

تشخیص شی با استفاده از گوشه یابی Harris و شبکه عصبی در فضای رنگ HSV

میثم شهسوار^۱، محمد رضا دلیری^۲

^۱ دانشجوی کارشناسی ارشد مکترونیک دانشگاه آزاد قزوین، m.shahsavari@qiau.ac.ir

^۲ استادیار دانشکده برق دانشگاه علم و صنعت، daliri@iust.ac.ir

چکیده - در این مقاله با استفاده از روش گوشه یابی Harris یک بردار ویژگی از یک تصویر ورودی تهیه می شود که شامل رنگ پیکسل های هشت همسایگی نقاط کلیدی در فضای رنگ HSV استخراج شده می باشد. این بردار ویژگی به یک هیستوگرام پنجاه ستونی تبدیل می شود و یک شبکه عصبی را تغذیه می کند، سپس از این شبکه عصبی برای کلاس بندی تصاویر ورودی استفاده می شود. کلید واژه - شبکه عصبی، Harris، HSV

می شود و در نهایت از این شبکه برای تشخیص تصویر ورودی استفاده می شود، حسن این روش سادگی و سرعت بالای آن به سبب استفاده از روش گوشه یابی Harris است. در این مقاله از ۳۲۰ تصویر در چهار کلاس برای آموزش شبکه استفاده شده که از تارنمای "آزمایشگاه بینایی ماشین" دپارتمان فن آوری اطلاعات و مهندسی برق دانشگاه ETH زوریخ استخراج شده است [6] و در هر کلاس ۸۰ تصویر وجود دارد. در این مقاله از رنگ نقاط ۸ همسایگی نقاط کلیدی در فضای رنگ HSV برای تشکیل بردار ویژگی استفاده شده است که حساسیت کمتری به تغییر نور نسبت به فضای RGB دارد. فضای HSV در یک مختصات استوانه ای تعریف می شود که زاویه از محور عمودی بیانگر رنگ، فاصله از محور عمودی اشباع را مشخص می کند و ارتفاع در طول محور عمودی مقدار یا روشنایی را نشان می دهد.

۲- گوشه یابی Harris

فرم کاشف نقطه کلیدی در تصویر به صورت زیر است:

$$R = \text{Det}(M) - k \times \text{Tr}(M)^2 \quad (1)$$

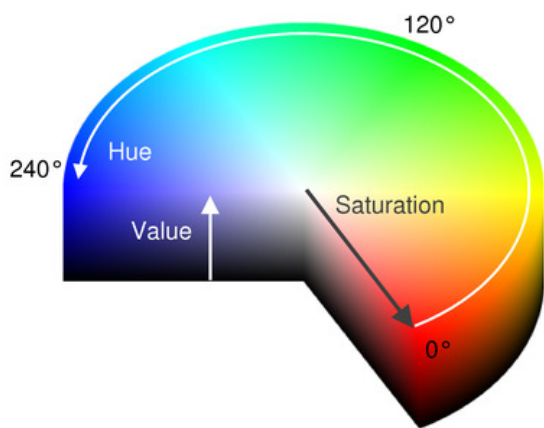
که k پارامتر از پیش تعیین شده و M اثر درجه دوم ماتریس است.

$$M = \begin{bmatrix} \left(\frac{\partial I}{\partial x}\right)^2 \otimes w & \left(\frac{\partial I}{\partial x}\right)\left(\frac{\partial I}{\partial y}\right) \otimes w \\ \left(\frac{\partial I}{\partial x}\right)\left(\frac{\partial I}{\partial y}\right) \otimes w & \left(\frac{\partial I}{\partial y}\right)^2 \otimes w \end{bmatrix} \quad (2)$$

۱- مقدمه

شناسایی اشیای یکی از پرکاربردترین مباحث حوزه بینایی ماشین است، استفاده از ویژگی های محلی که نسبت به تغییر نور، چرخش و بزرگ نمایی ثابت هستند برای تشخیص می تواند نتایج خوبی را تهیه کند [1]. برای انجام این کار نقاطی از تصویر انتخاب می شوند و سپس برداری از ویژگی های مربوط به آن نقاط (keypoints) تهیه می شود که می توان از آن ها برای تشخیص شباهت دو تصویر استفاده کرد.

