

## Design, Fabrication and Evaluation of an Eye Tracker to Help Disabled People to Communicate with Computer

D. Mardanbeigi<sup>1</sup>, M. R. Mallakzadeh<sup>2\*</sup>

<sup>1</sup>M.Sc Student, Biomechanics Department, Faculty of Mechanical Engineering, Iran University of Science and Technology, Tehran, Iran, dmardanbeigi@yahoo.com

<sup>2</sup>Assistant Professor, Biomechanics Department, Faculty of Mechanical Engineering, Iran University of Science and Technology, Tehran, Iran.

---

### Abstract

This paper investigates prototyping an online, low-cost, video based and applicable eye tracker, which is called "Dias Eye Tracker". Disabled people can use the proposed system to communicate with computer. What have made the system different from the other low-cost eye trackers, are the accuracy of gaze estimation, the different application parts of the software and the lightweight wireless hardware, which can be mounted on the user's head. This paper introduces the software/hardware and the methods of the system. In addition, two methods of pupil tracking have been compared together, and an uncertainty analysis on the mapping function of the system has been done. The performance of the designed eye tracker has been evaluated by analyzing the answers to the three questionnaires, which were filled by disabled people after performing three specific tasks. The results show that the system performs well for interaction with computer.

**Key words:** Human-computer interaction, Gaze estimation, Eye tracking.

---

\*Corresponding author

**Address:** Mohammad Reza Mallakzadeh, Biomechanics Department, School of Mechanical Engineering, Iran University of Science and Technology, postal code: 16846-13114, Tehran, Iran  
**Tel:** +982177240540-2935  
**Fax:** +982177240488  
**E-mail:** mmallak@iust.ac.ir

## طراحی، ساخت و ارزیابی یک ردیاب چشم جهت توانبخشی معلولان حرکتی در برقراری ارتباط با کامپیوتر

دیاکو مردانگی<sup>۱</sup>، محمدرضا ملاک‌زاده<sup>۲\*</sup>

<sup>۱</sup> دانشجوی کارشناسی ارشد، گروه بیومکانیک، دانشکده مهندسی مکانیک، دانشگاه علم و صنعت ایران، تهران

dmardanbeigi@yahoo.com

<sup>۲</sup> استادیار، گروه بیومکانیک، دانشکده مهندسی مکانیک، دانشگاه علم و صنعت ایران، تهران

### چکیده

در این مقاله به ساخت یک ردیاب حرکات چشم با ویژگی‌های تصویری، بی‌درنگ، ارزان و در عین حال کاربردی با نام **Dias Eye Tracker** اشاره می‌شود. این سیستم برقراری ارتباط معلولان با رایانه را بهبود می‌بخشد. آنچه سیستم مذکور را نسبت به سایر ردیاب‌های تحقیقاتی و ارزان قیمت مخصوص چشم متمایز ساخته، دقت در تشخیص نقطه دید، تنوع بخشهای کاربردی نرم‌افزار و همچنین سخت‌افزار سبک و بی‌سیم آن با قابلیت نصب بر روی سر است. در مقاله حاضر روشهای استفاده شده در سیستم و همچنین نرم‌افزار و سخت‌افزار آن، معرفی شده است. همچنین دو روش استفاده شده برای ردیابی مردمک با یکدیگر مقایسه شده، عدم قطعیت بر روی نگاشت بکار رفته در این سیستم نیز تحلیل شده است. عملکرد این ردیاب چشم برای معلولان، با کمک سه پرسش‌نامه و بررسی پاسخ‌های افراد پس از اجرای سه فعالیت بخصوص با دستگاه، ارزیابی شده است. نتایج این ارزیابی عملکرد خوب این سیستم را در کمک به برقراری ارتباط با کامپیوتر نشان می‌دهد.

کلیدواژه‌گان: ارتباط متقابل با کامپیوتر، تعیین نقطه دید، ردیابی حرکات چشم

\*عهده‌دار مکاتبات

نشانی: تهران، دانشگاه علم و صنعت ایران، دانشکده مهندسی مکانیک، گروه بیومکانیک، کدپستی ۱۳۱۱۴-۱۶۸۴۶

تلفن: ۰۲۹۳۵-۰۲۱۷۷۲۴۰۵۴۰، دورنگار: ۰۲۱۷۷۲۴۰۴۸۸، پیام نگار: mmallak@iust.ac.ir

## ۱- مقدمه

اندازه‌گیری و دنبال کردن حرکات چشم از اواخر قرن نوزدهم میلادی ذهن محققان را به خود مشغول کرده است. امروزه با پیشرفت سیستم‌های کامپیوتری، این امر بسیار دقیق‌تر از قبل محقق شده است. هر چند فناوری ردیابی چشم هنوز در مرحله تحقیقاتی است و سیستم‌های موجود دارای قیمت بالایی هستند، امروزه شاهد کاربردهای قابل توجه این فناوری در زمینه‌های گوناگون پزشکی، توانبخشی، ارگونومی [۱] و بازاریابی هستیم. یکی از کاربردهای فناوری ردیابی چشم در زمینه ارتباط متقابل میان انسان و کامپیوتر<sup>۱</sup> است.

فرایندی که حرکات چشم شخص نسبت به سر او را اندازه‌گیری می‌کند و متعاقباً نقطه دید ناظر را می‌یابد؛ ردیابی حرکات چشم<sup>۲</sup> نام دارد. همچنین وسیله‌ای که فرایند مذکور با آن انجام می‌شود، ردیاب حرکات چشم و یا بطور اختصار ردیاب چشم<sup>۳</sup> نامیده می‌شود [۲]. یک ردیاب چشم را می‌توان براحتی در کنار ابزارهای ارتباطی همچون موشواره و صفحه کلید، بکار برد و با کمک آن بر سرعت اجرای برخی فعالیت‌ها افزود. کار با مرورگرهای اینترنت و دیدن صفحات وب از جمله این فعالیت‌ها هستند [۳].

با توجه به خطای کنونی سیستم‌های ردیاب چشم و طراحی نرم‌افزارهایی با واسط مناسب، می‌توان این سیستم‌ها را جایگزینی مناسب برای موشواره و حتی صفحه کلید دانست. این امر می‌تواند به افراد با محدودیت‌های حرکتی در دست و یا پا کمک کرده؛ آنها را قادر به برقراری ارتباط با رایانه و متعاقباً برقراری ارتباط با دیگران نماید. در سالهای اخیر نیز تحقیقاتی درباره طراحی نرم‌افزارهایی با واسط‌های ویژه‌ای انجام شده است به طوری که کاربر با کمک سیستم ردیاب چشم بتواند در محیط این نرم‌افزارها فعالیت‌های بخصوصی را انجام دهد. از جمله این تحقیقات میتوان به مطالعه اسکوسگارد<sup>۴</sup> و وارد<sup>۵</sup> و همکارانشان [۵، ۴] با هدف امکان تایپ با چشم و همچنین مطالعه تال<sup>۶</sup> و همکارانش [۶] برای ایجاد محیط‌های خاص چندرسانه‌ای اشاره کرد.

