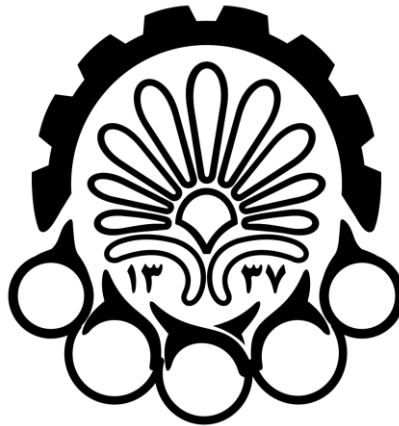


«*In The Name Of GOD*»



دانشگاه صنعتی امیر کبیر
(پلی تکنیک تهران)

[HW-03-Report]

[DIGITAL IMAGE PROCESSING]

Hasan Masroor | [403131030] | May 16, 2025

"فهرست مطالب تمرین 03"

Problem 1	3
A)	3
B)	4
Problem 2	5
A)	5
B)	6
C)	8
D)	9
E)	11
F)	12
G)	14
H)	15
I)	16
J)	19
K)	20
L)	21

M)	22
Problem 3	14
Problem 4	24
A)	26
B)	28
C)	29
Problem 5	34
A)	34
B)	35
C)	36
D)	37
E)	38
F)	40
G)	42

Problem 1: Decoding Identities – A Frequency Domain

Approach to Face Recognition

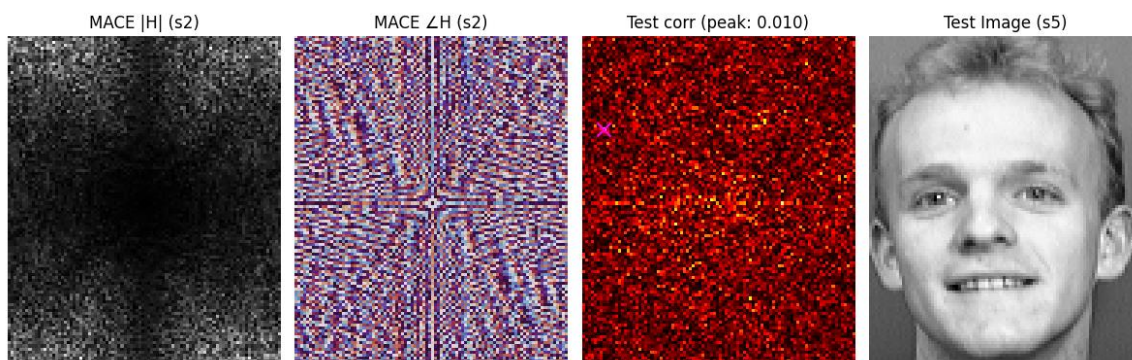
A.

این تمرین پیاده‌سازی یک سیستم تشخیص چهره با فیلتر MACE در حوزه فرکانس را بررسی می‌کند. از تصاویر خاکستری ۵ فرد (هر فرد ۷ تصویر آموزشی) استفاده می‌شود. فیلتر MACE با فرمول خاصی طراحی شده که در صورت تطابق تصویر تست، قله‌ای در صفحه همبستگی ایجاد می‌کند. دقت سیستم با نرخ تشخیص (تعداد تشخیص‌های صحیح تقسیم بر کل تست‌ها) سنجیده می‌شود. این روش کاربرد پردازش فرکانسی در تشخیص چهره را نشان می‌دهد.

در ابتدا کتابخانه‌های موردنیاز را import کردیم. در این بخش ما یک سیستم تشخیص چهره مبتنی بر فیلتر MACE در حوزه فرکانس پیاده‌سازی کردیم. ابتدا تصاویر آموزشی و تست نرمالایز شدند. سپس برای هر شخص، فیلتر MACE مربوطه با استفاده از تبدیل فوری تصاویر آموزشی و فرمول اختصاصی محاسبه گردید. در مرحله بعد، یک تصویر تست به صورت تصادفی انتخاب شده و با محاسبه همبستگی آن با فیلترهای ساخته‌شده، تشخیص چهره انجام شد. نتیجه بر اساس وجود قله واضح در خروجی همبستگی تعیین گردید. در نهایت، ویژگی‌های فیلتر (دامنه و فاز)، نقشه همبستگی و تصویر تست به همراه نتیجه تشخیص به صورت بصری نمایش داده شدند. این پیاده‌سازی، کارایی روش‌های حوزه فرکانس در تشخیص چهره و مکانیزم عملکرد فیلتر MACE را به خوبی نشان می‌دهد.

در نهایت خروجی این قسمت به صورت زیر در آمده است:

Test image from subject s5 - Predicted as s2 - Decision: Unmatched (peak = 0.010)



خروجی کد نشان می‌دهد که تصویر آزمایشی از فرد (s5) به صورت تصادفی انتخاب شده و با فیلتر MACE مربوط به کلاس (s2) پیش‌بینی شده است. با این حال، قله همبستگی (0.010) بسیار کوچک و کمتر از آستانه (0.5) است و چون کلاس پیش‌بینی شده با کلاس واقعی مطابقت ندارد، نتیجه "Unmatched" اعلام شده است. این نشان‌دهنده عدم تطبیق تصویر با کلاس پیش‌بینی شده است.

B.

در این بخش ما نرخ تشخیص سیستم تشخیص چهره مبتنی بر فیلتر MACE را محاسبه کردیم. ابتدا برای هر کلاس (شخص) فیلتر MACE مربوطه را با استفاده از تصاویر آموزشی آن کلاس محاسبه کردیم. سپس برای هر تصویر تست، پاسخ فیلترهای مختلف را بررسی کرده و کلاسی که بیشترین پاسخ (قله) را ایجاد کرده به عنوان تشخیص نهایی انتخاب شد. در نهایت با مقایسه تشخیص‌های سیستم با برچسب‌های واقعی، تعداد تشخیص‌های صحیح شمارش شده و نرخ تشخیص کلی به صورت درصد محاسبه و نمایش داده شد:

Test # 1	True: 0 - Pred: 4	Peak: 5.752	X
Test # 2	True: 0 - Pred: 0	Peak: 5.944	✓
Test # 3	True: 0 - Pred: 1	Peak: 6.279	X
Test # 4	True: 1 - Pred: 1	Peak: 23.137	✓
Test # 5	True: 1 - Pred: 1	Peak: 11.179	✓
Test # 6	True: 1 - Pred: 1	Peak: 20.602	✓
Test # 7	True: 2 - Pred: 2	Peak: 10.816	✓
Test # 8	True: 2 - Pred: 2	Peak: 8.180	✓
Test # 9	True: 2 - Pred: 0	Peak: 7.618	X
Test #10	True: 3 - Pred: 0	Peak: 7.679	X
Test #11	True: 3 - Pred: 1	Peak: 7.577	X
Test #12	True: 3 - Pred: 3	Peak: 5.146	✓
Test #13	True: 4 - Pred: 2	Peak: 6.287	X
Test #14	True: 4 - Pred: 4	Peak: 6.011	✓
Test #15	True: 4 - Pred: 0	Peak: 7.237	X
Recognition Rate: 53.33% (8/15)			

نتایج نشان می‌دهد سیستم با دقت 53.33% (8 تشخیص صحیح از 15 تست) عمل کرده است. بررسی خروجی‌ها نشان می‌دهد سیستم برای برخی افراد (مثل کلاس 1 با 3 تشخیص صحیح) عملکرد خوبی داشته اما در تشخیص برخی دیگر (مثل کلاس 3 با تنها 1 تشخیص صحیح از 3 تست) دچار مشکل شده است. مقادیر Peak در موارد خطا عموماً پایین‌تر است که نشان می‌دهد فیلترها برای برخی کلاس‌ها بهینه نیستند و نیاز به بهبود در طراحی فیلتر یا پیش‌پردازش تصاویر وجود دارد.

Problem 2: Image Inpainting: Interpolation vs. PDE-Based Inpainting

A.

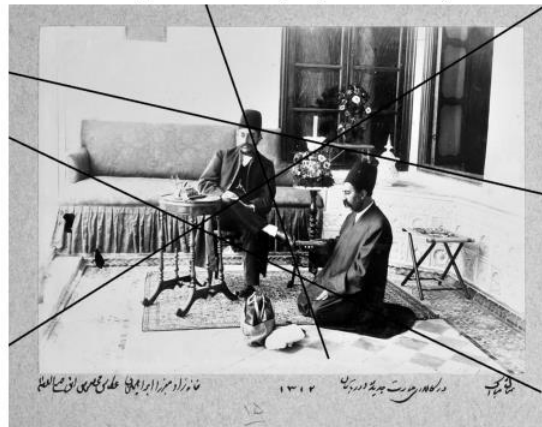
در این تمرین ما یک سیستم Inpainting تصویر را با استفاده از معادلات دیفرانسیل جزئی (PDE) و روش‌های درونیایی پیاده‌سازی می‌کنیم تا بخش‌های آسیب‌دیده یا حذف‌شده تصاویر را ترمیم کنیم. سپس ما عملکرد این دو روش را از نظر کیفیت بصری، سرعت پردازش و تأثیر پارامترها مقایسه و تحلیل می‌کنیم تا بفهمیم چطور مدل‌های ریاضی می‌توانند برای بازیابی تصاویر به کار روند.

در ابتدا کتابخانه‌های موردنیاز را import کردیم. در این بخش، ما تصاویر اصلی و تخریب‌شده (هم خاکستری و هم رنگی) را بارگذاری کردیم، آنها را به فرمت مناسب تبدیل نمودیم و مقادیر پیکسل‌ها را به بازه $[0,1]$ نرمالایز کردیم. سپس برای بررسی اولیه، تصاویر را در خروجی به صورت زیر نمایش دادیم:

Grayscale Image (Normalized)



Grayscale Destroyed (Normalized)



Color Image (Normalized)



Color Destroyed (Normalized)



❖ چرا نرمال سازی تصاویر به بازه $[0,1]$ برای الگوریتم های بازسازی بعدی مهم است؟

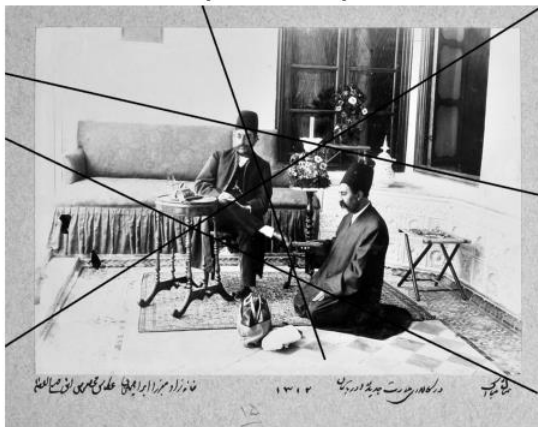
برای اجرای صحیح الگوریتم های بازسازی تصویر، مخصوصاً روش هایی که مبتنی بر معادلات دیفرانسیل یا محاسبات عددی هستند، نرمال سازی تصاویر به بازه $[0,1]$ نقش بسیار مهمی دارد. در ادامه به برخی از این دلایل می پردازیم:

- I. **سازگاری عددی:** بسیاری از الگوریتم های بازسازی مثل روش های PDE یا درونیایی فرض می کنند که مقادیر پیکسل ها در یک بازه استاندارد معمولاً $[0,1]$ قرار دارند. این کار باعث می شود محاسبات عددی پایدارتر و قابل اعتمادتر باشند.
- II. **کاهش خطاهای مقیاس:** اگر مقادیر پیکسل ها در بازه های مختلفی (مثل $[0,255]$ یا بالاتر) باشند، ممکن است در محاسبات مثل مشتق گیری در PDE یا وزن دهی در درونیایی (خطاهای مقیاس به وجود آید. نرمال سازی این مشکل را برطرف می کند.
- III. **تسهیل مقایسه:** نرمال سازی باعث می شود که مقادیر پیکسل ها در تصاویر خاکستری و رنگی قابل مقایسه باشند، به خصوص وقتی معیارهای کمی مثل MSE یا PSNR محاسبه می شوند.
- IV. **سازگاری با روش های PDE:** در روش های PDE مثل معادله انتشار حرارت، پارامترهایی مثل ضریب انتشار (λ) و مقادیر اولیه معمولاً برای بازه $[0,1]$ تنظیم می شوند. اگر مقادیر پیکسل ها خارج از این بازه باشند، ممکن است نیاز به تنظیم مجدد پارامترها باشد که پیچیدگی را افزایش می دهد.
- V. **مدیریت کانال های رنگی:** در تصاویر رنگی، نرمال سازی هر کانال به صورت مستقل تضمین می کند که هر کانال (قرمز، سبز، آبی) به طور یکسان در فرآیند بازسازی مشارکت کند و از تأثیر نامتعادل یک کانال جلوگیری می شود.
- VI. **پایداری الگوریتم ها:** در کنار مواردی که ذکر کردید، نرمال سازی به پایداری الگوریتم ها در فرآیندهای بهینه سازی هم کمک می کند. زمانی که مقادیر پیکسل ها در یک بازه استاندارد قرار دارند، الگوریتم ها می توانند سریع تر همگرا شوند و از مشکلاتی مانند افتادن در مینیمم های محلی جلوگیری کنند.

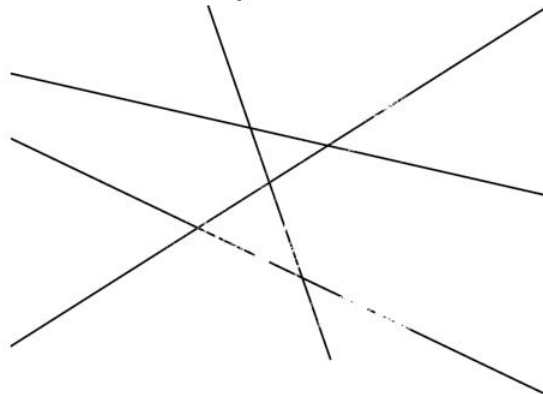
B

در این مرحله، ما ناحیه های تخریب شده تصاویر را با مقایسه تفاضل پیکسل های تصاویر سالم و تخریب شده شناسایی و ماسک های مربوطه را ایجاد کردیم. برای این کار، پیکسل هایی که اختلاف معناداری (بیش از ۰.۰۵) داشتند به عنوان مناطق آسیب دیده در نظر گرفته شدند. سپس ماسک های ایجاد شده برای تصاویر خاکستری و رنگی به همراه تصاویر تخریب شده نمایش داده شدند تا صحت تعریف مناطق آسیب دیده بررسی شود. خروجی این ماسک ها به همراه تصاویر اصلی به شرح زیر هستند:

Grayscale Destroyed



Grayscale Mask



Color Destroyed



Color Mask

عمارت عالی قاپو

❖ چگونه ماسک باینری را ذخیره و استفاده می‌کنید تا نواحی تخریب‌شده را شناسایی کنید؟ چه ساختارهای داده‌ای برای این کار لازم است؟

برای شناسایی نواحی آسیب‌دیده در تصاویر، از ماسک‌های باینری استفاده می‌شود که به صورت آرایه‌های NumPy ذخیره می‌شوند. این ماسک‌ها نواحی خراب‌شده را از نواحی سالم تصویر جدا می‌کنند. در تصویر خاکستری، ماسک باینری به صورت یک آرایهٔ دوبعدی با همان ابعاد تصویر است، که در آن مقدار ۱ نشان‌دهندهٔ پیکسل‌های سالم و مقدار ۰ نشان‌دهندهٔ پیکسل‌های آسیب‌دیده است. در تصاویر رنگی نیز از همین رویکرد استفاده می‌شود، با این تفاوت که تفاوت بین تصویر اصلی و تصویر خراب در هر سه کانال رنگی (قرمز، سبز، آبی) بررسی شده و اگر مقدار تفاوت در هر کانالی از یک آستانه مشخص (مثلاً 0.05) بیشتر باشد، آن پیکسل به عنوان ناحیه سالم علامت‌گذاری می‌شود.

