دانشکده مهندسی برق

پایان نامه کارشناسی ارشد

عنوان پایان‌نامه:

تشخیص حرکت در اجزای چهره و بازسازی رفتاری

دانشجو:

آمنه شبیری‌نژاد

استادان راهنما :

دکتر شهریاربرادران شکوهی و دکتر علی صدر

آبان‌ماه 1384

تشکر و قدردانی:

با تشکر از زحمات آقای دکتر شکوهی و آقای دکتر صدر که در طول انجام پروژه از کمک‌ها و راهنمایی‌های ایشان بهره بردم و همچنین زحمت مطالعه و تصحیح پروژه را تقبل نمودند.

این پروژه با حمایت مرکز تحقیقات مخابرات ایران انجام گردیده است.

با تشکر و قدردانی از مسئولان محترم

تقدیم به

پدر و مادر مهربانم

و

همسر عزیزم

چکیده

حالات چهره یکی از قوی‌ترین، متداول‌ترین و سریع‌ترین روشهای انتقال احساسات و عواطف و منظورهای انسان‌ها به یکدیگر می‌باشد. شناسایی حالات چهره توسط کامپیوتر از موضوعات جدیدی است که نظر محققین را در دهه اخیر به خود جلب نموده است. بدلیل رفتار متنوع انسان‌ها و ابعاد وسیع مربوط به حالات چهره در افراد مختلف، شناخت و بررسی این حالات در انسان بسیار پیچیده می‌باشد. لذا با وجود تحقیقات وسیعی که در زمینه‌های مختلف پردازش انجام شده است، تاکنون تشخیص خودکار حالات چهره بطور کامل امکان‌پذیر نشده است.

در حال حاضر تشخیص حالت‌های چهره کاربردهای فراوانی پیدا کرده است که از جمله آنها می‌توان به کاربرد آن در روانشناسی و استفاده در طراحی پویانمایی و فیلم‌های کارتونی اشاره نمود.

موضوع مورد بحث در این پایان‌نامه شناسایی حالات چهره می‌باشد. در ابتدا چهره و اجزای آن مورد بررسی قرار گرفته و روشهای مختلف شناسایی حالات چهره ارائه شده است. پس از بررسی این روشها، شناسایی حالات چهره به کمک دو روش مبتنی بر مشخصه و مبتنی بر الگو مورد بررسی قرار گرفته و با استفاده از بانک اطلاعاتی تصاویر پیاده‌سازی شده‌اند. در پیاده‌سازی روش اول، دو چهره تمام‌رخ و نیم‌رخ مورد آزمایش قرار گرفته‌اند و در پیاده‌سازی روش دوم، از یک طبقه‌بندی‌کننده شبکه عصبی استفاده شده است. این شبکه عصبی یک پرسپترون سه لایه با یک لایه مخفی می‌باشد و توسط الگوریتم یادگیری پس انتشار خطا آموزش داده شده است. میانگین نرخ تشخیص صحیح حالات مختلف چهره با روش‌های مبتنی بر مشخصه و مبتنی بر الگو به ترتیب ۹۲ و ۸۱ درصد بدست آمده است.

