

Ben Gurion University of the Negev

Department of Electrical and Computer Engineering

Multimedia Compression SEM 1 – Final Project

Image Compression – From DCT to Wavelets

Tal Shkuri: 316411974

Itamar Guerchon: 209703941

Guy Mayost: 316414697

מוטיבציה

עם התקדמות הטכנולוגיה, איקות המולטימדיה הולכת ומשתפרת - תמונות וסרטוני וידאו מצולמים באיכותות גבוהה יותר, הרזולוציות מתחדשות וכתוואה מכך כמוות המידע בכל תמונה או סרטון וידאו שכזה גדולה גם כן. הדבר מעלה קושי לאחסן ולשדר את נתוני המולטימדיה ברוחב פס בגודל נתון, עם כמוות אחסון מוגבלת שלא גדול באופן פרופורציונלית לכמות המידע. לפיכך, נרצה למצוא דרך אפקטיבית יותר לדחוס נתונים מולטימדיה.

רקע תאורטי

כפי שלמדנו בכיתה, טרנספורמציה ה-(DCT) Discrete Cosine Transform שימושה כבסיס לדחיסת JPEG מאז 1992. התמורה זו מאופיינת בקרוב נתונים שכן היא ממירה ממוקם למרחב התדר באמצעות סדרת קוסינוסים. התמונה מחולקת ל- 8×8 בלוקים עליהם מתבצעת התמורה ה-DCT. מאפיין עיקרי נוסף הוא שרב האנרגיה של האות מרווחת בתדרים הנמוכים, מה שמאפשר לדחוס נתונים ע"י התעלמות מהתדרים הגבוהים. את התמורה ה-DCT ניתן ליישם על ידי אלגוריתמים עם סיבוכיות זמן נמוכה (FFT) מה שהופך אותה ליעילה בשימוש בזמן אמיתי.

Subband Coding היא שיטה לדחיסת אותות מוחולק בתחוםי תדר שונים הנקראים subbands. בתדרים הנמוכים מוכל רב המידע לעומת התדרים הגבוהים שמכילים פרטים עדינים יותר הנקראים Discrete Wavelet Transform (DWT) הוא סוג של Subband Coding המבצע את החלוקה בתחוםי התדרים השונים. ה-DWT מתאפיין בתכונת הרבה רזולוציה – הוא מפרק את אותות לرمות תדר שונות ומאפשר שליטה על הרזולוציה. תכונה נוספת של התמורה זו היא הפירוק היררכי שהיא מבצעת. ה-DWT מחלק את אותות ל-Subbands ברמות שונות כאשר התדרים הנמוכים נקראים approximation והתדרים הגבוהים הם details. כמו כן, ההתמורה מותאמת למערכת הראייה האנושית משום שהאחרונה מסוגלת להבחין בטווח תדרים סופי. היכולת של ה-DWT לפפרק ולבורר רמות תדרים מסוימות מייעלת את הדחיסה ברמות דחיסה אגרסיביות כאשר תבחר להישאר עם התדרים הרלוונטיים בלבד.

הצגת הבעיה

המאמר מפרט על כך שישיות מבוססת DCT מוגבלות ביכולתן להתמודד עם מספר בעיות. בעיה ראשונה היא שכשר דוחסים בקצבוי שידור ביטים נמוכים, כולם כאשר אנו נאלצים לדחוס באופן אגרסיבי יותר – ככל הנראה בשל חוסר זמינות ברוחב הפס, אנו נתקלים בתופעת ה-Block-based artifacts. אלה נגזרים בגל החלוקה לאזורים קטנים (בלוקים בגודל 8×8 פיקסלים) ושיטת הקוונטיזציה שבוצעת על כל בלוק בנפרד. הסיבה העיקרית לכך היא שיטתה ה-DCT אוספת מידע על תדרים עבור כל בלוק בנפרד ולא על התמונה כולה, מה שעלול להוביל לאובדן קשרים בין הפיקסלים בקצוות של הבלוקים. בעיה שנייה היא העמידות המוגבלת שיש לאותם מקודדים מבוססי DCT. עמידות מתיחסת ליכולת של מערכת דחיסה לתמונות לתקן בתפקידו ולהשמור על איקות תמונה גבוהה גם בתנאים קשים כגון שגיאות בשידור, רעש נתונים או מחיקות חלקיות של מידע. משתמש מכך, שמערכות דחיסה עמידות יודעות להתמודד עם אובדן נתונים ויודעות לשחזר ולקבל תמונה איקותית גם אם חלק מהמקטעים נפגעים. בעיה נוספת שלאותם מקודדים מבוססי DCT אין תמייה ניהולית טובה של דרגות דחיסה שונות וקבالت רזולוציות שונות של התמונה בקלות. במקרים אחרים, הם בעלי סקלאbilיות מוגבלת.