تا کنون روشهای فراوانی برای ردیابی حرکات چشم بکار برده شده است. یکی از آنها که امروزه روش مناسبی در این زمینه شناخته شده؛ روش تصویری است. اساس این روش استفاده از تصاویر برداشته شده از چشم و تعیین موقعیت چشم با پردازش این تصاویر است. با استفاده از تصاویر برداشته شده می‌توان برخی مشخصه‌های چشم را اندازه‌گیری کرد و سپس با بررسی تغییرات این مشخصه‌ها مکان چشم و نقطه دید شخص را بدست آورد.

نمونه‌های تجاری ردیاب‌های چشم موجود است؛ اما قیمت بسیار بالای این سیستم‌ها و نبودن نرم‌افزاری یک‌پارچه که قادر به رفع نیازهای اولیه کاربر در برقراری ارتباط با کامپیوتر باشد، از مشکلاتی است که باعث شده امکان استفاده از چنین سیستم‌هایی بسیار محدود شود و نیاز به طراحی ردیاب‌های ارزان قیمت و در عین حال کارا احساس شود.

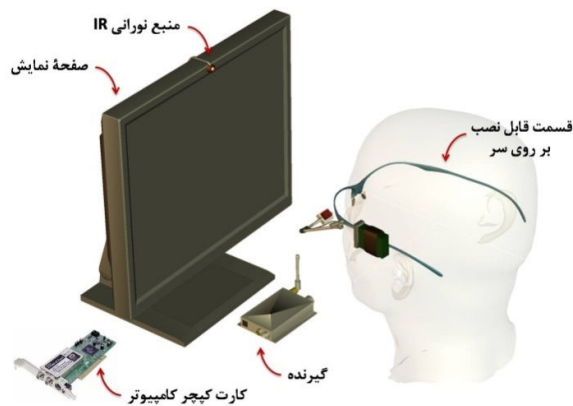
بابکوک<sup>۷</sup> و پلز<sup>۸</sup> [۷] یک ردیاب ارزان قیمت مخصوص چشم را که شامل دو دوربین نصب شده بر روی یک فریم عینک بود؛ مطرح کردند. لی<sup>۹</sup> و همکاران وی [۸] این سیستم را که می‌تواند برای تعیین نقطه دید شخص بکار رود؛ با نام OpenEye گسترش داده‌اند. از آنجایی که پردازش و ردیابی چشم در سیستم مذکور به صورت آفلاین بر روی تصاویر ضبط شده انجام می‌شود و خروجی سیستم، نقطه دید در فضای تصویر میدان دید است؛ امکان استفاده از آن برای برقراری ارتباط آنلاین با رایانه وجود ندارد. زلینسکی<sup>۱۰</sup> [۹] ردیابی برای حرکات چشم با نام Opengazer را مطرح کرده که شامل یک وب‌کم ساده نصب شده بر روی صفحه نمایش است و جهت برقراری ارتباط با رایانه از آن استفاده می‌شود. از آنجایی که روش استفاده شده در سیستم مذکور در برابر حرکات سر کاربر، نسبت به صفحه نمایش مقاوم نیست، سر کاربر باید حین کالیبراسیون و استفاده از ردیاب چشم کاملاً بی‌حرکت باشد. هر دو پروژه مذکور در سیستم عامل لینوکس اجرا می‌شوند و برای استفاده در ویندوز طراحی نشده‌اند.

این مقاله به معرفی یک ردیاب حرکات چشم مبتنی بر روش تصویری اختصاص یافته است که بطور خلاصه در مراجع [۱۰] و [۱۱] به بخشهای مختلف و عملکرد آن اشاره

<sup>۱</sup>Human-Computer Interaction<sup>۵</sup>Ward<sup>۹</sup>Li<sup>۲</sup>Eye Tracking<sup>۶</sup>Tall<sup>۱۰</sup>Zielinski<sup>۳</sup>Eye Tracker<sup>۷</sup>Babcock<sup>۴</sup>Skovsgaard<sup>۸</sup>Pelz

### ۳- بخش‌های سخت‌افزاری ردیاب ساخته شده برای چشم

مهمترین بخش سخت‌افزاری ردیاب چشم، دوربین آن است. دوربین استفاده شده در این سیستم، علاوه بر داشتن مداری کوچک و سبک، به شکل بی‌سیم بوده؛ تصویر را از طریق یک فرستنده کوچک به گیرنده‌ای که معمولاً روی میز قرار داده می‌شود؛ ارسال می‌کند.



شکل ۱- بخش‌های مختلف سخت‌افزاری سیستم

به طور کلی اجزای اصلی ردیاب ساخته شده برای چشم- که در شکل (۱) نشان داده شده است- عبارتند از: دوربین بی‌سیم، گیرنده، کارت کچر رایانه، منبع نورانی مادون قرمز و یک بخش قابل نصب بر روی سر که دوربین را در فاصله مشخصی از چشم نگاه می‌دارد.

برای نصب دوربین بر روی سر، از یک فریم سبک عینک استفاده شده است. اتصال دوربین به قاب عینک بسیار حائز اهمیت است؛ چرا که موقعیت دوربین نسبت به چشم بایستی قابل تنظیم باشد و علاوه بر آن مکان دوربین نیز در برابر حرکات سر، باید ثابت و بدون لرزش باشد. مکانیزم نگهدارنده دوربین به گونه‌ای است که ۵ درجه آزادی را برای آن فراهم می‌کند. بدین ترتیب براحتی می‌توان موقعیت دوربین را مقابل چشم تنظیم کرد تا تصویر مناسبی از چشم بدست آید. شکل ۲ بخش قابل نصب بر سر را نشان می‌دهد.

شده است.

راحتی استفاده از یک سیستم سخت‌افزاری و پایین بودن هزینه تمام شده آن، دو مزیت اصلی در طراحی آن سیستم محسوب می‌شوند. قابلیت استفاده در سیستم عامل ویندوز، سادگی، سرعت عمل و در عین حال کارایی آن، سه مشخصه مهمی هستند که در روش‌ها و الگوریتم‌های استفاده شده در نرم‌افزار سیستم مذکور در نظر گرفته شده‌اند. به‌علاوه سعی شده بخش‌های کاربردی متنوعی در نرم‌افزار این سیستم گنجانده شود. ارزیابی عملکرد ردیاب ساخته شده نیز بخش مهمی از این کار تحقیقاتی است که در ادامه توضیح داده می‌شود.

سه بخش اصلی این مقاله عبارتند از: طراحی و ساخت سخت‌افزار ردیاب چشم، معرفی روش‌های استفاده شده در سیستم و نرم‌افزار طراحی شده آن و ارزیابی سیستم.

