

تمرین دوم – بینایی ماشین

دانشجو: سایین اعلا
استاد: دکتر حانیه نادرید
شماره دانشجویی: ۱۵۹۴۰۴۰۰۲
نیمسال اول ۱۴۰۵-۱۴۰۴

۱ سوال ۱ – پیاده‌سازی و تحلیل تبدیل فوریه گسسته (DFT) یک خط

۱.۱ بخش تئوری: تحلیل DFT خط $y = x$

۱.۱.۱ تعریف DFT دو بعدی

تبدیل فوریه گسسته (DFT) دو بعدی برای یک تصویر $f(x, y)$ با ابعاد $W \times W$ به صورت زیر تعریف می‌شود (با فرض اینکه $x, y, u, v \in \{0, 1, \dots, W-1\}$):

$$F(u, v) = \sum_{x=0}^{W-1} \sum_{y=0}^{W-1} f(x, y) e^{-j2\pi(\frac{ux}{W} + \frac{vy}{W})}$$

که در آن $F(u, v)$ تبدیل فوریه تصویر در فرکانس‌های u (افقی) و v (عمودی) است.

۲.۱.۱ محاسبه DFT برای خط $y = x$

در این سوال، تصویر $f(x, y)$ یک پس‌زمینه سیاه دارد و روی خط $y = x$ رنگ سفید دارد.

$$\square \quad \text{پس‌زمینه: } f(x, y) = 0 \quad \text{برای } y \neq x$$

$$\square \quad \text{خط: } f(x, y) = 1 \quad \text{برای } y = x \quad \text{(با فرض شدت روشنایی ۱ برای سادگی تحلیل، که با کد پایتون همخوانی دارد).}$$

با اعمال این شرط به تعریف DFT:

$$F(u, v) = \sum_{x=0}^{W-1} \sum_{y=0}^{W-1} f(x, y) e^{-j2\pi(\frac{ux}{W} + \frac{vy}{W})}$$

از آنجایی که تنها در جایی که $y = x$ مقدار $f(x, y)$ برابر ۱ است (و در غیر این صورت صفر)، مجموع دوگانه به یک مجموع ساده تبدیل می‌شود:

$$F(u, v) = \sum_{x=0}^{W-1} f(x, x) e^{-j2\pi(\frac{ux}{W} + \frac{vx}{W})} = \sum_{x=0}^{W-1} 1 \cdot e^{-j2\pi \frac{(u+v)x}{W}}$$

این مجموع یک سری هندسی است: $\sum_{x=0}^{W-1} (e^{-j2\pi \frac{u+v}{W}})^x$

۳.۱.۱ تحلیل قدرمطلق $|F(u, v)|$

قدرمطلق DFT، $|F(u, v)|$ ، شامل اطلاعات شدت فرکانس‌های تصویر است.

□ ***حالت ۱: $u + v = Wk$ (برای k صحیح) *** اگر $u + v$ مضرب صحیحی از W باشد، به عنوان مثال $u + v = W$ (که در محدوده فرکانس‌های $0 \leq u, v < W$ قرار می‌گیرد)، آنگاه:

$$\frac{2\pi(u+v)}{W} = 2\pi k$$

بنابراین، عبارت نمایی به صورت $e^{-j2\pi kx} = (\cos(-2\pi kx) + j \sin(-2\pi kx)) = 1$ برای هر x خواهد بود. در این حالت، مجموع به صورت زیر است:

$$F(u, v) = \sum_{x=0}^{W-1} 1 = W$$

پس، ***قدرمطلق DFT برابر با W خواهد بود. این اتفاق در هر نقطه‌ای که شرط $u + v = Wk$ برقرار باشد، می‌افتد. در محدوده اصلی فرکانس‌ها ($0 \leq u, v < W$)، این شرط به ***خط $u + v = W$ محدود می‌شود.