یکی از روش های استخراج نقاط کلیدی ها استفاده از تکنیک SIFT است، این تکنیک نتایج خوبی را برای اشیای مشخص ایجاد می کند اما نمی تواند این نتایج را برای یک دسته از اشیای به خوبی تعمیم دهد [2,3]. در روش SIFT یک عکس به مجموعه ای از بردار های ویژگی محلی تبدیل می شود که به چرخش عکس، تغییر اندازه و محدوده ای از تغییرات نور حساس نیستند [4]. SIFT کارایی خوبی دارد ولی قدرت محاسباتی بالایی را نیاز دارد. گوشه یابی Harris هم یکی از روش های استخراج نقاط کلیدی است که مبتنی بر دترمینان و اثر درجه دوم ماتریس است که می تواند با سرعت و دقت بالایی گوشه ها را پیدا کند [5]. در این مقاله از روش گوشه یابی Harris به همراه شبکه عصبی استفاده شده تا سیستم بتواند تصاویری را که ابتدا به آن آموزش داده شده تشخیص دهد. روند کار به این صورت بعد از یافتن نقاط کلیدی ها یک بردار ویژگی از هر تصویر ساخته می شود که توسط آن شبکه عصبی آموزش داده

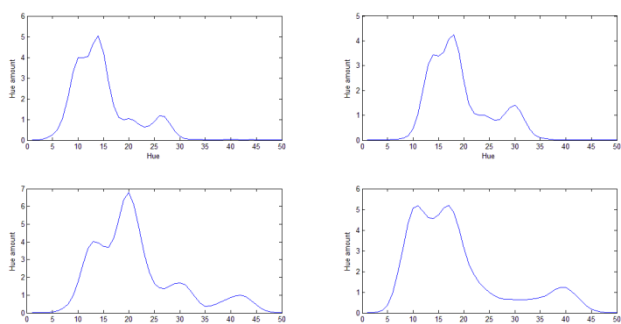


شکل ۱: سیلندر HSV

۴- بردار ویژگی

همان طور که در مقدمه اشاره شد هنگامی که از نقاط کلیدی استفاده می شود عکس به مجموعه ای از بردارهای ویژگی تبدیل می شود. در این مقاله ما با استفاده از ۸ همسایگی نقاط کلیدی استخراج شده از روش گوشه یابی Harris یک ماتریس که شامل ویژگی رنگ (Hue) آن نقاط است تشکیل دادیم. از آنجا که هدف ما آموزش یک شبکه عصبی بود ما نیاز داشتیم تعداد مشخصی ورودی برای شبکه عصبی مهیا کنیم ولی ماتریس ما دارای تعداد مشخصی عضو نبود، چون تعداد نقاط کلیدی در هر تصویر متغیر است. برای رفع این مشکل یک هیستوگرام ۵۰ ستونی از ماتریس ایجاد کردیم که بازه رنگ ها را به ۵۰ قسمت تقسیم می کرد و دارای فرمی مشابه برای هر کلاس از تصاویر بود.

زمان لازم برای استخراج هیستوگرام از ابتدای عملیات تا انتها بین ۳۰۰ تا ۳۵۰ میلی ثانیه با یک پردازنده Intel Core i7 720QM بود. با استفاده از روش SIFT پیاده سازی شده [8] نقاط کلیدی دو عکس از عکس های مورد استفاده در این مقاله استخراج شد که حدود ۱,۴ ثانیه با همان کامپیوتر قبلی زمان صرف شد.



شکل ۲: چهار نمونه از بردارهای ویژگی کلاس A

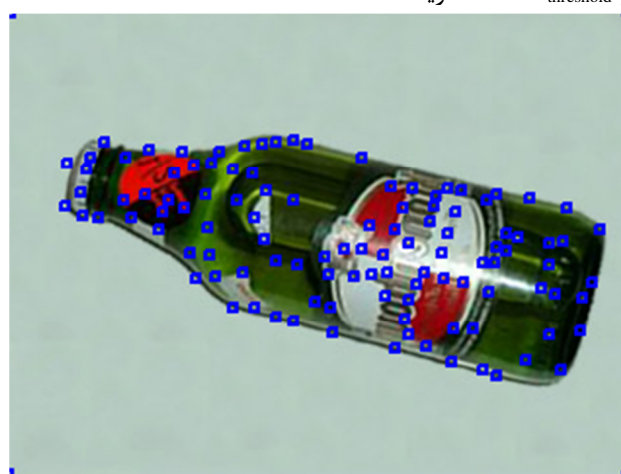
w وزن فیلتر گوسین را نشان می دهد و I بیان گر تابع شدت نور عکس است و \otimes بیانگر کانولوشن است.

