



دانشگاه صنعتی امیرکبیر
(پلی تکنیک تهران)

درس بینایی کامپیوتر

استاد درس جناب آقای دکتر صفابخش

(تمرین سری چهارم)

محسن عبادپور | ۴۰۰۱۳۱۰۸۰ | m.ebadpour@aut.ac.ir

نیمسال اول سال تحصیلی ۱۴۰۱-۱۴۰۲



فهرست پاسخ ها

- مسئله اول: تقطع تصویر با K-Means و نتیجه غیر مناسب آن (الف) ۳
- مسئله دوم: تقطع تصویر با GLCM و مبتنی بر K-Means (ب) ۵
- مسئله سوم: تقطع تصویر با Gabor و مبتنی بر K-Means (ج) ۱۲
- مسئله چهارم: مقایسه و بررسی تقطع تصاویر با Gabor و GLCM مبتنی بر K-Means (د) ۱۷

مسئله اول

مسئله اول. تقطع تصویر با K-Means و نتیجه غیر مناسب آن (الف)

استفاده از الگوریتم های خوشه‌بندی خصوصا خوشه‌بند های Partitioning مانند K-Means به تنهایی نمیتوانند تصاویر را تقطیع و مخصوصا تقطیع معنایی کنند و این نقطه ضعف جایی بیشتر نمایان می شود که گستره‌ی رنگی اشیا متفاوت بوده یا یک شی در بافت خود دارای چندین رنگ مختلف باشد.

دلیل این ناتوانی در distance محور بودن الگوریتم و عدم تاثیر دادن بافت و ساختار مکانی پیکسل ها در امر تقطیع می باشد چرا که اگر دو یا چند رنگ متضاد در بافت یک شی وجود داشته باشد، فاصله‌ی محاسبه شده برای آن دو رنگ (دو پیکسل) از هم زیاد حاصل شده و هر یک در خوشه‌ی جداگانه قرار خواهد گرفت و بدین صورت پیوسته شی کاملا زیر سوال خواهد رفت. در هر کدام از دو تصویر داده شده نیز این نقص به صورت در صورت تقطیع کاملا قابل رویت است.

Segmentation using K-Means with Normalization | Time:1.0



مثلا اگر بخواهیم تصاویر را در سه خوشه‌ی مجزا تقطیع معنایی کنیم (آسمان، زمین و حیوان) خواهیم دید که نتیجه کاملاً به دور از انتظار واقع می‌شود. در مثال گور خر بواسطه‌ی فاصله‌ی زیاد رنگ‌های بافت بدن آن (سیاه و سفید)، به عنوان یک شی قطعاً بندی نشده است. در مثال ببر نیز، نتیجه کاملاً بدور از انتظار واقع شده است بطوریکه حتی با آگاهی قبلی به سختی می‌توان وجود حیوان را شناسایی کرد و ملاحظه نمود.

