, י במערכת ללא הקצאת סיביות מתקבלת ירידה של כ-7dB (!) בבצועים האוביקטיביים, ומורגשת דגרדציה חזקה גם מבחינה סוביקטיבית. הקצאת הסיביות הינה לפיכך תכונה חשובה של המערכת ה"אפקית", ואין לוותר עליה.

גם זאת, התוצאות שמערכת זו הניבה, עדין לא היו משביעות רצון. לפיכך הוחלט נסות ולשפר את המערכת ע"י שימוש בטכניקה של בקרת הגבר, באותו אופן שהדבר עשה במערכת ה"אנכית".

© Technion - Israel Institute of Technology, Elyachar Central Library

4.2.1 הוספת בקרת הגבר

ם במערכת זו, כבמערכת ה"אנכית", נעשה השימוש בטכניקה של בקרת הגבר בדיוק מתואר בציור 3.6, אלא שהפעם מישמים שיטה זו בכל פס בנפרד. ההגבר משוערך תור הנורמה הממוצעת של בלוק בן ארבעה וקטורים (כלומר 16 דגמים), מקוונטט וגריתמית כמו במערכת ה"אנכית", ומשמש לנרמול הוקטורים. הוקטורים עצמם קודדים בעזרת מילון, המתוכנן כמוסבר בסעיף 3.4.3. האינפורמציה המשודרת ל הערוץ נחלקת גם הפעם לאינפורמציה עקרית (הוקטורים) ואינפורמצית צד ההגבר). ס"ה יש להקדיש עבור שדור ההגבר 250bps (4 סיביות כל 16msec), בור כל פס אפקטיבי (כלומר פס עם הקצאה של יותר מאפס סיביות). בהתאם לכך ש לתכנן את חלוקת הסיביות הנותרות, עבור האינפורמציה העקרית בפסים.

טבלה 4.6 מוצגות הקצאות סיביות המתאימות למערכות עם קצבי שידור מעשיים ל 8Kbps, 9.6Kbps ו-16Kbps בהתאמה. הקצבים האמיתיים הינם 7.5Kbps, 9Kbps, 15.5Kbps, כאשר ה"עודף" מיועד למטרות הגנה מפני רעשי ערוץ. כזכור, על אינפורמציה העקרית, יש להוסיף גם 0.25Kbps למטרות שדור אינפורמצית צד כל פס אפקטיבי. בטבלה מפורטים גם גדלי מילונים, הבנויים כמו במערכת חסרת קרת ההגבר.

מספר הפס	7.5Kbps			9Kbps			15.5Kbps		
	R_1	גודל מילון 1	גודל מילון 2	R_1	גודל מילון 1	גודל מילון 2	R_1	גודל מילון 1	גודל מילון 2
1	3.5	256	64	3.5	256	64	4.0	256	256
2	1.75	128	-	2.0	256	-	3.5	256	64
3	0.75	8	-	0.75	8	-	2.5	256	4
4	0.5	4	-	0.5	4	-	2.0	256	-
5	0	1	-	0.5	4	-	1.0	16	-
6	0	1	-	0.25	2	-	1.0	16	-
7	0	1	-	0	1	-	0	1	-
8	0	1	-	0	1	-	0	1	-

ובלה 4.6: פרמטרי מערכת SBC עם קוואנטיזציה וקטורית "אפקית" ובקרת הגבר.
Table 4.6: Parameters of a SBC system with "horizontal" vector quantization and gain adaptation.

© Technion - Israel Institute of Technology, Elyachar Central Library

בצועים האוביקטיביים של המערכת נתונים בטבלה 4.7:

SNRSEG (dB)	SNR (dB)	
13.9	13.2	7.5Kbps
14.4	14.0	9Kbps
18.6	19.2	15.5Kbps

טבלה 4.7: בצועי מערכת SBC עם קוואנטיזציה וקטורית "אפקית" ובקרת הגבר.
Table 4.7: Performance of a SBC system with "horizontal" vector quantization and gain adaptation.

פי שניתן לראות, התוצאות האוביקטיביות טובות מאלה של המערכת ללא בקרת הגבר. אולם הבדלים אלה מתגמדים לעומת ההשג הסוביקטיבי: איכות הדיבור משוחזר טובה במערכת זו בצורה משמעותית יחסית לכל המערכות שנבחנו עד תה. ב-15.5Kbps מתקבלת איכות קרובה לאיכות מקור, וב-7.5Kbps מתקבלים מנם אפקטים של צרידות וגרגורים, אך האיכות עדיפה לאין ערוך על מה שהושג מערכות האחרות שנבדקו בקצב כזה. עם זאת נראה כי בתור קצב-נמוך ריאלי דאי לבחור קצב של 9Kbps (מעשית - 9.6Kbps) ולא 7.5Kbps (מעשית - 8Kbps). קצב זה מתקבלת איכות טובה למדי (למרות שגם כאן ניתן להבחין ברעשים ונים): מהשוואה בין שני הקצבים הנמוכים ניתן לראות שיש בפרוש חשיבות קידוד הפסים בתדר גבוה, אפילו במשאבים זעומים, מבחינת האוזן.

של התוצאות הטובות שהניבה מערכת זו, הוחלט לנסות ולבדוק את תוצאותיה על בצי בדיקה חיצוניים לסדרת הלימוד. קובץ בדיקה שהכיל משפטים שונים מפי ותם דוברים כבסדרת הלימוד, נתן ירידה של כ-2dB כתוצאות האוביקטיביות ולם באוזן הדבר לא הורגש. קובץ זר לחלוטין נתן ירידה של כ-1dB נוסף, עם לקול מסוים, אך לא גדול, באיכות. המערכת המתוארת כאן, אם כן, ראויה ללא פק להמלצה כמערכת SBC עם קוואנטיזציה וקטורית.

© Technion-Israel Institute of Technology, Elyachar Central Library

:סעיף הבא ידונו נסויים שהתמקדו בהוספת הקצאת משאבים דינמית למערכת זו,
:מטרה לשפרה עוד. הסעיף האחרון מוקדש לנסיון לשפור סיבוכיות המערכת, ע"י
:מוש במילוני עץ.

4.2.2 הקצאה דינמית של מילונים

:מערכת שתוארה בסעיף הקודם תואמת למעשה את המערכת שתוארה בסעיף 2.3
:מערכת SBC עם קוואנטיזציה סקלרית והקצאת סיביות סטטית) פרט לקוואנטיזציה
:היא וקטורית. איכות הדיבור שהיא מפיקה בקצב של כ-16Kbps טובה יותר מזו
:ל המערכת ה"סקלרית", אולם העליה בסיבוכיות (שתידון בפרק 5) לא מצדיקה
:כל הנראה את השפור באיכות (איכות המערכת ה"סקלרית" היא בהחלט טובה גם
:ן). עם זאת, בקצב של כ-9.6Kbps, המערכת ה"וקטורית" עדיפה בהרבה על
:"סקלרית" (וזו למעשה הסיבה להכנסת קוואנטיזציה וקטורית למערכת, כפי
:נצפו הדברים מראש).

:"י שלוב של הקצאה דינמית של מילונים (השקולה להקצאת סיביות דינמית
:מערכת ה"סקלרית"), נביא את המערכת לנקודת השוואה עם המערכת ה"סקלרית"
:ממולצת. אם אכן יושג שיפור בדרך זו, הרי השגתו אינה עולה בסיבוכיות רבה,
:כן הקצאת המילונים נעשית לפי וריאנס משוערך, שאפשר לחשבו ע"י ההגבה
:משוזר בכל מקרה בערוץ.

:ם זאת מציבה המערכת ה"דינמית" בעיה חמורה בשלב תכנון המילונים. הדרך
:אופטימלית לתכנן כל מילון איננה בעזרת סדרת הלימוד כולה, אלא רק בעזרת
:ותו חלק שלה, שלקידודו נוצר מילון זה. לשם כך יש להפריד את סדרת הלימוד
:מספר סדרות לימוד, ע"פ חשוב מוקדם של הקצאות המבוסס כאמור על נתוני הגבר
:כאשר אלגוריתם ההקצאה נבחר בדיוק כמו במערכת הסקלרית, מלבד העובדה
:רזולוצית החלוקה איננה בסיביות שלמות אלא ברבעי סיביות). אולם הפרדה זו
:ותנת סדרות למוד מאורכים לא מספקים, אפילו כאשר מגדילים מאוד את גודל
:דרת הלימוד ההתחלתית.

:פיכך הוחלט לנסות ולעבוד בשיטה תת-אופטימלית, לפיה משמשת כל סדרת הלימוד
:יצור כל המילונים במערכת. בשיטה זו לא קיימות הבעיות של מילונים לא
:יצגים, המבוססים על סדרות לימוד קטנות, וכן היא איננה מחייבת תכנון
:חדש של מילוני המערכת לכל קצב, כמו השיטה האופטימלית. מכיוון שנמצא כי

© Technion - Israel Institute of Technology, Elyachar Central Library

שוני בין הוקטורים המקודדים הינו בעקר בואריאנס (וזאת מהעובדה הפשוטה, הכנסת בקרת ההגבר למערכת משפרת אותה באופן משמעותי), הרי נראה כי לאחר רמול, איננו מאבדים הרבה מהאופטימליות אם נתכנן את המילונים בשיטה תת-אופטימלית הנ"ל.

מילונים תוכננו אם כן ע"פ סדרת הלמוד כולה, והמערכת זהה למערכת שהוצגה סעיף קודם, פרט לעובדה שהקצאת המילונים מתבצעת דינמית, יחד עם שערוך הגבר, לפי אלגוריתם הקצאת הסיביות שצוין. בנוסף לכך יש להניח כעת שכל פסים אפקטיביים (כלומר - אין פס שכל הזמן מקודד באפס סיביות). עבור קצב ל 9Kbps, שהוא המועמד העקרי לשפור, התקבלו SNR=15.39dB ו-SNRSEG=15.82dB הם שפור של כ-1.5dB ע"פ המערכת ה"סטטית". עם זאת, סוביקטיבית, לא ניתן הצביע בברור על שפור משמעותי באיכות.

רעיון של הקצאת מילונים דינמית ננטש, אם כן, והמערכת המומלצת עבור וואנטיזציה וקטורית נשארת כמתואר בסעיף 4.2.1. עם זאת, לפני שנעבור תאור מדויק של מאפיני המערכת (עמידות לרעשים, סיבוכיות וכד') שיפורט פרק 5, יש לציין כי השיפור באיכות במערכת ה"וקטורית" לעומת המערכת "סקלרית" הינו יקר מבחינת סיבוכיות. לאור זאת הוחלט לבדוק את הדגרדציה איכות המערכת אם תמומש ע"י מילונים במבנה עץ (ר' סעיף 3.3.2), החוסכים אוד בסיבוכיות הזמן (שהיא ככל הנראה המשאב הקריטי במערכות כאלה). נסויים לה יפורטו בסעיף הבא.

4.2. מערכת עם מילונים במבנה עץ

פי שצויין בסעיף 3.3.2, המעבר למילונים במבנה עץ מקטין לוגריתמית את יבוכיות הזמן של הקידוד, תמורת תשלום במקום (דרישות המקום מוכפלות) באיכות (דגרדציה, קטנה בד"כ, הנובעת מהאילוץ על מבנה המילון).

כנון המילונים במבנה עץ נעשה בדיוק כמפורט בסעיף 3.3.2, פרט לעובדה שולבה במערכת בקרת הגבר, כך שחישוב הצנטרואידים היה כנתון בנוסחה (3.7). שינוי היחיד במערכת ה"וקטורית" המומלצת (סעיף 4.2.1) היה אם-כן החלפה של מקודד במקודד מטפוס עץ, הנעזר בעץ החיפוש שתוכנן עבורו כדי למצוא מהירות את מילת הקוד הרצויה. התוצאות האוביקטיביות של המערכת

© Technion - Israel Institute of Technology, Elyachar Central Library

מוצגות בטבלה 4.8:

SNRSEG (dB)	SNR (dB)	
13.02	12.76	9Kbps
16.79	17.63	15.5Kbps

טבלה 4.8: בצועי מערכת SBC עם קוואנטיזציה וקטורית "אפקית" במבנה עץ ובקרת הגבר.

Table 4.8: Performance of a SBC system with tree-structured "horizontal" vector quantization and gain-adaptation.

