

استخراج خودکار ویژگی‌های چهره، برای شناسایی هویت مستقل از زاویه در تصاویر سه بعدی

علی رضا بساق زاده، محمدعلی دوستاری و علی رضا بهراد

تهران، ابتدای آزاد راه خلیج فارس، روبروی حرم مطهر حضرت امام خمینی (ره)،

دانشگاه شاهد، دانشکده فنی و مهندسی

E-mail: Bossaghzadeh@shahed.ac.ir

Doostari@shahed.ac.ir

Behrad@shahed.ac.ir

چکیده

یکی از مشکلات الگوریتم‌های شناسایی بر مبنای تصاویر دو بعدی، شناسایی در حالاتی است که تصویر موجود برای شناسایی از روبرو نباشد. استفاده از پارامترهای مستقل از زاویه ی تصاویر سه بعدی چهره این امکان را به ما می دهد تا بتوانیم یک شناسایی مستقل از زاویه ی چهره انجام دهیم. تعیین محل نقاط ویژه ی چهره به عنوان یک مرحله مرتبط با تنظیم چهره از نظر زاویه و اندازه، لازم است. الگوریتم تعیین نقاط چهره ی پیشنهادی در این تحقیق بر اساس انحنای موجود در تصویر عمل می کند. بدین صورت که ابتدا میانگین انحنای موجود در تصویر را محاسبه کرده و پس از اعمال یک فیلتر برای حذف نویز محل بینی پیدا می شود. با پیدا شدن مکان نوک بینی، در یک ناحیه ی محدودتر به دنبال محل گوشه‌های داخلی چشم ها می گردیم. این روش بر روی پایگاه داده‌ی FRAV3D پیاده سازی شده است. نتایج بدست آمده نشان می دهد که روش پیشنهادی دقت قابل قبولی در تعیین مکان نقاط چهره دارد و به زمان اجرای بسیار کمی نیاز دارد.

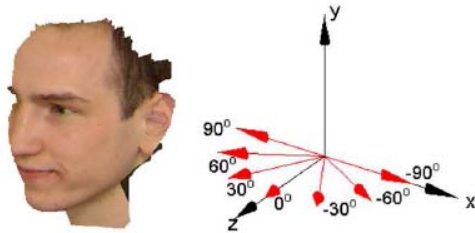
کلمات کلیدی - تصاویر سه بعدی چهره، انحنای میانگین، شناسایی چهره

هویت افراد استفاده کنند. در میان این مشخصه های زیستی، استفاده از تصویر چهره انسان بدلیل ماهیت غیر دخالت کننده گی¹ آن و نیاز کمتر به همکاری افراد مورد توجه قرار گرفته است. روشهای شناسایی چهره ی بر مبنای تصاویر دو بعدی در شرایط مناسب به نتایج مطلوبی رسیده اند [1]. مشکلات اصلی

۱- مقدمه

در طی سالهای گذشته گرایش به سمت شناسایی با کمک مولفه های زیستی افزایش یافته است. سیستم های بوجود آمده می توانند اطلاعات فیزیکی افراد از قبیل تصویر چهره، عنبیه، اثر انگشت و صدا را ذخیره کرده و از آنها برای تایید یا تشخیص

¹ non-intrusive



شکل 1: تصویر چهره با چرخش حول محور y ها [8]