Segmentation using K-Means with Normalization | Time:1.59



مسئله دوم

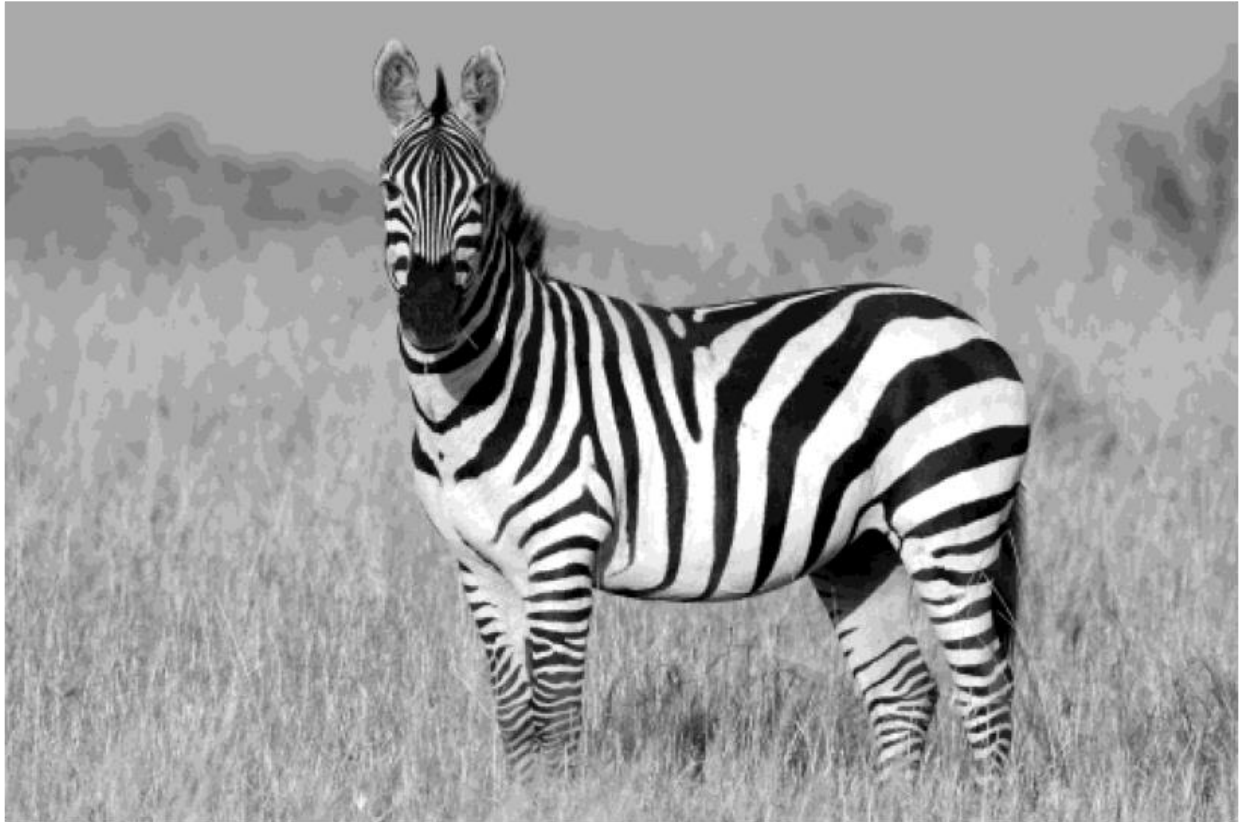
مسئله دوم: تقطع تصویر با GLCM و مبتنی بر K-Means (ب)

برای پیاده سازی این بخش سه تابع پیاده سازی شده است که به ترتیب استفاده مورد شرح قرار می گیرد؛ تابع اول PatchMakers می باشد که تصویر ورودی و ابعاد مورد نظر را دریافت کرده و با استفاده از تابع view_as_windows اقدام به ایجاد patches کرده و نتایج را بر میگرداند. تابع دوم GLCMCalculator می باشد که با استفاده از تابع greycmatrix و ورودی گرفتن پارامترهای زاویه (که ۰-۳۶۰ در نظر گرفته شده) و ورودی گرفتن فاصله از پیکسل کنونی اقدام به محاسبه و ایجاد GLCM می کند؛ سپس با استفاده از چهار معیار ("energy", "correlation", "contrast", "dissimilarity") و تابع greycoprops اقدام به تولید بردار ویژگی به ازای هر پیکسل می کنیم.

