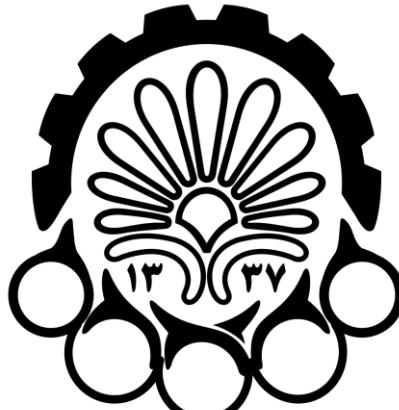


«In The Name Of GOD»



دانشگاه صنعتی امیرکبیر
(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

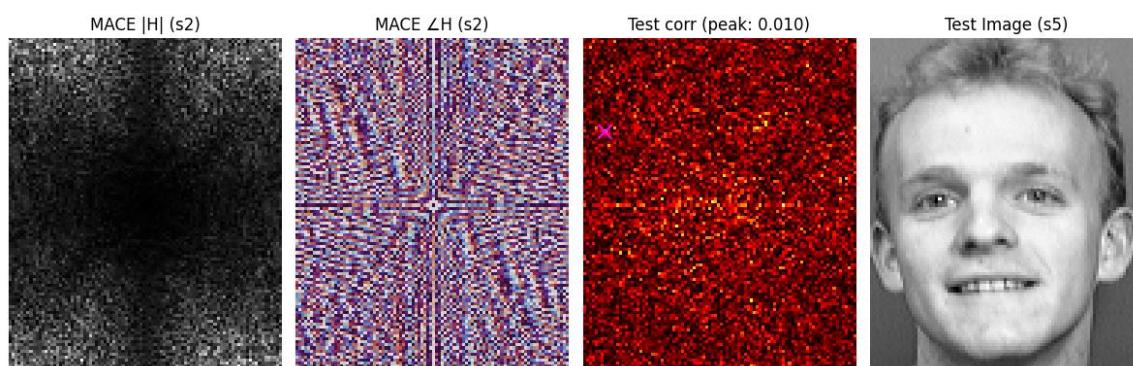
.A

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 مورد نیاز `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 فوریه 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 ساخته شده، 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 درآمده 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 دقیق 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 خطأ عموماً پایین‌تر است که نشان می‌دهد فیلترها برای کلاس‌ها بهینه نیستند و نیاز به بیرون 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 درونیابی پیاده‌سازی می‌کنیم تا بخش‌های آسیب‌دیده یا حذف شده 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 بهمیم چطور مدل‌های ریاضی می‌توانند برای بازیابی تصاویر به کار روند.

در ابتدا کتابخانه‌های موردنیاز import کردیم. در این بخش، ما تصاویر اصلی و تخریب شده (هم خاکستری و هم رنگ) 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)



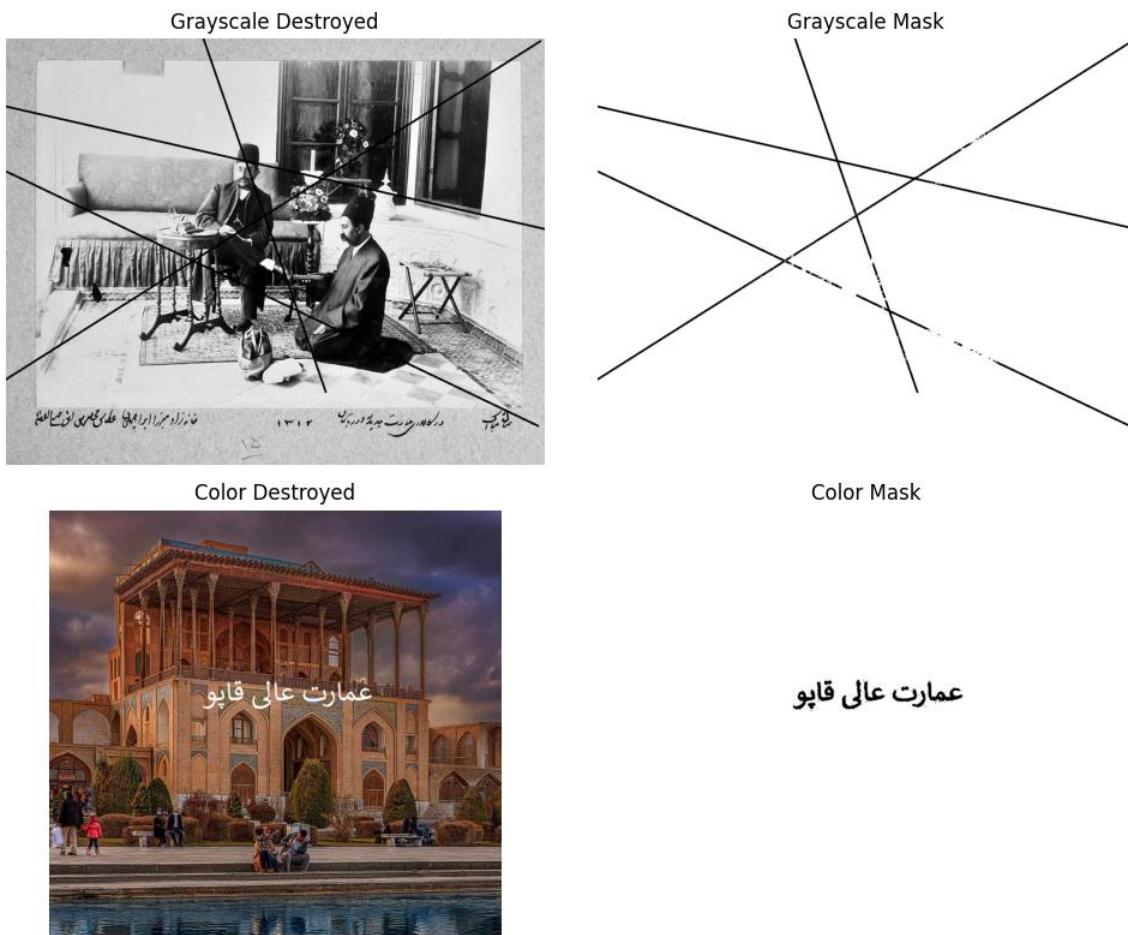
❖ چرانمآل‌سازی تصاویر به بازه $[0,1]$ برای الگوریتم‌های بازسازی بعدی مهم است؟

برای اجرای صحیح الگوریتم‌های بازسازی تصویر، مخصوصاً روش‌هایی که مبتنی بر معادلات دیفرانسیل یا محاسبات عددی هستند، نرم‌السازی تصاویر به بازه $[0,1]$ نقش بسیار مهمی دارد. در ادامه به برخی از این دلایل می‌پردازیم:

1. سازگاری عددی: بسیاری از الگوریتم‌های بازسازی مثل روش‌های PDE یا درونیابی فرض می‌کنند که مقادیر پیکسل‌ها در یک استاندارد معمولاً $[0,1]$ قرار دارند. این کار باعث می‌شود محاسبات عددی پایدارتر و قابل اعتمادتر باشند.
2. کاهش خطاهای مقیاس: اگر مقادیر پیکسل‌ها در بازه‌های مختلفی (مثل $[0,255]$ یا بالاتر) باشند، ممکن است در محاسبات (مثل مشتق‌گیری در PDE یا وزن‌دهی در درونیابی) (خطاهای مقیاس به وجود آید. نرم‌السازی این مشکل را برطرف می‌کند.
3. تسهیل مقایسه: نرم‌السازی باعث می‌شود که مقادیر پیکسل‌ها در تصاویر خاکستری و رنگ قابل مقایسه باشند، بهخصوص وقتی معیارهای کمی مثل MSE یا PSNR محاسبه می‌شوند.
4. سازگاری با روش‌های PDE: در روش‌های PDE (مثل معادله انتشار حرارت)، پارامترهایی مثل ضریب انتشار (λ) و مقادیر اولیه معمولاً برای بازه $[0,1]$ تنظیم می‌شوند. اگر مقادیر پیکسل‌ها خارج از این بازه باشند، ممکن است نیاز به تنظیم مجدد پارامترها باشد که پیچیدگی را افزایش می‌دهد.
5. مدیریت کانال‌های رنگ: در تصاویر رنگی، نرم‌السازی هر کanal به صورت مستقل تضمین می‌کند که هر کanal (قرمز، سبز، آبی) به طور یکسان در فرآیند بازسازی مشارکت کند و از تأثیر نامتعادل یک کanal جلوگیری می‌شود.
6. پایداری الگوریتم‌ها: در کنار مواردی که ذکر کردیم، نرم‌السازی به پایداری الگوریتم‌ها در فرآیندهای بهینه‌سازی هم کمک می‌کند. زمانی که مقادیر پیکسل‌ها در یک استاندارد قرار دارند، الگوریتم‌ها می‌توانند سریع‌تر همگرا شوند و از مشکلاتی مانند افتادن در مینیمم‌های محلی جلوگیری کنند.