כפי שניתן לראות (מהשוואה לטבלה 4.7) יש ירידה של כ- 1÷2dB בבצועים האוביקטיביים. עם זאת, מבחינה סוביקטיבית, הדגרדציה קטנה מאוד (כמעט לא מורגשת). מערכת זו נבדקה גם על קבצים חיצוניים לסדרת הלמוד, וכמו במערכת עם מילונים אופטימליים, התקבלה ירידה אוביקטיבית של כ- 1÷2dB (נוספים) אך כמעט לא הורגשה דגרדציה סוביקטיבית.

המערכת עם מילונים במבנה עץ היא, אם כן, מועמדת ריאלית למימוש, בשל הירידה בסיבוכיות הקידוד. עם זאת אין לשכוח, כי דרישות המקום מוכפלות, ולכן כשנגשים ליישם את המערכת יש לוודא היטב מהם שיקולי הפשרה בין זמן למקום. חישוב מדויק יותר ומפורט של הסיבוכיות עבור המערכות המומלצות יוצג בפרק הבא, יחד עם מאפיינים נוספים שלהן (כמו עמידות לרעשים, עבודה עם data וכד').

© Technion - Israel Institute of Technology, Elyachar Central Library

פרק 5 - מאפייני המערכות המומלצות

5.1 המערכת המומלצת עם קוואנטיזר סקלרי

לפני שניגש לדון במאפייני המערכת ה"סקלרית" המומלצת, המיועדת לעבודה בקצב של 16Kbps, נזכיר את מבנה המערכת והפרמטרים שלה (ר' גם סעיף 2.4.2).

הסינון במערכת מתבצע ע"י מערך QMF מלא, הבנוי במבנה עץ בן שלוש רמות (כלומר שמונה פסי תדר ברוחב 500Hz כ"א). הנסויים שנערכו במערכת זו השתמשו במסננים בני 64 מקדמים בכל רמות העץ, כדי להפחית את השפעת השגיאות בסנון ולהתרכז בבעית הקידוד, אולם דבר זה איננו הכרחי כמובן.

הקידוד בכל פס הוא בעזרת מקודד סקלרי, יוניפורמי, אדפטיבי וללא חזאי. הקידוד מתבצע ב-5÷0 סיביות לדגם, כאשר קידוד בסיבית אחת הוא אדפטיבי בבלוקים, ומבוסס על הנחת פילוג ז בפס (נוסחה (2.15)) והקידוד ב-5÷2 סיביות הוא אדפטיבי דגעית, ומבוסס על שיטת הכופלים של Jayant (נוסחה (2.13)). האדפטציה של הקידוד בשני המקרים היא אחורית.

הקצאת הסיביות למקודדים במערכת היא דינמית, בבלוקים בני 32 דגמים, ומתבצעת לפי אלגוריתם הקצאה תת-אופטימלי: ראשית מחושב וריאנס הדגם של הפס (נוסחה (2.9)), לאחר מכן מחושבת הקצאה שבורה אופטימלית (נוסחה (2.7)), ואז מתבצעת הפיכת ההקצאה לשלמה, ע"י קיצוץ, הגבלה וחלוקת עודפים בשיטת ה"פוקר", מתדר נמוך לגבוה. אדפטצית ההקצאה היא קדמית, עם שידור ערכי ההקצאה כאינפורמצית צד.

המערכת מיועדת לעבודה בקצב כולל של 15.75Kbps, כאשר 15Kbps מיועדים לאינפורמציה המרכזית: תדר העבודה בכל פס הוא 1KHz (לאחר הדצימציה), לכן לקבוצת דגמים המופיעים במקביל, אחד בכל פס, יש להקצות בס"ה 15 סיביות. 750bps מיועדים לשידור אינפורמצית צד, המכילה 3 סיביות עבור ערך ההקצאה בכל פס, פעם בבלוק אדפטציה (בן 32msec). 250 ה-bps הנותרים מיועדים לקוד גלוי ותקון שגיאות, שיגן על אינפורמצית הצד משגיאות ערוץ (ר' גם סעיף 5.1.2).

© Technion-Israel Institute of Technology, Elyachar Central Library

תוצאות האוביקטיביות המתקבלות במערכת זו הן: SNR=18.3dB, SNRSEG=19.4dB.
ובאיקטיבית מתקבל דיבור באיכות טובה מאוד, עם מעט רעשי רקע עדינים (הדהוד
על גוון של צלצול).

פנה כעת לתאור מספר אפיונים מעשים של המערכת המוצעת, שיבהירו כמה פרטים
ושובים, הנוגעים לשילובה במערכות תקשורת דיבור ספרתית.

5.1. סיבוכיות המימוש

חד הפרטים המעניינים לגבי מימוש של מערכת כגון זו, הינו חשוב סיבוכיות,
מהווה אינדיקציה למעשיות המימוש. מקובל לחלק את חישוב הסיבוכיות לפי
יבוכיות זמן וסיבוכיות מקום, וכן לבצע הפרדה של המערכת לפי משדר ומקלט.
מקרה של מערכת SBC, נוח לבצע הפרדה נוספת, לפי סיבוכיות הסינון
סיבוכיות הקידוד. סיבוכיות הסינון מופיעה במשדר (אנליזה) ובמקלט
סינתיזה), ואילו סיבוכיות הקידוד/פענוח שונה כמובן בין המשדר והמקלט.

חשב כעת את סיבוכיות הזמן של בצוע הסינון (אנליזה) במערכת QMF בסיסית,
מופיעה בציור (1.2) והמשתמשת במסננים בני L מקדמים. לאחר מכן נשתמש
תוצאות למציאת הסיבוכיות עבור עץ QMF מלא בן M פסי תדר, שברמתו ה-i-ית
וסננים עם L_i מקדמים (נסמן ע"י $m=\log_2 M$ את מספר הרמות בעץ).

סמן ב-x_0 ו-x_1 את מוצאי הדצימטורים בערוץ העליון והתחתון של ציור (1.2),
התאמה. ניתן לרשום:

$$x_k(l) = \sum_{n=-\infty}^{\infty} (-1)^{kn} h(n) x(2l-n) \qquad k=0,1 \tag{5.1}$$

אם נבצע כעת את חילוף המשתנים $n=2r+\beta$ נקבל:

$$x_k(l) = \sum_{\beta=0}^{1} \sum_{r=-\infty}^{\infty} (-1)^{k(2r+\beta)} h(2r+\beta) x(2(l-r)-\beta)) \qquad k=0,1 \tag{5.2}$$

© Technion - Israel Institute of Technology, Elyachar Central Library

גדיר כעת מסננים ואותות ע"י:

$$p_m(l) = h(2l+\beta) \qquad \beta=0,1 \tag{5.3}$$
$$y_m(l) = x(2l-\beta)$$

נוסחה (5.2) הופכת ל-

$$x_k(l) = p_0(l) * y_0(l) + (-1)^k p_1(l) * y_1(l) \qquad k=0, \tag{5.4}$$

ז נשים לב לכך, ש-$p_m(l)$ (הקרוי גם "מסנן פוליפאזי") הוא בעל L/2 מקדמים, קבל מיד שעבור שני דגמי כניסה למערכת, אנו צריכים לבצע L/2 כפלים ו-L/2 חיבורים לקבלת המחובר הראשון ב-(5.4), וכמספר הזה כפלים וחיבורים קבלת המחובר השני. עוד חיבור וחיסור יתנו לנו דגם אחד מ-$x_0(m)$ ודגם אחד $x_1(m)$- בהתאמה. בס"ה בצענו L כפלים ו-L חיבורים לזוג דגמי כניסה, או L/2 פלים ו-L/2 חיבורים לדגם כניסה.

עבור כעת ממערכת ה-QMF הבסיסית למערך QMF מלא במבנה עץ בעל m רמות $M=2^m$- פסי תדר במוצא, שברמתו ה-i-ית משתמשים במסננים בני L_i מקדמים.

בור הרמה הראשונה של העץ, יש לבצע $L_1/2$ חיבורים ו-$L_1/2$ כפלים לכל דגם ניסה, לפי החישוב שעשינו עבור מערכת ה-QMF הבסיסית. ברמה השניה, לפי ותו הפיתוח, יש לבצע $L_2/2$ חיבורים ו-$L_2/2$ כפלים בכל זוג פסי-תדר, וזאת גם בדגם הנכנס לזוג זה. מכיוון שכעת יש לנו שני זוגות, יש להכפיל את סיבוכיות. אולם כל דגם כניסה לזוג הפסים מופיע אחת לשני דגמי מקור, בשל וצימציה, ולכן הסיבוכיות נשארת כשהיתה. באותו אופן נקבל, שהסיבוכיות כוללת עבור הרמה ה-i-ית היא של $L_i/2$ כפלים ו-$L_i/2$ חיבורים לדגם, סיבוכיות עבור הסינון כולו היא כמובן של $\sum_{i=1}^{m} L_i/2$ כפלים, ואותו מספר יבורים, לדגם.

יבוכיות המקום של הסינון נובעת מהצורך להחזיק את כל מקדמי המסננים. מתוך שימוש במבנה הפוליפאזי, מונקבל כי יש להחזיק L_i מקדמים ברמה ה-i-ית, ולכן סה"כ $\sum_{i=1}^{m} L_i$ מקדמים (שהם מספרים בנקודה צפה).

© Technion - Israel Institute of Technology, Elyachar Central Library

סיבוכיות הזמן והמקום של הסינתיזה זהה לזו שחושבה עבור האנליזה, משיקולי סימטריה. לפיכך יש להביא בחשבון חישובים אלה של סיבוכיות גם עבור המשדר וגם עבור המקלט.

חישוב סיבוכיות הזמן עבור הקידוד (והפיענוח) הוא הרבה פחות נוח מאשר החישוב עבור הסינון, וזאת מכיוון שתהליך הקידוד הוא אדפטיבי. לבצוע החישוב נסמן ב-$N_i(l)$ את מספר הרמות החיוביות בקוואנטייזר של הפס ה-i-י, עבור דגם הכניסה ה-l-י שלו. $N_i(l)=2^{R_i(l)-1}$, כאשר $R_i(l)$ היא הקצאת הסיביות לפס ה-i-י באותו הרגע (סיבית אחת מוקדשת לסימן).

ביצוע הקוואנטיזציה עצמה כרוך, עבור כל דגם בפס, בפעולות הבאות: חישוב זימן, חלוקה בגודל צעד הקוואנטיזציה ו-$N_i(l)-1$ השוואות. אדפטצית הקוואנטייזר, עבור המקרה $R_i(l)>1$, כרוכה בכפל אחד לדגם, ואילו במקרה $R_i(l)=1$ יש לבצע כפל והוצאת שורש פעם בבלוק ("קבוע-וריאנס") בן K דגמים. זכיוון שמספר דגמי היציאה במערכת הסינון שווה למספר דגמי הכניסה, הרי אלו גם הפעולות המתבצעות לדגם מקור.

בנוסף לכך יש לבצע הקצאת סיביות פעם בבלוק בן KM דגמים. (המכיל K דגמים לכל פס-תדר). לשם כך יש לבצע, עבור כל דגם, כפל וחיבור (שערוך הוריאנס), וכן, פעם בבלוק, חישוב של ההקצאה, כמפורט בסעיף 2.4.2.

בפענוח, עבור כל דגם, מבצעים הכפלה בגודל צעד הקוואנטיזציה, וכן אדפטציה של הקוואנטייזר. עבור המקרה $R_i(l)>1$ יש לבצע כפל אחד לדגם לשם האדפטציה, ואילו עבור המקרה $R_i(l)=1$ יש לבצע כפל וחיבור לכל דגם (שערוך הוריאנס), וכן כפל והוצאת שורש פעם בבלוק בן K דגמים.

בנאשר לסיבוכיות המקום, הרי שמספיק להחזיק (במקודד ובמפענח) את הכופלים הדרושים לאדפטצית הקוואנטייזרים עבור המקרה $R_i(l)>1$. למעשה, מספיק להחזיק, עבור כל ערך אפשרי של $R_i(l)$, זוג מספרים, מהם ניתן לשחזר את הכופלים.

