

کتابخانه

بینایی ماشین و کاربردهای آن

استاد راهنما:

خانم عسکری

تهیه کننده:

مجتبی پور تیمور

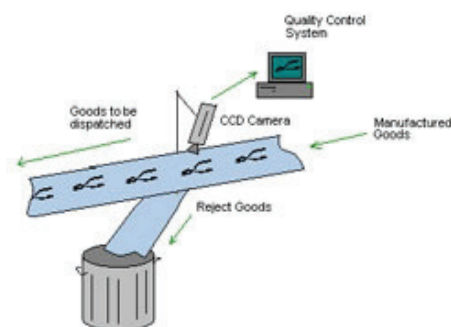
84143210

بینایی ماشین (Machine Vision)

از میان همه شاخه‌های هوش مصنوعی، شاید کاربردی‌ترین آن‌ها کامپیوتری و مکانیزه کردن سیستم‌های بینایی باشد. دامنه کاربرد این شاخه از فناوری در حال رشد، بسیار وسیع است و از کاربردهای عادی و معمولی مثل کنترل کیفیت خط تولید و نظارت ویدئویی گرفته تا تکنولوژی‌های جدید مثل اتومبیل‌های بدون راننده را دربر گرفته است. دامنه کاربردهای این تکنولوژی براساس تکنیک‌های مورد استفاده در آن‌ها تغییر می‌کند. در این مقاله سعی داریم به شما نشان دهیم که سیستم‌های بینایی ماشین چگونه کار می‌کنند و مروری کوتاه بر اهداف، تکنیک‌ها و تکنولوژی‌های موجود داشته باشیم و سعی داریم با نحوه کار بینایی ماشین و پیشرفت آن‌ها که مطابق با سیستم بینایی انسان است، آشنا شویم. در این متن، بررسی خود را با دو مثال انجام می‌دهیم. اولی سیستم کنترل کیفیت خط تولید است که شامل نحوه عکس برداری و ذخیره و شیوه تفسیر عکس‌های گرفته شده به صورت خودکار است و دیگری به عنوان یک مثال پیچیده‌تر، چگونگی بینایی یک ربات را توضیح می‌دهد، و در آخر نیز به کاربردهای پردازش تصویر در دنیای امروز می‌پردازیم.



کنترل کیفیت خط تولید



شکل 1- نمای ساده شده‌ای از یک سیستم بینایی کنترل کیفیت خط تولید غیر واقعی

یکی از کاربردهای بینایی ماشین در کنترل کیفیت خروجی کارخانه‌ها می‌باشد. شکل 1 مثالی بسیار ساده از چنین سیستمی است. اجناس تولید شده در کارخانه که بر روی یک نوار نقاله قرار گرفته‌اند و توسط یک دوربین CCD برای آزمایش دیده

می‌شوند و محصولات با کیفیت مناسب اجازه عبور پیدا خواهند کرد. چنانچه محصولی دارای استانداردهای مناسب نباشد از ادامه مسیر حذف می‌شود. معیار این استانداردها می‌تواند لبه‌های زائد، خراشیدگی و بادکردگی و تورم روی فلزات و بسیاری چیزهای دیگر باشد. در این مثال ما در پی یافتن مکانیزم خط تولید نیستیم و فقط می‌خواهیم ببینیم که یک شی تولیدشده چگونه استاندارد تشخیص داده شده و اجازه عبور می‌یابد و برعکس چگونه به بعضی از اشیاء اجازه عبور و ادامه دادن داده نمی‌شود .

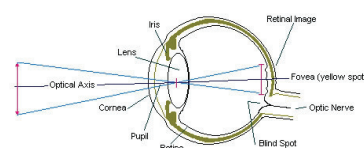
عکس برداری

در این مثال ما سعی در مکانیزه کردن فرآیندی یکنواخت داریم که به صورت معمول و تکراری توسط انسان انجام می‌شود. اولین مسأله و مشکل ما این است که چگونه عکس‌های تهیه شده از اشیایی که در حال حرکت بر روی نوار نقاله هستند را تبدیل به داده‌های قابل فهم و تفسیر برای سیستم نماییم، که این مشکل توسط دوربین CCD حل می‌شود. عملکرد این دوربین را می‌توان به عملکرد چشم انسان که قادر است سطوح مختلف نور را تشخیص دهد تشبیه نمود .

چشم انسان

چشم انسان که در شکل 2 نشان داده شده است، تقریباً یک عدسی کروی با قطر $2/5$ سانتی‌متر می‌باشد که از چندین لایه مختلف که درونی‌ترین آن‌ها شبکیه نام دارد تشکیل شده است . ماهیچه‌های اطراف چشم اندازه لنز را تنظیم می‌کنند که این کار چشم را قادر به زوم (zoom) کردن روی اشیاء می‌کند .

وظیفه عدسی چشم، فرم و شکل دادن به تصویری است که توسط میلیون‌ها سلول گیرنده مخروطی (Cone) و میله‌ای (rod) گرفته شده و بر روی پرده شبکیه افتاده است، می‌باشد. سلول‌های میله‌ای به یک عصب معمولی که از انتها به شبکیه ختم می‌شود و فقط در سطح نور پایین فعال است متصلند و سلول‌های مخروطی هر کدام به یک عصب اتصال دارند. آن‌ها در نورهای شدیدتر، بیشتر فعالند و میزان درک ما از رنگ‌ها را نوع فعالیت این مخروط‌ها مشخص می‌کند .



