بسم الله الرحمن الرحیم

تمرین سری ۲

بینایی ماشین

خانم دکتر هدی محمد زاده

امیرحسین رستمی ۹۶۱۰۱۶۳۵

بهار ۹۹

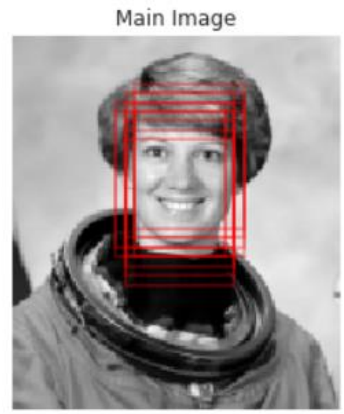
بخش اول:

ـ تمارین کامپیوتری:

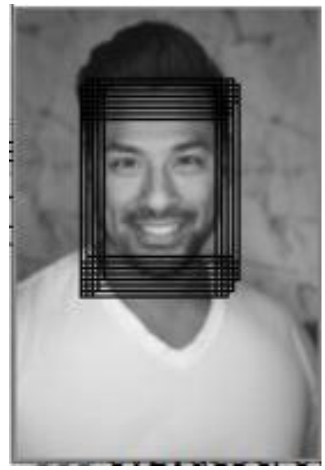
از آنجاکه احتمالا محوریت تی ای محترم در بخش اول پیاده سازی "کدی" است و از آنجاکه مباحث استفاده شده کاملا در کلاس درس تدریس شده بودند(ازجمله pyramid) لذا از توضیح دادن آن ها پرهیز می‌کنم و مستقیم سراغ خروجی های‌کد می روم.در جز به جز کد کامنت های زیادی گذاشته ام‌که کاملا هر مرحله و الزام وجودش رو توضیح داده است و لذا مصحح محترم را به خواندن‌کامنت های‌کد سوق می دهم و در اینجا فقط خروجی هارا می آورم.

خروجی اصلی فضانورد و عکس های ضمیمه شده:

در سطر اول خروجی بدون اعمال non-maximum-suppression و در خط بعدی خروجی آن پس از اعمال این فیلتر است.



ادامه ی عکس ها:

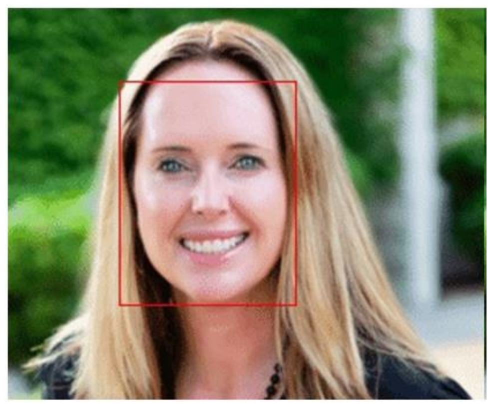


اطلاعی راجع به متد non-maximum-suppression:

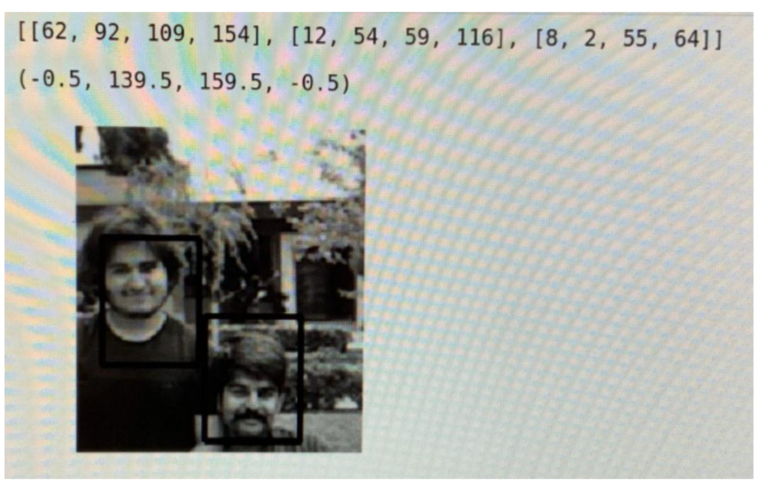
1ـ بنده برای این متد 2 پیاده سازی مختلف در کد ضمیمه شده ارایه کردم که یکی 100 برابر دیگری سرعت داشت!!!

