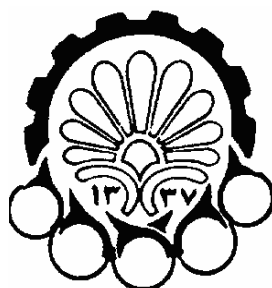


به نام گانه مهربان

ن و القلم و ما یسطرون ...



دانشگاه صنعتی امیرکبیر
دانشکده مهندسی پزشکی

پایان نامه کارشناسی ارشد مهندسی پزشکی - بیوالکترونیک

پردازش تصویر به منظور تعقیب حرکت چشم برای کمک به افراد ناتوان در اجرای کامپیوتری فرامین

استاد راهنما:

دکتر منصور وفادوست

نگارش:

مریم بیگزاده

زمستان ۱۳۸۷



دانشگاه صنعتی امیرکبیر
(پلی تکنیک تهران)

بسمه تعالی

فرم اطلاعات پایان نامه
کارشناسی - ارشد و دکترا

تاریخ:
شماره:

معاونت پژوهشی
فرم پروژه تحصیلات تکمیلی ۷

مشخصات دانشجو:

نام و نام خانوادگی: مریم بیگزاده
شماره دانشجویی: ۸۵۱۳۳۰۵۲
دانشجوی آزاد بورسیه معادل
دانشکده: مهندسی پزشکی رشته تحصیلی: مهندسی پزشکی گروه: بیوالکترونیک

مشخصات استاد راهنما:

نام و نام خانوادگی: دکتر منصور وفادوست
نام و نام خانوادگی:
درجه و رتبه: استادیار
درجه و رتبه:

مشخصات استاد مشاور: -

نام و نام خانوادگی:
نام و نام خانوادگی:
درجه و رتبه:
درجه و رتبه:

عنوان پایان نامه به فارسی: پردازش تصویر به منظور تعقیب حرکت چشم برای کمک به افراد ناتوان در اجرای کامپیوتری فرامین

عنوان پایان نامه به انگلیسی: Image Processing for Eye tracking To help disabled people

نوع پروژه: کارشناسی ارشد
کاربردی بنیادی
سال تحصیلی: دکترا سال تحصیلی: توسعه‌ای نظری

تاریخ شروع: ۱۳۸۶/۹/۲۰ تاریخ خاتمه: ۱۳۸۷/۱۱/۲۰ تعداد واحد: ۶ سازمان تأمین کننده اعتبار: -

واژه‌های کلیدی به فارسی: تعقیب چشم، تعقیب مسیر نگاه، تشخیص اجزاء چهره در تصویر

واژه‌های کلیدی به انگلیسی: Eye Tracking, Gaze Tracking, Facial Features Detection

مشخصات ظاهری	تعداد صفحات	تصویر <input checked="" type="radio"/> جدول <input checked="" type="radio"/> نمودار <input checked="" type="radio"/> نقشه <input type="radio"/> واژه‌نامه <input checked="" type="radio"/>	تعداد مراجع	تعداد صفحات ضمیمه
زبان متن	فارسی <input checked="" type="radio"/> انگلیسی <input type="radio"/>	چکیده	فارسی <input checked="" type="radio"/> انگلیسی <input checked="" type="radio"/>	۱۲
یادداشت				

نظرها و پیشنهادهای به منظور بهبود فعالیت‌های پژوهشی دانشگاه

استاد:

دانشجو:

امضاء استاد راهنما: تاریخ:

چکیده

بیش از سی سال است که سیستم‌های تعقیب حرکت چشم به منظور برقراری ارتباط بین افراد ناتوان حرکتی و کامپیوتر، طراحی و روانه بازار شده‌اند. در یک سیستم تعقیب‌گر حرکت چشم، کاربر در مقابل یک صفحه نمایش قرار می‌گیرد و یک (یا چند) دوربین ویدئویی از چشمان وی تصویربرداری می‌کند. دنباله تصاویر به دست آمده، برای تخمین مسیر نگاه کاربر پردازش می‌شود. تعقیب‌گرهای تجاری موجود در بازار، به دلیل استفاده از دوربین‌های حساس به مادون قرمز و دارای قابلیت بزرگنمایی زیاد، بسیار گران‌قیمت بوده و قیمتی بین چند هزار تا چند ده هزار دلار دارند. از طرفی، در این سیستم‌ها به منظور کنترل شرایط نورپردازی و استفاده از ویژگی خاص مردمک در برابر اشعه مادون قرمز از نورپردازی ویژه‌ای استفاده می‌شود. بسیاری از شرکتهای تولیدکننده دستگاه‌های تعقیب چشم به دنبال بررسی راه‌هایی هستند که امکان استفاده از تکنولوژی ارزان‌قیمتی را در شرایط نوری معمولی فراهم نماید. در این پروژه سیستم ارزان‌قیمتی مبتنی بر وب‌کم طراحی شده است که در شرایط نوری معمولی تصویربرداری می‌کند و نیازی به نصب سخت‌افزار بر روی سر کاربر ندارد و دوربین بر روی میز و در مقابل کاربر قرار می‌گیرد. هدف این پروژه، ارائه روشی به منظور بالا بردن دقت و کاهش قیمت تعقیب‌گر طراحی شده، با هدف کمک به افراد ناتوان حرکتی در کار با کامپیوتر است.