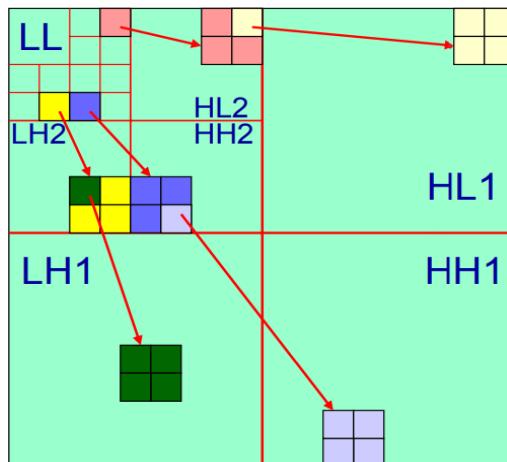
הפתרון

הפתרון שהמאמר מציג הוא שימוש בשיטות דחיסה מבוססות DWT. טרנספורמציה ה-wavelet, לעומת ה-DCT, מספקת דחיסה איכוטית יותר גם כאשר מושדרים בקצבם המקוריים ונאלצים לדחוס באופן אגרסיבי. הדבר בא לידי ביטוי בכך שאוותן שיטות מבוססות DWT נמנעות מטופעת ה-Block-based artifacts וبنוסף תומכות בשידור פרוגרסיבי – המידע הבסיסי נשלח תחילת, ומידע נוסף לשיפור איכוט התמונה נשלח בהמשך במידה ואנחנו רואים שהטפנה מקום ברוחב פס השידור. לדוגמה, שידור תמונה ברזולוציה נמוכה שמושלמת עם הזמן עד לכדי רזולוציה גבוהה יותר.

המאמר מציג מספר טכניות מתקדמות שמנצלות את התכונות של קידוד באמצעות DWT.

Embedded Zerotree Wavelet (EZW)

השיטה ראשונה היא שיטת ה- Embedded Zerotree Wavelet (EZW) שכאמור פותחה על ידי שאפיירו בשנת 1993. תחילתה, מקדמי ה-DWT מפרקם באמצעות מבנה פס-אוקטבה(Octave-Band) שמחلك את המקדים ל- תתי-פסים (Octave-Bands) בرمמות תדר שונות. פירוק זה מאפשר לכלוד גם פרטיים כליליים וגם פרטיים עדינים בתמונה.