ساختار داده‌ای که برای ذخیره‌سازی این ماسک‌ها استفاده می‌کنیم، آرایه‌ی NumPy است که در آن نوع داده به‌طور معمول np.uint8 است. این نوع داده برای ذخیره‌سازی مقادیر باینری (۰ و ۱) بهینه است و امکان پردازش سریع و بدون هدر دادن حافظه را فراهم می‌کند. این ماسک‌ها در مراحل بازسازی تصویر (Inpainting)

استفاده می‌شوند تا تنها روی نواحی آسیب‌دیده بازسازی انجام شود و بقیه نواحی بدون تغییر باقی بمانند. همچنین این ماسک‌ها به راحتی قابل نمایش، ذخیره‌سازی یا استفاده در سایر الگوریتم‌های پردازش تصویر هستند.

در نتیجه، مقادیر ۰ در ماسک‌ها نمایانگر نواحی تخریب‌شده و مقادیر ۱ نمایانگر نواحی سالم هستند. این ماسک‌ها کمک می‌کنند که تنها نواحی آسیب‌دیده در فرآیند بازسازی تصویر هدف قرار گیرند.

C.

در این مرحله، ما مختصات و مقادیر پیکسل‌های سالم تصاویر (هم خاکستری و هم رنگی) را استخراج کردیم تا به عنوان داده‌های پایه برای درون‌یابی مورد استفاده قرار گیرند. برای این کار، ابتدا با استفاده از ماسک‌های ایجادشده، موقعیت مکانی تمام پیکسل‌های سالم را شناسایی کرده و سپس مقادیر شدت روشنایی آنها را استخراج نمودیم. در تصویر رنگی، این فرآیند به صورت جداگانه برای هر کانال رنگ (قرمز، سبز و آبی) انجام شد. در نهایت، اطلاعات استخراج شده شامل تعداد پیکسل‌های سالم، مختصات و مقادیر شدت روشنایی آنها را در ادامه نشان دادیم:

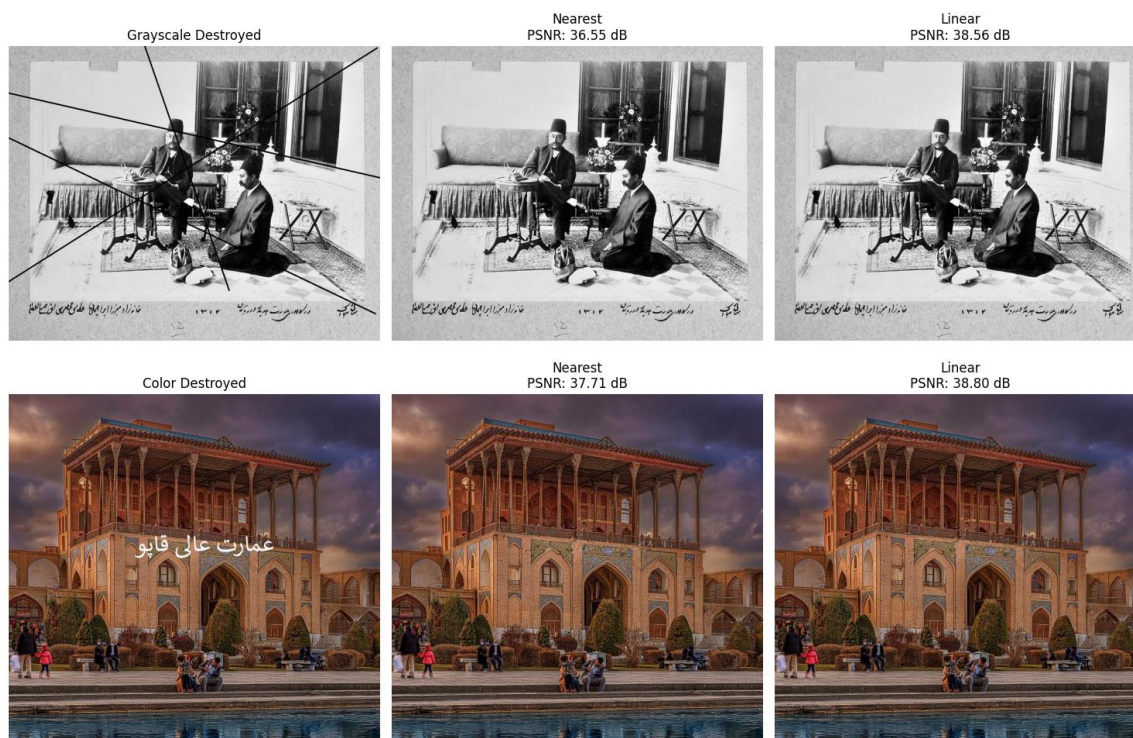
```
- Grayscale Image:
Number of Intact Pixels: 1357555
Coordinates Shape: (1357555, 2)
Intensities Shape: (1357555,)
First 5 Coordinates:
[[0 0]
 [0 1]
 [0 2]
 [0 3]
 [0 4]]
First 5 Intensities:
[0.7529412  0.7490196  0.72156864  0.6901961  0.69803923]

- Color Image:
Number of Intact Pixels: 790080
Coordinates Shape: (790080, 2)
Intensities Shapes: R=(790080,), G=(790080,), B=(790080,)
First 5 Coordinates:
[[0 0]
 [0 1]
 [0 2]
 [0 3]
 [0 4]]
First 5 Intensities (R):
[0.38431373  0.38431373  0.39215687  0.39607844  0.3882353 ]
First 5 Intensities (G):
[0.26666668  0.26666668  0.27450982  0.2784314  0.27058825]
First 5 Intensities (B):
[0.27450982  0.27450982  0.28235295  0.28627452  0.2784314 ]
```

تحلیل خروجی‌ها نشان می‌دهد که در تصویر خاکستری حدود ۱,۳۵۷,۵۵۵ پیکسل سالم وجود دارد که نشان‌دهنده ناحیه تخریب نسبتاً کوچکی است (چون تصاویر معمولاً میلیون‌ها پیکسل دارند). مقادیر شدت روشنایی پیکسل‌های اول مثلاً ۰.۷۵ بیانگر مناطق روشن در تصویر است. در تصویر رنگی نیز ۷۹۰,۰۸۰ پیکسل سالم شناسایی شد که کاهش قابل توجه نسبت به تصویر خاکستری احتمالاً ناشی از تفاوت در ابعاد یا میزان تخریب تصویر است. مقادیر کانال‌های رنگی (R, G, B) در پیکسل‌های ابتدایی مثلاً ۰.۳۸ R و ۰.۲۷ G و ۰.۲۷ B نشان‌دهنده رنگ‌های تیره با غلبه جزئی بر کانال قرمز است و به همین صورت می‌توانیم باقی موارد را نیز تفسیر کنیم.

D

در این مرحله، ما دو روش درونیایی نزدیک‌ترین همسایه و خطی را برای ترمیم تصاویر تخریب‌شده پیاده‌سازی کردیم. برای هر تصویر (هم خاکستری و هم رنگی)، با استفاده از مختصات و مقادیر پیکسل‌های سالم، مناطق تخریب‌شده را پر کردیم. سپس کیفیت نتایج را با معیار PSNR (نسبت سیگنال به نویز پیک) ارزیابی نمودیم. نتایج نشان داد که روش درونیایی خطی (Linear) معمولاً عملکرد بهتری نسبت به روش نزدیک‌ترین همسایه (Nearest) دارد، چرا که PSNR بالاتری تولید می‌کند. در نهایت، تصاویر ترمیم‌شده به همراه مقادیر PSNR مربوطه نمایش داده شدند تا مقایسه بصری و کمی بین روش‌ها بهتر ملموس باشند:



همانطور که می‌دانیم هرچه مقدار PSNR بالاتر باشد نشان‌دهنده ترمیم بهتر در این تصاویر است و تصویر رنگی با روش خطی با PSNR برابر 38.8 Db بهترین عملکرد را از خود نشان داد و بعد از آن هم تصویر خاکستری با روش خطی عملکرد بهتری داشت.

❖ هر روش درونیابی از نظر حفظ لبه‌ها (Edge Preservation) و صافی (Smoothness) چه رفتار متفاوتی دارد؟ دو روش (مثلاً درونیابی نزدیک‌ترین همسایه و درونیابی خطی) را برای هر دو حالت ناحیه‌های تخریب‌شده‌ی کوچک و ناحیه‌های تخریب‌شده‌ی بزرگ مقایسه کنید.

حفظ لبه‌ها (Edge Preservation) به توانایی روش در بازسازی لبه‌های تیز و جزئیات تصویر، مانند مرز بین نواحی روشن و تیره، اشاره دارد و روشی که بتواند این لبه‌ها را بدون محو شدن (blurring) یا ایجاد مصنوعات (artifacts) بازسازی کند، در حفظ لبه‌ها بهتر عمل می‌کند، در حالی که صافی (Smoothness) به پیوستگی و نرمی تغییرات شدت در تصویر بازسازی‌شده مربوط است و روشی که گرادیان‌های نرم و طبیعی تولید کند (بدون مصنوعات بلوکی یا تکه‌تکه) صافی بهتری خواهد داشت؛ همچنین نواحی کوچک (Small Destroyed Areas) معمولاً به چند پیکسل تخریب‌شده (مثلاً ۱ تا ۱۰ پیکسل) مانند خطوط نازک یا نقاط کوچک اشاره دارد، در حالی که مناطق بزرگ گمشده (Large Missing Regions) به نواحی بزرگ‌تر (مثلاً بلوک‌هایی با صدها یا هزاران پیکسل) مانند تکه‌های بزرگ تصویر که تخریب شده‌اند، مربوط می‌شود.

ما دو روش رو مقایسه می‌کنیم یکی نزدیک‌ترین همسایه (Nearest Neighbor) و خطی (Linear). این روش‌ها در کد با استفاده از `scipy.interpolate.griddata` پیاده‌سازی شدند:

نزدیک‌ترین همسایه: برای هر پیکسل تخریب‌شده، شدت نزدیک‌ترین پیکسل سالم روکی می‌کند.

خطی: شدت پیکسل‌های تخریب‌شده رو با میانگین وزن‌دار شدت‌های پیکسل‌های سالم اطراف (بر اساس فاصله) محاسبه می‌کند.

• **مقایسه روش‌ها:**

- **نزدیک‌ترین همسایه (Nearest Neighbor)**

در روش نزدیک‌ترین همسایه، حفظ لبه‌ها در نواحی کوچک به‌خوبی انجام می‌شود، زیرا این روش با کپی مستقیم شدت پیکسل سالم نزدیک، لبه‌های تیز و جزئیات ریز تصویر، مانند مرزهای بین نواحی روشن و تیره، را بدون محو شدن بازسازی می‌کند؛ به‌عنوان مثال، در بازسازی یک خط نازک تخریب‌شده، این روش قادر است با استفاده از شدت پیکسل سالم مجاور، شکل خط را با دقت بالایی بازسازی نماید؛ با این حال، در مناطق بزرگ، عملکرد این روش ضعیف است، زیرا با اعمال شدت یکنواخت یک پیکسل سالم به کل ناحیه تخریب‌شده، لبه‌ها به‌صورت تکه‌تکه و غیرطبیعی بازسازی می‌شوند، به‌طوری‌که در صورت وجود یک بلوک بزرگ تخریب‌شده، ممکن است کل بلوک با شدت ثابت پر شود و این امر منجر به غیریپوسته به نظر آمدن لبه‌های اطراف ناحیه شود؛ از منظر صافی، این روش در نواحی کوچک مشکلی ایجاد نمی‌کند، زیرا پیکسل‌های سالم نزدیک معمولاً شدت‌های مشابهی دارند و کپی شدت یک پیکسل به پیکسل‌های اطراف نتیجه‌ای نسبتاً پیوسته می‌دهد، اما در مناطق بزرگ، صافی ضعیفی دارد، زیرا شدت ثابت به کل ناحیه اعمال می‌شود و این امر منجر به ایجاد مصنوعات بلوکی می‌شود که تصویر را به‌ویژه در نواحی با تغییرات تدریجی شدت، مانند گرادیان‌های نرم، غیرطبیعی جلوه می‌دهد. روش

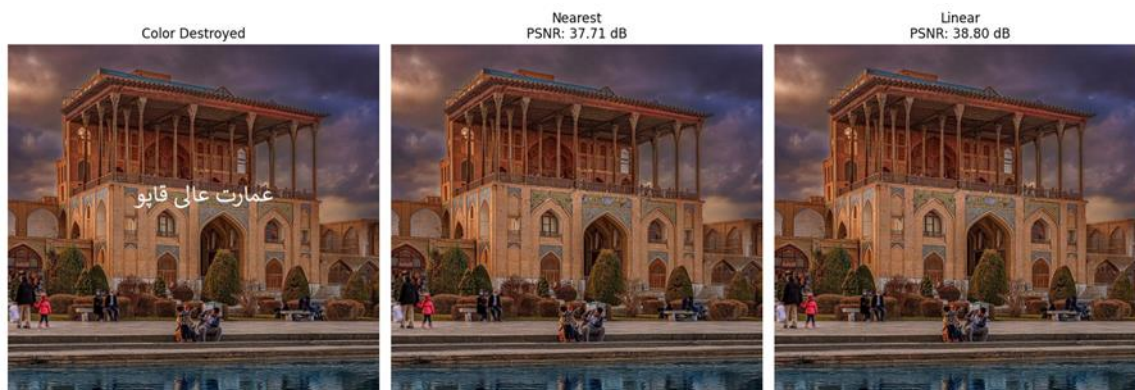
نزدیک‌ترین همسایه از مزایای نظیر سرعت بالا، سادگی اجرا و عملکرد عالی در نواحی کوچک با تغییرات کم شدت برخورداریست، اما در عین حال معایبی نیز دارد، از جمله کیفیت بصری پایین در مناطق بزرگ و ایجاد مصنوعات بلوکی که ظاهر تصویر را تحت تأثیر قرار می‌دهد.

- خطی (Linear)

در روش خطی، حفظ لبه‌ها در نواحی کوچک ممکن است با چالش‌هایی همراه باشد، زیرا محاسبات شدت‌ها بر اساس میانگین وزن‌دار با توجه به فاصله انجام می‌شود که می‌تواند لبه‌های تیز را نرم‌تر کند؛ به‌عنوان مثال، در بازسازی یک خط نازک تخریب‌شده، لبه‌های خط ممکن است کمی محو شوند، زیرا شدت پیکسل‌های اطراف با یکدیگر ترکیب می‌شوند، اما در مناطق بزرگ، این روش نسبت به نزدیک‌ترین همسایه عملکرد بهتری در حفظ لبه‌ها دارد، چرا که با ایجاد یک گرادیان پیوسته بین پیکسل‌های سالم اطراف، لبه‌های اطراف ناحیه تخریب‌شده را به‌صورت طبیعی‌تر بازسازی می‌کند؛ برای نمونه، در صورت وجود یک بلوک بزرگ تخریب‌شده، لبه‌های اطراف آن به‌صورت نرم و پیوسته بازسازی می‌شوند؛ از نظر صافی، در نواحی کوچک، این روش ممکن است باعث محو شدن جزئیات ریز شود، زیرا میانگین‌گیری شدت‌ها می‌تواند تغییرات تیز شدت، مانند لبه‌ها، را از بین ببرد، اما در مناطق بزرگ، صافی به‌صورت برجسته‌ای مطلوب است، زیرا گرادیان‌های نرم و پیوسته‌ای ایجاد می‌کند که تصویر بازسازی‌شده را طبیعی‌تر جلوه می‌دهد؛ به‌عنوان مثال، در نواحی بزرگ با تغییرات تدریجی شدت، مانند آسمان یا پوست، روش خطی نتایج بسیار بهتری ارائه می‌دهد؛ این روش از مزایای نظیر عملکرد عالی در مناطق بزرگ و نواحی با تغییرات تدریجی شدت، مانند گرادیان‌های نرم، برخوردار است، اما معایبی نیز دارد، از جمله محاسبات سنگین‌تر و احتمال محو شدن لبه‌های تیز که می‌تواند دقت بازسازی را در برخی موارد تحت تأثیر قرار دهد.

