



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 מקדמי התמרה של תמונות באמצעות דיאגרמת שבכה

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

מאת

שמואל פרקש

חל 952

2058695



000000869874

הוגש לסנט הטכניון - מכון טכנולוגי לישראל

9.1.89

המחזקר נעשה בהנחיתם של פרופ' ד. מלאך מהפקולטה להנדסת חשמל - טכניון חיפה, ופרופ' ויליאם פרלמן מ RPI טרוי ניו-יורק.

ברצוני להביע את תודתי לפרופ' דוד מלאך על הנחייתו המסורה, ועל תרומתו להשגת תוצאות המחקר. כן ברצוני להודות לפרופ' ויליאם פרלמן על הצעת הרעיון ועצותיו המועילות.

תודות לזיוה אבני, לנמרוד פלג, ליורם אור-חן וליתר אנשי מעבדת "עיבוד אותות" על עזרתם הרבה במהלך התקופה בה עבדתי במעבדה.

המחזקר מומן בחלקו ע"י הקרן המשותפת משרד התעשייה והמסחר (מת"מ) - טכניון (מחזקר 050-0520).

לזיכרו של אבי
מנחם פרקש.

תוכן הענינים

תקציר

רשימת סמלים וקיצורים

פרק 1 - מבוא

- 1.1 הצגת הבעיה - קידוד מקדמי ההתמרה בעזרת דיאגרמת שבכה (TRELLIS)
- 1.2 שיטות מקובלות לקידוד תמונות
- 1.3 מטרת המחקר ותוצאותיו העיקריות
- 1.4 מבנה העבודה

פרק 2 - התמרות אורתוגונליות - רקע עיוני

- 2.1 מבוא
- 2.2 התמרות אורתוגונליות
- 2.3 התמרת קרונן לוחב (KLT)
- 2.4 התמרות תת-אופטימליות
- 2.5 המוטיבציה לשמוש בהתמרת Cosine בסכמה המוצעת

פרק 3 - קוונטיזציה וקטורית

- 3.1 הקדמה
- 3.2 קוונטיזציה וקטורית באמצעות ספר קוד
- 3.3 קוונטיזציה וקטורית באמצעות דיאגרמת שבכה
 - 3.3.1 קוונטיזציה וקטוריים במבנה עץ
 - 3.3.2 קוונטיזציה וקטוריים במבנה שבכה
- 3.4 המוטיבציה לשמוש בדיאגרמת שבכה בסכמה המוצעת.
 - 3.4.1 הקדמה
 - 3.4.2 השוואת ספר קוד ושבכה בלא אילוצי סיבוכיות

תוכן עיניים (המשך)

- 36 3.4.3 השוואה תחת אילוצים של כמות חשובים וזכרון זהה בשני הקוונטזרים
- 37 3.4.4 בדיקת השפעת אורכי הבלוקים על ביצועי שני הקוונטזרים
- 40 3.4.5 בדיקת ביצועי הקוונטזרים תחת אילוצים של כמות חשובים או כמות זכרון זהה
- 42 3.4.6 בדיקת ביצועי הקוונטזרים תחת אילוצים של כמות חשובים וכמות זכרון זהה
(איכלוס ציקלי של השבכה)
- 44 3.4.7 סכום

פרק 4 - תאור סכמת הקידוד הבסיסית

- 46 4.1 כללי
- 46 4.2 תאור הסכמה האופטימלית
- 47 4.3 בנית דיאגרמת השבכה
- 47 4.3.1 פונקצית הקצב עוות של מקור גאוסי מרקובי
- 49 4.3.2 בנית הדיאגרמה
- 51 4.4 אכלוס הדיאגרמה
- 51 4.4.1 איכלוס אקראי
- 53 4.4.2 איכלוס איטרטיבי
- 54 4.5 אלגוריתמי חיפוש בדיאגרמה
- 54 4.5.1 האלגוריתם של ויטרבי
- 55 4.5.2 אלגוריתם M
- 55 4.5.3 אלגוריתם ML
- 56 4.6 שיקולי ממוש מעשיים
- 56 4.6.1 התמרת DCT - Cosine
- 58 4.6.2 בנית דיאגרמת השבכה ואכלוסה

פרק 5 - בעית אי התאמת התמונת המודל ושיטות לפתרון

- 61 5.1 אי התאמת התמונות למודל - תאור הבעיה
- 62 5.2 שמוש ב Clustering לפתרון בעית חוסר הסטציונריות
- 64 5.3 אדפטציה של הקוד

תוכן עיניים (המשך)

5.4 שמוש במקודד סקלרי לקידוד תמונת השארית

5.5 תאור הסכמה הכוללת

פרק 6 - ביצועי המערכת

6.1 ביצועי הסכמה המוצעת והשוואה לסכמות קידוד אחרות

פרק 7 - סיכום ומסקנות

7.1 סיכום העבודה

7.2 הצעות להמשך מחקר

נספח א - השוואת ספר קוד ושבכה בעלי כמות חישובים וזכרון זהה

מקורות

תקציר באנגלית

תקציר

בעבודה זו מוצגת סכימת קידוד אשר מבוססת על קדוד מקדמי התמרת Cosine דיסקרטית (DCT) של תמונות באמצעות קוונטיזציה וקטורי הממומש ע"י דיאגרמת שבכה. סכמת קדוד הכוללת התמרה אופטימלית (KLT) וקוונטיזציה וקטורית בעזרת דיאגרמת שבכה הינה אופטימלית, דהיינו בצוניה שואפים לפונקצית הקצב עוות עבור מודל אות גאוסי-מרקובי, אשר נחשב כמודל טוב לייצוג תמונות. התמרת Cosine משמשת אותנו כקרוב ל- KLT עקב העובדה שביצועיה קרובים לביצועי KLT והסבוכיות שלה קטנה יותר. הסיבה לשמוש בדיאגרמת שבכה ולא בספר קוד לממוש הקוונטיזציה והקטורי, נובעת מהעובדה שכמות החישובים וכמות הזכרון הנדרשים לממושה תלויים לינארית באורך הקטור המקודד, בהשוואה לתלות האקספוננציאלית הקיימת בספר קוד. פשטות יחסית זו של השבכה אינה פוגעת בהכרח בביצועי הסכמה, וכפי שאנו מראים בעבודה זו, ביצועי השבכה עולים על ביצועי ספר הקוד תחת אילוף של כמות זכרון זהה בשני הקוונטיזצורים בתחומי הפרמטרים שנבדקו.

בסכמת הקדוד המוצעת, תמונת המקור מחולקת לבלוקים אשר עוברים כל אחד התמרת Cosine דו-ממדית. ביצוע ההתמרה גורם להפחתה של הקורלציה בין האלמנטים השונים בבלוק, ולפיכך מתאפשרת דחיסה יעילה יותר. הבלוקים המותמרים מסווגים ע"י מנגנון הקבצה (Clustering) לקבוצות. מנגנון ההקבצה מבוסס על אלגוריתם LBG אשר משמש בד"כ לבניה של ספר קוד מתוך סדרת לימוד. בסיום תהליך הסווג מתקבלות קבוצות המכילות בלוקים בעלי תכולה ספקטרלית דומה. לכל קבוצה משוער ספקטרום מייצג אשר משמש לבנית דיאגרמת שבכה אשר מהווה את הקוונטיזציה והקטורי לקידוד הבלוקים השייכים לקבוצה. ספקטרום זה נשלח כאינפורמציה צד למקלט ומשמש שם לבנית דיאגרמת שבכה זהה לצורך הפיענוח. בנוסף, משמש הספקטרום המשוער להקצאה של קצבים בין הקבוצות השונות. האדפטציה הנ"ל של הקצבים בין הבלוקים השונים מאפשרת ניצול יעיל יותר של הקצב הכולל ע"י הקצאה של קצב גבוה יותר לקבוצה בעלת אנרגיה AC גבוהה יותר.

הבלוקים של כל אחת מהקבוצות מקודדים באמצעות דיאגרמת השבכה המתאימה לאותה קבוצה. לצורך שיפור נוסף של ביצועי הסכמה, ובמיוחד לצורך הפחתת תופעת ה"בלוקיות" (Blocking Effect), המאפיינת סכמות קידוד המבוססות על התמרה והפעולות בקצב נמוך, נעשה שימוש בקוונטיזציה סקלרית אשר מקודד את המקדמים בעלי העוצמה הגבוהה של תמונת השגיאה במישור ההתמרה. שרטי המקדמים האלה ומיקומם בבלוק מקודדים באמצעות קודי הופמן (Huffman Codes) ונשלחים כאינפורמציה צד למקלט.

בתהליך הפיענוח, מופענח כל בלוק בעזרת דיאגרמת השבכה אשר נבנתה במקלט, תוך שמוש באינפורמציה הצד. מתבצע תיקון של המקדמים אשר עבורו קוונטיזציה בעזרת הקוונטיזציה הסקלרית וכל בלוק מותמר בהתמרת Cosine הפוכה (IDCT) לקבלת הבלוק המשוחזר. התהליך הנ"ל חוזר על עצמו לגבי כל הבלוקים בתמונה.

עקב אי הסטציונריות של תמונות, הנחת המודל הגאוסי מרקובי אינה מתאימה בדרך כלל לייצוג תמונות. הסטייה של התמונות מהמודל מקוזזת בסכמה המוצעת ע"י השימוש בהקבצה, וע"י קידוד המקדמים הגדולים של תמונת השגיאה בעזרת הקוונטיזציה הסקלרית. ביצועי המערכת אופיינו על ידי שני מדדים כדלקמן:

א. יחס האות לדעש בין התמונה המקורית לתמונה המשוחזרת.

ב. האיכות הסובייקטיבית של התמונה המשוחזרת.

הביצועים נבדקו בקצבים שונים והתוצאות שהתקבלו היו טובות מאוד ואף מצויינות. עבור תמונת אשה "LENA" בגודל $[256 \times 256]$ התקבלה דחיסה לקצבים הקטנים מ 1 bit/pel עם יחסי אות לדעש גבוהים (מעל 32 dB ב 0.6 bit/pel ועד 35 dB ב 0.9 bit/pel) ובאיכות שיחזור טובה מאוד. ברזולוציה גבוהה (גודל תמונה $[512 \times 512]$) התקבלה דחיסה לקצבים הקטנים מ 0.5 bit/pel בביצועים מצויינים (37 dB). לפי מיטב ידיעתנו לא דווח עד כה בספרות על סכמת קידוד המשיגה עבור אותה תמונת מבחן ביצועים כה טובים כדוגמת אלה המושגים על ידנו.

הסיבוכיות הכרוכה במימוש הסכמה הינה גבוהה יחסית, אולם עיקר הסיבוכיות מרוכז בחלק המבצע את הקידוד. עובדה זו מאפשרת יישום של הסכמה למערכת אכסון ואחזור (Storage and Retrieval) של תמונות, שבחן שלב הקידוד יכול להתבצע באופן איטי, ולעומתו שלב הפיענוח חייב להיות מהיר. פתרון לבעיית הסיבוכיות יהיה קרוב לוודאי ישים עם התפתחות טכנולוגיית ה-VLSI.

בסיכום העבודה מובאות הצעות לשיפור המערכת בכוונים של שיפור הדיאגרמה ע"י איכלוסה בצורה יעילה יותר, שיפור מנגנון ההקבצה והתאמתו לתוכן התמונה והתאמת סכמת הקידוד לתכונות העין.

רשימת סמלים וקיצורים

Mean Square Error	-	MSE
Differential Pulse Code Modulation	-	DPCM
אלמנט תמונה	-	א"ת
גודל התמונה	-	$[N \times N]$
גודל בלוק ההתמרה	-	$[M \times M]$
Karhunen Loeve Transform	-	KLT
קצב	-	R
עוות	-	D
Discrete Cosine Transform	-	DCT
Signal to Noise Ratio	-	SNR
Vector Quantization	-	VQ
וקטור כניסה להתמרה	-	\underline{u}
וקטור יציאה מההתמרה	-	\underline{v}
סדרת וקטורי בסיס	-	$\{a_i\}$
Kronecker של הדלתה	-	δ_{ij}
התמרה אורתוגונלית	-	A
התמרה אורתוגונלית הפוכה	-	B
מטריצת קוואריאנס של הוקטור \underline{u}	-	$R_{\underline{u}\underline{u}}$
מדד יעילות הדה-קורלציה של ההתמרה	-	η_c
מדד יעילות רכוז האנרגיה של ההתמרה	-	η_E
הוקטור העצמי ה- i של מטריצת הקוואריאנס	-	ϕ_i
הערך העצמי המתאים לוקטור ϕ_i	-	λ_i
Discrete Fourier Transform	-	DFT
Discrete Walsh Hadamard Transform	-	DWHT
מקדם קורלציה	-	ρ
Quantizer	-	Q
מילון וקטורים	-	V
הקבוצה ה- k -ית מתוך המילון V	-	S_k
מדד עוות בין הוקטור \underline{u} לוקטור \underline{v}	-	$d(\underline{u}, \underline{v})$
פונקציית צפיפות ההסתברות של הוקטור \underline{u}	-	$P(\underline{u})$
סף עוות	-	ϵ
העוות באיטרציה m	-	$D^{(m)}$
וקטור פרטור בציה קבוע	-	δ

רשימת סמלים וקיצורים (המשך)

אורך הוקטור המקודד בספר הקוד	-	N_B
הקצב הנדרש בספר הקוד	-	R_B
גודל ספר הקוד	-	K_B
אורך הוקטור המקודד בשבכה	-	N_T
הקצב הנדרש בשבכה	-	R_T
אורך האילוך של השבכה	-	K_T, K
מספר ענפים המתפצלים מצומת בעץ או בשבכה	-	q
אורך העץ או השבכה	-	L
מספר מילות הקוד ברמה l של השבכה	-	n_l
כמות זכרון	-	Me
כמות חישובי עוות לאות מקור	-	Co
כמות זכרון בספר קוד	-	M_B
כמות חישובי עוות לאות מקור בספר קוד	-	C_B
כמות זכרון בשבכה	-	M_T
כמות חישובי עוות לאות מקור בשבכה	-	C_T
סדרת מדגם מתוך מקור	-	Z
מטריצת KLT	-	Γ
האינפורמציה ההדדית בין הוקטור \underline{x} לוקטור \underline{y}	-	$I(\underline{x}, \underline{y})$
פונקציית צפיפות ההסתברות המותנית של הוקטור \underline{y} בוקטור \underline{x}	-	$P(\underline{y} \underline{x})$
פונקציית הקצב עוות הפרמטרית	-	$R_N(D_\theta)$
הקצב ההמתאים לאות ה n -ית של וקטור המקור	-	r_θ^n
העוות ההמתאים לאות ה n -ית של וקטור המקור	-	d_θ^n
הקצב של מילות הקוד ברמה l של השבכה	-	ρ_l
ההפרש בין הקצב המעשי לקצב האופטימלי בשבכה	-	ϵ_N
מספר מילות הקוד אשר קיבלו קצב חיובי בהקצאה	-	N_c
סט האינדקסים ברמה l של השבכה	-	J_{N_l}
העוות המצטבר לצומת j ברמה l של השבכה	-	$d(j, l+1)$
העוות המתאים לענף ה k היוצא מצומת j בשבכה	-	$d(j_k)$
הואריאנט המנורמל של המקדם l בוקטור	-	\bar{S}_l
מספר הוקטורים באזור	-	M_a
ההספק הממוצע של הבלוק i	-	P_i
הואריאנט המנורמל המשוער של המקדם l בוקטור	-	\bar{S}_l

רשימת סמלים וקיצורים (המשך)

מקדם ההגבר של הבלוק i	-	G_i
הואריאנס המשווער של המקדם i בוקטור i	-	\bar{S}_i^i
מספר האברים במיצוע	-	N_s
הקבוצה ה- k ת-ת בהקבוצה	-	C_k
וקטור ממוצע של הקבוצה ה- k ת-ת	-	m_k
הקרדינליות של הקבוצה ה- k ת-ת	-	γ_k
פונקצית המדד של הקבוצה ה- k ת-ת	-	$J_k(C)$
פונקצית המדד	-	$J(C)$
עוות בין שני וקטורים	-	$\delta(\cdot, \cdot)$
אינפורמציות צד	-	SI_A, SI_B

פרק 1 - מבוא

1.1 הצגת הבעיה

אכסון תמונה בזכרון המחשב או שידורה בערוץ תקשורת דורש בד"כ כמות רבה של סיביות. השימוש ההולך וגדל במידע האצור בתמונות ספרתיות מעלה את הצורך בטכניקות יעילות לדחיסה (קידוד) של תמונות. המטרה המרכזית של דחיסת תמונות הינה יצוג התמונה ע"י מספר קטן ככל האפשר של סיביות, תוך שמירה על ביצועים רצויים, או לחילופין קבלת ביצועים טובים ככל האפשר עבור כמות סיביות נתונה (קצב נתון). המונח "ביצועים" אשר משמש אותנו בהגדרת הבעיה הינו מושג מעורפל במידה מסויימת בהקשר של דחיסת תמונות. טיב התמונה המשוחזרת כפי שהיא נתפסת במערכת הראיה שלנו אינו ניתן לאפיון ע"י מדד כמותי מובהק מחד גיסא, ומאידך האיכות הסובייקטיבית של התמונה המשוחזרת אינה מאפשרת השוואה כמותית בין תמונות שונות או אף בין אותה התמונה בתנאים שונים. הבעיות הנ"ל באפיון ביצועי הדחיסה מהווה את הקושי המרכזי בפיתוח שיטות דחיסה. מקובל בספרות לאפיין את הביצועים של מקודדים ע"י השגיאה הרבועית הממוצעת ($MSE - Mean Square Error$) בין תמונת המקור לתמונה המשוחזרת, למרות העובדה שאין מדד זה משקף נאמנה את האיכות הסובייקטיבית של התמונה. הסיבות לשימוש במדד זה הינן:

1. עבור מדד זה קיימים פתרונות לבעיות אופטימיזציה רבות.
2. קיימים חסמים תאורטיים (פונקצית הקצב-עוות) מתורת האינפורמציה אשר באמצעותם ניתן להעריך את הביצועים.
3. השגיאה הרבועית הממוצעת מבטאת למעשה את הספק שגיאת הקידוד.

החסרון העיקרי בהערכת ביצועי דחיסה בעזרת מדד MSE , הינו חוסר הרגישות שלו לבעיות שפה בתמונה ולתופעת הבלוקיות המופיעה בד"כ בדחיסה ע"י התמרות. תופעות אלה, למרות העובדה שאנרגית השגיאה הכרוכה בהן קטנה, פוגמות בצורה משמעותית באיכות התמונה.

בעבודה זו מיושמת סכמת דחיסה המבוססת על קידוד וקטורי של מקדמי התמרה של תמונה בעזרת דיאגרמת שבכה (Trellis) Mazor & Pearlman [1] הוכיחו שהסכמה הנ"ל הינה אופטימלית. דהיינו, ביצועיה עבור אות גאומטרי-סטטיסטי ומדד עוות MSE שואפים אסימפטוטית לפונקצית הקצב-עוות. בסכמה שמומשה על ידינו נעשו מספר שינויים בהשוואה לסכמה המקורית שהוצעה ב-[1]. כדי להתגבר על הבעיה של אי התאמת התמונות להנחות מודל אות גאומטרי-סטטיסטי, נעשה שימוש במנגנון הקבצה (*Clustering*) אשר מחלק את התמונה לאזורים הומוגניים, ומאפשר קידוד אופטימלי של כל אזור בנפרד. חלוקה כזו לאזורים מאפשרת בנוסף הקצאה של קצב שונה לכל אזור, בהתחשב בספקטרים המאפיין שלו, לקבלת ביצועים כוללים טובים יותר. כדי להתגבר על בעיית העוותים הנוצרים באזורי הקצה שבתוך התמונה, וכדי להתגבר על תופעת הבלוקיות המאפיינת מקודדי התמרה בקצב נמוך שולב במערכת מקודד סקלרי אשר מקודד את המקדמים האנרגטיים ביותר בתמונה השגיאה. תוספת הקצב והסיבוכיות המתלווים להוספת המנגנונים הנ"ל קטנה יחסית לסיבוכיות הקידוד בעזרת דיאגרמת שבכה בלבד. השיפור בביצועים האובייקטיביים (MSE) ובאיכות התמונה המשוחזרת הינם משמעותיים. בהמשך פרק זה נסקור שיטות מקובלות לדחיסת תמונות על מעלותיהן וחסרונותיהן, נפרט את מטרות המחקר ונסכם בקצרה את מבנה החיבור.

1.2 שיטות מקובלות לקידוד תמונות

בסעיף זה נתאר בקצרה את הגישות השונות המקובלות לפתרון הבעיה של דחיסת תמונות. רוב הטכניקות המקובלות לדחיסת תמונות מבוססות על שמוש באחת משתי הגישות הבאות: קידוד המשלב חזאי (Predictive Coding) או קידוד התמרה (Transform Coding) [2,3]. קיימת גם הגישה של קידוד מעורב (Hybrid Coding) [7] הכוללת קידוד המשלב חזאי עם קידוד ההתמרה, אולם אנו לא נתעמק בה בעבודה זו.

בשתי הגישות הנדונות מנצלים לצורך הדחיסה את היתירות (Redundancy) הקיימת בתמונה המקורית. בשלב ראשון מייצגים את תמונת המקור ע"י סדרת אותות חסרי קורלציה, ובשלב שני מייצגים כל אלמנט בסדרה בכמות סיביות (קצב) היחסית למידת התרומה שלו לאיכויות האובייקטיביות והסובייקטיביות של התמונה המשוחזרת. הייצוג הנ"ל של אלמנטים בסדרה, ונקרא קוונטיזציה, יכול להתבצע לגבי כל אלמנט בנפרד (קוונטיזציה סקלרית), או עבור וקטור של אלמנטים (קוונטיזציה וקטורית).

הביצועים השונים המצוטטים בסעיף זה הינם בד"כ ביצועים עבור קידוד תמונת אשה המכונה ("LENA"), אשר השימוש בה נפוץ מאוד וביצועי קידוד עבורה מדווחים בהרחבה בספרות. כל מקרה של ביצועים עבור תמונת מבחן אחרת יצוין במפורש.

1.2.1 מקודדי חיזוי - PREDICTIVE CODERS

טכניקות המשתמשות במקודדי חיזוי מנצלות את הקורלציה הקיימת בין אלמנטי תמונה (א"ת) שכנים בתמונה (תלות בתחום המרחב) לצורך יצירת חיזוי של אלמנט תמונה (א"ת) נוכחי על סמך הערך המקודד של שכניו הקודמים לו בשורות ו/או בעמודות. במידה וידוע המודל הסטטיסטי של התמונה ניתן להגיע לחיזוי טוב של ערכי א"ת ע"י העברתם דרך מסנן לנארי שנבנה ע"פ התכונות הסטטיסטיות הנ"ל. חיזוי טוב משמעו אנרגיה קטנה של אות השגיאה הנוצר בין א"ת המקוריים לחיזוי שלהם. אות השגיאה, אשר הקורלציה שבו קטנה מהקורלציה הקיימת בתמונה המקורית, עובר קוונטיזציה, מקודד ומשודר למקלט (או נשמר בזכרון). שחזור התמונה מתבצע ע"י פענוח אות השגיאה, העברתו דרך מסנן שחזור המבוסס על המסנן הנמצא במשדר וסיכומו עם אות השגיאה העוקב לקבלת האות המשוחזר.

השיטה הנ"ל של קידוד בשילוב חזאי הנקראת DPCM היא פשוטה וקלה יחסית לשימוש וניתן להגיע בעזרתה לשחזור טוב בקצב של כ- 3 סיביות לא"ת. החסרון העיקרי של שיטת ה-DPCM שהוצגה הינו רגישות המקודד לשינוי במודל האות הקובע את מסנן החיזוי, חסרון זה משמעותי במיוחד בתמונות עקב אי הסטציונריות שקיימת בד"כ בתמונות טבעיות. בעזרת אדפטציה של הקוונטיזציה ו/או של מסנן החיזוי ניתן להפחית את בעיית שינוי המודל ולהגיע לדחיסה בקצבים של 1-2 סיביות לא"ת באיכות שחזור טובה. תוספת הטיבויות הנובעת מאדפטציה של הסכמה הינה גדולה יחסית ולעיתים מונעת ממוש בזמן אמיתי. להרחבת ההסבר הנ"ל ולתאור סכמות קידוד המשלבות חזאי ראה

[2, 3, 4 pp 637-657].

1.2.2 מקודדי התמרה - TRANSFORM CODERS

הגישה השניה לדחיסת תמונות, שאליה משתייכת המערכת המוצעת בעבודה זו, מבוססת על מקודדי התמרה. בטכמה הבסיסית, התמונה בגודל $[N \times N]$ מחולקת לבלוקים בגודל $[M \times M]$, כל בלוק עובר התמרה דו-ממדית לקבלת בלוק של מקדמי התמרה. מקדמי ההתמרה עוברים קוונטיזציה, מקודדים ומשודרים למקלט. במקלט מתבצעת פעולה הפוכה הכוללת, פיענוח של המקדמים המקוונטים והתמרתם בהתמרה הפוכה לקבלת הבלוק המשוחזר. הפעולה הנ"ל נעשית ע"פ כל הבלוקים בתמונה.

בפעולת ההתמרה הדו-ממדית מתקבלת תמונה בעלת קורלציה נמוכה מזו הקיימת בתמונת המקור (התמרת KLT שתתואר בפרק הבא מבצעת דה-קורלציה מחולטת של תמונת המקור). בפעולת הקוונטיזציה מתבצעת חלוקת סיביות בין מקדמי ההתמרה לקבלת תמונה משוחזרת בעלת איכות טובה.

קוונטיזציה של האות וקידודו במישור ההתמרה מקנים יתרונות על פני קוונטיזציה וקידוד במישור התמונה. ב- [5 פרק 3] מובאים 2 מודים להערכת יעילות התמרה, המדד הראשון מבטא את מידת אי הקורלציה של האות, ואילו המדד השני מבטא את המידה בה האנרגיה מרוכזת במספר מועט של מקדמים. כפי שנראה בפרק 2 תכונות האות הרלוונטיות לדחיסה במישור התמונה, כפי שהוערכו ע"פ המדדים הנ"ל, פחות טובות מתכונותיו במישור ההתמרה. Pearl, Andrews & Pratt [6] מפתחים במאמרם מדד ליעילות התמרה, ומביאים השוואה בין התמרת הזהות (האות נשאר בתחום התמונה) לבין התמרות אורתוגנליות שונות. המדד שפותח, הדומה במהותו לפונקציית הקצב-עוות הידועה מתורת האינפורמציה, משווה עבור התמרות שונות (פורייה, KLT והדמדד) את הקצב R אליו ניתן להגיע תוך שמירה על עוות קבוע D עבור אות גאוסי - מרקובי. תוצאות המאמר מצביעות על עדיפות מכרעת של כל ההתמרות שנדונו על פני התמרה הזהות.

הזכרון העיקרי של מקודדי התמרה הינו הסיבוכיות שלהם יחסית למקודדי חיזוי. זכרון זה קטן בהדרגה עם התפתחות המעבדים המהירים, אשר יכולים לבצע התמרות שונות בזמן אמיתי.

השמוש במקודדי התמרה לקידוד תמונות התבסס בתחילה על סכמות קבועות אשר לא תלויות בשינויים הקיימים בסטטיסטיקת האות. הסכמות הללו (ראה למשל [8]) כללו בד"כ התמרה דו-ממדית מנגנון לבחירת מקדמים לקידוד ומנגנון להקצאת סיביות בין המקדמים. המקדמים שקודדו היו מקדמים באזור קבוע מראש (Zonal Sampling), הקצאת הסיביות בין המקדמים שקודדו היתה קבועה ע"פ כל התמונה. כצפוי, עקב אי הסטצינריות של התמונה, הביצועים של הטכמה הנ"ל לא היו מספקים, ולמעשה השתוו לביצועים של מקודדי חיזוי אדפטיביים.

כדי לשפר את הביצועים של מקודדי התמרה יש צורך להשתמש בשיטות אדפטציה שונות, אשר יתאימו את סכימת הקידוד לשינויים המקומיים של סטטיסטיקת האות. קיימות שלש שיטות עיקריות לאדפטציה של מקודדי ההתמרה:

- א. אדפטציה של ההתמרה.
- ב. אדפטציה של הקצאת הסיביות בין הבלוקים.
- ג. אדפטציה של הקוונטיזציה (הקצאת סיביות בין מקדמים השייכים לאותו בלוק).

אדפטציה של ההתמרה הינה השיטה הקשה ביותר למימוש, ומשמעותה הינה מציאה של סט חדש של וקטורי בסיס (מטריצת Karhunen-Loeve) עבור כל שינוי בסטטיסטיקת התמונה. לפיכך הגישה הנ"ל אינה ישימה.

השיטה השניה של אדפטציה, הקצאת סיביות בין הבלוקים בתמונה, הינה מעשית הרבה יותר ולכן גם נעשה בה שימוש רב יותר. השיטה הפשוטה ביותר השייכת לקבוצה זו הינה השיטה של קביעת סף קבוע לכל תמונה של יחידת המקדמים אשר עבורו את הסף בקצב מלא ואפוס המקדמים אשר קטנים ממנו. ב-[10] מצוטטות מספר מערכות כנ"ל. הביצועים המושגים טובים מביצועי המערכות ללא אדפטציה, אולם אין ירידה לקצבים הנמוכים מ-1 ביט לא"ת. בשיטה זו נדרשת אינפורמציה צד נוספת להעברת מיקום המקדמים שעברו את הסף. Chen & Smith [9] ממשו סכמה אדפטיבית ע"י חלוקה הבלוקים בתמונה, לאחר שעברו התמרת Cosine (DCT), ל-4 קבוצות על פי רמת הפעילות שלהם כפי שנמצאה ע"י אנרגיה ה-AC של כל בלוק. האדפטציה התבצעה ע"י הקצאת סיביות בין הקבוצות, תוך העדפה של הקבוצות בעלות רמת הפעילות הגבוהה על פני אלו בעלות רמת הפעילות הנמוכה יותר. התקבלה דחיסה לקבצים של 1 - 0.5 ביט לא"ת עבור תמונה בגודל [512x512]. כפי שמדווח במאמר, התמונה המשוחזרת היתה מצויינת. ה-MSE המנורמל המדוד עבור קצב של 1.0 ביט לא"ת הינו 0.05% (ערך המתאים ל-SNR של 33 dB כאשר $SNR \triangleq 10 \log \frac{255^2}{MSE}$). ה-MSE המדוד עבור 0.6 ביט לא"ת הינו 0.074% (המתאים ל-SNR של 31.3 dB).

השיטה השלישית לאדפטציה של מקודדי התמרה הינה אדפטציה של הקוונטיזר, ומשמעותה הקצאת סיביות אדפטיבית בין מקדמי ההתמרה השייכים לאותו בלוק. הקצאת סיביות בין מקדמי ההתמרה נעשית בעיקר על פי הסטטיסטיקה שלהם. רוב העבודות שנעשו עד כה הניחו מודל גאוסי עבור מקדמי ההתמרה. הקצאת הסיביות האופטימלית עבור מקדמים בעלי פילוג גאוסי מתקבלת כאשר שגיאת הקוונטיזציה של כל המקדמים זהה [11 פרק 5, 5 פרק 4.7]. משמעות הדבר היא ששגיאת השחזור (רעש הקוונטיזציה) של הבלוק הינה רעש לבן. קיום תנאי זה דורש שהקצאת הסיביות בין המקדמים תהיה יחסית ללוגריתם של הוואריאנס שלהם.

Chen & Smith [9] השתמשו בהקצאת סיביות אשר מבוססת על פונקציית הקצב - עוות של מ"א גאוסיים עבור מקדמי ה-DCT, למעט איבר ה-DC. לתאור מפורט של בעית הקצאת הסיביות בין מקדמי ההתמרה ולתוצאות נוספות שהושגו ראה פרק 3 ומקורות [2, 3, 5, 10, 12].

קוונטיזרים וקטוריים

כל הקוונטיזרים שהוזכרו בסקירה זו הן בהקשר למקודדי חיזוי והן בהקשר למקודדי התמרה היו קוונטיזרים סקלריים, דהיינו א"ת או מקדם במישור ההתמרה עברו קוונטיזציה ע"י קוונטיזר סקלרי. תוצאה חשובה של תורת האינפורמציה, הדנה בדחיסת אותות, אומרת שניתן להשיג ביצועים טובים יותר ע"י שמוש בקוונטיזר וקטורי במקום בקוונטיזר סקלרי. הביצועים הטובים יותר מושגים גם עבור סדרת א"ת או מקדמים בלתי תלויים סטטיסטית [13]. סקירה מקיפה של קוונטיזרים וקטוריים ניתן למצוא ב-[15].

קוונטיזציה וקטורי ניתן לאפיון בעזרת ספר קוד (Codebook) אשר מכיל K_B וקטורים באורך K_B כ"א הנקראים וקטורי קוד (Code Vectors). תהליך הקוונטיזציה הינו תהליך חיפוש בו משווה וקטור המקור לכ"א מוקטורי הקוד, האינדקס של וקטור הקוד הקרוב ביותר לוקטור המקור, בהתייחס לקריטריון מדויק נתון, נשלח בערוץ או נשמר בזכרון.

הצעד הראשון במימוש קוונטיזציה וקטורי הינו אכלוס ספר הקוד בוקטורי קוד. קיימות שתי גישות לאכלוס ספר קוד, אכלוס באקראי ואכלוס בוקטורים אופייניים הנבנים מתוך סדרת למוד. בשיטת האכלוס באקראי, מאכלטים את ספר הקוד בוקטורים אקראיים מתוך פילוג היציאה המתקבל מהפילוג המגשים של פונקציית הקצב - עוות [13 עמ 66].

השיטה השנייה לאכלוס ספר קוד הינה אכלוס של ספר הקוד בוקטורים אופייניים אשר נוצרים בעזרת סדרת לימוד. סדרת לימוד הינה סדרה גדולה יותר של וקטורים אשר נלקחים מתוך מדגם מייצג של המקורות אותם אנו רוצים לדחוס. בדחיסת תמונות, סדרת הלימוד הינה סט תמונות בעלות מאפיינים שייכסו את כל מגוון התמונות אותם אנו רוצים לדחוס. בניית ספר הקוד מתוך סדרת לימוד מתבצעת בעזרת אלגוריתם LBG המתואר ב- [14].

האכלוס בעזרת סדרת לימוד עדיף על פני האכלוס האקראי, עקב העובדה שהוא מקנה דרך נוחה ויעילה לתכנון ספר הקוד עבור מקור בעל פילוג כלשהו. האכלוס האקראי לעומת זאת, מאפשר בניה פשוטה של ספר קוד עבור מקורות סטטיסטיים ובעלי פונקציית פילוג ידועה.

הצעד הבא במימוש קוונטיזציה וקטורי הינו הבחירה של וקטור הקוד המייצג את וקטור המקור. חיפוש על פני כל הוקטורים בספר הקוד (Exhaustive Search) יתן כמובן את ההתאמה הטובה ביותר, אולם עבור מספר וקטורי קוד גדול שיטה זו אינה מעשית. שיטות אחרות לחיפוש בספר הקוד הינן יצוג בעזרת דיאגרמת עץ (Tree) או דיאגרמת שבכה (Trellis), וחיפוש יעיל עליהן. שיטות אלה מתוארות ב- [11, 15], תאור מפורט של דיאגרמת השבכה והשמוש בה יובאו בפרקים 3 ו-4.

בקידוד תמונות בעזרת קוונטיזציה וקטורי התמונה מחולקת לבלוקים. כל בלוק מאורגן כוקטור, ומתבצע חיפוש בספר הקוד למציאת וקטור הקוד הקרוב ביותר לוקטור המקור. האינדקס של וקטור הקוד שנבחר נשלח בערוץ או נשמר בזכרון. בתהליך הפענוח האינדקס הנ"ל משמש ככתובת לטבלה, אשר תוכנה מהווה את הוקטור המשוחזר. הקוונטיזציה הוקטורית שתוארה יכולה להתבצע בתחום המרחב (Spatial VQ) או בתחום ההתמרה (Transform VQ).

ב- [16] מתוארת סכמת קידוד במישור התמונה המשלבת מסנן חזיוני וקוונטיזציה וקטורי. מסנן החזיוני משמש להורדת היתירות וליצירת אות השגיאה, ואילו הקוונטיזציה וקטורי, הממומש בעזרת דיאגרמת עץ, מקוונטט את אות השגיאה הנוצר. יחס האות לרעש (SNR) שהתקבל בסכמה הנ"ל עבור דחיסה לקצב של 1 ביט לא"ת, היה 30.9 dB. תוצאה זו מהווה שיפור של כ-5 dB לתוצאות המתקבלות בסכמת DPCM עם קוונטיזציה סקלרית. ה- SNR שהתקבל עבור דחיסה לקצב של 0.5 ביט לא"ת היה 27.5 dB.

