

گزارش پروژه پایانی مبانی بینایی ماشین

مهدی نادری - ۹۶۳۶۷۱۳

برای حل مسئله مربوط به پروژه از روش Linear Programming استفاده شده است که ایده اصلی آن، استفاده از تشابه کناره‌های تصاویر می‌باشد. برای پیاده‌سازی هرچه بهتر راه حل مسئله، از مقاله‌هایی استفاده کرده‌ام که فایل pdf آن‌ها را ضمیمه پروژه کرده‌ام.

هر قسمت مورد نیاز از کدهای راه حل مسئله، در فایل جداگانه‌ای پیاده شده است اما برای دیدن نتیجه بر روی یک سری پازل لازم است تا فایل main.m اجرا شود. مسیر پازل‌ها نیز باید در خط ۷ همین فایل، به صورت دستی وارد شود. فرض را بر این گذاشته‌ام که نام‌گذاری فایل‌های مربوط به تصاویر که در این مسیر وجود دارند، به همان شکلی است که در پازل‌های مربوط به تست در اختیار داشتیم و تصویر اصلی و تصویر خروجی که قطعه‌های گوشه در آن قرار دارند نیز با همان شکل نام گذاری در این مسیر قرار دارند، در غیر این صورت عملکرد برنامه با مشکل مواجه خواهد شد.

در برخی قسمت‌ها، از منابع موجود در اینترنت برای پیاده سازی کد استفاده کرده‌ام.

نحوه عملکرد پروژه به شکل خیلی کلی به این صورت است که پس از خواندن تصاویر قطعات پازل‌ها، یک متود Linear Programming روی آن‌ها اجرا می‌شود که با کمینه سازی یک عبارت که به نوعی تشابه تصاویر مربوط به قطعات پازل را هنگامی که در حالت‌های مختلف در کنار همدیگر قرار می‌گیرند، می‌سنجد، سعی می‌کند تا قطعاتی که باید در جهت مناسب در کنار همدیگر باشند را در کنار یکدیگر قرار دهد. من نیز خروجی نهایی کد را همین خروجی که از این مرحله به دست می‌آید در نظر گرفته‌ام و در پایان، مقدار MSE بین این خروجی و تصویر اصلی را گزارش کرده‌ام.

برای کمینه سازی آن عبارت، از پکیج cvx استفاده کرده‌ام که همراه پروژه آپلود شده است و لازم است از این مسیر فایل cvx_setup اجرا شود و برای اینکه به درستی کار کند ظاهراً از نسخه متلب ۲۰۱۷ و بالاتر باید استفاده شود.

بلوک دیاگرام روش پیاده شده را مشاهده می کنید:

