

تخمین زاویه‌ی صورت در تصاویر سه‌بعدی چهره

علی‌رضا بساق‌زاده

علی‌رضا بهراد

محمدعلی دوستاری

تهران، ابتدای بزرگراه خلیج فارس، دانشگاه شاهد، دانشکده فنی و مهندسی

bossaghzadeh@shahed.ac.ir

Behrad@shahed.ac.ir

doostari@shahed.ac.ir

چکیده: کانتورهای IRAD به عنوان ویژگی‌های مستقل از زاویه برای تخمین زاویه‌ی چهره در تصاویر سه‌بعدی مطرح شده‌اند. با استفاده از این کانتورها و یک روش دقیق در تخمین موقعیت بینی به سیستمی برای تخمین زاویه‌ی چهره در تصاویر سه بعدی دست یافته‌ایم. این سیستم نیاز به هیچ نوع داده‌ی آموزشی ندارد و بدلیل استفاده از داده‌های عمق چهره نسبت به تغییر شدت روشهایی محیط مقاوم می‌باشد. نتایج پیاده‌سازی این روش بر روی پایگاه داده‌ی FRAV3D نشان می‌دهد که روش پیشنهادی در تخمین زاویه دارای دقتی بین درصد ۹۷٪-۹۹٪ با خطای ۹ درجه در سه محور مختصات می‌باشد.

واژه های کلیدی: کانتورهای هم‌شعاع، تخمین بیضی، تصاویر سه‌بعدی

و همکاران از ویولت و SVR¹ استفاده کرده است، روش آنها برای آموزش نیاز به تصاویری دارد که زوایای آنها از قبل معین شده باشند [7]. در [8] نیز در فاز ابتدایی یک سیستم با تصاویر سه‌بعدی در زوایای مرجع آموزش داده می‌شود. برای تصویر ورودی به منظور یک تخمین تقریبی ابتدا با این تصاویر مقایسه شده و تصویر با کمترین خطا انتخاب می‌شود. در فاز بعد زاویه‌ی دقیق تصویر با یک تابع خطا بدست می‌آید. ادامه‌ی این مقاله به این صورت است. در زیربخش ۲.۱ کلیات الگوریتم پیشنهادی توضیح داده شده است در زیر بخش ۲.۲ توضیح کوتاهی در مورد پایگاه داده‌ی مورد استفاده آمده است. در زیر بخش ۲.۳ درباره‌ی کانتورهای IRAD توضیح داده شده است. قسمتهای ۲.۴ ۲.۵ ۲.۶ ۲.۷ به توضیح نحوه‌ی محاسبه‌ی زوایا پرداخته است. در بخش ۳ نتایج تجربی بدست آمده از پایگاه داده‌ی FRAV3D آمده و در بخش ۴ به جمع‌بندی نتایج بدست آمده پرداخته شده است.

۱ مقدمه

تخمین زاویه چهره به عنوان یک مرحله‌ی میان پردازشی در خلال اجرای برنامه‌های دیگر مانند درک میزان توجه راننده، تعامل بین انسان و کامپیوتر، شناسایی چهره تحت زوایای مختلف، کاربرد دارد. بسیاری از این کاربردها نیاز به روشی با سرعت بالا دارند که به طور همزمان دقت مناسبی نیز در زوایای مختلف داشته باشد. تخمین زاویه‌ی چهره در کاربردهایی مانند شناسایی چهره امکان شناسایی بهتر و دقیقتری را برای ما فراهم می‌سازد.

روشهای مبتنی بر تصاویر دوبعدی بر روی تمام تصویر پردازش کرده و نیاز به تعیین مکان چهره دارند [2]. برخی از این روشها برای زوایای مختلف از کلاس‌بندهای متفاوت استفاده می‌کنند [3] که برای این کار نیاز به تعداد زیادی داده‌ی آموزشی دارند [4]. این روشها را می‌توان با استفاده از تنظیم یک مدل عمومی بر تصویر ورودی بهبود بخشید [6][5]. از آنجایی که اغلب این سیستمها خود قسمتی از سیستم بزرگتر مثل شناسایی چهره هستند، بطور جداگانه مورد بررسی قرار نگرفته‌اند. در کل برخی از این سیستمها نیاز به تعیین مکان اولیه یا داده‌های آموزشی دارند و از تغییرات شدت و زاویه نور رنج می‌برند. اخیرا با کاهش قیمت و افزایش کیفیت جاروبگرهای سه بعدی استفاده از این داده‌های عمق چهره برای تعیین زاویه گسترده‌تر شده است. Ajit