סכמת קידוד המנצלת קוונטיזציה וקטורי אדפטיבי במישור התמונה מתוארת ב- [17]. התמונה מחולקת לאזורים המכילים א"ת שכנים בתמונה ולכל אזור נבנה ספר קוד, כאשר בתור סדרת לימוד משתמשים בוקטורים השייכים לאזור המקודד. הוקטורים מהווים את ספר הקוד נשלחים למקלט כאינפורמציה צד. התוצאות המדווחות עבור הסכמה הנ"ל הינה עבור תמונת מבחן שונה מתמונת המבחן

אשר משמשת אותנו בעבודתנו, ולפיכך לא מצאנו טעם לצטט אותן. לפי המדווח במאמר, התוצאות המתקבלות הן בנות השוואה עם התוצאות המושגות בטכמת קידוד המשתמשת בהתמרה DCT ובקוונטוט מקדמיה בעזרת קוונטיזר סקלרי.

שמוש במסווג (Classifier) לשם חלוקת בלוקים בתמונה לקבוצות בהתאם למאפייני מערכת הראיה השונים, וקידוד הבלוקים באמצעות ספר קוד המתאים לאותה קבוצה מתואר ב- [20]. בטכמה זו חולקה התמונה לבלוקים אשר סווגו לאחת מתוך M קבוצות. הסווג נעשה בהתאם למאפיינים כפי שנתפסים ע"י מערכת הראיה, כגון: אזורי חלקים, אזורי בעלי שפוע מתון, אזורי בעלי קצוות בכוונים שונים וכו'. עבור כל קבוצה כזו נבנה ספר קוד באמצעות אלגוריתם LBG [14]. הבלוקים בתמונה קוונטוטו בעזרת ספר הקוד המתאים. הביצועים המדווחים עבור תמונת המבחן הסטנדרטית בגודל 512×512 הינם כדלקמן: SNR של 29.95 dB הושג בדחיסה לקצב של 0.65 ביט לא"ת. SNR של 28.55 dB הושג בדחיסה לקצב של 0.5 ביט לא"ת.

טכמת קידוד המשלבת קוונטיזר וקטורי במישור ההתמרה של תמונות מתוארת ב- [18]. רמות האנרגיה של בלוקים בתמונה, שעברו התמרת DCT, משמשות לסווג הבלוקים ל-8 קבוצות. לכל קבוצה נבנה ספר קוד מתוך סט תמונות המשמש כסדרת לימוד. כל מקדמי ההתמרה בבלוק, למעט איבר ה-DC, עוברים קוונטיזציה וקטורית בעזרת ספר הקוד של הקבוצה אליה שייך הבלוק. לקידוד מקדמי ה-DC יוצרים סדרת הפרשים של מקדמי DC שכנים. סדרת הפרשים מקוונטטת באמצעות קוונטיזר אחיד של סיביות (בין 8 ל +8), ערכי סדרת הפרשים שנמצאים מחוץ לתחום בקוונטיזר מקודדים ב-8 סיביות ונשלחים בתוספת סמון מיוחד המעיד על כך. שיטה זו לקידוד אברי ה-DC מורידה את הקצב הממוצע שלהם ל-4.5 סיביות לא"ת (במקום 8 סיביות לא"ת כמקובל). הקלקול המתקבל בתמונה כתוצאה מקידוד אברי ה-DC הינו זניח. התוצאות המדווחות במאמר הינן עבור תמונות אשר אינן מוכרות לנו. מדווח על דחיסה לקצב של 0.75 ביט לא"ת ושגיאת שחזור רבועית ממוצעת של 7.1 (מתאים ל- SNR של 31.1 dB).

לסיום סעיף זה, נתאר טכמת קידוד המשלבת קידוד של מקדמי DCT בעזרת דיאגרמת עץ. הטכמה הנ"ל דומה במבנה הבסיסי לטכמה המוצעת על ידינו. השוני העיקרי הינו במימוש של הקוונטיזר הוקטורי (דיאגרמת עץ לעומת דיאגרמת שבכה). הטכמה, אשר ביצועיה עבור מקור גאוסי - מרקובי שאפים לפונקציית הקצב-עוות כאשר גדל בלוק ההתמרה שואף לאינסוף, נוצלה ב- [19] לדחיסת תמונות.

התמונה המקודדת חולקה לבלוקים אשר עברו התמרת DCT לצורך הורדת הקורלציה בהם. התמונה חולקה לאזורי המכילים בלוקים שכנים, ועבור כל אזור שוערך ספקטרום מייצג. הבלוקים השייכים לאזור מסוים נורמלו ע"י הספקטרום המשוערך של אותו אזור וקוונטוטו בעזרת דיאגרמת עץ אשר אוכלסה באקראי בוקטורים אקראיים גאויסיים מנורמלים. ה-SNR שהושג בשחזור תמונת המבחן היה 32.4 dB בקצב של 0.74 ביט לא"ת.

1.3 מטרות המחקר ותאור התוצאות העיקריות שהושגו

מטרת המחקר הינה בחינה והערכה של סכמת קידוד לתמונות, המבוססת על התמרה וקוונטיזציה וקטורי הממומש באמצעות דיאגרמת שבכה. סכמה זו הוכחה כאופטימלית עבור סיגנלים גאומטריים - מרקוביים, אשר נחשבים כמודל טוב של תמונות אמיתיות. המודל הגאומטרי - מרקובי אינו מתאר במדויק תמונות אמיתיות, הבעיה העיקרית הקיימת בתמונות הינה חוסר הסטציונריות הקיים בהם. תופעה זו מתאפיינת בעיקר בבלוקיות הקיימת בתמונה המשוחררת. במסגרת המחקר פותחו לפיכך מספר מנגנונים כדי להתגבר על הבעיה הנ"ל. הקו אשר הנחה אותנו במימוש המערכת היה קבלת הביצועים הטובים ביותר הנתנים להשגה בלא תלות בסיבוכיות. הביצועים שהושגו היו אכן גבוהים בהשוואה לכל התוצאות שדווחו בספרות.

עיקרי התוצאות הינם דחיסה של התמונה (LENA) ברזולוציה של 256×256 לקצב של 0.61 ביט לא"ת, עם SNR של 32.6 dB ואיכות שחזור יותר מסבירה. דחיסה לקצב של 0.92 ביט לא"ת עם SNR של 35.2 dB ואיכות שחזור מושלמת. אותה תמונה ברזולוציה של 512×512 נדחסה בקצב של 0.41 ביט לא"ת עם SNR של 37.4 dB ואיכות שחזור מושלמת.

1.4 מבנה העבודה

פרק 2 מתאר את ההתמרות האורתוגונליות האופטימליות והתת-אופטימליות ואת התמרת DCT שמומשה בסכמה המוצעת ואת הסיבות לשמוש בה. פרק 3 מתאר קידוד וקטורי באמצעות ספר קוד ובאמצעות דיאגרמת שבכה, משווה ביניהם ונותן את המוטיבציה לשמוש בדיאגרמת שבכה בסכמה המוצעת. פרק 4 מתאר את סכמת הקידוד המוצעת ואת צורת הקידוד באמצעותה. פרק 5 מתאר את הבעיות הקיימות עקב אי התאמת תמונות אמיתיות למודל על פיו תוכננה סכמת הקידוד, ואת האמצעים המוצעים על ידנו לפתרון הבעיה. פרק 6 מתאר את ביצועי המערכת בהשוואה לביצועי מערכות קידוד אחרות. ולסיום, פרק 7 מסכם את העבודה ומתווה כווני המשך למחקר.

פרק 2 - התמרות אורתוגונליות.

2.1 מבוא.

השלב הראשון במימוש מערכת קידוד הינו הורדת הקורלציה באות ושיפור תכונותיו המאפשרות דחיסה יעילה. פעולה זו מבוצעת בסכמה המוצעת על ידנו באמצעות התמרה אורתוגונלית. פרק זה מתאר את התכונות הבסיסיות של ההתמרות אורתוגונליות (יוניטריות במקרה הקומפלקסי) ואת המדדים להשוואת יעילות ההתמרות בהקשרים של קידוד. בהמשך, מתוארת התמרת KLT אשר הינה ההתמרה האופטימלית, דהיינו היא גורמת לדה-קורלציה מלאה של האות. לטיוס מתוארות ההתמרות תת אופטימליות ובמיוחד התמרת DCT אשר נבחרה ליישום בסכמה המוצעת. לשם פשטות, הדיון בפרק זה יתבצע עבור התמרות חד-ממדיות ועבור אותות ממשיים בלבד. בעת הצורך נרחיב את הדיון למקרה החד-ממדי, הרחבה נוספת על האמור בפרק זה ניתן למצוא ב-[5, 11].

2.2 התמרות אורתוגונליות - רקע.

יהי \underline{u} וקטור באורך N מתוך מקור סטציונרי. התמרה אורתוגונלית של הוקטור \underline{u} הינה ייצוג הוקטור במרחב, אשר סדרת וקטורי הבסיס שלו $\{a_i\}$ ממלאת את תנאי האורתוגונליות הבא:

$$(2.1) \quad \underline{a}_j^T \cdot \underline{a}_i = \alpha \delta_{ij} = \begin{cases} \alpha, & i = j \\ 0, & i \neq j \end{cases}$$

כאשר δ_{ij} הינה הדלתה של Kronecker המרחב הנ"ל נקרא מרחב אורתוגונלי. כאשר $\alpha = 1$ סדרת וקטורי הבסיס $\{a_i\}$ הינה אורתונורמלית. ההתמרה הישירה וההתמרה ההפוכה מבוטאים ע"י הקשר הבא:

$$(2.2a) \quad v(k) = \sum_{n=0}^{N-1} u(n) a_k(n), \quad k=0,1,\dots,N-1$$

$$(2.2b) \quad u(n) = \sum_{k=0}^{N-1} v(k) b_n(k), \quad k=0,1,\dots,N-1$$

כאשר $b_n(k)$ נתון ע"י $b_n(k) = a_k(n)$ בצורה וקטורית מבוטאים ההתמרה הישירה והפוכה ע"י

$$(2.3a) \quad \underline{v} = A \cdot \underline{u}$$

$$(2.3b) \quad \underline{u} = B \cdot \underline{v} = A^{-1} \underline{v}$$

כאשר A היא מטריצה אשר שורותיה מורכבות מסדרת וקטורי הבסיס. במקרה זה תנאי האורתונורמליות נתון ע"י

$$(2.4) \quad B = A^{-1} = A^T$$

האורתוגונליות של וקטורי הבסיס המייצגים את הוקטור מאפשרת חישוב של כ"א מרכיבי הוקטור \underline{u} במרחב האורתוגונלי בצורה בלתי תלויה בשאר המרכיבים. בנוסף, כאשר רוצים לייצג את הוקטור ע"י חלק מרכיביו במרחב האורתוגונלי (למשל למטרות דחיסה), שגיאת השחזור הנמדדת ע"י ה-MSE בין הוקטור המקורי לוקטור המשוחזר, תהיה שווה לסכום האנרגיות של המרכיבים האורתוגונליים שלא נכללו בייצוג. תנאי הכרחי לייצוג הוקטור בלא שגיאה באמצעות סדרת הוקטורים האורתוגונלית הינו תנאי השלמות (Completeness) של הסדרה. אנו לא נרחיב את הדיון בנקודה זו, להרחבה ראה [5].

קיימות התמרות רבות אשר וקטורי הבסיס שלהן מקיימים את תנאי האורתוגונליות (הזונטריות) ואת תנאי השלמות, למשל: התמרה פורייה (DFT), התמרת קוסינוס (DCT) וכו'. נשאלת השאלה מהי ההתמרה היעילה ביותר לצורך הורדת הקורלציה, ומהם הקריטריונים ליעילות של התמרה לשימושים אלה. שלב ראשון בהשוואות ביצועים בין התמרות הינו הגדרת מודל האות עליו מבצעים את ההתמרה. מודל מקובל לייצוג תמונות הינו המודל המרקובי הדו-ממדי מסדר ראשון אשר מייצג בצורה סבירה תמונות, ומאפשר ביצוע חישובים בצורה פשוטה. כאמור, בדיון זה נתמקד במיקרה החד-ממדי. מודל מרקובי חד ממדי מסדר ראשון מאופיין ע"י פונקציית קורלציה מן הצורה $\sigma^2 \rho^{|k|}$ כאשר σ^2 הינה השונות של האות ו ρ הינו מקדם הקורלציה. בדיון שלהלן נניח מודל אות מרקובי מסדר ראשון. מקדמי הקורלציה בין המרכיבים השונים בוקטור \underline{u} מתוארים ע"י מטריצת הקואריאנס שלו.

$$(2.5) \quad R_{\underline{u}\underline{u}} = E[(\underline{u} - \bar{\underline{u}})(\underline{u} - \bar{\underline{u}})^T] = E[\underline{u}\underline{u}^T] - \bar{\underline{u}}\bar{\underline{u}}^T$$

כאשר $\bar{\underline{u}}$ הינו וקטור הממוצעים של \underline{u} . באופן דומה מטריצת הקואריאנס של הוקטור \underline{v} במרחב ההתמרה נתונה ע"י:

$$(2.6) \quad R_{\underline{v}\underline{v}} = E[\underline{v}\underline{v}^T] - \bar{\underline{v}}\bar{\underline{v}}^T$$

מתוך נוסחת ההתמרה (2.3) מתקבל הקשר הבא:

$$(2.7) \quad R_{\underline{v}\underline{v}} = A R_{\underline{u}\underline{u}} A^T$$

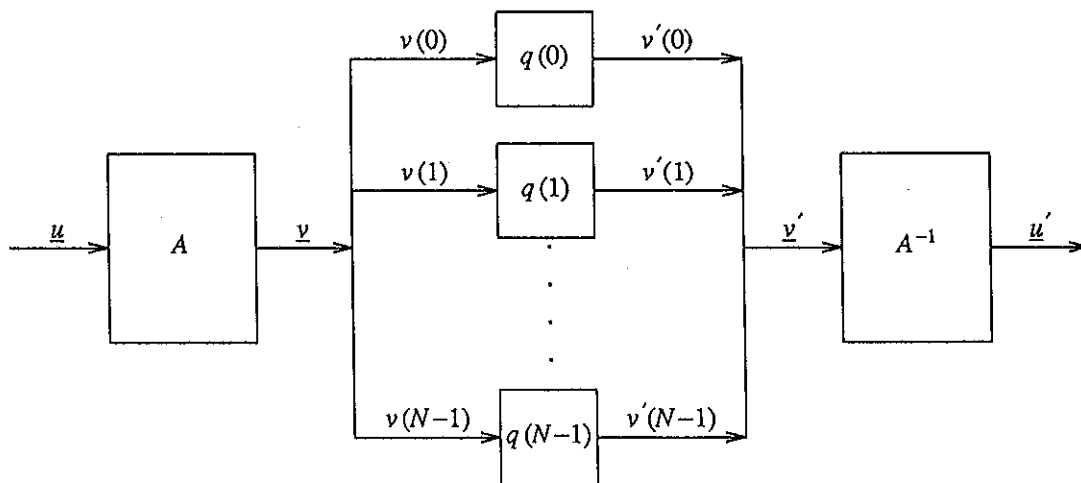
מידת הקורלציה בין רכיבי הוקטור מבטאת ע"י הערכים מחוץ לאלכסון במטריצת הקואריאנס שלו. לפיכך קריטריון לבדיקת מידת הדחיה - קורלציה של התמרה יכול להיות מידת ההפחתה בקורלציה בין רכיבי הוקטור עקב ביצוע ההתמרה. התמרה אופטימלית, כפי שנראה בסעיף 2.2 גורמת לדחיה - קורלציה מלאה של הוקטור, דהיינו מטריצת הקואריאנס של הוקטור במישור ההתמרה הינה אלכסונית. נגדיר את היעילות של ההתמרה האופטימלית כ- 100%. התמרת הזהות אינה גורמת כלל להפחתה בקורלציה בין רכיבי הוקטור ולפיכך נגדיר את יעילות ההתמרה הנ"ל כ- 0%. מדד יעילות הדחיה - קורלציה של התמרה יוגדר לפיכך [5]

$$(2.8) \quad \eta_c = \left[1 - \frac{\sum_{\substack{j,k=0 \\ j \neq k}}^{N-1} R_{\underline{v}\underline{v}}(j,k)}{\sum_{\substack{j,k=0 \\ j \neq k}}^{N-1} R_{\underline{u}\underline{u}}(j,k)} \right] \times 100 \%$$

תכונה חשובה אחרת המבטאת את טיב ההתמרה הינה המידה בה ההתמרה מרכזת את אנרגיית האות במספר קטן של מקדמים ("Energy Compaction Property"). אברי האלכסון של מטריצת הקוואריאנס של הוקטור מבטאים את האנרגיה של רכיביו. כאמור מודל האות המשמש אותנו לייצוג התמונה הינו מודל מרקובי, ולכן, עקב סטטיזונריות האות אברי האלכסון של מטריצת הקוואריאנס שלו יהיו זהים. עקב העובדה שהתמרות אורתוגונליות אינן גורמות לשינוי באנרגיית האות (פרסוול), סכום אברי האלכסון נשמר לאחר ביצוע ההתמרה. עובדה זו מאפשרת הגדרה של יעילות ההתמרה מבחינת "רכוז אנרגיה" אשר תבטא את הכמות היחסית של אנרגיה ב- M המקדמים הראשונים מתוך N מקדמי ההתמרה [5]. בהנחה שאנרגיית האות מנורמלת ליחידה נקבל:

$$(2.9) \quad \eta_E = \frac{\sum_{j=0}^{M-1} R_{v,v}(j,j)}{\sum_{j=0}^{N-1} R_{v,v}(j,j)} = \frac{1}{N} \sum_{j=0}^{M-1} R_{v,v}(j,j)$$

מדד שלישי להשוואת ביצועים בין התמרות שונות פותח ע"י Pearl et al [6]. מדד זה דומה במהותו לפונקציית הקצב - עוות הידועה מתורת האינפורמציה, והוא משווה את הקצב המינימלי R אליו ניתן להגיע בהתמרות השונות תוך שמירה על עוות קבוע D . הסכמה הבסיסית המשמשת להשוואה בין התמרות מתוארת בציור 2.1.



ציור 2.1 - סכמת התמרה בסיסית.

Fig. 2.1 - Basic Transform scheme.

בסכמה זו וקטור הכניסה u עובר התמרה A לקבלת וקטור v . כל רכיב של הוקטור v עובר קוונטיזציה סקלרית, הרכיבים ביציאות הקוונטיזרים יוצרים את הוקטור v' המותמר בהתמרה הפוכה לקבלת הוקטור המשוחזר u' . מדד ההשוואה בין ההתמרות הינו הפונקציה $R(A, D)$ המתארת את הקצב המינימלי אליו ניתן להגיע עבור עוות נתון D והתמרה נתונה A . נסמן ב- D_i את העוות בענף ה- i של הסכמה וב- D^* את מרחב ה- N יות- $D_i, i=1,2,\dots,N-1$ כך שהעוות הכולל $\sum D_i = D$ ז"א:

$$(2.10) \quad D^* = \{ (D_0, D_1, \dots, D_{N-1}) : \sum_{i=0}^{N-1} D_i = D \}$$

אזי עבור אות בעל מודל גאוסי מרקובי מדד הביצועים הוא [6]:

$$(2.11) \quad R(A, D) = \min_{(D_0, \dots, D_{N-1}) \in D^*} \frac{1}{N} \sum_{i=0}^{N-1} \max\{0, \frac{1}{2} \log \frac{\sigma_i^2}{D_i}\}$$

כאשר σ_i^2 הינו הוריאנס של מילת הכניסה v_i ייצוג פרמטרי של $R(A, D)$ נתון במקרה זה ע"י:

$$(2.12a) \quad R(A, D) = \frac{1}{N} \sum_{i=0}^{N-1} \max\{0, \frac{1}{2} \log \frac{\sigma_i^2}{i} \theta\}$$

$$(2.12b) \quad D = \frac{1}{N} \sum_{i=0}^{N-1} \min\{\theta, \sigma_i^2\}$$

בהמשך הפרק נשווה את ביצועי ההתמרות השונות על סמך המדדים שהוצגו.

2.3 התמרת Karhunen Loeve (KLT).

בסעיף הקודם סקרנו מדדים שונים להערכת איכות התמרות אורתוגונליות. המדדים הללו פרטו באיזו מידה התמרה גורמת לדה-קורלציה של האות, באיזה מידה היא מרכזת את האנרגיה שלו במספר מועט של מקדמים ומהו הקצב המינימלי אליו ניתן להגיע בסכמת קידוד התמרה בסיסית (ציור 2.1) עבור עוות נתון. בסעיף זה נתאר את התמרת Karhunen Loeve (KLT), אשר ביצועיה, כפי שמתקפים באמצעות המדדים השונים שתארנו, הינם הטובים ביותר מבין כל ההתמרות האורתוגונליות. בהמשך הסעיף נתאר את הבניה של התמרת KL, נתאר את עדיפות ביצועיה יחסית להתמרות אורתוגונליות אחרות ואת חסרונותיה אשר מונעים ישום שלה בסכמות מעשיות.

כאמור, אנו מניחים שהאותות בהם אנו עוסקים מקיימים את המודל המרקובי. יהי \underline{u} וקטור מדגם של המקור $R_{\underline{u}\underline{u}}$ מטריצת הקואריאנס שלו. יהיו $\{\phi_i\}_{i=0}^{N-1}$ ו $\{\lambda_i\}_{i=0}^{N-1}$ הוקטורים העצמיים המנורמלים והערכים העצמיים של $R_{\underline{u}\underline{u}}$ בהתאמה. כלומר, ϕ_i מקיימים את הקשר הבא:

$$(2.13) \quad R_{\underline{u}\underline{u}} \phi_i = \lambda_i \phi_i$$

כאשר ϕ_i מקיימים את תנאי האורתונורמליות

$$(2.14) \quad \phi_i^T \phi_j = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

המטריצה Φ אשר עמודותיה הינם הוקטורים העצמיים ϕ_i הינה מטריצה אורתוגונלית דהיינו:

$$(2.15) \quad \Phi^{-1} = \Phi^T$$

כמו כן Φ מלכסנת את מטריצת הקואריאנס של \underline{u} ע"י:

$$(2.16) \quad \Phi^{-1} R_{\underline{u}\underline{u}} \Phi = \begin{bmatrix} \lambda_0 & & \\ & \ddots & \\ & & \lambda_{N-1} \end{bmatrix} = \Lambda$$

התמרת KL מזגדרת ע"י המטריצה

$$(2.17) \quad A \triangleq \Phi^T$$

וקטור \underline{y} במישור ההתמרה מתקבל ע"י

$$(2.18) \quad \underline{y} = A\underline{u} = \Phi^T \underline{u}$$

מטריצת הקוואריאנס של \underline{y} הינה אלכסונית ונתונה ע"י

$$(2.19) \quad R_{\underline{y}\underline{y}} = \Phi^{-1} R_{\underline{u}\underline{u}} \Phi = \Lambda = \text{diag} \{ \lambda_i \}_{i=0}^{N-1}$$

תוצאה נוספת המתקבלת מהביטוי האחרון הינה העובדה שהאנרגיה של האיבר ה- i במישור ההתמרה נתונה ע"י הערך העצמי המתאים לוקטור העצמי ϕ_i דהיינו

$$(2.20) \quad \sigma_{v_i}^2 \triangleq E[v_i^2] = \lambda_i$$

ביצועיה של התמרת KL עדיפים על פני הביצועים של כל ההתמרות האורתוגונליות האחרות, והם מהווים למעשה חסם לביצועים של התמרות אורגנליות. מסיבה זו התמרת KL נקראת ההתמרה האופטימלית, האופטימליות של התמרת KL מתבטאת בנקודות הבאות:

א. התמרת KL מבצעת דה - קורלציה מושלמת של וקטור המקור, דהיינו מדד יעילות הדה - קורלציה שלה η_c (2.8) שווה ל 100%. אם וקטור המקור לקוח מאות גאוסי אזי רכיבי הוקטור במישור ההתמרה יהיו בנוסף גם בלתי תלויים סטטיסטית.

ב. מבין כל ההתמרות האורתוגונליות, התמרת KL מביאה למינימום את השגיאה הרבועית הממוצעת (MSE) בין וקטור המקור באורך N לבין וקטור השחזור המתקבל ממספר מוגבל $M \leq N$ של רכיבים [11, p 544]. משמעות התכונה הנ"ל הינה שהתמרת KL מביאה למקסימום את המדד η_E המתואר בנוסחה (2.9).

ג. מבין כל ההתמרות האורתוגונליות, התמרת KL מביאה למינימום את הפונקציה $R(A, D)$ המתוארת בנוסחה (2.12) [6, 21].

בניגוד לאופטימליות של התמרת KL כפי שתוארה לעיל, סובלת התמרה זו ממספר חסרונות בולטים.

א. אין אלגוריתם מהיר לממושה.

ב. היא דורשת חשוב של מטריצת הקוואריאנס (טנזור הקוואריאנס במקרה הדו-ממדי), חשוב זה דורש איסוף של כמות רבה של נתונים, דהיינו נדרשת כמות גדולה של זכרון.

ג. התמרת KL דורשת חשוב של הוקטורים העצמיים (מטריצות עצמיות במקרה הדו - ממדי) והערכים העצמיים של מטריצת הקוואריאנס, פעולות הדורשות עומס חישובי רב.

ד. התמרת KL אינה קבועה אלא תלויה בסטטיסטיקת האות, עובדה הדורשת חשוב מחדש של ההתמרה עבור על שינוי בפילוג האות.

החסרונות שתוארו לעיל מונעים שמוש בהתמרת KL ביישומים מעשיים, כתחליף משתמשים בהתמרות תת-אופטימליות שיתוארו בסעיף הבא.

2.4 התמרות תת - אופטימליות.

עקב הסיבוכיות הכרוכה במימוש התמרת KL שתוארה בסעיף הקודם, היא משמשת אותנו רק לצורך קביעת חסמים לביצועים של התמרות תת-אופטימליות כגון:

(DWT) Discrete Walsh Hadamard Transform

(DFT) Discrete Fourier Transform

(DCT) Discrete Cosine Transform

ההתמרות התת - אופטימליות הנ"ל אינן מלכסנות את מטריצת הקוואריאנס, דהיינו רכיבי הוקטור במישור ההתמרה אינם חסרי קורלציה. למרות התת - אופטימליות שבמתמרות הללו השמוש בהן הוא נפוץ עקב פשטות המימוש שלהן ועקב אי תלותן בסטטיסטיקת האות, עובדה המאפשרת חשוב מראש של סדרת וקטורי הבסיס.

כפי שנראה בהמשך הפרק עבור סיגל מרקובי מסדר ראשון, הביצועים של התמרות DFT ו-DCT שואפים לביצועים של התמרת KL עם הגדלת אורך הוקטור המקודד.

התמרת DWHT הינה ההתמרה הפשוטה ביותר למומש מבין כל ההתמרות האורתוגונליות המקובלות, כיוון שוקטורי הבסיס שלה מכילים רכיבים שערבם +1 או -1 בלבד. עובדה זו מאפשרת ביצוע של ההתמרה ע"י פעולות חבור וחיסור בלבד.

התמרת פורייה (DFT) הינה ההתמרה הידועה ביותר. וקטורי הבסיס שלה הינם פונקציות אקספוננט קומפלקסי בתדרים שונים. ניתן להראות [11, P 554] שהביצועים של התמרת DFT עבור סיגל מרקובי מסדר ראשון שואפים אסימפטוטית לביצועי התמרת KL עם הגדלת אורך הבלוק המותמר לאינסוף. המימוש של התמרת DFT הינו באמצעות אלגוריתם ה-FFT המאפשר חישוב ה-DFT של וקטור באורך N ברמת סיבוכיות של $O(N \log(N))$ פעולות. התמרת DCT הינה ההתמרה היעילה ביותר מבין ההתמרות האורתוגונליות הנדונות בעבודה זו והיא מפורטת בסעיף הבא.

2.4.1 התמרת (DCT) COSINE

התמרת (DCT) Cosine הוגדרה לראשונה ע"י [21] Ahamed et al. ההתמרה והתמרה ההפוכה

נתונות ע"י:

$$v_c(k) = \alpha(k) \sum_{n=0}^{N-1} u(n) \cos \frac{2n+1}{2N} k \pi, \quad k=0, \dots, N-1$$

$$u(n) = \sum_{k=0}^{N-1} \alpha(k) v_c(k) \cos \frac{2n+1}{2N} k \pi, \quad n=0, \dots, N-1$$

כאשר:

$$\alpha(k) = \begin{cases} \frac{1}{\sqrt{N}}, & k=0 \\ \frac{2}{\sqrt{N}}, & k=1, \dots, N-1 \end{cases}$$

התמדת Cosine הינה התמרה קבועה ואינה תלויה בסטטיסטיקת האות. כמו שנראה בפרק 4 קיים אלגוריתם מחיר המבוסס על אלגוריתם ה-FFT לחישוב מקדמי ה-DCT. כדוגמת התמרת ה-DFT הזו-ממדית התמדת Cosine הזו-ממדית הינה התמרה ספרבילית, לפיכך היא נתנת לחשוב ע"י הפעלת DCT חד-ממדי על השורות של הבלוק הזו-ממדי, ולאחר מכן הפעלה של ה-DCT הזו ממדי על עמודות הבלוק שנוצר בפעולה הראשונה.

בסעיף הבא מתוארת העליונות של DCT ע"פ התמרות תת-אופטימליות אחרות. כן מתוארות התכונות האסימפטוטיות של DCT אשר נותנות את המוטיבציה לשימוש בה בסכמות קידוד התמרה.

2.5 המוטיבציה לשמוש ב-DCT בסכמה המוצעת.

התמרת KL שהוצגה בסעיף 2.2 הינה ההתמרה האופטימלית, ולפיכך ביצועיה ישמשו אותנו כחסם לביצועי ההתמרות התת אופטימליות. המוטיבציה של Ahamed et al [21] להגדרת ה DCT הייתה העובדה שסדרת וקטורי הבסיס של DCT, המורכבת מפונקציות \cos , מהווה קרוב טוב לוקטורים העצמיים של מטריצות מסוג Toeplitz. מטריצות מסוג זה מתארות את מטריצת הקוואריאנס של אותות המקיימים את המודל המרקובי, אשר משמשים מודל טוב לייצוג תמונות אמיתיות.

Yemini & Pearl [22] הוכיחו שהביצועים של התמרת DCT עבור אותות מרקוביים הינם אופטימליים כאשר גודל הוקטור המותמר שואף לאינסוף. בנוסף לתכונה הנ"ל של התאמה טובה, עבור סיגנלים מרקוביים, להתמרת KL בהסתמך על מדדי היעילות שתוארו בסעיף 2.1 התמרת DCT הינה היעילה ביותר מבין כל ההתמרות התת - אופטימליות שנדונו בעבודה זו.

טבלה 2.1 מתארת את מדד יעילות הדה - קורלציה η_c (נוסחה 2.8), של ההתמרות השונות עבור אות מרקובי בעל מקדם קורלציה $\rho = 0.91$ המיוצג ע"י וקטור באורך 8.

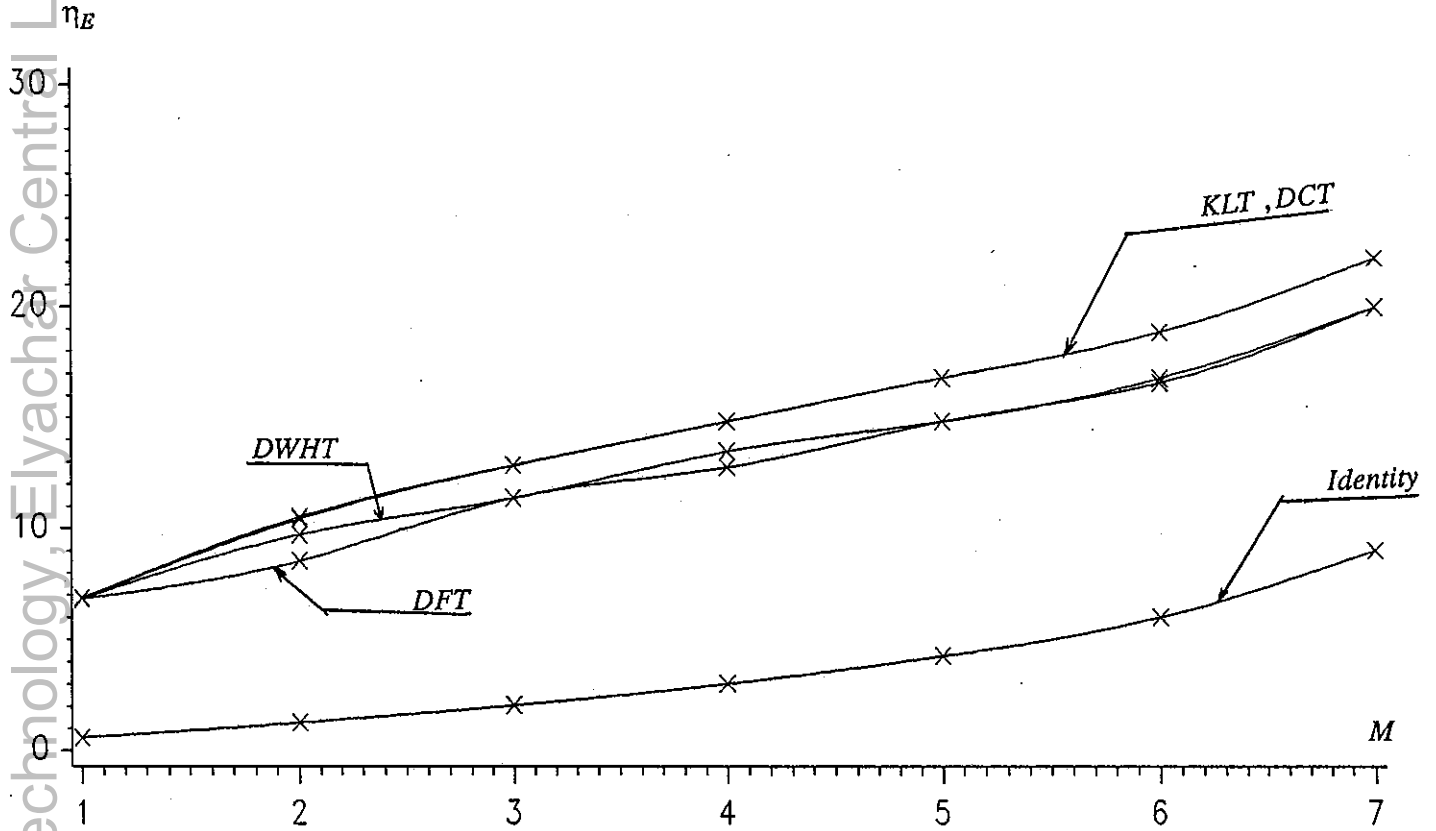
התמרה	η_c
KLT	100%
DCT	98.05%
DFT	89.48%
DWHT	94.86%
זהות	0%

טבלה 2.1 - מדד יעילות הדה-קורלציה η_c של התמרות שונות (מתוך [5]).

Table 2.1 - Decorrelation efficiency factor η_c for various transforms (from [5]).

ציור 2.2 מתאר את המדד η_E (נוסחה 2.9) כתלות במספר המקדמים בשחזור של וקטור באורך 8

שנלקח מתוך מקור מרקובי בעל מקדם קורלציה $\rho = 0.91$. הנתונים לציור נלקחו מתוך [5 ch 3].