2ـ توضیحات دقیق این متد و نحوه ی عملکردی آن در کد خط به خط آورده شده است و برای اینکه حوصله ی مصحح از خواندن گزارش سر نرود(با توجه به فیدبک مصححین) بنده در این بار مستقیم سراغ خروجی های پرسیده شده رفتم و از دادن توضیحات اضافه پرهیز کردم.(توضیحات کد به صورت کامل کامل در قالب کامنت در کد ضمیمه شده موجود است).

ادامه ی تصاویر در صفحه ی بعد.



و یه عکس دلخواه دیگر از گالری گوشی!!



همانطور که مشاهده می کنید جزییات پیاده سازی های انجام شده به خوبی انجام شده و خروجی گرفته شده نیز از دقت و کیفیت خوبی برخوردار است.

بخش اول:

ـ تمارین دستی:

پاسخ سوال اول:

Face alignment که به تشخیص نقاط اختصاصی صورت (face landmark) اشاره دارد که برای خیلی از کاربرد های مربوط به موارد صورت کاربرد دارد به مانند تشخیص چهره و تایید چهره اشاره کرد.

تشخیص چهره مفهوم ساده ای دارد, تایید چهره امری است که ما دو عکس صورت را با هم مقایسه کرده و به این بررسی می پردازیم که آیا این دو چهره یکی می باشند یا خیر. در face frontalization عکسی که در اختیار ما قرار گرفته است ممکن است در هر جهتی باشد هدف این است که صورت ها را تشخیص داده و با نگاشتی اثر دوران را از بین باید ببریم.

این روش برای تشخیص جنسیت و صورت کاربردهای فراوانی دارد و یک پیش پردازش برای این امور محسوب می شود.

به طور کلی اگر بخواهیم چند روش را نام برده برای این کار استفاده از مرز ها برای اینکار، متد های برپایه هیت مپ کانولوشنی و روش های مبتنی بررگرسیون را نام برد.

روش مبتنی بر لبه یابی:براساس صحبت های گفته شده مرزهای صورت یک ساختار جغرافیایی خوش حالت صورت است که به سادگی میتوانیم آنرا معرفی کنیم و خیلی تشخیص آن نسبت به تشخیص facial landmark آسان تر است.

روش مبتنی برلبه یابی در مقاله دارای دو مرحله است ابتدا ما لبه های صورت را heatmap آنرا مشخص کرده و سپس از روی آن مکان های facial landmark را استخراج می کنیم.

برای اینکه اثر لبه ها مشخص شود می توان به این اشاره کرد که در آزمایش های انجام شده هرچه دقت لبه یابی بهتر باشد facial landmarkها به صورت بهتری مشخص می شوند.

اگر بخواهیم کمی دقیق تر شویم بعد تولید هیت مپ لبه ها، در گام بعدی که استخراج landmark میباشد ما از هیت مپ به عنوان یک کمک کننده ویژگی برای landmark regressor استفاده می کنیم. برای استفاده حداکثری از آن ما هیپ مپ لبه ای را در چند مرحله در شبکه landmark regression استفاده می کنیم تحقیقات نشان داده است هرچه که در مراحل بیشتری از هیت مپ لبه ای استفاده شود در یادگیری ویژگی نتیجه خروجی بهتر می شود.

مدل رگرسیون هماهنگ:

در این روش ما به صورت مستقیم از مپ کردن تصاویر ورودی به بردار های مرتبط به landmark ها استفاده می کنیم. اقای zhang (طبق اشاره ی مقاله)این مسئله را به عنوان یک مسئله یادگیری multi task مطرح کرده است، که همزمان مختصات landmarkهای صورت و ویژگی های آن را همزمان یاد می گیرد.این شبکه اولین شبکه عصبی بازگشتی end to end برای face alignment از ریز تا درشت است.

TSRصورت را به چندین قسمت تقسیم کرده برای این که این کار تمیز دادن قسمت ها و regress کردت مختصات قسمت های مختلف صورت را راحت می کند. هرچند که مدل های رگرسیون هماهنگ برتری واضح بودن پیش پردازش های مربوط به مختصات landmark هارا دارند، ولی به اندازه مدل های رگرسیون هیت مپ خوب عمل نمی کنند.