تعقیب‌گرهایی که از کل چهره تصویربرداری می‌کنند، نیاز به انجام مراحل پردازشی متعددی دارند که عبارتند از تشخیص موقعیت چهره در تصویر، تشخیص محل اجزاء چهره (چشم‌ها، لبها و ...)، تشخیص نقاط مشخصه‌ای مانند مختصات مرکز مردمک، گوشه‌های چشم، روش‌های مطرح در زمینه تشخیص و تعقیب چهره و اجزاء آن، به دو دسته سرتاسری و محلی (سلسله مراتبی) تقسیم می‌شوند و در هر مورد ایده‌های مختلفی وجود دارد. در این پروژه دو روش محلی و سرتاسری، پیاده‌سازی و نتایج آنها با یکدیگر مقایسه شده است. همچنین برای ایجاد نداشت از فضای تصویر به مختصات مورد نظر روی نمایشگر از دو شبکه عصبی پرسپترون با یک لایه پنهان استفاده شد که ویژگی‌های در نظر گرفته شده و مدل تعلیم یافته موفق شد به دقت تشخیص ۶,۵۲ درصد دست یابد. این در حالی است که نمونه‌های مشابه مبتنی بر وب‌کم دقت‌های حدود ۸ درصد را گزارش کرده‌اند. مجموعه دادگان مورد استفاده در این پروژه عبارت است از دنباله‌های ویدئویی ضبط شده از ۱۰ فرد مختلف که هر یک در شرایط خاصی (به فاصله ۶۰ سانتی‌متری و در نمای رو به رو) نسبت به دوربین و نمایشگر قرار گرفته‌اند و به نقاط خاصی از صفحه نمایش نگریسته‌اند. در مورد یکی از افراد، علاوه بر نمای رو به رو، دادگانی از دوران‌های مختلف سر نیز ثبت شده تا مقاومت روش را به تغییر وضعیت سر نیز نشان دهد.

کلمات کلیدی: تعقیب چشم، تعقیب مسیر نگاه، تشخیص اجزاء چهره در تصویر

Key Words: Eye Tracking, Gaze Tracking, Facial Features Detection.

فهرست مطالب

مقدمه	۱
تعییب چشم و کاربردهای آن	۱
اجزاء یک سیستم تعقیب گر چشم مبتنی بر ویدئو	۳
اهداف پروژه	۸
ساختار پایان نامه	۱۰
۱- فصل اول - انواع سیستم های تصویری تعقیب چشم	۱۱
۱-۱- مقدمه	۱۱
۲-۱- ویژگی های یک سیستم تعقیب گر چشم	۱۵
۳-۱- تعقیب گرهایی که در آنها حرکت سر مجاز نیست	۲۱
۴-۱- تعقیب گرهایی که در آنها حرکت سر مجاز است	۲۴
۴-۱-۱- سیستم های چند دوربینه	۲۴
۴-۱-۲- سیستم های تک دوربینه	۲۶
۴-۱-۳- سیستم های مبتنی بر نصب دوربین درست در مقابل چشم کاربر	۲۹
۵-۱- جمع بندی مطالب فصل	۳۱
۲- فصل دوم - روش های استخراج نقاط مشخصه از تصاویر چهره و چشم	۳۲
۱-۲- مقدمه	۳۲
۲-۲- روش های استخراج ویژگی از تصاویر چشمی	۳۳
۲-۲-۱- ساختمان و انواع حرکت های چشم	۳۴
۲-۲-۲- مردمک	۳۶
۲-۲-۳- عنبیه	۴۰
۲-۲-۴- گوشه های چشمها	۴۱
۲-۲-۵- پلکها	۴۵
۲-۲-۶- سفیدی چشمها و الگوی تغییر شکل پذیر متناظر با آنها	۴۶
۳-۲- روش های تشخیص چهره و اجزاء آن در تصویر و ویدئو	۴۷

۴۹	۲-۳-۱- تشخیص محدوده احتمالی صورت در تصویر
۵۱	۲-۳-۲- روش‌های محلی تشخیص اجزاء صورت
۵۷	۲-۳-۳- روش‌های سرتاسری تشخیص اجزاء صورت
۷۱	۲-۴-۴- ایجاد نگاهت بین اطلاعات تصویر و موقعیت مورد نظر روی نمایشگر (تخمین مسیر نگاه)
۷۱	۲-۴-۱- روش‌های مبتنی بر مدلسازی سه بعدی چشم در فضا
۷۲	۲-۴-۲- روش‌های مبتنی بر نقاط روشن مینا با بهره‌گیری از تکنولوژی مادون قرمز
۷۳	۲-۴-۳- روش‌های تشخیص مسیر نگاه به کمک تصاویر اخذ شده از کل چهره
۷۴	۲-۵- جمع‌بندی مطالب فصل
۷۵	۳- فصل سوم - روش پیاده‌سازی شده
۷۵	۳-۱- مقدمه
۷۸	۳-۲- تشخیص محل چهره و موقعیت کلی چشم‌ها و لبها
۷۸	۳-۲-۱- تشخیص ماسک صورت
۸۱	۳-۲-۲- تشخیص محل چشم‌ها
۸۳	۳-۲-۳- تشخیص محل لب‌ها
۸۴	۳-۲-۴- رأی‌گیری و انتخاب بهترین کاندیداها
۸۸	۳-۳- تشخیص اجزاء چهره با روش مستقل از مدل
۸۸	۳-۳-۱- استخراج مرکز و شعاع عنبیه در هر چشم
۹۲	۳-۳-۲- استخراج گوشه‌های چشم
۹۵	۳-۳-۳- استخراج گوشه‌های لبها
۹۶	۳-۳-۴- نتایج اجرای متوالی مراحل الگوریتم محلی (سلسله‌مراتبی)
۹۷	۳-۴- تشخیص اجزاء چهره با روش مبتنی بر مدل
۹۸	۳-۴-۱- تعلیم
۱۰۲	۳-۴-۲- تطبیق شکل مناسب به تصویر ورودی
۱۰۵	۳-۵- تعریف ویژگی‌های مناسب و تعلیم مدل کالیبراسیون برای مشخص کردن مسیر نگاه
۱۰۶	۳-۶- جمع‌بندی مطالب فصل
۱۰۸	۴- فصل چهارم - نتایج پیاده‌سازی
۱۰۸	۴-۱- مقدمه
۱۰۹	۴-۲- شرایط تصویربرداری

