INTRODUCTION TO DIGITAL IMAGE PROCESSING

361.1.4751

EXERCISE 3 - Transformations

Submission Date: 18.11.2022

Noam Atias – 311394357

Chanel Michaeli - 208491787

תוכן עניינים

1. 2D – FOURIER TRANSFORM 3

1.1 WRITE YOUR OWN FUNCTIONS 3

1.1.1פונקציה המבצעת התמרת פורייה דו ממדית 3

1.1.2 Shift zero frequency component to center 4

1.1.3Read the 'beatles' image 5

1.1.4 2D – FFT and shift 5

1.1.5 Reconstruct the original image 7

1.2 TRANSFORMATION PROPERTIES 8

1.2.1 Linearity (Free Willy) 8

(a) Display image 8

(b) Spatial frequency of prison bars - fx 9

(c) Prison bars image 10

(d) 2D – FFT of the prison bars 11

(e) Free Willy 12

1.2.2 Scaling, translation, and separability 13

(a) Box $40X40$ in the center 13

(b) Box $40X40$ in location $M128 \times 12864{:}\,103$ת$64{:}\,103$ 15

(c) Box 80X20 in the center 17

(d) Represent the image using 1D vector 18

(e) Compute 2D – FFT by using 1D – FFT 19

2. DISCRETE COSINE TRANSFORM (DCT) 20

2.1 READ THE 'BEATLES' IMAGE 20

2.2 DISPLAY THE DCT OF THE IMAGE 20

2.3 RANDOMLY SET TO ZERO 50% OF THE DCT VALUES 21

2.4 SET TO 0 THE LOWEST 50% ABSOLUTE VALUES OF THE DCT 22

2.5 SET TO 0 THE VALUES IN THE RANGE OF $-a, a$ 23

2.6 EXPLANATION OF THE OBTAINED RESULTS FROM SECTIONS 3 – 5 25

3. WAVELET TRANSFORM 26

3.1 READ THE 'BEETLE.JPG' IMAGE 26

3.2 EXTRACT THE WAVELET DECOMPOSITION, LEVEL BETWEEN 3 – 5 26

3.3 FIND THE DETAIL AND THE APPROXIMATION COEFFICIENTS FOR EACH LEVEL 27

3.4 DISPLAY THE RESULTS 27

3.5 EXPLANATION FOR: 28

(a) What are the differences between the details and the approximation coefficients? 28

(b) What are the differences between the horizontal, vertical, or diagonal coefficients? 28

(c) What are the differences between each level? 28

1. 2D – Fourier Transform

1.1 Write your own functions

1.1.1 פונקציה המבצעת התמרת פורייה דו ממדית

התבקשנו לכתוב פונקציות משלנו, כאשר הפונקציה הראשונה מבצעת התמרת פורייה דו ממדית ופונקציה השניה היא ההופכית שלה, כלומר התמרת הפורייה ההופכית.

לפונקציה הראשונה נקרא: *'dip_fft2()'* ונממש אותה על פי ההגדרה להתמרת הפורייה הדו ממדית כך:

$$F(u+1,v+1)=\sum_{m=0}^{M-1}\sum_{n=0}^{N-1} I(m+1,n+1)e^{-i2\pi\left(\frac{um}{M}+\frac{vn}{N}\right)}$$

<u>כאשר:</u>

- I – התמונה שלה נרצה לבצע את ההתמרה הדו ממדית.
- גודל התמונה הוא: $M\times N$, כלומר $I_{M\times N}$

לצורך בניית הפונקציה נסתכל על ההתמרה הדו ממדית באופן כללי:

$$F(u,v)=\sum_{m=0}^{M-1}\left[\sum_{n=0}^{N-1} I(m,n)e^{-i2\pi\frac{vn}{N}}\right]e^{-i2\pi\frac{um}{M}}=\left\{I'(m,v)=\sum_{n=0}^{N-1} I(m,n)e^{-i2\pi\frac{vn}{N}}\right\}=$$

$$=\sum_{m=0}^{M-1} I'(m,v)e^{-i2\pi\frac{um}{M}}$$

כלומר $I'(m,v)$ הוא ההתמרה החד ממדית כאשר m מיוחס כפרמטר ולכן נוכל לרשום:

$$I(m,:)=I'_m=(W_N^*\cdot I_m)^T$$

<u>כאשר:</u>

- $W_N\in\mathbb{C}^{N\times N}$, ו - $W_N^{vn}=e^{i2\pi\frac{vn}{N}}$, כאשר n הוא השורה ו – v הוא העמודה.
- $I_m\in\mathbb{R}^{1\times N}$ – הוא וקטור השורה ה – m ית של התמונה.
- אופרטור ה - $(\cdot)^*$ (כוכבית) מסמל צמוד.

כאשר נרוץ על כל השורות של התמונה I אשר נוכל להגדיר:

$$I'=[I'_0\ ,I'_1\ ,I'_2\ ,\dots,I'_{M-1}]^T=[I_0\ ,I_1\ ,I_2\ ,\dots,I_{M-1}]^T\cdot W_N^{*T}=\left[W_N^{*T}=W_N^*\right]=$$

$$=\begin{bmatrix} I_{0\ (1\times N)}\\ I_{1\ (1\times N)}\\ \vdots\\ I_{M-1\ (1\times N)}\end{bmatrix}_{M\times N}\begin{bmatrix} e^{-i2\pi\frac{0\cdot 0}{N}} & e^{-i2\pi\frac{1\cdot 0}{N}} & \cdots & e^{-i2\pi\frac{(N-1)\cdot 0}{N}}\\ e^{-i2\pi\frac{0\cdot 1}{N}} & e^{-i2\pi\frac{1\cdot 1}{N}} & & e^{-i2\pi\frac{(N-1)\cdot 1}{N}}\\ \vdots & & \ddots & \vdots\\ e^{-i2\pi\frac{0\cdot(N-1)}{N}} & e^{-i2\pi\frac{1\cdot(N-1)}{N}} & \cdots & e^{-i2\pi\frac{(N-1)\cdot(N-1)}{N}}\end{bmatrix}_{N\times N}=I\cdot W_N^*$$

כעת נבצע את ההתמרה החד הממדית השניה על $I'\in\mathbb{C}^{M\times N}$:

$$F(u,v)=\sum_{m=0}^{M-1} I'(m,v)e^{-i2\pi\frac{um}{M}}$$

באופן דומה כעת הסכימה תתבצע על m. נוכל לרשום זאת בכתיב מטרציוני כך:

$$F(:,v) = F_v = W_M^* \cdot I'_v$$

<u>כאשר:</u>

- $W_M \in \mathbb{C}^{M\times M}$, ו - $W_M^{um} = e^{i2\pi\frac{um}{M}}$, כאשר m הוא העמודה ו – v הוא שורה.
- $I'_v \in \mathbb{R}^{M\times 1}$ – הוא וקטור העמודה ה – v ית של I'.
- אופרטור ה - $(\cdot)^*$ (כוכבית) מסמל צמוד.

כאשר נרוץ על כל העמודות של I' אשר נוכל להגדיר:

$$F = [F_0 , F_1 , F_2 , \dots , F_{N-1}] = W_M^* \cdot [I'_0 , I'_1 , I'_2 , \dots , I'_{N-1}] = W_M^* \cdot I'$$

לסיכום נוכל לרשום את ההתמרה הדו ממדית בצורה מטרציונית כך

$$F = W_M^* \cdot I \cdot W_N^*$$

לפונקציה השניה נקרא: *'dip_ifft2()'* ונממש אותה לפי ההגדרה להתמרת הפורייה ההופכית הדו ממדית כך:

$$I(m+1,n+1) = \frac{1}{MN} \cdot \sum_{u=0}^{M-1}\sum_{v=0}^{N-1} F(u+1,v+1)e^{i2\pi\left(\frac{um}{M}+\frac{vn}{N}\right)}$$

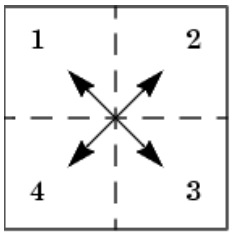
בצורה דומה לפיתוח שהראינו עבור הפונקציה להתמרה הדו ממדית נוכל לרשום את הקשר באופן מטרציוני המתאר את ההתרה ההופכית הדו ממדית. כעת נשים לב שישנו הכפלה בפקטור. נציג את הקשר באופן הבא:

$$I = \frac{1}{MN} W_M \cdot F \cdot W_M$$

בקוד ה – MATLAB מימשנו את התוצאות הסופיות של הפיתוחים הנ"ל.

1.1.2 Shift zero frequency component to center

נבנה פונקציה אשר תמרכז את תדר האפס בשני הממדים במרכז התמונה, נקרא לפונקציה: 'dip_fftshift()' . לאחר הפעלת הפונקציה נקבל את האופרציה הבאה:



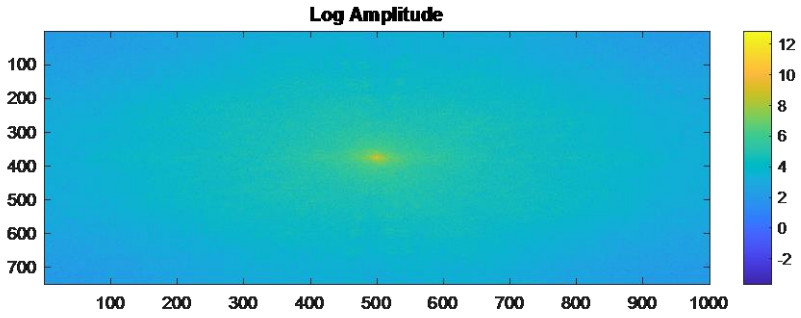
הפונקציה מוצאת את הנקודה מרכזית של ההתמרה של התמונה וביחס אליה מבצעת החלפה בין האזורים 1 ו – 3 ובין האזורים 2 ו – 4 של ההתמרה כמתואר באיור למעלה.

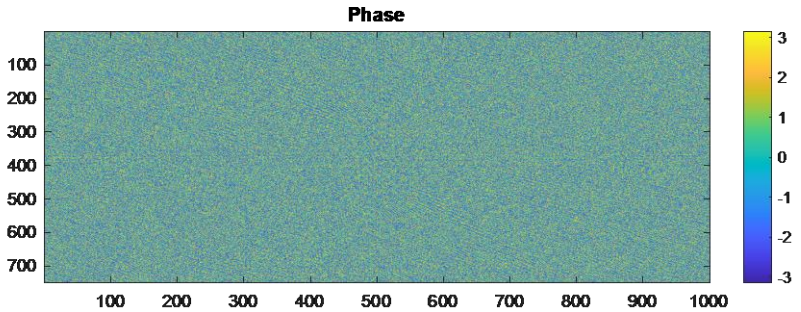
1.1.3Read the 'beatles' image

נקרא את התמונה 'beatles.png' ונמיר אותה ל – Grayscale וננרמל אותה. נקבל את התמונה הבאה:

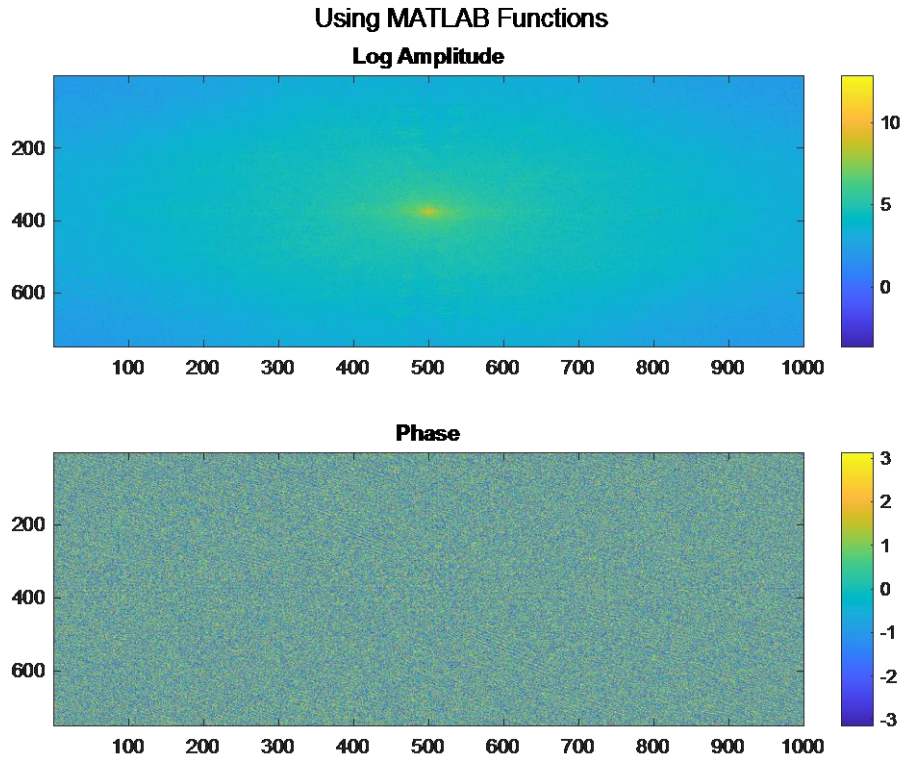
1.1.4 2D – FFT and shift

נבצע התמרה דו ממדית בעזרת הפונקציה שבנינו עבור התמונה שהמרנו ל Grayscale ונרמלנו בסעיף הקודם, ולאחר מכן נבצע FFT Shift בעזרת הפונקציה שבנינו. נציג את הלוגריתם של האמפליטודה והפאזה של ההתמרה הדו ממדית המתקבלת:





נשווה את התוצאות שקיבלנו לתוצאות המתקבלות עבור שימוש בפונקציות הבנויות של MATLAB (*'fft2()' ,'fftshift()'*). נציג את התמונות שמתקבלות:



נוכל לראות שהתוצאות המתקבלות מאוד דומות. בנוסף נוכל להסתכל על השוואה מתמטית בין התמונות שהתקבלו:

```
ssim_amp_FFT2D =

     1


mse_amp_FFT2D =

   3.3637e-21


ssim_phase_FFT2D =

    1.0000


mse_phase_FFT2D =

   5.2638e-05
```

נשים לב שקיבלנו במדד של ה – SSIM ערך של 1 עבור האמפליטודה ועבור הפאזה, כלומר הדמיון בין התמונות גבוה, בנוסף אם הסתכל על מדד ה – MSE שעבור האמפליטודה וגם עבור הפאזה נקבל ערך שקרוב מאוד ל – 0.

1.1.5 Reconstruct the original image

נשחזר את התמונה המקורית על ידי שימוש בפונקציה שבנינו המבצעת התמרת פורייה דו ממדית הופכית. נשתמש בחלק הממשי של התוצאה, ונשווה את התמונה המתקבלת לתמונה המקורית. בנוסף נשווה את התוצאה המתקבלת לתמונה המשוחזרת על ידי שימוש בפונקציה הבנויה של MATLAB:

Original image - gray scale

Reconstrunct image

Reconstrunct image - using MATLAB function

מהתוצאות ניתן לראות שהשחזור על ידי הפונקציה שבנינו דומה מאוד לתמונה המקורית, ובנוסף גם דומה מאוד לשחזור של התמונה על ידי שימוש בפונקציה המובנית של MATLAB. נציג את הדמיון בין השחזורים ובין השחזור שלנו לתמונה המקורית על ידי מדדים כמותיים:

```
ssim_recon =

    1.0000


mse_recon =

   5.7530e-27


ssim_recon_func =

    1.0000


mse_recon_func =

   5.7536e-27
```

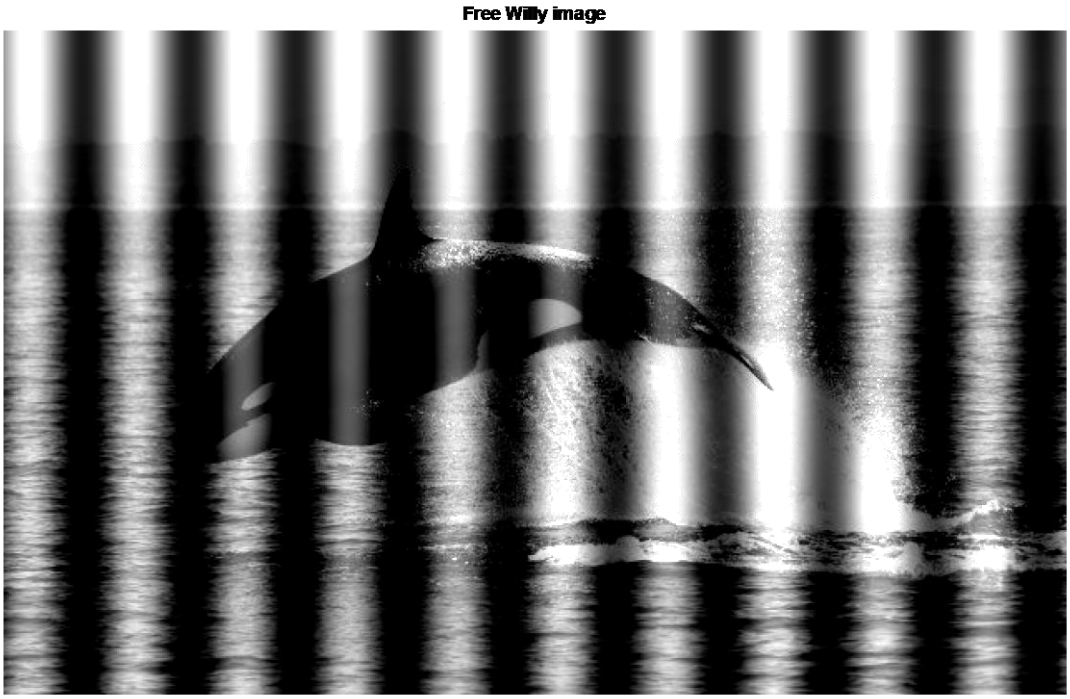
גם כאן נוכל לראות שמדד ה – SSIM הוא 1 בשני המקרים כלומר התמונה המשוחזרת והתמונה המקורית דומות מאוד, וכך גם השחזורים התקבלים על ידי הפונקציה שבנינו והפונקציה הבנויה של MATLB. במדד ה – MSE נוכל לראות שגיאה קטנה מאוד שקרובה ל – 0 בשני המקרים.

1.2 Transformation properties

1.2.1 Linearity (Free Willy)

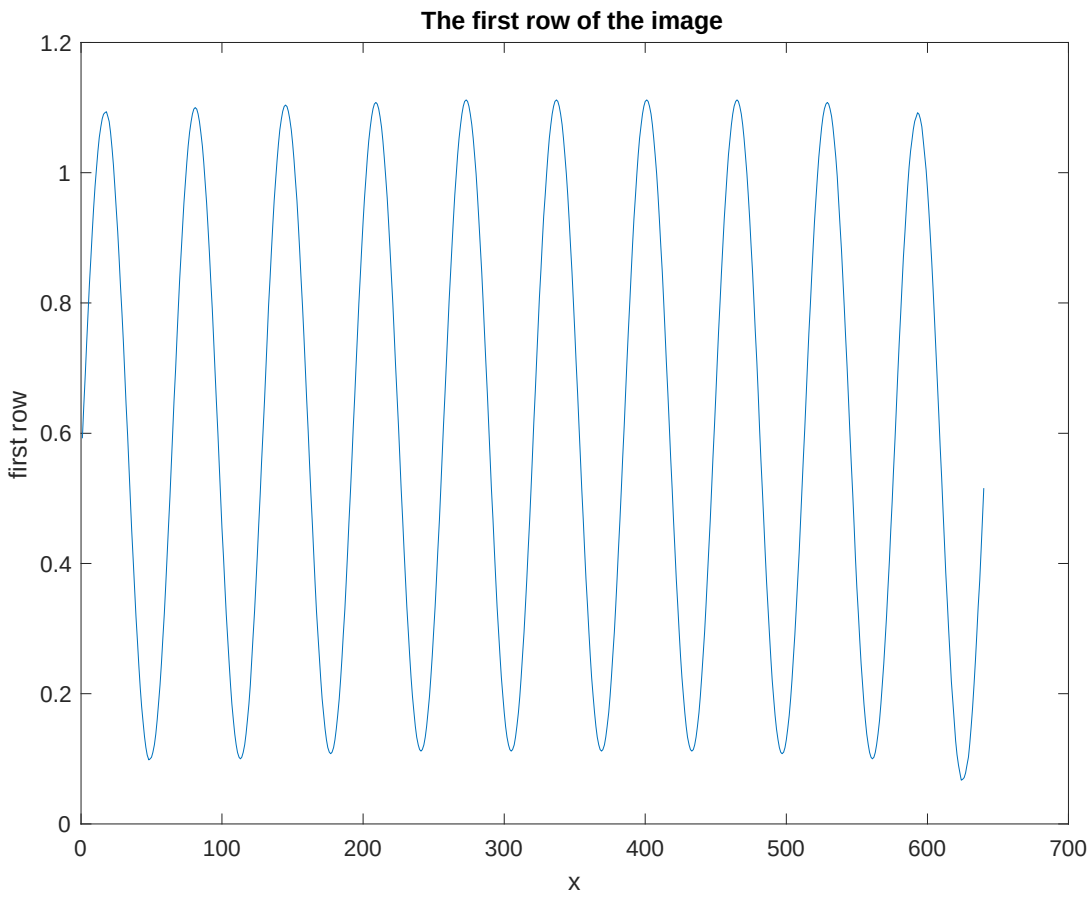
(a) Display image

נציג את התמונה *'free willy'* :



(b) Spatial frequency of prison bars - f_x

בסעיף זה התבקשנו למצוא את התדר f_x של האות הסינוסואידלי המתאר את הסוגרים של הכלא בו נמצא ווילי. ראשית, נשתמש ברמז ונציג את השורה הראשונה של התמונה:



נשים לב כי קיבלנו אות סינוסי עם הסטה. לכן בהתמרת פורייה של אות זה נקבל 3 דלתאות, כאשר אחד מהם יהיה בתדר אפס המתאר את ווילי ו-2 הדלתאות האחרים מתארים את התמרת פורייה של אות סינוסי. כידוע, התמרת פורייה של אות סינוסי הוא 2 דלתאות סימטריים מסביב לתדר אפס:

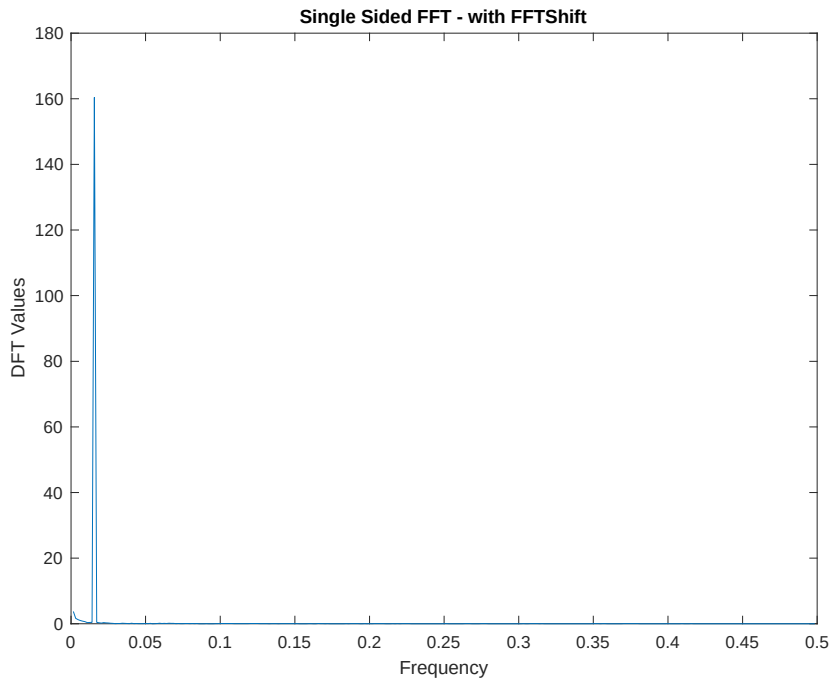
$$\mathcal{F}\left\{0.5\sin\left(\frac{2\pi}{N}f_x\cdot x\right)\right\}=\frac{0.25\cdot\pi}{j}\left[\delta\left(f-\frac{f_x}{N}\right)-\delta\left(f+\frac{f_x}{N}\right)\right]$$

לכן כדי למצוא את f_x מספיק למצוא את התדר של אחת הדלתאות f_0 של ההתמרה של הסינוס.

נשים לב כי ציר התדר מנורמל לפי מספר הפיקסלים שבשורה $N=640$ ולכן נקבל דלתא בתדר $f_0=\frac{f_x}{N}$.

נמצא את התמרת פורייה החד צדדית של האות, כך שבמישור התדר נשאר רק עם הדלתא של האות הסינוס בחלק החיובי של ציר התדר. מתוך התמרה זו נמצא את התדר של הדלתא f_0.

נציג את ההתמרה החד צדדית של השורה הראשונה:



מהתמרה זו נוכל למצוא את התדר f_0 שבו נמצא הערך הגדול ביותר של ההתמרה ולהציג את התדר שלו.

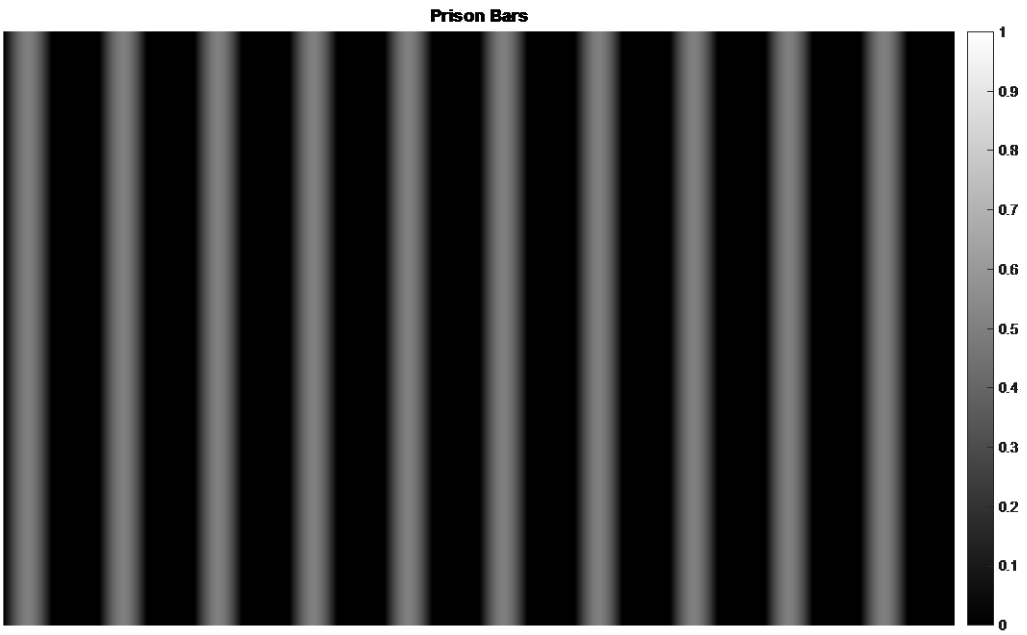
נקבל שהתדר בו נמצא הדלתא הוא : $f_0 = \frac{1}{64}$.

לכן נוכל למצוא את f_x :

$$f_x = f_0 \cdot N = \frac{1}{64} \cdot 640 = 10$$

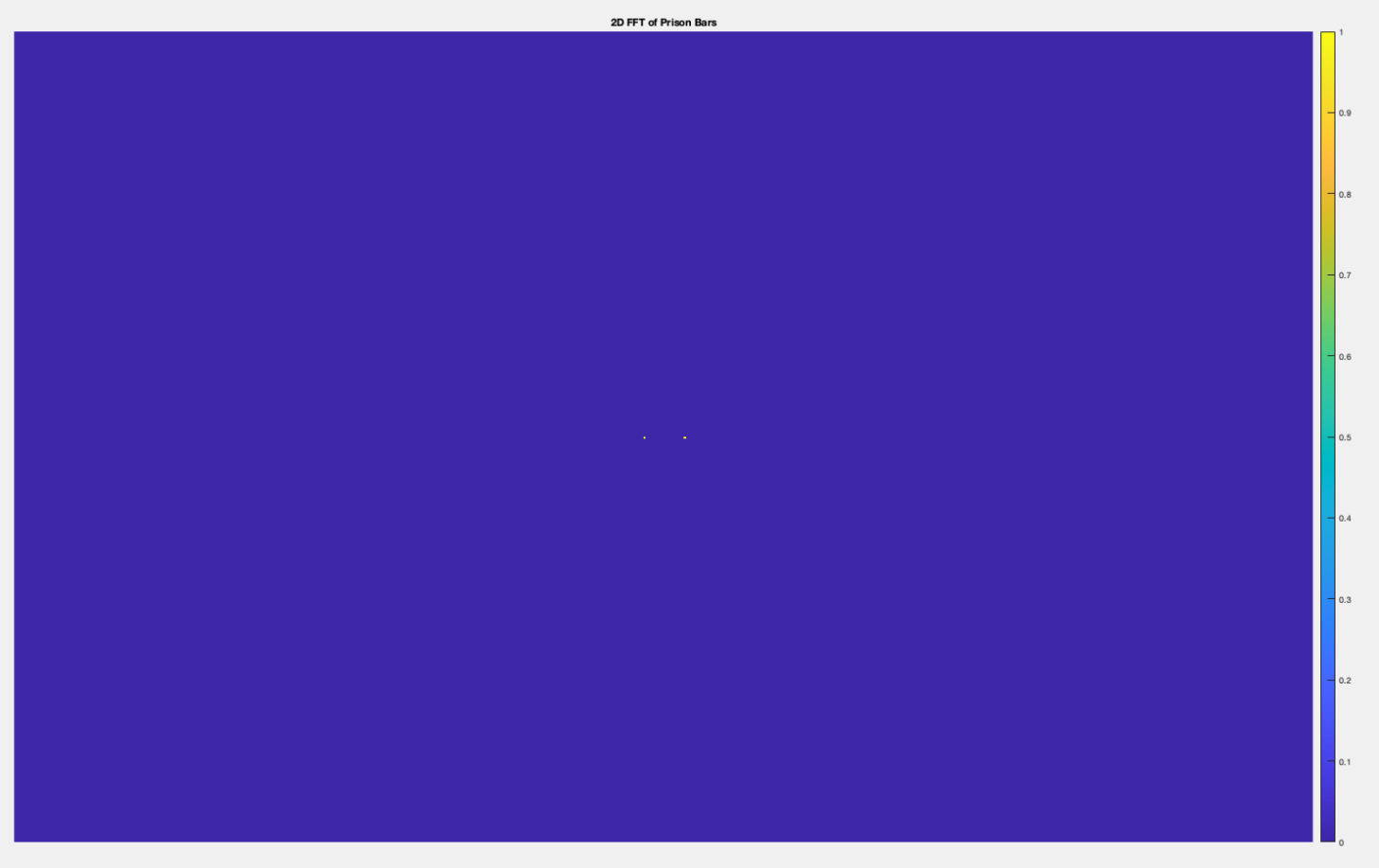
(c) Prison bars image

באמצעות התדר שחישבנו בסעיף הקודם נוכל להציג את הסורגים של הכלא שבו נמצא ווילי:



(d) 2D – FFT of the prison bars

נציג את האמפליטודה של ההתמרה הדו ממדית של תמונת הסורגים:



נשים לב כי קיבלנו 2 נקודות הממורכזות סביב המרכז לאורך הציר x , נקודות אלה מתארות את הדלתאות שקיבלנו בהתמרה החד ממדית לאורך ציר x .

נוכיח באופן מתמטי שקיבלנו את התוצאה שלה ציפינו:

$$F(u+1,v+1) = \sum_{m=0}^{M-1}\sum_{n=0}^{N-1} I(m+1,n+1)e^{-i2\pi\left(\frac{um}{M}+\frac{vn}{N}\right)}$$

$$= \sum_{m=0}^{M-1}\sum_{n=0}^{N-1} 0.5\sin\left(\frac{2\pi}{N}10\cdot(n+1)\right)e^{-i2\pi\left(\frac{um}{M}+\frac{vn}{N}\right)}$$

$$= \sum_{m=0}^{M-1} e^{-i2\pi\frac{um}{M}}\sum_{n=0}^{N-1} 0.5\sin\left(\frac{2\pi}{N}10\cdot(n+1)\right)e^{-i2\pi\frac{vn}{N}}$$

$$= \sum_{m=0}^{M-1} W_M^{um*}\sum_{n=0}^{N-1} 0.5\sin\left(\frac{2\pi}{N}10\cdot(n+1)\right)W_N^{vn*}$$

מתכונות ידועות של טווידל נקבל:

$$\sum_{m=0}^{M-1} W_M^{um^*} = \sum_{m=0}^{M-1} (W_M^{u^*})^m = \frac{W_M^{uM^*} - 1}{W_M^{u^*} - 1} \overset{mod(u,M)\neq 0}{\cong} 0$$

וכאשר $mod(u,M) = 0$ נקבל $\sum_{m=0}^{M-1} W_M^{um^*} = \sum_{m=0}^{M-1} 1 = M$

לכן סך הכל נקבל:

$$\sum_{m=0}^{M-1} W_M^{um^*} = M \cdot \delta(u) \quad , \quad 0 \leq u \leq M-1$$

מכאן נקבל כי רק בשורה הראשונה של התמונה יכולים להיות ערכים שונים מאפס.

לאחר שנבצע fft_shift נקבל כי רק בשורה האמצעית יכולים להיות ערכים שונים מאפס.

כעת נותר לבצע את ההתמרה החד ממדית על העמודות של התמונה, וכמו שהסברנו ההתמרה החד ממדית של אות סינוסי הוא 2 דלתאות מוזזות מסביב למרכז ולכן הגיוני שקיבלנו 2 נקודות סביב המרכז.

(e) Free Willy

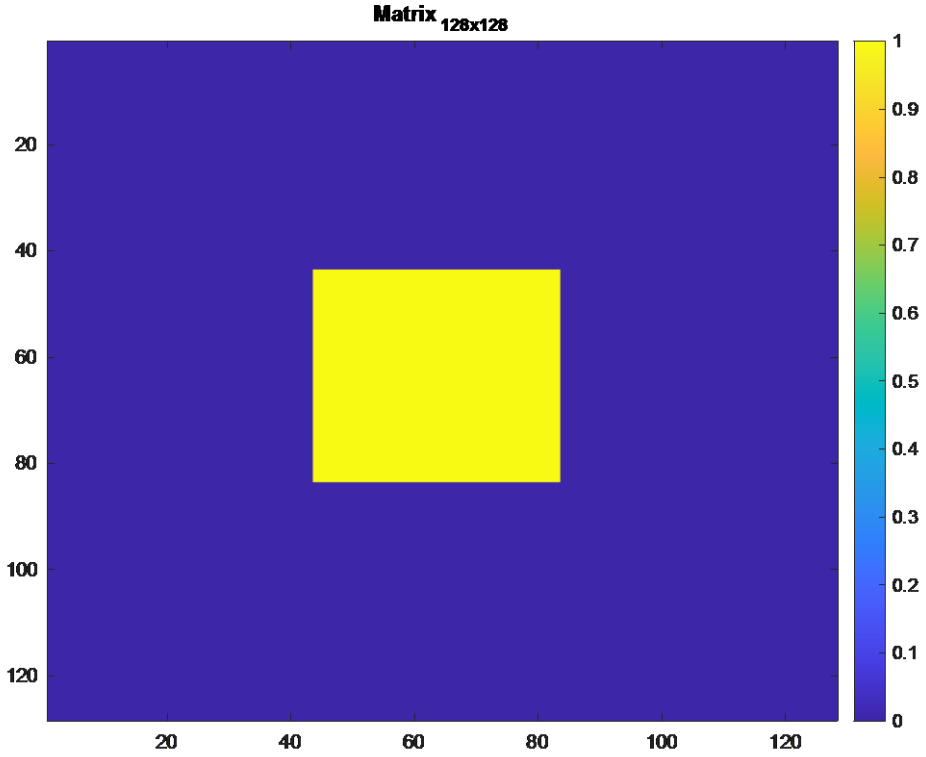
כדי לשחרר את וילי מהכלא, לאחר שמצאנו את התדר של אות הסינוס המתאר את הסורגים, נוכל לעבור למרחב התדר, לבצע התמרה עבור התמונה המקורית של וילי בכלא והתמרה עבור התמונה של הסורגים בלבד. כאשר נחסר בין התמרות אלו, נקבל את ההתמרה של וילי בלבד ואז נוכל לחזור למרחב המקורי ולקבל את וילי משוחרר כך:

כתבנו פונקציה שנקראת Free_Willy(Willy) שמקבלת תמונה של וילי בכלא, מחשבת את התדר של אות הסינוס המתאר את הסורגים של הכלא ואז מבצעת התמרה לתמונה של הכלא בלבד והתמרה לתמונה של וילי בכלא. נחסיר בין ההתמרות ונקבל התמרה של וילי. נבצע התמרה הפוכה ונקבל את וילי משוחרר

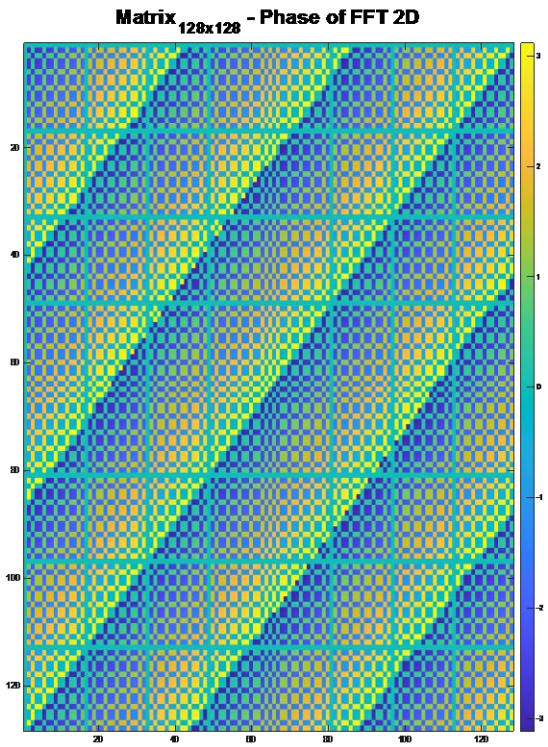
1.2.2 Scaling, translation, and separability

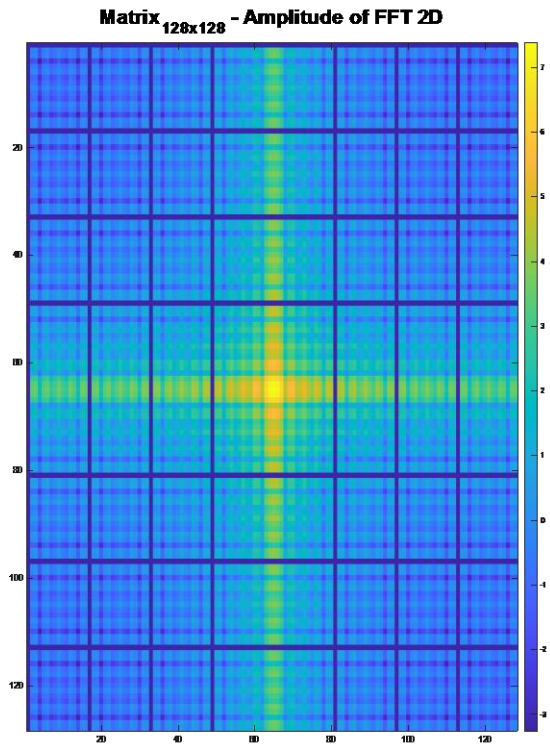
(a) Box $40X40$ in the center

נציג את ההתמרה הדו ממדית של מטריצת האפסים בגודל $128X128$ כאשר יש במרכזה מטריצה של אחדות בגודל $40X40$. נציג את המטריצה:



כעת נציג את ההתמרה הדו ממדית של המטריצה:





ניתן הסבר מתמטי לתוצאות:

נשים לב שנוכל לייצג את התמונה על ידי הווקטור v אשר נגדיר אותו להיות כך:

$$v = \left[0\,,0\,,0\,,\dots,\underbrace{1\,,1\,,1\,,\dots,1}_{(44:83)},0\,,0\,,\dots,0\right] \quad , \quad v \in \mathbb{R}^{128}$$

נסמן את המטריצה כך: $M_{128\times128}$. לכן נוכל לרשום כך:

$$M_{128\times128} = v^T \cdot v$$

כעת נפתח את ההתמרה הדו ממדית של המטריצה בצורה מתמטית:

$$F(u+1,v+1) = \sum_{m=0}^{M-1}\sum_{n=0}^{N-1} I(m+1,n+1)e^{-i2\pi\left(\frac{um}{M}+\frac{vn}{N}\right)}$$

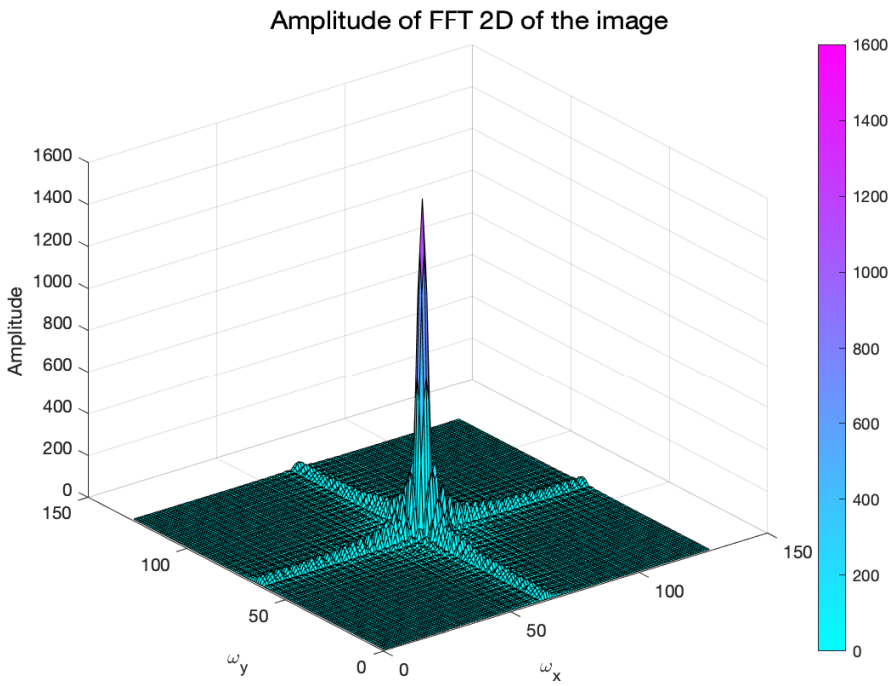
כעת נציב את תמונת המטריצה:

$$F(u+1,v+1) = \sum_{m=0}^{M-1}\sum_{n=0}^{N-1} M_{128\times128}(m+1,n+1)e^{-i2\pi\left(\frac{um}{M}+\frac{vn}{N}\right)} = [N,M=128]$$

$$= \sum_{m=0}^{127}\sum_{n=0}^{127} v(m+1)\cdot v(n+1)\cdot e^{-i2\pi\left(\frac{um}{M}+\frac{vn}{N}\right)} =$$

$$= \sum_{m=0}^{127} v(m+1)\cdot e^{-i2\pi\frac{um}{M}} \cdot \sum_{n=0}^{127} v(n+1)\cdot e^{-i2\pi\frac{un}{N}}$$

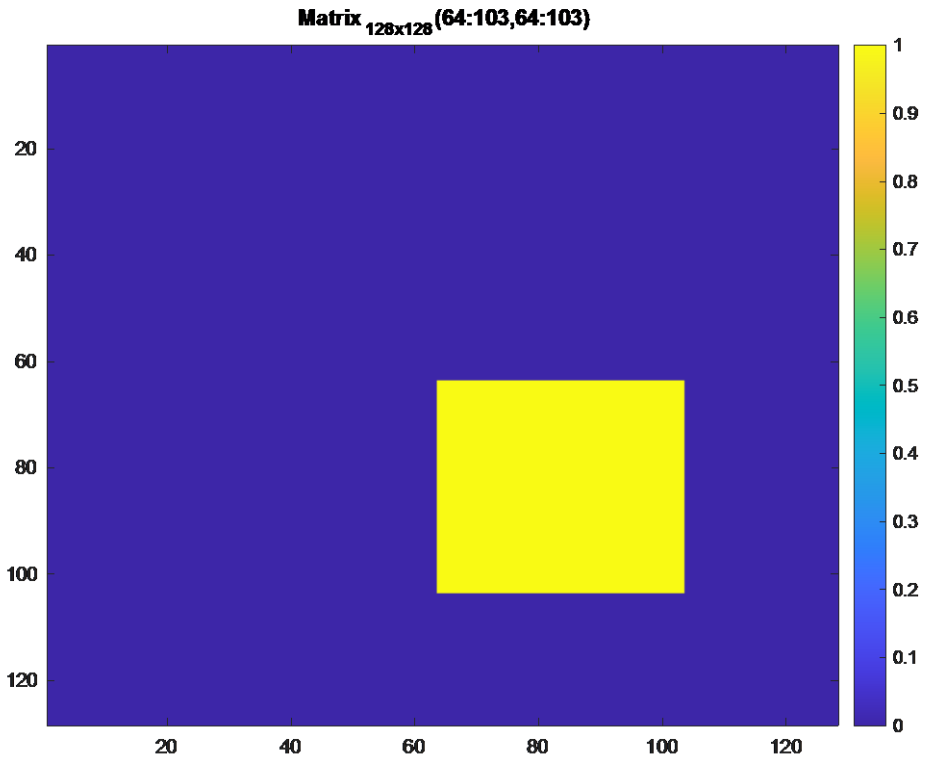
קיבלנו 2 התמרות DFT חד ממדיות של חלונות מוזזים. נזכור שהתמרת ה – DTFT של חלון הינה $sinc(\cdot)$, כלומר נקבל הכפלה בתדר בין $sinc(\cdot)$ אופקי לבין $sinc(\cdot)$ אנכי. נוכל לראות זאת בברור בתמונה הבאה:



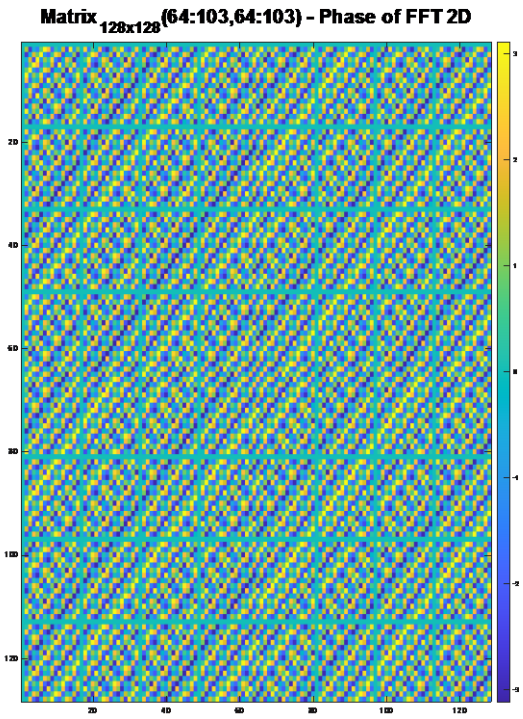
נוכל לראות שההסבר המתמטי תואם את התוצאות שקיבלנו.