۲ الگوریتم پیشنهادی

۲.۱ کلیات الگوریتم

روش پیشنهادی ما یک روش کاملا اتوماتیک بوده و در آن نیاز به مقدار دهی اولیه نمی‌باشد، همچنین نسبت به چرخش زیاد تصویر در زوایای مختلف مقاوم بوده و نیاز به داده‌های آموزشی ندارد. مراحل انجام کار بدین صورت است:

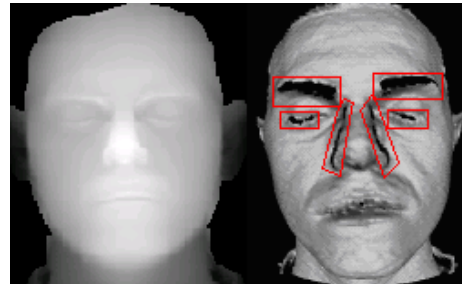
- ۱- بینی با استفاده از یک روش ابتکاری در تصویر پیدا می‌شود
- ۲- کانتورهای هم‌فاصله در تصویر استخراج می‌شوند
- ۳- بر روی یکی از این کانتورها یک بیضی برازش می‌شود
- ۴- با استفاده از معادله‌ی بیضی زاویه‌ی چهره حول محور Zها بدست می‌آید
- ۵- با چرخش نقاط یکی از کانتورها حول محور Z و اندازه‌گیری اختلاف نقاط راست و چپ تصویر زاویه حول محور Y بدست می‌آید
- ۶- با استفاده از دو ناحیه در قسمت‌های بالا و پایین بینی و هم‌سطح کردن آنها، زاویه سر حول محور X بدست می‌آید.

¹ Support Vector Regression

۲.۲ پایگاه داده و استخراج تصاویر

پایگاه داده مورد استفاده FRAV3D می باشد. این پایگاه داده حاوی تصاویر ۱۰۶ نفر می باشد، که تقریباً به ازای هر سه مرد تصویر یک زن موجود است [10]. تصاویر موجود در پایگاه داده با استفاده از دوربین Minolta VI-700 [11] تصویربرداری شده، که اطلاعات خروجی بدست آمده از این دوربین شامل اطلاعات بافت چهره است که در یک تصویر دوبعدی و اطلاعات عمق چهره است که در یک فایل VRML ذخیره شده اند. به ازای هر فرد ۱۶ تصویر گرفته شده است که سعی شده تمامی حالات را از قبیل چرخش در تمامی جهات، حالت چهره^۲ و تغییرات شدت روشنایی را شامل شود. در هر تصویر فقط یکی از پارامترهای اشاره شده تغییر داده شده است.

در تصویر برداری سه بعدی به دلیل وجود نویزهای تصویر برداری دو مشکل نویزهای پرت^۳ و حفره بوجود می آید. تصاویر دارای نویز پرت دارای جهش^۴ هستند و در تصویر دارای حفره، به ازای نقاطی در تصویر هیچ مقداری موجود نمی باشد. پس از استخراج داده ها به صورت نقشه عمقی^۵ برای حذف نویزهای پرت از یک فیلتر میانه و برای پر کردن حفره ها از روش درون یابی مکعبی استفاده شده است. در شکل 1 تصویر دارای حفره و معادل درون یابی شده آن آورده شده است.

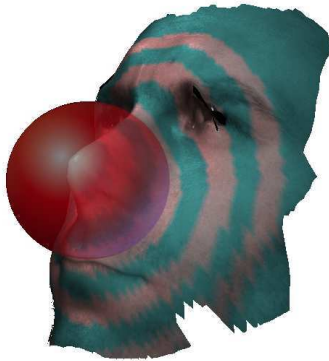


شکل 1: تصویر دارای حفره (سمت راست) و معادل درون یابی شده آن (سمت چپ)

۲.۳ کانتورهای IRAD و استخراج آنها

ایدهی استفاده از کانتورهای IRAD^۶ در شناسایی چهره توسط Hesseltnie و همکاران ارائه شده است [9]. کانتورهای IRAD مکانهایی در یک رویه سه بعدی هستند که از یک نقطه‌ی مورد نظر در همان سطح (در شناسایی چهره بینی) فاصله‌ی ثابتی دارند. این نقاط را می توان به عنوان تقاطع یک کره به شعاع r و به مرکزیت نقطه‌ی مورد نظر p با سطح در نظر گرفت. اگر رویه مورد نظر مسطح^۷ باشد این کانتورها به صورت دایره‌ی به مرکزیت نقطه مورد نظر خواهند بود، اما

برای سطوح غیرمسطح^۸ این کانتورها شکل‌های دیگری پیدا خواهند کرد [9]. در شکل 2 نمونه‌ای از این کانتورها که با استفاده از تقاطع یک کره با صورت بدست آمده، آورده شده است.

