



دانشگاه صنعتی شاهرود



دانشکده برق و رباتیک

گروه الکترونیک

تشخیص چهره با استفاده از مدل‌های سه‌بعدی

دانشجو:

مهدی تاجیانی

استاد راهنما:

دکتر علی سلیمانی

استاد مشاور:

دکتر حسین مروی

پایان نامه ارشد جهت اخذ درجه کارشناسی ارشد

شهریور ماه ۱۳۸۷

#### چکیده

مطالعه‌ی روش‌های جدید تشخیص بیومتریک جهت دستیابی به کاربردهای بیشتر و عملکرد بهینه از چالش‌های مهم پژوهشگران این حوزه می‌باشد. تشخیص چهره به عنوان مهمترین روش تشخیصی انسان‌ها به علت قابلیت‌های بالا از موارد مورد توجه در این زمینه است. با توجه به ماهیت سه‌بعدی چهره روش‌های تشخیص چهره‌ی سه‌بعدی نسبت به روش‌های دوبعدی از پتانسیل بیشتری جهت رفع موانع موجود برخوردارند. در این میان روش‌های تشخیص چهره‌ی سه‌بعدی مبتنی بر مدل، به علت قابلیت سنتز چهره و همچنین عدم نیاز به تصاویری ورودی سه‌بعدی - که تهیه‌ی آنها در حال حاضر فرایندی پیچیده، زمانبر و پرهزینه است - از اهمیت بیشتری برخوردار هستند.

تشخیص چهره‌ی سه‌بعدی مبتنی بر مدل به دو روش اصلی تقسیم می‌شود. روش اول روش مبتنی بر مدل عمومی بوده و روش دوم استفاده از مدل میان‌ساز می‌باشد. جهت ساخت مدل میان‌ساز به یک پایگاه داده‌ی سه‌بعدی نیاز داریم که به علت در اختیار نداشتن چنین پایگاه داده‌ای، ساخت پایگاه داده‌ی مدل‌های سه‌بعدی چهره از روی تصاویر دوبعدی مورد توجه قرار گرفت. برای این منظور از الگوریتم هندسه‌یابی تصویری که روشی مبتنی بر مدل عمومی است، استفاده شد. بدین ترتیب مطالعه و پیاده‌سازی هر دو روش اصلی تشخیص چهره‌ی مبتنی بر مدل در این پایان نامه صورت گرفته است. در مطالعه‌ی حاضر، دو پایگاه داده‌ی سه‌بعدی چهره از روی تصاویر دوبعدی ایجاد شدند. پایگاه داده اول با استفاده از تصاویر دوبعدی چهره‌ی پایگاه داده‌ی MPI و پایگاه داده‌ی بعدی بر اساس تصاویر چهره‌های قومیت‌های مختلف ایرانی تهیه شدند.

تشخیص چهره مبتنی بر مدل میان‌ساز به علت ویژگی‌های قابل توجه و مزیت‌های نسبی به عنوان مهمترین روش تشخیص چهره‌ی سه‌بعدی مبتنی بر مدل مورد مطالعه و بررسی قرار گرفته است. مدل میان‌ساز در واقع ترکیبی از مدل‌های ریخت‌پذیر و تکنیک‌های ارتباط تصویری کامپیوتری جهت شبیه‌سازی اثرات نوردهی می‌باشد. در این مدل تغییر زاویه‌ی دید و شرایط نورپردازی به سادگی صورت گرفته و می‌تواند تنها با یک تصویر ورودی و بدون نیاز به دخالت کاربر تخمین قابل قبولی از بافت، شکل، زاویه‌ی دید، و مشخصات نوردهی مدل چهره ایجاد کند. سنتز چهره و پیاده‌سازی حالات مختلف بر روی مدل سنتز شده نیز از مزایای این روش می‌باشد. مهمترین بخش مدل‌سازی میان‌ساز تعیین تناظر نقطه به نقطه بین مدل‌های نمونه می‌باشد. در این پژوهش جهت پیاده‌سازی این بخش، از روشی کاملاً جدید و مبتنی بر گراف به منظور بهینه‌سازی نتایج انطباق استفاده شده است. نتایج بهینه‌سازی توسط این روش بهبود قابل توجه در عملکرد مدل میان‌ساز، کاهش حجم محاسبات و کیفیت بالاتر سنتز چهره را نشان می‌دهد. این الگوریتم یک الگوریتم بهینه‌سازی گراف و مستقل از روش انطباق می‌باشد. یک‌به‌یک‌سازی نتایج انطباق نیز علاوه بر بهینه‌سازی آن از دیگر مزایای این روش جدید است. تعیین درجه‌ی یک‌به‌یک‌سازی بهینه در این روش می‌تواند، بین پوشایی انطباق و یک‌به‌یک‌سازی آن توازن مطلوبی برقرار سازد. پس از پیاده‌سازی مدل

میان‌ساز تاثیر پایگاه داده‌ی سه‌بعدی در کیفیت سنتز و تشخیص توسط این الگوریتم مورد بررسی قرار گرفت. کیفیت قابل توجه تصاویر سنتز شده توسط مدل میان‌ساز مبتنی بر پایگاه داده‌ی سه‌بعدی چهره‌ی ایرانی، حاکی از نقش موثر پایگاه داده در ساخت مدل میان‌ساز می باشد. در این پژوهش، تشخیص چهره با استفاده از مدل میان‌ساز با نرخ برابر ۱۰۰٪ صورت پذیرفت با وجود حجم کم تصاویر آزمایشی ورودی، این نرخ بالا، بیانگر کارایی این مدل در زمینه‌ی تشخیص چهره می‌باشد.

کلمات کلیدی: تشخیص چهره‌ی سه بعدی، مدل‌سازی سه‌بعدی چهره، هندسه‌یابی تصویری، مدل میان‌ساز، انطباق تصاویر

## فصل اول: کلیات

- ۱-۱- مقدمه..... ۱
- ۲-۱- مشکلات و موانع تشخیص چهره..... ۲
- ۳-۱- مزایا و معایب تشخیص چهره سه‌بعدی..... ۵
- ۴-۱- روش‌های تشخیص چهره‌ی سه‌بعدی ..... ۹
  - ۱-۴-۱- روش‌های مبتنی بر رویه..... ۱۰
  - ۲-۴-۱- روش‌های آماری..... ۱۶
- ۵-۱- روش‌های مدل‌سازی سه‌بعدی چهره و تشخیص مبتنی بر مدل..... ۱۸
  - ۱-۵-۱- مدل‌های چهره‌ی ماهیچه‌ای..... ۲۱

۲۲	.....۱-۶- مزایای روش‌های تشخیص چهره‌ی سه‌بعدی مبتنی بر مدل
۲۴	.....۱-۶-۱- مزایای تشخیص چهره‌ی مبتنی بر مدل میان‌ساز
<b>فصل دوم: ارتباط تصویری کامپیوتری سه‌بعدی</b>	
۲۵	.....۱-۲- مقدمه
۲۵	.....۲-۲- تبدیلات هندسی سه‌بعدی
۲۶	.....۱-۲-۲- انتقال
۲۶	.....۲-۲-۲- دوران
۳۴	.....۳-۲-۲- تغییر مقیاس
۳۵	.....۳-۲- دید سه‌بعدی
۳۵	.....۱-۳-۲- مختصات دید
۳۸	.....۲-۳-۲- تصویرسازی
۴۲	.....۴-۲- تعیین سطوح قابل رؤیت
۴۳	.....۱-۴-۲- تعیین وجه پسین
۴۴	.....۲-۴-۲- روش بافر عمق

- ۴۵.....A-۳-۴-۲- روش بافر
- ۴۵.....۵-۲- مدل‌های نوردهی
- ۴۸.....۱-۵-۲- نور محیطی و بازتابش پخشی
- ۵۱ .....۲-۵-۲- بازتابش آینه‌ای و مدل فانگ
- ۶۰ .....۳-۵-۲- پیاده‌سازی مدل نوردهی بر روی چندضلعی‌ها

### فصل سوم: پایگاه داده‌ی سه‌بعدی

- ۶۴ .....۱-۳- مقدمه
- ۶۴ .....۲-۳- روش‌های تهیه تصاویر سه‌بعدی چهره
- ۶۵.....۱-۲-۳- تهیه‌ی مدل با استفاده از پویش لیزری
- ۶۸ .....۲-۲-۳- تهیه‌ی مدل با استفاده از هندسه‌یابی تصویری
- ۷۰ .....۳-۳- روش‌های ذخیره‌سازی اطلاعات سه‌بعدی
- ۷۰ .....۱-۳-۳- زبان مدل‌سازی حقایق مجازی (VRML)

۷۱	..... ۲-۳-۳- فایل‌های شیئی
۷۴	..... ۴-۳- ساخت مدل‌های سه‌بعدی چهره با استفاده از روش هندسه‌یابی تصویری
۷۵	..... ۱-۴-۳- تطبیق مدل
۸۴	..... ۲-۴-۳- استخراج بافت
۸۶	..... ۵-۳- پایگاه‌های داده‌ی چهره‌ی سه‌بعدی
۸۷	..... ۱-۵-۳- پایگاه داده MPI
۸۷	..... ۲-۵-۳- پایگاه داده چهره سه‌بعدی ارائه شده
۸۹	..... ۳-۵-۳- پایگاه داده‌ی Gavab3D

### فصل چهارم: سنتز و تشخیص چهره با استفاده از مدل میان‌ساز

۹۱	..... ۱-۴- مقدمه
۹۳	..... ۲-۴- مدل چهره‌ی میان‌ساز سه‌بعدی
۹۹	..... ۱-۲-۴- مدل میان‌ساز تقطیع شده
۱۰۲	..... ۲-۲-۴- خصوصیات چهره
۱۰۳	..... ۳-۲-۴- تعیین تناظر بین مدل‌ها

- ۱۰۴ ..... ۴-۲-۳-۱- روش‌های انطباق تصاویر
- ۱۰۵ ..... ۴-۲-۳-۲- الگوریتم شارش نوری (Optical Flow)
- ۱۱۰ ..... ۴-۲-۳-۳- انطباق تصاویر با استفاده از همبستگی محلی
- ۱۱۲ ..... ۴-۲-۳-۴- بهینه‌سازی نتایج انطباق
- ۱۲۳ ..... ۴-۳-۲- تطبیق تصاویر ورودی و مدل میان‌ساز
- ۱۲۵ ..... ۴-۳-۱- سنتز تصویر
- ۱۲۷ ..... ۴-۳-۲- مراحل تطبیق مدل و تصویر ورودی
- ۱۲۸ ..... ۴-۳-۲-۱- تابع هزینه
- ۱۳۱ ..... ۴-۳-۲-۲- فرایند بهینه‌سازی
- ۱۴۰ ..... ۴-۳-۲-۳- تطبیق مدل میان‌ساز و مدل‌های سه‌بعدی
- ۱۴۱ ..... ۴-۳-۳- ساخت مدل میان‌ساز

### فصل پنجم: بررسی نتایج

- ۱۴۴ ..... ۵-۱- مقدمه
- ۱۴۶ ..... ۵-۲- نتایج سنتز چهره



۱۵۰.....۳-۵- نتایج تشخیص چهره.

۱۵۲ .....۴-۵- روش تشخیص چهره.

### فصل ششم: نتیجه‌گیری و پیشنهادات

۱۵۵ .....۱-۶- نتیجه‌گیری

۱۵۶ .....۲-۶- پیشنهادات

### منابع و مراجع

۱۵۸.....منابع و مراجع انگلیسی.

### پیوست‌ها

۱۶۴.....پیوست الف.

### فصل اول

۸ .....۱-۱- دو نمونه از نقایص ممکن در تصاویر سه‌بعدی

۱۱ .....۲-۱- نمایش یک چندوجهی محدب با استفاده از EGI

۱۲ .....۳-۱- امضای نقطه واقع بر روی بینی

۱۱ .....۴-۱- صفحه تقارن چهره و ویژگی استخراج شده

۵-۱- سنتز حالات چهره بر روی مدل واترز ..... ۲۲

## فصل دوم

- ۲۹ ..... ۱-۲- مراحل اجرای دوران حول محور دوران موازی یک محور مختصات.....
- ۳۱ ..... ۲-۲- مراحل دوران سه بعدی، (الف): موقعیت اولیه، (ب) تا (و): مراحل ۱ تا ۵.....
- ۳۲ ..... ۳-۲- مراحل انطباق محور دوران بر محور Z.....
- ۳۳ ..... ۴-۲- ارتباط زوایای  $\alpha$  و  $\beta$  با تصاویر بردار U.....
- ۳۷ ..... ۵-۲- نحوه تشکیل دستگاه مختصات دید.....
- ۳۹ ..... ۶-۲- تصویر سازی بر روی صفحه‌ی دید، (الف): موازی، (ب): پرسپکتیو.....
- ۴۰ ..... ۷-۲- موقعیت تصویر در تصویرسازی اریب.....
- ۴۰ ..... ۸-۲- تصویر یک مکعب با  $\phi = 45$ .....
- ۴۱ ..... ۹-۲- - نگاشت پرسپکتیو نقطه‌ی  $(X, Y, Z)$  بر صفحه‌ی دید.....
- ۴۴ ..... ۱۰-۲- استفاده از خط پویش جهت تعیین عمق نقاط روی صفحه.....
- ۴۶ ..... ۱۱-۲- بافرهای موجود در روش بافر A.....
- ۴۹ ..... ۱۲-۲- جهت تشعشع انرژی از سطح دیفرانسیلی dA.....

- ۵۰ ..... ۱۳-۲- بین زاویه‌ی تابش و شدت بازتابش پخشی
- ۵۱ ..... ۱۴-۲- تأثیر بازتابش پخشی بر چهره
- ۵۲ ..... ۱۵-۲- بردارها و زوایای مهم در بازتابش آینه‌ای
- ۵۲ ..... ۱۶-۲- نمودار پارامتر فانگ به ازای مقادیر مختلف  $n_s$
- ۵۳ ..... ۱۷-۲- ضریب بازتابش آینه‌ای برای چند ماده‌ی مختلف
- ۵۵ ..... ۱۸-۲- تاثیر بازتابش آینه‌ای بر چهره
- ۵۹ ..... ۱۹-۲- شکست نور
- ۶۳ ..... ۲۰-۲- نتایج سه روش سایه‌گذاری

## فصل سوم

- ۶۶ ..... ۱-۳- تصاویر سه‌بعدی چهره شماره‌ی ۵ در دستگاه مختصات استوانه‌ای توسط پویشگر لیزری
- ۶۷ ..... ۲-۳- حفره‌های پر شده در تصویر شماره‌ی ۵ پایگاه داده‌ی MPI
- ۷۷ ..... ۳-۳- محل قرارگیری دوربین‌ها و تبدیلات هندسی در هندسه‌یابی تصویری
- ۸۶ ..... ۴-۳- نتایج مراحل تطبیق مدل
- ۸۸ ..... ۵-۳- نمونه‌ای از مدل‌های ۳ بعدی موجود در پایگاه داده ارائه شده

۸۹ ..... ۳-۶- نمایش دو بعدی ماتریس‌های شکل و بافت.....

۹۰ ..... ۳-۷- نمونه‌ای از تصاویر موجود در پایگاه داده‌ی Gavab3D.....

## فصل چهارم

۹۵ ..... ۴-۱- میان‌سازی بین دو مدل چهره.....

۹۷ ..... ۴-۲- مدل‌های میانگین پایگاه داده‌ی MPI در دستگاه مختصات استوانه‌ای.....

..... ۴-۳- تقطیع مدل میان‌ساز به ۴ ناحیه.....

۱۰۱

..... ۴-۴- مقایسه مدل میان‌ساز تقطیع شده و تقطیع نشده.....

۱۰۱

..... ۴-۵- پیاده‌سازی خصوصیات چهره بر روی مدل سنتز شده در شکل ۵-۴.....

۱۰۳

..... ۴-۶- انطباق مدل‌های چهره در مختصات استوانه‌ای با استفاده از روش همبستگی محلی.....

۱۱۲

..... ۴-۷- یک درخت انطباق نمونه.....

۱۱۴

۱۱۶	۸-۴ پیوندهای بهینه درخت انطباق شکل ۴-۷.....
	۹-۴ مراحل تعیین پیوندهای بهینه.....
	۱۲۰
	۱۰-۴ پیوندهای بهینه گسترش یافته زیردرخت گسترش یافته.....
	۱۲۰
	۱۱-۴ پیوندهای بهینه گسترش نیافته زیردرخت گسترش یافته و انتخاب‌های ممکن.....
	۱۲۱
	۱۲-۴ تصاویر ورودی.....
۱۲۲	۱۳-۴ نتیجه انطباق ۱ به ۲ توسط همبستگی محلی.....
۱۲۳	۱۴-۴ ساخت مدل میان‌ساز و سنتز چهره ورودی.....
۱۲۳	۱۵-۴ ابزار تعاملی جهت اختصاص مقادیر اولیه‌ی زاویه‌ی دید و مقیاس.....
۱۲۵	۱۶-۴ تأثیر استخراج بافت بر سنتز چهره.....
۱۴۰	

## فصل پنجم

	۱-۵ مقایسه نتایج سنتز چهره.....
	۱۴۸

۵-۲- مقایسه نتایج سنتز چهره بدون استخراج بافت.....۱۴۹

۵-۳- سنتز یک چهره‌ی مؤنث در دو زاویه‌ی دید متفاوت.....

۱۴۹

۵-۴- سنتز چهره‌های دارای حالت.....

۱۵۰

۵-۵- ۱۲ تصویر موجود برای فرد شماره‌ی ۱ در پایگاه داده‌ی چهره‌ی دانشگاه صنعتی امیرکبیر...۱۵۱

۵-۶- مراحل تشخیص چهره توسط مدل میان‌ساز.....۱۵۳

# فصل اول

# کلیات

تشخیص چهره به عنوان یکی از تشخیص‌های بیومتریک در دو دهه‌ی اخیر بسیار مورد توجه واقع شده است. یکی از دلایل مهم برای این موضوع پیشرفت بسیار زیاد سیستم‌های کامپیوتری و تجهیزات تصویر برداری است که تهیه‌ی تصاویر با کیفیت و دقت بالا و پیاده‌سازی الگوریتم‌های پیچیده را ممکن ساخته است.

به عنوان دلیل مهم دیگری می‌توان به وسعت کاربردهای ممکن برای سیستم‌های تشخیص چهره اشاره کرد. برای سیستم‌های تشخیص چهره کاربردهایی در زمینه‌های تجاری، امنیتی و قانونی قابل تصور می‌باشد. با وجود این میزان توجه، سیستم‌های تشخیص چهره‌ی امروزی تنها در شرایط کنترل شده و همسان تصویر برداری، عملکرد مناسبی از خود نشان می‌دهند و هنوز به سطح لازم جهت حضور در کاربردهای فوق دست پیدا نکرده‌اند. بنابراین توسعه الگوریتم‌های تشخیص چهره‌ی مقاوم<sup>1</sup> نسبت به شرایط کنترل نشده هنوز از موارد قابل پژوهش می‌باشد.

از دیگر دلایل اهمیت تشخیص چهره در بین سایر تشخیص‌های بیومتریک، عدم نیاز به همکاری افراد در فرایندهای جمع‌آوری اطلاعات و تشخیص می‌باشد. این خاصیت در تکنیک‌های بسیار مطمئن بیومتریک، همچون تشخیص اثر انگشت و عنبیه وجود ندارد. علاوه بر این در تشخیص چهره حتی نیازی به آگاهی افراد نیز وجود نداشته و پذیرش و اقبال عمومی بیشتری نسبت به این روش مشاهده می‌شود.

---

<sup>1</sup> Robust



## ۲-۱ مشکلات و موانع تشخیص چهره

چهره یک شیء سه بعدی است که تحت تأثیر مواردی چون تغییرات نور، زاویه دید، حالات چهره، انسداد<sup>۲</sup> و سایر تغییرات وابسته به زمان همچون افزایش سن و تغییرات مربوط به آرایش چهره دچار تغییر می شود. طراحی الگوریتم های تشخیص چهره ای که در مقابل این تغییرات مقاوم باشند، چالش اصلی پژوهشگران در این زمینه می باشد.

اطلاعات چهره در تصاویر دو بعدی تنها در قالب شدت نور و یا رنگ پیکسل ها و همچنین دو بعد از سه بعد چهره نمایش داده می شود. بنابراین اطلاعات این نوع از تصاویر به طور ذاتی وابسته به تغییرات تابش نور و زاویه دید می باشد. هیل<sup>۳</sup> و همکاران در سال ۱۹۹۷ نشان دادند که مغز انسان نیز در تشخیص چهره تحت تأثیر شدت نور و زاویه تابش نور می باشد [هیل و همکاران، ۱۹۹۷]. علاوه بر این مغز انسان در تشخیص چهره های آشنا زمانی که تصویر دو بعدی آنها تحت تابش از بالا تهیه شده باشد، دچار اشکال و سردرگمی می شود [جانستون<sup>۴</sup> و همکاران، ۱۹۹۲]. همچنین مغز انسان در تشخیص چهره های آشنا از تصاویر نگاتیو نیز پراشتباه و تأثیرپذیر ظاهر می شود [لیو<sup>۵</sup> و همکاران، ۱۹۹۹]. تمامی این آزمایشات بیانگر نقص ذاتی اطلاعات در تصاویر دو بعدی می باشد، به گونه ای که مغز انسان به عنوان قوی ترین ابزار تشخیص چهره ای شناخته شده نیز تحت تأثیر تغییر مشخصات تابش نور قرار می گیرد.

از دیگر مشکلات اساسی موجود در تشخیص چهره، تغییرات زاویه دید یا همان تغییرات زاویه تصویربرداری می باشد. همانطور که اشاره شد اطلاعات موجود در تصاویر دو بعدی شامل دو بعد از سه بعد چهره می باشد. بنابراین با تغییر زاویه تصویر برداری این بخش از اطلاعات موجود در تصویر دو بعدی تغییر

---

<sup>2</sup> Occlusion

<sup>3</sup> Hill

<sup>4</sup> Johnston

<sup>5</sup> Liu

پیدا می‌کند. به عبارت دیگر اختلاف بین دو تصویر چهره از دو شخص مختلف تحت زوایای تصویربرداری یکسان، بسیار کمتر از اختلاف بین دو تصویر چهره از یک شخص تحت زوایای تصویربرداری متفاوت، می‌باشد. انسان نیز در تشخیص چهره‌های ناآشنایی که تصاویر آموزشی و آزمایشی آن تحت زوایای مختلف تهیه شده باشد، دچار اشکال می‌شود [بروس<sup>۶</sup>، ۱۹۸۲]. با این وجود در چند سال اخیر همواره مناظراتی بر سر میزان وابستگی تشخیص شیء توسط انسان به زاویه‌ی دید صورت گرفته است [تار و بالتھوف<sup>۷</sup>، ۱۹۹۵]. آزمایشات بیانگر این است که مغز انسان در تعمیم زاویه‌ی دید هنگامی که تغییرات زاویه زیاد نباشد، عملکرد مناسبی دارد. برای مثال تطبیق تصاویر تمام‌رخ<sup>۸</sup> و نیم‌رخ<sup>۹</sup> توسط انسان امری مشکل و پراشتباه می‌باشد، در حالی که تطبیق تصاویر سه‌ربع و نیم‌رخ دارای خطا و پیچیدگی کمتری است [هیل و همکاران، ۱۹۹۷]. در مورد نحوه‌ی تشخیص مغز در تغییرات زاویه‌ی دید نظرات مختلفی ارائه شده است. اما تمامی آزمایشات مربوط به زاویه‌ی دید و تغییرات تابش نور نظریه ذخیره‌سازی تصاویر سه‌بعدی چهره در مغز را رد می‌کنند. این آزمایشات این سؤال اساسی را ایجاد می‌کنند که آیا می‌توان سیستم تشخیص چهره‌ای مبتنی بر تصاویر دوبعدی طراحی کرد که مستقل از تغییرات تابش نور و زاویه تصویربرداری باشد.

یکی دیگر از موانع تشخیص چهره مسئله‌ی انسداد می‌باشد. انسداد چهره به معنای دیده نشدن بخشی از چهره بر اثر وجود مانع می‌باشد. انسداد چهره معمولاً بر اثر وجود ریش، کلاه و یا عینک بوجود می‌آید.

اثرات ناشی از تغییر حالات چهره از دیگر موانع موجود در تشخیص دقیق چهره می‌باشند. چهره ساختاری پویا است که به علت تحریک ماهیچه‌های صورت و حرکت استخوانها به صورت غیرصلب<sup>۱۰</sup> تغییر

---

<sup>6</sup> Bruce

<sup>7</sup>Tarr, Balthoff

<sup>8</sup> Frontal

<sup>9</sup> Profile

<sup>10</sup> Non-Rigid

شکل می‌دهد. حالاتی همچون، گریه، خنده، اخم و تعجب از مواردی هستند که یک سیستم تشخیص چهره‌ی کارآمد، می‌بایست مستقل از آنها چهره را تشخیص دهد. تعیین حالت چهره نیز از موارد قابل پژوهش در این زمینه می‌باشد. علاوه بر این تغییرات وابسته به زمان نیز مشکلاتی اساسی در تشخیص چهره بوجود می‌آورد. ظاهر چهره بر اثر افزایش سن و شیوه‌ی زندگی دچار تغییر می‌شود. کاهش خاصیت کشسانی پوست و افتادگی آن، کاهش حجم لبها، ریزش مو، تغییر رنگ پوست بر اثر افزایش سن و تغییر در آرایش صورت، مدل مو و ریش از جمله تغییرات وابسته به زمان در چهره می‌باشند. این موارد می‌توانند منجر به تغییرات فاحش در ظاهر چهره شده و عملیات تشخیص را بسیار پیچیده و مشکل سازند.

### ۱-۳ مزایا و معایب تشخیص چهره‌ی سه‌بعدی

تشخیص چهره‌ی سه‌بعدی به نسبت تشخیص چهره‌ی دوبعدی از سابقه کمتری برخوردار است. در حال حاضر روش‌های تشخیص چهره‌ی دوبعدی به نسبت روش‌های تشخیص چهره‌ی سه‌بعدی از عملکرد بهتری برخوردار هستند. البته نقص اطلاعات موجود در تصاویر دوبعدی، به نوعی برای میزان موفقیت روش‌های دوبعدی، حد بیشینه ایجاد می‌کند. در این بخش با بررسی معایب و مزایای تشخیص چهره‌ی سه‌بعدی به امکان‌سنجی بهبود و غلبه‌ی روش‌های سه‌بعدی می‌پردازیم.

همانطور که اشاره شد اطلاعات چهره در تصاویر دو بعدی در قالب شدت نور و یا رنگ پیکسل‌ها و همچنین دو بعد از سه بعد چهره نمایش داده می‌شود. اطلاعات مربوط به بعد سوم نیز در اطلاعات مربوط به شدت نور و یا رنگ پیکسل‌ها نهفته است. اما رنگ و یا شدت نور پیکسل‌ها تابعی از شکل چهره، مشخصات نور تابیده شده، مشخصات نور محیطی، بافت چهره، میزان بازتابش چهره و مشخصات و محل قرارگیری دوربین می‌باشد. ساخت مدل نورپردازی‌ای که بتواند تمام این موارد را در نظر گرفته و آنها را از هم تفکیک

کند، کاری بسیار مشکل می‌باشد و تاکنون چنین مدلی ارائه نشده است. بنابراین استخراج اطلاعات مربوط به بعد سوم و در نتیجه تعیین شکل چهره با دقت بالا از روی تصاویر دوبعدی تاکنون امکان‌پذیر نگشته و با توجه به نقص اطلاعات در تصاویر دوبعدی، غیرممکن به نظر می‌رسد. در تصاویر سه‌بعدی نیز تغییرات مربوط به تابش نور در بافت چهره تأثیر می‌گذارد ولی به علت عدم تأثیرپذیری اطلاعات شکل چهره، وابستگی تصاویر سه‌بعدی به مشخصات تابش نور بسیار کمتر می‌باشد.

از دیگر تفاوت‌های بین تشخیص چهره‌ی دوبعدی و سه‌بعدی می‌توان به عدم وابستگی تشخیص چهره‌ی سه‌بعدی به زاویه‌ی دید یا زاویه‌ی چهره اشاره کرد. در مقابل همانطور که اشاره شد تشخیص چهره‌ی دوبعدی به شدت وابسته به زاویه‌ی چهره می‌باشد. جهت رفع این مشکل، تلاش‌هایی در جهت تغییر زاویه‌ی چهره در تصاویر دو بعدی به یک زاویه‌ی خاص و مقایسه‌ی چهره‌ها در این زاویه‌ی یکسان صورت گرفته است [کیم و کیتلر<sup>۱۱</sup>، ۲۰۰۵]، ولی این روش مستلزم تعیین دقیق نقاط راهنما<sup>۱۲</sup> بر روی چهره بوده و جهت رفع انسداد بخش‌هایی از چهره که در زاویه اولیه دیده نمی‌شوند، راه‌حلی در نظر نگرفته است. همچنین این روش یک روش تقریبی بوده و در مواردی که تغییرات زاویه بزرگ باشد، روش کارآمدی نمی‌باشد. ذخیره‌سازی تعدادی از تصاویر چهره در زوایای مختلف از دیگر تکنیک‌هایی است که جهت رفع این مشکل در تصاویر دوبعدی در نظر گرفته شده است [لی<sup>۱۳</sup> و همکاران، ۲۰۰۰]. مشکل این روش نیز نیاز به ذخیره‌سازی تعداد بالای تصاویر به ازای هر فرد می‌باشد. در تصاویر سه‌بعدی جهت دستیابی به زاویه‌ی چهره‌ی مورد نظر کافی است تا یک نگاهت از فضای سه‌بعدی به دوبعدی<sup>۱۴</sup> تحت زاویه‌ی مورد نظر صورت پذیرد. نگاهت‌های سه‌بعدی به دوبعدی در فصل ۲ مورد بررسی قرار خواهند گرفت. از دیگر مسائل مربوط به

---

<sup>11</sup> Kim, Kittler

<sup>12</sup> Landmark

<sup>13</sup> Li

<sup>14</sup> Projection

زاویه‌ی چهره در تصاویر دوبعدی، عدم قابلیت تعیین فواصل بین نقاط می‌باشد. فاصله‌ی واقعی بین دو نقطه تحت تأثیر مقیاس تصویر و زاویه‌ی چهره تغییر پیدا می‌کند. در تصاویر سه‌بعدی فواصل بین تمامی نقاط بدون تأثیر پذیری از مقیاس و زاویه‌ی چهره قابل محاسبه هستند. البته در مورد عدم تأثیرپذیری از مقیاس، می‌بایست به تفاوت بین تصاویر سه‌بعدی و مدل‌های سه‌بعدی ساخته‌شده از تصاویر دوبعدی توجه کرد. نمونه‌ای از این مدل‌ها در فصل ۳ بررسی خواهد شد.

از دیگر مزایای تشخیص چهره‌ی سه‌بعدی می‌توان به پراکندگی اطلاعات در تمام بخش‌های چهره اشاره کرد. از آنجایی که شکل چهره در تصاویر دوبعدی مشخص نمی‌باشد، الگوریتم‌های تشخیص چهره‌ی دوبعدی به پردازش بخش‌های دارای کنتراست بالا همچون چشم‌ها، بینی، دهان و مرز چهره می‌پردازند، زیرا تصاویر دوبعدی، حاوی اطلاعات اندکی در مورد بخش‌های دارای کنتراست پایین همچون پیشانی، گونه‌ها و آرواره‌ها می‌باشند [گوردون<sup>۱۵</sup>، ۱۹۹۲]. در مقابل، در تصاویر سه‌بعدی تفاوتی بین بخش‌های دارای کنتراست بالا و پایین از لحاظ اطلاعات مربوط به شکل وجود ندارد.

تشخیص چهره‌ی سه‌بعدی دارای معایبی نیز در مقایسه با تشخیص چهره‌ی دو بعدی می‌باشد. با وجود وابستگی بسیار کمتر تشخیص چهره سه بعدی به مشخصات تابش نور در مرحله تشخیص، این روش در مرحله‌ی تهیه‌ی تصاویر سه‌بعدی مستقل از چگونگی تابش نور نمی‌باشد. نقایص تصاویر سه بعدی عمدتاً شامل زائده‌ها<sup>۱۶</sup> و حفره‌ها<sup>۱۷</sup> می‌شوند. نحوه‌ی تابش نور و چگونگی باز تابش آن تحت شرایط خاص همچون چرب بودن بخش‌هایی از صورت می‌تواند منجر به ایجاد زائده در تصاویر سه‌بعدی شود. نقص لحظه‌ای حسگرهای تهیه‌ی تصاویر سه‌بعدی و یا هر مشکل دیگر در جمع‌آوری اطلاعات نیز به ایجاد حفره در تصاویر

---

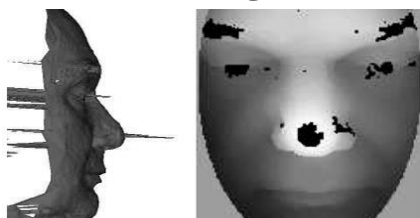
<sup>15</sup> Gordon

<sup>16</sup> Spikes

<sup>17</sup> Holes

منجر می‌شود. شکل ۱-۱ دو نمونه از تصاویر سه‌بعدی دارای نقص را نشان می‌دهد. انواع روش‌های تهیه‌ی

تصاویر سه‌بعدی و شیوه نمایش آنها در فصل ۳ بررسی خواهند شد.



شکل ۱-۱ - دو نمونه از نقایص ممکن در تصاویر سه‌بعدی [بویر<sup>۱۸</sup> و همکاران، ۲۰۰۶]

با وجود کیفیت بسیار بالای تصاویر دوبعدی و پیشرفت روزافزون فناوری تهیه‌ی تصاویر دیجیتال دوبعدی، تصاویر سه‌بعدی از کیفیت به مراتب پایین‌تری برخوردار هستند. همچنین هزینه‌ی تصویربرداری سه‌بعدی نیز به مراتب بیشتر از تصویربرداری دوبعدی می‌باشد. این دو مورد کاربرد فعلی تشخیص چهره‌ی سه‌بعدی را محدود ساخته‌اند. با توجه به این شرایط حتی در صورت ارائه الگوریتم‌های بسیار کارآمد و قابل عرضه به صورت تجاری، امکان ورود تشخیص چهره‌ی سه‌بعدی به عرصه‌ی تجاری بسیار پایین می‌باشد. تجهیزات تصویربرداری سه‌بعدی با توجه به فناوری مورد استفاده، جهت تصویربرداری به زمانی در حدود چند ثانیه تا چند ده ثانیه نیاز دارند. در تمام این مدت شخص مورد نظر می‌بایست بی حرکت در مقابل دستگاه تصویربرداری قرار گیرد. این الزام در تصویربرداری و در نتیجه تشخیص چهره‌ی سه‌بعدی، گویای نیازمندی این روش تشخیص، به همکاری و آگاهی افراد است. عدم برخورداری از خاصیتی که به عنوان یکی از مزایای مهم تشخیص چهره به نسبت سایر روش‌های بیومتریک مطرح می‌شود نیز، از موارد محدودکننده‌ی کاربرد تشخیص چهره‌ی سه‌بعدی می‌باشد.

پس از بررسی مزایا و معایب تشخیص چهره‌ی سه‌بعدی توجه به این نکته حائز اهمیت است که هیچ یک از معایب مذکور، معایبی بنیادین نیستند. بر خلاف معایب تشخیص چهره‌ی دو بعدی که ناشی از

<sup>18</sup> Bowyer

نقص اطلاعات در تصاویر دوبعدی است، مشکلات موجود در تشخیص چهره‌ی سه‌بعدی از نوع مشکلات مربوط به فناوری بوده و با توجه به روند رشد فناوری‌های مختلف، تصور رفع تمامی معایب و نواقص تشخیص چهره‌ی سه‌بعدی در آینده‌ی نزدیک، کاملاً واقع‌گرایانه به نظر می‌رسد. بنابراین چیرگی روش‌های تشخیص چهره‌ی سه‌بعدی بر روش‌های دوبعدی، در آینده امری کاملاً محتمل است.

#### ۴-۱ روش‌های تشخیص چهره‌ی سه‌بعدی

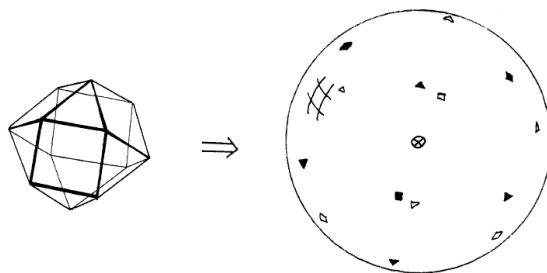
شروع پژوهش‌های مربوط به تشخیص چهره‌ی سه‌بعدی به کمتر از سه دهه‌ی پیش باز می‌گردد. در این قسمت تاریخچه‌ی پژوهش‌های صورت گرفته در این زمینه را مرور خواهیم کرد. روش‌های تقسیم چهره‌ی سه‌بعدی به سه دسته تقسیم می‌شوند. روش‌های مبتنی بر رویه که خود شامل دو دسته‌ی محلی و فراگیر می‌شوند، به همراه روش‌های آماری و روش‌های مبتنی بر مدل، این سه دسته را تشکیل می‌دهند. تاریخچه‌ی روش‌های تشخیص چهره‌ی مبتنی بر رویه و آماری، در این بخش تشریح شده و روش‌های مبتنی بر مدل به علت نیاز به آشنایی با تاریخچه‌ی مدل‌سازی چهره، در بخش ۱-۵ جداگانه بررسی می‌شوند.

##### ۱-۴-۱ روش‌های مبتنی بر رویه

این روش‌ها از مشخصات موجود در هندسه‌ی چهره به عنوان ویژگی‌های تشخیص استفاده می‌کنند. بر حسب این که مشخصات مورد استفاده محلی یا فراگیر هستند، این روش‌ها به دو دسته‌ی محلی و فراگیر تقسیم می‌شوند.

##### - روش‌های مبتنی بر رویه‌ی محلی

یکی از اولین روش‌های مبتنی بر رویه‌ی محلی، روشی است که از نمایش مشخصات چهره در قالب تصاویر گوسی گسترش یافته<sup>۱۹</sup> (EGI) استفاده می‌کند [لی و میلیوس<sup>۲۰</sup>، ۱۹۹۰]. EGI یکی از روشهای نمایش رویه‌های سه‌بعدی است که به علت عدم وابستگی به زاویه‌ی دید نسبت به سایر روش‌های نمایش سه‌بعدی همچون نمایش در دستگاه‌های مختصات متعامد برتری دارد. EGI نگاهی است که رویه‌های محدب و مقعر را به کره‌ی گوسی انتقال می‌دهد. این روش حالت گسترش یافته‌ی تصاویر گوسی است که جهت نمایش چندوجهی‌های محدب مورد استفاده قرار می‌گیرند. در این روش موقعیت هر وجه بر روی کره‌ی گوسی نمایش داده می‌شود. جهت حفظ مقیاس مساحت وجه نیز مشخص می‌شود. باید دقت کرد که نیازی به مشخص کردن شکل وجوه، وجود نداشته و مشخص بودن جهت بردار متعامد<sup>۲۱</sup> وجوه، جهت تعیین چندوجهی محدب کافی است. شکل ۱-۲ نحوه‌ی نمایش یک چندوجهی محدب را بر روی کره گوسی نمایش می‌دهد. نمایش چندوجهی‌های غیرمحدب همچون چهره با استفاده از EGI در مرجع [هورن<sup>۲۲</sup>، ۱۹۸۴] تشریح شده است.



شکل ۱-۲ - نمایش یک چندوجهی محدب با استفاده از EGI [هورن، ۱۹۸۴]

روش دیگر استفاده از میزان خمیدگی‌های<sup>۲۳</sup> محلی رویه‌ی چهره به عنوان ویژگی‌های مورد استفاده در تطبیق رویه‌ها می‌باشد [گوردون، ۱۹۹۲]. خمیدگی بر روی یک خم در فضا به صورت اندازه‌ی مشتق

<sup>19</sup> Extended Gaussian Images

<sup>20</sup> Lee, Milios

<sup>21</sup> Normal Vector

<sup>22</sup> Horn

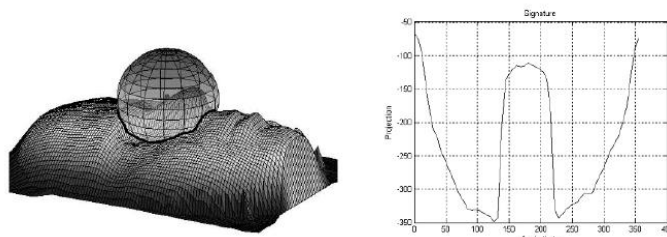
<sup>23</sup> Curvature



بردار مماس بر خم نسبت به فاصله در هر نقطه تعریف می‌شود. برای رویه‌ها در هر نقطه بیشمار میزان خمیدگی قابل تعریف است. مورنو<sup>۲۴</sup> و همکاران، از میزان خمیدگی‌های گوسی و میانگین رویه‌ی چهره در نقاط مهم چون مرکز چشم، خط بینی، نوک بینی، گونه‌ها به عنوان ویژگی استفاده کرده‌اند [مورنو و همکاران، ۲۰۰۳]. خمیدگی‌های گوسی و میانگین، خمیدگی‌های دوبعدی بوده و بر حسب خمیدگی‌های کمینه و بیشینه به صورت زیر تعریف می‌شوند.

$$\kappa_{gaussian} = \kappa_{min} * \kappa_{max} \quad \kappa_{mean} = \frac{\kappa_{min} + \kappa_{max}}{2} \quad (1-1)$$

نسبت فواصل و زوایای بین نقاط ویژگی<sup>۲۵</sup> همچون گوشه‌های چشم، نوک بینی، گوشه‌های لب و انتهای ابروها نیز از ویژگی‌های مورد استفاده در تشخیص چهره می‌باشند [لی<sup>۲۶</sup> و همکاران، ۲۰۰۵]. از دیگر روش‌های تشخیص چهره‌ی سه‌بعدی استفاده از الگوریتم امضای نقاط<sup>۲۷</sup> است [چوا و ژارویس<sup>۲۸</sup>، ۱۹۹۷]. این الگوریتم شیوه‌ای جهت تشریح رویه‌های پیچیده بوده و به هر رویه تعدادی تابع یک‌بعدی نسبت می‌دهد که بیانگر ساختار محلی رویه در همسایگی نقطه‌ی مورد نظر است. امضای نقاط مهم چهره به عنوان ویژگی در تشخیص چهره بکار می‌روند. نحوه محاسبه‌ی امضای یک نقطه در شکل ۳-۱ نشان داده شده است.



شکل ۳-۱ - امضای نقطه واقع بر روی بینی [چوا و ژارویس، ۱۹۹۷]

<sup>24</sup> Moreno

<sup>25</sup> Feature Points

<sup>26</sup> Lee

<sup>27</sup> Points Signatures Algorithm

<sup>28</sup> Chua, Jarvis

به ازای نقطه  $a$  بر روی رویه، یک کره به مرکز  $a$  و شعاع  $R$  در نظر گرفته می‌شود. محل برخورد کره با سطح رویه، منحنی فضایی  $C$  را ایجاد می‌کند. یک صفحه‌ی  $P$  بر روی منحنی منطبق می‌گردد. صفحه‌ی  $P$  در جهت بردار نرمال  $(n1)$  به نقطه  $a$  انتقال می‌یابد و صفحه‌ی  $P'$  را تشکیل می‌دهد. تصویر منحنی فضایی  $C$  بر روی صفحه‌ی  $P'$ ، منحنی مسطح  $C'$  را ایجاد می‌کند. هر نقطه روی  $C'$  با یک فاصله‌ی جهت دار تا نقطه متناظرش بر روی  $C$  مشخص می‌شود. نقطه‌ی  $a$  بر روی  $C'$  که بیشترین فاصله را با نقطه‌ی متناظرش بر روی  $C$  دارد، با نقطه‌ی  $a$  بردار مرجع  $n2$  را تشکیل می‌دهند. حال هر نقطه  $C'$  دارای زاویه‌ی  $\theta$  با این نقطه‌ی مرجع می‌باشد. فاصله جهت دار هر نقطه روی  $C'$  بر حسب زاویه‌ی  $\theta$  تابع  $d(\theta)$  را ایجاد می‌کند. این تابع امضای نقطه‌ی  $a$  می‌باشد. امضای نقطه در واقع دنباله‌ای از فواصل جهت‌دار می‌باشد. جهت مقایسه‌ی دو امضای نقاط متناظر در دو رویه‌ی چهره‌ی متفاوت از معیار فاصله زیر استفاده شده است.

$$D = \min \left\{ \sqrt{\sum_{i=1}^N (d_s(\theta_i - \phi) - d_m(\theta_i))^2} \right\} \quad (2-1)$$

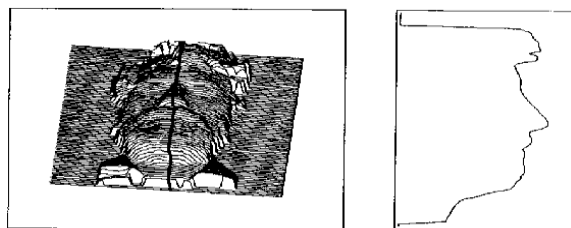
قابل توجه است که الگوریتم امضای نقطه مستقل از چرخش حول محورهای  $x$  و  $y$  بوده و تنها به چرخش حول محور  $z$  وابسته است که توسط بکارگیری  $\phi$  در رابطه‌ی فوق این وابستگی نیز از بین می‌رود. جهت حذف تغییرپذیری نسبت به تغییرات حالت چهره، امضای نقاط صلب چهره محاسبه و به عنوان ویژگی در نظر گرفته می‌شود.