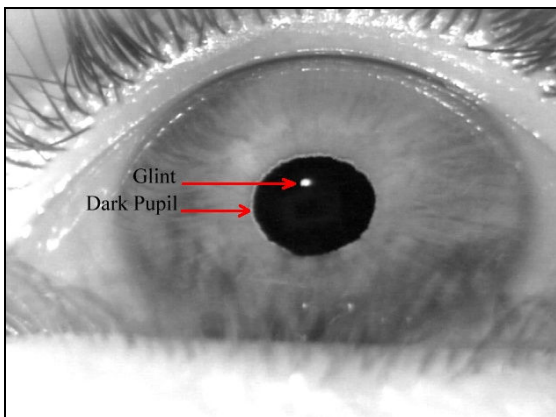
### ۲- ردیابی حرکات چشم با روش تصویری

به طور کلی سه روش برای اندازه‌گیری حرکات چشم<sup>۱۱</sup> بکار برده می‌شود [۱۲]:

- (۱) اندازه‌گیری به روش الکتریکی؛
- (۲) استفاده از لنزهای تماسی؛
- (۳) اندازه‌گیری با کمک تصویر.

دو روش اول دقت بالایی دارند؛ اما کاملاً تهاجمی‌اند و استفاده از آنها تقریباً محدود به فعالیت‌های آزمایشگاهی است. روش سوم بر اساس استفاده از تصاویر برداشته شده از چشم و تعیین موقعیت چشم با پردازش این تصاویر است. با کمک این تصاویر می‌توان برخی مشخصه‌های چشم را اندازه‌گیری کرد و سپس با بررسی تغییرات این مشخصه‌ها میزان حرکت چشم بدست می‌آید. از مشخصه‌های قابل اندازه‌گیری چشم می‌توان به مرکز مردمک، شکل مردمک و موقعیت لیمبوس (مرز میان عنبیه و صلبیه) اشاره کرد.

<sup>۱۱</sup> OculoGraphy

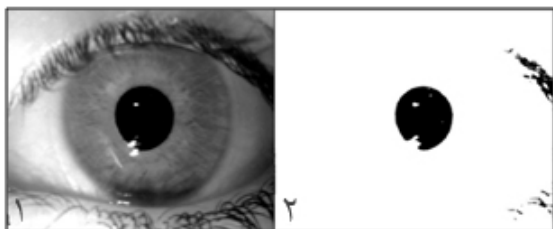


شکل ۳- گلینت و مردمک تاریک در تصویر چشم

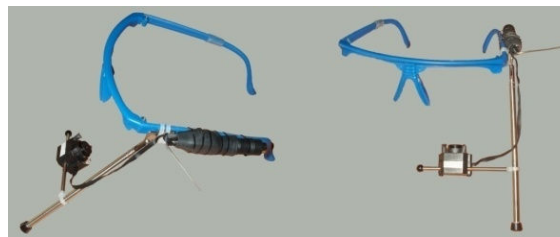
بازتاب اول را گلینت<sup>۱۳</sup> گویند که در تصویر به صورت یک نقطه نورانی بر سطح قرنیه دیده می‌شود (شکل ۳). وضعیت مرکز مردمک نسبت به گلینت تعیین کننده نقطه دید ناظر است؛ زیرا وقتی سر حرکت کند تا زمانی که چشم به یک نقطه مشخص خیره شده باشد مکان گلینت و مردمک نسبت به یکدیگر ثابت خواهد بود؛ اما اگر شخص به نقطه دیگری خیره شود، موقعیت نقطه نورانی نسبت به مرکز مردمک تغییر می‌کند. بنابراین بردار واصل مرکز مردمک به مرکز گلینت، معیار اندازه‌گیری نقطه دید است.

#### ۲-۴- ردیابی مردمک

در ردیاب ساخته شده پس از گرفتن تصویر، باید مرکز مردمک در تصویر مشخص شود. برای یافتن ناحیه حدودی مردمک، ابتدا از تصویر آستانه‌گیری می‌شود. با فرض آنکه مردمک در تصویر بدست آمده به صورت یک دایره و یا یک بیضی تیره‌رنگ در زمینه‌ای روشن دیده می‌شود، می‌توان ناحیه حدودی مردمک را، مطابق شکل ۴، تعیین کرد. پس از آن مرکز چهارضلعی محیط بر مردمک به جای مرکز مردمک در تصویر، شناخته شده؛ در مراحل بعدی استفاده می‌شود.



شکل ۴- (چپ) تصویر چشم (راست) تصویر آستانه‌گیری شده از چشم



شکل ۲- بخش قابل نصب بر سر

تنظیم دوربین با نیروی اندک دست امکان‌پذیر بوده؛ این مرحله بسیار سریع انجام می‌شود. وزن این سیستم بدون باتری تقریباً ۴۵ گرم است.

#### ۴- روش

در یک سیستم بر پایه روش تصویری (VOG) لازم است در هر لحظه، چشم در میدان دید دوربین قرار داشته باشد. این مسأله در ردیاب‌های قابل نصب بر روی سر خود به خود حل می‌شود؛ چرا که با حرکت سر، دوربین مقابل چشم نیز حرکت می‌کند و مکان چشم در میدان دید دوربین تغییر نمی‌کند. در ردیاب‌هایی که بر روی میز نصب می‌شوند، معمولاً از دو دوربین و اجرای مکانیزم حرکتی برای یکی از دوربین‌ها استفاده می‌شود. برای مثال می‌توان به مراجع [۱۳، ۱۴] اشاره کرد. نکته مهم دیگر کاهش تأثیر حرکات سر در طول ردیابی چشم است. یکی از تکنیک‌هایی که به این منظور استفاده می‌شود، تکنیک بازتاب قرنیه- مردمک است؛ که در ادامه توضیح داده خواهد شد.

در ردیاب ساخته شده، از روش تصویری و ردیابی مادون قرمز با مردمک تاریک استفاده شده است؛ به طوری که برای کاهش تأثیر حرکات سر در آن نیز امکان استفاده از ایده بازتاب قرنیه- مردمک وجود دارد.

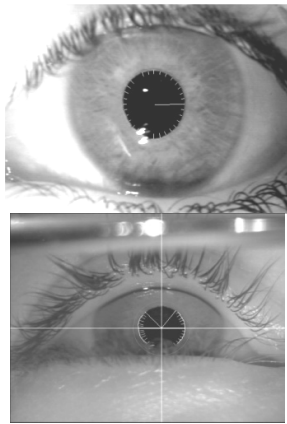
#### ۱-۴- روش بازتاب قرنیه- مردمک

برای کاستن تأثیر حرکات سر در روش‌های تصویری از ایده بازتاب قرنیه- مردمک استفاده می‌شود. هنگامی که پرتو نور یک منبع روشنایی به چشم می‌تابد چندین نوع بازتاب نور در لایه‌های مختلف چشم انجام می‌شود که تصاویر پورکنیا<sup>۱۲</sup> نام دارند.