כדי לקבל תחושה מעשית לגבי סיבוכיות הזמן (שהיא בד"כ בעייתית יותר מסיבוכיות המקום), נציין כי עבור המערכת ששמשה בסימולציות, ועבור הקצאת הסיביות הממוצעת, התקבלו התוצאות הבאות: 96 פעולות לדגם לבצוע הסנון, כ-8

© Technion - Israel Institute of Technology, Elyachar Central Library

עולות לדגם לבצוע הקידוד, וכ-2 פעולות לדגם לשם הפענוח (כדי להמנע סרבול נניח שכל הפעולות מתבצעות בזמן שווה).

פי שניתן לראות, הסיבוכיות של הקידוד והפענוח היא זניחה לעומת סיבוכיות סינון. מכיוון שכך, יש צורך לחפש פתרונות, ע"מ להקל על עומס הסינון. תרון אפשרי אחד, הבא על חשבון ייקור המערכת, הוא להקדיש מעבד נפרד לבצוע סינון בכל פס (עבוד במקביל). באופן זה קטן משך ביצוע הסינון פי M (למשל קבל 12 פעולות לדגם ולמעבד במערכת הדוגמה). אפשרות אחרת, הבאה על חשבון איכות, היא להקטין את ארכי המסננים. פתרון זה אינו מוסיף רעש למערכת, ולם גורם לתופעה של התחזות (שכן ביטול ההתחזות במערכות QMF הוא, כאמור, ד כדי הקוואנטיזציה). לדוגמא - הורדת מספר המקדמים מ-64 ל-8 במערכת רוגמה תקטין את הסיבוכיות ל-12 פעולות לדגם.

זך הכל, אם משתמשים במעבדי אותות הקיימים כיום בשוק, והמסוגלים לבצע -50 פעולות לדגם, נראה בהחלט כי אין מניעה לממש את המערכת ה"סקלרית" זביבה מעשית, בעזרת ההמלצות דלעיל.

5.1. עמידות לרעשים

בונה נוספת של המערכת, החשובה עבור הממוש המעשי היא עמידותה לרעשים. ד"כ מחלקים נושא זה לשניים: עמידות לרעשי ערוץ ספרתיים (כלומר - להיפוך כן של סיביות, הקורה בהסתברות שגיאה מסוימת), ועמידות לרעשי רקע נלוגיים (בד"כ רעש לבן גאוסי המצטרף לאות המקור - נושא זה חשוב יותר זקודדי מקור, ולכן יטופל כאן בקצרה).

פני שנפנה לבדיקת עמידות המערכת לרעש ערוץ, נגדיר את האופן בו משודרת אינפורמציה על הערוץ במערכת זו. בכל 32msec משודרת מסגרת בת 512 סיביות, מכילה אינפורמציה מרכזית (קידוד של דגמי דיבור), אינפורמצית צד (הקצאת סיביות) וסיביות הגנה, כאשר החלוקה היא כדלקמן: 480 סיביות עבור 32 מיניות "אנכיות" של דגמים (כל שמיניה מקודדת ב-15 סיביות), 24 סיביות בור ההקצאה (3 סיביות לכל אחד משמונת הפסים) ו-8 סיביות הגנה.

© Technion - Israel Institute of Technology, Elyachar Central Library

:תחילה נבדקו ביצועי המערכת בלי שמוש בסיביות ההגנה. נוסו הסתברויות
'שגיאה (BER או Bit Error Rate) בערכים 10^{-3} ו-10^{-4}, והתקבלו תוצאות
:וביקטיביות כמפורט בטבלה 5.1 (עם השוואה לביצועים ללא שגיאות).

SNRSEG (dB)	SNR (dB)	BER
19.40	18.30	0
6.01	-5.36	10^{-3}
14.55	10.82	10^{-4}

טבלה 5.1: בצועי המערכת ה"סקלרית" תחת השפעת שגיאות ערוץ.
Table 5.1: "Scalar" system performance with channel errors.

פי שניתן לראות, השפעת שגיאות הערוץ ב-$BER=10^{-3}$ (נחשב למעשי) היא
טסטרופאלית, וגם ב-$BER=10^{-4}$ מתקבלת איכות נמוכה. תוצאות אלה מתאימות
'תוצאות הסוביקטיביות: ב-$BER=10^{-3}$ האיכות גרועה מאוד ונשמעים צפצופים
הפרעות שונות; ב-$BER=10^{-4}$ האיכות טובה בהרבה, למרות שעדין יש "קליקים"
א נעימים לאוזן. משום כך נעשו הבדיקות שלאחר מכן עם שימוש בסיביות
:הגנה, המיועדות לשפור האיכות בתנאים של שגיאות ערוץ.

יביות הגנה, או, למעשה, קודים לגילוי ותיקון שגיאות, הן אחת מהשיטות
מקובלות כדי להלחם בהשפעת שגיאות ערוץ במערכות תקשורת ספרתית. קיימת
פרות ענפה בנושא זה, לדוגמא [26], ממנו הוצא הרקע ששמש לפרק זה.

וד (ליניארי) לתקון וגלוי שגיאות מאופיין ע"י השלישיה (n,k,d), כאשר n
:וא אורך מילת קוד (בסיביות), k היא כמות האינפורמציה במילה (בסיביות)
כלומר n-k סיביות הן למטרות הגנה) ו-d הוא מרחק האמינג המינימלי בין שתי
ילות קוד, כך ש-$t=[1/2(d-1)]$ הוא מספר השגיאות שהקוד מסוגל לתקן.

:קצב של הקוד מוגדר ע"י $R=k/n$. לפי משפט של Shannon, כל זמן ש-$R<C(BER)$
:נו יכולים להקטין את הסתברות השגיאה כרצוננו, עם הגדלת n (בתשלום
יבוכיות והשהיה), כאשר C(BER) הוא קיבול הערוץ הבינארי הסימטרי עם

© Technion - Israel Institute of Technology, Elyachar Central Library

סתברות BER לשגיאה, ונתון ע"י הביטוי:

$$C(BER) = 1 + BER \log_2 BER + (1-BER) \log_2(1-BER) \tag{5.5}$$

ובן, שבמערכות מעשיות עלינו לעבוד בתנאי סיבוכיות והשהיה סבירים, ולכן סתברות השגיאה איננה קטנה כרצוננו, אבל בד"כ זניחה. בסוף [26] נתונה בלה של קודים מעשיים לגילוי ותקון שגיאות ותכונותיהם.

מקובל במערכות מסוג זה, נבחר להגן על אינפורמצית הצד בלבד, שכן פענוח כון שלה הוא קריטי לתפקוד המערכת. לפיכך עלינו לבחור קוד שייקיים $n=32$, $k=2$, ומסתבר (ע"פ הטבלה בסוף [26]) כי קוד כזה הוא לכל היותר בעל $d=3$, לומר $t=1$.

ישוב של "BER אפקטיבי", שיציין את ההסתברות המעשית לשגיאה בחלק המוגן, וא לרוב בלתי אפשרי. כיוון שכך, משווים Block Error Rates, הקלים יותר חישוב, תוך הנחה שמספיקה שגיאה אחת שלא ניתן לתקנה בבלוק (=מילת קוד), די להרסו לגמרי. ההסתברות לשגיאה בבלוק בגודל n, עם קוד המסוגל לתקן t גיאות היא כמובן:

$$BLER = 1 - \sum_{i=0}^{t} \binom{n}{i} BER^i (1-BER)^{n-i} \tag{5.6}$$

רק זה יש להשוות להסתברות השגיאה ללא הגנה (בבלוק באורך k), שהיא מובן:

$$BLER^* = 1 - (1-BER)^k \tag{5.7}$$

© Technion - Israel Institute of Technology, Elyachar Central Library

כור המערכת הסקלרית אנו מקבלים, בערכי ה-BER שנוסו, את BLER ו-BLER*
זפורטים בטבלה 5.2.

BLER	BLER*	BER
$5x10^{-4}$	$2.4x10^{-2}$	10^{-3}
$5x10^{-6}$	$2.4x10^{-3}$	10^{-4}

בלה 5.2: ערכי הסתברות לשגיאה בבלוק, עבור הסתברות נתונה לשגיאה בסיבית,
עם ובלי הגנה.

Table 5.2: BLER values for given BER, with and without protection.

פי שניתן לראות, ההסתברות לשגיאה בבלוק עם הגנה היא זניחה לעומת
הסתברות ללא הגנה, וזאת למרות העובדה, שניתן לתקן רק שגיאה אחת. לדוגמא,
כור $BER=10^{-3}$ מתקבלת שגיאת בלוק פעם ב-1.33 sec ללא הגנה, ופעם ב-640
se עם הגנה (לשם השוואה, משכי המשפטים בסימולציות היו כ- 10-20sec).
איכך, בסימולציות, "מומשה" ההגנה פשוט ע"י אי הכנסת שגיאות לאינפורמציה
נד, כלומר שגיאות הוכנסו למידע העיקרי בלבד. התוצאות שהתקבלו מפורטות
זבלה 5.3:

SNRSEG (dB)	SNR (dB)	BER
9.53	2.58	10^{-3}
18.33	15.08	10^{-4}

טבלה 5.3: ביצועי המערכת ה"סקלרית" עם השפעת שגיאות ערוץ והגנה.

Table 5.3: "Scalar" system performance with channel errors and protection.

© Technion - Israel Institute of Technology, Elyachar Central Library

פי שניתן לראות, התוצאות משתפרות ללא הכר. סוביקטיבית, התקבלה
-BER=10-³ איכות טובה למדי (עם הפרעות קלות), וב-BER=10-⁴ לא מרגישים
ל בשגיאות. ניתן לומר, לפיכך, שהמערכת (הכוללת הגנה) בנויה לעבודה
-BER של 10-⁴.

זיכום נושא זה יש לציין, כי הבעייתיות בתגובת המערכת ה"סקלרית" המוצעת
שגיאות ערוץ, נובעת מתוך העובדה, שאדפטצית הקוואנטייזרים היא אחורית,
לכן יש התפשטות של שגיאות. כדי לטפל בכך, יש להכניס למערכת גורמי דעיכה,
ו לחילופין לעבור למערכת עם אדפטציה קדמית. שתי הצעות אלה חורגות מתחום
בודה זו.

ציין עוד בקצרה בדיקות שנעשו בענין השפעת רעש רקע על בצועי המערכת.
כניסת המערכת הוסף רעש רקע גאוסי לבן ב-SNR-ים של 25dB (חלש), 15dB
בינוני), ו-5dB (חזק). האפקט שהתקבל היה צביעה של הרעש, ומכיוון שהרעש
וסך לחלוטין את פעולת מערכת הקידוד, התקבל ביציאה רושם של צירוף הרעש
צבוע לדיבור הנכנס. לא חל שינוי סוביקטיבי בעצמת הרעש (למרות
אובייקטיבית היה שינוי קטן בערכי ה-SNR). לא חל גם שינוי במובנות
דיבור, והרעש הצבוע הוא בברור פחות נעים לאוזן מהרעש המקורי.

5.1. Tandeming

ות דיבור העובר במערכת תקשורת, בה נעשה שמוש במקודד דיבור ספרתי כמו זה
לפנינו, יכול לעבור בדרכו מספר פעמים דרך מקודדים מאותו סוג, בתלות
מסלול שלו מהמקור ליעד. משום כך מענין לדעת, מה ההשפעה של מעבר אות
יבור דרך מספר מערכות מהטיפוס המוצג, המחוברות בטור (או, כפי שמקובל
ומר, ב-tandem). אנו מעוניינים כמובן שהירידה באיכות תהיה מתונה ככל
אפשר.

בדיקות שנעשו מוצגות בטבלה 5.4. עבור כל רמת Tandeming (כלומר, מספר
מערכות המחוברות בטור), נמדדה הדגרדציה האוביקטיבית ביחס לאות המקורי.

© Technion - Israel Institute of Technology, Elyachar Central Library

SNRSEG (dB)	SNR (dB)	רמת Tandeming
19.38	18.26	1
15.05	15.37	2
14.04	14.17	3
13.48	13.28	4
12.82	12.35	5

טבלה 5.4: הדגרדציה המתקבלת באות העובר מספר מערכות "סקלריות" ב-tandeming.

Table 5.4: Degradation in a signal passing through several tandemed "scalar" systems.