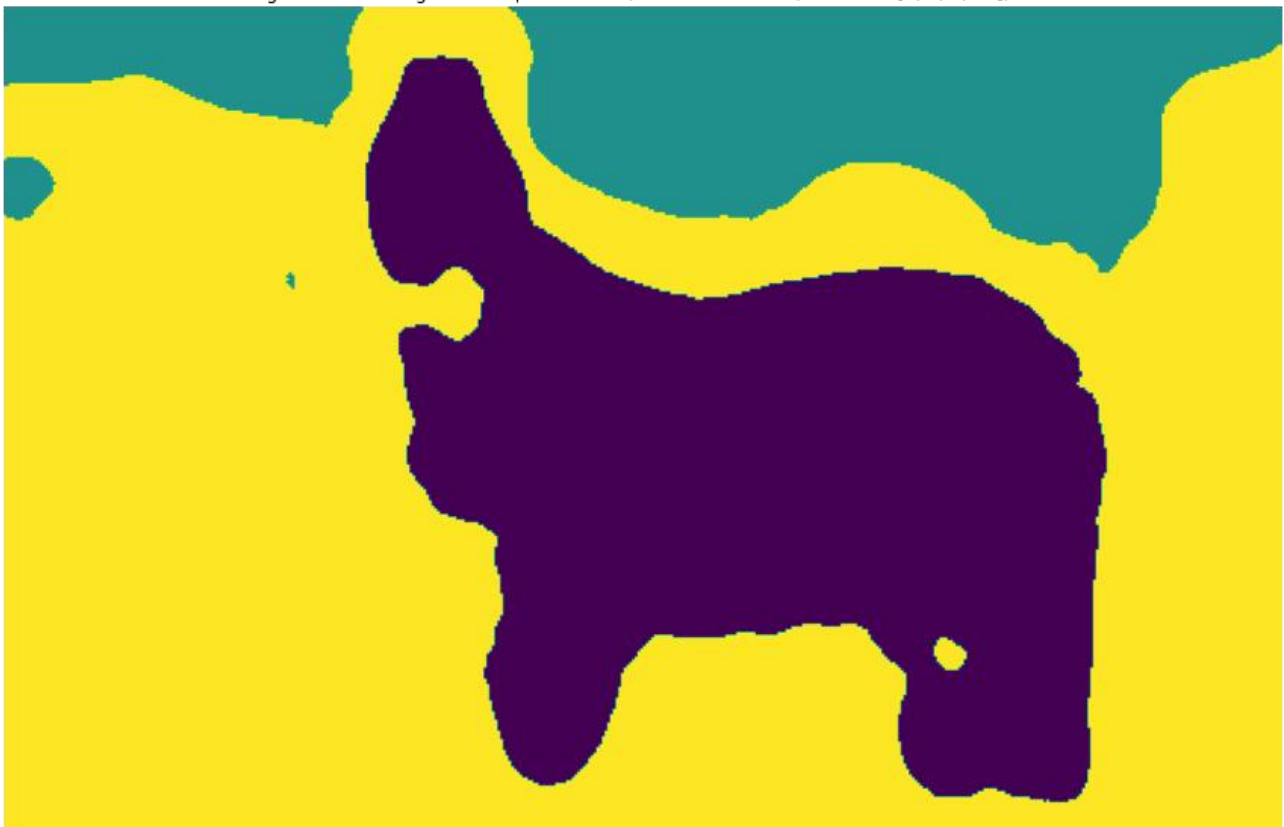
تابع سوم Segmentation می باشد که تصویر ورودی را با نسبت نصف کاهش سایز داده و سپس سطوح خاکستری آن را به ۱۶ سطح کاهش می دهد که هدف بالا رفتن سرعت پردازشی و مصرف کم منابع می باشد؛ سپس با استفاده از دو تابع فوق بردارهای ویژگی را تولید می کند ولی از جایی که فرآیند تولید بردار ویژگی به ازای هر پیکسل از هم مستقل بوده و قابلیت موازی سازی دارد، از تابع Parallel استفاده کرده و این پروسه بصورت موازی انجام داده می شود. در نهایت بردارهای ویژگی نرمال شده (حساس بودن الگوریتم های خوشه بندی مبتنی بر فاصله به مقیاس) و سپس با استفاده از الگوریتم K-Means تصویر ورودی تقطیع معنایی می شود. همچنین چندین فعالیت پیش پردازشی نیز میتواند انجام داد؛ مانند هموارسازی تصویر با هدف ایجاد پیوستگی در تقطیع و حذف نویز یا افزودن حاشیه برای اجتناب از عدم تقطیع پیکسل های مرزی در تصویر اصلی.

برای هر کدام از تصاویر چندین آزمایش انجام شده است تا بهترین نتیجه حاصل شود که در قسمت عنوان تصاویر مقادیر بدست آمده برای پارامترها درج شده است. برای هر کدام از تصاویر و پارامترها دو بار تقطیع شده است که یکی خوشه بندی سه کلاسی (زمین، آسمان و حیوان) می باشد و یکی هم خوشه بندی دو کلاسی (حیوان و غیر حیوان) در نظر گرفته شده است. برای تصویر گورخر دو نتیجه برتر آورده شده است که یکی شبیه ترین حالت به نمونه آورده شده در متن سوال می باشد و دومی تقطیعی است که به جهت بصری توانسته بدن گورخر را کامل تر تقطیع کند. (فارغ از سایر درستی سایر اجزا)