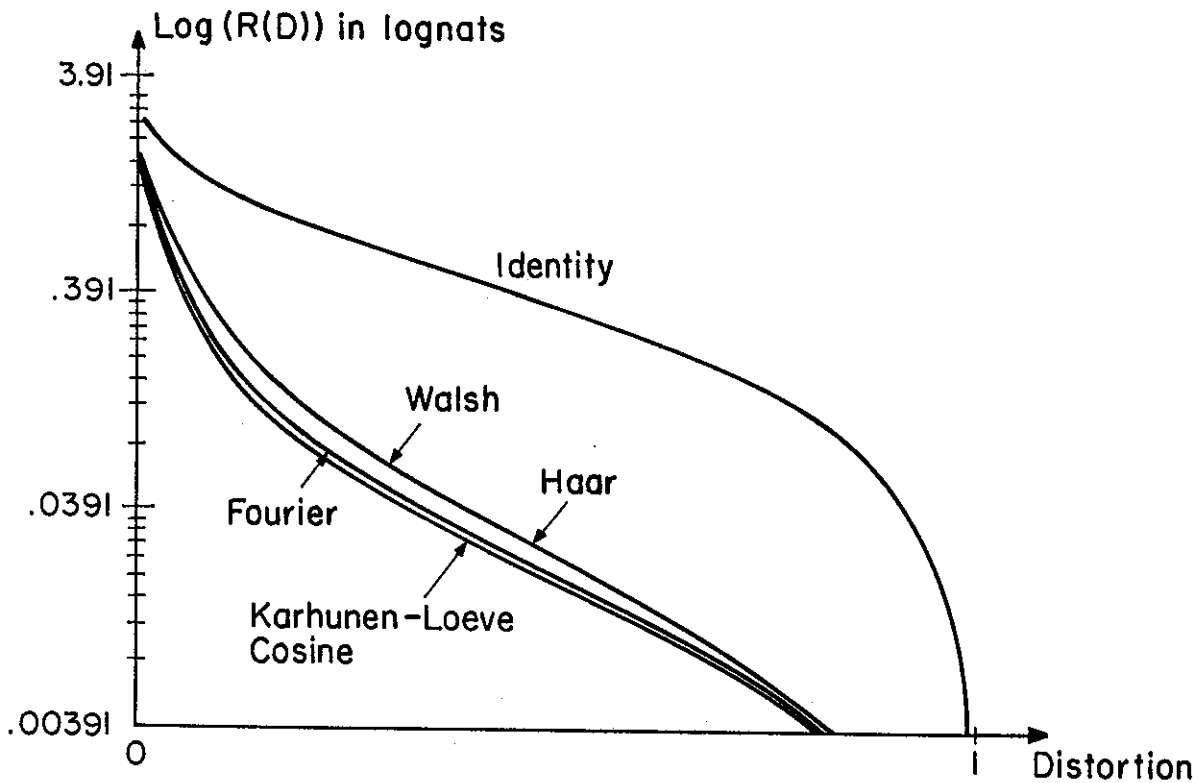


ציור 2.2 - η_E כפונקציה של מספר מקדמי השחזור עבור התמרות שונות.

Fig 2.2 - η_E vs the number of retained coefficients for various transforms.

ציור 2.3 מתאר את פונקציית $R(A, D)$ (נוסחה 2.12) של התמרות KLT, DCT והתמרת הזהות,

עבור וקטור מדגם באורך 16 שנלקח מתוך מקור מרקובי מסדר ראשון בעל מקדם קורלציה $\rho = 0.9$.



ציור 2.3 - קצב כתלות בעוות עבור התמרות שונות (מתוך [21]).

Fig 2.3 - Rate vs Distortion for various transforms (from [21]).

בכל שלושת המדדים שתוארו לעיל רואים את העדיפות של התמרת DCT ע"פ כל שאר ההתמרות התת - אופטימליות. בנוסף, רואים שהביצועים של התמרת DCT קרובים לביצועי KLT. בכל המדדים שנבדקו, ביצועי התמרת הזהות - דהיינו קידוד במישור התמונה, גרועים בהרבה מהביצועים במישור ההתמרה.

התמרת Cosine מביאה להפחתה של תופעת Gibbs אשר מתגלית בהתמרה DFT, ונוצרת עקב אי רציפיות בין קצוות בלוק ההתמרה [12 p 158]. הסיבוכיות של התמרת DCT משתווה לסיבוכיות של DFT. בפרק 4 תואר אלגוריתם לחשב ה-DCT של וקטור באורך N באמצעות אלגוריתם ה-FFT באורך N .

לאור העדיפות של התמרת DCT ע"פ ההתמרות התת-אופטימליות האחרות, עקב הקרוב הטוב שהיא מהווה להתמרת KL עבור סיגנלים מקרוביים ועקב פשטות המימוש שלה, בחרנו לממש את סכמת הקידוד באמצעותה.

פרק 3 - קוונטיזציה וקטורית

3.1 הקדמה

כפי שציינו בפרק המבוא, דחיסת תמונות מתבצעת בד"כ בשני שלבים. בשלב הראשון מבצעים דה-קורלציה של איברי התמונה (בסכמה המוצעת על ידנו הדבר מתבצע באמצעות התמרה אורתוגונלית), ואילו בשלב השני מבצעים חלוקת סיביות בין האברים חסרי הקורלציה, ומבצעים קוונטיזציה. אחת התוצאות הבסיסיות של תורת האינפורמציה הינה שניתן להשיג ביצועי דחיסה טובים יותר כאשר מבצעים קוונטיזציה של וקטור מקדמים במקום קוונטיזציה של כל מקדם בנפרד, גם כאשר המקדמים הינם בלתי תלויים סטטיסטית [13]. פרק זה דן בקוונטיזציה הוקטורית. סעיף 3.2 מתאר את עקרונות הקוונטיזציה הוקטורית באמצעות ספר קוד. סעיף 3.3 מתאר את דיאגרמת העבכה. ולסיום, סעיף 3.4 מפרט את המוטיבציה לשמוש בדיאגרמת שבכה כקוונטייזר וקטורי בסכמה המוצעת על ידנו.

3.2 קוונטיזציה וקטורית באמצעות ספר קוד

קוונטייזר וקטורי מוגדר ע"י המיפוי $\{Q: R^N \rightarrow V\}$ כאשר $V = \{v_k\}_{k=1}^K$ הינה קבוצה של K וקטורים במימד N הקרויה מילון או ספר - קוד (Code-Book). כדי להגדיר את המיפוי יש צורך לבצע חלוקה של המרחב R^N ל- K אזורים $\{S_k\}_{k=1}^K$, אשר מקיימים

$$(3.1) \quad \bigcup_{k=1}^K S_k = R^N ; \quad \bigcap_{\substack{i,j \\ i \neq j}} S_i S_j = 0$$

אזורים אלה נקראים אזורי החלטה. הגדרת המיפוי תהיה

$$(3.2) \quad Q(u) = v_i \quad \text{if} \quad u \in S_i$$

כאשר u הינו וקטור דגמי מקור במימד N . הקצב המתקבל בקוונטייזר וקטורי כנ"ל בסיביות לדגם מקור נתון ע"י:

$$(3.3) \quad R = \frac{\log_2 K}{N} \quad [\text{bits/sample}]$$

המטרה של קוונטיזציה וקטורית כפי שתוארה לעיל הינה יצירה של סט וקטורי ייצוג $\{v_k\}_{k=1}^K$ אשר יהווה את הסט "הטוב ביותר" עבור קצב נתון R . לשם כך יש להגדיר מדד לביצועי הקוונטייזר. אם נגדיר את $d(u, v)$ כמידת עוות בין וקטור המקור u לוקטור הקוד (וקטור השחזור) v , אזי מדד לביצועי הקוונטייזר יכול להיות העוות הממוצע D המוגדר ע"י:

$$(3.4) \quad D \triangleq E [d(u, v)]$$

כאשר המקור המקודד הינו סטציונרי וארגודי העוות הממוצע D נתון ע"י:

$$(3.5) \quad D = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=0}^{n-1} d(u_j, v_j)$$

כאשר $\{\underline{u}_j\}_{j=0}^{\infty}$ הינם וקטורים המיוצרים ע"י סדרת דגם של המקור. בהתאם למדד הנ"ל קוונטיזר יחשב כטוב אם העוות הממוצע D המתאים הינו קטן.

הבעיה העיקרית הנובעת משמוש בקוונטיזרים וקטוריים הינה התכנון של קוונטיזר אופטימלי, דהיינו, אותו קוונטיזר אשר יתן עוות ממוצע מינימלי. שני תנאים הכרחיים (אולם לא מספיקים) לאופטימליות של קוונטיזר וקטורי הינם התנאים הבאים:

א. בהנתן המילון $V = \{\underline{v}_k\}_{k=1}^K$, איזורי החלטה $\{S_k\}_{k=1}^K$ חייבים לקיים את התנאי

$$(3.6) \quad S_k = \{\underline{u} \mid d(\underline{u}, \underline{v}_k) \leq d(\underline{u}, \underline{v}_j), \quad \forall j \neq k\}$$

דהיינו, הקבוצה S_k מכילה את כל הוקטורים \underline{u} במרחב אשר קרובים ל \underline{v}_k יותר מאשר לכל וקטור קוד אחר. תנאי זה נקרא תנאי השכן הקרוב. וקטורים גבוליים ב V משוייכים בצורה שרירותית כלשהי.

ב. בהנתן קבוצת איזורי החלטה $\{S_k\}_{k=1}^K$ אזי וקטורי המלון $\{\underline{v}_k\}_{k=1}^K$ מקיימים את התנאי

$$(3.7) \quad \underline{v}_k = \text{cent}\{S_k\} = \arg \left\{ \min_{\underline{v} \in S_k} E[d(\underline{u}, \underline{v})] \mid \underline{u} \in S_k \right\}$$

כלומר הוקטור \underline{v}_k הינו אותה נקודה בתוך האזור S_k עבורה מקבלים עוות מינימלי עבור וקטורי המקור $\underline{u} \in S_k$. תנאי זה נקרא תנאי הצנטרואיד. קיימת הוכחה [14] ששני התנאים שפורטו לעיל הינם תנאים מספיקים לקבלת אופטימום מקומי. מדד העוות הנפוץ ביותר להערכת ביצועי קוונטיזרים וקטוריים הינו מדד העוות הרבועי:

$$(3.8) \quad d(\underline{u}, \underline{v}) = \|\underline{u} - \underline{v}\|^2$$

הפופולריות של המדד הנ"ל נובעת ממספר סיבות:

- א. האופטימיזציה של העוות הרבועי מובילה במקרים רבים לפתרון מערכת משוואות לינאריות.
- ב. קיימים חסמים תיאורטיים מתורת האינפורמציה (פונקציית הקצב - עוות) עבור מדד העוות הריבועי, אשר ניתן באמצעותם להעריך ביצועים של אלגוריתמים.
- ג. לעוות הריבועי הממוצע יש מובן פיזיקלי של הספק שגיאת הקוונטיזציה, ועל כן הוא קשור ישירות ליחס אות לרעש (SNR).

עבור עוות רבועי ניתן להראות שתנאי הצנטרואיד הוא:

$$(3.9) \quad \underline{v}_k = E[\underline{u} \mid \underline{u} \in S_k]$$

במקרה זה, העוות הממוצע D של הקוונטיזר נתון ע"י:

$$(3.10) \quad D \triangleq E[d(\underline{u}, \underline{v})] = \sum_{k=1}^K \text{prob}\{\underline{u} \in S_k\} \cdot \int_{\underline{u} \in S_k} d(\underline{u}, \underline{v}_k) p(\underline{u}) d\underline{u}$$

כאשר:

$$\text{prob}\{\underline{u} \in S_k\} \quad - \quad \text{הינה ההסתברות ש } \underline{u} \in S_k$$

הינה פונקציית צפיפות ההסתברות של הוקטור \underline{u} - $p(\underline{u})$

על בסיס שני התנאים שהוצגו לעיל הוצע ב [14] אלגוריתם לתכנון קוונטיזר וקטורי הידוע בשם LBG (ע"ש מחברי המאמר). במימוש האלגוריתם הנ"ל נעשית הבחנה בין שני מקרים:
 א. סטטיסטיקת המקור $p(\underline{u})$ ידועה.
 ב. סטטיסטיקת המקור אינה ידועה.

להלן נתאר את האלגוריתם עבור המקרה השני שהוא המקרה הרלוונטי בד"כ בדחיסת תמונות עקב העובדה שפונקציית הפילוג של תמונות אינה ידועה. במקרה זה נתונה סדרת וקטורים מהמקור הנקראת סדרת אמון (Training Sequence). האלגוריתם משתמש בוקטורים הללו לבניית הקוונטיזר.

אלגוריתם LBG [14]

(0) נתונים:

- K - גודל המילון.
- $\epsilon > 0$ - סף עוות.
- $V^{(0)}$ - מילון התחלתי בגודל K .
- $\{\underline{u}_i\}_{i=1}^L$ - סדרת אימון של וקטורים במימד N .

קובעים:

- $m = 0$ - אינדקס צעד האיטרציה.
- $D_{-1} = \infty$ - עוות ממוצע התחלתי.

(1) בהנתן מילון $V^{(m)}$ חלק את וקטורי סדרת הלמוד $\{\underline{u}_i\}_{i=1}^L$ ל- K תת קבוצות $\{S_k\}_{k=1}^K$ ע"י מיפוי כל וקטור מסדרת הלימוד לוקטור במילון אליו הוא קרוב ביותר. כלומר:

$$\underline{u}_i \in S_k^{(m)} \text{ if } d(\underline{u}_i, \underline{v}_k^{(m)}) < d(\underline{u}_i, \underline{v}_j^{(m)}) \quad \forall k \neq j$$

חשב את העוות הממוצע באיטרציה זו ע"י:

$$D^{(m)} = \frac{1}{L} \sum_{k=1}^K \sum_{\underline{u}_j \in S_k^{(m)}} d(\underline{u}_j, \underline{v}_k^{(m)})$$

(2) אם $\frac{D^{(m-1)} - D^{(m)}}{D^{(m)}} \leq \epsilon$ עצור. $V^{(m)}$ הוא המילון הסופי.

אחרת, המשך ל-(3).

(3) קבע את המילון $V^{(m+1)}$ מתוך $S_k^{(m)}$ לפי כלל הצנטרואיד (נוסחה 3.7).

$$\underline{v}_k^{(m+1)} = \text{cent}(S_k^{(m)}), \quad k = 1, 2, \dots, K$$

(4) קבע $m = m + 1$ וחזור ל-1.

אלגוריתם LBG שתואר לעיל מבטיח התכנסות למינימום מקומי בלבד של מדד העוות, ולפיכך טיב המלון המתקבל תלוי בבחירת המלון ההתחלתי $v^{(0)}$. נבחנו בספרות [15] מספר שיטות לבחירת מלון התחלתי:

- א. בחירת מלון התחלתי בגודל K באקראי מתוך סדרת הלמוד.
 - ב. בחירת K הוקטורים המרוחקים ביותר זה מזה בסדרת הלימוד.
 - ג. מילון מכפלה - המילון ההתחלתי הוא מכפלה קרטזית של מילונים במימדים נמוכים יותר (למשל K מילונים סקלדיים).
 - ד. שיטת הפיצול (splitting Method) - בנית מילון התחלתי בגודל K מתוך המילון האופטימלי בגודל $\frac{K}{2}$ ע"י פיצול כל וקטור v_i כמתואר בהמשך.
- בנית המילון ההתחלתי בשיטת הפיצול נתנה באופן אמפירי תוצאות טובות, כמו כן בנית המילון בשיטה זו מאפשרת קבלת מילונים אופטימליים במימדים שהם חזקות של 2 עד לגודל K .
- להלן תאור אלגוריתם LBG תוך שמוש בשיטת הפיצול למציאת המילון ההתחלתי.

אלגוריתם LBG עם שיטת הפיצול

- (0) אתחול: קבע $M=1$ והגדר את V_0^M כוקטור v_1 הנתון ע"י $v_1 = \text{cent} \{u_m\}_{l=1}^L$.
- (1) בחינת $V_0^M = \{v_m\}_{m=1}^M$ פצל כל וקטור קוד v_m לשני וקטורים $v_m(1 \pm \delta)$, כאשר δ הינו וקטור פרטורבציה קבוע, לקבלת מילון התחלתי בגודל $2M$, הצב $M = 2M$.
- (2) מצא באמצעות אלגוריתם LBG שתואר לעיל את המילון האופטימלי בגודל M .
- (3) אם $M = K$ המילון שנמצא בצעד (2) הינו ספר הקוד הנדרש ; עצור.
אחרת חזור ל (1).

3.3 קוונטיזציה וקטוריים הממומשים באמצעות דיאגרמת שבכה

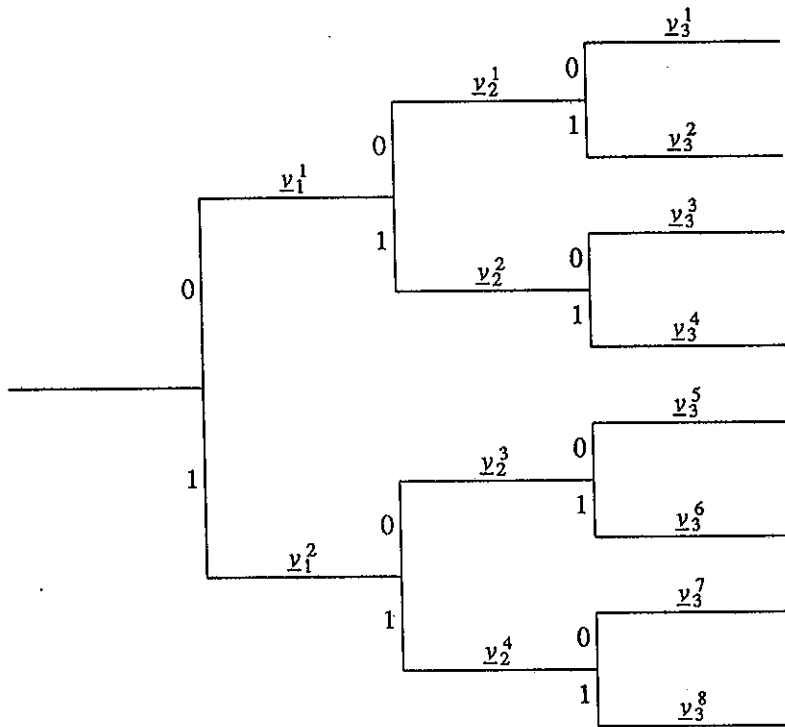
הקוונטיזציה הוקטוריים הממומש באמצעות ספר קוד, שתואר בסעיף הקודם, הינו קוונטיזציה אופטימלי, דהיינו ניתן להתקרב באמצעותו כרצוננו לפונקציית הקצב-עוות עם עלית גודל הוקטור המקודד. כמו כן, כפי שראינו, קיים אלגוריתם מעשי לבנית ספר קוד (אלגוריתם LBG) אשר מבחינה אמפירית הוכח כבעל ביצועים טובים. אולם, קיימת עדיין בעיה בשימוש מעשי בספר קוד כקוונטיזציה וקטוריים, והיא הסיבוכיות הגבוהה עמה יש להתמודד. סיבוכיות זו מתבטאת הן בכמות הזכרון הרבה הדרושה לשמירת ספר הקוד במקודד ובמפענח, והן בכמות החישובים הנדרשת במהלך חיפוש הוקטור המייצג. למשל עבור וקטור באורך $N = 256$ אלמנטים (מתאים לגודל בלוק התמרה של 16×16 א"ת) וקצב נדרש של $R = 1$ bits/pel, כמות הזכרון הנדרשת היא של $2^{NR} = 2^{256} \approx 10^{76}$ אלמנטים. כמות החישובים לדגם במקרה זה תהיה גם היא מאותו סדר גודל. כמובן שטיפול בכמות זכרון וחישובים רבה כזו אינו מעשי. לפיכך נשאלת השאלה האם ניתן לממש קוונטיזציה וקטוריים אופטימליים שלא באמצעות ספר קוד, אשר ניתן יהיה ליישם באופן

מעשי.

קוונטיזרים כאלה הינם קוונטיזרים הממומשים באמצעות דיאגרמת שבכה. המבנה של דיאגרמת שבכה יוצר תלות בין וקטורי הקוד, עובדה המאפשרת חסכון בכמות הזכרון והחישובים. בתחילה נתאר קוונטיזרים וקטוריים הממומשים באמצעות דיאגרמת עץ ובאמצעותם נסביר את המבנה של דיאגרמת השבכה.

3.3.1 קוונטיזרים וקטוריים במבנה עץ

דוגמה לקוונטיזר וקטורי במבנה עץ מתוארת בציור 3.1.



ציור 3.1 - קוונטיזר וקטורי במבנה עץ.

Fig. 3.1 - Vector Quantizer implemented as a Tree .

העץ מתחיל מצומת יחיד ממנו מתפצלים q ענפים ל- q צמתים בדרגה הבאה. מכל צומת בדרגה הבאה מתפצלים q ענפים ל- q^2 צמתים, וכן הלאה מדרגה לדרגה עד לדרגה L המציינת את קצה העץ והמכילה q^L צמתים. כל ענף בדרגה l של העץ מאוכלס בוקטור קוד v_l המורכב מ n_l מילות קוד (וקטור קוד שונה לכל ענף). בדוגמה המתוארת בציור 3.1 מתואר עץ בעל $q=2$ ו $L=3$.

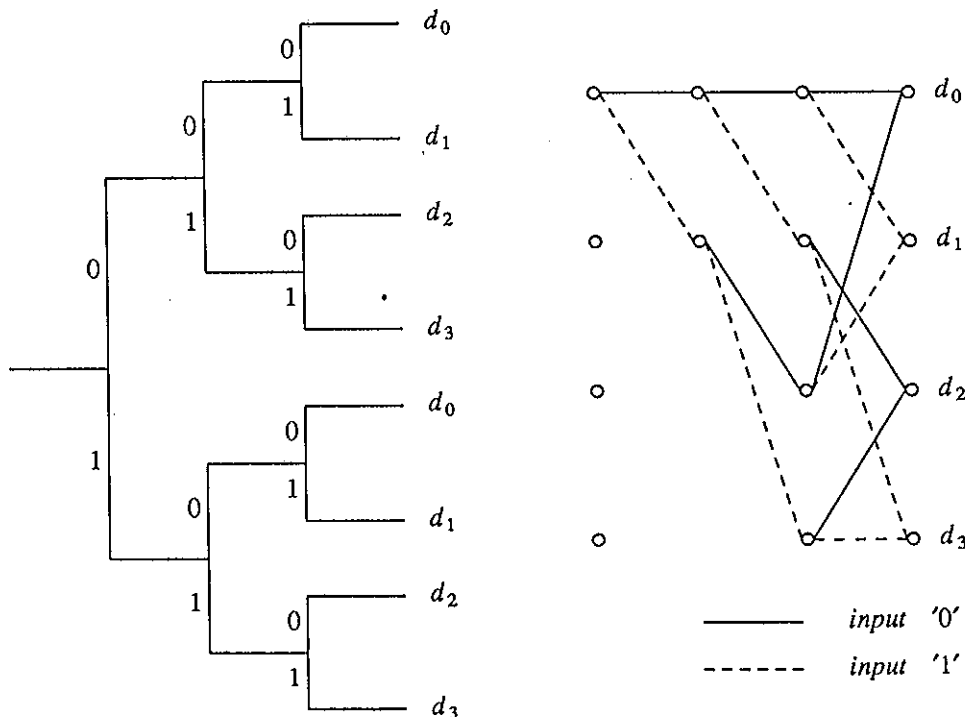
בתהליך הקידוד מחפשים את המסלול בעץ אשר עבורו העוות בין וקטור המקור u לבין סדרת הוקטורים $\{v_l\}_{l=1}^L$ המאכלסת את ענפי המסלול הנבחר והמהווה את וקטור השחזור \hat{u} , הוא מינימלי. סדרת האינדקסים ה- q -ארית משודרת דרך ערוץ (אשר נחשב כחסר רעש) למקלט. המקלט אשר מכיל דיאגרמת עץ זהה לזו הנמצאת במשדר, משחזר מהמסלול הנקלט את וקטור השחזור \hat{u} .

קוונטייזר וקטורי במבנה עץ הינו אופטימלי, דהיינו ביצועיו שואפים לפונקציית הקצב-עוות עם הגדלת אורך הבלוק המקודד [13].

השימוש בדיאגרמת עץ כפי שתואר לעיל מאפשר הקטנה במספר החישובים הנדרש, וזאת מהסיבה שאם שני וקטורי קוד מאכלסים מסלול משותף המתחיל בראשית העץ עד לנקודה כלשהי בעץ, ניתן לזוּעב את העוות המתאים לקטע המשותף הנ"ל רק פעם אחת. מספר החישובים בעץ הינו מסדר גודל $O(NR)$. בהשוואה לסדר גודל של $O(2^{NR})$ הקיים בספר קוד. כאשר, R הינו הקצב הנדרש, ו- N הינו אורך הוקטור המקודד. בניגוד למספר החישובים, כמות הזכרון בעץ דומה לזו הקיימת בספר קוד והינה מסדר גודל של $O(2^{NR})$, דבר המקשה על מימוש מעשי של קוונטייזרים וקטוריים יעילים במבנה העץ.

3.3.2 קוונטייזרים וקטוריים במבנה שבכה.

כפי שראינו קוונטייזר במבנה עץ אינו ניתן למימוש מעשי, עקב התלות האקספוננציאלית של כמות הזכרון הנדרשת באורך הבלוק המקודד. כדי להפחית את כמות הזכרון יש צורך לייצור תלות נוספת במבנה הקוונטייזר אשר תאפשר הפחתה בכמות הזכרון. דיאגרמת השבכה המוצגת בצירור 3.3a מתארת קוונטייזר כזה. דיאגרמת השבכה מתקבלת מדיאגרמת העץ ע"י אחוד כל הצמתים אשר סדרת האינדקסים (המסלול) המוליכה אליהם זהה $K-1$ דרגות ויותר אחורה בדיאגרמה להזגמה ראה ציור 3.2.



ציור 3.2 - הקשר בין דיאגרמת השבכה לדיאגרמת העץ.

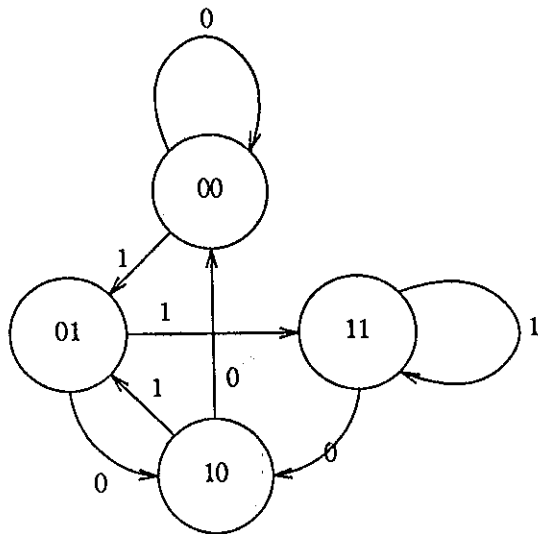
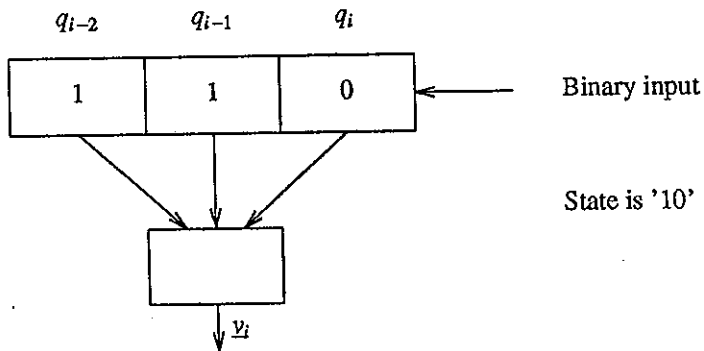
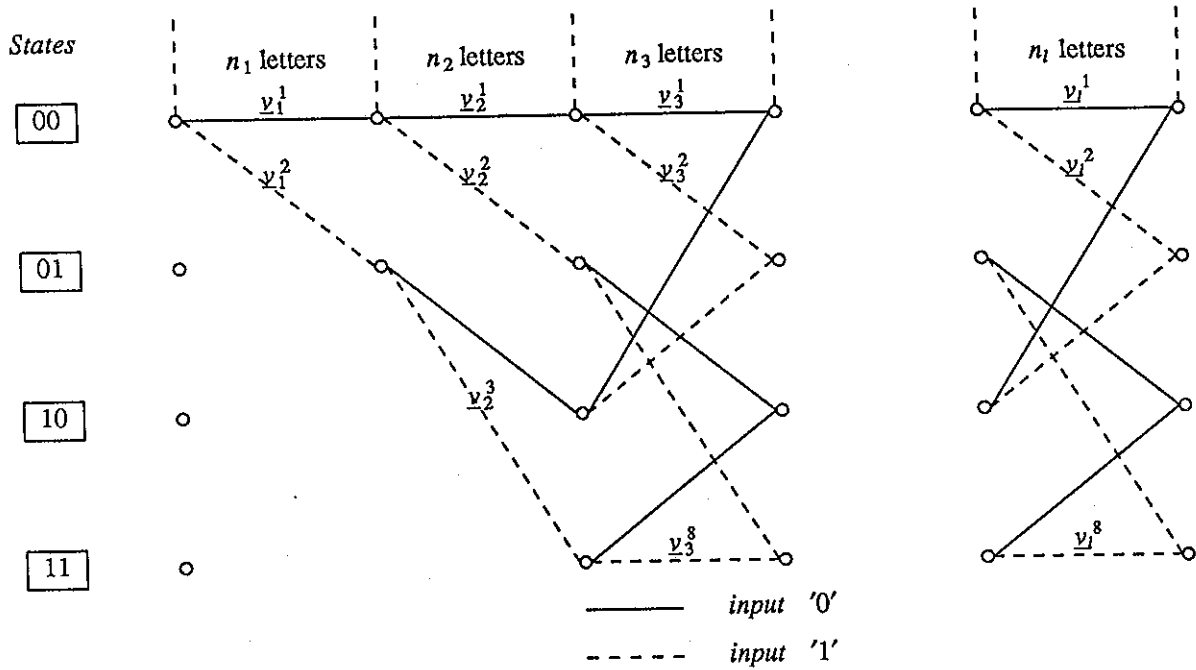
Fig 3.2 - The relation between trellis and tree diagrams.

התלות המבנית הנוספת שהוכנסה יצרה דיאגרמה אשר מספר הצמתים בדרגה שלה מוגבל ואינו גדל האקספוננציאלית עם אורך הדיאגרמה, כדוגמת מס' הצמתים בדרגה של העץ. מס' הצמתים בדרגה של

השבכה, פרט לדרגות בתחילת הדיאגרמה, קבוע ושוה ל q^{K-1} . לפיכך גם כמות הזכרון בדרגה קבועה ולא תלויה באורך הבלוק המקודד N .

המפגנח של דיאגרמת השבכה שתוארה לעיל נתנת לאפיון באמצעות מכונת מצבים סופית הממומשת באמצעות Shift Register (SR) באורך K הפועל בבסיס q (כלומר הא"ב שלו כולל ספרות בתחום $0 : q-1$). המצב הקודם של ה-SR נתון ע"י $K-1$ האברים השמאליים בו, ומספר המצבים האפשריים הינו q^{K-1} . המצב של ה-SR נקבע על פי המצב הקודם והכניסה החדשה. ציור מס 3.3b מתאר את ה-SR המתאים לדיאגרמת השבכה המתוארת בציור 3.3a. ציור 3.3c מתאר את דיאגרמת המעבר בין המצבים עבור $K=3$, $q=2$. כדוגמת העץ, כל ענף בדרגה l של דיאגרמת השבכה מאוכלס בוקטור y_l המורכב מ n_l מילות קוד (וקטור שונה לכל ענף). השבכה נקראת קבועה אם ערכי מילות הקוד ומספרן זהה בכל דרגות הדיאגרמה. כאשר מילות הקוד שונות מדרגה לדרגה הדיאגרמה משתנה בזמן.

בתהליך הקידוד על גבי הדיאגרמה מחפשים את המסלול בדיאגרמה כך ש- העוות בין וקטור המקור u לבין סדרת הוקטורים $(y_l)_{l=1}^L$ המאכלסת את המסלול הנבחר, והמהווה את וקטור השחזור y , הוא מינימלי. סדרת האינדקסים ה- q -ארית משודרת דרך ערוץ חסר רעש למקלט. המקלט אשר מכיל דיאגרמת שבכה זהה לזו הנמצאת במשדר, משחזר מהמסלול הנקלט את וקטור השחזור y . קוונטיזציה וקטורי במבנה שבכה הינו אופטימלי דהיינו ביצועיו שואפים לפונקציית הקצב-עוות עם עלית אורך האילוף K [24]. בפרק 4 יתוארו ביתר פירוט דרכי בניית השבכה, שיטות לאכלוס יעיל של מילות הקוד על גבי השבכה ושיטות לחיפוש בדיאגרמה. בסעיף הבא נפרט את המוטיבציה לשימוש בשבכה בסכמה המוצעת.



ציור 3.3 - דיאגרמת שבכה (a) מקודד (b) מפענח (c) מכונת מצבים.
 Fig 3.3 - Trellis Diagram (a) Encoder (b) Decoder (c) State machine.

3.4 המוטיבציה לשימוש בדיאגרמת השבכה

3.4.1 הקדמה

בקידוד וקטורי באמצעות דיאגרמת שבכה קיימת תלות בין וקטורי הקוד השונים. עובדה זו מאפשרת הפחתה של כמות הזכרון והחשובים הנדרשים לקידוד וקטור באורך N לסדר גודל של $O(NR)$ בהשוואה לסדר גודל של $O(2^{NR})$ הנדרשת בספר קוד. התלות הקיימת בין וקטורי הקוד בדיאגרמת השבכה פוגעת כמובן במספר מילות הקוד השונות זו מזו, ולפיכך נפגעים ביצועי מקודד השבכה בהשוואה למקודד ספר הקוד כאשר לא נלקחת בחשבון הסיבוכיות. באופן מעשי, שיקולי סיבוכיות הינם פרמטר חשוב בתכנון מקודדים, ולכן נשאלת השאלה איזה הוא המקודד הוקטורי העדיף תחת אילוף של סיבוכיות נתונה. סעיף זה משווה בין מקודדים וקטוריים הממומשים באמצעות ספר קוד למקודדים וקטוריים הממומשים באמצעות דיאגרמת שבכה.

לצדכיי ההשוואות נגדיר בטבלה 3.1 את הפרמטרים הממששים אותנו לתאור ספר הקוד ודיאגרמת השבכה.

פרמטרי ספר הקוד		פרמטרי השבכה	
N_B	אורך הוקטור המקודד	N_T	אורך הוקטור המקודד
R_B	הקצב הנדרש	R_T	הקצב הנדרש
K_B	גודל המילון	K_T	אורך האילוף
		q	מספר הענפים היוצאים מצומת
		L	אורך השבכה
			מספר מילות הקוד
		n_l	ברמה l של השבכה

טבלה 3.1 - פרמטרי השבכה וספר הקוד.

Table 3.1 - Trellis and Codebook parameters.

הממדים לפיהם נשווה את שתי צורות המימוש של הקוונטיזר הוקטורי הינם:

א. כמות הזכרון - Me .

ב. כמות החישובים (מספר חישובי העוות) לאות מקוד - Co .

ג. הביצועים המושגים - SNR .

טבלה 3.2 מפרטת את כמות הזכרון, החישובים והקצב עבור ספר הקוד והשבכה בעלי הפרמטרים המפורטים בטבלה 3.1.

ספר קוד	שבכה*	
$K_B \cdot N_B$	$q^{K_T} \cdot N_T$	Me
K_B	q^{K_T}	Co
$\frac{1}{N_B} \log K_B$	$\frac{L}{N_T} \log q$	R

טבלה 3.2 - כמות הזכרון והחישובים עבור ספר הקוד ודיאגרמת שבכה.

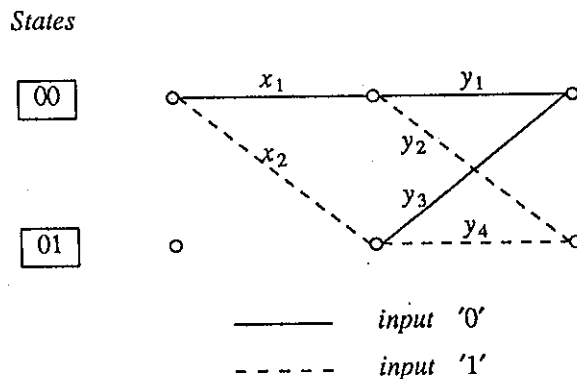
Table 3.2 - Memory requirements and amount of Computations for Codebook and Trellis.

* השבכה הביטוי עבור כמות החישובים עבור השבכה אינו לוקח בחשבון את התחלת הדיאגרמה בה כמות החישובים קטנה יותר אלא מניח דיאגרמה המתחילה מ q^{K-1} מצבים (דיאגרמה מלאה).

הביצועים המושגים ע"י המקודדים הנ"ל אינם נתנים לחשוב אנליטי, לפיכך קבענו אותם בצורה אמפירית ע"י סימולציות. בתור אות מקור לקידוד השתמשנו במקור גאוסי הפולט אותיות מקור בלתי תלויות סטטיסטית בעלות פלוג סטנדרטי $N(0,1)$.

3.4.2 השוואה בלא אילוצי סיבוכיות

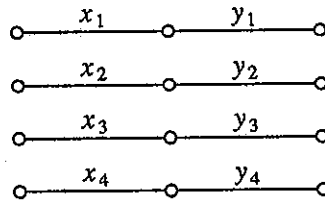
כאמור, בלי להתחשב כלל בסיבוכיות, הביצועים של דיאגרמת השבכה נחותים יחסית לאלה המושגים ע"י ספר קוד. הסיבה לכך הינה המספר הקטן יותר של מילות הקוד במילון הממומש ע"י דיאגרמת השבכה הנובע מהתלות הקיימת בין וקטורי הקוד. נתיחס לדוגמה לוקטור באורך $N = N_T = N_B = 2$ ציור 3.4 מתאר דיאגרמת שבכה בעלת הפרמטרים הבאים: $L=2, K_T=2, q=2$ ו $n_l = 1, l=1,2$.