این عملگر از تغییر متوسط سطح خاکستری یک پنجره مربعی که بین دو جهت عمودی حرکت می کند بدست می آید. گوشه ها بیشینه های محلی مقیاس Harris هستند که دارای آستانه ای بیشتر از مقدار مشخص هستند.

$$\{x_c\} = \{x_c | R(x_c) > R(x_i), \forall x_i \in w(x_c)\} \quad (3)$$

$$R(x_c) > t_{threshold}$$

$\{x_c\}$ مجموعه ای از تمام گوشه هاست، $R(x)$ مقیاس Harris است در x است، $W(x_c)$ هشت همسایگی حول x_c است، و $t_{threshold}$ آستانه تعریف شده است.



شکل ۱: یکی از عکس های مورد استفاده که توسط روش گوشه یابی Harris گوشه یابی شده. نقطه های آبی رنگ نقاط کلیدی استخراج شده هستند.

۳- فضای رنگ HSV

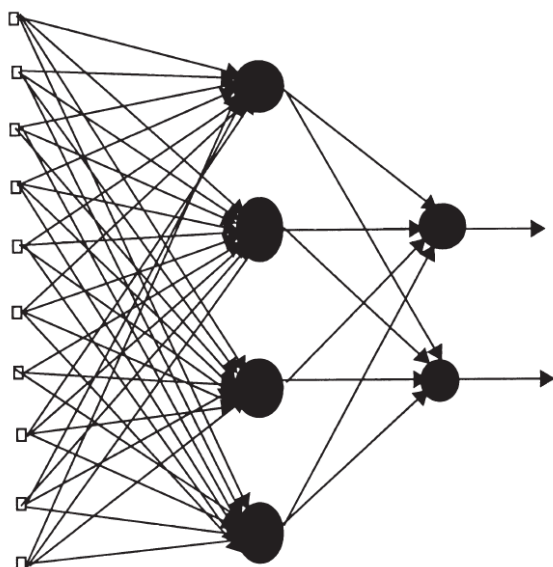
فضای رنگی HSV در یک مختصات استوانه ای تعریف می شود که رنگ در این مختصات متناسب با زاویه تغییر می کند، به این صورت که رنگ قرمز از زاویه ۰، سبز از زاویه ۱۲۰، آبی از ۲۴۰. محور عمودی رنگ های خاکستری را تشکیل می دهد، که از رنگ سیاه در پایین محور (مقدار ۰) تا رنگ سفید در بالای محور را شامل می شود.

فاصله شعاعی از مرکز مقدار اشباع را تعیین می کند [7].

انجام می شود. مرحله رو به جلو که در این مرحله پارامترهای آزاد شبکه مقدار دهی می شوند و سیگنال ورودی در شبکه منتشر می شود و سپس مقدار خطا از رابطه (۴) محاسبه می شود.

$$e_i = d_i - y_i \quad (4)$$

که در آن d_i خروجی مطلوب و y_i خروجی واقعی است. در مرحله دوم، مقدار خطا e_i در شبکه پس انتشار میابد و مقدار خطاها را تنظیم می کند [10,11].



شکل ۵: ساختار یک شبکه چند لایه.

برای پیاده سازی در این مقاله از یک شبکه عصبی با یک لایه مخفی استفاده شده است. ماتریس ورودی دارای ۳۲۰ ستون و ۵۰ ردیف، که ۷۰ درصد این مقدار برای آموزش، ۱۵ درصد برای آزمایش و ۱۵ درصد هم برای اعتبارسنجی استفاده شده است. معیار انتخاب تعداد نرون های شبکه عصبی آزمون و خطا بوده است. [12]. جدول زیر میانگین تشخیص شبکه عصبی را به ازای چهار مقدار متفاوت برای شبکه مخفی نشان می دهد.

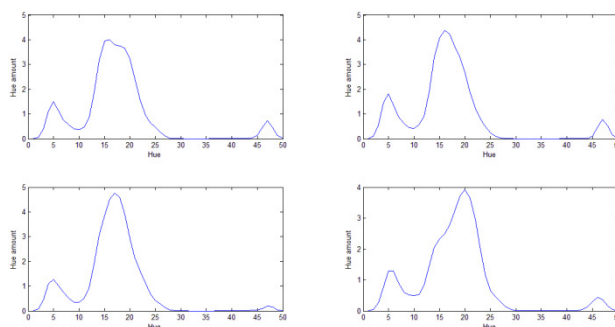
تعداد نرون ها	۸۰ نرون	۷۰ نرون	۶۵ نرون	۶۰ نرون
خطای شبکه	۲۵,۷٪	۱۳,۵٪	۶,۴۲٪	۱۹,۵٪
انحراف از معیار	۴,۵۸	۱,۶۷	۲,۱	۲,۱۷

جدول ۱

بر اساس جدول بدست آمده تعداد ۶۵ نرون برای لایه مخفی انتخاب شد که دارای کمترین خطا بود.