Input Image | Levels:16, Resize Ratio:0.5



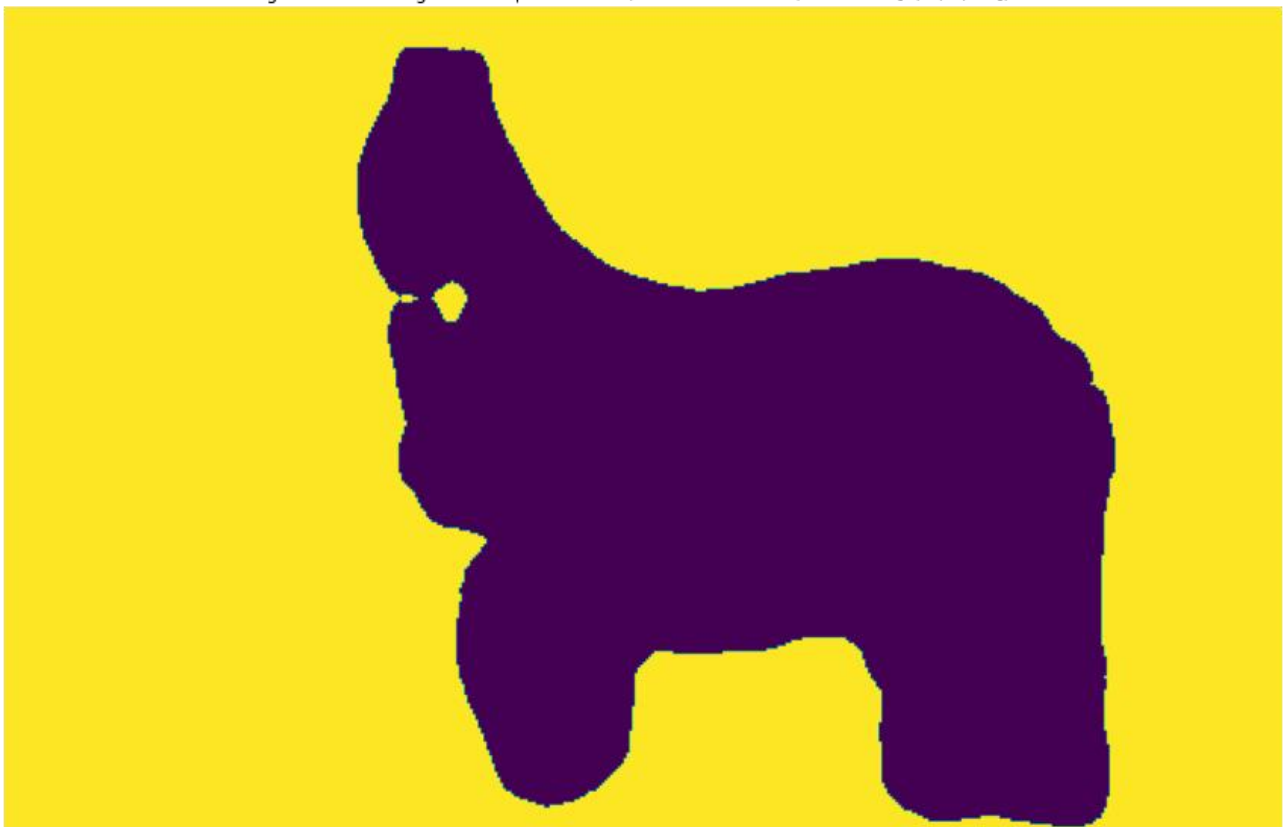
GLCM based Segmentation using K-Means | Workers:60, Patch Size:35x35, Distances[1, 5, 9, 17], Duration:19.02s



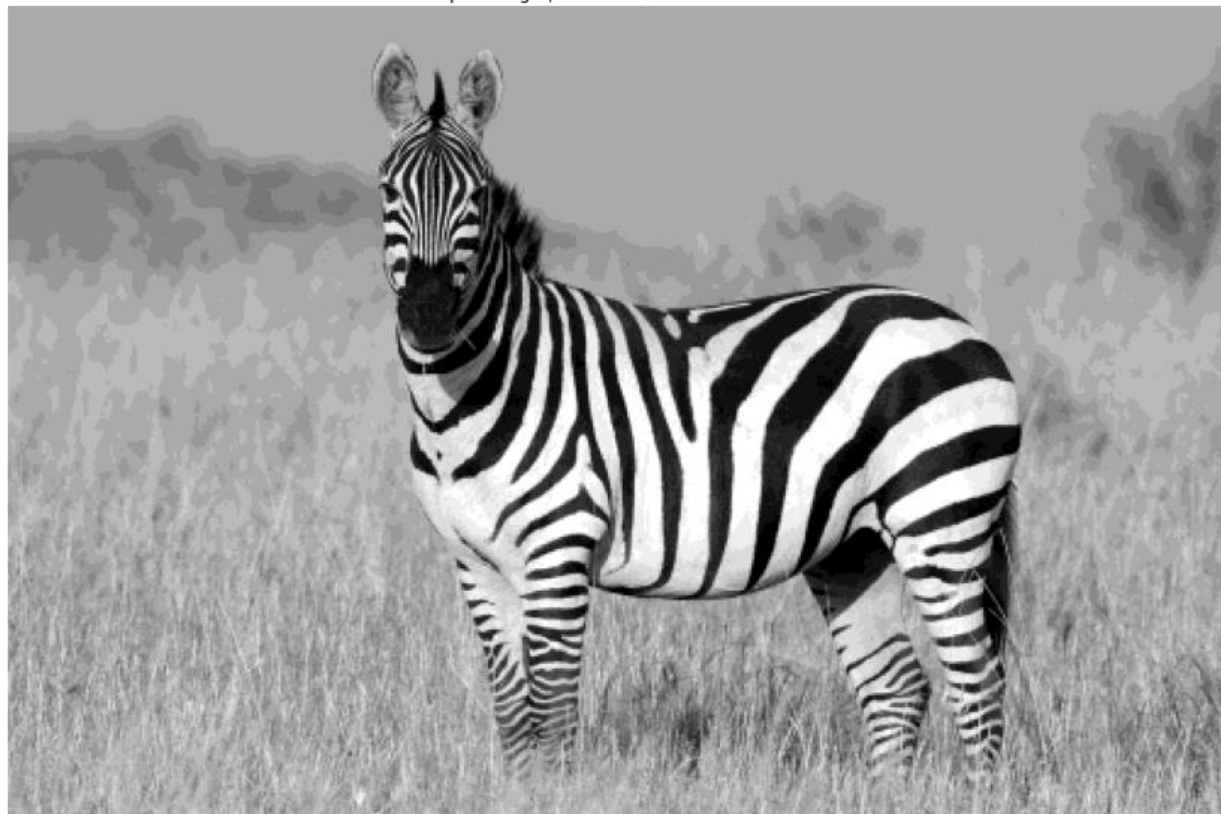
Input Image | Levels:16, Resize Ratio:0.5



GLCM based Segmentation using K-Means | Workers:60, Patch Size:35x35, Distances[1, 5, 9, 17], Duration:19.37s



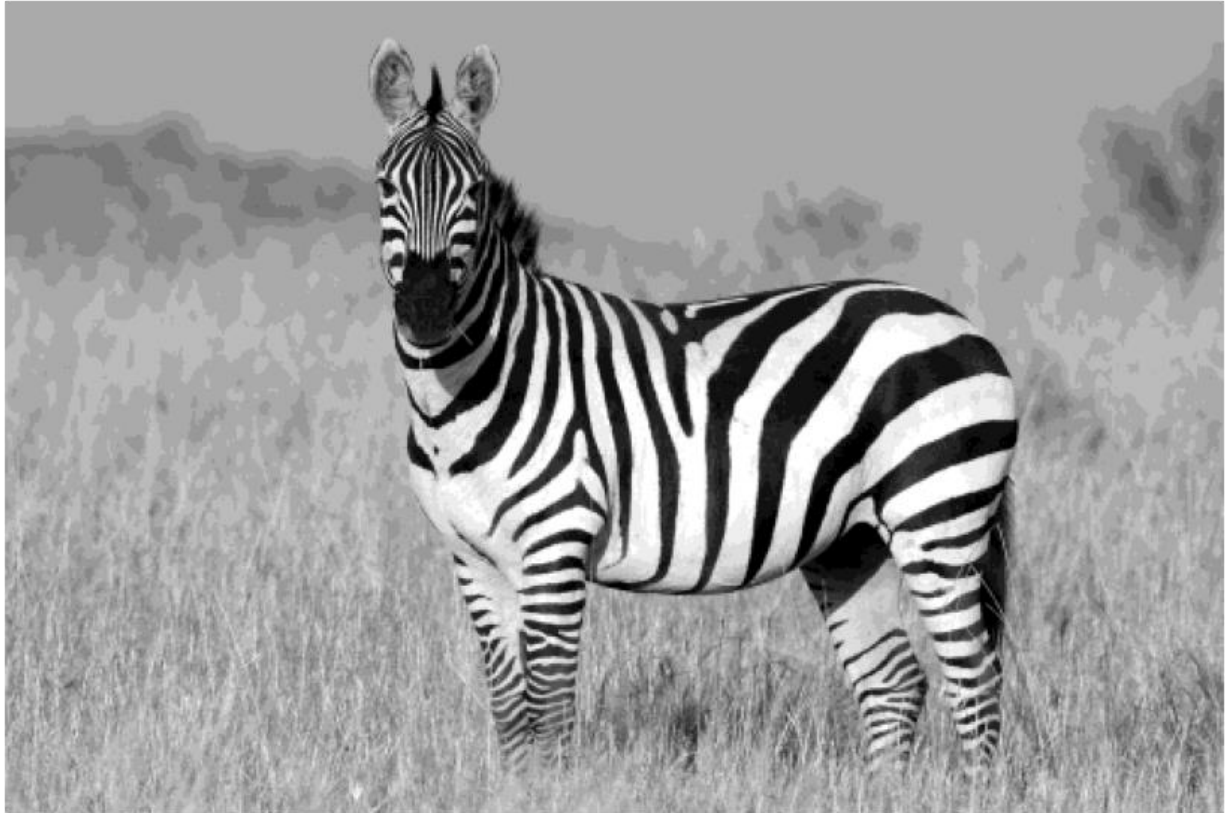
Input Image | Levels:16, Resize Ratio:0.5