E

این پارت را در بخش قبلی همراه با تصویر خاکستری انجام دادیم و نتیجه زیر را مشاهده کردیم و توانستیم با هر دو روش عملکرد نسبتاً خوبی بدست آوریم و روش خطی عملکرد بهتری از خود نشان داد:



❖ چطور بازسازی را برای تصاویر رنگی انجام می‌دهید؟ چه مشکلاتی ممکن است هنگام درونیابی جداگانه‌ی هر کانال رنگی به وجود بیاید؟

در این بخش، برای بازسازی تصاویر رنگی هر یک از کانال‌های R، G و B به صورت مستقل و جداگانه درونیابی شدند. ابتدا مختصات و شدت پیکسل‌های سالم برای هر کانال با استفاده از ماسک استخراج شد و سپس از تابع griddata برای تخمین مقادیر تخریب‌شده استفاده شد. مقادیر پیش‌فرض در نواحی بدون داده با میانگین شدت‌ها جایگزین شد و در پایان، تنها پیکسل‌های تخریب‌شده به روزرسانی شدند و مقادیر سالم دست‌نخورده باقی ماندند. این روش ساده و کارآمد است و نتایج خوبی در نواحی کوچک ارائه می‌دهد، اما چون کانال‌ها مستقل بازسازی می‌شوند، ممکن است در نواحی بزرگ باعث بروز ناسازگاری رنگی یا ناهماهنگی در لبه‌ها شود. برای بهبود کیفیت در چنین مواردی، استفاده از روش‌های برداری یا فضاهای رنگی ادراکی می‌تواند مفید باشد. با وجود سادگی پیاده‌سازی، این روش ممکن است با مشکلات زیر مواجه شود:

یکی از مهم‌ترین چالش‌ها ناسازگاری رنگی (Color Inconsistency) است؛ چون هر کانال به‌طور جداگانه بازسازی می‌شود، ممکن است تفاوت‌های جزئی بین آن‌ها منجر به تولید رنگ‌های غیرواقعی یا نویزهای رنگی شود. مشکل دوم، عدم تطابق لبه‌ها (Edge Misalignment) است. اگر لبه‌ها در هر کانال به‌طور دقیق و هماهنگ بازسازی نشوند، ممکن است مرزهای تصویر دچار انحراف یا تاری شوند، که معمولاً به‌صورت هاله‌های رنگی یا لبه‌های ناهم‌راستا بروز می‌کند. همچنین، این روش می‌تواند منجر به کاهش انسجام ادراکی (Perceptual Incoherence) شود؛ زیرا چشم انسان نسبت به هماهنگی رنگی بسیار حساس است، و اختلافات جزئی میان کانال‌ها ممکن است باعث ناهماهنگی در بافت‌ها و گرادیان‌های رنگی شده و تصویر بازسازی‌شده را مصنوعی جلوه دهد.

F

همانطور که در ابتدا نشان دادیم تصاویر تخریب شده به صورت زیر بودند:

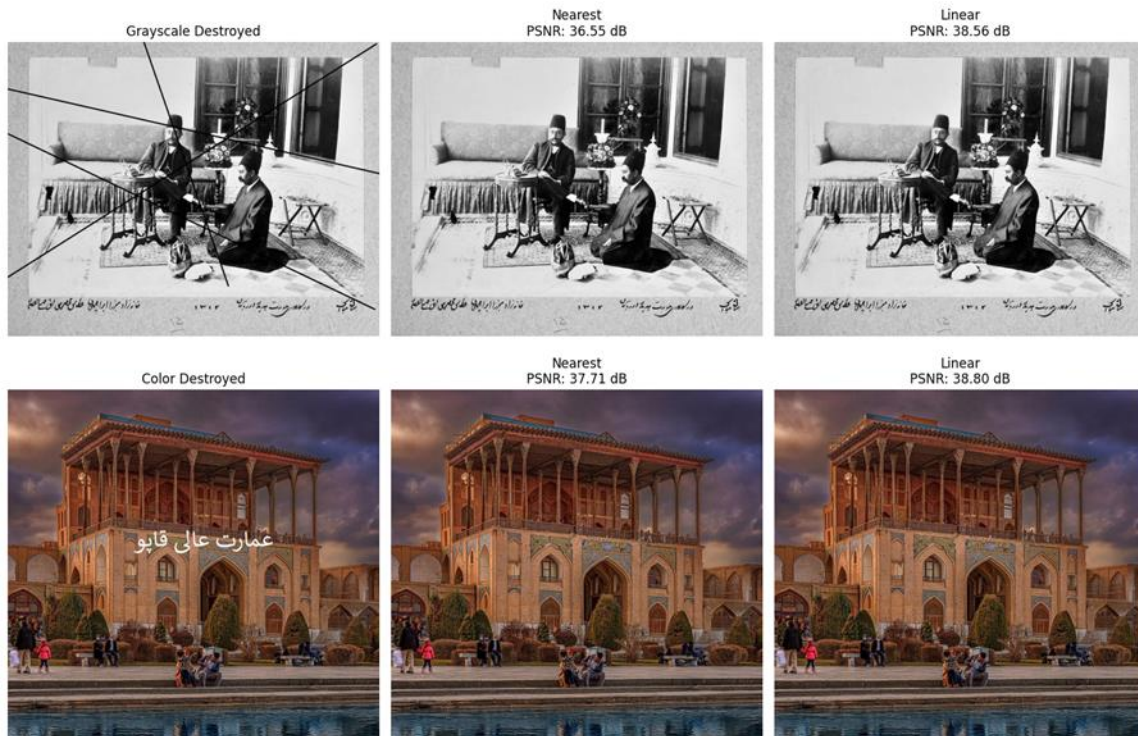
Color Destroyed (Normalized)



Grayscale Destroyed (Normalized)



در بخش D آمدم و با دو روش مختلف تلاش کردیم که این تصاویر را ترمیم کنیم و به نتایج زیر رسیدیم:



هر دو روش برای دو تصویر رنگی و خاکستری توانستند به مقدار PSNR بالایی دست پیدا کنند و عملکرد خیلی خوبی در ترمیم تصاویر داشتند و همانطور که در پارت های قبلی اشاره کردیم روش خطی کمی بهتر از روش دیگر عمل کرده است اما در کل هر دو به خوبی توانستند این کار را انجام دهند.

❖ مقدار PSNR (نسبت سیگنال به نویز اوج) تصاویر بازسازی شده را نسبت به تصاویر اصلی محاسبه کنید. روش درونیایی چگونه بر مقدار PSNR تأثیر می گذارد؟

مقدار PSNR نشان دهنده میزان نزدیکی تصویر بازسازی شده به تصویر مرجع است و هرچه این مقدار بالاتر باشد، کیفیت بازسازی بهتر و خطای کمتری وجود دارد. با مقایسه روش های درونیایی نزدیک ترین همسایه و خطی مشاهده شد که روش خطی به دلیل میانگین گیری وزن دار شدت پیکسل های سالم اطراف، تصاویر بازسازی شده نرم تر و یکنواخت تر تولید می کند که معمولاً باعث افزایش مقدار PSNR و بهبود کیفیت بصری می شود. در مقابل، روش نزدیک ترین همسایه به دلیل حفظ کمتر صافی و وجود لبه های پله ای، معمولاً PSNR پایین تری دارد. البته لازم به ذکر است که PSNR تنها یک معیار کمی است و همیشه بازتاب دهنده تفاوت های بصری به طور کامل نیست؛ بنابراین در کنار آن باید مقایسه بصری نیز انجام شود تا کیفیت واقعی بازسازی به درستی ارزیابی گردد.

در این بخش، ما یک روش مبتنی بر معادلات دیفرانسیل جزئی (PDE) برای ترمیم تصاویر تخریب شده پیاده سازی کردیم. این روش با استفاده از عملگر لاپلاسین، اطلاعات پیکسل های سالم را به تدریج به مناطق تخریب شده منتشر می کند. برای تصاویر رنگی، این فرآیند به صورت مستقل روی هر یک از کانال های R ، G و B اعمال شد. پارامترهای مهم مانند تعداد تکرارها (5000 بار)، گام زمانی (0.05) و ضریب (0.05) λ به دقت تنظیم شدند تا هم پایداری الگوریتم حفظ شود و هم بهترین نتیجه حاصل گردد. در نهایت، کیفیت تصاویر ترمیم شده با معیار PSNR ارزیابی و نتایج به صورت زیر حاصل شدند و این روش به خصوص برای ترمیم نواحی بزرگ تر تخریب شده مناسب است، چرا که با حفظ پیوستگی و همواری تصویر، نتایج طبیعی تری تولید می کند:



❖ نقش های λ ، Δu و fidelity را توضیح دهید. تغییر در λ چه تأثیری روی فرآیند پخش دارد؟

معادله ی مورد استفاده در بازسازی تصویر مبتنی بر PDE شامل سه جزء اصلی است که هر کدام نقش مهمی در فرآیند ترمیم دارند. نخست، ضریب پخش λ که تعیین کننده سرعت و شدت پخش اطلاعات شدت پیکسل ها در تصویر است. مقدار بالاتر این ضریب موجب می شود که فرآیند پخش سریع تر انجام شود و نواحی تخریب شده با سرعت بیشتری پر شوند، اما اگر مقدار λ بیش از حد زیاد باشد، ممکن است باعث ناپایداری عددی و محوشدگی بیش از حد لبه ها و جزئیات تصویر شود. از طرف دیگر، مقدار کم λ فرآیند پخش را کند می کند و بازسازی تصویر به آرامی صورت می گیرد. دوم، لاپلاسیان تصویر Δu است که گرادینان دوم شدت تصویر را نشان می دهد و معیاری برای تغییرات محلی شدت محسوب می شود. این عبارت تفاوت بین مقدار شدت یک پیکسل و میانگین شدت پیکسل های همسایه را اندازه گیری کرده و نقش اصلی آن هموارسازی و پراکندگی اطلاعات از نواحی سالم به نواحی

تخریب شده است. سوم $\chi_{\Omega \setminus D}(f - u)$ است که تضمین می‌کند مقدار شدت پیکسل‌های سالم حفظ شود و در طول فرآیند بازسازی تغییر نکند؛ این امر به حفظ جزئیات اصلی تصویر کمک می‌کند و مانع از تغییر ناخواسته نواحی سالم می‌شود. بنابراین، تغییر مقدار λ تأثیر مستقیم بر سرعت و کیفیت بازسازی دارد؛ مقدار زیاد آن موجب سرعت بالاتر اما احتمال افت کیفیت می‌شود و مقدار کم باعث حفظ بهتر جزئیات ولی سرعت کمتر در ترمیم است. به طور کلی، تنظیم مناسب این پارامترها برای دستیابی به تعادل بین سرعت ترمیم و حفظ کیفیت تصویر حیاتی است.

H.

در بخش قبلی، ما به طور کامل پیاده‌سازی شرایط مرزی نیومن (Neumann Boundary Conditions) را در تابع Laplacian انجام دادیم که یکی از اجزای کلیدی روش مبتنی بر معادلات دیفرانسیل جزئی (PDE) برای ترمیم تصاویر است. در این پیاده‌سازی، با استفاده از یک PDE، با استفاده از explicit finite difference سه ناحیه مختلف از تصویر را به صورت مجزا پردازش کردیم: برای نقاط داخلی تصویر، از عملگر استاندارد لاپلاسین با الگوی ۵-نقطه‌ای استفاده شد که مشتق دوم را در جهت‌های x و y محاسبه می‌کند. در لبه‌های تصویر (بالا، پایین، چپ و راست)، شرایط مرزی نیومن را با تقریب مشتق نرمال صفر پیاده‌سازی کردیم که در آن فقط از سه همسایه موجود استفاده شده است. این کار از انتشار ناخواسته اطلاعات از مرزها به داخل تصویر جلوگیری می‌کند. برای گوشه‌های تصویر، یک نسخه ساده‌شده از شرایط نیومن را اعمال کردیم که فقط از دو همسایه مجاور بهره می‌برد.

❖ شرایط مرزی را چگونه به صورت عددی پیاده‌سازی می‌کنید؟ استفاده از (ghost points) یا تغییر

مستقیم ماتریس را توضیح دهید

در کد بخش G که پیاده‌سازی تابع Laplacian را داشتیم، شرایط مرزی Neumann به صورت عددی با استفاده از جایگزینی مقادیر پیکسل‌های مرزی انجام شده است. به این شکل که برای محاسبه لاپلاسیان در نقاط داخلی تصویر از اختلاف محدود مرکزی استفاده می‌کنیم، اما در نقاط مرزی و گوشه‌ها به جای دسترسی به پیکسل‌های خارج از تصویر که وجود ندارند، مقادیر پیکسل‌های همسایه‌ی داخل تصویر را به عنوان جایگزین لحاظ می‌کنیم. این روش معادل فرض مشتق نرمال صفر (شرایط مرزی Neumann) است، زیرا با این کار تغییرات شدت در جهت عمود بر مرز صفر می‌شود و از نفوذ پخش به بیرون تصویر جلوگیری می‌شود. بنابراین، این پیاده‌سازی در کد به طور مستقیم شرایط مرزی را به صورت عددی مدیریت کرده و به حفظ ثبات و صحت حل PDE کمک می‌کند.

برای مدیریت شرایط مرزی Neumann به صورت عددی در حل معادلات دیفرانسیل جزئی مانند معادله پخش تصویر، دو روش معمول وجود دارد: استفاده از ghost points و تغییرات مستقیم ماتریسی. در روش نقاط شبی، در اطراف مرزهای تصویر نقاط فرضی (نقاط شبی) تعریف می‌شوند که مقادیر آن‌ها به گونه‌ای تعیین می‌شوند تا مشتق نرمال (مشتق در جهت عمود بر مرز) صفر باشد. این معمولاً با تنظیم مقدار نقاط شبی برابر مقدار پیکسل‌های مجاور در داخل تصویر انجام می‌شود که به معنای بازتاب داده شدن مقادیر است. به این ترتیب تغییر شدت در مرز صفر نگه داشته شده و پخش به خارج از تصویر رخ نمی‌دهد. در روش تغییرات

مستقیم ماتریسی، هنگام ساخت ماتریس سیستم معادلات گسسته شده، ضرایب معادلات در مرزها به گونه ای اصلاح می شوند که معادل شرایط Neumann باشند. به طور خاص، برای نقاط مرزی، رابطه ای ایجاد می شود که مشتق نرمال صفر را تضمین می کند و معمولاً این باعث می شود مقادیر مربوط به نقاط خارج از مرز در معادلات حذف یا جایگزین با مقادیر داخل مرز شوند. در هر دو روش، هدف حفظ پیوستگی و جلوگیری از پخش سیگنال به خارج از مرز است. این رویکردها باعث می شوند که حل عددی معادلات با دقت بالا و پایداری در شرایط مرزی انجام شود و تصویر بازسازی شده از نظر فیزیکی و بصری واقعی باقی بماند.

I.

در این بخش، ما به بررسی حساسیت الگوریتم PDE به پارامترهای λ و Δt پرداختیم. در دو پارت قبلی این دو پارامتر را با مقادیر 0.05 بررسی کردیم و نمایش دادیم و حالا می خواهیم دو مجموعه پارامتر مختلف را آزمایش کنیم و هر بار یکی از پارامترها را ثابت و دیگری را تغییر دهیم تا مقایسه بهتری بتوانیم داشته باشیم:

1. آزمایش اول ($\lambda=0.05, \Delta t=0.02$): این تنظیمات با مقدار λ بالاتر، انتشار سریع تر اطلاعات را موجب شد و منجر به ترمیم هموارتر نواحی تخریب شده می گردد.