فهرست مطالب

عنوان فصل | شماره صفحه

۱- مقدمه ۲

۲- بررسی روشهای شناسایی حالات چهره ۷

۲-۱- مقدمه ۷

۲-۲- آشکارسازی چهره ۹

۲-۳- استخراج مشخصه‌های صورت ۱۲

۲-۴- طبقه‌بندی حالات ۱۵

۲-۵-نتیجه گیری ۱۷

۳- تشخیص حالات چهره با استفاده از روش مبتنی بر مشخصه‌ها ۱۹

۳-۱- مقدمه ۱۹

۳-۲-شناسایی چهره ۲۰

۳-۲-۱- سیستم کدینگ حرکت چهره (FACS) ۲۱

۳-۲-۲-تحلیل خودکار حالات چهره ۲۷

۳-۳- شناخت واحد حرکت برپایه مشخصات چند حالته ۳۰

۳-۴- استخراج مشخصه‌های چهره ۳۲

۳-۴-۱- مدلهای چند حالته اجزای چهره ۳۳

۳-۴-۲- مشخصه‌های پایدار ۳۳

۳-۴-۳- مشخصه‌های گذرا ۳۶

۳-۴-۴- نمونه‌هایی برای استخراج مشخصه ۳۸

فهرست مطالب

عنوان فصل | شماره صفحه

۳-۵- تعیین نقاط مشخص چهره ۴۸

۳-۵-۱- استفاده از چهره تمام‌رخ ۴۸

۳-۵-۲ - استفاده از چهره نیم‌رخ ۵۹

۳-۶- نتیجه گیری ۶۲

۴- تشخیص حالات چهره با استفاده از روش مبتنی بر الگو ۶۴

۴-۱- مقدمه ۶۴

۴-۲- انتخاب یک مدل از حالات ۶۵

۴-۲-۱- حالات پایه ۶۶

۴-۳- جمع‌آوری تصاویر لازم ۶۷

۴-۴- استفاده از یک طبقه‌بندی کننده ۶۸

۴-۴-۱- شبکه عصبی MLP ۶۸

۴-۵- انتخاب یک محیط برنامه نویسی ۷۶

۴-۶- پیاده سازی برنامه ۷۷

۴-۶-۱- بازکردن تصاویر ۷۷

۴-۶-۲- آموزش طبقه‌بندی‌کننده ۸۰

۴-۶-۳- آزمایش طبقه‌بندی‌کننده ۸۱

۴-۷- نتیجه گیری ۸۸

۵- نتیجه گیری و پیشنهادات ۹۰

ضمیمه ۹۴

فهرست منابع و مراجع ۱۰۵

فهرست اشکال

عنوان شکل | شماره صفحه

شکل ۱-۱- حالات چهره و عوامل مؤثر بر آن ۳

شکل ۲-۱- مراحل اصلی شناسایی حالات چهره..... ۷

شکل ۲-۲- مراحل شناسایی حالات چهره..... ۸

شکل ۲-۳- یک چهره انسان در دو سطح خاکستری مختلف..... ۱۰

شکل ۲-۴- یک چهره انسان با دقت‌های مختلف..... ۱۰

شکل ۲-۵- نقاط مشخص چهره (Kobayashi & Hara) ۱۳

شکل ۲-۶- نقاط مرزی چهره (kearney&McKenzie) ۱۳

شکل ۳-۱- سیستم تحلیل خودکار حرکت چهره بر اساس مشخصه‌های چهره..... ۳۱

شکل ۳-۲- ماسک نیمدایره عنبیه.. ..... ۳۶

شکل ۳-۳- محل چین‌های بالای بینی، بین لب و بینی، گوشه‌های بیرونی چشم..... ۳۸

شکل ۳-۴- نتایج ردیابی مشخصه‌های پایدار برای حالات متفاوت از یک فرد. ۴۰

شکل ۳-۵- نتایج ردیابی مشخص‌های پایدار برای افراد متفاوت..... ۴۱

شکل ۳-۶- نتایج ردیابی مشخصه‌های پایدار همراه با حرکت سر ۴۲

شکل ۳-۷- نتایج آشکارسازی چین لب و بینی..... ۴۳

شکل ۳-۸- آشکارسازی چین بالای بینی و چین پاکلاغی. ۴۴

شکل ۳-۹- مشخصه‌های چهره بالایی..... ۴۶

شکل ۳-۱۰- مراحل تشخیص حالات در چهره بالایی ۴۶

شکل ۳-۱۱- مشخصه‌های چهره پایینی. ۴۷

فهرست اشکال

عنوان شکل | شماره صفحه

شکل ۳-۱۲- مراحل تشخیص حالات در چهره پایینی ۴۸