ترکیب تکنیک‌های محلی و فراگیر نیز از روش‌های تشخیص چهره‌ی سه‌بعدی می‌باشد. در یکی از این روش‌ها اطلاعات هندسی محلی به همراه نقاط تشکیل‌دهنده‌ی رویه، بردار ویژگی اولیه را ایجاد کرده که

پس از کاهش ابعاد توسط الگوریتم PCA، جهت مقایسه و تشخیص چهره استفاده شده است [ز و<sup>۲۹</sup> و همکاران، ۲۰۰۴].

### - روش‌های مبتنی بر رویه‌ی فراگیر

روش‌های مبتنی بر رویه‌ی فراگیر، کل رویه‌ی چهره را به عنوان ورودی سیستم تشخیص چهره در نظر می‌گیرند. قدیمی‌ترین روش تشخیص چهره‌ی سه‌بعدی که در سال ۱۹۸۹ ارائه شده است، در زمره‌ی این روش‌ها قرار می‌گیرد. همان‌طور که در شکل ۴-۱ نشان داده شده است این الگوریتم با استفاده از تقارن چهره، به همترازسازی<sup>۳۰</sup> رویه‌های چهره پرداخته و پروفایل ناشی از برخورد صفحه‌ی تقارن و رویه‌ی چهره را به عنوان ویژگی استخراج می‌کند [همکاران، ۱۹۸۹].



شکل ۴-۱ - صفحه تقارن چهره و ویژگی استخراج شده [کارتوکس و همکاران، ۱۹۸۹]

یکی دیگر از روش‌های فراگیر مبتنی بر رویه استفاده از خمیدگی‌های کمینه و بیشینه‌ی رویه‌ی چهره و جهت آنها به عنوان ویژگی‌های چهره است. در این الگوریتم برای غلبه بر مشکلات ناشی از تغییر زوایای خمیدگی‌های کمینه و بیشینه با تغییر زاویه تصویر برداری، از EGI استفاده شده است [تاناکا<sup>۳۲</sup> و همکاران، ۱۹۹۸]. بردارهای متعامد رویه‌ی چهره به همراه EGI نیز جهت تشخیص چهره‌ی سه‌بعدی بکار گرفته شده‌اند [ونگ<sup>۳۳</sup> و همکاران، ۲۰۰۴].

<sup>29</sup> Xu  
<sup>30</sup> Alignment  
<sup>31</sup> Cartoux  
<sup>32</sup> Tanaka  
<sup>33</sup> Wong

روش‌های مبتنی بر فاصله نیز در تشخیص چهره‌ی سه‌بعدی استفاده شده‌اند. یکی از این روش‌ها، محاسبه‌ی شباهت بین رویه‌های چهره با استفاده از فاصله‌ی هاسدورف<sup>۳۴</sup> [اکرمان و بونک<sup>۳۵</sup>، ۲۰۰۰ - پن<sup>۳۶</sup> و همکاران، ۲۰۰۳] و یا نگارش‌های تغییر یافته‌ی آن است [الی و شیم<sup>۳۷</sup>، ۲۰۰۴ - راس<sup>۳۸</sup> و همکاران، ۲۰۰۵]. مزیت قابل توجه فاصله‌ی هاسدورف این است که این معیار فاصله برای رویه‌های متشابه با همپوشانی کم نیز مقدار کوچکی است و نیازی به همترازسازی رویه‌ها وجود ندارد. الگوریتم تکراری نزدیکترین نقطه<sup>۳۹</sup> (ICP) نیز جهت همترازسازی رویه‌های چهره بکار گرفته شده است. در این الگوریتم دو رویه بر روی هم منطبق شده و از فاصله باقی مانده بین دو رویه به عنوان معیار شباهت استفاده می‌شود [مدیونی و واپوتیتش<sup>۴۰</sup>، ۲۰۰۳]. بافت چهره نیز علاوه بر هندسه‌ی رویه‌ی چهره می‌تواند در تعیین فاصله بین دو رویه جهت همترازسازی و انطباق استفاده شود [پاپاتئودورو و روکرت<sup>۴۱</sup>، ۲۰۰۴]. ترکیب سایر مشخصات چهره همچون بردارهای متعامد و هندسه‌ی رویه‌ی چهره نیز در تعیین فاصله‌ی بین دو رویه و انطباق آنها استفاده شده است [لو<sup>۴۲</sup> و همکاران، ۲۰۰۴]. فرض صلب بودن رویه‌ی چهره و در نظر نگرفتن تغییر حالت چهره و استفاده از الگوریتم‌های انطباق صلب همچون ICP از نقایص این روش‌ها محسوب می‌شود. این ایراد با بکارگیری روش‌های انطباق غیرصلب همچون TPS<sup>۴۳</sup> تا حد زیادی رفع شده است [لو و جین<sup>۴۴</sup>، ۲۰۰۵].

---

<sup>34</sup> Hausdorff Distance

<sup>35</sup> Ackermann, Bunke

<sup>36</sup> Pan

<sup>37</sup> Shim

<sup>38</sup> Russ

<sup>39</sup> Iterative Closest Point

<sup>40</sup> Medioni, Waupotitsch

<sup>41</sup> Papatheodorou, Rueckert

<sup>42</sup> Lu

<sup>43</sup> Thin Plate Spline

<sup>44</sup> Jain

بخش دیگری از روش‌های فراگیر مبتنی بر رویه، روش‌هایی هستند که با استفاده از پروفایل‌ها و منحنی‌های بسته‌ی<sup>۴۵</sup> رویه به انطباق چهره و محاسبه‌ی ویژگی‌ها می‌پردازند [ناگامین<sup>۴۶</sup> و همکاران، ۱۹۹۲ – بومیر و آکروی<sup>۴۷</sup>، ۲۰۰۰ – وو<sup>۴۸</sup> و همکاران، ۲۰۰۳]. پروفایل‌ها و منحنی‌های بسته‌ی بافت نیز جهت انطباق چهره و محاسبه‌ی ویژگی‌ها بکار گرفته شده‌اند [بومیر و آکروی، ۲۰۰۱].

#### ۲-۴-۱ روش‌های آماری

روش‌های آماری همچون تحلیل جزء اصلی<sup>۴۹</sup> (PCA)، تحلیل جداساز خطی<sup>۵۰</sup> (LDA) و تحلیل جزء مستقل<sup>۵۱</sup> (ICA) در تشخیص چهره‌ی دوبعدی بسیار مورد استفاده قرار گرفته‌اند. روش‌هایی بر اساس PCA نیز در تشخیص چهره سه‌بعدی استفاده شده‌اند. مرجع [سالاکانیدو<sup>۵۲</sup> و همکاران، ۲۰۰۳] استفاده از رنگ و عمق و یا ترکیب این دو را در محاسبه‌ی PCA پیشنهاد کرده است. مرجع [سالاکانیدو<sup>۵۳</sup> و همکاران، ۲۰۰۴] ترکیب این روش‌های مبتنی بر PCA با روش‌های طبقه‌بندی همچون مدل پنهان مارکوف<sup>۵۴</sup> (HMM) را معرفی کرده است. استفاده از LDA [گوکبرک<sup>۵۵</sup> و همکاران، ۲۰۰۵] و ICA [سریواستوا<sup>۵۶</sup> و همکاران، ۲۰۰۳] نیز جهت تشخیص و تحلیل چهره‌ی سه‌بعدی گزارش شده‌اند.

---

<sup>45</sup> Contour

<sup>46</sup> Nagamine

<sup>47</sup> Beumier, Achery

<sup>48</sup> Wu

<sup>49</sup> Principal Component Analysis

<sup>50</sup> Linear Discriminant Analysis

<sup>51</sup> Independent Component Analysis

<sup>52</sup>Tsalakanidou

<sup>53</sup>Tsalakanidou

<sup>54</sup> Hidden Markov Model

<sup>55</sup> Gokberk

<sup>56</sup> Srivastava

هر یک از روش‌های PCA، LDA و ICA یک نگاهت خطی از فضای داده‌ها به یک زیرفضا می‌باشند. عملکرد این روش‌ها با توجه به ابعاد فضا و مشخصات داده‌ها همچون میزان پراکندگی آنها متغیر بوده و تعیین این که کدام یک عملکرد بهتری دارد مشکل می‌باشد. مرجع [لیو<sup>۵۷</sup> و همکاران، ۲۰۰۴] روشی را جهت پیدا کردن نگاهت بهینه به جای استفاده از یک نگاهت ثابت در تصاویر دوبعدی پیشنهاد داده است. این روش در [سریواستاوا و همکاران، ۲۰۰۳] بر روی تصاویر چهره‌ی سه‌بعدی پیاده‌سازی شده است. اساس عملکرد این روش یافتن زیرفضایی بهینه از فضای نمونه‌ها به کمک روش‌های بهینه‌سازی است. بکارگیری این روش و مقایسه‌ی آن با نتایج روش‌های مبتنی بر نگاهت‌های ثابت، بهبود قابل توجهی را در نرخ تشخیص نشان می‌دهد. با این حال این الگوریتم به علت استفاده از روش‌های بهینه‌سازی و نداشتن جواب تحلیلی مشخص، نیاز به حجم محاسباتی بسیار بالایی داشته و از سرعت مناسبی در مقایسه با نگاهت‌های ثابت برخوردار نیست.

مزیت قابل توجه روش‌های آماری کاهش ابعاد داده‌ها می‌باشد. در مقابل، این روش‌ها به علت اعمال بر تمام اطلاعات رویه‌ی چهره و در نظر گرفتن هر چهره به صورت یک بردار در فضا، به تغییراتی همچون تغییر حالت چهره وابسته بوده و در این موارد عملکرد مناسبی نشان نمی‌دهند. علاوه بر این، این روش‌ها در صورت وجود نمونه‌های دارای انحراف زیاد نیز عملکرد مناسبی ندارند.

## ۵-۱ روش‌های مدل‌سازی سه‌بعدی چهره و تشخیص مبتنی بر مدل

از آغاز پژوهش‌های مربوط به مدل‌سازی سه‌بعدی چهره، این دانش بیشتر در کاربرد سنتز و پویانمایی<sup>۵۸</sup> کامپیوتری چهره مورد استفاده قرار گرفته و کمتر در تشخیص چهره کاربرد داشته است. به

---

<sup>57</sup> Liu

<sup>58</sup> Animation

عبارت دیگر به جز چند کاری که در چند سال اخیر در زمینه‌ی تشخیص چهره‌ی سه‌بعدی مبتنی بر مدل صورت گرفته است، پژوهش دیگری در این زمینه گزارش نشده است. بنابراین جهت آشنایی با انواع روش‌های مدل‌سازی چهره می‌بایست تاریخچه‌ی پویانمایی کامپیوتری سه‌بعدی چهره را مورد توجه قرار دهیم. پیشینه‌ی مدل‌سازی و پویانمایی سه‌بعدی چهره به بیش از سه دهه‌ی پیش باز می‌گردد. با این وجود مدل‌های واقعی‌تر چهره به علت افزایش قابل توجه سرعت و توان محاسباتی کامپیوترها در ۱۵ سال اخیر ارائه شده‌اند. البته هنوز راه زیادی جهت دستیابی به مدلی واقعی برای چهره که از کنترل‌پذیری کافی برخوردار باشد، در پیش است. علت پیچیدگی ساخت یک مدل چهره‌ی واقعی، حساسیت بسیار زیاد چهره می‌باشد. هر تغییر کوچک در مختصات و مشخصات چهره بیانگر حالتی در چهره می‌باشد که ساخت مدل‌های نزدیک به واقعیت را مشکل می‌سازد.

اولین مدل‌سازی سه‌بعدی چهره در اوایل دهه‌ی ۷۰ توسط پارک<sup>۵۹</sup> ارائه شد [پارک، ۱۹۷۲]. او چهره‌ی سه‌بعدی را توسط چند ضلعی‌ها مدل‌سازی کرد و موفق به ساخت مدلی از چهره شد که قابلیت باز و بسته کردن چشم‌ها و دهان را داشت. در همین زمان گوراود<sup>۶۰</sup> با ارائه‌ی الگوریتم سایه‌گذاری نرم خود (سایه‌گذاری گوراود<sup>۶۱</sup>) و پیاده‌سازی آن بر روی مدل چهره‌ی پارک بهبود قابل توجهی در این مدل به علت تأثیر نور و سایه‌ها ایجاد کرد. ما مدل سایه‌گذاری گوراود را در فصل ۲ بررسی خواهیم کرد. در ۱۹۷۴ پارک مدل پارامتری خود را تکمیل کرد [پارک، ۱۹۷۴]. در ۱۹۸۰ پلات<sup>۶۲</sup> یک مدل فیزیکی ماهیچه‌ای برای چهره ارائه کرد [پلات، ۱۹۸۰]. مدل او قابلیت ایجاد حالات مختلف در چهره را توسط کنترل ماهیچه‌های صورت دارا بود. در ۱۹۸۷ واترز<sup>۶۳</sup> مدل چهره‌ی ماهیچه‌ای جدیدی را ارائه کرد [واترز، ۱۹۸۷]. مدل او به مراتب

---

<sup>59</sup> Parke

<sup>60</sup> Gouraud

<sup>61</sup> Gouraud Shading

<sup>62</sup> Platt

<sup>63</sup> Waters

دقیق‌تر، پیچیده‌تر و کامل‌تر از مدل ارائه شده توسط پلات بود. واترز در مدل خود حالات چهره را به هندسه‌ی صورت و ماهیچه‌ها مرتبط ساخته بود. ۱۹۹۳ مصادف با ساخت اولین پویشگر لیزری تجاری تحت عنوان Cyberware™ بود که تحولی در ساخت مدل‌های سه‌بعدی چهره بوجود آورد. در ۱۹۹۸ پیگهین<sup>۶۴</sup> روشی را معرفی کرد که در آن با استفاده از تصاویر دو بعدی، مدل چهره‌ی سه‌بعدی را ایجاد کرده و به میان‌سازی<sup>۶۵</sup> حالت‌های مختلف چهره پرداخت [پیگهین و همکاران، ۱۹۹۸]. این روش در فصل ۳ مورد بررسی قرار خواهد گرفت.

در سال ۱۹۹۸ وتر و بلانز<sup>۶۶</sup> اولین بار از مدل میان‌سازی<sup>۶۷</sup> در تشخیص چهره استفاده کردند. این مدل مبتنی بر ترکیب خطی مدل‌های چهره‌ی سه‌بعدی تناظر یافته، می‌باشد [وتر و بلانز، ۱۹۹۸]. در سال‌های ۱۹۹۹ و ۲۰۰۳، بلانز و وتر به تکمیل مدل مذکور پرداخته و از این مدل در تشخیص چهره استفاده کردند [بلانز و وتر، ۱۹۹۹- بلانز و وتر، ۲۰۰۳]. انصاری و عبدالمطلب<sup>۶۸</sup> نیز در سال ۲۰۰۳ با استفاده از یک مدل چهره‌ی سه‌بعدی عمومی<sup>۶۹</sup> و تغییر آن بر اساس موقعیت نقاط راهنمای مشخص شده در دو تصویر تمام‌رخ و نیم‌رخ، مدل چهره سه بعدی مربوط به هر فرد را ایجاد کردند [انصاری و عبدالمطلب، (الف) ۲۰۰۳]. با استفاده از پارامترهای تغییر دهنده‌ی مدل عمومی، تشخیص چهره توسط این روش نیز صورت پذیرفت [انصاری و عبدالمطلب، (ب) ۲۰۰۳].

در این بخش به بررسی اجمالی روش‌های مدل‌سازی سه‌بعدی ماهیچه‌ای چهره می‌پردازیم. به عنوان نمونه‌ای از مدل‌سازی چهره‌ی مبتنی بر مدل عمومی، روش ارائه شده توسط پیگهین را در فصل ۳ بررسی

---

<sup>64</sup> Pighin

<sup>65</sup> Morphing

<sup>66</sup> Vetter, Blanz

<sup>67</sup> Morphable Model

<sup>68</sup> Ansari, Abdel-mottaleb

<sup>69</sup> Generic Model



خواهیم کرد. مدل میان‌ساز نیز که بخش اساسی این پایان‌نامه را تشکیل می‌دهد در فصل ۴ و مطالب مربوط به آن در سایر فصول بررسی خواهند شد.

#### ۱-۵-۱ مدل‌های چهره‌ی ماهیچه‌ای

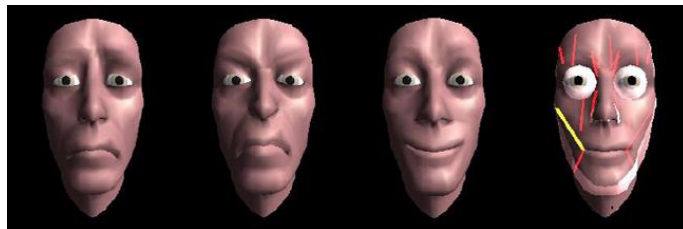
آناتومی کامل سر و صورت شامل ترکیب پیچیده‌ای از استخوان‌ها، ماهیچه‌ها، رگها، بافت‌های پیوندی، پوست و مو می‌شود. تاکنون هیچ مدل چهره‌ی فیزیکی‌ای بر اساس آناتومی کامل صورت ارائه نشده است. با این وجود چند مدل فیزیکی بر اساس ساختارهای ساده شده‌ی استخوان‌ها، ماهیچه‌ها، بافت‌های پیوندی و پوست صورت، گزارش شده‌اند. مدل ارائه شده توسط پلات [پلات، ۱۹۸۰] یک مدل دینامیکی است که همانند مدل اولیه‌ی پارک از چندضلعی‌ها جهت نمایش رویه‌ی چهره استفاده کرده است. در این مدل فرم‌هایی بین برخی از رؤس چندضلعی‌ها به صورت کشسانی متصل شده‌اند که رفتار ماهیچه‌های صورت را شبیه‌سازی می‌کنند. مدل ارائه شده توسط واترز نیز یک مدل دینامیکی چهره است که شامل دو دسته ماهیچه‌ی خطی<sup>۷۰</sup> و مدور<sup>۷۱</sup> می‌شود. ماهیچه‌های خطی ماهیچه‌هایی هستند که به صورت خطی قرار گرفته و در راستای قرارگیری ماهیچه کشیده و باز می‌شوند. ماهیچه‌های مدور ماهیچه‌هایی هستند که به صورت دایره‌ای منقبض و منبسط می‌شوند. برای هر ماهیچه یک ناحیه‌ی تأثیرگذاری تعریف می‌شود. ناحیه تأثیرگذاری ماهیچه‌های خطی در اطراف نقاط اتصال بوده و در ماهیچه‌های مدور در اطراف ماهیچه می‌باشد. علاوه بر این واترز ماهیچه‌هایی را که جهت ایجاد یک حالت خاص همچون لبخند در چهره می‌بایست تحریک شوند، مشخص کرد و این ماهیچه‌ها و میزان تحریک آنها را به این حالت چهره انتساب داد. مدل‌های دیگری که در آن‌ها بافت چهره توسط یک ساختار شبکه‌ای سه لایه شبیه‌سازی می‌شوند، به عنوان مدل‌های

---

<sup>70</sup> Linear Muscle

<sup>71</sup> Sphincter Muscle

گسترش یافته‌ی مدل واترز در نظر گرفته می‌شوند [ترزوپولوس<sup>۷۲</sup> و واترز، ۱۹۹۰]. شکل ۱-۵ پیاده‌سازی حالات مختلف چهره را توسط مدل واترز نمایش می‌دهد. به علت پیچیدگی رفتار ماهیچه‌ها و تعداد زیاد آنها، نمایش و پیاده‌سازی یک چهره واقعی توسط مدل ماهیچه‌ای بسیار مشکل و پرهزینه می‌باشد. به همین دلیل از این مدل تنها در ساخت شخصیت‌های نسبتاً ساده‌ی سینمای متحرک‌سازی استفاده می‌شود. البته همانطور که ملاحظه می‌شود قابلیت پیاده‌سازی حالات چهره با دقت بالا در این شخصیت‌ها از مزایای مدل‌سازی سه‌بعدی چهره‌ی ماهیچه‌ای می‌باشد.



شکل ۱-۵ - سنتز حالات چهره بر روی مدل واترز [ترزوپولوس<sup>۷۳</sup> و واترز، ۱۹۹۰]

## ۱-۶ مزایای روش‌های تشخیص چهره‌ی سه‌بعدی مبتنی بر مدل

از مهمترین مزایای روش‌های تشخیص چهره‌ی سه‌بعدی مبتنی بر مدل می‌توان به قابلیت کاربرد این روش‌ها در سنتز چهره و حالات چهره اشاره کرد. از آنجایی که در این روش‌ها یک مدل سه‌بعدی تهیه می‌شود که قابلیت پیاده‌سازی حالات مختلف چهره را دارد، تشخیص چهره توسط این مدل‌ها تا حد قابل توجهی مستقل از حالت چهره می‌باشد.

علاوه بر این با توجه به کاهش پارامترهای تعیین‌کننده‌ی مدل نسبت به حجم اطلاعات چهره، می‌توان از مدل‌سازی چهره در کاربردهای فشرده‌سازی چهره و کدینگ استفاده کرد. استفاده از سنتز چهره

<sup>72</sup> Terzopoulos

<sup>73</sup> Terzopoulos

با استفاده از مدل‌های سه‌بعدی در بازی‌های رایانه‌ای و مواردی از این دست نیز از کاربردهای سنتی و مرسوم آن به شمار می‌رود.

مزیت بسیار مهم تشخیص چهره با استفاده از مدل‌های سه‌بعدی این است که تصاویر ورودی جهت تشخیص، دو بعدی می‌باشند و نیازی به تصاویر سه‌بعدی جهت تشخیص وجود ندارد. در واقع تشخیص چهره‌ی مبتنی بر مدل نیازی به همکاری و آگاهی شخص و تجهیزات گران قیمت تصویربرداری سه‌بعدی نداشته و در عین برخورداری از تمام مزایای تشخیص چهره سه‌بعدی همچون کاهش وابستگی به جهت و شدت تابش نور و زاویه دید، هیچ یک از معایب آن را شامل نمی‌شود.

از بین تمام روش‌های مدل‌سازی چهره تنها دو روش مدل عمومی آنصاری و عبدالمطلب، (ب) [۲۰۰۳] و مدل میان‌ساز [بلانز و وتر، ۲۰۰۳] به همراه چند برداشت تغییر یافته از این دو روش، در کاربرد تشخیص چهره استفاده شده‌اند. علت این امر این است که سایر مدل‌ها یا بسیار ابتدایی بوده و قابلیت سنتز چهره‌های واقعی و در نتیجه تشخیص چهره را نداشته و یا پیچیدگی و حجم بالای پارامترهای تعریف کننده‌ی مدل‌ها، مانع از استفاده‌ی آنها به عنوان روش‌های تشخیص چهره شده است. ما یک نمونه از مدل عمومی را جهت مدل‌سازی چهره در فصل ۳ و مدل میان‌ساز را در فصل ۴ بررسی خواهیم کرد.

#### ۱-۶-۱ مزایای تشخیص چهره‌ی مبتنی بر مدل میان‌ساز

تشخیص چهره‌ی مبتنی بر مدل میان‌ساز نسبت به مدل عمومی دارای چند مزیت اساسی می‌باشد. مزیت اول این است که این روش بر خلاف مدل عمومی، روشی کاملاً خودکار بوده و در هیچ یک از مراحل تشخیص نیازی به دخالت کاربر ندارد. مزیت دیگر این روش این است که تنها به یک تصویر ورودی نیاز داشته و بدون تأثیرپذیری از شرایط مختلف نورپردازی و زاویه‌ی تصویربرداری به تشخیص چهره می‌پردازد.

علاوه بر این پس از ساخت مدل میان‌ساز، سنتز هر چهره، ساخت مدل مربوطه و تشخیص آن در زمان بسیار کمتری نسبت به سایر روش‌های مدل‌سازی چهره صورت می‌گیرد. نرخ بالای تشخیص توسط این روش [بلانز و وتر، ۲۰۰۳] نیز بیانگر قابلیت پیاده‌سازی این روش به صورت یک روش تجاری تشخیص چهره می‌باشد.

در مقابل، روش تشخیص چهره‌ی مبتنی بر مدل عمومی نیازمند به دخالت مستقیم کاربر جهت تعیین نقاط راهنما بوده و حداقل به دو تصویر متعامد از چهره به عنوان تصاویر ورودی جهت تشخیص چهره نیاز دارد. در این روش تمام مراحل ساخت مدل در زمان تشخیص چهره صورت می‌گیرد که تشخیص چهره توسط این روش را به الگوریتمی کند تبدیل می‌کند. دلایل فوق، علل انتخاب روش تشخیص چهره‌ی مبتنی بر مدل میان‌ساز را جهت بررسی و ارائه در این پایان‌نامه بیان می‌کنند.

## ۷-۱ مقدمه

جهت بررسی و پیاده‌سازی الگوریتم‌های سه‌بعدی، آشنایی با ارتباط تصویری سه‌بعدی کامپیوتری<sup>۷۴</sup> شامل مفاهیمی چون تبدیلات هندسی سه‌بعدی، دید سه‌بعدی، تعیین سطوح قابل رویت و مدل‌های نوردهی

---

<sup>74</sup> 3D Computer Graphics

الزامی است. در کاربرد مدل سازی سه بعدی چهره نیز این مفاهیم در بخش های مختلف مورد استفاده واقع می شود. بنابراین در این فصل به بررسی اجمالی این بخش از ارتباط تصویری کامپیوتری می پردازیم. [فولی<sup>75</sup> و همکاران، ۱۹۹۶] مرجع مورد استفاده جهت مطالب این فصل می باشد. تشریح کامل و گسترده تری از مفاهیم مذکور در [فولی و همکاران، ۱۹۹۶] ارائه شده است.

## ۸-۱ تبدیلات هندسی سه بعدی

تبدیلات هندسی سه بعدی گسترشی بر تبدیلات هندسی دو بعدی هستند که در آن ها مؤلفه ی z نیز در نظر گرفته شده است. در این بخش سه تبدیل اساسی انتقال، دوران و تغییر مقیاس و نحوه ی ترکیب تبدیلات فوق را بررسی می کنیم.

### ۱-۸-۱ انتقال

در یک دستگاه مختصات متعامد سه بعدی، هر نقطه ی  $P=(x, y, z)$  به نقطه  $P'=(x', y', z')$  توسط

ماتریس زیر انتقال می یابد.

$$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & t_x \\ 0 & 1 & 0 & t_y \\ 0 & 0 & 1 & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (1-2)$$

$$P' = T.P \quad (2-2)$$

ماتریس T در رابطه ی (۲-۲) ماتریس انتقال نامیده می شود. مقادیر  $t_x, t_y, t_z$  بیانگر میزان انتقال در مولفه های x, y و z هستند. جهت انتقال یک شیء در فضای سه بعدی می بایست تمام نقاط تشکیل دهنده ی شیء را با استفاده از ماتریس فوق انتقال دهیم. جهت انتقال چندضلعی ها، روش مرسوم انتقال رئوس و

<sup>75</sup> Foley

ترسیم مجدد چندضلعی است. معکوس ماتریس انتقال و در نتیجه معکوس عمل انتقال با قرینه کردن مقادیر  $t_x, t_y, t_z$  بدست می آید که معادل انتقالی به میزان برابر و در جهت عکس انتقال اولیه است.

### ۱-۸-۲ دوران

برای ایجاد یک دوران به محور دوران و زاویه‌ی دوران نیاز داریم. ساده‌ترین حالت، دوران حول محورهای مختصات می‌باشد. با ترکیب انتقال و دوران حول محورها، می‌توان دوران حول هر محور دلخواه را محاسبه کرد. در محاسبه‌ی دوران بنابر قرارداد، زاویه را کمیتی جهت‌دار و جهت مثبت آن را هنگامی که از سمت مثبت محور مختصات به مرکز نگاه می‌کنیم به صورت پادساعتگرد در نظر می‌گیریم.

#### دوران حول محورهای مختصات :

معادلات دوران حول مبداء در فضای دوبعدی به سادگی به معادلات دوران حول محور  $z$  در فضای

سه‌بعدی تبدیل می‌شوند.

$$\begin{aligned} x' &= x \cos \theta - y \sin \theta \\ y' &= x \sin \theta + y \cos \theta \\ z' &= z \end{aligned} \quad (۳-۲)$$

که در آن  $\theta$  زاویه‌ی دوران را مشخص می‌کند. نمایش ماتریسی رابطه‌ی فوق به صورت زیر است.

$$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} \cos \theta & -\sin \theta & 0 & 0 \\ \sin \theta & \cos \theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (۴-۲)$$

$$P' = R_z(\theta) \cdot P \quad (۵-۲)$$

معادلات مربوط به دوران حول سایر محورهای مختصات با استفاده از جایگزینی چرخشی پارامترهای  $x, y$  و

$z$  در معادله‌ی (۳-۲) بدست می‌آید:

$$x \rightarrow y \rightarrow z \rightarrow x \quad (۶-۲)$$

در نتیجه روابط مربوط به دوران حول محور  $x$  به صورت زیر بدست می‌آیند.

$$\begin{aligned} y' &= y \cos \theta - z \sin \theta \\ z' &= y \sin \theta + z \cos \theta \\ x' &= x \end{aligned} \quad (7-2)$$

و نمایش ماتریسی آن به صورت زیر می‌باشد.

$$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta & 0 \\ 0 & \sin \theta & \cos \theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (8-2)$$

$$P' = R_x(\theta) \cdot P \quad (9-2)$$

با جایگزینی چرخشی پارامترها طبق رابطه‌ی (۲-۶)، ماتریس دوران حول محور  $y$  به صورت زیر بدست می‌آید:

$$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} \cos \theta & 0 & \sin \theta & 0 \\ 0 & 1 & 0 & 0 \\ -\sin \theta & 0 & \cos \theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (10-2)$$

$$P' = R_y(\theta) \cdot P \quad (11-2)$$

ماتریس‌های  $R_x(\theta)$ ،  $R_y(\theta)$  و  $R_z(\theta)$  ماتریس دوران نامیده می‌شوند. ماتریس دوران معکوس با تبدیل

زاویه‌ی دوران  $\theta$  به  $-\theta$  در ماتریس‌های دوران بدست می‌آید. از آنجایی که با این تغییر، تنها مقادیر  $\sin \theta$

قرینه می‌شوند، ماتریس دوران معکوس برابر با ترانزپوز ماتریس دوران است.

$$R^{-1}(\theta) = R^T(\theta) \quad (12-2)$$

رابطه‌ی فوق در مورد دوران حول تمام محورهای مختصات و هر محور دلخواه صادق است.

دوران سه‌بعدی تعمیم یافته:

ماتریس دوران حول هر محور دلخواه، توسط یک تبدیل مرکب شامل ماتریس‌های انتقال و ماتریس دوران حول محورهای مختصات بدست می‌آید. این ماتریس مرکب در سه مرحله قابل محاسبه است. ابتدا محور دوران را با استفاده از تبدیلات هندسی سه‌بعدی به یکی از محورهای مختصات منتقل می‌کنیم. سپس نقطه P را حول محور مختصات دوران می‌دهیم. در آخر نیز معکوس تبدیلات هندسی را جهت بازگرداندن محور دوران به محل ابتدایی خود اجرا می‌کنیم. در حالت خاص هنگامی که محور دوران موازی یکی از محورهای مختصات باشد، سه مرحله‌ی فوق به شرح زیر خواهند بود.

۱- انتقال نقطه‌ی P به نحوی که محور دوران بر محور مختصات موازی منطبق شود.

۲- اجرای دوران به اندازه‌ی زاویه‌ی مورد نظر حول محور

۳- انتقال نقطه‌ی P به نحوی که محور دوران به محل اصلی خود بازگردد.

شکل ۱-۲ مراحل انجام این دوران را نشان می‌دهد. هر نقطه‌ی P روی شیء به نقطه P' تحت تبدیل زیر نگاشته می‌شود.

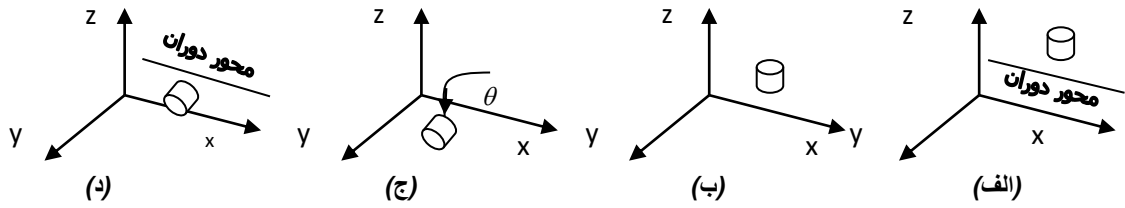
$$P' = T^{-1} \cdot R_x(\theta) \cdot T \cdot P \quad (۱۳-۲)$$

بنابراین ماتریس دوران مرکب برای تبدیل هندسی مورد نظر برابر است با:

$$R(\theta) = T^{-1} \cdot R_x(\theta) \cdot T \quad (۱۴-۲)$$

ماتریس فوق معادل با ماتریس دوران دوبعدی حول هر نقطه‌ی دلخواه می‌باشد.





شکل ۱-۲ - مراحل اجرای دوران حول محور دوران موازی یک محور مختصات

اگر محور دوران مورد نظر موازی هیچ یک از محورهای مختصات نباشد، جهت بدست آوردن ماتریس دوران به چند تبدیل دیگر نیز نیاز داریم. با داشتن مشخصات محور و زاویه ی دوران، پیاده سازی دوران مورد نظر در پنج مرحله صورت می گیرد.

۱- انتقال شیء به گونه ای که محور دوران از مبدا مختصات بگذرد.

۲- دوران شیء به نحوی که محور دوران انتقال یافته، بر یکی از محورهای مختصات منطبق شود.

این مرحله خود نیازمند حداکثر ۲ بار پیاده سازی دوران است.

۳- انجام دوران مورد نظر حول محور مختصات

۴- اعمال دوران معکوس به نحوی که محور دوران به محل قبل از مرحله ی ۲ باز گردد.

۵- اعمال انتقال معکوس به گونه ای که محور دوران به محل اولیه ی خود منتقل شود.

برای بدست آوردن ماتریس دوران فرض می کنیم در مرحله ی ۲، محور دوران بر محور  $Z$  منطبق شود. شکل

۲-۲ پنج مرحله ی مذکور، جهت اعمال دوران مورد نظر را نشان می دهد.

فرض می‌کنیم محور دوران توسط دو نقطه  $P_1$  و  $P_2$  مشخص شده است. همچنین فرض می‌کنیم

در صورت مثبت بودن زاویه  $\theta$ ، هنگامی که از  $P_2$  به  $P_1$  نگاه می‌کنیم جهت دوران مورد نظر پادساعتگرد

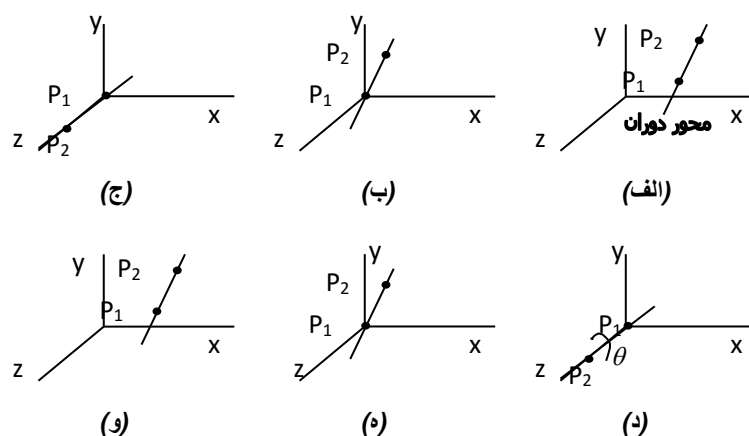
می‌باشد (شکل ۲-۲). بردار دوران و بردار دوران واحد به ترتیب به صورت زیر تعریف می‌شوند.

$$V = P_2 - P_1 = (x_2 - x_1, y_2 - y_1, z_2 - z_1) \quad (۱۵-۲)$$

$$U = \frac{V}{|V|} = (a, b, c) \quad (۱۶-۲)$$

اگر فرض دوم ما در مورد پادساعتگرد بودن جهت دوران برقرار نباشد، کفایت در روابط فوق مقادیر بردار

دوران و بردار واحد دوران را قرینه در نظر بگیریم.



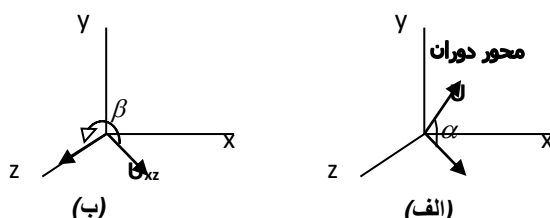
شکل ۲-۲ - مراحل دوران سه بعدی، (الف): موقعیت اولیه، (ب) تا (و): مراحل ۱ تا ۵

جهت انتقال محور دوران به مبدا مختصات کفایت نقطه  $P_1$  را به مبدا منتقل کنیم. ماتریس انتقال

مورد نظر برابر است با:

$$T = \begin{bmatrix} 1 & 0 & 0 & -x_1 \\ 0 & 1 & 0 & -y_1 \\ 0 & 0 & 1 & -z_1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (۱۷-۲)$$

در این مرحله بردار دوران را بر محور  $Z$  منطبق می‌کنیم. برای این منظور ابتدا با یک دوران حول محور  $X$  و به اندازه  $\alpha$ ، بردار دوران را بر صفحه  $XZ$  منتقل کرده (شکل ۳-۲ الف) و سپس با اعمال دوران حول محور  $Y$  و به اندازه  $\beta$ ، این بردار را بر محور  $Z$  منطبق می‌کنیم (شکل ۳-۲ ب).



شکل ۳-۲ - مراحل انطباق محور دوران بر محور  $Z$

از آنجایی که محاسبات مربوط به ماتریس دوران نیازمند در اختیار داشتن مقادیر سینوس و کسینوس زاویه‌ی دوران می‌باشد، بنابراین در این مرحله این مقادیر را با استفاده از ضرب نقطه‌ای و برداری محاسبه می‌کنیم. زاویه‌ی  $\alpha$  در شکل ۳-۲ الف، در واقع زاویه‌ی بین تصویر بردار  $U$  بر صفحه  $YZ$  و محور  $Z$  می‌باشد (شکل ۴-۲ الف). تصویر بردار  $U$  بر محور  $YZ$  را  $U_{yz}$  و تصویر این بردار بر محور  $Z$  را  $U_z$  می‌نامیم. از آنجایی

که  $U_z = (0, 0, 1)$  و  $U_{yz} = (0, b, c)$ ، بنابراین:

$$\cos \alpha = \frac{U_z \cdot U_{yz}}{|U_z| |U_{yz}|} = \frac{a}{d} \quad (18-2)$$

$d$  در رابطه‌ی فوق به صورت زیر تعریف می‌شود:

$$d = |U_{yz}| = \sqrt{b^2 + c^2} \quad (19-2)$$

سینوس این زاویه نیز به صورت زیر محاسبه می‌شود:

$$\sin \alpha = \frac{U_z \times U_{yz}}{|U_z| |U_{yz}|} = \frac{b}{d} \quad (20-2)$$

بنابر روابط (۸-۲)، (۱۸-۲) و (۲۰-۲) ماتریس دوران حول محور  $X$  به صورت زیر می‌باشد.

$$R_x(\alpha) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & c/d & -b/d & 0 \\ 0 & b/d & c/d & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (21-2)$$

نحوه‌ی محاسبه‌ی دوران حول محور  $y$  در مرحله‌ی دو نیز مشابه محاسبه‌ی دوران حول محور  $x$

می‌باشد. ماتریس دوران حول این محور و به اندازه‌ی  $\beta$  به صورت زیر محاسبه می‌شود.

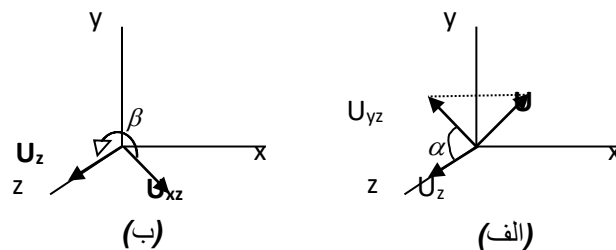
$$R_y(\beta) = \begin{bmatrix} d & 0 & -a & 0 \\ 0 & 1 & 0 & 0 \\ a & 0 & d & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (22-2)$$

ماتریس دوران مرحله‌ی ۳ نیز یک ماتریس دوران ساده حول محور  $z$  و طبق رابطه (۲-۴) می‌باشد.

ماتریس‌های مراحل ۴ و ۵ نیز به ترتیب برابر ماتریس‌های معکوس مراحل ۲ و ۱ می‌باشند. بنابراین ماتریس

دوران نهایی، طبق ترکیب تبدیلات به صورت زیر بدست می‌آید.

$$R(\theta) = T^{-1} \cdot R_x^{-1}(\alpha) \cdot R_y^{-1}(\beta) \cdot R_z(\theta) \cdot R_y(\beta) \cdot R_x(\alpha) \cdot T \quad (23-2)$$



شکل ۲-۴ - ارتباط زوایای  $\alpha$  و  $\beta$  با تصاویر بردار  $U$

#### انطباق دستگاه‌های مختصات

اگر سه بردار متعامد  $u$ ،  $v$  و  $w$  در دستگاه مختصات مرجع را بخواهیم بر محورهای مختصات مرجع

منطبق کنیم، کفایت از ماتریس تبدیل زیر استفاده شود.

$$R = \begin{bmatrix} u_1 & u_2 & u_3 & 0 \\ v_1 & v_2 & v_3 & 0 \\ w_1 & w_2 & w_3 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (24-2)$$

ماتریس فوق بردار  $u$  را بر محور  $x$ ، بردار  $v$  را بر محور  $y$  و بردار  $w$  را بر محور  $z$  منطبق می‌کند. به عبارت دیگر اگر  $p$  نقطه‌ای در دستگاه مختصات  $xyz$  باشد، مختصات این نقطه در دستگاه  $uvw$  توسط ماتریس فوق بصورت  $R.P$  بدست می‌آید. از این رابطه جهت تغییر دستگاه مختصات در دید سه‌بعدی استفاده می‌شود. باید توجه کرد که ۹ پارامتر موجود در ماتریس  $R$ ، به علت وجود الزام تعامد بردارها و راستگرد بودن دستگاه مختصات، پارامترهایی وابسته می‌باشند. در واقع در این ماتریس ۳ پارامتر مستقل وجود دارد که توسط زوایای بین بردارهای  $u$ ،  $v$ ،  $w$  و محورهای متناظر در دستگاه  $xyz$  نیز قابل محاسبه می‌باشند.

### ۱-۸-۳ تغییر مقیاس

ماتریس تغییر مقیاس نسبت به مبداء و با پارامترهای  $s_x$ ،  $s_y$  و  $s_z$  به صورت زیر می‌باشد.

$$S = \begin{bmatrix} s_x & 0 & 0 & 0 \\ 0 & s_y & 0 & 0 \\ 0 & 0 & s_z & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (25-2)$$

از آنجایی که تغییر مقیاس نسبت به یک نقطه‌ی اختیاری برابر با انتقال نقطه به مبداء، تغییر مقیاس نسبت به مبداء و انتقال نقطه به محل اولیه می‌باشد، بنابراین ماتریس تغییر مقیاس نسبت به نقطه‌ی  $(x_p, y_p, z_p)$  به صورت زیر محاسبه می‌شود.

$$S = \begin{bmatrix} s_x & 0 & 0 & (1-s_x)x_p \\ 0 & s_y & 0 & (1-s_y)y_p \\ 0 & 0 & s_z & (1-s_z)z_p \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (26-2)$$

### ۱-۹ دید سه‌بعدی

دید سه‌بعدی به مفهوم نمایش اشیاء سه‌بعدی در سیستمهای نمایشی دوبعدی است. به علت عدم وجود سیستم نمایشی سه بعدی کارآمد، این بخش از ارتباط تصویری کامپیوتری، از بخش‌های بسیار

پر کاربرد در اغلب مباحث سه بعدی است. در این بخش دو مبحث مختصات دید و نگاشت تصاویر مورد بررسی قرار می‌گیرد.

### ۱-۹-۱ مختصات دید

در کاربردهای دید سه بعدی دو دستگاه مختصات مهم وجود دارد. دستگاه مختصات اول، دستگاه مختصات فراگیر یا مرجع<sup>۷۶</sup> می‌باشد که موقعیت اشیاء سه بعدی در این دستگاه تعریف می‌شود. دستگاه مختصات دوم، دستگاه مختصات دید<sup>۷۷</sup> است. ایجاد یک نما از اشیاء سه بعدی مشابه تصویربرداری از اشیاء است. موقعیت نسبی دوربین و شیء، تعیین کننده‌ی تصویر می‌باشد. برای سهولت انجام محاسبات، دستگاه مختصات جدیدی بر حسب موقعیت نسبی دوربین و شیء تعریف می‌شود که مختصات دید نام دارد. صفحه‌ی دید<sup>۷۸</sup> یا صفحه‌ی تصویر<sup>۷۹</sup>، صفحه‌ای متعامد بر محور  $Z$  دستگاه مختصات دید است. به طور خلاصه عملیات دید سه بعدی شامل دو مرحله‌ی انتقال موقعیت نقاط از مختصات مرجع به مختصات دید و نگاشت مختصات دید به صفحه‌ی دید می‌باشد.