פי שניתן לראות, הדגרדציה היא אכן הדרגתית, כאשר השינוי הוא החריף ביותר זעבר ממערכת אחת לשתיים, ומתמתן עם חיבור מערכות נוספות. גם באופן וביקטיבי התקבלו תוצאות דומות, כאשר האפקט של "צלצולים" ברקע מתגבר זהדהד יותר ככל שמחברים מערכות נוספות ל"שרשרת". עם זאת, הירידה החריפה זתקבלת אוביקטיבית בהגדלת מספר המערכות מאחת לשניים, איננה מורגשת וביקטיבית, ולמעשה כמעט ולא מורגש כלל הבדל במקרה זה. בס"ה הדגרדציה זוביקטיבית היא מתונה יותר מהאובייקטיבית.

יתן לומר אם כן, כי הניסוים במערכת המכילה מספר אבני בנין "סקלריות" זור, נתנו בהחלט תוצאות משביעות רצון. באות דיבור העובר בדרכו מספר לא דול של אבני בנין כאלה תחול דגרדציה, אולם באופן הדרגתי ובמתינות.

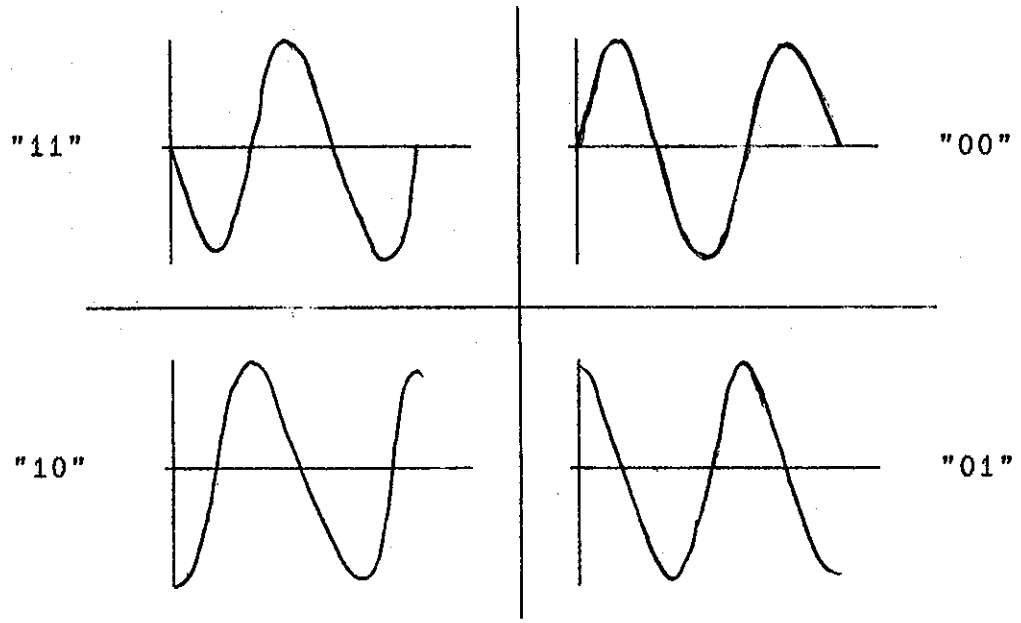
5.1. טיפול באותות נתונים (data)

ערכות תקשורת דיבור משמשות כיום במקרים רבים גם לתקשורת נתונים, מועברים באותו תחום תדרים כמו אות הדיבור (Voice-Band-Data - VBD). פיכך רצוי לבדוק כיצד מתפקדת המערכת, כאשר אות הכניסה שלה הוא מטיפוס ה.

צורך בדיקות אלה נעשתה סימולציה של modem המקובל בהעברת VBD. זהו modem טיפוס QPSK (Quadrature Phase Shift Keying), המיועד להעברת סיביות מידע קצב של 2400 bps. הגל הנושא הוא בתדר 1800 Hz, ומכאן שכל זוג סיביות

© Technion - Israel Institute of Technology, Elyachar Central Library

ויוצג ע"י מחזור וחצי של סינוס, בארבע פאזות אפשריות. אותות ה-modem
תונים בציור 5.1:



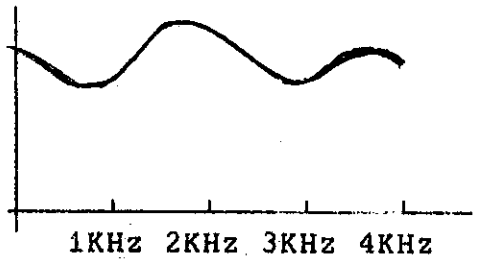
ציור 5.1: אותות QPSK ומשמעויותיהם; קצב שידור 2400bps, תדר גל נושא 1800Hz.

Figure 5.1: QPSK signals and symbols; bit rate 2400bps, carrier freq 1800Hz.

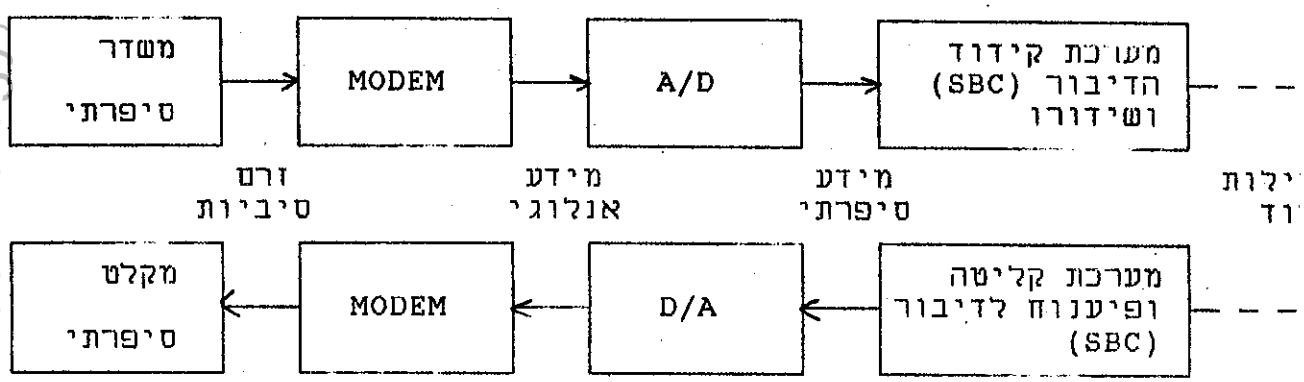
של קצב הדגימה במערכת (8KHz), מכיל כל סימבול בממוצע 6.66 דגימות. למעשה מחזור של 20 דגימות מתקבלים שלושה סימבולים רצופים, המכילים 6 דגימות, 7 גימות ו-7 דגימות בהתאמה.

דה-מודולציה במערכת נעשית בפשטות ע"י ארבעה קורלטורים המתאימים לארבעת ותות ה-QPSK, בשיטה של "החזק מנצח". בציור 5.2 נתון ספקטרום טיפוסי של ות ה-VBD. ציור 5.3 מתאר דיאגרמת בלוקים פשוטה של מערכת התקשורת במקרה ה.

© Technion - Israel Institute of Technology, Elyachar Central Library



ציור 5.2: ספקטרום טיפוסי של אות ה-VBD ששמש בסימולציות.
Figure 5.2: Typical spectrum of the VBD signal used in the simulations.



ציור 5.3: דיאגרמת בלוקים המתארת מעבר אות נתונים במערכת המומלצת.
Figure 5.3: A block diagram depicting passage of a VBD signal through the proposed system.

ניסויים שנערכו שמש אות ה-modem כקלט למערכת הקידוד ה"סקלרית", ועל הפלט וצעה דה-מודולציה, והשוואה לזרם הסיביות המקורי. הסתבר, כי המערכת עבירה את זרם הסיביות ללא שום טעות, גם בתנאים רגילים, וגם כאשר מוסיפים אות ה-modem רעש גאוסי לבן (הפרעה אפינית לתקשורת מידע) ביחס אות לרעש ל 16dB. לעומת זאת, רעש גאוסי לבן ב-SNR של 10dB נותן שגיאות ביציאה, הסתברות של $7x10^{-5}$. לשם השוואה נציין כי בסביבת רעש גאוסי לבן בעוצמות נ"ל, ללא שימוש במערכת קידוד הדיבור, לא מתקבלות שגיאות ביציאה כלל

ם כאשר מוסיפים במערכת שגיאות ערוץ (ספרתיות) ב-BER של 10^{-4} (ובתוספת

© Technion Israel Institute of Technology, Elyachar Central Library

גנה, כמתואר בסעיף 5.1.2), עובר זרם הסיביות ללא כל שגיאה, בלי רעש אוסי לבן ועם רעש ב-SNR של 16dB. רעש ב-SNR של 10dB נותן שוב שגיאות יציאה, בהסתברות של $7x10^{-5}$, כמו במקרה של ערוץ ללא שגיאות.

סה"כ ניתן לראות כי VBD עובר היטב במערכת, אפילו בתנאי רעש גאוסי לבן דיטיבי בעצמה בינונית, ורעש ערוץ ספרתי ה"מותר" עבור אות דיבור. דבר זה ובע בראש ובראשונה מהיתירות העצומה שבאות ה-modem, הנשמרת למרות הדחיסה, כן עוזרת ההקצאה הדינמית של סיביות, שבפרוש מתאימה עצמה לספקטרום שבציור 5., כפי שנמצא בנסויים שנערכו.

© Technion - Israel Institute of Technology, Elyachar Central Library

5. המערכת המומלצת עם קוואנטיזר וקטורי

אחר שנידונו אפייניה של המערכת ה"סקלרית", המיועדת לעבודה בקצב של 16Kbps, יש לפרט אפיינים אלה גם לגבי המערכת ה"וקטורית", המיועדת לעבודה בקצב של 9.6Kbps. לפני שנגש לפרוט, נזכיר בקצרה את מבנה המערכת (ר' גם סעיף 4.2.1).

מינון במערכת ה"וקטורית" מתבצע,כמו ב"סקלרית", ע"י מערך QMF מלא, הבנוי ובנה עץ בן שלוש רמות (כלומר 8 פסי-תדר). הניסויים נערכו בעזרת מסננים ל 64 מקדמים, במטרה להמנע משגיאות הנובעות מהסנון, אולם דבר זה אינו רחי.

קידוד בכל פס מתבצע בנפרד, בעזרת מקודד וקטורי מבוקר הגבר. ראשית, קובצים הדגמים בפס התדר לוקטורים ממימד 4, ולאחר מכן משוערך ההגבר, לכל וק של ארבעה וקטורים, בתור הנורמה הממוצעת של וקטורי הבלוק. ההגבר קוונטט לוגריתמית בארבע סיביות, משמש לנרמול הוקטורים, ונשלח אינפורמצית צד. הוקטורים עצמם מקודדים בעזרת מילון אופטימלי, המתוכנן ווסבר בסעיף 3.4.3.

קצאת הסיביות במערכת היא קבועה (זו למעשה הקצאה של מילוני קידוד). ערכי הקצאה, מתדר נמוך לגבוה בסיביות לדגם הם 3.5, 2, 0.75, 0.5, 0.5, 0.25, ו-0 בהתאמה. המילון בפס הראשון ממומש ע"י קטקדה של שני מילונים, בני 2 וקטורים ו-64 וקטורים בהתאמה. גדלי המילונים הבאים הם, כסדרם, 256, 4, 4, 2, 1 ו-1.

י שניתן לראות, כמות הסיביות הכוללת המיועדת לאינפורמציה המרכזית היא 7.5Kbps. בנוסף לכך יש להקצות 0.25Kbps להעברת אינפורמצית הצד בכל אחד משת הפסים ה"אפקטיביים", כלומר בס"ה יש ליחד 1.5Kbps לאינפורמצית הצד. שארים, לכן, 0.6Kbps למטרות הגנה (ר' גם סעיף 5.2.2).

נוצאות האוביקטיביות המתקבלות במערכת הן של SNR=14dB ו-SNRSEG=14.4dB. ביקטיבית האיכות טובה, למרות שמורגשים מעט רעשים.

© Technion - Israel Institute of Technology, Elyachar Central Library

פנה לכן לתאור המאפיינים המרכזיים של המערכת, הנוגעים לישומה בסביבה נשית של תקשורת דיבור.

5.2. סיבוכיות המימוש

פי שצוין בעת הדיון במערכת הסקלרית (סעיף 5.1.1), מקובל לדון בנפרד מקולי סיבוכיות זמן ובשיקולי סיבוכיות מקום, וכן להפריד את סיבוכיות זשדר מסיבוכיות המקלט במערכת. בנוסף לכך, מפרידים את סיבוכיות הסינון זיבוכיות הקידוד, כמקובל במערכות SBC.

כיוון שמנגנון הסינון במערכת זו זהה לחלוטין לזה שתואר בסעיף 5.1.1, בור המערכת ה"סקלרית", לא נפרטו שנית, ונזכיר רק את התוצאה הסופית: יבוכיות הזמן של האנליזה היא: $\sum_{i=1}^{m} L_i/2$ כפלים, ואותו מספר חיבורים, לדגם; יבוכיות המקום היא: $\sum_{i=1}^{m} L_i$ מקדמים ($m=3$ הוא מספר הרמות בעץ, L_i היא מספר קדמי המסנן הבסיסי ברמה ה-i). סיבוכיות הזמן והמקום של הסינתיזה זהות זו של האנליזה, ולפיכך העומס המוטל על המשדר והמקלט מבחינת הסינון הוא זה. סיבוכיות הזמן של הסינון במערכת הנסוי היתה של 96 פעולות לדגם, אך יתן להורידה בערך פי 8, כמפורט בסעיף 5.1.1.

זמקד עתה בסיבוכיות הזמן של הקידוד במערכת. באופן דומה להגדרות שבסעיף .5.1, נגדיר בתור $N_i=2^{R_i}$ את מספר הוקטורים במילון של הפס ה-i-י (R_i היא קצאת הסיביות לוקטור בפס זה), וזאת מלבד במילון הראשון, שם המערכת ורכבת מקסקדה של שני מילונים.

חיפוש (האופטימלי) במילון אופטימלי בן N_i וקטורים, המבוסס על פונקצית נוות הרבועית, דורש N_i כפלים ו-N_i חיבורים לדגם כניסה, לפי הפרוט הבא: ל וקטור המקודד בעזרת המילון, יש להשוות ל-N_i וקטורי קוד; השוואה לוקטור וד אחד דורשת k כפלים ו-k (למעשה k-1) חיבורים, כאשר k הוא מימד הוקטור; שוואה זו מתבצעת פעם ב-k דגמי כניסה (המהווים וקטור), ולכן דרוש כפל חיבור לכל דגם; מכיוון שיש N_i וקטורי קוד, עלינו לבצע N_i כפלים ו-N_i יבורים לכל דגם.

© Technion - Israel Institute of Technology, Elyachar Central Library

:וסף לכך יש לבצע N_1 (למעשה N_1-1) השוואות פעם בוקטור כניסה, למציאת
ורחק המינימלי. יש להוסיף לכן לסיבוכיות הזמן N_1/k השוואות לדגם.

ז'יבוכיות עבור פס-תדר (כדוגמת הראשון), שבו משתמשים בקסקדה של שני
'לונים, עם N_1^1 ו-N_1^2 וקטורי קוד בהתאמה, נתונה כמובן ע"י $N_1^1+N_1^2$ כפלים
ובורים ו-$(N_1^1+N_1^2)/k$ השוואות - כל זאת לדגם. בנוסף לכך יש לבצע חיסור
וד עבור כל דגם (ניתן להזניחו).

'י לקבל את הסיבוכיות הכוללת של הקידוד במערכת, יש לסכם את הסיבוכיות
:ל הפסים (מתקבלת הסיבוכיות עבור M דגמים "אנכיים"), ולחלק ב-M, לקבלת
ויבוכיות לדגם כניסה.

:ור בקרת ההגבר, אם נגדיר בתור K את מספר הדגמים ב"בלוק-קבוע-הגבר", אזי
ן לבצע ל-Kk הדגמים בבלוק זה Kk כפלים וחבורים, K הוצאות שורש, והזזה
ות (חלוקה ב-K), ובס"ה כפל, חבור, $1/k$ הוצאות שורש ו-$1/Kk$ הזזות לדגם.
ואנטיזצית ההגבר מתבצעת, פעם ב"בלוק קבוע הגבר", ע"י מספר השוואות
שוואה אחת לדגם בס"ה). יש גם להוסיף חלוקה אחת בהגבר המקוונטט, פעם
'גם.

'ענוח במערכת מתבצע בפשטות ע"י פניה לטבלת המילון, ועל כך יש להוסיף את
ענוח ההגבר (ג"כ פניה לטבלת קוואנטיזציה) והכפלה בהגבר המפוענח, עבור כל
.ם.

'בוכיות המקום במקודד ובמפענח זהה, והיא דורשת החזקתם של כל המילונים,
'ומר N_1 וקטורים בני k דגמים בכל מילון.

'י שעשינו בסעיף 5.1.1, נפרט גם עתה את הערכים המעשיים של סיבוכיות הזמן
וערכת המומלצת כדי לסבר את העין. סיבוכיות הזמן של הקידוד עצמו היא 75
נולות לדגם (בהזנחת ההשוואות לעומת הכפלים והחיבורים). עבור בקרת ההגבר
ן להוסיף כשלוש פעולות לדגם במקודד ובמפענח. סיבוכיות הפענוח (פניה
ובלאות) היא זניחה.

'י שניתן לראות, וכצפוי, העומס החישובי במערכת אינו נובע עתה רק מהסינון
:מערכת ה"סקלרית", כי אם גם מהקידוד. לפיכך, בנוסף להצעות שהוצעו בסעיף

© Technion - Israel Institute of Technology, Elyachar Central Library

5.1 לחסכון בסיבוכיות הסינון (עבוד מקבילי ו/או הורדת סדרי המסננים).
ן להציע במערכת זו גם שיטות לחסכון בסיבוכיות הקידוד.

ו במקרה של הסינון, ניתן להציע גם כאן עבוד מקבילי שיביא לחסכון ע"ח
ויר המערכת. בשיטה זו נבצע את הקידוד בכל פס ע"י מקודד נפרד, כאשר צוואר
קבוק של המערכת הוא המקודד בפס הראשון, שהוא התובעני מכולם. בהתחשב
ך, נצליח להוריד את זמן הקידוד בערך פי שניים. באופן כללי, אלגוריתמים
סכון סיבוכיות במערכות קידוד וקטורי ע"י תוספת חמרה, הנם נושא בעל
יבות רבה כיום, ולאחרונה התפרסמו מאמרים רבים הדנים בו, ויכולים לעזור
מוש קונקרטי של המערכת (ר' למשל [27]).

טה שניה לחסכון, הפעם על חשבון איכות, דווחה בסיעף 4.2.3 ומתייחסת
ילונים במבנה עץ. כפי שצויין שם, הדגרדציה באיכות איננה משמעותית,
חסכון בסיבוכיות הזמן של הקידוד הוא רב; במקום N_1 פעולות לדגם עבור
ידוד, יש לבצע רק $R_1=\log_2 N_1$ פעולות, ובס"ה נקבל כ-30 פעולות לדגם.
שלום, בנוסף לדגרדציה באיכות, הוא כמובן, גם בסיבוכיות המקום הגדלה פי
יים (במקרה של המערכת שלפנינו, יש להחזיק 4800 מספרים בנקודה צפה במקום
24).

ה"כ, כפי שניתן לראות, למרות העובדה שהמערכת ה"וקטורית", הפועלת בקצב
ווך מה"סקלרית", מסובכת יותר ממנה למימוש, ניתן להביאה, בשיטות השונות
ורטו, לרמת הטכנולוגיה הקיימת כיום בשוק. נראה, אם כן, כי אין מניעה
מש גם את המערכת ה"וקטורית" בסביבה מעשית.

5.2 עמידות לרעשים

ו במערכת ה"סקלרית" (סעיף 5.1.2), יש לדון גם בעמידותה של המערכת
יוקטורית" לשגיאות ערוץ ולרעש רקע. מכיוון שהרקע בנושא זה ניתן כבר
רחבה בסעיף 5.1.2, נגש מיד לפירוט העובדות הנוגעות למערכת הספציפית.

נורה בה משודרת האינפורמציה על הערוץ במערכת זו היא הבאה: בכל 32msec
שודרת מסגרת, בת 307 סיביות, המכילה אינפורמציה מרכזית (קידוד וקטורים
נורמלים), אינפורמצית צד (הגבר מקוונטט) וסיביות הגנה. הזולוקה היא

© Technion - Israel Institute of Technology, Elyachar Central Library

לקמן: האינפורמציה המרכזית מכילה אינווקטים של שמונה וקטורים מכל פס, ומר 112 סיביות (64+48) עבור הפס הראשון, 64 סיביות עבור הפס השני, 24 ביות עבור השלישי, 16 סיביות לרביעי, 16 לחמישי ו-8 לשישי - בסה"כ 240 ביות לאינפורמציה המרכזית; אינפורמצית הצד מכילה שני ערכי הגבר (8 ביות) לכל פס מהששה, כלומר 48 סיביות בס"ה; נותרו 19 סיביות להגנה.

חילה נבדקו בצועי המערכת ללא שימוש בסיביות ההגנה. נוסו ערכי BER של 10, 10^{-3} ו-10^{-4}, והתקבלו תוצאות אובייקטיביות, כמפורט בטבלה 5.5 (עם וואה לביצועים ללא שגיאות).

SNRSEG (dB)	SNR (dB)	BER
14.40	14.00	0
8.29	7.37	10^{-2}
11.79	10.68	10^{-3}
12.29	11.44	10^{-4}

טבלה 5.5: בצועי המערכת ה"וקטורית" תחת השפעת שגיאות ערוץ.
Table 5.5: "Vector" system performance with channel errors.

י שניתן לראות באופן אובייקטיבי, גם ב-BER גבוה של 10^{-2}, השפעת השגיאות ננה קטסטרופאלית. מבחינה סוביקטיבית, התקבלו ב-$BER=10^{-2}$ הפרעות מעותיות למדי, למרות שבאופן אחיד ולא פתאומי; ב-$BER=10^{-3}$ התקבלו הפרעות טות וב-$BER=10^{-4}$ כמעט שלא הורגשו הפרעות כלל.

מולציה של השימוש בסיביות הגנה במקרה זה, נעשתה, כבמערכת ה"סקלרית", י אי הכנסת שגיאות לאינפורמצית הצד, שעליה חשוב ביותר להגן. במקרה של ערכת ה"וקטורית" ההצדקה לכך היא אף גדולה מבמקרה ה"סקלרי", משום שבקוד ו זה שלפנינו, המקיים $k=48, n=67$, ניתן להשיג (לפי הטבלה הנתונה בסוף 2]) $d=7$, כלומר תיקון של שלוש שגיאות. משיקולים שפורטו בסעיף 5.1.2 בל לכן כי ההסתברות שתקרה שגיאה באינפורמצית הצד, למרות ההגנה, היא יחה. התוצאות שהתקבלו עבור ניסויים אלה מפורטות בטבלה 5.6.

© Technion - Israel Institute of Technology, Elyachar Central Library

SNRSEG (dB)	SNR (dB)	BER
9.70	8.15	10^{-2}
12.02	11.05	10^{-3}
12.29	11.40	10^{-4}

טבלה 5.6: ביצועי המערכת ה"וקטורית", עם השפעת שגיאות ערוץ והגנה.
Table 5.6: "Vector" system perfomance, with channel errors and protection.

ני שניתן לראות, חל שיפור ניכר בתוצאות בזכות ההגנה, כאשר סוביקטיבית
תן לומר כי ב-$BER=10^{-2}$ התקבלו הפרעת קטנות, בעוד כ-BER-ים של 10^{-3}
10^{-4}. לא נשמעות כלל הפרעות. ניתן לומר, לפיכך, כי למערכת זו "מותר"
נבוד ב-$BER=10^{-3}$.

יבות לתוצאות טובות אלה במערכת ה"וקטורית" (לעומת התוצאות הטובות פחות
ערכת ה"סקלרית") הן כדלקמן: חוסר האופטציה האחורית מונע הפרעות דרסטיות
אלה שנוצרות במערכת ה"סקלרית"; גם ללא הגנה, שגיאות הנופלות באינפורמציית
ד הן הרבה פחות קריטיות מאשר במקרה ה"סקלרי", שכן השפעתן מוגבלת לפסים
ודים בבלוק ולא לכל הבלוק; בנוסף לכך המערכת ה"וקטורית" רועשת יותר
ה"סקלרית" מלכתחילה (בשל הקצב הנמוך יותר) ורעש זה יוצר רמובן אפקט
זוים של מיסוך.