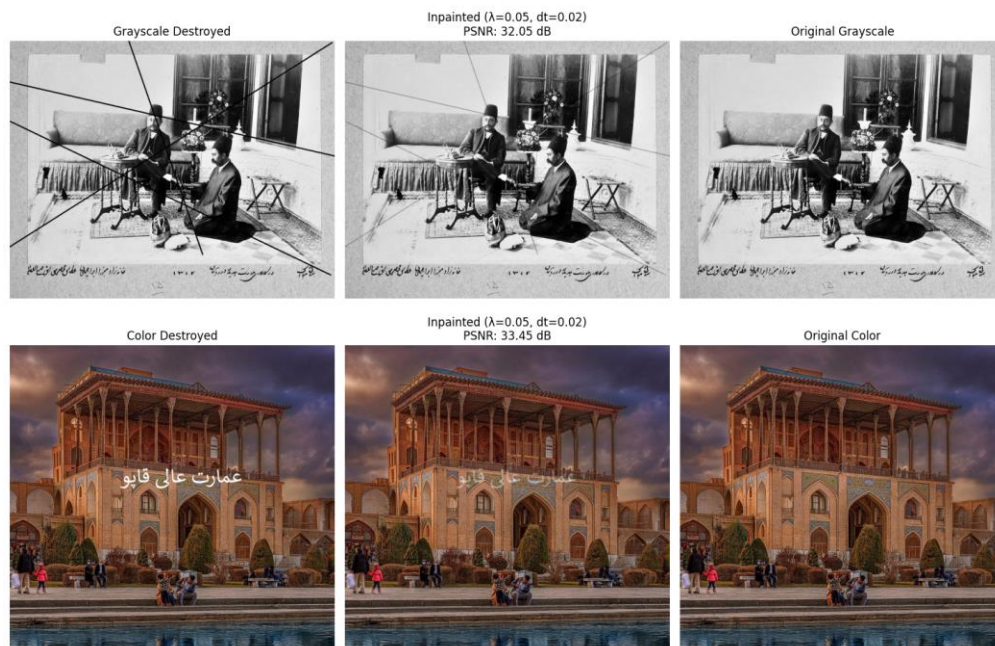
2. آزمایش دوم ($\lambda=0.01, \Delta t=0.05$): با کاهش λ ، فرآیند انتشار کندتر شد و جزئیات و لبه ها بهتر حفظ شدند، هرچند ممکن است نواحی بزرگ تر به طور کامل ترمیم نشده باشند و گام زمانی بزرگ تر باعث شد الگوریتم در همان تعداد تکرار به همگرایی کمتری برسد.

3. نتایج مقارنه دو پارت قبلی و این دو آزمایش را در زیر نمایش دادیم:

■ $\lambda=0.05, dt=0.05$

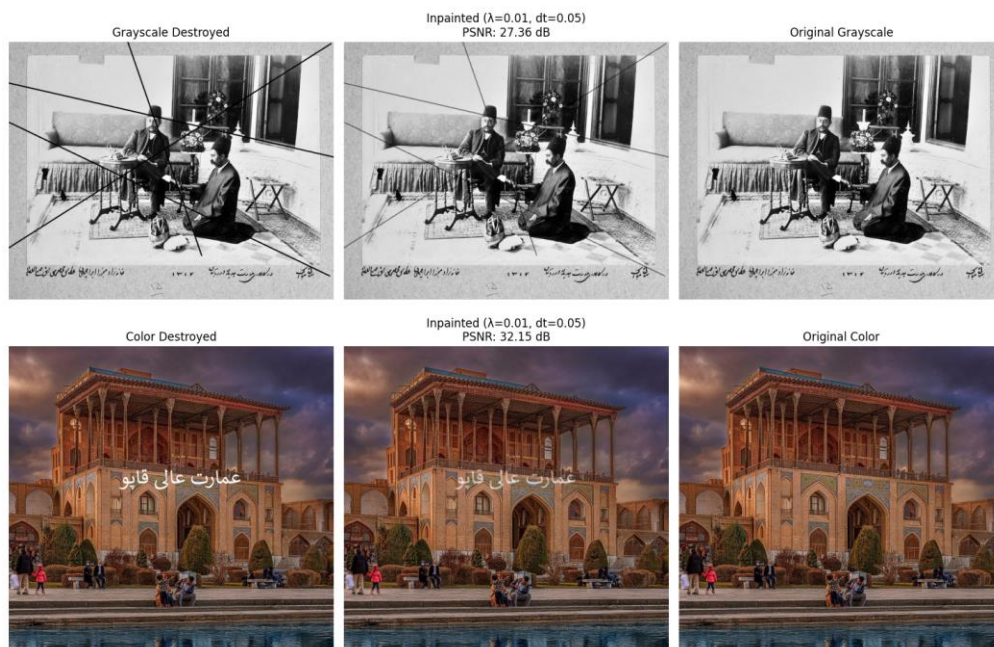


- $\lambda=0.05, dt=0.02$



همانطور که می بینیم دو تصویر ترمیم یافته با این مقداردهی نسبت به مقدار دهی هر دو 0.05 کمی عملکردش بدتر شده است و هم مقدار PSNR آن کمتر شده است و هم بخش های تخریب شده مثل نوشته و خطوط را در خروجی همچنان با اثر کم می توانیم مشاهده کنیم.

- $\lambda=0.01, dt=0.05$



در آزمایش سوم هم می بینیم که باز هم نتیجه از قسمت قبلی بدتر شد و هم مقدار PSNR کاهش پیدا کرده است و هم بخش های تخریب شده نسبت به تصاویر آزمایش قبلی بیشتر مشخص هستند و عملکرد آنچنان خوبی از خود برای بازسازی و ترمیم نشان نداده است.

❖ تغییر λ چه تأثیری در صاف سازی لبه ها دارد؟ اگر Δt خیلی زیاد یا خیلی کم باشد چه اتفاقی می افتد؟ با تحلیل پایداری و نمایش تصویری توضیح دهید.

در معادله بازسازی بر پایه PDE (پخش گرما)، پارامتر λ نقش ضریب پخش را دارد و مستقیماً سرعت و میزان صاف سازی تصویر را کنترل می کند. مقدار بزرگ تر λ باعث پخش سریع تر شدت پیکسل ها در نواحی تخریب شده می شود که به نرمی بیشتر تصویر منجر می شود، اما در عوض می تواند لبه های تیز و جزئیات ساختاری را محو کند. در مقابل، مقدار کوچک تر λ باعث پخش آهسته تر می شود که لبه ها را بهتر حفظ می کند، ولی ممکن است نواحی تخریب شده به طور کامل بازسازی نشوند یا الگوی غیرطبیعی در تصویر باقی بماند. با $\lambda = 0.05$ و $\Delta t = 0.02$ ، بازسازی هموار و طبیعی انجام شد و لبه ها به طور قابل قبول حفظ شدند. مقادیر PSNR برابر با 32.05 (خاکستری) و 33.45 (رنگی) نشان دهنده تعادل مناسب بین صافی و دقت بازسازی بود. در مقابل، با $\lambda = 0.01$ و $\Delta t = 0.05$ ، سرعت پخش بسیار پایین تر بود. در نتیجه تصویر خاکستری با PSNR پایین تر (27.36 dB) بازسازی شد و لبه ها بهتر حفظ شدند، اما نواحی بازسازی شده نسبت به حالت اول پراکنده تر و کم تر هموار بودند. در تصویر رنگی، تغییر زیادی ایجاد نشد (32.15 dB) ولی تفاوت های ظریفی در یکنواختی رنگ ها دیده می شد.

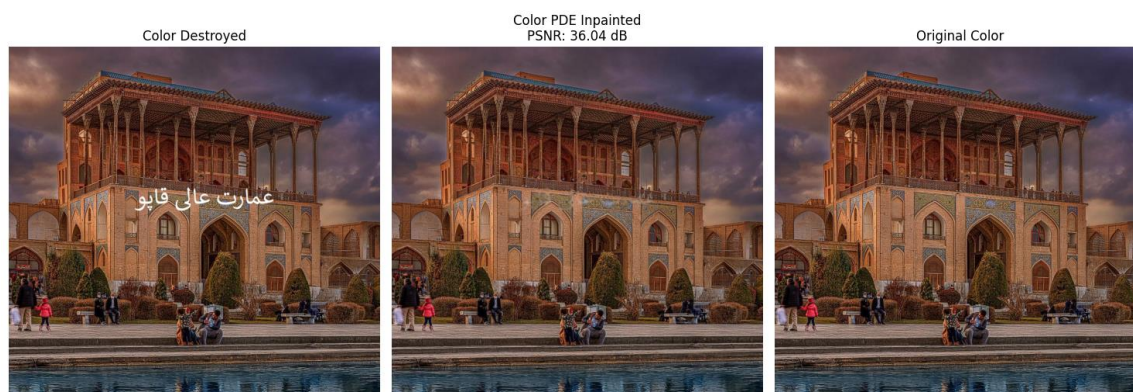
اگر Δt خیلی بزرگ باشد: الگوریتم ممکن است ناپایدار شود و نوسانات عددی ایجاد کند (پخش بیش از حد، نویز، یا ناپایداری محاسباتی). در آزمایش دوم با $\lambda = 0.01$ و $\Delta t = 0.05$ مقدار حاصل 0.004 بود که در محدوده پایدار قرار دارد، ولی اگر λ افزایش می یافت، پایداری به خطر می افتاد. به همین دلیل باید شرط پایداری CFL رعایت شود:

$$\Delta t \cdot \lambda \cdot \lambda_{\max} < 0.5$$

اگر Δt خیلی کوچک باشد: پایداری بالا می ماند، اما تعداد تکرارهای لازم برای رسیدن به بازسازی مؤثر افزایش می یابد، بنابراین زمان اجرا بیشتر می شود بدون بهبود چشمگیر در کیفیت.

نتایج به دست آمده نشان می دهد که پارامترهای λ و Δt تأثیر مستقیمی بر کیفیت بازسازی تصویر دارند. انتخاب مقدار مناسب λ باعث ایجاد تعادل بین حفظ ساختارهای لبه ای و صافی نواحی تخریب شده می شود. مقدار زیاد λ بازسازی را سریع تر و هموارتر می کند اما می تواند به محو شدن لبه ها منجر شود، در حالی که مقدار کم آن باعث حفظ لبه ها می شود ولی صافی تصویر را کاهش می دهد. همچنین، انتخاب Δt باید با توجه به λ انجام شود تا شرط پایداری عددی رعایت شود و از نوسانات ناخواسته جلوگیری گردد. در مجموع، تنظیم صحیح این دو پارامتر نقش کلیدی در رسیدن به بازسازی بصری قابل قبول و عددی پایدار دارد.

این پارت را هم در قسمت D به همراه تصویر خاکستری انجام دادیم و نتیجه زیر حاصل شد و خروجی را به طور کامل بررسی کردیم:



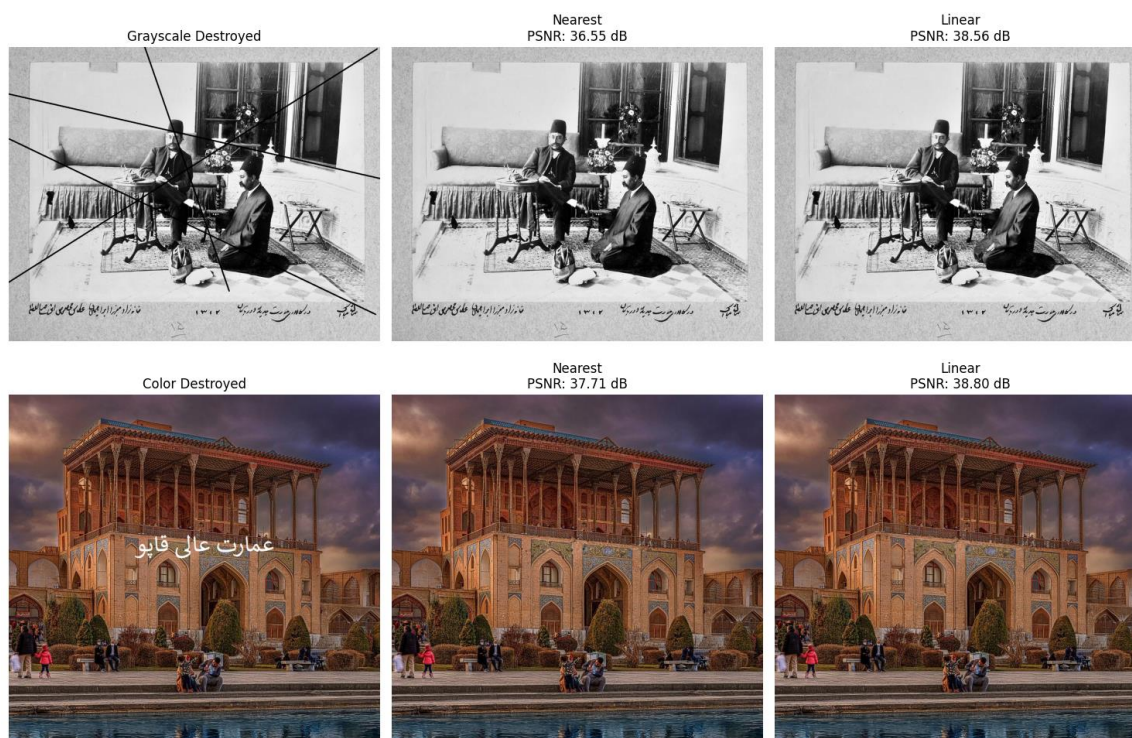
❖ اگر PDE را به صورت جداگانه روی هر کانال اجرا کنیم، چه تأثیری روی سازگاری رنگ‌ها دارد؟ آیا رویکرد coupled می‌تواند بهتر باشد؟

اگر PDE را به صورت جداگانه روی هر کانال رنگی (R، G، B) اعمال کنیم، هر کانال بدون در نظر گرفتن وابستگی‌های ساختاری یا رنگی نسبت به کانال‌های دیگر پردازش می‌شود. اگرچه این رویکرد ساده و در بسیاری از موارد مؤثر است، اما ممکن است باعث بروز مشکلاتی در color consistency شود، به‌ویژه در نواحی دارای لبه‌های تیز یا بافت‌های پیچیده. دلیل این امر آن است که بازسازی در هر کانال مسیر متفاوتی را طی می‌کند و حتی اختلاف‌های جزئی در بازسازی لبه‌ها یا گرادیان‌ها بین کانال‌ها می‌تواند منجر به ایجاد مصنوعات رنگی مانند هاله‌های رنگی، لبه‌های ناهم‌راستا یا تغییر رنگ‌های غیرطبیعی شود. این مسئله اهمیت ویژه‌ای دارد زیرا سیستم بینایی انسان نسبت به ناهماهنگی در گذارهای رنگی بسیار حساس است، حتی زمانی که شدت روشنایی تغییر چندانی نداشته باشد. برای رفع این مشکلات، می‌توان از رویکرد coupled استفاده کرد که در آن سه کانال رنگی به صورت هم‌زمان و وابسته به یکدیگر بازسازی می‌شوند. در این روش‌ها، فرآیند diffusion بر پایه ساختار مشترک تصویر در تمام کانال‌ها هدایت می‌شود تا اطمینان حاصل شود که گرادیان‌ها و لبه‌ها در بین کانال‌ها هماهنگ باقی می‌مانند. همچنین، یک رویکرد جایگزین، تبدیل تصویر به فضای رنگی ادراکی مانند CIE Lab یا YUV و اعمال بازسازی PDE فقط روی کانال روشنایی L یا Y است که بیشترین اطلاعات ساختاری را در خود دارد. سپس کانال‌های رنگی به صورت ساده‌تر بازسازی شده و تصویر به فضای RGB بازگردانده می‌شود. به‌طور کلی، رویکرد به هم پیوسته یا استفاده از فضای رنگی ادراکی می‌تواند انسجام رنگی و یکپارچگی ساختاری بهتری نسبت به پردازش مستقل کانال‌ها ارائه دهد.