شکل۳-۱۳- نقاط مشخص چهره تعریف شده توسط Kobayashi ۴۹

شکل ۳-۱۴- نرم افزار FCP ۴۹

شکل ۳-۱۵- مدل نقطه‌ای تمام‌رخ چهره ۵۲

شکل ۳-۱۶- دوران سر در تصویر ۵۴

شکل۳-۱۷- مدل نقطه‌ای نیم‌رخ چهره ۵۹

شکل ۴-۱- شش گروه حالات اصلی ۶۵

شکل ۴-۲- نمودار حالات راسل ۶۷

شکل ۴-۳- پرسپترون سه لایه ۶۹

شکل ۴-۴- مراحل اجرای نرم‌افزار Java ۷۶

شکل ۴-۵- پنجره برنامه FER ۷۷

شکل ۴-۶- پنجره نشان دهنده انتخاب مشخصه‌های چهره ۷۸

شکل ۴-۷- پنجره مربوط به تصاویر و مشخصات آن ۷۹

شکل ۴-۸- پنجره آموزش طبقه‌بندی‌کننده ۸۱

شکل ۴-۹- پنجره قسمت آزمایش ۸۲

شکل ۴-۱۰- ترسیم گراف خروجی برنامه ۸۲

شکل ۴-۱۱- جدول خروجی برای تصاویر ۸۳

شکل ۴-۱۲- نمونه‌هایی از خروجی شبکه عصبی برای شناسایی حالات چهره ۸۵

شکل ۵-۱- مدل دو منظره‌ای چهره ۹۳

فهرست جداول

عنوان جدول | شماره صفحه

جدول ۲-۱- خلاصه‌ای از روش‌های آشکارسازی خودکار موقعیت چهره ۱۱
جدول ۲-۲- خلاصه مدلهای چهره استفاده شده در روش‌های شناسایی حالات چهره ۱۴
جدول ۲-۳- خلاصه روشهای استخراج اطلاعات حالات چهره..... ۱۶
جدول ۳-۱- واحدهای حرکت تک در سیستم کدینگ حرکت چهره ۲۲
جدول ۳-۲- سایر واحدهای حرکت چهره..... ۲۳
جدول ۳-۳- واحدهای حرکت چهره بالایی و برخی ترکیبات آنها..... ۲۵
جدول ۳-۴- واحدهای حرکت چهره پایینی و برخی ترکیبات..... ۲۶
جدول ۳-۵- مدلهای چند حالته اجزای چهره در یک چهره تمام‌رخ ۳۵
جدول ۳-۶- نمایش مشخصه‌های چهره بالایی برای شناسایی واحد حرکت ۴۵
جدول ۳-۷- نمایش مشخصه‌های چهره پایینی برای شناسایی واحدهای حرکت ۴۷
جدول ۳-۸- پارامترهای خروجی در برنامه FCP ۵۰
جدول ۳-۹- ارتباط واحدهای حرکت و پارامترهای خروجی برنامهFCP ۵۱
جدول ۳-۱۰- تعریف حالات پایه بر اساس واحدهای حرکت ۵۲
جدول ۳-۱۲- نمایش واحدهای حرکت با مدل تمام رخ چهره ۵۸
جدول ۳-۱۴- نمایش واحدهای حرکت با مدل نیم‌رخ چهره ۶۱
جدول ۴-۱- نتایج نرخ شناسایی به ازای مقادیر مختلف نرخ یادگیری ۷۱
جدول ۴-۲- درصد تشخیص حالات مختلف چهره با استفاده از برنامه FER..... ۸۴
جدول ۵-۱- درصد تشخیص حالات مختلف چهره با استفاده از روش مبتنی بر الگو..... ۹۱
جدول ۵-۲- درصد تشخیص حالات مختلف با استفاده از روش مبتنی بر مشخصه از تصویر تمام رخ ۹۱
جدول ۵-۳- درصد تشخیص حالات مختلف با استفاده از روش مبتنی بر مشخصه از تصویر نیم‌رخ ۹۱

فصل اول

مقـــــدمـــه

مقدمه

ارتباطات انسان‌ها شامل دو بخش اساسی است : گفتاری و غیرگفتاری (بصری). حالات چهره، حرکات بدن و واکنش‌های فیزیولوژیکی نمونه هایی از بخش دوم می‌باشند. تمامی این مثالهای بیان شده، اطلاعات مهمی درمورد حالت مشخص می‌دهند :