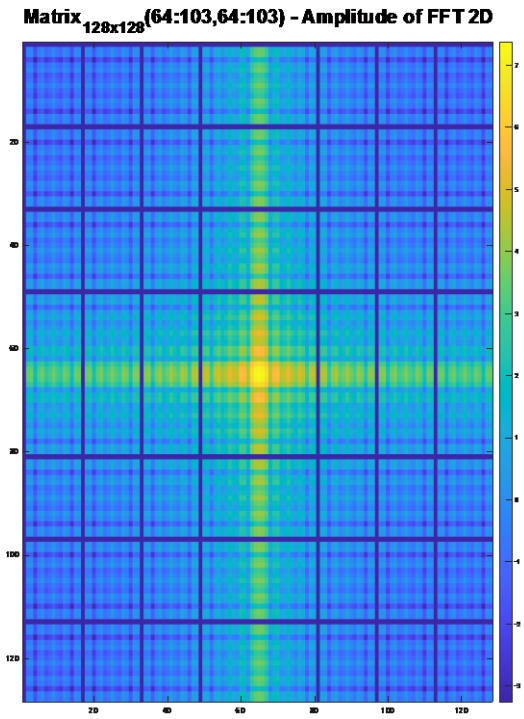
(b) Box $40X40$ in location $M_{128\times128}(64:103,64:103)$

נציג את המטריצה בדומה לסעיף הקודם רק שכעת המיקום של מטריצת ה - $40X40$ היא במיקום הפיקסלים $M_{128\times128}(64:103,64:103)$. נציג את המטריצה:



כעת נציג את ההתמרה הדו ממדית של המטריצה הנ"ל:



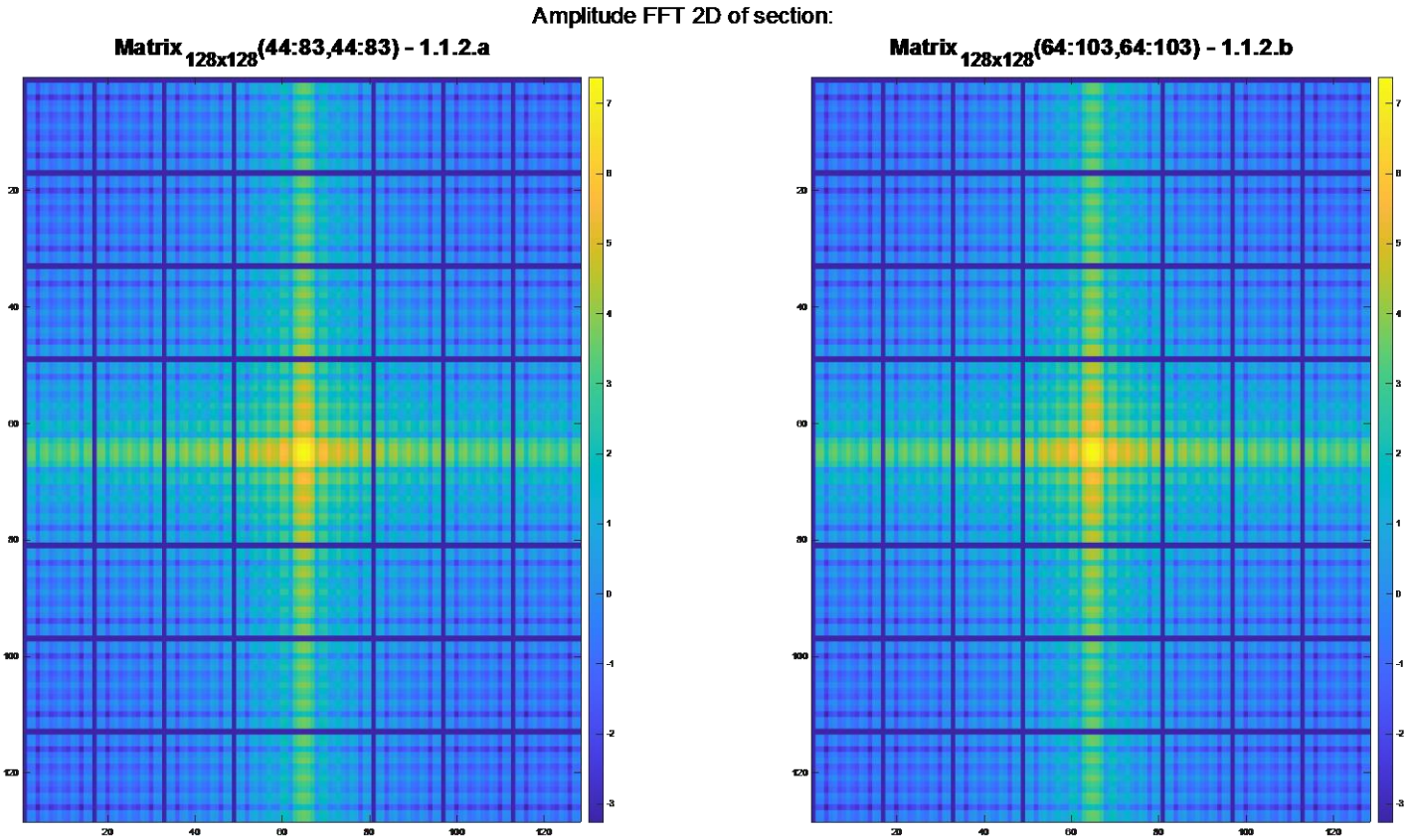


נשים לב שכעת רק הפאזה השתנתה והאמפליטודה נשארה זהה. בהמשך להסבר מהסעיף הקודם נשים לב שכעת התמונה שקיבלנו היא הזזה של מטריצת האחדות בציר ה $X-$ וציר ה $Y-$ של התמונה הקודמת.

מכיוון שהזזה במרחב תתבטא בהכפלה של אקספוננט מרוכב בצורה הבאה: $e^{-i\omega k}$, כאשר k הוא מספר ההזזה במרחב. מכיוון שבסעיף זה ישנה הזזה זהה בציר האופקי וגם בציר האנכי נקבל שסך כל ההזזה במישור התדר תבוטא בהכפלה באקספוננט כך: $e^{-i\frac{2\pi}{128}kv} \cdot e^{-i\frac{2\pi}{128}ku}$. לכן הפאזה החדשה שנקבל היא לאחר הכפלת האקספוננט הקודם באקספוננט ההזה בתדר. לכן נקבל שינוי בפאזה.

לעומת זאת נזכור שמתקיים: $\left|e^{-i\omega}\right| = 1$, כלומר נצפה שהאמפליטודות של ההתמרות יהיו זהות.

נבצע השוואה של האמפליטודה מהסעיף הקודם לאמפליטודה מהסעיף הנ"ל. נציג את התוצאות:



מהתוצאות נוכל לראות שאכן כמו שהסברנו האמפליטודות של שני המקרים זהות.

נראה שהסברים אלו ותוצאות אלו מתקבלות גם על ידי השוואה כמותית:

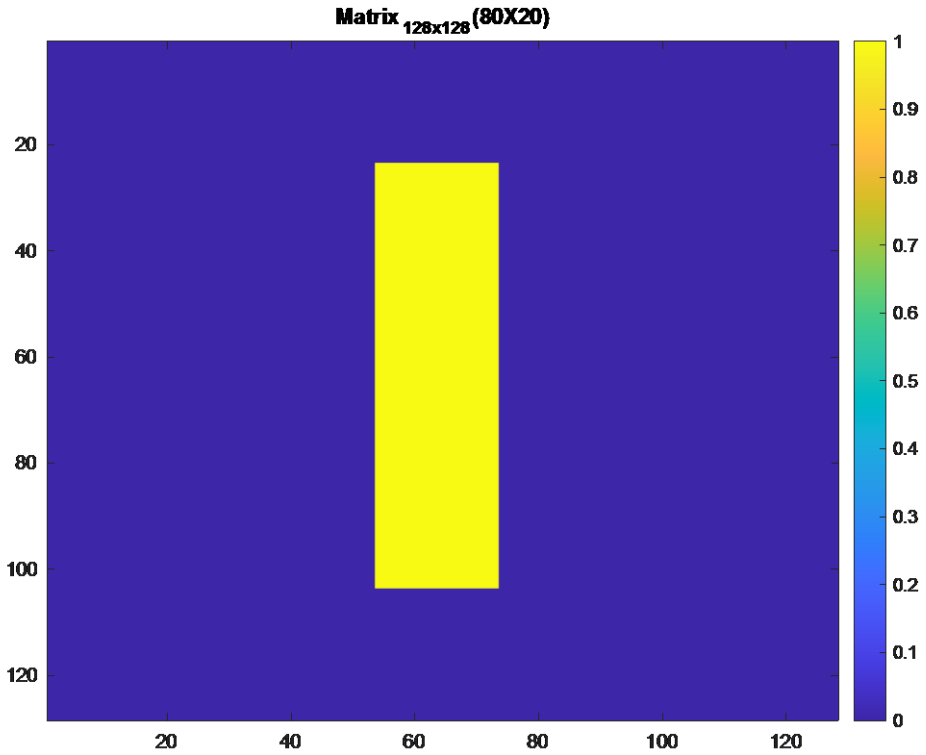
```
ssim_amp =

     1

mse_amp =

   8.2294e-29
```

```
ssim_phase =

    0.0203

mse_phase =

    6.7007
```

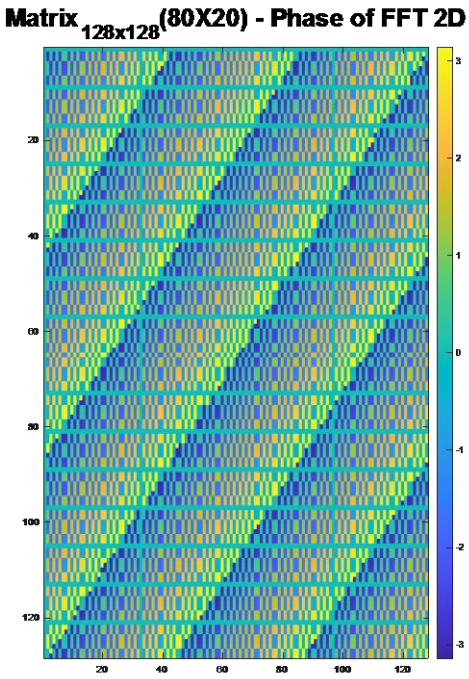
מההשוואה הכמותית נוכל לראות שאכן ההבדל בין התוצאות של הפאזה גדול והשגיאה הריבועית הממוצעת גדולה לעומת האמפליטודה שישנו דמיון רב בין התוצאות והשגיאה קטנה מאוד וקרובה ל – 0.

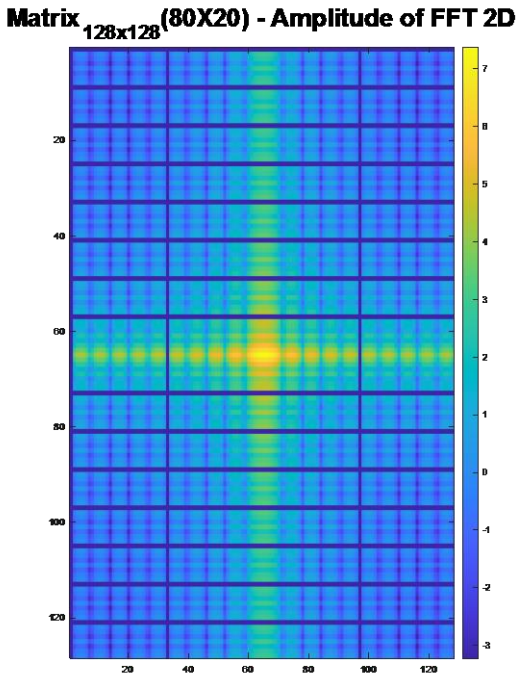
(c) Box 80X20 in the center

נציג את המטריצה בדומה לסעיף (a) רק שכעת במרכזה המטריצה יהיה מלבן של אחדות בגודל של $80X20$ במקום ריבוע. נציג את המטריצה:



כעת נציג את ההתמרה הדו ממדית של המטריצה הנ"ל:





נציג קשר מתמטי בין התמונה שבסעיף (a) לתמונה מהסעיף הזה (c).

נסמן את התמונה מסעיף (a) כך: $M_{(a)}$, ואת התמונה מהסעיף הזה כך: $M_{(c)}$.

נוכל להציג את התמונה מסעיף (a) על ידי 2 משתנים כך:

$$M_{(a)}(x,y) = rect\left(\frac{x}{a}\right) \cdot rect\left(\frac{y}{a}\right)$$

לכן נקבל שההתמרה היא:

$$F_a(f_x, f_y) = a \cdot sinc\left(\frac{f_x}{a}\right) \cdot a \cdot sinc\left(\frac{f_y}{a}\right)$$

נציג את התמונה מהסעיף הנ"ל כך:

$$M_{(c)}(x,y) = rect\left(\frac{x}{a/2}\right) \cdot rect\left(\frac{y}{2a}\right) = rect\left(\frac{2x}{a}\right) \cdot rect\left(\frac{y/2}{a}\right)$$

לכן נקבל שההתמרה היא:

$$F_c(f_x, f_y) = \frac{a}{2} \cdot sinc\left(\frac{\frac{f_x}{2}}{a}\right) \cdot 2a \cdot sinc\left(\frac{2f_y}{a}\right)$$

נשים לב שבתמונה של האמפליטודה בציר ה $X-$ נקבל שיש מעט התאפסויות , כלומר "ציר איטי" ואילו בציר ה $Y-$ יש הרבה התאפסויות, כלומר "ציר מהיר".

נציג את ההשוואה הכמותית בין התמונות וההתמרות הדו ממדיות שלהן:

```
ssim_images =

    0.8082
```

```
ssim_FFT2D_images =

    0.0076
```

```
mse_images =

    0.0977
```

```
mse_FFT2D_images =

    1600
```

נשים לב שהתמונות משני הסעיפים יחסית דומות אחת לשנייה, אך לעומת זאת נוכל לראות שישנו הבדל גדול בין ההתמרות הדו ממדיות של התמונות.