در تمامی روشهایی برای تعیین موقعیت نقاط خاص در تصاویر سه بعدی چهره پیشنهاد شده است ابتدا بینی به عنوان بارزترین نقطه پیدا شده و بر اساس موقعیت آن بقیه نقاط مورد نیاز پیدا می شوند. برخی برای تعیین مکان بینی در تصاویر، نزدیکترین نقطه به دوربین را به عنوان بینی در نظر گرفته اند [9] و از روی بینی سایر نقاط مورد نیاز را استخراج می کرده اند، که این فرض در حالتی صادق است که چرخش زیادی حول محور x و y نداشته باشیم و در واقع تصاویر روبرو باشند. در [10] از تصویر چرخشی⁴ برای پیدا کردن بینی و سایر نقاط استفاده شده است. استفاده از تصویر چرخشی بار محاسباتی زیادی دارد که برای بهبود عملکرد ابتدا یک سری مکان ها که احتمال حضور بینی و چشم در آنها زیاد است تعیین شده و سپس برای تعیین موقعیت دقیق در آن نقاط تصویر چرخشی استخراج شده و با کمک کلاس بند SVM ⁵ محل نقاط مورد نظر تعیین شده اند. استفاده از این روش نیاز به داده های اولیه و آموزش یک شبکه عصبی دارد. در [8] برای تعیین بینی، زوایای حول محور y را کوانتیزه کرده و سپس تصاویر را حول این محور چرخانده و در این چرخش های متفاوت نزدیکترین نقطه به دوربین به عنوان کاندید در نظر گرفته می شود. پس از چرخش کامل تصویر حول محور y از بین این نقاط کاندید، بینی تشخیص داده می شود. این روش تنها در حالتی که تصویر چرخش حول محور y ها داشته باشد می تواند بینی را پیدا کند و در صورتی که تصویر چرخش تحت زوایای دیگر داشته باشد استفاده از این روش مقرون به صرفه نخواهد بود. در [3] برای پیدا کردن بینی و گوشه های درونی چشم از روش قطعه بندی HK استفاده کرده است. برای یک سطح، با داشتن انحنای میانگین $-H$ و انحنای گوسی $-K$ می توان وضعیت ظاهری آن را با توجه به شکل 2 تعیین کرد. تعیین کرد که به صورت انحنای روبه بالا، روبه پایین، نقطه ی زین اسبی، ماکزیمم، مینیمم یا سطح صاف است. در این حالت بینی نقطه ماکزیمم و گوشه های درونی چشم حالت مینیمم را دارند. روش پیشنهادی ما قادر به تعیین موقعیت بینی تحت هر زاویه دلخواه حول محور x ، y و z می باشد و تنها محدودیت آن در حالتی است که بیش از نیمی از بینی در چرخش حول محور

این روشها، تاثیر تغییرات شدت روشنایی و تغییرات در زاویه ی چهره است که می تواند به شناسایی اشتباه منجر شود. با توجه به اینکه چهره انسان سه بعدی است، استفاده از اطلاعات دو بعدی باعث از بین رفتن بخشی از اطلاعات تصویر می شود. با توجه به پیشرفتهای انجام شده در زمینه ی تصویر برداری سه بعدی و کاهش قیمت جاروب گر های¹ لیزری، شناسایی با استفاده از تصاویر سه بعدی چهره مورد توجه خاص قرار گرفته است. در هر دو نوع شناسایی، با استفاده از تصاویر دو بعدی و سه بعدی، تنظیم تصاویر "آموزشی و آزمایشی"² پیش از انجام شناسایی لازم می باشد. شناسایی با استفاده از تصاویر سه بعدی چهره این امکان را دارد که تغییرات زاویه ی چهره را جبران کند. تصاویر سه بعدی را می توان در فضای سه بعدی چرخاند، به طوری که این تصاویر پیش از شناسایی تنظیم شوند. یکی از روشهای تنظیم چهره استفاده از تبدیل هتلینگ³ است که در [2] پیشنهاد شده است. در [3] برای تصحیح زاویه از تنظیم تصویر ورودی با یک سطح عمومی استفاده کرده است. تنظیم تصاویر بر اساس ویژگی های چهره یکی دیگر از روشهای متداول است [4]. تصاویر دوبعدی با استفاده از دو نقطه پیش از شناسایی حول محور Z و نسبت به اندازه تنظیم می شوند. تصاویر سه بعدی را می توان با استفاده از سه نقطه حول هر سه محور مختصات و نسبت به اندازه تنظیم کرد.