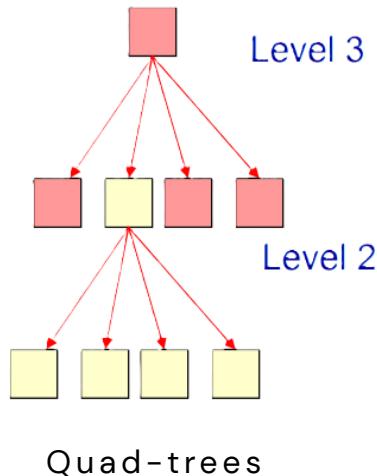


Octave-Band Wavelet Decomposition

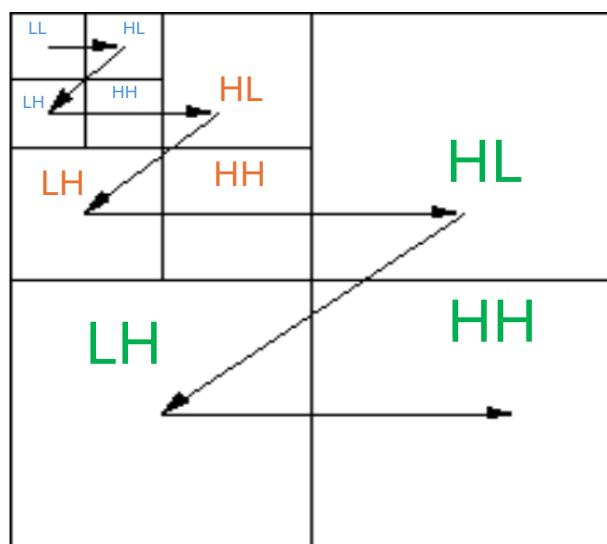
בפירוק Octave-Band Wavelet Decomposition כפי שניתן לראות באירור שלעיל, מתקבל שכל level בניו מה- approximation coefficients היכן שהמידע הכללי, כדוגמת קווי המתאר של עצמים בתמונה שומר. מידע זה שומר בתדרים הנמוכים – LL (Low Low). מסבבים נמצאת details – פרטים הגבוהים, היכן ששמורים הפרטים העדינים יותר כמו הקצוט והtekstורה של התמונה. אלה הם ה-HL, HH, LH.

התচום כאן הוא שההיליך הפירוק עצמו אינו דורש ידע מקדים על סוג התמונה או מאפייניה, אלא הוא מבוסס על התכונות המובנות של ה-DWT שמייצרות הפרדה טבעית בין מבנה כללי של התמונה לבין פרטיה. המקדים החשובים ביותר ממוקמים בתדרים הנמוכים ונשמרים בעדיפות גבוהה, בעוד שהמקדים בתדרים הגבוהים מזונחים אם אינם שימושתיים. ניתן לתמוך הרב-רזולוציה של ה-DWT שקיים מלכתחילה, והוא זה שמאפשר לייצג את התמונה בرمמות פירוט שונות באופן שכזה.

עם זאת, מבנה הפס-אוקטבה מצריך שיטה חכמה יותר לקידוד מקדמי ה-wavelet על מנת להשיג תוצאות דחיפה טובות יותר. בשנת 1992 לואיס ונואלס הציגו לראשונה לבנייה נתוניים דמו-עץ לייצוג המקדים של פירוק פס-אוקטבה. תוכנות ההיררכיה המובנת של ה-DWT מארגנת את המקדים כך שאלה שנמצאים בתתי הפסים ברזולוציה הנמוכה פועלים כ"הוריות" למקדים שבתתי הפסים ברזולוציה הגבוהה יותר, ובכך מאפשרים ייצוג קומפקטי.



ברגע שהערך של אחד הצלאים של תען עץ מסוים הוא מתחת לסקף threshold מסויים T שנקבע מראש, סבירות גבוהה ששם הילדים של אותו צאצא יהיו מתחת לאותו T . במצב זה, נוכל להזין ערך End-of-block ולא נצטרך להמשיך לرمות הנמוכות יותר בתת העץ שמתאימות לצאצא ממנו יצאתי. כתוצאה לכך אנו מקבלים חישכון במידע מהCHASE ויזואלית לכך ניתן לראות באירוע שמצוג לעיל.



אופן השחזר מתחילה מה- level (LL) של ה- approximation coefficients (כלומר מה-LL) של ה- level ה- details, את ה- details, על פי הצורך לפירוט נוסף, בסדר הבא : HH → LH → LL. במידה וסימן לפענה את כל ה- details של אותו ה- level אז ימשיך ל- LL של ה- level הבא באותו האופן הניל. המשחה לכך ניתן לראות באIOR שבסוף העמוד הקודם : השחזר מתחילה מה- LL של 3 level וועובר ל- details שלו. כאשר סיים, הוא ממשיך ל- 2 level בדומה לכך שככל 3 level מהויה בעצם את ה- LL של 2 level, וכך ממשיך להלאה ל- 1 עד שמסיים את תהליך השחזר.

היתרון באופי תהליכי השחזר הניל הוא שאנו יכולים לעצור אותו בכל עת כרצוננו, על פי זמינות רוחב הפס. כך אנו מאפשרים בעצם את תוכנת השידור הגרפי שהסבירנו קודם לכן.