(d) Represent the image using 1D vector

נוכל להציג את התמונה על ידי 2 וקטורים חד ממדיים בצורה הבאה:

$$v_1 = \left[0\,,0\,,0\,,\dots,\underbrace{1\,,1\,,1\,,\dots,1}_{(24:103)},0\,,0\,,\dots,0\right] \quad , \quad v \in \mathbb{R}^{128}$$

$$v_2 = \left[0\,,0\,,0\,,\dots,\underbrace{1\,,1\,,1\,,\dots,1}_{(54:73)},0\,,0\,,\dots,0\right] \quad , \quad v \in \mathbb{R}^{128}$$

כעת נוכל להציג את התמונה באופן הבא:

$$M_{(80X20)} = {v_1}^T \cdot v_2$$

(e) Compute 2D – FFT by using 1D – FFT

נציג את החישוב של ההתמרה הדו ממדית על ידי שימוש בהתמרה החד ממדית.

ראשית נרשום את ההתמרה הדו ממדית של התמונה:

$$F(u+1,v+1) = \sum_{m=0}^{M-1}\sum_{n=0}^{N-1} I(m+1,n+1)e^{-i2\pi\left(\frac{um}{M}+\frac{vn}{N}\right)}$$

כעת נציב את התמונה כאשר הצגנו אותה על ידי 2 וקטורים חד ממדיים כך:

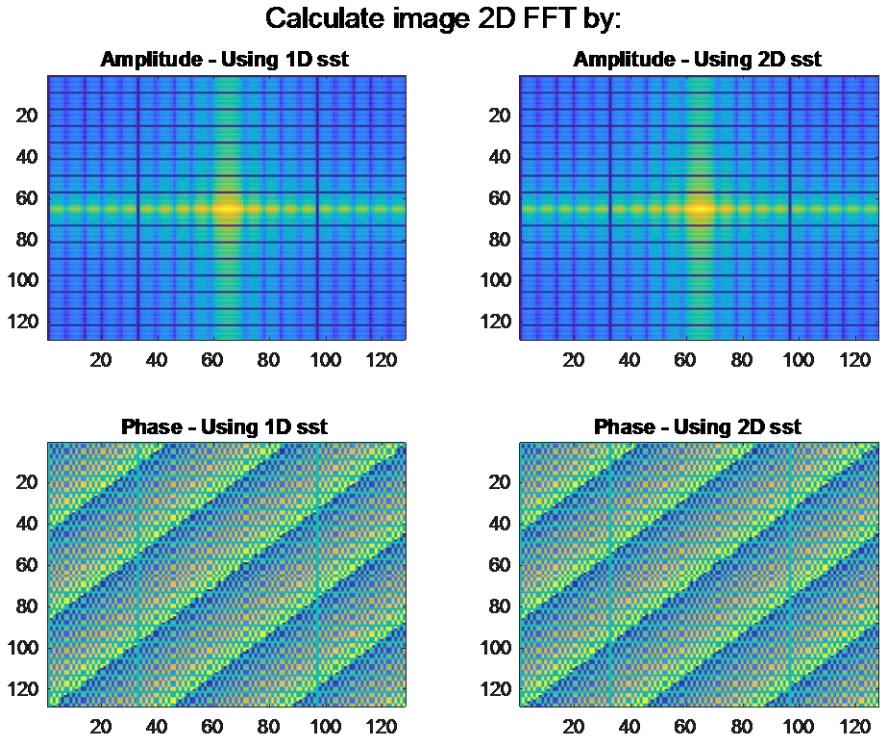
$$\sum_{m=0}^{M-1}\sum_{n=0}^{N-1} v_1(m+1)\cdot v_2(n+1)\cdot e^{-i2\pi\left(\frac{um}{M}+\frac{vn}{N}\right)} =$$

$$= \sum_{m=0}^{M-1} v_1(m+1)e^{-i2\pi\frac{um}{M}} \cdot \sum_{m=0}^{M-1} v_1(m+1)e^{-i2\pi\frac{vn}{N}}$$

נשים לב שקיבלנו הכפלה בין שתי התמרות חד ממדיות, כלומר:

$$F(u+1,v+1) = [FFT(v_1)]^T \cdot FFT(v_2)$$

נציג את ההתמרות הדו ממדיות המתקבלות מחישוב על ידי שני האופנים ונשווה ביניהם בצורה כמותית:



במבט על התוצאות נוכל לראות שהתוצאות שקיבלנו כשחישבנו את ההתמרה הדו ממדית ישירות או דרך התמרה חד ממדית זהות, נראה זו גם בצורה כמותית:

```
ssim_FFT2D_1D =

1


mse_FFT2D_1D =

7.2057e-29
```

גם כאן נוכל לראות ממדד ה – SSIM שערכו 1 כלומר ישנו דמיון רב בין התמונות, ובנוסף נוכל לראות ששגיאת ה – MSE קטנה מאוד ומאוד קרובה ל – 0.

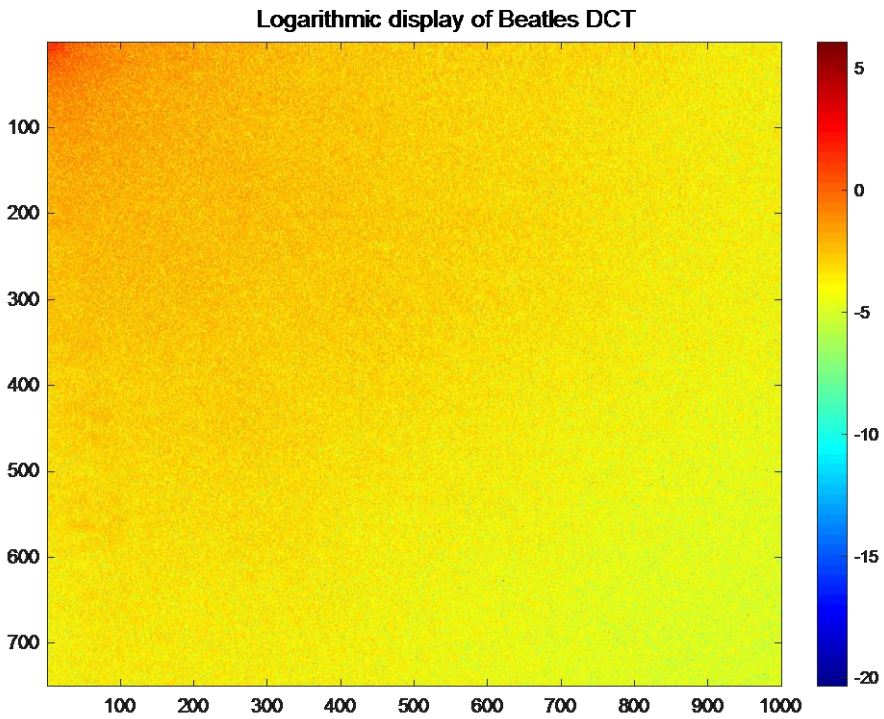
2. Discrete Cosine Transform (DCT)

2.1 Read the 'beatles' image

נקרא את התמונה 'beatles.png', נמיר אותה ל – Grayscale וננרמל אותה. נקבל את התמונה שהצגנו בסעיף (1.1.3).

2.2 Display the DCT of the image

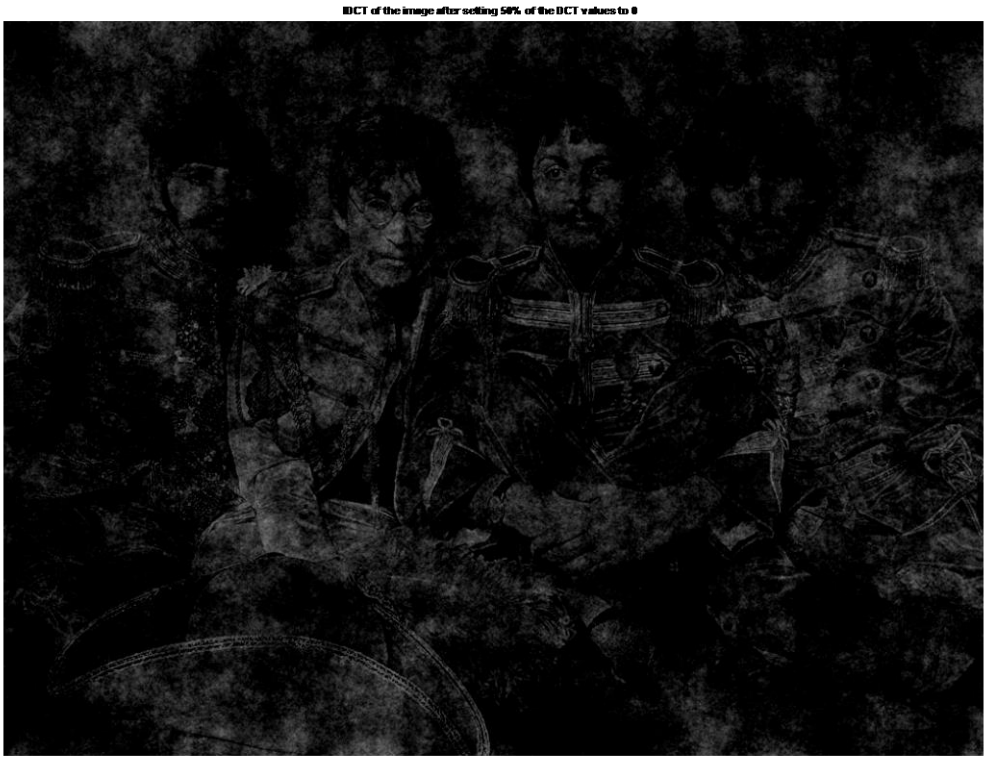
נפעיל את הפונקציה 'dct2(Img)' על התמונה מהסעיף הקודם. נציג את התוצאות המתקבלות על ידי שימוש ב: *imagesc(log(dct2(DCT))); , colormap(jet(64)); , colorbar;*



כפי שציפינו ניתן לראות משקלים גדולים יותר בפינה השמאלית העליונה , מכיוון שעיקר המשקל נמצא בהרמוניות הנמוכות המיוצגות בצד השמאלי העליון של התמונה.

2.3 Randomly set to zero 50% of the DCT values

כעת נאפס בצורה רנדומלית 50 אחוז מהערכים של ה – DCT של התמונה, ולאחר מכן נבצע שחזור ונקבל:



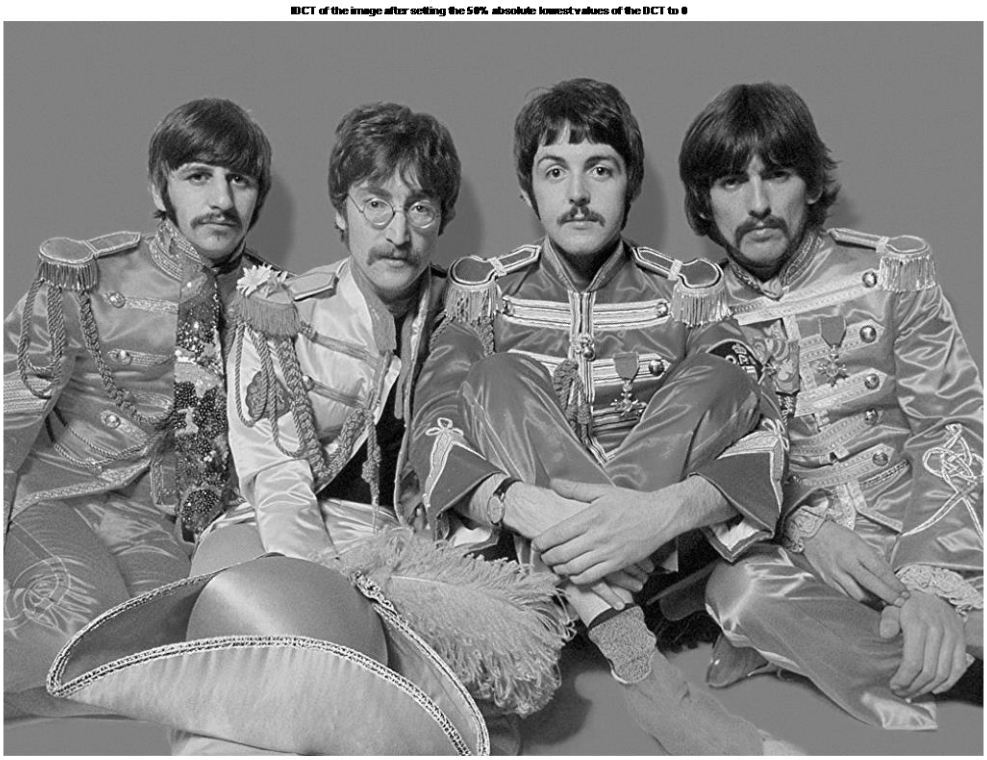
למרות שבשחזור נוכל לראות (בקושי רב אבל עדיין) 4 אנשים השחזור של התמונה לא מוצלח מכיוון שכאשר אנחנו מאפסים 50 אחוז מהערכים בצורה רנדומלית אנחנו מאבדים מידע חשוב. נוכל לראות שהשחזור לא מוצלח גם על ידי מדדים כמותיים:

```
ssim_rand =

-0.0285


mse_rand =

0.2714
```

גם כאן נוכל לראות שמדד ה – SSIM קרוב מאוד ל – 0 , כלומר אין כלל דמיון בין התמונה המשוחזרת לתמונה המקורית, ובנוסף נוכל לראות שגיאה יחסית גדול במדד ה – MSE.

