**دانشگاه زنجان**

دانشکده مهندسی

**پایان نامه کارشناسی**

گرایش : الکترونیک

**عنوان**:

# هیستوگرام گرادیان های شیب گرا برای آشکارسازی انسان

(Histogram of oriented gradients for human detection)

استاد راهنما :

دکتر مصطفی چرمی

نگارش :

پریناز اسکندری

بهار 93

## فهرست مطالب:


چکیده ..... 1

کلمات کلیدی ..... 2

مقدمه ..... 4

گام های محاسبه بردار ویژگی تصویر ..... 5

بردار ویژگی نهایی ..... 9

ماشین بردار پشتیبان ..... 10

شبیه سازی ها با استفاده از نرم افزار متلب..... 12

نتیجه گیری ..... 21

منابع و مراجع ..... 22

## چکیده :

برای آشکارسازی انسان درتصویر روش های متعددی وجود دارد،که یکی از این روش ها مبتنی بر هیستوگرام گرادیان های شیب گرا است. در این مقاله نشان می دهیم که چگونه میتوان حضور یا عدم حضور انسان را البته در تصاویری با تعداد اندک، با این روش تشخیص داد. ایده کلی بدین صورت است که با گرفتن گرادیان از شدت نور تک تک پیکسل های تصویر می توان یک بردار ویژگی از تصویر را تهیه کرد. این تکنیک را می توان روی تعدادی تصویر اعمال نموده و ماتریس ویژگی تصاویر را بدست آورد و با اعمال این ماتریس ویژگی به ماشین بردار پشتیبان، آن را آموزش داده و برای آزمودن آماده نمود. پس از آموزش ماشین بردار پشتیبان، ماتریس ویژگی تصاویری را که به منظور آزمایش در نظر گرفته ایم، به ماشین آموزش داده شده، اعمال میکنیم. حال ماشین بردار پشتیبان با بکارگیری اطلاعات حاصله از آموزش، میتواند تصاویر شامل انسان را از تصاویر فاقد انسان متمایز کند. البته عیبی که این مقاله دارد این است که از تصاویر زیادی در بخش آموزش استفاده نکرده است، بنابراین ماشین بردار پشتیبان خطای زیادی دارد . البته راهکاری که می توان  برای کاهش خطا، ارایه داد این است که تعداد تصاویری را که برای آموزش ماشین بردار پشتیبان اعمال میکنیم، افزایش دهیم. از طرفی نرم افزار مورد استفاده، قابلیت پردازش تعداد تصاویر بالا را ندارد. لذا میتوان با استفاده از دستورات مربوطه ، کدها را به زبان دلخواه که قابلیت پردازش حجم بالا را داشته باشد، تبدیل کرد. البته این راهکار در این مقاله منظور نشده است.

## کلمات کلیدی:

1) **هیستوگرام تصویر**: نموداری است که توسط آن تعداد پیکسل های هر سطح روشنایی در تصویر ورودی مشخص می شود، در واقع هر تصویر از نظر روشنایی به 5 قسمت : خیلی تاریک،تاریک،متوسط،روشن،خیلی روشن تقسیم می شود و هر سطح روشنایی خود شامل 50 سطح مجزای روشنایی است که هر پیکسل مقداری بین 0 (خیلی تاریک) تا 255 (خیلی روشن) می تواند داشته باشد. لذا این نمودار مشخص می کند که در تصویر مورد نظر در هر سطح روشنایی چه تعداد پیکسل موجود است. در واقع هیستوگرام فقط یک معیار تقریبی به ما نشان می دهد، مثل یک نگاه گذرا به ساعت مچی که حدود زمان را به ما نشان می دهد ، نه مقادیر دقیق ثانیه و دقیقه ها را. با کسب مهارت در خواندن هیستوگرام می توان کیفیت نور سنجی دوربین را با یک نگاه تخمین زد.