.B

در این مرحله، ما ناحیه‌های تخریب‌شده تصاویر را با مقایسه تفاضل پیکسل‌های تصاویر سالم و تخریب‌شده شناسایی و ماسک‌های مربوطه را ایجاد کردیم. برای این کار، پیکسل‌هایی که اختلاف معناداری (بیش از ۰۰۵) داشتند به عنوان مناطق آسیب‌دیده در نظر گرفته شدند. سپس ماسک‌های ایجاد شده برای تصاویر خاکستری و رنگ به همراه تصاویر تخریب‌شده نمایش داده شدند تا صحت تعریف مناطق آسیب دیده بررسی شود. خروجی این ماسک‌ها به همراه تصاویر اصلی به شرح زیر هستند:



❖ چگونه ماسک باینری را ذخیره و استفاده می‌کنید تا نواحی تخریب شده 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 خراب شده را از نواحی سالم تصویر جدا می‌کنند. در تصویر خاکستری، ماسک باینری به صورت یک آرایه دو بعدی با همان ابعاد تصویر است، که در آن مقدار ۱ نشان‌دهنده پیکسل‌های سالم و مقدار ۰ نشان‌دهنده 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 کanal رنگی (قرمز، سبز، آبی) برسی شده و اگر مقدار تفاوت در هر کanal از یک آستانه مشخص (مثلًاً 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 به طور 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

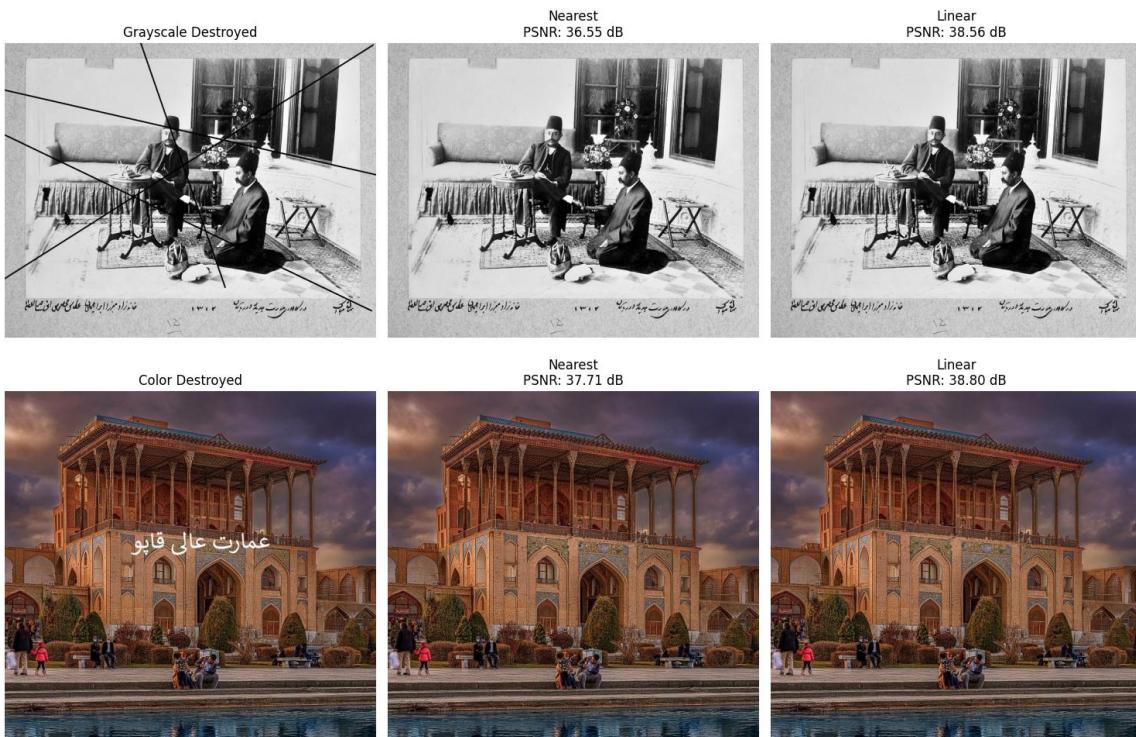
در این مرحله، ما مختصات و مقادیر پیکسل‌های سالم تصاویر (هم خاکستری و هم رنگ) را استخراج کردیم تا به عنوان داده‌های پایه برای درونیابی مورد استفاده قرار گیرند. برای این کار، ابتدا با استفاده از ماسک‌های ایجاد شده، موقعیت مکانی تمام پیکسل‌های سالم را شناسایی کرده و سپس مقادیر شدت روشنایی آنها را استخراج نمودیم. در تصویر رنگی، این فرآیند به صورت جداگانه برای هر کanal 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 ۷۵.. بیانگر مناطق روشن در تصویر است. در تصویر رنگی نیز ۷۹۰، ۸۰۰ پیکسل سالم شناسایی شد که کاهش قابل توجه نسبت به تصویر خاکستری احتمالاً ناشی از تفاوت در ابعاد یا میزان تخریب تصویر است. مقادیر کanal‌های رنگی (R, G, B) در پیکسل‌های ابتدایی مثلاً R..۲۷ G..۳۸ و B..۲۷ نشان‌دهنده رنگ‌های تیره با غلبه جزئی بر کanal قرمز است و به همین صورت می‌توانیم باقی موارد را نیز تفسیر کنیم.

.D

در این مرحله، ما دو روش درونیابی 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 درونیابی خطی (Linear) معمولاً عملکرد بهتری نسبت به روش نزدیک‌ترین همسایه PSNR (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 درون‌بایی از نظر حفظ لبه‌ها (Edge Preservation) و صاف (Smoothness)** چه رفتار متفاوتی دارد؟ دو روش (مثلاً درون‌بایی نزدیک‌ترین همسایه و درون‌بایی 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 صاف (Smoothness) به پیوستگی و نرمی تغییرات شدت در تصویر بازسازی شده 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 بلوک‌های 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 را کپی می‌کنه.

خطی: شدت پیکسل‌های تخریب‌شده را با میانگین وزن دار شدت‌های پیکسل‌های سالم اطراف (بر اساس فاصله) محاسبه می‌کنه.

• مقایسه روش‌ها:
-

(Nearest Neighbor همسایه)

در روش نزدیک‌ترین همسایه، حفظ لبه‌ها در نواحی کوچک به خوبی 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؛ به عنوان مثال، در بازسازی یک خط نازک تخریب‌شده، این روش قادر است با استفاده از شدت پیکسل سالم مجاور، شکل خط را با دقت بالای 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 به صورت تکه‌تکه و غیرطبیعی بازسازی می‌شوند، به طوری که در صورت وجود یک بلوک بزرگ تخریب‌شده، ممکن است کل بلوک با شدت ثابت پر شود و این امر منجر به غیرپیوسته به نظر آمدن لبه‌های اطراف ناحیه شود؛ از منظر صاف، این روش در نواحی کوچک مشکلی ایجاد نمی‌کند، زیرا پیکسل‌های سالم نزدیک معمولاً شدت‌های مشابهی دارند و کپی شدت یک پیکسل‌های اطراف نتیجه‌ای نسبتاً پیوسته می‌دهد، اما در مناطق بزرگ، صاف ضعیفی دارد، زیرا شدت ثابت به کل ناحیه اعمال می‌شود و این امر منجر به ایجاد مصنوعات بلوکی می‌شود که تصویر را به ویژه در نواحی با تغییرات تدریجی شدت، مانند گرادیان‌های نرم، غیرطبیعی جلوه می‌دهد. روش

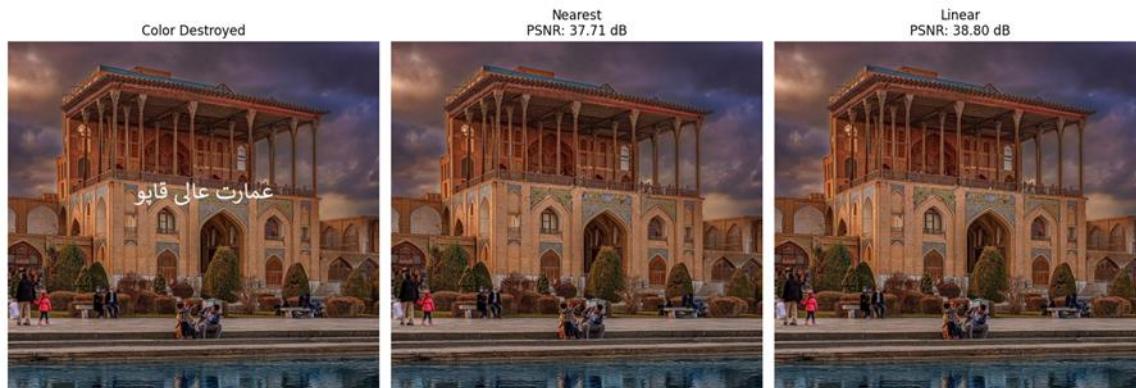
نزدیک‌ترین همسایه از مزایایی نظیر سرعت بالا، سادگی اجرا و عملکرد عالی در نواحی کوچک با تغییرات کم شدت برخوردار است، اما در عین حال معایی نیز دارد، از جمله کیفیت بصری پایین در مناطق بزرگ و ایجاد مصنوعات بلوکی که ظاهر تصویر را تحت تأثیر قرار می‌دهد.