جهت تعیین مختصات دید در مرحله‌ی اول می‌بایست یک نقطه در مختصات مرجع تحت عنوان نقطه‌ی دید مرجع انتخاب کرد. این نقطه مبداء مختصات دید می‌باشد (شکل ۲-۵). اگر نقطه‌ی مرجع بر روی شیء قرار گرفته باشد، این نقطه را می‌توان به عنوان نقطه‌ای که دوربین به سمت آن تنظیم شده است در نظر گرفت. در صورتی که نقطه‌ی مرجع از شیء فاصله داشته باشد، این نقطه را می‌توان به عنوان محل قرارگیری دوربین در نظر گرفت. محور  $Z$  دستگاه مختصات دید یا همان  $z_v$ ، در راستای بردار متعامد بر

---

<sup>76</sup> World Coordinate

<sup>77</sup> Viewing Coordinate

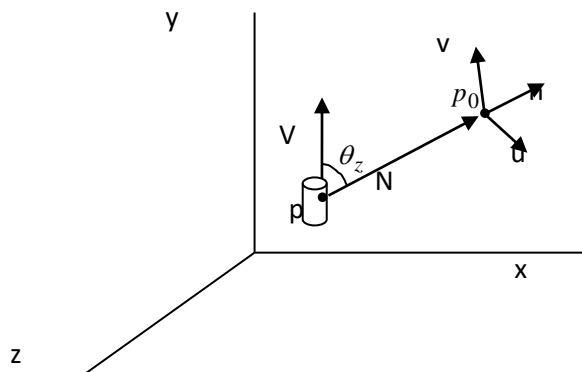
<sup>78</sup> View Plane

<sup>79</sup> Projection Plane

صفحه‌ی دید (  $N$  ) قرار می‌گیرد. بردار  $N$  نیز که راستای قرار گیری دوربین را نشان می‌دهد، به عنوان یک بردار از پیش تعیین شده در نظر گرفته می‌شود. جهت بالا را توسط بردار  $V$  نمایش می‌دهیم. زاویه‌ی بین این بردار که بردار نمای بالا<sup>80</sup> نامیده می‌شود و  $N$  را زاویه‌ی پیچش<sup>81</sup> یا  $\theta_z$  می‌نامیم. حال با داشتن  $N$  و  $V$  سه بردار واحد و متعامد دستگاه مختصات دید را به صورت زیر تعریف می‌کنیم.

$$\begin{aligned} n &= \frac{N}{|N|} = (n_1, n_2, n_3) \\ u &= \frac{V \times N}{|V \times N|} = (u_1, u_2, u_3) \\ v &= n \times u = (v_1, v_2, v_3) \end{aligned} \quad (27-2)$$

شکل ۲-۵ نحوه‌ی قرار گیری این سه بردار را در دستگاه مختصات مرجع پس از انتقال به محل قرار گیری دوربین در نقطه‌ی  $p_0 = (x_0, y_0, z_0)$  نشان می‌دهد. محورهای دستگاه مختصات دید  $x_v$ ،  $y_v$  و  $z_v$  نامیده شده و به ترتیب در راستای بردارهای واحد و متعامد  $u$ ،  $v$  و  $n$  قرار می‌گیرند. مبداء این دستگاه در نقطه‌ی  $p_0$  یا نقطه‌ی مرجع دید قرار داشته و صفحه‌ی دید، معادل با صفحه‌ی  $uv$  می‌باشد.



شکل ۲-۵ - نحوه تشکیل دستگاه مختصات دید

همانطور که اشاره شد عملیات دید سه بعدی شامل دو مرحله‌ی انتقال موقعیت نقاط از مختصات مرجع به مختصات دید و نگاشت مختصات دید به صفحه‌ی دید می‌باشد. نگاشت مختصات دید به صفحه‌ی دید

<sup>80</sup> View-Up Vector

<sup>81</sup> Twist Angle

دید در بخش بعدی تشریح خواهد شد. پس از تشکیل دستگاه مختصات دید، می‌بایست موقعیت نقاط را از دستگاه مرجع به دستگاه مختصات دید انتقال دهیم. برای این منظور در مرحله‌ی اول با استفاده از ماتریس انتقال، مبداء مختصات دید را به مبداء مختصات مرجع منتقل کرده و سپس با اعمال ماتریس دوران طبق مطالب ذکر شده در مبحث "انطباق دستگاه‌های مختصات" در بخش ۲-۱-۲، دستگاه مختصات دید را به دستگاه مرجع منطبق می‌کنیم. ترکیب این دو ماتریس معادل تبدیل موقعیت نقاط از مختصات مرجع به مختصات دید می‌باشد.

$$T = \begin{bmatrix} 1 & 0 & 0 & -x_0 \\ 0 & 1 & 0 & -y_0 \\ 0 & 0 & 1 & -z_0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (28-2)$$

$$R = \begin{bmatrix} u_1 & u_2 & u_3 & 0 \\ v_1 & v_2 & v_3 & 0 \\ n_1 & n_2 & n_3 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (29-2)$$

$$M_{WC2VC} = R.T \quad (30-2)$$

### ۲-۹-۱ تصویرسازی<sup>۸۲</sup>

تصویرسازی در دید سه‌بعدی به نگاشت اشیاء از دستگاه مختصات دید به صفحه‌ی دید اطلاق می‌شود. این نگاشت یک نگاشت از فضای سه‌بعدی به فضای دوبعدی بوده و به دودسته‌ی موازی و پرسپکتیو تقسیم می‌شود. در تصویرسازی موازی نقاط تحت خطوط موازی به صفحه‌ی دید تصویر می‌شوند (شکل ۲-۶ الف)، در حالی که این خطوط در تصویرسازی پرسپکتیو همگرا به نقطه‌ی مرجع تصویرسازی<sup>۸۳</sup> می‌باشند

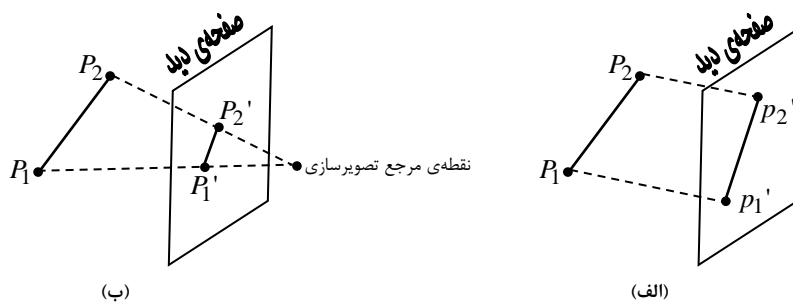
<sup>82</sup> Projection

<sup>83</sup> Projection Reference Point



(شکل ۲-۶ ب). این خطوط، خطوط تصویرسازی نامیده می‌شوند. تصویر دوبعدی از محل برخورد خطوط

تصویرسازی و صفحه‌ی دید بوجود می‌آید.



شکل ۲-۶ - تصویر سازی بر روی صفحه‌ی دید، (الف): موازی، (ب): پرسپکتیو

تصویرسازی موازی نسبت‌ها را حفظ کرده و بیشتر در ترسیم فنی اشیاء سه‌بعدی کاربرد دارد. این

روش تصویر واقع‌گرایانه و ملموسی از اشیاء ایجاد نمی‌کند. در مقابل، تصویرسازی پرسپکتیو نسبت‌ها را حفظ

نکرده و تصاویری ملموس و نزدیک به واقعیت از اشیاء ارائه می‌کند.

### تصویرسازی موازی

هنگامی که در تصویرسازی موازی، خطوط تصویرسازی بر صفحه‌ی دید عمود باشند، تصویرسازی از

نوع موازی عمودی<sup>۸۴</sup> خواهد بود. در غیر این صورت تصویرسازی موازی اریب<sup>۸۵</sup> خواهیم داشت. از آنجایی که

صفحه‌ی دید بر محور  $z_v$  متعامد است، روابط تصویرسازی عمودی به سادگی به صورت زیر بدست می‌آیند.

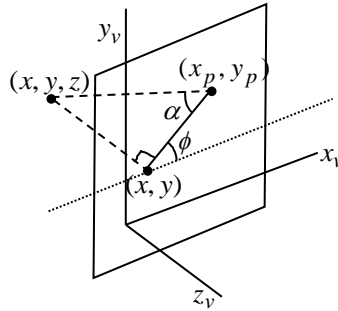
$$x_p = x \quad y_p = y \quad (۲-۳۱)$$

تصویرسازی موازی اریب، با دو زاویه‌ی  $\alpha$  و  $\phi$  همانند شکل ۲-۷ مشخص می‌شود. نقطه‌ی  $(x, y, z)$

به موقعیت  $(x_p, y_p)$  بر روی صفحه‌ی دید یا همان صفحه‌ی  $uv$  نگاشته می‌شود.

<sup>84</sup> Orthographic Parallel Projection

<sup>85</sup> Oblique Parallel Projection



شکل ۷-۲ - موقعیت تصویر در تصویرسازی اریب

با توجه به شکل ۷-۲ و با استفاده از روابط هندسی ساده خواهیم داشت.

$$x_p = x + \frac{z}{\tan \alpha} \cos \phi \quad (۳۲-۲)$$

$$y_p = y + \frac{z}{\tan \alpha} \sin \phi$$

بنابراین ماتریس نگاشت موازی به صورت زیر می‌باشد.

$$M_{parallel} = \begin{bmatrix} 1 & 0 & \frac{z}{\tan \alpha} \cos \phi & 0 \\ 0 & 1 & \frac{z}{\tan \alpha} \sin \phi & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (۳۳-۲)$$

اگر  $\tan \alpha = 1$ ، تصویر سازی را کوالیر<sup>۸۶</sup> و اگر  $\tan \alpha = 2$ ، تصویر سازی را کابینت<sup>۸۷</sup> می‌نامند. شکل ۸-۲

تصویر یک مکعب را تحت نگاشتهای کوالیر و کابینت نشان می‌دهد.



شکل ۸-۲ - تصویر یک مکعب با  $\phi = 45$ ، (الف): تصویرسازی کوالیر، (ب): تصویرسازی کابینت [فولی و همکاران، ۱۹۹۶]

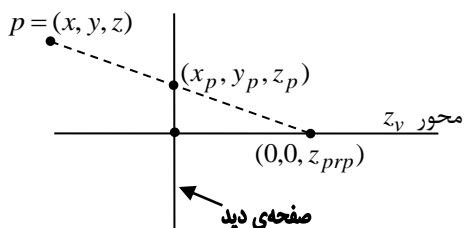
تصویرسازی پرسپکتیو

<sup>86</sup> Cavalier  
<sup>87</sup> Cabinet

جهت بدست آوردن موقعیت تصویر نقاط در تصویرسازی پرسپکتیو، معادلات پارامتری خط را برای

شکل ۹-۲ ایجاد می‌کنیم. صفحه دید، عمود بر محور  $z_v$  و در فاصله‌ی  $z_{vp}$  از مبدا مختصات واقع شده

است. مرجع پرسپکتیو نیز بر روی محور  $z_v$  و در نقطه‌ی  $z_{prp}$  قرار گرفته است.



شکل ۹-۲ - نگاشت پرسپکتیو نقطه‌ی  $(X, Y, Z)$  بر صفحه‌ی دید

با استفاده از پارامتر  $u$  که مقداری بین صفر و یک بوده و بیانگر موقعیت نسبی صفحه‌ی دید در فاصله‌ی بین

نقطه‌ی مرجع تصویرسازی و نقطه‌ی  $p$  می‌باشد، معادلات پارامتری خط متصل کننده‌ی این دو نقطه به

صورت زیر بدست می‌آیند.

$$\begin{aligned} x_p &= x(1-u) \\ y_p &= y(1-u) \\ z_p &= z(1-u) + z_{prp}u \end{aligned} \quad (۳۴-۲)$$

حال با توجه به این که بر روی صفحه‌ی دید،  $z_p = z_{vp}$  از رابطه‌ی فوق مقدار  $u$  بدست می‌آید.

$$u = \frac{z_{vp} - z}{z_{prp} - z} \quad (۳۵-۲)$$

با جایگزینی معادله‌ی (۳۵-۲) در (۳۴-۲) معادلات تصویرسازی پرسپکتیو بدست می‌آیند.

$$\begin{aligned} x_p &= x \left( \frac{z_{prp} - z_{vp}}{z_{prp} - z} \right) \\ y_p &= y \left( \frac{z_{prp} - z_{vp}}{z_{prp} - z} \right) \\ z_p &= z_{vp} \end{aligned} \quad (۳۶-۲)$$

بنابراین ماتریس تصویرسازی پرسپکتیو به صورت زیر است.

$$M_{perspective} = \begin{bmatrix} \frac{z_{prp} - z_{vp}}{z_{prp} - z} & 0 & 0 & 0 \\ 0 & \frac{z_{prp} - z_{vp}}{z_{prp} - z} & 0 & 0 \\ 0 & 0 & 0 & z_{vp} \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (37-2)$$

## ۱۰-۱ تعیین سطوح قابل رؤیت

یکی از بخش‌های مهم در ارتباط تصویری سه‌بعدی، تعیین بخش‌های قابل رؤیت اشیاء می‌باشد. بدین منظور روش‌های مختلفی با توجه به نوع کاربرد، میزان حافظه و سرعت پردازش ارائه شده‌اند. این الگوریتمها به دو دسته تقسیم می‌شوند. دسته‌ی اول الگوریتم‌هایی هستند که سطوح قابل رؤیت را از مشخصات شیء در فضای سه‌بعدی استخراج می‌کنند و روش‌های فضای شیء<sup>۸۸</sup> نامیده می‌شوند. دسته‌ی دوم، سطوح قابل رؤیت را از مشخصات تصویر در فضای دوبعدی استخراج می‌کنند و روش‌های فضای تصویر<sup>۸۹</sup> نامیده می‌شوند. بیشتر الگوریتم‌های تعیین سطوح، از نوع روش‌های فضای تصویر می‌باشند. با این حال روش‌های دسته‌ی اول نیز در بعضی از کاربردها عملکرد بهتری نشان می‌دهند. با وجود تفاوت عملکرد روش‌های مختلف تعیین سطوح، استفاده از یکنواختی<sup>۹۰</sup> و مرتب‌سازی<sup>۹۱</sup> داده‌ها در اغلب الگوریتم‌ها منجر به افزایش چشمگیر سرعت پردازش می‌شود. در این بخش به بررسی چند روش مهم و پراستفاده جهت تعیین سطوح قابل رؤیت می‌پردازیم.

<sup>88</sup> Object-Space Methods

<sup>89</sup> Image-Space Methods

<sup>90</sup> Coherence

<sup>91</sup> Sorting

## ۱-۱۰-۱ تعیین وجه پسین

این روش یکی از روش‌های بسیار سریع و ساده‌ی فضای شیء می‌باشد. کاربرد این روش در تعیین سطوح قابل رؤیت چندوجهی‌های محدب است. جهت بردار عمود بر سطح هر یک از وجوه ( $N$ ) چندوجهی نسبت به جهت بردار دید ( $z_v$ )، تعیین‌کننده‌ی رؤیت‌پذیری وجه می‌باشد. وجه مورد نظر یک وجه پسین یا غیر قابل رؤیت است، اگر رابطه‌ی زیر برقرار باشد.

$$V \cdot N > 0 \quad (38-2)$$

در صورتی که انتقال به دستگاه مختصات دید صورت گرفته باشد، تعیین سطوح پسین از این هم راحت‌تر می‌شود. در این حالت  $V = (0, 0, V_z)$  است و در صورت هم‌علامت بودن  $V_z$  و  $N_z$  وجه مورد نظر یک وجه پسین می‌باشد. این روش جهت تعیین سطوح قابل رؤیت در چندوجهی‌های غیر محدب، نیازمند محاسبات مربوط به انسداد وجوه<sup>۹۲</sup> نیز می‌باشد. با این حال این روش جهت تعیین سطوح غیرقابل رؤیت در تمامی چندوجهی‌ها، روشی کامل بوده و به همین جهت، روش تعیین سطوح پسین نامیده می‌شود. این روش معمولاً به عنوان پیش‌پردازشی جهت کاهش تعداد نمونه‌های مورد بررسی استفاده می‌شود.

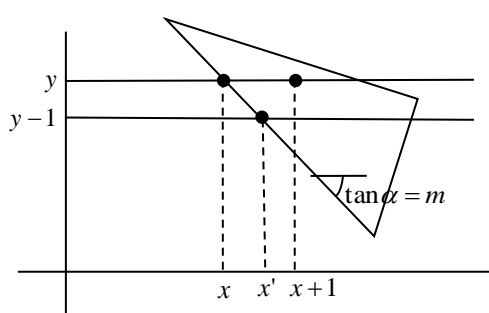
## ۱-۱۰-۲ روش بافر عمق

روش بافر عمق<sup>۹۳</sup> یا روش بافر  $Z$  یکی از روش‌های پرکاربرد در حوزه‌ی روش‌های فضای تصویر می‌باشد. در این روش، در هر پیکسل تصویر نگاشته شده بر صفحه‌ی دید، عمق رویه‌ی مربوطه محاسبه می‌شود و تعیین سطوح قابل رویت برحسب عمق رویه‌ها صورت می‌گیرد. در صورت انتقال نقاط به دستگاه مختصات دید، عمق هر نقطه معادل با مؤلفه‌ی  $Z$  نقاط می‌باشد. این روش در محاسبه‌ی اشیاء متشکل از

<sup>92</sup> Occlusion

<sup>93</sup> Depth Buffer Method

چندضلعی‌ها نیز عملکرد سریعی دارد. در این روش برای هر پیکسل در تصویر، دو بافر در نظر گرفته می‌شود. بافر عمق که جهت حفظ عمق نقاط تصویرشده بر روی پیکسل استفاده شده و بافر بافت که مقدار بافت نقاط را ذخیره می‌کند. مقادیر اولیه‌ی بافر عمق و بافت به ترتیب برابر صفر و رنگ پس‌زمینه در نظر گرفته می‌شوند. پس از محاسبه‌ی مقدار عمق مربوط به هر رویه، این مقدار با مقدار ذخیره شده در بافر عمق مقایسه شده و در صورت بزرگتر بودن در بافر عمق ذخیره می‌شود. رنگ این نقطه نیز در بافر بافت ذخیره می‌شود.



شکل ۲-۱۰ - استفاده از خط پوش جهت تعیین عمق نقاط روی صفحه

جهت تعیین عمق نقاط داخلی چندضلعی‌ها از روش خط پوش<sup>۹۴</sup> همانند شکل ۲-۱۰ استفاده

می‌کنیم. با داشتن معادله‌ی صفحه، عمق هر نقطه روی آن به صورت زیر بدست می‌آید.

$$z = \frac{-Ax - By - D}{C} \quad (۲-۳۹)$$

برای هر خط پوش همچون خط  $y$  در شکل ۲-۱۰، با داشتن مقدار  $z$  برای نقطه‌ی  $(x, y)$ ، مقدار  $z'$  برای نقطه‌ی  $(x+1, y)$  از رابطه‌ی ساده‌ی زیر محاسبه می‌شود.

$$z' = \frac{-A(x+1) - By - D}{C} \Rightarrow z' = z - \frac{A}{C} \quad (۲-۴۰)$$

از آنجایی که  $A/C$  مقداری ثابت است محاسبه‌ی عمق برای نقطه‌ی روی خط پوش با انجام یک بار عمل جمع صورت می‌گیرد. محاسبه‌ی عمق نقاط ابتدایی هر خط پوش نیز با معلوم بودن شیب اولین ضلع

<sup>94</sup> Scan Line

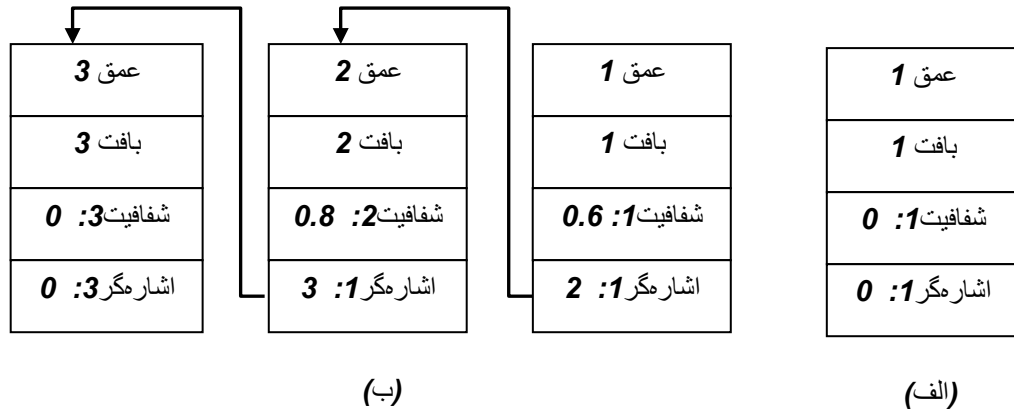
مقاطع با خطوط پویس به سادگی صورت می‌گیرد. در صورتی که شیب این خط برابر  $m$  باشد، مقدار  $x'$  و عمق نقطه‌ی  $(x', y-1)$  در شکل ۲-۱۰ به صورت زیر محاسبه می‌شود.

$$x' = x - \frac{1}{m} \Rightarrow z' = z + \frac{A/m + B}{C} \quad (۲-۴۱)$$

### ۱-۱۰-۳ روش بافر A

این روش حالت گسترش یافته‌ی روش بافر عمق می‌باشد. نقص مهم روش بافر عمق عدم امکان تعیین سطوح قابل رؤیت در صورت وجود رویه‌های شفاف می‌باشد. روش بافر A<sup>۹۵</sup> با در نظر گرفتن آرایه‌های پیاپی از رویه‌ها به رفع این نقیصه می‌پردازد. در این روش به ازای هر پیکسل در تصویر، چهار بافر عمق، بافت، ضریب شفافیت و اشاره‌گر در نظر گرفته شده است. ضریب شفافیت عددی بین صفر و یک می‌باشد. به ازای هر پیکسل، عمق نقطه تصویرشده با مقدار بافر عمق مقایسه می‌شود. در صورت بزرگتر بودن این مقدار اگر نقطه‌ی موردنظر مربوط به یک رویه‌ی غیرشفاف باشد، جایگزین مقدار موجود در بافر عمق می‌شود. مشخصات بافت و شفافیت نقطه نیز در بافرهای بافت و شفافیت ذخیره شده و بافر اشاره‌گر نیز برابر صفر در نظر گرفته می‌شود (شکل ۲-۱۱ الف). اما در حالتی که نقطه‌ی تصویرشده مربوط به یک رویه‌ی شفاف باشد، مقدار بافر عمق، بافت، شفافیت و اشاره‌گر در یک آرایه ذخیره شده و سپس عمق و بافت و شفافیت این نقطه جایگزین مقادیر موجود در بافرهای مربوطه می‌شود. بافر اشاره‌گر نیز آدرس آرایه‌ی مذکور را ذخیره می‌کند. در صورت کوچکتر بودن مقدار عمق نقطه نسبت به مقدار بافر عمق، مراحل فوق برای تمامی آرایه‌های اشاره شده توسط بافرهای اشاره‌گر ادامه پیدا می‌کند (شکل ۲-۱۱ ب).

<sup>95</sup> A-Buffer Method



شکل ۱۱-۲- بافرهای موجود در روش بافر A، (الف): حالت بدون رویه‌ی شفاف قابل رویت، (ب): حال دارای دو رویه شفاف قابل رویت

با توجه به شکل ۱۱-۲ مقدار واقعی بافت هر پیکسل با میانگین‌گیری وزن‌دار از مقادیر بافت ذخیره

شده در آرایه‌های پیاپی بدست می‌آید. اگر طول آرایه‌ی پیاپی برابر n باشد و بافت و شفافیت را به ترتیب با

C و t نمایش دهیم، آنگاه مقدار بافت برای هر پیکسل به صورت زیر بدست می‌آید.

$$C_{pixel} = \frac{1}{n} \sum_{i=1}^n (1-t_i) C_i \quad (۴۲-۲)$$

### ۱۱-۱ مدل‌های نوردهی

نمایش واقعی اشیاء سه‌بعدی نیازمند تصویرسازی پرسپکتیو و پیاده‌سازی نوردهی طبیعی می‌باشد.

یک مدل نوردهی<sup>۹۶</sup> که به آن مدل سایه‌گذاری<sup>۹۷</sup> نیز گفته می‌شود، مدلی است که جهت محاسبه‌ی شدت

نوری که باید در هر نقطه از رویه توسط ناظر دیده شود، می‌پردازد. ارائه و پیاده‌سازی مدل کاملی که تمام

شرایط نوری را در نظر بگیرد، بسیار پیچیده و تقریباً غیرممکن است. با این حال با ساده‌سازی‌های صورت

گرفته در فیزیک نور، مدل‌های ساده و قابل قبولی ارائه شده‌اند. در این بخش روش‌های محاسبه‌ی شدت نور

در یک نقطه از سطح ارائه می‌شوند. این روش‌ها ساده و سریع بوده و برای اغلب کاربردها نتایج مناسبی

<sup>96</sup> Illumination Model

<sup>97</sup> Shading Model



ایجاد می‌کنند. پارامترهای تعیین‌کننده در محاسبات نوری شامل خواص نوری سطوح، شرایط نوری پس‌زمینه و مشخصات منبع نور می‌باشند. خواص نوری سطوح همچون میزان براقی، ماتی و شفافیت توسط پارامترهای معرفی‌کننده میزان بازتابش، جذب و عبوردهی نور در معادلات وارد می‌شوند. تمام منابع نور نیز از نوع نقطه‌ای با موقعیت مشخص می‌باشند.

### ۱-۱۱-۱ نور محیطی و بازتابش پخشی

اشیایی که تحت تابش مستقیم نور نیز قرار ندارند، در حالتی که اشیاء مجاور آنها تحت تابش باشند، قابل رویت هستند. علت این امر وجود نور محیطی یا نور پس‌زمینه می‌باشد. نور محیطی در واقع برآیند تمام بازتابش‌هایی از اشیاء است که منجر به ایجاد نوری یکنواخت می‌شود. نور محیطی جهت و یا هیچ مشخصه فضایی دیگری ندارد. میزان نور محیطی تابیده‌شده به تمام اشیاء و در تمام جهات مقداری ثابت می‌باشد. شدت نور محیطی تابیده شده را با  $I_d$  نشان می‌دهیم. شدت بازتابش نور محیطی نیز مقداری مستقل از جهت و موقعیت فضایی سطوح بوده ولی تابعی از خواص نوری سطح یعنی میزان جذب و بازتابش سطح می‌باشد.

بازتابش نور توسط سطوح به دو دسته‌ی بازتابش پخشی<sup>۹۸</sup> و بازتابش آینه‌ای<sup>۹۹</sup> تقسیم می‌شود. بازتابش پخشی مستقل از جهت و زاویه دید بوده و تنها به جهت بردار تابش بستگی دارد. نور محیطی نیز تقریبی از بازتابش پخشی تمام اشیاء موجود در محیط می‌باشد. بخشی از نور تابیده شده به هر سطح توسط آن بازتابش شده و بخشی نیز جذب می‌شود. میزان نسبت شدت نور بازتابش شده به کل نور را که از خواص نوری سطح می‌باشد، با  $k_d$  نشان داده و ضریب بازتابش پخشی نامیده می‌شود. مقدار  $k_d$  با توجه به میزان

---

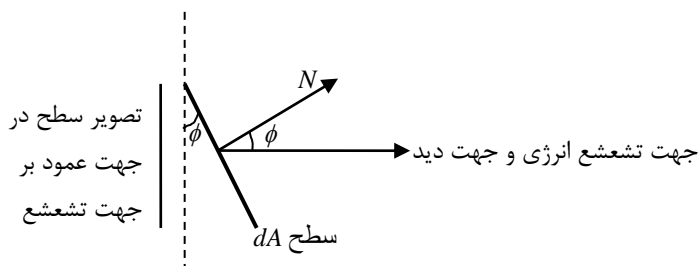
<sup>98</sup> Diffused Reflection

<sup>99</sup> Specular Reflection

خاصیت بازتابشی سطح، عددی بین صفر و یک می‌باشد. در واقع پارامتر  $k_d$  تابعی از رنگ سطح می‌باشد. در

صورتی که فقط نور محیطی به سطح اعمال شود، میزان بازتابش نور توسط رابطه‌ی زیر مشخص می‌شود.

$$I_{\text{ambdiff}} = k_d I_a \quad (۴۳-۲)$$



شکل ۲-۱۲ - جهت تشعشع انرژی از سطح دیفرانسیلی  $dA$

میزان نور بازتابش شده به صورت پخشی از یک منبع نقطه‌ای نیز در تمام جهات یکسان بوده و

تابعی از زاویه‌ی دید نمی‌باشد. این نکته براساس قانون کسینوسی لامبرت<sup>۱۰۰</sup>، اثبات می‌شود. طبق این قانون

و با توجه به شکل ۲-۱۲، میزان انرژی تشعشع یافته از سطح دیفرانسیلی  $dA$  در جهت  $\phi_N$  نسبت به بردار

عمود بر سطح  $N$ ، با  $\cos \phi_N$  متناسب است. اما شدت نور بازتابیده شده نیز متناسب با نسبت انرژی تشعشع

یافته به مساحت تصویر سطح در جهت عمود بر  $\phi_N$  می‌باشد. بنابراین شدت نور بازتابشی مستقل از  $\phi_N$  و

زاویه‌ی دید می‌باشد.

با وجود این که شدت نور بازتابش پخشی ناشی از یک منبع نقطه‌ای، مستقل از زاویه‌ی دید و در

تمام جهات یکسان می‌باشد، اما وابسته به جهت تابش نور بوده و سطوح عمود بر جهت تابش نور روشن‌تر از

سطوح اریب می‌باشند. علت این امر کاهش پرتوهای ورودی به سطح بر اثر کاهش مساحت تصویر سطح در

جهت عمود بر بردار تابش نور می‌باشد. اگر زاویه‌ی بین پرتو تابش و بردار عمود بر سطح را  $\theta$  بنامیم، در این

صورت با توجه به شکل ۲-۱۳ مساحت تصویر سطح در جهت عمود بر بردار تابش، متناسب با  $\cos \theta$  بوده که

<sup>100</sup> Lambert's Cosine Law

متناسب بامیزان پرتوهای ورودی است. اگر شدت نور منبع نقطه‌ای را  $I_t$  بنامیم، با توجه به موارد فوق شدت

نور بازتابش پخش‌ناشی از تابش منبع نقطه‌ای نور، از رابطه‌ی زیر بدست می‌آید.

$$I_{t,diff} = k_d I_t \cos \theta \quad (44-2)$$

بردار منبع نور را در جهت عکس بردار تابش تعریف کرده و آن را با  $L$  نمایش می‌دهیم. بنابراین رابطه‌ی (۲)-

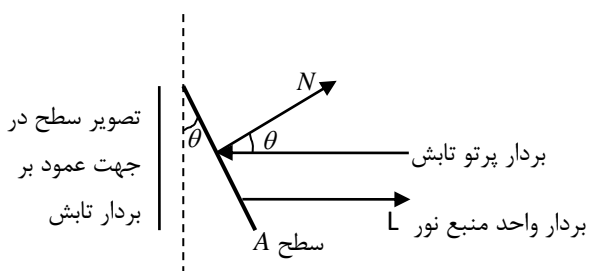
(۴۴) به رابطه‌ی برداری زیر تبدیل می‌شود.

$$I_{t,diff} = k_d I_t (N \cdot L) \quad (45-2)$$

با ترکیب بازتابش پخش‌ناشی از نور محیطی و بازتابش پخش‌ناشی از منبع نقطه‌ای نور، مولفه‌ی بازتابش

پخش‌ناشی نور در هر نقطه به صورت زیر محاسبه می‌شود.

$$I_{diff} = k_d I_a + k_d I_t (N \cdot L) \quad (46-2)$$



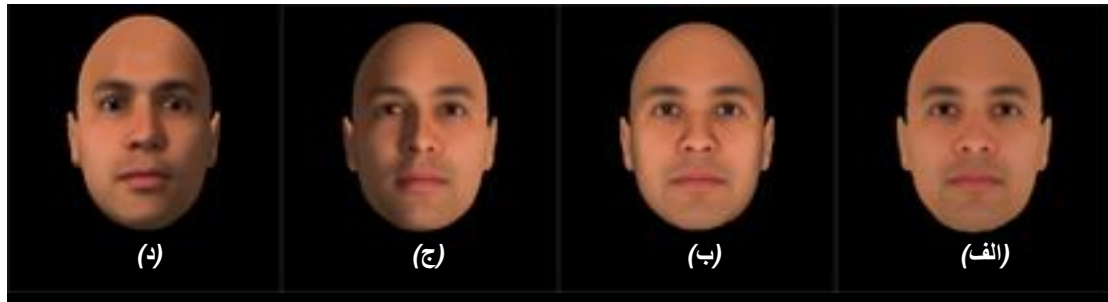
شکل ۲-۱۳ - ارتباط بین زاویه‌ی تابش و شدت بازتابش پخش‌ناشی

شکل ۲-۱۴ تأثیر تغییر پارامترهای بازتابش پخش‌ناشی را بر روی تصویر سه‌بعدی چهره نشان می‌دهد. دقت

شود که در این تصاویر، اثر سایه پیاده‌سازی نشده و تنها بازتابش پخش‌ناشی مدل فانگ استفاده شده است.

روشن بودن بخشی از گوش در تصاویر (ج) و (د) که دارای بردار  $N$  هم‌جهت با  $L$  می‌باشد، مؤید این مطلب

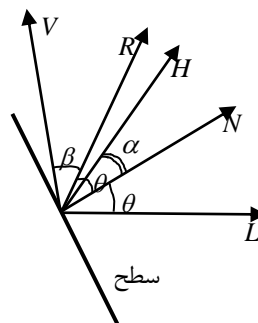
است.



شکل ۲-۱۴ - تأثیر بازتابش پخشی بر چهره، (الف): نور محیطی، در سایر موارد نور محیطی بعلاوه‌ی (ب): منبع نقطه‌ای در  $0^\circ$ ، (ج): منبع نقطه‌ای در  $45^\circ$ ، (د): منبع نقطه‌ای در  $45^\circ$

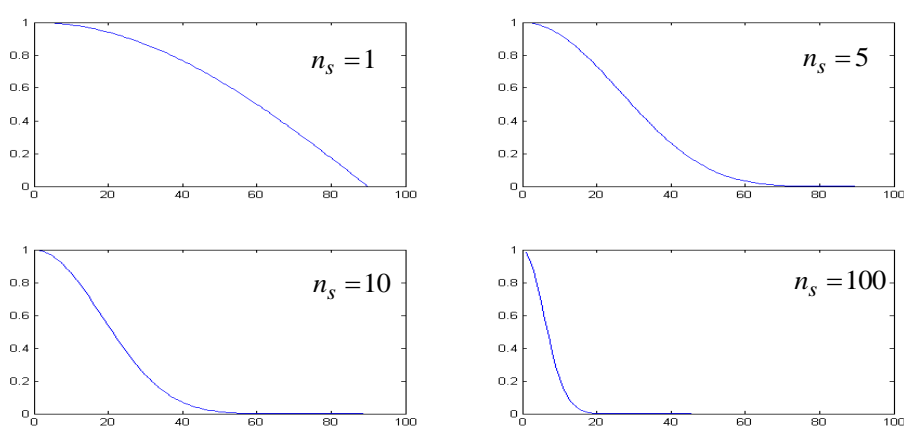
### ۱-۱۱-۲ بازتابش آینه‌ای و مدل فانگ

اشیاء براق بر اثر تابش نور در بعضی از زوایای خاص، نور بیشتری بازتابش می‌کنند. این نوع بازتابش که وابسته به زاویه‌ی دید است، بازتابش آینه‌ای نامیده می‌شود. بردار بازتابش آینه‌ای برداری است که جهت بیشترین بازتابش آینه‌ای را نشان می‌دهد. این بردار را با  $R$  نشان می‌دهیم. به عبارت دیگر اگر سطح به اندازه‌ی کافی شفاف باشد و ناظر در راستای این بردار به سطح نگاه کند، پرتوهای بازتابش آینه‌ای را خواهد دید. اگر زاویه‌ی بین بردار عمود بر سطح ( $N$ ) و بردار منبع نور ( $L$ ) را  $\theta$  بنامیم، آنگاه بردار بازتابش آینه‌ای، برداری است در صفحه‌ی  $NL$  که با  $N$  زاویه‌ای برابر  $\theta$  و با  $L$  زاویه‌ای برابر  $2\theta$  می‌سازد (شکل ۲-۱۵). زاویه‌ی بین بردار دید ( $V$ ) و بردار بازتابش آینه‌ای ( $R$ ) را  $\beta$  می‌نامیم. اگر سطح یک بازتابنده‌ی کامل همچون آینه باشد، پرتوهای بازتابش آینه‌ای تنها در جهت  $R$ ، یعنی به ازای  $\beta = 0$  منتشر خواهند شد. در غیر این صورت به ازای مقادیر غیر صفر  $\beta$  نیز مقداری بازتابش آینه‌ای خواهیم داشت.



شکل ۲-۱۵ - بردارها و زوایای مهم در بازتابش آینه‌ای

مدل بازتابش آینه‌ای فانگ یا مدل نوردهی فانگ، اثر بازتابش آینه‌ای را متناسب با  $\cos^{n_s} \beta$  در نظر می‌گیرد. با تغییر  $\beta$  بین ۰ و ۹۰، مقدار پارامتر فانگ نیز بین ۱ و ۰ تغییر می‌کند. برای سطوح شفاف همچون فلزات، عددی بزرگ (نزدیک به ۱۰۰) و برای سطوح کدر عددی کوچک (کمتر از ۱) در نظر گرفته می‌شود. برای بازتابنده‌ی کامل معادل بی‌نهایت می‌باشد. نمودار  $\cos^{n_s} \beta$  به ازای مقادیر مختلف  $n_s$  بر حسب  $\beta$  در شکل ۱۶-۲ آورده شده است.



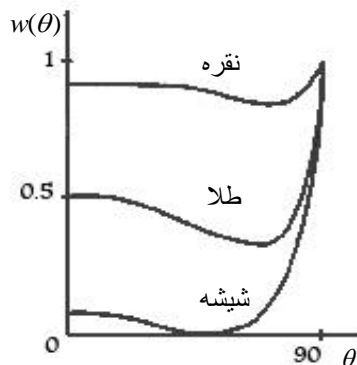
شکل ۱۶-۲ - نمودار پارامتر فانگ به ازای مقادیر مختلف  $n_s$

شدت نور بازتابش آینه‌ای علاوه بر وابستگی به زاویه‌ی دید، به زاویه‌ی تابش یا همان  $\theta$  نیز وابسته می‌باشد. هرچه مقدار  $\theta$  به ۹۰ درجه نزدیک‌تر باشد درصد بیشتری از نور تابیده شده به صورت آینه‌ای بازتابش می‌شود. در مقابل، به ازای  $\theta$ ‌های نزدیک به صفر بیشتر پرتوهای تابش، جذب شده و در زوایای بین این دو مقدار بخش قابل توجهی از پرتوهای تابشی به صورت پخش بازتاب می‌شوند. ارتباط بین شدت بازتابش آینه‌ای و زاویه‌ی تابش توسط قانون فرسnel<sup>۱۰۱</sup> بیان می‌شود. بنابراین به منظور اعمال این وابستگی در معادلات بازتابش آینه‌ای ضریب بازتابش آینه‌ای به صورت تابعی از  $\theta$  در نظر گرفته می‌شود. این ضریب

<sup>101</sup> Fresnel's Law

که با  $w(\theta)$  نمایش داده می‌شود، با توجه به خواص نوری سطح و طبق قانون فرسnel تعیین می‌شود. شکل

۱۷-۲، نمودار ضریب بازتابش آینه‌ای را برای چند ماده‌ی مختلف نشان می‌دهد.



شکل ۱۷-۲ - ضریب بازتابش آینه‌ای برای چند ماده‌ی مختلف

با توجه به مطالب مربوط به ضریب بازتابش آینه‌ای و ارتباط بین  $\beta$  و بازتابش آینه‌ای، فانگ مدل بازتابش خود را به صورت زیر ارائه داد.

$$I_{spec} = w(\theta) I_t \cos^{n_s} \beta \quad (47-2)$$

در رابطه‌ی فوق  $I_{spec}$  شدت روشنایی ناشی از بازتابش آینه‌ای و  $I_t$  شدت نور تابیده شده می‌باشد.

همانطور که در شکل ۱۷-۲ مشخص است برای موادی چون شیشه، بازتابش آینه‌ای تنها در زوایای

تابش نزدیک به  $90^\circ$  از شدت قابل توجهی برخوردار است. اما برای مواد کدر چنین تغییراتی در نمودار

$w(\theta)$  وجود نداشته و ضریب بازتابش آینه‌ای مقدار تقریباً ثابتی دارد. بنابراین در رابطه‌ی (۴۷-۲) می‌توان

بجای  $w(\theta)$  از مقدار ثابت  $k_s$  استفاده کرد. علاوه بر این با توجه به شکل ۱۵-۲،  $\cos^{n_s} \beta$  را نیز می‌توان بر

حسب بردارهای دید و بازتابش آینه‌ای نوشت.

$$I_{spec} = k_s I_t \cos^{n_s} \beta = k_s I_t (V \cdot R)^{n_s} \quad (48-2)$$

حجم محاسبات مربوط به شدت روشنایی بازتابش آینه‌ای را می‌توان با استفاده از یک تقریب ساده، کاهش

داد. بردار  $H$  را به صورت  $H = \frac{L+V}{|L+V|}$  که معادل با نیم‌ساز دو بردار  $L$  و  $V$  است، تعریف می‌کنیم. همانطور

که در شکل ۲-۱۵ مشخص است، زاویه‌ی بین  $H$  و  $N$  ( $\alpha$ )، تقریباً با زاویه‌ی بین  $V$  و  $R$  ( $\beta$ ) برابر است. بنابراین رابطه (۲-۴۸) را میتوان با تقریب به صورت زیر نوشت.

$$I_{spec} = k_s I_t (N.H)^{n_s} \quad (۲-۴۹)$$

در حالتی که منبع نور و ناظر از شیء دور باشند، دو بردار  $V$  و  $L$ ، به ازای تمام نقاط شیء ثابت می‌باشند. بنابراین تنها به یک بار محاسبه‌ی  $H$  نیاز داریم. در حالی که بردار  $R$  تابعی از  $N$  بوده و می‌بایست برای تمام نقاط شیء محاسبه شود. بردار  $N$  نیز در مرحله محاسبه بازتابش پخششی، بدست آمده و نیازی به تکرار محاسبات مربوط به این بردار نیست. شکل ۲-۱۸ تأثیر افزایش  $k_s$  را در یک مدل چهره‌ی سه‌بعدی نمایش می‌دهد. این حالت براق و روغنی چهره در واقع ناشی از نوعی خطای دید می‌باشد. مشخصات بافت چهره به جز در نقاط روشن ناشی از بازتابش آینه‌ای، در دو تصویر کاملاً یکسان است. در حالی که به علت خطای دید ما تمام نقاط چهره‌ی (ب) را براق و روغنی تصور می‌کنیم.



شکل ۲-۱۸ - تأثیر بازتابش آینه‌ای بر چهره،  $n_s$  برابر ۸۰ و منبع نقطه‌ای در موقعیت ۰-۰، (الف)  $k_s=0.05$ ، (ب)  $k_s=0.6$

حال می‌توان با توجه به مطالب مطرح شده در دو بخش اخیر، شدت نور بازتابشی به ازای یک منبع

نقطه‌ای را طبق رابطه‌ی زیر محاسبه کرد.

$$I_{ref} = I_{diff} + I_{spec} = k_a I_a + k_d I_t (N.L) + k_s I_t (N.H)^{n_s} \quad (۲-۵۰)$$

رابطه فوق را در صورت وجود  $n$  منبع نقطه‌ای می‌توان به صورت زیر تغییر داد.

$$I_{ref} = k_a I_a + \sum_{i=1}^n I_{t_i} [k_d (N.L) + k_s (N.H)^{n_s}] \quad (۲-۵۱)$$

## تضعیف شدت نور

دامنه‌ی انرژی تشعشعی دریافتی در یک نقطه با فاصله‌ی نقطه تا منبع انرژی به صورت مربع معکوس ارتباط دارد. این بدین معنا است که سطحی که در فاصله‌ی نزدیک‌تری نسبت به منبع نور قرار گرفته است، شدت نور تابشی بیشتری را نسبت به سطوح دورتر دریافت می‌کند. بنابراین جهت ایجاد یک مدل نوردهی واقعی تر باید تضعیف شدت نور بر اثر افزایش فاصله به نسبت  $\frac{1}{d^2}$  را نیز در نظر گرفت. اما این نسبت در مدل نوردهی ارائه شده که از منابع نقطه‌ای استفاده می‌کند، مشکلاتی را بوجود می‌آورد. نسبت  $\frac{1}{d^2}$  در فواصل کوچک تغییرات زیادی داشته و در فواصل بزرگ تغییرات کمی دارد که منجر به ایجاد تصاویر غیر واقعی می‌شود. علت این امر سادگی بسیار زیاد مدل نوردهی می‌باشد. به عبارت دیگر استفاده از یک منبع نقطه‌ای به عنوان تنها منبع تأمین نور در شرایط واقعی، بسیار کم پیش می‌آید و مدل ارائه شده نیز قابلیت مدل‌سازی منابع نور پیچیده‌ی واقعی را ندارد. جهت جبران این مشکل به جای استفاده از  $\frac{1}{d^2}$  از یک چند جمله‌ای معکوس درجه‌ی دو استفاده می‌شود که با تغییر ضرایب آن می‌توان تصاویر واقعی‌تری ایجاد کرد. بنابراین مدل نوردهی ارائه شده در رابطه‌ی (۲-۵۱) با اعمال اثر تضعیف شدت نور، به رابطه‌ی زیر تبدیل خواهد شد.

$$I_{ref} = k_a I_a + \sum_{i=1}^n f(d_i) I_i [k_d (N \cdot L_i) + k_s (N \cdot H_i)^{n_s}] \quad (۲-۵۲)$$
$$f(d) = \frac{1}{a_2 d^2 + a_1 d + a_0}$$

هنگامی که فاصله‌ی منابع از شیء به نسبت ابعاد شیء زیاد باشد استفاده از رابطه فوق تأثیر قابل توجهی در نتیجه‌ی نهایی نخواهد داشت. در اغلب تصاویر سه بعدی چهره نیز می‌توان این حالت را متصور بود.