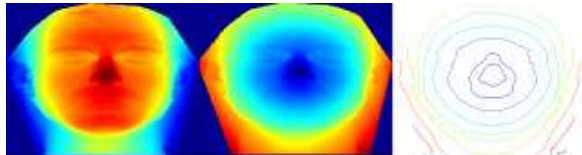


شکل 2: تولید کانتور با استفاده از کره [9]

برای محاسبه‌ی این کانتورها -در تصاویر چهره- نقطه‌ی مرکزی، بینی در نظر گرفته شده است. دلیل استفاده از بینی به عنوان نقطه‌ی مرکز این است که بینی به دلیل داشتن ویژگی‌های منحصر به فرد، امکان تعیین مکان آن در تصاویر سه بعدی با دقت و سرعت بالایی امکان پذیر است [۱۱]. پس از تعیین نقطه‌ی مرکزی کره، فاصله‌ی شعاعی را با استفاده از معادله

$$I = (X - X_{nose})^2 + (Y - Y_{nose})^2 + (Z - Z_{nose})^2 \quad (1)$$

بدست می آوریم که در آن X موقعیت ستونی نقطه، Y موقعیت سطری نقطه و Z ارتفاع نقطه مورد نظر می باشد. تصویری که از اجرای این معادله به دست می آید در شکل 3 آورده شده است.



شکل 3: تصویر اصلی (سمت چپ)، تصویر فاصله تا بینی (وسط)،

کانتورهای هم شعاع

پس از محاسبه‌ی این پارامتر، کانتوری که دارای مقادیر معین یا به عبارتی فاصله‌ی معینی از بینی است را استخراج می کنیم. انتخاب مقدار این کانتور در عملکرد سیستم تاثیر زیادی دارد. اگر فاصله زیاد در نظر گرفته شود، در تصاویری که دارای چرخش زیاد در جهت محور Y هستند برخی از نقاط این کانتور در بیرون از چهره و تصویر قرار می گیرند که باعث می شود نقاط کافی برای محاسبه‌ی زاویه را در اختیار نداشته باشیم. همچنین در صورت کوچک گرفتن این فاصله بدلیل وجود پنهان سازی^۹ امکان استفاده‌ی از برخی نقاط را نداریم. در شکل 4 تصاویری که در آنها شعاع کره بزرگ و کوچک می باشد، آورده

² gesture

³ Outlier

⁴ Spike

⁵ Depth Map

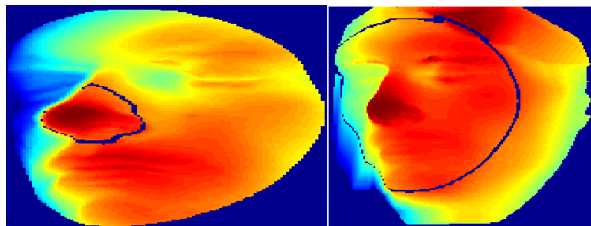
⁶ IsoRADious

⁷ Planar

⁸ non-planar

⁹ occlusion

شده است. برای تعیین مقدار مناسبی برای شعاع کانتور مورد نظر، اندازه‌ی شعاع در فاصله‌ی تقریبی ۲.۵ سانتی‌متری در دو طرف بینی را بدست آورده و بزرگترین مقدار به عنوان شعاع کانتور در نظر گرفته شده است. با انتخاب این مقدار تضمین می‌شود که نقاط کانتور مورد نظر در تصویر وجود دارند.



شکل 4: دو تصویر که در آنها فاصله کم (سمت چپ) و فاصله زیاد (سمت راست) باعث شده یک بیضی کامل استخراج نشود.