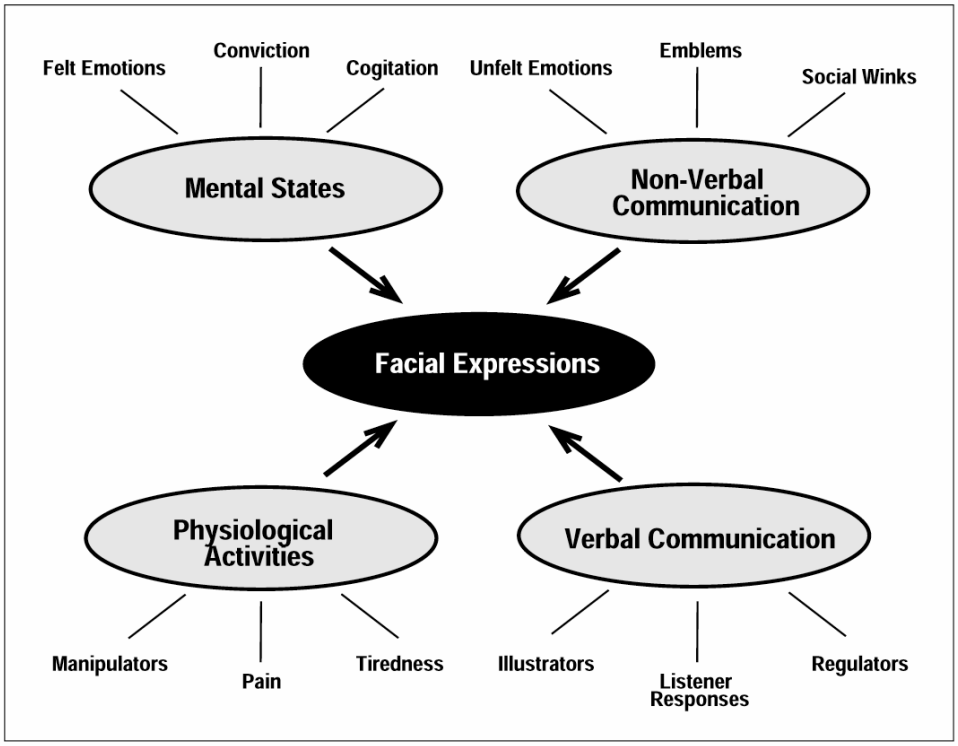
- حالات مؤثر[1] که شامل احساسات مانند ترس، تعجب، عصبانیت، تنفر، ناراحتی و یا حالات پر دوام‌تر مانند احساس رضایت، احساس ملالت یا زودرنجی، می‌باشد.
- فعالیت شناسنده[2] مانند بهت و حیرت، خستگی، و یا تمرکز؛ امور فطری و شخصیتی که شامل ویژگی‌هایی مانند خصومت، اجتماعی بودن یا خجالتی بودن است.
- صداقت و راستگویی که نشان‌دهنده احساسات پنهان شده است و نیز کلیدی برای یافتن حقیقت احساس است هنگامی که کلمه‌ها و لغات حالات درونی انسان را پنهان می‌کند و دروغ می‌گوید.
- آسیب شناسی روانی که نه تنها اطلاعات تشخیصی مربوط به افسردگی و پریشانی، دیوانگی و جنون و سایر اختلالات سبک‌تر روانی را دربرمی‌گیرد بلکه شامل اطلاعات مربوط به پاسخ‌های قابل رؤیت به درمان نیز می‌باشد.

شکل ۱-۱ حالات چهره و عوامل موثر بر آن را نمایش می‌دهد.

---

[1] Affective States

[2] Cognitive Activity



شکل ۱-۱- حالات چهره و عوامل مؤثر بر آن [۱]

تحقیقات قابل ملاحظه‌ای در روانشناسی اجتماعی انجام گرفته که نشان‌دهنده این مطلب است که حالات چهره به هماهنگ‌سازی مکالمات کمک می‌کند و حتی در احساس خوش یا بد آمدن شنونده از سخنان گوینده، حالات چهره بیش از خود کلمات تأثیر گذارند، تا جایی که بخش گفتاری یک پیام تنها ۷٪ و لحن و صدای گوینده ۳۸٪ از تأثیر کل پیام را می‌رسانند، درحالی‌که حالات چهره ۵۵٪ اثر پیام گوینده را در خود دارد [۱]. بنابراین شناسایی حالات چهره بخش مهمی در تحقیقات بینایی ماشین ایفا می‌کند.