- ۳-۴- کالیبراسیون ۱۱۱
- ۴-۴- شرایط تست و دقت نهایی ۱۱۵
- ۴-۴-۱- ارزیابی دقت تشخیص برای ۸ نفر ۱۱۵
- ۴-۴-۲- ارزیابی میزان تعمیم‌پذیری روش ۱۱۷
- ۴-۴-۳- ارزیابی مقاومت روش در برابر تغییر وضعیت سر ۱۱۸
- ۴-۴-۴- مقایسه روش محلی (مبتنی بر ظاهر- سلسله مراتبی) با روش مبتنی بر مدل در تشخیص نقاط مشخصه .. ۱۱۹
- ۴-۵- سرعت ۱۲۰
- ۴-۶- جمع‌بندی مطالب فصل ۱۲۲
- ۵- فصل پنجم - جمع‌بندی و ارائه پیشنهادات ۱۲۴
- پیوست ۱ - محاسبات مربوط به روش ASM ۱۲۹
۱. آنالیز مؤلفه‌های اساسی ۱۲۹
۲. هم‌راستا کردن شکل‌ها در مجموعه تعلیم ۱۳۱
۳. نحوه محاسبه وضعیت سر به کمک اطلاعات مدل سه بعدی ۱۳۳
- پیوست ۲ - برخی نمونه‌های تجاری سیستم‌های تعقیب چشم ۱۳۸
- پیوست ۳ - واژه‌نامه ۱۴۵
- فهرست مراجع ۱۴۸

فهرست علائم اختصاری

AAM	Active Appearance Model
ACM	Active Contour Model
ASM	Active Shape Model
CCD	Charge Coupled Device
CR	Corneal Reflection
DCT	Discrete Cosine Transform
EOG	Electro-Oculo-Gram
ET	Eye Tracker
HCI	Human Computer Interaction
HMD	Head Mounted Display
IR	Infra Red
LCD	Liquid Crystal Display
LED	Light Emitting Diode
LoG	Line of Gaze
LoS	Line of Sight
LOO	Leave One Out
NA	Narrow Angle
PCA	Principal Component Analysis
WA	Wide Angle
wpm	Word per minute

فهرست شکل‌ها

- شکل ۱-۰: یک نمونه سیستم تایپ چشمی [13] ۳
- شکل ۲-۰: بلوک دیاگرام مراحل اجرای عملیات تعقیب چشم ۵
- شکل ۳-۰: انواع سیستم‌های تعقیب مسیر نگاه ۵
- شکل ۴-۰: محدوده مجاز حرکت انتقالی سر در محصول EyeFollower شرکت EyeGaze ۸
- شکل ۵-۰: انواع دوران‌های سر ۸
- شکل ۱-۱: تفاوت سیستم‌های تعقیب‌گر از نظر میزان آزادی سر ۱۳
- شکل ۲-۱: انواع صفحه‌کلیدهای طراحی شده ۱۹
- شکل ۳-۱: پدیده مردمک تیره/ روشن در شرایط مادون قرمز ۱۹
- شکل ۴-۱: سیستم تعقیب چشم معرفی شده در مرجع [20] ۲۲
- شکل ۵-۱: ایجاد لکه‌های روشن بر روی چشم به کمک منابع مادون قرمز ۲۳
- شکل ۶-۱: روش تعقیب به کار رفته در مرجع [22] ۲۳
- شکل ۷-۱: استفاده از دو دوربین برای تعقیب مسیر نگاه با در نظر گرفتن حرکت سر [23] ۲۶
- شکل ۸-۱: استفاده از اطلاعات لبه برای تشخیص و تعقیب پلک در دنباله‌های ویدئویی ۲۸
- شکل ۹-۱: تشخیص چشم‌ها در دنباله‌های ویدئویی شامل تصاویر کل چهره در شرایط نوری مادون قرمز [15] ۲۸
- شکل ۱۰-۱: تشخیص موقعیت چشم‌ها در دنباله‌های ویدئویی شامل تصاویر چهره برای کنترل هوشیاری راننده [2] ۲۹
- شکل ۱۱-۱: نصب تعقیب‌گر در مقابل چشمان کاربر ۳۰
- شکل ۱۲-۱: نمونه‌هایی از نمایشگرهای عینکی ۳۰
- شکل ۱۳-۱: سیستم تصویری طراحی شده در مرجع [29] برای تعقیب‌گر و نمایشگر نصب شونده بر روی سر ۳۱
- شکل ۱-۲: ساختمان چشم ۳۶
- شکل ۲-۲: مراحل مختلف روش تطبیق بیضی به مرز عنبیه [42] ۳۹
- شکل ۳-۲: یافتن موقعیت مرکز مردمک با استفاده از اطلاعات لبه [41] ۴۰
- شکل ۴-۲: اعمال تبدیل ویولت گابور به چشم (در ۸ جهت و ۵ مقیاس مختلف) [51] ۴۲
- شکل ۵-۲: فیلترهای گوشه‌یاب به دست آمده از تبدیل ویولت گابور [51]. الف) ۵*۵، ب) ۷*۷ ۴۲
- شکل ۶-۲: استخراج گوشه‌های چشم به کمک فیلترهای گوشه‌یاب. [51] ۴۳
- شکل ۷-۲: استخراج گوشه‌های چشم با روش تعقیب پلکها بر اساس شدت روشنایی [52] ۴۴