אשר למדידות שנעשו עם רעש רקע לבן, התקבלו תוצאות דומות לאלה שבמערכת
סקלרית" - האפקט שנוצר הוא צביעה של הרעש, אם כי בגוון אחר מאשר במקרה
סקלרי", ללא שינוי ניכר בעצמת הרעש או במובנות הדיבור; ניתן לומר בפרוש
הרעש הצבוע נעים פחות לאוזן מהמקורי.

© Technion - Israel Institute of Technology, Elyachar Central Library

5.2 Tandeming

בנושא זה ניתן חומר הרקע בסעיף המקביל עבור המערכת ה"סקלרית" (5.1.3)
פיכך נגש ישר לפירוט תוצאות הנסויים, המופיעות בטבלה 5.7.

SNRSEG (dB)	SNR (dB)	רמת Tandeming
14.40	14.00	1
10.40	9.62	2
9.26	8.46	3
8.19	7.64	4
7.59	7.09	5

טבלה 5.7: הדגרדציה המתקבלת באות העובר מספר מערכות "וקטוריות" ב-tandeming.

Table 5.7: Degradation in a signal passing through several tandemed "vector" systems.

ו במקרה ה"סקלרי", גם כאן יש ירידה חריפה, אובייקטיבית, במעבר ממערכת
ת לשתיים, וירידה מתונה יותר בהמשך. עם זאת גם כאן, מבחינה סוביקטיבית,
רידה באיכות היא מתונה, ולא חריפה אפילו במעבר ממערכת אחת לשתיים.
פקט העקרי בשמיעה הוא לאו דווקא התגברות של רעש הרקע, כי אם אפקט כללי
עמעום באות, אולם, כאמור, באופן מתון.

5.2 טיפול באותות נתונים

תום את הדיון באפיניה המעשיים של המערכת ה"וקטורית", בהתייחסות להשפעת
ערכת על אותות VBD. גם במקרה זה ניתן הרקע העיוני בסעיף המקביל עבור
ערכת ה"סקלרית" (5.1.4), ולכן נפרט מיד את התוצאות.

ור אות modem נקי, ללא תוספת רעש לבן וללא התחשבות בשגיאות ערוץ,
קבלו שגיאות בפענוח בהסתברות של $4x10^{-3}$. כפי שניתן לראות, אין טעם
משיך ולבדוק את השפעת המערכת על אותות נתונים בתנאי רעש, שכן תוצאה זו
ר גרועה למדי לכשעצמה.

© Technion - Israel Institute of Technology, Elyachar Central Library

סברים לקבלת תוצאה כה גרועה, לעומת התוצאות הטובות במערכת ה"סקלרית",
ו הבאים: ראשית, המערכת ה"וקטורית" עובדת בקצב של 9.6Kbps (ולא 16Kbps)
וכאן שדחיסת אות ה-modem היא רבה, וקשה יותר לנצל את היתירות שבו. לשם
זוואה, במערכת "סקלרית" העובדת בקצב של 9.6Kbps באותם תנאים, מתקבלת
תברות לשגיאה ביציאה של $4x10^{-4}$, לעומת מידע נקי משגיאות ב-16Kbps.
ית, המערכת עובדת עם הקצאה קבועה של סיביות המותאמת לספקטרום הדיבור,
יננה מותאמת לספקטרום ה-VBD (ר' ציור 5.2). המערכת לא בנויה, לכן, לטפל
ורה טובה באותות נתונים. על בעיות אלו ניתן להתגבר למשל ע"י תכנון
ודדים נפרדים, המותאמים לדיבור ולנתונים, שיעבדו בשילוב עם מערכת זיהוי
בור/נתונים; המלצה זו חורגת מגבולות העבודה.

י לסכם כעת את אפיניה של המערכת ה"וקטורית", כמו גם את אלה של
סקלרית", נפנה לסעיף הבא, העורך השוואה ממצה בין שתי המערכות.

© Technion - Israel Institute of Technology, Elyachar Central Library

5. השוואה בין קוואנטיזציה סקלרית ווקטורית במערכת SBC

ערכות המוצעות, עם קידוד סקלרי ווקטורי, לא ניתנות למעשה להשוואה
ידית, שכן כל אחת מהן מיועדת לעבודה בקצב אחר. עם זאת, ישנן מספר
קודות חשובות שיש להדגישן, לגבי ההבדלים בין שתי המערכות.

אלה המרכזית היא שאלת הסיבוכיות. המערכת ה"וקטורית" היא בעליל מסובכת
בה יותר למימוש מאשר ה"סקלרית", על אף העובדה שהיא מיועדת לעבודה בקצב
וך יותר. למרות שהוצעו בסעיפים הקודמים מספר פתרונות להקלת עומס
יבוכיות, ולמרות העובדה שכיום מתכננים כבר שבבים מיוחדים, המיועדים
קידוד וקטורי מהיר, יש עדיין לראות את גורם הסיבוכיות כמרכזי, כאשר
ונינים לממש את המערכת. עם זאת, דווקא בקצב כמו זה המוצע (9.6Kbps) קשה
צוא פתרונות טובים, הזולים יותר מהמערכת ה"וקטורית".

יה נוספת היא בעית סדרת הלימוד. בעוד המערכת ה"סקלרית" איננה מתוכננת
י סוג דיבור מסוים, ואמורה להתאים את עצמה לכל סביבה, הרי המערכת
וקטורית" מתוכננת ע"ס סדרת למוד, האמורה לאפיין את הדיבור שיקודד
זרתה. למרות אורכה של סדרת הלימוד, ולמרות מנגנון האדפטציה ששולב
ערכת, יש עדין הבדלים בתוצאות בין קידוד דיבור מאותו טיפוס כמו סדרת
מוד, לבין קידוד דיבור ממקור זר לחלוטין. יש, אם כן, להבין כי מערכת זו
ועדת לעבוד בסביבה מסוימת, ויש להתאימה מראש לסביבה זו כדי להפיק ממנה
צאות טובות.

ודות נוספות שהועלו בפרק זה הן כדלקמן: המערכת ה"וקטורית" רגישה פחות
גיאות ערוץ מה"סקלרית", ובתוספת ההגנה המוצעת מסוגלת לעבוד ב-$BER=10^{-3}$
עומת הסקלרית, שלה מומלץ $BER=10^{-4}$). הדבר נובע, קרוב לודאי, מהעובדה
מערכת זו אין אדפטציה אחורית משום סוג. במערכת ה"וקטורית" אפשר גם לשלב
לות מנגנון הגנה מתוחכם יותר מאשר ב"סקלרית", שכן עומד לרשותנו יותר
ב "שיורי". נקודה נוספת היא ענין העברת אות VBD: המערכת ה"וקטורית"
קשה בכך יותר, גם בשל הקצב הנמוך שלה, וגם בשל מנגנון האדפטציה הלא
פק. שתי המערכות מגיבות באופן סביר ל-tandeming.

יכות הסוביקטיבית של הדיבור הנשמע בשתי שיטות הקידוד היא כמובן שונה.
ערכת ה"סקלרית", העובדת בקצב של 16Kbps, נשמעת איכות טובה למדי,

© Technion - Israel Institute of Technology, Elyashar Central Library

מורגשים ברקע מעט צלצולים עדינים. ניתן להשוות איכות זו לאיכות הסוביקטיבית של מערכת μ-law-PCM עם בין 6 ל-7 סיביות, המערכת ה"וקטורית", המיועדת לעבוד בקצב של 9.6Kbps, רועשת כמובן מעט יותר, והרעש הוא בעל אופי של צרידות ועמעום. ניתן להשוות את איכותה לזו של מערכת μ-law-PCM עם בין 5 ל-6 סיביות. יש לציין, כי ההשוואה ל-μ-law-PCM אינה פשוטה, בשל האופי השונה של הרעש בכל מערכת, אולם בצורה פשטנית אלה התוצאות המתקבלות.

בשלב זה ניתן, אם כן, לסכם את תוצאות העבודה, לאור המטרות שהעמדנו לעצמנו, וכן לדון במספר הצעות למחקר המשך בנושא זה. נושאים אלה ידונו בפרק הבא.

© Technion - Israel Institute of Technology, Elyachar Central Library

פרק 6 - סיכום

6. סיכום העבודה

פי שצוין בראשית העבודה, מטרת מחקר זה היתה שילוב של קוואנטיזציה קטורית במערכת SBC, והשוואתה למערכות SBC קיימות עם קוואנטיזציה קלרית.

מהלך העבודה תוכננה מערכת SBC "סקלרית", נעשתה סימולציה שלה ובוצעו מספר סויים שנועדו לקבוע את הפרמטרים עמם היא עובדת. גובשה הצעה למערכת כזו, מומלצת לעבוד בקצב של 16Kbps, ומאופיינת בעקר ע"י הקצאה דינמית של יביות וקוואנטיזרים אדפטיביים (אחורית) בכל פס תדר. מאפייני המערכת תאור מפורט שלה נתונים בסעיף 5.1 בעבודה זו.

אחר שנלמד על בוריו נושא הקוואנטיזציה הוקטורית, שהינו אחד הכלים אטרקטיביים ביותר כיום להורדת קצב התמסורת במערכות קידוד דיבור, הועלו ספר הצעות לשילובה במערכת SBC. שני הכיוונים עליהם הושם דגש במהלך סימולציות היו: מחד, נסיון למצוא מערכת שתתמודד עם ה"סקלרית" ב-16Kbps, לא תשלום רב מדי בסיבוכיות, ומאידך, נסיון למצוא מערכת, שתאפשר עבודה קצבים של 8-9.6Kbps, בהם המערכת הסקלרית אינה מסוגלת לפעול באיכות טובה יה, גם אם ידרש תשלום גבוה יחסית בסיבוכיות.

וסו שתי שיטות שונות לקידוד וקטורי במערכת. האחת, "אנכית", שבה מורכב וקטור לקידוד ממספר רכיבי תדר המופיעים בו זמנית במוצא מערך המסננים. שניה - "אפקית", בה מקודד כל פס תדר בנפרד, כשהדגמים בו מקובצים וקטורים. בכל אחת משתי השיטות נוסה מנגנון אדפטציה קדמית, המבוסס על רמול לפי הגבר, במטרה לשפר את ביצועי המערכת.

שיטה ה"אנכית" הניבה תוצאות טובות למדי, אולם השיטה ה"אפקית" עלתה עליה הרבה (בעיקר לאחר הוספת מנגנון האדפטציה). בשיטה זו נוסו גם מילוני ידוד תת-אופטימליים במבנה עץ, המיועדים לשפר את סיבוכיות הזמן של קידוד, והם הניבו תוצאות טובות. בקצב של 16Kbps מפיקה המערכת איכות צוינת, אולם הסיבוכיות הגבוהה, והאיכות הטובה של המערכת ה"סקלרית" בקצב

© Technion - Israel Institute of Technology, Elyachar Central Library

זה, מורידים את אטרקטיביות המערכת ה"וקטורית". בקצב של 8Kbps המערכת ה"וקטורית" איננה באיכות מספקת, אולם בקצב של 9.6Kbps האיכות טובה למדי והיא בפרוש מוממדת מתאימה למימוש, ללא תחרות מצד המערכת ה"סקלרית", וזאת גם בהתחשב בשיקולי סיבוכיות.

נאורים מפורטים של שתי המערכות המוצעות, "סקלרית" (לקצב של 16Kbps) ו-"וקטורית" (לקצב של 9.6Kbps), נתונים כאמור בפרק 5, יחד עם פירוט של ואפיינים מעשיים, כגון סיבוכיות, עמידות לרעשים וכדומה. עם זאת, נראה כי יש עדין מספר כוונים אפשריים למחקר, במטרה לשפר את הסיבוכיות, האיכות וקצב העבודה. נושאים אלה ידונו בסעיף הבא.

© Technion - Israel Institute of Technology, Elyachar Central Library

6.2 כוונים למחקר המשך

המטרה העקרית של מחקר המשך בתחום זה, היא כמובן להמשיך ולהוריד את קצב התמסורת במערכת, מבלי לפגוע במידה ניכרת באיכות, ומבלי שידרש תשלום רב ודי בסיבוכיות. נראה אם כן כי המערכת שעליה יש לבצע נסויים נוספים היא המערכת ה"וקטורית", שהוכיחה את עצמה בהשגת מטרות אלה. נפרט מספר כוונים אפשריים, הנראים מבטיחים.