۲.۴ برازش کردن بیضی

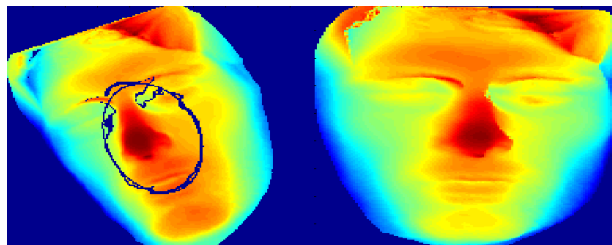
پس از استخراج نقاط برای محاسبه‌ی زاویه‌ی چهره در جهت محور Zها از یک بیضی استفاده می‌کنیم. برای این کار ابتدا نقاط را به صفحه‌ی $Z=0$ نگاشت کرده، سپس یک بیضی بر روی نقاط نگاشت یافته برازش می‌کنیم. برای محاسبه‌ی یک بیضی به ۵ پارامتر نیاز داریم که شامل قطر بزرگ و کوچک بیضی، مرکز بیضی بر روی محور X و Y و زاویه‌ی بیضی است. از این ۵ پارامتر، هدف بدست آوردن زاویه‌ی بیضی است. معادله‌ی بیضی در صفحه از رابطه‌ی

$$aX^2 + bXY + cY^2 + dX + eY = 1 \quad (2)$$

بدست می‌آید. زاویه بیضی که نقاط آن در معادله بالا صدق کنند، از رابطه‌ی $teta = \tan^{-1}\left(\frac{b}{c-a}\right)$ بدست می‌آید.

۲.۵ زاویه در جهت محور Zها

چرخش چهره حول محور Z همان زاویه‌ی بیضی بدست آمده در مرحله ۲.۴ است. در شکل ۵ یک تصویر با نقاط استخراج شده و بیضی برازش شده را مشاهده می‌کنید. با محاسبه‌ی زاویه‌ی چهره در جهت محور Zها، تصویر را به میزان منفی این زاویه چرخش داده تا محور بینی در جهت موازی با محور Yها قرار گیرد. در شکل ۵ یک تصویر قبل و بعد از چرخش را مشاهده می‌کنید.



شکل ۵: تصویر با چرخش حول محور Z با نقاط هم‌فاصله و بیضی برازش شده بر آن (سمت چپ)، تصویر چرخش یافته (سمت راست)

۲.۶ زاویه چرخش حول محور Yها

برای محاسبه‌ی زاویه در جهت محور Yها از این فرض استفاده می‌شود که چهره انسان تقریباً در جهت محور Yها دارای تقارن است. این تقارن به ما کمک می‌کند که پس از بدست آوردن نقاط هم‌فاصله با داشتن اختلاف ارتفاع نقاط زاویه را در جهت محور Y برای هم‌سطح کردن نقاط محاسبه کنیم. برای بدست آوردن این فاصله یکسان می‌توان فاصله‌ی پیکسلی معینی را در طرفین چهره بدست آورده و مقدار اختلاف ارتفاع را در آنها محاسبه کرد. در این صورت پس از چرخش چهره بدلیل اینکه نقاط تغییر مکان پیدا می‌کنند باید این الگوریتم را مجدد تکرار کرد تا خط کاهش یابد. در اینجا برای تعیین نقاط هم‌فاصله با استفاده از کانتورهای IRAD از فاصله‌ی شعاعی استفاده شده است که نسبت به چرخش مقاوم بوده و تغییر پیدا نمی‌کند. لذا فاصله‌ی بدست آمده پس از چرخش چهره ثابت مانده و نیازی به محاسبه‌ی مجدد نمی‌باشد. بدین منظور پس از چرخش نقاط حول محور Zها و موازی شدن قطر بزرگ بیضی با محور Yها قسمت راست و چپ بیضی را جدا می‌کنیم. برای داشتن دقت مناسب در محاسبه‌ی زاویه از تعدادی نقاط در سمت راست و چپ بینی استفاده شده است. بدین صورت که میانگین مقادیر Z در قسمتی از نیمه راست و چپ بیضی را محاسبه کرده سپس اختلاف مقادیر ارتفاع را بر فاصله نقاط از هم تقسیم می‌کنیم. رابطه ریاضی آن به صورت

$$teta_y = \tan^{-1}\left(\frac{Z_2 - Z_1}{X_2 - X_1}\right) \quad (3)$$

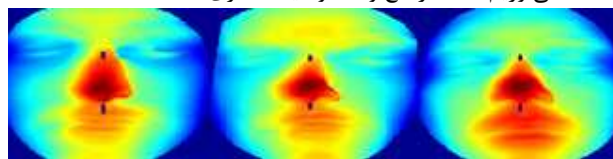
می‌باشد که Z اندازه ارتفاع و X موقعیت ستونی مقدار مورد استفاده است.