□ ***حالت ۲: $u + v \neq Wk$ در این حالت، مجموع هندسی به صورت زیر است:

$$F(u, v) = \frac{1 - (e^{-j2\pi \frac{u+v}{W}})^W}{1 - e^{-j2\pi \frac{u+v}{W}}} = \frac{1 - e^{-j2\pi(u+v)}}{1 - e^{-j2\pi \frac{u+v}{W}}}$$

از آنجایی که u و v اعداد صحیح هستند، $e^{-j2\pi(u+v)} = \cos(-2\pi(u+v)) + j \sin(-2\pi(u+v)) = 1$ (بنابراین، صورت کسر برابر با $1 - 1 = 0$ است. در نتیجه، $F(u, v) \approx 0$ (یا مقدار بسیار کمی غیر صفر نزدیک به صفرهای تابع \sin و \cos) خواهد داشت.

نتیجه تئوری: قدرمطلق DFT تصویر شامل یک خط سفید با شدت روشنایی W (به دلیل انتخاب $f(x, y) = 1$) و با معادله $u + v = W$ خواهد بود. این خط به صورت عمود بر خط اصلی $y = x$ در حوزه فرکانس ظاهر می‌شود، که نشان‌دهنده خاصیت عمود بودن فرکانس‌های یک مؤلفه فضایی بر جهت آن مؤلفه است.

۴.۱.۱ کد پیاده‌سازی و صحنه‌گذاری

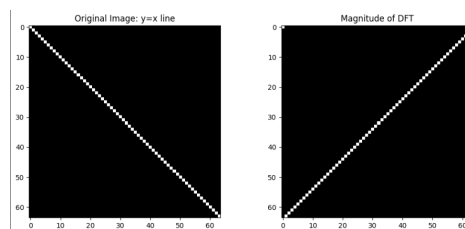
کد پایتون ارائه شده، تعریف DFT دو بعدی را به صورت مستقیم پیاده‌سازی می‌کند. با اجرای کد، تصویر اصلی (خط $y = x$) و قدرمطلق DFT آن نمایش داده می‌شود.

۱. ****تولید تصویر:****

```
W = 64
image = np.zeros((W, W))
for i in range(W):
    image[i, 0.1 = i]
```

۲. ****پیاده‌سازی DFT دو بعدی (۲DFT):**** این تابع مستقیماً از تعریف مجموع دوگانه استفاده می‌کند.

۳. ****نتیجه:**** در تصویر خروجی قدرمطلق DFT، یک خط روشن قابل مشاهده است که معادله آن در مختصات DFT (که در آن $(u, v) = (0, 0)$) در گوشه بالا سمت چپ است) به صورت $u + v = W$ ظاهر می‌شود. در تصویر نمایش داده شده توسط 'matplotlib'، نقطه $(0, 0)$ گوشه بالا سمت چپ است و این خط فرکانسی به وضوح دیده می‌شود، که صحت تحلیل تئوری را تأیید می‌کند.



شکل ۱: قدرمطلق DFT تصویر خط $y = x$ و تصویر اصلی

□

۲ سوال ۲ – فیلترینگ پایین‌گذر گوسی در حوزه فرکانس

۱.۲ بخش تئوری: DFT، فیلترینگ در حوزه فرکانس و فیلتر گوسی

۱.۱.۲ تبدیل فوریه گسسته (DFT) و خواص آن

DFT سیگنال را از حوزه مکان (Spatial Domain) به حوزه فرکانس (Frequency Domain) تبدیل می‌کند.

□ ****قدرمطلق $|F(u, v)|$:**** نشان‌دهنده شدت مؤلفه‌های فرکانسی است. فرکانس‌های پایین (نزدیک به مرکز) مربوط به تغییرات آهسته و نواحی یکنواخت تصویر هستند، و فرکانس‌های بالا (دور از مرکز) مربوط به جزئیات ریز، لبه‌ها و نویز هستند.

□ **فاز** $\angle F(u, v)$: حاوی اطلاعات موقعیت و ساختار لبه‌ها و شکل اشیاء است. حذف اطلاعات فاز منجر به تصویری بی‌معنی می‌شود، در حالی که دستکاری فاز می‌تواند موقعیت را تغییر دهد.