پاسخ سوال دوم:

این پاسخ از مقاله ی دکتر Bruno Tunjic که در فایل پاسخ تمرین ضمیمه شده است استخراج شده است.

در این مقاله روش های متعددی از قبیل:

1- Bilateral filtering

2- modified bilateral filtering

3- recursive bilateral filtering

4- joint bilateral filtering

5- Adaptive bilateral filtering

6- Guided image filtering

7- Adaptive guided image filtering

8- Anisotropic diffusion

9- kuwahara filter

و ...

مطرح شده است و در این چند صفحه بنده به توضیح اجمالی تعدادی از آن ها می پردازم.

روش های Edge preserving smoothing به طور کلی باعث می شوند که تصویر در مکان هایی که لبه نیستند نرم شود،این روش ها ناحیه های همگن را نرم می کنند در حالی که لبه ها بدون تغییر باقی می مانند که این باعث می شود اشیا تصویر حتی بعد از انجام فیلترینگ قابلیت تمایز قرار دادن از هم را داشته باشند،برای مثال فیلتر های میانگین گیر با وزن کاملا ثابت را در نظر بگیرید،از آنجا که وزن نسبت داده شده به پیکسل ها برابر است لذا این چنین فیلتر ها باعث Bluring می گردند و مرز ناحیه هارا از بین می برند و لذا فیلترهای میانگین گیر با وزن برابر موردی مناسبی برای این کار نیستند.حال با استناد به روش های مطرحی در مقاله به بیان چند فیلتر که این کار را انجام می دهند می پردازیم:

بدیهتا در هر روشی که چنین کاری انجام می گردد باید به طریقی وزن های نایکسان به پیکسل های رنگی نسبت داده شود.چون اگر وزن برابر نسبت داده شود بدتر باعث blurring می گردد.

Bilateral filtering_1 :

در این روش کرنل فیلتر برخلاف کرنل میانگین گیر که وزن برابری به پیکسل ها نسبت می دهد، به نوعی وزن دار عمل می کند و اوزان نسبت داده شده به دو پارامتر شدت و فاصله از پیکسل مرکزی وابسته اند، فرض کنید که I_p شدت مقدار پیکسل p باشد و w_k پنجره کرنل ما میباشد که به مرکزیت پیکسل k می باشد،حال فیلتر bilateral را به صورت زیر تعریف می گردد:

$$BLF(I)_p = \frac{1}{\sum_{q\in w_k} W_{BLF_{pq}}(I)} \sum_{q\in w_k} W_{BLF_{pq}}(I) I_q$$

که مقدار پارامتر $W_{BLF_{pq}}$ (که تابع وزن ما است)به صورت زیر تعریف می گردد:

$$W_{BLF_{pq}}(I) = exp\left(\frac{-\|p-q\|^2}{2\sigma_s^2}\right) exp\left(\frac{-|I_p-I_q|^2}{2\sigma_r^2}\right)$$

همانطور که از ظاهر شبه گوسی این وزن ها برداشت می شود داریم که وزن ها با مقدار نابرابر و متاثر از فاصله ی رنگی و اقلیدسی پیکسل ها است.سیگما های موجود در مخرج کسر ها نیز برای کنترل سرعت کاهش در فضای فاصله پیکسل ها اند.به نوعی انگار ما در این سوال متریک های 5 بعدی که سه متر از آن فاصله ی رنگی پیسکل ها و دو متر از آن فاصله ی اقلیدسی دو پیکسل درنظرگرفتیم.در حالی کهσ_r همین کار را در کنترل تاثیر اختلاف شدت دارد.

به این پارامتر ها پارامتر های "tuning" نیز به خاطر اینکه در هر دو حوضه تاثیر دارند می گویند. عملگر ||.|| نشانگر فاصله Euclidean میباشد که این نشان می دهد هرچه فاصله دو پیکسل بیشتر باشد تاثیر آن کمتر شود. برای شدت پیکس ها هم هرچقدر اختلاف آنها بیشتر باشد تاثیر کمتر میباشد که در نتیجه در لبه ها که اختلاف شدت به شدت زیاد می باشد با اینکار باعث میشود که پیکسل ها دست نخورده باقی بماند و لبه ها حفظ بشوند.

بدین ترتیب با نسبت دادن اوزان با ویژگی بالا به پیکسل ها به ویژگی Edge Preserving Smoothing مدنظر،نزدیک تر می شویم.