## نوردهی رنگی



برای مدل‌های سه‌بعدی دارای بافت رنگی همانند مدل‌های چهره‌ی سه‌بعدی می‌بایست اثر بافت چهره را بر مدل نوردهی در نظر بگیریم. روابط مربوط به مدل نوردهی فانگ که تاکنون بررسی شدند، جهت ایجاد تصاویر تکفام مناسب هستند. روابط فوق به سادگی به روابط مربوط مدل نوردهی رنگی تبدیل می‌شوند. برای این منظور کافی است تا مدل را بر روی هر یک از اجزاء تشکیل دهنده‌ی رنگ همچون  $G, R$  و  $B$  اعمال کنیم. رابطه‌ی زیر شدت بازتابش رنگ آبی را در اثر نوردهی نشان می‌دهند.

$$I_{B,ref} = k_{aB}I_{aB} + \sum_{i=1}^n f(d_i)I_{tB_i} [k_{dB}(N.L_i) + k_{sB}(N.H_i)^{n_s}] \quad (53-2)$$

فانگ در مدل نوردهی خود مقدار  $k_s$  را برای تمام اجزای رنگ مقداری یکسان فرض کرده است. این امر باعث می‌شود که سطوح براق بدون در نظر گرفتن رنگ خود به بازتابش آینه‌ای نور تابشی بپردازند که باعث می‌شود سطوح، ظاهری پلاستیکی داشته باشند. در واقع ضرایب  $k_a, k_d$  و  $k_s$  هر سه متناسب با مؤلفه‌های رنگ نقاط می‌باشند. برای مثال می‌توان ضریب  $k_{aB}$  را برای هر نقطه به صورت  $k_a B$  در نظر گرفت.  $B$  مؤلفه‌ی رنگ آبی نقطه‌ی مورد نظر است. بنابراین روابط مدل نوردهی رنگی فانگ به صورت زیر می‌باشند.

$$\begin{aligned} I_{R,ref} &= k_a R I_{aR} + \sum_{i=1}^n f(d_i) R I_{tR_i} [k_d(N.L_i) + k_s(N.H_i)^{n_s}] \\ I_{G,ref} &= k_a G I_{aG} + \sum_{i=1}^n f(d_i) G I_{tG_i} [k_d(N.L_i) + k_s(N.H_i)^{n_s}] \\ I_{B,ref} &= k_a B I_{aB} + \sum_{i=1}^n f(d_i) B I_{tB_i} [k_d(N.L_i) + k_s(N.H_i)^{n_s}] \end{aligned} \quad (54-2)$$

یک سطح شفاف سطحی است که دو دسته نور بازتابشی و انتقالی<sup>۱۰۳</sup> را تولید می‌کند. به عبارت دیگر یک سطح شفاف هم نور تابیده شده را بازتابش می‌کند و هم نور اشیاء پس‌زمینه را عبور می‌دهد. عبور نور اشیاء پس‌زمینه از داخل سطح شفاف منجر به جابجایی بردار جهت نور می‌شود. به عبارت دقیق‌تر طبق قانون اسنل<sup>۱۰۴</sup> پرتو نور در محل جداکننده‌ی دو محیط به علت تغییر سرعت دچار شکست می‌شود. در صورتی که نور از محیط رقیق وارد محیط غلیظ شود زاویه‌ی پرتو شکست نسبت به بردار عمود بر سطح کمتر می‌شود و در عبور از محیط غلیظ به رقیق، زاویه‌ی پرتو شکست بیشتر می‌شود. شکل ۲-۱۹ (الف) بردارهای منبع نور، بازتابش و شکست را نشان می‌دهد. با توجه به این شکل و طبق قانون اسنل رابطه‌ی بین زاویه‌ی تابش و زاویه شکست به صورت زیر می‌باشد.

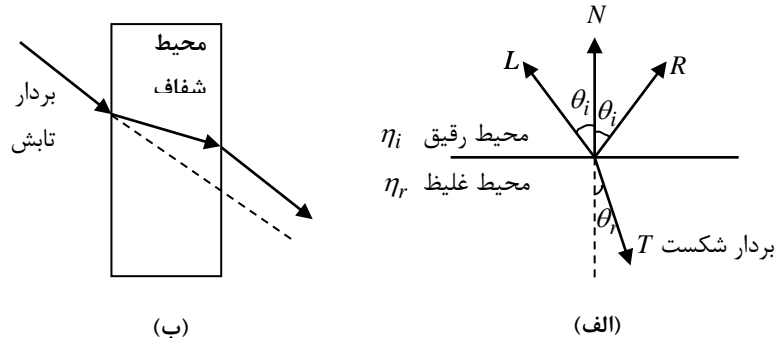
$$\frac{\sin \theta_r}{\sin \theta_i} = \frac{\eta_i}{\eta_r} \quad (۵۵-۲)$$

۱۱ در رابطه‌ی فوق ضریب شکست محیط بوده و برابر با نسبت سرعت نور در خلاء به سرعت نور در ماده می‌باشد. با توجه به رابطه‌ی فوق جهت بردار شکست ( $T$ ) به صورت زیر قابل محاسبه است.

$$T = \left( \frac{\eta_i}{\eta_r} \cos \theta_i - \cos \theta_r \right) N - \frac{\eta_i}{\eta_r} L \quad (۵۶-۲)$$

اما در اغلب موارد، شیء شفاف موجود در صحنه یک تیغه‌ی شفاف می‌باشد. در این حالت نیازی به محاسبات پیچیده‌ی فوق نبوده و همانطور که در شکل ۲-۱۹ (ب) نشان داده شده است، با جابجایی بردار تابش، بردار شکست خروجی بدست می‌آید. میزان جابجایی به ضریب شکست و ضخامت محیط شفاف وابسته است.

<sup>103</sup> Transmitted Light  
<sup>104</sup> Snell's Law



شکل ۲-۱۹ - شکست نور، (الف): بردارهای شکست و بازتابش، (ب): عبور نور از تیغه‌ی شفاف

اما شدت نور در هنگام حضور جسم شفاف از دو مؤلفه‌ی نور انتقالی و بازتابشی تشکیل شده است. نسبت این دو مؤلفه به میزان شفافیت محیط بستگی دارد. هر چه محیط شفاف‌تر باشد نسبت نور انتقالی به نور بازتابشی بیشتر است. این نسبت توسط ضریب شفافیت ( $k_t$ ) نشان داده می‌شود. هر یک از دو مؤلفه‌ی بازتابشی و انتقالی می‌بایست توسط رابطه‌ی (۲-۵۴) و یا روابط مشابه محاسبه شوند. با مشخص بودن جهت بردار شکست ( $T$ ) و تعیین بردار تابش بر حسب آن و در نتیجه مشخص کردن نقاط متناظر سطوح شفاف و پس‌زمینه، می‌توان الگوریتم تعیین سطوح قابل رؤیت بافر  $A$  را پیاده‌سازی کرد. تعیین نقاط متناظر باعث مشخص شدن مقادیر بافر اشاره‌گر در این الگوریتم می‌شود. در نهایت با استفاده از رابطه‌ی زیر مقدار نهایی

شدت نور برای هر یک از مؤلفه‌های رنگ پیکسل مورد نظر در تصویر شیء محاسبه می‌شود.

$$\begin{aligned}
 I_R &= (1 - k_{t_R}) I_{ref} + k_{t_R} I_{trans} \\
 I_G &= (1 - k_{t_G}) I_{ref} + k_{t_G} I_{trans} \\
 I_B &= (1 - k_{t_B}) I_{ref} + k_{t_B} I_{trans}
 \end{aligned}
 \tag{۲-۵۷}$$

باید دقت کرد که محاسبه‌ی تناظر بین نقاط سطوح شفاف و پس‌زمینه همچون بازتابش آینه‌ای وابسته به بردار دید بوده و می‌بایست در فضای تصویر صورت گیرد. الگوریتم تعیین سطوح قابل رؤیت بافر  $A$  یک الگوریتم فضای تصویر می‌باشد.

### ۱-۱۱-۳ پیاده‌سازی مدل نوردهی بر روی چندضلعی‌ها

چندضلعی‌ها در نمایش اشیاء سه بعدی دو کاربرد متفاوت دارند. کاربرد اول نمایش رویه‌های منحنی است که در آن با استفاده از چندضلعی‌ها تقریبی از رویه نمایش داده می‌شود و کاربرد دوم کاربرد بدون تقریب نمایش چندوجهی‌ها می‌باشد. نمایش رویه‌ها با استفاده از چندضلعی‌ها مرسوم‌ترین شیوه‌ی نمایش رویه‌ها است. در این شیوه مختصات رئوس چند ضلعی‌ها مشخص شده و با استفاده از آن، بردار متعامد بر چندضلعی بدست می‌آید. هدف این بخش تعیین شدت نور نقاط درونی چند ضلعی‌ها با استفاده از مدل نوردهی می‌باشد که به سه روش مهم قابل انجام است. این سه روش که روش‌های سایه‌گذاری<sup>۱۰۵</sup> نامیده می‌شوند، در این بخش تشریح می‌شوند.

#### سایه‌گذاری ثابت

سریع‌ترین و ساده‌ترین روش تعیین شدت نور نقاط درونی چند ضلعی‌ها، استفاده از روش سایه‌گذاری ثابت و یا سایه‌گذاری مسطح<sup>۱۰۶</sup> می‌باشد. در این روش تمام نقاط روی سطح چندضلعی با شدت نور یکسان نشان داده می‌شوند. شدت نور نقطه‌ی میانی چندضلعی گزینه‌ی مناسبی جهت گسترش روی سطح تمام چندضلعی می‌باشد. این روش جهت نمایش سریع رویه‌های منحنی مورد استفاده قرار می‌گیرد. استفاده از این روش تنها زمانی نتایج دقیقی ایجاد می‌کند که شیء مورد نظر یک چند وجهی باشد و فاصله‌ی منبع نور و ناظر از شیء در مقایسه با ابعاد چند ضلعی‌ها بسیار زیاد باشد تا بتوان بردارهای  $R$  و  $V$  را برای تمام نقاط روی چندضلعی ثابت فرض کرد.

#### سایه‌گذاری گوراود

---

<sup>105</sup> Shading

<sup>106</sup> Flat Shading

این روش مبتنی بر درون‌یابی شدت نور در نقاط روی سطح چند ضلعی می‌باشد. روش سایه‌گذاری گوراود دارای سه مرحله می‌باشد.

- تعیین بردار متعامد میانگین در رئوس چندضلعی
  - اعمال مدل نوردهی در رئوس چند ضلعی و تعیین شدت نور در رئوس
  - درون‌یابی خطی شدت نور نقاط روی سطح چندضلعی با استفاده از شدت نور رئوس
- بردار متعامد میانگین در هر رأس چندضلعی با استفاده از میانگین‌گیری از بردارهای متعامد تمامی چندضلعی‌های مجاور رأس بدست می‌آید. پس از محاسبه‌ی بردار متعامد هر رأس می‌توان با اعمال مدل نوردهی، شدت نور رأس را محاسبه کرد. در این مرحله با مشخص بودن شدت نور رئوس، شدت نور سایر نقاط روی سطح چند ضلعی با استفاده از درون‌یابی خطی بدست می‌آید. درون‌یابی با استفاده از روش خط پویش صورت می‌گیرد. این روش در بخش ۲-۳-۲ تشریح شده است. کافی است بجای عمق هر نقطه، از شدت نور استفاده شود.

#### سایه‌گذاری فانگ

روش دقیق‌تر تعیین شدت نور نقاط روی سطح چندضلعی، بر اساس درون‌یابی بردارهای متعامد به ازای تمامی نقاط روی سطح و اعمال مدل نوردهی می‌باشد. این روش که توسط فانگ ارائه شده است به نام روش "سایه‌گذاری درونیابی بردارهای متعامد" نیز شناخته می‌شود. روش سایه‌گذاری فانگ نیز دارای سه مرحله می‌باشد.

- تعیین بردار متعامد میانگین در رئوس چندضلعی
- درون‌یابی خطی بردار متعامد نقاط روی سطح چندضلعی با استفاده از بردارهای رئوس

- اعمال مدل نوردهی بر روی تمام نقاط روی سطح چند ضلعی و تعیین شدت نور در هر نقطه

مرحله اول این روش برابر با مرحله اول روش سایه‌گذاری گوراود می‌باشد. درون‌یابی بردارهای متعامد نیز با استفاده از روش خط پویش مطرح شده در بخش ۲-۳-۲ به سادگی صورت می‌گیرد. اما اعمال مدل نوردهی بر روی تمامی نقاط درون‌یابی شده، حجم محاسبات این روش دقیق را به شدت افزایش می‌دهد. با این حال یکنواختی و نرمی قابل ملاحظه در نتایج این روش، انگیزه‌ی لازم برای استفاده از این روش در کاربردهای دقیق را ایجاد می‌کند. شکل ۲-۲۰ نتایج پیاده‌سازی سه روش سایه‌گذاری مذکور بر روی یک مدل سه‌بعدی چهره را نشان می‌دهد. همانطور که ملاحظه می‌شود تفاوت قابل توجهی بین نتایج روش فانگ و روش گوراود وجود ندارد.



شکل ۲-۲۰ - نتایج سه روش سایه‌گذاری، (الف): سایه‌گذاری مسطح، (ب): سایه‌گذاری گوراود، (ج) سایه‌گذاری فانگ

همانطور که در فصل ۱ اشاره شد، به علت مزایای نسبی مدل میان‌ساز، این مدل و تشخیص چهره بر اساس آن، بخش اصلی این پایان‌نامه را شامل می‌شود. پایگاه داده‌ی سه‌بعدی چهره نیز یکی از مهمترین و تأثیرگذارترین بخش‌ها در ساخت مدل میان‌ساز به شمار می‌رود. نحوه‌ی ساخت و ذخیره‌سازی داده‌های موجود در پایگاه، از موارد تعیین‌کننده‌ی کیفیت پایگاه داده و در نتیجه کیفیت مدل میان‌ساز است. این فصل از پایان‌نامه به بررسی انواع تصاویر سه‌بعدی چهره، روش‌های ذخیره‌سازی اطلاعات سه‌بعدی، بررسی یک روش ساخت مدل‌های سه‌بعدی چهره از تصاویر دوبعدی و بررسی چند پایگاه داده‌ی چهره می‌پردازد. اغلب مطالب این فصل در پیاده‌سازی سایر روش‌های سه‌بعدی تشخیص و سنتز چهره نیز، قابل استفاده می‌باشند.

### ۱۳-۱ روش‌های تهیه تصاویر سه‌بعدی چهره

تهیه‌ی تصاویر سه‌بعدی چهره شامل دو بخش تهیه‌ی اطلاعات هندسی چهره و تهیه‌ی اطلاعات مربوط به بافت یا رنگ چهره می‌شود. علاوه بر این، تکنیک‌های مربوط به مثلث‌سازی<sup>۱۰۷</sup> و یا چندضلعی‌سازی<sup>۱۰۸</sup> اطلاعات سه‌بعدی از اهمیت بالایی به خصوص در کاربردهایی چون پویانمایی چهره، برخوردار هستند. گذشته از روش‌های کم‌دقت ابتدایی، امروزه دو روش مهم جهت تهیه‌ی تصاویر سه‌بعدی مورد استفاده قرار می‌گیرند. روش اول روش مبتنی بر لیزر بوده و روش دوم روش هندسه‌یابی تصویری<sup>۱۰۹</sup> می‌باشد. تهیه‌ی تصاویر با استفاده از روش مبتنی بر لیزر همچون استفاده از پویسگر لیزری Cyberware™ [Cyberware] از دقت بسیار بالایی برخوردار بوده و می‌تواند اطلاعات هندسی و بافت چهره را همزمان و

<sup>107</sup> Triangulation

<sup>108</sup> Polygonizing

<sup>109</sup> Photogrammetry

بدون جابجایی در چند ده ثانیه تولید کند. البته کاربرد اصلی این روش در تهیه تصاویر سه بعدی اجسام غیرمتحرک می باشد و در تهیه تصاویر چهره، مقداری خطا به علت حرکت ناخواسته سر وجود دارد. در روش هندسه یابی تصویری، اطلاعات هندسی و بافت چهره در مراحل جداگانه با استفاده از حداقل دو تصویر دوبعدی چهره بدست می آید. در این روش تصاویر دوبعدی به طور همزمان و در زوایای مختلف برداشته شده و با استفاده از روش های ارتباط تصویری کامپیوتری بیان شده در فصل دو، مختصات فضایی نقاط بدست می آید. جهت دستیابی به دقت مناسب، اغلب این روش ها نیازمند مشارکت کاربر جهت تعیین تناظر بین نقاط می باشند.

### ۱-۱۳-۱ تهیه مدل با استفاده از پویش لیزری

پویشگرهای لیزری مشخصات یک شیء را توسط حرکت سنسورها در یک مسیر دایره ای در اطراف شیء اندازه گیری می کنند. این فرایند منجر به تولید یک توریننه<sup>۱۱۰</sup> منظم بسیار بزرگ از مقادیر اندازه گیری شده رویه ی چهره، شامل رنگ و مختصات هندسی در دستگاه مختصات استوانه ای می شود. نمونه ای از اطلاعات چهره های پویش شده توسط این روش در شکل ۱-۳ نشان داده شده است. تصویر (الف) اطلاعات مربوط به عمق در دستگاه مختصات استوانه ای را شامل می شود. تصویر (ب) اطلاعات مربوط به رنگ را در همین دستگاه نشان می دهد. محورهای مختصات افقی و عمودی به ترتیب پارامترهای  $\phi$  و  $h$  در دستگاه مختصات استوانه ای هستند.

---

<sup>110</sup> Mesh





شکل ۳-۱ - تصاویر سه بعدی چهره شماره ۵ در دستگاه مختصات استوانه‌ای توسط پوشگر لیزری [MPI]، (الف): تصویر عمق، (ب): تصویر بافت

یکی از مشکلات پوشگرهای لیزری، وجود حفره در اطلاعات خروجی است. این حفره‌ها که بر اثر انسداد یا مشکلات فنی سنسورها و یا بافت خاص بخشی از صورت بوجود می‌آیند، بیشتر در بخش‌هایی چون زیر چانه، اطراف چشم، زیر بینی و بخش‌های دارای مو مشاهده می‌شوند. به منظور استفاده از این روش می‌بایست حفره‌ها در تصاویر خروجی با استفاده از درون‌یابی پر شوند که البته به علت بزرگی ابعاد حفره‌ها، نتایج متفاوتی نسبت به حالت واقعی خواهد داشت. شکل ۳-۲ نمونه‌ای از حفره‌های پر شده در تصاویر موجود در پایگاه داده‌ی MPI [MPI] را نشان می‌دهد.



شکل ۳-۲ - حفره‌های پر شده در تصویر شماره ۵ پایگاه داده‌ی MPI

مشکل دیگر در تصاویر پوشگرهای لیزری، تولید تورینه‌های منظم می‌باشد. پوشگرهای لیزری حجم بسیار بالایی از اطلاعات را در یک رزولوشن فضایی ثابت تولید می‌کنند. معمولاً داده‌هایی بیش از میزان مورد نیاز در بخش‌های مسطح چهره همچون گونه‌ها، پیشانی و پشت سر تولید می‌شوند. در صورت

کم کردن رزولوشن، داده‌های تولید شده برای بخش‌های مهم چهره همچون چشم‌ها و اطراف آن، بینی و لب‌ها ناکافی می‌باشد. روش مورد استفاده جهت حل این مشکل، کم کردن حجم داده‌ها با توجه به میزان خمیدگی<sup>۱۱۱</sup> در روبه‌ی چهره می‌باشد. با این وجود جهت استفاده در کاربردهایی چون پویانمایی و پیاده‌سازی حالات چهره، داده‌های پویش شده می‌بایست بر توپولوژی چهره منطبق باشند. در غیر این صورت اطلاعاتی در مورد این که هر چندضلعی فضایی مربوط به کدام بخش از صورت می‌باشد، وجود نخواهد داشت. این انطباق بر توپولوژی می‌بایست هم در مورد هندسه‌ی چهره و هم در مورد بافت چهره مد نظر قرار گیرد. برای مثال اگر یک چندضلعی واقع در چشم شامل بخش‌هایی از عنبیه و کره‌ی چشم شود، بنا بر مطالب مذکور در بخش ۲-۴-۳، امکان ایجاد لبه‌ی تیز در این ناحیه وجود نخواهد داشت. جهت ایجاد چندضلعی‌های منطبق بر توپولوژی چهره روش‌هایی خودکار بر اساس شناخت از آناتومی چهره ارائه شده‌اند [الی و همکاران، ۱۹۹۵].

### ۱-۱۳-۲ تهیه‌ی مدل با استفاده از هندسه‌یابی تصویری

در این روش که یک نمونه از آن را به تفصیل در بخش ۴-۳ تشریح خواهیم کرد، تصاویر دوبعدی و یا تصاویر ویدئویی به عنوان ورودی استفاده می‌شوند. تناظر بین نقاط موجود در تصاویر ورودی می‌بایست با دقت بالا تعیین شود. بر اساس این تناظر و در صورت مشخص بودن پارامترهای تصویربرداری همچون موقعیت دقیق دوربین و فاصله‌ی کانونی عدسی، تخمینی از هندسه‌ی فضایی چهره بدست می‌آید. البته باید توجه کرد که در روبه‌های نرمی همچون چهره، تعیین تناظر بین نقاط تصاویر دوبعدی به صورت خودکار،

---

<sup>111</sup> Curvature

فرایندی بسیار مشکل و دارای حجم محاسباتی بالا بوده و نتایجی با دقت پایین تولید می‌کند. برای رفع این مشکل معمولاً از همکاری کاربر جهت تعیین تناظر بین چند نقطه‌ی راهنما استفاده می‌شود.

جهت ساخت مدل سه بعدی با استفاده از روش هندسه‌یابی تصویری ساده، تنها به دو تصویر دوبعدی همزمان از چهره نیاز داریم. معمولاً این دو تصویر در زوایای دید متعامد انتخاب می‌شوند و تصاویر نیم‌رخ و تمام‌رخ بهترین گزینه‌های ممکن هستند. هر یک از این دو تصویر دو بعد از سه بعد لازم جهت ساخت هندسه‌ی فضایی را در اختیار می‌گذارند. البته این روش به علت نداشتن اطلاعات مربوط به بخش‌هایی از چهره و نیاز به ساخت بافت چهره‌ی سه‌بعدی تنها از دو تصویر دوبعدی دارای نواقصی نیز می‌باشد.

روش هندسه‌یابی تصویری در مقایسه با روش‌های مبتنی بر پویش لیزری به علت عدم نیاز به تجهیزات گران‌قیمت و سرعت بالاتر در تهیه‌ی اطلاعات اولیه‌ی مورد نیاز دارای مزیت می‌باشد. دقت بالاتر در اطلاعات مربوط به بافت چهره به علت رزولوشن نسبی بالاتر دوربین‌های تصویربرداری نیز از دیگر مزایای این روش می‌باشد. با این حال دقت در تعیین مختصات فضایی نقاط به علت نیاز به تعیین دقیق تناظر بین نقاط در تصاویر مختلف، در این روش نسبت به روش پویش لیزری پایین‌تر می‌باشد.

نکته‌ی قابل توجه در این روش کاهش مشکلات مربوط به چندضلعی‌سازی در صورت استفاده از مدل عمومی منطبق بر توپولوژی چهره می‌باشد. از آنجایی که در برخی از الگوریتم‌های هندسه‌یابی تصویری، مدل چهره از تغییر مختصات نقاط یک مدل عمومی ساخته می‌شود، در صورت انطباق این مدل عمومی بر توپولوژی چهره، نیاز به چندضلعی‌سازی مجدد در مدل چهره‌ی جدید وجود نداشته و در صورت برخورداری از دقت بالا در تعیین تناظر بین نقاط در تصاویر دوبعدی، تورینده‌ی مدل چهره‌ی جدید نیز بر توپولوژی چهره

منطبق خواهد بود. البته باید توجه کرد که بالا بودن رزولوشن فضایی نیز تا حد قابل توجهی مشکلات مربوط به چندضلعی سازی و عدم انطباق بر توپولوژی چهره را کاهش می دهد.

## ۱-۱۴ روش های ذخیره سازی اطلاعات سه بعدی

جهت ذخیره سازی اطلاعات سه بعدی روش های گوناگونی پیشنهاد شده اند. اغلب این روش ها توسط شرکت های تولیدکننده نرم افزارهای سه بعدی ارائه شده اند. جهت استفاده از پایگاه های داده سه بعدی موجود همچون MPI و GavabDB می بایست با فرمت های ذخیره سازی اطلاعات سه بعدی آشنایی داشت. علاوه بر این تولید پایگاه های داده ی سه بعدی قابل استفاده برای عموم نیز نیاز به بکارگیری فرمت های استاندارد ذخیره سازی سه بعدی دارد. دو فرمت بسیار مهم ذخیره سازی اطلاعات سه بعدی، فرمت های "زبان مدل سازی حقایق مجازی"<sup>۱۱۲</sup> یا VRML و فرمت فایل های شیئی<sup>۱۱۳</sup> می باشند. تصاویر موجود در پایگاه داده ی سه بعدی GavabDB تحت فرمت VRML و تصاویر موجود در پایگاه داده ی MPI تحت فرمت فایل های شیئی ذخیره سازی شده اند. در این بخش، قسمت هایی از این دو فرمت ذخیره سازی اطلاعات سه بعدی که در پایگاه های داده ی چهره ی سه بعدی استفاده می شوند، معرفی خواهند شد.

### ۱-۱۴-۱ زبان مدل سازی حقایق مجازی (VRML)

همان طور که از نام این روش ذخیره سازی اطلاعات سه بعدی مشخص است، این روش تنها یک فرمت ذخیره سازی نبوده بلکه زبانی برای نمایش دنیای مجازی می باشد. این زبان که مطرح ترین و پرکاربردترین زبان نمایش اطلاعات سه بعدی است، دارای جزئیات زیادی می باشد که ما در این بخش به تشریح موارد مورد استفاده در نمایش اشیاء ایستا بسنده می کنیم. قابل توجه است که این زبان که بخش

---

<sup>112</sup> Virtual Reality Modeling Language

<sup>113</sup> Object Files

اصلی اغلب نرم‌افزارهای گرافیکی سه‌بعدی را شامل می‌شود، قابلیت پیاده‌سازی مدل‌های مختلف نوردهی، ایجاد سایه‌ها، ایجاد انواع مختلف حرکت اشیاء و دوربین، تعامل با کاربر از طریق ماوس و صفحه‌کلید، تشخیص رویدادهای مختلف و تمامی قابلیت‌های مورد نیاز جهت نمایش دنیای سه‌بعدی را دارا می‌باشد. این زبان که اولین بار در ۱۹۹۴ ارائه شد، امروزه به فرمت استاندارد نمایش اطلاعات سه‌بعدی در اینترنت تبدیل شده و معادل گرافیکی زبان HTML قلمداد می‌شود. مرجع کامل این زبان در [VRML] موجود می‌باشد. مبانی مورد نیاز از این زبان جهت نمایش مدل‌های ایستا در پیوست (الف) ارائه شده است.

#### ۱-۱۴-۲ فایل‌های شیئی

این نوع فایل‌ها بر خلاف VRML تنها یک فرمت ذخیره‌سازی اطلاعات سه‌بعدی بوده و از ساختار به مراتب ساده‌تری برخوردار می‌باشد. در حال حاضر دو ساختار استاندارد برای این نوع فایل‌ها موجود می‌باشد. ساختار استاندارد اول Wavefront™ و استاندارد دوم Java3D نام دارند. فایل‌های شیئی نوع Wavefront علاوه بر ذخیره‌سازی سطوح مبتنی بر چندضلعی‌ها، قابلیت ذخیره‌سازی و بازیابی رویه‌ها و خم‌های پارامتری را نیز دارا می‌باشند، در حالی که فایل‌های شیئی نوع دوم، تنها سطوح مبتنی بر چندضلعی‌ها را ذخیره و بازیابی می‌کنند. تمام داده‌های موجود در پایگاه‌های داده‌ی چهره‌ی سه‌بعدی مورد مطالعه که از فایل‌های شیئی جهت ذخیره‌سازی اطلاعات استفاده کرده‌اند، از نوع سطوح مبتنی بر چندضلعی‌ها می‌باشند. بنابراین در این بخش تنها مطالب مربوط به این قسمت بررسی می‌شود.

در فایل‌های شیئی نمایش هر بخش از اطلاعات توسط حروف نشانه‌ای صورت می‌گیرد که در ذیل به آنها اشاره می‌شود.

# : این علامت جهت معرفی آغاز توضیحات متنی بکار می‌رود. تمام عبارات بعد از این علامت تا

انتهای خط به عنوان توضیحات متنی در نظر گرفته می‌شوند.

float float float : حرف v جهت معرفی یک مختصات فضایی یک رأس بکار می‌رود. سه عدد

اعشاری بعد از حرف v مختصات این رأس در دستگاه مختصات کارتزین را بیان می‌کنند. اولین رأسی که در

فایل به این صورت نشان داده می‌شود دارای اندیس ۱ می‌باشد و سایر رئوس به ترتیب اندیس‌های بعدی را

به خود اختصاص می‌دهند.

float float float vn : عبارت vn جهت معرفی یک بردار متعامد در فضای سه بعدی بکار می‌رود.

سه عدد اعشاری بعد از عبارت vn مؤلفه‌های این بردار را در دستگاه مختصات کارتزین بیان می‌کنند. اولین

بردار متعامد نیز دارای اندیس ۱ بوده و سایر بردارها به ترتیب اندیس‌های بعدی را به خود اختصاص خواهند

داد.

float float vt : این عبارت جهت معرفی یک پیکسل از فایل تصویری مرتبط با فایل شیئی مورد

استفاده قرار گرفته و مختصات بافت نامیده می‌شود. دو اعداد اعشاری بعد از عبارت vt بیانگر موقعیت

پیکسل مورد نظر در فایل تصویری دوبعدی می‌باشد. موقعیت پیکسل به صورت یک عدد نرمالیزه شده بین

۰ و ۱ مطرح می‌شود. اولین مختصات بافت دارای اندیس ۱ و سایر بافت‌ها به ترتیب اندیس‌های بعدی را به

خود اختصاص می‌دهند.

... int int int f : این عبارت جهت معرفی یک وجه و یا یک چندضلعی به کار می‌رود. اعداد بعد از

حرف f بیانگر اندیس نقاطی هستند که رئوس این چندضلعی را تشکیل می‌دهند. هیچ محدودیتی در حد

بیشینه‌ی تعداد رئوس تشکیل دهنده‌ی چندضلعی‌ها وجود ندارد. تنها الزامات مهم تحدب چندضلعی‌ها و قرارگیری تمام رئوس بر روی یک صفحه می‌باشند.

$f \text{ int/int int/int int/int } \dots$  : این عبارت نیز مانند عبارت قبلی مشخص کننده‌ی یک وجه در فضا

می‌باشد. تفاوت موجود در مشخص بودن بافت رئوس می‌باشد. اعداد اول اندیس نقاط تشکیل دهنده‌ی رئوس و اعداد دوم اندیس مختصات بافت رئوس می‌باشند.

$f \text{ int/int/int int/int/int int/int/int } \dots$  : در این عبارت اعداد سوم بیانگر بردارهای متعامد رئوس

می‌باشند. این بردارها که به سادگی قابل محاسبه نیز می‌باشند، موجب تسریع پیاده‌سازی مدل نوردهی بر روی چندضلعی می‌شوند.

بخشی از فایل شیئی یکی از داده‌های پایگاه داده‌ی MPI به صورت زیر می‌باشد. ملاحظه می‌شود که برای هر رأس هر سه مقدار مختصات فضایی، مختصات بافت و بردار متعمد مشخص شده است.

```
#Converter by Volker Blanz, 25.10.2000 mtl lib average.mtl use mtl ss g
v -7749.206055 118080.000000 89986.953125
vn -0.142694 0.604331 0.783851
vt 0.486328 511.875000
v -6649.048828 118080.000000 90139.101562
vn -0.111444 0.638840 0.761225
vt 0.488281 511.875000
v 5542.415527 118080.000000 90213.906250
vn 0.116233 0.621705 0.774579
vt 0.482422 511.873047
```

...

```
# 75872 vertices. g all_triangles
```

```
f 1/1/1 7/7/7 8/8/8
```

```
f 1/1/1 8/8/8 9/9/9
```

```
f 1/1/1 9/9/9 2/2/2
```

## ۱۵-۱ ساخت مدل‌های سه‌بعدی چهره با استفاده از روش هندسه‌یابی تصویری

در این بخش روش تهیه‌ی مدل‌های سه‌بعدی چهره با استفاده از تصاویر دوبعدی را که توسط پیگهین و همکاران [پیگهین و همکاران، ۱۹۹۸] ارائه شده است، و نتایج حاصل از پیاده‌سازی آن را بررسی می‌کنیم. بررسی این روش به دو دلیل صورت گرفته است. اول آن که این روش نمونه‌ای از روش‌های مبتنی بر مدل عمومی می‌باشد. همانطور که در فصل ۱ اشاره شد، این روش‌ها یکی از دو دسته‌ی روش‌های تشخیص چهره‌ی سه‌بعدی مبتنی بر مدل را تشکیل می‌دهند [انصاری و عبدالمطلب، (ب) ۲۰۰۳]. بنابراین بررسی این روش، منجر به آشنایی با این دسته‌ی مهم از روش‌های تشخیص چهره‌ی مبتنی بر مدل می‌شود. دلیل دوم نیاز به در اختیار داشتن یک پایگاه داده‌ی چهره‌ی سه‌بعدی دارای بافت جهت استفاده در مدل میان‌ساز می‌باشد. پایگاه داده‌ی MPI، یک پایگاه داده‌ی چهره‌ی سه‌بعدی دارای بافت می‌باشد. اما از این پایگاه داده به ازای هر شخص، ما تنها ۷ تصویر دوبعدی در زوایای  $0, \pm 30, \pm 60, \pm 90$  در اختیار داشتیم. بنابراین جهت ساخت مدل سه‌بعدی چهره نیازمند استفاده از روش هندسه‌یابی تصویری هستیم. علاوه بر این، این روش امکان ساخت مدل‌های سه‌بعدی جدید از چهره‌های ایرانی را نیز فراهم می‌کند که در انتهای این فصل، این پایگاه داده و تأثیر آن بر کیفیت سنتز را تشریح خواهیم کرد.

روش ساخت مدل چهره با استفاده از تصاویر دوبعدی در مراجع [ایپ و یین<sup>۱۱۴</sup>، ۱۹۹۶ - کوریهارا و آرای<sup>۱۱۵</sup>، ۱۹۹۱] نیز ارائه شده است. این روش‌ها از تعداد کمی نقطه راهنما جهت تعیین تناظر بین نقاط استفاده کرده‌اند که کیفیت مدل‌های تولید شده را تحت تأثیر قرار داده است. پیگهین و همکاران با استفاده از منحنی‌ها تعداد این نقاط را افزایش داده و مدلی با کیفیت بالا تولید کرده‌اند. علاوه بر این تخمین

---

<sup>114</sup> Ip, Yin

<sup>115</sup> Kurihara, Arai



پارامترهای تصویر برداری با استفاده از نقاط راهنما باعث می‌شود تا به تنظیم دقیق موقعیت دوربین‌ها نیاز نباشد. این روش شامل دو مرحله‌ی اساسی تطبیق مدل و استخراج بافت می‌باشد. هر یک از این مراحل با جزئیات مربوط به پیاده‌سازی آن در این بخش تشریح می‌شود.

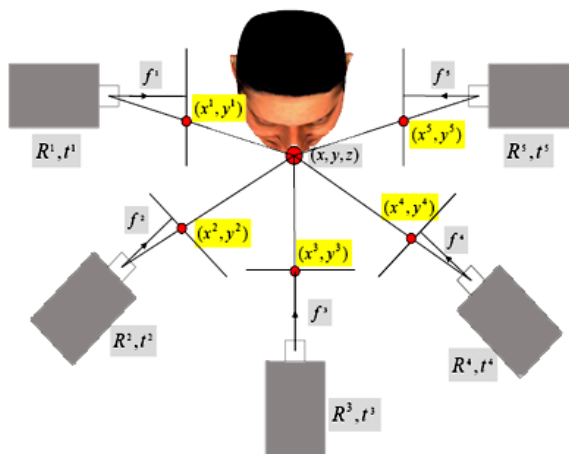
### ۱-۱۵-۱ تطبیق مدل

تطبیق مدل فرایندی است که در آن یک مدل عمومی به گونه‌ای تغییر می‌یابد که متشابه مدل چهره‌ی ناشی از تصاویر ورودی باشد. به عبارت دیگر هدف تغییر مدل عمومی به نحوی است که تصاویر دوبعدی استخراج شده از آن مشابه تصاویر ورودی باشد. روش کار به این ترتیب است که ۱۳ نقطه راهنما بر روی مدل عمومی مشخص می‌شوند. کاربر بر روی تمام تصاویر ورودی این ۱۳ نقطه را در صورت عدم انسداد مشخص می‌کند. تعداد تصاویر ورودی ۵ عدد در نظر گرفته شده است (شکل ۳-۳). البته با افزایش تعداد نقاط راهنما می‌توان تعداد تصاویر ورودی را کاهش داد. ما با توجه به این نکته از دو تصویر به عنوان تصاویر ورودی استفاده کرده‌ایم. این موضوع در بخش بعدی تشریح می‌شود. با تشکیل معادلات هندسی مربوط به تصویرسازی موقعیت فضایی نقاط راهنما به همراه پارامترهای مربوط به موقعیت دوربین‌ها معین می‌شوند. این مرحله تخمین موقعیت نامیده می‌شود. سپس نقاطی از مدل عمومی که در محل نقاط راهنمای مشخص شده واقع شده‌اند، به موقعیت بدست آمده در مرحله‌ی قبل انتقال پیدا می‌کنند. در مرحله‌ی بعدی سایر نقاط مدل عمومی نیز با استفاده از روش درون‌یابی داده‌های پراکنده<sup>۱۱۶</sup> به موقعیت فضایی جدید انتقال پیدا کرده و مدل چهره‌ی تطبیق یافته ساخته می‌شود. اما به علت پایین بودن تعداد نقاط راهنما اختلاف این مدل جدید با مدل سه‌بعدی واقعی زیاد می‌باشد. این نقص در مرحله‌ی پالایش

---

<sup>116</sup> Scattered Data Interpolation

شکل ۱۱۷ با استفاده از منحنی‌های راهنما رفع می‌شود. مراحل تطبیق مدل با جزئیات مربوط به پیاده‌سازی در این بخش ارائه می‌شوند.



شکل ۳-۳ - محل قرارگیری دوربین‌ها و تبدیلات هندسی در هندسه‌یابی تصویری

### تخمین موقعیت

مسئله‌ی تخمین موقعیت با مشخص بودن روابط مربوط به انتقال دستگاه‌های مختصات و تصویرسازی پرسپکتیو، به یک مسئله‌ی بهینه‌سازی تبدیل می‌شود که روش حداقل مربعات<sup>۱۱۸</sup> جهت حل آن بکار گرفته شده است.

طبق مطالب مطرح شده در فصل ۲، ساخت یک تصویر دوبعدی از یک مدل سه‌بعدی در دو مرحله صورت می‌گیرد. مرحله‌ی اول انتقال دستگاه مختصات مرجع به دستگاه مختصات دید، و مرحله‌ی دوم تصویرسازی جهت نگاشت شیء سه‌بعدی به صفحه‌ی دید می‌باشد. صفحه‌ی دید در تصویربرداری، صفحه‌ای است که شامل فیلم یا CCD دوربین بوده و دستگاه مختصات دید دستگاهی است که دارای محور z متعامد بر صفحه‌ی دید باشد. مختصات نقاط در دستگاه مختصات مرجع با استفاده از تبدیلات متوالی انتقال و دوران به مختصات در دستگاه دید تبدیل می‌شوند. انتقال برابر فاصله‌ی مبداء دستگاه مختصات دید تا

<sup>117</sup> Shape Refinement

<sup>118</sup> Least Squares

دستگاه مختصات مرجع و دوران برابر با میزان دوران دستگاه مختصات دید نسبت به دستگاه مختصات مرجع می‌باشد. در صورتی که دوران را با  $R$  و انتقال را با  $T$  نشان دهیم، مختصات نقطه‌ی  $p_i$  در دستگاه مرجع به صورت زیر به نقطه‌ی  $p'_i$  در دستگاه مختصات دید تبدیل می‌شود.

$$p'_i = R.T.p_i \quad (2-3)$$

ماتریس  $t$  را به صورت یک ماتریس  $3 \times 1$  و برابر با درایه‌های واقع بر سطرهای ۱ تا ۳ در ستون ۴ ماتریس  $T$  در نظر می‌گیریم. ماتریس  $r$  را نیز به صورت یک ماتریس  $3 \times 3$  و برابر با درایه‌های واقع در ستون‌های ۱ تا ۳ و سطرهای ۱ تا ۳ در ماتریس  $R$  در نظر می‌گیریم. حال می‌توان معادله‌ی (۲-۳) را به صورت زیر بیان کرد.

$$p'_i = r.p_i + t \quad (3-3)$$

روابط حاکم بر رفتار عدسی‌های محدب بیانگر تصویرسازی پرسپکتیو در این عدسی‌ها و در نتیجه در دوربین‌های تصویربرداری می‌باشد.

$$\left. \begin{array}{l} \frac{1}{p} + \frac{1}{q} = \frac{1}{f} \\ \frac{q}{p} = \frac{A'B'}{AB} \\ p = z, A'B' = x^k, AB = x \end{array} \right\} \Rightarrow q = \frac{x^k z}{x} \Rightarrow \frac{1}{z} + \frac{x}{zx^k} = \frac{1}{f} \Rightarrow x^k = -\frac{f}{f-z}x \quad (4-3)$$

با در نظر گرفتن  $f = z_{prp} - z_{vp}$  و  $z_{vp} = 0$  و طبق رابطه‌ی (۲-۳۶)، رابطه‌ی فوق برابر با تصویرسازی پرسپکتیو می‌باشد. علامت منفی نیز به علت وارونگی تصویر در عدسی‌های محدب می‌باشد. بنا بر معادلات (۳-۳) و (۴-۳)، دو معادله‌ی زیر ارتباط بین مختصات نقطه در تصویر و مختصات فضایی نقطه را مشخص می‌کنند.

$$x_i^k = f^k \frac{r_x^k \cdot p_i + t_x^k}{r_z^k \cdot p_i + t_z^k} \quad y_i^k = f^k \frac{r_y^k \cdot p_i + t_y^k}{r_z^k \cdot p_i + t_z^k} \quad (5-3)$$

در معادلات فوق بالانویس‌های  $k$  بیانگر شماره‌ی تصویر دوبعدی مورد نظر هستند. حال با در نظر گرفتن

$$\eta^k = \frac{1}{t_z^k} \text{ و } s^k = \frac{f^k}{t_z^k} \text{ و جایگزینی در روابط فوق، دو رابطه‌ی زیر بدست می‌آیند.}$$

$$x_i^k = s^k \frac{r_x^k \cdot p_i + t_x^k}{1 + \eta^k r_z^k \cdot p_i} \quad y_i^k = s^k \frac{r_y^k \cdot p_i + t_y^k}{1 + \eta^k r_z^k \cdot p_i} \quad (6-3)$$

در روابط فوق  $s^k$ ، نسبت تغییر مقیاس شیء به تصویر می‌باشد. مزیت استفاده از این تغییر متغیر این است

که تخمین مقدار  $s^k$  جهت تعیین مقدار اولیه به مراتب راحت‌تر از تخمین  $f^k$  صورت می‌گیرد. حال با در

نظر گرفتن  $w_i^k = \frac{1}{1 + \eta^k r_z^k \cdot p_i}$  و چند محاسبه‌ی ساده به روابط نهایی زیر می‌رسیم.

$$\begin{cases} w_i^k (x_i^k + x_i^k \eta^k (r_z^k \cdot p_i) - s^k (r_x^k \cdot p_i + t_x^k)) = 0 \\ w_i^k (y_i^k + y_i^k \eta^k (r_z^k \cdot p_i) - s^k (r_y^k \cdot p_i + t_y^k)) = 0 \end{cases} \quad (7-3)$$

در دو معادله‌ی فوق پارامترهای  $p_i$ ،  $t^k$ ،  $\eta^k$  و  $s^k$  و  $R^k$  مجهول بوده و تنها پارامترهای معلوم  $x_i^k$  و  $y_i^k$  می

باشند که بر حسب پیکسل بر روی تصاویر دو بعدی اندازه‌گیری می‌شوند. علاوه بر این، این دو معادله بر

حسب هر یک از پارامترهای مجهول، معادلاتی خطی می‌باشند. با در نظر گرفتن ۵ تصویر ورودی، تعداد

معادلات و مجهولات را می‌توان به صورت زیر محاسبه کرد.

با فرض عدم انسداد ۱۳ نقطه‌ی راهنما در تمام تصاویر (شکل ۳-۴ الف)،  $5 \times 13 \times 2 = 130$  معادله

داریم. البته به علت انسداد برخی از نقاط راهنما در تصاویر، تعداد واقعی معادلات برابر  $45 \times 2 = 90$  معادله

می‌باشد. به ازای هر تصویر ۴ مجهول  $t_x^k$ ،  $t_y^k$ ،  $\eta^k$  و  $s^k$  و سه مجهول مربوط به پارامترهای  $R^k$  یعنی (

$\theta_x, \theta_y, \theta_z$ ) وجود دارند.  $13 \times 3$  مجهول نیز مربوط به موقعیت فضایی نقاط راهنما می‌شوند. بنابراین مجموع

مجهولات برابر  $74 = 7 \times 5 + 13 \times 3 = 74$  عدد می‌باشد. از آنجایی که تعداد معادلات بیشتر از تعداد مجهولات است،

این دستگاه معادلات، یک دستگاه فوق معین<sup>۱۱۹</sup> می باشد که جوابهای آن توسط حل یک مسئلهی بهینه سازی بدست می آید. علاوه بر این باید توجه کرد که حذف پارامتر  $w_i^k$  در معادلات (۳-۶)، ممکن است پاسخهای مسئلهی بهینه سازی را تغییر دهد.