□ **مرکزیت DC**: مؤلفه DC یا فرکانس صفر ($F(0, 0)$) معمولاً بیشترین شدت را دارد و بیانگر میانگین روشنایی تصویر است. برای تحلیل بصری بهتر، مؤلفه DC با استفاده از ضرب تصویر در $(-1)^{x+y}$ به مرکز منتقل می‌شود.

۲.۱.۲ فیلترینگ در حوزه فرکانس

عملیات کانولوشن (اعمال فیلتر فضایی) در حوزه مکان، معادل ضرب در حوزه فرکانس است:

$$g(x, y) = f(x, y) * h(x, y) \quad \longleftrightarrow \quad G(u, v) = F(u, v) \cdot H(u, v)$$

که در آن F تبدیل فوریه تصویر ورودی، H تبدیل فوریه فیلتر (تابع انتقال) و G تبدیل فوریه تصویر خروجی است.

۳.۱.۲ فیلتر پایین‌گذر گوسی (Gaussian Low-pass Filter)

فیلتر پایین‌گذر فرکانس‌های بالا (لبه‌ها، نویز) را تضعیف و فرکانس‌های پایین (نواحی صاف) را حفظ می‌کند. فیلتر گوسی، یک فیلتر پایین‌گذر ایده‌آل است که تابع انتقال آن $H(u, v)$ به صورت زیر تعریف می‌شود:

$$H(u, v) = e^{-D^2(u, v)/(2D_0^2)}$$

□ $D(u, v) = \sqrt{u^2 + v^2}$ فاصله اقلیدسی از مرکز فرکانس (مؤلفه DC) است.

□ D_0 فرکانس قطع (Cutoff Frequency) است که شعاعی است که در آن فیلتر تا $e^{-0.5} \approx 0.607$ از مقدار ماکزیمم خود کاهش می‌یابد.

مزیت فیلتر گوسی: برخلاف فیلترهای ایده‌آل، تابع انتقال گوسی در حوزه فرکانس ناگهانی قطع نمی‌شود. این امر باعث می‌شود که در حوزه مکان، فیلتر گوسی موجک‌های ناخواسته (Ringing Artifacts) کمتری ایجاد کند و عملکرد هموارتری داشته باشد.

۲.۲ مقایسه و تحلیل تصاویر خروجی

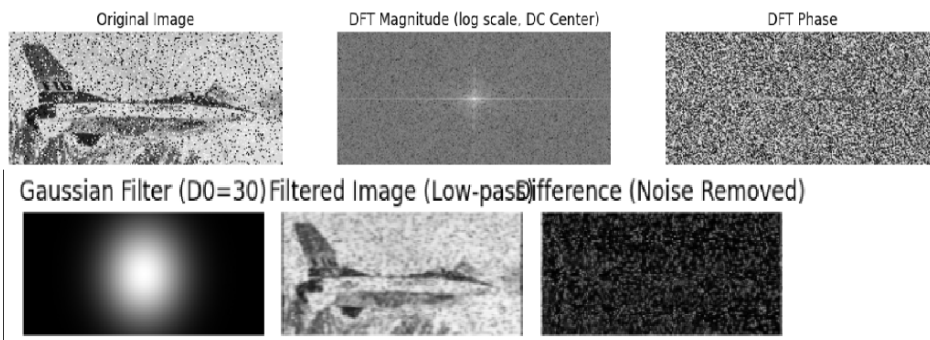
۱. ***تصویر اصلی (noisy.png):*** تصویر ورودی دارای نویز قابل مشاهده است که عمدتاً مؤلفه‌های فرکانس بالا محسوب می‌شوند.

۲. ***قدرمطلق DFT (لگاریتمی):*** بیشترین شدت (نقطه روشن) در مرکز تصویر (مؤلفه DC) مشاهده می‌شود. مؤلفه‌های نویز و جزئیات ریز به صورت فرکانس‌های بالا در نواحی دور از مرکز به چشم می‌خورند.