ציור 3.4 - דיאגרמת שבכה בעלת קצב $R=1bit/pel$ עבור וקטורים באורך 2.
Fig 3.4 - Trellis diagram with rate of $R=1bit/pel$ for vectors of length 2.

כל ענף בדיאגרמה מאוכלס במילת קוד שונה. נסמן את מילות הקוד ברמה הראשונה של השבכה ב x_1, x_2 ואת מילות הקוד ברמה השנייה של הדיאגרמה ב- y_1, y_2, y_3, y_4 . מספר מילות הקוד השונות זו מזו המאכלסות את הדיאגרמה הינו לפיכך 6. סט הוקטורים האפשריים מכיל 4 וקטורים כדלקמן $\{x_1y_1, x_1y_2, x_2y_1, x_2y_2\}$

ציור 3.5 מתאר ספר קוד לקידוד וקטור בעל אורך זהה באותו קצב.



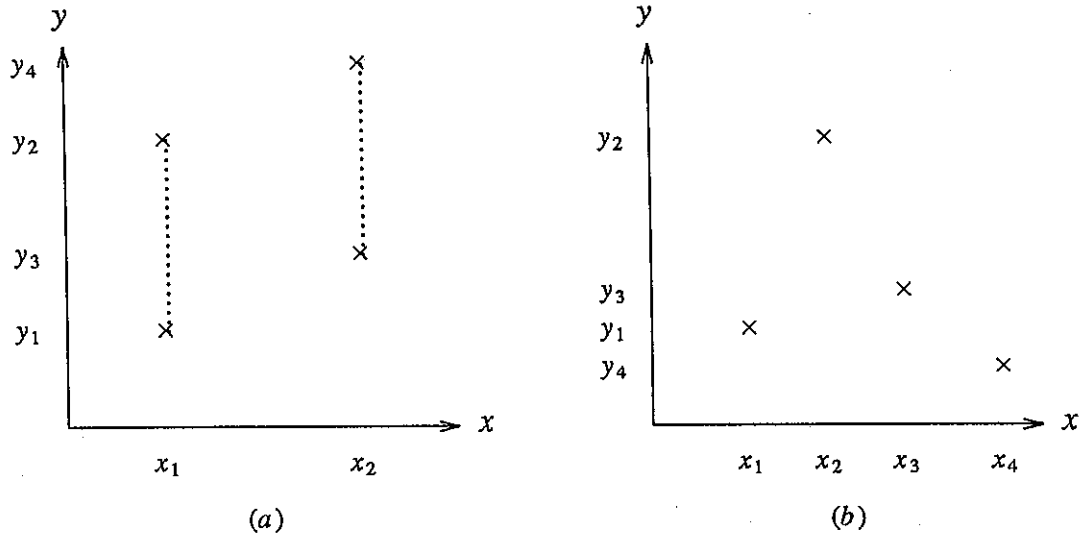
ציור 3.5 - ספר קוד בעל קצב של $R=1 \text{ bit/pel}$ לקידוד וקטורים באורך 2.

Fig 3.5 - Code-book for coding vectors of length 2, at the rate of $R = 1 \text{ bit/pel}$.

במקרה זה המילון מכיל $K_B = 2^{NR} = 4$ וקטורים המורכבים מ-2 מילות קוד כ"א, סך כל מילות הקוד השונות במקרה זה הינו 8, סט הוקטורים האפשריים מכיל 4 וקטורים כדלקמן

$$\{x_1y_1, x_2y_2, x_3y_3, x_4y_4\}$$

כפי שרואים סט מילות הקוד השונות זו מזו, עבור קצב ואורך וקטור נתונים, גדול יותר בספר הקוד מאשר בשבכה, עובדה המצביעה על עדיפות ספר הקוד. ציור 3.6 מתאר בצורה גרפית על המישור הדו-ממדי את סט הוקטורים האפשריים עבור ספר הקוד ועבור השבכה שתוארו לעיל.



ציור 3.6 - וקטורי הקוד על המישור הדו-ממדי (a) עבור שבכה (b) עבור ספר קוד.

Fig 3.6 - The code vectors on the 2-D plane (a) trellis (b) code-book.

בשני המקרים, כמות הוקטורים האפשריים זהה ושווה ל-4. אולם, עבור סט הנקודות המרכיב את השבכה קיים אילוץ שהרכיב הראשון של כל זוג וקטורים מתלכד. ניתן להסתכל על סט הוקטורים הנוצר בשבכה כעל תת-קבוצה של סט הוקטורים הנוצר בספר הקוד, אשר נמצאת תחת אילוץ של זהות בין רכיבים בוקטורים השונים.

לסיכום, ניתן להסתכל על דיאגרמת השבכה כעל מקרה פרטי של ספר קוד, ולפיכך ביצועי השבכה ללא שום אילוץ על הסיבוכיות יהיו נחותים מביצועי ספר הקוד.

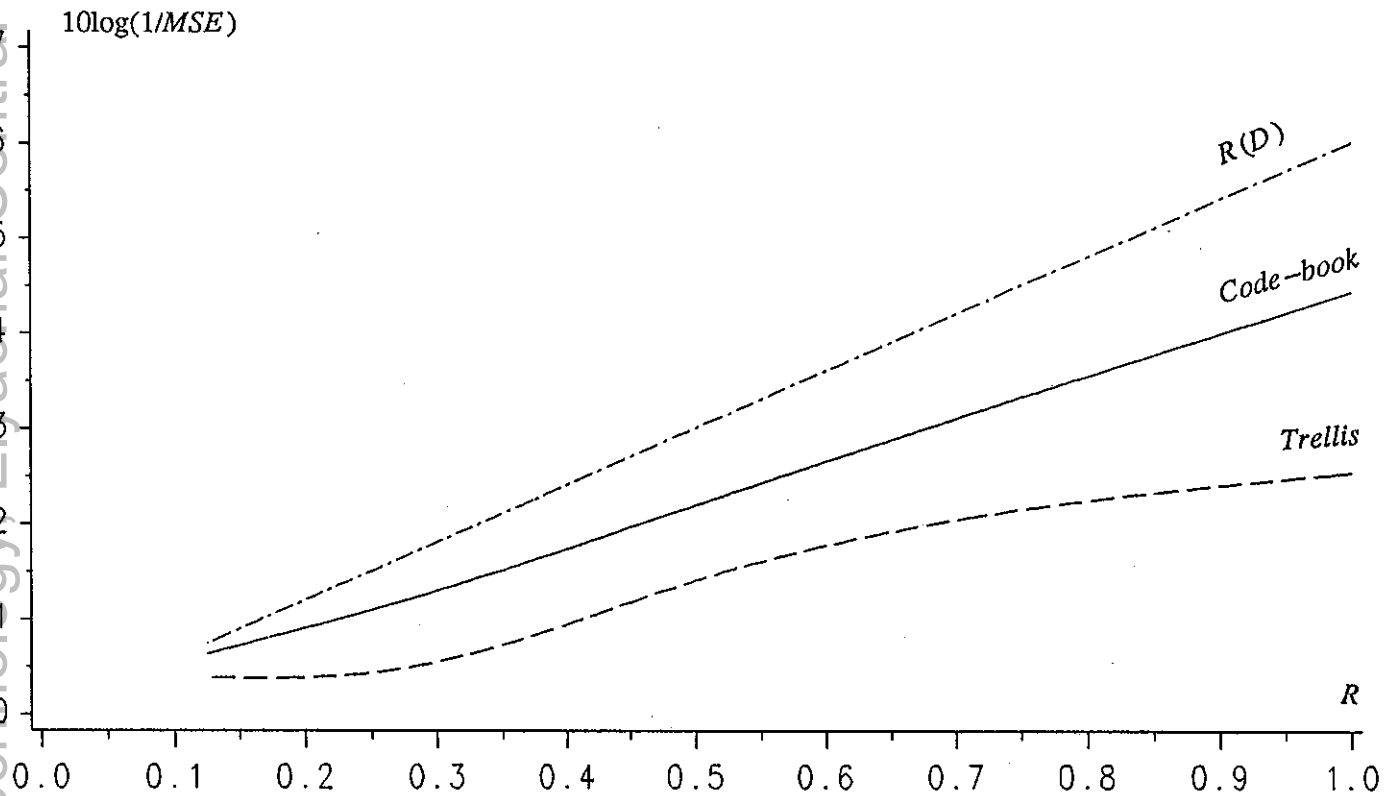
כדי לאמת את ההשערה הנ"ל קודדנו וקטור באורך נתון מתוך מקור גאוסי באמצעות שני המקודדים. בסימולציה השתמשנו בדיאגרמת שבכה בעלת הפרמטרים $q=2$ $K_T=3$ 1. טבלה 3.3 מתארת את התוצאות שהתקבלו עבור שני המקודדים ואת הסיבוכיות הכרוכה במימושם.

R	N	$10\log\frac{1}{MSE_B}$	$10\log\frac{1}{MSE_T}$	$10\log\frac{1}{MSE_{D(R)}}$	M_B	M_T	C_B	C_T
1	10	4.45	2.55	6.02	10240	80	1024	8
1/2	20	2.20	1.41	3.01	20480	160	1024	8
1/4	30	1.10	0.44	1.51	40960	320	1024	8
1/8	40	0.64	0.39	0.75	81920	640	1024	8

טבלה 3.3 - ביצועי ספר קוד ושבכה עבור וקטור באורך זהה ללא אילוצי סיבוכיות.

Table 3.3 - Code-book and Trellis performance for the same length vector with no constraints on the complexity.

ניתן לראות שהביצועים של המקודד שמומש באמצעות ספר קוד עדיפים על אלה של מקודד השבכה. לעומת זאת, סיבוכיות המימוש בעזרת שבכה, דהיינו כמות הזכרון והחישובים, קטנה בהרבה מזו של ספר הקוד. גרף המתאר את ביצועי שני המקודדים בהשוואה לפונקציית הקצב-עוות של המקודד הגאוסי מובא בצירוף 3.7.



ציור 3.7 - השוואה בין שבכה לספר קוד בלא אילוצים.

Fig 3.7 - Comparison between trellis and code-book with no constraints.

הגרף מתאר את הביצועים של כל מקודד ואת הביצועים המתאימים לפונקציית הקצב-עוות בתלות בקצב. הגרף מביט את העדיפות של מקודד הממומש בעזרת ספר קוד ללא אלוץ הסיבוכיות.

3.4.3 השוואה תחת אילוצים של כמות זיכרון וזכרון זהה בשני הקוונטיזרים

באופן מעשי, הסיבוכיות של הקוונטיזר הינה פרמטר חשוב בשיקולי התכנון שלו, ולפיכך יש צורך להתחשב בהם בעת התכנון. כפי שראינו, הפרמטרים להשוואה בין ספר הקוד לשבכה הינם רבים. לצורך הבהרת הנושא ננתח בשלב הראשון את המצב בו רוצים לקודד וקטור באורך N ובקצב R בשתי הסכמות תחת האלוץ של כמות זכרון וכמות זיכרון זהה. התוצאה הצפויה הינה שוויון בביצועי שתי הסכמות, כיוון שמספר מילות הקוד השונות זו מזו ומספר הקוטורים השונים האפשריים זהה בשתי הסכמות. לצורך פשטות הניתוח נניח דיאגרמת שבכה המתחילה מ q^{K-1} מצבים אפשריים (דיאגרמה מלאה). במקרה זה קצב השבכה נתון ע"י הביטוי הבא:

$$(3.11) \quad R_T = \frac{L}{N} \log(q) + \frac{1}{N} \log(q^{K_T-1})$$

האיבר השני בביטוי הינו הקצב הנדרש להעברת המידע על המצב ההתחלתי. הקצב של ספר הקוד נתון ע"י

$$(3.12) \quad R_B = \frac{1}{N} \log(K_B)$$

הביטויים לכמות הזכרון ולכמות החישובים בשני הקוונטיזורים נתונה בטבלה 3.2. השוואה של כמות הזכרון, החישובים והקצב בין שני המקודדים מביאה לדרישה הבאה

$$(3.13) \quad \frac{K_T + L - 1}{K_T} = 1$$

פיתוח מפורט של התוצאה הנ"ל מובא בנספח א'. דרישת השיוויון מתקיימת בשני מקרים:

א. $L = 1$, K_T כלשהו - במקרה זה מבנה השבכה מתלכד בדיוק עם מבנה ספר הקוד וברור שמתקבלים ביצועים זהים לביצועי ספר הקוד דבר המאשר את ציפיותינו.

ב. L כלשהו, $K_T \rightarrow \infty$ - מקרה זה אינו מעשי כיוון שהוא גורר בעקבותיו את התוצאה

$$q = 2^{\frac{NR}{K_T}} \rightarrow 0$$

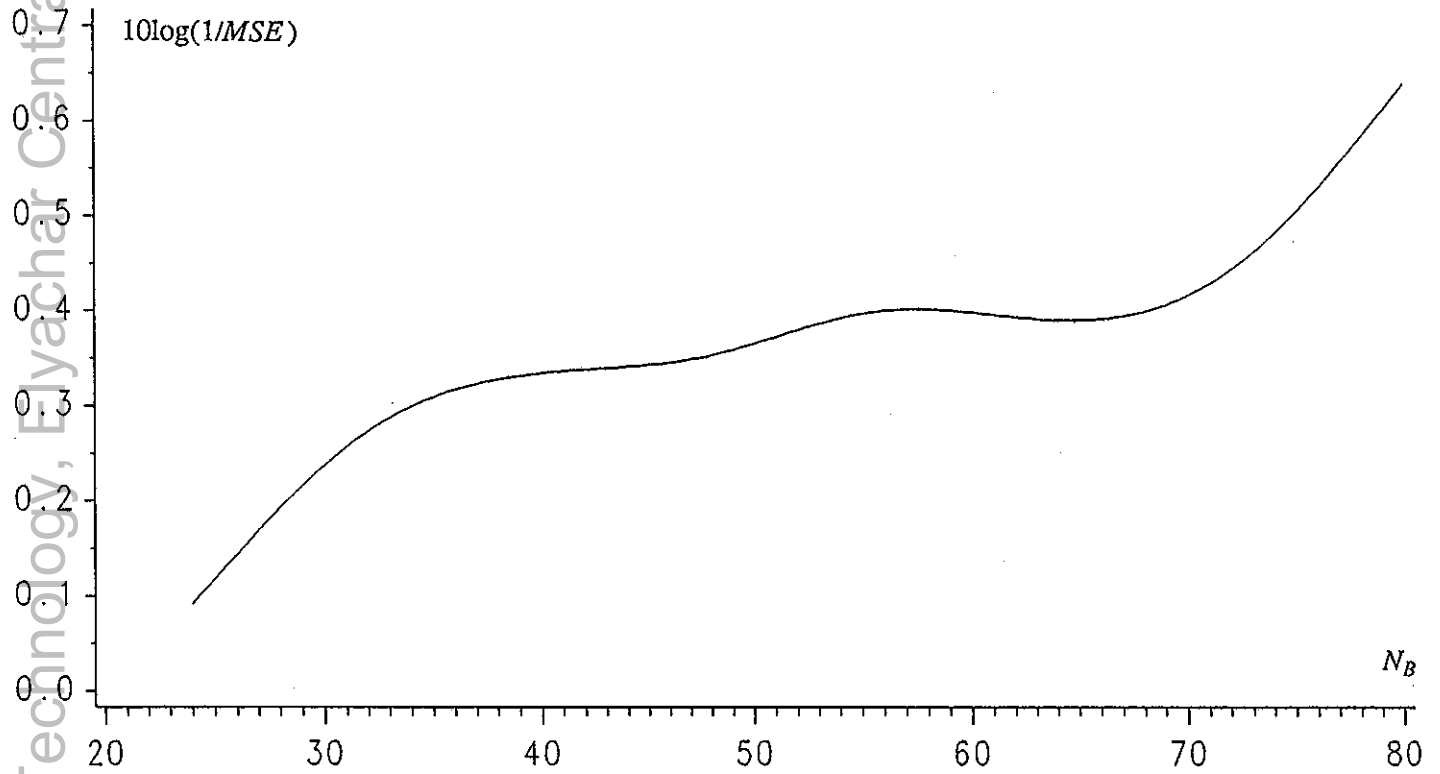
וכמובן שלא ניתן לבנות דיאגרמת שבכה כזו.

כפי שרואים, לאור הניתוח הנ"ל, השוואה בו זמנית של כמות הזכרון ושל כמות החישובים בשני המקודדים עבור אותו קצב ואותו אורך בלוק אינה תורמת מידע על עדיפות אחד הקוונטיזורים, כיוון שבמקרה זה דיאגרמת השבכה הופכת לספר קוד.

3.4.4 בדיקת השפעת אורכי הבלוקים על ביצועי שני הקוונטיזורים

תוצאה בסיסית מתורת האינפורמציה אומרת שהחסס על ביצועי מקודד וקטורי הדוק יותר עם עלית גודל הבלוק המקודד. מתוצאה זו ניתן להסיק שגם ביצועי הקוד עצמו משתפרים עם עלית גודל הבלוק המקודד. השערה זו נבדקה בסימולציה שבמסגרתה קודדו באמצעות ספר קוד ובאמצעות שבכה וקטורים באורכים שונים ובקצב קבוע ($R = \frac{1}{8}$ במקרה שלנו). תוצאות הסימולציה עבור ספר קוד מתוארות

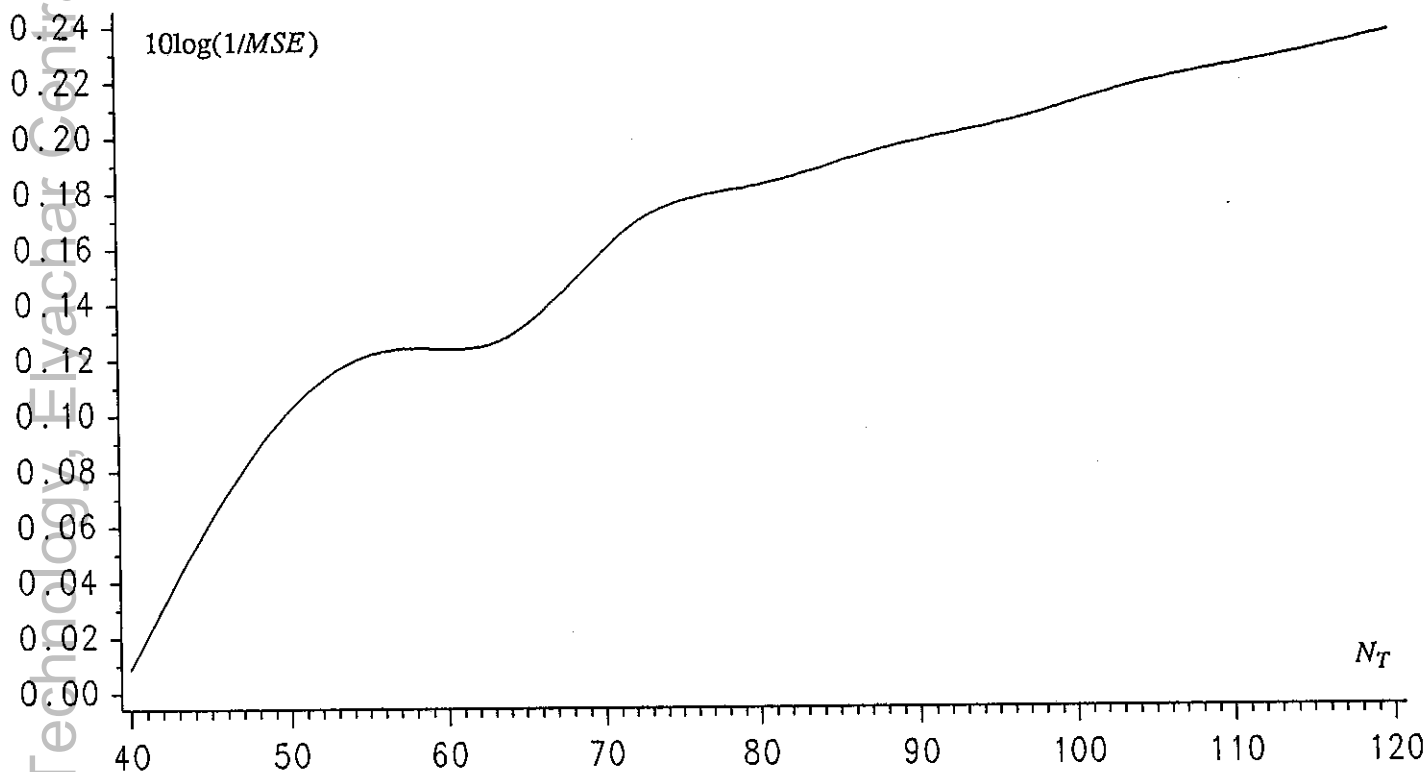
בציור 3.8.



ציור 3.8 - ביצועי ספר קוד כתלות באורך הקטור.

Fig 3.8 - Code-book performance vs vector length.

כפי שרואים ביצועי הקוד משתפרים עם עלית אורך הבלוק. תוצאה דומה התקבלה בסימולציה עבור דיאגרמת השבכה והיא מתוארת בציור 3.9.



ציור 3.9 - ביצועי שבכה כתלות באורך הוקטור.

Fig 3.9 - Trellis performance vs vector length.

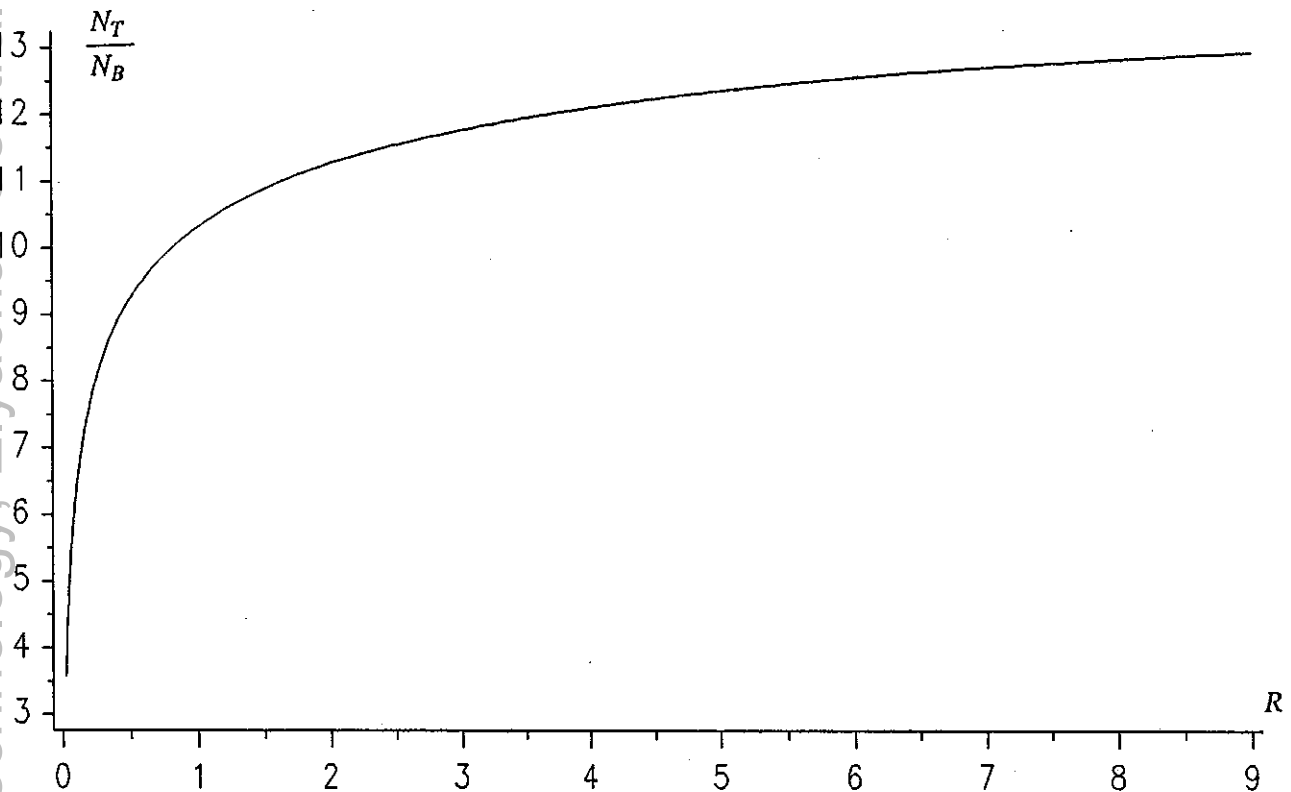
בנקבות התוצאות הללו בדקנו מהו אורך הבלוק אותו יש לקודד בכ"א מהקוונטיזרים, לשם קבלת כמות זכרון זהה. דרישה לכמות זכרון זהה נותנת את המשוואה הבאה:

$$(3.14) \quad M_T = N_T q^{K_T} = N_T 2^{\frac{N_T R K_T}{L + K_T - 1}} = N_B 2^{RN_B} = M_B$$

אשר ניתנת לפיתוח כמשוואה סתומה כדלקמן:

$$(3.15) \quad \frac{N_T}{N_B} = 2^{R(N_B - \frac{K_T}{L + K_T - 1} N_T)}$$

ציור 3.10 מתאר בצורה גרפית את היחס $\frac{N_T}{N_B}$ עבור הפרמטרים $K_T = 3$, $L = 64$



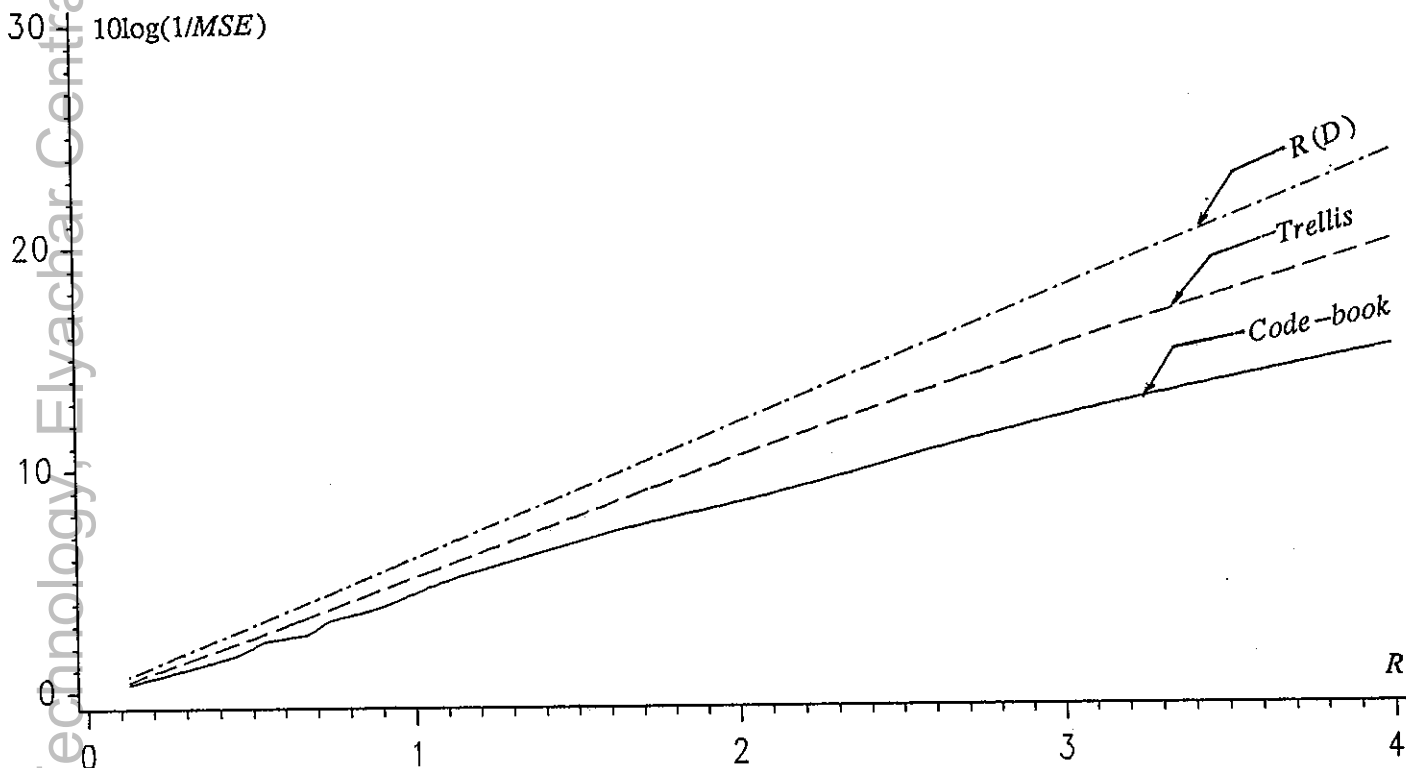
ציור 3.10 - יחס אדכי הוקטורים הניתנים לקידוד בשבכה ובספר קוד תחת אילוך של כמות זכרון זהה

Fig 3.10 - The ratio of vector lengths which can be coded on the trellis and code book under the constraint of same amount of memory.

כפי שרואים מהתוצאות הנ"ל, עבור אילוך של כמות זכרון זהה בשני הקוונטיזרים מתקבלת דרישה לאורך וקטור ארוך יותר בשבכה. בסיום הסעיף, ננתח את המשמעות של תוצאה זו.

3.4.5 בדיקת ביצועי הקוונטיזרים תחת אילוך של כמות חישובים או כמות זכרון זהה

להלן נתאר תוצאות ניסויים בהם נקבעו כמות הזכרון או כמות החישובים כגדלים זהים בשני הקוונטיזרים. יש לשים לב שתוצאות הניסויים הללו מעידות על התנהגות הקוונטיזרים רק בתחומים של הפרמטרים בהם הם נבדקו, דהיינו בתחומי אורכי הבלוקים, הקצבים והסיבוכיות שנבדקו. בסכום לסעיף ננתח את משמעות התוצאות ונדון בתקפותן מחוץ לתחום הפרמטרים שנבדקו. בתחילה נעשה השוואת ביצועים בין שני הקוונטיזרים תחת אילוך של כמות חישובים זהה (256 - בסימולציות שהורצו). תחת אילוך זה קודדנו וקטורים מן המקור הגאוסי בקצבים שונים. ציור 3.11 מתאר בצורה גרפית את הביצועים המושגים ע"י כ"א מהקוונטיזרים בהשוואה לביצועים המתאימים לפונקציית הקצב-עוות של המקור.



ציור 3.11 - השוואת ביצועי ספר קוד ושבכה תחת אילוף של כמות חישובים זהה.

Fig 3.11 - Comparison of code-book and trellis performances under the constraint of same amount of computations.

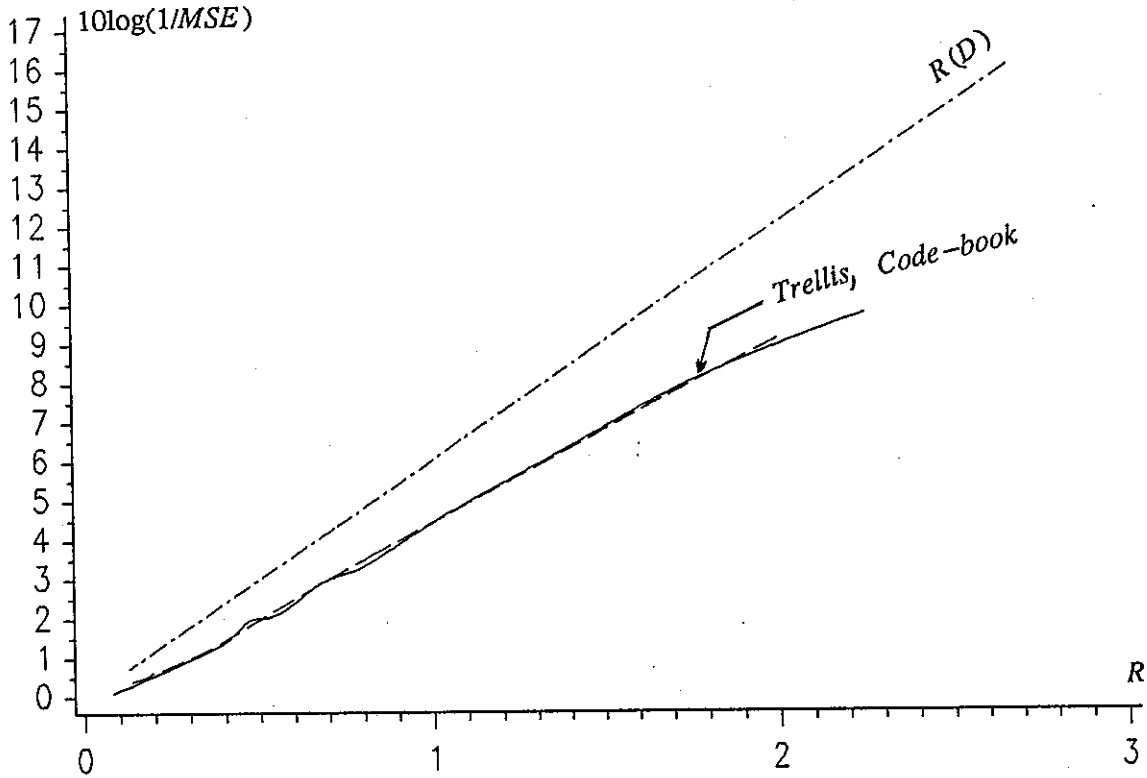
כפי שרואים, הביצועים המושגים באמצעות דיאגרמת השבכה עדיפים על אלה המושגים באמצעות ספר קוד תחת האילוף הנ"ל. בהשוואה האחרונה לא נלקחה כלל בחשבון כמות הזכרון הנדרשת בכ"א

מהקוונטיזורים. עבור אילוף של כמות חישובים זהה בשני הקוונטיזורים, מתקבל הקשר $\frac{N_T}{N_B} = \frac{L}{K_T}$,

ויחס זה גדול מ-1 בשבכה. דהיינו, כמות הזכרון עבור השבכה בהשוואה הנ"ל גדולה מזו הנדרשת עבור ספר הקוד, כך שלא ניתן לקבוע בצורה חד-משמעית את העדיפות הכוללת של השבכה.

בניסוי הבא קבענו כמות זכרון זהה בשבכה ובספר הקוד (2048-בסימולציות שחורצו). תחת

האילוף הזה קוודדנו וקטורים מן המקור הגאוסי בקצבים שונים. ציור 3.12 מתאר בצורה גרפית את הביצועים המושגים ע"י כ"א מהקוונטיזורים בהשוואה לביצועים המתאימים לפונקציית הקצב-עוות של המקור



ציור 3.12 - השוואת ביצועי ספר קוד ושכבה תחת אילוף של כמות זכרון זהה.

Fig 3.12 - Comparison of code-book and trellis performances under the constraint of same memory requirements.

כפי שרואים, הביצועים עבור שני הקוונטיזרים מזדהים. תוצאה זו אכן סבירה עקב העובדה שהמשמעות של האילוף שקבענו הינה כמות זהה של מילות קוד בשני הקוונטיזרים. היחס בין כמויות החישובים

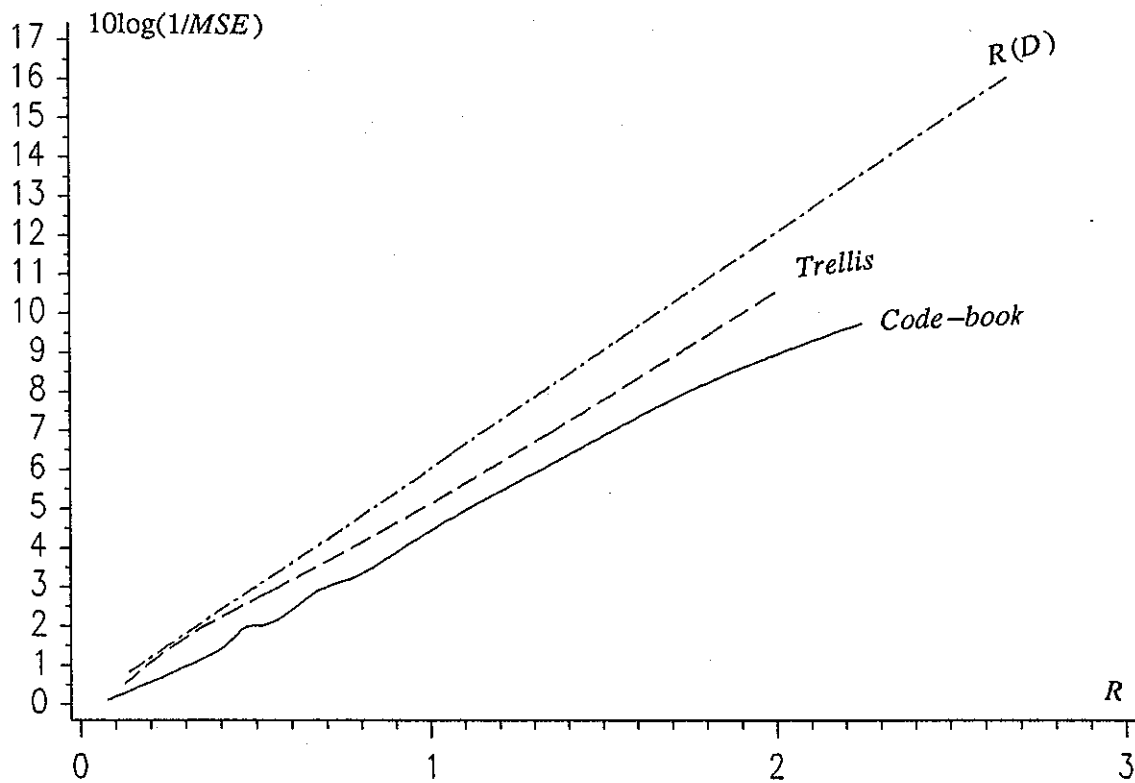
לאות מקור תחת האילוף של כמות זכרון זהה בשכבה ובספר הקוד נתון ע"י: $\frac{C_B}{C_T} = \frac{N_T}{N_B}$. כפי שראינו

בנוסחה 3.27 ובציור 3.10 תחת האילוף הנ"ל מקבלים $\frac{N_T}{N_B} > 1$. לפיכך, מתקבל $\frac{C_B}{C_T} > 1$.