- خطی (Linear) -

در روش خطی، حفظ لبه‌ها در نواحی کوچک ممکن است با چالش‌های همراه باشد، زیرا محاسبات شدت‌ها بر اساس میانگین وزن‌دار با توجه به فاصله انجام می‌شود که می‌تواند لبه‌های تیز را نرم‌تر کند؛ به عنوان 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 به صورت طبیعی تر بازسازی می‌کند؛ برای نمونه، در صورت وجود یک بلوک بزرگ تخریب‌شده، لبه‌های اطراف آن به صورت 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 به صورت برجسته‌ای مطلوب است، زیرا گرادیان‌های نرم و پیوسته‌ای ایجاد می‌کند که تصویر بازسازی‌شده را طبیعی‌تر جلوه می‌دهد؛ به عنوان مثال، در نواحی بزرگ با تغییرات تدریجی شدت، مانند آسمان یا پوست، روش خطی نتایج بسیار بهتری ارائه می‌دهد؛ این روش از مزایایی نظیر عملکرد عالی در مناطق بزرگ و نواحی با تغییرات تدریجی شدت، مانند گرادیان‌های نرم، برخوردار است، اما معایی 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 جدآگاهی 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 جدآگاهه 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 تخریب شده استفاده شد. مقادیر پیش‌فرض در نواحی بدون داده با میانگین شدت‌ها جایگزین شد و در پایان، تنها پیکسل‌های تخریب شده 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 ناهماهنگ 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:

یک از مهم‌ترین چالش‌ها ناسازگاری رنگی (Color Inconsistency) است؛ چون هر کانال به طور جدآگاهه 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 به صورت 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 بازسازی شده را مصنوعی جلوه دهد.

F

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

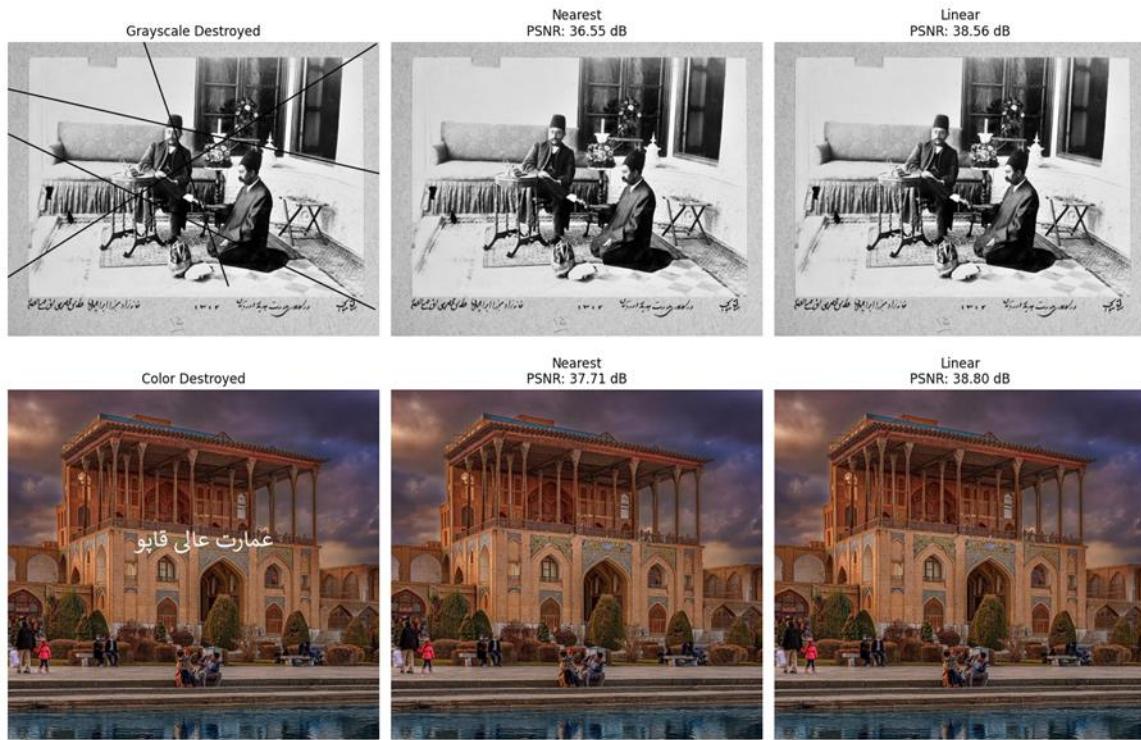
Color Destroyed (Normalized)



Grayscale Destroyed (Normalized)



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



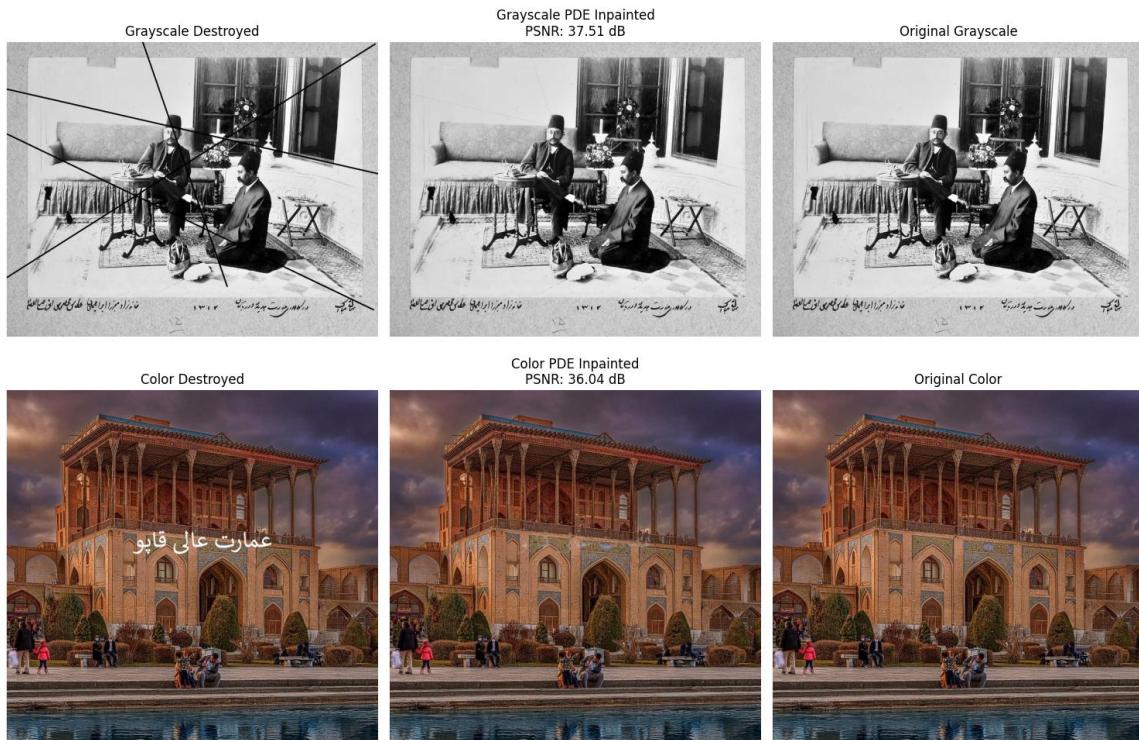
هر دو روش برای دو تصویر رنگی و خاکستری توانستند به مقدار PSNR بالایی دست پیدا کنند و عملکرد خیلی خوبی در ترمیم تصاویر داشتند و همانطور که در پارت های قبلی اشاره کردیم روش خطی کمی بهتر از روش دیگر عمل کرده است اما در کل هر دو به خوبی توانستند این کار را انجام دهند.

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

مقدار PSNR نشان‌دهنده 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 روش‌های درونیابی نزدیک‌ترین همسایه و خطی مشاهده شد که روش خطی به دلیل میانگین‌گیری وزن‌دار شدت پیکسل‌های سالم اطراف، تصاویر بازسازی شده نرم‌تر و یکنواخت‌تری تولید می‌کند که معمولاً باعث افزایش مقدار PSNR و بهبود کیفیت بصری می‌شود. در مقابل، روش نزدیک‌ترین همسایه به دلیل حفظ کمتر صافی و وجود لبه‌های پله‌ای، معمولاً PSNR پایین‌تری دارد. البته لازم به ذکر است که PSNR تنها یک معیار کمی است و همیشه بازنگردی تفاوت‌های بصری به طور کامل نیست؛ بنابراین در کنار آن باید مقایسه بصری نیز انجام شود تا کیفیت واقعی بازسازی به درستی ارزیابی گردد.

.G

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 (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 دهد. تغییر در λ چه تأثیری روی فرآیند پخش دارد؟ ❖

معادله‌ی 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 سریع‌تر 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 محوشدنگی بیش از حد لبه ها و جزئیات تصویر شود. از طرف دیگر، مقدار کم λ فرآیند پخش را کند و بازسازی تصویر به آرامی صورت می گیرد. دوم، لاپلاسیان تصویر Δu است که گرادیان دوم شدت تصویر را نشان می دهد و معیاری برای تغییرات محلی شدت محسوب می شود. این عبارت تفاوت بین شدت یک پیکسل و میانگین شدت پیکسل های همسایه را اندازه‌گیری کرده و نقش اصلی آن هموارسازی و پراکندگی اطلاعات از نواحی سالم به نواحی

تخریب شده است. سوم $(f - u)_{\Omega \setminus D} \chi$ 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 لaplاسین 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 مرزی نیومن را اعمال کردیم که فقط از دو همسایه مجاور بهره می‌برد.