۲.۷ زاویه در جهت محور Xها

برای تخمین زاویه در جهت محور Xها از این فرض استفاده شده است که ناحیه بین بینی و لب و ناحیه‌ی زین اسبی بالای بینی تقریباً در یک مقدار ارتفاع قرار گیرند. تصویر این دو ناحیه در شکل 6 آمده است. برای بدست آوردن این دو ناحیه می‌توان همانند روش ذکر شده در مرحله قبل به صورت تکرار دو نقطه را در بالا و پایین بینی با فاصله‌ی پیکسلی در نظر گرفت و زاویه را با استفاده از آنها محاسبه کرد. برای جلوگیری از تکرار الگوریتم از فاصله‌ی شعاعی با بدست آمده از کانتورهای IRAD استفاده می‌کنیم. بدین صورت که در ناحیه‌ی بالا و پایین بینی به دنبال دو ناحیه می‌گردیم که اولاً فاصله شعاعی آنها از بینی حد معینی باشد، ثانیاً موقعیت ستونی این نقاط در نزدیکی بینی قرار داشته باشد. چند نمونه از این تصاویر که نقاط بالا و پایین بینی در آنها مشخص شده اند را در شکل 6 مشاهده می‌کنید. پس از بدست آوردن این نقاط، با استفاده از اختلاف ارتفاع و فاصله‌ی آنها، زاویه را از رابطه‌ی

$$x_teta = \tan^{-1}\left(\frac{Z_2 - Z_1}{Y_2 - Y_1}\right) \quad (4)$$

بدست می‌آوریم که Z ارتفاع و Y موقعیت سطری نقطه است.



شکل 6: نقاط انتخاب شده در قسمت پایین و بالای بینی

داده‌ای برای آموزش ندارد. بدلیل ساختار سلسه مراتبی در صورت تخمین اشتباه زاویه در یک مرحله امکان تخمین اشتباه زوایای دیگر نیز وجود دارد. بدلیل نگاشت داده‌ها به صفحه‌ی $Z=0$ امکان بروز خطا در محاسبه‌ی زاویه در جهت محور Z وجود دارد و در صورتی که بتوان با استفاده از بیضی در فضای سه‌بعدی پارامترها را محاسبه کرد می‌توان دقت روش را افزایش داد. همچنین در صورتی که زوایای در راستای محور Y و X را بتوان از روی تصویر اولیه و همزمان با تعیین زاویه در راستای محور Z بدست آورد، سرعت و خطای روش کاهش خواهد یافت. همچنین استفاده از نقطه‌ی زین اسبی بالای بینی بدلیل داشتن انحنا و تغییر اندازه زیاد در اطراف خود در صورت تخمین اشتباه منجر به خطای زیادی در سیستم می‌شود و بهتر است از نقاط دیگر مانند پیشانی برای این منظور استفاده شود.

سیاسگزاری

این پژوهش از حمایت مرکز تحقیقات مخابرات ایران برخوردار بوده است که بدین وسیله از این مرکز تشکر و قدردانی به عمل می‌آید

مراجع

[1] بساق‌زاده، دوستاری، بهراد، "استخراج خودکار ویژگی‌های چهره، برای شناسایی هویت مستقل از زاویه در تصاویر سه بعدی"، MVIP2008، تبریز

[2] L. M. Brown and Y.-L. Tran. *Comparative study of coarse head pose estimation. In MOTION*, 2002.

[3] M. Jones and P. Viola. *Fast multi-view face detection*. Technical Report TR2003-096, Mitsubishi Electric Research Laboratories, 2003.

[4] M. Osadchy, M. L. Miller, and Y. LeCun. *Synergistic face detection and pose estimation with energy-based models*. NIPS, pages 1017–1024, 2005

[5] V. Blanz and T. Vetter. *Face recognition based on fitting a 3d morphable model*. PAMI, 25(9):1063–1074, 2003.