2.4 Set to 0 the lowest 50% absolute values of the DCT

כעת נאפס את הערכים הנמוכים ביותר בערכם המוחלט של ה – DCT של התמונה, ולאחר מכן נשחזר את התמונה ונקבל:



בשחזור של התמונה במקרה הזה נוכל לראות שיפור משמעותי לעומת השחזור בסעיף הקודם. מכיוון שכעת אנחנו מאפסים את 50 אחוז הערכים הנמוכים היותר בערך המוחלט אנחנו בעצם שומרים את הערכים הגבוהים המיוחסים לתדרים הגבוהים וכך שמרנו את רוב המידע החשוב ביותר של התמונה. לכן כאשר נבצע שחזור נוכל לראות שחזור של התמונה בצורה טובה. נראה את איכות השחזור מול התמונה המקורית בצורה כמותית:

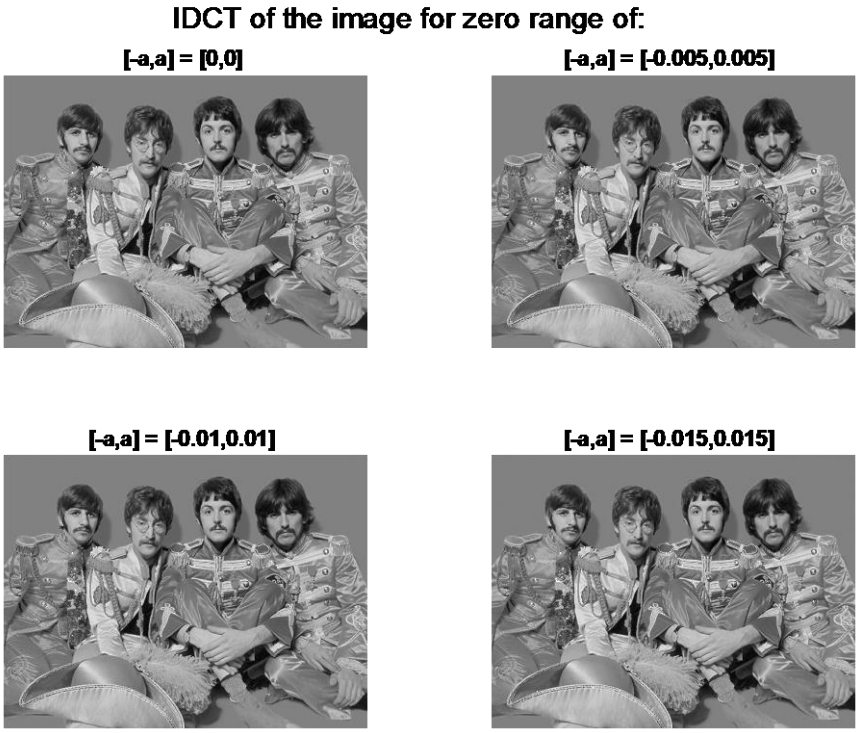
```
ssim_lowest =

    0.9462


mse_lowest =

    1.8348e-04
```

נוכל לראות שגיאה קטנה ב MSE וקרבה לערך 1 במדד ה – SSIM, כלומר הדבר מעיד על דמיון רב בין השחזור לבין התמונה המקורית.

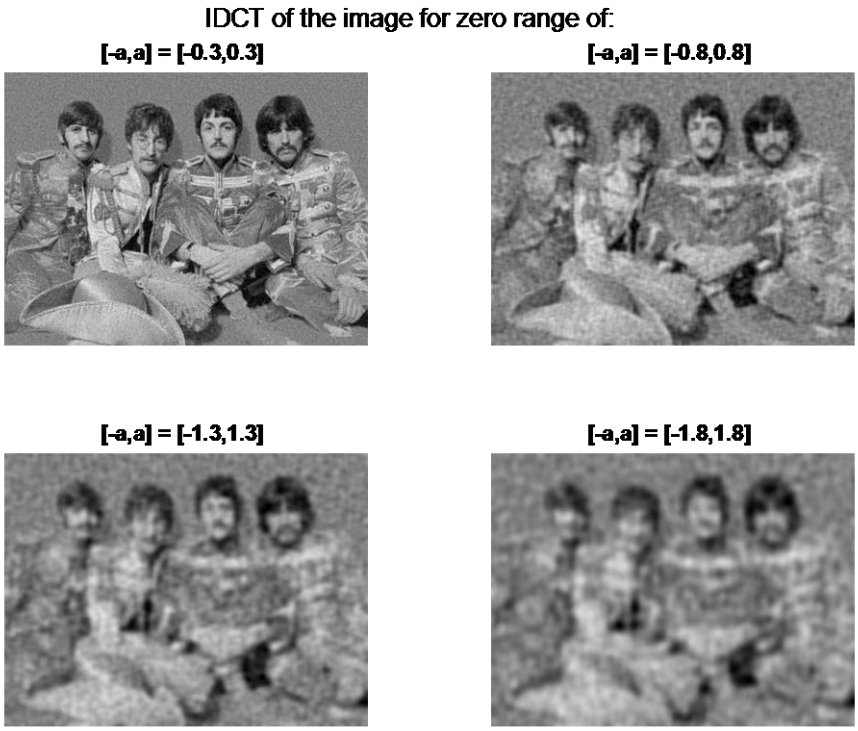
2.5 Set to 0 the values in the range of $(-a, a)$

כעת נאפס את הערכים של ה – DCT אשר נמצאים בטווח הערכים של $(-a, a)$ ולאחר מכן נבצע שחזור. נראה כמה קבוצות של ערכים, ונסביר בהמשך את ההבדלים.

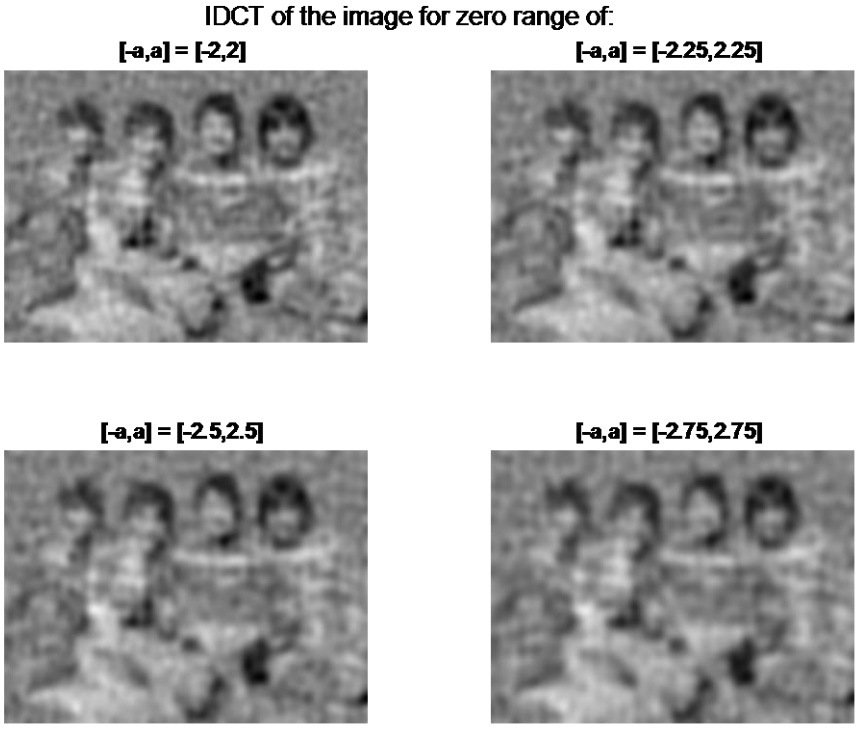
- עבור $a = [0 - 0.015]$ נקבל:



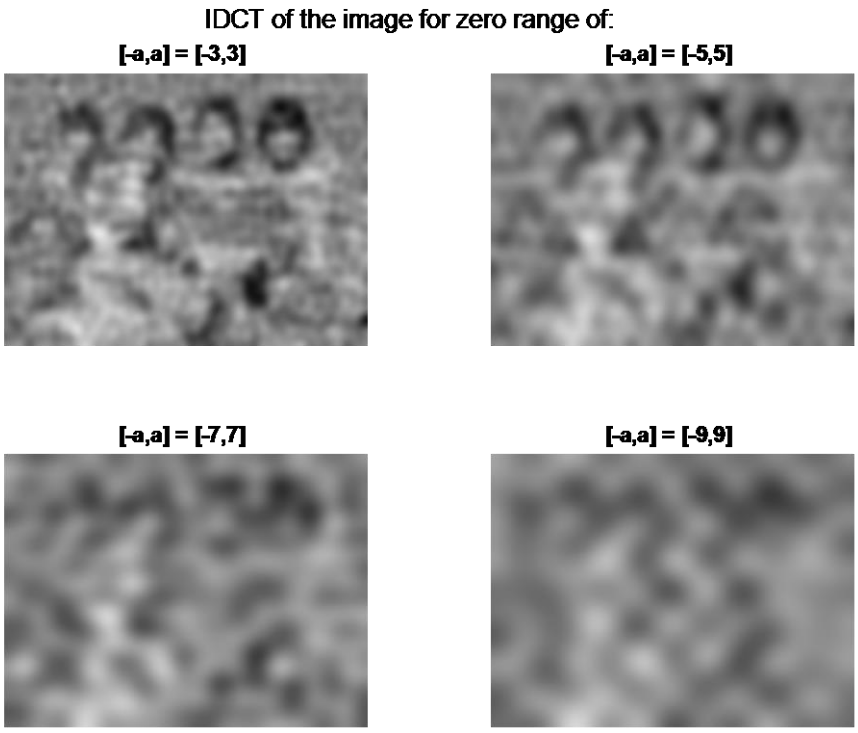
- עבור $a = [0.3 - 1.8]$ נקבל:



- עבור $a = [2 - 2.75]$ נקבל:



- עבור $a = [3 - 9]$ נקבל:



מהתוצאות נוכל לראות שכאשר $(-a, a) = (-0.015\,, 0.015)$ הוא הערך המקסימלי עבורו נוכל לשחזר את התמונה בצורה איכותית. במקרה זה אחוז הערכים של ה – DCT אותם איפסנו הוא 24.71%

הסבר לתוצאות:

- ראשית נוכל לראות שעבור ערך של $(-a, a) = (-0.015\,, 0.015)$ הוא הערך המקסימלי עבורו נוכל לשחזר את התמונה בצורה טובה ללא איבוד של מידע כמעט כלל.
- נשים לב שככל שנגדיל את הערך של a השחזור פחות טוב מכיוון שככל ש a גדול יותר אנחנו מאבדים יותר מידע.
- עבור ערכים נמוכים של a כלומר בטווח ערכים של $[0 - 0.015]$ נוכל לראות את שיחזור התמונה בבירור כמעט ללא איבוד מידע. עבור הטווח של $[0.3 - 1.8]$, כבר נוכל לראות שהשחזור של התמונה עדיין נראה טוב אך כעת נוסף לתמונה קצת רעש.
- עבור ערכים גבוהים של a , נוכל לראות שהשחזור לא באיכות טובה, כלומר בטווח הערכים של $[2 - 2.75]$, עדיין נוכל לראות את מבנה התמונה (4 אנשים שיושבים) אך התמונה המשוחזרת מטושטשת נורא ובאיכות גרועה. עבור ערכים ממש גדולים כמו טווח הערכים של $[3 - 9]$ ממש נראה שהשחזור של התמונה הופך ממטושטש לתמונה שהשחזור שלה כלל לא מובן ולא ברור כך שלא נדע מה הייתה התמונה המקורית כלל.