ما در این بخش نتایج دو روش درونیایی و PDE را مقایسه کردیم. نتایج نشان داد که هر یک از این روش‌ها در شرایط خاصی عملکرد بهتری دارند. روش PDE با مقدار PSNR برابر 37.51 دسی‌بل برای تصویر خاکستری و 36.04 دسی‌بل برای تصویر رنگی، در حفظ لبه‌ها و جزئیات تصویر عملکرد بهتری دارد، چرا که با استفاده از عملگر لاپلاسین، اطلاعات را به صورت ساختاریافته و با حفظ پیوستگی منتشر می‌کند. این ویژگی باعث می‌شود لبه‌ها به طور طبیعی‌تر و با وضوح بالاتر ترمیم شوند. از سوی دیگر، روش درونیایی خطی (Linear Interpolation) با مقادیر PSNR بالاتر (38.56 دسی‌بل برای تصویر خاکستری و 38.80 دسی‌بل برای تصویر رنگی) برای ترمیم نواحی بزرگ‌تر تخریب‌شده مناسب‌تر است، زیرا با برونمایی داده‌های سالم، پیوستگی رنگ و روشنایی را در نواحی وسیع بهتر حفظ می‌کند. در مقابل، روش درونیایی نزدیک‌ترین همسایه (Nearest Neighbor) با مقادیر PSNR پایین‌تر (36.55 دسی‌بل برای تصویر خاکستری و 37.71 دسی‌بل برای تصویر رنگی) به دلیل ایجاد اثرات بلوکی و تار شدن لبه‌ها، عملکرد ضعیف‌تری دارد. در نتیجه، اگر هدف ترمیم تصاویر با حفظ جزئیات و لبه‌های واضح باشد، روش PDE گزینه بهتری است، اما برای بازیابی نواحی بزرگ‌تر تخریب‌شده، روش درونیایی خطی نتایج مطلوب‌تری ارائه می‌دهد. برای دستیابی به بهترین نتیجه، می‌توان از ترکیب این دو روش به صورت هیبریدی استفاده کرد تا مزایای هر دو روش را همزمان داشته باشیم.

برای درک بهتر، تجزیه و تحلیل و مقایسه نتایج این دو روش را مجدد در ادامه نمایش دادیم:

■ Interpolation:



▪ PDE:



❖ بر اساس مشاهدات بصری، کدام روش را برای نواحی بزرگتر ترجیح می‌دهید؟ چرا؟

همانطور که بالاتر گفتیم بر اساس مشاهدات بصری، روش درونیایی خطی (Linear Interpolation) برای نواحی بزرگتر مناسب‌تر است، زیرا با برونمایی یکنواخت داده‌های سالم، پیوستگی رنگ و روشنایی را بهتر حفظ می‌کند و نتایج طبیعی‌تری در مقایسه با روش PDE (که ممکن است باعث همواری بیش از حد یا محو شدن جزئیات شود) تولید می‌نماید.

L

در این بخش، ما از معیارهای کمی MSE و PSNR (که قبلاً PSNR را برای هر تصویر محاسبه و نمایش دادیم و به طور کامل تفسیر کردیم) برای ارزیابی دقیق‌تر عملکرد روش‌های ترمیم استفاده کردیم. همانطور که در مراحل قبل مشاهده شد، روش درونیایی خطی بالاترین مقدار PSNR را داشت و برای سایر موارد هم کامل بررسی کردیم. حالا می‌خواهیم در این پارت با محاسبه MSE (میانگین مربعات خطا) این معیار را هم محاسبه و نمایش دهیم. مقادیر پایین‌تر MSE نشان‌دهنده خطای کمتر و کیفیت بالاتر ترمیم است. انتظار داریم روش‌هایی که PSNR

بالتری داشتند (مثل درونیایی خطی)، MSE کمتری نیز نشان دهند که این موضوع به صورت کمی دقت هر روش را تأیید می‌کند. خروجی این بخش به صورت زیر حاصل شده است:

```
MSE Grayscale Nearest Interpolation: 0.000221
MSE Grayscale Linear Interpolation: 0.000139
MSE Color Nearest Interpolation: 0.000170
MSE Color Linear Interpolation: 0.000132
MSE Grayscale PDE: 0.000178
MSE Color PDE: 0.000249
```

تحلیل نتایج نشان می‌دهد که روش درونیایی خطی (Linear Interpolation) با کمترین مقادیر MSE (0.000139 برای خاکستری و 0.000132 برای رنگی) بهترین عملکرد را داشته است که این با مقادیر PSNR بالاتر آن در مراحل قبل نیز همخوانی دارد. روش درونیایی نزدیک‌ترین همسایه (Nearest) با MSE های بالاتر (0.000221 و 0.000170) کمی عملکرد ضعیف‌تری نشان داد. روش PDE برای تصویر خاکستری (MSE=0.000178) نسبت به روش Nearest بهتر عمل کرد اما برای تصویر رنگی (MSE=0.000249) ضعیف‌ترین نتیجه را داشت.

❖ مقایسه‌ی MSE و PSNR بین دو روش چه نتایجی می‌دهد؟ این اعداد چه چیزی درباره کیفیت بازسازی نشان می‌دهند؟

مقایسه‌ی MSE و PSNR بین دو روش نتایج کاملاً همسوئی ارائه می‌دهد و هرچه MSE کمتر باشد، PSNR بالاتر است که نشان‌دهنده‌ی کیفیت بهتر بازسازی است. به‌طور مشخص، روش درونیایی خطی با کمترین MSE و بالاترین PSNR بهترین عملکرد را داشت، یعنی خطای کمتری دارد و تصویر را با دقت بالاتری بازسازی کرده است و می‌توانیم همان تفسیرهای قبلی را برای اینجا شرح دهیم.

M.

نتایج مقایسه کارایی محاسباتی دو روش نشان می‌دهد که روش درونیایی با زمان اجرای حدود 1 دقیقه و 40 ثانیه به طور قابل توجهی سریع‌تر از روش PDE با زمان 15 دقیقه و 59 ثانیه عمل کرده است. این تفاوت فاحش در زمان پردازش عمدتاً به دلیل ماهیت تکراری و پیچیده‌ی روش PDE است که نیاز به 5000 تکرار برای همگرایی دارد و در هر تکرار، محاسبات سنگین عملگر لاپلاسین و به‌روزرسانی پیکسل‌ها را انجام می‌دهد. در مقابل، روش درونیایی (به‌ویژه خطی) با محاسبات مستقیم و غیرتکراری، بسیار بهینه‌تر عمل می‌کند. این نتایج نشان می‌دهد که اگرچه روش PDE ممکن است در برخی موارد (مانند حفظ لبه‌ها) کیفیت بهتری ارائه دهد اما هزینه محاسباتی بسیار بالایی دارد که آن را برای پردازش تصاویر بزرگ یا بلادرنگ نامناسب می‌سازد.

❖ کدام روش برای نواحی کوچک‌تر سریع‌تر است؟ در نواحی مخدوش بزرگ، کدام روش کندتر می‌شود؟

بر اساس یافته‌های این تمرین، روش درونیایی (به ویژه درونیایی خطی) برای نواحی کوچک‌تر مخدوش بسیار سریع‌تر عمل می‌کند و زمان پردازش آن حدود ۱ تا ۲ دقیقه است. این سرعت بالا ناشی از ماهیت غیرتکراری و محاسبات مستقیم این روش می‌باشد که آن را به گزینه‌ای ایده‌آل برای پردازش‌های بلادرنگ و کاربردهای عملیاتی تبدیل کرده است. در مقابل، روش مبتنی بر معادلات دیفرانسیل جزئی (PDE) با افزایش اندازه ناحیه مخدوش به شدت کند می‌شود و زمان اجرای آن به ۱۵ تا ۲۰ دقیقه می‌رسد. این کندی محسوس عمدتاً به دلیل ماهیت تکراری الگوریتم (۵۰۰ تکرار)، محاسبات سنگین عملگر لاپلاسین در هر مرحله و نیاز به به‌روزرسانی تدریجی پیکسل‌هاست. جالب توجه اینکه سرعت روش درونیایی تقریباً مستقل از اندازه ناحیه تخریب شده است، در حالی که روش PDE با بزرگ شدن ناحیه مخدوش به صورت نمایی کندتر می‌شود. این ویژگی باعث می‌شود روش درونیایی به ویژه برای تصاویر با نواحی تخریب گسترده، کارآمدتر و عملی‌تر باشد، هرچند که ممکن است در برخی موارد از نظر کیفیت بازسازی اندکی ضعیف‌تر عمل کند. بنابراین انتخاب نهایی روش بهینه باید با در نظر گرفتن توازن بین سه فاکتور سرعت پردازش، کیفیت بازسازی و منابع سخت‌افزاری موجود انجام پذیرد.

Problem 3: Innovative Augmentation with FFT



در این تمرین به بررسی روش‌های نوین افزایش داده مبتنی بر تبدیل فوریه می‌پردازیم. بر اساس مطالعات اخیر می‌دانیم که اطلاعات معنایی تصاویر عمدتاً در فاز تبدیل فوریه قرار دارد، در حالی که دامنه بیشتر ویژگی‌های ظاهری را نمایندگی می‌کند. در این تحقیق، دو تکنیک اصلی را پیاده‌سازی می‌کنیم: روش AmpSwap و روش AmpMix. با استفاده از این روش‌ها، تصاویر جدیدی تولید می‌کنیم که از نظر معنایی با تصاویر اصلی یکسان هستند، اما از نظر ظاهری تنوع قابل توجهی دارند.

-پاسخ Second step:

تعمیم دامنه‌ای (Domain Generalization) یکی از چالش‌های اصلی در یادگیری ماشین است که در آن مدل روی یک یا چند دامنه‌ی منبع آموزش داده می‌شود، اما باید بدون هیچ‌گونه تنظیم مجدد (fine-tuning)، در دامنه‌ای جدید و نادیده عمل‌کرد مناسبی داشته باشد. این مسئله به دلیل تفاوت ظاهری میان تصاویر با معنای یکسان در دامنه‌های مختلف (مانند تغییر سبک، روشنایی، بافت یا نویز) پیچیده است. تکنیک افزایش تنوع مبتنی بر FFT این چالش را با تفکیک طیف دامنه (amplitude) و طیف فاز (phase) تصویر هدف قرار می‌دهد. از آن‌جا که اطلاعات معنایی تصویر عمدتاً در طیف فاز نهفته است، این روش با حفظ فاز و دستکاری دامنه، تصاویر جدیدی ایجاد می‌کند که از نظر مفهومی ثابت اما از نظر ظاهری متنوع‌اند.

دو روش اصلی استفاده‌شده در این تکنیک عبارت‌اند از:

در AmpSwap، دامنه‌ی تصویر منبع با دامنه‌ی تصویر دوم جایگزین می‌شود.

در AmpMix، دامنه‌ی هر دو تصویر به صورت وزن‌دار ترکیب می‌شود تا دامنه‌ای جدید و کنترل‌پذیر ایجاد گردد.

برای مثال، یک تصویر چهره در نور کم و نور زیاد ممکن است فاز یکسانی داشته باشد، اما دامنه‌ی آن‌ها متفاوت باشد. این تکنیک می‌تواند چنین تفاوت‌هایی را شبیه‌سازی و مدل را نسبت به آن‌ها مقاوم کند. این فرایند به مدل کمک می‌کند تا از بیش‌برازش (overfitting) به ویژگی‌های ظاهری دامنه‌ی منبع جلوگیری کرده و به جای آن، بر روی ویژگی‌های معنایی و غیروابسته به دامنه تمرکز کند. در نتیجه، مدل قادر می‌شود در مواجهه با دامنه‌های جدید نیز تعمیم‌پذیری مناسبی داشته باشد و انعطاف‌پذیری بالایی در سناریوهای واقعی مانند تشخیص چهره تحت شرایط نوری مختلف از خود نشان دهد.

- پاسخ (bonus) Third step:

پس از آزمایش الگوریتم FFT پیشنهادی Xu و همکاران که فاز تصویر را برای حفظ محتوای معنایی و دامنه را برای ایجاد تنوع ظاهری دستکاری می‌کند، دو روش جدید و الهام گرفته از سایر تبدیلات مطرح می‌شود: Wavelet Phase Fusion (WPF) و DCT Texture Transfer (DTT). در روش اول WPF، تصویر با استفاده از Wavelet Transform به مؤلفه‌های فرکانسی low-pass و high-pass تجزیه می‌شود. فاز مؤلفه low-pass از تصویر اصلی که حاوی ساختار معنایی است، حفظ می‌گردد. سپس دامنه‌ها از تصویر دوم با وزن‌های قابل تنظیم (w_1, w_2) ترکیب شده و از طریق Inverse Wavelet Transform تصویر نهایی بازسازی می‌شود. این روش با حفظ محتوای اصلی و ایجاد تنوع در جزئیات، مقاومت خوبی در برابر نویز نشان می‌دهد.

در روش دوم DTT، ابتدا تصویر به بلوک‌های DCT تقسیم شده و سپس در هر بلوک، فاز مؤلفه‌های low-frequency (مرتبط با ساختار) حفظ شده و دامنه مؤلفه‌های high-frequency (مرتبط با texture یا سبک بصری) از تصویر دوم جایگزین یا ترکیب می‌شود. با انجام Inverse DCT، تصویر نهایی با محتوای معنایی پایدار و بافت‌های جدید تولید می‌گردد. این دو تکنیک، برخلاف روش‌هایی مانند CutMix، محتوای معنایی تصویر را حفظ کرده و تنها در سطوح ظاهری آن تغییر ایجاد می‌کنند.

.....

Problem 4: Practicing the Basics of 1-D Fourier Transform

.A

3. $x(t) = \left(\frac{1}{1+t^2}\right) \cdot e^{i\pi t}$ 07
فرض می کنیم $\omega = 1$ است، داریم:

$g(t) = \frac{1}{1+t^2}$, $F(g(t)e^{j\omega_0 t}) = G(j(\omega - \omega_0))$ 08

حال تبدیل فوریه $g(t)$ تبدیل فوریه استاندارد است و داریم: 09

$G(j\omega) = \int_{-\infty}^{\infty} \frac{1}{1+t^2} e^{-j\omega t} dt = \pi e^{-|\omega|}$ 10

و می دانیم که ω_0 همان π است پس داریم:

$X(j\omega) = G(j(\omega - \pi)) = \pi e^{-|\omega - \pi|}$ 11

4. $x(t) = \left(\frac{\sin(\pi t)}{\pi t}\right) \cdot e^{-\alpha t}$, $\alpha > 0$ 12

چون تابع برای $t < 0$ رشد نمایی دارد و $u(t)$ هم در صورت این سوال نیامده که تبدیل فوریه وجود ندارد و داریم: 13

$X(j\omega) = \int_{-\infty}^{\infty} \frac{\sin(\pi t)}{(\pi t)} e^{-\alpha t} e^{-j\omega t} dt$ 14

این انتگرال به دلیل رشد نمایی در $t < 0$ همگرا نیست و تبدیل فوریه ندارد. 15

5. $\pi(n, y) = \pi(n) \pi(y)$ 16
 $F(\pi(y))$ را هم به همین صورت می توانیم بنویسیم

$F(u, v) = F\{\pi(n)\}(u) \cdot F\{\pi(y)\}(v)$

$F(\pi(n)) = \int_{-\frac{1}{2}}^{\frac{1}{2}} e^{-j\omega n} dn = \frac{e^{-j\omega/2} - e^{j\omega/2}}{-j\omega} = \frac{2 \sin(\omega/2)}{\omega}$ 17

$F(u, v) = \frac{2 \sin(u/2)}{u} \cdot \frac{2 \sin(v/2)}{v}$ 18

$$(2.) x(t) = \delta(t-3) + \delta(t+3) - 2\delta(t)$$

$$\delta(t-t_0) \Leftrightarrow e^{-j\omega t_0} \text{ (تبدیل فوری)}$$

$$\delta(t-3) \rightarrow e^{-j3\omega}$$

$$\delta(t+3) \rightarrow e^{j3\omega} \quad 12(\delta t) = 2 \frac{e^{-j0\omega}}{1} = 2$$

$$\rightarrow F(\omega) = e^{-j3\omega} + e^{j3\omega} - 2 \Rightarrow \boxed{2\cos(3\omega) - 2}$$

$$* \cos(3\omega) = \frac{e^{j3\omega} + e^{-j3\omega}}{2} \Rightarrow e^{j3\omega} + e^{-j3\omega} = 2\cos(3\omega)$$

$$(1.) x(t) = (e^{-\alpha t})u(t); \alpha > 0$$

$$X(j\omega) = \int_{-\infty}^{\infty} e^{-\alpha t} u(t) e^{-j\omega t} dt = \int_0^{\infty} e^{-\alpha t} e^{-j\omega t} dt$$

$$= \int_0^{\infty} e^{-(\alpha + j\omega)t} dt$$

از طرفی این یک انتگرال استاندارد به فرم زیر است:

$$\int_0^{\infty} e^{-kt} dt = \left[-\frac{1}{k} e^{-kt} \right]_0^{\infty} = 0 - \left(-\frac{1}{k} \right) = \frac{1}{k}$$

در اینجا $k = \alpha + j\omega$ است پس داریم:

$$(6.) f(x, y) = \pi(x) \delta(x-y)$$

$$F(u, v) = \iint \pi(x) \delta(x-y) e^{-j(ux+vy)} dx dy$$

برای حل این انتگرال از دلتا یکدما گیریم و سپس یکباریم:

$$\int \delta(x-y) e^{-jvy} dy = e^{-jvx}$$

1403

$$F(u, v) = \int \pi(x) e^{-j(ux+vx)} dx = \int_{-1/2}^{1/2} e^{-j(u+v)x} dx$$

$$\int_{-1/2}^{1/2} e^{-j(u+v)x} dx = \frac{2\sin((u+v)/2)}{u+v} = \frac{\sin((u+v)/2)}{(u+v)/2}$$

07. با. در این سؤال ابتدا 6 ویژگی گفته شده را بررسی می کنیم و بعد به سوال می پردازیم.