GLCM based Segmentation using K-Means | Workers:60, Patch Size:45x45, Distances[1, 2, 3, 4, 5, 6, 7, 8, 9], Duration:47.28s



Input Image | Levels:16, Resize Ratio:0.5



GLCM based Segmentation using K-Means | Workers:60, Patch Size:45x45, Distances[1, 2, 3, 4, 5, 6, 7, 8, 9], Duration:41.6s



Input Image | Levels:16, Resize Ratio:0.5



GLCM based Segmentation using K-Means | Workers:60, Patch Size:40x40, Distances[1, 2, 3, 4, 5, 6, 7, 8, 9], Duration:45.53s



Input Image | Levels:16, Resize Ratio:0.5



GLCM based Segmentation using K-Means | Workers:60, Patch Size:40x40, Distances[1, 2, 3, 4, 5, 6, 7, 8, 9], Duration:43.3s



مسئله سوم: تقطع تصویر با Gabor و مبتنی بر K-Means (ج)

برای پیاده سازی این بخش نیز مانند قسمت قبل سه تابع پیاده سازی شده است که به ترتیب استفاده مورد شرح قرار می گیرد؛ تابع اول GaborFilters می باشد که تصویر ورودی و پارامترهای مد نظر را دریافت کرده و فیلترهای دو بعدی Gabor نرمال شده را ایجاد میکند. این پارامترها عبارت اند از درجه چرخش (۰-۱۸۰)، سایز فیلتر (برای گورخر: ۳۰، ۳۵، ۴۰ و برای ببر: ۵۰، ۵۵، ۶۰)، سیگما، لاندا و گاما. تابع دوم GenerateOutput نام دارد که هر یک از فیلترهای تولید شده در قسمت قبل را بروی تصویر ورودی اعمال کرده و بردار ویژگی هر پیکسل را به وجود می آورد.

تابع سوم GaborSegmentation نام دارد که از دو تابع فوق و هموارسازی گوسی استفاده کرده و بردارهای ویژگی نرمال شده (حساس بودن الگوریتم های خوشه بندی مبتنی بر فاصله به مقیاس) را به وجود می آورد؛ سپس با استفاده از الگوریتم K-Means تصویر ورودی تقطیع معنایی می شود. برای هر کدام از تصاویر چندین آزمایش انجام شده است تا بهترین نتیجه حاصل شود که در قسمت عنوان تصاویر مقادیر بدست آمده برای پارامترها درج شده است. برای هر کدام از تصاویر و پارامترها دو بار تقطیع شده است که یکی خوشه بندی سه کلاسی (زمین، آسمان و حیوان) می باشد و یکی هم خوشه بندی دو کلاسی (حیوان و غیر حیوان) در نظر گرفته شده است.