Set Partitioning in Hierarchical Trees (SPIHT)

שיטת השנייה שפותחה בעקבות ה- EZW היא שיטת ה- Set Partitioning in Hierarchical Trees (SPIHT). שיטת ה- SPIHT לוקחת את עקרונות ההיררכיה והיעילות של ה- EZW ומשפרת אותן באמצעות מיון וחלוקת חכמים יותר של קבוצות המקדים. החלוקה מתבצעת בשני שלבים : זיהוי מקדים משמעוניים ולאחר מכן חלוקתם לקבוצות. בשלב הראשון, מקדים בעלי ערך גדול מסף threshold מסוימים ממשמעוניים. בשלב השני, הקבוצות ההיררכיות (trees) מחולקות שוב ושוב, כאשר כל שלב מתמקד במקדים החשובים ביותר.

לשיטה ה- SPIHT קיימים שיפורים ניכרים על פני שיטת ה- EZW. הראשון מביניהם הוא סדר המיון הדינמי. בעוד EZW משתמש על "עצים אפסים סטטיים" (Zerotrees), ה- SPIHT ממין מקדים לפי חישובותם באופן דינמי. כמו כן, ה- SPIHT יעיל יותר בשל חלוקת הקבוצות שנעשה באלגוריתם שזו מצמצמת את הצורך במידע נוסף למידוד ומפחיתה את כמות הביטים הנדרשת. יתרה מזאת, שיטת ה- SPIHT משפרת את יכולת השידור הגרפי בכך שמאפשרת שידור מדורג שבו התמונה משתפרת ככל שמתקבלים יותר נתוניים.

Scalable Image Compression with EBCOT

השיטה השלישית שפותחה בעקבות ה- EZW ושיטת ה- SPIHT הוא אלגוריתם EBCOT (Embedded Block Coding with Optimized Truncation).

ה- EBCOT משתמש בטכנולוגיית ה- Wavelet הבדידה לפירוק התמונה, ולאחר מכן מבצע קוונטיזציה וקידוד של מקדמי ה- Sub-bands. האלגוריתם פועל על ידי חלוקת כל Sub-band לבלוקים קטנים (Code-Blocks), כאשר כל בלוק מקודד בנפרד, מה שמאפשר גמישות גבוהה בדחסיה ושיפור הסקלאביליות. נזכיר כי סקלאביליות (מדרגיות) מתייחסת ליכולת של מערכת לתמוך במגוון רזולוציות או קצבי ביטים, כך שהתמונה יכולה להישמר ונitin לשחזר אותה באיכות שונות ללא צורך בדחסיה חוזרת. המשמעות היא שניתן להפיק מהתמונה גרסאות שונות בהתאם לצורך, למשל לשימושים בעלי מגבלות רוחב פס או דרישות רזולוציה שונות.

אחד היתרונות של ה-EBCOT הוא הסקלאלibilitות מלאה. ככלمر האפשרות לדוחס ולשדר תמונה ברזולוציות שונות וביחס דחיסה מגוונים ללא צורך בדחיסה מחדש. זה מאפשר שליטה גמישה יותר על איכות ורזולוציה. יתר על כן קיימת הגישה אקראית. ככלמר, ניתן לגשת לבlokים בודדים באופן עצמאי, מה שמקל על עיבוד ותצוגה של חלקים בתמונה. שליטה מדויקת יותר בשחרור נתונים מתאפשר בשל החלוקה של כל Sub-band לבlokים קטנים. ושיפור השידור הプログסיבי. בזכות הגישה האקראית ניתן לשדר רק את הזורמים הרלוונטיים, מה שמוסיף דיוק לבחירת הנתונים ומיעיל את קצב השידור בהתאם לדרישות הרזולוציה והאיכות.

הSHIPORIM העיקריים ביחס לשיטות הקודמות הם מצויים בזבוז הביטים ושהזר תמונה איכותית יותר בرمאות שונות. מצויים בזבוז הביטים מתאפשר בזכות חלוקה חכמתה של הזרם הדוחס ויכולת התאמת מדויקת של הדחיסה לכל בלוק. שהזר תמונה איכותית יותר בرمאות שונות מתאפשר בזכות חלוקה לבlokים עצמאיים ושימוש בטרנספורם ה-Wavelet הבודד. זה מאפשר קוונטיזציה מדויקת ושיפור יחס אות לרעש (SNR).

