به نام خدا

پردازش تصویر

تمرین شماره‌ی ۴

حوزهٔ فرکانسی تصویر، نگاهی بیشتر به شیوهٔ نمایش تصویر و

آشنایی با فرمت تصویری DICOM

تاریخ تحویل: ۱۴۰۰/۳/۱

ارشین سلطان بایزیدی

۹۷۳۳۰۳۷

استاد درس: دکتر حامد آذرنوش

نیمسال بهار ۹۹–۰۰

سؤال ١

سؤال (١) :

$$\text{first-order derivative (gradient)} = \begin{cases} \dfrac{\partial f(x,y)}{\partial x} = f(x+1,y) - f(x,y) \\ + \\ \dfrac{\partial f(x,y)}{\partial y} = f(x,y+1) - f(x,y) \end{cases}$$

$$\xleftrightarrow{F.T.} juF(u,v) + jvF(u,v) = e^{ju}F(u,v) - F(u,v) + e^{jv}F(u,v) - F(u,v)$$

$$= F(u,v)\left[e^{ju} + e^{jv} - 2\right]$$

$$\text{second-order derivative} = \begin{cases} \dfrac{\partial^2 f(x,y)}{\partial x^2} = [f(x+1,y) - f(x,y)] - [f(x) - f(x-1)] = f(x+1,y) + f(x-1,y) - 2f(x,y) \\ + \\ \dfrac{\partial^2 f(x,y)}{\partial y^2} = f(x,y+1) + f(x,y-1) - 2f(x,y) \end{cases}$$

$$\xleftrightarrow{F.T.} \begin{cases} -u^2F(u,v) = e^{ju}F(u,v) + e^{-ju}F(u,v) - 2F(u,v) \\ + \\ -v^2F(u,v) = e^{jv}F(u,v) + e^{-jv}F(u,v) - 2F(u,v) \end{cases}$$

$$\Rightarrow -F(u,v)(u^2+v^2) = F(u,v)\left[\underbrace{e^{ju} + e^{-ju}}_{=2\cos u} + \underbrace{e^{jv} + e^{-jv}}_{2\cos v} - 4\right]$$

سؤال ۲

الف) اگر آرایش مکانی پیکسل‌ها تغییر کند، باتوجه به تفاوت دوره‌ی تناوب تصویر و فیلتر، بخش‌هایی از تصویر و فیلتر که روی‌هم قرار می‌گیرند دچار تغییر شده و مکان پیکسل‌های تصویر نهایی تغییر می‌کند.

ب) اعوجاج یا wraparound در تصویر یا سیگنال هنگامی اتفاق می‌افتد که دو سیگنال متناوب با یکدیگر کانوالو شده و سیگنال متناوب دیگری به‌وجود آورند و چون دوره‌ی تناوب نزدیکی دارند، روی‌هم افتادن این سیگنال‌های کانوالوشده باعث اعوجاج می‌شود. برای مثال اگر دو تابع داشته باشیم که دارای نمونه‌های A و B هستند، با اضافه‌کردن صفر به آن‌ها طوری‌که هردو طول یکسان P را بیابند، داریم:

$$P \geq A + B - 1$$

این کار درواقع پدینگ از نوع صفر است. پس برای اعمال فیلترهای فرکانسی، قبل از آن با استفاده از zero padding فیلتر را در تصویر کانوالو یا cross-correlation می‌کنیم.

رابطه‌ی زیر برای دو آرایه‌ی دوبعدی f(x,y) و h(x,y) با ابعاد AxB و CxD است. می‌خواهیم برای اینکه هنگام کانوالوشدن مشکل اعوجاج رخ ندهد، از پدینگ صفر استفاده کنیم:

$$f_p(x, y) = \begin{cases} f(x, y) & 0 \leq x \leq A-1 \quad \text{and} \quad 0 \leq y \leq B-1 \\ 0 & A \leq x \leq P \quad \text{or} \quad B \leq y \leq Q \end{cases}$$

$$h_p(x, y) = \begin{cases} h(x, y) & 0 \leq x \leq C-1 \quad \text{and} \quad 0 \leq y \leq D-1 \\ 0 & C \leq x \leq P \quad \text{or} \quad D \leq y \leq Q \end{cases}$$

$$P \geq A + C - 1$$

$$Q \geq B + D - 1$$

تصویر نهایی پس از پدینگ، ابعاد PxQ دارد.