<sup>12</sup> Purkinje Images

<sup>13</sup> Glint

نقطه‌ای یافت نشود، جستجو را از مرکز آغاز می‌کنیم. در شکل (۵ چپ) نمونه‌ای از این فرایند نشان داده شده است. نکته دیگری که لازم است به آن اشاره شود، این است که برخی افراد دارای چشمان باریک هستند و یا مژه‌های بلندی دارند و برخی دیگر نیز از عینک استفاده می‌کنند و ما ناچاریم زاویه دوربین را با خط دید بیشتر کنیم و تصویر را از پایین‌تر برداشت کنیم. در این افراد مرز بالا و پایین مردمک در تصویر معمولاً در زیر پلک‌ها و یا مژه‌ها پنهان می‌شود و یافتن نقاط مرزی در این نواحی دشوار است. برای اجتناب از چنین حالت‌هایی زاویه جستجوی نقاط مرزی را همانند شکل ۵ (راست)، محدود کرده، از یافتن نقاط بالا و پایین مرز مردمک صرف نظر می‌کنیم.



شکل ۵- (چپ) روش بهینه شده جستجوی شعاعی (راست) جستجوی بهینه شعاعی نقاط مرزی مردمک با زاویه محدود در دو طرف مردمک

### ۳-۴- تشخیص گلینت در تصویر

پس از تشخیص مرکز مردمک به تعیین مرکز گلینت می‌پردازیم (شکل ۶). برای تشخیص گلینت و تعیین مرکز آن در تصویر، همانند روش ردیابی مردمک عمل کرده؛ به گونه‌ای که ابتدا مکان حدودی گلینت را در تصویر شناسایی می‌کنیم و سپس به شناسایی دقیق‌تر مرکز آن می‌پردازیم. برای تعیین ناحیه حدودی گلینت نیز باید محدوده جستجو را به اطراف مردمک محدود کنیم.

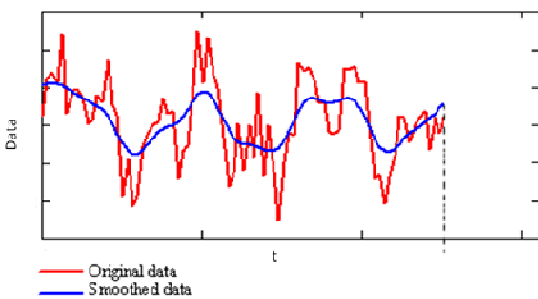
در استفاده از این روش به دو مسأله باید توجه قرار کرد. یکی تغییر قطر مردمک بوده؛ دیگری وجود نواحی روشن در تصویر است که باعث یک تغییر جزئی چند پیکسلی در مختصات نقطه مرکزی مردمک می‌شود. تغییر قطر مردمک قابل تشخیص است و در فریم‌هایی که قطر مردمک در مقایسه با فریم قبلی تغییر کرده، مرکز مردمک را تغییر نداده و مقدار آن از فریم قبلی برداشت می‌کنیم. برای رفع عامل دوم خطا نیز موقعیت اضلاع چهارضلعی محیطی مردمک در هر فریم، ذخیره شده و در فریم بعدی با مکان جدید آنها مقایسه می‌شود؛ چنانچه تنها یکی از دو ضلع افقی و یا عمودی جابجایی داشته باشد، مقدار قبلی موقعیت آنها در نظر گرفته می‌شود. این امر موجب حذف بخشی از جابجایی‌های غیر واقعی مرکز مردمک می‌شود.

برای بهبود دقت اندازه‌گیری مرکز مردمک می‌توان از یک گام تکمیلی استفاده کرد. در این روش از آرایه نقاط مرزی استفاده کرده؛ به طور قراردادی میانگین آنها را به جای مرکز مردمک در نظر می‌گیریم. در این حالت جابجایی چند نقطه در آرایه نقاط، تغییرات کمتری را نسبت به جابجایی چهار نقطه کناری در روش قبلی ایجاد می‌کند. برای تعیین آرایه نقاط مرزی مردمک، از روش استاربرست<sup>۱۴</sup> [۱۵] استفاده می‌شود، با این تفاوت که برای پیدا کردن نقاط مرزی مردمک با دقت مطلوب، می‌بایست از جستجوی نقاط در محدوده نقاط نورانی بر روی مردمک صرف نظر کرد. بدین منظور از شعاع‌هایی که در مسیر جستجو به پیکسلی با سطح روشنایی بیش از ۲۰۰ برخورد کنند؛ صرف نظر خواهد شد. در این فرایند باید نقاط زیادی کنترل شوند، بدین ترتیب زمان پردازش افزایش می‌یابد. البته که با کمک ایده‌ای زمان پردازش را تا حدی بهینه کرد. بدین صورت که اگر در شعاع اول نقطه مرزی در فاصله  $r_1$  از مرکز قرار داشته باشد، در شعاع دوم دامنه جستجو را به  $[r_1 - \delta, r_2 + \delta]$  کاهش می‌دهیم، چنانچه در این بازه نقطه مرزی در فاصله  $r_2$  یافت شود، در شعاع سوم نیز بازه جستجو را  $[r_1 - \delta, r_2 + \delta]$  قرار می‌دهیم. به این ترتیب زمان جستجو بسیار کمتر و در حدود  $\frac{1}{4}$  حالت پیش می‌شود. البته در هر شعاع چنانچه در بازه مربوطه

<sup>14</sup> STARBURST

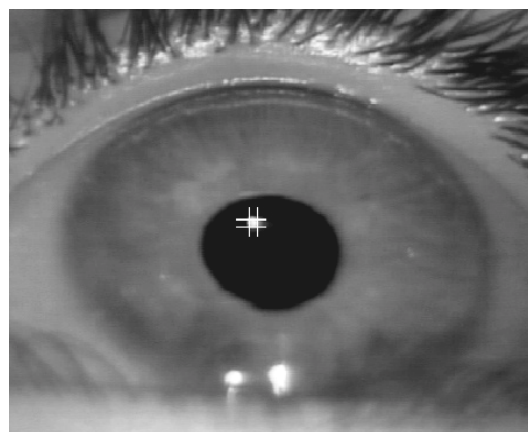
نقطه، دو معادله بدست می‌آید. در پایان با کمک روش کمترین مربعات، بهترین مدل دو جمله‌ای برازش شده محاسبه می‌شود.

آنچه در بخش تشخیص مرکز مردمک و گلینت بدست آمد، مختصات دو نقطه بود. این نقاط مستقیماً از هر فریم تصویر بدست می‌آمد. این داده‌ها در واحد زمان همچون یک سیگنال خشک<sup>۱۵</sup> و پرافت و خیز هستند. علت این امر وجود نویزهایی در حد پیکسل، اختلافات جزئی در حد پیکسل میان دو فریم و همچنین خطاهای اندازه‌گیری و تشخیص اجزا در تصویر است. همین وضع برای خروجی نگاشت، یعنی نقاط دید محاسبه شده نیز وجود دارد. بنابراین آنچه در مرحله اول ضروری به نظر می‌رسد اعمال صافی‌هایی بر این داده‌ها است تا از یک سیگنال خشک به سیگنالی نرم تبدیل شوند (شکل ۷).