3.4.6 בדיקת ביצועי הקוונטיזרים תחת אילוף של כמות חישובים וכמות זכרון זהה (אכלוס ציקלי של השכבה)

כצעד אחרון בתהליך ההשוואה השוונג את הביצועים של ספר קוד עם דיאגרמת שכבה תחת האילוף של מספר חישובים קבוע וכמות מילות קוד קבועה. הדבר התאפשר ע"י שמוש בדיאגרמת שכבה אשר כמות מילות הקוד השונות זו מזו שבה זהה לכמות מילות הקוד בספר הקוד. עקב העובדה שלאכלוס השכבה נדרשים יותר מילות קוד, משתמשים באותן מילות קוד שימוש חוזר בצורה ציקלית. אורך הדיאגרמה נקבע כך שכמות החישובים לאות מקור תהיה זהה לכמות החישובים לאות מקור בספר

הקוד. ניתן להסתכל על דיאגרמה כזו כשילוב של דיאגרמת שבכה קבועה בזמן עם דיאגרמת שבכה משתנה בזמן. הדיאגרמה משתנה על פרקי זמן קצרים אולם חוזרת על עצמה בפרקי זמן ארוכים. דיאגרמה כזו לא נדונה כלל בספרות ואין לגביה משפטי קידוד, אולם היא מדגימה היטב את יתרון הקוונטיזר אשר ממומש באמצעות שבכה ע"פ זה הממומש באמצעות ספר קוד. ציור 3.13 מתאר את ביצועי השבכה הנ"ל, את ביצועי ספר הקוד ואת הביצועים המתאימים לפונקציית הקצב עוות עבור קצבים שונים.



ציור 3.13 - השוואת ביצועי ספר קוד ושבכה תחת אילוף של כמות זכרון וחישובים זהה (איכלוס ציקלי).

Fig 3.13 - Comparison of code-book and trellis performances under the constraint of same memory requirements and computational load (cyclic population).

כפי שרואים, בתנאים אלו, הביצועים של השבכה עדיפים על אלה המושגים ע"י ספר הקוד.

3.4.7 טכנס

לסיכום הסעיף נסכם בטבלה 3.4 את ההשוואות שבוצעו בין ספר הקוד לשבכה. האילוצים תחתם נערכו ההשוואות מפורטים אף הם בטבלה, וכל ההשוואות הינן עבור קצב זהה בשני הקוונטיזרים.

הערות	מקום התוצאות	תוצאת הבדיקה	תאור הבדיקה	
אורך בלוק זהה בשני הקוונטיזרים	טבלה 3.3 ציור 3.7	עדיפות לספר קוד	השוואת ביצועים ללא אילוצים	1
תחת אילוצים אלה דיאגרמת השבכה הופכת לספר קוד	פיתוח תיאורטי נספח א'	שוויון בביצועים	השוואת ביצועים עם כמות חישובים וזכרון זהה	2
---	ציור 3.10	אורך בלוק בשבכה גדול מאורך בלוק בספר קוד	השוואת אורך בלוק לקידוד; כמות זכרון זהה	3
כמות זכרון בשבכה גדולה מזו בספר קוד	ציור 3.11	עדיפות לשבכה	השוואת ביצועים עם כמות חישובים זהה	4
כמות חישובים בשבכה קטנה מזו בספר קוד	ציור 3.12	שוויון בביצועים	השוואת ביצועים עם כמות זכרון זהה	5
שבכה מחזורית	ציור 3.13	עדיפות לשבכה	השוואת ביצועים עם כמות זכרון וחישובים זהה	6

טבלה 3.4 - סיכום השוואות בין דיאגרמת שבכה לספר קוד.

Table 3.4 - Summary of comparisons between trellis diagram and code-book.

בניסוי ללא כל אילוצים (סודר 1 בטבלה) התקבלה עדיפות בביצועי ספר הקוד בהשוואה לביצועי השבכה. וקטורי הקוד בשבכה מאולצים להיות בעלי רכיבים זהים עקב המבנה של השבכה, ולפיכך, כמות מילות הקוד במילון הנוצר באמצעות השבכה קטנה מכמות מילות הקוד במילון הנוצר ע"י ספר הקוד. עובדה זו גוררת בעקבותיה, כצפוי, עדיפות בביצועי ספר הקוד.

השוואה בין הקוונטיזרים עבור כמות חישובים וזכרון זהה (סודר 2 בטבלה) לא הוסיפה מידע נוסף, כיוון שהתקבלה דרישה שאורך השבכה יהיה $L=1$. משמעותה של דרישה זו הינה שהשבכה מהווה ספר קוד, ואין תלות בין וקטורי הקוד שלה.

בניסוי שתוצאותיו מתוארות בסודר 4 הושו ביצועי שני הקוונטיזרים תחת אילוף של כמות חישובים זהה. התוצאות שהתקבלו הצביעו על עדיפות בביצועי השבכה. היחס בין כמות הזכרון הנדרשת

$$\frac{M_T}{M_B} = \frac{L}{K_T}$$

יחס זה גדול מ 1 בשבכה. לפיכך, מתוצאות הניסוי הנ"ל, לא ניתן להסיק מסקנה חד משמעית לגבי עדיפות של אחד מן הקוונטיזרים. מצד אחד ביצועי השבכה עדיפים, ואילו, מצד שני, כמות הזכרון הנדרשת למימוש השבכה גדולה יותר.

בניסוי הבא (סודר 5) הושוּוּ ביצועי השבכה וביצועי ספר הקוד תחת אילוף של כמות זכרון זהה. בתחום הפרמטרים שנבדק על ידנו, (כמות זכרון של 2048 מילים וקצבים בתחום שבין 125 bit/element לבין 2 bits/element), התגלה שוויון בביצועי שתי הסכמות. היחס בין הסיבוכיות של שני הקוונטיזרים נתון במקרה זה ע"י:

$$\frac{C_T}{C_B} = \frac{N_T}{N_B}$$

כפי שרואים בניסוי המתואר בסודר 3, יחס אורכי הוקטורים תחת האילוף של כמות זכרון זהה גדול מ 1. לפיכך מתקבל, שעבור כמות זכרון זהה, בתחומי הפרמטרים שנבדקו, ביצועי שני הקוונטיזרים מזדהים ואילו סיבוכיות השבכה קטנה יותר, עובדה המצביעה על עדיפות השבכה בתחומים אלה של פרמטרים.

כאמור, כפי שהדגשנו במהלך פרק זה, השוואת הביצועים הנ"ל תקפה אך ורק בתחום הפרמטרים בו היא בוצעה, ולא ניתן להסיק ממנה על ביצועי הקוונטיזרים עבור ערכים אחרים של הפרמטרים. דוגמה לסט פרמטרים בו ההשוואה אינה תקפה הינה הדוגמה בה אורכי הוקטורים של הקוונטיזרים N_T , N_B שואפים ל ∞ . כידוע [13], ביצועי ספר הקוד שואפים לפונקציה הקצב-עוות עם עליית אורך הוקטור ל ∞ . לעומת זאת עבור אורך אילוף K_T סופי ביצועי השבכה אינם שואפים לפונקציה הקצב-עוות עם עליית אורך הוקטור המקודד ל ∞ [24].

בעבודה זו לא נעסוק בחקירה נוספת של ההשוואה בין הקוונטיזרים. העובדה שביצועי השבכה עדיפים, בתחום פרמטרים רחב יחסית (אורכי וקטורים $N_B = 60$ וקצבים בתחום (2 - 0.125)), על ביצועי ספר הקוד נותנת את המוטיבציה לשימוש בשבכה למימוש הקוונטיזר הוקטורי בסכמה המוצעת.

פרק 4 - תיאור סכמת הקידוד הבסיסית

4.1 כללי

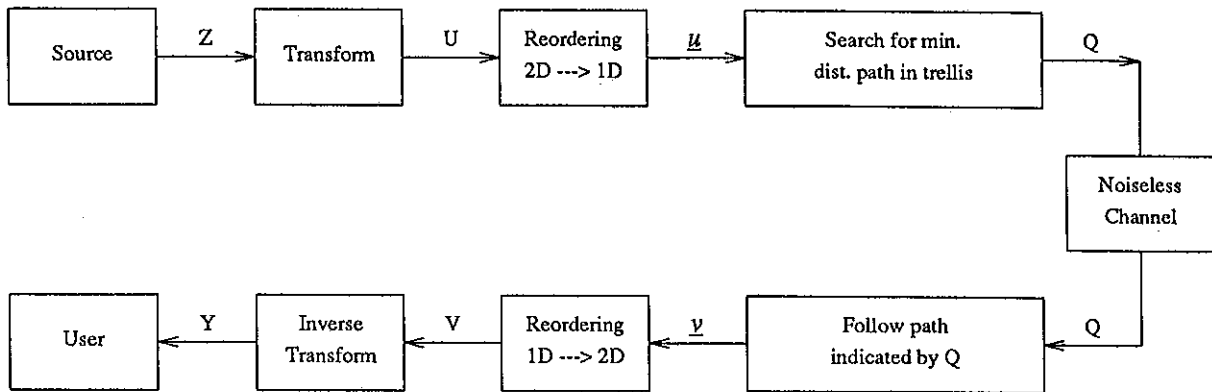
בפרקים 2 ו-3 תארנו את אבני הבנין הנדרשות ליישום סכמת הקידוד המוצעת. סכמת הקידוד המבוססת על התמרה אופטימלית (KLT) וקידוד מקדמי התמונה במישור ההתמרה באמצעות דיאגרמת שבכה הוכחה ע"י Mazor & Pearlman [1] כאופטימלית עבור מקור גאוסי מרקובי וקריטריון שגיאה ריבועית ממוצעת. ב-[27] הופעלה הסכמה בצורתה הבסיסית לקידוד אותות דיבור. הסכמה בצורתה הבסיסית הופעלה על ידינו לצורך קידוד תמונות והתוצאות שהתקבלו היו בלתי מספקות. הבעיה העקרית שהתגלתה היתה "תופעת הבלוקיות" אשר מתגלית בד"כ בסכמות קידוד המבוססות על התמרה והפעולות בקצב נמוך. הסיבה לביצועים הנחותים ולתופעת הבלוקיות הבלתי רצויה שהתגלתה נעוצה בעובדה שהתמונות המקודדות אינן מתאימות למודל הגאוסי-מרקובי שעל פיו תוכננה הסכמה.

הסכמה המוצעת על ידינו הינה סכמה משופרת אשר מקודדת תמונות בקצבים נמוכים מ- 1 bit/pel ומבטלת כמעט לחלוטין את תופעת הבלוקיות. בפרק זה מתוארת סכמת הקידוד הבסיסית אשר מהווה את הבסיס לסכמה המוצעת על ידינו. פרק 5 מתאר את הסכמה המוצעת על ידינו.

4.2 תיאור הסכמה האופטימלית

סכמת הקידוד האופטימלית מתוארת בצירוף 4.1. הסכמה מבצעת שני מעברים על תמונת המקור. במעבר הראשון מחולקות התמונה ל- s בלוקים, כל בלוק עובר התמרה אופטימלית דו-מימדית (KLT) לקבלת בלוק מקדמים חסרי קורלציה. במעבר השני, כל אחד מ- s הבלוקים במישור ההתמרה מאורגן כוקטור המשמש כוקטור כניסה למקודד השבכה. דיאגרמת השבכה מאוכלסת באקראי במילות קוד הנלקחות מפילוג גאוסי.

בשלב הקידוד מחפשים על פני הדיאגרמה את המסלול עבורו השגיאה הריבועית הממוצעת, בין וקטור המקודד לבין וקטור הקוד המורכב ממילות הקוד המאכלסות את המסלול, הינה מינימלית. מפת המסלול המתאימה למסלול הנבחר נשלחת דרך ערוץ (הנחשב כחסר רעש) למקלט. המקלט אשר מכיל דיאגרמת שבכה זהה לזו הנמצאת במשדר מפענח על סמך מפת המסלול הנקלט את וקטור הקוד. וקטור הקוד מאורגן חזרה כבלוק מקדמים במישור ההתמרה, אשר עובר התמרה הפוכה לקבלת הבלוק המשוחזר במישור התמונה. התהליך הנ"ל חוזר על עצמו לגבי כל הבלוקים בתמונה.



ציור 4.1 - סכמת הקידוד האופטימלית.
 Fig 4.1 - The optimal coding scheme.

4.3 בנית דיאגרמת השבכה

המבנה של דיאגרמת השבכה מאופיין ע"י ארבעה פרמטרים כדלקמן:

- q מספר הענפים היוצאים מצומת (Branching Factor).
- K גודל ה SR אשר מאפיין את השבכה - אורך האילוף (Constraint Length).
- n_l מספר מילות הקוד המאכלסות את הענפים בדרגה l של השבכה.
- L אורך דיאגרמת השבכה.

לצורך הבנה מלאה של האופן בו נקבעים הפרמטרים הנ"ל נצטט בתחילה מספר תוצאות מתורת האינפורמציה העוסקות בפונקציות הקצב-עוות של מקור גאוסי מרקובי. מקור זה משמש אותנו כמודל לתמונת, ועבודו הסכמה המוצעת ב [1], המשמשת בסיס לסכמה המוצעת על ידנו, הינה אופטימלית.

4.3.1 פונקציות הקצב-עוות של מקור גאוסי-מרקובי

המודל המשמש אותנו לתאור תמונות הינו מודל של מקור גאוסי-מרקובי. בסעיף זה נביא את תמצית התוצאות הקשורות בקידוד מקור גאוסי מרקובי, אשר נדרשות להבנת אופן בנית דיאגרמת השבכה. תאור מפורט של התוצאות הנ"ל ניתן למצוא ב [13, pp 111-113].
 תהי $z = z(t), \{t = \dots, -1, 0, 1, \dots\}$ סדרת מדגם מתוך מקור סטציונרי, גאוסי בעל ממוצע אפס ופונקציות ספקטרום הספק $S_z(\omega)$ רציפה וחטומה. תהי ϕ_z מטריצת הקואריאנס של וקטור z המורכב מ- N מדידות של המקור Z ו Γ מטריצת KLT אשר גורמת לדה-קורלציה של הוקטור Z ע"י $\underline{u} = \Gamma^T \underline{z}$ מטריצת הקואריאנס של \underline{u} הינה המטריצה המלוכסנת $\Lambda_u = \Gamma^T \phi_z \Gamma = [\lambda_l \delta_{lk}]$ הוקטור \underline{u} במישור ההתמרה הינו וקטור אקראי גאוסי בעל ממוצע אפס ווואריאנסים $\{\lambda_l\}_1^N$. וקטור השחזור \underline{y} של הוקטור \underline{z} מתקבל מוקטור השחזור במישור ההתמרה \underline{y} ע"י $\underline{y} = \Gamma \underline{z}$.
 פונקציות העוות ה- N ממדית, אשר מהווה חסם לביצועי הקודים באורך N , תלויה בשלושת הפרמטרים הבאים:

- $I(\underline{u}, \underline{y})$ האינפורמציה ההדדית הממוצעת בין המקור \underline{u} לשחזור \underline{y} במישור ההתמרה.

- $d(\underline{u}, \underline{v})$ - מדד העוות.
 $P(\underline{v} | \underline{u})$ - פונקציית הצפיפות המותנית של וקטור השחזור \underline{v} בהנתן וקטור המקור \underline{u} .

כיוון שההתמרה Γ הינה יוניטרית האינפורמציה ההדדית $I(\underline{u}, \underline{v})$ נשמרת. כמו כן עבור מדד עוות רבועי ההתמרה Γ משמרת את מדד העוות הרבועי הנתון ע"י:

$$(4.1) \quad d(\underline{u}, \underline{v}) = \|\underline{u} - \underline{v}\|^2$$

שתי העובדות הנ"ל מאפשרות חשוב של פונקציית הקצב-עוות במישור ההתמרה ולפיכך ביצוע של הקידוד במישור ההתמרה בצורה אופטימלית.

פתרון פרמטרי של פונקציית הקצב-עוות במישור ההתמרה נתון ע"י המשוואות הבאות:

$$(4.2a) \quad D_\theta = \frac{1}{N} \sum_{n=1}^N d_\theta^n = \frac{1}{N} \sum_{n=1}^N \min(\theta, \lambda_n)$$

$$(4.2b) \quad R_N(D_\theta) = \frac{1}{N} \sum_{n=1}^N r_\theta^n = \frac{1}{N} \sum_{n=1}^N \max\left\{0, \frac{1}{2} \log\left(\frac{\lambda_n}{\theta}\right)\right\}$$

כאשר:

- D_θ - העוות הממוצע בין וקטור המקור \underline{u} לוקטור הקוד \underline{v} עבור פרמטר נתון θ .
 d_θ^n - העוות בין האות ה- n של וקטור המקור לאות ה- n של וקטור הקוד.
 r_θ^n - הקצב המתאים לאות ה- n של וקטור המקור \underline{u} .

המימוש של קוד אופטימלי אשר יגשים את פונקציית הקצב-עוות דורש ידיעה של פונקציית הפילוג של וקטור הקוד \underline{v} , $P_N(\underline{v})$ לצורך אכלוס אקראי של המקודד. הפילוג $P_N(\underline{v})$ נתון ע"י:

$$(4.3a) \quad P_N(\underline{v}) = \prod_{n=1}^N P_n(v_n)$$

$$(4.3b) \quad P_n(v_n) = \begin{cases} \frac{1}{\sqrt{2\pi(\lambda_n - \theta)}} \exp\left\{-\frac{v_n^2}{2(\lambda_n - \theta)}\right\} & , \lambda_n > \theta \\ \delta(v_n) & , \lambda_n \leq \theta \end{cases}$$

כאשר v_n הינה האות ה- n של וקטור השחזור \underline{v} .

בסעיף הבא יתואר התהליך בו קובעים את פרמטרי השבכה על סמך פונקציית הקצב-עוות של האות המקודד במישור ההתמרה, ועל סמך פונקציית הפילוג $P_N(\underline{v})$ שתוארה לעיל.

4.3.2 בניית דיאגרמת השבכה

לצורך בניית קוונטיזצור אופטימלי באמצעות דיאגרמת שבכה יש לאכלס את הדיאגרמה בצורה אקראית במילות קוד מתוך פילוג $P_n(v_n)$ (נוסחה 4.2). כאמור, הדרגה ה- l ית של הדיאגרמה מאוכלסת בוקטור קוד v_l המכיל n_l מילות קוד מהפילוג הנ"ל. מילות הקוד המהוות את הוקטורים בענפי השבכה הינן בלתי תלויות סטטיסטית.

הקצב המעשי של כל מילת קוד ברמה l של השבכה נתון ע"י

$$(4.3) \quad \rho_l = \frac{\log(q)}{n_l}$$

הקצב הממוצע הכולל של השבכה הנמדד ב *bits per pel* נתון לפיכך ע"י:

$$(4.4) \quad R = \frac{1}{N} \sum_{l=1}^L n_l \rho_l = \frac{L}{N} \log(q)$$

הפרמטר הראשון הנקבע בדיאגרמה הינו אורך האלוק K . הפרמטר הזה נקבע באופן שרירותי בהתחשב בסיבוכיות המותרת הרחבה בנושא זה הובאה בסעיף 3.4. בהנתן q מסויים אורך הדיאגרמה L נקבע חד משמעית מתוך נוסחה 4.4 כתלות בקצב R הנדרש ובאורך הוקטור המקודד N . בנוסף, כאשר q נתון, סט הקצבים המעשיים האפשרי למימוש בדיאגרמה נתון ע"י

$$(4.5) \quad \rho_l = \frac{\log(q)}{n_l}, \quad l = 1, 2, \dots, N$$

המטרה בקביעת הקצב של כל דרגה בשבכה הינה לקרב את הקצב הנ"ל לקצב האופטימלי, המתאים לאותיות המקור אשר מקודדות בדרגה, אשר נקבע ע"י פונקצית הקצב-עוות (נוסחה 4.1b). דהיינו, אנו מעוניינים להקטין את הביטוי

$$(4.6) \quad \epsilon_N = \frac{1}{N} \sum_{n=1}^{N_c} \epsilon_n = \frac{1}{N} \sum_{l=1}^L \sum_{n \in J_{N_l}} (\rho_l - r_n^*)$$

כאשר:

- N_c הינו מספר מילות הקוד אשר הקצב האופטימלי שנקבע להן, על סמך פונקצית הקצב עוות, גדול מ-0.
- J_{N_l} הינו סט האינדקסים של מילות הקוד ברמה l של השבכה, הנתון ע"י:

$$(4.7a) \quad J_{N_l} = \{N_{l-1} + 1, N_{l-1} + 2, \dots, N_{l-1} + n_l\}$$

$$(4.7b) \quad N_{l-1} = \sum_{i=1}^{l-1} n_i, \quad N_0 = 0, \quad N_L = N_c = \sum_{i=1}^L n_i$$

האלגוריתם המדוייק לקביעת ערכי n_l , שהוא למעשה האלגוריתם להקצאת ביטים בשבכה, מובא

להלן:

אלגוריתם הקצאת הביטים בשבכה:

(1) קבע את q שרירותית אך גדול מספיק כך שיתקיים

$$(4.8) \quad \log(q) \geq \max_n \{r_\theta^n\}$$

(2) קבע את סט מספרי האכלוס $\{n_i\}$ האפשריים:

א. n_{\min} נקבע כמספר השלם החיובי הגדול ביותר המקיים

$$(4.9) \quad \frac{\log(q)}{n_{\min}} > \max_n \{r_\theta^n\}$$

ב. n_{\max} נקבע כמספר השלם הקטן ביותר המקיים

$$(4.10) \quad \frac{\log(q)}{n_{\max} + 1} < \min_n \{r_\theta^n\}$$

הקבוצה של מספרי האכלוס $\{n_i\}$ כוללת את כל המספרים השלמים בין n_{\min} ו- n_{\max} כולל. בהתאמה, קבוצת הקצבים האפשריים נקבעת עפ"י נוסחה 4.3.

ג. קבע את הקצב המעשי של אות המקור u_n , בעלת קצב תיאורטי r_θ^n , כ- ρ_l אם מתקיים $\rho_l > r_\theta^n > \rho_{l+1}$. הדבר נעשה ע"י אכלוס של n_l מילות קוד בדרגה המתאימה בשבכה. כאשר קיימות פחות מ- n_l מילות מקור בעלות קצב אופטימלי בתחום הנ"ל, אזי מילות מקור עם קצב אופטימלי $r_\theta^n < \rho_{l+1}$ יקודדו עם קצב ρ_l , ומילות הקוד המתאימות להן יאוכלסו באותה דרגה.

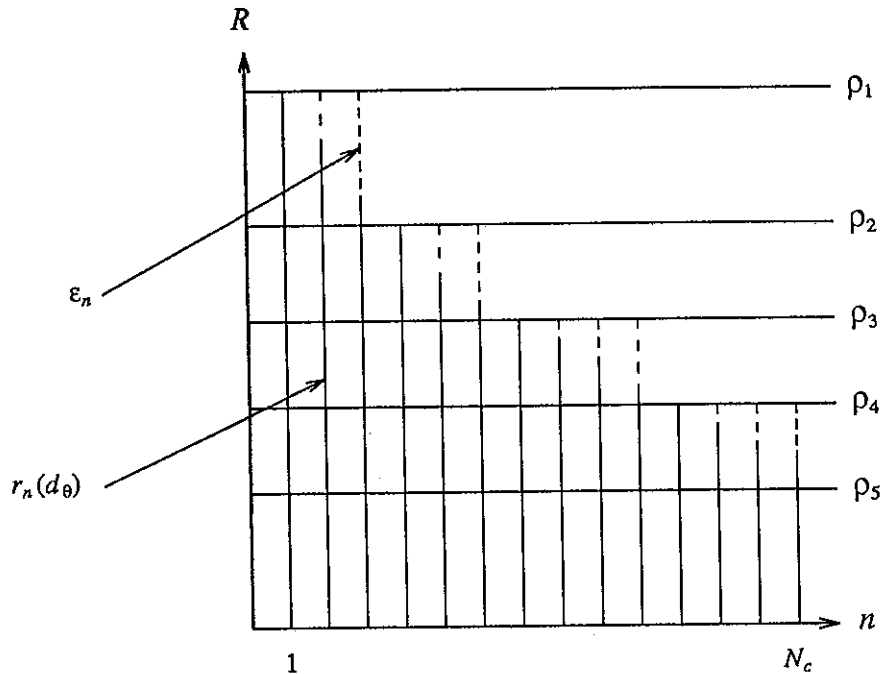
כאשר קיימות יותר מ- n_l מילות מקור עם קצב אופטימלי בתחום הנ"ל כל n_l מילות מקור עוקבות יאוכלסו בדרגה נפרדת עד שנקבל את המקרה הקודם.

אלגוריתם הקצאת הביטים שתואר לעיל קובע את סט הקצבים לכל דרגה בשבכה $\{\rho_l\}_{l=1}^L$. תאור

סכמתי של הקצאת הביטים בשבכה מובא בציור 4.2.

הקטנה של הסטייה ε_N (נוסחה 4.6) יכולה להתבצע ע"י הגדלת מספר הקצבים בסט האפשרי או ע"י הגדלת הוקטור N . הגדלת סט הקצבים מתאפשרת ע"י הגדלת q אשר גוררת הגדלת קבוצת מספרי האכלוס ובעקבות כך הגדלת קבוצת הקצבים האפשריים. הגדלת אורך הוקטור N גורמת לעידון של קבוצת הקצבים האופטימליים r_θ^n הנוצרת מפונקציית הקצב-עוות עובדה המקטינה את כמות המקרים בהם אותיות מקור בעלות קצב אופטימלי $r_\theta^n < \rho_{l+1}$ יאוכלסו בדרגה בעלת קצב ρ_l .

בהתאם למשפט שהוכח ב [1] קיימת דיאגרמת שבכה שנבנתה בדרך שתוארה לעיל ואוכלסה באקראי מפילוג היציאה של הפילוג המגשים של פונקציית הקצב-עוות, אשר עבור עוות נתון D_θ כאשר אורך האילוף K שואף לאינסוף הקצב R של הדיאגרמה קרוב כרצוננו לערך $R(D_\theta)$ המתקבל מפונקציית הקצב - עוות.



ציור 4.2 - הקצאת הביטים בשבכה.
 Fig 4.2 - Bit allocation in the trellis.

4.4 איכלוס הדיאגרמה

בסעיף הקודם תארנו כיצד קובעים את המבנה הפיזי של דיאגרמת השבכה. השלב הבא בתהליך
 בנית המקודד הינו איכלוס הדיאגרמה במילות קוד מתאימות, דהיינו מציאת המילון המאכלט את
 הדיאגרמה. במקרה האופטימלי היינו מעוניינים לאכלט את הדיאגרמה, כך שנקבל את מינימום תוחלת
 העוות בין וקטורי המקוד (וקטורי הכניסה) לבין וקטורי השחזור (וקטורי היציאה) המתאימים להם,
 המאכלטים את הדיאגרמה. התהליך של מציאת אכלוס אשר יביא למינימום את תוחלת העוות אינו ידוע.
 בסעיף זה נתאר שתי שיטות לאכלוס דיאגרמת השבכה. שיטות אלו אמנם אינן מבטיחות מינימום של
 תוחלת העוות, אולם הן מאפשרות פתרון סביר של הבעיה.
 השיטה הראשונה נקראת אכלוס אקראי והיא מבוססת על משפט צפינת המקורות באמצעות דיאגרמת
 שבכה. השיטה השניה מנצלת אלגוריתם איטרטיבי בדומה לאלגוריתם LBG (סעיף 3.3) לצורך אכלוס
 השבכה באמצעות סדרת לימוד. בסכמה שמומשה על ידנו בחרנו להשתמש באכלוס אקראי עקב פשטותו
 היחסית בהשוואה לאכלוס באמצעות אלגוריתם איטרטיבי.

4.4.1 אכלוס אקראי

משפט צפינת המקורות באמצעות דיאגרמת שבכה [24] מוכיח שלכל מקור סטציונרי וחסר זכרון
 בעל מידת עוות לאות מקור חסומה, קיים מקודד שבכה משתנה בזמן, בעל הפרמטרים K, q , שעבורו
 תוחלת העוות לאות מקור D חסומה. הוכחת המשפט מתבססת על מצוע ע"פ כל הקודים המוגרלים

באקראי מתוך פילוג היציאה $P(\underline{y})$, המתקבל מתוך הפילוג המגשים $P(\underline{y} | \underline{u})$ של פונקצית הקצב-עוות ע"י:

$$P(\underline{y}) = \int P(\underline{y} | \underline{u}) \cdot P(\underline{u}) d\underline{u}$$

מתוך העובדה שהמשפט מתקיים עבור ממוצע הקודים המוגרלים באקראי נובע, שקיים לפחות קוד אחד המקיים אף הוא את המשפט. מתוך שיטת ההוכחה נובעת המוטיבציה לשיטת האיכלוס האקראי.

בסכמה המוצעת על ידנו אנו מאכלסים את ענפי השבכה בפילוג $P(\underline{y})$ הנתון ב-4.2, המהווה את פילוג היציאה המתקבל מהפילוג המגשים לפונקצית הקצב עוות עבור מקור גאוסי-מרקובי. השיטה של אכלוס באקראי טובלת ממספר חסרונות כדלקמן:

- א. אין בטחון שהקוד הנבחר באקראי אכן מקיים את משפט הצפינה ומהווה קוד טוב.
- ב. בהרבה מקרים ובמיוחד בתמונות, סטטיסטיקת המקור אינה ידועה, ולפיכך אין אפשרות לחשב עבודה פונקצית קצב-עוות ופילוג יציאה לאכלוס.
- ג. תנאי המשפט אשר מהווה מוטיבציה לאכלוס האקראי, דורשים שמידת העוות לאות מקור תהיה חסומה. לפיכך עבור מקורות גאויסיים ולפלס, אשר מהווים בד"כ מודלים לאיפיון תמונות, ומדד עוות שגיאה רבועית או שגיאת ערך מוחלט התנאי הנ"ל אינו מתקיים. תוצאות אמפיריות שהתקבלו ב- [25] תומכות בהשערה שהמשפט תקף גם עבור מקור גאוסי ומדד עוות שגיאה רבועית.

4.4.2 אֶלְגוֹרִיתֵם אִיטְרָטִיבִי לֶאכְלוֹס הַשֶּׁבֶכָה

שיטה אחרת לאכלוס השבכה, אכלוס איטרטיבי באמצעות סדרות לימוד מוצגת ב [26]. שיטה זו דומה במהותה לשיטת איכלוס ספר קוד מתוך סדרות לימוד שהוצגה בסעיף 3.3. האלגוריתם מורכב משני שלבים עיקריים. בשלב הראשון מקודדות סדרת הלימוד ע"י דיאגרמת שבכה בעלת אכלוס נתון כלשהו. בשלב השני מחושב האכלוס האופטימלי מתוך סידרת הלימוד והקידוד שנמצא בשלב הראשון. התהליך הנ"ל מבוצע בצורה איטרטיבית לקבלת עוות ממוצע קטן יותר בכל שלב בין סדרת הלימוד והשחזור שלה באמצעות דיאגרמת השבכה. תאור של אופן בחירת האכלוס ההתחלתי ותאור תוצאות סימולציות מובאות בהרחבה ב - [26].

החסרון העיקרי באיכלוס באמצעות האלגוריתם הנ"ל טמון בסיבוכיות הגבוהה הכרוכה במימושו. בכל איטרציה יש לבצע קידוד של סדרת לימוד ארוכה באמצעות שבכה, הסיבוכיות של קידוד כזה גבוהה ביותר. היתרון העיקרי של איכלוס בשיטה הנ"ל הוא שאין צורך לדעת את סטטיסטיקת המקור. כאמור עקב הסיבוכיות הגבוהה של שיטה זו בחרנו להשתמש בשיטה של אכלוס אקראי בסכמה המוצעת על ידנו.

4.5 אלגוריתמי חיפוש בשבכה

הקידוד של וקטורים באמצעות דיאגרמת שבכה דורש חיפוש של מסלול על פני הדיאגרמה, אשר מילות הקוד המאכלסות אותו יהוו קרוב טוב לוקטור המקור. הפרמטרים החשובים בהערכת טיב אלגוריתם חיפוש הינם מידת העוות המתקבלת בין וקטור המקור לבין וקטור השחזור והסיבוכיות של האלגוריתם. הסיבוכיות של אלגוריתם חיפוש מאופיינת באמצעות שני פרמטרים: סיבוכיות חישוב (Co) וכמות זכרון (Me) [11]. סיבוכיות החישוב מוגדרת כמספר הענפים עבורם צריך לחשב את מדד העוות כדי לקודד אות מקור בודדת. כמות הזכרון הינה מספר המילים הכולל אותן יש לשמור בזכרון. חיפוש מלא (Exhaustive Search) על פני כל המסלולים האפשריים יתן כמובן את המסלול האופטימלי, דהיינו את המסלול אשר העוות בין וקטור השחזור המאכלס אותו לבין וקטור המקור הוא המינימלי האפשרי עבור דיאגרמת שבכה נתונה. הסיבוכיות של חיפוש מלא הינה גבוהה מאוד, סיבוכיות החישוב עבור וקטור באורך N נתונה ע"י:

$$(4.11) \quad Co = 2^{RN}$$

כמות הזכרון הנדרשת במקרה זה נתונה ע"י:

$$(4.12) \quad Me = N2^{RN}$$

בסעיף זה נתאר את האלגוריתם של ויטרבי (Viterbi) אשר מאפשר מציאת המסלול האופטימלי בסיבוכיות קטנה מזו של החיפוש המלא. כן נתאר אלגוריתמי חיפוש תת-אופטימליים אשר משיגים ביצועים אשר מתקרבים לביצועים האופטימליים תוך הפחתה נוספת בסיבוכיות החיפוש.

4.5.1 האלגוריתם של ויטרבי

האלגוריתם של ויטרבי מבוסס על עקרון האופטימליות בתכנות דינמי האומר, שאם מסלול אופטימלי מנקודה A לנקודה C עובר דרך הנקודה B, אזי קטע המסלול בין הנקודה B לנקודה C מתלכד עם המסלול האופטימלי בין הנקודה B לנקודה C. מספר המצבים האפשריים עבור דיאגרמה בעלת אורך אילוף K ו- q מצבים היוצאים מצומת הינו q^{K-1} . נניח כי ידועים המסלולים האופטימליים מהדרגה הראשונה של השבכה לכל אחד מ- q^{K-1} המצבים בדרגה l של השבכה, כמו כן נניח שידוע העוות המצטבר של כל אחד מהמסלולים הנ"ל שייסומן כ- $d(j, l)$, לכל אחד מ- q^{K-1} המצבים בדרגה $l+1$ נתן להגיע מ- q מצבים בדרגה l שייסומנו כ- j_k כאשר $k = 1, 2, \dots, q$. העוות המתאים לכל אחד מהמעברים הנ"ל מהדרגה l לדרגה $l+1$ יסומן כ- $d(j_k)$. המסלול האופטימלי לכל אחד מ- q^{K-1} המצבים בדרגה $l+1$ הינו המסלול אשר העוות המצטבר המתאים לו הוא מינימלי מבין כל המסלולים המגיעים אליו, דהיינו:

$$(4.13) \quad d(j, l+1) = \min_k \{d(j_k, l) + d(j_k)\}$$

נסמן ב- k^* את האינדקס k שמקיים את המינימום ב-4.13. המסלול האופטימלי למצב j בדרגה $l+1$ של השבכה מורכב מהמסלול האופטימלי למצב j_{k^*} בדרגה l ומהענף המחבר את המצב j_{k^*} עם המצב j בדרגה $l+1$. האלגוריתם שתארנו הוא רקורסיבי וניתן להפעילו מדרגה 1 בשבכה עד הדרגה האחרונה L . בסיום התהליך מגיעים לדרגה L בשבכה עם q^{K-1} מסלולים ומתוכם בחורים את המסלול בעל העוות

המינימלי, סדרת האינדקסים המתאימה למסלול הנבחר נשלחת דרך הערוץ למקלט. במהלך האלגוריתם יכול להתרחש מאורע בו מגיעים למצב j בדרגה מסויימת שני מסלולים בעלי עוות זהה מינימלי, במקרה זה בוחרים ביניהם שרירותית (למשל המסלול העליון בשבכה).