هدف بینایی کامپیوتر، نسخه‌برداری و تقلید کردن اثر بینایی انسان توسط دریافت و درک الکترونیکی یک تصویر است. تکنیک‌های بینایی کامپیوتر از نتایج

ریاضی، الگوشناسی، هوش مصنوعی، روانشناسی، علم کامپیوتر، الکترونیک و سایر اصول علمی استفاده می کند[۲].

پیشرفت‌های اخیر در تحلیل تصویر و الگوشناسی، آشکارسازی و طبقه‌بندی خودکار سیگنالهای احساسی و کلامی چهره را ممکن می‌سازد. خودکارسازی تحلیل حالات چهره می‌تواند این حالات را به‌عنوان یک کیفیت جدید و نو به عرصه تقابلات انسان و ماشین وارد نموده و روابط متقابل انسان و ماشین را محکم‌تر و کارآمدتر نماید. چنین سیستمی همچنین می‌تواند طبقه‌بندی حالات چهره را بسیار گسترده و بعنوان ابزاری برای تحقیقات در علوم رفتاری و پزشکی به محققین عرضه کند.

اکثر سیستم‌های تحلیل خودکار حالات چهرۀانسان، سعی در شناخت تعداد محدودی از حالات اصلی مانند خوشحالی، تعجب و ترس را دارد. البته چنین حالات اصلی کمتر اتفاق می‌افتند. روانشناسان اعتقاد دارند که احساسات و مقاصد انسان‌ها اغلب توسط تغییرات در یک یا چند عضو صورت نشان داده می‌شوند[۳].

در این پایان‌نامه پس از بررسی روشهای انجام شده، دو روش شناسایی حالات چهره مبتنی بر مشخصه‌های چهره تمام رخ و نیم رخ پیاده‌سازی می‌شوند، سپس روش سومی به کمک استخراج اطلاعات مبتنی بر الگو و طبقه‌بندی حالات توسط شبکه عصبی پیشنهاد و پیاده سازی می گردد و نتایج سه روش بررسی و مقایسه می شود.

ساختار پایان نامه

این پایان نامه در پنج فصل تنظیم شده است. در فصل دوم پایان نامه مروری بر مراحل انجام شناسایی حالات چهره انجام گرفته است و روشهای موجود برای انجام هر مرحله بررسی شده است. در فصل سوم روش‌های استخراج تشخیص حالات چهره با

استفاده از روش مبتنی بر مشخصه‌ها مورد بررسی و پیاده سازی قرار گرفته است. در این فصل ابتدا توضیحی در مورد روش ارائه شده و سپس روش مبتنی بر مشخصه در رشته تصاویر که کاربرد فراوانی دارد مورد مطالعه قرار گرفته است. در ادامه این روش در تک تصویر در دو چهره تمام رخ و نیم رخ پیاده سازی شده و نتایج کار بررسی و مقیاسه شده اند. در فصل چهارم به معرفی و پیاده‌سازی روشی با استخراج اطلاعات مبتنی بر الگو و طبقه‌بندی کننده شبکه عصبی پرداخته شده است. در فصل پنجم نتیجه‌گیری و پیشنهادات مطرح خواهد شد.در نهایت فهرست منابع ارائه می‌شود.

فصل دوم

بررسی روشهای شناسایی حالات چهره

بررسی روشهای شناسایی حالات چهره

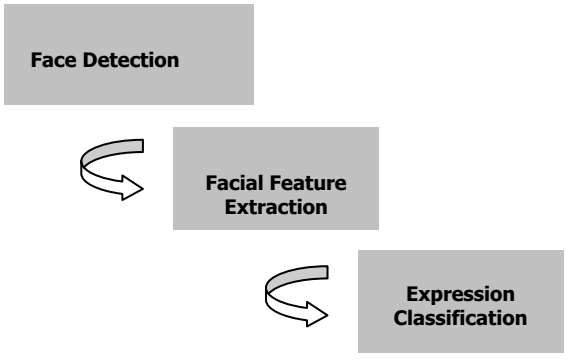
۲-۱- مقدمه

دراینجا مروری بر روش‌های موجود برای شناسایی حالات چهره داریم. در این بررسی بعضی از روش‌های موجود برای تحلیل خودکار حالات چهره در تصاویر استاتیک، تحلیل و با هم مقایسه می‌شوند.