شکل ۷- یک نمونه از داده اولیه به همراه داده نرم شده آن

یکی از نرم‌کننده‌های معروف داده‌های رقمی، نرم‌کننده کرنل<sup>۱۶</sup> است. یکی از این نوع نرم‌کننده‌ها، نرم‌کننده کرنل گاوسی<sup>۱۷</sup> نام دارد و برای داده‌هایی که خطای آنها در یک محدوده مشخص است و مدل قابل پیش‌بینی ندارند، بسیار مناسب است. میزان نرم کردن نیز با یک پارامتر ویژه، قابل تغییر است. در اینجا تنها به بیان رابطه این نرم‌کننده و نحوه اعمال آن بر داده‌ها بسنده می‌کنیم. نرم‌کننده کرنل گاوسی از داده‌های قبلی برای تخمین داده کنونی استفاده می‌کند. از آنجایی که تعداد داده‌ها در یک ردیاب چشم بی‌درنگ، از لحظه شروع ردیابی دائماً در حال افزایش است و نامحدود فرض می‌شود، همواره داده فریم کنونی و  $n$  فریم



شکل ۶- تشخیص گلینت در اطراف مرکز مردمک

## ۵- کالیبراسیون

در ردیاب ساخته شده از نگاشت غیر خطی مرتبه دو استفاده شده است. این نگاشت را اولین بار موریموتو [۱۶] استفاده کرد و بصورت زیر است:

$$\begin{cases} S_x = a_0 + a_1x + a_2y + a_3xy + a_4x^2 + a_5y^2 \\ S_y = b_0 + b_1x + b_2y + b_3xy + b_4x^2 + b_5y^2 \end{cases} \quad (1)$$

که در آن،  $(S_x, S_y)$  مکان نشانگر در صفحه یا همان نقطه دید شخص و  $(x, y)$  پارامتر اندازه‌گیری شده در تصویر چشم است.  $(x, y)$  می‌تواند مختصات مرکز مردمک در تصویر یا مؤلفه‌های بردار مردمک- گلینت در روش بازتاب قرنیه- مردمک باشد.

طبق رابطه (۱)، تعداد ۱۲ پارامتر مجهول در این نگاشت برداری دوبعدی دیده می‌شود که با روش کمترین مربعات طی مرحله کالیبراسیون تعیین می‌شوند. در نرم‌افزار سیستم امکان کالیبراسیون با ۹، ۱۶ و ۲۵ نقطه وجود دارد. در مرحله کالیبراسیون نقاطی در مکان‌های مشخص از صفحه برای مدت زمان کوتاهی نشان داده می‌شوند و کاربر باید به آنها خیره شود. بردار مردمک- گلینت در مدت زمان نمایش نقطه بر روی صفحه، اندازه‌گیری و ذخیره شده؛ در پایان میانگین آن به عنوان  $(x, y)$  در نظر گرفته می‌شود. مختصات نقطه‌ای که نمایش داده می‌شود نیز در واقع همان نقطه دید واقعی یا  $(S_x, S_y)$  در رابطه (۱) است. بنابراین با نگاه کردن به هر

<sup>15</sup> Rough

<sup>16</sup> kernel smoother

<sup>17</sup> Gaussian kernel smoother

بسیک نوشته شده و دارای امکانات متنوعی است. شش بخش اصلی در نرم افزار گنجانده شده که این شش بخش عبارتند از: تنظیمات تصویربرداری، ردیابی چشم، کالیبراسیون، حرکت دادن نشانگر با چشم، کاربردها و ترسیم نمودار.

بخش های ویژه تایپ، بازی و سرگرمی، نمایش آلبوم عکس و پخش فایل های صوتی و امکانات جانبی همچون رسم نمودار از داده های مربوط به ردیابی چشم و برخی مشخصات ظاهری چشم و همچنین رسم الگوهای حرکتی چشم بر حسب زمان، هنگام دیدن فایل های ویدئویی و عکس، همگی از جمله امکانات این نرم افزار به شمار می رود.

در بخش مربوط به کاربردها، یک صفحه مخصوص تایپ وجود دارد. در این صفحه حروف به ترتیب و با نظم خاصی چیده شده اند. فرد با مشاهده حروف و انتخاب آنها می تواند جمله مورد نظر را تایپ کند. این بخش نرم افزار علاوه بر زبان انگلیسی از زبان فارسی نیز پشتیبانی می کند (شکل ۸). در ردیاب ساخته شده نحوه کلیک کردن و یا انتخاب با چشم، با دو روش مختلف امکان پذیر است. کلیک کردن با روش مکث Dwell-time و همچنین کلیک کردن به کمک ترکیب پلک زدن و مکث. در حالت کلیک به روش مکث، چنانچه کاربر برای مدتی مشخص و قابل تنظیم به یک آیتم خیره شود، آن آیتم انتخاب می شود. در حالت دوم کلیک کردن، هنگامی که شخص به نقطه ای نگاه کند و پلک بزند، یک علامت دایره ای کوچک در آن نقطه از صفحه ظاهر می شود. چنانچه مرکز دایره مذکور همان نقطه مورد نظر کاربر برای کلیک کردن باشد، کاربر باید برای مدت زمانی به دایره خیره شود تا آن را انتخاب کند؛ در غیر این صورت با خروج نقطه دید از دایره مذکور، دایره حذف شده و به منزله انصراف از کلیک کردن است.

بخش دیگر مربوط به بازی و سرگرمی بوده که در این بخش تعداد زیادی بازی در رابطه با هوش، حافظه و غیره گنجانده شده است و همگی تنها با کمک چشم قابل اجرا هستند. با کمک فهرست اصلی این صفحه، کاربر می تواند به

قبلی را بررسی می کنیم.  $N$  نیز بصورت اختیاری قابل تنظیم است. مقادیر بیشتر از ۵۰ برای  $n$  در تخمین داده جدید تقریباً بی تأثیر بوده و تنها زمان محاسبه را افزایش می دهد، بنابراین مقدار آن کمتر از ۵۰ در نظر گرفته می شود.

داده تخمین زده شده  $i$  ام از رابطه زیر بدست می آید:

$$y'_i = \frac{\sum_{j=1}^n K\left(\frac{x_i - x_j}{b}\right) \cdot y_j}{\sum_{j=1}^n K\left(\frac{x_i - x_j}{b}\right)} \quad (2)$$

که در رابطه بالا  $x_i$  شماره داده  $i$  ام،  $y_i$  داده واقعی  $i$  ام و  $b$  پارامتر تعیین کننده میزان نرم شدن داده، است.  $K(t)$  نیز «تابع کرنل» نامیده می شود و یک مدل معروف آن کرنل گوس است و به صورت زیر تعریف می شود:

$$K(t) = \frac{1}{\sqrt{2\pi} \cdot (\sigma)} \cdot e^{-\left(\frac{t^2}{2\sigma^2}\right)} \quad (3)$$

که در آن  $\sigma$  انحراف معیار و یا به اصطلاح «عرض کرنل» نام دارد. با انتخاب مقداری بین ۰/۳ تا ۰/۵ برای عرض کرنل، نتیجه مطلوبی در حذف اغتشاش های<sup>۱۸</sup> چند پیکسلی از داده های ورودی با فرکانس حدودی ۲۰ هرتز در کاربرد مورد نظر، مشاهده شده است.