08. $\text{Re}\{X(j\omega)\} = 0 \Leftrightarrow$ وقتی بخش حقیقی یک سیگنال صفر باشد

09. یعنی سیگنال ما فرد است و داریم: $x(t) = -x(-t)$

10. $\text{Im}\{X(j\omega)\} = 0 \Leftrightarrow$ وقتی بخش موهومی یک سیگنال صفر باشد

یعنی سیگنال ما زوج است و داریم: $x(t) = x(-t)$

11. ویژگی سوم به این اشاره دارد که سیگنال را با یک شیفต์ بتوان به سیگنال زوج تبدیل کرد.

12. این ویژگی می گوید که سیگنال در زمان صفر برابر صفر شود.

13. یعنی مشتق سیگنال در صفر برابر صفر باشد.

14. یعنی سیگنال به صورت زمان گسسته باشد.

در ادامه 6 سیگنال a تا f را بررسی می کنیم:

15. a سیگنال فردی که با یک شیفต์ به زوج تبدیل می شود و در زمان صفر برابر صفر است \Leftrightarrow ویژگی ها $(1, 3, 4)$

16. این شکل یک سیگنال گسسته در زمان است و می توان با یک شیفต์ آن را به یک سیگنال زوج تبدیل کرد $\Leftrightarrow (3, 6)$

17. c می توان با یک شیفต์ آن را به یک سیگنال زوج تبدیل کرد $\Leftrightarrow (3)$

18. d یک سیگنال فرد است که در زمان صفر برابر صفر است $\Leftrightarrow (1, 4)$

07. e یک سیگنال زوج است که مشتق آن در صفر برابر صفر است $\Leftrightarrow (2, 5)$

08. f این سیگنال نیز یک سیگنال زوج است که مقدار در صفر نیز برابر صفر می شود $\Leftrightarrow (2, 4)$

09.

$$f = [1, -1, 2, 3, 0, 0, 0, 0] \quad (1 \text{ C } 07)$$

$$g = [1, -2, -1, 0, 0, 0, 0, 0] \quad 08$$

$$f * g = \sum_{k=0}^{n-1} f[k] \cdot g[n-k] \quad 09$$

$$n=0 \Rightarrow f[0] \cdot g[0] = 1 \quad 10$$

$$n=1 \Rightarrow f[0] \cdot g[1] + f[1] \cdot g[0] = -2 - 1 = -3$$

$$n=2 \Rightarrow f[0] \cdot g[2] + f[1] \cdot g[1] + f[2] \cdot g[0] = 3$$

$$n=3 \Rightarrow f[0] \cdot g[3] + f[1] \cdot g[2] + \dots + f[3] \cdot g[0] = 0 \quad 11$$

$$n=4 \Rightarrow f[0] \cdot g[4] + \dots + f[4] \cdot g[0] = -8 \quad 12$$

$$n=5 \Rightarrow f[0] \cdot g[5] + \dots + f[5] \cdot g[0] = -3 \quad 13$$

$$[1, -3, 3, 0, -8, -3] \rightarrow \text{بخش اصلی و غیر صفر کانفلوئنس} \quad 14$$

$$\hookrightarrow [1, -3, 3, 0, -8, -3, 0, 0] \rightarrow \text{یا پدینگ 0 به طول 8}$$

2. برای این بخش سیگنال های f و g را که در قسمت قبل ¹⁵

پیدا کردیم، با استفاده از DFT تبدیل فوریه هر کدام را حساب می

کنیم و بعد با ضرب نقطه ای DFT های f و g حاصل را بدست می ¹⁶

آوریم و بعد با IDFT نتیجه را در حوزه زمان محاسبه می کنیم و ¹⁷

دریقا با نتیجه کانفلوئنس مرحله قبل یعنی $[1, -3, 3, 0, -8, -3, 0, 0]$

تکسان حاصل می شود.

15 پنجشنبه 12 شهریور 1402

$$\text{DFT: } X[K] = \sum_{n=0}^{N-1} x[n] e^{-j \frac{2\pi}{N} Kn}, \quad K = 0, 1, \dots, N-1$$

$$\text{IDFT: } x[n] = \frac{1}{N} \sum_{K=0}^{N-1} X[K] e^{j \frac{2\pi}{N} Kn}$$

$$F[K] = \sum_{n=0}^7 f[n] e^{-j \frac{2\pi}{8} Kn} = 1 \cdot e^{-j \frac{\pi}{4} K \cdot 0} + (-1) \cdot e^{-j \frac{\pi}{4} K \cdot 1} +$$

$$2 \cdot e^{-j \frac{\pi}{4} K \cdot 2} + 3 \cdot e^{-j \frac{\pi}{4} K \cdot 3}$$

$$F[0] = 1 - 1 + 2 + 3 = 5$$

$$F[1] = 1 - e^{-j \frac{\pi}{4}} + 2e^{-j \frac{\pi}{2}} + 3e^{-j \frac{3\pi}{4}}$$

$$= (1 - 2\sqrt{2}) + j(-2 - \sqrt{2})$$

$$F[2] = 1 - e^{-j \frac{\pi}{2}} + 2e^{-j \pi} + 3e^{-j \frac{3\pi}{2}} = -1 + 4j$$

$$F[3] = 1 - e^{-j \frac{3\pi}{4}} + 2e^{-j \frac{6\pi}{4}} + 3e^{-j \frac{9\pi}{4}} = (1 + 2\sqrt{2}) + j(2 - \sqrt{2})$$

26 بهمن 15 شهریور 1402 Feb Fri 14

$$F[4] = 1 - 1 + 2(1) + 3(-1) = -1$$

$$F[5] = F[3]^* = (1 + 2\sqrt{2}) - j(2 - \sqrt{2})$$

$$F[6] = F[2]^* = -1 - 4j$$

$$F[7] = F[1]^* = (1 - 2\sqrt{2}) - j(-2 - \sqrt{2})$$

به همین صورت DFT مربوط به g را هم در ادامه محاسبه می کنیم:

سه شنبه ۱۲ شهریور ۱۳۹۴ 2025

$$G[k] = 1 - 2e^{-j\frac{\pi}{4}k} - e^{-j\frac{\pi}{2}k}$$

$$G[0] = 1 - 2 - 1 = -2$$

$$G[1] = (1 - \sqrt{2}) + j(1 + \sqrt{2})$$

$$G[2] = 1 - 2(-j) - (-1) = 2 + 2j$$

$$G[3] = (1 + \sqrt{2}) + j(1 - \sqrt{2})$$

$$G[4] = 1 - 2(1) - 1 = -2$$

$$G[5] = G[3]^* = (1 + \sqrt{2}) - j(1 - \sqrt{2})$$

$$G[6] = G[2]^* = 2 - 2j$$

$$G[7] = G[1]^* = (1 - \sqrt{2}) - j(1 + \sqrt{2})$$

13 حالا در ادامه ضرب نقطه DFT مربوط به $G[k]$ و $F[k]$ را حساب

$$H[k] = F[k] \cdot G[k]$$

می کنیم:

$$H[0] = 5 \cdot (-2) = -10$$

$$H[1] = [(1 - 2\sqrt{2}) + j(-2 - \sqrt{2})] \cdot [(1 - \sqrt{2}) + j(1 + \sqrt{2})]$$

و به همین صورت تا $H[7]$ را محاسبه می کنیم و بعد حالتی آرایه

$$h[n] = \frac{1}{8} \sum_{k=0}^7 H[k] e^{j\frac{\pi}{4}kn}$$

DFT معکوس را حساب می کنیم:

17 که حاصل نهایی برابر همان مقادیر کانولوشن خطی می آید

می شود و داریم:

$$f * g = [1, -3, 3, 0, -8, -3, 0, 0]$$

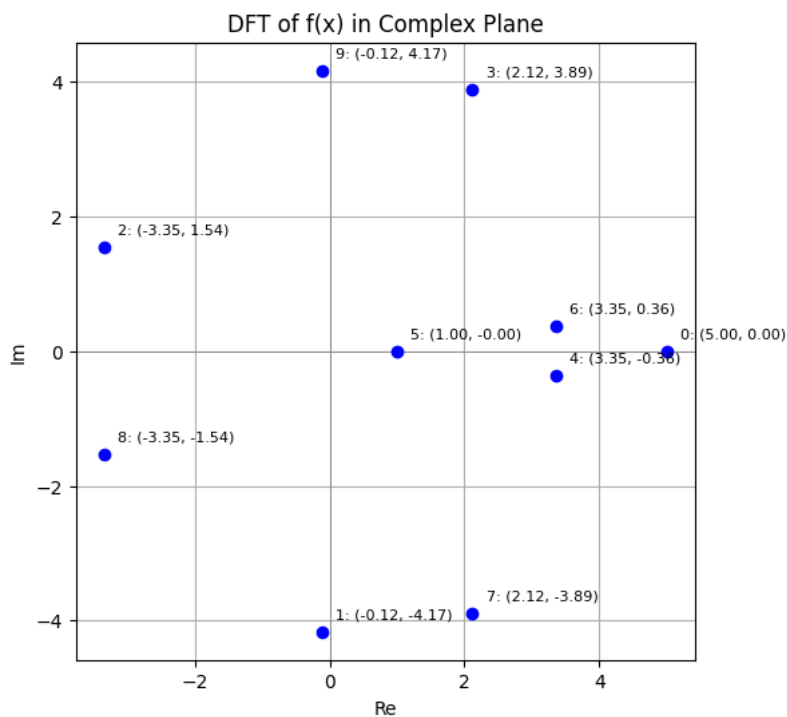
4. در این قسمت نیز باید انرژی سیگنال f و g را حساب کنیم:

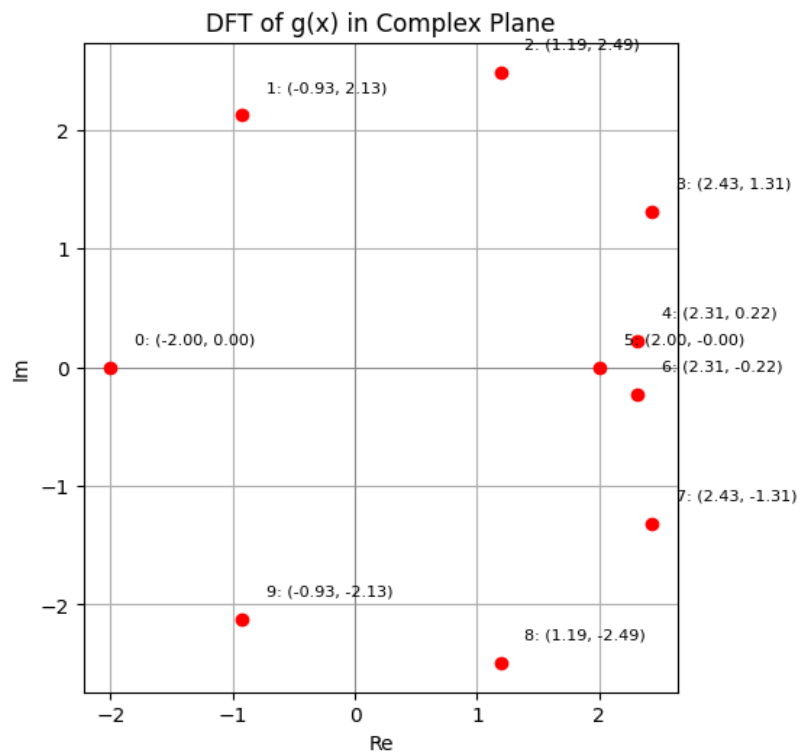
$$\sum_n |f(n)|^2 = 1^2 + (-1)^2 + (2)^2 + (3)^2 = 15$$

$$\sum_n |g(n)|^2 = 1^2 + (-2)^2 + (-1)^2 = 6$$

طبق قضیه پارسوال انرژی در حوزه زمان برابر با انرژی در حوزه فرکانس است و این انرژی باید در حوزه فرکانس نیز حفظ شده باشد.

3. این پارت را به دلیل دشواری رسم دستی در پایتون رسم کردیم:





.....

Problem 5: Some Explanatory Questions

A.

چرخش یک تصویر در حوزه فضایی تأثیر مستقیمی بر نمایش آن در حوزه فوریه (Fourier domain) دارد. به طور مشخص، یکی از خواص مهم تبدیل فوریه دوبعدی بیان می‌کند که چرخش تصویر در حوزه فضایی منجر به چرخش طیف فوریه تصویر به همان اندازه در حوزه فرکانسی می‌شود. این ویژگی که به عنوان ویژگی چرخش (Rotation Property) شناخته می‌شود و بیانگر این نکته کلیدی است که ساختارهای فرکانسی تصویر نسبت به چرخش در فضای مکانی حساس هستند و این چرخش به طور منظم در فضای فرکانسی بازتاب می‌یابد.

از منظر magnitude و phase تبدیل فوریه، چرخش تصویر موجب چرخش کل طیف فرکانسی (شامل هر دو مؤلفه دامنه و فاز) به زاویه‌ای معادل می‌گردد، بدون آنکه تغییری در مقادیر آن‌ها ایجاد شود. در واقع، توزیع انرژی فرکانسی تصویر، که در تابع دامنه نمایش داده می‌شود، به همان شکل باقی می‌ماند و صرفاً جهت‌گیری آن در صفحه فرکانسی تغییر می‌کند. به همین ترتیب، نقشه فاز نیز که مسئول ثبت اطلاعات ساختاری و مکانی تصویر است، با حفظ مقادیر اولیه خود، به همان میزان چرخش پیدا می‌کند. این خاصیت نقش مهمی در بسیاری از کاربردهای پردازش تصویر دارد، از جمله در image registration یا image alignment که در آن نیاز است تصاویر با زوایای مختلف به صورت دقیق بر یکدیگر منطبق شوند. استفاده از ویژگی چرخش تبدیل فوریه، امکان مقایسه و هم‌ترازسازی تصاویر با زوایای متفاوت را فراهم می‌سازد، چرا که چرخش تصویر منجر به تغییری سازمان‌یافته و قابل پیش‌بینی در حوزه فرکانسی می‌شود.