به منظور داشتن یک سیستم شناسایی اتوماتیک به یک سیستم نیاز داریم که ویژگی های مورد نیاز را به طور دقیق تعیین کرده و خطای کمی داشته باشد. در شناسایی با کمک تصاویر سه بعدی به دلیل اینکه اطلاعات موجود داده های هندسی چهره می باشند تنظیم چهره به عنوان یک عملیات پیش پردازشی اهمیت بیشتری می یابد. ویژگی های مورد استفاده ی متفاوتی برای تنظیم تصاویر دو بعدی و سه بعدی چهره وجود دارد، از جمله ی این ویژگی ها، خصوصیات محلی چهره مورد استفاده در [5]، [6] و نقاط خاص در چهره می باشد [7]. به طور کلی روش استفاده از نقاط خاص دقت بالاتری نسبت به خصوصیات محلی چهره دارد [8]. لذا در این تحقیق هدف ما پیدا کردن یک مجموعه نقاط خاص که شامل نوک بینی و گوشه های داخلی چشم است می باشد. دلیل استفاده از این نقاط این است که استفاده از این نقاط برای تنظیم چهره کافی بوده و در تمامی حالاتی که چرخش حول محور Z یا x داشته باشیم این نقاط در تصویر قابل رویت هستند و پنهان شدن آنها فقط در حالتی است که چرخش زیادی حول محور y ها داشته باشیم. در شکل 1 یک تصویر با چرخش حول محور y ها آورده شده است.

⁴ Spin Image

⁵ Support Vector Machine

¹ Scanner

² Train, Test

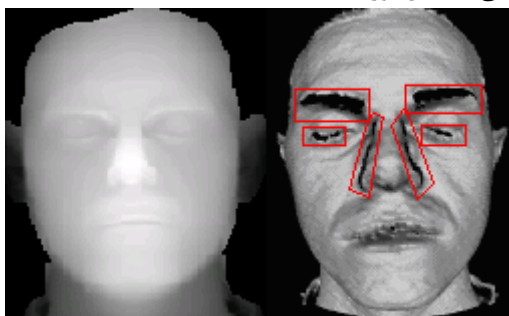
³ Hotelling

افراد خواسته شده است که چشمهای خود را بسته نگه دارند. به ازای هر فرد ۱۶ تصویر گرفته شده است که سعی شده تمامی حالات را از قبیل چرخش در تمامی جهات، حالت چهره^۱ و تغییرات شدت روشنایی را شامل شود [11]. در هر تصویر فقط یکی از پارامترهای اشاره شده تغییر داده شده است. در شکل 3 چند نمونه از تصاویر سه بعدی و معادل دو بعدی یک شخص آورده شده است.



شکل 3: تصاویر سه بعدی و معادل دو بعدی آنها

در تصویر برداری سه بعدی به دلیل وجود نویز هایی که در هنگام تصویر برداری بوجود می آید همچنین مشکلاتی که در تصویر برداری از نقاط حاوی مو وجود دارد دو مشکل نویزهای پرت^۲ و حفره بوجود می آید. تصاویر دارای نویز پرت دارای جهش^۳ هستند و در حالتی که تصویر دارای حفره است به ازای نقاطی در تصویر هیچ مقداری موجود نمی باشد. پس از استخراج داده ها به صورت نقشه عمقی برای حذف نویزهای جهش از یک فیلتر میانه و برای پر کردن حفره ها از روش درون یابی مکعبی استفاده شده است. در شکل 4 تصویر دارای حفره و معادل درون یابی شده آن آورده شده است.



شکل 4: تصویر دارای حفره و معادل درون یابی شده آن

¹ gesture
² Outlier
³ Spike

پنهان شده باشد. این روش فقط از داده های میانگین انحنای چهره استفاده می کند و نیازی به استفاده از داده های آموزشی و کلاس بند ندارد، همچنین، نقاط چهره را در همان زاویه ای که تصویر قرار دارد پیدا میکند و نیازی به چرخش تصاویر ندارد. به همین دلیل این روش سرعت محاسباتی بالاتری نسبت به روشهایی دارد که بینی را در زوایای مختلف پیدا میکنند.