جهت یافتن پاسخهای این مسئلهی بهینه سازی از الگوریتم تکراری حداقل مربعات استفاده می کنیم. در این الگوریتم یکی از پارامترها را بهینه سازی کرده و سایر پارامترها را ثابت در نظر می گیریم. در هر مرحلهی تکرار، این فرایند را برای تمام پارامترهای مجهول تکرار می کنیم. ترتیب بهینه سازی پارامترها در هر مرحلهی تکرار به صورت  $s^k, p_i, R^k, t_x^k, t_y^k$  و  $\eta^k$  می باشد. این ترتیب به علت ایجاد بیشترین پایداری عددی بر اساس تخمینهای اولیه، انتخاب شده است.

الگوریتم بهینه سازی حداقل مربعات از ساده ترین روشهای بهینه سازی می باشد. در حالت کلی، مجموعه ای از معادلات خطی به صورت زیر می باشند.

$$a_j \cdot x - b_j = 0 \quad (۸-۳)$$

نمایش ماتریسی معادله فوق برابر است با:

$$A \cdot x = b \quad (۹-۳)$$

در رابطهی فوق  $A$  یک ماتریس  $m \times n$  می باشد. در یک دستگاه فوق معین،  $m > n$  است.  $x$  یک بردار  $n \times 1$  و  $b$  یک بردار  $m \times 1$  می باشد. با در نظر گرفتن ۵ تصویر ورودی،  $m$  برابر تعداد معادلات و  $n$  برابر تعداد مجهولات است. حال جهت بدست آوردن بردار مجهولات  $x$ ، معادلهی زیر را کمینه می کنیم.

$$\sum_j (a_j \cdot x - b_j)^2 \quad (۱۰-۳)$$

کمینه سازی معادلهی فوق معادل با برابر صفر قرار دادن مشتق آن می باشد.

<sup>119</sup> Over Determined

$$\left( \sum_j a_j a'_j \right) x = \sum_j b_j a_j \quad (11-3)$$

$$A'Ax = A'b$$

از آنجایی که دستگاه معادلات نسبت به هر یک از پارامترهای مجهول با ثابت در نظر گرفتن سایر مجهولات یک دستگاه خطی است، در هر یک از مراحل تکرار پارامترهای مجهول طبق ترتیب اشاره شده بر اساس رابطه‌ی فوق مقدار یابی می‌شوند. مراحل تکرار تا رسیدن جواب‌ها به پایداری ادامه پیدا می‌کند. برای مثال دستگاه معادلات برای پارامتر  $s^k$  به صورت زیر می‌باشد.

$$w_i^k (r_x^k \cdot p_i + t_x^k) s^k = w_i^k (x_i^k + x_i^k \eta^k (r_z^k \cdot p_i)) \quad (12-3)$$

$$w_i^k (r_y^k \cdot p_i + t_y^k) s^k = w_i^k (y_i^k + y_i^k \eta^k (r_z^k \cdot p_i))$$

در رابطه‌ی فوق تعداد مجهولات برابر ۱ و تعداد معادلات برابر  $2 \times 13$  می‌باشد. برای پارامتر  $p_i$  تعداد مجهولات برابر ۳ و تعداد معادلات برابر  $2 \times 5$  می‌باشد.

$$w_i^k (x_i^k \eta^k r_z^k - s^k r_x^k) p_i = w_i^k (s^k t_x^k - x_i^k) \quad (13-3)$$

$$w_i^k (y_i^k \eta^k r_z^k - s^k r_y^k) p_i = w_i^k (s^k t_y^k - y_i^k)$$

#### درون‌یابی داده‌های پراکنده

پس از تعیین موقعیت نقاط راهنما، نقاط راهنمای مدل عمومی به محل تعیین شده منتقل می‌شوند (شکل ۳-۴ (د)). حال می‌بایست سایر نقاط مدل عمومی را نیز بر اساس تغییر موقعیت نقاط راهنما، تغییر دهیم. این فرایند توسط یک روش درون‌یابی تحت عنوان درون‌یابی داده‌های پراکنده صورت می‌گیرد. مختصات نقاط راهنما در مدل عمومی را  $p_i^0$  می‌نامیم. این نقطه پس از بهینه‌سازی و تخمین موقعیت به نقطه‌ی  $p_i$  منتقل می‌شود. میزان جابجایی نقاط راهنما برابر  $u_i = p_i - p_i^0$  می‌باشد. هدف یافتن تابع  $f$  به نحوی است که به ازای تمام نقاط  $p_j$  مدل عمومی، میزان جابجایی  $u_j$  را معین کند. توابع متفاوتی همچون

روش شیپارد<sup>۱۲۰</sup> برای این منظور پیشنهاد شده‌اند. در این کاربرد از توابع پایه‌ی شعاعی<sup>۱۲۱</sup> به صورت زیر جهت درون‌یابی پخش شده استفاده شده است.

$$f(p) = \sum_i c_i \phi(\|p - p_i\|) + Mp + t \quad (۱۴-۳)$$

در رابطه‌ی فوق تابع پایه‌ی شعاعی  $\phi(r)$  می‌باشد. انتخاب  $\phi(r) = e^{\frac{-r}{64}}$  نتایج مناسبی ایجاد کرده است.  $c_i$  و  $t$  بردارهایی  $3 \times 1$  و  $M$  ماتریسی  $3 \times 3$  می‌باشد. بنابراین با فرض ۱۳ نقطه‌ی راهنما تعداد مجهولات برابر  $3 \times 13 + 12 = 3(13+4)$  می‌باشد. رابطه‌ی (۱۴-۲)، ۱۳ معادله‌ی ۳ پارامتری ایجاد می‌کند. سایر معادلات از دو شرط زیر تأمین می‌شوند.

$$\sum_i c_i = 0 \quad \sum_i c_i p'_i = 0 \quad (۱۵-۳)$$

باید دقت کرد که شرط اول ۳ معادله و شرط دوم ۹ معادله ایجاد می‌کند.

پس از تعیین پارامترهای مجهول و بدست آوردن  $f(p)$  تمام نقاط مدل عمومی بر اساس این تابع تغییر موقعیت داده و مدل چهره‌ی مورد نظر ایجاد می‌شود. ما این مدل را جهت مشخص بودن جزئیات پس از استخراج بافت در شکل ۳-۴ (د) نشان داده‌ایم.

### پالایش شکل

همانطور که در شکل ۳-۴ (د) مشخص است به علت پایین بودن تعداد نقاط راهنمای انتخاب شده، مدل تطبیق یافته اختلاف قابل توجهی از لحاظ مشخصات هندسی با مدل واقعی دارد. جهت رفع این مشکل می‌بایست تعداد نقاط راهنما را افزایش داد تا درون‌یابی داده‌های پراکنده با دقت بالاتری صورت گیرد. اما افزایش تعداد نقاط راهنما، به علت مشخص نبودن موقعیت دقیق آنها برای کاربر ایجاد مشکل می‌کند. به

<sup>120</sup> Shepard's Method

<sup>121</sup> Radial Basis Functions

همین دلیل از منحنی‌های راهنما به جای نقاط راهنما استفاده می‌شود. برای این منظور بر روی مدل عمومی تعدادی منحنی راهنما در نظر گرفته می‌شود. کاربرد می‌بایست این منحنی‌ها را بر روی هر یک از تصاویر ورودی مشخص کند (شکل ۳-۴ (ب)). جهت ساده‌سازی فرایند تعیین منحنی‌ها از توابع چندجمله‌ای درجه‌ی ۳ و روش منحنی‌های بی‌زیر<sup>۱۲۲</sup> استفاده شده است. در این روش هر منحنی درجه‌ی ۳ به جای استفاده از ضرایب چندجمله‌ای‌ها توسط چهار نقطه مشخص می‌شود. این چهار نقطه، نقاط ابتدایی، انتهای و دو نقطه جهت تعیین شیب منحنی در ابتدا و انتها را شامل می‌شوند. پس از مشخص شدن موقعیت منحنی‌های راهنما بر روی تصاویر ورودی، به ازای هر منحنی تعدادی نقطه‌ی راهنمای جدید استخراج شده و مراحل تخمین موقعیت و درونیابی داده‌های پراکنده بر اساس تعداد زیادی نقطه‌ی راهنما صورت می‌پذیرد. ما در پیاده‌سازی این روش تعداد ۱۱۰ نقطه‌ی راهنما از این منحنی‌ها استخراج کرده‌ایم. مدل تطبیق یافته‌ی نهایی پس از استخراج بافت در شکل ۳-۴ (ه) نشان داده شده است.

## ۱-۱۵-۲ استخراج بافت

در این مرحله می‌بایست برای هر نقطه‌ی  $p$  در مدل تطبیق یافته، بافت  $T(p)$  محاسبه شود. در مرحله‌ی اول می‌بایست تناظر بین پیکسل‌های تصاویر ورودی و مدل چهره‌ی تطبیق یافته مشخص شود. این تناظر با استفاده از موقعیت نقاط راهنما بر روی تصاویر ورودی و موقعیت نقاط راهنما بر روی مدل تطبیق یافته و استفاده از روش درونیابی داده‌های پراکنده حاصل می‌شود. کفایت تا تابع  $(x, y) = g(p)$  به نحوی انتخاب شود که هر نقطه‌ی  $p$  بر روی مدل را به نقطه‌ی متناظر بر روی تصویر ورودی تطبیق دهد. پس از تعیین تناظر به هر نقطه روی تصاویر ورودی ضریب وزنی جهت تعیین میزان مشارکت نقطه در تعیین بافت

<sup>122</sup> Bézier Curves



$T(p)$  انتساب داده می‌شود. این ضریب وزن را  $m(g(p))$  می‌نامیم. بافت هر نقطه از مدل به صورت میانگین

وزنی بافت نقاط متناظر در تصاویر ورودی محاسبه می‌شود.

$$T(p) = \frac{\sum_k m^k(g^k(p)) I^k(g^k(p))}{\sum_k m^k(g^k(p))} \quad (16-3)$$

$k$  در رابطه‌ی فوق بیانگر شماره‌ی تصویر ورودی و  $I(g^k(p))$  نیز بیانگر رنگ هر پیکسل در تصاویر ورودی

می‌باشد.  $g^k(p)$  پیکسل متناظر با  $p$  در تصویر ورودی است.

برای بدست آوردن  $m^k(g^k(p))$  دو شرط مهم را در نظر می‌گیریم.  $m^k(g^k(p))$  می‌بایست برای

نقاطی که در تصاویر ورودی دچار انسداد شده‌اند، برابر صفر باشد. در صورتی که از دو تصویر ورودی جهت

ساخت مدل استفاده شود، نمی‌توان این شرط را در نظر گرفت. علاوه بر این  $m^k(g^k(p))$  می‌بایست برابر با

حاصل ضرب بین بردار متعامد بر رویه‌ی چهره در نقطه‌ی  $p$  و بردار  $n$  در دستگاه مختصات دید تصویر

ورودی، در نظر گرفته شود. بردار  $n$  در شکل ۲-۵ نشان داده شده است. در صورتی که تعداد تصاویر ورودی

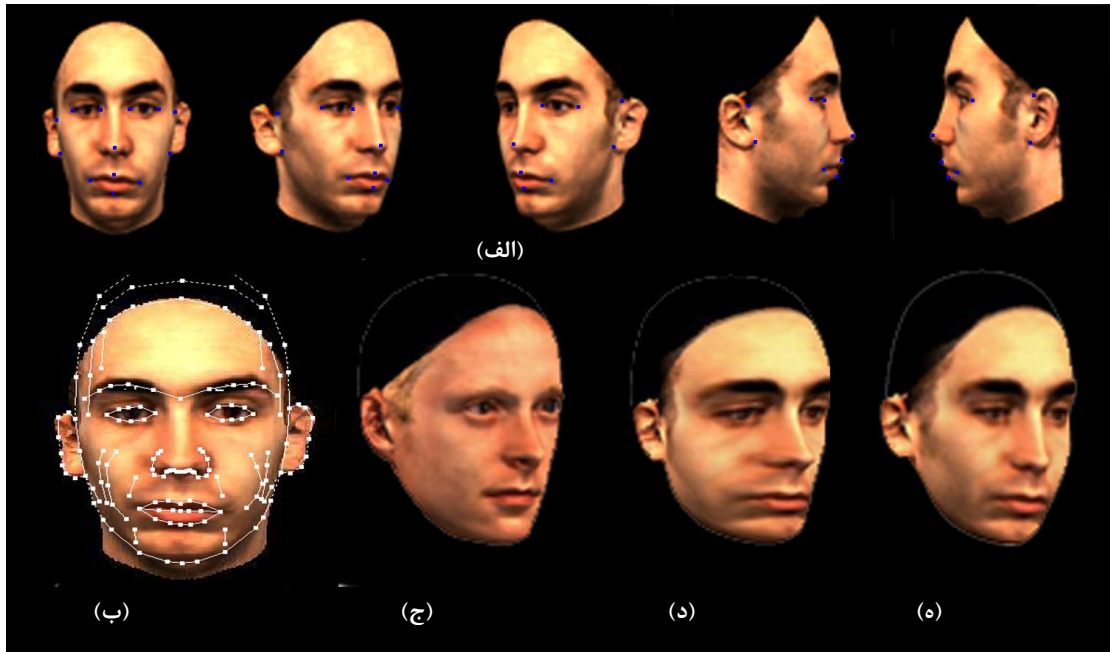
برابر ۲ باشد، جهت تعیین  $m^k(g^k(p))$  در تصویر نیم‌رخ، به ازای هر نقطه‌ی  $p$  دو بردار متعامد در نظر گرفته

شده و  $m^k(g^k(p))$  به صورت زیر محاسبه می‌گردد.

$$\begin{aligned} N_{p1} &= N_p \\ N_{p2} &= (-N_{p_x}, N_{p_y}, N_{p_z}) \\ m(g(p)) &= \max(n \cdot N_{p1}, n \cdot N_{p2}) \end{aligned} \quad (17-3)$$

پس از تعیین  $m^k(g^k(p))$  برای تمام نقاط  $p$  با استفاده از رابطه‌ی (۱۶-۳)، بافت استخراج می‌شود. شکل ۳-

۴ (د) و (ه) نتایج استخراج بافت بر روی مدل‌های تطبیق‌یافته‌ی پالایش نشده و پالایش شده نشان می‌دهد.



شکل ۳-۴ - نتایج مراحل تطبیق مدل، (الف): تصاویر ورودی و نقاط راهنما (ب): منحنی‌های راهنما (ج): مدل عمومی (د): مدل چهره‌ی تطبیق‌یافته قبل از پالایش شکل (ه): مدل چهره‌ی نهایی

### ۱-۱۶ پایگاه‌های داده‌ی چهره‌ی سه‌بعدی

با توجه به تأثیرپذیری مدل میان‌ساز از پایگاه داده، در این پایان نامه یک پایگاه داده‌ی چهره‌ی سه‌بعدی از چهره‌های ایرانی تهیه‌شده‌است که در این بخش به معرفی و مقایسه‌ی آن پایگاه داده‌ی MPI می‌پردازیم. علت تأثیرپذیری مدل میان‌ساز از پایگاه داده و نتایج حاصل از سنتز چهره‌های ایرانی بر اساس پایگاه داده‌ی ارائه شده در فصول ۴ و ۵ مورد بررسی قرار می‌گیرد. علاوه بر این پایگاه داده‌ی Gavab3D نیز به عنوان یک پایگاه داده‌ی چهره‌ی سه‌بعدی بدون بافت مورد بررسی قرار می‌گیرد.

### ۱-۱۶-۱ پایگاه داده‌ی MPI

این پایگاه داده شامل ۱۰۰ مدل سه‌بعدی مذکر و ۱۰۰ مدل چهره سه‌بعدی مونث می‌باشد. سن افراد در فاصله ۱۸ تا ۴۵ سال در نظر گرفته شده است. تمامی آنها به جز ۱ مورد از نژاد قزاقی<sup>۱۲۳</sup> بوده و

<sup>123</sup> Caucasian

اغلب چهره‌های موجود در این پایگاه داده از نواحی غربی‌تر این نژاد انتخاب شده‌اند. در این پایگاه داده یک مدل چهره از نژاد آسیای شرقی نیز وجود دارد. جمع‌آوری اطلاعات سه‌بعدی چهره، توسط پوششگر لیزری Cyberware™ صورت گرفته که هر مدل چهره را توسط ۷۰۰۰۰ نقطه با دقت بالایی اندازه‌گیری می‌کند. اطلاعات موجود در هر نقطه شامل موقعیت فضایی و رنگ نقاط می‌شود. فرمت ذخیره‌سازی اطلاعات در این پایگاه داده به صورت فایل‌های شیئی می‌باشد.

#### ۱-۱۶-۲ پایگاه داده‌ی چهره سه‌بعدی ارائه شده

بر اساس روش مذکور در بخش ۴-۳ پایگاه داده‌ی سه‌بعدی از چهره‌های ایرانی تهیه شد. پایگاه داده‌ی ارائه شده متشکل از ۱۰۰ چهره سه‌بعدی مذکر می‌باشد. در حال حاضر مدل‌های چهره مؤنث این پایگاه داده با مدل‌های مؤنث موجود در پایگاه داده MPI مشترک می‌باشند. سن افراد بین ۱۷ تا ۵۰ سال در نظر گرفته شده است. سعی شده است تا از تمام اقوام ایرانی، چهره‌هایی در پایگاه داده گنجانده شود. از هر شخص دو تصویر تمام‌رخ و نیم‌رخ دوبعدی تحت شرایط نور استاندارد (تابش نور از روبرو) تهیه شده است. پیش‌پردازشی شامل حذف مو و گردن در این تصاویر صورت گرفته تا وابستگی مدل ایجاد شده به این بخش‌ها از بین برود. مدل‌های سه‌بعدی چهره با استفاده از این دو تصویر و روش ارائه شده در ۴-۳، با مشخص کردن بیش از ۱۰۰ نقطه راهنما به ازای هر تصویر، تهیه شده است. شکل ۴-۴ نمونه‌ای از مدل‌های چهره سه‌بعدی موجود در این پایگاه داده را نشان می‌دهد.



شکل ۳-۵ - نمونه‌ای از مدل‌های ۳ بعدی موجود در پایگاه داده ارائه شده

هر مدل از ۷۰۰۰ مثلث تشکیل شده است که تقریباً شامل ۹۰۰۰ رأس می‌شود. مختصات هر مثلث در دستگاه مختصات استوانه‌ای مشخص شده (مؤلفه‌های  $h$ ،  $\phi$  و  $r$ ) و به ازای هر رأس سه مؤلفه‌ی  $R$ ،  $G$  و  $B$  نیز جهت مشخص شدن بافت بیان شده است. با درون‌یابی و افزایش تعداد رئوس تا  $2^{18}$  رأس برای هر مدل سه‌بعدی، ماتریسهای عمق و بافت با دقت قابل قبولی محاسبه می‌شوند. چهار ماتریس  $r$ ،  $R$ ،  $G$  و  $B$  با اندازه‌های  $512 \times 512$  معرفی‌کننده‌ی مدل سه‌بعدی ایجاد شده هستند. سطرها و ستون‌های ماتریس‌های مذکور مؤلفه‌های  $h$  و  $\phi$  می‌باشند. بنابراین هر مدل سه‌بعدی چهره در دستگاه مختصات استوانه‌ای به صورت زیر تعریف می‌شود.

$$I(h, \phi) = (r(h, \phi), R(h, \phi), G(h, \phi), B(h, \phi))^T; \quad h, \phi \in \{0, \dots, 511\} \quad (18-3)$$

نمایش دوبعدی ماتریس‌های RGB و  $r$  در شکل ۳-۶ آورده شده است.



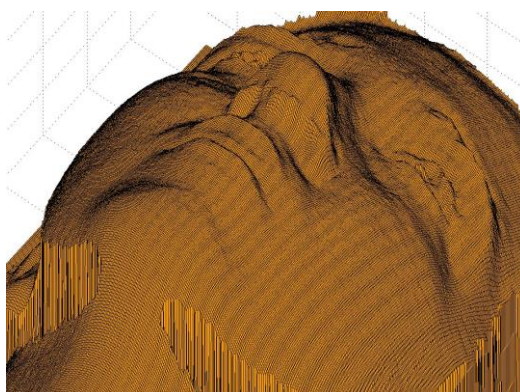
شکل ۳-۶ - نمایش دو بعدی ماتریس‌های شکل و بافت

### ۱-۱۶-۳ پایگاه داده‌ی Gavab3D

این پایگاه داده که توسط دانشگاه اسپانیایی "ری خوان کارلوس"<sup>۱۲۴</sup> ارائه شده است، دارای ۴۲۷ روبه‌ی سه بعدی چهره می‌باشد که تحت فرمت VRML ذخیره شده‌اند. این روبه‌ها مربوط به ۶۱ شخص مختلف شامل ۴۵ مرد و ۱۶ زن در سنین بین ۱۸ تا ۴۰ سال می‌شوند. تمام افراد در این پایگاه داده نیز از نژاد قزاقی می‌باشند. به ازای هر فرد ۹ تصویر در این پایگاه داده تحت زوایای دید مختلف و حالات چهره‌ی

<sup>124</sup> University of Rey Juan Carlos

مختلف موجود می‌باشد. ۵ تصویر تمام رخ، ۲ تصویر نیم‌رخ و ۲ تصویر سر متمایل به پایین این ۹ تصویر را تشکیل می‌دهند. به تصاویر سه‌بعدی چهره‌ای که تنها از یک زاویه‌ی خاص تصویر برداری شده باشند، تصاویر 2.5 بعدی اطلاق می‌شود. البته در اغلب اوقات تفاوتی بین تصاویر سه‌بعدی و 2.5 بعدی قایل نمی‌شوند. تصاویر موجود در این پایگاه داده از نوع 2.5 بعدی و بدون بافت می‌باشند. بدون بافت بودن این پایگاه داده مانع از کاربرد آن در مواردی چون مدل‌سازی میان‌ساز می‌شود. در شکل ۳-۷ یک رویه‌ی چهره از این پایگاه داده به نمایش گذاشته شده است.



شکل ۳-۷ - نمونه‌ای از تصاویر موجود در پایگاه داده‌ی Gavab3D

پس از بررسی مراحل ساخت مدل‌های سه‌بعدی چهره و تهیه‌ی یک پایگاه داده‌ی سه‌بعدی، شرایط لازم جهت ساخت مدل میان‌ساز و تشخیص چهره بر اساس آن فراهم می‌باشد. همانطور که در فصل ۱ اشاره شد، به علت مزایای نسبی مدل میان‌ساز، تشخیص چهره بر اساس این مدل به عنوان مهمترین و پرکاربردترین روش تشخیص چهره‌ی مبتنی بر مدل شناخته می‌شود. در این بخش مدل چهره‌ی میان‌ساز و تشخیص بر اساس آن مورد بررسی قرار می‌گیرد.

ساخت مدل میان‌ساز چهره، اولین بار توسط بیر<sup>۱۲۵</sup> و نیلی<sup>۱۲۶</sup> در ۱۹۹۲ بر روی تصاویر چهره‌ی دوبعدی صورت گرفت. در این روش با استفاده از میان‌سازی تصاویر چهره انیمیشنی ارائه شد که در آن چهره‌ی شخصیت مورد نظر به آرامی تغییر پیدا کرده و به چهره‌ی جدید تبدیل می‌شد [بیر و نیلی، ۱۹۹۲]. چن<sup>۱۲۷</sup> و همکاران در ۱۹۹۵ روش ارائه شده توسط بیر و نیلی را بر روی مدل‌های چهره‌ی سه‌بعدی پیاده‌سازی کردند [چن و همکاران، ۱۹۹۵]. در واقع این روش اولین میان‌سازی تصاویر سه‌بعدی چهره به شمار می‌رود. در ۱۹۹۷ برگلر<sup>۱۲۸</sup> و همکاران از میان‌سازی ناحیه‌ی اطراف دهان جهت هماهنگ‌سازی لب<sup>۱۲۹</sup> در تصاویر ویدئویی موجود با اطلاعات ورودی گفتار استفاده کردند [برگلر و همکاران، ۱۹۹۷]. پیگهین و همکاران نیز در ۱۹۹۸ با استفاده از مدل‌های چهره‌ی سه‌بعدی ساخته شده توسط روش هندسه یابی تصویری به میان‌سازی حالات چهره و ایجاد حالت‌های جدید پرداختند [پیگهین و همکاران، ۱۹۹۸].

---

<sup>125</sup> Beier

<sup>126</sup> Neely

<sup>127</sup> Chen

<sup>128</sup> Bregler

<sup>129</sup> Lip Synchronization

کاربرد مدل میان‌ساز سه بعدی در سنتز چهره از تصاویر دوبعدی توسط جونز<sup>۱۳۰</sup> و پوگیو<sup>۱۳۱</sup> در ۱۹۹۶ [جونز و پوگیو، ۱۹۹۶]، وتر و همکاران در ۱۹۹۷ [وتر و همکاران، ۱۹۹۷] و وتر و بلانز در ۱۹۹۸ [وتر و بلانز، ۱۹۹۸] ارائه شد که شباهت‌های ساختاری و الگوریتمی بسیاری در این روش‌ها مشاهده می‌شود. در ۱۹۹۹ بلانز و وتر از مدل میان‌ساز جهت تشخیص چهره استفاده کردند [بلانز و وتر، ۱۹۹۹]. برخی از پژوهشگران نیز به ترکیب مدل میان‌ساز با سایر روش‌های تشخیص چهره پرداخته‌اند. از این دسته می‌توان به هیسل<sup>۱۳۲</sup> و بلانز اشاره کرد که از مدل میان‌ساز جهت آموزش یک سیستم تشخیص چهره‌ی مبتنی بر اجزا استفاده کرده‌اند [هیسل و بلانز، ۲۰۰۶].

مدل میان‌ساز در واقع ترکیبی از مدل‌های ریخت‌پذیر<sup>۱۳۳</sup> و تکنیک‌های ارتباط تصویری کامپیوتری جهت شبیه‌سازی اثرات نوردهی می‌باشد. در این مدل تغییر زاویه‌ی دید و شرایط نورپردازی به سادگی صورت می‌گیرد. همچنین با بکارگیری روش‌های ایجاد سایه، می‌توان سایه‌ها را نیز در مدل پیاده‌سازی کرد. این روش می‌تواند با استفاده از تنها یک تصویر ورودی و بدون نیاز به دخالت کاربر، تخمین قابل قبولی از بافت، شکل، زاویه‌ی دید و مشخصات نورپردازی مدل چهره ایجاد کند. بخش اصلی این روش ساخت مدل میان‌ساز و تخمین تابع توزیع احتمال آن می‌باشد [بلانز و وتر، ۲۰۰۳]. این مدل ریخت‌پذیر، ترکیب خطی‌ای از بردارهای شکل و بافت مدل‌های چهره‌ی سه‌بعدی است. این فصل شامل دو مبحث مدل میان‌ساز چهره‌ی سه‌بعدی و تطبیق تصاویر ورودی و مدل میان‌ساز می‌باشد.

---

<sup>130</sup> Jones

<sup>131</sup> Poggio

<sup>132</sup> Heisele

<sup>133</sup> Deformable Models

## ۱۸-۱ مدل چهره‌ی میان‌ساز سه‌بعدی

مدل چهره‌ی میان‌ساز، یک ترکیب خطی از بردارهای چهره در فضایی با ابعاد بالاست به نحوی که منجر به تولید چهره‌هایی واقعی شود. این ترکیب خطی بر بردارهای شکل و بافت چهره اعمال می‌شود. هر چهره از تعداد زیادی نقطه که بیانگر رئوس چندضلعی‌های سازنده‌ی چهره هستند تشکیل شده است. هر رأس دارای مختصات فضایی و مشخصات مربوط به رنگ یا بافت می‌باشد. حال اگر موقعیت فضایی رئوس چهره را در دستگاه مختصات کارتزین بیان کنیم، می‌توان مشخصات هندسی چهره را با استفاده از یک بردار متشکل از موقعیت تمام رئوس نشان داد. این بردار، بردار شکل<sup>۱۳۴</sup> چهره نامیده می‌شود. مشخصات رنگ یا بافت هر رأس را نیز با استفاده از مؤلفه‌های رنگی  $R$ ،  $G$  و  $B$  می‌توان نشان داد. بنابراین برداری متشکل از مؤلفه‌های رنگی تمام رئوس چهره، بیانگر مشخصات بافت چهره بوده و بردار بافت<sup>۱۳۵</sup> نامیده می‌شود. این دو بردار بیانگر یک چهره در فضای سه‌بعدی می‌باشند. از آنجایی که ابعاد بردارها با توجه به تعداد زیاد رئوس، بزرگ می‌باشد، این بردارها، بردارهای چهره در فضای بالا نامیده می‌شوند. در صورتی که تعداد رئوس تشکیل‌دهنده‌ی چهره برابر  $n$  باشد، بردارهای شکل و بافت به صورت زیر می‌باشند.

$$\begin{aligned} S &= (x_1, y_1, z_1, x_2, y_2, \dots, x_n, y_n, z_n)^T \in \mathbb{R}^{3n} \\ T &= (R_1, G_1, B_1, R_2, G_2, \dots, R_n, G_n, B_n)^T \in \mathbb{R}^{3n} \end{aligned} \quad (1-4)$$

به عبارت دیگر با استفاده از این نحوه‌ی نمایش، هر چهره توسط دو نقطه در دو فضای  $3n$  بعدی متمایز نمایش داده می‌شود.  $n$  در مدل‌های چهره‌ی سه‌بعدی در پایگاه داده‌ی MPI در حدود ۷۰۰۰۰ می‌باشد.

حال استفاده از ترکیب خطی بردارهای شکل و بافت چهره‌ها منجر به ساخت نقاطی جدید در دو

فضای  $3n$  بعدی می‌شود.

<sup>134</sup> Shape Vector

<sup>135</sup> Texture Vector



$$S_{\text{model}} = \sum_{i=1}^m a_i S_i \quad T_{\text{model}} = \sum_{i=1}^m b_i T_i \quad (2-4)$$

در رابطه‌ی فوق جهت جلوگیری از تغییرات کلی در اندازه و رنگ چهره‌ها از شرط  $\sum_{i=1}^m a_i = \sum_{i=1}^m b_i = 1$  استفاده

می‌کنیم. باید توجه کرد تنها در صورتی  $S_{\text{model}}$  و  $T_{\text{model}}$  بیانگر چهره‌های واقعی هستند که بردارهای  $S_i$  و

$T_i$  بردارهایی تناظر یافته باشند. به عبارت دیگر نقاط متناظر در مدل‌های سه‌بعدی چهره می‌بایست با

یکدیگر جمع شوند. برای مثال مختصات و رنگ نقطه‌ی واقع در نوک بینی در هر تصویر می‌بایست با

مختصات و رنگ نقطه‌ی واقع در نوک بینی در سایر تصاویر جمع شود. در غیر این صورت روابط (2-4) به

جای ایجاد یک مدل چهره‌ی سه‌بعدی جدید، تنها یک تصویر سه‌بعدی نامفهوم ناشی از میانگین‌گیری وزنی

بین مدل‌های سه‌بعدی ایجاد می‌کنند. ایجاد تناظر بین مدل‌ها در ادامه تشریح خواهد شد. در این قسمت

فرض می‌کنیم که تناظر نقطه به نقطه بین تمام مدل‌ها برقرار شده است. در این صورت رابطه‌ی (2-4)

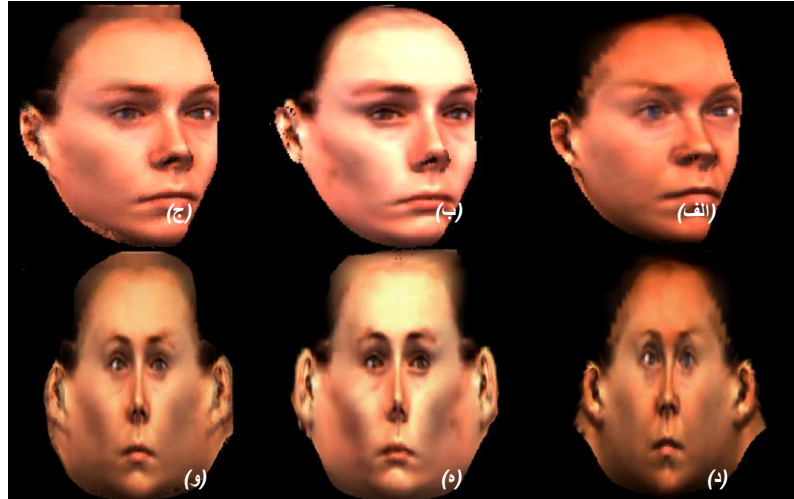
بیانگر یک مدل میان‌ساز می‌باشد. شکل 4-1 نتیجه‌ی میان‌سازی بین دو مدل چهره را نشان می‌دهد.

هر مدل چهره‌ی سه‌بعدی جدید را می‌توان به صورت  $(S_{\text{mod}}(\bar{a}), T_{\text{mod}}(\bar{b}))$  نمایش داد که در آن

بردارهای  $\bar{a} = (a_1, a_2, \dots, a_m)^T$  و  $\bar{b} = (b_1, b_2, \dots, b_m)^T$  ضرایب مدل نامیده شده و تعریف‌کننده مدل سه‌بعدی

جدید هستند. تغییر بردارهای  $\bar{a}$  و  $\bar{b}$  منجر به ایجاد مدل‌های جدید می‌شود.  $m$  مدل چهره‌ی سه‌بعدی نیز

که مدل‌های جدید بر اساس ترکیب خطی آنها ساخته می‌شوند، مدل‌های نمونه نامیده می‌شوند.



شکل ۴-۱ - میان‌سازی بین دو مدل چهره، (الف): مدل چهره‌ی اول، (ب): مدل چهره‌ی دوم (ج): میان‌سازی چهره‌های اول و دوم، (د، ه، و): تصاویر (الف)، (ب) و (ج) در دستگاه مختصات استوانه‌ای

یکی از نکات مهم در سنتز چهره‌ی مبتنی بر مدل‌های سه‌بعدی، یافتن روشی جهت جلوگیری از

تولید چهره‌های غیرواقعی می‌باشد. در مراجع [جونز و پوگیو، ۱۹۹۶- وتر و همکاران، ۱۹۹۷] روش خاصی

برای این منظور پیشنهاد نشده و در مقابل، یک الزام محدودکننده برای بردارهای  $\bar{a}$  و  $\bar{b}$  در نظر گرفته شده

است. این الزام در شرط تحدب ترکیب خطی نهفته است که القاکننده‌ی الزام زیر می‌باشد.

$$a_i, b_i \in [0,1]; \quad i = 1 \dots m \quad (3-4)$$

اما در مرجع [بلانز و وتر، ۱۹۹۹] جهت جلوگیری از تولید چهره‌های غیرواقعی، استفاده از تابع

توزیع بردارهای  $\bar{a}$  و  $\bar{b}$  و تنظیم تابع درست‌نمایی<sup>۱۳۶</sup> این بردارها پیشنهاد شده است. بردارهای  $\bar{a}$  و  $\bar{b}$

معادل بردارهای  $\bar{a}$  و  $\bar{b}$  هستند که با بکارگیری روش PCA محاسبه شده‌اند. مراحل ایجاد این بردارها و

روش PCA در ادامه تشریح خواهد شد. در مرحله‌ی اول یک تابع گوسی  $3n$  متغیره بر روی  $m$  مدل سه‌بعدی

چهره منطبق می‌شود. در پیاده‌سازی این روش تعداد مدل‌های چهره برابر ۲۰۰ در نظر گرفته شده است.

تابع گوسی مذکور، به صورت زیر می‌باشد.

<sup>136</sup> Likelihood

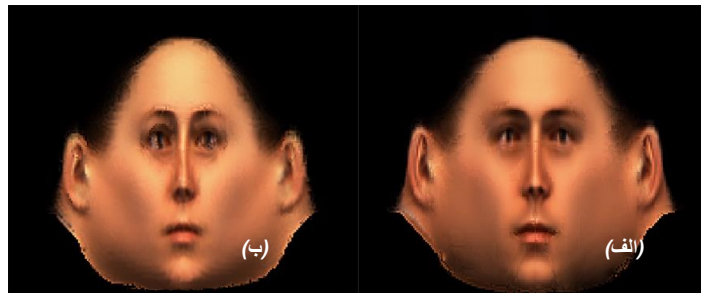
$$P(S) \propto \exp\left(-\frac{1}{2}(S-\bar{S})^T C_S^{-1}(S-\bar{S})\right) \quad (4-4)$$

$$P(T) \propto \exp\left(-\frac{1}{2}(T-\bar{T})^T C_T^{-1}(T-\bar{T})\right)$$

که در آن  $\bar{S}$  و  $\bar{T}$  به ترتیب میانگین بردارهای  $S_i$  و  $T_i$  و  $C_S$  و  $C_T$  ماتریس‌های کواریانس بردارهای فوق می‌باشند. بردارهای  $\bar{S}$  و  $\bar{T}$  بردارهایی  $3n$  بعدی بوده و به صورت زیر محاسبه می‌شوند.

$$\bar{S} = \sum_{i=1}^m \frac{1}{m} S_i \quad \bar{T} = \sum_{i=1}^m \frac{1}{m} T_i \quad (5-4)$$

باید توجه کرد که به علت برقراری تناظر نقطه به نقطه بین مدل‌ها، بردارهای  $\bar{S}$  و  $\bar{T}$  یک مدل چهره‌ی سه‌بعدی تشکیل می‌دهند که مدل چهره‌ی میانگین نامیده می‌شود. مدل میانگین چهره‌های مذکر و مؤنث پایگاه داده‌ی MPI در دستگاه مختصات استوانه‌ای در شکل ۴-۲ نشان داده شده است.



شکل ۴-۲ - مدل‌های میانگین پایگاه داده‌ی MPI در دستگاه مختصات استوانه‌ای، (الف): مذکر، (ب): مؤنث

$C_S$  و  $C_T$  نیز ماتریس‌هایی  $3n \times 3n$  بوده و به صورت زیر محاسبه می‌شوند.

$$C_S = \frac{\sum_{i=1}^m (S_i - \bar{S})(S_i - \bar{S})^T}{m-1} \quad C_T = \frac{\sum_{i=1}^m (T_i - \bar{T})(T_i - \bar{T})^T}{m-1} \quad (6-4)$$

در این مرحله جهت دسترسی به اجزای اصلی مدل‌های چهره، از انتقال مدل‌های سه‌بعدی چهره به یک فضای پایه‌ی متعامد دارای ابعاد کمتر با بکارگیری روش PCA استفاده می‌شود. طبق این روش هر یک از بردارهای نمونه  $S_i$  را می‌توان به صورت زیر در نظر گرفت.

$$S_i = \bar{S} + \sum_{j=1}^d c_j s_j \quad (7-4)$$

در رابطه فوق  $s_j$  ها  $d$  بردار ویژه ماتریس پخش کننده<sup>۱۳۷</sup>  $S$  هستند که دارای بیشترین مقادیر ویژه می باشند. ماتریس پخش کننده، برابر  $(n-1)C_S$  می باشد [دودا و همکاران، ۲۰۰۱]. حال با توجه به معادلات (۲-۴) و (۷-۴) داریم:

$$S_{model} = \sum_{i=1}^m a_i (\bar{S} + \sum_{j=1}^d c_{ji} s_j) = \bar{S} + \sum_{j=1}^d \alpha_j s_j \quad (۸-۴)$$

در رابطه فوق  $\alpha_j = \sum_{i=1}^m a_i c_{ji}$  است. با در نظر گرفتن  $d = m-1$  روابط اساسی زیر بوجود می آید.

$$S_{model} = \bar{S} + \sum_{i=1}^{m-1} \alpha_i s_i \quad T_{model} = \bar{T} + \sum_{i=1}^{m-1} \beta_i t_i \quad (۹-۴)$$

بنابر روابط (۴-۴) و (۹-۴) و با فرض استقلال مؤلفه های بردارها از یکدیگر، تابع توزیع بردارهای  $\alpha$  و  $\beta$  به صورت زیر بیان می شود.

$$P(\bar{\alpha}) \propto \exp\left(-\frac{1}{2} \sum_{i=1}^m \left(\frac{\alpha_i}{\sigma_{S,i}}\right)^2\right) \quad (۱۰-۴)$$

$$P(\bar{\beta}) \propto \exp\left(-\frac{1}{2} \sum_{i=1}^m \left(\frac{\beta_i}{\sigma_{T,i}}\right)^2\right)$$

در رابطه فوق  $\sigma_{S,i}$  و  $\sigma_{T,i}$  مقادیر ویژه متناظر با بردارهای ویژه  $s_i$  و  $t_i$  هستند. با مشخص بودن تابع توزیع بردارهای  $\alpha$  و  $\beta$  در مراحل بعدی از تابع درستنمایی این بردارها استفاده کرده و از انتخاب مقادیر غیرمحمتم  $\alpha$  و  $\beta$  جلوگیری می کنیم.

مزیت استفاده از رابطه (۹-۴) به جای استفاده از رابطه (۲-۴) در مفهوم PCA نهفته است. در رابطه (۹-۴) به جای ترکیب خطی مدل های چهره ی سه بعدی نمونه، از ترکیب خطی اجزای اصلی مدل ها

استفاده می‌کنیم. این امر باعث می‌شود که با کم کردن تعداد نمونه‌های مدل میان‌ساز (کاهش  $m$ )، به علت استفاده از اجزای اصلی کیفیت مدل کمتر تحت تأثیر قرار بگیرد.

از آنجایی که هر چهره به صورت دو بردار ویژگی در فضای  $3n$  بعدی در نظر گرفته شده است، جهت پیمایش<sup>۱۳۸</sup> این فضا توسط ترکیب خطی بردارها به  $3n$  بردار مستقل نیاز داریم. اما در این روش تنها از ترکیب خطی  $m$  بردار جهت ساخت مدل‌های جدید استفاده می‌شود. با توجه به این که  $n$  در حدود  $70000$  و  $m$  برابر  $200$  می‌باشد، مشخص است که این تعداد بردار تنها یک زیرفضای بسیار کوچک از فضای  $3n$  بعدی را پیمایش می‌کنند. اما باید به این نکته توجه کرد که تمام مدل‌های چهره‌ی  $3$  بعدی  $n$  رأسی ممکن نیز تنها یک زیرفضای کوچک از فضای  $3n$  بعدی را شامل می‌شوند. (تعداد رویه‌های  $n$  رأسی غیرچهره به مراتب بیشتر از چهره‌های  $n$  رأسی می‌باشد). بنابراین ترکیب خطی  $m$  بردار بخش قابل توجهی از زیرفضای چهره‌های ممکن را پیمایش می‌نماید.

#### ۱-۱۸-۱ مدل میان‌ساز تقطیع شده<sup>۱۳۹</sup>

جهت افزایش درجه‌ی آزادی مدل میان‌ساز و قابلیت پیمایش بخش بیشتری از فضای تصاویر چهره، از مدل میان‌ساز تقطیع شده استفاده می‌کنیم. برای این منظور کفایت تا هر مدل نمونه را به چند بخش تقسیم کنیم و هر بخش را به طور جداگانه میان‌سازی نماییم. در نهایت کفایت تا با استفاده از الگوریتم ارائه شده در مرجع [برت و ادلسون<sup>۱۴۰</sup>، ۱۹۸۵] به یکنواخت‌سازی مرزها پردازیم. تقسیم مدل‌های چهره به چند بخش معادل با افراز بردارهای شکل و بافت چهره به چند بخش می‌باشد. در صورتی که تعداد مدل‌های قابل تولید توسط مدل میان‌ساز تقطیع نشده برابر  $Q$  و تعداد بخشهای مدل میان‌ساز تقطیع شده برابر  $r$

<sup>138</sup> Span

<sup>139</sup> Segmented Morphable Model

<sup>140</sup> Burt, Adelson

باشد، تعداد مدل‌های چهره‌ی قابل تولید توسط مدل میان‌ساز تقطیع‌شده برابر  $Q^r$  می‌باشد. باید توجه کرد که  $Q$  از لحاظ تئوری بینهایت است ولی به علت محدودیت‌های کوانتیزاسیون دارای مقداری محدود می‌باشد. این روش یک ایراد مهم ایجاد می‌کند که افزایش تعداد بردارهای  $\alpha$  و  $\beta$  می‌باشد. در مدل میان‌ساز تقطیع شده به جای یک بردار  $\alpha$  و  $\beta$ ،  $r$  بردار  $\alpha$  و  $\beta$  با همان ابعاد تولید می‌شود. این امر منجر به افزایش حجم محاسبات و زمان لازم برای ساخت مدل می‌شود. بدون در نظر گرفتن حجم محاسباتی مربوط به یکنواخت‌سازی مرزها، محاسبات در مدل میان‌ساز تقطیع شده  $r$  برابر می‌شود (در تکرارهای آخر الگوریتم بهینه‌سازی). شکل ۳-۴ نحوه‌ی تقطیع مدل چهره را نشان می‌دهد. در این شکل  $r$  برابر ۴ در نظر گرفته شده است. نواحی چشم‌ها، بینی، دهان و نواحی اطراف، این ۴ بخش را تشکیل می‌دهند. در پیاده‌سازی این ۴ ناحیه بر روی مدل نمونه توسط یک ابزار ساده و با کمک کاربر معین می‌شوند. بر روی تصاویر ورودی نیز این نواحی با استفاده از تخمین اولیه‌ی موقعیت و زاویه‌ی دید معین می‌شوند. تخمین اولیه‌ی موقعیت و زاویه‌ی دید در بخش بعدی تشریح خواهد شد.