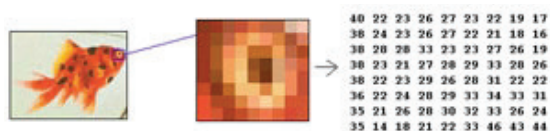
شکل 2- نمای چشم انسان

در میان شبکیه ناحیه‌ای به نام نقطه کور وجود دارد که در آن هیچ گیرنده‌ای موجود نیست. در این ناحیه اعصاب به صورت جداگانه به عصب بینایی که سیگنال‌های دریافت شده را به قشر بینایی مخ انتقال می‌دهند، وصل می‌شود .

دوربین CCD

CCD از جهت عملکرد تقریباً مانند چشم انسان کار می‌کند . نور از طریق یک عدسی وارد دوربین و بر روی یک پرده مخصوص تصویر می‌شود که تحت عنوان تراشه CCD شناخته می‌شود. تراشه CCD (Charge Coupled Device) که تصاویر با استفاده از آن گرفته می‌شوند از تعداد زیادی سلول تشکیل شده که همگی در یک تراشه با الگوی خاصی مرتب شده‌اند و تحت عنوان پیکسل (pixels) شناخته می‌شوند .

زمانی که تراشه CCD این اطلاعات را دریافت می‌کند، آن‌ها را به شکل سیگنال‌های دیجیتالی از طریق کابل‌هایی به سیستم دریافت‌کننده می‌فرستد و بعد تصاویر در این سیستم به صورت مجموعه‌ای از اعداد ذخیره می‌شوند. همان‌طور که در شکل 3 می‌بینید هر عدد نماینده یک پیکسل است.



شکل 3- تصویر وسط یک نمای نزدیک از چشم ماهی را نمایش می‌دهد و نشان می‌دهد که هر قسمت از یک تصویر چگونه با تعدادی مقادیر عددی ذخیره می‌شود. به تعداد داده‌های عددی مورد نیاز برای ذخیره یک فضای کوچک از تصویر توجه کنید.

پردازش تصویر:

در این قسمت مقدماتی در مورد تکنیک‌های اصلی که در سیستم‌های بینایی ماشین به منظور پردازش تصویر استفاده می‌شود مطرح می‌گردد. پروسه تشخیص تصویر به چند مرحله اصلی زیر شکسته می‌شود:

1. دریافت تصویر
2. تشخیص لبه
3. قطعه بندی
4. تشخیص و آنالیز

1- دریافت تصویر :

در مورد چگونگی دریافت تصویر توسط دوربین CCD صحبت شد. در ادامه سه مورد بعدی را توضیح می‌دهیم.

2- تشخیص لبه:

با هر تصویر، چه با دوربین گرفته شود و چه با چشم انسان، مقداری تحریف و تغییر شکل و به عبارتی "نویز" (noise) وجود دارد. البته در مورد مثال ما در سیستم خط تولید این مسأله چندان اهمیت ندارد اما در موقعیت‌هایی که نیاز به دقت بالا وجود دارد باید از نورپردازی خاصی برای تصویربرداری استفاده شود .

انسان برای درک تصاویری که می‌بیند نیازی ندارد هیچ کاری در مورد فیلتر کردن و از بین بردن نویزهای یک تصویر انجام دهد. مثلاً در یک روز ابری که همه جا را فرا گرفته، دید ما به شدت ضعیف و دچار مشکل می‌شود. اما هر آنچه را که قادر به دیدنش باشیم درک می‌کنیم. یعنی برای درک اشیاء نیازی به حذف نویزهای تصویر نیست. مثلاً اگر در این روز در حال رانندگی در یک جاده باشید و تصویر مبهمی از یک ماشین را مقابل خود ببینید، بالطبع عکس‌العمل نشان می‌دهید و به عبارتی سرعت خود را کم می‌کنید .

این یعنی ما هنوز تصویر ماشین را علیرغم وجود مه می‌توانیم تشخیص دهیم و در مقابل آن عکس‌العمل نشان‌دهیم. و یا مثلاً زمانی که دچار سرگیجه می‌شوید، علیرغم این که تصاویر اطراف خود را تار و مبهم می‌بینید اما قادر به درک و تشخیص وسایل و تصاویر اطراف خود هستید. یعنی ابتدا صبر نمی‌کنید تا سرگیجه‌تان به پایان برسد و بعد تصاویر را تشخیص دهید و این یعنی با

قدرت بینایی انسان، علیرغم خراب شدن تصاویر اطراف، می‌توانیم متوجه فضای اطراف خود بشویم. اما برای بینایی ماشین ابتدا باید این نویزها طی فرآیندی که تصفیه کردن یا فیلترینگ نامیده می‌شود، از بین برود و بعد هر آنچه برای پردازش عکس لازم است انجام شود.

خوشبختانه در حال حاضر تکنیک‌هایی برای انجام این کار وجود دارد. از بین بردن نویزها به صورت نرمال توسط تعدادی از توابع ریاضی یا الگوریتم‌هایی که تحت عنوان 'treshholding' یا 'quantizing' نامیده می‌شود انجام می‌گردد. این فرآیند بسیار حرفه‌ای و پیچیده‌ای است و نیاز به دانش و پشتوانه بالای ریاضی دارد. زمانیکه خرابی‌ها از بین رفت، می‌توانیم پردازش عکس‌ها را ادامه دهیم که این کار با استخراج صورت‌ها و حالت‌ها از یک تصویر انجام می‌شود. یک شیوه معمول که غالباً مورد استفاده قرار می‌گیرد استخراج لبه‌ها که در شکل 4 دیده می‌شود، می‌باشد.



شکل 4- تشخیص لبه. شکل اول تصویر اصلی می‌باشد. در شکل وسط نویزها با فیلتر نمودن از بین رفته است. در نهایت شکل 3، تصویر حاصل از به کار بردن الگوریتم تشخیص لبه می‌باشد.