[6] L. Gu and T. Kanade. *3d alignment of face in a single image*. CVPR, 2006

[7] Ajit Rajwade, Martin D. Levine. *Facial pose from 3D data*, Image and Vision Computing 2006

[8] Michael D. Breitenstein, *Real-Time Face Pose Estimation from Single Range Images*

[9] Heseltine T., *Face Recognition: Two Dimensional and Three Dimensional Techniques*, PhD. Dissertation, University of York, 2005

[10] <http://www.frav.es/database/FRAV3D/index.html>

[11] Minolta VIVID 700 3D laser scanner" <http://kmpi.konicaminolta.us/vivid>

۳ نتایج تجربی

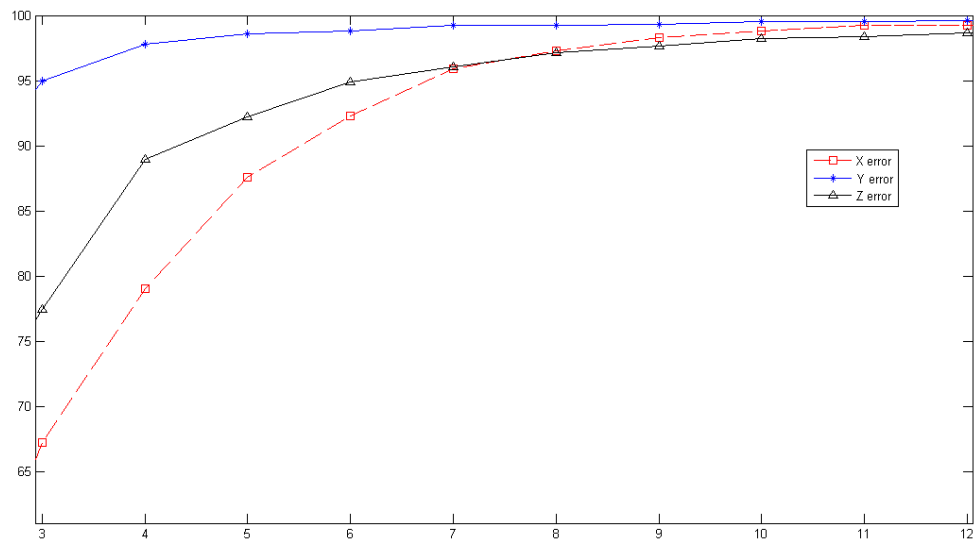
الگوریتم ذکر شده بر روی پایگاه داده‌ی FRAV3D پیاده‌سازی شده است و نتایج آن در جدول 1 آورده شده است. برای سنجش میزان دقت روش پیشنهادی پس از تخمین زاویه‌ی چهره، تصویر به همان اندازه چرخش داده شده تا تصویر روبرو استخراج گردد. پس از آن با استفاده از چهار نقطه در اطراف بینی که به صورت دستی تعیین شده‌اند زاویه‌ی چهره محاسبه شده است. در این مرحله انتظار داریم که به خاطر تراز شدن چهره این زوایا صفر باشند، بنابراین اندازه زوایه‌ای که این چهار نقطه بیان کرده‌اند به عنوان خطای سیستم در نظر گرفته‌ایم. این چهار نقطه بصورت دو نقطه در طرفین بینی و دو نقطه در بالا و پایین بینی می‌باشند. زاویه‌ی خط واسط دو نقطه‌ی بالا و پایین بینی با محور عمودی خطا در راستای محور Z در نظر گرفته شده‌است، خطای صفر به معنی عمود بودن خط واسط دو نقطه است. با استفاده از اختلاف ارتفاع بین نقطه‌ی زیر و بالای بینی چرخش حول محور X محاسبه شده است، هم ارتفاع بودن این دو نقطه به معنی خطای صفر در راستای این محور است. با استفاده از دو نقطه‌ی اطراف بینی و محاسبه‌ی اختلاف ارتفاع میان آنها خطا در راستای محور Y محاسبه می‌شود. نتایج پیاده سازی الگوریتم بر روی 14 گروه تصاویر پایگاه داده در جدول 1 آمده است. در شکل 7 نمودار دقت روش به ازای خطاهای متفاوت رسم شده است.

جدول 1: خطای تخمین زاویه در سه راستای محور مختصات

	راستای محور Z	راستای محور Y	راستای محور X	
خطای 10 درجه	%98.21	%99.52	%98.81	
خطای 6 درجه	%94.88	%98.81	%92.33	
خطای 3 درجه	%77.46	%95.00	%67.24	

۴ نتیجه‌گیری

زمان اجرای الگوریتم بدون در نظر گرفتن زمان مورد نیاز برای تعیین مکان بینی، کمتر از 1 ثانیه است. همچنین این روش نیاز به هیچ



شکل 7: دقت تخمین زاویه به ازای خطاهای مختلف