

سوال اول) عوامل موثر روی روشنایی پیکسل:

میزان روشنایی منبع نوری که به صحنه می تابد: $r(x, y)$

میزان روشنایی که توسط اشیا موجود در صحنه بازتاب میشود: $i(x, y)$

$F(x, y) = i(x, y) * r(x, y)$ مشخصه ی محیط عددی بین ۰ تا ۱ و میزان بازتاب عددی بین ۰ تا بی نهایت است.

منبع: جزوه ی استاد

سوال دوم:

یه ایده ای که فکر میکنم توی ایران انجام نشده اما در خارج از ایران در مسابقات بوده کار کردن روی تصاویر پزشکی مثل رادیولوژی و فیزیوتراپی و ... هست که در مرحله ی ساده تر خیلی میتونه به بیماران کمک کنه تا در شناخت بیماری و عارضه ای که دارن اطلاعات کسب کنند و در مراحل بالاتر و پیشرفته تر به تشخیص پزشک کمک کنه. این مسابقات تحت عنوان ImageCLEF در زمینه ی VQA-Med هستند که اگر دیتاستش رو در ایران تهیه کنیم، با توجه به قدرت فوق العاده ی پزشکان ایرانی میتونیم پیشرو باشیم. در واقع به این صورت هستش که ما باید یه سری تصاویر همراه با پرسش و پاسخ های مشخص تهیه کنیم که این سوالات به چند دسته تقسیم میشن. سوالاتی که در مورد نوع بیماری هستند، سوالاتی که در مورد نوع تصویر برداری هستند، سوالاتی که در مورد جهت تصویر برداری هستند مثلاً از بالا گرفته شده یا از پهلو، سوالاتی که در مورد عضو بدن که درگیر بیماری شده هستند. با توجه به اینکه نوع سوالات و توقع ما چی باشه سیستمی که طراحی میکنیم میتونه متفاوت باشه و برای تولید پاسخ علاوه بر بینایی ماشین به دانش NLP هم نیاز هست.

یک ایده هم که در مورد موارد قضایی به ذهنم میرسه این هست که یک مجموعه داده از تصاویر مجرم ها در زمان بازجویی اولیه و نوع جرمی که انجام دادن تهیه بشه و هنگامی که پلیس به یه نفر مذنون هست با دادن تصویرش و جرمی که مدنظرش هست به این برنامه بتونه بفهمه حدسش درسته یا نه.

سوال سوم:

الف) در حسگر آرایه ای در یک لحظه تمام مقادیر صحنه ی مورد نظر عکاس ثبت میشود. در حسگر خطی چندین سنسور به اندازه ی پیکسل های یک ردیف رزولوشن مورد نظر قرار دارد که اطلاعات یک خط از نقاط کنار هم را همزمان ثبت میکنند و با چند درجه چرخیدن به اندازه ی یک پیکسل پایین می آید و دیتای ردیف بعدی را ثبت میکند. عملکرد حسگر آرایه ای بهتر است ولی گرانتر است. البته ضعف عملکرد حسگر خطی را میتوان تا حدودی با پردازش تصویر برطرف کرد و حرکت مکانیکی آن هم با دقت بالا قابل کنترل هست. تعداد جابجایی حسگر آرایه ای برای تصویر ۱۰۰۰ در ۱۰۰۰ یک مرتبه و برای حسگر خطی ۱۰۰۰ یعنی به تعداد ردیف ها می باشد. سرعت تصویربرداری با حسگر آرایه ای بیشتر از حسگر لحظه ای است. با توجه به نکاتی که بالاتر ذکر شد.

ب) تصویر آسیاب بادی در حال چرخش با حسگر آرایه ای با توجه به اینکه در یک لحظه کل نقاط ثبت میشوند مثل تصویر داده شده می افتاد یعنی انگار ساکن است و حرکتی ندارد.

اما با حسگر خطی نقاط اطراف پره ها تار و نا واضح می افتاد چون پره ها در حال چرخش هستند و حین ثبت هر ردیف تا ردیف بعدی نقاط پره حرکت میکنند برای همین یک نقطه از ممکن است در چند ردیف پیکسل ثبت شود و در تصویر مشخص میشود که آسیاب در حال حرکت است.

منبع: جزوه و صحبت های استاد در جلسه ی اول

سوال چهارم: OpenCV

۴,۱ نصب OpenCV:

مرحله ی اول نصب numpy:

```
E:\01021\ComputerVision\HW1>pip install numpy
Requirement already satisfied: numpy in c:\users\lenovo\appdata\local\programs\python\python38\lib\site-packages (1.22.2)
```

مرحله ی بعد نصب خود opencv:

```
E:\01021\ComputerVision\HW1>pip install opencv-python
Collecting opencv-python
  Downloading opencv_python-4.6.0.66-cp36-abi3-win_amd64.whl (35.6 MB)
    ----- 35.6/35.6 MB 1.6 MB/s eta 0:00:00
Requirement already satisfied: numpy>=1.14.5 in c:\users\lenovo\appdata\local\programs\python\python38\lib\site-packages (from opencv-python) (1.22.2)
Installing collected packages: opencv-python
Successfully installed opencv-python-4.6.0.66
```

من چون از jupyter notebook استفاده میکنم در آن هم این دستور را اجرا کردم:

```
[12] !pip install opencv-python

Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-wheels/public/simple/
Requirement already satisfied: opencv-python in /usr/local/lib/python3.7/dist-packages (4.6.0.66)
Requirement already satisfied: numpy>=1.14.5 in /usr/local/lib/python3.7/dist-packages (from opencv-python) (1.21.6)
```

۴,۲ خواندن تصویر background.png با تابع cv2.imread در ۳ کانال و توضیح پارامتر های تابع:

پارامتر اول اسم فایل تصویر همراه پسوند آن است البته اگر تصویر در محل فایل پایتون برنامه نباشد باید آدرس آن را بدهیم.

پارامتر دوم یک flag است که تعیین میکند تصویر را باید چگونه بخواند. این پارامتر به صورت پیش فرض برابر با مقدار cv2.IMREAD_COLOR است که در واقع تصویر را رنگی و در سه کانال طوری که عمق هر درایه ۸ بیت است نشان میدهد اگر flag بزرگتر از ۱ باشد هم معادل همین است. شفافیت تصویر را هم نادیده میگیرد.

اگر مقدار آن برابر cv2.IMREAD_GRAYSCALE یا صفر باشد هم تصویر درحالت خاکستری و یک کانال با عمق ۸ بیت میباشد.

اگر مقدار آن برابر cv2.IMREAD_UNCHANGED یا منفی باشد هم تصویر را به عنوان کانال آلفا بارگیری میکند. یعنی تصویر را ۴ کاناله میکنه اگر تصویر ۱۶ یا ۳۲ بیت عمق داشته باشد آن را با عمق ۱۶ یا ۳۲ برمیگرداند وگرنه به ۸ بیتی تبدیل میکند.

بقیه ی سوالات در فایل HW1 پاسخ داده شده اند. از منابع قرار داده شده در فایل سوالات برای پاسخ ها استفاده شد.