هیستوگرام تصویر

2) **ماسک** (mask) : ماسک را می توان پنجره ای در نظر گرفت که بر روی تک تک پیکسل های تصویر حرکت کرده و در هر نقطه با توجه به مقادیر ماسک ، مقدار شدت روشنایی پیکسل مرکزی در تصویر خروجی محاسبه می شود. البته مقادیری که در ماسک قرار میگیرند، ضرایبی می باشند که نشان می دهند هریک از پیکسل های همسایه تا چه حد در تعیین مقدار شدت روشنایی پیکسل مرکزی تاثیر گذارند. در این مقاله به علت استفاده از گرادیان شدت روشنایی تصویر از ماسک هایی به صورت
-1 0 1 استفاده می شود که داریم:

$$(-1) \times i\,(x-1) + (0) \times i\,(x) + (1) \times i\,(x+1) = i\,(x+1) - i\,(x-1)$$

که همان تعریف مشتق (گرادیان) می باشد.

3) **کوانتیده کردن**(quantization): یک تصویر را می توان توسط تابع دو بعدی $f(x,y)$ که در آن $x,y$ را مختصات مکانی و مقدار $f$ در هر نقطه را شدت روشنایی تصویر در آن نقطه می نامند،توصیف کرد. دیجیتال کردن $x,y$ را sampling و دیجیتال کردن مقدار $f$ را quantization یا کوانتیده کردن، می نامند.

## مقدمه :

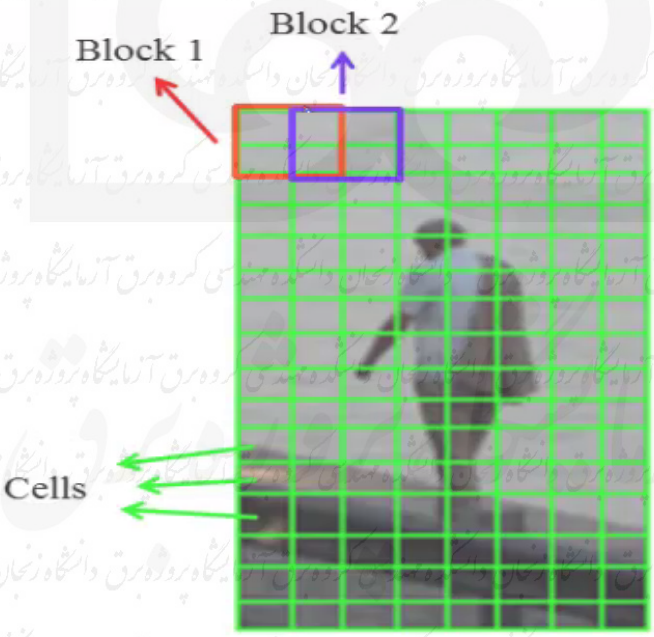
تشخیص انسان در تصاویر، همواره چالش برانگیز بوده است. این روش اولین بار توسط N.Dalal and B.Triggsدر سال 2005 مطرح گردید.

امروزه از روش HOG به جای روش self descriptor استفاده می شود. اگرچه روش self descriptor نیز بر مبنای محاسبه گرادیان های شیب گراست و مشابه روش HOG است ، اما مزایایی که روش HOG دارد، این روش را تبدیل به یک روش متداول کرده است. به طور مثال در روش self descriptor هر تصویر 16 هیستوگرام دارد، اما در روش HOG ، هرتصویر 105 هیستوگرام دارد که این تقسیم بندی به ابعاد و سلول های کوچکتر می تواند باعث افزایش دقت عملکرد شود و از طرفی در روش HOG ، window (پنجره) بزرگتر است، چرا که در روش self descriptor ابعاد تصاویر 16×16 است ، اما در روش HOG ابعاد 128×64 می باشد. علاوه بر افزایش تعداد سلولهای هر تصویر می توان با افزایش تعداد تصاویر استفاده شده برای آموزش ماشین بردار پشتیبان ، به نحوی دقت را بالا برده و از خطا کاست. همانطور که اشاره شد، تا کنون به طرق مختلف سعی بر این مهم شده است.در ادامه یکی از این روش ها موسوم به HOG را تشریح می کنیم.

## گام های محاسبه بردار ویژگی تصویر :

- گام اول،استخراج ویژگی است، که این استخراج ویژگی شامل محاسبه محور افقی و عمودی گرادیان ها وهمچنین محاسبه اندازه و جهت گرادیان ها است.