تمام روش‌های موجود برای شناسایی خودکار حالات چهره بطور اساسی دارای سه مرحله می‌باشند:

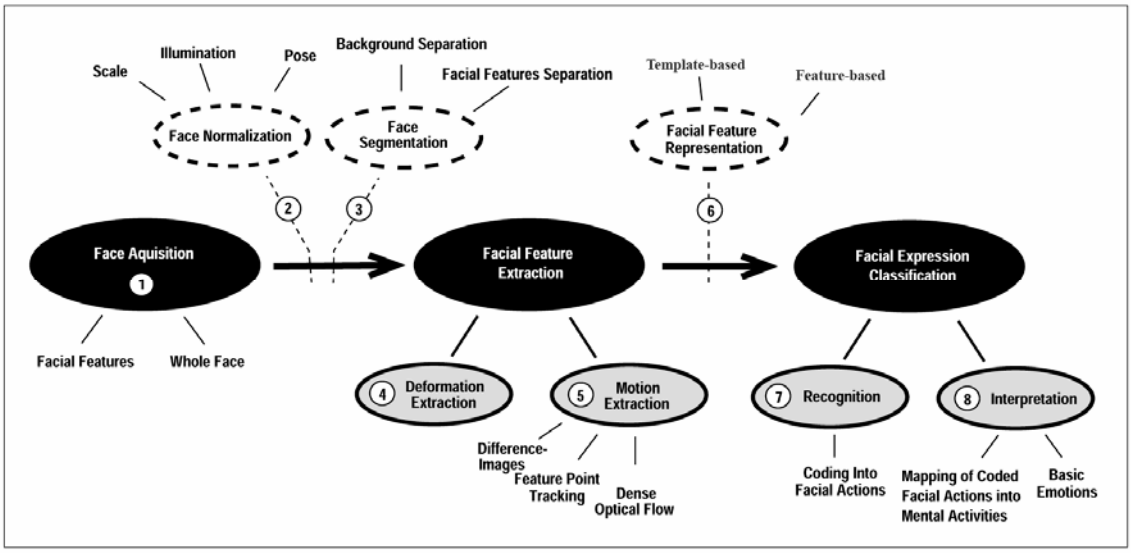
- آشکارسازی چهره[1]
- استخراج داده‌های حالات چهره[2]
- طبقه‌بندی حالات چهره[3]

شکل ۲-۱ مراحل کلی کار و شکل ۲-۲ مراحل را با جزئیات بیشتر بیان می‌کند.



شکل ۲-۱- مراحل اصلی شناسایی حالات چهره

[1] Face detection
[2] Facial expression data extraction
[3] Facial expression Classification



شکل ۲-۲- مراحل شناسایی حالات چهره

داده‌ها از تصویر چهره موجود یا از یک رشته تصویر استخراج می‌شوند. درحالت تصاویر استاتیک، روند استخراج اطلاعات در مورد حالات چهره بعنوان قرارگیری چهره و مشخصه‌های آن در صحنه بیان می‌شود. درحالت رشته تصاویردینامیک ، این روند بعنوان ردیابی چهره و مشخصه‌های در صحنه بیان می‌شود. تشخیص بین مشخصه‌های چهره و مشخصه‌های مدل چهره بسیار دارای اهمیت است. مورد اول یعنی مشخصه‌های چهره، به مشخصه‌های پایدار چهره گفته می‌شود : ابروها، چشمها، بینی، دهان و چانه. درحالی که مشخصه‌های مدل چهره برای نمایش چهره استفاده می‌شوند. چهره می‌تواند به روش‌های مختلفی نمایش داده شود : بعنوان یک واحد کلی، که نمایش Holistic است؛ بصورت مجموعه‌ای از مشخصه‌ها، که نمایش Analytic است؛ و درنهایت بصورت مجموعه‌ای از ترکیب آنها، که روش Hybrid است. آنچه مکانیزم این استخراج، را تعیین می‌کند نمایش چهره اعمال شده و تصاویر ورودی است. پیش از آنکه یک تحلیل‌کننده خودکار حالات چهره طراحی شود، نحوه عملکرد سیستم باید مشخص گردد. و در نهایت هم کار طبقه‌بندی حالات انجام می‌شود و تعیین می‌گردد