כמות החישובים הנדרשת לקידוד וקטור באורך N באמצעות שבכה בעלת אורך אילוף K , אורך L ו- q עפים היוצאים מצומת, תוך שמוש באלגוריתם של ויטרבי נתונה ע"י $q^K = 2^{RK}$. כמות הזכרון הנדרשת במקרה זה הינה Nq^K . ניתן לראות שכמות החישובים לאות מקור תלויה בפרמטרי השבכה K ו- q בלבד, ואילו כמות הזכרון תלויה לינארית באורך הקטור N וזאת בהשוואה לתלות האקספוננציאלית (נוסחאות 4.11 4.12) הקיימת בחיפוש מלא.

4.5.2 אלגוריתם M

אלגוריתם M הינו אלגוריתם תת-אופטימלי לחיפוש בשבכה, אלגוריתם M מבוסס על אלגוריתם ויטרבי שתואר בסעיף הקודם. השוני בינו לבין האלגוריתם של ויטרבי הינו העובדה שבאלגוריתם M הפיצול של הענפים מהדרגה l לדרגה $l+1$ הינו מתוך M מצבים בעלי מדד עוות מצטבר הקטן ביותר מתוך q^{K-1} המצבים הקיימים, כפי שמתבצע באלגוריתם של ויטרבי. סיבוכיות החישוב באלגוריתם M הינה M ואינה תלויה כלל באורך האילוף K . כמות הזכרון הנדרשת באלגוריתם M הינה $MNq = MN2^R$. Wilson et al. השוו ב [25] את הביצועים של קודד שבכה תוך שמוש באלגוריתם ויטרבי ובאלגוריתם M. במאמר נבדקו שתי דיאגרמות עם אורכי אילוף K_1, K_2 כאשר $K_1 < K_2$. הדיאגרמה הראשונה קודדה באמצעות האלגוריתם של ויטרבי, ואילו הדיאגרמה השנייה קודדה באמצעות אלגוריתם M באותה סיבוכיות חשוב דהיינו $q^{K_1} = Mq$. התוצאות שהתקבלו הצביעו על עדיפות של הסכמה שקודדה באמצעות אלגוריתם M. למעשה, לקבלת ביצועים זהים, נדרש בסכמה הממומשת באמצעות האלגוריתם של ויטרבי, אורך אילוף הגורם לסיבוכיות חישוב של פי 2-2.5 מזו הקיימת בסכמה הממומשת באמצעות אלגוריתם M.

4.5.3 אלגוריתם ML

באלגוריתם ויטרבי ובאלגוריתם M שתוארו לעיל ההחלטה על המסלול הנבחר בדיאגרמה התקבלה לאחר סיום החיפוש על פני כל הדיאגרמה. עובדה זו גוררת בעקבותיה את הצורך לשמור את מילות השחזור של המסלולים האופטימליים על פני אורך כל הדיאגרמה. בנוסף, העובדה הנ"ל גורמת השהיה בשידור המסלול הנבחר עד לאחר גמר הקידוד. אלגוריתם ML פותר חלקית את הבעיות הנ"ל בכך שהחל מהדרגה $L_D + 1$ מוצאים את המסלול הטוב ביותר מבין Mq המסלולים שהגיעו לדרגה הנדונה והאינדקס המתאים לענף הנמצא L_D דרגות אחורה במסלול משודר למקלט. כמות הזכרון במקרה זה מוגבלת לכמות הזכרון הנדרשת לשמירת המסלולים באורך L_D בלבד. ההשהיה המתקבלת במקרה זה קטנה אף היא מ- L ל- L_D . הסיבוכיות הכרוכה במימוש המעשי של האלגוריתם ML גדולה מזו הקיימת באלגוריתם M ובאלגוריתם ויטרבי. הדגמה ראה במאמר [25].

בסכמה הממומשת על ידנו אנו משתמשים באלגוריתם M עקב העובדה שביצועיו מזדהים למעשה אם אלה המתגים באלגוריתם ויטרבי ופשטות המימוש שלו יחסית לאלגוריתם של ויטרבי ואלגוריתם ML.

4.6 שיקולי ממוש מעשיים

הממוש של סכמת הקידוד שהוצגה בסעיף הקודם הינו בעייתי עקב הבעיות הבאות:

א. החישוב של התמרת KL הינו בעיה מסובכת ויקרה מבחינת זכרון וסיבוכיות החישוב, עקב העובדה שנדרש חישוב של מטריצת הקואריאנס ושל הוקטורים והערכים העצמיים שלה (ראה סעיף 2.2).

ב. סכמת הקידוד מיועדת לתמונות סטציונריות, ולכן האופי הלא סטציונרי של התמונות המקודדות יוצר בעיה.

בסעיף זה מתוארים שיקולי מימוש מעשיים אשר מאפשרים יישום של סכמת הקידוד ברמת סיבוכיות וביצועים סבירים.

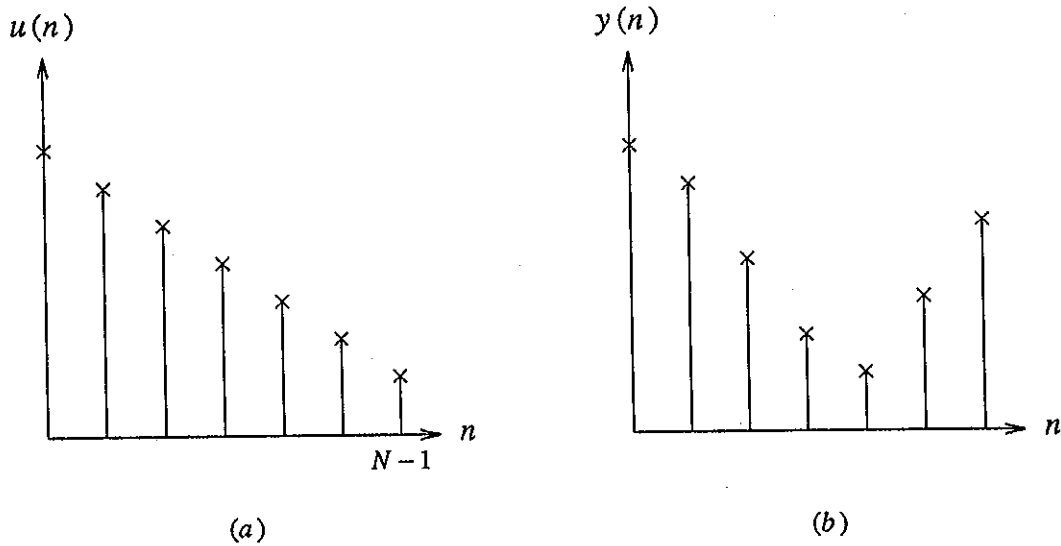
4.6.1 התמרת DCT - Cosine

הסכמה האופטימלית ממומשת באמצעות התמרת KL. עקב הסיבוכיות של התמרת KL, ועקב העובדה שהיא תלויה בסטטיסטיקת האות, אנו משתמשים ביישום המעשי של הסכמה בהתמרת DCT המוגדרת בסעיף 2.3. כפי שראינו, DCT הינה ההתמרה בעלת הביצועים הקרובים ביותר לביצועי KLT מבין כל ההתמרות התת-אופטימליות עבור אות בעל מודל מרקובי.

התמרת DCT הינה התמרה קבועה אשר נתנת לממוש באמצעות אלגוריתם מהיר. האלגוריתם המשמש אותנו לחישוב DCT הינו האלגוריתם שהוצג ב-[23] והמבוסס על אלגוריתם ה-FFT. האלגוריתם מבוסס על ביצוע FFT על סדרה $y(n)$ המתקבלת מהסדרה המקורית $u(n)$ באורך N ע"י:

$$(4.14) \quad y(n) = \begin{cases} u(2n), & 0 \leq n \leq \left\lfloor \frac{N-1}{2} \right\rfloor \\ u(2N-2n-1), & \left\lfloor \frac{N+1}{2} \right\rfloor \leq n \leq N-1 \end{cases}$$

כאשר $[a]$ מסמל את הערך השלם של a . תאור גרפי של הסדרה המקורית $u(n)$ והסדרה המסודרת מחדש $y(n)$ מובא בציר 4.3.



ציור 4.3 - (a) הסדרה המקורית $u(n)$ (b) הסדרה המסודרת $y(n)$.
 Fig 4.3 - (a) The original sequence $u(n)$ (b) The reordered sequence $y(n)$.

נסמן את ה DFT של $y(n)$ ב $Y(k)$. ואת ה DCT של $u(n)$ ב $U_C(k)$.
 $U_C(k)$ ו $Y(k)$ מקיימות את הקשר הבא:

$$(4.15) \quad Y(k) = \frac{1}{2} W_{4N}^{-k} [U_C(k) - jU_C(N-k)]$$

כאשר:

$$W_{4N}^{-k} = \exp\{-j\frac{2\pi k}{4N}\}$$

מתוך נוסחה 4.15 ניתן לראות שהתמרת DCT של $u(n)$ מתקבלת מהתמרת פורייה של הסדרה $y(n)$ ע"י:

$$(4.16) \quad \begin{cases} U_C(k) = \text{Re} \{2W_{4N}^k Y(k)\} \\ U_C(N-k) = -\text{Im} \{2W_{4N}^k Y(k)\} \end{cases}$$

כאשר:

$$0 \leq k \leq \left\lfloor \frac{N}{2} \right\rfloor$$

התמרת IDCT מתקבלת ע"י חשוב $Y(k)$ לפי נוסחה 4.14, חשוב IDFT של $Y(k)$ לקבלת $y(n)$ וחשוב $u(n)$ מתוך $y(n)$ לפי נוסחה 4.13. הרחבה של ההסבר הנ"ל למקרה הדו- ממדי מובאת ב [23].

4.6.2 בניית דיאגרמת השבכה ואכלוסה

בניית דיאגרמת השבכה ואיכלוסה כפי שתוארו בסעיפים הקודמים נעשו תחת הנחה שהתמונה המקודדת הינה סטציונרית ומקיימת את המודל הגאומי - מרקובי. בניית הדיאגרמה נעשתה תוך שמוש בערכים העצמיים של מטריצת הקוואריאנס של התמונה. אחת הבעיות החמורות בהן נתקלים בדיחיסת תמונות הינה חוסר הסטציונריות שלהן, אשר מקשה על שיערוך מטריצת (פונקצית) הקוואריאנס והערכים העצמיים של התמונות. כדי להתגבר על הבעיה נעשתה הנחה בתכנון הסכמה הבסיסית שהתמונות הינן קוואזי-סטציונריות, דהיינו מקיימות תכונה של סטציונריות באזור מקומי. הנחה זו מאפשרת שיערוך של מטריצת הקוואריאנס וחשוב הערכים העצמיים שלה עבור כל אזור.

כאמור בסעיף 4.6.1 התמרת DCT משמשת אותנו במימוש סכמת הקידוד במקוט התמרת KLT. במקרה זה הערכים העצמיים המשמשים לקביעת מבנה השבכה משוערכים מתוך הוואריאנסים של מקדמי התמונה במישור ההתמרה. הערכים העצמיים המשוערכים נשלחים בתור אינפורמציה צד למקלט, ומשמשים שם לבניית הדיאגרמה ואכלוסה.

סעיף זה מתאר את התהליך של שיערוך הערכים העצמיים. אלגוריתם השיערוך מבוסס על ההנחה שאזור המורכב מבלוקים הנמצאים בשכנות בתמונה הינו סטציונרי. בשלב ראשון יש לחלק את התמונה לאזורים הכוללים M_a בלוקים כ"א. שיערוך הערכים העצמיים של אזור כזה בתמונה מתואר באלגוריתם הבא:

אלגוריתם שיערוך הערכים העצמיים:

(1) סרוק את M_a הבלוקים באזור בשיטה כלשהי (למשל Raster Scan), לקבלת M_a וקטורים חז- ממדיים באורך N כ"א.

(2) קבע את סדרת הוואריאנסים המנורמלים \tilde{S}_l עבור M_a הבלוקים בקבוצה לפי הנוסחה הבאה:

$$(4.17) \quad \tilde{S}_l = \frac{1}{M_a} \sum_{i=1}^{M_a} \frac{(u_i^l)^2}{P_i}, \quad l=1,2,\dots,N$$

כאשר:

- \tilde{S}_l - הוואריאנס המנורמל של המקדם l בוקטור.
- M_a - מספר הוקטורים באזור.
- N - אורך כל וקטור.
- u_i^l - המקדם ה- l בוקטור ה- i מתוך M_a הוקטורים.
- P_i - ההספק הממוצע של הוקטור i הנתון ע"י:

$$(4.18) \quad P_i = \sum_{l=1}^N (u_i^l)^2, \quad l=1,2,\dots,N$$

(3) מצע את הוואריאנטים המנורמלים \tilde{S}_l של M_s אברים סמוכים לקבלת $N_s = \frac{N}{M_s}$ ואריאנטים משוערכים מנורמלים \bar{S}_l .

$$(4.19) \quad \bar{S}_l = \frac{1}{M_s} \sum_{i=l}^{l+M_s-1} \tilde{S}_i, \quad l = 1, 1+M_s, 1+2M_s, \dots, 1+N-M_s$$

(4) ערכי הוואריאנטים המשוערכים המנורמלים, $\{\bar{S}_l\}_{l=1}^N$ המשמשים לבנית השבכה של האזור הנדון מתקבלים ע"י אינטרפולציה לינארית בין N_s ערכי \tilde{S}_l שהתקבלו בטעיף ג.

(5) חשב את מקדם ההגבר לוקטור ה- i :

$$(4.20) \quad G_i = \frac{P_i}{\sum_{l=1}^N \bar{S}_l}, \quad i = 1, 2, \dots, M_a$$

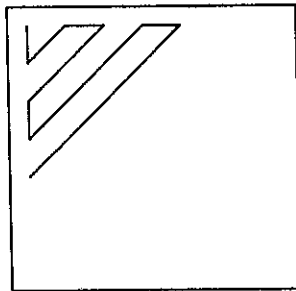
(6) ערכי הוואריאנטים המשוערכים (הספקטרום המשוערך) הנדרשים לצורך אכלוס השבכה עבור הבלוק ה- i מתקבלים ע"י:

$$(4.21) \quad \{\hat{S}_i\}_{i=1}^N = G_i \cdot \{\bar{S}_l\}_{l=1}^N$$

כאשר $\{\bar{S}_l\}_{l=1}^N$ נקבעו בשלב (4) ו G_i נקבעו לפי נוסחה 4.21

התהליך שתואר לעיל חוזר על עצמו עבור כל אזור.

באלגוריתם שמומש על ידינו, סריקת הבלוק הדו-ממדי לשם הפיכתו לוקטור, כפי שנדרש בשלב (1) של האלגוריתם, בוצע בצורת זיגזג המתוארת בציור 4.4. הסריקה הנ"ל נתנה את הביצועים הטובים ביותר בהשוואה לשיטות סריקה אחרות שנוסו (כגון: סריקה לפי שורות וסריקה לפי עמודות).



ציור 4.4 - סריקת בלוק בצורת זיגזג.

Fig 4.4 - Zigzag scanning of a block.

החלקה בשלבים (3) ו (4) באלגוריתם בוצעה כדי להוריד את כמות אינפורמציה הצד הנשלחת למקלט, והן כדי לקבל ספקטרום מוחלק המשפר את ביצועי הקוד. אינפורמציה הצד הנשלחת לצורך בנית השבכה במקלט כוללת את הנתונים הבאים:

- א. N_s ערכי הוואריאנט המנומל המשוערך \bar{S}_l הנתונים בנוסחה 4.19.
- ב. M_a מקדמי ההגבר G_l הנתונים בנוסחה 4.7.
- ג. הפרמטר θ הנדרש לצורך בניית השבכה, המתקבל מפונקצית הקצב-עוות נוסחה 4.1.
- הוואריאנטים המשוערכים המנומלים $\{\bar{S}_l\}_{l=1}^N$ המתקבלים בשלב (4) של האלגוריתם משמשים לבניית השבכה עבור האזור הנדון כפי שמתואר בסעיפים 4.3 ו 4.4. בניית השבכה מתבצעת מחדש עבור כל M_a בלוקים המהווים אזור סטציונרי.
- איכלוס השבכה עבור כל בלוק מתבצע באקראי כמתואר בסעיף 4.5.1 לפי פילוג היציאה הנתון בנוסחה 4.2 כאשר הוואריאנטים $\{\hat{S}_l\}_{l=1}^N$ ששוערכו לפי נוסחה 4.21 משמשים כוואריאנטים המופיעים בפילוג. ביצועי המערכת הבסיסית מודגמים ונבחנים בהשוואה לסכמה הסופית בפרק 6.

פרק 5 - בעיית אי התאמת התמונות למודל ושיטות לפתרונה

5.1 אי התאמת התמונות למודל - תאור הבעיה

סכמת הקידוד אשר הוצגה בפרק הקודם הינה אופטימלית עבור אותות המקיימים את המודל הגאוסי - מרקובי. רוב התמונות הטבעיות אינן מקיימות את המודל הנ"ל. יתרה מכך, גם שמוש בהנחה מקילה של קוואזי - סטציונריות ומימוש הסכמה על תת-אזורים של התמונה בנפרד, כפי שנדונו בפרק הקודם, לא פתרו את הבעיה. לפיכך עלה הצורך בשיפור הסכמה כך שתוכל לטפל בעילות בתמונות החורגות מהמודל הגאוסי מרקובי.

הנחת הגאוסיות של מקדמי ההתמרה נבעה מהעובדה שמקדמי ה-DCT, ההתמרה אשר משמשת אותנו כקרוב להתמרת KL, מתקבלים כסכום משוקלל של משתנים אקראיים. ב- [4 p 275] נטען שלמרות העובדה שאברי התמונה אינם בהכרח בלתי תלויים סטטיסטית, ולכן אינם מקיימים את תנאי משפט הגבול המרכזי, מקדמי ההתמרה המתקבלים כסכום משוקלל של אברי התמונה נוטים להיות בעלי פילוג גאוסי. עבודות מאוחרות יותר [28, 29] מראות שההנחה שהמקדמים הינם בעלי פילוג לפלס מתאימה יותר מאשר הנחת הפילוג הגאוסי. בעבודות הנ"ל בוצעו ונבדקו ביצועי סכמות קידוד אשר תכנון מבוסס על פילוג התמונה. התוצאות שהתקבלו הראו שהנחת פילוג לפלס של מקדמי ההתמרה עדיפה על פני הנחת הפילוג הגאוסי. בעבודתנו נשארנו צמודים להנחת המודל הגאוסי עקב העובדה שאופטימליות הסכמה בה אנו משתמשים הוכחה עבור מודל זה בלבד.

הבעיה המהותית יותר באי התאמת התמונות למודל הינה חוסר הסטציונריות האינהרנטית הקיימת בתמונות טבעיות. תופעה זו גורמת לירידה באיכות התמונות המשוחזרות, ואחד ממאפייניה העיקריים הינו אפקט ה"בלוקיות" המופיע בין היתר בתמונות אשר קודדו באמצעות מקודדי התמרה הפועלים בקצב נמוך. יתרה מכך, כפי שראינו בפרק הקודם, לא ניתן להניח קוואזי סטציונריות באזורים. לפיכך, היה צורך לבצע חלוקה אדפטיבית (משתנה מתמונה לתמונה) של התמונה לקבוצות (Clusters) של בלוקים בעלי תכולה ספקטרלית דומה, ולהתייחס אל אותם בלוקים, בלי חלות במקומם הפיזי בתמונה, כאל בלוקים השייכים לאותו מקור סטציונרי. האפיון של בעיית ההקבצה (Clustering) הנ"ל ופתרונה יתוארו בסעיף 5.2. פועל יוצא של החלוקה לקבוצות הינו הצורך בהקצאת קצבים שונים לקבוצות השונות, ובניית דיאגרמת שבכה שונה לכל קבוצה. נושא זה יתואר בסעיף 5.3. השימוש בהקבצה, אשר שיפר במידה משמעותית את ביצועי הסכמה, לא פתר לחלוטין את בעיית אי התאמת התמונות למודל, ובתמונות שקודדו עדין התגלו תופעות של בלוקיות. בסעיף 5.4 נתאר מנגנון העושה שמוש בקוונטיזר סקלרי לקידוד במישור ההתמרה של תמונת השארית בין תמונת המקור לתמונה המשוחזרת. השימוש במנגנון הנ"ל פתר כמעט לחלוטין את בעיית הבלוקיות בתוספת סיבוכיות יחסית זניחה. לסיום הפרק, בסעיף 5.5, נתאר את הסכמה הכוללת שמומשה על ידינו.

5.2 שמוש בהקבצה לפתרון בעיית חוסר הסטציונריות

חוסר הסטציונריות הקיימת בתמונת, ואף בין בלוקים שכנים בתמונה, אילץ אותנו למצוא תהליך אשר יבצע חלוקה של התמונה לקבוצות המכילות בלוקים בעלי תכולה ספקטרלית דומה. בלוקים אלה ישמשו אותנו כמקור סטציונרי ועבורם תבנה דיאגרמת שבכה מתאימה. עקב השוני בין תמונות, תהליך החלוקה הנ"ל צריך להתבצע עבור כל תמונה מחדש, ומסיבות מעשיות התהליך הנ"ל אמור להיות אוטומטי לחלוטין.

בעיית ההקבצה דנה בחלוקה של קבוצה W המכילה s וקטורים $\{w_i \mid i=1,2,\dots,s\}$ ל c קבוצות $C_k, k=1,2,\dots,c$ המיוצגות כל אחת ע"י וקטור הממוצע m_k של הוקטורים בקבוצה הנתון ע"י

$$(4.1) \quad m_k \triangleq \frac{1}{\gamma_k} \sum_{i=1}^{\gamma_k} w_i \quad s.t. \quad w_i \in C_k$$

כאשר: γ_k הינו מספר הוקטורים בקבוצה C_k (הקרדינליות של הקבוצה C_k) יהי $\delta(w_i, m_k)$ מדד מרחק בין הוקטור w_i והממוצע m_k המייצג את הקבוצה C_k (מרחק אאוקלידי במקרה שלנו). תהי C קבוצת כל החלוקות של W ל c קבוצות הנתונה ע"י:

$$(5.2) \quad C = \{ C_k \mid k=1, \dots, c; \bigcap_{k=1}^c C_k = \emptyset; \bigcup_{k=1}^c C_k = W \}$$

מבין כל החלוקות האפשריות C , יש למצוא את החלוקה המביאה למינימום את סך כל המרחקים בין הוקטורים $w_i \in C_k$ לבין הוקטור המייצג m_k עבור כל הקבוצות. דהיינו, מציאת החלוקה המביאה למינימום את פונקציית המדד $J(C)$ הנתונה ע"י:

$$(5.3) \quad J(C) = \sum_{k=1}^c \sum_{i=1}^{\gamma_k} \delta(w_i, m_k) = \sum_{k=1}^c J_k(C), \quad w_i \in C_k$$

נושא ההקבצה מכוסה בהרחבה במקור [30] אשר מתאר שתי גישות בסיסיות לביצוע הקבצה:

א. שיטת ה- (DC) Dynamic Clustering אשר משתמשת באלגוריתם איטרטיבי לביצוע אופטימיזציה של פונקציית המדד $J(C)$ הנתונה ב 5.3

ב. שיטת ה- (HC) Hierarchical Clustering אשר משתמשת ב s הוקטורים המרכיבים את הקבוצה W כחלוקה התחלתית, וע"י אחד של שתי הקבוצות הקרובות ביותר מתבצעת הקטנה של מספר הקבוצות. תהליך האחד נמשך עד לקבלת מספר הקבוצות הנדרש.

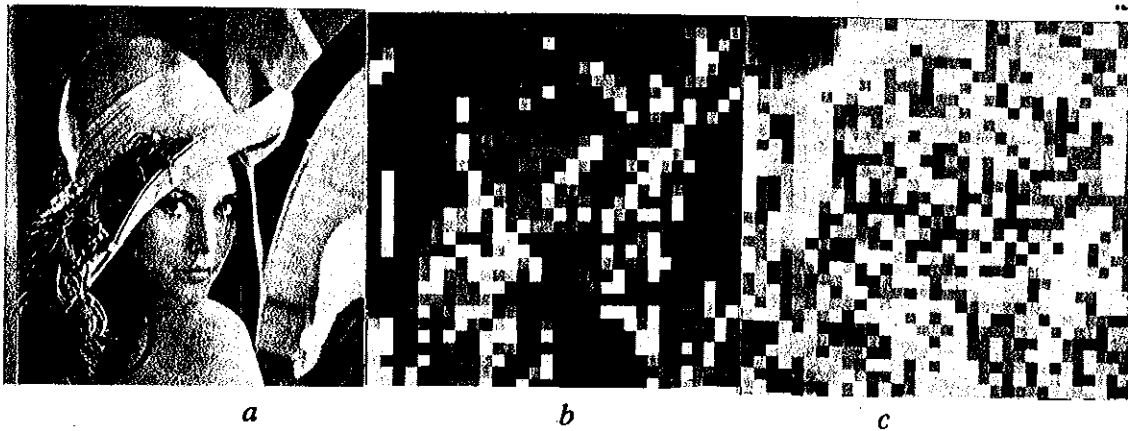
שיטת ה HC הינה תת-אופטימלית, דהיינו היא אינה מתכנסת בהכרח למינימום של פונקציית המדד $J(C)$. מבחינת סיבוכיות שיטה זו מבטיחה לכאורה סיבוכיות קטנה יותר, עקב העובדה שהיא סדרתית ובכל איטרציה האלגוריתם עובר פעם אחת בלבד ע"פ הוקטורים. למעשה התברר שסיבוכיות השיטה גדולה מזו הנדרשת בשיטת ה DC . הסיבה לכך הינה המספר הקטן יחסית של קבוצות הנדרש בהשוואה למספר הכולל של וקטורים (למשל $s=1024$ ו- $C=16, 32, 64$). עובדה זו גרמה למספר גדול של איטרציות בעת הפעלת האלגוריתם. שיטת ה- DC מתכנסת למינימום של הפונקציה $J(C)$ (מינימום זה עלול להיות מקומי בלבד).

בנוסף, כמות האיטרציות הנדרשת להתכנסות האלגוריתם הינה קטנה יחסית.

עקב הסיבות הנ"ל בחרנו להשתמש בשיטת ה- Dynamic Clustering בסכמה המוצעת על ידינו.

אלגוריתם ה- DC שמומש על ידינו מזדהה למעשה עם אלגוריתם LBG לבנית ספר קוד עבור קוונטיזציה וקטורי שתואר בסעיף 3.1. בניגוד לאלגוריתם לבנית ספר קוד בו תוצר האלגוריתם הוא סדרת הוקטורים המייצגים, בעת ביצוע ההקבצה תוצר האלגוריתם הינו החלוקה האופטימלית המתקבלת בצעד מס' (1) באיטרציה האחרונה באלגוריתם LBG . השיטה בה בחרנו את המילון ההתחלתי לאלגוריתם היתה שיטת הפיצול שתוארה גם היא בסעיף 3.1. שיטה זו בנוסף לעילותה, שהזכחה בצורה אמפירית, נתנה לנו גם חלוקה ל $c = 1, 2, 4, 8, \dots, 2^j$ קבוצות.

התוצאה של פעולת ההקבצה הינה תמונה המחולקת ל c קבוצות המכילות בלוקים בעלי תכולה ספקטרלית דומה. דוגמה לתוצאה של פעולת ההקבצה עבור $c=16$ ו- $c=64$ קבוצות מובאת בציור מס' 5.1.



ציור 5.1 - תוצאת הקבצה עבור 16 ו 64 קבוצות.

Fig. 5.1 - Clustering results for 16 and 64 clusters.

ציור 5.1a מתאר את התמונה המקורית עליה הפעלנו את האלגוריתם, ואילו ציורים 5.1b ו- 5.1c מתארים את תוצאת ההקבצה עבור $c=16$ ו- $c=64$ קבוצות בהתאמה. בתמונות המתארות את תוצאות ההקבצה כל רמת אפור מתאימה לקבוצה שונה שנמצאה. ניתן להבחין שהאזורים החלקים בתמונה שוייכו כולם לקבוצה אחת, ושהאזורים בעלי שיפוע מתון לכיוון מסויים שוייכו גם הם לקבוצה אחת. תוצאות אלה מעידות על הצלחת האלגוריתם במימוש החלוקה בצורה מתאימה לחלוקה שהיה מבצע מוח אנושי לפי אותם קריטריונים.

5.3 אדפטציה של הקוד

לאחר ביצוע ההקבצה התקבלו c קבוצות בעלות מאפיינים ספקטראליים שונים זה מזה. השונות הנ"ל במאפיינים הספקטראליים גוררת בעקבותיה את הצורך בבניית שבכה שונה לכל קבוצה, והקצאה של קצב שונה לכל קבוצה. הקצאה של קצב שווה לכל קבוצה לא תהיה אופטימלית, עקב העובדה שקבוצות המכילות בלוקים בעלי אנרגיה גבוהה דורשות קצב גבוה יותר מאשר קבוצות המכילות בלוקים בעלי אנרגיה מועטה. הן לצורך בניית דיאגרמת השבכה והן לצורך קביעת הקצאת הבלוקים נדרשת ידיעת הערכים העצמיים לכל קבוצה. אלגוריתם שיערוך הערכים העצמיים שתואר בפרק 4 משמש אותנו לשיערוך הערכים העצמיים גם בסכמה המוצעת בפרק זה. התחומים, עבורם משוערכים הערכים העצמיים, מורכבים מהבלוקים השייכים לכל אחת מהקבוצות שנקבעו בתהליך ההקבצה, כאשר מספר הבלוקים בכל תחום אינו קבוע כפי שמתואר באלגוריתם המתואר בפרק 4 אלא משתנה מאיזור לאיזור כפי שנקבע בתהליך ההקבצה.

כאמור, יש צורך לבצע הקצאת קצבים בין הקבוצות השונות. ערכי הערכים העצמיים המשוערכים (הספקטרום המשוערך) המשמשים להקצאת הקצבים הינם הסדרות $\{\bar{S}_l^k\}_{l=1}^N$, $k=1,2,\dots,c$ אשר התקבלו מאלגוריתם שיערוך הערכים העצמיים (פרק 4.2 צעד 4) (באלגוריתם). הבעיה של הקצאת הקצבים בין הקבוצות השונות נתנת לאפיון פורמלי בצורה הבאה:

נתונים c וקטורים \bar{S}^k , $k=1,2,\dots,c$ אשר מהווים כ"א את הספקטרום המשוערך המנורמל של הקבוצה ה- k ית. יש צורך לחלק R ביטים לכל הבלוקים בתמונה כך שה-MSE בין התמונה המקורית לתמונה המשוחזרת יהיה מינימלי.

האלגוריתם בו השתמשנו להקצאת הקצבים מבוסס על ההנחה שהתמונה במישור ההתמרה הינה גאוסית, והמוטיבציה במימוש נובעת מפונקציית הקצב - עוות של מקור גאוזי כפי שנתונה בנוסחה 4.2.

אלגוריתם הקצאת קצבים בין הקבוצות.

(1) עבור קצב נתון R קבע את הפרמטר θ כך ש

$$(5.4) \quad R = \frac{1}{P} \sum_{k=1}^c \sum_{l=1}^N \gamma_k \max \left\{ 0, \frac{1}{2} \log \left[\frac{\bar{S}_l^k}{\theta} \right] \right\}$$

כאשר:

- R הקצב הכולל הנדרש.
- P מספר הפיקסלים הכולל בתמונה.
- N מספר הפיקסלים בבלוק.
- \bar{S}_l^k הערך העצמי המשוערך המנורמל ה- l של הקבוצה ה- k ית.
- γ_k מספר הוקטורים בקבוצה ה- k ית.

(2) הקצב R_k של כל בלוק בקבוצה ה- k ית נתון ע"י

$$(5.5) \quad R_k = \frac{1}{N} \sum_{l=1}^N \max \left\{ 0, \frac{1}{2} \log \left[\frac{\bar{S}_l^k}{\theta} \right] \right\}$$

כאשר θ הינו הפרמטר שנקבע בצעד (1) לעיל.

הקצב R_k והספקטרום המשוערך \hat{S}^k שתהליך קביעתם תואר בסעיף זה, משמשים לבנית הדיאגרמה המתאימה לקבוצה ה- k ית ואכלוסה בהתאם למתואר בפרק 4 בהסתייגות הבאה. עקב העובדה שוורסיה זהה של הספקטרום המשוערך נמצאת הן במקלט והן במשדר, ניתן לסדר את הספקטרום בסדר עולה של אנרגיות המקדמים המדכיבים אותו, ולהשתמש בספקטרום המסודר הנ"ל לבנית השבכה. סדור מחדש זה יבטיח שסדרת הקצבים המתאימה למקדמים המקודדים תהיה מונוטונית עולה. במקרה זה אלגוריתם הקצאת הביטים שתואר בסעיף 4.3 יהיה יעיל יותר, עקב העובדה שברמה של השבכה יקודדו יחדיו מקדמים בעלי קצב קרוב. אם לא היינו מסדרים את הספקטרום בסדר עולה, אזי עלול להוצר מקרה בו מקדם הדורש קצב גבוה ומקדם הדורש קצב נמוך יקודדו באותה רמה, דהיינו שניהם יקודדו באותו קצב. במקרה זה יעילות הקידוד הכוללת נפגמת.

5.4 קידוד תמונת השארית באמצעות מקודד סקלרי

שמוש במנגנון ההקבצה לחלוקת התמונה לתחומים, וקידוד כל תחום באמצעות דיאגרמת שבכה המותאמת לו, לא פתר לחלוטין את בעיית הבלוקיות. עדין הופיעו בלוקים בתמונה אשר בשחזור שלהם הופיעה תופעת הבלוקיות. הבלוקים הללו היו בעיקר אותם בלוקים אשר ממוקמים במעברים בין אזורים שונים בתמונה. דהיינו הפלוג או ההיסטוגרמה שלהם עבור אותו אזור מכילה יותר מ "mode" אחד. בלוקים אלו בברור אינם מקיימים את המודל הגאוסי אשר פונקציית הפלוג שלו הינה בעלת "mode" בודד. כדי להתגבר על התופעה הנ"ל נעשה שמוש בקוונטיזר סקלרי, אשר יקודד את תמונת השארית במישור ההתמרה (תמונת השגיאה), הנוצרת ע"י הפחתת התמונה המשוחזרת במישור ההתמרה מהתמונה המקורית באותו מישור. ניתן להסתכל על הסכמה החדשה שייצרנו כעל מקודד רב-דרגתי (Multi Stage) המורכב מקוונטיזר וקטורי הממומש באמצעות דיאגרמת שבכה, ובהמשכו קוונטיזר סקלרי של אות השארית. תאור סכמתי של המקודד הרב-דרגתי מובא בציור הבא:

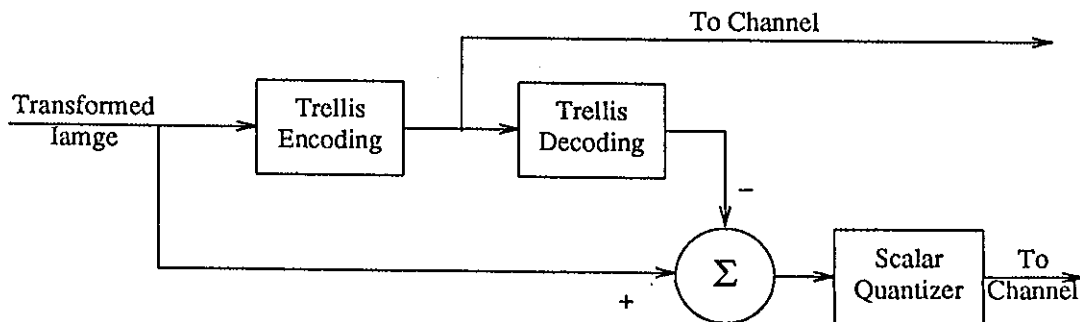


Fig 5.2 - Multi-stage coder.