2.6 Explanation of the obtained results from sections 3 – 5

כפי שהצגנו בסעיפים הקודמים, בהתמרת DCT שמירת הערכים הגבוהים חיונית על מנת לקבל שחזור טוב של התמונה המקורית.

לכן התוצאות שקיבלנו תואמות את ההיגיון שעבור ערכים גבוהים של a קיבלנו שחזור פחות מוצלח, השחזור היה מטושטש ולא ברור ואילו עבור ערך a יחסית נמוך קיבלנו שחזור מוצלח על ידי ההתמרה.

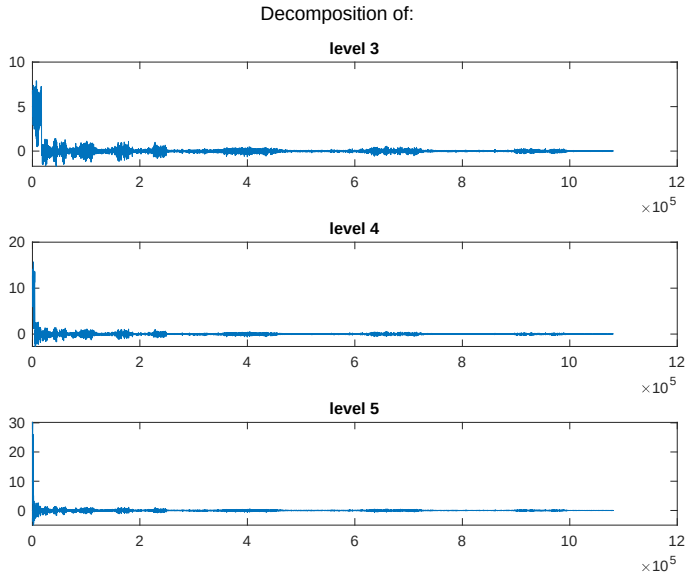
3. Wavelet Transform

3.1 Read the 'beetle.jpg' image

נקרא את התמונה 'beatles.png', נמיר אותה ל – Grayscale וננרמל אותה. נקבל את התמונה הבאה:

3.2 Extract the wavelet decomposition, level between 3 – 5

נשתמש ב wavelet *'haar'*. בשיטה זו התמונה מיוצגת באמצעות קומבינציה של מטריצות Haar, כך שניתן לשמור רק את המקדמים של בסיס מטריצות אלה, כלומר פעולה ששומרת פחות מידע של התמונה וחוסכת מקום בזיכרון במחשב. כל פעם שנבצע את הפעולה הזו נחסוך עוד מקום אחסון אך נאבד יותר מידע. בתמונה הבאה נוכל לראות את רמות הפירוק $3-5$, כאשר ניתן לראות אזורים רבים בהם המידע הוא אפסי ולכן נוכל לדחוס את המידע. ככל שרמת הפירוק גדלה כך נוכל לדחוס יותר מידע.



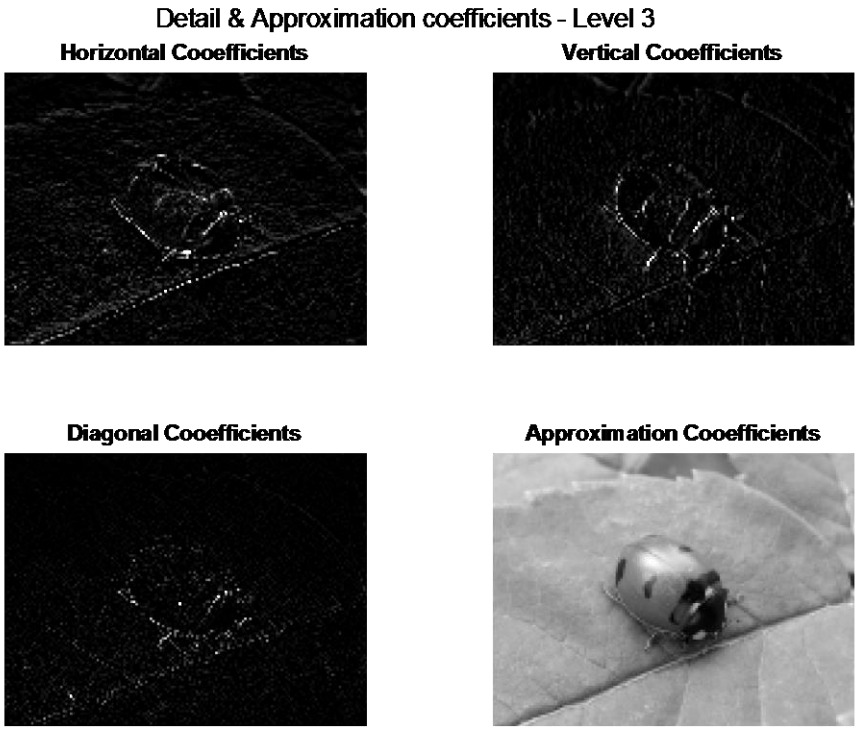
3.3 Find the detail and the approximation coefficients for each level

נשתמש בפונקציות של MATLAB : *'detcoef2()'* ו – *'appcoef()'* על מנת למצוא את המקדמים המשוערכים של כל רמה. בסעיף הבא נציג את התוצאות.

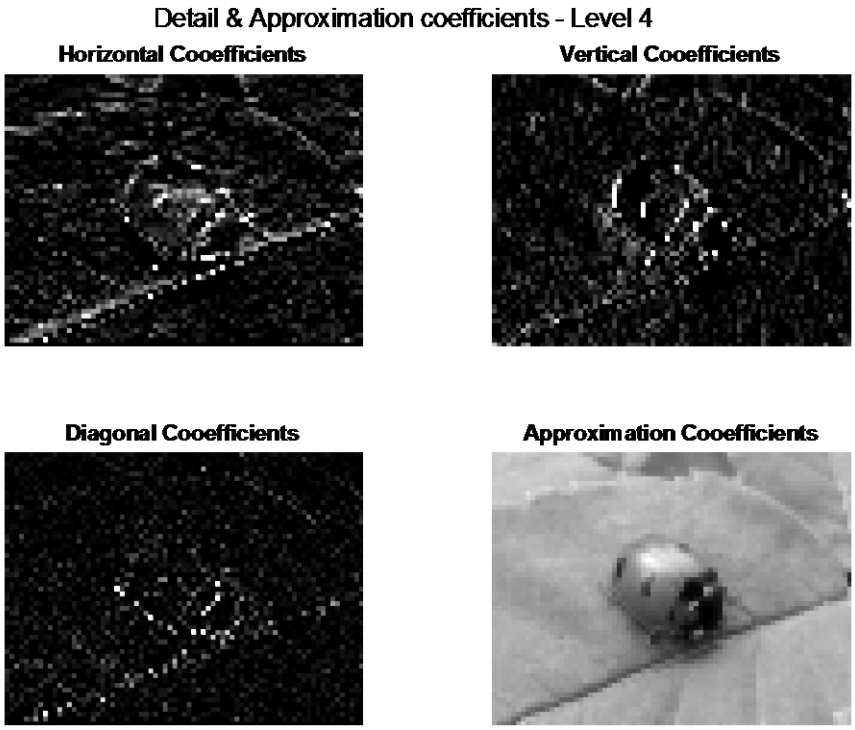
3.4 Display the results

נציג את המקדמים והתמונה המשוחזרת עבור כל אחד מהשלבים:

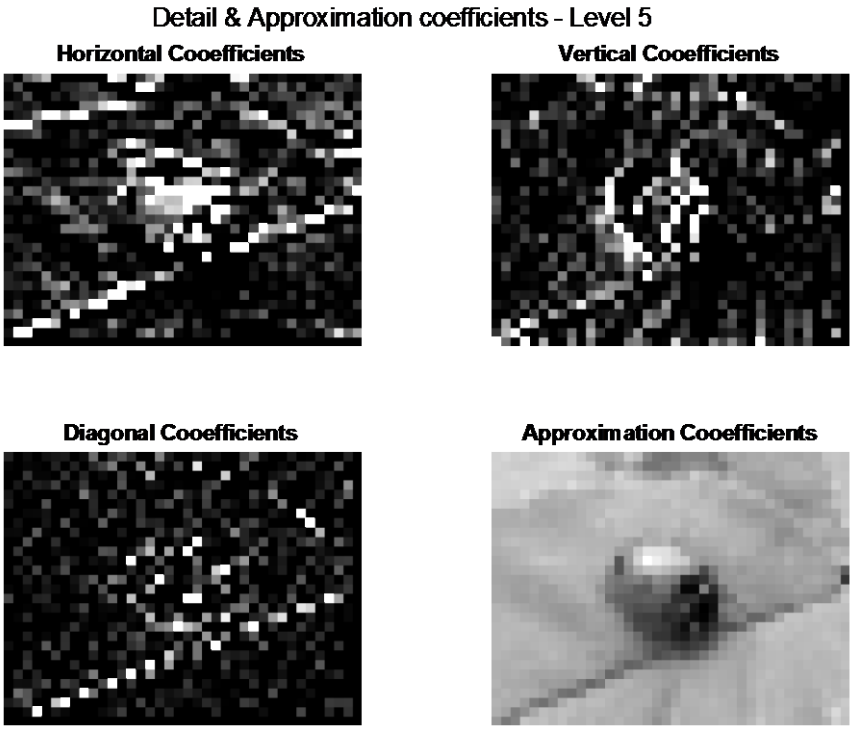
- $Level = 3$:



- $Level = 4$:



- $Level = 5$:



נוכל לראות שככל שנעלה ברמה של הפירוק נקבל שחזור פחות איכותי עקב איבוד מידע. ניתן הסברים בסעיף הבא.

3.5 Explanation for:

(a) What are the differences between the details and the approximation coefficients?

מקדמי הקירוב ("approximation coefficients") מתארים את המידע של התמונה שנמצא בתדרים הנמוכים, כלומר מייצגים את התמונה המסוננת על ידי מסנן מעביר נמוכים.

מקדמי הפרטים ("details coefficients") מתארים את המידע של התמונה שנמצא בתדרים הגבוהים, כלומר מייצגים את התמונה המסוננת על ידי מסנן מעביר גבוהים.

(b) What are the differences between the horizontal, vertical, or diagonal coefficients?

כדי למצוא את המקדמים האנכיים או האופקיים, מעבירים את התמונה במספר מסנני מעבירים גבוהים עבור העמודות או השורות בהתאמה, ומהקומבינציה שלהם יוצרים את המקדמים האלכסוניים.

מקדמים אלה מתארים את המידע בתדרים הגבוהים של התמונה, כלומר המקדמים האנכיים מתארים את התדרים הגבוהים לאורך העמודות של התמונה ולכן נוכל לראות כי השפות האנכיות בתמונה מודגשות, המקדמים האופקיים מתארים את המידע בתדרים הגבוהים לאורך השורות של התמונה ולכן נראה את השפות האופקיות מודגשות, והמקדמים האלכסוניים מתארים את המידע בתדרים הגבוהים לאורך האלכסונים של התמונה ולכן נראה את השפות האלכסוניות מודגשות.

(c) What are the differences between each level?

ככל שהרמה גדולה יותר, דחיסת המידע תהיה גדולה יותר:

עבור מקדמי הקירוב, נקבל שככל שהרמה גדולה יותר, כך נשאר עם תדרים נמוכים יותר, עד שנשאר עם התדרים הנמוכים ביותר כך שנאבד מידע רב על התמונה ולכן נקבל תמונה מטושטשת מאוד שנראית ברזולוציה נמוכה ללא השפות של התמונה.

עבור מקדמי הפרטים, נקבל שככל שהרמה גדולה יותר, כך נשאר עם תדרים גבוהים יותר, עד שנשאר עם התדרים הגבוהים ביותר, כך שנקבל את השפות החדות ביותר והמצומצמות ביותר.

כלומר ככל שהרמה גדלה, כך נשאר עם המידע על התדרים הגבוהים והנמוכים ביותר, ונאבד מידע רב של התמונה המקורית.