این رابطه برای داده تخمین زده شده در فریم کنونی به کمک  $n$  داده قبلی بصورت زیر درمی آید:

$$y'_i = \frac{\sum_{j=1}^n K\left(\frac{n-j}{b}\right) \cdot y_j}{\sum_{j=1}^n K\left(\frac{n-j}{b}\right)} \quad (4)$$

$y_j$  داده  $j$  ام و  $y_n$  در واقع داده کنونی بوده که  $y'_i$  مقدار تخمین زده آن است.

از این صافی نرم کننده در دو مورد استفاده شده است: یکی برای نرم کردن موقعیت مرکز مردمک و گلینت و دیگری برای نرم کردن خروجی نگاشت یا همان مختصات نشانگر در صفحه.

## ۶- نرم افزار

نرم افزار طراحی شده برای این سیستم ردیاب با زبان ویژوال

<sup>18</sup> Noise



حرکات پرشی چشم و ترتیب آنها را هنگام تماشای یک عکس و یا یک فایل ویدئویی با قالب‌های مختلف به نمایش درآورد. این بخش کاربردهای بسیار فراوانی دارد و بویژه در زمینه‌های تبلیغات، بازاریابی و تحقیقات مربوط به روانشناسی به کار برده می‌شود. در شکل (۱۰) یک عکس و مسیر حرکت چشم و دید شخص را مشاهده می‌کنیم که نرم‌افزار آن را در قالب ScanPath ترسیم کرده است..



شکل ۱۰- نمایش مسیر حرکت چشم و نقاط مکث بر روی یک تصویر به صورت ScanPath

## ۷- ارزیابی دو روش استفاده شده برای ردیابی مردمک

در این بخش از مقاله به ارزیابی دو الگوریتم استفاده شده برای ردیابی مردمک، از لحاظ دقت و سرعت می‌پردازیم. برای نشان دادن دقت عملکرد دستگاه، پس از مرحله کالیبراسیون، از ۵ شخص سالم خواسته می‌شود که به ۲۰ نقطه با مکان‌های مشخصی در صفحه، برای مدت زمان کوتاهی خیره شوند. در مدت زمان خیره شدن افراد به هر نقطه، مکان نشانگر در صفحه نمایش- همان نقطه دیدی که سیستم آن را اندازه‌گیری کرده است- در یک فایل متنی ذخیره می‌شود. در حالت ایده‌آل، زمانی که شخص به یک نقطه خیره می‌شود، نشانگر در صفحه بر روی همان نقطه قرار می‌گیرد و ثابت می‌ماند؛ اما عملاً بدلیل خطای سیستم در اندازه‌گیری نقطه دید (مکان نشانگر)، ممکن است نقطه اندازه‌گیری شده در هر لحظه در فاصله‌ای جزئی از نقطه دید واقعی قرار گیرد. در این حالت، فاصله نشانگر تا نقطه دید

بازی‌های قبل و بعد برود و یا از این بخش خارج شود.



شکل ۸- صفحه مربوط به تایپ فارسی نرم‌افزار

بخش دیگری که در نرم‌افزار گنجانده شده، یک محیط چندرسانه‌ای است که امکان استفاده از فایل‌های صوتی و تصویری را برای کاربر فراهم می‌کند. در این محیط کاربر می‌تواند آلبوم‌های تصویری، صوتی و فایل‌های موجود در آنها را انتخاب کند (شکل ۹).



شکل ۹- محیط چندرسانه‌ای نرم‌افزار

در بخش ترسیم نمودار این نرم‌افزار نیز، سه بخش گنجانده شده است. در بخش اول می‌توان نمودار داده‌هایی چون قطر و مختصات مردمک را در طول زمان رسم کرده؛ تغییرات آن را مشاهده کرد. در بخش دوم می‌توان داده‌های مربوط به مختصات نقاط دید شخص و مختصات نشانگر در صفحه را بصورت یک فایل استاندارد در بازه‌ای از زمان ذخیره کرد. این فایل به شکل یک فایل متنی است و با قالب استاندارد که نرم‌افزارهای تحلیل داده‌های ردیابی چشم از آن استفاده می‌کنند؛ استخراج می‌شود. در بخش سوم می‌توان داده‌های مذکور یعنی مسیر حرکت چشم، نقاط مکث،

است. در اینجا، می‌خواهیم بدانیم عدم قطعیت در اندازه‌گیری مرکز مردمک، چه میزان خطایی در تعیین نقطه دید ایجاد می‌کند. بدین منظور یک تحلیل عدم قطعیت عمومی بر نگاشت مورد نظر انجام می‌دهیم.

به طور کلی میزان عدم قطعیت در نتیجه تابع  $f$  با  $n$  متغیر  $X_i$  یعنی  $f = f(X_1, X_2, \dots, X_n)$  از ضابطه زیر بدست می‌آید:

$$U_f^2 = \left(\frac{\partial f}{\partial X_1}\right)^2 U_{X_1}^2 + \left(\frac{\partial f}{\partial X_2}\right)^2 U_{X_2}^2 + \dots + \left(\frac{\partial f}{\partial X_n}\right)^2 U_{X_n}^2 \quad (5)$$

که برای مثال در آن  $U_{X_i}$  میزان عدم قطعیت در اندازه‌گیری متغیر  $X_i$  است.

نگاشت مورد نظر ضابطه (۵) دو جمله‌ای است و باید عدم قطعیت کل را جداگانه برای هر جمله، یعنی برای  $S_x$  و  $S_y$  محاسبه می‌کنیم. هر ضابطه تابعی از دو متغیر  $x$  و  $y$  است؛ اما ۶ پارامتر مجهول نیز در آنها وجود دارد. با انجام کالیبراسیون و با فرض آنکه در این مرحله اندازه‌گیری متغیرهای  $x$  و  $y$  قطعی و بدون خطا است، ۶ پارامتر مذکور را پیدا می‌کنیم. پس از آن با یافتن رابطه عدم قطعیت برای نگاشت خود، تأثیر عدم قطعیت به اندازه ۱ پیکسل در تشخیص مرکز مردمک را بر نقطه دید اندازه‌گیری شده، محاسبه می‌کنیم.