❖ شرایط مرزی را چطور به صورت عددی پیاده‌سازی می‌کنید؟ استفاده از (ghost points) با تغییر مستقیم ماتریس را توضیح دهید

در کد بخش G که پیاده‌سازی تابع Laplacian، شرایط مرزی Neumann به صورت عددی با استفاده از جایگزینی مقادیر پیکسل‌های مرزی انجام شده است. به این شکل که برای محاسبه لaplاسین 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 شبیه، در اطراف مرزهای تصویر نقاط فرضی (نقطه شبیه) 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 شبیه برابر مقدار پیکسل‌های مجاور در داخل تصویر انجام می‌شود که به معنای بازنگاری 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

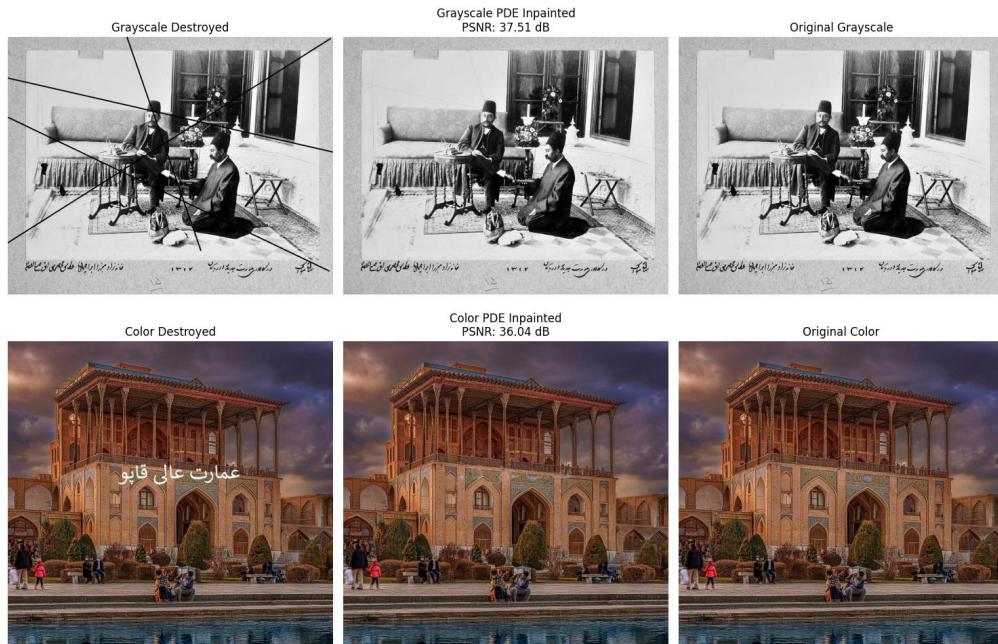
مستقیم ماتریسی، هنگام ساخت ماتریس سیستم معادلات گسسته شده، ضرایب معادلات در مرزها به گونه‌ای اصلاح می‌شوند که معادل شرایط Neumann باشند. به طور خاص، برای نقاط مرزی، رابطه‌ای ایجاد می‌شود که مشتق نرمال صفر را تضمین می‌کند و معمولاً این باعث می‌شود مقادیر مربوط به نقاط خارج از مرز در معادلات حذف یا جایگزین 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 می‌شوند 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.

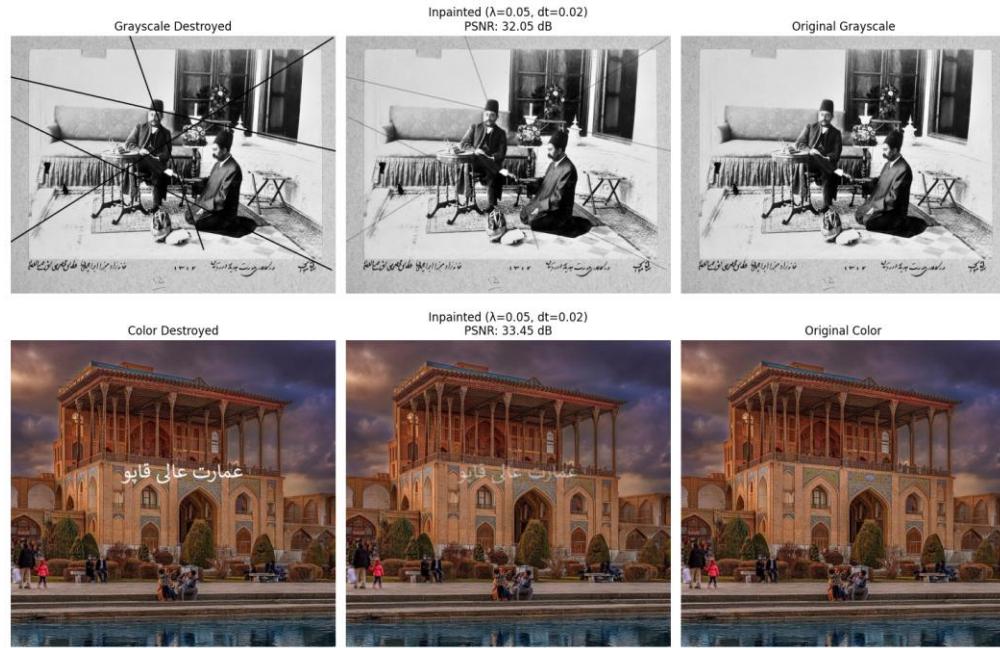
در این بخش، ما به بررسی حساسیت الگوریتم PDE به پارامترهای λ و Δt پرداختیم. در دو پارت قبلی این دو پارامتر را با مقدادیر 0.05 بررسی کردیم و نمایش دادیم و حالا می‌خواهیم دو مجموعه پارامتر مختلف را آزمایش کنیم و هر بار یکی از پارامترها را ثابت و دیگری را تغییر دهیم تا مقایسه بهتری بتوانیم داشته باشیم:

1. آزمایش اول ($\lambda=0.05$, $\Delta t=0.02$) : ین تنظیمات با مقدار λ بالاتر، انتشار سریع تراطلاعات را موجب شد و منجر به ترمیم هموارتر نواحی تخریب شده می‌گردد.
2. آزمایش دوم ($\lambda=0.01$, $\Delta t=0.05$) : با کاهش λ ، فرآیند انتشار کندر شد و جزئیات و لبه‌ها بهتر حفظ شدند، هرچند ممکن است نواحی بزرگ‌تر به طور کامل ترمیم نشده باشند و گام زمانی بزرگ‌تر باعث شد الگوریتم در همان تعداد تکرار به همگرایی کمتری برسد.
3. نتایج مقداردهی دو پارت قبلی و این دو آزمایش را در زیر نمایش دادیم:

- $\lambda=0.05$, $dt=0.05$

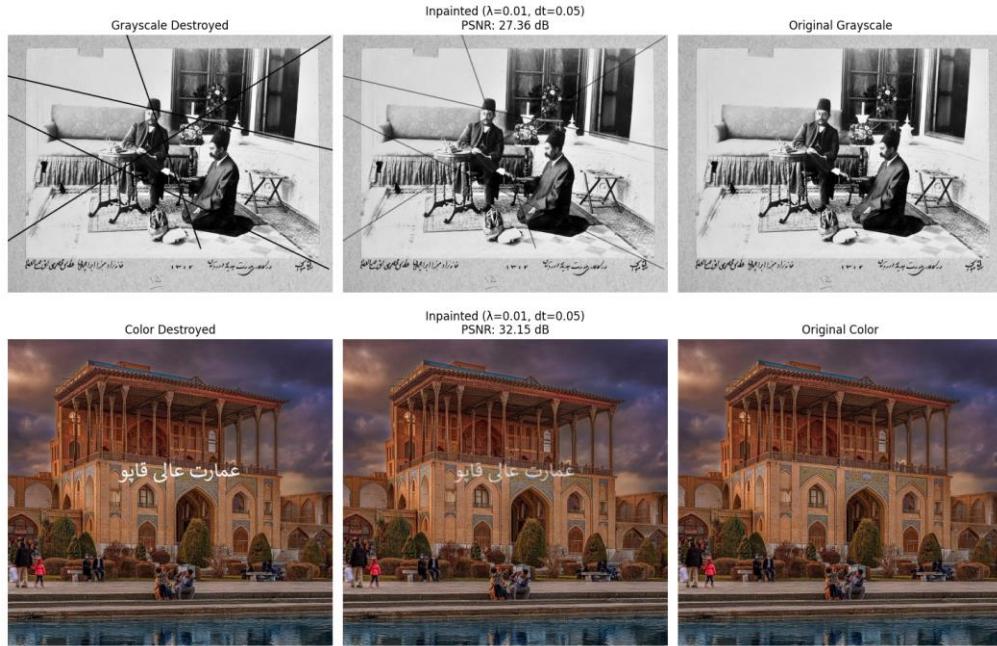


- $\lambda=0.05$, $dt=0.02$



همانطور که می بینیم دو تصویر ترمیم یافته با این مقداردهی نسبت به مقداردهی هردو 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 می‌بینیم که باز هم نتیجه از قسمت قبلی بدتر شد و هم مقدار PSNR کاهش پیدا کرده است و هم بخش‌های 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 صاف‌سازی لبه‌ها دارد؟ اگر Δt خیلی زیاد یا خیلی کم باشد چه اتفاق می‌افتد؟ 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 صاف‌سازی تصویر را کنترل می‌کند. مقدار بزرگ‌تر λ باعث پخش سریع‌تر شدن پیکسل‌ها در نواحی تخریب شده می‌شود که به نرمی بیشتر تصویر منجر می‌شود، اما در عوض می‌تواند لبه‌های تیز و جزئیات ساختاری را محو کند. در مقابل، مقدار کوچک‌تر λ باعث پخش آهسته‌تر می‌شود که لبه‌ها را بهتر حفظ می‌کند، ولی ممکن است نواحی تخریب شده به‌طور 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 لبه‌ها به‌طور قابل قبول حفظ شدند. مقادیر PSNR برابر با 32.05 (خاکستری) و 33.45 (رنگی) نشان‌دهنده تعادل مناسب بین صافی و دقت بازسازی بود. در مقابل، با $\lambda = 0.01$ و $\Delta t = 0.05$ ، سرعت پخش بسیار پایین‌تر بود. در نتیجه تصویر خاکستری با PSNR پایین‌تر (dB 27.36) بازسازی شد و لبه‌ها بهتر حفظ شدند، اما نواحی بازسازی شده نسبت به حالت اول پراکنده‌تر و کمتر هموار بودند. در تصویر رنگی، تغییر زیادی ایجاد نشد (dB 32.15) ولی تفاوت‌های ظرفی در یکنواختی رنگ‌ها دیده می‌شد.