- شکل ۸-۲: نتایج استخراج گوشه‌های چشم [52]، الف) از تصویر تیز شده، ب) از تصویر نرم شده. ۴۴
- شکل ۹-۲: تخمین سهمی به پلکها [42]. ۴۶
- شکل ۱۰-۲: تطبیق سهمی به پلکها [51]. ۴۶
- شکل ۱۱-۲: یافتن کانتور مربوط به سفیدی چشمها [52]. ۴۶
- شکل ۱۲-۲: امکان بروز خطا در روش‌های تشخیص چهره مبتنی بر رنگ [60]. ۵۰
- شکل ۱۳-۲: وابستگی رنگ به شدت روشنایی در فضاهاى رنگى مختلف. [57]. ۵۰
- شکل ۱۴-۲: استفاده از مفهوم بافت‌گونه در تشخیص چشمها [62]. ۵۳
- شکل ۱۵-۲: نقشه سطح روشنایی چشمها که با استفاده از عملیات گسترش و فرسایش به دست آمده است [57]. ۵۴
- شکل ۱۶-۲: استفاده از اطلاعات رنگ برای تشخیص لبها [57]. ۵۴
- شکل ۱۷-۲: استفاده از اطلاعات رنگ در تشخیص چشمها [57]. ۵۵
- شکل ۱۸-۲: اعمال عملگر سوبل عمودی به تصویر چهره. لبه‌های عمودی در حوالی چشمها قوی هستند [4]. ۵۶
- شکل ۱۹-۲: حاصل جمع نقاط لبه عمودی در راستای پروفایل‌های عمودی [27]. ۵۶
- شکل ۲۰-۲: نقاط مشخصه چهره. [54]. ۶۰
- شکل ۲۱-۲: نقاط مشخصه مربوط به تصویر چهره. [70]. ۶۰
- شکل ۲۲-۲: استخراج پروفایل سطوح روشنایی در امتداد عمود بر مرز شکل، در نقاط مشخصه [70]. ۶۳
- شکل ۲۳-۲: بررسی تصویر در مقیاس‌های مختلف برای تطبیق بهترین شکل. [70]. ۶۵
- شکل ۲۴-۲: نتایج اجرای الگوریتم جستجوی ASM چند مقیاسه در یک تصویر چهره. [70]. ۶۵
- شکل ۲۵-۲: ایجاد تناظر بین تصویر چهره و مدل سه بعدی سر برای تخمین وضعیت سر در فضای سه بعدی (pose) به کمک نقاط مشخصه [54]. ۶۸
- شکل ۲۶-۲: ماسک‌های مختلف ویولت در شش جهت انتخابی در هر دو دستگاه ویولت موازی برای تشخیص اجزاء چهره [75]. ۶۹
- شکل ۲۷-۲: مراحل روش استفاده از ویولت در تشخیص اجزاء چهره [75]. ۷۰
- شکل ۲۸-۲: استفاده از ویولت در شناسایی اجزاء چهره. [76]. ۷۰
- شکل ۲۹-۲: مدل سه بعدی چشم [42]. ۷۲
- شکل ۱-۳: نمونه‌ای از مجموعه تصاویر تعلیم بر چسب زده شده شامل چهره فرد. ۷۷
- شکل ۲-۳: مدل رنگی مورد استفاده برای تشخیص پوست [57]. ۸۰
- شکل ۳-۳: تعیین محدوده صورت بر اساس فیلتر رنگ پوست. ۸۱
- شکل ۴-۳: استخراج کاندیداهای چشم. ۸۳
- شکل ۵-۳: تشخیص موقعیت چشمها به کمک نقشه لبه‌های عمودی. لبه‌های عمودی در حوالی چشمها تعداد و شدت بیشتری دارند. ۸۴

- شکل ۳-۶: تشخیص کاندیداهای لب‌ها به کمک اطلاعات رنگ..... ۸۴
- شکل ۳-۷: وضعیت‌های مختلف سر و تغییر فواصل چشم‌ها تا لب در اثر دوران جانبی (pan)..... ۸۶
- شکل ۳-۸: نتایج نهایی تشخیص چشم‌ها و لب در تصاویر چهره..... ۸۷
- شکل ۳-۹: استخراج محدوده صورت بر اساس فواصل نسبی چشم‌ها و لب‌ها..... ۸۸
- شکل ۳-۱۰: مراحل تشخیص عنیبه در تصویر چشم..... ۹۱
- شکل ۳-۱۱: مراحل استخراج عنیبه با استفاده از اطلاعات لبه..... ۹۱
- شکل ۳-۱۲: نتیجه نهایی تشخیص عنیبه در تصویر چشم..... ۹۲
- شکل ۳-۱۳: استفاده از اطلاعات گرادیان در تخمین موقعیت گوشه‌های چشم..... ۹۴
- شکل ۳-۱۴: مثال دیگری از استفاده از اطلاعات گرادیان در تخمین موقعیت گوشه‌های چشم..... ۹۴
- شکل ۳-۱۵: ناتوانی عملگر *canny* در تصاویر با کیفیت پایین و عملکرد مناسب روش گرادیان منفی در استخراج پلک‌ها..... ۹۵
- شکل ۳-۱۶: کاهش قدرت روش تشخیص گوشه در تصاویر با کیفیت پایین..... ۹۵
- شکل ۳-۱۷: نتایج روش تخمین دایره به عنیبه و تخمین موقعیت گوشه‌های چشم بر اساس اطلاعات گرادیان..... ۹۵
- شکل ۳-۱۸: نتایج استخراج نقاط مشخصه به روش سلسله مراتبی..... ۹۶
- شکل ۳-۱۹: بروز خطا در مراحل مختلف تشخیص نقاط مشخصه..... ۹۶
- شکل ۳-۲۰: برچسب زدن به تصاویر پایگاه داده برای تعلیم مدل شکل‌های فعال..... ۹۹
- شکل ۳-۲۱: ماسک‌های گابور مورد استفاده در سه فرکانس و شش جهت..... ۱۰۱
- شکل ۳-۲۲: نتیجه اعمال شش فیلتر (در یک فرکانس مشخص و با شش راستای مختلف) گابور به یک تصویر..... ۱۰۱
- شکل ۳-۲۳: چند نمونه از تصاویر تعلیم و نقشه MinMax متناظر با آنها..... ۱۰۲
- شکل ۳-۲۴: نمونه‌هایی از عملکرد روش ASM تغییر یافته در تشخیص نقاط مشخصه چهره..... ۱۰۲
- شکل ۳-۲۵: فلوجارت روش محلی..... ۱۰۳
- شکل ۳-۲۶: فلوجارت روش سرتاسری مبتنی بر الگوریتم شکل‌های فعال..... ۱۰۴
- شکل ۴-۱: صفحه کلید طراحی شده در آزمایش اول..... ۱۱۱
- شکل ۴-۲: صفحه کلید طراحی شده در آزمایش دوم..... ۱۱۱
- شکل ۴-۳: فواصل و زوایای مورد استفاده در تعریف ویژگی‌ها و مفهوم دوران جانبی سر..... ۱۱۳
- شکل ۴-۵: نمونه‌هایی از صفحه کلیدهای سلسله مراتبی..... ۱۲۸
- شکل پ-۱: سیستم تصویری مفروض..... ۱۳۵
- شکل پ-۲: یافتن نقطه پرسپکتیو..... ۱۳۷
- شکل پ-۳: محصولات مختلف شرکت *tobii*..... ۱۳۹
- شکل پ-۴: روش‌های مختلف بهره‌گیری از سیستم تعقیب‌گر مدل X60 و X120..... ۱۳۹