3- قطعه بندی:

بعد از اینکه لبه‌ها در تصویر تشخیص داده شدند این اطلاعات می‌توانند به منظور قطعه بندی تصویر به نواحی متشابه مورد استفاده واقع شوند. وقتی که می‌گوییم یک ناحیه از تصویر متشابه است منظورمان این است که رنگ یا شدت سایه‌ها در آنها زیاد با هم متفاوت نمی‌باشند. به عبارت دیگر لبه‌ای در آن ناحیه وجود ندارد.

روش‌های دیگری نیز جهت قطعه بندی تصویر وجود دارند که در آنها از لبه‌یابی استفاده نمی‌شود یکی از روش‌های ساده در این زمینه استفاده از حد آستانه است در این روش رنگ هر کدام از نقاط تصویر بدست آمده و نقاط همسایه در یک منطقه از تصویر تا جایی که رنگشان به حد کافی شبیه یکدیگر باشند به عنوان یک قطعه در نظر گرفته می‌شوند.

این روش خیلی شبیه لبه‌یابی می‌باشد اما به منظور قطعه بندی استفاده می‌شود و فقط به یافتن لبه‌ها در تصویر نمی‌پردازد.

4- تشخیص و آنالیز:

در مرحله آخر هم با استفاده از تکنیک‌های خاصی مثل بافت‌شناسی و آنالیز بافت و دیتابیسی که در این ماشین طراحی شده می‌تواند اشیاء مورد نظر را شناسایی کند.

در مورد مثال ما در سیستم خط تولید، وظیفه اصلی یک اپراتور کنترل کیفیت این است که به سرتاسر محصول تولید شده نگاه کرده و با مقایسه آن با استانداردهای مورد قبول، برای محصول تولید شده جواز عبور یا عدم عبور صادر کنند.

اگر این کار با استفاده از بینایی ماشین صورت گیرد باید عکس گرفته شده از محصول تولید شده با عکسی که از یک محصول استاندارد وجود دارد مقایسه شود. یکی از روش‌های انجام این کار به این صورت است:

برای انجام این کار، یک تصویر از محصول استاندارد در کامپیوتر ذخیره می‌شود و سپس از محصولاتی که از خط تولید عبور می‌کنند. تصویر گرفته می‌شود و به عنوان نقشه لبه ذخیره می‌شود. و بعد سیستم، تصویر گرفته شده را از چپ به راست و از بالا به

پایین به گونه‌ای که در هر زمان فقط یک رقم عبور کند، می‌لغزاند و عدد ظاهر شده در هر موقعیت را با عدد همان موقعیت در تصویر اصلی مقایسه می‌کند و در صورت تفاوت آن را اعلام می‌نماید.

لذا عملیات بینایی کامپیوتر در حقیقت مقایسه دو مجموعه عدد است که اگر تفاوت این دو مجموعه از یک محدوده خاص فراتر برود، از پذیرفتن محصول امتناع شده و در غیر این صورت محصول پذیرفته می‌شود.

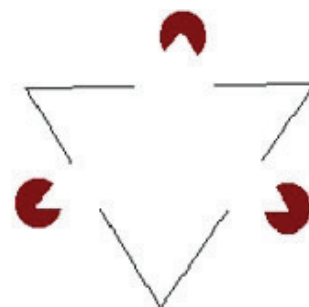
یک مثال پیچیده‌تر

در مثال قبل سیستم مورد مطالعه بسیار محدود بود و فقط یک تصویر دو بعدی از یک محصول را با تصویر اصلی و ایده‌آل مقایسه می‌کرد و احتیاجی به بررسی مقادیر اندازه و زاویه نبود.

در این مثال می‌خواهیم به سیستم بینایی که برای یک ربات خانگی که قادر به تمیز کردن خانه، پختن غذا و ... طراحی شده نگاهی بیاندازیم. این مثال بسیار پیچیده‌تر از مثال قبلی است و نیاز به آگاهی از تغییرات محیط دارد. به عبارت بهتر نیاز به یک سیستم هوشمندتر داریم. قبلاً دیدیم که تصاویر چگونه ذخیره و تفسیر می‌شوند. غالب تکنیک‌های گذشته دوباره در این مثال به کار گرفته می‌شوند. تفاوت اصلی در تفسیر تصاویر گرفته شده است. در مثال قبلی فضای سیستم بینایی فقط متشکل از یک سری محصول بود، اما در این مثال ربات باید از همه آنچه در اطرافش می‌گذرد باخبر بوده و این یعنی با دنیای وسیع‌تر و بزرگ‌تری روبروست. بدین منظور نیاز به تکنیک‌های تشخیص پیچیده‌تری وجود دارد.

یعنی در این مثال، یک صحنه فرضی شامل ده‌ها یا حتی صدها شیء مختلف در معرض دید است. این اشیاء در اندازه‌ها و تحت زوایای مختلف که متأثر از نوع نورپردازی هم هستند به نمایش درمی‌آیند و به همین دلیل برای تشخیص این اشیاء نیاز به تکنیک‌های هوش مصنوعی (Ai) می‌باشد.

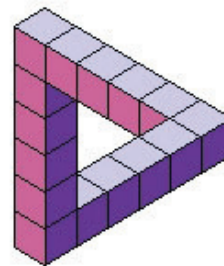
پردازش اطلاعات در مغز



شکل 5 - تصویر به نظر دو مثلث می‌آید که در جهت عکس هم و بر روی هم قرار داده شده‌اند، ولی از نظر ریاضی هیچ مثلثی رسم نشده است و فقط سه دسته از خطوط وجود دارند.