شکل ۳-۴ - تقطیع مدل میان‌ساز به ۴ ناحیه [بلانز و وتر، ۱۹۹۹]

به علت افزایش حجم محاسبات تا ۴ برابر، از این روش تنها در سنتز حالت چهره به صورت دقیق استفاده می‌کنیم. در کاربرد تشخیص چهره استفاده از این روش مگر در مواردی که در بین چهره‌های ورودی و مورد آزمایش، چهره‌های شبیه به هم وجود داشته باشد، الزامی ندارد. وجود چهره‌های شبیه به هم معمولاً در کتابخانه‌های بسیار بزرگ اتفاق می‌افتد. شکل ۴-۴ تفاوت بین سنتز یک چهره بر اساس مدل میان‌ساز

تقطیع شده و تقطیع نشده را نشان می دهد. در مدل تقطیع نشده، حالت ابروها و شکل چشم با تصویر ورودی کمی متفاوت بوده و بینی نیز کوچکتر می باشد. تاریک تر بودن چهره ناشی از عدم تنظیم صحیح پارامترهای نوردهی بوده و ارتباطی به مدل تقطیع نشده ندارد.



شکل ۴-۴ - مقایسه مدل میان ساز تقطیع شده و تقطیع نشده، (الف): تصویر ورودی از پایگاه داده‌ی دانشگاه صنعتی امیرکبیر، (ب): چهره‌ی سنتز شده با استفاده از مدل میان ساز تقطیع شده، (ج) چهره‌ی سنتز شده با استفاده از مدل میان ساز تقطیع نشده

## ۱-۱۸-۲ خصوصیات چهره

خصوصیات قابل پیاده سازی در سنتز چهره به دو دسته تقسیم می شوند. دسته اول خصوصیات هستند که امکان مشاهده آن در هر چهره‌ای وجود دارد. همچون حالات خنده، گریه، اخم و ترس. دسته دوم خصوصیات هستند که در برخی از چهره‌ها مشاهده می شوند. همچون چاقی و یا لاغری صورت، انحنای بینی، مردانگی و یا زنانگی چهره. از آنجایی که بردارهای  $\alpha$  و  $\beta$  ارتباط منطقی با خصوصیات چهره ندارند، جهت سنتز خصوصیات چهره از تکنیکهای آماری استفاده می شود.

به منظور پیاده سازی دسته اول خصوصیات چهره، تکنیکهای مبتنی بر عملکرد [ادواردز<sup>۱۴۱</sup> و همکاران، ۱۹۹۸] به کار گرفته می شوند. در این شیوه دو مدل سه بعدی از یک شخص، یکی در حالت عادی و دیگری در حالت دارای خصوصیت مورد نظر، تهیه می شود. با افزودن تفاضل این دو مدل به یک مدل

<sup>141</sup> Edwards

چهره‌ی سه‌بعدی جدید، مدلی دارای خصوصیت مورد نظر بوجود می‌آید. در شکل ۴-۵ حالت خنده در یک مدل چهره سه‌بعدی سنتز شده است. جهت سنتز خصوصیات دسته دوم، دو بردار  $\Delta S$  و  $\Delta T$  به صورت زیر تعریف می‌شوند.

$$\Delta S = \sum_{i=1}^m \mu_i (S_i - \bar{S}) \quad \Delta T = \sum_{i=1}^m \mu_i (T_i - \bar{T}) \quad (11-4)$$

در رابطه فوق  $\mu_i$  ضریبی است که به هر مدل چهره عددی متناسب با میزان خاصیت مورد نظر نسبت می‌دهد. برای مثال اگر خاصیت مورد نظر چاقی صورت باشد، به چهره‌های چاق اعداد نزدیک به ۱ و به چهره‌های لاغر اعداد نزدیک به ۰ انتساب داده می‌شود. حال با افزودن و یا کم کردن ضریبی از دو بردار  $\Delta S$  و  $\Delta T$  به هر مدل می‌توان خاصیت مذکور را در مدل ایجاد و یا حذف کرد. دقت شود که بردارهای  $\Delta S$  و  $\Delta T$  نیز دارای تناظر نقطه‌به‌نقطه با سایر بردارها می‌باشند و افزودن و یا کم کردن این بردارها از سایر مدل‌ها منجر به ایجاد یک مدل جدید می‌شود. در شکل ۴-۵ خاصیت چاقی صورت در یک مدل سه‌بعدی ایجاد و حذف شده است. تصویر ورودی تصویر شماره‌ی ۱ از پایگاه داده‌ی چهره‌ی دوبعدی دانشگاه صنعتی امیرکبیر می‌باشد.



شکل ۴-۵ - پیاده‌سازی خصوصیات چهره بر روی مدل سنتز شده در شکل ۴-۵، از چپ به راست: حالت عادی، خنده، چاقی صورت، لاغری صورت

### ۳-۱۸-۱ تعیین تناظر بین مدل‌ها

همانطور که در ابتدای این بخش اشاره شد، هسته‌ی اساسی مدل‌سازی میان‌ساز سه‌بعدی چهره، تعیین تناظر نقطه به نقطه بین مدل‌های نمونه می‌باشد. روش‌های متعددی جهت تعیین تناظر و انطباق



تصاویر معرفی شده‌اند. در مدل میان‌ساز ارائه شده توسط وتر و همکاران [وتر و همکاران، ۱۹۹۷] و بلانز و وتر [بلانز و وتر، ۱۹۹۹]، انطباق بین مدل‌های چهره بوسیله‌ی روش شارش نوری<sup>۱۴۲</sup> صورت پذیرفته است. در این بخش پس از بررسی روش‌های مطرح در تعیین تناظر بین نقاط به بررسی الگوریتم شارش نوری می‌پردازیم. همچنین الگوریتمی مبتنی بر همبستگی محلی را نیز جهت تعیین تناظر معرفی خواهیم کرد. علاوه بر این روشی جدید و مبتنی بر گراف جهت بهینه‌سازی نتایج روش‌های انطباق ارائه شده است که در بخش بعدی آن را معرفی خواهیم کرد.

#### ۱-۱۸-۳-۱ روش‌های انطباق تصاویر

روش‌های انطباق تصاویر در مدل‌سازی سه بعدی را می‌توان به دو دسته‌ی روش‌های مبتنی بر نقاط راهنما<sup>۱۴۳</sup> و روش‌های مبتنی بر رویه تقسیم کرد. در روش‌های دسته‌ی اول تعدادی نقطه‌ی راهنما همچون گوشه‌های چشم و نوک بینی توسط کاربر در هر دو تصویر انتخاب می‌شود. این روش‌ها خود به دو دسته‌ی روش‌های صلب و غیرصلب تقسیم می‌شوند. در روش‌های صلب یک چهره‌ی قالب<sup>۱۴۴</sup> در نظر گرفته شده و انطباق توسط کمینه کردن فاصله بین نقاط راهنمای تصویر جدید و نقاط راهنمای چهره قالب صورت می‌گیرد [آرون<sup>۱۴۵</sup> و همکاران، ۱۹۸۷]. در این روش‌ها عملگرهای انتقال، چرخش و تغییر مقیاس بر روی کل تصویر اعمال می‌شوند و به همین دلیل این روش‌ها، صلب نامیده می‌شوند. در دسته‌ی روش‌های غیرصلب انطباق مبتنی بر نقاط راهنما، عملگرهای انتقال، چرخش و تغییر مقیاس به صورت محلی اعمال می‌شوند.

---

<sup>142</sup> Optical Flow

<sup>143</sup> Landmark Based Approaches

<sup>144</sup> Template

<sup>145</sup> Arun

الگوریتم‌های <sup>۱۴۶</sup>TPS [بوکشتاین<sup>۱۴۷</sup>، ۱۹۸۹]، تغییر شکل آزاد<sup>۱۴۸</sup> [سدربرگ و پاری<sup>۱۴۹</sup>، ۱۹۸۶]، تبدیلات BSpline [لی و همکاران، ۱۹۹۷] و درون‌یابی داده‌های پراکنده از جمله روش‌هایی است که در این دسته قرار می‌گیرند.

از مهمترین روش‌های انطباق مبتنی بر رویه می‌توان به الگوریتم‌های تطبیق گراف‌های دسته‌ای کشسان<sup>۱۵۰</sup> [ویسکات<sup>۱۵۱</sup> و همکاران، ۱۹۹۷]، الگوریتم تکراری نزدیکترین نقطه<sup>۱۵۲</sup> [بسل و مک‌کی<sup>۱۵۳</sup>، ۱۹۹۲] و انطباق با استفاده از شارش نوری<sup>۱۵۴</sup> [بلانز و وتر، ۱۹۹۹] اشاره کرد. همچنین الگوریتم‌های مبتنی بر انواع مختلف همبستگی محلی نیز در این دسته قرار می‌گیرند.

#### ۱-۱۸-۳-۲ الگوریتم شارش نوری (Optical Flow)

تخمین شارش نوری روشی است که جهت اندازه‌گیری بردار سرعت و تناظریابی در فریم‌های تصاویر ویدئویی استفاده می‌شود. مهمترین تکنیک مورد استفاده در این روش، بکارگیری روش‌های دیفرانسیلی و استفاده از بسط تیلور جهت کاهش مرتبه‌ی چندجمله‌ای‌های دیفرانسیلی می‌باشد. روش‌های تخمین شارش نوری را می‌توان به دو دسته‌ی محلی و فراگیر تقسیم کرد. محلی و یا فراگیر بودن تابع انرژی‌ای که می‌بایست بهینه‌سازی شود، معیار تقسیم‌بندی این روش‌هاست. مرجع [لوکاس و کاناده<sup>۱۵۵</sup>، ۱۹۸۱] معرفی کننده مهمترین روش محلی تخمین شارش نوری است. بنیادی‌ترین روش فراگیر تخمین شارش نوری نیز

---

<sup>146</sup> Thin Plate Spline

<sup>147</sup> Bookstein

<sup>148</sup> Free Form Deformation

<sup>149</sup> Sederberg, Parry

<sup>150</sup> Elastic Bunch Graph Matching

<sup>151</sup> Wiskott

<sup>152</sup> Iterative Closest Point

<sup>153</sup> Besl, McKay

<sup>154</sup> Optical Flow

<sup>155</sup> Lucas, Kanade

در [هورن و شانک<sup>۱۵۶</sup>، ۱۹۸۱] ارائه شده است. روش‌های محلی مقاومت بیشتری در مقابل نویز داشته اما میدان شارشی با چگالی پایین تولید می‌کنند. در مقابل، روش‌های فراگیر، میدان شارشی با چگالی ۱۰۰٪ تولید کرده اما حساسیت بیشتری نسبت به نویز دارند [براهن<sup>۱۵۷</sup> و همکاران، ۲۰۰۲].

اگر تصاویر را به صورت دو تابع  $I_1(x, y)$  و  $I_2(x, y)$  در نظر بگیریم، رابطه زیر الزام اساسی در ایجاد شارش نوری است.

$$I_1(x, y) = I_2(x + \Delta x, y + \Delta y) \quad (۱۲-۴)$$

به عبارت دیگر می‌بایست دو مقدار  $\Delta x$  و  $\Delta y$  را به گونه‌ای بیابیم که الزام فوق برآورده شود. با استفاده از بسط تیلور سمت راست معادله (۱۲-۴) و صرف نظر کردن از عبارات درجه دوم و بالاتر، رابطه‌ی زیر بدست می‌آید.

$$E_2 = I_1(x, y) - I_2(x, y) + \Delta x \frac{\partial I_2(x, y)}{\partial x} + \Delta y \frac{\partial I_2(x, y)}{\partial y} \quad (۱۳-۴)$$

پس الزام معادله (۱۲-۴) به کمینه‌سازی سمت چپ معادله (۱۳-۴) تبدیل می‌شود. معادله فوق دارای دو مجهول  $\Delta x$  و  $\Delta y$  می‌باشد. علاوه بر این می‌بایست نرم کردن تغییرات  $\Delta x$  و  $\Delta y$  مد نظر قرار گیرد. بنابراین به یک الزام دیگر و در نتیجه یک معادله‌ی دیگر نیاز داریم. الزام دیگر با در نظر گرفتن تغییرات آهسته‌ی بردار  $(\Delta x, \Delta y)$  بوجود می‌آید. هورن و شانک [هورن و شانک، ۱۹۸۱] این الزام را با کمینه کردن لاپلاسیان بردارهای سرعت اعمال کرده‌اند. برای این منظور به جای استفاده از دو تصویر  $I_1(x, y)$  و  $I_2(x, y)$ ، دو فریم متوالی از یک رشته‌ی ویدئویی در نظر گرفته شده و معادله (۱۳-۴) به صورت زیر در نظر گرفته می‌شود.

<sup>156</sup> Horn, Schunck  
<sup>157</sup> Bruhn

$$\frac{dI}{dt} = 0 \Rightarrow \frac{\partial I}{\partial x} \frac{dx}{dt} + \frac{\partial I}{\partial y} \frac{dy}{dt} + \frac{\partial I}{\partial t} = 0 \quad (14-4)$$

در رابطه‌ی فوق توجه به تفاوت و مفهوم دو عبارت  $\frac{\partial I}{\partial t}$  و  $\frac{dI}{dt}$  مهم می‌باشد. بیانگر تغییرات یک نقطه‌ی

مشخص مثلاً نوک بینی بر حسب زمان (در دو فریم متوالی) بوده در حالی که  $\frac{\partial I}{\partial t}$  تغییرات یک پیکسل مثلاً

پیکسل (۱۰، ۱۸) بر حسب زمان (در دو فریم متوالی) می‌باشد. با در نظر گرفتن بردارهای سرعت  $u$  و  $v$  به

صورت  $u = \frac{dx}{dt}$  و  $v = \frac{dy}{dt}$  و جایگزینی در رابطه‌ی (۱۴-۴) خواهیم داشت.

$$\frac{\partial I}{\partial x} u + \frac{\partial I}{\partial y} v + \frac{\partial I}{\partial t} = 0 \quad (15-4)$$

در این رابطه بردارهای  $u$  و  $v$  بردارهای سرعت بوده و معادل با  $\Delta x$  و  $\Delta y$  در رابطه‌ی (۱۳-۴) می‌باشد. الزام

دوم نیز به صورت کمینه‌سازی عبارت زیر مطرح می‌شود.

$$\nabla^2 u + \nabla^2 v = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} \quad (16-4)$$

لوکاس و کاناده [لوکاس و کاناده، ۱۹۸۱] الزام دوم را به صورت کمینه کردن تفاوت بین بردار

$(\Delta x, \Delta y)$  در هر نقطه و متوسط بردار  $(\Delta x, \Delta y)$  در یک همسایگی نقطه‌ی مذکور در نظر گرفته‌اند. مراجع

[وتر و بلانز، ۱۹۹۸ - بلانز و وتر، ۱۹۹۹ - بلانز و وتر، ۲۰۰۳] نیز جهت تعیین تناظر نقطه به نقطه بین

مدل‌های چهره‌ی سه‌بعدی، از همین روش استفاده کرده‌اند. این الزام معادل با ثابت فرض کردن بردارهای

سرعت در یک همسایگی نقطه‌ی مذکور می‌باشد. البته باید توجه کرد که به علت همپوشانی همسایگی‌ها

تغییرات بردارهای سرعت کاملاً تدریجی می‌باشد. برای این منظور به جای کمینه کردن معادله‌ی (۱۳-۴)،

معادله‌ی زیر کمینه می‌شود.

$$E(x_0, y_0) = \sum_{x, y \in R(x_0, y_0)} \left( v_x \frac{\partial I(x, y)}{\partial x} + v_y \frac{\partial I(x, y)}{\partial y} + \Delta I(x, y) \right)^2 \quad (17-4)$$

در رابطه‌ی فوق  $R(x_0, y_0)$  یک همسایگی نقطه‌ی  $(x_0, y_0)$  می‌باشد. در پیاده‌سازی این روش ابعاد همسایگی  $5 \times 5$  در نظر گرفته شده‌است. جهت کمینه‌سازی رابطه‌ی (۴-۱۷) کفایت تا مشتق آن را برابر ۰ قرار دهیم.

$$\begin{aligned} \frac{\partial E(x_0, y_0)}{\partial v_x} &= \sum_{x, y \in R(x_0, y_0)} 2 \left( \frac{\partial I(x, y)}{\partial x} \right) (v_x \frac{\partial I(x, y)}{\partial x} + v_y \frac{\partial I(x, y)}{\partial y} + \Delta I(x, y)) \\ &= \sum_{x, y \in R(x_0, y_0)} (2v_x \frac{\partial I(x, y)}{\partial x} \frac{\partial I(x, y)}{\partial x} + 2v_y \frac{\partial I(x, y)}{\partial x} \frac{\partial I(x, y)}{\partial y} + 2 \frac{\partial I(x, y)}{\partial x} \Delta I(x, y)) = 0 \quad (۴-۱۸) \\ \Rightarrow v_x \sum \left( \frac{\partial I(x, y)}{\partial x} \right)^2 + v_y \sum \frac{\partial I(x, y)}{\partial x} \frac{\partial I(x, y)}{\partial y} + \sum \frac{\partial I(x, y)}{\partial x} \Delta I(x, y) &= 0 \end{aligned}$$

با اعمال همین مراحل برای  $v_y$  خواهیم داشت.

$$\begin{aligned} \frac{\partial E(x_0, y_0)}{\partial v_y} &= 0 \quad (۴-۱۹) \\ \Rightarrow v_x \sum \frac{\partial I(x, y)}{\partial x} \frac{\partial I(x, y)}{\partial y} + v_y \sum \left( \frac{\partial I(x, y)}{\partial y} \right)^2 + \sum \frac{\partial I(x, y)}{\partial y} \Delta I(x, y) &= 0 \end{aligned}$$

بیان ماتریسی معادلات (۴-۱۸) و (۴-۱۹) به صورت زیر می‌باشد.

$$\begin{bmatrix} \sum \left( \frac{\partial I(x, y)}{\partial x} \right)^2 & \sum \frac{\partial I(x, y)}{\partial x} \frac{\partial I(x, y)}{\partial y} \\ \sum \frac{\partial I(x, y)}{\partial x} \frac{\partial I(x, y)}{\partial y} & \sum \left( \frac{\partial I(x, y)}{\partial y} \right)^2 \end{bmatrix} \begin{bmatrix} v_x \\ v_y \end{bmatrix} = - \begin{bmatrix} \sum \frac{\partial I(x, y)}{\partial x} \Delta I(x, y) \\ \sum \frac{\partial I(x, y)}{\partial y} \Delta I(x, y) \end{bmatrix} \quad (۴-۲۰)$$

در معادله‌ی فوق بدست آوردن جواب‌ها و در نتیجه تعیین مقدار شارش نوری به سادگی صورت می‌گیرد.

جهت پیاده‌سازی الگوریتم شارش نوری بر روی مدل‌های سه‌بعدی چهره، می‌بایست ابتدا با نمایش

مدل چهره‌ی سه‌بعدی در دستگاه مختصات استوانه‌ای، از هر مدل، دو تصویر دوبعدی استخراج کنیم. برای

این منظور کفایت تا از مدل چهره در مقادیر مختلف  $h$  و  $\phi$  نمونه‌برداری شود. همانطور که در بخش ۳-

۴-۲ اشاره شد،  $h$  و  $\phi$  هر یک به ۵۱۲ بخش کوانتیزه شده و مقدار  $r$  و بافت چهره در این نقاط استخراج

می‌شوند. بدین ترتیب می‌توان چهره را به صورت ۴ بردار تابع  $h$  و  $\phi$  نشان داد.

$$I(h, \phi) = (r(h, \phi), R(h, \phi), G(h, \phi), B(h, \phi))^T; \quad h, \phi \in \{0, \dots, 511\} \quad (۴-۲۱)$$

جهت بدست آوردن مقادیر  $h$ ،  $\phi$  و  $r$  در رابطه‌ی فوق از معادلات ساده‌ی تبدیل دستگاه مختصات کارتزین به دستگاه مختصات استوانه‌ای استفاده می‌کنیم.

$$\begin{aligned} h &= y \\ \phi &= \text{atan}(x/z) \\ r &= \sqrt{(x^2 + z^2)} \end{aligned} \quad (22-4)$$

شکل ۳-۶، مقادیر  $r$  و بافت یک مدل چهره‌ی سه‌بعدی را در دستگاه مختصات استوانه‌ای نشان می‌دهد. در این مرحله رابطه‌ی (۴-۱۷) می‌بایست بر حسب  $h$  و  $\phi$  بیان شود. علاوه بر این، از آنجایی که  $I(h, \phi)$  تابع چهار پارامتر  $r(h, \phi)$ ،  $R(h, \phi)$ ،  $G(h, \phi)$  و  $B(h, \phi)$  است، خواهیم داشت:

$$E(h_0, \phi_0) = \sum_{h, \phi \in R(h_0, \phi_0)} \left\| v_h \frac{\partial I(h, \phi)}{\partial h} + v_\phi \frac{\partial I(h, \phi)}{\partial \phi} + \Delta I(h, \phi) \right\|^2 \quad (23-4)$$

که در رابطه‌ی فوق نُرم به صورت زیر تعریف می‌شود.

$$\|I\|^2 = w_r r^2 + w_R R^2 + w_G G^2 + w_B B^2 \quad (24-4)$$

وزن‌های  $w_r$ ،  $w_R$ ،  $w_G$  و  $w_B$  جهت جبران‌سازی اندازه‌ی متفاوت مقادیر  $r$ ،  $R$ ،  $G$  و  $B$  در نظر گرفته شده‌اند. ما این مقادیر را متناسب با معکوس متوسط تغییرات پارامترهای مربوطه بر حسب تغییرات  $h$  و  $\phi$  در نظر گرفته‌ایم.

با توجه به موارد فوق، معادل رابطه‌ی (۴-۲۰) نیز به صورت زیر بدست می‌آید.

$$\begin{bmatrix} \sum \left\| \frac{\partial I(h, \phi)}{\partial h} \right\|^2 & \sum \left\langle \frac{\partial I(h, \phi)}{\partial h}, \frac{\partial I(h, \phi)}{\partial \phi} \right\rangle \\ \sum \left\langle \frac{\partial I(h, \phi)}{\partial h}, \frac{\partial I(h, \phi)}{\partial \phi} \right\rangle & \sum \left\| \frac{\partial I(h, \phi)}{\partial \phi} \right\|^2 \end{bmatrix} \begin{bmatrix} v_h \\ v_\phi \end{bmatrix} = - \begin{bmatrix} \sum \left\langle \frac{\partial I(h, \phi)}{\partial h}, \Delta I(h, \phi) \right\rangle \\ \sum \left\langle \frac{\partial I(h, \phi)}{\partial \phi}, \Delta I(h, \phi) \right\rangle \end{bmatrix} \quad (25-4)$$

در رابطه‌ی فوق باید دقت کرد که  $I(h, \phi)$  یک بردار  $4 \times 1$  بوده و به همین دلیل در رابطه‌ی فوق از ضرب نقطه‌ای بین بردارها و نُرم درجه‌ی دو استفاده شده است.

روش شارش نوری به علت استفاده از تقریب مرتبه‌ی اول، در مواردی که تفاوت مدل‌های  $I_1$  و  $I_2$  زیاد باشد و یا مشتقات فضایی مرتبه دوم و بالاتر  $I_1$  و  $I_2$  بزرگ باشند، روش مناسبی نیست. در پایگاه داده‌ی MPI به علت دقت بالای پویسگر لیزری Cyberware™، مدل‌های ۳ بعدی تفاوت زیادی با هم ندارند و روش شارش نوری خطای قابل توجهی ایجاد نمی‌کند. با این حال به علت وجود همین میزان خطا و وجود چهره‌های غیر معمول، در [بلانز و وتر، ۲۰۰۳] جهت بدست آوردن تناظری قابل قبول بین نقاط و ساخت مدل میان‌ساز از یک الگوریتم تکراری و با حجم محاسباتی بالا استفاده شده است. این الگوریتم در بخش ۴-۳-۲ تشریح خواهد شد.

در پایگاه داده‌ی ارائه شده، به علت تفاوت بیشتر بین مدل‌های سه‌بعدی ناشی از روش تهیه‌ی مدل‌ها، استفاده از روش شارش نوری مناسب نبوده و استفاده از الگوریتم تکراری مذکور نیز نیازمند تعداد زیادی تکرار می‌باشد. به همین دلیل جهت بدست آوردن تناظر، از روش‌های بدون تقریب همچون روش‌های مبتنی بر همبستگی محلی استفاده کرده‌ایم. در این روش سعی بر این است که تابع خطایی همچون فاصله اقلیدسی کمینه شود.

$$E = \sqrt{(\|I_2(x + \Delta x, y + \Delta y)\| - \|I_1(x, y)\|)^2} \quad (۴-۲۶)$$

در معادله فوق  $\bar{I}_1$  و  $\bar{I}_2$  یک همسایگی از مدل‌های  $I_1$  و  $I_2$  در اطراف نقطه  $(x, y)$  هستند. به عبارت دیگر هدف این است که همبستگی بین دو همسایگی  $\bar{I}_1$  و  $\bar{I}_2$  توسط انتخاب مقادیر مناسب  $\Delta x$  و  $\Delta y$  بیشینه شود. ماتریس‌های  $\Delta x$  و  $\Delta y$  ماتریس‌های جابجایی و تعیین کننده تناظر بین نقاط دو مدل  $I_1$  و  $I_2$  می‌باشند.  $\|I\|$  نیز با در نظر گرفتن بردار متعامد بر هر نقطه از رویه چهره به صورت زیر تعریف می‌شود.

$$\|I\| = \frac{1}{2} (w_r r^2 + w_R R^2 + w_G G^2 + w_B B^2 + w_{nx} N_x^2 + w_{ny} N_y^2 + w_{nz} N_z^2) \quad (4-27)$$

در رابطه فوق  $N_x$ ،  $N_y$  و  $N_z$  مؤلفه‌های بردار متعامد بر رویه چهره می باشند. استفاده از بردار

متعامد منجر به بهبود نتایج تناظر می‌شود. ضرایب  $w_r$ ،  $w_R$ ،  $w_G$ ،  $w_B$  و  $w_N$  نیز با توجه به متوسط

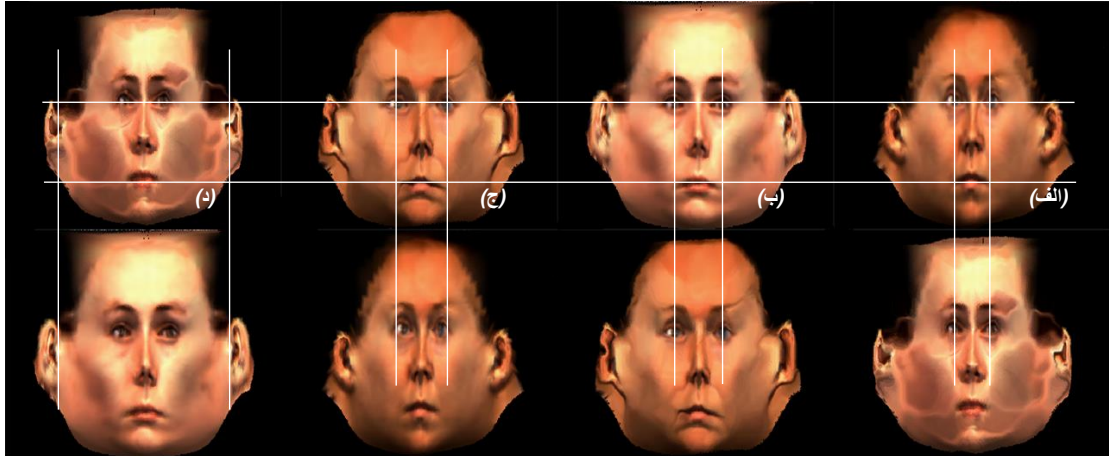
تغییرات پارامترهای مربوطه بر حسب تغییرات  $h$  و  $\phi$  تعیین می‌شوند.

در شکل ۴-۶ نتیجه‌ی انطباق دو مدل چهره در دستگاه مختصات استوانه‌ای با استفاده از این روش

نشان داده شده‌است. مشاهده می‌شود که با وجود تفاوت بسیار زیاد بین دو مدل، این روش تنها در یک

مرحله، انطباق مناسبی ایجاد کرده است. خطوط، جهت مقایسه‌ی فواصل بخش‌های مختلف چهره، ترسیم

شده است.



شکل ۴-۶- انطباق مدل‌های چهره در دستگاه مختصات استوانه‌ای با استفاده از روش همبستگی محلی،

(الف): مدل چهره اول، (ب): مدل چهره دوم، (ج): انطباق چهره‌ی اول بر چهره‌ی دوم، (د): انطباق مدل چهره‌ی دوم بر اول،

(ردیف دوم): تصاویر ردیف اول به ترتیب عکس جهت مقایسه



در این بخش به ارائه‌ی الگوریتمی جدید و مبتنی بر گراف جهت بهینه‌سازی و یک‌به‌یک‌سازی نتایج انطباق تصاویر می‌پردازیم. در ابتدا به بررسی نواقص موجود در الگوریتم‌های انطباق پرداخته و سپس جهت رفع آنها الگوریتم مورد نظر را ارائه می‌کنیم. در اغلب موارد نواقص الگوریتم‌های انطباق یا به علت استفاده از تقریب‌ها و یا به علت استفاده کردن از همسایگی‌ها ایجاد می‌شود. همین دو عامل در الگوریتم‌های شارش نوری و همبستگی محلی باعث ایجاد نواقصی می‌شوند که به بررسی آن‌ها می‌پردازیم.

در الگوریتم شارش نوری، خطای خطی‌سازی  $E_2$  که در معادله (۴-۱۳) از آن صرف نظر شده است،

برابر است با :

$$E_2 = \frac{\Delta x^2}{2!} \frac{\partial^2 I_2(x, y)}{\partial x^2} + \frac{\Delta y^2}{2!} \frac{\partial^2 I_2(x, y)}{\partial y^2} + \dots \quad (۴-۲۸)$$

این خطا تابعی از  $I_2(x, y)$  است. حال اگر جای  $I_1$  و  $I_2$  را عوض کنیم، یعنی تصویر اول را بر روی تصویر دوم منطبق کنیم، خطای مذکور تابعی از  $I_1(x, y)$  خواهد بود که با توجه به تفاوت بین  $I_1$  و  $I_2$ ، این دو مقدار نیز متفاوت می‌باشند. در نتیجه ماتریس‌های  $\Delta x$  و  $\Delta y$  به دست آمده از انطباق تصویر اول بر روی تصویر دوم، با ماتریس‌های  $\Delta x$  و  $\Delta y$  ناشی از انطباق تصویر دوم بر روی تصویر اول متفاوت هستند. ما این دو حالت انطباق را به ترتیب انطباق ۱ به ۲ و انطباق ۲ به ۱ می‌نامیم.

به عبارت دیگر اگر نقطه‌ی  $a_1$  از تصویر اول در نتیجه‌ی انطباق ۱ به ۲ با نقطه‌ی  $b_1$  از تصویر دوم متناظر در نظر گرفته شود، انطباق ۱ به ۲ نقطه‌ی  $b_1$  را الزاماً با  $a_1$  متناظر نمی‌داند. علاوه بر این امکان دارد که چند نقطه از تصویر اول به یک نقطه در تصویر دوم منطبق شوند. این مسئله یعنی غیریک‌به‌یک بودن انطباق نیز منجر به ایجاد برخی مشکلات در بعضی از کاربردها همچون ماتشدگی تصاویر سنتز شده در مدل‌سازی سه

بعدی می‌شود. هدف ما در ارائه‌ی این الگوریتم انتخاب بهینه بین انطباق ۱ به ۲ و انطباق ۲ به ۱ در هر نقطه از

تصاویر می‌باشد. علاوه بر این یک‌به‌یک‌سازی انطباق نیز توسط الگوریتم ارائه شده قابل انجام می‌باشد.

در الگوریتم مبتنی بر همبستگی محلی، با توجه به این که همسایگی در تصویر دوم حول نقطه‌ی

$(x+\Delta x, y+\Delta y)$  در نظر گرفته می‌شود، با وجود این که از تقریب‌های خطی استفاده نشده است، به علت

استفاده از همسایگی‌ها، نتایج انطباق ۱ به ۲ با نتایج انطباق ۲ به ۱ برابر نیست.

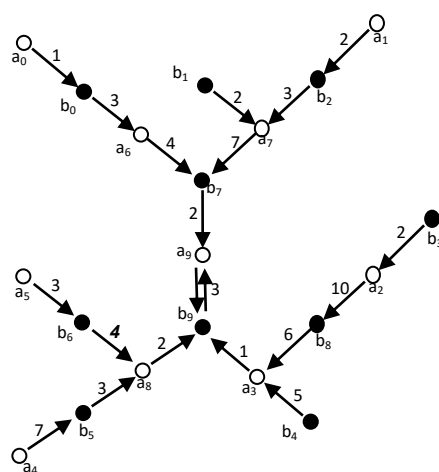
### گراف‌ها و بهینه‌سازی انطباق تصاویر:

اگر نقاط دو تصویر ۱ و ۲ را که توسط انطباق‌های ۱ به ۲ و ۲ به ۱ به هم مربوط می‌شوند به صورت

گراف نمایش دهیم، یک درخت جهت‌دار و وزن‌دار مشابه شکل ۴-۷ بوجود می‌آید که آن را درخت انطباق

می‌نامیم. در این درخت گره‌های سفید مربوط به پیکسل‌های تصویر اول و گره‌های سیاه مربوط به

پیکسل‌های تصویر دوم می‌باشند.



شکل ۴-۷ - یک درخت انطباق نمونه

فلش‌های جهت‌دار متصل‌کننده‌ی دو گره که پیوند نامیده می‌شود، بیانگر انطباق بین دو پیکسل

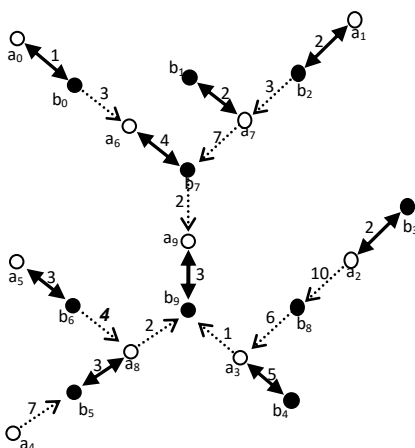
از تصاویر مورد مطالعه می‌باشد. همانطور که اشاره شد جهت در این پیوند بیانگر انطباق یکطرفه می‌باشد.

وزن هر پیوند برابر مقدار خطا و یا فاصله‌ای است که در محاسبه‌ی انطباق نقطه مربوطه بوجود آمده است.

همان‌طور که در شکل ۴-۷ مشخص است، درخت از برگ‌ها (گره‌های انتهایی) شروع شده و به یک پیوند دوطرفه ختم می‌شود. دو گره مربوط به پیوند دوطرفه، گره‌هایی هستند که در دو انطباق ۱ به ۲ و ۲ به ۱ به یکدیگر اشاره می‌کنند. به عبارت دیگر گرهی  $a_0$  از تصویر اول توسط انطباق ۱ به ۲ به گره  $b_0$  از تصویر دوم منطبق شده و گره  $b_0$  از تصویر دوم نیز توسط انطباق ۲ به ۱ به گره  $a_0$  از تصویر اول منطبق می‌شود. انطباق یک‌به‌یک انطباقی است که در آن تمام پیوندهای درخت انطباق مربوطه از نوع دوطرفه بوده و هیچ گره‌ای بیش از یک پیوند ورودی نداشته باشد.

همان‌طور که اشاره شد وزن هر پیوند مقدار خطا محاسبه شده در انطباق نقاط می‌باشد. برای مثال در روش شارش نوری وزن هر پیوند، مقدار  $E_2$  ارائه شده در معادلات (۴-۱۳) و (۴-۲۸) می‌باشد. در روش انطباق مبتنی بر همبستگی وزن هر پیوند، مقدار  $E$  کمینه محاسبه شده در معادله (۴-۲۶) است. بنابراین هر چه وزن پیوند کمتر باشد، بیانگر این است که انطباق گره‌های مربوط به آن پیوند، انطباق مناسب‌تری است.

زمانی که فقط انطباق ۱ به ۲ را در نظر می‌گیریم، در درخت انطباق شکل ۴-۷، از تمامی پیوندهایی که گره‌های سیاه را به گره‌های سفید متصل می‌کنند، صرف نظر کرده‌ایم. در حالی که بعضی از این پیوندها خطای کمتری نسبت به پیوندهای گره‌های سفید به سیاه ایجاد می‌کنند. در نتیجه استفاده همزمان از انطباق‌های ۱ به ۲ و ۲ به ۱ منجر به کاهش خطا و بهینه‌سازی انطباق می‌شود. همچنین می‌توان پیوندها را به گونه‌ای انتخاب کرد که هر گره تنها دارای یک ورودی باشد. بدین ترتیب علاوه بر کاهش خطای انطباق به انطباقی یک‌به‌یک دست پیدا کرد. البته باید توجه کرد که یک‌به‌یک‌سازی انطباق منجر به از دست دادن خاصیت پوشایی و یا کاهش چگالی انطباق می‌شود. به عبارت دیگر بعضی از گره‌ها همچون گره  $b_8$  در شکل ۱ به هیچ گره‌ای از تصویر اول منطبق نمی‌شوند.



شکل ۴-۸ - پیوندهای بهینه درخت انطباق شکل ۴-۷

بنابراین جهت بهینه‌سازی و یک‌به‌یک‌سازی انطباق تصاویر، پیوندهایی از درخت انطباق را انتخاب می‌کنیم که دارای کمترین مجموع وزن بوده و هر گره فقط دارای یک پیوند ورودی و یا خروجی باشد. این پیوندها را پیوندهای بهینه می‌نامیم. پیوندهای بهینه مربوط به درخت انطباق شکل ۴-۷، در شکل ۴-۸ نمایش داده شده است. پس از انتخاب پیوندهای بهینه، آن‌ها را از نوع دوطرفه در نظر می‌گیریم. بنابراین ماتریس انطباق جدیدی تشکیل می‌شود که بهینه، یک‌به‌یک و دوطرفه می‌باشد.

#### تعاریف مهم

قبل از تشریح روش انتخاب پیوندهای بهینه و یک‌به‌یک‌سازی، عبارات مهم مورد استفاده را تعریف می‌کنیم.

زیردرخت: زیرمجموعه‌ای از گره‌های درخت است که بین هر دو گره آن مسیری متشکل از گره‌های عضو زیرمجموعه وجود داشته باشد.

زیردرخت سرشاخه: زیردرختی است که از برگ‌های درخت شروع شده و تا اولین گره چند ورودی

ادامه پیدا می‌کند.

زیردرخت کامل: زیردرختی است که دارای یک گره انتهایی بوده و تمام گره‌های پیشین را که با گره انتهایی تشکیل مسیر می‌دهند، شامل می‌شود.

گره گسترش‌دهنده: گره بعد از آخرین گره زیردرخت کامل گره گسترش‌دهنده نامیده می‌شود. پیوند ورودی به این گره، پیوند گسترش‌دهنده نامیده می‌شود.

زیردرخت گسترش‌یافته: زیردرختی است که از افزودن گره گسترش‌دهنده به زیردرخت کامل ایجاد می‌شود.

پیوندهای بهینه گسترش‌یافته: زیرمجموعه‌ای از پیوندهای یک زیردرخت گسترش‌یافته و یا سرشاخه است که دارای کمترین مجموع وزن بوده و پیوند انتهایی را نیز شامل شود. پیوندها باید به گونه‌ای انتخاب شوند که هر گره فقط دارای یک پیوند ورودی و یا خروجی باشد.

پیوندهای بهینه گسترش‌نیافته: زیرمجموعه‌ای از پیوندهای یک زیردرخت گسترش‌یافته و یا سرشاخه است که دارای کمترین مجموع وزن بوده و پیوند انتهایی را شامل نشود. پیوندها باید به گونه‌ای انتخاب شوند که هر گره فقط دارای یک پیوند ورودی و یا خروجی باشد.

پیوند دوطرفه: در هر درخت انطباق یک پیوند دوطرفه وجود دارد که همواره به عنوان یکی از پیوندهای بهینه در نظر گرفته می‌شود. دو گره تشکیل‌دهنده این پیوند گره‌های پایانی نامیده می‌شوند.

#### روش انتخاب پیوندهای بهینه

جهت انتخاب پیوندهای بهینه ابتدا از زیردرخت‌های سرشاخه شروع کرده و پیوندهای بهینه گسترش‌یافته و گسترش‌نیافته هر زیردرخت سرشاخه را بدست می‌آوریم. نحوه انتخاب پیوندهای بهینه زیردرخت سرشاخه در ادامه این بخش تشریح می‌شود.

در این مرحله زیردرخت‌های سرشاخه‌ای که دارای گره انتهایی یکسان هستند، یک زیردرخت کامل را تشکیل می‌دهند که با افزودن یک گره گسترش‌دهنده به این زیردرخت، یک زیردرخت گسترش‌یافته بوجود می‌آید. در این مرحله پیوندهای بهینه گسترش‌یافته و گسترش‌نیافته زیردرخت گسترش‌یافته فوق را با استفاده از پیوندهای بهینه زیردرخت‌های سرشاخه تشکیل‌دهنده، ایجاد می‌کنیم. نحوه انتخاب پیوندهای بهینه زیردرخت گسترش‌یافته نیز در ادامه این بخش تشریح می‌شود.

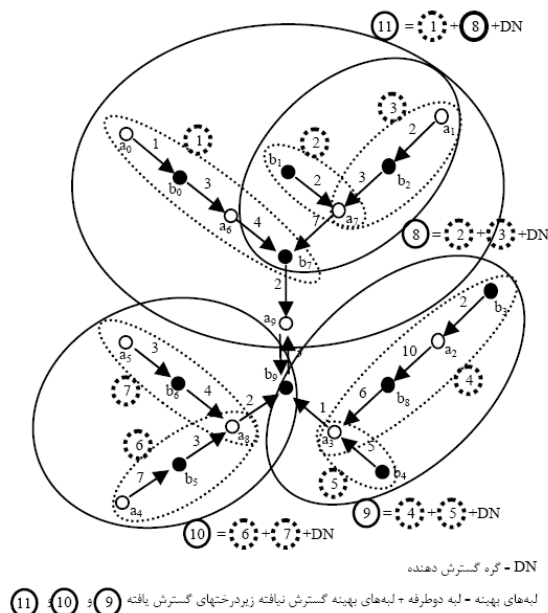
زیردرخت‌های گسترش‌یافته و سرشاخه دارای گره انتهایی یکسان به همراه گره گسترش‌دهنده، یک زیردرخت گسترش‌یافته جدید تشکیل می‌دهند. در این مرحله پیوندهای بهینه زیردرخت گسترش‌یافته فوق را با استفاده از پیوندهای بهینه زیردرخت‌های سرشاخه و گسترش‌یافته تشکیل‌دهنده، بدست می‌آوریم. عملیات فوق را تا رسیدن به هر یک از گره‌های پایانی ادامه می‌دهیم. در پایان، پیوندهای بهینه، برابر با مجموع پیوندهای گسترش‌نیافته زیردرخت‌های مختوم به گره‌های پایانی، به اضافه پیوند دوطرفه در نظر گرفته می‌شود. شکل ۴-۹ مراحل تعیین پیوندهای بهینه زیردرخت‌های گسترش‌یافته و سرشاخه را در درخت انطباق شکل ۴-۷ نشان می‌دهد. در این شکل، زیردرخت‌های سرشاخه توسط نقطه‌چین و زیردرخت‌های گسترش‌یافته توسط خطوط ممتد مشخص شده‌اند.

حال برای هر زیردرخت سرشاخه نحوه بدست آوردن پیوندهای بهینه گسترش‌یافته و گسترش‌نیافته را تشریح می‌کنیم. از آنجایی که زیردرخت‌های سرشاخه از گره‌های تک‌ورودی تشکیل شده‌اند، انتخاب پیوندهای بهینه به سادگی صورت می‌گیرد. جهت انتخاب پیوندهای بهینه گسترش‌یافته، ابتدا پیوند انتهایی را بدون توجه به وزن آن انتخاب می‌کنیم و تمام پیوندهای متصل به دو گره انتخاب شده را حذف می‌کنیم. در مرحله بعدی پیوند دارای کمترین وزن را انتخاب کرده و پیوندهای متصل به دو گره انتخاب شده را حذف

می‌کنیم. این عملیات را تا اتمام همه پیوندهای زیردرخت سرشاخه ادامه می‌دهیم. جهت انتخاب پیوندهای

بهینه گسترش نیافته زیردرخت سرشاخه، همانند نحوه انتخاب پیوندهای بهینه گسترش یافته عمل می‌کنیم.

با این تفاوت که در مرحله اول گره و پیوند انتهایی حذف می‌شوند.



شکل ۴-۹ - مراحل تعیین پیوندهای بهینه

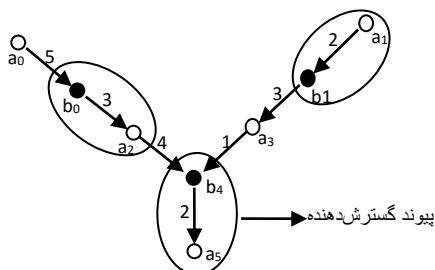
زیردرخت گسترش یافته از یک گره گسترش دهنده و یک زیردرخت کامل تشکیل شده است.

زیردرخت کامل نیز خود از مجموع چند زیردرخت گسترش یافته و یا سرشاخه دارای گره انتهایی یکسان

تشکیل شده است. این زیردرخت‌های سرشاخه و گسترش یافته را، زیردرخت‌های تشکیل دهنده می‌نامیم.