	K	-	0	+
H				
-				
0				
+				

شکل 2: نوع انحنای سطح با توجه به علامت انحنای میانگین و گوسی

ادامه ای این مقاله به این صورت است: در بخش ۲-۱- توضیحاتی در مورد پایگاه داده ی FRAV3D و روش استخراج داده ها و پردازشهای مورد نیاز آورده شده است. در بخش ۲-۲- در مورد انحنای تصویر و روش محاسبه آن توضیح داده شده است. در بخش ۲-۳- روش استخراج بینی و چشمها آورده شده است. در بخش ۳- نتایج تجربی و در قسمت ۴- نتیجه گیری آمده است.

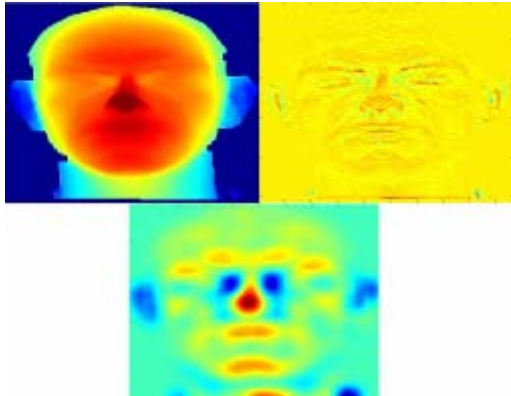
۲- استخراج ویژگی ها

۲-۱- 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 [11]. تصاویر موجود در پایگاه داده با استفاده از دوربین Minolta VI-700 [12] تصویربرداری شده که اطلاعات خروجی این دوربین شامل اطلاعات بافت چهره، تصویر دوبعدی، و یک فایل VRML، تصویر سه بعدی، می باشد. در صورت نیاز می توان تصویر معادل ۲،۵ بعدی را از تصویر VRML استخراج کرد. در تمامی حالات شدت روشنایی محیط کنترل شده بوده و در هیچ یک از تصاویر از کلاه، عینک استفاده نشده است و نیز هنگام تصویربرداری از

۲-۲- انحنا های تصویر

که بینی حاوی بیشترین انحنا مثبت و چشم ها حاوی بیشترین انحنا منفی در تصویر می باشند. برای اینکه این نقاط در تصویر بارزتر شده و استخراج آن ساده تر باشد از یک فیلتر استفاده شده که تاثیر نویز را کم کرده و موقعیت نقاط مورد نظر را در تصویر بهتر نشان می دهد. این فیلتر یک فیلتر جمع کننده است که همزمان با کاهش اثر نویز باعث می شود که محدوده بینی و چشم که حاوی انحنا بیشینه هستند، در تصویر واضح تر و برجسته تر نشان داده بشوند. در شکل 5 یک تصویر نقشه عمقی و تصویر میانگین انحنا قبل و بعد از فیلتر شدن آورده شده است.