در مورد بینایی انسان متذکر شدیم که شاید یکی از قسمت‌هایی از مغز که بیشتر فعالیت درک تصویر را انجام می‌دهد ناحیه Visual Cortex باشد. همان‌طور که دیده‌ایم، این‌جا ناحیه‌ای است که اطلاعات منتقل شده در طول عصب بینایی در آن پردازش می‌شود. البته این را هم مدنظر داشته باشید که قسمتی از فعالیت پردازش اطلاعات در ناحیه شبکه چشم قبل از این که اطلاعات به مغز برسند، انجام می‌شود.

البته خود ناحیه شبکه به عنوان قسمتی از مغز شمرده می شود. در ضمن این مسأله هم قبلاً مشخص شده است که نواحی مختلف قشر بینایی مخ در مقابل تصاویر مختلف عکس العمل نشان می دهد. به عبارت دیگر هر قسمت از این ناحیه مربوط به یک حالت خاص است. مثلاً نواحی معین و مشخصی در مقابل رنگ های متفاوت عکس العمل نشان می دهند و یا مثلاً نواحی وجود دارند که سلول هایشان در مقابل جزئیات دقیق موجود در یک تصویر عکس العمل نشان می دهند. مثلاً در شکل 5 شما می توانید یک مثلث سفید که به طور واضح سه راس آن مشخص نشده است را ببیند، علیرغم این که به طور واضح و مجزا این مثلث مشخص نشده است. اما سلول های موجود در قشر بینایی می توانند تکه های این خطوط را کنار هم گذاشته و از آن یک مثلث استنتاج کنند.



شکل 6- مثلث Penrose: عدم هماهنگی هندسی در لبه ها باعث عدم درک صحیح شکل می شود

زمانی که اطلاعات مربوط به چیزی که دیده می شود را دریافت کردیم، مثل شکل ظاهری، لبه ها و غیره، مغز نیاز به رمزگشایی و کشف اطلاعات به دست آمده دارد تا دقیقاً نتیجه گیری کند که آنچه در صحنه می گذرد چیست. اینجا قسمتی است که ما خیلی کم در رابطه با آن می دانیم و مشخص نیست که دقیقاً مغز چگونه این کار را انجام می دهد. اما مسأله مشخص این است که سیستم بینایی ما چندان مبرا از خطا نیست و به راحتی دچار اشتباه می شود. موارد بسیار زیادی وجود دارد که اطلاعات کسب شده می توانند مبهم و یا حتی اشتباه باشند، مانند مثال هایی که در شکل 6 و 7 مشاهده می کنید.



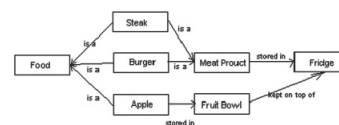
شکل 7- خطای دید: در شکل سمت چپ علیرغم نحوه ظهور خطوط همه آن ها با هم موازیند. در شکل سمت راست که به خطای Muller-Lyer معروف است دو خط ظاهر شده به نظر دارای طول های متفاوت می آیند. اما در حقیقت هم اندازه اند.

نمایش دانش

متخصصین هوش مصنوعی انواع سیستم ها را با استفاده از برنامه های معمول کامپیوتری و تفاوت قایل شدن بین پردازش اطلاعات (information) و پردازش دانش (knowledge)، شناسایی می کنند. این کار منجر به ایجاد سیستم های مبتنی بر دانش که کاربرد بسیار زیادی در هوش مصنوعی دارد می شود.

ربات ما نیاز به داشتن اطلاعات یا "دانش" از دنیای اطراف خود دارد تا سیستم بینایی اش به درستی عمل کند. ما به طور خلاصه بعضی تکنیک هایی که توسط مهندسین دانش (knowledge engineer) برای نمایش و پردازش این اطلاعات به کار می رود را آزمایش می کنیم.

یک شبکه معنایی (Semantic network) یک تکنیک خوش ساخت برای نمایش اشیاء یا ماهیت موجود در یک دنیا و ارتباطاتی که بین شان موجود است، می باشد. مثلاً قسمتی از دنیای ربات ما آشپزخانه است.



شکل 8- یک شبکه معنایی برای نمایش قسمتی از آشپزخانه.

استفاده از یک شبکه معنایی می تواند راه مناسبی برای نمایش اطلاعات باشد. بدین ترتیب آشپزخانه را تحت عنوان یک سری از اشیاء و روابط بین شان فرض کنیم. یک شبکه می تواند جزئیات موجود را به خوبی نمایش داده و ما را قادر به کاوش بیشتر و بهتر در روابط موجود بین اشیاء کند. برای مثال، همان طور که از شکل شماره 8 می توان دید، گوشت های سرد در یخچال نگهداری می شوند. این آیتم ها تحت عنوان غذاها می توانند دسته بندی شوند. اگر چه همه غذاها در یخچال نگهداری نمی شوند، اما همه محصولات گوشتی ملزم به نگهداری یخچال هستند.

مشخص است که این شبکه باید برای نمایش دنیای اطراف ربات ما به خوبی گسترش یابد و نیاز به نمایش دقیق تر و جزئی تری از اشیاء و ارتباطات وجود دارد و همچنین این شبکه باید اتاق ها، افراد، اشیاء موجود و غیره را هم نمایش دهد.

پس از ساخته شدن شبکه معنایی ما قادر به ساخت یک پایگاه داده از دانش حاصل از شبکه معنایی در مورد اشیاء موجود در دنیای مربوطه هستیم. در مورد مثال ما، ورودی های مناسب می تواند به شکل زیر باشد :

(سیب - شکل - گرد) (سیب - رنگ - قرمز) (سیب - نما - براق)

(همبرگر - شکل - گرد) (همبرگر - رنگ - قهوه ای) (همبرگر - نما - گوشتی)

(یخچال - شکل - مستطیل) (یخچال - رنگ - سفید) (یخچال - نما - فلزی)

(ظرف میوه - شکل - بیضی) (ظرف میوه - رنگ - نقره ای) (ظرف میوه - نما - فلزی)

((میوه) نگهداری - در (ظرف میوه))

((سیب) هست - یک (خوردنی))

توجه کنید که ورودی های این پایگاه دانش باید به گونه ای باشد که منجر به سردرگمی ماشین نشود. مثلاً در مورد مثال ما، سیب یک شی براق قرمز گرد معرفی شده است. اشیاء زیادی هستند که ممکن است چنین مشخصه ای داشته باشند و اگر مثلاً شما ربات خود را بفرستید که برای شما یک سیب تازه بیاورد امکان دارد ربات با یک توپ نزد شما باز گردد.