شکل پ-۵: محل قرار گیری نسبی دوربین، نمایشگر و کاربر و همچنین طراحی صفحه کلید در یکی از محصولات شرکت Tobii	۱۳۹
شکل پ-۶: انواع محصولات شرکت EyeGaze	۱۴۱
شکل پ-۷: انواع محصولات شرکت EyeLink و نحوه قرار گیری سخت افزار	۱۴۱
شکل پ-۸: محصولات شرکت EyeTech	۱۴۳
شکل پ-۹: محصول شرکت VisionKey که در آن دوربین در مقابل چشم کاربر نصب شده و صفحه کلید سلسله مراتبی طراحی شده است	۱۴۳

فهرست جدول‌ها

- جدول ۱-۴: میانگین‌های خطا در تشخیص نقطه مورد نظر روی نمایشگر برای ۸ نفر با استفاده از دادگان برچسب دار ... ۱۱۶
- جدول ۲-۴: میانگین‌های خطا در تشخیص نقطه مورد نظر روی نمایشگر در ارزیابی تعمیم پذیری روش کالیبراسیون با استفاده از تصاویر برچسب‌دار چهره ۵ فرد مختلف. ۱۱۷
- جدول ۳-۴: ارزیابی دقت روش‌های محلی و سرتاسری در تشخیص نقاط مشخصه در ۲۸۰ تصویر. ۱۲۱
- جدول ۴-۴: خطای نهایی سیستم بر حسب درجه. ۱۲۳
- جدول ۵-۴: ارزیابی و مقایسه سیستم طراحی شده با نمونه‌های مشابه از نظر دقت. ۱۲۳
- جدول پ-۱: مقایسه محصولات مختلف شرکت Tobii [35]. ۱۴۲

مقدمه

تعقیب چشم و کاربردهای آن

بسیاری از افراد ناتوان حرکتی به دلایلی نظیر بیماری‌های عصبی-عضلانی قادر به استفاده و کنترل اندامهای خود نیستند. در مورد این افراد نگرستن تنها راه ارتباط با دنیای خارج محسوب می‌شود. خوشبختانه امروزه به کمک روشهای پردازش سیگنال و تصویر، ابزارهایی طراحی شده‌اند که از سیگنال نگاه این افراد به عنوان ورودی سیستم برای اجرای برخی دستورات استفاده می‌کنند. یکی از مهم‌ترین سیستم‌ها در این زمینه "سیستم تایپ به کمک تعقیب حرکت چشم"^۱ است. تایپ کردن چشمی^۲ عبارت است از "تولید یک متن به کمک متمرکز کردن چشم روی حروف صفحه کلید".

حدود صد سال است که مطالعه در زمینه تعقیب مسیر نگاه^۳، نه تنها با هدف توان‌بخشی بلکه در سایر زمینه‌ها نیز توجه محققین را به خود جلب کرده است. در حقیقت، بحث تعقیب حرکت چشم با برخی حوزه‌های دیگر نیز در ارتباط بوده و بحث‌های نظری مشترک بسیاری در این میان وجود دارد. برخی از این حوزه‌ها عبارتند از تشخیص و تعقیب چهره در (تصویر) و ویدئو [1]، بازشناسی چهره، سیستم پایش میزان هوشیاری راننده^۴ در حین رانندگی [2-3]، ایجاد ارتباط بین انسان و کامپیوتر به منظور شبیه‌سازی حرکات

1. Eye tracking

2. Eye typing

3. Gaze Tracking

4. driver alertness

چهره در انیمیشن [4-5]، انجام تحقیقات روانشناختی و بررسی میزان توجه در حین مطالعه یک متن یا تماشای یک برنامه تلویزیونی [6-9]. در عین حال، حدود بیست سال است که تعقیب مسیر نگاه با هدف استفاده در کاربردهای توانبخشی و ایجاد ارتباط بین انسان و ماشین^۱ (HCI) مورد توجه محققین قرار گرفته است [10].

باید خاطر نشان کرد که روش‌های تعقیب مسیر نگاه تنها منحصر به رویکردهای مبتنی بر ویدئو نیستند و روش‌هایی وجود دارند که از لنزهای تماسی^۲ و یا از سیگنال‌های زیستی (مانند EOG)^۳ برای اندازه‌گیری میزان حرکت چشم بهره می‌گیرند. اگر چه این روشها به دلیل تماس مستقیم با کاربر، به نوعی تهاجمی^۴ محسوب می‌شوند، اما دقت قابل ملاحظه‌ای دارند و اغلب در مطالعات روانشناختی، سنجش میزان توجه و هوشیاری به کار می‌روند [11]. باید توجه داشت که در این تحقیق منظور از یک سیستم تعقیب‌گر مسیر نگاه، سیستمی مبتنی بر ویدئو است. بنابراین از آنجا که سایر انواع تعقیب‌گرهای چشم از محدوده مطالعه این پروژه خارج هستند، در ادامه این گزارش، سیستم‌های مبتنی بر ویدئو (تصویر) را مختصراً با عنوان سیستم‌های تعقیب چشم (نگاه) یاد خواهیم کرد.