شکل 5: تصویر اصلی، تصویر انحنا متوسط، تصویر فیلتر شده

پس از اعمال فیلتر و بدست آوردن تصویر بهبود یافته، ابتدا موقعیت بینی را در تصویر تعیین می کنیم. برای این کار در تصویر فیلتر شده به دنبال بیشترین مقدار انحنا می گردیم. همانطور که در شکل 5 مشخص است در محدوده ی صورت ناحیه بینی بیشترین میزان انحنا را دارد. پس از تعیین موقعیت بینی، به سراغ موقعیت گوشه های درونی چشم می رویم. برای جستجوی بهتر و دقیق تر ابتدا محدوده بینی را در تصویر تعیین می کنیم. محدوده تقریبی چشم ها در حالتی که چرخش چهره حول محور Z بیشتر از ۹۰ درجه نباشد، بالای بینی قرار دارد. به همین دلیل ابتدا محدوده بینی را در تصویر جدا می کنیم. در این حالت به دنبال کمترین مقدار انحنا در تصویر فیلتر شده می گردیم. به دلیل اینکه در اینجا دو کمینه محلی داریم که ممکن است با هم برابر نباشند، ابتدا تصویر را به باینری تبدیل کرده و سپس در اجزای به هم مرتبط^۴ به دنبال مقادیر کمینه می گردیم. پس از پیدا کردن مینیمم ها، آنها را به ترتیب نزولی به صعودی مرتب می کنیم. دو مقدار اول مینیمم ها که معادل بیشترین اندازه انحنا مینیمم هستند را به عنوان مکان گوشه های درونی چشم در نظر می گیریم. در شکل 6 یک نمونه تصویر بالای بینی که چشمها در آن قرار دارند و معادل باینری شده ی آن را مشاهده می کنید.

تمام منحنی هایی را که بر نقطه P در سطح S می گذرند را C_i در نظر بگیریم. هر کدام از این C_i ها دارای یک انحنا κ_i می باشند. از بین تمامی این κ_i ها یکی کمینه κ_1 و یکی بیشینه κ_2 است که به این دو مقدار، انحناهای اصلی^۱ سطح S می گویند. انحنا میانگین و انحنا گوسی برای P را با H و K نمایش می دهند که از رابطه ی آنها در معادلات (۲و۱) آمده است.

$$K = \kappa_1 * \kappa_2 \quad (1)$$

$$H = \frac{1}{2}(\kappa_1 + \kappa_2) \quad (2)$$

اگر نقطه P در سطح S باشد انحنا متوسط توسط رابطه ی معادله ی (۳) محاسبه می شود.

$$H = \frac{eG - 2fF + gE}{2(EG - F^2)} \quad (3)$$

E ، F و G ضرایب شکل اصلی اولیه^۲ و e ، f و g ضرایب شکل اصلی ثانویه^۳ هستند. شکل اصلی اولیه از حاصل ضرب بردارهای متعامد و شکل اصلی ثانویه از حالت متقارن درجه دو بردارهای متعامد بدست می آید [13]. رابطه ضرایب مورد نیاز در معادله ی (۳) از رابطه های معادلات (۴) تا (۹) قابل محاسبه هستند.

$$E = 1 + h_u^2 \quad (4)$$

$$F = h_u * h_v \quad (5)$$

$$G = 1 + h_v^2 \quad (6)$$

$$e = \frac{h_{uu}}{\sqrt{1 + h_u^2 + h_v^2}} \quad (7)$$

$$f = \frac{h_{uv}}{\sqrt{1 + h_u^2 + h_v^2}} \quad (8)$$

$$g = \frac{h_{vv}}{\sqrt{1 + h_u^2 + h_v^2}} \quad (9)$$

در صورتی که در بخشی از سطح انحنا زیادی داشته باشیم، هر دو انحنا کمینه و بیشینه و در نتیجه انحنا میانگین مقادیر بزرگی خواهند داشت. بنابراین پس از محاسبه انحنا میانگین باید به دنبال نقاط با انحنا بالا بگردیم.

۲-۳- استخراج بینی و چشم ها

پس از محاسبه انحنا تصویر باید به دنبال نقاط با انحنا بیشینه بگردیم ولی داده های حاصل حاوی نویز هستند و کار با آنها سخت می باشد لذا نیاز به فیلتر شدن دارند. می دانیم