این پایگاه دانش تحت عنوان حافظه کاری سیستم شناخته می شود. به منظور استفاده کارا و مناسب از اطلاعات موجود در این پایگاه، سیستم نیاز به داشتن تعدادی قانون دارد. بدین ترتیب پس از ساخت پایگاه دانش نیاز به ساخت یک پایگاه از قوانین برای برخورد مناسب با اطلاعات موجود در پایگاه دانش هستیم. با عبارتی مثل **If Statement Then** آشنا هستید مثلاً **If cold Then wear coat** ما از چنین ساختاری برای ایجاد پایگاه قوانین استفاده می کنیم. قسمت **IF** عبارت تحت عنوان راس قانون و قسمت عملیات یعنی قسمت **Then** تحت عنوان بدنه قانون شناخته می شوند.

و بدین ترتیب ربات مثلاً می داند که یک سیب چیست و در کجا قرار دارد. مثلاً فرض کنید یکی از کارهایی که قرار است ربات انجام دهد رفتن به سمت یخچال و برداشتن چیزی از یخچال است، در این صورت می توانیم مجموعه ای از قوانین را به این منظور بسازیم.

اکنون هم دانش و هم قوانین مربوطه را داریم که نیاز به روش‌هایی برای تفسیرشان داریم. این کار توسط سیستم مفسر انجام می‌شود. مفسر برنامه‌ای است که قوانین را پردازش می‌کند. وظیفه آن خواندن هر قانون و بررسی این که موقعیت‌ها با هم همخوانی دارند یا نه می‌باشد. اگر همخوانی داشتند عمل موردنظر را انجام می‌دهد. این عمل آن‌قدر انجام می‌شود تا به عملی که دستور پایان داده‌است، برسد.

گفتیم برای استفاده از تکنیک همسان‌سازی الگو (Template matching) نیاز به ذخیره مقادیر بسیار زیادی از دانش در رابطه با هر آیتم موجود در دنیای ربات وجود دارد. با استفاده از قوانین می‌توانیم استفاده از این اطلاعات را از بین ببریم و دیگر نیازی به آن‌ها حس نمی‌شود. مثلاً در مورد یک صندلی، به جای ذخیره یک صندلی از تعدادی زوایا و اندازه‌ها می‌توانیم اطلاعاتی را جمع‌به برخی خواص و ظواهر صندلی نظیر جنس، دسته‌ها، پایه‌ها و ... که یک سیستم مبتنی بر دانش از آن‌ها به منظور تشخیص صندلی استفاده می‌کند، ذخیره کنیم. یک صندلی دارای چهار پایه، یک نشیمنگاه و یک تکیه‌گاه است.

از طرف دیگر میز هم چهار پایه و یک رویه دارد. بدین ترتیب به جای ذخیره تعداد زیادی عدد برای هر شی، از خواص آن شی برای توصیفش استفاده کنیم. بدین ترتیب نیاز به یک پایگاه دانش هست که چیزهایی نظیر اجزا (دسته‌ها، پایه‌ها، دیواره‌ها و ...) خواصی که این اجزا را توصیف می‌کند (ظاهر و غیره)، و حقایقی در رابطه با اشیایی که توسط این اجزا ساخته می‌شوند (آشپزخانه چهار دیوار و یک در دارد و غیره) و همچنین یک سری قوانین به منظور پردازش این اطلاعات را نگهداری نماید.

خلاصه

همان‌طور که دیدیم، عملیات زیادی باید به منظور پردازش اطلاعات تصویری انجام شود. تصاویر غالباً دارای نویز هستند که باعث خرابی و تحریفشان می‌شود. هر چند این مشکل چندان حادی برای سیستم بینایی انسان به حساب نمی‌آید اما برای سیستم‌های بینایی ماشین این نویزها به‌طور کامل باید فیلتر و پاکسازی شوند که این عمل با استفاده از تعدادی الگوریتم انجام می‌شود.

همچنین دیدیم که اشیاء چگونه با استفاده از تکنیک همسان‌سازی الگو دیده می‌شوند هر چند پس از این که اشیاء داخل صحنه توسط سیستم بینایی مشخص شدند هنوز کار تمام نشده است. اطلاعات حاصل باید به یک نمایش دانش مناسب ترجمه شوند به‌طوری که قابل پردازش باشند.

نتیجه‌گیری

مثال اول یک مثال محدود شده از نوعی سیستم بود که در کاربردهای روزانه مشابه آن را فراوان می‌توان یافت. سیستم‌های بینایی اولیه روز به روز در حال گسترش هستند و هر روز کاربرد جدیدی برای آن‌ها یافت می‌شود. با توسعه این سیستم‌ها، کاربردهایی برای سایر محیط‌ها نظیر توسعه پردازش نیز یافت می‌شود. سیستم‌های بینایی مصنوعی از کاربردهای پزشکی و نظامی گرفته تا کاربردهای ساده و معمول روزمره مورد استفاده هستند. و این شاخه از هوش مصنوعی از کاربردی‌ترین شاخه‌ها می‌باشد که هنوز مسایل حل نشده زیادی در خود دارد.

کاربردهای پردازش تصویر در دنیای امروز

زمینه‌های مختلف کاربرد پردازش تصویر عبارتند از:

صنعت، هواشناسی، شهرسازی، کشاورزی، علوم نظامی و امنیتی، نجوم و فضا نوردی، پزشکی، فناوری‌های علمی، باستان‌شناسی، تبلیغات، سینما، اقتصاد، روانشناسی و زمین‌شناسی که در ادامه درباره هر کدام مختصراً بحث شده است.

1- صنعت

امروزه کمتر کارخانه پیشرفته ای وجود دارد که بخشی از خط تولید آن توسط برنامه های هوشمند بینایی ماشین کنترل نشود. خطای بسیار کم، سرعت زیاد، هزینه نگهداری بسیار پایین، عدم نیاز به حضور اپراتور 24 ساعته و خیلی مزایای دیگر باعث شده که صنایع و کارخانه ها به سرعت به سمت پردازش تصویر و بینایی ماشین روی بیاورند. دستگاهی ساخته شده که قادر است کیک های پخته را از کیک هایی که نیاز به پخت مجدد دارند، تشخیص دهد و آنها را به صورت اتوماتیک به بسته بندی بفرستد و کیک هایی که نیاز به پخت دارند را دوباره برای پختن ارسال کند.

یکی دیگر از دلایل استفاده از بینایی ماشین قابلیت دیدن و اندازه گیری محصولاتی است که دیدن یا اندازه گیری آنها با چشم غیر مسلح غیر ممکن است. عناصر تشکیل دهنده یک سیستم بینایی ماشین نرم افزار هوشمند بینایی است که ورودی خود را از دوربین های نصب شده در بخش های مختلف خط تولید می گیرد و بر اساس تصاویر دریافتی دستورات لازم برای کنترل ماشین های صنعتی را صادر می کند. پردازش تصویر در تشخیص دمای کوره هایی که هیچ وسیله ی مکانیکی و الکترونیکی تحمل دمای آنها را ندارد، کاربرد دارد. دوربین های حرارتی می توانند مشکل بخشی از سازه ی مورد نظر را تشخیص دهند.

2- هواشناسی

از آنجایی که در علم هواشناسی تشخیص و پیش بینی آب و هوا اکثراً از طریق تصاویر هوایی و ماهواره ای انجام می گیرد، پردازش تصویر در این علم کاربرد زیادی دارد و دقت و سرعت پیش بینی آب و هوا و طوفان ها را بسیار بالا می برد. جبهه های پرفشار، کم فشار، گردبادها و گرداب های بوجود آمده در سطح کره زمین را می توان مشاهده کرد.

3- شهرسازی

با مقایسه عکس های مختلف از سال های مختلف یک شهر می توان میزان گسترش و پیشرفت آن را مشاهده کرد.

کاربرد دیگر پردازش تصویر می تواند در کنترل ترافیک باشد. با گرفتن عکس های هوایی از زمین ترافیک هر قسمت از شهر مشخص می شود.

قبل از ساختن یک شهر می توان آن را توسط کامپیوتر شبیه سازی کرد که به صورت دو بعدی از بالا و حتی به صورت سه بعدی از دید های مختلف، یک شهرک چطور ممکن است به نظر برسد. تصاویر ماهواره ای که از شهرها گرفته می شود، می تواند توسط فیلترهای مختلف پردازش تصویر فیلترشود و اطلاعات مختلفی از آن استخراج شود. به طور مثال این که شهر در چه قسمت هایی دارای ساختمان ها، آب ها یا راه های بیشتری است و همین طور می توان جاده هایی که داخل یا خارج از شهر کشیده شده اند را تحلیل کرد.

4- کشاورزی

این علم در بخش کشاورزی معمولاً در دو حالت کاربرد دارد. یکی در پردازش تصاویر گرفته شده از ارتفاعات بالا مثلاً از هواپیما و دیگری در پردازش تصاویر نزدیک به زمین.

در تصاویر دور به عنوان مثال می توان تقسیم بندی اراضی را تحلیل کرد. همچنین می توان با مقایسه تصاویر دریافتی در زمان های متفاوت میزان صدمات احتمالی وارد به محیط زیست را دید. به عنوان مثال می توان برنامه ای نوشت که با توجه به محل رودخانه ها و نوع خاک مناطق مختلف، به صورت اتوماتیک بهترین نقاط برای کشت محصولات مختلف را تعیین می کند.

تصاویر نزدیک هم در ساخت ماشین های هرز چین اتوماتیک کاربرد دارد. امروزه ماشین های بسیار گران قیمت کشاورزی وجود دارند که می توانند علف های هرز را از گیاهان تشخیص بدهند و به صورت خودکار آنها را نابود کنند.

برای مثال یکی از پروژه های جالب در بخش کشاورزی، تشخیص خودکار گل زعفران برای جداسازی پرچم قرمز رنگ آن بوده است. این پردازش که توسط نرم افزار Stigma detection® انجام گرفته است.

5- علوم نظامی و امنیتی

پردازش تصویر بخصوص بینایی هوشمند، کاربردهای بسیاری را در علوم نظامی و امنیتی دارند و این کاربرد برای دولت اکثر کشورها بسیار مهم است. به عنوان مثال موشک هدایت شونده خودکاری وجود دارد که می تواند روی یک ساختمان قفل کند و حتی می تواند به درز بین در و دیوار آن ساختمان که حساس ترین جای ساختمان است به راحتی نفوذ کند. این موشک به صورت اتوماتیک این قسمت را شناسایی کرده و به سمت آن حمله می کند.

در مسائل امنیتی هم کاربرد پردازش تصویر کاملاً در زندگی ما مشهود است. دوربین های که به صورت اتوماتیک از ماشین هایی که تخلف رانندگی انجام می دهند عکس برداری می کند.