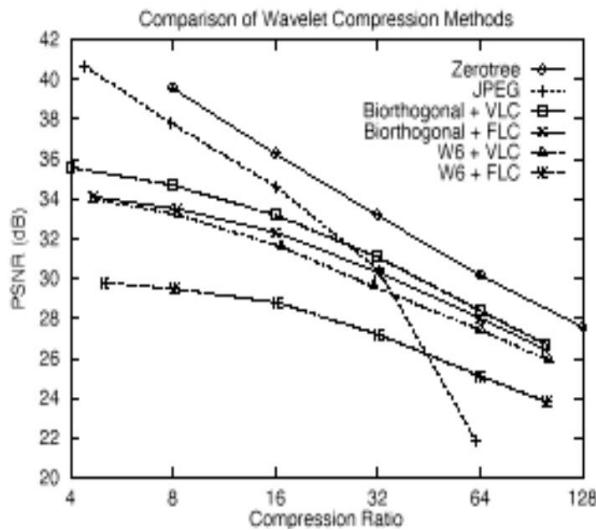
Lossless Image Compression using Integer to Integer WT

השיטה הרביעית שפותחה היא דחיסה לא איבוד נתונים (Lossless Compression) חיונית לתמונות רפואיות, נתוני סייסמוגרפיה (מדידות של גלי סייסמיים הנוצרים בפעילות גיאולוגית כמו ברעידות אדמה), תמונות לוויין וטכנאי וידאו באיכות גבוהה. ה-JPEG מציע קידוד Lossless על בסיס קידוד הפרשי פיקסלים, אך עקב גודל התמונות הגדול, יש צורך בשיטות המאפשרות אחסון ושירור פרוגרסיבייעיל.

למרות שהשימוש ב-Wavelet מתאים לדחיסה פרוגרסיבית, אחת הביעות המרכזיות היא שה-Output שעבר סינון מתקיים בערכים עם נקודה צפה (Floating Point). זה מקשה על שהזר Lossless של הנתונים. הפתרון לכך הוא טרנספורם Wavelet הפיך מסוג Integer-to-Integer, שמיר נתונים שלמים לערכים שלמים תוך שמירה על היכולת לפענן את התמונה באופן מדויק לחלוטין.

שיטה זו מtabסת על שלבי הרמה (Lifting Steps) המאפשרים בניה גמישה של טרנספורמים שלמים לכל תמונה. מספר טרנספורמים כאלה יושמו ונבדקו, והמסקנה היא שהם מאפשרים דחיסת תמונות ללא איבוד נתונים תוך תמייה בשידור פרוגרסיבי. ככלמר, שידור הדרוגתי שבו ניתן להעביר גרסה ברזולוציה נמוכה של התמונה תחילתה ולאחר מכן לשדר פרטים נוספים לשיפור האיכות.

השוואה בין DCT ל-DWT



Comparison of Wavelet Compression

יחס דחיסה :

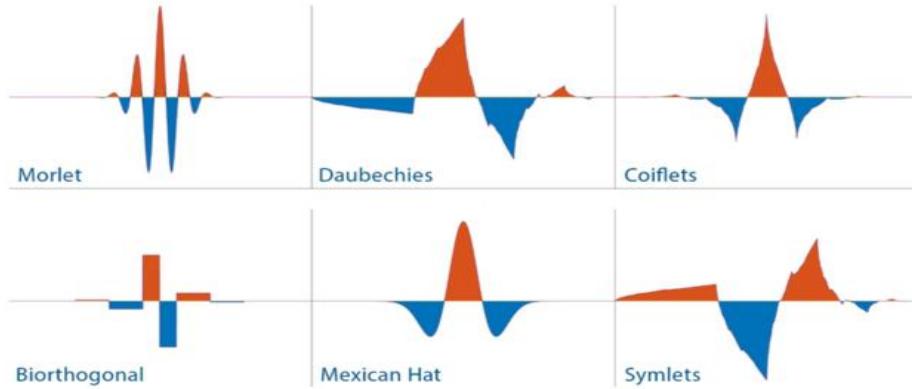
ביחס דחיסה נמוך של עד 1:25 ניתן לראות כי דחיסת JPEG מבוססת על DCT מציגה ביצועים טובים יותר מאשר דחיסת DWT (פרט-ל-פרט). ביחס דחיסה גבוהה של מעל 30 ניתן לראות כי הביצועים של ה-DCT מדרדרים במהירות בשל מגבלות הבלוקים הקבועים של DCT, מה שיוצר את בעיית artifacts-block-based. לעומת זאת, קידודים המבוססים על ה-Wavelet הבודד מדרדרים באופן הדרגתי ומתוון יותר גם ביחס דחיסה של 1:100, תוך שמירה על איכות התמונה טוביה יותר. ה-Zerotree ודחיסתWavelet מתאימים יותר לדחיסה גבוהה בזכות עמידותם ותוכנת השחזור האיכותית שלהם גם בرمמות דחיסה גבוהות מאוד.