اگر تصویر اصلی را با $f(x,y)$ نمایش دهیم و تصویر چرخانده‌شده به اندازه زاویه θ را با $f_\theta(x,y)$ نشان دهیم، رابطه بین تبدیل فوریه این دو به صورت زیر بیان می‌شود:

$$\mathcal{F}[f_\theta(x,y)](u,v) = F_\theta(u,v) = F(u',v')$$

که در آن:

$$(u',v') = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix}$$

یعنی طیف فرکانسی تصویر چرخانده‌شده برابر است با طیف تصویر اصلی، که به همان اندازه در صفحه فرکانسی چرخانده شده است. در نتیجه:

دامنه تبدیل فوریه (Magnitude) حفظ می‌شود و صرفاً جهت آن تغییر می‌کند:

$$|F_\theta(u,v)| = |F(u',v')|$$

فاز تبدیل فوریه نیز چرخش می‌یابد، اما اطلاعات ساختاری تصویر همچنان در آن باقی می‌ماند:

$$\angle F_{\theta}(u, v) = \angle F(u', v')$$

این روابط به وضوح نشان می‌دهند که چرخش تصویر در فضای مکانی منجر به چرخشی متناظر و منظم در فضای فرکانسی می‌شود، بدون آن که تغییری در میزان انرژی طیفی یا توزیع اطلاعات ساختاری تصویر ایجاد شود. این ویژگی، مبنای کلیدی برای بسیاری از تکنیک‌های پیشرفته در تحلیل ساختار تصویر، شناسایی الگوهای چرخیده، و الگوریتم‌های هم‌راستاسازی (registration) در پردازش تصویر دیجیتال به شمار می‌رود، چراکه امکان مقایسه، تطبیق و ترکیب تصاویر با زوایای مختلف را در حوزه فرکانسی فراهم می‌سازد.

B.

تبدیل کسینوسی گسسته یا DCT یکی از پرکاربردترین تبدیل‌ها در فشرده‌سازی تصویر است، به خصوص در الگوریتم‌های معروفی مثل JPEG. این تبدیل مزایای مختلفی برای فشرده‌سازی نسبت به DFT دارند و در ادامه به طور کامل آنها را بررسی می‌کنیم:

تراکم انرژی بالا (Energy Compaction):

یکی از مهم‌ترین ویژگی‌های DCT این است که می‌تواند بیشتر انرژی سیگنال (یعنی اطلاعات تصویری) را در تعداد کمی از ضرایب، به ویژه ضرایب مربوط به فرکانس‌های پایین، متمرکز کند. این موضوع باعث می‌شود اطلاعات اصلی تصویر در چند مقدار محدود حفظ شده و سایر ضرایب را بتوان حذف یا با دقت پایین‌تری ذخیره کرد، بدون افت محسوس در کیفیت.

کاهش همبستگی بین پیکسل‌ها:

در تصاویر طبیعی، پیکسل‌های مجاور معمولاً با هم همبستگی دارند. DCT، این همبستگی را کاهش می‌دهد و باعث می‌شود ضرایب خروجی مستقل‌تر از هم باشند. این ویژگی کوانتیزاسیون و فشرده‌سازی را بسیار مؤثرتر می‌سازد.

سازگاری با ویژگی‌های سیستم بینایی انسان:

چشم انسان نسبت به جزئیات کم‌فرکانس حساس‌تر است تا فرکانس‌های بالا. DCT، با جدا کردن مؤلفه‌های فرکانسی، این امکان را می‌دهد که ضرایب با فرکانس بالا (که اهمیت بصری کمتری دارند) با دقت پایین‌تری ذخیره شوند، بدون اینکه کیفیت تصویر به طور قابل توجهی کاهش یابد.

خروجی حقیقی (Real-Valued Output):

بر خلاف DFT که خروجی آن شامل اعداد مختلط (complex) است، DCT فقط مقادیر حقیقی تولید می‌کند. این موضوع هم نیاز به حافظه را کاهش می‌دهد و هم زمان پردازش را کمتر می‌کند.

کاهش آرتیفکت‌های مرزی (Boundary Artifacts):

DFT فرض می‌کند تصویر ورودی تناوبی است، که این فرض باعث ایجاد ناپیوستگی در مرز بلوک‌ها و

آرتیفکت‌های قابل مشاهده می‌شود DCT. به جای آن، سیگنال را نسبت به مرکز متقارن فرض می‌کند که باعث مرزهای نرم‌تر و تصویر نهایی طبیعی‌تر می‌شود.

ساده‌سازی مرحله کوانتیزاسیون (Quantization):

ضرایب DCT، به خصوص در فرکانس‌های بالا، را می‌توان به راحتی کوانتیزه کرد (گرد کرد) بدون اینکه کیفیت تصویر افت شدیدی پیدا کند. این ویژگی، فشردگی با اتلاف (lossy) را کارآمدتر می‌سازد

امکان پیاده‌سازی سریع (Fast Algorithms):

برای DCT الگوریتم‌های سریع مشابه FFT در DFT وجود دارد که اجرای DCT را بر روی بلوک‌های تصویر بسیار سریع و بهینه می‌سازد.

در مجموع، استفاده از DCT در فشردگی تصویر، به ویژه در استاندارد JPEG، به دلیل توانایی بالا در تراکم انرژی، حذف همبستگی، سادگی محاسبات و سازگاری با سیستم بینایی انسان است. این مزایا باعث شده‌اند که DCT نسبت به DFT گزینه بسیار مناسب‌تری برای فشردگی تصویر باشد.

C.

در پردازش تصویر، عملیات بلور (Blurring) در حوزه مکانی به صورت کانولوشن (Convolution) تصویر با یک هسته صاف‌کننده (مانند فیلتر گاوسی) انجام می‌شود. بر اساس قضیه کانولوشن، کانولوشن در حوزه مکانی معادل با ضرب نقطه‌ای (element-wise multiplication) در حوزه فرکانسی است. این قضیه به این معنی است که هنگامی که یک تصویر با یک فیلتر صاف‌کننده در حوزه مکانی کانولوشن می‌شود، در حوزه فرکانسی همان عملیات ضرب بین تبدیل فوری تصویر و تبدیل فوری فیلتر اتفاق می‌افتد. به طور ریاضی، این ارتباط به شکل زیر بیان می‌شود:

$$H(u, v) \cdot F(u, v) \xrightarrow{\text{Fourier}} h(x, y) * f(x, y)$$

که در آن $f(x, y)$ تصویر اصلی، $h(x, y)$ فیلتر، $F(u, v)$ تبدیل فوری تصویر و $H(u, v)$ تبدیل فوری فیلتر هستند. در نتیجه، بلور کردن تصویر در حوزه مکانی منجر به ضرب تبدیل فوری تصویر در یک فیلتر پایین‌گذر (Low-Pass Filter) در حوزه فرکانسی می‌شود. این فیلترهای پایین‌گذر مؤلفه‌های فرکانسی پایین را حفظ می‌کنند و مؤلفه‌های فرکانسی بالا، که معمولاً شامل جزئیات و لبه‌های تصویر هستند، را تضعیف یا حذف می‌کنند. از آنجایی که فرکانس‌های بالا نمایانگر تغییرات سریع در شدت روشنایی تصویر هستند، حذف آن‌ها باعث کاهش جزئیات و صاف شدن نواحی تصویر می‌شود. بنابراین، بلور کردن تصویر در حوزه مکانی منجر به حذف اطلاعات فرکانسی بالا در حوزه فرکانسی می‌شود، که دقیقاً همان عملکرد یک فیلتر پایین‌گذر است.

تأثیر این فرآیند بر تبدیل فوری تصویر به دو بخش اصلی تقسیم می‌شود: اندازه (Magnitude) و فاز (Phase). اندازه تبدیل فوری، میزان انرژی هر مؤلفه فرکانسی را نشان می‌دهد. در اثر اعمال فیلتر پایین‌گذر، مقادیر مربوط به فرکانس‌های بالا در طیف فرکانسی کاهش می‌یابند، در حالی که مؤلفه‌های مربوط به فرکانس‌های پایین حفظ می‌شوند. در نتیجه، توزیع انرژی در حوزه فرکانسی به ناحیه مرکزی (فرکانس‌های پایین) محدود می‌شود و انرژی

موجود در فرکانس‌های بالا به شدت کاهش می‌یابد. این تمرکز انرژی در فرکانس‌های پایین، بازتابی از فرآیند بلورشدگی در تصویر است.

در مورد فاز، باید توجه داشت که فاز تبدیل فوریه اطلاعات مکانی مربوط به ساختار تصویر را در خود دارد. فیلتر پایین‌گذر با حذف فرکانس‌های بالا، دقت مکانی موجود در تصویر را کاهش می‌دهد. اگرچه فاز تا حدی حفظ می‌شود، اما حذف مؤلفه‌های فرکانسی بالا باعث از دست رفتن جزئیات دقیق مکان‌یابی اجزای تصویر مانند لبه‌ها و بافت‌ها می‌شود. بنابراین، ساختار کلی تصویر تا حدی حفظ می‌شود، اما جزئیات ریز آن از بین می‌رود.

در مجموع، عملیات بلور در حوزه مکانی با ضرب طیف تصویر در یک فیلتر پایین‌گذر در حوزه فرکانسی معادل است. این فرآیند باعث کاهش اندازه مؤلفه‌های فرکانس بالا در تبدیل فوریه شده و فاز را نیز تا حدودی تحت تأثیر قرار می‌دهد، به گونه‌ای که جزئیات مکانی تضعیف می‌شوند، اما ساختار کلی تصویر باقی می‌ماند. این ویژگی‌ها باعث می‌شود بلور یکی از ابزارهای مهم در بیش‌پردازش تصاویر و کاهش نویز محسوب شود.

D

در پردازش تصویر، فیلتر میان‌گذر (Band-Pass Filter) به منظور نگه‌داشتن مؤلفه‌های فرکانسی خاص از یک تصویر و سرکوب مؤلفه‌های دیگر استفاده می‌شود. این فیلتر معمولاً به عنوان ترکیبی از یک فیلتر پایین‌گذر (Low-Pass Filter) و یک فیلتر بالاگذر (High-Pass Filter) عمل می‌کند. در اینجا به طور کامل توضیح خواهیم داد که این فیلتر چگونه در حوزه فرکانسی کار می‌کند و چه تأثیری بر ظاهر تصویر در حوزه مکانی خواهد داشت.

فیلتر بند-پس با تعریف دو مقدار آستانه (یک فرکانس پایین و یک فرکانس بالا) تنها به مؤلفه‌های فرکانسی بین این دو مقدار اجازه عبور می‌دهد. به عبارت دیگر، این فیلتر فرکانس‌هایی را که بین دو مقدار low و high قرار دارند حفظ کرده و سایر فرکانس‌ها را تضعیف یا حذف می‌کند. به صورت ریاضی، فیلتر بند-پس را می‌توان به شکل زیر مدل کرد:

$$H_{HP}(u, v) - H_{LP}(u, v) = H_{BP}(u, v)$$

$$H_{BP}(u, v) = \begin{cases} 1 & \text{if } D_1 \leq \sqrt{u^2 + v^2} \leq D_2, \\ 0 & \text{otherwise.} \end{cases}$$

$H_{LP}(u, v)$ تبدیل فوریه فیلتر پایین‌گذر است که فرکانس‌های پایین را عبور می‌دهد.

$H_{HP}(u, v)$ تبدیل فوریه فیلتر بالاگذر است که فرکانس‌های بالا را عبور می‌دهد.

$H_{BP}(u, v)$ نتیجه حاصل از ترکیب این دو است و یک باند خاص از فرکانس‌ها را عبور می‌دهد.

این فیلتر در حوزه فرکانسی معمولاً به صورت یک ناحیه حلقوی در طیف فرکانسی نمایش داده می‌شود که فقط انرژی فرکانسی موجود در این ناحیه را حفظ می‌کند.

وقتی فیلتر بند-پس در حوزه فرکانسی روی تصویر اعمال می‌شود، اثرات آن در حوزه مکانی به صورت تغییر در ظاهر تصویر دیده می‌شود. این تغییرات شامل موارد زیر است:

تقویت جزئیات میانه‌بردار: چون این فیلتر فرکانس‌های میانه را عبور می‌دهد، جزئیاتی مثل لبه‌ها، خطوط و بافت‌هایی که در این بازه فرکانسی قرار دارند، واضح‌تر و برجسته‌تر دیده می‌شوند.

کاهش نویز و حذف جزئیات بسیار ریز یا خیلی بزرگ: فرکانس‌های بالا که معمولاً شامل نویز و جزئیات بسیار ریز هستند حذف می‌شوند، در حالی که فرکانس‌های بسیار پایین مثل نواحی یکنواخت یا گرادیان‌های ملایم نیز سرکوب می‌شوند. نتیجه آن یک تصویر با ظاهری تمیزتر و متمرکز بر ساختارهای اصلی است.

افزایش کنتراست موضعی: نواحی‌ای از تصویر که دارای تغییرات روشنایی در بازه فرکانسی میانه هستند، کنتراست بیشتری پیدا می‌کنند که می‌تواند تشخیص ویژگی‌ها را آسان‌تر کند.

ایجاد جلوه بصری خاص: در برخی موارد، اعمال فیلتر بند-پس می‌تواند تصویری با ظاهر هنری یا استایل خاص ایجاد کند، چراکه فقط بخش خاصی از اطلاعات تصویری باقی مانده است.

فیلتر بند-پس یکی از ابزارهای مهم در پردازش تصویر است که با حذف فرکانس‌های بسیار پایین و بسیار بالا و تمرکز بر فرکانس‌های میانه، می‌تواند ویژگی‌های خاص تصویر را تقویت کرده و نویز را کاهش دهد. تأثیر نهایی این فیلتر در حوزه مکانی باعث افزایش وضوح لبه‌ها و بافت‌ها، کاهش نویز و بهبود نمایش جزئیات کلیدی تصویر می‌شود. این ویژگی‌ها باعث می‌شود از فیلتر بند-پس در کاربردهایی مثل تشخیص الگو، تحلیل بافت و فشرده‌سازی تصویر استفاده گسترده‌ای شود.

E

الیاسینگ (Aliasing) پدیده‌ای است که در فرآیند نمونه‌برداری از سیگنال‌های آنالوگ به دیجیتال رخ می‌دهد، و یکی از چالش‌های کلیدی در پردازش سیگنال محسوب می‌شود. این پدیده زمانی اتفاق می‌افتد که فرکانس نمونه‌برداری به اندازه‌ای کافی نباشد تا تمام اطلاعات فرکانسی موجود در سیگنال آنالوگ را به درستی ثبت کند. به زبان ساده، سیگنال‌هایی با فرکانس بالا اشتباهاً به شکل سیگنال‌هایی با فرکانس پایین‌تر ظاهر می‌شوند، و این منجر به اعوجاج در بازسازی سیگنال اصلی می‌گردد. قضیه‌ی نایکوئیست-شانون بیان می‌کند که برای جلوگیری از الیاسینگ، باید نرخ نمونه‌برداری حداقل دو برابر بزرگ‌ترین فرکانس موجود در سیگنال آنالوگ باشد و وقتی نرخ نمونه‌برداری پایین‌تر از نرخ نایکوئیست باشد، فرکانس‌های بالاتر از نصف نرخ نمونه‌برداری (که به آن فرکانس نایکوئیست نیز گفته می‌شود) به اشتباه به محدوده‌ی پایین‌تر انتقال می‌یابند. به بیان دقیق‌تر، در حوزه‌ی فرکانس، نمونه‌برداری باعث تکرار طیف سیگنال اصلی در بازه‌های منظم حول مضرب‌های نرخ نمونه‌برداری می‌شود. اگر این تکرارها با هم همپوشانی پیدا کنند، مؤلفه‌های فرکانسی سیگنال‌ها با هم مخلوط می‌شوند و نتیجه‌ی آن سیگنالی است که از نظر اطلاعات فرکانسی غیرقابل اعتماد است.

در تئوری برای جلوگیری از الیاسینگ دو روش اصلی وجود دارد. اولین روش، استفاده از نرخ نمونه‌برداری بالاتر از نرخ نایکوئیست است. بر اساس قضیه‌ی نایکوئیست-شانون، برای جلوگیری از الیاسینگ، نرخ نمونه‌برداری باید

حداقل دو برابر بزرگ‌ترین فرکانس سیگنال آنالوگ باشد. دومین روش، استفاده از فیلتر پایین‌گذر (Anti-) است که پیش از انجام نمونه‌برداری، فرکانس‌های بالاتر از نصف نرخ نمونه‌برداری را حذف می‌کند. این فیلتر باعث می‌شود که فقط فرکانس‌هایی که می‌توانند به درستی نمونه‌برداری شوند، به سیگنال دیجیتال وارد شوند و از ایجاد الیاسینگ جلوگیری می‌کند.