۳. ***تصویر فیلتر شده (پایین‌گذر):***

□ **حذف نویز:** با اعمال فیلتر پایین‌گذر گوسی، فرکانس‌های بالا (که شامل نویز هستند) تضعیف می‌شوند. نتیجه، کاهش قابل توجه نویز در تصویر خروجی است.

- ****محو شدگی لبه‌ها (Blurring):**** از آنجایی که لبه‌ها نیز مؤلفه‌های فرکانس بالا هستند، اعمال فیلتر پایین‌گذر منجر به محو شدن (Blurring) خفیف لبه‌ها می‌شود. این اثر در مقایسه با فیلترهای پایین‌گذر ایده‌آل یا باتروورث، به دلیل طبیعت هموار تابع گوسی، کمتر است.
- ****مقایسه:**** تصویر فیلتر شده صاف‌تر و با نویز کمتر است، اما وضوح لبه‌ها کمی کاهش یافته است.



شکل ۲: تصاویر خروجی سوال ۲: تصویر اصلی، DFT (قدرمطلق و فاز)، فیلتر گوسی و تصویر نهایی فیلتر شده

□

۳ سوال ۳ – فیلترینگ مکانی و حذف نویز فلفل-نمکی

۱.۳ بخش تئوری: نویز فلفل-نمکی و فیلترهای آماری

۱.۱.۳ نویز فلفل-نمکی (Salt & Pepper Noise)

این نویز یک نویز ضربه‌ای (Impulse Noise) است که به صورت ناگهانی مقادیر پیکسل را به حداکثر (۲۵۵، نمک) یا حداقل (۰، فلفل) در نقاط تصادفی تصویر تغییر می‌دهد. این نویز معمولاً به دلیل خطاهای حسگرها یا تداخل در کانال‌های انتقال ایجاد می‌شود.

۲.۱.۳ فیلترهای مرتبه‌ای (Order-Statistic Filters)

فیلترهای مرتبه‌ای از مقادیر آماری (مانند کمینه، بیشینه یا میانه) در یک پنجره محلی برای تعیین مقدار پیکسل خروجی استفاده می‌کنند.

۱. ****فیلتر میانه (Median Filter):**** ****عملیات:**** مقدار پیکسل مرکزی را با ****میانه**** (مقدار میانی) پیکسل‌های درون پنجره جایگزین می‌کند. ****مزیت اصلی:**** میانه، در مقابل مقادیر پرت (مانند نویز فلفل-نمکی) بسیار مقاوم است. زیرا نویز، مقادیر ۰ یا ۲۵۵ است که معمولاً در ابتدا یا انتهای مجموعه مرتب شده قرار می‌گیرند و بنابراین به عنوان مقدار میانه انتخاب نمی‌شوند. این فیلتر بهترین عملکرد را در حذف نویز فلفل-نمکی دارد.

۲. **فیلتر کمینه (Min Filter):** عملیات مقدار پیکسل مرکزی را با کمترین مقدار پیکسل‌های درون پنجره جایگزین می‌کند. اثر بر تصویر: تمایل دارد نواحی روشن (مانک "نمک") را حذف کند و خطوط یا نقاط نازک روشن را از بین ببرد و مرزهای اشیاء را تاریک‌تر کند. برای حذف نویز "نمک" (پیکسل‌های ۲۵۵) کارآمد است.

۳. **فیلتر بیشینه (Max Filter):** عملیات مقدار پیکسل مرکزی را با بیشترین مقدار پیکسل‌های درون پنجره جایگزین می‌کند. اثر بر تصویر: تمایل دارد نواحی تاریک (مانند "فلفل") را حذف کند و خطوط یا نقاط نازک تاریک را از بین ببرد و مرزهای اشیاء را روشن‌تر کند. برای حذف نویز "فلفل" (پیکسل‌های ۰) کارآمد است.