اما...

ولی این روش یک مشکلی دارد این است که اگر نویز ما به صورت نمک فلفل طور باشد (یعنی یک پیسکل روشن در بین پیکسل هایی تاریک) پس با روش گفته شده این مشکل می باشد که اگر در مرکزیت فیلتر این نویز قرار بگیرد نویزحفظ می شود و ممکن است تشدید یابد،لذا نیاز است که به طریقی این مشکل را برطرف کنیم.

2_1:فیلتر اصلاح شده:

در این اصلاحیه طبق گفته ی مقاله هر پیکسل در پنجره فیلتری W_x با یک تابع مینیمم هزینه که مسیر آن ها را به پیکسل x متصل می کند معرفی شده است. هزینه مسیر برابر است با مجموع هزینه خونه های همسایه که یک مسیر را تشکیل می دهند است.
در اینجا هزینه اتصال قدرمطلق تفاصل دو پیکسل میاشد. مینیمم هزینه هر مسیر توسط روش dijksta algorithm محاسبه میشود این یعنی هر پیکسل در کرنل با یک مسیر با مینیمم هزینه به پیسکل مرکزی x متصل میاشند بعد تابع هزینه برای این استفاده میشود که وزن آنها در فیلترینگ مشخص شود و خروجی فیتلر به صورت یک فیلتر وزن دار از پیکسل های احاطه کننده

پیکسل x می باشند.هم چنین در این متد، وزن ها به جای اینکه به اختلاف شدت دو پیکسل مربوط باشد به روشی بازگشتی براساس احتلاف شدت هر دو پیکسل مجاور در مسیر محاسبه می شوند.

2_1:در ادامه نیز یک روش دیگر به نام joint bilateral filtering در مقاله معرفی می گردد که با روش هایی،جزییات بیشتری از آن استخراج می کنیم و مسیر یابی و وزن یابی مان را بهبود می دهیم.

روش دوم:Guided image filtering

در این روش ما از دو تصویر(یکی تصویر اصلی و دیگری تصویر هدایت گر) نیز استفاده می کنیم. در این روش ما دو تصویر داریم یک تصویر اصلی که با I آنرا معرفی میکنیم و یک توصیر هدایت گر که آن را با G نمایش می دهیم.
I_p , G_p برابر شدت پیکسل رنگی p در این دو تصویر است.فیلترینگ را در پنجره w_k که یک پنجره با مرکزw است انجام می دهیم. فیلترینگ به صورت زیر بدست می آید

$$GIF(I)_p = \frac{1}{\sum_{q \epsilon w_k} W_{GIF_{pq}}(G)} \sum_{q \epsilon w_k} W_{GIF_{pq}}(G)\, I_q$$

و ماتریس اوزان نیز به شرح زیر است:

$$W_{GIF_{pq}}(G) = \frac{1}{|w|^2} \sum_{k:(p,q) \epsilon w_k} \left(1 + \frac{(Gp-\mu_k)(Gq-\mu_k)}{\sigma_k^2+\varepsilon}\right)$$

که μ_k, σ_k واریانس و میانگین تصویر هدایت گر در پنجره مشخص شده است و w تعداد پیکسل های در پنجره میباشد. به ϵ پارامتر نرم کننده می گویند و همانند کمک کننده ی واریانس برای نرم تر کردن خروجی است. در این روش هم به دلیل حفظ کردن فاصله ی 5 بعدی پیکسل ها،داریم که ویژگی Edge preserving Smoothing حفظ می گردد.

نکته:برای اعمال روش فوق برای ماتریس های رنگی،لازم است تا این فیلترینگ هارا جداگانه روی تک تک کانال ها انجام بدهیم و در نهایت حاصل به دست آمده از سه کانال را با هم به نحوی درست ترکیب کنیم تا در نهایت خروجی برای حالت رنگی نیز محاسبه گردد.