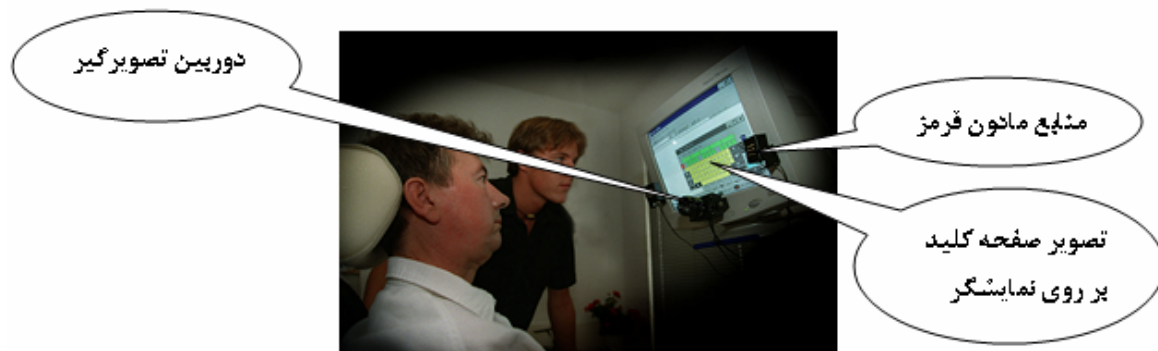
برای تایپ چشمی، تصویر یک صفحه کلید بر نمایشگر نقش می‌بندد و کاربر با نگاه به نقطه مشخصی که تصویر کلید موردنظر وی در آن نقش بسته است، آن را انتخاب کرده و حرف مربوطه تایپ می‌شود (شکل ۰-۱). سیستم‌های دیگری نیز وجود دارند که اختصاصاً با هدف تایپ چشمی طراحی نشده‌اند اما کارکردی شبیه به آن دارند. در این سیستم‌ها صفحه کلید طراحی شده حاوی دستورات مختلفی (نظیر روشن شدن لامپ، شماره‌گیری با تلفن، درخواست آب، غذا و ...) است و کاربر با انتخاب آنها قادر است با دنیای خارج ارتباط برقرار کند. در این تحقیق سیستمی طراحی شده است که می‌تواند در هر دو کاربرد مذکور، مورد استفاده قرار گیرد. در ادامه مطالب این فصل، اجزاء یک سیستم تعقیب‌گر و نحوه عملکرد آن معرفی خواهد شد. سپس مختصراً به روش‌های پردازشی مطرح در این زمینه اشاره خواهیم داشت و با بیان اهداف پروژه و معرفی ساختار پایان‌نامه مطالب این فصل را به پایان خواهیم برد.

¹ . Human Computer Interaction

² . contact lens

³ . Eletro Occulo Gram (EOG)

⁴ . intrinsic



شکل ۱-۰: یک نمونه سیستم تایپ چشمی [13]

اجزاء یک سیستم تعقیب گر چشم مبتنی بر ویدئو

در یک سیستم تعقیب مسیر نگاه، اجزاء و مراحل پردازشی مختلفی مورد نیاز است. بلوک دیاگرام شکل ۲-۰، مراحل اصلی فرآیند تعقیب مسیر نگاه را برای تایپ چشمی نشان می‌دهد. به طور خلاصه، آنچه در یک سیستم تعقیب گر چشم اتفاق می‌افتد، شامل مراحل زیر است:

✓ ابتدا کاربر در مقابل یک صفحه نمایش قرار گرفته و به تصویر صفحه کلید که روی نمایشگر نقش بسته است می‌نگرد. در بیشتر موارد، نگاه کاربر به یک نقطه خاص به معنای انتخاب کلید متناظر با آن موقعیت است.

✓ هم‌زمان یک (یا چند) دوربین ویدئویی از وضعیت چشم (و سر) کاربر تصویربرداری می‌کند. وضعیت نسبی کاربر، نمایشگر و دوربین می‌تواند کاملاً متنوع بوده و این امر در محاسبات و الگوریتم‌های پردازشی مورد استفاده بسیار تأثیرگذار است. به عنوان مثال یکی از حالات ممکن (که در این پروژه نیز اجرا شده است) این است که دوربین در مکانی در مقابل کاربر کنار صفحه نمایش نصب شود و تصویربرداری از چشم و یا کل چهره فرد صورت می‌گیرد (شکل ۳-۰-الف)^۱. حالت دیگر، نصب دوربین روی قابی شبیه به عینک است که روی سر کاربر و درست در مقابل چشمان وی قرار می‌گیرد (شکل ۳-۰-ب)^۲. جزئیات بیشتر در این باره، مزایا و معایب هر سیستم در فصل اول ارائه خواهد شد.

^۱ . Remote Mounted

^۲ . Head Mounted

✓ دنباله‌های ویدئویی ضبط شده به صورت زمان واقعی^۱ پردازش می‌شوند و در هر قاب از ویدئو، پارامتر (یا پارامترهایی) که مشخص کننده مسیر نگاه کاربر است، استخراج می‌شود. این پارامتر می‌تواند شامل مختصات مرکز مردمک باشد. توضیحات بیشتر در این باره نیز به فصل دوم موكول خواهد شد.