ضابطه عدم قطعیت برای نگاشت دو جمله‌ای رابطه (۱) بصورت زیر خواهد بود:

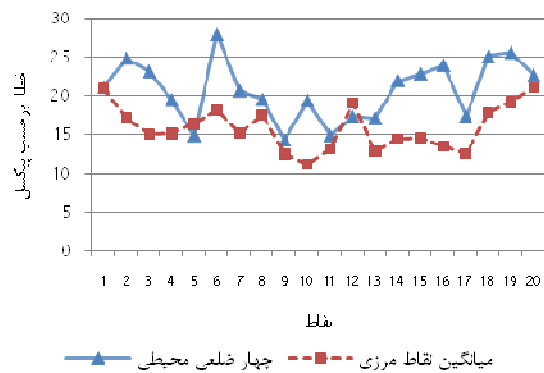
$$\begin{cases} \frac{U_{S_x}^2}{S_x^2} = \left(\frac{x}{S_x}(a_1 + a_3y + 2a_4x)\right)^2 \left(\frac{U_x}{S_x}\right)^2 + \left(\frac{y}{S_x}(a_2 + a_3y + 2a_4x)\right)^2 \left(\frac{U_y}{S_y}\right)^2 \\ \frac{U_{S_y}^2}{S_y^2} = \left(\frac{x}{S_y}(b_1 + b_3y + 2b_4x)\right)^2 \left(\frac{U_x}{S_x}\right)^2 + \left(\frac{y}{S_y}(b_2 + b_3y + 2b_4x)\right)^2 \left(\frac{U_y}{S_y}\right)^2 \end{cases} \quad (6)$$

که در آن  $U_x$  و  $U_y$  میزان عدم قطعیت در اندازه‌گیری متغیرهای  $x$  و  $y$  است و  $U_{S_x}$  و  $U_{S_y}$  میزان عدم قطعیت در مکان نقطه دید یعنی  $S_x$  و  $S_y$  است.

با فرض اینکه در اندازه‌گیری مرکز مردمک به اندازه ۱ پیکسل خطا داشته باشیم، میزان عدم قطعیت در اندازه‌گیری نقطه دید در ۹ نقطه از صفحه محاسبه شد. نتیجه آن، ۱۵ الی ۳۰ پیکسل عدم قطعیت در اندازه‌گیری نقطه دید در صفحه نمایشی با رزولوشن  $1024 \times 768$  بود.

واقعی بر حسب پیکسل اندازه‌گیری شده؛ با نام «خطا» معرفی می‌شود. این آزمایش را ۵ کاربر برای هر کدام از دو روش تکرار کرده‌اند. در نهایت برای هر یک از نقاط میانگین خطا در ۵ بار آزمایش بدست می‌آید و با این معیار دو روش را با یکدیگر مقایسه می‌کنیم (شکل ۱۱).

در طول این آزمایش، فاصله سر اشخاص تا صفحه نمایشی با قدرت تفکیک  $1024 \times 768$ ، تقریباً ۷۰ سانتی‌متر است. همچنین از ۱۶ نقطه برای کالیبراسیون استفاده شده است. شایان ذکر است که آزمایش بر روی یک سیستم با پردازنده اینتل دو هسته‌ای ۳ گیگاهرتز با رم ۳ گیگا بایت انجام شده است.



شکل ۱۱- مقایسه میزان خطا در دو روش ردیابی مردمک همانطور که گفته شد، روش تکمیلی میانگین نقاط مرزی تا حدودی میزان تغییرات کوچک مرکز مردمک در فریم‌های متوالی را کاهش داد؛ در نتیجه سبب کاهش جابجایی نقطه دید اندازه‌گیری شده نیز می‌شود. همین امر خطای سیستم را کاهش می‌دهد.

زمان کل عملیات برای یک فریم در روش چهارضلعی محیطی ۱۳ میلی‌ثانیه بوده؛ که در روش تکمیلی میانگین نقاط مرزی به ۲۵ میلی‌ثانیه می‌رسد.

## ۸- تحلیل حساسیت نگاشت

اندازه‌گیری مؤلفه‌های افقی و عمودی مرکز مردمک،  $(X, Y)$ ، در تصویر برداشته شده از چشم، ممکن است در حدود ۱-۲ پیکسل با خطا همراه باشد. این میزان خطا بدلیل نوع روش استفاده شده برای تشخیص مرکز مردمک اجتناب‌ناپذیر

## ۹- ارزیابی عملکرد سیستم بر معلولان

در این آزمایش از ۹ فرد معلول استفاده کردیم. این افراد ۱۵ تا ۳۵ سال داشته و همگی با رایانه و مفاهیم اولیه آن آشنا بودند. از افراد خواسته شد که سه فعالیت بخصوص را با کمک سیستم مذکور انجام دهند. این فعالیت‌ها در سه بخش تایپ در محیط نرم افزار، بازی و کار با محیط چندرسانه‌ای نرم‌افزار اجرا شدند. پس از آن عملکرد سیستم با بررسی پاسخ افراد به سه پرسش‌نامه، ارزیابی شد. یکی از پرسش‌نامه‌ها بر اساس پرسش‌نامه<sup>۱۹</sup> (TLX) [۱۷] تنظیم شده بود و برای ارزیابی میزان پیچیدگی محیط نرم‌افزار به کار برده شد. دو پرسش‌نامه دیگر که بر پایه Q.U.I.S Interface Evaluation [۱۸] و IBM Psychometric Evaluation [۱۹] تنظیم شده بودند نیز میزان رضایت کاربر از سیستم و میزان کارایی آن برای فرد معلول از دیدگاه او، را ارزیابی کردند.

نتایج بررسی‌ها، کارا بودن این سیستم را در یاری رساندن به معلولان در انجام فعالیت‌هایی همچون تایپ و صحبت از طریق کامپیوتر، بازی و سرگرمی و به طور کلی برقراری ارتباط با کامپیوتر نشان می‌دهد، به طوری که در پایان ارزیابی، بیشتر افراد از ورود چنین سیستمی به مجتمع توانبخشی خود، اعلام رضایت کردند.

برای نمونه، در ادامه ۱۱ سؤال از پرسش‌نامه‌ها و پاسخ‌های افراد، همراه نمودار آورده شده است.

۱. در مجموع من از سادگی کار با دستگاه ساخته شده، راضی هستم.

۲. کار با دستگاه بسیار راحت است.

۳. من فعالیت‌های مورد نظر را می‌توانم بصورت مؤثر با این سیستم انجام دهم.

۴. من می‌توانم با این سیستم به سرعت کارم را به اتمام برسانم.

۵. من در هنگام استفاده از دستگاه احساس راحتی می‌کنم.

۶. یادگیری نحوه کار کردن با سیستم راحت است.

۷. در صورت بروز اشتباه به سرعت می‌توانم آن را برطرف

سازم.