כל בלוק שקודד ע"י דיאגרמת השבכה מופענח ומופחת מהבלוק המתאים לו במישור ההתמרה של התמונה המקורית. מקדמי בלוק השגיאה שנוצר מקוונטטים באמצעות קוונטיזר סקלרי, תרכיחם ומקומם בבלוק נשמרים. עם סיום הפעולה הנ"ל על פני כל הבלוקים בתמונה, הערכים המקוונטטים ומקומם שנאגרו מכל הבלוקים בתמונה, מקודדים באמצעות קודי הפמן (Huffman codes), ונשלחים כאינפורמציה צד למקלט. קודי הפמן שנוצרו נשלחים אף הם למקלט כאינפורמציה צד.

המקודד הסקלרי שהוספנו מתקן את המקדמים הגדולים בתמונת השגיאה ומגביל אותם לתחום השווה ל $\pm \frac{1}{2}$ גודל צעד הקוונטיזציה שלו. ניתן להסתכל על פעולת הקוונטיזר הסקלרי בעל הלבנת תמונת השגיאה. פעולה זו גרמה להקטנת תופעת הבלוקיות לרמה זניחה. גודל צעד הקוונטיזציה נקבע בהתחשב בכמות הבלוקים בתמונה אשר סוטים מהמודל על פיו תוכננה דיאגרמת השבכה, ככל שכמות הבלוקים הסוטים מהמודל רבה כן נקטין את צעד הקוונטיזציה דהיינו נקצה יותר קצב לקוונטיזר הסקלרי. הטיבוכיות שנוספה לסכמה עקב הוספת מנגנון הקוונטיזר הסקלרי הינה מועטה יחסית. תהליך הפענוח הנדרש הינו תוצר לוואי של תהליך הקידוד בשבכה, ואילו פעולת החיסור והקוונטיזציה הסקלרית הינן פעולות פשוטות ביותר. הטיבוכיות בבנית קודי הפמן עבור ערכי המקדמים ומקומם וההשהיה הנגרמת עקב כך הינם קטנים יחסית לטיבוכיות וההשהיה הכוללים של הסכמה. עקב העובדה שהמוטיבציה העיקרית שהנחתה אותנו ביישום הסכמה היתה השגת הביצועים הטובים ביותר הנתנים להשגה, לא השקענו מאמצים רבים בהקטנת הטיבוכיות.

5.5 תאור הסכמה הכוללת

תרשים עקרוני של הסכמה הכוללת הממומשת על ידינו מופיע בציור 5.3.

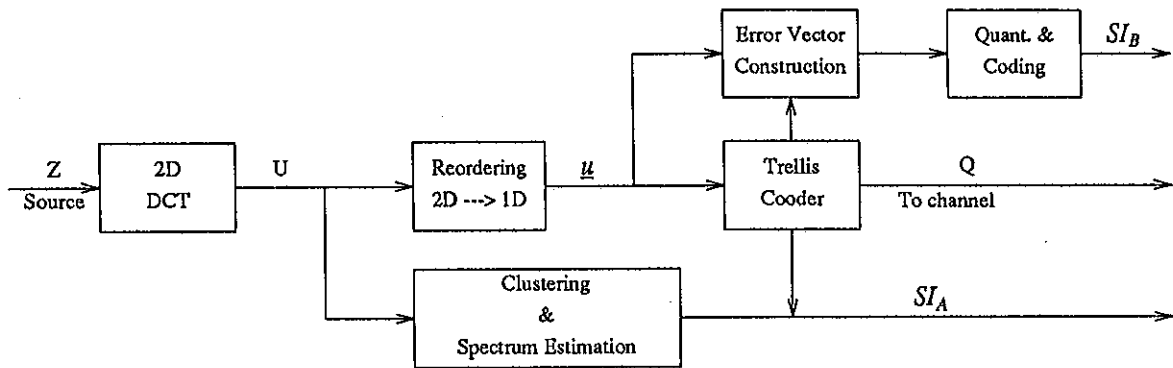


Fig 5.3a - TRANSMITER

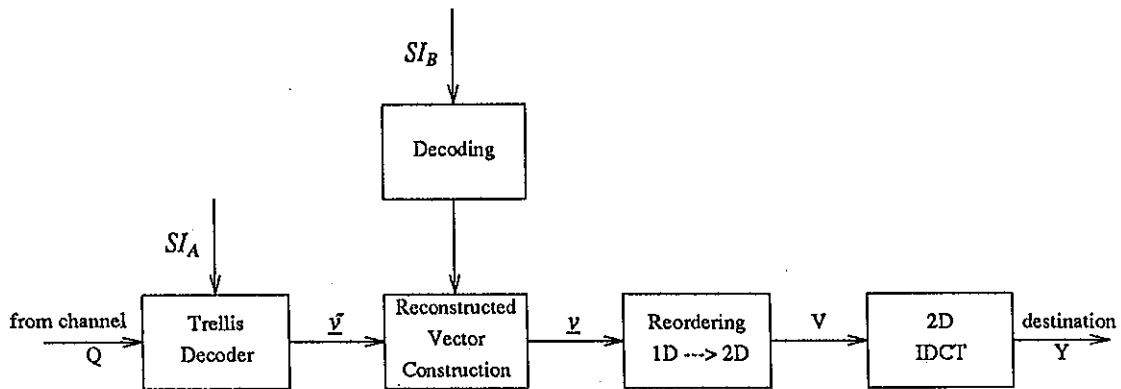


Fig 5.3b - RECEIVER

ציור 5.3 - תאור הסכמה הכוללת (a) משדר (b) מקלט.

Fig. 5.3 - The Final coding scheme (a) transmitter (b) receiver.

הסכמה מבצעת 2 מעברים על תמונת הכניסה. במעבר הראשון התמונה בגודל $[P' \times P']$ (סך כולל של P פיקסלים) מחולקת ל s בלוקים בגודל $[N' \times N']$ (סך כולל של N פיקסלים לבלוק). כל בלוק Z עובר התמרת DCT דו-מימדית לקבלת בלוק מקדמים במישור ההתמרה U . s הבלוקים מסווגים ל c קבוצות באמצעות מנגנון ההקבצה שתואר בסעיף 5.2. לכל אחד מ c הקבוצות משוערן ספקטרום מנורמל $\bar{\sigma}^k$ אשר ורסיה שלו המסודרת בסדר אנרגיות עולה משמשת לקביעת מבנה השבכה והקצאת הקצבים בין הקבוצות כמתואר בסעיף 5.3. הספקטרום המשוערן $\bar{\sigma}^k$ משמש לאיכלוס השבכה ולחלוקת הקצבים בתוך הוקטור. הספקטרום $\bar{\sigma}^k$ ומקדמי ההגבר G_i של כל בלוק נשלחים כאינפורמציה צד המסומנת ב

SI_A

במעבר השני כל בלוק U מסודר כוקטור \underline{u} לפי אותו סדר בו סודר הספקטרום המשוער של הקבוצה אליה הוא שייך. הוקטור \underline{u} מקודד ע"י השבכה ומפת הקוד המתאימה לו Q נשלחת לערוץ. במקביל כל בלוק שקודד מפתנח ומופחת מוקטור המקור \underline{u} ליצירת וקטור שגיאה. המקדמים בעלי העוצמה הגבוהה בוקטור השגיאה מקוונטים ע"י הקוונטיזציה הסקלרית, מקודדים ונשלחים למקלט כאינפורמציה צד SI_B , כפי שמתואר בסעיף 5.4.

המקלט (ציור 5.3b) משתמש באינפורמציה הצד SI_A לבניית דיאגרמת שבכה זהה לזו הנמצאת במשדר. מפת הקוד הנקלטת Q משמשת לקביעת הוקטור \underline{y} . וקטור השגיאה משוחזר אף הוא מתוך אינפורמציה הצד SI_B , ומסוכם עם הוקטור \underline{y} ליצירת וקטור השחזור \underline{z} . הוקטור \underline{z} מאורגן חזרה כבלוק דו-ממדי V אשר מותמר באמצעות התמרת IDCT לקבלת הבלוק המשוחזר X . תהליך בניית השבכה מתבצע עבור כל קבוצה שנקבעה ע"י מנגנון ההקבצה. תהליך הקודד והפענוח שתואר לעיל מתבצע עבור כל הבלוקים בתמונה.

פרק 6 - ביצועי המערכת

6.1 ביצועי הסכמה המוצעת והשוואה לסכמות קידוד אמרות

סכמת הקידוד שהוצגה בעבודה זו פותחה בשלושה שלבים עקריים. בשלב ראשון כמתואר בפרק 4, נבנתה סימולציה לסכמה הבסיסית כפי שהוצגה ע"י Mazor & Pearlman ב [1], ויושמה על ידס לקידוד אותות דיבור ב [27]. בשלב שני הוסף לסכמה מנגנון הקבצה (ראה סעיף 5.2) אשר בצע בצורה אוטומטית חלוקה של התמונה לקבוצות. לכל קבוצה כזו הוקצה קצב מתאים והיא קודדה ע"י דיאגרמת שבכה נפרדת. בשלב שלישי הוסף לסכמה מנגנון קוונטיזר סקלרי (ראה סעיף 5.4) אשר הפחית כמעט לחלוטין את תופעת הבלוקיות בתחום הקצבים שנבדק. תאור ביצועי המערכת יבוצע בהתאם לשלבים הנ"ל.

ביצועי המערכות שתוארו בפרקים הקודמים נבדקו בסימולציות על גבי מחשב VAX 750 והוצגו באמצעות מערכת Gould De-Anza IP8500. תמונת המבחן אשר שמשה אותנו לצורך הערכת ביצועי הסכמה היתה תמונת המבחן LENA בשתי רזולוציות [256×256] ו [512×512] תמונה זו משמשת להערכת הביצועים בסכמות קידוד רבות, ולפיכך ניתן להשוות בעזרתה בין הביצועים של סכמת הקידוד שלנו לביצועים של סכמות אחרות.

ביצועי המערכת נבדקו באמצעות שני מדדים:

א. איכות הסוביקטיבית של התמונה המשוחזרת.

ב. יחס האות לדעש המוגדר ע"י:

$$SNR = 10 \log \frac{255^2}{MSE} \quad [dB]$$

כאשר MSE הינו ה- Mean Square Error בין התמונה המקורית לתמונה המשוחזרת.

כל תמונות התוצאה המובאות בעבודה זו הינן בעלות מבנה זהה הנתון בציר 6.1

Reconst. Image	Error Image
Original Image	Cluster Image

ציר 6.1 - מבנה תמונות התוצאה.

Fig. 6.1 - Resultant Images Structure.

תמונות ההקבצה מתארות את שייכות הבלוקים בתמונה לקבוצות השונות בתמונה, כאשר כל רמת אפור בתמונה מגדירה קבוצה. תמונות השגיאה המופיעות בפרק זה נבנו כערך מוחלט של ההפרש בין תמונת המקור לתמונת השיחזור מוכפל ב - 10.

סכמת הקידוד הבסיסית שמומשה על ידינו, והמתוארת בפרק 4, לא נתנה תוצאות משביעות רצון. ציורים 6.2 עד 6.5 מציגים את התוצאות שהתקבלו עבור 4 סטים של פרמטרים (לצורכי השוואה כל התמונות מרוכזות בסוף הפרק). טבלה 6.1 מסכמת את נתוני המערכות והתוצאות האובייקטיביות (קצב R ויחס האות לרעש - SNR) שהושגו.

SNR [dB]	R [bits/pel]	גודל בלוק	רזולוציה	ציור	סדר
30.4	0.78	8×8	256×256	6.2	1
32.4	1.01	8×8	256×256	6.3	2
31.7	0.91	16×16	256×256	6.4	3
35.6	0.55	16×16	512×512*	6.5	4

* ברזולוציה הגבוהה קודד, מסיבות מעשיות, רק הרבע השמאלי התחזן של התמונה. טבלה 6.1 - ביצועי הסכמה הבסיסית.

Table 6.1 - Basic Scheme Performance.

כפי שרואים מהתמונות, ביצועי הסכמה הבסיסית אינם טובים מספיק. כל התמונות המשוויות טובלות מפגמים בעיקר בשפות בין אזורים בעלי סטטיסטיקה שונה בתמונה. (למשל השפה בין הכתף לרקע וכו') תמונות השגיאה באזורים אלו הינן רועשות מאוד. ניתן לראות שביצועי הסכמה עבור התמונה בעלת הרזולוציה הגבוהה (Fig 6.5) הינם טובים בהרבה מהביצועים המושגים עבור התמונות בעלות הרזולוציה הנמוכה. הסיבה לכך הינה הקורלציה הגבוהה הקיימת ברזולוציה גבוהה, המאפשרת דחיסה טובה יותר.

השלב הבא בפיתוח הסכמה היה הוספה של מנגנון הקבצה וביצוע אדפטציה של הסכמה ע"י חלוקת קצבים שונים לקבוצות בהתאם לספקטרוס המייצג של כל קבוצה. ציורים 6.6 עד 6.9 מציגים את התוצאות שהתקבלו. טבלה 6.2 מסכמת את נתוני המערכות ואת התוצאות האובייקטיביות שהתקבלו. מסיבות מעשיות לא היה נוח לקבוע קצב זהה בהשוואה בין סכמות הקידוד השונות. בכל מקרה השתדלנו לקבוע למערכת בעלת הביצועים הפוטנציאליים הטובים יותר קצב נמוך יותר, כדי למנוע אי הבנות בפרוש התוצאות.

השוואה של ביצועי הסכמה הבסיסית לסכמה הכוללת הקבצה ואדפטציה מבליטה את העדיפות המכרעת של הסכמה השנייה. השיפור באיכות תמונת השיחזור ברור לעין, והשיפור ב SNR נע בין 2dB ל 3dB. נסיון לשפור את ביצועי הסכמה הבסיסית בעזרת אדפטציה של הקצב בלבד לא עלה יפה. עבור תמונה ברזולוציה של 256×256 וגודל בלוק התמורה של 16×16 התקבל בקצב של 0.92 bit/pel SNR של 30.35 dB, תוצאה הגרועה מזו שהושגה עבור אותה תמונה באותם תנאים בלא אדפטציה (סדר 3 טבלה 6.1). הסיבה לתוצאות הנחותות שהתקבלו, כאשר בוצעה חלוקת קצבים אדפטיבית, נעוצה בעובדה שהקבוצות ששימשו לחלוקת הקצבים לא היו סטציונריות, ולפיכך לא היתה כלל משמעות לחלוקת הקצבים.

SNR [dB]	R [bits/pel]	גודל בלוק	רזולוציה	ציור	סודר
33.16	0.75	8×8	256×256	6.6	1
34.76	0.96	8×8	256×256	6.7	2
32.92	0.88	16×16	256×256	6.8	3
37.78	0.49	16×16	512×512	6.9	4

טבלה 6.2 - תוצאות הסכמה הכוללת הקבצה ואדפטציה.

Table 6.2 - The Performance of the scheme with clustering and adaptation.

השיפור שהתקבל עקב הוספת מנגנון ההקבצה והאדפטציה הינו משמעותי, אולם עדיין ניתן להבחין בתופעות של "בלוקיות" במיוחד בבלוקים הכוללים מעברים בין אזורי התמונה השונים. תוספת של מנגנון קוונטיזציה סקלרי, לקידוד המקדמים הגבוהים של תמונת הפרש במישור ההתמרה, הפחיתה בד"כ את התופעה הנ"ל. ציורים 6.10 עד 6.13 מתארים את התוצאות שהתקבלו בסכמה הסופית הכוללת גם קוונטיזציה סקלרי. טבלה 6.3 מסכמת את נתוני המערכות ואת התוצאות שהתקבלו.

SNR [dB]	R [bits/pel]	גודל בלוק	רזולוציה	ציור	סודר
33.64	0.75	8×8	256×256	6.10	1
35.2	0.92	8×8	256×256	6.11	2
33.28	0.88	16×16	256×256	6.12	3
37.44	0.41	16×16	512×512	6.13	4

טבלה 6.3 - ביצועי הסכמה הסופית.

Table 6.3 - Final Scheme Performance.

בהשוואה לסכמה הכוללת מנגנון הקבצה ואדפטציה בלבד התקבל שיפור של כ- 0.5 dB (בסודר 4 השיפור מתבטא בקצב נמוך יותר בסכמה הכוללת קוונטיזציה סקלרי). ניתן להבחין בשיפור באיכות התמונות הסובייקטיביות. כמו כן ניתן לראות שתמונות השגיאה הינן בעלות אופי לבן והרעש מפוזר בצורה יחסית אחידה ע"פ התמונה ולא מרוכז באזורים מסויימים.

הרצה של אותה הסכמה על תמונה נוספת (Tiffany) ברזולוציה של 256×256 נתנה יחסי אות לרעש של 32.1 dB ו 33.26 dB וקצבים של כ 1 bit/pel עבור גדלי בלוקים של 8 ו 16 בהתאמה. לא מצאנו בספרות תוצאות להשוואה עבור תמונת המבחן הנ"ל.

הרצה של תמונת המבחן Lena ברזולוציה של 512×512 בקצב של 0.31 bit/pel נתנה SNR של 35.9 dB ואיכות שחזור טובה, אך נפלת בהרבה מזו המושגת בקצב של 0.41 bit/pel (סודר 4 טבלה

הסיבוכיות הנוספת לסכמה הבסיסית עקב הוספת מנגנוני ההקבצה, האדפטציה והקוונטיזציה הסקלרי, מתבטאת בקצב נוסף הנדרש לשליחת אינפורמציה צד למקלט לצורך פיענוח נכון. הקצב הנדרש לצורך מימוש מנגנון ההקבצה הינו שליחה של סמן המציין את ה Cluster אליו שייך כל בלוק. הקצב הנ"ל קטן מ $\log_2 c$ (כאשר c הינו מספר הקבוצות) עקב העובדה שהשתמשנו בקוד הפמן לקידוד סמני הקבוצות. הקצב הנ"ל בתוספת הקצב הנדרש לשליחת ערכי הספקטרום של כל Cluster והחגבר של כל בלוק (ראה פרק 4.3) מהווים את SI_A המתואר בסעיף 5.6.

הקצב המתאים לאינפורמציה הצד של המקודד הסקלרי הינו הקצב הנדרש לשליחת מיקום המקדמים הגדולים וערכם לאחר קוונטיזציה, והוא מהווה את SI_B בתאור הסכמה בפרק 5.6. ערכים אופייניים של אינפורמציה הצד שהתקבלה בסימולציה מובאים בטבלה 6.4.

טורד	ציור	R [bits/pel]	SI_A [bits/pel]	SI_B [bits/pel]
1	6.11	0.92	0.185	0.117
3	6.12	0.88	0.172	0.062
4	6.13	0.41	0.062	0.091

טבלה 6.4 - ערכי אינפורמציה צד אופייניים.

Table 6.4 - Typical Side-Information values.

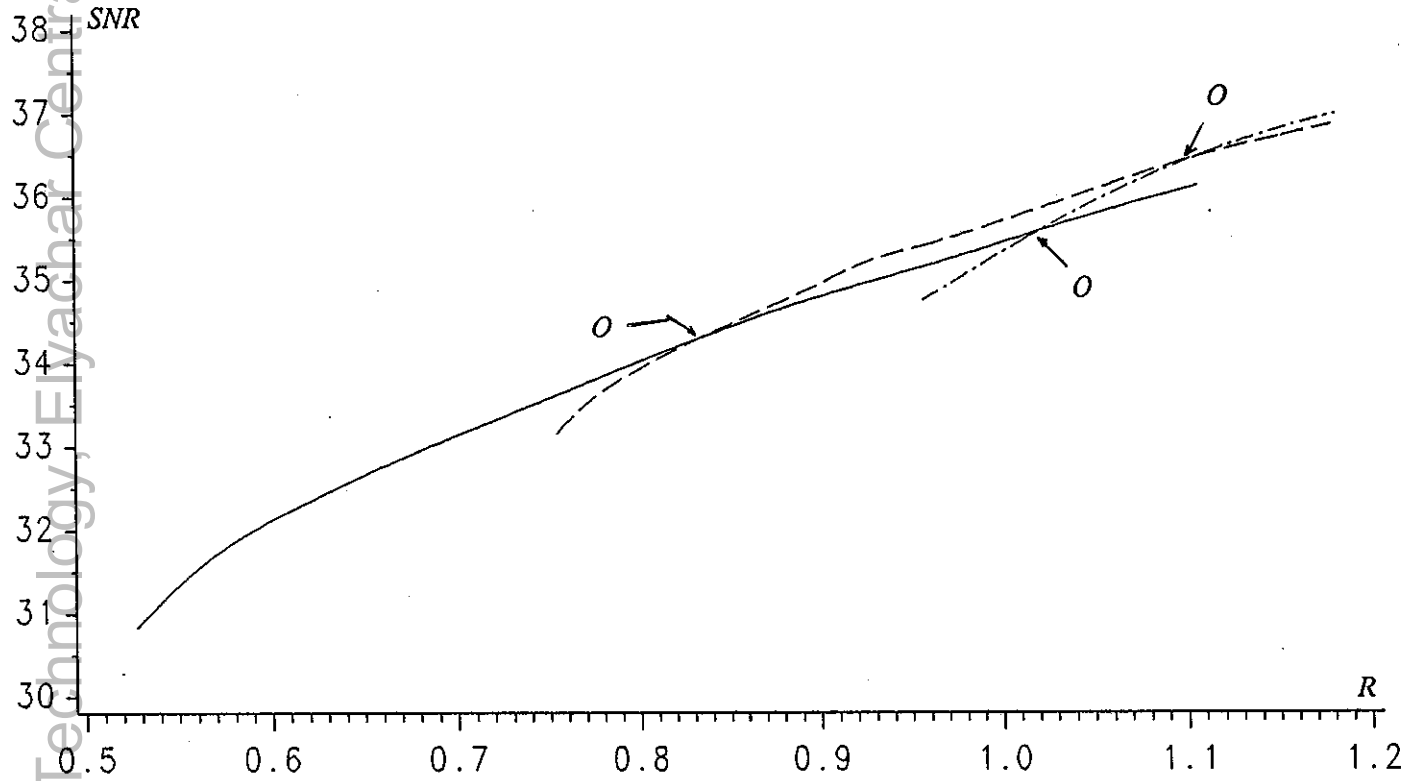
כאמור הסכמה הסופית מהווה מקודד דו-דרגתי המורכב מקוונטיזציה וקטורי, ואחריו קוונטיזציה סקלרי לקידוד תמונת השגיאה במישור ההתמרה. אם נסמן את הקצב המתאים לקוונטיזציה הווקטורית ב R_v ואת הקצב המתאים לקוונטיזציה הסקלרית ב R_s , הקצב הכולל R_t ניתן ע"י

$$R_t = R_v + R_s$$

ציור 6.15 מתאר את יחס האות לרעש (SNR) המושג בתלות בקצב הכולל R_t . כל גרף מתאים לערך R_v קבוע וערך R_s גדל בהדרגה. הגרפים מתאימים לגודל בלוק התמרה 8×8 ול ערכי R_v .

ניתן להבחין שהגרף בעל R_v גדול יותר (קצב גדול יותר לקוונטיזציה הווקטורית) מתחיל ב SNR נמוך יותר מה-SNR המושג עבור אות קצב כולל R_t ע"י המערכת בעלת R_v קטן יותר אשר משלבת קוונטיזציה סקלרית. דהיינו מערכת הכוללת מקודד וקטורי בלבד הינה בעלת ביצועים אובייקטיביים נחותים מהמערכת המשלבת קוונטיזציה וקטורי וקוונטיזציה סקלרית הפועלת באותו קצב כולל. הוספת מקודד סקלרי לסכמה הווקטורית בלבד, ועידונו בהדרגה, מקטינה את הפער בין שתי המערכות עד להשוואתו (הנקודות 0 בגרף) המשך עידון הקוונטיזציה הסקלרית מביא למצב שבו המערכת בעלת R_v גבוה יותר הינה בעלת ביצועים טובים יותר.

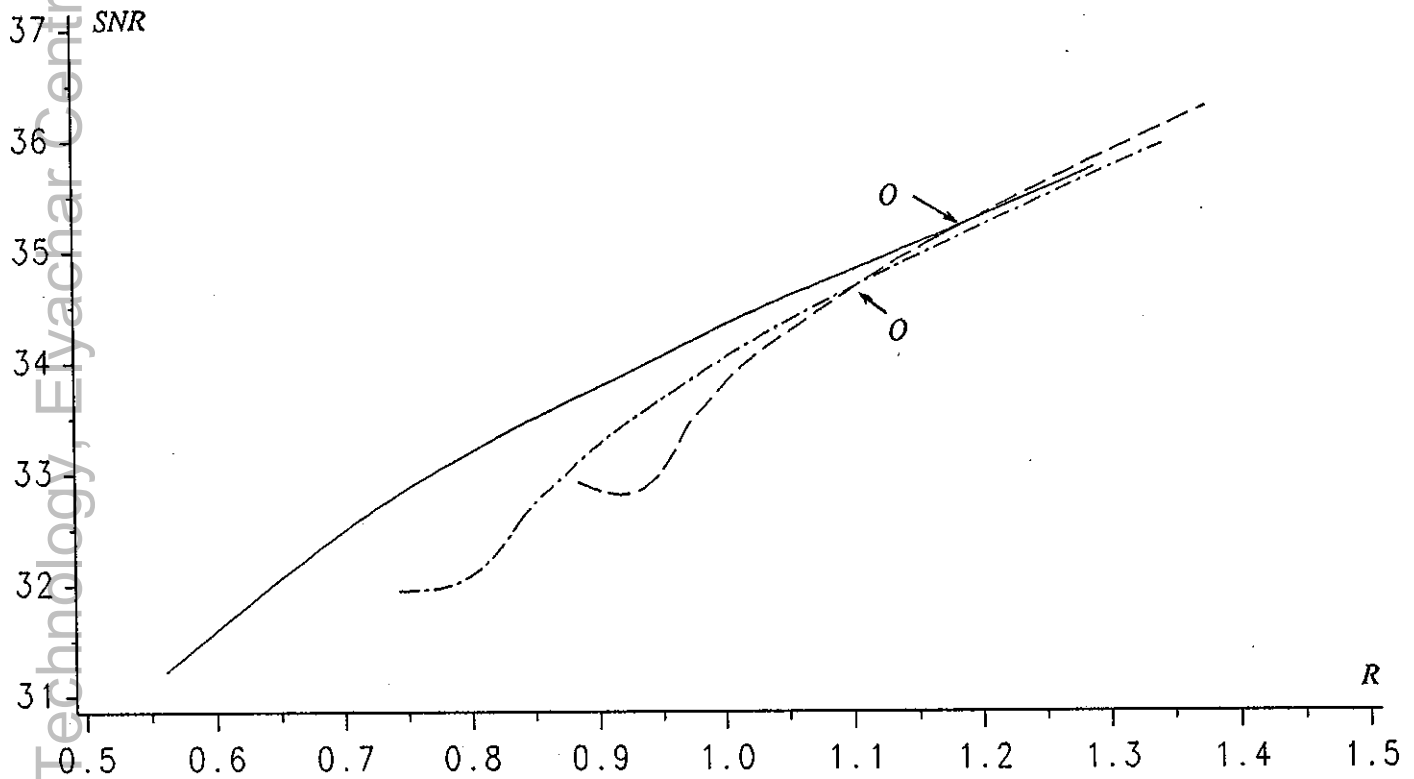
הסיבה לתופעה הנ"ל נעוצה בעובדה שקיימים בלוקים בתמונה אשר אינם מתאימים למודל לפיו נבנה הקוונטיזציה הווקטורית. ולפיכך, הקצב שהושקע בקידוד אותם בלוקים בעזרת הקוונטיזציה הווקטורית, לא היה מסוגל לתרום לשפור משמעותי בביצועים. תוספת הקוונטיזציה הסקלרית, אשר הופעל רק על המקדמים



ציור 6.15 - השוואת ביצועים עבור ערכי R_v שונים (גודל בלוק 8×8).

Fig 6.15 - Performance comparison for various values of R_v (8×8 block size).

בעלי העוצמה הגדולה בתמונת השגיאה, גרמה לתיקון אותם מקדמים בתוספת קצב קטנה יחסית. תוספת הקוונטיזר הסקלרי ועידונו בהדרגה שיפרה את הביצועים עד למצב בו לא היו מקדמי שגיאה חריגים (רעש לבן). במצב זה השקעה של קצב נוסף בקוונטיזר הסקלרי תהיה פחות יעילה מהשקעה של אותו הקצב בקוונטיזר הוקטורי. מסיבה זו מקבלים את נקודות החיתוך בגרף. תופעה דומה, אך פחות מודגשת, ניתן לראות בציור 6.16 עבור תמונה בגודל בלוק 16×16 .



ציור 6.16 - השוואת ביצועים עבור ערכי R , שונים (גודל בלוק 16×16).

Fig 6.16 - Performance comparison for various values of R , (16×16 block size).

טבלה 6.5 משווה את ביצועי הסכמה המוצעת על ידינו על התמונה ברזולוציה הנמוכה $[256 \times 256]$

עם התוצאות עבור אותה תמונה שדווחו לאחרונה בספרות $[32,33,34]$ כמו שמתבטא בטבלה, הביצועים המועטים על ידינו עדיפים על הביצועים המדווחים ב $[32,33,34]$.

טבלה 6.6 משווה את ביצועי הסכמה המוצעת על ידינו על התמונה ברזולוציה הגבוהה (512×512)

עם התוצאות שדווחו לאחרונה $[35,36]$. גם במקרה זה מתבטאת בבירור עדיפות הסכמה המוצעת על ידינו.

מקור ספרות	SNR [dB]	R bits/pel	סודר
הסכמה המוצעת	35.2	0.92	1
	33.6	0.76	2
	32.6	0.61	3
[31]	30.9	0.67	4
	32.5	1.	5
[33]	32.4	0.74	6
	32.8	1.	7
[32]	27.5	0.5	8
	30.9	1.	9

טבלה 6.5 - השוואת ביצועים עם סכמות קידוד אחרות (רזולוציה נמוכה 256×256).

Table 6.4 - Performance comparison with other coding schemes (low resolution image 256×256).

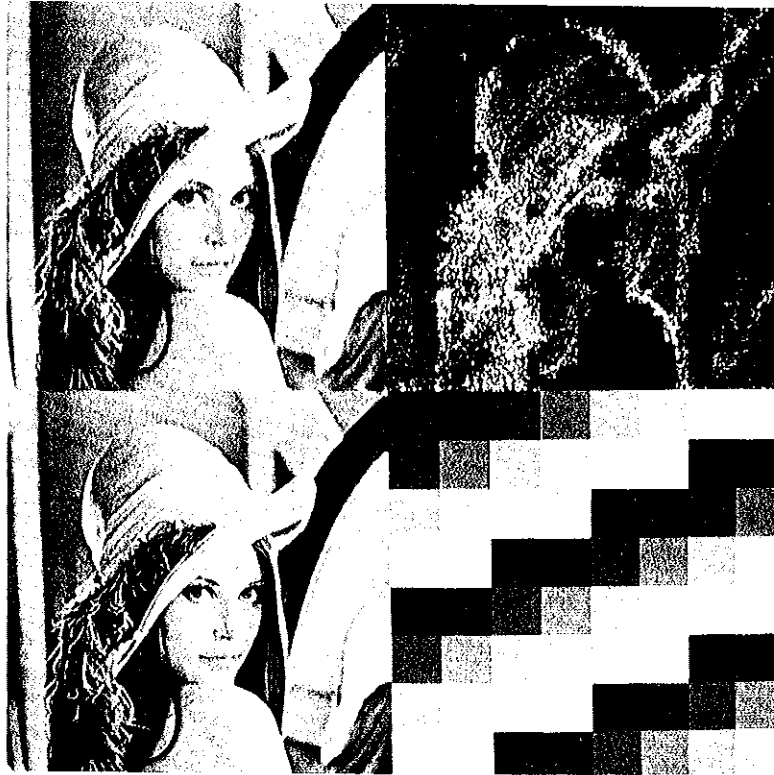
מקור ספרות	SNR [dB]	R bits/pel	סודר
הסכמה המוצעת	37.44	0.41	1
[34]	32.19	0.557	2
[35]	31.3	0.65	3

טבלה 6.6 - השוואת ביצועים עם סכמות קידוד אחרות (רזולוציה גבוהה).

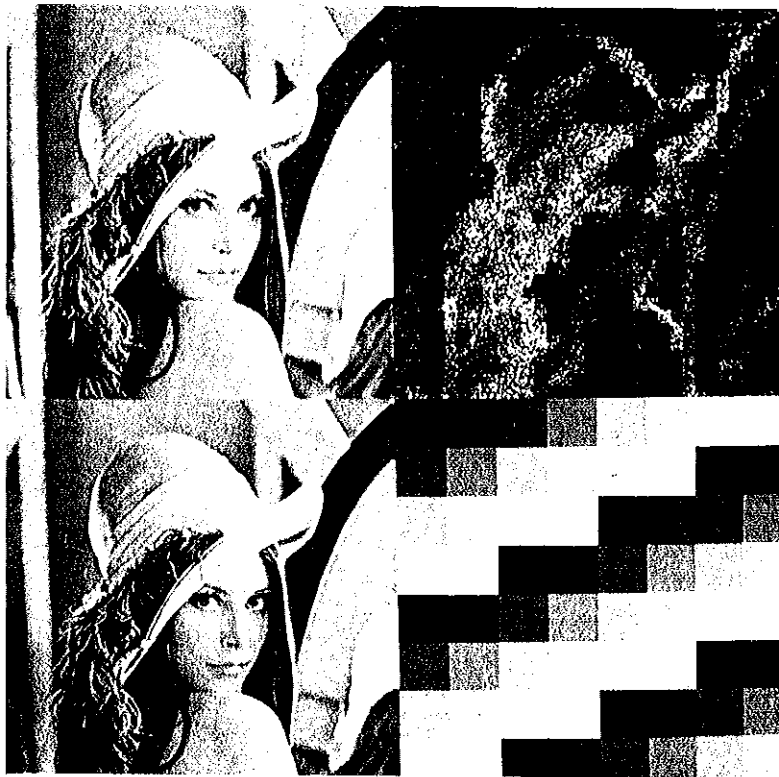
Table 6.4 - Performance Comparison with other coding schemes (high resolution image).