תלות במקדים :

מקדים ה-DCT קבועים ותדריו מיוצגים במטריצה של 8×8 לעומת ה-DWT שהוא תלוי בסוג המسان. מכיוון שאין מסן יחיד שמתאים לכל התמונות באופן אופטימלי, אז קידוד תמונות מבוסס DWT דורש התאמת של מסן ה-Wavelet בהתאם למאפיינים הסטטיסטיים של התמונה. בנוסף ניסויים הראו שביצועי הדחיסה (Lossy ו-Lossless) תלויים בסוג התמונה. לעומת זאת, מסננים מסוימים עובדים טוב יותר על תמונות ספציפיות, אך אין מסן אחד שמתפרק בצורה מיטבית על כל סוגי התמונות. עופר גם בהרצאה שלנו הוסיף הערת ידע כללי ו אמר שבאופן כללי המسان של Daubechies עובד היטב כיום עבור תמונות רנדומליות אך זה לא חלק מהחומר של המאמר, סתם הערת ידע כללי).

לכן, בחירה אדפטטיבית של מסן ה-Wavelet חיונית. בנוסף, גם שימוש בקונטיזציה ומודלים סטטיסטיים מתקדמים לא יפצה על מסן לא מתאים.

תמונה של דוגמאות של משפחות אם של Wavelet Daubechies ושל Wavelet Daubechies עליה עופר נתן את הערת הידע הכללי.



Mother wavelet functions

השורה התחתונה של ההשוואה והמאמר הוא :

DWT אינו מהוות תחליף ל-DCT אלא משמש לשיפורו, במיוחד כאשר יש צורך להפחית ארטיפקטים ולהתמודד עם דחיסות גבהות יותר.

ביקורת על המאמר

המאמר מספק סקירה מצוינת על המעבר מ-DCT ל-DWT, אבל סובל ממיקוד חד-צדדי ומחוסר ניתוח מעשי מספק. הוא מועיל להבנת השיטות החדשות, אך חסר נתונים אמפיריים מעמיקים, ניתוח חישובי, והשוואה לטכניקות דחיסה אחרות. היה ניתן לשפר אותו למשל עם הוספים השוואות מקיפות יותר, ניתוח של חסרונות ה-DWT, וධינן על היישום המעשי של השיטות השונות.

להלן ליתרונות וחסרונות :

• יתרונות :

- סקירה מקיפה ומעמיקה – יש תיאור מפורט על האבולוציה של שיטות הדחיסה, כולל היתרונות והחסרונות של טכניקות שונות, תוך שילוב רקע מתמטי ואנליזה של הביצועים.
- פרקטיות ויישומים מגוונים – המאמר מדגיש כיצד שיטות כמו SPIHT ו-EBCOT מתאימות ליישומים קרייטיים, כולל שידור דיגיטלי, דחיסה סקלאלרית, דחיסת תמונות רפואיות ולוויניית, ומסביר כיצד תוכנות אלו מספקות גמישות רבה יותר לעומת JPEG המסורתית.