از سیستم های امنیتی دیگر می توان سیستم تشخیص اثر انگشت اتوماتیک را نام برد. در لپ تاپ های جدید قابلیت finger print به آنها اضافه شده و می تواند صاحب لپ تاپ را توسط اثر انگشت شناسایی کند.

کد امنیتی دیگری که همیشه همراه انسان حمل می شود، چشم انسان است. دانشمندان ثابت کرده اند که پترن های (Pattern) موجود در مردمک چشم هر انسان منحصر به فرد است و هیچ دو فردی در دنیا وجود ندارند که پترن هایی که در مردمک چشم آنها وجود دارد دقیقاً مثل هم باشد. از همین روش برای شناخت افراد و سیستم های امنیتی استفاده می شود.

در کل این خواص بیومتریک در انسان بسیار زیاد است. عرض و طول صورت، فاصله بین انگشتان دست، طول و عرض انگشت ها، فاصله ی بندها از یکدیگر و حتی خط های کشیده شده کف دست و هزاران خاصیت دیگر، تماماً خصوصیتی هستند که برای انسان ها منحصر به فرد هستند. دوربین هایی وجود دارند که به صورت دید در شب، قادر هستند چیزهایی را که ما نمی بینیم، ببینند و پردازش کنند. اسلحه های خودکاری ساخته شده اند که به صورت اتوماتیک و دقیق نشانه گیری می کنند.

پردازش تصویر همینطور با پردازش تصاویر گرفته شده از فاصله های دور هم می تواند در علوم نظامی و امنیتی کمک کند. به عنوان مثال دوربینی قادر است با سرعت بسیار زیاد یک توپ را دنبال کند. این مسئله کاربرد بسیاری در مسائل نظامی دارد.

6- نجوم و فضا نوردی

ساخت دستگاه های اتوماتیک رصد آسمان و ثبت وقایع آسمانی به صورت خودکار از کاربردهای پردازش تصویر است که امروزه روی آن کار می شود.

از پروژه های جدید در بخش نجوم که بخشی از آن توسط سیستم پردازش تصویر انجام می شود، تهیه نقشه سه بعدی از کل عالم کائنات است!

پردازش تصویر در فضانوردی هم کاربرد زیادی دارد. در تصاویر دور می توان سطح سیارات و همچنین سطح قمرها را اسکن کرده و اطلاعات بسیار ریزی از آنها استخراج کنیم.

کاربرد دیگر پردازش تصویر در فیلتر کردن عکس هایی است که توسط تلسکوپ های فضایی مختلف از جمله هابل (Hubble Space Telescope)، از فضا گرفته می شود.

کاربرد دیگر آن حذف گرد و خاک و جو سیاره ها از تصاویر به کمک تصویربرداری IR و X-RAY به صورت همزمان و ترکیب این تصاویر است.

در تصاویر نزدیک هم کاربرد دارد، از جمله هدایت مریخ نوردها، فرود فضاپیماهای بدون سرنشین و الصاق تجهیزات جدید به ایستگاههای فضایی به صورت خود کار.

از امکانات سایت گوگل، امکاناتی است به نام Google Mars که این برنامه دقیقاً مانند Google Earth عمل می کند با این تفاوت که Google Earth سطح زمین را در هر زمان که بخواهید و در هر نقطه ای از زمین و از ارتفاع های بسیار پائین هم نشان می دهد ولی Google Mars دقیقاً همین کار را برای سطح سیاره مریخ انجام می دهد.

7- پزشکی

یکی از مهمترین کاربردهای پردازش تصویر در علم پزشکی است. در جایی که ما نیاز داریم تمام عکس ها با نهایت شفافیت و وضوح گرفته شوند زیرا دیدن تمام جزئیات لازم است. جراحی های ریز microsurgery با ایجاد یک سوراخ کوچک و فقط دیدن محل جراحی توسط پزشک، از راه دور و توسط بازوهای رباتیک بسیار دقیق انجام می شوند.

8- فناوری های علمی

پردازش تصویر در افزایش سرعت پیشرفت های علمی تأثیر فوق العاده داشته است. اولین و مشخص ترین تأثیر آن را می توان در علم عکاسی یا هنر دید. شکار لحظه های شگفت آوری که در کسری از ثانیه اتفاق می افتد، بالا بردن وضوح عکس های گرفته شده و ایجاد افکت های خیره کننده، از دستاوردهای پردازش تصویر است.

همچنین در توسعه تکنولوژی پیشرفته (Global Positioning Systems) gps کمک زیادی داشته و تهیه نقشه های سه بعدی از جاده ها در تمام نقاط جهان، از کاربردهای دیگر آن است. با به وجود آمدن این علم، مسابقات ربات های فوتبالیست به صورت جدی دنبال شد.

این علم در پیشرفت علوم پایه فیزیک، شیمی و مخصوصاً تحقیقات فیزیکی و مکانیکی، کمک فراوانی کرده است. به عنوان مثال وسیله ای برای حمل و نقل کالاها در مسیرهای صعب العبور ساخته شده است. قبل از ساخت آن، رفتار چهارپایان در حالت های مختلف توسط کامپیوتر تحلیل و عیناً به دستگاه آموزش داده شده است. در کل پردازش تصاویر به علت سرعت زیاد آن، در ساخت وسایل مکانیکی پر سرعت، کاربرد زیادی دارد. وسیله ای وجود دارد که قادر است، تویی که با سرعت بسیار زیاد به سمت پائین می آید را مهار کند.