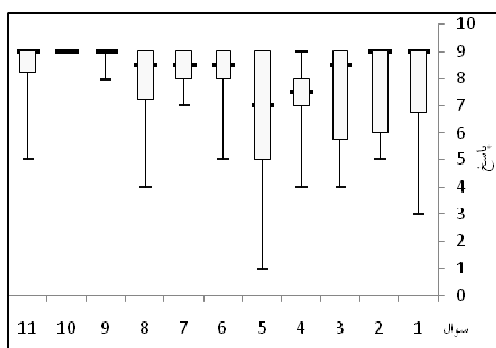
۸. چینه اطلاعات در صفحه بسیار واضح است.

۹. محیط نرم‌افزار خوشایند و جذاب است.

۱۰. من از کار کردن با این دستگاه لذت می‌برم.

۱۱. در مجموع من از سادگی کار با دستگاه ساخته شده راضی هستم.

پاسخ این سؤالات با گزینه‌های: به شدت موافق (اختصاص مقدار ۹) و به شدت مخالف (اختصاص مقدار صفر) ارزیابی شد. شایان ذکر است سؤال ۱ به عنوان یک سؤال کلی مجدداً در پایان پرسش‌نامه مطرح شده است.



شکل ۱۲- پاسخ کلی افراد به پرسش‌نامه سوم درباره ردیاب ساخته شده

همانطور که در شکل ۱۲ مشاهده می‌شود شاخص میانه در پاسخ افراد، بالاتر از ۵ است. این موضوع مثبت بودن نظر افراد درباره ردیاب مذکور را بیان می‌کند.

## ۱۰- نتیجه‌گیری

در این مقاله به طراحی و ساخت سخت‌افزار و نرم‌افزار یک سیستم ردیاب حرکات چشم که مبتنی بر روش تصویری و قابل نصب بر روی سر است؛ اشاره شد. از ویژگی‌های مهم سخت‌افزاری این سیستم می‌توان به سبکی، بی‌سیم بودن، ارزان بودن آن و راحتی استفاده از آن اشاره کرد. نرم‌افزار سیستم قابلیت‌های کاربردی متنوعی همچون تایپ با چشم و تکلم جملات، دارد و همچنین دارای محیط ویژه بازی و سرگرمی و یک محیط چندرسانه‌ای به منظور استفاده از آلبوم‌های صوتی و تصویری با حرکات چشم است. دقت عملکرد سیستم ردیاب ساخته شده نیز در حدود ۱ درجه

<sup>19</sup> NASA Task Load Index

[۱۰] مردانگی دیاکو، ملاک زاده محمدرضا؛ طراحی و ساخت یک ردیاب چشم به منظور توانبخشی معلولان در برقراری ارتباط با کامپیوتر؛ پانزدهمین کنفرانس مهندسی پزشکی ایران؛ مشهد؛ ۱۳۸۷.

[۱۱] مردانگی دیاکو، ملاک زاده محمدرضا؛ ارزیابی عملکرد یک ردیاب حرکات چشم در زمینه توانبخشی معلولان کوادریلیژیک به منظور برقراری ارتباط با رایانه؛ شانزدهمین کنفرانس مهندسی پزشکی ایران؛ تهران؛ ۱۳۸۸.

[12] Duchowski A.T., Eye Tracking Methodology: Theory and Practice; Springer, London; in press.

[13] Nouredin B., Lawrence P.D., Man C.F., A non-contact device for tracking gaze in a human computer interface.; In Computer Vision and Image Understanding; 2005; 98: 52-82.

[14] Beymer D., Flickner M., Eye gaze tracking using an active stereo head; IEEE Conf. Comput. Vis. Pattern Recogn; 2003; 2: 451-458.

[15] Li D., Win\_eld D., Parkhurst D.J., Starburst: A hybrid algorithm for video-based eye tracking combining feature-based and model-based approaches; In Proceedings of the IEEE Vision for Human-Computer Interaction Workshop at CVPR; 2005; pp. 1.8.

[16] Morimoto C.H., Marcio R.M., Mimica, Eye gaze tracking techniques for interactive applications; Computer Vision and Image Understanding; 2005; 98: 4-24.

[17] NASA, Task Load Index (TLX): Computerized version (Version 1.0). Moffett Field, CA: Human Research Performance Group; NASA Ames Research Center; 1986.

[18] Chin J. P., Diehl V. A, Norman K., Development of an instrument measuring user satisfaction of the human-computer interface; Proc. ACM CHI; 1987; 88: 213-218.

[19] Lewis J. R., IBM Computer Usability Satisfaction Questionnaires: Psychometric Evaluation and Instructions for Use; International Journal of Human-Computer Interaction; 1995; 7:1, 57-78.

میدان دید بوده؛ این میزان با دقت اکثر ردیاب‌های تجاری موجود برابر است.

موفقیت روش طراحی و ساخت و همچنین درستی عملکرد نرم‌افزار، زمینه را برای تبدیل سیستم به یک ردیاب قابل نصب بر روی میز فراهم کرده، به گونه‌ای که دوربین به جای نصب بر سر شخص، بر روی میز و مقابل صفحه نمایش نصب می‌شود. لازمه این کار استفاده از تجهیزات، دوربین‌های گران‌تر و الگوریتم‌های جانبی پردازش تصویر همچون ردیاب صورت و چشم است.

## قدردانی

در پایان از آقای رضایی، مدیر محترم مجتمع توانبخشی ملائک صمیمانه تشکر می‌کنیم که امکانات مرکز را بی‌دریغ برای انجام مرحله ارزیابی این کار پژوهشی در اختیار ما قرار دادند.

## مراجع

[۱] پرنیان پور محمد، مردانگی دیاکو، سرشار محمد؛ ردیابی حرکات چشم (Eye Tracking) و کاربردهای آن در ارگونومی؛ اولین کنفرانس بین‌المللی ارگونومی ایران؛ تهران؛ ۱۳۸۷.

[2] Eye tracking; <http://www.wikipedia.com/2008>.

[3] Kumar M., Paepcke A., Winograd T., EyePoint: Practical Pointing and Selection Using Gaze and Keyboard; In Proceedings of CHI. San Jose, California, USA: ACM Press; 2007.

[4] Skovsgaard H., Noise Tolerant Selection by Gaze Controlled Pan and Zoom; 2008.

[5] Ward D.J., Blackwell A.F., MacKay D.J.C., Dasher: a gesture driven data entry interface for mobil computing; Hum Comput. Interact; 2002; 17: 199-228.

[6] Tall M., EyeTube, Gaze Group; IT University of Copenhagen; 2008.

[7] Babcock J.S., Pelz J.B., Building a lightweight eyetracking headgear; symposium on Eye tracking research & applications, San Antonio, Texas, ACM; 2004; 109-114.

[8] Li D., Babcock J.S., Parkhurst D.J., openEyes: a low-cost head-mounted eye-tracking solution; symposium on Eye tracking research & applications, San Diego, California, ACM; 2006; 95-100.

[9] Zielinski P., Opengazer: open-source gaze tracker for ordinary webcams; <http://www.inference.phy.cam.ac.uk/opengazer/>.