• חסרונות :

- מיעוט דוגמאות מעשיות – חסר נתונים אמפיריים רחבים או הדגומות מעשיות של ביצועי השיטות השונות על תמונות שונות. השוואות-PSNR מוצגות באופן תיאורטי בלבד, אך לא מספקות ניתוח מפורט של התוצאות בפועל.
- מרכיבות חישובית לא מוסברת – אין התייחסות מספקת לדרישות החישוביות בשיטות כמו SPIHT ו-EBCOT. אין ניתוח לגבי צרכות הזיכרון, זמן הריצה או ההשפעה של פרמטרים שונים על הביצועים, מה שעלול להקשות על הבנת היישום המעשי של השיטות.

- היעדר ביקורת על DWT - המאמר משבח את יתרונות ה-DWT אך אינו מציג את האתגרים הטמונהים בו, כמו הצורך בבחירה פילטרים מתאימים והשפעתם על איכות השחזר. כמו כן, לא מוזכרת בעיתת המורכבות של התמרת ה-Wavelet הבדיד לעומתWavelet DCT, מה שמנע נקודת מבט מאוזנת יותר.
- מיקוד חד-צדדי – המאמר מתמקד כמעט לחלוטין בטכניקות מבוססות Wavelet, ואינו משווה אותן לטכניקות דחיסה מודרניות אחרות, כגון למידת מכונה או אלגוריתמים היברידיים המשלבים ו-DWT. כתוצאה לכך, הקורא אינו מקבל תמונה רחבה של התחום כולו.

המלצות להמשך העבודה

- 1) שיפור טכניקות ה-DWT – ניתן לשפר את בחירת מסנני ה-Wavelet ואת התאמת שלהם לסוגי תמונות שונים (לוויין, רפואיות, טבעי וכו'), להפחית artifacts ולשפר את ה-PSNR. ניתן לעשות זאת על ידי שימוש במידה מוגנה לבחירת מסננים דינמיים שמתאימים לכל תמונה באופן אדפטיבי ונitinן לביצוע אופטימיזציה של המסננים באמצעות מאגרי תמונות גדולים ולבוחן השפעת פילטרים שונים על איכות התמונה לאחר דחיסה ושחזר.
- 2) אינטגרציה עם למידת מכונה – ניתן לשפר את יעילות הדחיסה והתאמת האוטומטית בהתאם למאפייני התמונה. ונitinן לשפר את השיטות כמו ה-H-T ו-SPIHT-EBCOT באמצעות אלגוריתמים חכמים יותר. ניתן לעשות זאת על ידי אימון רשתות נוירוניות לניטוח מאפייני התמונה ולבחרור את פרמטרי הדחיסה האופטימליים. בנוסף, לשלב במידה מוגנה עם השיטות כמו ה-H-T ו-SPIHT-EBCOT כדי ליעיל את הדחיסה בזמן אמת ולשפר את השחזר תוך שמירה על איכות גבוהה.
- 3) העברת מידע ועמידות לשגיאות בסביבות רועשות – ניתן לשפר את עמידות הדחיסה לשגיאות בתנאי תקשורת לא יציבים, למשל, העברת תמונות בלויינים, טלפונים סולריים ורשתות אלחוטיות על ידי שילוב דחיסת Wavelet עם פרוטוקולי תיקון שגיאות כדי לשמר על איכות התמונה גם בתנאים של איבוד נתונים או רעשים חיצוניים. ונitinן גם לפתח אלגוריתמים מותאמים לשידור תמונות בסביבות רועשות.
- 4) דחיסה רב-מודאלית – ניתן לישם טכניקות דחיסה גמישות יותר שיכולים להתמודד עם סוגים מודדים שונים כמו טקסט, תמונות, וידאו ותמונות מושלבות ונitinן לשפר את האחסון והשידור של מידע מגוון בצורה אופטימלית. ניתן לעשות זאת על ידי פיתוח מודלים המשלבים דחיסת Wavelet עם טכניקות כמו קידוד הופמן (Huffman Coding) או דחיסה פרדיקטיבית. בנוסף, ניתן לישם שיטות דחיסה שמתאימות למידע רב-מודאלי, למשל עיבוד תמונה וtekst יחד, מה שיכול לשמש ליישומים מתקדמים כמו זיהוי טקסט מתוך תמונות.