9- باستان شناسی

در علم باستان شناسی تنها مدارک باقی مانده از دوران باستان، دست نوشته ها، نقاشی ها و غارنگاری های قدیمی است. تهیه تصاویر از بناهای گذشته و بازسازی مجازی این بناهای تاریخی یکی از کاربردهای پردازش تصویر در این علم است. همچنین می توان نقاشی ها و غارنگاری ها را مورد پردازش دقیق قرار داد و شکل آنها را همان طور که در ابتدا بوده اند، شبیه سازی کرد. حتی می توان مکانهای باستانی را از زوایایی که تصاویر مستندی از آنها وجود ندارد، شبیه سازی کرد.

امروزه یکی از پروژه های پر سر و صدای بازسازی بناهای باستانی، بازسازی شهر روم باستان توسط دانشمندان ایتالیایی است. هم اکنون توریست ها با زدن عینک های مخصوص می توانند در خیابان های شهر روم باستان قدم بزنند.

10- تبلیغات

از مقایسه تبلیغات دهه ی 70 و 80 میلادی با تبلیغات امروزی می توان تأثیر تکنولوژی را در تبلیغات کاملاً درک کرد. تغییر شکل تبلیغات از اشکال مربع و زاویه دار به شکل های دایره ای، تغییر رنگ تبلیغات و هزاران تغییر دیگر. یکی از مهمترین فاکتورهای فروش و دلایل بالا رفتن یا پایین آمدن فروش، شکل و نحوه ی بسته بندی کالا است. پردازش تصویر می تواند به ما کمک کند تا قبل از تولید یک بسته بندی آن را شبیه سازی کنیم. با ادغام کردن علم الگوریتم ژنتیک با پردازش تصویر می توان برنامه ای را نوشت که به صورت اتوماتیک به ساختن بسته بندی های مختلف پردازد و آنهایی که از نظر کاربران زیباتر و جالب تر به نظر خواهند آمد را به ما معرفی نماید.

11- سینما

اولین علمی که پردازش تصویر در آن مورد استفاده قرار گرفت، هنر و سینما بود. یکی از تکنولوژی های برتر دنیا motion capture است که در آن یک کاراکتر انیمیشنی قادر است حرکات دست انسان را تقلید کند. امروزه این سیستم جهت ساخت فیلم ها و بازی های کامپیوتری مورد استفاده قرار می گیرد.

در پردازش تصویر قابلیت به نام هیستوگرام (Histogram) وجود دارد که با آن قادرند تصاویر را شفاف یا تیره تر کرده و یا هر تغییر مورد نیاز دیگری را روی تصاویر با توجه به منحنی ها و نمودارهای هیستوگرام بدهند.

در سینما برای اینکه تصویری شفاف به نظر آید، با استفاده از یک کره ی نقره ای رنگ، تصاویر اطراف دوربین را هم ثبت می کنند. بنابراین تصویر نسبت به محیط اطراف خود شفافیت غیر قابل تصویری پیدا می کند.

12- اقتصاد

در دنیای امروز تمام نوآوری ها، به نوعی مستقیم یا غیر مستقیم باعث تغییراتی در اقتصاد گروهی از کشورها و یا کل دنیا می شوند. پردازش تصویر هم، به صورت مستقیم و غیر مستقیم در اقتصاد تأثیر گذار است. در تبلیغات، سیاست، فضانوردی، کشاورزی، شهرسازی، سینما، پزشکی و علوم نظامی می تواند تأثیر غیر مستقیمی در اقتصاد کشورها داشته باشد. همچنین از تأثیر مستقیم آن در اقتصاد، می توان به وجود شعبه های بانک بدون کارمند اشاره کرد. این شعبه ها قادرند به صورت خودکار سربال چک ها و قبوض پرداختی را بخوانند، نوع اسکناس ها را تشخیص دهند و تا حد زیادی از کارهای یک بانک عادی را انجام دهند.

13- روانشناسی

بحث تأثیر رنگ در روحیه انسان اهمیت بسیار زیادی دارد به طوری که در روانشناسی گرایشی به نام روانشناسی رنگ وجود دارد. در این علم در مورد رنگ ها و تأثیر هر یک بر روح و جسم انسان صحبت می شود. به عنوان مثال رنگ قرمز بیشتر تأثیر را در چشم انسان دارد. در حالی که رنگ سبز بیشترین تأثیر را در مغز انسان دارد.

همچنین رنگ آبی باعث ایجاد حس آرامش و اطمینان در انسان می شود. به همین دلیل در سخنرانی های اکثر سیاستمداران دنیا از پرده آبی رنگ در پشت سر آن ها استفاده می شود.

با پردازش تصویر می توان به راحتی تصاویر ثابت و متحرک را ویرایش کرد. به طور مثال رنگ آبی را برای ایجاد حس اطمینان یا رنگ سبز را برای حس زیبایی و قرمز را برای ایجاد هیجان در تصاویر پر رنگ تر کرد.

14- زمین شناسی

با پردازش تصویر می توان کانی های مختلف را از روی رنگ و اندازه آن ها شناسایی و دسته بندی کرد. همچنین در زمین شناسی برای پی بردن به مواد تشکیل دهنده کانی ها از روش پرتونگاری (tomography) استفاده می کنند و پردازش تصویر در این بخش می تواند سرعت و دقت این روش را بسیار بالا ببرد. کاربرد دیگر آن این است که دانشمندان با مقایسه کردن ارتفاع آب در سال های مختلف، در واقع روند تند شدن یا کند شدن کاهش آب در سطح زمین را مورد بررسی قرار می دهند.

نتیجه گیری

رد پای پردازش تصویر در بسیاری از علوم و صنایع مشاهده می شود و بعضی از این کاربردها آنچنان به پردازش تصویر وابسته هستند که بدون آن، اساساً قابل استفاده نمی باشند. کاربرد پردازش تصویر در هر یک از زمینه هایی که بحث شد، بسیار گسترده است