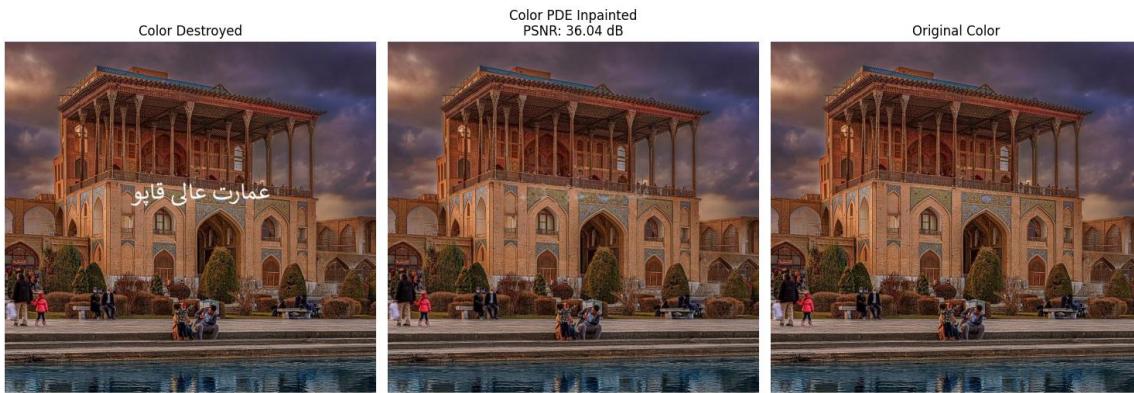
اگر Δt خیلی بزرگ باشد: الگوریتم ممکن است ناپایدار شود و نوسانات عددی ایجاد کند (پخش بیش از حد، نویز، یا ناپایداری محاسباتی). در آزمایش دوم با $\lambda = 0.01$ و $\Delta t = 0.05$ مقدار حاصل 0.004 بود که در محدوده پایدار قرار دارد، ولی اگر λ افزایش می‌یافتد، پایداری به خطر می‌افتد. به همین دلیل باید شرط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

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

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

نتایج به دست آمده نشان می‌دهد که پارامترهای λ و Δt تأثیر مستقیمی بر کیفیت بازسازی تصویر دارند. انتخاب مقدار مناسب λ باعث ایجاد تعادل بین حفظ ساختارهای لبه‌ای و صاف نواحی تخریب شده می‌شود. مقدار زیاد λ بازسازی را سریع‌تر و هموارتر می‌کند اما می‌تواند به محوشدن لبه‌ها منجر شود، در حالی که مقدار کم آن باعث حفظ لبه‌ها می‌شود ولی صاف تصویر را کاهش می‌دهد. 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 کanal اجرا کنیم، چه تأثیری روی سازگاری رنگ‌ها دارد؟ آیا رویکرد coupled می‌تواند بهتر باشد؟

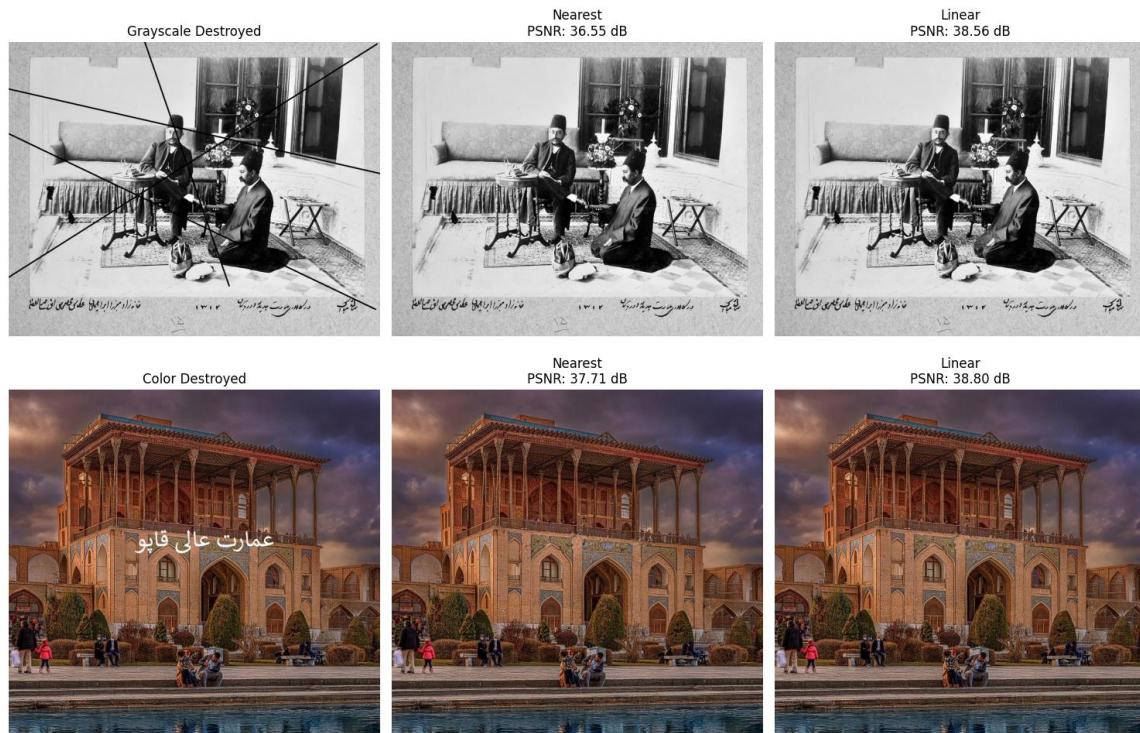
اگر PDE را به صورت جداگانه روی هر کanal رنگی (R, G, B) اعمال کنیم، هر کanal بدون در نظر گرفتن وابستگی‌های ساختاری یا رنگی نسبت به کanal‌های دیگر پردازش می‌شود. اگرچه این رویکرد ساده و در بسیاری از موارد مؤثر است، اما ممکن است باعث بروز مشکلاتی در color consistency شود، به ویژه در نواحی دارای لبه‌های تیز یا بافت‌های پیچیده. دلیل این امر آن است که بازسازی در هر کanal مسیر متفاوتی را طی می‌کند و حتی اختلاف‌های جزئی در بازسازی لبه‌ها یا گرادیان‌ها بین کanal‌ها می‌تواند منجر به ایجاد مصنوعات رنگی مانند هاله‌های رنگی، لبه‌های ناهم‌راستا یا تغییر رنگ‌های غیرطبیعی شود. این مسئله اهمیت ویژه‌ای دارد زیرا سیستم بینیای 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 کanal رنگی به صورت همزمان و وابسته به یکدیگر بازسازی می‌شوند. در این روش‌ها، فرآیند diffusion بر پایه ساختار مشترک تصویر در تمام کanal‌ها هدایت می‌شود تا اطمینان حاصل شود که گرادیان‌ها و لبه‌ها در بین کanal‌ها همانهنج باقی مانند. همچنین، یک رویکرد جایگزین، تبدیل تصویر به فضای رنگی ادراکی مانند CIELab یا YUV و اعمال بازسازی PDE فقط روی کanal روشنایی L یا Y است که بیشترین اطلاعات ساختاری را در خود دارد. سپس کanal‌های رنگی به صورت ساده‌تر بازسازی شده و تصویر به فضای RGB بازگردانده می‌شود. به طور کلی، رویکرد به هم پیوسته یا استفاده از فضای رنگی ادراکی می‌تواند انسجام رنگی و یکپارچگی ساختاری بهتری نسبت به پردازش مستقل کanal‌ها ارائه دهد.