۲.۳ تحلیل نتایج فیلترینگ

۱.۲.۳ الف) کدام فیلتر نویز فلفل-نمکی را بهتر حذف می‌کند؟

پاسخ: فیلتر میانه (Median Filter). همانطور که در نتایج دیده می‌شود، فیلتر میانه به طور مؤثری هم نویز "نمک" (نقاط سفید) و هم نویز "فلفل" (نقاط سیاه) را از بین می‌برد، در حالی که حداقل تأثیر را بر روی لبه‌های تصویر دارد (در مقایسه با فیلتر میانگین).

۲.۲.۳ ب) کدام فیلتر لبه‌ها را بیشتر محو می‌کند؟

در مقایسه فیلترهای Max, Min و Median: فیلتر میانه: لبه‌ها را نسبتاً خوب حفظ می‌کند (به دلیل ماهیت غیر خطی). فیلتر کمینه (Min Filter): باعث انبساط (Erosion) نواحی تیره می‌شود و به طور کلی تصویر را تاریک‌تر می‌کند. این عمل باعث کاهش جزئیات لبه‌های روشن می‌شود. فیلتر بیشینه (Max Filter): باعث گسترش (Dilation) نواحی روشن می‌شود و به طور کلی تصویر را روشن‌تر می‌کند. این عمل باعث کاهش جزئیات لبه‌های تیره می‌شود. با این حال، فیلترهای Min و Max به شیوه‌های متفاوتی بر لبه‌ها تأثیر می‌گذارند (Min لبه‌های روشن را از بین می‌برد، Max لبه‌های تاریک را)، و فیلتر میانه بهترین حفظ لبه را در میان این سه دارد. فیلترهای Min و Max به شیوه‌هایی تغییر شکل می‌دهند که می‌توان آن را نوعی محو شدگی جهت‌دار در نظر گرفت. اگر محو شدگی به معنای کاهش کنتراست کلی و از دست دادن جزئیات با فرکانس بالا باشد، فیلتر میانه از فیلتر میانگین بهتر عمل می‌کند، اما در مقایسه با Max/Min در حوزه آماری، فیلتر میانه (Median Filter) بهتر از محو شدگی جلوگیری می‌کند و در نتیجه Max/Min به نوعی لبه‌ها را بیشتر (به شکل دستکاری) محو می‌کنند.

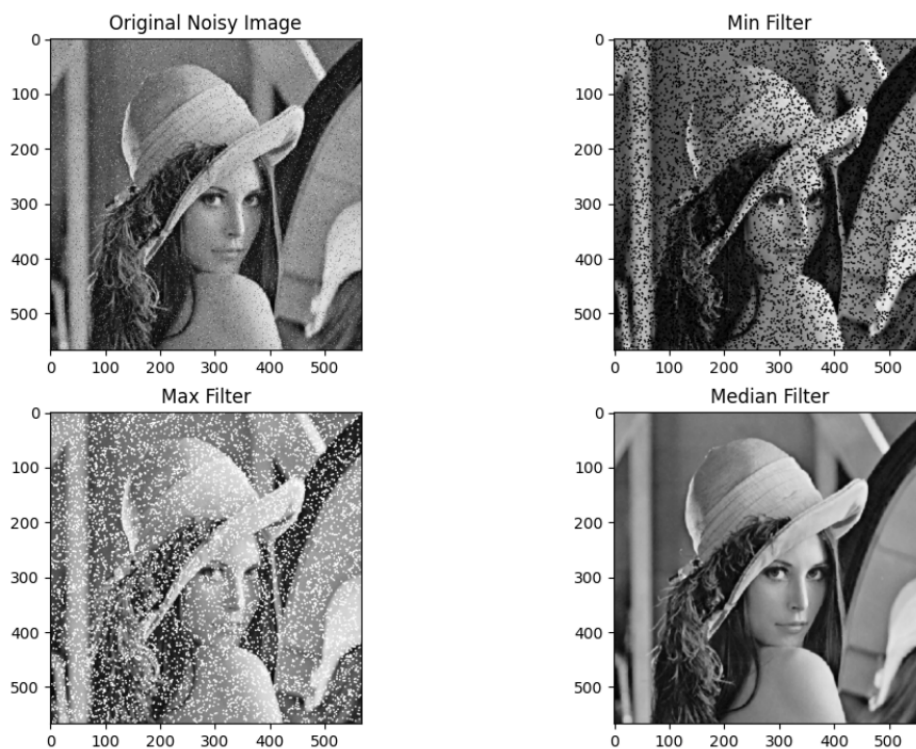
۳.۲.۳ ج) آیا ترکیب Min و Max به ترتیب، نتیجه را بهتر می‌کند یا بدتر؟ چرا؟