Input Image | Gaussian Blur(7x7,1)



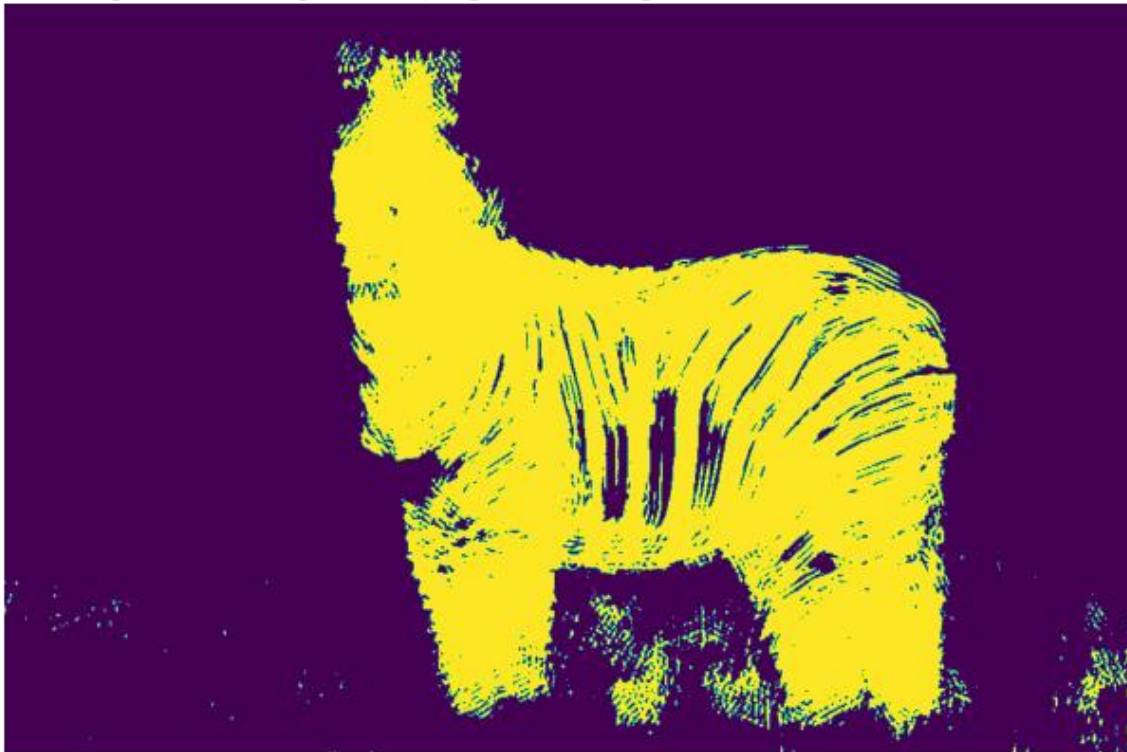
Gabor based Segmentation using K-Means | Degrees:0-180, Sigma:40, Lambda:10, Gamma:1.2, Duration:39.2s



Input Image | Gaussian Blur(7x7,1)



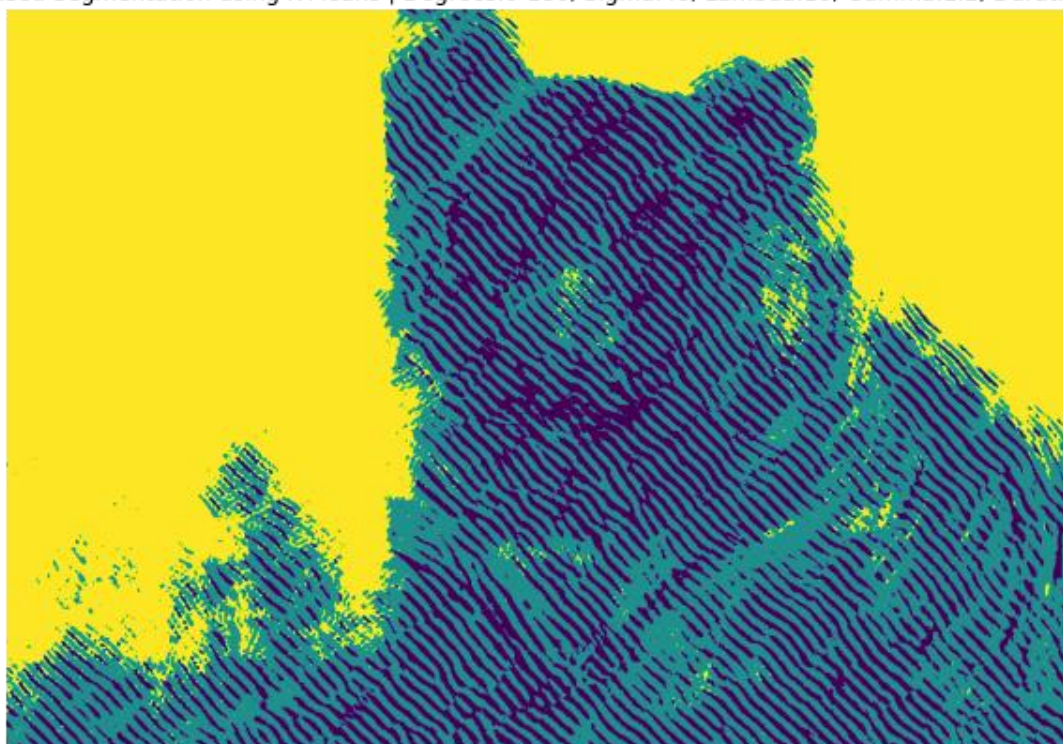
Gabor based Segmentation using K-Means | Degrees:0-180, Sigma:40, Lambda:10, Gamma:1.2, Duration:39.2s