سؤال ۳

۱. با استفاده از دستور dcmread در کتابخانه‌ی dicom فایل‌ها را می‌خوانیم. نوع تصویر این فایل‌ها pixel data است و نمی‌توان عملیاتی را که روی تصاویر (که نوع‌شان آرایه است) انجام داد. پس با دستور pixel_array آن‌ها را به آرایه تبدیل می‌کنیم.

۲ و ۳. قبل از تبدیل داده‌های فایل‌ها به آرایه، از آن‌ها خروجی می‌گیریم تا اطلاعات درون‌شان را مشاهده کنیم. یکی از اطلاعات درون فایل‌ها، بیت‌های اختصاص‌یافته و ذخیره‌شده است:

```
(0028, 0100) Bits Allocated                    US: 16
(0028, 0101) Bits Stored                       US: 12
```

همان‌طور که دیده می‌شود، بیت‌های اختصاص‌یافته ۱۶ هستند (مطابق با نوع داده‌ی uint16) و ساختار آرایه ۱۶ بیتی است اما تصاویر ما ۱۲ بیتی هستند. پس L برابر با $۲^{۱۲}$ خواهد بود و بیشترین میزان شدت روشنایی $۲^{۱۲}-۱$.

٤. تبدیل فوریه‌ی یک‌بعدی و دوبعدی متناوب هستند. هنگامی که روی یک تصویر تبدیل فوریه اعمال می‌کنیم، بخش‌هایی از طیف که فرکانس صفر دارند در ابتدا و انتهای آرایه و بخش‌هایی که بیشترین فرکانس را دارند در وسط آرایه قرار می‌گیرند. چنین تصویری اطلاعات و دید مناسبی برای درک و تفسیر تصویر به ما نمی‌دهد؛ برای همین نیاز داریم تا بخش‌های با فرکانس صفر را در وسط تصویر قرار دهیم. اگر M و N ابعاد تصویر ما باشند، ما می‌خواهیم فرکانس‌های صفر (یا همان مقدار DC سیگنال) را در نقطه‌ی M/2 و N/2 قرار دهیم. این کار را باید با خاصیت شیفت در تبدیل فوریه انجام دهیم:

$$f(x.y)e^{j2\pi(\frac{u_0x}{M}+\frac{v_0y}{N})} \Leftrightarrow F(u-u_0.v-v_0)$$

می‌بینیم که در حوزه‌ی فرکانس، نقاط u و v منتقل شده‌اند.

آرایه‌ی mask آرایه‌ای متشکل از اعداد ۱- و ۱ است که در هر ردیف به‌صورت یکی‌درمیان قرار گرفته‌اند. می‌دانیم که فرم نمایی عدد مختلط $e^{j\pi x}$ همان $(-1)^x$ است. پس درواقع آرایه‌ی mask با ضرب‌شدن در تصویر ما، نقطه‌ی F(0,0) را با شیفت به‌اندازه‌ی M/2 و N/2، در مرکز تصویر قرار می‌دهد. یعنی مقدار dc در مرکز قرار می‌گیرد.

$$f(x.y)(-1)^{x+y} \Leftrightarrow F(u-M/2.v-N/2)$$

۵. شدت‌های img1 همگی مثبت هستند پس وقتی در محدوده‌ی L-1 – 0 نشان می‌دهیم تصویر درست را به ما می‌دهد. وقتی تصویر در mask ضرب می‌شود، شدت‌های منفی خواهیم داشت. پس باید در محدوده‌ای آن را نشان دهیم تا مقادیر منفی را هم شامل شود (L-1- تاL-1). بنابراین عکس پایین سمت چپ stretched و عکس بالا سمت راست clipped شده است.

٦. با استفاده از دستورهای np.abs و np.angle به‌ترتیب دامنه و فاز تصویرمان را که اکنون در حوزه‌ی فرکانس است به‌دست می‌آوریم.