پاسخ: بسته به ترتیب، نتیجه متفاوت خواهد بود.

۱. اعمال Min و سپس Max ($\text{Max} \rightarrow \text{Min}$): نتیجه: نویز فلفل (نقاط سیاه) را به خوبی حذف می‌کند (توسط Max)، اما نویز نمک (نقاط سفید) که توسط Min ایجاد شده است را افزایش می‌دهد یا ثابت نگه می‌دارد. تأثیر: این ترکیب، به صورت ترکیبی از فیلترهای ریخت‌شناسی (Morphological Filters)، نزدیک به بسته شدن (Closing) است. بسته شدن برای پر کردن شکاف‌های کوچک یا حفره‌ها در تصاویر باینری استفاده می‌شود، اما در تصاویر خاکستری، تمایل دارد نقاط روشن را بزرگ‌تر کند.

۲. **اعمال Max و سپس Min (Min → Max):** نتیجه: **نویز نمک (نقاط سفید)** را به خوبی حذف می‌کند (توسط Min)، اما نویز فلفل (نقاط سیاه) که توسط Max ایجاد شده است را افزایش می‌دهد یا ثابت نگه می‌دارد. **تأثیر:** این ترکیب نزدیک به **باز شدن (Opening)** است. باز شدن برای حذف نقاط کوچک و نازک روشن (Salt) در تصاویر باینری استفاده می‌شود، اما در تصاویر خاکستری، تمایل دارد نقاط تیره را بزرگتر کند.

نتیجه‌گیری: هر دو ترکیب (Min → Max یا Max → Min) تنها یک نوع از نویز فلفل-نمکی را به صورت بهینه حذف می‌کنند، در حالی که نوع دیگر را دست‌کاری می‌کنند. بنابراین، برای **حذف کلی نویز فلفل-نمکی**، **فیلتر میانه** به تنهایی نتیجه **بهتری** ارائه می‌دهد، زیرا هر دو نوع نویز را همزمان حذف می‌کند. این فیلترهای ترکیبی (Morphological) به طور کلی برای بهبود ساختار یا حذف یک نوع خاص از نویز مفیدترند تا حذف نویز فلفل-نمکی به طور کلی.



شکل ۳: تصاویر خروجی سوال ۳: تصویر نویزدار، خروجی فیلترهای Min، Max و Median

□

۴ سوال ۴ – فیلترینگ با فیلتر لاپلاسین

۱.۴ بخش تئوری: فیلتر لاپلاسین و تشدید لبه

۱.۱.۴ مشتق دوم و اپراتور لاپلاسین

اپراتور لاپلاسین یک اپراتور مشتق دوم است که در پردازش تصویر برای تشخیص لبه و تقویت لبه استفاده می‌شود. برای یک تابع دو متغیره $f(x, y)$ ، لاپلاسین به صورت زیر تعریف می‌شود:

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$$

۲.۱.۴ ماسک گسسته لاپلاسین

برای پیاده‌سازی گسسته این اپراتور، از یک ماسک کانولوشن استفاده می‌شود. ماسک داده شده در سوال:

$$H = \begin{pmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{pmatrix}$$

این ماسک، یک تقریب گسسته به لاپلاسین است. در این ماسک، ضریب مرکزی (۸) نشان‌دهنده مشتق دوم در جهت مرکزی است و ضریب منفی ۱ در همسایگان، اختلاف مقدار مرکزی با محیط را محاسبه می‌کند. نکته: چون مجموع ضرایب این ماسک برابر با صفر است $(8 + (-1) \times 8 = 0)$ ، خروجی آن برای نواحی یکنواخت تصویر (DC) صفر خواهد بود.