¹ Principal Curvature

² First Fundamental Form

³ Second Fundamental Form

⁴ Connected Component

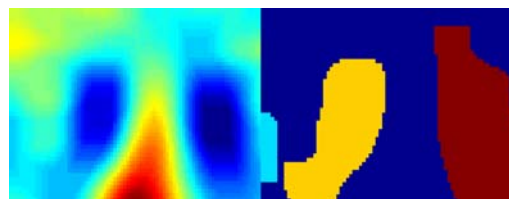
			حول محور Y
%۹۶	%۱۰۰	%۱۰۰	چرخش ۵" حول محور Y
%۱۰۰	%۱۰۰	%۱۰۰	چرخش ۵" حول محور Y
%۹۸	%۱۰۰	%۱۰۰	چرخش کم حول محور Z
%۹۸	%۱۰۰	%۱۰۰	چرخش زیاد حول محور Z
%۹۷	%۹۹	%۱۰۰	لیخند
%۹۷	%۹۸	%۹۸	دهان باز
%۹۸	%۱۰۰	%۱۰۰	چرخش حول محور X، بالا
%۹۶	%۹۹	%۹۹	چرخش حول محور X، پایین

نتایج بدست آمده نشان می‌دهد که روش پیشنهادی دارای %۹۹,۶ تشخیص صحیح موقعیت بینی، %۹۹,۲ تشخیص صحیح موقعیت یکی از چشم‌ها و %۹۶,۶ تشخیص صحیح موقعیت چشم دوم را دارد.

روش پیشنهاد شده بدلیل اینکه نیازی به چرخش تصویر ندارد، نسبت به روش [8] سرعت بالاتری دارد. این روش از فرض ساده‌شده‌ی نزدیکترین نقطه به دوربین که فقط در زوایای کم صادق است و در [9] پیشنهاد شده نیز استفاده نمی‌کند. همچنین نسبت به روش ارایه شده در [10] این مزیت را دارد که نیاز به داده‌های پیشین برای آموزش ندارد و از کلاس بند استفاده نمی‌کند. در جدول 2 مقایسه‌ی بین روش پیشنهادی با روشهای موجود آمده است.

جدول 2: مقایسه‌ی روش پیشنهادی با روشهای موجود

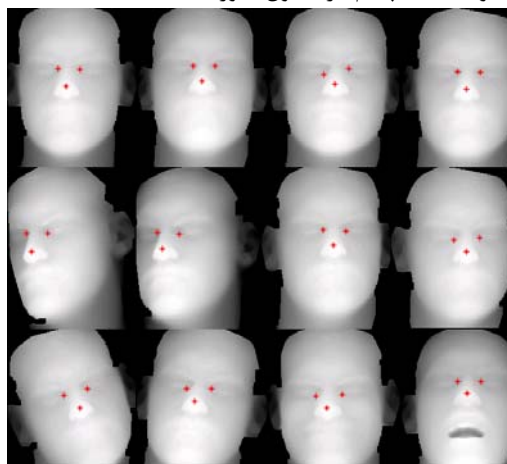
Spin Image [10]	نزدیکترین نقطه [9]	چرخش تصاویر [8]	روش پیشنهادی	
تحت هر زاویه‌ی دلخواه	تصاویر روبرو با چرخش بسیار کم	فقط چرخش حول محور Yها	تحت هر زاویه‌ی دلخواه	زوایایی که نقاط را پیدا میکند
حدود ۱۰ ثانیه	بسیار کم		حدود ۱,۵ ثانیه	زمان اجرا
دارد	ندارد	ندارد	ندارد	نیاز به داده‌های قبلی



شکل 6: بخش جداشده حاوی چشم و تصویر باینری شده آن

۳- نتایج تجربی

برای پیاده‌سازی این روش از نرم‌افزار Matlab استفاده شده است. روش پیشنهادی بر روی پایگاه داده‌ای که در بخش ۲-۱- توضیح داده شد پیاده‌سازی شده است. تصاویر با توجه به تقسیم بندی درون پایگاه داده به ۱۶ دسته تقسیم می‌شدند، که برای پیاده‌سازی از ۱۴ دسته استفاده شده است. در هر دسته تصاویر تمامی ۱۰۶ نفر بررسی شده‌اند. در شکل 7 تصاویر مربوط به یک نفر که مکان بینی و چشم‌ها در آن مشخص شده است آورده شده است. نتایج شبیه‌سازی با توجه به درست تعیین شدن موقعیت بینی و گوشه‌های چشم در جدول 1 آورده شده است.