Input Image | Gaussian Blur(7x7,1)



Gabor based Segmentation using K-Means | Degrees:0-180, Sigma:40, Lambda:10, Gamma:1.2, Duration:37.32s



Input Image | Gaussian Blur(7x7,1)



Gabor based Segmentation using K-Means | Degrees:0-180, Sigma:40, Lambda:10, Gamma:1.2, Duration:37.32s



مسئله چهارم

مسئله چهارم: مقایسه و بررسی تقطع تصاویر با Gabor و GLCM مبتنی بر K-Means (د)

در مرحله‌ی اول بررسی، ابتدا زمان محاسبات هر یک را مقایسه می‌کنیم؛ در بحث زمان پردازشی فیلتر Gabor با اختلاف بهتر از GLCM ظاهر شده است و میتواند بردار ویژگی را سریع تر خروجی دهد لذا در تصاویر با ابعاد بزرگ و با منابع پردازشی محدود، انتخاب Gabor در اولویت قرار دارد. (فارغ از دقت و نتیجه)

مورد بعدی بحث پیوستگی اشیا تقطیع شده در تصویر می‌باشد که GLCM به خاطر استفاده از ویژگی های محلی پیرامون هر پیکسل توانسته است پیوستگی را حفظ کند و این در وضعیتی است که در Gabor به هیچ عنوان پیوستگی قابل رویت نیست.

حال به ازای هر یک از تصاویر، دو روش یاد شده و نتایج شان را مقایسه می‌کنیم؛ در تصویر گورخر GLCM توانسته تقطیع معنایی مناسبی به ازای هر سه شی ارائه دهد و بصورت فاحش یکی فدای دیگری نشده است (زمین، آسمان و گورخر) و نواحی داخلی هر کدام از اشیا نیز پیوسته حاصل شده است و این در حالی است که در Gabor نتیجه حاصل شده به هیچ عنوان نتوانسته است آسمان و زمین را از هم تقطیع کند و بدن خود گورخر نیز پیوسته حاصل نشده است و در قسمت زیری نیز نوبز هایی نیز در تقطیع دیده می شود که حاکی از عدم محلی بودن تقطیع می باشد. در تصویر ببر نتیجه متفاوت است؛ Gabor تا حدودی بهتر عمل کرده و توانسته است صرفا محدوده‌ی کلی ببر را مشخص کند (هر چند غیر پیوسته) اما با این وجود در تقطیع آسمان و زمین مانند GLCM ناتوان بوده است. از جایی که GLCM بصورت محلی و از هر patch ویژگی استخراج می‌کند، با بزرگ کردن سائز پنجره به اندازه کافی میتوان ساختار بافت شی را از تصویر استخراج و برای تقطیع به کار برد لذا در عکس گورخر روش مذکور توانسته بهتر عمل کند؛ ولی در مقابل و در تصویر ببر چون بافت و ویژگی های محلی شی مورد نظر (خود ببر) بسیار شبیه به محیط پیرامون می‌باشد، ویژگی های استخراجی مناسب نبوده لذا تقطیع معنایی مناسب صورت نگرفته است.