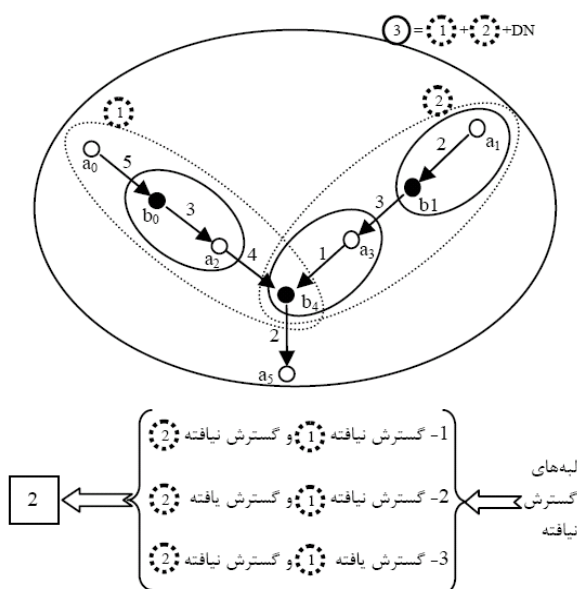
پیوندهای بهینه گسترش یافته زیردرخت گسترش یافته از مجموع پیوند گسترش دهنده و پیوندهای

بهینه گسترش نیافته تمامی زیردرخت‌های تشکیل دهنده بدست می‌آید. (شکل ۴-۱۰)



شکل ۴-۱۰ - پیوندهای بهینه گسترش یافته زیردرخت گسترش یافته

پیوندهای بهینه گسترش نیافته زیردرخت گسترش یافته از مجموع پیوندهای بهینه گسترش یافته حداکثر یکی از زیردرخت‌های تشکیل دهنده و پیوندهای بهینه گسترش نیافته سایر زیردرخت‌های تشکیل دهنده بدست می‌آید. اگر تعداد زیردرخت‌های تشکیل دهنده برابر  $n$  باشد،  $n+1$  انتخاب جهت پیوندهای بهینه گسترش نیافته خواهیم داشت. انتخاب نهایی انتخابی است که کمترین نسبت مجموع وزن به تعداد پیوندهای بهینه را ایجاد کند. برای ۲ زیردرخت تشکیل دهنده، ۳ انتخاب ممکن در شکل ۴-۱۱ نمایش داده شده است.



شکل ۴-۱۱ - پیوندهای بهینه گسترش نیافته زیردرخت گسترش یافته و انتخاب‌های ممکن

### پیاده‌سازی و بررسی نتایج:

روش تطبیق مبتنی بر همبستگی محلی روشی کندتر ولی دقیق‌تر نسبت به روش تخمین شارش نوری است. بنابراین در این بخش، بهینه‌سازی انطباق بر روی نتایج الگوریتم همبستگی محلی ارائه می‌شود. تصاویر ورودی مدل‌های سه‌بعدی چهره در دستگاه مختصات استوانه‌ای هستند که در شکل ۴-۱۲ نمایش داده شده‌اند. با اجرای این الگوریتم، درخت‌های انطباقی تا ۳۰۰۰ گره، مورد پردازش و بهینه‌سازی قرار



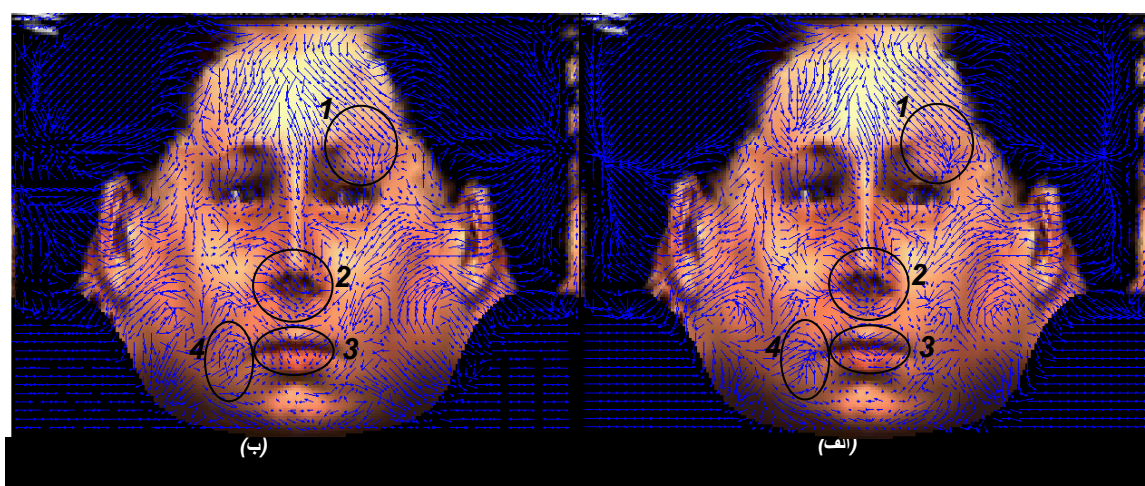
گرفتند. درخت‌هایی با این ابعاد در نقاط دارای اطلاعات کم همچون روی گونه‌ها، صورت و پشت سر بوجود می‌آیند، اما متوسط ابعاد درخت‌های انطباق در این تصاویر، حدود ۳۰ گره می‌باشد.



شکل ۴-۱۲ - تصاویر ورودی

با بهینه‌سازی انطباق توسط الگوریتم ارائه شده، نتایج بهبود قابل توجهی پیدا کرده است. در شکل ۴-۱۳ (الف) نتایج انطباق مبتنی بر همبستگی محلی نشان داده شده است. نتایج یک‌به‌یک و بهینه‌سازی شده نیز در شکل ۴-۱۳ (ب) ارائه شده است.

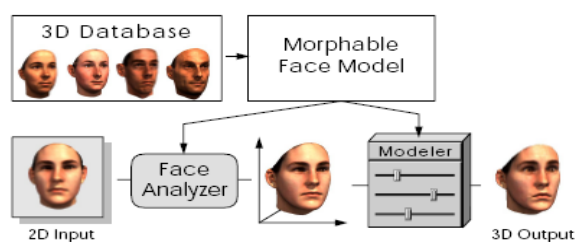
بهینه‌سازی انطباق در نواحی ۲ و ۳ در شکل ۴-۱۳ کاملاً مشخص می‌باشد. در این نواحی در شکل ۴-۱۳ (ب) درهم‌ریختگی در بردارهای جابجایی مشاهده نمی‌شود، که بیانگر بهبود عملکرد انطباق است. همچنین نواحی ۱ و ۴ در این شکل‌ها، مثال‌های خوبی از یک‌به‌یک‌سازی انطباق هستند. در این نواحی در شکل ۴-۱۳ (ب) به جای اشاره بردارها به یک نقطه، بردارها به صورت منظم پراکنده شده‌اند.



شکل ۴-۱۳ - (الف): نتیجه انطباق ۱ به ۲ توسط همبستگی محلی، (ب): بهینه‌سازی و یک‌به‌یک‌سازی انطباق ۱ به ۲

## ۱۹-۱ کاربرد تصاویر ورودی و مدل میان‌ساز

بخش بسیار مهم تشخیص چهره با استفاده از مدل میان‌ساز یافتن الگوریتمی خودکار جهت تطبیق مدل چهره‌ی میان‌ساز با تصاویر ورودی است. برای این منظور این الگوریتم می‌بایست بردارهای  $\bar{\alpha}$  و  $\bar{\beta}$  را به گونه‌ای بیابد که تصویرسازی دوبعدی مدل ایجاد شده، تصویری مشابه با تصویر چهره‌ی ورودی ایجاد کند. بنابراین علاوه بر بردارهای  $\bar{\alpha}$  و  $\bar{\beta}$ ، تمام پارامترهای مربوط به مدل نوردهی، دید سه‌بعدی و تصویرسازی پرسپکتیو همچون جهت تابش نور، شدت و رنگ منبع نور نقطه‌ای، زاویه‌ی دید، موقعیت دوربین و فاصله‌ی کانونی دوربین می‌بایست تخمین زده شوند.



شکل ۴-۱۴ - ساخت مدل میان‌ساز و سنتز چهره ورودی [بلانز و وتر، ۱۹۹۹]

این الگوریتم توسط بلانز و وتر [بلانز و وتر، ۱۹۹۹ - بلانز و وتر ۲۰۰۳] به صورت یک "حلقه‌ی تحلیل بوسیله‌ی سنتز"<sup>۱۵۸</sup> ارائه شده است. در این روش در هر مرحله پارامترهای مذکور تخمین زده شده و تصویری دوبعدی بر اساس این پارامترها سنتز می‌شود. سپس تصویر سنتز شده با تصویر ورودی مقایسه شده و با توجه به میزان خطا پارامترهای مورد نظر به‌روز<sup>۱۵۹</sup> می‌شوند. این روند تا به حداقل رسیدن خطا و رسیدن به حد مطلوب بهینه‌سازی ادامه پیدا می‌کند. شکل ۴-۱۴ مراحل مختلف این الگوریتم را نشان می‌دهد. به منظور اختصاص مقادیر اولیه کاربر می‌بایست موقعیت ۷ نقطه راهنما را بر روی تصویر ورودی معین کند. این ۷ نقطه با توجه به زاویه‌ی تصویربرداری در تصویر ورودی معین می‌شوند. بنابراین کاربر

<sup>158</sup> Analysis by Synthesis Loop

<sup>159</sup> Update

می‌بایست این ۷ نقطه را بر روی تصویر ورودی و مدل چهره معین کند. موقعیت این نقاط مقدار اولیه‌ی زاویه‌ی دید و مقیاس تصویر ورودی را مشخص می‌کند. این مقادیر اولیه بر حسب زاویه‌ی بین نقاط راهنما و فواصل نسبی آنها تعیین می‌شود. ما در پیاده‌سازی این مدل میان‌ساز از روشی سریع و ساده‌تر جهت تعیین مقادیر اولیه‌ی زاویه‌ی دید و مقیاس تصویر ورودی استفاده کرده‌ایم. در این روش یک مدل نمونه توسط کاربر از لحاظ زاویه‌ی دید و مقیاس بر روی تصویر ورودی منطبق می‌شود. انجام این کار برای کاربر، عملی به مراتب سریع‌تر و ساده‌تر بوده و محاسبه‌ی زاویه‌ی دید و مقیاس نیز به سادگی صورت می‌گیرد. انطباق مدل نمونه و تصویر ورودی توسط یک ابزار تعاملی انجام شده که در شکل ۴-۱۵ نشان داده شده است. در این بخش ابتدا مراحل سنتز تصویر و سپس مراحل تطبیق مدل و چهره مطرح می‌شود.



شکل ۴-۱۵ - ابزار تعاملی جهت اختصاص مقادیر اولیه‌ی زاویه‌ی دید و مقیاس

## ۱-۱۹-۱ سنتز تصویر

در این بخش نحوه‌ی ساخت تصویر دوبعدی از روی مدل سه‌بعدی تشریح می‌شود.

تصویرسازی پرسپکتیو:

موقعیت هر رأس  $k$  در فضا را با  $X_k = (x_k, y_k, z_k)^T$  نمایش می‌دهیم. در این صورت طبق مطالب

مطرح شده در فصل ۲ و بخش ۳-۳-۱، مختصات این نقطه در دستگاه مختصات دید به صورت زیر می‌باشد.

$$(w_{x,k}, w_{y,k}, w_{z,k}) = R_\gamma R_\theta R_\phi X_k + t_w \quad (28-4)$$

در رابطه‌ی فوق ماتریس‌های R ماتریس‌های دوران و بردار  $t_w$  بردار انتقال می‌باشد. تصویر سازی پرسپکتیو نیز به صورت زیر در نظر گرفته می‌شود.

$$p_{x,k} = P_x + f \frac{w_{x,k}}{w_{z,k}} \quad p_{y,k} = P_y + f \frac{w_{y,k}}{w_{z,k}} \quad (29-4)$$

رابطه‌ی فوق در واقع همان رابطه‌ی (3-5) می‌باشد. با این تفاوت که در (4-29) فرض قرارگیری نقطه‌ی مرجع پرسپکتیو بر روی محور  $Z_v$  در نظر گرفته نشده است.  $P_x$  و  $P_y$  در واقع همان متغیرهای  $X_{prp}$  و  $Y_{prp}$  می‌باشند که در شکل 2-9 و رابطه‌ی (2-36) برابر صفر در نظر گرفته شده‌اند.

نوردهی و رنگ:

مشخصات رنگ در هر رأس طبق رابطه‌ی (2-54) محاسبه شده و به پیکسل مربوطه بر روی تصویر

دو بعدی سنتز شده اعمال می‌شود.

$$\begin{aligned} I_{R,ref} &= k_a R I_{aR} + \sum_{i=1}^n R I_{iR_i} [k_d (N \cdot L_i) + k_s (N \cdot H_i)^{n_s}] \\ I_{G,ref} &= k_a G I_{aG} + \sum_{i=1}^n G I_{iG_i} [k_d (N \cdot L_i) + k_s (N \cdot H_i)^{n_s}] \\ I_{B,ref} &= k_a B I_{aB} + \sum_{i=1}^n B I_{iB_i} [k_d (N \cdot L_i) + k_s (N \cdot H_i)^{n_s}] \end{aligned} \quad (30-4)$$

اما هدف روش مدل سازی میان ساز تنها استفاده از مدل‌های نوردهی واقعی نیست. امکان سنتز تصاویر چهره در تون‌های مختلف رنگی همچون تصاویر تکفام (برای مثال تصاویر سیاه و سفید) و نقاشی‌های چهره نیز از اهداف این روش می‌باشد. برای این منظور بهره‌های  $g_r$ ،  $g_r$ ،  $g_r$ ، مقادیر ثابت  $o_r$ ،  $o_r$ ،  $o_r$  و کنتراست رنگ

c به هر مولفه‌ی رنگ اعمال می‌شوند. لومینانس<sup>۱۶۰</sup> کلی تصویر نیز به صورت زیر تعریف می‌شود [فولی و همکاران، ۱۹۹۶].

$$L = 0.3 \cdot I_{R,ref} + 0.59 \cdot I_{G,ref} + 0.11 \cdot I_{B,ref} \quad (۳۱-۴)$$

کنتراست رنگ نسبت رنگ اصلی و لومینانس را مشخص می‌کند. بنابراین میزان نهایی رنگ برای هر پیکسل در تصویر دوبعدی سنتز شده به صورت زیر بدست می‌آید.

$$\begin{aligned} I_R &= g_r (cI_{R,ref} + (1-c)L) + o_r \\ I_G &= g_g (cI_{G,ref} + (1-c)L) + o_g \\ I_B &= g_b (cI_{B,ref} + (1-c)L) + o_b \end{aligned} \quad (۳۲-۴)$$

تعیین سطوح قابل رویت نیز توسط روش بافر عمق صورت می‌گیرد. این روش در ۲-۳-۲ تشریح شده است. جهت پیاده‌سازی سایه‌ها نیز می‌توان از روش بافر عمق و تعیین سطوح قابل رویت توسط منبع نور استفاده کرد.

## ۱-۱۹-۲ مراحل تطبیق مدل و تصویر ورودی

تطبیق مدل با تصویر ورودی در واقع بهینه‌سازی پارامترهای مربوطه جهت کاهش اختلاف بین تصویر سنتز شده و تصویر ورودی می‌باشد. پارامترهایی که در بهینه‌سازی مورد استفاده قرار می‌گیرند شامل دو بردار  $\alpha$  و  $\beta$  و ۲۲ پارامتر مربوط به سنتز تصویر دوبعدی از مدل سه‌بعدی می‌شود. بردارهای  $\alpha$  و  $\beta$ ، بردارهایی دارای ابعاد حداقل  $100 \times 1$  می‌باشند. این ۲۲ پارامتر را توسط بردار  $\rho$  نمایش می‌دهیم. زوایای  $\gamma, \theta, \phi$  مربوط به زاویه‌ی تصویربرداری،  $t_{wx}, t_{wy}, t_{wz}$  مربوط به فصله‌ی دوربین نسبت به مبداء، فاصله‌ی کانونی  $f$ ،  $I_{a,R}, I_{a,G}, I_{a,B}$  مربوط به نور محیطی، پارامترهای رنگ منبع نور  $I_{t,R}, I_{t,G}, I_{t,B}$  بردارهای  $L$  و  $H$  که توسط زوایای منبع نقطه‌ای یعنی  $\theta_t$  و  $\phi_t$  قابل تعیین هستند، کنتراست رنگ  $c$  و بهره‌ها و مقادیر<sup>۱۶۰</sup> لومینانس مقدار انرژی‌ای است که یک ناظر از منبع نور دریافت می‌کند می‌توان بجای منبع نور، یک شیء با خاصیت بازتابشی را در نظر گرفت.

ثابت رنگ یعنی  $g_r, g_b, g_g, o_r, o_g, o_b$  ، پارامترهای تشکیل دهنده‌ی بردار  $\rho$  می‌باشند. بنابراین بردار  $\rho$  به صورت زیر تعریف می‌شود.

$$\rho = (\gamma, \theta, \phi, t_{wx}, t_{wy}, t_{wz}, f, I_{a,R}, I_{a,G}, I_{a,B}, I_{t,R}, I_{t,G}, I_{t,B}, \phi_t, \theta_t, c, g_r, g_b, g_g, o_r, o_g, o_b)^T \quad (33-4)$$

#### ۱-۱۹-۲-۱ تابع هزینه

هدف در تطبیق مدل و تصویر ورودی ساخت مدلی است که تصاویر دوبعدی ناشی از آن کمترین اختلاف ممکن را با تصویر ورودی داشته باشد. اگر تصویر سنتز شده را  $I_{model}$  بنامیم و تصویر ورودی به صورت  $I_{input}(x, y) = (I_r(x, y), I_g(x, y), I_b(x, y))^T$  تعریف شود آنگاه تابع خطا به صورت زیر در نظر گرفته می‌شود.

$$E_I = \sum_{x,y} \|I_{input}(x, y) - I_{model}(x, y)\|^2 \quad (34-4)$$

در چند مرحله‌ی تکرار اول طبق مطالب ابتدای بخش ۴-۲، تعیین پارامترها بر اساس موقعیت نقاط راهنمای مشخص شده بر روی مدل چهره و تصویر ورودی صورت می‌گیرد. نقاط راهنمای مشخص شده بر روی مدل در رئوس  $k_j$  واقع شده و موقعیت این نقاط بر روی تصویر  $I_{model}$  با توجه به روابط مربوط به بخش ۴-۲-۱ به صورت  $(p_{x,k_j}, p_{y,k_j})$  بدست می‌آید. موقعیت نقاط راهنما بر روی تصویر ورودی نیز به صورت  $(q_{x,j}, q_{y,j})$  در نظر گرفته می‌شود. حال تابع خطای مربوط به نقاط راهنما را نیز به صورت زیر در نظر می‌گیریم.

$$E_F = \sum_j \left\| \begin{pmatrix} q_{x,j} \\ q_{y,j} \end{pmatrix} - \begin{pmatrix} p_{x,k_j} \\ p_{y,k_j} \end{pmatrix} \right\|^2 \quad (35-4)$$

ما در پیاده‌سازی این روش همانطور که اشاره شد از نقاط راهنما صرف نظر کرده و مستقیماً تخمین زاویه‌ی دید و مقیاس تصویر توسط کاربر صورت گرفته است. با این حال جهت تشریح کامل روش، موارد مربوط به نقاط راهنما را نیز مطرح می‌کنیم.

کمینه‌سازی توابع فوق نسبت به پارامترهای  $\alpha$ ،  $\beta$  و  $\rho$ ، به علت در نظر نگرفتن حدود قابل قبول برای  $\alpha$  و  $\beta$  و در نظر نگرفتن اثر نویز در تصویر ورودی، ممکن است منجر به سنتز تصاویری شود که از لحاظ فاصله به تصویر ورودی نزدیک بوده ولی شباهتی به چهره‌های واقعی نداشته باشد. این پدیده معادل با تعیین نقاط کمینه‌ی محلی به جای نقاط کمینه‌ی مطلق می‌باشد. برای رفع این مشکل از تخمین‌گر احتمال پسین بیشینه<sup>۱۶۱</sup> استفاده می‌کنیم. این روش تخمین پارامتر در مرجع [دودا و همکاران، ۲۰۰۱] به طور کامل تشریح شده است. در واقع این روش همان روش تخمین‌گر درست‌نمایی بیشینه<sup>۱۶۲</sup> است که در آن اثر احتمالات پیشین<sup>۱۶۳</sup> نیز در نظر گرفته شده است. به همین دلیل این تخمین‌گر مانع تعیین پارامترهای  $\alpha$  و  $\beta$  در خارج از محدوده‌ی محتمل می‌شود. در تخمین‌گر احتمال پسین بیشینه، هدف یافتن مقادیری از پارامتر  $\theta$  است که مقدار  $l(\theta)p(\theta)$  را بیشینه کند.  $l(\theta)$  نیز به صورت  $\sum_{k=1}^n \ln(p(x_k | \theta))$  تعریف می‌شود که در آن  $x_k$ ها ویژگی‌های ورودی می‌باشند. در مسئله‌ی تطبیق مدل و تصویر پارامتر مورد تخمین یعنی  $\theta$  برابر  $\alpha$ ،  $\beta$  و  $\rho$  بوده و نقاط راهنما نقش ویژگی‌های ورودی را بر عهده دارند. در روش ارائه شده توسط بلانز و وتر [بلانز و وتر، ۲۰۰۳]، تابع  $\text{Ln}(\cdot)$  بر روی احتمالات پیشین نیز اعمال شده است. در

این روش با توجه با قاعده‌ی بیز<sup>۱۶۴</sup> احتمال پسین به صورت زیر تعریف می‌شود.

$$p(\alpha, \beta, \rho | I_{input}, F) = \frac{p(I_{input}, F | \alpha, \beta, \rho) \cdot P(\alpha, \beta, \rho)}{p(I_{input}, F)} \sim p(I_{input}, F | \alpha, \beta, \rho) \cdot P(\alpha, \beta, \rho) \quad (۳۶-۴)$$

<sup>161</sup> Maximum a Posteriori (MAP)

<sup>162</sup> Maximum Likelihood

<sup>163</sup> Prior Probability

<sup>164</sup> Bayes

با در نظر گرفتن استقلال پارامترها (به صورت تقریبی) احتمال پسین به صورت زیر مطرح می‌شود.

$$p(I_{input} | \alpha, \beta, \rho) \cdot p(F | \alpha, \beta, \rho) \cdot P(\alpha) \cdot P(\beta) \cdot P(\rho) \quad (37-4)$$

توابع توزیع  $P(\alpha)$  و  $P(\beta)$  با توجه به توزیع مدل‌های نمونه و روش PCA در رابطه‌ی (4-10) ارائه شدند.

برای  $P(\beta)$  نیز یک توزیع نرمال گوسی در نظر می‌گیریم که در آن مقدار میانگین برابر مقادیر اولیه بوده و

واریانس نیز به صورت ساده و با توجه به مولفه‌های بردار  $\rho$  تخمین زده می‌شود.

درست‌نمایی مشاهده‌ی  $I_{input}$  با توجه به مقادیر  $\alpha$ ،  $\beta$  و  $\rho$ ، تابعی از خطای  $E_I$  و نویز اعمال شده

به تصویر ورودی می‌باشد. با فرض این‌که نویز تصویر ورودی از نوع گوسی و با واریانس  $\sigma_I^2$  باشد، درست‌نمایی

مذکور به صورت زیر می‌باشد.

$$p(I_{input} | \alpha, \beta, \rho) \sim \exp\left(\frac{-E_I}{2\sigma_I^2}\right) \quad (38-4)$$

به همین ترتیب انتخاب نقاط راهنما نیز تحت تأثیر نویزی گوسی با واریانس  $\sigma_F^2$  می‌باشد. باید توجه کرد که

بخش اصلی این نویز مربوط به عدم دقت تعیین نقاط راهنما می‌باشد. بنابراین درست‌نمایی مشاهده‌ی نقاط

راهنما با توجه به مقادیر  $\alpha$ ،  $\beta$  و  $\rho$ ، به صورت زیر تعریف می‌شود.

$$p(F | \alpha, \beta, \rho) \sim \exp\left(\frac{-E_F}{2\sigma_F^2}\right) \quad (39-4)$$

با توجه به موارد فوق بیشینه‌سازی احتمال پسین معادل با کمینه‌سازی عبارت زیر می‌باشد.

$$E = -2 \ln(p(\alpha, \beta, \rho | I_{input}, F)) \\ = \frac{1}{\sigma_I^2} E_I + \frac{1}{\sigma_F^2} E_F + \sum_i \frac{\alpha_i^2}{\sigma_{S,i}^2} + \sum_i \frac{\beta_i^2}{\sigma_{T,i}^2} + \sum_{i=1}^{22} \frac{(\rho_i - \bar{\rho}_i)^2}{\sigma_{R,i}^2} \quad (40-4)$$

معادله‌ی فوق تابع هزینه‌ی مورد نظر بوده و در مرحله‌ی بعدی الگوریتمی تکراری جهت کمینه‌سازی این

تابع بررسی می‌شود. باید توجه کرد که در تکرارهای اولیه‌ی الگوریتم، به علت اهمیت بیشتر نقاط راهنما

جهت تعیین مقادیر اولیه‌ی مناسب،  $\sigma_F^2$  کوچک در نظر گرفته شده و در تکرارهای بعدی با بزرگتر نمودن



این ضریب، از اثر نقاط راهنما بر تابع هزینه کاسته خواهد شد. به عبارت دیگر با کوچک کردن توان نویز در تکرارهای اولیه، مقادیر نقاط راهنما دارای اهمیت شده و پس از آن با زیاد کردن توان نویز، نقاط راهنما نقاطی نویزی در نظر گرفته شده و از اهمیت آنها کاسته می‌شود. علاوه بر این در تکرارهای اولیه  $\sigma_I^2$  نیز مقداری بزرگ بوده و به تدریج کاهش می‌یابد.

#### ۱-۱۹-۲-۲ فرایند بهینه سازی

جهت کمینه سازی تابع هزینه ی (۴-۴۰) از روش نیوتن-رافسن با اعمال تغییراتی آماری استفاده شده است [بلانز و وتر، ۲۰۰۳]. روش نیوتن-رافسن<sup>۱۶۵</sup> از جمله ی روش‌های بهینه‌سازی گرادیانی است. در نتیجه این روش به محاسبه مشتقات تحلیلی تابع هدف نسبت به پارامترهای بهینه‌سازی نیاز دارد. نحوه ی محاسبه ی مشتقات تحلیل در ادامه ی این بخش تشریح شده و پس از آن روش نیوتن-رافسن و پیاده‌سازی آن ارائه می‌گردد. در این قسمت به بیان تغییرات آماری ارائه شده و فرایند کلی بهینه‌سازی می‌پردازیم.

الگوریتم بهینه‌سازی ارائه شده از تقریبی از تابع هزینه ناشی از انتخاب تصادفی چندضلعی‌های رویه استفاده می‌کند. برای سادگی بیشتر محاسبات مربوط به موقعیت چندضلعی‌های رویه ی چهره و بافت آنها در مرکز مرکز چند ضلعی محاسبه شده و برای تمام چندضلعی یکسان در نظر گرفته می‌شود (سایه‌گذاری یکنواخت). بنابراین در مرکز مثلث  $k$ ، بافت برابر  $(\bar{R}_k, \bar{G}_k, \bar{B}_k)$  و مختصات فضایی برابر  $(\bar{X}_k, \bar{Y}_k, \bar{Z}_k)$  می‌باشد. علامت '۲' در بالای حروف بیانگر میانگین‌گیری از مقادیر مربوط به رئوس چندضلعی می‌باشد. مرکز چندضلعی‌ها بر اثر تصویرسازی پرسپکتیو به نقطه ی  $(\bar{p}_{x,k}, \bar{p}_{y,k})$  نگاشته می‌شود که برابر میانگین مختصات

<sup>165</sup> Newton-Rophson Method

تصویر رئوس چندضلعی می‌باشد. با توجه به مدل نوردهی فانگ مولفه‌های رنگ مرکز چندضلعی  $k$  به صورت زیر تبدیل می‌شوند.

$$\begin{aligned}
 I_{R,ref,k} &= k_a \bar{R}_k I_{aR} + \sum_{i=1}^n \bar{R}_k I_{tR_i} [k_d (N_k \cdot L_i) + k_s (N_k \cdot H_{k,i})^{n_s}] \\
 I_{G,ref,k} &= k_a \bar{G}_k I_{aG} + \sum_{i=1}^n \bar{G}_k I_{tG_i} [k_d (N_k \cdot L_i) + k_s (N_k \cdot H_{k,i})^{n_s}] \\
 I_{B,ref,k} &= k_a \bar{B}_k I_{aB} + \sum_{i=1}^n \bar{B}_k I_{tB_i} [k_d (N_k \cdot L_i) + k_s (N_k \cdot H_{k,i})^{n_s}]
 \end{aligned} \tag{۴۱-۴}$$

در تصاویر سنتز شده از تورینه<sup>۱۶۶</sup>های سه‌بعدی دارای وضوح بالا، تغییرات  $I_{model}$  در هر مثلث

$k \in \{1, \dots, n_t\}$  کوچک بوده و بنابر این می‌توان خطای  $E_I$  را به صورت زیر تقریب زد.

$$E_I \approx \sum_{k=1}^{n_t} a_k \cdot \|I_{input}(\bar{p}_{x,k}, \bar{p}_{y,k}) - I_{model,k}\|^2 \tag{۴۲-۴}$$

در واقع رابطه‌ی فوق به جای معیار قرار دادن مختصات تصویر ورودی، مثلث‌های مدل سه‌بعدی و تصویر حاصل از آنها را معیار قرار می‌دهد. این امر معادل با در نظر گرفتن تأثیر موقعیت پیکسل‌ها در  $I_{input}$  می‌باشد. در رابطه‌ی فوق  $a_k$  مساحتی از تصویر است که توسط چندضلعی  $k$  تحت پوشش قرار می‌گیرد. مقدار  $a_k$  در صورت انسداد چندضلعی  $k$  برابر صفر در نظر گرفته می‌شود. حال به‌جای در نظر گرفتن تمام چندضلعی‌ها، با استفاده از زیرمجموعه‌ی تصادفی  $\mathcal{K} \subset \{1, \dots, n_t\}$ ،  $40$  چندضلعی را انتخاب کرده و تقریب  $E_I$  را به صورت زیر در نظر می‌گیریم.

$$E_{\mathcal{K}} = \sum_{k \in \mathcal{K}} \|I_{input}(\bar{p}_{x,k}, \bar{p}_{y,k}) - I_{model,k}\|^2 \tag{۴۳-۴}$$

احتمال انتخاب  $k$  به صورت  $p(k \in \mathcal{K}) = a_k$  می‌باشد. به همین  $a_k$  در رابطه‌ی (۴۳-۴) حذف شده است. در نظر گرفتن تابع هزینه‌ی رابطه‌ی (۴۰-۴) با استفاده از  $E_{\mathcal{K}}$ ، علاوه بر کاهش حجم محاسبات، بوسیله‌ی

اعمال نویز در محاسبه‌ی گرادیان باعث جلوگیری از انتخاب نقاط کمینه‌ی محلی به عنوان جواب مسئله‌ی بهینه‌سازی می‌شود.

در پیاده‌سازی الگوریتم بهینه‌سازی، قبل از تکرار اول و پس از هر ۱۰۰۰ بار تکرار، محاسبات بر روی تمام چندضلعی‌ها صورت می‌گیرد. در این حالت مقادیر  $a_k$  محاسبه شده و سطوح قابل رؤیت و محل قرارگیری سایه‌ها تعیین می‌گردند. فرض بر این است که بخش‌های انسداد یافته و محل قرارگیری سایه‌ها در تکرارهایی از الگوریتم که بر روی  $K$  صورت می‌گیرد، ثابت می‌باشند.

جهت جلوگیری از انتخاب نقاط کمینه‌ی محلی، الگوریتم بهینه‌سازی از رزولوشن پایین به سمت رزولوشن بالا حرکت می‌کند. این شیوه در بخش‌های مختلف پیاده‌سازی می‌شود.

- تکرارهای اولیه بر روی تصاویر ورودی و مدل میان‌ساز با رزولوشن پایین اعمال می‌شود.
- تکرارهای اولیه با استفاده از ضرایب ابتدایی  $\alpha$  و  $\beta$  جهت میان‌سازی بین اجزای اصلی مدل پیاده‌سازی شده و با افزایش تعداد تکرارها بر تعداد ضرایب  $\alpha$  و  $\beta$  نیز افزوده می‌شود. مقدار اولیه برای ضرایب  $\alpha$  و  $\beta$  برابر صفر می‌باشد.
- همانطور که در تشریح رابطه‌ی (۴-۴۰) اشاره شد، در تکرارهای ابتدایی، مقدار  $\sigma_F^2$  کوچک و مقدار  $\sigma_I^2$  بزرگ انتخاب می‌شود تا اثر نقاط راهنما بیشتر در نظر گرفته شود.
- در تکرارهای نهایی مدل چهره‌ی سنتز شده را تقطیع کرده و برای هر یک از بخش‌های چهره، با ثابت در نظر گرفتن مقادیر  $\rho$ ، به بهینه‌سازی پارامترهای  $\alpha$  و  $\beta$  را می‌پردازیم. باید دقت کرد که با توجه به رابطه‌ی (۴-۴۲) و در نظر گرفتن مختصات تصویر چندضلعی‌ها، نیازی به

مشخص کردن محل قرارگیری بخش‌های چهره بر روی تصویر دوبعدی ورودی وجود نداشته و

مختصات تصویر این بخش‌ها با توجه به موقعیت چندضلعی‌های مربوطه تعیین می‌شوند.

## محاسبه‌ی مشتقات E

همانطور که اشاره شد الگوریتم بهینه‌سازی مورد استفاده یک نگارش تصادفی از الگوریتم

بهینه‌سازی نیوتن-رافسن می‌باشد. این الگوریتم نیازمند محاسبه‌ی مشتقات دوم تابع هزینه نسبت به

پارامترهای بهینه‌سازی می‌باشد. در روش ارائه شده توسط بلانز و وتر [بلانز و وتر، ۲۰۰۳]، مشتقات مرتبه‌ی

اول به صورت تحلیلی محاسبه شده و مشتقات مرتبه‌ی دوم با استفاده از مشتق‌گیری عددی از مشتقات

مرتبه‌ی اول بدست می‌آیند. در این بخش روش محاسبه‌ی تحلیلی چند نمونه از مشتقات مرحله‌ی اول را

بیان می‌کنیم.

یکی از مهمترین مشتقات مورد نیاز، محاسبه‌ی  $\frac{\partial X_k}{\partial \alpha_j}$  می‌باشد که بنا بر رابطه‌ی (۴-۹) به صورت

زیر بدست می‌آید.

$$S_{\text{mod el}} = \bar{S} + \sum_{i=1}^{m-1} \alpha_i s_i \Rightarrow \frac{\partial S_{\text{mod el}}}{\partial \alpha_j} = s_j \quad (44-4)$$

$$\Rightarrow \left( \frac{\partial X_1}{\partial \alpha_j} \quad \dots \quad \frac{\partial X_k}{\partial \alpha_j} \quad \dots \quad \frac{\partial X_n}{\partial \alpha_j} \right)^T = (s_{j,1} \quad \dots \quad s_{j,k} \quad \dots \quad s_{j,n})^T \Rightarrow \frac{\partial X_k}{\partial \alpha_j} = s_{j,k}$$

در رابطه‌ی فوق  $s_{j,k}$  مؤلفه‌ی  $k$  ام بردار  $s_j$  می‌باشد. در ادامه اجزاء مشتق  $\frac{\partial E_I}{\partial \alpha_j}$  را بررسی می‌کنیم. سایر

مشتقات نیز به روشی مشابه و با استفاده از قاعده‌ی زنجیری قابل محاسبه هستند. طبق رابطه‌ی (۴-۴۲) و

با در نظر گرفتن تمام رئوس  $k$  و مؤلفه‌های رنگ  $f = r, g, b$ ، به صورت زیر بدست می‌آید.

$$\frac{\partial E_I}{\partial \alpha_j} = 2 \sum_{k=1}^{n_i} \sum_{f=r,g,b} (I_{f,\text{input}}(p_{x,k}, p_{y,k}) - I_{f,\text{mod el},k}) \cdot \left( \frac{\partial I_{f,\text{input}}(p_{x,k}, p_{y,k})}{\partial \alpha_j} - \frac{\partial I_{f,\text{mod el},k}}{\partial \alpha_j} \right) \quad (45-4)$$

در رابطه‌ی فوق،  $\frac{\partial I_{f,input}(P_{x,k}, P_{y,k})}{\partial \alpha_j}$  به صورت زیر محاسبه می‌شود.

$$\frac{\partial I_{f,input}(P_{x,k}, P_{y,k})}{\partial \alpha_j} = \frac{\partial I_{f,input}}{\partial p_{x,k}} \frac{\partial p_{x,k}}{\partial \alpha_j} + \frac{\partial I_{f,input}}{\partial p_{y,k}} \frac{\partial p_{y,k}}{\partial \alpha_j} \quad (46-4)$$

با استفاده از روش‌های عددی (عملگر سوبل) قابل محاسبه می‌باشند. نیز با در  $\frac{\partial p_{x,k}}{\partial \alpha_j}$  و  $\frac{\partial I_{f,input}}{\partial p_{y,k}}$

نظر گرفتن دید سه‌بعدی و تصویرسازی پرسپکتیو در روابط (4-28) و (4-29) به صورت زیر بدست می‌آیند.

$$\frac{\partial w_k}{\partial \alpha_j} = R_\gamma R_\theta R_\phi \frac{\partial X_k}{\partial \alpha_j} \quad (47-4)$$

$$\frac{\partial p_{x,k}}{\partial \alpha_j} = f \frac{1}{w_{z,k}} \left( \frac{\partial w_{x,k}}{\partial \alpha_j} - (p_{x,k} - P_x) \frac{\partial w_{z,k}}{\partial \alpha_j} \right) \quad (48-4)$$

باید توجه داشت که تغییر  $\alpha_j$  به علت تأثیرگذاری بر بردار متعامد بر چندضلعی، بر مشخصات مدل نوردهی نیز تأثیر می‌گذارد. این مشتقات نیز با استفاده از قاعده‌ی زنجیری و روابط مربوطه به راحتی قابل محاسبه هستند.

بهینه‌سازی به روش نیوتن-رافسن:

روش نیوتن-رافسن جهت بهینه‌سازی توابع درجه‌ی دو بکار می‌رود. یک تابع درجه‌ی ۲ به صورت

زیر تعریف می‌شود.

$$f(\theta) = a + \theta^T b + \frac{1}{2} \theta^T B \theta \quad (49-4)$$

که در آن  $a$  مقداری اسکالر،  $b$  برداری با عناصر ثابت و  $B$  یک ماتریس متقارن معین مثبت است.  $f$  تابع هدف و  $\theta$  بردار پارامترهای بهینه‌سازی است. با استفاده از برابر صفر قرار دادن مشتق تابع فوق مقدار کمینه‌ی آن به صورت زیر محاسبه می‌شود.

$$\theta^* = -B^{-1}b \quad (50-4)$$

با استفاده از بسط تیلور می‌توان نشان داد که تابع  $f(\theta)$  در ناحیه‌ای در اطراف یک نقطه‌ی  $\theta_0$  توسط تابع  $g(\theta)$  به صورت زیر تقریب زده می‌شود.

$$g(\theta) = f(\theta_0) + (\theta - \theta_0)^T \nabla f|_{\theta_0} + \frac{1}{2} (\theta - \theta_0)^T H(\theta_0) (\theta - \theta_0)^T \quad (51-4)$$

در رابطه‌ی فوق  $\nabla f|_{\theta_0} = \left( \frac{\partial f}{\partial \theta_1}, \dots, \frac{\partial f}{\partial \theta_n} \right)^T$  بردار گرادیان تابع  $f(\theta)$  در نقطه‌ی  $\theta_0$  می‌باشد. ماتریس

$H(\theta_0)$  نیز ماتریس هسین<sup>۱۶۷</sup> تابع  $f(\theta)$  در نقطه‌ی  $\theta_0$  است. در صورتی که طول بردار  $\theta$  برابر  $n$  باشد،

ماتریس هسین ماتریسی مربعی و  $n \times n$  می‌باشد که عناصر آن به صورت زیر تعریف می‌شوند.

$$h_{i,j} = \frac{\partial^2 f}{\partial \theta_i \partial \theta_j} \quad (52-4)$$

حال مقدار کمینه‌ی تابع  $f(\theta)$  را با استفاده از مقدار کمینه‌ی  $g(\theta)$  به صورت تقریبی بدست می‌آوریم. با توجه به روابط (۴۹-۴)، (۵۰-۴) و (۵۱-۴) مقدار کمینه‌ی تابع  $g(\theta)$  به صورت زیر محاسبه می‌شود.

$$\theta^* = \theta_0 - H^{-1}(\theta_0) \nabla f|_{\theta_0} \quad (53-4)$$

روش نیوتن-رافسن  $\theta^*$  را به عنوان تقریب بعدی کمینه بکار می‌برد و بنابراین رابطه‌ی تکراری زیر بدست می‌آید.

$$\theta_{i+1} = \theta_i - \lambda_i H^{-1}(\theta_i) \nabla f|_{\theta_i} \quad (54-4)$$

پارامتر  $\lambda_i$  معمولاً به صورت یک پارامتر ثابت جهت پایداری جواب در نظر گرفته می‌شود.

همگرایی روش نیوتن-رافسن وقتی که  $\theta$  نزدیک مقدار کمینه باشد، بسیار سریع است. زیرا  $g(\theta)$

در این ناحیه تقریب مناسبی برای  $f(\theta)$  می‌باشد. عیب اصلی این روش این است که در هر مرحله‌ی تکرار

<sup>167</sup> Hessian

در صورت زیاد بودن تعداد پارامترها، محاسبه و معکوس کردن ماتریس هسین مستلزم محاسباتی سنگین می‌باشد. علاوه بر این در صورتی که  $\theta_i$  نزدیک نقطه‌ی کمینه نباشد، تابع هسین ممکن است معین منفی شود که منجر به واگرایی الگوریتم نیوتن-رافسن می‌شود.

جهت پیاده‌سازی این روش بر روی مسئله‌ی بهینه‌سازی تابع هزینه‌ی مدل میان‌ساز (رابطه‌ی ۴-)

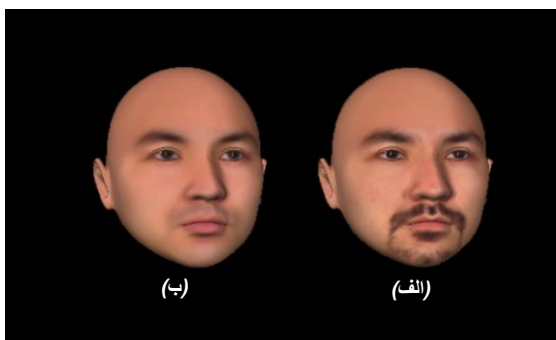
(۴۰)، ابتدا بردار  $\theta$  را به صورت  $\theta = (\alpha^T, \beta^T, \rho^T)^T$  تعریف می‌کنیم. با فرض ۹۹ مولفه برای هر یک از بردارهای  $\alpha$  و  $\beta$ ، بردار  $\theta$  یک بردار  $220 \times 1$  می‌باشد. البته باید توجه کرد که در تکرارهای اول طول این بردار به مراتب کمتر می‌باشد. در این مرحله با استفاده از مطالب مطرح شده در زیربخش "محاسبه‌ی مشتقات E" گرادینان تابع E را محاسبه می‌کنیم.

$$\nabla E = \left( \frac{\partial E}{\partial \alpha_1}, \frac{\partial E}{\partial \alpha_2}, \dots, \frac{\partial E}{\partial \alpha_{99}}, \frac{\partial E}{\partial \beta_1}, \frac{\partial E}{\partial \beta_2}, \dots, \frac{\partial E}{\partial \beta_{99}}, \frac{\partial E}{\partial \rho_1}, \frac{\partial E}{\partial \rho_2}, \dots, \frac{\partial E}{\partial \rho_{22}} \right)^T \quad (۴-۵۵)$$

گرادینان تابع E یک بردار  $220 \times 1$  می‌باشد. حال با استفاده از مشتقات مرتبه‌ی اول و دوم ماتریس هسین را طبق رابطه‌ی (۴-۵۲) محاسبه می‌کنیم. باید توجه داشت که مشتقات مرتبه‌ی اول به صورت تحلیل در زیربخش قبل محاسبه شده‌اند و مشتقات مرتبه‌ی دوم را با روش‌های عددی و به صورت مشتق مشتقات مرتبه‌ی اول محاسبه می‌کنیم. با توجه به این که ماتریس هسین جهت تخمین پارامترهای بهینه‌سازی در تکرار بعدی بکار می‌رود (رابطه‌ی (۴-۵۴)) و با توجه به استفاده از تقریب تابع E به علت بکارگیری رابطه‌ی (۴-۴۳)، نیازی به دقت بالا در محاسبه‌ی ماتریس هسین وجود نداشته و تنها محاسبه‌ی مقادیر واقع بر روی قطر اصلی این ماتریس کفایست [بلانز و وتر، ۲۰۰۳]. در این حالت معکوس ماتریس هسین نیز به سادگی و به صورت  $H^{-1} = \text{diag}(1/h_{ii})$  محاسبه می‌شود.

استخراج بافت از تصاویر ورودی:

آخرین مرحله در سنتز چهره با استفاده از مدل میان‌ساز استخراج بافت از تصویر ورودی می‌باشد. از آنجایی که جزئیات بافت چهره توسط مدل میان‌ساز با دقت بالا سنتز نمی‌شود، می‌توان طبق روش تشریح شده در بخش ۳-۳-۲ به استخراج بافت چهره از تصاویر ورودی و اعمال آن بر مدل پرداخت. با توجه به مشخص بودن زاویه‌ی دید تصویر دوبعدی و بردارهای متعامد بر رویه این فرایند دقیقاً مشابه مراحل ارائه شده در ۳-۳-۲ می‌باشد. شکل ۴-۱۶ اثر استخراج بافت را بر مدل چهره‌ی سه‌بعدی سنتز شده نشان می‌دهد. مشاهده می‌شود که موهای صورت توسط مدل میان‌ساز سنتز نشده است. باید توجه کرد که فرایند استخراج بافت هیچ تأثیری در پارامترهای  $\alpha$  و  $\beta$  نمی‌گذارد و فقط در مرحله‌ی نمایش با بافت حاصل از ضرایب  $\beta$  ترکیب می‌شود.