✓ برای تعیین و تعقیب مسیر نگاه، لازم است ابتدا کاربر به نقاط مشخصی بنگرد و اصطلاحاً فرآیند کالیبراسیون^۲ را انجام دهد. در هر سیستم، نقاط کالیبراسیون متعددی می‌تواند تعریف شود که مشخص کننده محدوده حرکت چشم هنگام مشاهده سراسر صفحه نمایش باشد؛ مانند نقاط چهار گوشه و مرکز نمایشگر.

طراحی و ساخت ابزارهای تعقیب چشم با هدف برقراری ارتباط بین انسان و کامپیوتر در حدود سی سال قدمت دارد. در این مدت، نمونه‌های تجاری تعقیب‌گرهای چشمی روز به روز تکامل یافته‌تر و با دقت‌های بهتر به بازار عرضه شده‌اند و تلاش در جهت بهبود دقت و کیفیت این محصولات هنوز ادامه دارد. مهم‌ترین ویژگی سیستم‌های تعقیب‌گر چشم تجاری، عبارت است از کاربرد دوربین‌های پر سرعت حساس به نور مادون قرمز^۳ (IR) که قابلیت بزرگنمایی بالایی داشته و قادرند در صورت تنظیم مناسب از فاصله ۴۰ الی ۷۰ سانتی‌متری نسبت به کاربر، تصاویری با ابعاد و قدرت تفکیک بالا از چشم (یا چشمان) وی تهیه کنند. این دستگاه‌ها به دلیل بهره‌گیری از دوربین‌های بسیار حساس دارای قیمت‌های بالایی بوده و استفاده از آنها برای بیشتر افراد ممکن نمی‌باشد. بر اساس قیمت‌های اعلام شده در پایان سال ۲۰۰۸ یک تعقیب‌گر چشم مبتنی بر ویدئو (در بیشتر شرکت‌های تولیدکننده مانند Tobii) قیمتی بین ۴۰۰۰ تا ۱۰۰۰۰۰۰ دلار دارد [12].

علاوه بر قیمت بالای این دستگاه‌ها ویژگی مشترک دیگر آنها این است که همگی آنها از تکنولوژی مبتنی بر تاباندن اشعه مادون قرمز به چشم کاربر استفاده می‌کنند. این کار که نوعی تصویربرداری در شرایط نوری فعال^۴ محسوب می‌شود، باعث می‌گردد که تشخیص موقعیت مردمک با دقت بسیار زیادی انجام شود. به این ترتیب با داشتن تصاویر بزرگ از چشم کاربر و تشخیص دقیق مردمک و مرز آن، این سیستم‌ها به دقت‌های بالایی در تشخیص مسیر نگاه رسیده‌اند که آنها را برای کاربردهایی مانند تایپ چشمی مناسب می‌کند. اگر چه استفاده از تکنولوژی مادون قرمز مزایایی به همراه دارد، اما مهم‌ترین مشکل آن این است که

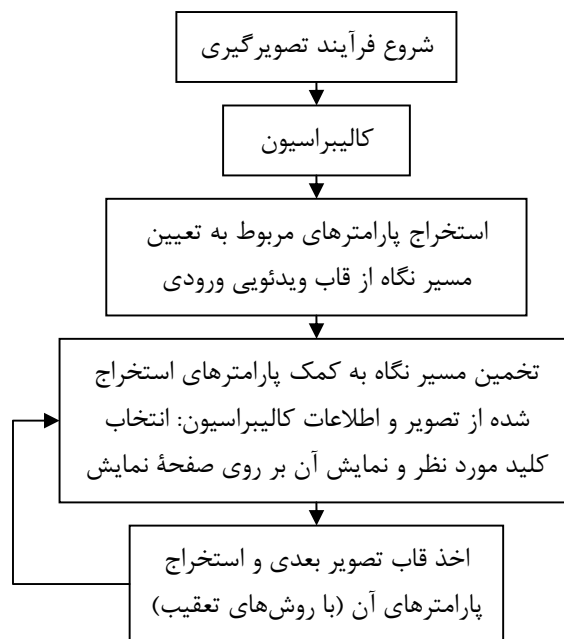
^۱ . real time

^۲ . Calibration

^۳ . Infra Red (IR)

^۴ . Active illumination

در نور آفتاب و به طور کلی در محیط‌هایی که منبع مادون قرمز خارجی وجود دارد، دقت این سیستم‌ها کاهش می‌یابد و بهترین شرایط عملکردی آنها در شرایط نوری فلورسنت و به دور از هر گونه منبع مادون قرمز دیگر می‌باشد. به همین جهت برخی شرکت‌ها به دنبال ارائه روش‌هایی برای کاهش اثرپذیری سیستم خود از شرایط نوری محیط هستند [13].



شکل ۲-۰: بلوک دیاگرام مراحل اجرای عملیات تعقیب چشم.



ب



الف

شکل ۳-۰: انواع سیستم‌های تعقیب مسیر نگاه. الف) نصب سخت‌افزار در مقابل کاربر و بدون ایجاد تماس با او [13]، ب) نصب دوربین تعقیب‌گر در مقابل چشم کاربر در ساختاری شبیه به کلاه یا عینک [36].

بهترین دقت‌های گزارش شده توسط این شرکت‌ها در حدود ۱ درجه می‌باشد. به عبارت دیگر برای کاربری که به فاصله ۶۰ سانتی‌متر از یک نمایشگر قرار گرفته باشد، سیستم قادر است نقاط به فاصله ۱ سانتی‌متری را بر روی نمایشگر از یکدیگر تفکیک نماید ($\tan^{-1}(1/60) \approx 1^\circ$). با وجود اینکه شرکت‌های تولید کننده دقتی در حد ۱ درجه را گزارش کرده‌اند اما متأسفانه اطلاعات چندانی از روش‌های پردازشی به کار رفته در این سیستم‌ها منتشر نشده است. بنابراین برای مقایسه روش‌های پردازشی ناچار به بررسی مقالات منتشر شده در این زمینه هستیم [11]. نمونه‌های تحقیقاتی موجود که با وب‌کم^۱ کار کرده‌اند، دقتی حدود ۸ درجه دارند.