.K

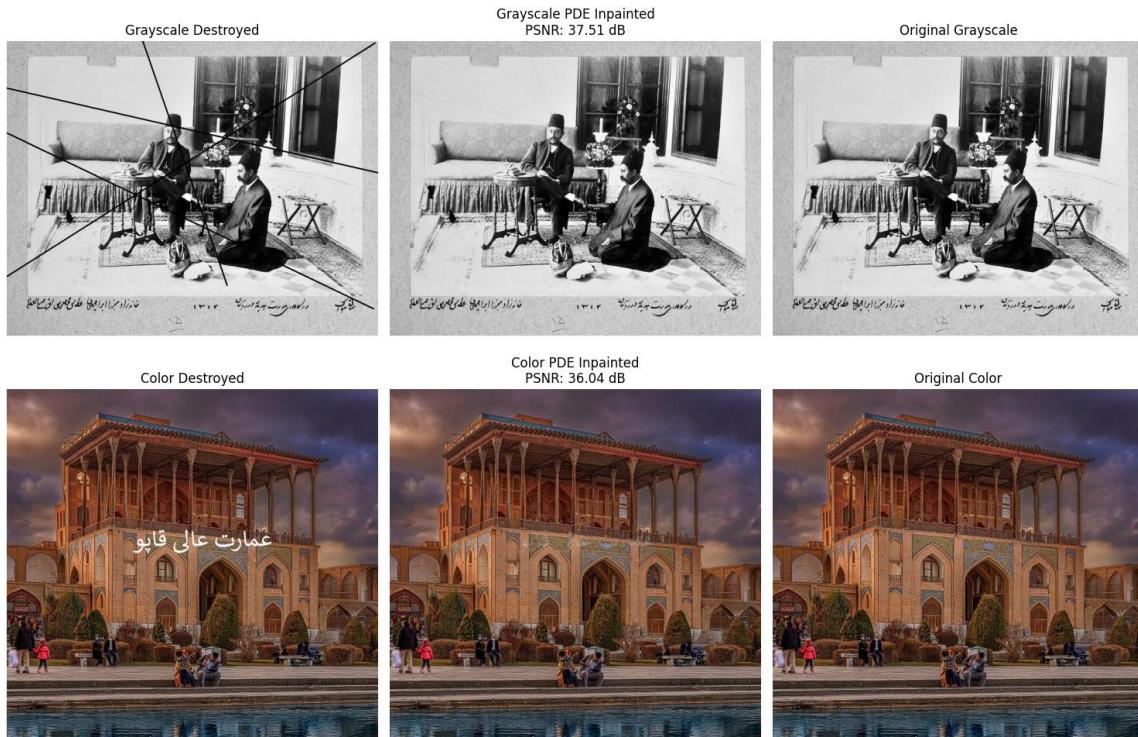
ما در این بخش نتایج دو روش درونیابی و PDE را مقایسه کردیم. نتایج نشان داد که هر یک از این روش‌ها در شرایط خاصی عملکرد بهتری دارند. روش PDE با مقدار PSNR برابر 37.51 دسی‌بل برای تصویر خاکستری و 36.04 دسی‌بل برای تصویر رنگی، در حفظ لبه‌ها و جزئیات تصویر عملکرد بهتری دارد، چرا که با استفاده از عملگر لاپلاسین، 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 درونیابی 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 برونویابی داده‌های سالم، پیوستگی رنگ و روش‌نمایی را در نواحی وسیع بهتر حفظ می‌کند. در مقابل، روش درونیابی 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 تارشدن 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 درونیابی 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 درونیابی خطی (Linear Interpolation) برای نواحی بزرگ‌تر مناسب‌تر است، زیرا با بروزیابی 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 درونیابی خطی بالاترین مقدار PSNR را داشت و برای سایر موارد هم کامل بررسی کردیم و حالا می‌خواهیم در این پارت با محاسبه MSE (میانگین مربعات خطأ) این معیار را هم محاسبه و نمایش دهیم. مقادیر پایین‌تر MSE نشان‌دهنده خطای کمتر و کیفیت بالاتر ترمیم است. انتظار داریم روش‌هایی که PSNR

بالاتری داشتند (مثل درونیابی خطی)، MSE کمتری نیز نشان دهنده این موضوع به صورت کمی دقت هر روش را تأیید می کند. 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 می دهد که روش درونیابی خطی (Linear Interpolation) با کمترین مقادیر MSE (0.000139 برای خاکستری و 0.000132 برای رنگ) بهترین عملکرد را داشته است که این با مقادیر PSNR بالاتر آن در مراحل قبل نیز همخوانی دارد. روش درونیابی 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 رنگ (MSE=0.000249) ضعیفترین نتیجه را داشت.

❖ مقایسه MSE و PSNR بین دو روش چه نتایجی می دهد؟ این اعداد چه چیزی درباره کیفیت بازسازی نشان می دهند؟

مقایسه MSE و PSNR بین دو روش نتایج کاملاً همسوی ارائه می دهد و هرچه MSE کمتر باشد، PSNR بالاتر است که نشان دهنده کیفیت بهتر بازسازی است. به طور مشخص، روش درونیابی خطی با کمترین MSE و PSNR بهترین عملکرد را داشت، یعنی خطای کمتری دارد و تصویر را با دقت بالاتری بازسازی کرده است و می توانیم همان تفسیرهای قبلی را برای اینجا شرح دهیم.

M

نتایج مقایسه کارای محاسباتی دو روش نشان می دهد که روش درونیابی 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 عمده ای به دلیل ماهیت تکراری و پیچیدهی روش PDE است که نیاز به 5000 تکرار برای همگرایی دارد و در هر تکرار، محاسبات سنگین عملگر لایل‌سین و بهره‌رسانی پیکسل‌ها را انجام می دهد. در مقابل، روش درونیابی (بهویژه خطی) با محاسبات مستقیم و غیرتکراری، بسیار بهینه‌تر عمل می کند. این نتایج نشان می دهد که اگرچه روش PDE ممکن است در برخی موارد (مانند حفظ لبه‌ها) کیفیت بهتری ارائه دهد اما هزینه محاسباتی بسیار بالای دارد که آن را برای پردازش تصاویر بزرگ یا بلادرنگ نامناسب می سازد.

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

بر اساس یافته‌های این تمرین، روش درونیابی (به ویژه درونیابی 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 کندر می‌شود و زمان اجرای آن به ۱۵ تا ۲۰ دقیقه می‌رسد. این کندر محسوس عمدتاً به دلیل ماهیت تکراری الگوریتم (۵۰۰۰ تکرار)، محاسبات سنگین عملگر لایپلاسین در هر مرحله و نیاز به بهروزرسانی تدریجی پیکسل‌هاست. جالب توجه اینکه سرعت روش درونیابی تقریباً مستقل از اندازه ناحیه تخریب شده است، در حالی که روش PDE با بزرگ شدن ناحیه مخدوش به صورت نمایی کندر می‌شود. این ویژگی باعث می‌شود روش درونیابی به ویژه برای تصاویر با نواحی تخریب گستردۀ، 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

.A

در این تمرین به بررسی روش‌های نوین افزایش داده مبتنی بر تبدیل فوریه می‌پردازیم. براساس مطالعات اخیر می‌دانیم که اطلاعات معنایی تصاویر عمده‌اً 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 عمل کرد 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 آنجاکه اطلاعات معنایی تصویر عمده‌اً 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 متتنوع‌اند.

دو روش اصلی استفاده شده 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 تعیین‌پذیری مناسبی داشته باشد و انعطاف‌پذیری بالایی در سناریوهای واقعی مانند تشخیص چهره تحت شرایط نوری مختلف از خود نشان دهد.

- پاسخ (Third step(bonus))

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 الهام‌گرفته از سایر تبدیلات مطرح می‌شود: Wavelet DCT Texture Transfer (DTT) و Phase Fusion (WPF). 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 بصیری) 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

$$(3) n(t) = \left(\frac{1}{1+t^2}\right) \cdot e^{j\pi t} \quad \text{فرضیه لیم} \quad \text{اسے، داریم:} \quad 07$$

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

حال بسطل فوریت $g(t)$ بدل فوریت استاندارد اسے و داریم:

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

و همان داشتم w اسے سُب داریم:

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

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

چون تابع برای $t < 0$ رسیدن تابعی دارد و $n(t)$ هم صورت این

سوال نیامده بدل فوریت وجود ندارد و داریم:

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

این انتقال بدل رسیدن تابعی در $t < 0$ هم ایست و بدل فوریت ندارد.

$$(5) \pi(u, v) = \pi(u) \pi(v) \quad F(\pi(v)) \quad \text{احمیه حسین} \quad 15$$

$$F(u, v) = F\{\pi(u)\}(u) \cdot F\{\pi(v)\}(v) \quad \text{طورت جوانش} \quad 16$$

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

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

١٢٥٣ | شنبه ۱۳۹۷ | ۲۰۲۵

② $n(t) = \delta(t-3) + \delta(t+3) - 2\delta(t)$ ٠٧

$\delta(t-t_0) \Leftrightarrow e^{-j\omega t_0}$ (بدل خود) ٠٨

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

$\delta(t+3) \rightarrow e^{j3\omega}$ $I_2(\delta t) = 2 e^{-j3\omega} = 2e^{j3\omega}$ ٠٩

$\Rightarrow F(u) = e^{-j3\omega} + e^{j3\omega} - 2 \Rightarrow [2 \cos(3\omega) - 2]$ ١٠

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

① $n(t) = (e^{-\alpha t}) u(t); \alpha > 0$ ١٢

$X(j\omega) = \int_{-\infty}^{\infty} e^{-\alpha t} u(t) e^{-j\omega t} dt = \int_0^{\infty} e^{-\alpha t} e^{-j\omega t} dt$ ١٣