با این حال، این راه‌حل‌ها در عمل همیشه به طور کامل قابل پیاده‌سازی نیستند. در دنیای واقعی، مشکلاتی مانند نویز در سیگنال آنالوگ، غیریاده‌آل بودن فیلترهای پایین‌گذر، اختلالات در نرخ نمونه‌برداری به دلیل مشکلات سخت‌افزاری و حتی محدودیت‌های پردازشی در سیستم‌های دیجیتال باعث می‌شود که الیاسینگ ممکن است حتی زمانی که اصول تئوریک رعایت شده باشد، همچنان رخ دهد.

- برخی از علل رخ دادن پدیده الیاسینگ را در ادامه بررسی خواهیم کرد:

نویز در سیگنال آنالوگ: سیگنال‌های طبیعی و واقعی همواره دارای نویز هستند. این نویز می‌تواند شامل مؤلفه‌هایی با فرکانس بسیار بالا باشد که در حوزه فرکانس، خارج از محدوده‌ی مورد انتظار قرار دارند. اگر فیلتر پایین‌گذر به درستی این مؤلفه‌های نویزی را حذف نکند، آن‌ها در فرآیند نمونه‌برداری باعث بروز الیاسینگ خواهند شد.

غیریاده‌آل بودن فیلترهای پایین‌گذر: هیچ فیلتر آنالوگی ایده‌آل نیست. همه‌ی فیلترهای فیزیکی دارای شیب محدود هستند و یک ناحیه گذار (transition band) بین ناحیه عبور (passband) و ناحیه توقف (stopband) دارند. در این ناحیه، فیلتر نمی‌تواند به طور کامل فرکانس‌ها را رد یا عبور دهد. اگر این ناحیه به خوبی طراحی نشود، فرکانس‌هایی که باید حذف شوند ممکن است از فیلتر عبور کنند و در نمونه‌برداری ظاهر شوند.

نویز سخت‌افزاری و اختلال در نرخ نمونه‌برداری: در عمل، نرخ نمونه‌برداری سیستم‌ها ممکن است دقیق نباشد. نویز الکتریکی، نوسانات کلاک یا جیتر (jitter) در سیستم‌های نمونه‌بردار باعث می‌شود تا فواصل بین نمونه‌ها دقیقاً برابر نباشد. این اختلالات می‌توانند منجر به الیاسینگ‌های پیش‌بینی‌نشده شوند، حتی اگر نرخ میانگین نمونه‌برداری بالاتر از نرخ نایکوئیست باشد.

پدیده‌ی آندرسمپلینگ در منابع چندگانه یا داده‌های پیچیده: در سناریوهایی مثل سیستم‌های چندحسگری یا داده‌های ترکیبی، ممکن است بخش‌هایی از داده با نرخ‌های مختلف نمونه‌برداری شده باشند. ناسازگاری بین این نرخ‌ها می‌تواند باعث تداخل فرکانسی شود، که به نوعی الیاسینگ میان‌منبعی منجر خواهد شد.

محدودیت‌های پردازشی یا حافظه‌ای سیستم: در برخی کاربردها، به دلیل محدودیت در توان پردازش یا حجم داده، نرخ نمونه‌برداری آگاهانه پایین نگه داشته می‌شود. در این موارد، خطر الیاسینگ آگاهانه پذیرفته می‌شود، و گاهی با تکنیک‌هایی مثل میانگین‌گیری یا downsampling هوشمندانه سعی در کاهش اثر آن می‌شود.

الیاسینگ می‌تواند اثرات منفی قابل توجهی بر تحلیل داده‌ها و بازسازی سیگنال‌ها داشته باشد. اولاً، باعث کاهش کیفیت سیگنال بازسازی‌شده می‌شود، به طوری که سیگنال دیجیتال دیگر تصویر دقیقی از سیگنال اصلی نیست. این مسئله می‌تواند باعث ایجاد اعوجاج‌های غیرقابل حذف شود؛ چرا که پس از وقوع الیاسینگ، بازسازی دقیق سیگنال اصلی عملاً امکان‌پذیر نخواهد بود. علاوه بر این، الیاسینگ می‌تواند باعث اشتباه در تحلیل فرکانسی یا ویژگی‌های زمانی داده‌ها شود. به ویژه در زمینه‌هایی مانند پردازش سیگنال‌های EEG، ECG یا تصاویر پزشکی، الیاسینگ می‌تواند منجر به تشخیص نادرست یا گمراه‌کننده‌ای شود که اثرات جدی در فرآیندهای تشخیصی و درمانی خواهد داشت.

الیاسینگ پدیده‌ای جدی و خطرناک در پردازش سیگنال است که در صورت عدم کنترل می‌تواند منجر به اعوجاج شدید، از بین رفتن اطلاعات و تحلیل اشتباه شود. اگرچه نظریه نایکوئیست راهکارهایی روشن ارائه داده است، اما در عمل با نویز، خطای فیلتر، اختلالات سخت‌افزاری و شرایط محیطی مواجه هستیم که می‌توانند باعث بروز الیاسینگ شوند. برای کاهش این خطر، باید از فیلترهای مناسب، نرخ نمونه‌برداری دقیق، کالیبراسیون سخت‌افزار و تحلیل مناسب نویز استفاده کرد.

F

قضیه شیفت تبدیل فوریه یکی از ویژگی‌های مهم در پردازش سیگنال‌ها است که می‌گوید اگر سیگنالی در دامنه فضایی (یا زمانی) جابجا شود، این جابجایی به یک تغییر فازی خطی در دامنه فوریه تبدیل می‌شود. به عبارت دیگر، اگر یک سیگنال به اندازه‌ای مشخص در فضای اصلی جابجا شود، این جابجایی در دامنه فوریه به صورت تغییر فاز برای اجزاء فرکانسی سیگنال ظاهر می‌شود، بدون اینکه مقدار آن‌ها تغییر کند.

به صورت ریاضی، اگر سیگنال $f(x)$ به اندازه a جابجا شود، تبدیل فوریه سیگنال جابجا شده به شکل زیر خواهد بود:

$$\mathcal{F}\{f(x - a, y - b)\} = F(u, v) \cdot e^{-j2\pi(au+bv)}$$

$f(x - a, y - b)$ سیگنال جابجا شده در دامنه فضایی است.

$F(u, v)$ تبدیل فوریه سیگنال اصلی است.

$e^{-j2\pi(au+bv)}$ تغییر فاز در دامنه فوریه است.

در ادامه دو کاربرد عمده از این ویژگی در پردازش تصویر ثبت تصویر و هماهنگ‌سازی تصویر را بررسی خواهیم کرد:

▪ ثبت تصویر (Image Registration) :

ثبت تصویر فرآیند هم‌راستا کردن دو یا چند تصویر است که ممکن است از شرایط مختلف یا زمان‌های متفاوت گرفته شده باشند. این شرایط می‌توانند شامل تغییرات در موقعیت دوربین، زاویه دید، نورپردازی و حتی تغییرات اندکی در شیئی که تصویر گرفته شده است باشند. هدف از ثبت تصویر این است که تصاویر مختلف به گونه‌ای تنظیم شوند که اجزاء مشترک آن‌ها به درستی هم‌راستا شوند.

توضیح روش کار: اگر یکی از تصاویر نسبت به دیگری جابجا شده باشد، طبق قضیه شیفت، جابجایی فضایی در دامنه زمانی/فضایی معادل یک تغییر فاز خطی در دامنه فرکانسی است. برای شناسایی میزان جابجایی، ابتدا تبدیل فوریه هر دو تصویر محاسبه می‌شود. سپس تفاوت تغییر فاز در دامنه فوریه به ما نشان می‌دهد که تصویر جابجا شده است و به‌طور دقیق می‌توان میزان جابجایی را تعیین کرد. با استفاده از این اطلاعات، می‌توان تصویر دوم را به گونه‌ای اصلاح کرد که با تصویر اول هم‌راستا شود.

کاربردها: این روش به‌ویژه در کاربردهایی مانند تصویربرداری پزشکی برای هم‌راستا کردن تصاویر اسکن‌شده از بدن بیمار مانند MRI یا CT و سنجش از راه دور برای هم‌راستا کردن تصاویر ماهواره‌ای یا هوایی در بازه‌های زمانی مختلف بسیار مفید است. در این کاربردها، ممکن است تصاویر از زوایای مختلف یا در زمان‌های مختلف گرفته شده باشند و با استفاده از تحلیل فوری، جابجایی‌های فضایی دقیقاً مشخص می‌شوند و هم‌راستایی دقیق صورت می‌گیرد.

▪ هماهنگ‌سازی تصویر (Image Alignment) :

در هماهنگ‌سازی تصویر، هدف این است که چندین تصویر را از لحاظ هندسی با یکدیگر هماهنگ کنیم، به‌طوری‌که تفاوت‌های موجود در جابجایی‌ها، چرخش‌ها یا تغییر مقیاس‌ها در تصاویر اصلاح شوند.

توضیح روش کار: زمانی که تغییرات هندسی در تصاویر به‌طور عمده جابجایی باشند، قضیه شیفت تبدیل فوری کمک می‌کند تا به‌طور سریع‌تر و دقیق‌تر این تغییرات اصلاح شود. با محاسبه تفاوت‌های فاز در دامنه فوری تصاویر، می‌توان جابجایی‌ها را شناسایی و به‌طور دقیق به تصاویر اعمال کرد. در واقع، پس از محاسبه تفاوت فاز بین تصاویر، می‌توانیم تصاویر را با توجه به تغییرات فاز در دامنه فرکانسی به هم‌راستایی برسانیم.

کاربردها: این تکنیک در فرآیندهایی مانند ساخت پانوراما که در آن تصاویر مختلف به یکدیگر متصل می‌شوند تا یک تصویر وسیع‌تر ایجاد کنند و یا بازسازی سه‌بعدی از تصاویر دوبعدی به کار می‌رود. در این فرآیندها، هماهنگ‌سازی دقیق تصاویر باعث می‌شود که کیفیت نهایی تصویر بدون هیچ گونه اعوجاج یا بهم‌ریختگی باشد.

قضیه شیفت تبدیل فوری این امکان را می‌دهد که به‌جای انجام محاسبات پیچیده در دامنه فضایی که ممکن است برای تصاویر بزرگ زمان‌بر و هزینه‌بر باشد، در دامنه فرکانسی عملیات مورد نظر را انجام دهیم. تغییرات فاز در دامنه فوری به راحتی قابل محاسبه است و به این ترتیب می‌توانیم تغییرات جابجایی یا چرخش را تنها با اعمال تغییر فاز در دامنه فرکانسی انجام دهیم. این ویژگی به‌ویژه در پردازش‌های زمان واقعی و سیستم‌های پردازش تصویر با حجم بالا که نیاز به سرعت بالا دارند، بسیار کارآمد است و می‌تواند زمان پردازش را به شدت کاهش دهد. این امر در سیستم‌های بینایی ماشین و پردازش تصویر در زمان واقعی که نیاز به هم‌راستایی سریع تصاویر دارند، بسیار حیاتی است.

در نتیجه قضیه شیفت تبدیل فوری ابزاری قدرتمند در پردازش سیگنال‌ها و تصاویر است که امکان تحلیل و اصلاح جابجایی‌ها و تغییرات هندسی را در دامنه فرکانسی فراهم می‌آورد. این ویژگی در کاربردهایی مانند ثبت تصویر و هماهنگ‌سازی تصاویر بسیار حیاتی است، زیرا با استفاده از تحلیل فاز در دامنه فوری می‌توان جابجایی‌های فضایی را شناسایی و اصلاح کرد. توانایی انجام این عملیات در دامنه فرکانسی نه تنها دقت پردازش را افزایش می‌دهد، بلکه باعث می‌شود که پردازش‌ها سریع‌تر و بهینه‌تر انجام شوند، به‌ویژه در سیستم‌های پیچیده و زمان واقعی. این ویژگی‌های کاربردی قضیه شیفت تبدیل فوری را به ابزاری اساسی در پردازش تصویر و بینایی ماشین تبدیل کرده است.

G

در بازسازی تصویر تار شده با استفاده از تبدیل فوری، ابتدا مدل (PSF یا Point Spread Function) را به صورت کانولوشن تصویر اصلی $f(x,y)$ با تابع $h(x,y)$ در نظر می گیریم که همراه با نویز $n(x,y)$ منجر به تصویر مشاهداتی $g(x,y)$ می شود:

$$g(x,y) = f(x,y) * h(x,y) + n(x,y)$$

با اعمال تبدیل فوری دو بعدی، کانولوشن در فضای فضایی به ضرب ساده در فضای فرکانسی تبدیل می گردد:

$$G(u,v) = F(u,v) H(u,v) + N(u,v)$$

هدف بازسازی $F(u,v)$ از روی $G(u,v)$ است. روش پایه، فیلتر معکوس ساده است که طی آن هر مؤلفه فرکانسی با تقسیم بر پاسخ PSF در آن فرکانس بازسازی می شود:

$$\hat{F}(u,v) = \frac{G(u,v)}{H(u,v)},$$

اما در عمل $H(u,v)$ ممکن است در برخی فرکانس ها به صفر نزدیک باشد و نویز $N(u,v)$ با تقسیم تقویت شود، بنابراین بازسازی ساده موجب برجسته شدن نویز و آرتیفکت می شود. برای رفع این مشکل معمولاً از فیلتر وینر استفاده می شود که با در نظر گرفتن نسبت طیف توان نویز به سیگنال $S_n(u,v)$ and $S_f(u,v)$ میان بازسازی جزئیات و سرکوب نویز تعادلی برقرار می کند:

$$\hat{F}(u,v) = \frac{H^*(u,v)}{|H(u,v)|^2 + \frac{S_n(u,v)}{S_f(u,v)}} G(u,v),$$

در این معادله $H^*(u,v)$ مزدوج مختلط H است و پارامتر $\frac{S_n}{S_f}$ می تواند به صورت میانگین یا تابعی از فرکانس برآورد شود تا نویز در ناحیه هایی که H ضعیف است تضعیف شود.

پس از به دست آوردن تخمین طیف بازسازی شده $\hat{F}(u,v)$ ، با اعمال تبدیل فوری معکوس، تصویر بازسازی شده $\hat{f}(x,y)$ در حوزه فضایی تولید می شود. استفاده از حوزه فرکانس این مزیت را دارد که کانولوشن و تفکیک نویز به اعمال ضرب و تقسیم ساده در فضای فرکانسی تبدیل می شود و می توان با تنظیم دقیق پارامترهای وینر یا افزودن منظم سازی (regularization) از آرتیفکت هایی مانند «رینگینگ» و تقویت نویز جلوگیری کرد. در عمل انتخاب مناسب مدل $H(u,v)$ ، برآورد صحیح نسبت نویز به سیگنال و اعمال پنجره گذاری برای کاهش ناهماهنگی های فرکانسی، کلید دستیابی به یک تصویر بازسازی شده با کیفیت بالا و کمترین اعوجاج است.

در فرآیند بازسازی تصویر تار، استفاده از تبدیل فوری امکان تفکیک ساده عملیات کانولوشن و فیلترینگ را فراهم می کند. ابتدا با مدل سازی مکانیسم تارشدگی (PSF) و محاسبه طیف فرکانسی تصویر تار و PSF، می توان با روش های فیلتر معکوس و سپس فیلتر وینر به بازسازی اولیه دست یافت. فیلتر وینر با در نظر گرفتن نسبت توان نویز به سیگنال از افزایش بی رویه نویز جلوگیری کرده و تعادلی میان حفظ جزئیات و سرکوب اعوجاج برقرار

می‌کند. در نهایت، بازگشت به حوزه فضایی از طریق تبدیل فوریه معکوس تصویر واضح‌تر و با آرتیفکت کمتر را به دست می‌دهد. کلید موفقیت در این روش در دقت مدل‌سازی PSF، برآورد مناسب نویز و انتخاب دقیق پارامترهای فیلتر است؛ به گونه‌ای که تصویر نهایی دارای بیش‌ترین سازگاری با تصویر اصلی و کم‌ترین اثرات ناخواسته باشد.

«... اردیبهشت‌ماه 1404 ...»

.....