۶- نتیجه گیری

در این مقاله به روشی سریع برای کلاس بندی تصاویر با استفاده از شبکه عصبی و گوشه یابی Harris اشاره که شد در



شکل ۳: چهار نمونه از بردارهای ویژگی کلاس B

همانطور که اشاره شد بردار ویژگی استفاده شده در این مقاله شامل ۵۰ ستون از رنگ های نقاط ۸ همسایگی نقاط کلیدی است که توسط گوشه یابی Harris استخراج شده اند. بر اساس مشاهدات از هر عکس حدود ۲۲۰۰ تا ۲۴۰۰ نقطه استخراج شد که برای نمونه برداری مجزور این تعداد نقطه حدود ۵۰ بدست می آید. عکس هایی که برای پیاده سازی استفاده شده دارای پس زمینه ساده هستند و به جز شی مورد نظر چیز دیگری در تصویر وجود ندارد. تصاویر از نماهای مختلف، دارای چرخش، و مقدار کمی بزرگ نمایی و شرایط نور متفاوت هستند.



شکل ۴: نمونه ای از عکس های یک کلاس مورد استفاده.

۵- شبکه عصبی

بعد از تهیه بردار ویژگی باید یک شبکه عصبی را با استفاده از این داده ها آموزش دهیم تا بتواند بعد از یادگیری تصویر ورودی را شناسایی کند.

شبکه عصبی یک پردازنده موازی بزرگ است که میل به ذخیره سازی اطلاعات تجربی و استفاده از آن ها دارد [9].

یکی از الگوریتم هایی که برای شبکه های عصبی چند لایه استفاده می شود، الگوریتم پس انتشار خطاست که در دو مرحله

کاربرد هایی که تصاویر ورودی محدود هستند شباهت رنگ گسترده ای بین کلاس های مورد نظر وجود ندارد کار برد دارد. استفاده از فضای رنگ HSV از حساسیت به تغییرات نور می کاهد و گوشه یابی Harris باعث افزایش سرعت استخراج بردار ویژگی می شود. برای بهبود این روش می توان از الگوریتم های تشخیص شی استفاده کرد تا تصاویر ورودی محدود به تصاویری بدون پس زمینه شلوغ نباشند.

مراجع

- [1] H.Bay, A. Ess, T. Tuytelaar, and L. Van Gool: "Speeded-Up Robust Features (SURF)", *Computer Vision and Image Understanding* 110(3): 346-359, 2008.
- [2] S. Batra, "Multi-Class Object Recognition Using Shared SIFT Features", *Learning Report, Machine Learning Cs229*, 2007.
- [3] David G. Lowe, "Object recognition from local scale-invariant features" *International Conference on Computer Vision, Corfu, Greece* (September 1999).
- [4] X. Wu, Z. Shi, Y. Zhong, "Detailed Analysis and Evaluation of Keypoint Extraction Methods" *International Conference on Computer Application and System Modeling*, 2010.
- [5] Harris, C., Stephens, M. "A Combined Corner and Edge Detector" *Proceedings of 4th Alvey Vision Conference, Manchester*, pp. 147-151, 1998.
- [6] <http://www.vision.ee.ethz.ch/datasets/index.en.html>
- [7] Alvy Ray Smith "Color gamut transform pairs," *Computer Graphics*, 12(3):12-25, 1978.
- [8] Demo code for detecting and matching SIFT features, <http://www.cs.ubc.ca/~lowe/keypoints/>
- [9] Haykin S., "Neural networks: A comprehensive foundation, 2nd ed. ", *Prentice-Hall*, 1998.
- [10] Werbos, P. J "Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences" PhD thesis, Harvard University, 1974.
- [11] Rumelhart, D. E., Hinton, G. E., and Williams, R. J. "Learning internal representations by error propagation. In *Parallel Distributed Processing*", volume 1, pages 318-362. MIT Press, 1986.
- [12] Shin-ike, K. "A two phase method for determining the number of neurons in the hidden layer of a 3-layer neural network" *SICE Annual Conference*, 2010.