$= \int_0^{\infty} e^{-(\alpha+j\omega)t} dt$ ١٤

از طریق این یک انتقال استاندار بفرم نیر است:

$\int_0^{\infty} e^{-kt} dt = \left[-\frac{1}{k} e^{-kt} \right]_0^{\infty} = 0 - \left(-\frac{1}{k} \right) = \frac{1}{k}$ ١٥

$\cdot \frac{1}{\alpha+j\omega} : \text{ارم} \omega \text{ باشد } K = \alpha + j\omega \text{ باشد}$ ١٦

٦. $f(n, y) = \pi(n) \delta(n-y)$ ١٧

$F(u, v) = \iint \pi(n) \delta(n-y) e^{-j(u n + v y)} dndy$

برای حل این انتقال از دو کم عکس و دو معمولی: ١٨

$\int \delta(n-y) e^{-jv y} dy = e^{-jv n}$

١٤٠٣

٢٩ ٢٨ ٢٧ ٢٦ ٢٥ ٢٤ ٢٣ ٢٢ ٢١ ٢٠ ١٩ ١٨ ١٧ ١٦ ١٥ ١٤ ١٣ ١٢ ١١ ١٠ ٠٩ ٠٨ ٠٧ ٠٦ ٠٥ ٠٤ ٠٣ ٠٢ ٠١

$F(u, v) = \int \pi(n) e^{-j(u n + v n)} dn = \int_{-1/2}^{1/2} e^{-j(u+v)n} dn$ ١٩

$\int_{-1/2}^{1/2} e^{-j(u+v)n} dn = \frac{2 \sin((u+v)/2)}{u+v} = \frac{\sin((u+v)/2)}{(u+v)/2}$ ١٢

B

۱۴۰۵ | ۱۳۹۶ | ۲۰۲۵

۰۷) در این سوال ابتدا ۶ ویژگی لغتنامه را بررسی کنید و بعد پرسش می‌پرسید.

۰۸) $RC\{X(j\omega)\} = 0 \Rightarrow$ وقتی بخوبی حقیقی یک سیگنال صفر باشد

۰۹) $n(t) = n(-t)$ است و داریم:

۱۰) $IM\{X(j\omega)\} = 0 \Rightarrow$ وقتی بخوبی موهومی یک سیگنال صفر باشد

۱۱) $n(t) = n(-t)$ است و داریم:

۱۲) این ویژگی سوم به این اشاره دارد که سیگنال را با یک سیستم بتعادل بدستگانل نزوح تبدیل کرد.

۱۳) یعنی متناسب سیگنال در صفر برابر صفر شود.

۱۴) در ادامه ۶ سیگنال آتا که بررسی کنید:

(a) سیگنال فردی که باید سیستم بتعادل نزوح تبدیل کنند و در زمان صفر

۱۵) برابر صفر است \Rightarrow ویژگی ها (۱, ۳, ۴)

(b) این سیگنال یک سیگنال کسنه در زمان است و در آن باید سیستم

۱۶) آن را به یک سیگنال نزوح تبدیل کرد \Rightarrow (۳, ۶)

۱۷) این سیگنال یک سیستم آن را به یک سیگنال نزوح تبدیل کرد \Rightarrow (۳)

(c) یک سیگنال فرد است که در زمان صفر برابر صفر است \Rightarrow (۱, ۴)

۱۴۰۵ | ۱۳۹۶ | ۲۰۲۵

(d) یک سیگنال نزوح است که متناسب آن در صفر برابر صفر است \Rightarrow (۲, ۵)

(e) این سیگنال نزوح است که مقدار در صفر نیز

۰۷) برابر صفر می‌شود \Rightarrow (۲, ۴)

۰۸)

۰۹)

.C

$$f = [1, -1, 2, 3, 0, 0, 0, 0]$$

(1) C⁰⁷

$$g = [1, -2, -1, 0, 0, 0, 0, 0]$$

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

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

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

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

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

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

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

بخش اصلی و غیر صفر کافل دوشن $\rightarrow [1, -3, 3, 0, -8, -3]$

$\hookrightarrow [1, -3, 3, 0, -8, -3, 0, 0]$ \rightarrow پایه نیگ ۰ بدهش ۸

2. هر کار این بخش سیگنال ها را که در قسمت قبل بدستوریم، با استفاده از DFT تبدیل فوریه هر کدام را حساب می‌کنیم و بعد با ضرب نقطه‌ای DFT های f و g حاصل را بدست می‌آوریم و بعد با IDFT نتیجه را در حوزه زمان محاسبه می‌کنیم و

دقیقاً با نتیجه کافل دوشن مرحله قبل برابر $[1, -3, 3, 0, -8, -3, 0, 0]$

پس از حاصل محسوب شود.

١١٦ پنجشنبه ٢٠٢٥ | ١١٠

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

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

$$F[K] = \sum_{n=0}^7 f[n] e^{-j \frac{2\pi}{8} Kn} = 1 \cdot e^{-j \frac{\pi}{4} K} + (-1) \cdot e^{-j \frac{3\pi}{4} K} +$$

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

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

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

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

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

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

٢٦ جمعه ١٥ شعبان ١٤٤٦ | Feb 2025

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

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

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

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

میکنیم: صورتی داریم که مربوط به DFT است.

$$\begin{aligned} e^{-j \frac{\pi}{4}} &= \frac{\sqrt{2}}{2} - j \frac{\sqrt{2}}{2} \\ e^{-j \frac{3\pi}{4}} &= -\frac{\sqrt{2}}{2} - j \frac{\sqrt{2}}{2} \\ e^{-j \frac{\pi}{2}} &= -j \end{aligned}$$

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

07

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

08

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

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

09

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

10

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

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

11

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

12

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

حالا در ادامه می‌بینیم که $G[k]$ و $G[k]$ را مربوط به DFT محاسبه کنیم.

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

محاسبه:

14

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

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

و بدین منظور می‌خواهیم $H[1]$ را محاسبه کنیم و بدلاً حالا محاسبه

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

16

DFT مکلفس را محاسبه کنیم:

که حاصل نتایج برابر همان مقادیر کانولوشن نتیجه خواهد بود.

محاسبه:

17

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

18

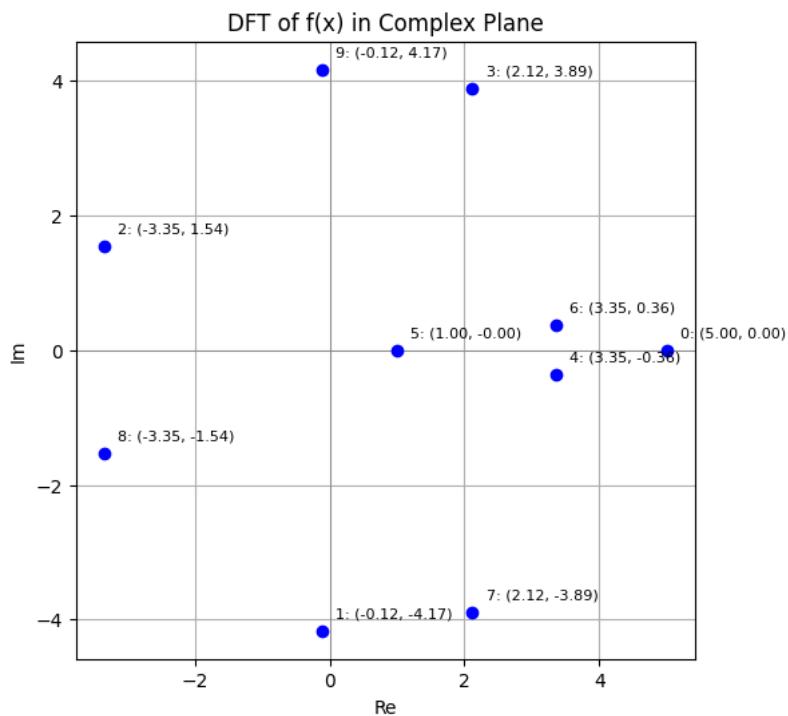
۴. در این فسیل نیز باید انحراف سلسله حاکم f و را حساب کنیم:

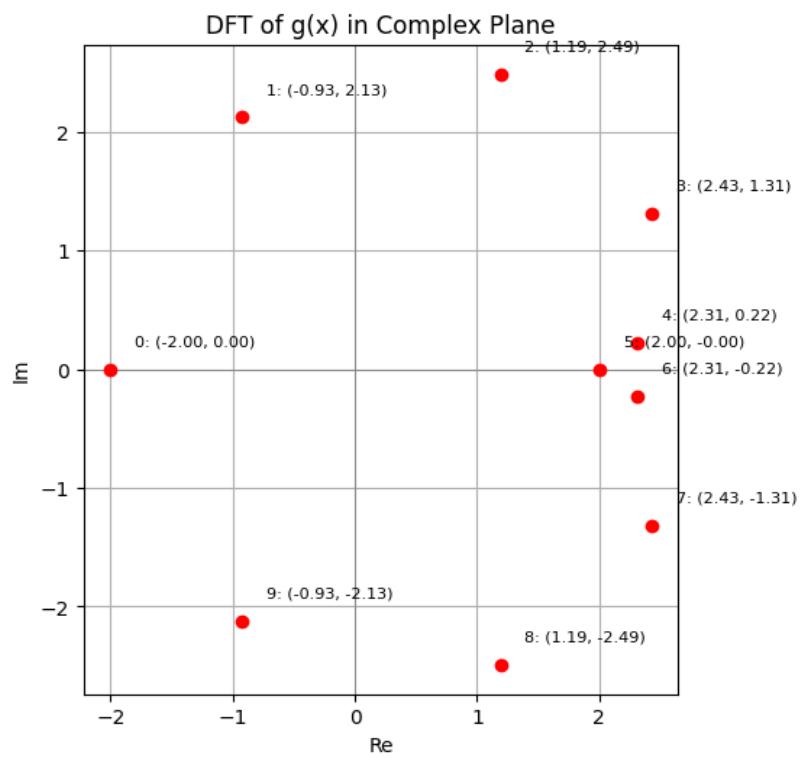
$$\sum_n |f(n)|^2 = 1^2 + (-1)^2 + (2)^2 + (3)^2 = 15 \quad [08]$$

$$\sum_n |g(n)|^2 = 1^2 + (-2)^2 + (-1)^2 = 6 \quad [09]$$

طبق قاعده پارسول انحراف در حوزه زمانی برابر با انحراف در حوزه فرکانس است و این انحراف باید در حوزه فرکانس نشانه شده باشد.

۵. این باره رابه دلیل دستواری رسم $f(x)$ 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، یک از خواص مهم تبدیل فوریه دو بعدی بیان می کند که چرخش تصویر در حوزه فضایی منجر به چرخش طیف فوریه تصویر به همان اندازه در حوزه فرکانسی می شود. این ویژگی که به عنوان ویژگی چرخش (Rotation Property) شناخته می شود 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 می یابد.

از منظر magnitude و phase تبدیل فوریه، چرخش تصویر موجب چرخش کل طیف فرکانسی (شامل هر دو مؤلفه دامنه و فاز) به زاویه ای معادل می گردد، بدون آنکه تغییری در مقادیر آنها ایجاد شود. در واقع، توزیع انرژی فرکانسی تصویر، که در تابع دامنه نمایش داده می شود، به همان شکل باقی می ماند و صرفاً جهت گیری آن در صفحه فرکانسی تغییر می کند. 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 می کند. این خاصیت نقش مهمی در بسیاری از کاربردهای پردازش تصویر دارد، از جمله در image alignment یا image registration 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 می سازد، چرا که چرخش تصویر منجر به تغییری سازمان یافته و قابل پیش بینی در حوزه فرکانسی می شود.

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

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

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

دامنه تبدیل فوریه (Magnitude) حفظ می شود و صرفاً جهت آن تغییر می کند:

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

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 بهوضوح 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 یک از پرکاربردترین تبدیل‌ها در فشرده‌سازی تصویر است، به خصوص در الگوریتم‌های معروفی مثل JPEG. این تبدیل مزایای مختلفی برای فشرده‌سازی نسبت به DFT دارد 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 سیگнал (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 کم فرکانس 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 آرتیفیکت‌های 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، فشرده‌سازی با اتلاف (lossy) را کارآمدتر می‌سازد

امکان پیاده‌سازی سریع (Fast Algorithms):

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

در مجموع، استفاده از DCT در فشرده‌سازی 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 نسبت به DFT گزینه بسیار مناسب‌تری برای فشرده‌سازی تصویری باشد.

C.

در پردازش تصویر، عملیات بلور (Blurring) در حوزه مکانی به صورت کانولوشن (Convolution) تصویر با یک هسته صاف کننده (مانند فیلتر گاوی) 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 صاف کننده در حوزه مکانی کانولوشن می‌شود، در حوزه فرکانسی همان عملیات ضرب بین تبدیل فوریه تصویر و تبدیل فوریه 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، $F(u, v)$ فیلتر، $h(x, y)$ تبدیل فوریه تصویر و $H(u, v)$ تبدیل فوریه فیلتر هستند. در نتیجه، بلور کردن تصویر در حوزه مکانی منجر به ضرب تبدیل فوریه تصویر در یک فیلتر پایین گذر-Low-Pass Filter در حوزه فرکانسی می‌شود. این فیلترهای پایین گذر مؤلفه‌های فرکانسی پایین را حفظ می‌کنند و مؤلفه‌های فرکانسی بالا، که معمولاً شامل جزئیات و لبه‌های تصویر هستند، را تضعیف یا حذف می‌کنند. از آن جایی 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 پایین گذر است.

تأثیر این فرآیند بر تبدیل فوریه تصویر به دو بخش اصلی تقسیم می‌شود: اندازه (Magnitude) و فاز (Phase). اندازه تبدیل فوریه، میزان انرژی هر مؤلفه فرکانسی را نشان می‌دهد. در اثر اعمال فیلتر پایین گذر، 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 بلورشدن 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 مکانی با جزای 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 پیش‌پردازش 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 ظاهري تمیزتر و متمرکز بر ساختارهای اصلی است.

افزایش کنتراست موضعی: نواحی ای 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-*High-pass*) 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 راه حل‌ها در عمل همیشه به طور کامل قابل پیاده‌سازی نیستند. در دنیای واقعی، مشکلات مانند نویز در سیگنال آنالوگ، غیرایده‌آل بودن فیلترهای پایین‌گذر، اختلالات در نرخ نمونه‌برداری به دلیل مشکلات سخت‌افزاری و حتی محدودیت‌های پردازشی در سیستم‌های دیجیتال باعث می‌شود که الیاسینگ ممکن است حق 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 حوزه‌ی 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.

غیرایده‌آل بودن فیلترهای پایین‌گذر: هیچ فیلتر آنالوگ ایده‌آل نیست. همه‌ی فیلترهای فیزیکی دارای شبیه 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 به‌خوبی 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 پیش‌بینی نشده 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. اولًا، باعث کاهش کیفیت سیگنال بازسازی شده می‌شود، به‌طوری 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. به‌ویژه در زمینه‌های 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 اشتباہ شود. اگرچه نظریه نایکوئیست راهکارهای 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، نخ 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:

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

سیگنال جابجا شده در دامنه فضایی است.

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 حق تغییرات اندکی در شئی 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 مانشان می‌دهد که تصویر جابجا شده است و به طور 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 کاربردهای مانند تصویربرداری پزشکی برای هم راستا کردن تصاویر اسکن شده از بدن بیمار مانند CT یا MRI و سنجش از راه دور برای هم راستا 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 فوریه، جابجایی‌های فضایی دقیقاً مشخص می‌شوند و هم راستایی 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 فوریه کمک می‌کند تا به‌طور سریع‌تر و دقیق‌تر این تغییرات اصلاح شود. با محاسبه تفاوت‌های فاز در دامنه فوریه 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 هم راستایی برسانیم.

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

قضیه شیفت تبدیل فوریه این امکان را می‌دهد که به‌جای انجام محاسبات پیچیده در دامنه فضایی که ممکن است برای تصاویر بزرگ زمان بر و هزینه بر باشد، در دامنه فرکانسی عملیات مورد نظر را انجام دهیم. تغییرات فاز در دامنه فوریه 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 هم راستایی سریع تصاویر دارند، بسیار حیاتی است.

در نتیجه قضیه شیفت تبدیل فوریه 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 کاربردهای مانند ثبت تصویر و هماهنگ‌سازی تصویر بسیار حیاتی است، زیرا با استفاده از تحلیل فاز در دامنه فوریه 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 فوریه را به ابزاری اساسی در پردازش تصویر و بینایی ماشین تبدیل کرده است.

.G

در بازسازی تصویر تارشده با استفاده از تبدیل فوریه، ابتدا مدل Point Spread Function یا PSF را به صورت کانولوشن تصویر اصلی $f(x,y)$ با تابع $h(x,y)$ در نظر می‌گیریم که همراه با نویز $n(x,y)$ منجر به تصویر مشاهداتی $g(x,y)$ می‌شود:

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

با اعمال تبدیل فوریه دو بعدی، کانولوشن در فضای فضایی به ضرب ساده در فضای فرکانسی تبدیل می‌گردد:

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

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

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

اما در عمل $H(u,v)$ ممکن است در برخی فرکانس‌ها به صفر نزدیک باشد و نویز $N(u,v)$ با تقسیم تقویت شود، بنابراین بازسازی ساده موجب برجسته شدن نویز و آرتیفیکت می‌شود. برای رفع این مشکل معمولاً از فیلتر وینر استفاده می‌شود 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 برقار می‌کند:

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

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

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

در فرآیند بازسازی تصویر تار، استفاده از تبدیل فوریه امکان تفکیک ساده عملیات کانولوشن و فیلترینگ را فراهم می‌کند. ابتدا با مدل‌سازی مکانیسم تارشدنگی (PSF) و محاسبه طیف فرکانسی تصویر تار و PSF، می‌توان با روش‌های فیلتر معکوس و سپس فیلتر وینر به بازسازی اولیه دست یافت. فیلتر وینر با در نظر گرفتن نسبت توان نویز به سیگنال از افزایش بی‌رویه نویز جلوگیری کرده و تعادلی میان حفظ جزئیات و سرکوب اعوجاج برقار

می‌کند. در نهایت، بازگشت به حوزه فضایی از طریق تبدیل فوریه معکوس تصویر واضح‌تر و با آرتیفیکت 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»