ציור 6.2 - תוצאות עבור רזולוציה [256×256] : $R = 0.78 \text{ bit/pel}$, $SNR = 30.4 \text{ dB}$

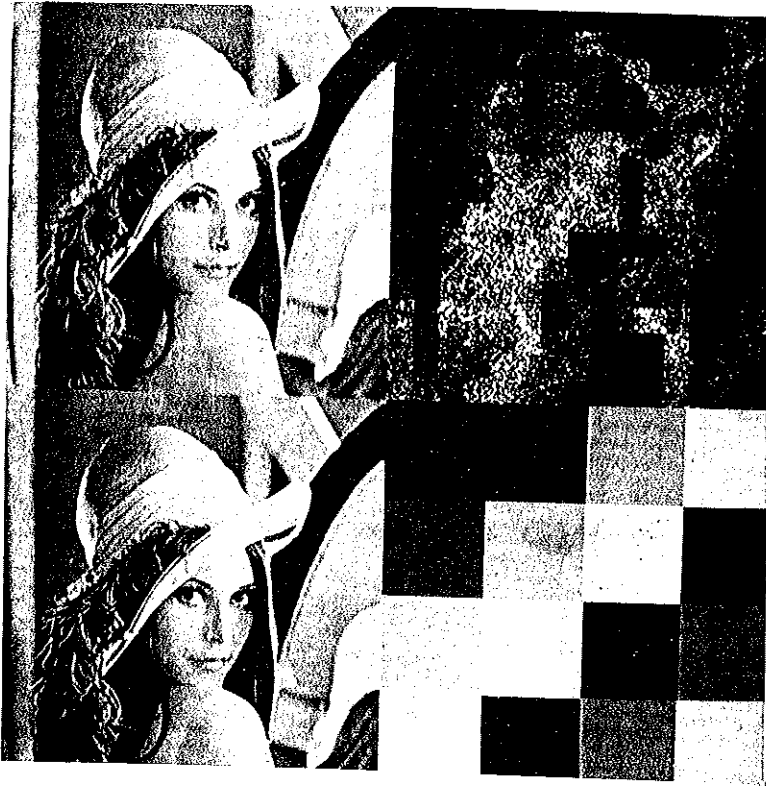
Fig. 6.2 - Results for [256×256] resolution image : $R = 0.78 \text{ bit/pel}$ $SNR = 30.4 \text{ dB}$



צירוף 6.2 - תוצאות עבור רזולוציה $[256 \times 256]$: $R = 0.78 \text{ bit/pel}$, $SNR = 30.4 \text{ dB}$
 Fig. 6.2 - Results for $[256 \times 256]$ resolution image : $R = 0.78 \text{ bit/pel}$ $SNR = 30.4 \text{ dB}$



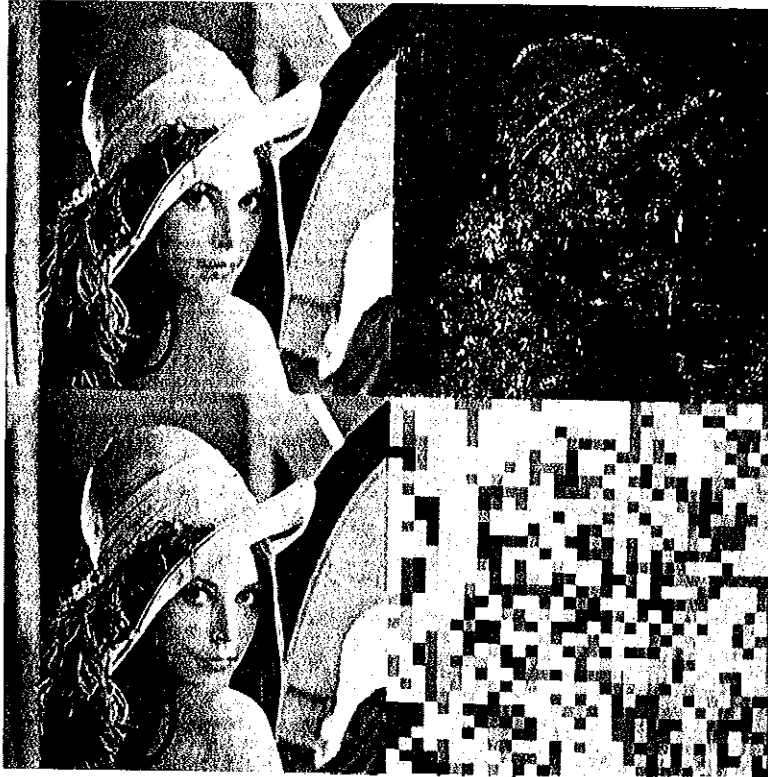
צירוף 6.3 - תוצאות עבור רזולוציה $[256 \times 256]$: $R = 1.01 \text{ bit/pel}$, $SNR = 32.4 \text{ dB}$
 Fig. 6.3 - Results for $[256 \times 256]$ resolution image : $R = 1.01 \text{ bit/pel}$ $SNR = 32.4 \text{ dB}$



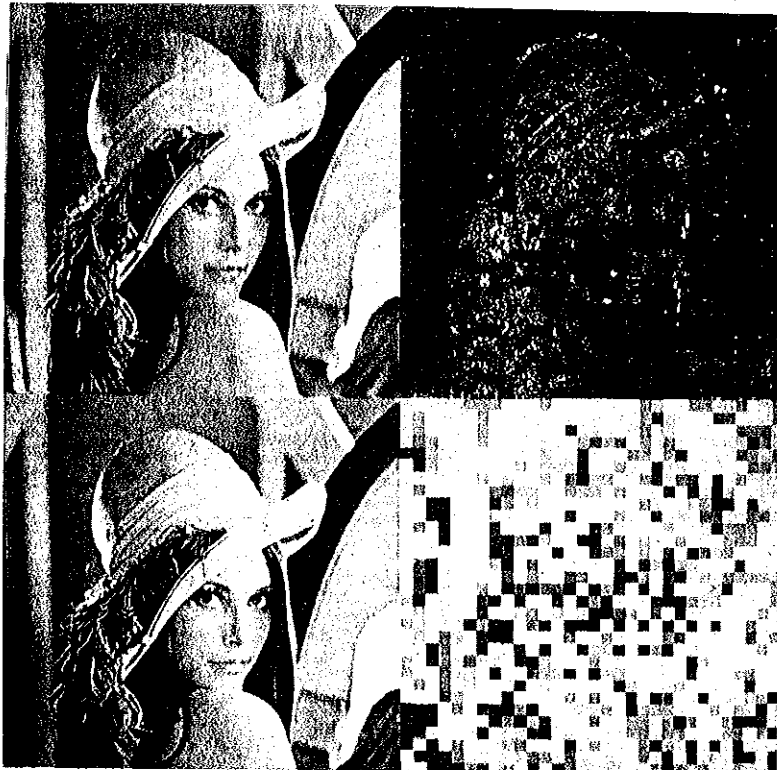
ציור 6.4 - תוצאות עבור רזולוציה [256x256] : $R = 0.91 \text{ bit/pel}$ $SNR = 31.7 \text{ dB}$
Fig. 6.4 - Results for [256x256] resolution image : $R = 0.91 \text{ bit/pel}$ $SNR = 31.7 \text{ dB}$



ציור 6.5 - תוצאות עבור רזולוציה [512x512] : $R = 0.55 \text{ bit/pel}$ $SNR = 35.6 \text{ dB}$
Fig. 6.5 - Results for [512x512] resolution image : $R = 0.55 \text{ bit/pel}$ $SNR = 35.6 \text{ dB}$



ציור 6.6 - תוצאות עבור רזולוציה [256x256] : $R = 0.75 \text{ bit/pel}$ $SNR = 33.16 \text{ dB}$
Fig. 6.6 - Results for [256x256] resolution image : $R = 0.75 \text{ bit/pel}$ $SNR = 30.16 \text{ dB}$

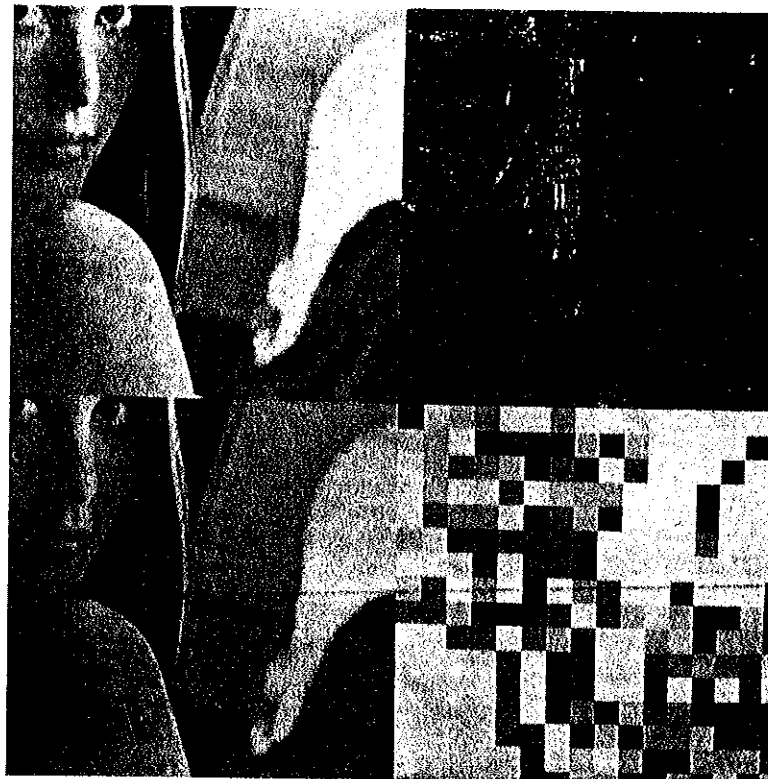


ציור 6.7 - תוצאות עבור רזולוציה [256x256] : $R = 0.96 \text{ bit/pel}$ $SNR = 34.76 \text{ dB}$
Fig. 6.7 - Results for [256x256] resolution image : $R = 0.96 \text{ bit/pel}$ $SNR = 34.76 \text{ dB}$



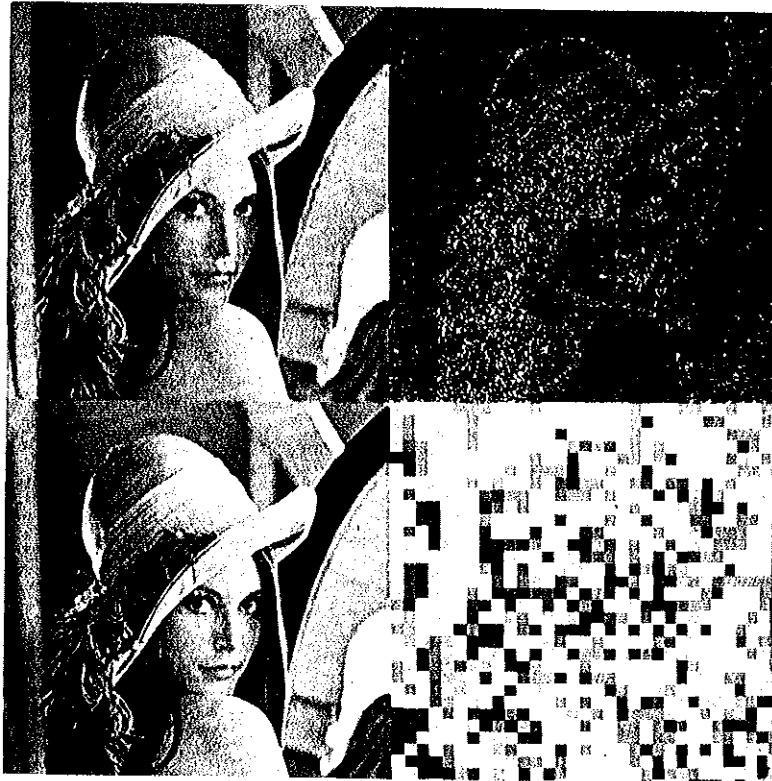
ציור 6.8 - תוצאות עבור רזולוציה [256×256] : $R = 0.88 \text{ bit/pel}$, $SNR = 32.92 \text{ dB}$

Fig. 6.8 - Results for [256×256] resolution image : $R = 0.88 \text{ bit/pel}$, $SNR = 32.92 \text{ dB}$



ציור 6.9 - תוצאות עבור רזולוציה [512×512] : $R = 0.49 \text{ bit/pel}$, $SNR = 37.78 \text{ dB}$

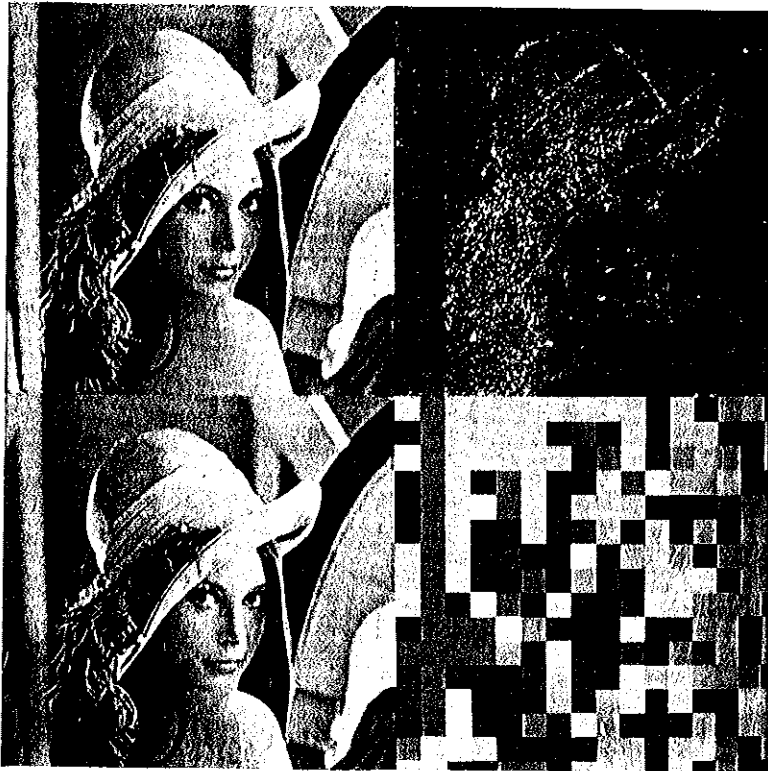
Fig. 6.9 - Results for [512×512] resolution image : $R = 0.49 \text{ bit/pel}$, $SNR = 37.78 \text{ dB}$



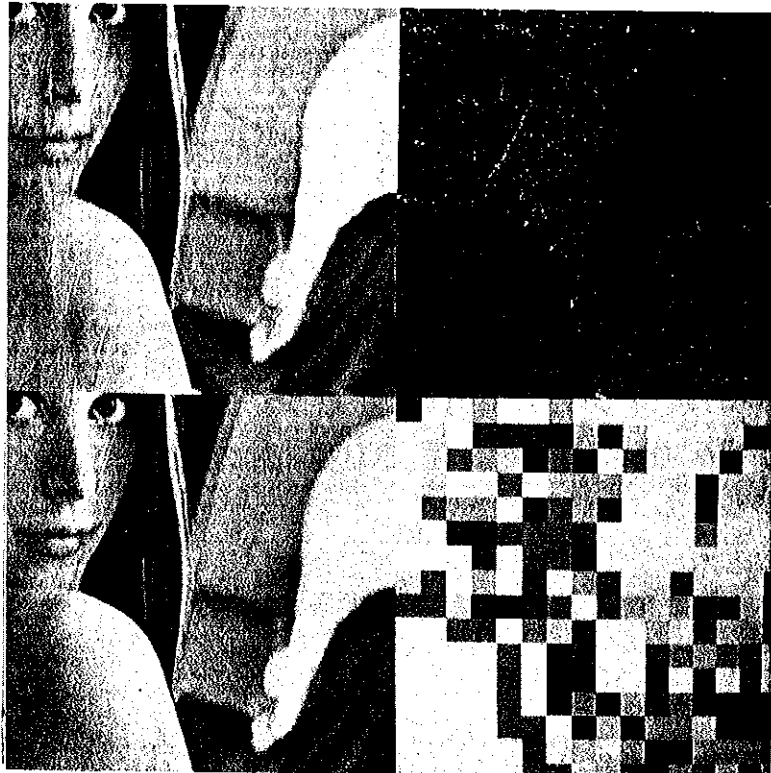
ציור 6.10 - תוצאות עבור רזולוציה [256x256] : $R=0.75 \text{ bit/pel}$, $SNR = 33.64 \text{ dB}$
Fig. 6.10 - Results for [256x256] resolution image : $R=0.75 \text{ bit/pel}$ $SNR = 30.64 \text{ dB}$



ציור 6.11 - תוצאות עבור רזולוציה [256x256] : $R=0.92 \text{ bit/pel}$, $SNR = 35.2 \text{ dB}$
Fig. 6.11 - Results for [256x256] resolution image : $R=0.92 \text{ bit/pel}$ $SNR = 35.2 \text{ dB}$



ציור 6.12 - תוצאות עבור רזולוציה [256x256] : $SNR = 33.28 \text{ dB}$, $R = 0.88 \text{ bit/pel}$
 Fig. 6.12 - Results for [256x256] resolution image : $R = 0.88 \text{ bit/pel}$ $SNR = 33.28 \text{ dB}$



ציור 6.13 - תוצאות עבור רזולוציה [512x512] : $SNR = 37.41 \text{ dB}$, $R = 0.41 \text{ bit/pel}$
 Fig. 6.13 - Results for [512x512] resolution image : $R = 0.41 \text{ bit/pel}$ $SNR = 37.41 \text{ dB}$

פרק 7 - סכום והמלצות

7.1 סכום

סכמת הקידוד שהוצגה בעבודה זו מקודדת את מקדמי התמרת DCT של תמונות באמצעות קוונטיזציה וקטורי הממומש ע"י דיאגרמת שבכה. הבלוקים בתמונה, לאחר שעברו התמרת DCT דו-ממדית, מסווגים ע"י אלגוריתם הקבצה (Clustering) לקבוצות. בסיוט הסוג מתקבלות קבוצות המכילות בלוקים בעלי תכולה ספקטרלית דומה. לכל קבוצה משוערך ספקטרום, אשר משמש לבניית הדיאגרמה, לקידוד הבלוקים השייכים לקבוצה ולאכלוסה. אותו ספקטרום נשלח כאינפורמציה צד למקלט ומשמש לבניית דיאגרמת שבכה זהה לצורך הפיענוח. בנוסף משמש הספקטרום המשוערך להקצאה אדפטיבית של קצבים בין הקבוצות השונות. אלגוריתם ההקבצה שמומש מבוסס על אלגוריתם LBG אשר משמש בד"כ לבניית ספר-קוד מתוך סדרת לימוד.

לצורך הפחתת "תופעת - הבלוקיות", אשר מאפיינת מקודדי התמרה בקצב נמוך, נעשה שימוש בקוונטיזציה סקלרית, אשר מקודדת את המקדמים בעלי העוצמה הגבוהה של תמונת השגיאה במישור ההתמרה. ערכי יציאת הקוונטיזציה הסקלרית ומיקום המקדמים שעברו קוונטיזציה בכל בלוק מקודדים באמצעות קודי הפמן (Huffman codes) ונשלחים כאינפורמציה צד למקלט.

המוטיבציה במימוש הסכמה המוצעת היתה השגת ביצועי קידוד טובים ככל האפשר, גם במחיר הגדלת סיבוכיות הסכמה. לשם כך נעשה שימוש במקודד המשלב התמרה וקידוד על שבכה. מקודד כזה הינו אופטימלי עבור אותות בעלי מודל גאוסי - מרקובי, אשר נחשב כמודל טוב לייצוג תמונות. הסטייה של התמונות המקודדות מהמודל קרוזה ע"י השימוש באלגוריתם ה Clustering ובקוונטיזציה הסקלרית.

התוצאות שהושגו באמצעות הסכמה המוצעת עדיפות על כל תוצאה אחרת שדווחה בספרות עד כה. הסיבוכיות במימוש הסכמה הינה גבוהה מאוד, אולם רובה מרוכזת בחלק המבצע את הקידוד. עובדה זו מאפשרת יישום של הסכמה במערכות לשמירה ואיחזור של תמונות, שבהן שלב הקידוד יכול להתבצע בצורה איטית, ואילו שלב הפיענוח חייב להיות מהיר. בעתיד צפוי פתרון לבעיית הסיבוכיות, עם התפתחות טכנולוגיית ה VLSI ופתוח מקודדי ויטרבי שיעבדו בזמן אמת [36].

במסגרת העבודה בוצעה השוואה אמפירית של ביצועי קוונטיזציה וקטורי הממומש באמצעות ספר קוד לקוונטיזציה וקטורי הממומש ע"י שבכה. תוצאות ההשוואה הראו שביצועי השבכה מזדהים עם ביצועי ספר הקוד בתחום רחב של אורכי וקטורים, בעוד שכמות החישובים הנדרשת לקידוד בעזרת השבכה, באותו תחום אורכי וקטורים, קטן. מכמות החישובים הנדרשת בספר קוד. תוצאה זו נותנת מוטיבציה לשימוש בשבכה לקוונטיזציה וקטורית.

7.2 כוונת המשך

אחת הבעיות העיקריות שעדיין קיימת בסכמה המוצעת הינה הטיפול בבלוקים אשר כוללים בתוכם את השפה בין אזורים שונים בתמונה. ניתן להבחין בבעיות הני"ל בתמונות התוצאה המופיעות בפרק הקודם. פתרון אפשרי לבעיה זו יכול להתממש ע"י הוספת מנגנון אשר יגלה את הבלוקים הני"ל (למשל גלאי שפה), וקידוד של הבלוקים באמצעים אחרים ולא באמצעות דיאגרמה שבכה. פתרון אפשרי אחר יכול להיות הפתרון הבא:

- (א) ביצוע חלוקה של התמונה לאזורים הומוגנים עד לרמת אלמנט תמונה.
 (ב) טריקת כל אזור כזה ויצירת וקטור המכיל את הפיקסלים של אותו אזור.
 (ג) ביצוע התמרת DCT חד-ממדית לכל וקטור
 (ד) קידוד הוקטור המותמר ע"י שבכה

בעיה אחרת אשר הוזכרה מוקדם יותר בחיבור זה הינה העובדה שמקדמי התמרת DCT של תמונות הינם בעלי פילוג לפלט ולא פלוג גאוס. שיפור נוסף בביצועים יכול להתקבל ע"י בניית שבכה אופטימלית לפילוג לפלט.

שיפור נוסף בביצועי הסכמה יכול להתקבל ע"י אכלוס השבכה מתוך סדרת לימוד כפי שמתואר בפרק 3.

בעיה קריטית בדחיסת תמונות הינה הערכת טיב הדחיסה. המדד הכמותי המשמש אותנו להערכת טיב הדחיסה הינו יחס האות לרעש (SNR), אשר לא תמיד מבטא נאמנה את איכות השחזור כפי שהיא נתפסת ע"י מערכת הראיה שלנו. למשל בעית ה"בלוקיות" שהוזכרה בעבודה זו הינה בד"כ בעלת אנרגיה נמוכה, ולעומת זאת היא פוגמת בצורה משמעותית באיכות התמונה כפי שהיא נתפסת במערכת הראיה. פתרון אפשרי לבעיה זו יכול להתממש ע"י בצוע ההקבצה וחלוקת הקצבים בין ובתוך הבלוקים בצורה שתתחשב במערכת הראיה.

כאמור בעבודה זו בוצעה השוואה אמפירית בין ביצועי ספר קוד וביצועי שבכה. ניתן להרחיב נושא זה בכונים של הגדרה מדויקת של התחזומים בהם עדיפים הביצועים הכוללים (SNR וסבוכיות) של השבכה על אלה של ספר הקוד. בנוסף, שאלה מעניינת אחרת הנובעת מהשוואה זו הינה השאלה הבאה: תחת אילו צי סבוכיות נתונים, אילו קוונטיזורים הינם בעלי הביצועים הטובים ביותר, מהם ביצועיהם ובאיזה מידה הם ממשים את אופטימום הביצועים תחת אילו צי אלה.

השיפורים וההמלצות שהוזכרו לעיל לא נבחנו במסגרת עבודה זו, אך להערכתנו הם עשויים לתרום לשיפור ביצועי המערכת, ולהוות בסיס להמשך המחקר בנושא.

נספח 1 - השוואת ספר קוד ושכבה בעלי כמות זכרון וחשובים זהה לאותו אורך וקטור מקודד, השוואה המובאת בנספח זה הינה עבור קוונטזרים הממומשים ע"י דיאגרמת שכבה וע"י ספר קוד, בעלי קצב זהה R הפועלים על וקטור באורך זהה N . הסימנים המופיעים בנספח זה מפורטים בטבלה 5.1.

נניח דיאגרמת שכבה מלאה, דהיינו דיאגרמה המתחילה מ q^{K_T-1} מצבים. הקצב במקרה זה נתון ע"י:

$$(1א) \quad R_T = \frac{L}{N} \log(q) + \frac{K_T}{N} \log(q)$$

האיבר השני בביטוי מסמן את הקצב הנדרש להעברת המידע על המצב ההתחלתי של השכבה. כמות הזכרון בשכבה נתונה ע"י:

$$(2א) \quad M_T = Nq^{K_T}$$

וכמות החשובים לאות מקוד נתונה ע"י:

$$(3א) \quad C_T = q^{K_T}$$

קצב ספר הקוד הינו

$$(4א) \quad R_B = \frac{L}{N} \log(K_B)$$

כמות הזכרון בספר קוד נתונה ע"י

$$(5א) \quad M_B = K_B N$$

כמות החשובים לאות מקוד בספר הקוד הינה

$$(6א) \quad C_B = K_B$$

מהביטוי עבור הקצב בשכבה (1א) מתקבל

$$(7א) \quad q^{K_T} = 2^{NR_B \frac{K_T}{L+K_T-1}} = 2^{NR \frac{K_T}{L+K_T-1}}$$

ומהביטוי עבור הקצב בספר קוד מתקבל

$$(8א) \quad K_B = 2^{NR_B} = 2^{NR}$$

השוואת (7א) ו (8א), דהיינו השוואת כמות הזכרון והחשובים נותנת

$$(9א) \quad 2^{NR} = 2^{NR \frac{K_T}{L+K_T-1}}$$

ממשוואה (9א) מתקבל התנאי המובא ב (5.8) ז"א

$$(10א) \quad \frac{K_T}{L+K_T-1} = 1$$

REFERENCES

- [1] B. Mazor & W. A. Pearlman, "A Trellis Code Construction & Coding Theorem for Stationary Gaussian Sources", IEEE Trans on Information Theory Vol IT-20 No. 3, pp 924-930, May 1983.
- [2] A. K. Jain "Image Data Compression : A Review ", Proceedings of the IEEE Vol. 69, No. 3, pp 349-389, March 1981.
- [3] A. N. Netravali & J. O. Limb " Picture Coding : A Review", Proceedings of the IEEE Vol. 68, No. 3, pp 366-406, March 1981.
- [4] W. K. Pratt, Digital Image Processing, New York: Wiley Interscience 1978.
- [5] R. J. Clarke, Transform Coding of Images, London Academic Press 1985.
- [6] J. Pearl, H. C. Andrews & W. K. Pratt "Performance Measures for Transform Data Coding", IEEE Trans. Comm., pp 411-415, June 1972.
- [7] A. Habibi, "Hybrid Coding of Pictorial Data", IEEE Trans. Comm., COM-22, pp 612-624, May 1974.
- [8] P. Wintz, "Transform Picture Coding", Proceedings of the IEEE Vol. 60, No. 7, pp 809-819, July 1972.
- [9] W. Chen & C. Smith, "Adaptive Coding of Monochrome and Colour Images", IEEE Trans. Comm., COM-25, pp 1285-1292 November 1977.
- [10] A. Habibi, "Survey of Adaptive Image Coding Techniques", IEEE Trans. Comm., Vol. COM-25, No 11, pp 1275-1284, November 1977.
- [11] N. S. Jayant & P. Noll "Digital Coding of Waveforms", Engwood Cliffs N.J. : Prentice Hall, 1971.
- [12] A. Rosenfeld & A. C. Kak, "Digital Picture Processing", London Academic Press 1985.
- [13] T. Berger, "Rate Distortion Theory - A Mathematical Basis for Data Compression", Engwood Cliffs N.J. : Prentice Hall, 1971.
- [14] Y. Linde, A. Buzo & R.M. Gray, "An Algorithm for Vector Quantization", IEEE Tran. Comm., Vol. COM-28, No. 1, Jan. 1980, pp. 84-95.
- [15] R. W. Gray "Vector Quantization", IEEE ASSP Magazine, April 1984.
- [16] H. Hang & J. W. Woods, "Predictive Vector Quantization of Images", IEEE Trans Comm., Vol, COM-33, No. 11, pp 1208-1219, Nov. 1985.
- [17] M. Goldberg, P. R. Boucher & S. Shlien, "Image Compression using Adaptive Vector Quantization", IEEE Trans Comm., Vol, COM-34, No. 2, pp 180-187, Feb. 1986.

REFERENCES (cont.)

- [18] A. A. Abdelwahab & S. C. Kwatra, "Image Data Compression with Vector Quantization in the Transform Domain", ICASSP pp 1285-1289, 1985.
- [19] W. A. Pearlman, M. M. Leung & P. Jakatdar, "Adaptive Transform Tree Coding of Images" ICASSP 85 pp. 145-148, Tampa, Florida.
- [20] B. Ramamurthi & A. Gersho, "Classified Vector Quantization of Images", IEEE Trans Comm., Vol, COM-34, No. 11, pp 1105-1115, Nov. 1986.
- [21] N. Ahamed, T. Natarajan, & K. R. rao, "Discrete Cosine transform", IEEE Tran. on Computers, C-23, pp 90-93, Jan 1973.
- [22] Y. Yemini & J. Pearl, "Asymptotic Properties of Discrete Unitary Transforms", IEEE Trans. on Pattern Analysis & Machine Intelligence, Vol. PAMI-1 No. 1, pp. 366-371, Oct. 1978.
- [23] J. Makhoul "A Fast Cosine Transform in One & Two Dimensions", IEEE Trans. Acoustics Speech and Signal Processing, Vol. ASSP-28, No. 1, pp. 27-34, Feb 1980.
- [24] A. J. Viterbi & J. K. Omura, "Trellis Encoding of Memory-less Discrete-Time Source with a Fidelity Criterion", IEEE Tran. on Information Theory, Vol. IT-20, No 3, pp 325-332, May 1974
- [25] S. G. Wilson & D. W. Lyle, "Trellis Encoding of Continues-Amplitude Memory less Sources", IEEE Trans. on Information Theory , pp 404-409, ay 1977.
- [26] L. C. Stewart, R. M. Gray & Y. Linde, "The Design of Trellis waveform Coders", IEEE Trans Comm., Vol, COM-30, No. 4, pp 702-710, Apr. 1982.
- [27] B. Mazor & W. A. Pearlman, "An Optimal Transform Trellis Code With Application To Speech ", IEEE Trans Comm., Vol. COM-33 No. 10, pp 1109-1116, October 1985.
- [28] R. C. Reininger & J. D. Gibson "Distribution of the Two-Dimensional DCT Coefficients for Images", IEEE Tran. Comm., Vol. COM-31, No. 6, June 1983, pp. 835-839.
- [29] J. D. Eggerton , "Optimal Visual Quantization for The Transform Source Encoding of Images", Ph.D. Dissertation, Southern methodist University, Nov 30, 1983.
- [30] P. A. Devijuer & J. Kittler "Pattern Recognition A Statistical Approach", Prentice-Hall London, 1982.
- [31] J. W. Woods & S. D. O'Neil, "Subband Coding of Images", IEEE Trans. Acoustics Speech and Signal Processing, Vol. ASSP-34, pp 1278-1288, Nov 1986.
- [32] H. Hang & J. W. Woods, "Predictive Vector Quantization of Images", IEEE Trans Comm., Vol, COM-33, No. 11, pp 1208-1219, Nov. 1985.

REFERENCES (cont.)

- [33] W. A. Pearlman, M. M. Leung & P. Jakatdar, "Adaptive Transform Tree Coding of Images" ICASSP 85 pp. 145-148, Tampa, Florida.
- [34] A. Tran & K. liu "An Efficient Pyramid Image coding System" ICCASP 87, pp 744-747.
- [35] V. J. Mathews, R. W. Waite & T. D. tran " Image Compression Using Vector Quantization of Linear (One-Step) prediction Error. ICCASP 87, pp 744-747.
- [36] S. Mohan & A. K. Sood, "A Multiprocessor Architecture for the (M,L)-Algorithm Suitable for VLSI Iplementation" IEEE Trans Comm., Vol, COM-34, No. 12, pp 1218-1224, Dec. 1986.

TRANSFORM TRELLIS CODING OF IMAGES

Research Thesis

Submitted in Partial Fulfillment of the requirement

for the degree of master of science

in

ELECTRICAL ENGINEERING

SHMUEL FARKASH

Submitted to the Senate of the Technion - Israel Institute of Technology

Adar 5748

Haifa

March, 1988

This paper was carried out in the Faculty of Electrical Engineering under the supervision of Prof. DAVID MALAH from the Technion Haifa and Prof WILLIAM PEARLMAN from RPI Troy N.Y.

I thank Prof. DAVID MALAH for his dedicated supervision, and his contribution to the Thesis achievements. I would also like to thank Prof WILLIAM PEARLMAN for proposing the subject and for his useful advice during the work.

I would like to thank Ziva Avni, Nimrod Peleg, Yoram Or-chen, and the other staff members of the Signal Processing Laboratory for their help and patience during the research.

This Research was partially supported by the Ministry of Industry and commerce

The generous financial help of the Guttwirth Fund is gratefully acknowledged.

CONTENTS

ABSTRACT	1
LIST OF ABBREVIATIONS	3
CHAPTER 1 - INTRODUCTION	6
1.1 Problem presentation - transform and trellis coding	6
1.2 Conventional coding schemes	7
1.3 Research topics and main results	12
1.4 Structure of the thesis	12
CHAPTER 2 - ORTHOGONAL TRANSFORMS - THEORETIC BACKGROUND	13
2.1 Introduction	13
2.2 Orthogonal transforms	13
2.3 Karhunen - Loeve transform	16
2.4 Sub-optimal transforms	18
2.5 The motivation of using DCT in the proposed scheme	20
CHAPTER 3 - VECTOR QUANTIZATION	23
3.1 Introduction	23
3.2 Vector quantization using code-book	23
3.3 Vector quantization using a trellis	26
3.2.1 Tree structured vector quantizers	27
3.2.2 Trellis structured vector quantizers	28
3.4 The motivation of using a trellis diagram in the proposed scheme	31
3.4.1 Introduction	31
3.4.2 Comparison with no constraints	32
3.4.3 Comparison under constraints of memory size and computations amount	36

CONTENTS (cont.)

3.4.4	The influence of the blocks length on the performance of the quantizers	37
3.4.5	Comparison under constraints of memory size or computations amount	40
3.4.6	Comparison under constraints of memory size and computations amount (cyclic population of the trellis)	42
3.4.7	Summary	44
CHAPTER 4 - BASIC CODING SCHEME		46
4.1	Introduction	46
4.2	Optimal scheme description	46
4.3	Trellis diagram construction	47
4.3.1	Rate distortion function of Gausse - Markov source	47
4.3.2	Diagram Construction	49
4.4	Diagram population	51
4.4.1	Random population	51
4.4.2	Iterative population	53
4.5	Encoding algorithms	54
4.5.1	The Viterbi algorithm	54
4.5.2	The M algorithm	55
4.5.3	The ML algorithm	55
4.6	Practical implementation considerations	56
4.6.1	Cosine transform - DCT	56
4.6.2	Trellis diagram construction and population	58
CHAPTER 5 - THE MISSMATCH PROBLEM OF IMAGES TO THE MODEL		61
5.1	Problem description	61
5.2	Using clustering to overcome the non-stationarity problem	62
5.3	Code adaptation	64
5.4	Using scalar quantizer to quantize the residual image	66

CONTENTS (cont.)	
5.5 Final scheme presentation	67
CHAPTER 6 - SYSTEM PERFORMANCE	69
6.1 Performance of the proposed scheme and comparison to other coding schemes	69
CHAPTER 7 - SUMMARY AND CONCLUSIONS	82
7.1 Thesis summary	82
7.2 Proposals for further research	82
Appendix A - Comparison of code-book and trellis with the same amount of memory and computations number	84
References	85
Abstract (in English)	I

ABSTRACT

In this research we present a coding scheme which encodes the Discrete Cosine Transform (DCT) coefficients of images using a vector quantizer based on a trellis diagram. A coding scheme composed of an optimal transform (KLT) and a trellis based vector quantizer is asymptotically optimal, i.e its performance converges to the rate distortion function for a Gausse-Markov signal. This source model is considered to be a good model for images. The reason for using a trellis diagram rather than a code book for vector quantization, stems from the linear dependence of its complexity on the coded vector length. This relative simplicity however, doesn't necessarily degrades its performance. As we show in this work, the performance of the trellis under computations' amount constraint is superior, in the range of the parameters that were checked, to those of the code book under the same constraint.

In the proposed scheme, the source image is divided into blocks, and each block undergoes a two dimensional discrete cosine transform (DCT). The transformed coefficients are less redundant and are almost uncorrelated, and hence an efficient compression is enabled. The transformed blocks are classified by a clustering algorithm into several classes. The clustering algorithm is based on the LBG algorithm which is usually used to construct code-books from a training sequence. At the end of the clustering process, each of the classes obtained include blocks having similar spectral content. A representative spectrum is estimated for each class and used to construct a trellis diagram for each class. The same estimated spectrum is sent as side information to the receiver, and is being used there to construct the same trellis diagram for the decoding process. The same spectrum is also used to to allocate rates among the classes. This rate allocation is essential for efficient exploitation of the allowed overall rate by allocating higher rates to classes with high AC energy.

The blocks in each class are encoded by the trellis diagram that corresponds to that class. To further improve the scheme's performance, particularly to reduce the blocking-effect problem, which characterizes transform coding schemes which operate at low bit rates, a scalar quantizer is used to encode the high magnitude coefficients of the error image in the transform domain. The values and locations of those coefficients are encoded by Huffman codes and sent as side information to the receiver.

In the decoding process, each block is decoded by the corresponding trellis diagram, and a correction of the coefficients that were quantized by the scalar quantizer is being done. The block is inverse transformed to obtain the reconstructed block. The same procedure is repeated for every block in the image.

The Gausse-Markov model assumption is not a very good one, because of the non-stationary nature of images. The deviation of the images from the assumed model is compensated in the proposed scheme by the clustering process and the scalar quantizer used. The

performance obtained with the proposed scheme is characterized by two measures as follows:

- a) The signal to noise ratio (SNR) computed from the error between the original image and the reconstructed one.
- b) The subjective quality of the reconstructed image.

The results obtained were very good and even excellent. For the the woman image ("LENA") compression to rates below 1 *bit/pel* with high SNR values and very good image quality is achieved (more than 32 *dB* at 0.6 *bit/pel* , and 35 *dB* at 0.9 *bit/pel*). For a high resolution image (image size of [512×512]) rates below .5 *bits/pel* were obtained with excellent performance (37 *dB*). To our knowledge such good results were not reported earlier in the literature with any coding scheme on the same test image.

The complexity of the proposed coding scheme is indeed very high. However, most of the complexity is in the encoder part of the scheme. This fact enables employment of the proposed scheme in image storage and retrieval systems, where the decoder's simplicity is a essential property. With the continued development of VLSI technology, this scheme could in the future be implemented also in real-time systems.