از طرف دیگر مطابق آنچه در مرجع [10] ارائه شده است، دقت سیستم‌های تعقیب‌گر چشم به حدی بیش از ۰٫۵ درجه نخواهد رسید و علت آن محدودیت‌های فیزیولوژیک چشم و ابعاد لکه زرد^۲ در چشم اعلام شده است. به این دلیل و همچنین اینکه قیمت دستگاه‌ها به دلیل تکنولوژی بالای به کار رفته در ساخت اجزاء آن (به خصوص دوربین‌ها) از حد معمول و قابل استفاده اکثریت افراد جامعه ناممکن است، محققین درصدد طراحی سیستم‌های ارزان قیمتی هستند که قادر باشد در شرایط نوری معمولی، تعقیب چشم را با دقت معقولی انجام دهد. اگر چه این امر هنوز کاملاً میسر نشده و دقت سیستم‌های موجود در بازار قابل مقایسه با این سیستم‌های ارزان قیمت مبتنی بر دوربین‌های معمولی نمی‌باشد، اما حتی شرکت‌های تولید کننده‌ای مانند Tobii گروه‌های تحقیقاتی را با همراهی دانشگاه‌ها^۳ تشکیل داده‌اند تا راه‌های ممکن برای طراحی سیستم‌های ارزان قیمت مبتنی بر دوربین‌های ساده وب‌کمرا بررسی نمایند.

در انواع مختلف تعقیب‌گرها بسته به این که تصویرگیری از کل چهره فرد باشد و یا تنها از چشم (های) او تصویربرداری انجام شود، مراحل پردازشی مختلفی انجام شده و برخی از مراحل افزوده و یا حذف می‌شوند. به عنوان مثال اگر تعقیب‌گر از کل چهره فرد تصویر بگیرد، مراحل پردازشی شامل استخراج محدوده صورت، تعیین موقعیت چشمها، تعیین نقاط مشخصه‌ای مانند مرکز مردمک (یا عنبیه)، گوشه‌های چشم، گوشه‌های لب و ... می‌شود. در حالی که در سیستم‌هایی که تنها از چشمان کاربر تصویربرداری می‌کنند، برای تعیین مسیر نگاه تنها کافی است ویژگی‌هایی مانند مرکز مردمک و یا نقاط روشن مبنای^۴ استخراج گردند. در سیستم‌هایی که از منبع مادون قرمز برای کمک به تشخیص مردمک استفاده می‌کنند، نقاط روشن مبنای به

¹. webcam (Web Camera)

². fovea

³. COGAIN group, Tampere University

⁴. glint

عنوان اطلاعات کمکی، به شیوه‌ای که در فصل اول توضیح داده خواهد شد، مورد استفاده قرار می‌گیرند. در این پروژه به دلیل استفاده از یک وب‌کم ساده و تصویربرداری از کل چهره فرد، مراحل پردازشی متعددی انجام می‌شود و نهایتاً با مقایسه دو رویکرد کلی برای تشخیص اجزاء چهره، روش دقیقی برای شناسایی موقعیت نقاط مشخصه در چهره پیشنهاد می‌شود.

تشخیص چهره و اجزاء آن در تصویر مبحث بسیار گسترده‌ای است که کاربردهای متعددی از جمله بازشناسی چهره، بازشناسی حالت چهره، و ... دارد. روشهای مطرح در این حوزه به طور کلی به دو دسته تقسیم می‌شوند: روش‌های محلی (سلسله مراتبی)^۱ و روش‌های سرتاسری (کلی)^۲. در روش‌های سلسله مراتبی ابتدا محدوده صورت تعیین می‌گردد و سپس هر یک از اجزاء آن (چشمها، بینی، لبها) شناسایی می‌شوند. در هر مرحله از ترکیب اطلاعاتی نظیر شدت روشنایی، رنگ، لبه و ... استفاده می‌شود. این در حالی است که در روش‌های سرتاسری تمام اجزاء چهره به صورت هم زمان استخراج می‌شوند. این روش‌ها که روش‌های مبتنی بر مدل^۳ نیز نامیده می‌شوند، از الگوریتم‌هایی مانند مدل شکل‌های فعال^۴ (ASM) استفاده می‌کنند. این رویکردها اگر چه نسبت به روش‌های سلسله مراتبی نسبتاً پرسرعت تر هستند، اما به طور کلی دقت آنها در تشخیص نقاط مشخصه چهره کمتر از روش‌های محلی است. علت این امر آن است که در روش‌های محلی اطلاعات کمتری از تصویر دور ریخته می‌شود. توضیحات بیشتر در رابطه با انواع روش‌های پردازشی به فصل دوم موكول خواهد شد.

امروزه مهم‌ترین اهداف محققین در سیستم‌های تعقیب چشم، حل مشکلات کالیبراسیون فرد به فرد و همچنین کنترل مسأله حرکت سر می‌باشد [11]. دقت یک سیستم تعقیب‌گر به شدت متأثر از حرکت سر کاربر است. در نمونه‌های تجاری موجود ادعا می‌شود که حرکت سر کاربر تا حدی مجاز و قابل جبران است. بسته به دوربین مورد استفاده و ناحیه تحت پوشش آن، بیشترین محدوده حرکت مجاز سر که توسط شرکت EyeGaze اعلام شده است (در محصول جدیدی تحت عنوان EyeFollower)، عبارت است از ۵۰ سانتیمتر به راست و چپ، ۳۰ سانتیمتر به بالا و پائین و ۳۰ سانتیمتر به عقب و جلو (شکل ۰-۴). لازم به تأکید است که تنها نوع حرکت مجاز در این سیستم‌ها حرکت انتقالی سر بوده و از انواع دورانهای سر به

^۱ . local (hierarchical) methods

^۲ . global methods

^۳ . model-based

^۴ . Active Shape Models (ASM)