Subject:
Date

✓ سؤال سوم:

می‌دانیم که هنگام کانوالو کردن ابتدا ماتریس فیلتر دوران سطری و سپس ستونی می‌گردد و سپس با انتقال مرکز ماتریس‌ها و ضرب داخلی آن‌ها عکس نتیجه محاسبه می‌گردد؛ داریم:

$$A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \xrightarrow[\text{ستونی}]{\text{دوران سطری}} \begin{bmatrix} i & h & g \\ f & e & d \\ c & b & a \end{bmatrix}$$

(حال جاروب می‌زنیم روی پیکسل‌ها و ضرب داخلی می‌کنیم)

ابتدا با اندکی بازی با محاسبات روی پیکسل‌های غیر مرزی درمی‌یابیم که فیلتر ما فیلتر تمام ۱ میانگین‌گیر است

۱۲	۶	۴	۲	۳
۴	۶	۴	۱۲	۶
۴۵	۴	۱	۶	۳
۳۴	۵	۱	۷	۳

اعمال فیلتر میانگین‌گیر →

$$\begin{bmatrix} 86 & 95 & 41 \\ 104 & 46 & 43 \end{bmatrix}$$

$$F = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

← حال به تست کردن انواع Border Handling ها می‌پردازیم و در نهایت داریم که روش استفاده شده Border_WRap است. (با آزمون و خطای روش‌های موجود برای Border-Type).

PAPCO

Subject:
Date

سؤال چهارم :

$$\begin{bmatrix} u \\ v \end{bmatrix} = \underbrace{\underbrace{\begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}}_{\text{ماتریس دوران}} \underbrace{\begin{bmatrix} x \\ y \end{bmatrix}}_{\text{مختصات اولیه}}}_{\text{مختصات دوران یافته}}$$

$$\begin{pmatrix} u = x\cos\theta + y\sin\theta \\ v = x(-\sin\theta) + y\cos\theta \end{pmatrix}$$

حکم: $\underbrace{\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}}_{\text{حالت اولیه}} = \overbrace{\frac{\partial^2 f}{\partial u^2} + \frac{\partial^2 f}{\partial v^2}}^{\text{دوران یافته}}$

$$\frac{\partial f}{\partial y} = \frac{\partial f}{\partial u} \times \frac{\partial u}{\partial y} + \frac{\partial f}{\partial v} \times \frac{\partial v}{\partial y} = \frac{\partial f}{\partial u}(\sin\theta) + \frac{\partial f}{\partial v}(\cos\theta)$$

$$\frac{\partial^2 f}{\partial y^2} = \frac{\partial}{\partial y}\Big(\underbrace{\frac{\partial f}{\partial u}(\sin\theta) + \frac{\partial f}{\partial v}(\cos\theta)}_{\checkmark}\Big) = \frac{\partial(\checkmark)}{\partial u} \times \frac{\partial u}{\partial y} + \frac{\partial(\checkmark)}{\partial v} \times \frac{\partial v}{\partial y}$$

$$= \left(\frac{\partial^2 f}{\partial u^2}\sin\theta + \frac{\partial^2 f}{\partial u \partial v}(\cos\theta)\right)(\sin\theta) +$$

$$\left(\frac{\partial^2 f}{\partial u \partial v}(\sin\theta) + \frac{\partial^2 f}{\partial v^2}(\cos\theta)\right)(\cos\theta)$$

$$= \frac{\partial^2 f}{\partial v^2}\cos^2\theta + \frac{\partial^2 f}{\partial u^2}\sin^2\theta + \frac{\partial^2 f}{\partial u \partial v} 2\sin\theta\cos\theta$$

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial u} \times \frac{\partial u}{\partial x} + \frac{\partial f}{\partial v} \times \frac{\partial v}{\partial x} = \underbrace{\frac{\partial f}{\partial u}(\cos\theta) + \frac{\partial f}{\partial v}(-\sin\theta)}_{\checkmark}$$

$$\frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x}(\checkmark) = \frac{\partial(\checkmark)}{\partial u} \times \frac{\partial u}{\partial x} + \frac{\partial(\checkmark)}{\partial v} \times \frac{\partial v}{\partial x}$$

$$\frac{\partial^2 f}{\partial x^2} = \left(\frac{\partial^2 f}{\partial u^2} \cos\theta + \frac{\partial^2 f}{\partial u \partial v} (-\sin\theta) \right) \cos\theta + \left(\frac{\partial^2 f}{\partial u \partial v} \cos\theta + \frac{\partial^2 f}{\partial v^2} (-\sin\theta) \right) (-\sin\theta)$$

$$= \frac{\partial^2 f}{\partial u^2} \cos^2\theta + \frac{\partial^2 f}{\partial v^2} \sin^2\theta + \frac{\partial^2 f}{\partial u \partial v} \, (-2\sin\theta\cos\theta)$$

$$\Rightarrow \left(\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = \frac{\partial^2 f}{\partial u^2} + \frac{\partial^2 f}{\partial v^2} \right)$$