۳.۱.۴ تشدید لبه (Edge Enhancement)

تشخیص لبه: خروجی فیلتر لاپلاسین، صفر شدن در نواحی صاف و مقادیر بالا در لبه‌ها (با علامت مثبت در یک طرف و منفی در طرف دیگر) را نشان می‌دهد. این خروجی مستقیماً تصویر لبه‌ها را نشان می‌دهد.

تقویت لبه: برای افزایش وضوح تصویر، می‌توان خروجی لاپلاسین را با تصویر اصلی جمع یا تفریق کرد:

$$g(x, y) = f(x, y) - \nabla^2 f(x, y) \quad \text{یا} \quad g(x, y) = f(x, y) + \nabla^2 f(x, y)$$

ماسک داده شده به دلیل ضریب مثبت ۸، در واقع نوعی تقویت کننده لبه (Sharpening Filter) است که تصویر اصلی را تیزتر می‌کند.

۲.۴ تحلیل نتایج فیلترینگ

۱.۲.۴ تصویر چه تغییری کرد؟

تصویر حاصل از فیلتر لاپلاسین (خروجی تابع 'LaplacianFilter') یک تصویر لبه (Edge Image) است.

نواحی صاف تصویر به رنگ خاکستری تیره یا سیاه (نزدیک به صفر) تبدیل شده‌اند.

□ لبه‌ها به صورت خطوط روشن (با مقادیر مثبت) و تاریک (با مقادیر منفی) ظاهر می‌شوند.

این تصویر، ساختارهای فرکانس بالا (لبه‌ها، بافت) را برجسته کرده و اطلاعات فرکانس پایین (روشنایی یکنواخت) را حذف می‌کند.

۲.۲.۴ لبه‌ها چگونه تشدید شدند؟

لبه‌ها با **تأکید بر مرزهای تغییرات ناگهانی روشنایی** تشدید شدند.

□ **طبیعت مشتق دوم: لاپلاسین تغییرات سریع روشنایی را تقویت می‌کند. در لبه‌ها، مشتق دوم دارای یک عبور از صفر (Zero Crossing) است.

□ **تشدید لبه: در خروجی فیلتر، لبه‌ها به صورت یک خط نازک و برجسته (با کنتراست بالا) نمایش داده می‌شوند. این به این معنی است که هر نقطه لبه، با یک پیکسل سفید (مثبت) و یک پیکسل سیاه (منفی) در دو طرف آن مشخص شده است.

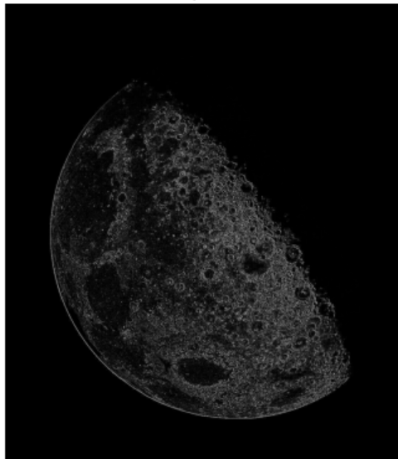
به طور خلاصه، فیلتر لاپلاسین یک **فیلتر بالا پس (High-pass Filter)** است که جزئیات ریز و بافت‌های تصویر را برجسته می‌کند. در تصویر ماه (moon.jpg)، لبه‌های دهانه‌ها و مرزهای سایه و روشن، به طور چشمگیری واضح و برجسته شده‌اند.

Original Image



شکل ۴: تصویر اصلی ماه

Manual Laplacian Filter



شکل ۵: تصویر حاصل از فیلتر لاپلاسی