- تصویر زیر تصویری با ابعاد 128 × 64 است. ما این تصویر را به سلول های نشان داده شده تقسیم می کنیم. این سلول ها را دسته بندی می کنیم و هر 4 سلول را یک بلوک می نامیم ،تعداد سلولها در هر بلوک 2×2می باشد. از طرفی سایز هر سلول 8×8 است . با این حساب ابعاد بلوکها 16×16 می باشد. در نتیجه 7 بلوک در جهت افقی و 15 بلوک در جهت عمودی داریم، لذا در کل در هر تصویر 7×15=105 بلوک خواهیم داشت. برای بهبود آشکارسازی ، بلوک ها را طوری در نظر می گیریم که هر بلوک با بلوک مجاور 50% هم پوشانی داشته باشد. ضرورت استفاده از این تکنیک افزایش دقت عملکرد است. به طور مثال اگر تصویر انسان مورد نظر دقیقا بین دو بلوک واقع شود ، تخمین دقیقی نخواهیم داشت.اما با این تکنیک می توانیم اطلاعاتی از بلوک قبلی داشته باشیم و تخمین دقیق تری داشته باشیم.



از طرفی طبق تعریف مشتق داریم :

$$f' = \lim_{h \to 0} \frac{f(x+h) - f(x-h)}{2h}$$

که در آن f اندازه شدت روشنایی هر پیکسل می باشد.

-ماسک هایی که در این روش مورد استفاده قرار می گیرند در جهت های افقی و عمودی به شکل زیر هستند.

-1 0 1

-1

0

1

نحوه عملکرد این ماسک ها به گونه ای است که با حرکت دادن این ماسک ها مثلا ماسک افقی روی تصویر در تمامی جهات به گونه ای که تمامی پیکسل های تصویر را پوشش دهد، داریم :

$$(-1) \times i(x-1) + (0) \times i(x) + (1) \times i(x+1) = i(x+1) - i(x-1)$$

که همان تعریف مشتق (گرادیان) است.

لذا با این تکنیک می توان به نحوی از تمامی پیکسل های تصویر مشتق (گرادیان) گرفت.

از طرفی می دانیم که اندازه و زاویه هر برداری به صورت زیر قابل محاسبه است:

اندازه :

$$S = \sqrt{S_x^2 + S_y^2}$$

زاویه :

$$\theta = \text{Arctg}\ \left(\frac{S_y}{S_x}\right)$$

**دانشجویان محترم:**

**جهت دسترسی به متن کامل پایان نامه‌ها به کتابخانه دانشکده مهندسی و یا آزمایشگاه پروژه گروه برق مراجعه فرمایید.**

## نتیجه گیری :

در این مقاله، برای آشکارسازی انسان در تصویر، روشی تحت عنوان هیستوگرام گرادیان های شیب گرا ارایه شد. به کمک این روش و با استفاده از نرم افزار متلب، می توان حضور یا عدم حضور انسان را در تعداد تصاویر اندک اما با خطای بالا تشخیص داد.

عیب این روش این است که نرم افزار مربوطه توانایی پردازش تصاویر زیادی را ندارد، پس تصاویر بکاربرده شده در قسمت آموزش کم بوده لذا خطا بالاست. البته میتوان راهکارهایی در این خصوص ارایه داد. به طور مثال با استفاده از دستوراتی خاص، نظیر mex کدهایی را که با نرم افزار متلب نوشته شده اند، به زبان دلخواه به عنوان مثال c++ تغییر داد واز قابلیت های آن برای پردازش تصاویر زیاد(حجم بالا) استفاده کرد که متاسفانه این راهکار در این مقاله به کار گرفته نشده است.

## منابع :

-INRIA Rh.one-Alps, 655 avenue de l'Europe, Montbonnot 38334, France {Navneet.Dalal,Bill.Triggsg}@inrialpes.fr, http://lear.inrialpes.fr

-C. Papageorgiou and T. Poggio. A trainable system for object detection. *IJCV*, 38(1):15.33, 2000.

-www.wikipedia.com

## مراجع :

[1]. http://lear.inrialpes.fr

[2]. MIT pedestrian database