۷. طبق رابطه‌ی زیر که عکس تبدیل فوریه است، اندازه‌ی f(x,y) را با قدرمطلق‌گرفتن محاسبه می‌کنیم. در این رابطه، دو مجموع داریم که از صفر تا M-1 و N-1 که طول و عرض تصویر هستند محاسبه می‌شوند که همان img.size است. با ضرب این اندازه در L-1 که بیشترین شدت روشنایی‌مان است، بیشترین شدت روشنایی برای اندازه‌ی طیف فرکانس تصویر به‌دست می‌آید.

$$f(x,y) = \frac{1}{MN} \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} F(u,v) \exp(j2\pi(\frac{ux}{M} + \frac{vy}{N}))$$
$$x = 0, 1, 2, \ldots, M\text{-}1 \qquad y = 0, 1, 2, \ldots, N\text{-}1$$

۸. مقدار dc طیف تصویر انرژی زیادی دارد و بخش غالب طیف فوریه را تشکیل می‌دهد. همان‌طور که در خروجی دیده می‌شود، تمام تصویر اندازه‌ی طیف فوریه تیره است. لگاریتم‌گرفتن از اندازه باعث می‌شود که تصویر از حالت فشرده‌ی شدت‌های تیره دربیاید و جزییات آن بهتر دیده شود.

در حالت کلی، تبدیل فوریه‌ی یک تصویر که شامل دامنه و فاز می‌شود، اطلاعاتی را درباره‌ی موج‌های سینوسی تشکیل‌دهنده‌ی تصویر به ما می‌دهد. طیف دامنه اطلاعات درمورد شدت و انرژی سینوس‌ها می‌دهد و طیف فاز مکان این سینوس‌ها را نشان می‌دهد. با استفاده از طیف فاز می‌توانیم تغییراتی (translation) روی تصویر ایجاد کنیم و آن تغییرات را روی طیف فاز هم مشاهده کنیم. برای مثال، اگر تصویر را a درجه بچرخانیم، طیف فاز هم همان‌قدر می‌چرخد و تغییر آن قابل‌مشاهده است.

۹. قرینه‌ی مزدوج یک عدد مختلط، بخش حقیقی آن را قرینه کرده و بخش موهومی آن را ثابت نگه می‌دارد. در تبدیل فوریه، مزدوج تابع در حوزه‌ی فرکانس باتوجه به مختلط‌بودن و زوج یا فرد بودن تابع قوانینی دارد.

۱۰. در اینجا مزدوج تصویر در حوزه‌ی فرکانس محاسبه شده است. یعنی مقادیر موهومی قرینه شده‌اند. طبق قانون تقارن هرمیت یا conjugate-symmetry، هنگامی که تابع ما در حوزه‌ی زمان (مکان) حقیقی باشد، مزدوج آن درحوزه‌ی فرکانس برابر است با:

$$F(-x.-y) = F^{*}(x.y)$$

وقتی که مزدوج تصویر را در حوزه‌ی فرکانس حساب می‌کنیم، علامت j که در رابطه‌ی عکس تبدیل فوریه مثبت بود، منفی می‌شود. طبق رابطه‌ی زیر می‌توان دید که این علامت منفی پشت ورودی‌ها قرار گرفته و رابطه‌ی بالا نتیجه می‌دهد:

$$\begin{aligned} F^{*}(u,v) = \left[\sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x,y)e^{-j2\pi(ux/M+vy/N)}\right] &= \sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f^{*}(x,y)e^{j2\pi(ux/M+vy/N)} \\ &= \sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x,y)e^{-j2\pi([-u]x/M+[-v]y/N)} \\ &= F(-u,-v) \end{aligned}$$

یعنی تصویر ما با برگشتن به حوزه‌ی مکان، نسبت به محورهای x و y قرینه می‌شود. حال تصویر حاصل (img1) در mask ضرب می‌شود و تمام شدت‌های تصویر مثبت می‌شود. سپس دستور clip مقادیر کمتر از صفر را صفر و بیشتر از L-1 را L-1 می‌کند. تصویر حاصل همان تصویر چرخانده‌شده است.

برای img2 مزدوج را قرینه می‌کنیم. تصویر همچنان طبق قوانین تقارن هرمیت نسبت به محورهای عمودی و افقی قرینه می‌شود اما در حوزه‌ی زمان، شدت‌های تصویر دقیقاً قرینه‌ی حالتی که img1 یافت خواهد بود. پس وقتی در mask ضرب می‌شود، تمام مقادیر آرایه منفی خواهد شد. با دستور clip مقادیر کمتر از L-1- به L-1– و مقادیر مثبت به صفر تنظیم می‌شوند. حال تصویر با L-1 جمع می‌شود. از آنجا که مقادیر شدت همه منفی هستند، با جمع‌شدن با L-1 مکمل آن‌ها نسبت به L-1 در حالتی که شدت بین صفر تا L-1 است تبدیل می‌شود؛ پس شدت‌های عکس همگی مکمل شده و تصویر نهایی ما نگاتیو تصویر اصلی خواهد شد.