אחת הטכניקות המקובלות לשפור איכות של מערכת קידוד דיבור, היא לשלב בה אלמנט של חיזוי. כפי שראינו, חיזוי פשוט (לא אדפטיבי ועם מעט מקדמים) איננו אפקטיבי במערכת ה"סקלרית", ונראה לכן שגם לא ב"וקטורית". עם זאת, יש בסיס להניח, כי חיזוי מתוחכם יותר, אדפטיבי ועם מספר גדול יחסית של מקדמים, עשוי לשפר את בצועי המערכת. טכניקות לשלוב חזאי במערכת עם קידוד וקטורי נסקרו בסעיף 3.4.2, ולא נחזור ונפרטן כעת.

שיטות להקטנת העומס החשובי של קידוד וקטורי הנן כיום נושא, לו מוקדשת תשומת לב רבה. בין השיטות שניתן ליישם במערכת בולטות בעיקר טכניקות חיפוש תת-אופטימלי במילון אופטימלי, וכן אלגוריתמים המשתמשים ב"סריגים". אלה האחרונים הם למעשה קוואנטייזרים וקטוריים יוניפורמיים, המאפשרים, כמו מקרה הסקלרי, בניה מהירה של מילוני קידוד, אכסון נוח וקידוד מהיר (על שבון האיכות כמובן). שיטות להקטנת הסיבוכיות יכולות לאפשר קידוד מטריצי" (שילוב של "אנכי" ו-"אפקי"), וע"י כך לשפר את איכות המערכת.

מידה ומעוניינים להוריד עוד את קצב העבודה, ניתן להשתמש במקודדי אנטרופיה ללא עוות) או דוחסים בעלי עוות למיניהם, וע"י כך להשיג קצב תמסורת נמוך ותר, תמורת תשלום לא גדול באיכות ובסיבוכיות. תוצאה דומה ניתן להשיג ע"י ילוב טכניקות כווץ רוחב הסרט, כגון TDHS (ר' [28]).

סיכום, יש לציין רעיון נוסף שעלה לאחרונה (ר' [29]), ומתיחס אופטימיזציה משולבת של המסננים והמקודדים במערכת. כפי שצוין כבר בעבודה, קובל בד"כ להפריד את מנגנוני הסינון והקידוד, אולם הפרדה זו אינה כרחית. כפי שראינו, הקידוד גורם בעקיפין גם לשגיאות הנובעות מהסינון, תכנון משולב של המסננים והמקודדים יכול לשפר את איכות המערכת, על חשבון בודת הכנה רבה יותר בזמן התכנון.

© Technion - Israel Institute of Technology, Elyachar Central Library

כפי שניתן לראות, למרות העובדה שבעבודה זו הוצגו מערכות מומלצות שלמות על כל מאפייניהן, יש עדין אפשרות למחקר נוסף בתחום זה, במטרה לשפר את איכות המערכת, להקטין את סיבוכיות מימושה, ולאפשר עבודה עמה בתחום קצבים נמוך והמוצע.

© Technion - Israel Institute of Technology, Elyachar Central Library

ספח א' - נוסחאות ה-QMF והוכחתן

ספח זה מכיל פתוח מפורט של נוסחאות ה-QMF (ר' גם סעיף 1.2.1). פתוח זה יתן גם למצוא במספר מקורות ספרות, לדוגמא [1] עמ' 496-500, [2] עמ' 378-38 וכד'.

פיתוח מתייחס לדיאגרמה שבציור 1.2. נסמן ב-$X(e^{jw})$ את התמרת פוריה של אות כניסה, וב-$X_1(e^{jw})$, $X_2(e^{jw})$ את התמרות האותות בפסי הנתר, הנמוך והגבוה התאמה. מסנני האנליזה יסומנו בתור $h_1(n)=h(n)$ ו-$h_2(n)=(-1)^n h(n)$, ומסנני סינתיזה יהיו $h_1(n)$ ו- $-h_2(n)$ (כאשר התמרותיהם יסומנו כרגיל באותיות דולות).

קשר בין אות הכניסה לאותות בתת הפסים הוא

$$X_1(e^{jw}) = X(e^{jw})H_1(e^{jw}) \qquad (A.1)$$
$$X_2(e^{jw}) = X(e^{jw})H_2(e^{jw})$$

נדיר בתור $Y_1(e^{jw})$ ו-$Y_2(e^{jw})$ את הספקטרה לאחר הדצימציה ונקבל:

$$Y_1(e^{jw}) = 1/2[X_1(e^{jw/2})+X_1(e^{j(w+2\pi)/2})] \qquad (A.2)$$
$$Y_2(e^{jw}) = 1/2[X_2(e^{jw/2})+X_2(e^{j(w+2\pi)/2})]$$

אינטרפולציה והסינתיזה יתנו אותות בעלי התמרות $U_1(e^{jw})$ ו-$U_2(e^{jw})$, זקיימים:

$$U_1(e^{jw}) = 2Y_1(e^{2jw})H_1(e^{jw}) \qquad (A.3)$$
$$U_2(e^{jw}) = -2Y_2(e^{2jw})H_2(e^{jw})$$

אות היציאה מקיים:

$$\hat{X}(e^{jw}) = U_1(e^{jw})+ U_2(e^{jw}) \qquad (A.4)$$

© Technion - Israel Institute of Technology, Elyachar Central Library

הצבת (A.3) ב-(A.4) נותנת:

$$\hat{X}(e^{jw}) = 2Y_1(e^{2jw})H_1(e^{jw})-2Y_2(e^{2jw})H_2(e^{jw}) \tag{A.5}$$

הצבת (A.2) ב-(A.5) נותנת:

$$\hat{X}(e^{jw})=[X_1(e^{jw})+X_1(e^{j(w+\pi)})]H_1(e^{jw})-[X_2(e^{jw})+X_2(e^{j(w+\pi)})]H_2(e^{jw}) \tag{A.6}$$

הצבת (A.1) ב-(A.6) נותנת:

$$\begin{aligned}\hat{X}(e^{jw}) &= [X(e^{jw})H_1(e^{jw})+X(e^{j(w+\pi)})H_1(e^{j(w+\pi)})]H_1(e^{jw}) - \\ &\quad [X(e^{jw})H_2(e^{jw})+X(e^{j(w+\pi)})H_2(e^{j(w+\pi)})]H_2(e^{jw}) \\ &= X(e^{jw})[H_1^2(e^{jw})-H_2^2(e^{jw})] + \\ &\quad X(e^{j(w+\pi)})[H_1(e^{jw})H_1(e^{j(w+\pi)})-H_2(e^{jw})H_2(e^{j(w+\pi)})]\end{aligned} \tag{A.7}$$

האיבר הראשון במשוואה האחרונה מייצג את הסינון הרצוי, ואילו השני את השפעת ההתחזות, הלא רצויה. בנקודה זו יש לזכור כי $h_2(n)=(-1)^n h_1(n)$, או במישור התדר - $H_1(e^{jw})=H_2(e^{j(w+\pi)})$. קשר זה מבטל כליל את האיבר הלא רצוי ב-(A.7). יש לשים לב כי ביטול ההתחזות מתקבל גם ללא הדרישה שהמסננים יהיו בעלי מעבר חד.

זכור כעת גם כי $h(n)$ הוא זוגי וסימטרי. לכן, אם L הוא סדר המסנן, ניתן לרשום:

$$H(e^{jw}) = |H(e^{jw})|e^{jw(L-1)/2} \tag{A.8}$$

הצבה ב-(A.7) תתן:

$$\hat{X}(e^{jw}) = X(e^{jw})[|H(e^{jw})|^2 + |H(e^{j(w+\pi)})|^2]e^{jw(L-1)} \tag{A.9}$$

לפיכך, כדי שכל ההבדל בין האות המקורי והמשוחזר יהיה בהשהיה (ההכרחית בחינה פיזיקלית) עלינו לדרוש:

$$|H(e^{jw})|^2 + |H(e^{j(w+\pi)})|^2 = 1 \tag{A.10}$$

לפי שצויין בגוף העבודה (משוואה (1.1)).

© Technion - Israel Institute of Technology, Elyachar Central Library

פח ב' - נוסחאות ההקצאה הדינמית של סיביות והוכחתן

ספח זה מפותחות נוסחאות ההקצאה הדינמית של סיביות, כפי שצויין בסעיף 2.2. גם במקרה זה ניתן למצוא את הפיתוחים (או לפחות חלקם) במקורות רות, למשל ב-[1] עמ' 493-494. נפתח כאן את הנוסחה הכללית, למקרה של פסי ר לא שווי רוחב. מתוכה ניתן, בתור מקרה פרטי, להוציא את הנוסחה עבור ים שווי רוחב.

ב מעבר האינפורמציה במערכת מוגדר ע"י:

$$I = \sum_{i=1}^{M} f_i R_i \quad \text{Kbps} \tag{B.1}$$

ור f_i הוא תדר הדגימה של הפס ה-i-י (KHz) ו-R_i היא הקצאת הסיביות לדגם ו זה.

יר את R כממוצע מס' הסיביות לפס, משוקלל לפי קצבי דגימה:

$$R = 1/2W \sum_{i=1}^{M} f_i R_i \tag{B.2}$$

ור W הוא רוחב הסרט של אות הדיבור. מתקיים הקשר $I=2WR$, האופייני רכות תקשורת דיבור.

קצית המטרה לבעית האופטימיזציה מוגדרת בנוסחה (2.5) ונחזור עליה כאן:

$$D = \varepsilon/M \sum_{i=1}^{M} 2^{-2R_i} \sigma_i^2 \tag{B.3}$$

ר σ_i^2 הוא וריאנס האות בפס ה-i-י.

יה היא לפיכך הבעיה הבאה:

$$\min_{R_i} \ \varepsilon/M \sum_{i=1}^{M} 2^{-2R_i} \sigma_i^2 \quad \text{s.t.} \quad \sum_{i=1}^{M} f_i R_i = 2WR = I \tag{B.4}$$

© Technion - Israel Institute of Technology, Elyachar Central Library

יר, לשם הפתרון, לגרנז'יאן, תוך שימוש בכופל ג:

$$\mathcal{L}(R_1,\ldots,R_M,\lambda) = \varepsilon/M \sum_{i=1}^{M} 2^{-2R_i}\,\sigma_i^2 - \lambda(I - \sum_{i=1}^{M} f_i\,R_i) \tag{B.5}$$

רה של (B.5) לפי R_i והשוואה לאפס יתנו:

$$\frac{\partial\mathcal{L}}{\partial R_i} = -2\ln 2\ \varepsilon/M\ \sigma_i^2\ 2^{-2R_i^*} + \lambda f_i = 0 \tag{B.6}$$

י מניפולציות אלגבריות פשוטות נקבל:

$$R_i^* = \log_2\left(\frac{2\ln 2\ \varepsilon/M\ \sigma_i^2}{\lambda f_i}\right)^{1/2} \tag{B.7}$$

ב את (B.7) באילוץ ונקבל:

$$\begin{aligned} I &= \sum_{i=1}^{M} f_i \log_2\left(\frac{2\ln 2\ \varepsilon/M\ \sigma_i^2}{\lambda f_i}\right)^{1/2} \\ &= \sum_{i=1}^{M} \log_2\left(\frac{2\ln 2\ \varepsilon/M\ \sigma_i^2}{\lambda f_i}\right)^{f_i/2} \\ &= \log_2 \prod_{i=1}^{M}\left(\frac{2\ln 2\ \varepsilon/M\ \sigma_i^2}{\lambda f_i}\right)^{f_i/2} \\ &= \log_2 \left[\left(\frac{2\ln 2\ \varepsilon/M}{\lambda}\right)^{1/2 \sum_{i=1}^{M} f_i} \cdot \prod_{i=1}^{M}\left(\sigma_i^2/f_i\right)^{f_i/2}\right] \end{aligned} \tag{B.8}$$

© Technion - Israel Institute of Technology, Elyachar Central Library

רת מעט אלגברה, ואם נזכור ש- $\sum_{i=1}^{M} f_i = 2W$, נקבל :

$$\frac{2\ln 2\ \varepsilon/M}{\lambda} = \left[\frac{2^{I}}{\prod_{i=1}^{M} (\sigma_i^2/f_i)^{f_i/2}} \right]^{1/W} \tag{B.9}$$

נציב זאת ב-(B.7) נקבל, בעזרת עוד מעט אלגברה:

$$R_i^* = I/2W + 1/2 \log_2 \left(\frac{\sigma_i^2/f_i}{\prod_{i=1}^{M} (\sigma_i^2/f_i)^{f_i/2W}} \right) \tag{B.10}$$

נוסחה (2.8). ניתן לראות בקלות, שהצבת המקרה הפרטי של פסי תדר שווי
ב (כלומר $f_i=2W/M$) מביאה נוסחה זו לצורה הפשוטה, הנתונה בנוסחה
(2.