که چهره آشکارسازی شده چه حالتی از ۶ حالت : غم، شادی، تنفر، تعجب، ترس و یا عصبانیت را داراست[۳].

۲-۲- آشکارسازی چهره

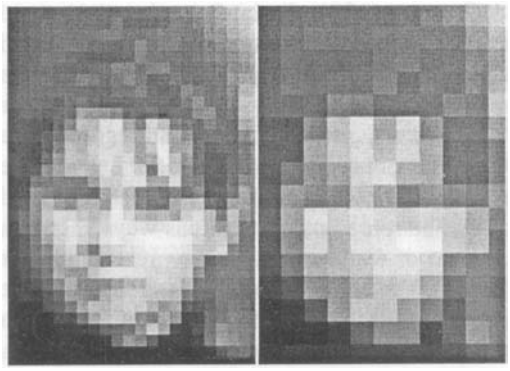
آشکارسازی چهره قدم ابتدائی و اساسی در سیستم‌های شناسایی چهره است. در این مرحله پیدا کردن مکان چهره و استخراج محل چهره از پس زمینه تصویر انجام می‌شود[۴].

آشکارسازی چهره آنگونه که به نظر می‌رسد، ساده نیست. باید مشکلاتی نظیر پیچیدگی زمینه در یک تصویر دیجیتال، مقیاس تصویر و چرخش سر، حضور نویز و غیره را درنظر داشت.

عملکرد سیستم بینائی انسان اینگونه است که تمام چهره را بعنوان یک واحد کلی می‌بیند، نه بصورت مجموعه‌ای از مشخصه‌های مجزای چهره. به نظر می‌رسد حضور مشخصه‌ها و ارتباط هندسی بین آنها مهمتراز جزئیات هر مشخصه باشد [۵]. نکته دیگر این است که انسان‌ها قادرند چهره‌ها را به راحتی و در حوزه وسیعی از شرایط متفاوت شناسایی کنند. مثلاً در شرایط نوری بد و یا از فاصله زیاد نیز بینایی انسان قادر به تشخیص سریع و آسان چهره می‌باشد. شکل ۲-۳ یک چهره را در دو سطح خاکستری با دقت‌های مختلف نشان میدهد. از تصویر چنین استنباط می‌گردد که برای شناسایی چهره یک تصویر با دو سطح خاکستری کافی می‌باشد. شکل ۲-۴ همان تصویر را با تعداد پیکسلهای متفاوت نشان می‌دهد. در تصویر 22x32 پیکسلی بینایی انسان به راحتی قادر به شناسایی یک چهره انسان است. درحالیکه تصویر 11x16 پیکسلی تنها شباهت کمی به چهره انسان دارد. چنین تصور می‌شود که تصاویر ۱۰۰ تا ۲۰۰ پیکسلی

با ۲ سطح خاکستری مرز پائین‌تری از آشکارسازی از یک چهره را برای یک بیننده انسان تشکیل می‌دهد[۳].

شکل ۲-۳- یک چهره انسان در دو سطح خاکستری مختلف
۲۵۶ سطح خاکستری و دو سطح خاکستری



شکل ۲-۴- یک چهره انسان با دقت‌های مختلف
تصاویر **22x32** و **11x16** پیکسلی از یک چهره

با این وجود، در زمینه بینایی ماشین، برخلاف بینایی انسان، آشکارسازی چهره کار آسانی نیست. آشکارسازی یک چهره در یک تصویر دیجیتال، شامل قسمت‌بندی تصویر، استخراج و تشخیص چهره‌ها و احتمالاً مشخصه‌های چهره از یک پس زمینه