سوال پنجم :

توجه کنید که هیستوگرام اطلاعات فضایی ماتریس را در اختیار ما نمی‌گذارد لذا قضاوت‌های سوال در نظر گرفتن شرایط خاص ممکن است ، اما اگر فرض کنیم که حداقل یکی از ماتریس‌ها ثابت باشد داریم که برای موارد مورد پرسش سوال می‌توانیم پاسخ را فراهم کنیم حال فرض می‌کنیم حداقل یکی از ماتریس‌ها ثابت باشد بدون از دست دادن کلیت مسئله فرض می‌کنیم $g(x,y)$ ثابت است و مقدار $g(x,y) = c$! دارد حال داریم که فرض می‌کنیم پیکسل رنگی r_k در $f(x,y)$ تعداد $h_f(r_k)$ را داشته باشد (ارتفاع ستون‌های هیستوگرام پیکسل رنگی r_k) حال با فرض‌های ذکر شده به حل هر قسمت می‌پردازیم:

اگر u_k را شدت رنگی پیکسل متناظر r_k در خروجی هر اپراتور بنامیم، داریم که:

در الف) $f(x,y) + g(x,y)$ داریم :

$$\begin{cases} u_k = r_k + c \\ h_{f+g}(u_k) = h_f(r_k) \end{cases}$$

Subject:
Date

در این حالت هیستوگرام $f+g$ همان هیستوگرام g است که به اندازه c مقدار محور شدت رنگ آن شیفت یافته است (به سمت راست)

* نکته: در حالت جمع دو تصویر با هم یا همان حالت الف دلیل آن این است که اگر pdf دو عکس f, g مستقل از هم باشند (pdf همان توزیع رنگ پیکسل های f, g مستقل از هم باشد) در این صورت داریم که طبق تئوری های آمار، برای جمع دو توزیع مستقل داریم که pdf حاصل برابر با کانولوشن pdf های f, g خواهد بود:

$$\begin{cases} f, g \text{ independent} \\ h_{sum} = h_f * h_g \end{cases}$$

$f, g \to$ تابع چگالی توزیع رنگی پیکسل‌ها در تصاویر

نکته مهم: این مسئله کانولوشن برای زمانی است که f, g مستقل از هم باشند.

ب) در این حالت اگر g را ثابت بگیریم داریم:

بجای $f(x,y)-g(x,y)$ داریم:

$$\begin{cases} u_k = r_k - c \\ h_{f-g}(u_k) = h_f(r_k) \end{cases}$$

توجه کنید همانند نکته خاص با ۲ در اینجا نیز اگر مستقل بگیریم f, g را (به عبارت دیگر تابع توزیع رنگی پیکسل‌ها f, g مستقل از هم باشند) در اینجا نیز داریم:

$$\begin{cases} f, g: \text{independent} \\ h_{sum} = h_f * (-h_g) \end{cases}$$

PAPCO

Subject:
Date

ادامه ب) در حالت ب هیستوگرام $f-g$ مانند هیستوگرام f خواهد بود و به اندازه

c در راستای محور شدت رنگی نمودارها به سمت چپ c واحد شیفت می‌خورند.

ج) در این حالت هیستوگرام $f*g$ فاصله‌ی بین مستطیل‌های هیستوگرام‌ها، c برابر می‌گردد.

$$\begin{cases} u_k = r_k \times c \\ h_{f*g} = h_f \end{cases}$$

د) در این حالت در هیستوگرام $f \div g$ فاصله‌ی بین مستطیل‌های هیستوگرام‌ها، $\frac{1}{c}$ برابر می‌گردد.

$$\begin{cases} u_k = r_k \div c \\ h_{f \div g} = h_f \end{cases}$$

در همه‌ی پاسخ سؤال ۵، u_k شدت رنگ پیکسل متناظر r_k پس از انجام تابع $\left.\begin{matrix} f+g \\ f-g \\ f*g \\ f \div g \end{matrix}\right\}$ در ماتریس رنگی حاصل است.

. ـــ . ـــ . ـــ . ـــ .

PAPCO