© Technion - Israel Institute of Technology, Elyachar Central Library

© Technion - Israel Institute of Technology, Elyachar Central Library

References

שימת מקורות

[1] Jayant N.S. and Noll P., "Digital Coding of Waveforms - Principles and Applications to Speech and Video," Englewoo Cliffs, NJ: Prentice Hall, 1984.

[2] Crochiere R.E. and Rabiner L.R., "Multirate Digital Signa Processing," Englewood Cliffs, NJ: Prentice Hall, 1983.

[3] Rabiner L.R. and Schafer R.W., "Digital Processing of speec Signals," Englewood Cliffs, NJ: Prentice Hall, 1978.

[4] Falangan J.L. et al,"Speech Coding," IEEE Trans. Comm., vo COM-27, pp.710-737, Apr. 1979.

[5] Tribolet J.M. and Crochiere R.E., "Frequency Domain Coding Speech," IEEE Trans. ASSP, vol. ASSP-27, pp.512-530, Oct. 1979.

[6] Esteban D. and Galand C., "Application of Quadrature Mirr Filters to Split Band Voice Coding Schemes," Proc. 1977 Int IEEE Conf. on ASSP, pp.191-195.

[7] Lloyd S.P., "Least Squares Quantization in PCM," IEEE Tran Info. Theory, vol. IT-28, pp.129-137, March 1982.

[8] Shoham Y. and Gersho A., " Efficient Codebook Allocation for Arbitary Set of Vector Quantizers," Proc. 1985 Int'l IEEE Con on ASSP, pp.1696-1699.

[9] Cox R.V. and Crochiere R.E., "Real-Time Simulation of Adapti Transform Coding," IEEE Trans. ASSP, vol. ASSP-29, pp.147-15 Apr. 1981.

© Technion - Israel Institute of Technology, Elyachar Central Library

[10] Ramstad T.A., "Subband Coders with a Simple Adaptive Bit-Allocation Algorithm - A Possible Candidate for Digital Mobile Telephony?" Proc. 1982 Int'l IEEE Conf. on ASSP, pp.203-207

[11] Gray R.M., "Vector Quantization," IEEE ASSP Magazine, pp.4-29 Apr. 1984.

[12] Makhoul J.,Roucous S. and Gish H., "Vector Quantization in Speec Coding," Proc. of the IEEE, vol. 73, pp.1951-1985, Nov. 1985.

[13] Linde Y., Buzo A. and Gray R.M., "An Algorithm for Vecto Quantizer Design," IEEE Trans. Comm, vol COM-28, pp.84-95, Jul 1980.

[14] Gray R.M. Kieffer J.C. and Linde Y., "Locally Optimal Bloc Quantizer Design," Information and Control 45, pp.178-198, 1980

[15] Buzo A., Gray A.H. Jr., Gray R.M. and Markel J.D., "Speech Codi Based Upon Vector Quantization," IEEE Trans. ASSP, vol ASSP-2 pp.562-574, Oct. 1980.

[16] Juang B.H. and Gray A.H. Jr., "Multiple Stage Vector Quantizati for Speech Coding," Proc. 1982 Int'l IEEE Conf. on ASS pp.597-600.

[17] Atal B.S. and Schroeder R.M., "Stochastic Coding of Spee Signals at Very Low Bit Rates," Proc. 1984 Int'l IEEE Conf. Comm., pp.1610-1613.

[18] Foster J. and Gray R.M., "Finite-State Vector Quantizatior Abstracts of the 1982 Intl. Symposium on Information Theory.

[19] Cuperman V. and Gersho A., "Vector Predictive Coding of Speech at 16 Kbits/s," IEEE Trans. Comm., vol COM-33,pp.685-696, July 1985.

[20] Chen J.H. and Gersho A., "Vector Adaptive Predictive Coding o: Speech at 9.6 Kbps," Proc. 1986 Int'l IEEE Conf. on ASSP pp.1693-1696.

[21] Chen J.H. and Gersho A., "Gain Adaptive Vector Quantization fo Medium-Rate Speech Coding," Proc. 1985 Int'l IEEE Conf. on Comm. pp.1456-1460.

[22] He N., Buzo A. and Kuhlmann F., "A Frequency Domain Wavefor Speech Compression System Based on Product Vector Quantizers, Proc. 1986 Int'l IEEE Conf. on ASSP, pp.3031-3034.

[23] Abut H. and Luse S.A., "Vector Quantization for Subband Code Waveforms," Proc. 1984 Int'l IEEE Conf. on ASSP, No.10.6.

[24] Gersho A., Ramstad T. and Versvik I., "Fully Quantized Subbar Coders with Adaptive Codebook Allocation," Proc. 1984 Int'l IEI Conf. on ASSP, No.10.7.

[25] Mensa G., Montagna R. and Rusina F., "Comparison Between Vect and Scalar Quantization in Variable Rate Subband Coders," Pro 1986 Int'l IEEE Conf. on ASSP, pp.2383-2385.

[26] McWilliams F.J. and Sloane N.J.A., "The Theory of Err Correcting Codes," North Holland Publishing Company,1981.

[27] Capello P., Davidson G.,Gersho A., Koc C. and Somayazulu V., Systolic Vector Quantization Processor for Real-Time Spee Coding," Proc. 1986 Int'l IEEE Conf. on ASSP, pp.2143-2146.

© Technion - Israel Institute of Technology, Elyachar Central Library

[28] Malah D., "Time Domain Algorithms for Harmonic Bandwidth Reduction and Time Scaling of Speech Signals," IEEE Trans. ASSP, vol. ASSP-27, pp.121-133, Apr. 1979.

[29] Dembo A., "Design of Digital FIR Filter Arrays," Ph.D. Thesis Technion I.I.T., May 1986 (in Hebrew).

© Technion - Israel Institute of Technology, Elyachar Central Library

Scalar and Vector Quantization in Subband Coding of Speech

Research Thesis

Submitted in Partial Fulfillment of the Requirements
for the Degree of
Master of Science
in Electrical Engineering

By
Ran Arad

Submitted to the Senate of the
Technion - Israel Institute of Technology
Tishrei 5747 Haifa October 1986

© Technion - Israel Institute of Technology, Elyachar Central Library

This research was carried out in the Faculty of Electrical Engineering under the Supervision of Prof. David Malah.

I wish to express my deep gratitude to Professor David Malah, for proposing the subject, and for his helpful and instructive guidance throughout this work.

I also wish to thank the Staff of the Signal Processing Lab. and especially Yoram Or-Chen, Zipi Portnoy, Ziva Avni and Baruch Kohavi, for helping me through some difficult hours.

echnion - Israel Institute of Technology, Elyachar Central Lib

Contents

Page

Abstract

List of Symbols and Abbreviations

Chapter 1 - Introduction

1.1 Frequency Domain Coding of Speech.....1

1.2 Subband Coding (SBC) of Speech.....3

1.2.1 The Filtering in SBC.....4

1.2.2 The Quantization in SBC.....7

1.3 Research Aims.....9

1.4 Structure of the Thesis.....10

Chapter 2 - A SBC System with Scalar Quantization

2.1 Scalar Quantization.....11

2.2 Bit Alloction Problem.....14

2.2.1 Theoretical Background.....14

2.2.2 Practical Considerations.....16

2.3 A System with Static Bit Allocation.....19

2.4 A System with Dynamic Bit Allocation.....24

2.4.1 Backward Adaptation.....26

2.4.2 Forward Adaptation.....27

Chapter 3 - Vector Quantization

3.1 Vector Quantization Problem.....30

3.2 The LBG Algorithm for Optimal Vector Quantizer Design.....32

3.2.1 Lloyd's Conditions.....32

3.2.2 The LBG Algorithm.....33

3.2.3 Initial Conditions for the Algorithm.....34

3.2.4 the Empty Cell Problem.....35

© Technion - Israel Institute of Technology, Elyachar Central Library

Contents (Cont.)

Page

3.3 Suboptimal Vector Quantizers....37
3.3.1 Product Codes....37
3.3.2 Tree-Structured Quantizers....38
3.3.3 Multistage Quantizers....41
3.4 Speech Compression Systems with Vector Quantization....43
3.4.1 Vector Quantization in Source Coding....43
3.4.2 Predictive Systems....44
3.4.3 Gain-Adaptive Systems....47
3.4.4 Frequency Domain Vector Coding....49

Chapter 4 - A SBC System with Vector Quantization

4.1 A System with "Vertical" Quantization....51
4.4.1 Incorporation of Gain Adaptation....53
4.1.2 Incorporation of Frequency Domain Weighting....55
4.2 A System with "Horizontal" Quantization....59
4.2.1 Incorporation of Gain Adaptation....62
4.2.2 Dynamic Codebook Allocation....64
4.2.3 A System with Tree-Structured Codebooks....65

chapter 5 - Characteristics of the Proposed Systems

5.1 The Proposed System with Scalar Quantization....67
5.1.1 Complexity of Implementaion....68
5.1.2 Robustness....71
5.1.3 Tandeming....75
5.1.4 Dealing with Data Signals....76
5.2 The Proposed System with Vector Quantization....80
5.2.1 Complexity of Implementaion....81
5.2.2 Robustness....83
5.2.3 Tandeming....86
5.2.4 Dealing with Data Signals....86
5.3 Comparison of Scalar and Vector Quantization in SBC Systems...88

© Technion - Israel Institute of Technology, Elyachar Central Library

Contents (Cont.)

Page

Chapter 6 - Summary

6.1 Thesis Summary.....90

6.2 Proposals for Further Research.....92

Appendix A - QMF Formulas and Proofs.....94

Appendix B - Dynamic Bit Allocation Formulas and Proofs.....96

References

Abstract (in English)

© Technion - Israel Institute of Technology, Elyachar Central Library

© Technion - Israel Institute of Technology, Eyshan Central Library

Abstract

This research deals with Subband Coding systems - SBC - for speech compression, and is mainly concerned with the quantization mechanism in those systems.

Subband coding of speech is a well-known frequency domain method for speech coding, and is especially efficient at medium-band rates (8-16Kbps). In this method, the speech signal is separated into several subbands, using an analysis filter bank, and each band is coded separately, according to analytical (e.g. SNR) or subjective (hearing) criteria. The coded signals are transmitted through a digital channel, and are decoded at the receiver and recombined to form the reconstructed speech signal.

The system design is conventionally seperated into two: filter bank (or analysis-synthesis system) design for transmitter and receiver, and the design of the quantizers, used in the subband signals' coding and decoding. This research, as mentioned above, is concerned with the latter problem.

The SBC systems that are in use today, usually employ quantizers such as ADPCM, that may be called "Scalar Quantizers", for they work upon the coded signal sample by sample. Nevertheless, the last few years have seen the emergence of Vector Quantizers - VQ - based upon simultanous coding of several adjacent samples. Vector quantizers are not new in the literature, but their popularity had grown as of late, following the publishing of a straight-forward design algorithm, and with the technological advancements permitting their real-time implementation.

Vector quantizers are used today as building blocks for many speech coding schemes, and, amongst others, were proposed to be incorporated into SBC systems. Nonetheless, a thorough work covering all the

aspects of such a system, with concrete propositions for applications, based upon simulations, was not published as yet. Such propositions and experimentations are included in this work

The scope of this research covers, on one hand, the conventional systems, with scalar quantization, and on the other hand, some algorithms for the incorporation of vector coders into the system. The proposed systems were tested, to ensure their applicability to real-time implementations, and all those tests are documented in the work as well. The main conclusion, regarding SBC-VQ systems, is that, although at higher medium-band rates (e.g. 16Kbps) the improvement in quality does not justify the computational burden, it seems that at lower rates (e.g. 9.6Kbps) the "vector" scheme can not be matched by the "scalar" one.

Concluding the work are proposals for further research in SBC-VQ systems, having the goal of enhancing their quality and decreasing their complexity, beyond the results described in this project.

Technion - Israel Institute of Technology, Eyadah Central Library