شکل 7: ۱۲ تصویر یک شخص به همراه موقعیت بینی و چشم

جدول 1: تعداد تصاویری که موقعیت بینی یا چشم در آنها درست پیدا شده است.

تعیین محل چشم دوم	تعیین محل چشم اول	تعیین محل بینی	زاویه‌ی تصویر
%۹۶	%۹۸	%۱۰۰	روبرو
%۹۵	%۹۹	%۹۹	روبرو
%۹۳	%۹۸	%۱۰۰	روبرو
%۹۶	%۹۸	%۹۹	روبرو
%۹۳	%۱۰۰	%۱۰۰	چرخش ۲۵" حول محور Y
%۹۶	%۱۰۰	%۹۹	چرخش ۲۵"

۴- نتیجه گیری

- Recognition" International Conference on Automatic Face and Gesture Recognition, Southampton, UK, April 2006
- [9] Heseltine T., Face Recognition: Two Dimensional and Three Dimensional Techniques, Ph. D. Dissertation, University of York, 2005
- [10] Cristina Conde et al, "3D Facial Feature Location with Spin Images", MVA2005, IAPR Conference on Machine Vision Applications, Japan.
- [11] <http://www.frav.es/database/FRAV3D/index.html>
- [12] Minolta VIVID 700 3D laser scanner" <http://kmpi.konicaminolta.us/vivid>
- [13] Modern Differential Geometry of Curves and Surfaces with Mathematica. 2nd ed, 1997 (p. 377).

این روش قابلیت پیاده سازی بر روی سیستمهایی را که به صورت برخط^۱ کار می کنند را بدلیل داشتن زمان اجرای بسیار کم، حداکثر ۱,۵ ثانیه برای هر تصویر، دارا می باشد. از نظر بار محاسباتی بدلیل اینکه فقط با انحنای متوسط سطح کار می کند، نیاز به محاسبات زیادی ندارد. خطای این روش بدلیل عملکرد سلسله مراتبی در صورت بروز اشتباه در بخش جستجوی بینی در بخش جستجوی چشم نیز دچار اشتباه می شود. خطای پیدا کردن بینی بیشتر ناشی از موهای اطراف صورت بوجود می آید و که در صورت حذف این موها دقت روش پیشنهادی بالاتر خواهد بود. در صورت بهبود روش پیدا کردن بینی و تمایز بین بینی و دهان در تصویر می توان دقت الگوریتم پیشنهادی را افزایش داد.

سپاسگزاری

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

مراجع

- [1] W. Zhao, R. Chellappa, P.J. Phillips, and A. Rosenfeld "Face Recognition : A Literature Survey " ACM Computing Surveys, Volume 35, Issue 4, December 2003
- [2] An Efficient Multimodal 2D-3D Hybrid Approach to Automatic Face Recognition. IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL. 29, NO. 11, NOVEMBER 2007
- [3] Multiple Nose Region Matching for 3D Face Recognition under Varying Facial Expression. IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL. 28, NO. 10, OCTOBER 2006
- [4] S. Li and A. Jain (Eds.). Handbook of Face Recognition. Springer, 2005.
- [5] D. Cristinacce and T. Cootes. Facial feature detection using adaboost with shape constraints. In Proc. 14th BMVC, pages 231-240, Norwich, UK, Sep. 2003.
- [6] F. Tsalakanidou, S. Malassiotis, and M. Strintzis. Face localization and authentication using color and depth images .IEEE Transactions on Image Processing, 14(2):152-168, 2005.
- [7] K. Toyama, R. Feris, J. Gemell, and V. Kruger. Hierarchical wavelet networks for facial feature localization. In Proc. IEEE International Conference on Automatic Face and Gesture Recognition, pages 118-123, Washington D.C., 2002.
- [8] Xiaoguang Lu and Anil K. Jain, "Automatic Feature Extraction for Multiview 3D Face

¹ Online