شکل ۴-۱۶ - تأثیر استخراج بافت بر سنتز چهره، (الف): با استخراج بافت، (ب): بدون استخراج بافت

#### ۳-۲-۱۹-۱ تطبیق مدل میان‌ساز و مدل‌های سه‌بعدی

روش ارائه شده در بخش قبلی جهت تطبیق مدل میان‌ساز با تصویر ورودی، به راحتی قابل استفاده در تطبیق مدل میان‌ساز و مدل‌های سه‌بعدی جدید می‌باشد. در تطبیق مدل میان‌ساز و تصاویر ورودی، جهت مقایسه‌ی بین مدل و تصویر از تصویرسازی پرسپکتیو استفاده شد تا مدل میان‌ساز از فضای سه‌بعدی به فضای دوبعدی نگاشته شود. این فرایند مستلزم استفاده از پارامترهای مربوط به مختصات دید و تصویرسازی پرسپکتیو مدل نوردهی فانگ می‌باشد. در واقع تمام ۲۲ پارامتر بردار  $\rho$  در جریان این فرایند بکار



گرفته می‌شوند. اما در تطبیق مدل میان‌ساز و مدل سه‌بعدی به علت این که هر دو مدل در فضای سه‌بعدی می‌باشند نیازی به نگاشت بین فضاها وجود ندارد. در واقع همان طور که قبلاً در بخش ۳-۴-۲ اشاره شد، مدل سه‌بعدی در دستگاه مختصات استوانه‌ای به صورت زیر نشان داده می‌شود.

$$I(h, \phi) = (r(h, \phi), R(h, \phi), G(h, \phi), B(h, \phi))^T ; \quad h, \phi \in \{0, \dots, 511\} \quad (56-4)$$

باید توجه کرد که تابع هزینه در این بخش شامل  $E_F$  نمی‌شود. در واقع نیازی به استفاده از نقاط راهنما به علت برابری تقریبی مقیاس مدل‌ها و عدم وابستگی به زاویه‌ی دید، وجود ندارد. علاوه بر این به علت مقایسه بین رویه‌ها نیازی به بکارگیری روش تخمین گر احتمالاً پسین بیشینه نیز وجود نداشته و بهینه‌سازی تابع خطا دچار نقاط کمینه‌ی محلی نمی‌شود. بنابراین کفایت تابع هزینه را به صورت زیر تعریف کنیم.

$$E = \sum_{h, \phi} \|I_{input}(h, \phi) - I_{model}(h, \phi)\|^2 \quad (57-4)$$

بهینه‌سازی تابع هدف مذکور به علت کمتر بودن تعداد پارامترها و ساده بودن تابع هدف، به مراتب سریع‌تر از بهینه‌سازی تابع (۴-۴۰) صورت می‌گیرد. فرایند بهینه‌سازی مشابه روند اشاره شده در ۴-۲-۲-۲ می‌باشد.

### ۱-۱۹-۳ ساخت مدل میان‌ساز

جهت ساخت مدل میان‌ساز می‌بایست تناظر بین مدل‌های چهره‌ی نمونه بر قرار شود تا با استفاده از روابط (۲-۴) و یا (۹-۴) مدل میان‌ساز ایجاد شود. محاسبه‌ی تناظر با استفاده از دو روش الگوریتم شارش نوری و روش مبتنی بر همبستگی محلی قابل انجام است. در این بخش ساخت مدل میان‌ساز با استفاده از هر دو روش تشریح می‌شود.

ساخت مدل میان‌ساز با استفاده از الگوریتم شارش نوری

جهت ساخت مدل میان‌ساز با استفاده از الگوریتم شارش نوری به ترتیب زیر عمل می‌کنیم. در ابتدا یک مدل چهره به عنوان مدل مرجع در نظر گرفته شده و تناظر بین سایر مدل‌های چهره و مدل مرجع با استفاده از الگوریتم شارش نوری محاسبه می‌شود. حال با استفاده از تجزیه به روش PCA، اجزای اصلی مدل‌های تناظر یافته استخراج شده و مدل میان‌ساز ناشی از این مدل‌ها، طبق رابطه‌ی (۴-۹) ساخته می‌شود. تعداد اجزای اصلی در مرحله‌ی اول تعداد کمی در نظر گرفته می‌شود. با اعمال روابط (۴-۵) مدل میانگین، بدست آمده و به عنوان مدل مرجع در نظر گرفته می‌شود. در این مرحله مدل میان‌ساز ساخته شده به هر یک از مدل‌های چهره طبق مطالب بخش ۴-۲-۳-۳ تطبیق شده و تقریبی برای مدل‌های چهره بوجود می‌آورد. حال مجدداً تناظر میان هر یک از مدل‌های چهره و مدل‌های تقریبی مربوطه ناشی از تطبیق مدل میان‌ساز، محاسبه می‌شود. ترکیب این تناظر و تناظر بدست آمده در مرحله‌ی قبلی، تناظری بهبود یافته بین هر یک از مدل‌های چهره و مدل مرجع ایجاد می‌کند. این عملیات مجدداً با افزایش تعداد اجزای اصلی تکرار می‌شود. پس از چند بار تکرار، تناظر و مدل میان‌ساز قابل قبولی ایجاد می‌شود.

در روش مذکور توجه به این نکته حائز اهمیت است که در مرحله‌ی محاسبه‌ی تناظر میان هر یک از مدل‌های چهره با مدل‌های سنتز شده ناشی از تطبیق مدل میان‌ساز، به علت تقریبی بودن تناظر در مرحله‌ی اول، مدل‌های سنتز شده، مدل‌هایی با کیفیت پایین و تقریبی می‌باشند. این مدل میان‌ساز توانایی تولید چهره‌های واقعی را نداشته و چهره‌هایی درهم‌ریخته و تار ناشی از عدم تناظر دقیق تولید می‌کند. پس این مدل میان‌ساز امکان تولید هیچ مدلی، حتی مدل‌های چهره‌ی موجود در پایگاه داده را ندارد. بنابراین محاسبه‌ی تناظر بین مدل‌های چهره و این مدل‌های تقریبی همواره بردارهای سرعت غیر صفر تولید می‌کند که منجر به بهبود تناظر می‌شوند.

باید توجه کرد که این الگوریتم در هر مرحله به ۲ بار میان‌سازی با بکارگیری PCA، 2m بار بکارگیری الگوریتم شارش نوری و m بار تطبیق بین مدل‌ها که فرآیندی با حجم محاسباتی بالا است، نیاز دارد. m تعداد مدل‌های چهره ۳ بعدی موجود در پایگاه داده می‌باشد.

ساخت مدل میان‌ساز با استفاده از تطبیق مبتنی بر همبستگی محلی:

جهت ساخت مدل میان‌ساز با استفاده از این روش، کافی است به ازای هر چهره ۳ بعدی جدید، تنها ۱ بار تناظر بین این چهره و چهره مرجع با استفاده از تطبیق مبتنی بر همبستگی محلی تشریح شده در بخش ۴-۱-۳ محاسبه شود. سپس مجدداً تناظر بین مدل‌های نمونه و مدل متوسط ایجاد شده را محاسبه می‌کنیم. در نتیجه تنها با 2m بار بکارگیری این روش مدل میان‌ساز ساخته می‌شود. مدل‌های میانگین نمایش داده در شکل ۴-۲ نتیجه‌ی میان‌سازی مرحله‌ی اول این روش می‌باشند.

## ۲۰-۱ مقدمه

همانطور که در قسمت‌های مختلف این گزارش اشاره شد، اغلب بخش‌های الگوریتم مدل میان‌ساز بر اساس روش ارائه شده در [وتر و بلانز، ۱۹۹۸ - بلانز و وتر، ۱۹۹۹ - بلانز و وتر، ۲۰۰۳] پیاده‌سازی شده است تا بستر مناسبی جهت پژوهش‌های بعدی در این زمینه فراهم آید. نتایج پیاده‌سازی این الگوریتم در

فصول مختلف این گزارش ارائه شده است. در این فصل نیز تعداد بیشتری از نتایج در بخش‌های سنتز و تشخیص چهره مورد بررسی قرار می‌گیرند. از آنجایی که کامل‌ترین الگوریتم سنتز و تشخیص چهره‌ی مبتنی بر مدل میان‌ساز، روش ارائه شده توسط بلانز می‌باشد، پیاده‌سازی مدل میان‌ساز بر اساس این الگوریتم مد نظر قرار گرفت. به علت وقوع برخی مشکلات، روش پیاده‌سازی شده دارای تفاوت‌های با روش ارائه شده در [بلانز و وتر، ۲۰۰۳] می‌باشد، که به آنها اشاره می‌کنیم.

- به علت تفاوت موجود بین پایگاه داده‌ی مورد استفاده و پایگاه داده‌ی MPI، استفاده از روش شارش نوری جهت یافتن تناظر بین نقاط، نتایج مناسبی ایجاد نکرده و الزاماً انطباق مبتنی بر همبستگی محلی مورد استفاده قرار گرفت (بخش ۴-۱-۳-۳). این روش نتایج نسبتاً مناسبی در تعیین تناظر نقطه به نقطه بین مدل‌های چهره ایجاد کرده است. یکنواخت و نرم نبودن نتایج، از مشکلات مربوط به این روش انطباق است که با اعمال فیلترهای پایین‌گذر، این مشکل تا حد قابل توجهی رفع شده است.

- بنابر دلیل تاکنون نامشخص، الگوریتم پیاده‌سازی شده در تشخیص موقعیت منبع نور دچار اشتباه شده و نتایج ضعیفی در مورد بهینه‌سازی پرامترهای مربوط به منبع نور ایجاد می‌کند. علت این امر احتمالاً مربوط به اشتباه در محاسبه‌ی مشتقات تابع هزینه نسبت به پارامترهای  $\phi_i$  و  $\theta_i$  و یا اشتباه در پیاده‌سازی آن‌ها می‌باشد. به همین دلیل در اجرای این الگوریتم، نیازمند تعیین موقعیت منبع نور توسط کاربر می‌باشیم. علاوه بر این سایه‌ها نیز در پیاده‌سازی این الگوریتم لحاظ نشده‌اند که در بعضی از موارد مشکلات بوجود می‌آورد.

- از دیگر تفاوت‌های موجود بین الگوریتم پیاده‌سازی شده و الگوریتم اصلی، نحوه‌ی تعیین مقادیر اولیه می‌باشد. در الگوریتم پیاده‌سازی شده کاربر با استفاده از یک مدل چهره‌ی نمونه زاویه‌ی دید و مقیاس مدل را تعیین می‌کند. این فرایند در ابتدای بخش ۴-۲ تشریح شده است.

زمان لازم جهت انجام محاسبات مربوط به سنتز یک تصویر ورودی در مرجع [بلانز و وتر، ۲۰۰۳] با استفاده از یک کامپیوتر شخصی مجهز به پردازنده‌ی پنتیوم ۴ با سرعت 2 GHz، در حدود ۴,۵ دقیقه گزارش شده است. اما زمان لازم جهت اجرای الگوریتم پیاده‌سازی شده در این پایان‌نامه، توسط یک کامپیوتر شخصی مجهز به پردازنده‌ی پنتیوم ۴ با سرعت 3.6 GHz، در حدود ۴۰ دقیقه می‌باشد. باید توجه کرد که دو پارامتر تأثیرگذار نیز از فرایند بهینه‌سازی حذف شده‌اند. این تفاوت فاحش در زمان اجرای این الگوریتم ناشی از پیاده‌سازی آن توسط نرم‌افزار MATLAB<sup>®</sup> می‌باشد. زبان برنامه‌نویسی MATLAB<sup>®</sup> علی‌رغم ساده‌سازی فرایند برنامه‌نویسی، زبانی بسیار کند بوده و جهت پیاده‌سازی نهایی الگوریتم‌های دارای حجم محاسباتی بالا پیشنهاد نمی‌شود. در مقایسه‌ی صورت گرفته بین نرم‌افزار MATLAB<sup>®</sup> و نرم‌افزار OpenCV<sup>®</sup> 168، سرعت پردازش توسط OpenCV<sup>®</sup> در حدود ۸ برابر سرعت نرم‌افزار MATLAB<sup>®</sup> تخمین زده شد. انجام عملیات جمع بین دو ماتریس دوبعدی به عنوان معیار مقایسه‌ی این دو نرم‌افزار در نظر گرفته شد.

در این فصل نتایج پیاده‌سازی الگوریتم مدل میان‌ساز را در دو بخش سنتز و تشخیص چهره بررسی می‌کنیم.

---

<sup>168</sup> OpenCV : Intel Open Source Computer Vision Library

OpenCV کتابخانه‌ای شامل توابع مورد نیاز جهت پیاده‌سازی الگوریتم‌های بینایی ماشین می‌باشد که به صورت سورس باز تحت زبان برنامه‌نویسی C و توسط شرکت اینتل ارائه شده است.

## ۱-۲۱ نتایج سنتز چهره

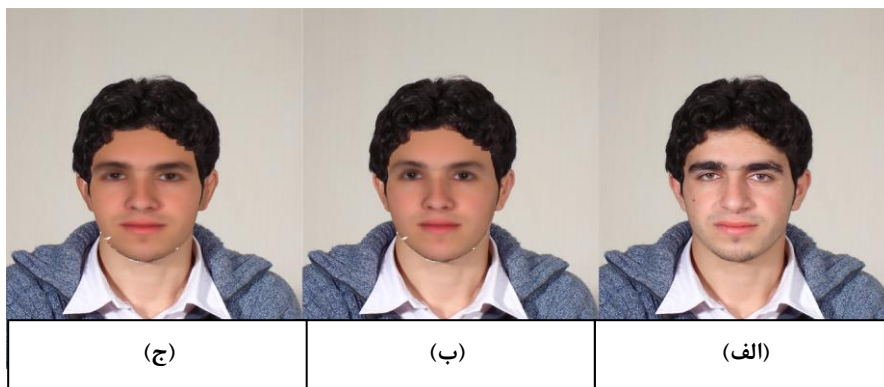
همانطور که در فصل ۲ اشاره شد، در این پایان نامه به ساخت یک پایگاه داده‌ی چهره‌ی سه‌بعدی بر اساس چهره‌های ایرانی اقدام کرده‌ایم. انگیزه‌ی ما جهت ساخت چنین پایگاه داده‌ای، وابستگی مدل میان‌ساز به پایگاه داده می‌باشد. از آنجایی که ترکیب خطی  $m$  بردار در فضای  $3n$  بعدی ( $n > m$ )، تنها می‌تواند یک زیرفضا از این فضای برداری را پوشش نماید، ساختار پایگاه داده و نژادهای موجود در آن در کیفیت سنتز چهره تاثیر قابل توجهی دارد. با توجه به تفاوت قابل توجه بین چهره‌های ایرانی - که از مردمان بخش‌های شرقی نژاد هندواروپایی می‌باشند- و چهره‌های بخش‌های غربی‌تر این نژاد، استفاده از پایگاه‌های داده‌ی سه‌بعدی چهره‌ی مرسوم همچون MPI جهت سنتز چهره‌های ایرانی، در بعضی از موارد نتایج نامناسبی ایجاد می‌کند. علاوه بر این در روش مدل‌سازی میان‌ساز به علت حجم بالای محاسبات، استفاده از پایگاه‌های داده بزرگ امکان‌پذیر نیست. بنابراین پایگاه‌های داده‌ی چهره‌ی مورد استفاده در این روش، از دربرگیرندگی نژادی کافی برخوردار نیستند. به همین دلیل جهت سنتز باکیفیت چهره‌های یک نژاد خاص بهترین راه استفاده از پایگاه داده‌ی چهره‌ی سه‌بعدی مربوط به همان نژاد است. مشخصات و ساختار پایگاه داده‌ی چهره‌ی سه‌بعدی ارائه شده، در بخش ۳-۴-۲ مورد بررسی قرار گرفت. در این بخش به بررسی و مقایسه‌ی مدل‌های سه‌بعدی چهره‌ی سنتز شده با استفاده از روش مدل سه‌بعدی میان‌ساز می‌پردازیم. این مدل‌ها با استفاده از مدل میان‌ساز مبتنی بر پایگاه داده‌ی ارائه شده در این مقاله و پایگاه داده‌ی MPI تولید شده‌اند. تصاویر ورودی جهت مقایسه دو مدل میان‌ساز، از پایگاه داده‌ی چهره‌ی دانشگاه صنعتی امیرکبیر استخراج شده‌اند. این دو چهره به علت داشتن ویژگی‌های مردمان شرقی نژاد هندواروپایی همچون ابروهای پرپشت، توسط مدل میان‌ساز مبتنی بر پایگاه داده‌ی MPI به خوبی سنتز نشده‌اند (ستون سوم شکل ۵-۱) و تفاوت

قابل توجهی بین تصاویر سنتز شده و تصاویر ورودی وجود دارد. در مقابل مدل میان‌ساز مبتنی بر پایگاه داده‌ی ارائه شده، با دقت مناسبی این نوع چهره‌ها را سنتز می‌کند (ستون دوم شکل ۵-۱). همان‌گونه که در شکل ۵-۱ مشخص است رنگ چهره و حالت ابروها در تصاویر ستون سوم شباهت کمتری به تصاویر ورودی (ستون اول) دارند. این نقص در تصاویر سنتز شده توسط پایگاه داده‌ی ارائه شده (ستون دوم)، به‌خوبی رفع شده است. شباهت بیشتر تصاویر ستون دوم به تصاویر ورودی، بیانگر تأثیر پایگاه داده بر کیفیت سنتز چهره با استفاده از مدل میان‌ساز می‌باشد.



شکل ۵-۱ - مقایسه نتایج سنتز چهره، از چپ به راست: (ستون اول): تصاویر ورودی، (ستون دوم): چهره‌های سنتز شده مبتنی بر پایگاه داده ارائه شده، (ستون سوم): چهره‌های سنتز شده مبتنی بر پایگاه داده MPI

باید توجه کرد که استخراج بافت در مدل‌های سنتز شده‌ی شکل ۵-۱ با ضریب کوچکی اعمال شده است. شکل ۵-۲ یک نمونه‌ی دیگر از تصاویر سنتز شده توسط دو پایگاه داده را نمایش می‌دهد. در این شکل استخراج بافت از تصاویر ورودی اعمال نشده است. به علت عدم وابستگی پارامترهای  $\alpha$  و  $\beta$  به استخراج بافت از تصاویر ورودی، بهبود هرچند اندک در سنتز مدل بدون استخراج بافت، منجر به بهبود قابل توجه در  $\alpha$  و  $\beta$  می‌شود.



شکل ۵-۲ - مقایسه نتایج سنتز چهره بدون استخراج بافت، (الف): تصویر ورودی، (ب): چهره‌ی سنتز شده مبتنی بر پایگاه داده MPI، (ج): چهره‌ی سنتز شده مبتنی بر پایگاه داده ارائه شده

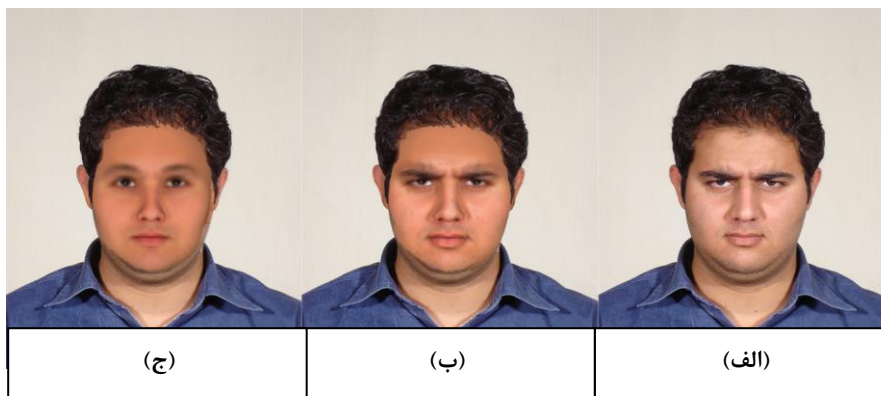
شکل ۵-۳ دو نمونه سنتز یک مدل چهره‌ی مؤنث موجود در پایگاه داده‌ی دانشگاه صنعتی امیرکبیر را در زوایای متفاوت نشان می‌دهد. همان طور که در بخش ۳-۴-۲ اشاره شد، مدل‌های چهره‌ی مؤنث موجود در پایگاه داده‌ی ارائه شده، با مدل‌های چهره‌ی مؤنث موجود در پایگاه داده‌ی MPI یکسان هستند.



شکل ۵-۳ - سنتز یک چهره‌ی مؤنث در دو زاویه‌ی دید متفاوت، (سمت چپ): تصاویر ورودی، (سمت راست): تصاویر سنتز شده

در شکل ۵-۴ نیز پایین بودن قابلیت سنتز خودکار حالات چهره نمایش داده شده است. همانطور که مشاهده می‌شود، حالت چهره تنها توسط استخراج بافت بر روی مدل پیاده‌سازی شده است. علت این امر عدم وجود مدل‌های دارای حالت چهره در پایگاه داده‌ی سه‌بعدی می‌باشد.





شکل ۴-۵ - سنتز چهره‌های دارای حالت، (الف): تصویر ورودی، (ب): چهره‌ی سنتز شده با استخراج بافت از ورودی، (ج): چهره‌ی سنتز شده بدون استخراج بافت

## ۲۲-۱ نتایج تشخیص چهره

جهت تشخیص چهره به عنوان داده‌های آموزشی و آزمایش از پایگاه داده‌ی چهره‌ی دوبعدی دانشگاه صنعتی امیرکبیر استفاده شده است. به علت حجم بالای محاسباتی این الگوریتم و نیاز به تنظیمات دستی نسبتاً زیاد، استفاده از این الگوریتم در کاربرد تشخیص چهره در حال حاضر مشکل می‌باشد. به همین دلیل از پایگاه‌های داده‌ی بزرگتر جهت آموزش و آزمایش استفاده نشده است.

پایگاه داده‌ی چهره‌ی دانشگاه صنعتی امیرکبیر دارای تصاویر دوبعدی از ۴۰ چهره‌ی مختلف می‌باشد. از این میان تنها ۶ چهره، مؤنث می‌باشند. محدوده‌ی سنی اغلب افراد، بین ۲۰ تا ۳۰ سال به نظر می‌رسد. در این بین، سن سه شخص - دو مذکر و یک مؤنث- بالاتر از محدوده‌ی سنی غالب اشخاص موجود در پایگاه داده به نظر می‌رسد. از هر فرد ۱۲ تصویر ارائه شده است. ۵ تصویر مربوط به تغییرات زاویه‌ی دید، ۱ تصویر با عینک، ۴ تصویر مربوط به حالات مختلف چهره و ۲ تصویر نیز مربوط به زوایای تابش نور می‌باشند. مجموعه‌ی این تصاویر برای یک شخص در شکل ۵-۵ نشان داده شده است.



شکل ۵-۱۲ تصویر موجود برای فرد شماره ۱ در پایگاه داده‌ی چهره‌ی دانشگاه صنعتی امیرکبیر

جهت تشخیص چهره از تصاویر تمام‌رخ، تابیده شده از روبرو و بدون حالت چهره، جهت آموزش

سیستم استفاده کردیم. بر اساس این ۴۰ تصویر، ۴۰ مدل چهره تولید شد که ضرایب  $\alpha$  و  $\beta$  مربوط به هر

یک از مدل‌ها به عنوان ویژگی نمونه‌ها استفاده شد. جهت تست، ۲۰ تصویر به صورت تصادفی انتخاب شدند.

این تصاویر شامل تصاویر تابیده شده از زوایای مختلف (۴ تصویر)، حالات مختلف چهره (۹ تصویر) و زوایای

دید متفاوت (۷ تصویر) می‌شوند. به علت تنظیم موقعیت منبع نور توسط کاربر، نتایج تشخیص چهره در

مورد تصاویر تابیده شده از زوایای مختلف، معتبر نمی‌باشد. تشخیص چهره در تمامی موارد به درستی صورت

گرفت و نرخ تشخیص ۱۰۰٪ حاصل شد.

از علل بالا بودن نرخ تشخیص می‌توان به موارد زیر اشاره کرد.

- پایین بودن تعداد نمونه‌های موجود در کتابخانه به علت زمان زیاد مورد نیاز جهت سنتز توسط

مدل میان‌ساز

- تفاوت زیاد بین نمونه‌ها که ناشی از تعداد کم آنها می‌باشد.

- عدم وجود اختلاف قابل توجه در زوایای دید تصاویر آموزشی و آزمایشی

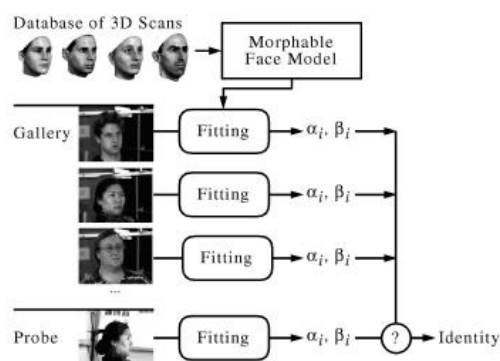
- کاهش تأثیر شرایط متفاوت نوردهی به علت تعیین موقعیت منبع نور توسط کاربر. تخمین اشتباه

موقعیت منبع منجر به تغییر پارامتر  $\beta$  و در نتیجه کاهش نرخ تشخیص می‌شود.

البته باید توجه کرد که در [بلانز و وتر، ۲۰۰۳] تشخیص بر روی نمونه‌های موجود در دو پایگاه داده‌ی بزرگ FERET و CMU-PIE صورت گرفته و نرخ تشخیصی در حدود ۹۵٪ گزارش شده است. با توجه به گستردگی داده‌های آموزشی و آزمایشی و تنوع تصاویر چهره از لحاظ زاویه‌ی دید و شرایط مختلف نورپردازی، این نرخ تشخیص نشان‌دهنده‌ی توان بالای این الگوریتم در تشخیص چهره می‌باشد.

### ۱-۲۳ روش تشخیص چهره

تشخیص چهره بر اساس روش ارائه شده در [بلانز و وتر، ۲۰۰۳] صورت گرفته که در این قسمت به معرفی آن می‌پردازیم. شکل ۵-۷ شمای کلی تشخیص چهره بوسیله‌ی مدل میان‌ساز را نشان می‌دهد.



شکل ۵-۶ - مراحل تشخیص چهره توسط مدل میان‌ساز [بلانز و وتر، ۲۰۰۳]

ویژگی‌های مورد استفاده جهت تشخیص چهره بردارهای ۹۹ عضوی  $\alpha$  و  $\beta$  می‌باشند. برای هر چهره ۵ زوج بردار  $\alpha$  و  $\beta$  در نظر گرفته می‌شود. یکی از این ۵ زوج بردار، بردارهای شکل و بافت کلی مدل چهره بوده و ۴ زوج دیگر مربوط به هر یک از ۴ بخش مدل میان‌ساز تقطیع شده می‌باشند (به بخش ۴-۱-۱ مراجعه شود). هر یک از ضرایب  $\alpha_i$  و  $\beta_i$  توسط انحراف استاندارد مربوطه تغییر مقیاس یافته و به صورت  $\frac{\alpha_i}{\sigma_{S,i}}$  و  $\frac{\beta_i}{\sigma_{T,i}}$  در نظر گرفته می‌شوند. از آنجایی که  $\sigma_{S,i}$  و  $\sigma_{T,i}$  در واقع مقادیر ویژه در روش PCA می‌باشند، بزرگ بودن آنها بیانگر اصلی‌تر بودن جزء مربوطه و در نتیجه نزدیکی آن به میانگین داده‌ها

می‌باشد (به ۱-۴ مراجعه شود). بنابر این با تقسیم ضرایب  $\alpha_i$  و  $\beta_i$  به این مقادیر، از اهمیت اجزاء نزدیک به میانگین و تأثیر آنها بر فرایند تشخیص کاسته و به اهمیت اجزای دور از میانگین و تأثیر آنها بر فرایند تشخیص افزوده می‌شود.

با ترکیب ۵ زوج بردار مذکور به ازای هر مدل چهره یک بردار ویژگی  $c$  شامل  $2 \times 5 \times 99 = 990$  مولفه به صورت  $\frac{\beta_i}{\sigma_{T,i}}$  و  $\frac{\alpha_i}{\sigma_{S,i}}$  خواهیم داشت که در تشخیص از آن استفاده می‌شود. معیار فاصله بین ویژگی‌ها، فاصله‌ی ساده‌ی اقلیدسی و یا کسینوس زاویه‌ی بین دو بردار ویژگی می‌باشد. این دو فاصله به صورت زیر تعریف می‌شوند.

$$d_{Euclidian} = \|c_1 - c_2\|^2 \quad d_A = \frac{\langle c_1, c_2 \rangle}{\|c_1\| \|c_2\|} \quad (1-5)$$

به علت دستیابی به نرخ تشخیص بالاتر توسط فاصله‌ی  $d_A$  در [بلانز و وتر، ۲۰۰۳]، ما نیز از این معیار فاصله جهت تشخیص استفاده کرده‌ایم.

## ۱-۲۴ نتیجه گیری

در جریان انجام این پایان نامه مباحثی چون بررسی تاریخی روش های تشخیص چهره ی سه بعدی و روش های مدل سازی چهره، مباحث مرتبط با ارتباط تصویری سه بعدی، روش ساخت مدل چهره با استفاده از شیوه ی هندسه یابی تصویری بررسی شدند. علاوه بر این دو پایگاه داده ی چهره ی سه بعدی با توجه به روش هندسه یابی تصویری ایجاد شده و مورد استفاده قرار گرفت. همچنین بررسی مدل میان ساز و سنتز و تشخیص چهره بر اساس آن، بخش اصلی مباحث ارائه شده در این پایان نامه را به خود اختصاص داد. پیاده سازی کامل این الگوریتم نسبتاً گسترده، به عنوان یکی از اهداف اصلی در انجام این پژوهش مد نظر قرار گرفت تا بدین وسیله بستر مناسبی جهت انجام پژوهش های مرتبط با این موضوع فراهم شود. نتایج پیاده سازی سنتز و تشخیص چهره با استفاده از مدل میان ساز، بیانگر کارایی این مدل در زمینه ی سنتز و تشخیص چهره می باشد. علاوه بر این بهره مندی این الگوریتم از مزایای تشخیص چهره ی سه بعدی و دوبعدی به طور همزمان، کاربرد این روش را بیش از سایر روش های تشخیص چهره ی سه بعدی، ممکن می سازد. توانایی در ساخت و سنتز مدل های سه بعدی چهره نیز از مزایای قابل توجه این روش جهت بکارگیری در کاربردهایی چون پویانمایی کامپیوتری می باشد.

## ۱-۲۵ پیشنهادات

با وجود نتایج قابل قبول در تشخیص و سنتز چهره با استفاده از مدل میان ساز، پیشنهادات زیر می توانند به عنوان زمینه ی پژوهش های آتی و در جهت بهبود پژوهش فعلی مورد توجه قرار گیرند.

بررسی و پیاده سازی روش خودکار تشخیص و سنتز حالات چهره، از موارد بهبود دهنده ی عملکرد مدل میان ساز می باشد. برای این منظور می توان مدل های نمونه ی دارای حالات چهره ی مختلف را به پایگاه

داده موجود اضافه کرد. یافتن روشی جهت جلوگیری از افزایش بیش از اندازه‌ی حجم پایگاه داده، چالش اساسی مرتبط با این روش می‌باشد.

صرف زمان زیاد جهت سنتز و تشخیص چهره در الگوریتم پیاده‌سازی شده‌ی فعلی، مانع اعمال روش‌های جدید و بررسی نتایج آن‌ها می‌شود. برای رفع این مشکل می‌بایست به بهینه‌سازی الگوریتم و پیاده‌سازی آن در نرم‌افزاری سریع‌تر همچون OpenCV پرداخت. البته به علت پیچیدگی پیاده‌سازی الگوریتم تحت زبان‌های برنامه‌نویسی سریع‌تر، انجام این فرایند مستلزم حصول اطمینان از رفع نواقص فعلی الگوریتم پیاده‌سازی شده می‌باشد.

جهت بهینه‌سازی روش‌های تشخیص، می‌توان به ازای هر فرد از چند تصویر و در نتیجه چند مدل چهره در کتابخانه استفاده کرد. ای تصاویر و مدل‌های چهره تحت شرایط مختلف نورپردازی و حالات مختلف چهره و در زوایای دید متفاوت تهیه می‌شوند. علت پیشنهاد استفاده از چنین کتابخانه‌ای، وابستگی ضرایب مدل‌های سنتز شده‌ی چهره به تصاویر ورودی می‌باشد. این وابستگی ناشی از نقص مدل نوردهی، خطاهای مربوط به الگوریتم بهینه‌سازی، وجود انواع مختلف نویز در تصاویر ورودی و نفس تقریبی بودن سنتز چهره بر اساس مدل میان‌ساز می‌باشد. بنابراین با تهیه‌ی مدل‌های مختلف از هر شخص می‌توان به تخمین تابع توزیع ضرایب مدل برای هر شخص اقدام کرد. وجود تابع توزیع بردار ویژگی‌ها و توابع درست‌نمایی این بردار، امکان بکارگیری روش‌های مؤثرتر طبقه‌بندی و تشخیص را فراهم می‌کند.

تکمیل مدل میان‌ساز جهت در نظر گرفتن بخش‌های بیشتری از مدل سر همچون شکل و حالت مو منجر به افزایش کاربرد این مدل و سادگی استفاده از آن می‌شود.

در مدل میان‌ساز امکان استفاده از چند تصویر ورودی در زوایای مختلف جهت سنتز بهتر چهره وجود دارد. ایجاد ضرایبی برای تعیین میزان تأثیرپذیری مدل از تصاویر مختلف ورودی همچون ضرایب ناشی از زاویه‌ی بین بردارهای متعامد رویه‌ی چهره و زاویه‌ی دید تصاویر ورودی، و بررسی تأثیر این ضرایب بر عملکرد این روش، از جمله موارد قابل بررسی می‌باشد.

## اصول زبان VRML

VRML زبانی است که در آن اشیائی تعریف می‌شوند که قابلیت خواندن و ایجاد خود را دارا هستند. این اشیاء می‌توانند شامل اطلاعات مختلفی همچون اطلاعات هندسی، بافت، تصویر و صوت باشند. این اشیاء در زبان VRML گره نامیده می‌شوند. گره‌ها در یک ساختار سلسله‌مراتبی تحت عنوان گراف صحنه<sup>۱۶۹</sup> قرار می‌گیرند. این گراف علاوه بر نشان دادن گره‌های عضو، ترتیب آنها را نیز مشخص می‌کند. به گونه‌ای که گره‌های بعدی تحت تأثیر گره‌های ماقبل خود واقع می‌شوند. جهت محدود کردن تأثیر گره‌ها از گره‌ی خاصی به نام گره‌ی جداساز<sup>۱۷۰</sup> استفاده می‌شود. این گره به بخش‌هایی از گراف صحنه اجازه می‌دهد تا عملکردی مستقل از سایر بخش‌ها داشته باشد.

هر گره می‌تواند شامل بخش‌های زیر باشد:

- نوع شیء: یک گره می‌تواند شامل یک شیء همچون کره، مکعب، تبدیل هندسی، مجموعه‌ای از چندضلعی‌ها باشد.
- پارامترهای گره: هر گره می‌تواند پارامترهایی جهت تشریح مشخصات گره دارا باشد. این پارامترها فیلد<sup>۱۷۱</sup> نامیده می‌شوند.

---

<sup>169</sup> Scene Graph

<sup>170</sup> Separator

<sup>171</sup> Field

- نام گره: توانایی نام‌گذاری گره‌ها و فراخوانی آنها از نکات برجسته‌ی این زبان سه‌بعدی است که به کاربر امکان ساخت دنیایی مجازی و نزدیک به واقعیت را می‌دهد.

- گره‌های مشتق: ساختار سلسله‌مراتبی این زبان توسط گره‌های مشتق، پیاده‌سازی می‌شود. برخی از گره‌ها می‌توانند شامل گره‌های دیگر شوند. به این دسته از گره‌ها، گره‌های گروهی<sup>۱۷۲</sup> گفته می‌شود.

ساختار نوشتاری اطلاعات فوق در مورد گره‌ها به صورت زیر می‌باشد.

*DEF* { "گره‌های مشتق" "فیلدها" } "نوع شیء" "نام شیء"

در ساختار فوق تنها نوع شیء و علامت‌های قلاب الزامی هستند. در صورت استفاده از "نام شیء"، استفاده از کلمه کلیدی *DEF* الزامی می‌باشد.

### ساختار نوشتاری

برای سادگی در شناسایی فایل‌های مربوط به زبان VRML و نگارش<sup>۱۷۳</sup> مربوط به آن، تمام این فایل‌ها با عبارت زیر آغاز می‌شوند.

*#VRML V1.0 ascii*

علامت # معرفی کننده توضیحات متنی<sup>۱۷۴</sup> بوده و تمام عبارات بعد از این علامت تا پایان خط به عنوان توضیحات قلمداد شده و در برنامه‌ی نوشته شده تأثیری نخواهند گذاشت. تنها استثنا استفاده از این علامت در رشته‌های کاراکتری می‌باشد. بعد از عبارت فوق، تمام فایل‌های VRML شامل یک گره خواهند بود. این گره از نوع گروهی بوده و تمامی گره‌های موجود در فایل، گره‌های مشتق این گره‌ی مبنا می‌باشند.

---

<sup>172</sup> Group Nodes

<sup>173</sup> Version

<sup>174</sup> Comments



VRML حساس به حروف بزرگ و کوچک می‌باشد. از ارقام نمی‌توان به عنوان حرف ابتدایی "نام شیء" استفاده کرد و هیچ یک از نام‌های مورد استفاده نیز نباید شامل حروف " " " " \ { } . # و & بشوند. "فیلدها" با حروف کوچک شروع شده و "نوع شیء" می‌بایست با حروف بزرگ آغاز شود.

#### دستگاه مختصات

VRML از یک دستگاه مختصات کارتزین راستگرد سه‌بعدی استفاده می‌کند. به صورت پیش‌فرض اشیاء، بر صفحه‌ی متعامد بر محور Z که در جهت عمود بر نمایشگر می‌باشد، تصویر می‌شوند. در این صفحه محور X در جهت راست و محور Y در جهت بالا قرار دارند. با استفاده از جابجایی جهت دوربین و یا اعمال تبدیلات هندسی می‌توان این تنظیمات پیش‌فرض را تغییر داد. واحد استاندارد برای اندازه‌گیری فاصله، متر و واحد اندازه‌گیری زاویه، رادیان می‌باشد.

علاوه بر دستگاه مختصات فراگیر مذکور، یک فایل VRML می‌تواند هر تعداد دستگاه مختصات محلی داشته باشد. این دستگاه‌ها توسط تبدیلات فضایی همچون انتقال، دوران و تغییر مقیاس پیاده‌سازی می‌شوند. اگر  $V$  یک نقطه در فضای سه‌بعدی باشد دستورات زیر این نقطه را در فضای محلی جدید بیان می‌کنند.

*Translation { translation T }*  
*Rotation { rotation R }*  
*Scale { scaleFactor S }*  
*Coordinate3 { point V } PointSet { numPoints 1 }*

همانطور که در ابتدای این بخش اشاره شد هر گره همچون گره‌های انتقال، دوران و تغییر مقیاس فوق بر روی تمامی گره‌های بعدی همچون *Coordinate3* اعمال می‌شوند. بنابر نکات مطرح شده در فصل ۲، توالی تبدیلهای هندسی فوق معادل با رابطه‌ی زیر می‌باشد.

$$V' = T.R.S.V \quad (1-3)$$

که در رابطه‌ی فوق  $V'$  بیانگر نمایش نقطه‌ی  $V$  در دستگاه مختصات جدید ناشی از تبدیلات مذکور می‌باشد. گره‌های  $Coordinate3$  و  $PointSet$  در بخش‌های بعد تشریح خواهند شد.

## فیلدها

فیلدها پارامترهای تعیین‌کننده‌ی گره‌ها می‌باشند و نوع آنها معادل با نوع متغیر در سایر زبان‌ها می‌باشد. در این بخش نوع فیلدهای مهم و کاربردی در نمایش اشیاء ایستا مورد بررسی قرار می‌گیرند. دو نوع فیلد در VRML وجود دارد که شامل فیلدهای تک‌مقدرای و فیلدهای چندمقداری می‌شوند. فیلدهای تک‌مقداری با عبارت "SF" و فیلدهای چندمقداری با عبارت "MF" آغاز می‌شوند. فیلد تک‌مقداری، فیلدی است که تنها شامل یک مقدار شود. این یک مقدار می‌تواند یک عدد، یک بردار و حتی یک تصویر باشد. اجزای یک فیلد تک‌مقداری با استفاده از کاراکتر فاصله از هم جدا می‌شوند. فیلدهای چندمقداری، دنباله‌ای از مقادیر هستند که توسط علامت ',' از هم متمایز شده و توسط '[' ]' احاطه می‌شوند. فیلدهای چندمقداری نیز می‌توانند تنها شامل یک مقدار شوند. در این صورت نیازی به علامت ',' و '[' ]' نمی‌باشد. برای مثال اگر یک فیلد چندمقداری برابر عدد ۱ باشد، نمایش این مقدار به هر سه صورت ۱، [1,] و [1] صحیح می‌باشد. در این بخش چند فیلد پرکاربرد را تشریح می‌کنیم.

- SFCColor/MFCColor: این فیلد جهت معرفی رنگ در بعضی از گره‌ها بکار می‌رود. فیلد تک‌مقداری شامل سه عدد اعشاری بین ۰ و ۱ می‌باشد. فیلد چندمقداری نیز از فیلدهای تک‌مقداری طبق قواعد مذکور تشکیل می‌شود برای مثال [1.0 0.0, 0 1.0 0.0, 0 1.0 0.0] یک فیلد سه‌مقداری شامل رنگ‌های قرمز، سبز و آبی می‌باشد.

- SFFloat/MFFloat : این فیلد جهت معرفی اعداد اعشاری بکار می‌رود. دو روش نوشتن اعداد اعشاری در فیلد دو متغیره‌ی  $[3.141596, 1.5e-2]$  نشان داده شده است.

- SFImage : این فیلد شامل یک تصویر دوبعدی رنگی و یا سیاه‌وسفید می‌شود. این فیلد با دو عدد بیانگر عرض و ارتفاع تصویر شروع می‌شود. تعداد اجزای تصویر که بیانگر تعداد اجزای هر پیکسل تصویر می‌باشد، پس از این دو عدد قرار می‌گیرد. برای مثال برای تصاویر سیاه‌وسفید تعداد اجزا برابر ۱، برای تصاویر سیاه‌وسفید با پارامتر شفافیت تعداد اجزا برابر ۲، برای تصاویر رنگی تعداد اجزا برابر ۳ و برای تصاویر رنگی با پارامتر شفافیت تعداد اجزا برابر ۴ می‌باشد. در نهایت نیز به تعداد (عرض  $\times$  ارتفاع) عدد در مبنای ۱۶ که بیانگر مقادیر پیکسل‌ها می‌باشند، قرار می‌گیرند. تعداد بایت هر یک از این اعداد برابر با تعداد اجزای تصویر می‌باشد. برای مثال برای تصاویر سیاه‌وسفید از اعداد مبنای ۱۶ یک‌بایتی استفاده می‌شود. 0xff بیانگر سطح خاکستری سفید و 0x00 بیانگر سطح سیاه می‌باشد. به همین ترتیب برای تصاویر رنگی دارای پارامتر شفافیت از اعداد مبنای ۱۶ چهاربایتی استفاده می‌شود. برای مثال 0x00ff0080 بیانگر یک پیکسل نیمه شفاف دارای رنگ سبز می‌باشد. از آنجایی که VRML با این اعداد مبنای ۱۶ به صورت عدد برخورد می‌کند، نوشتن صفرهای سمت چپ عدد الزامی نداشته و رنگ مذکور را می‌توان به صورت 0xff0080 نیز نمایش داد. عبارت زیر یک تصویر رنگی  $4 \times 2$  را با ترکیبی از رنگهای قرمز، سبز، سیاه، سفید و زرد نمایش می‌دهد.

2 4 3 0xFF0000 0xFF00 0 0 0 0 0xFFFFF 0xFFFF00

- SFMatrix : این فیلد بیانگر یک ماتریس  $4 \times 4$  می‌باشد. از ۱۶ عدد اعشاری تشکیل شده و اعداد به ترتیب قرارگیری در ستون‌ها بیان می‌شوند. عبارت زیر یک ماتریس انتقال در جهت محور x به اندازه‌ی ۲,۳ واحد را نشان می‌دهد.

1 0 0 0 0 1 0 0 0 0 1 0 2.3 0 0 1

- SFRotation : این فیلد به عنوان پارامتر دوران در بعضی از گره‌ها استفاده می‌شود. این فیلد از چهار عدد اعشاری تشکیل می‌شود. ۳ عدد اول محور دوران را مشخص کرده و عدد چهارم میزان زاویه‌ی دوران معین می‌کند. عبارت زیر دوران حول محور x به اندازه ی ۱۸۰ درجه را نشان می‌دهد.

1 0 0 3.14159265

## گره‌ها

اغلب گره‌ها در VRML را می‌توان به سه دسته‌ی گره‌های شکل، خاصیت و گروهی تقسیم کرد. گره‌های شکل گره‌های معرف هندسه هستند. به عبارت دیگر تنها گره‌های شکل می‌توانند به ترسیم اشیاء بپردازند. گره‌های خاصیت نحوه ترسیم اشیاء را تحت تأثیر قرار می‌دهند. گره‌های گروهی نیز دسته‌بندی اشیاء و تعیین گره‌های مشتق و مبنا را بر عهده دارند. فیلدهای مربوط به هر گره به صورت ترکیبی از نام فیلد و مقدار فیلد مشخص می‌شوند. در صورتی که فیلدی مقداردهی نشود، مقدار پیش‌فرض برای آن در نظر گرفته می‌شود. ترتیب نوشتاری فیلدها نیز اهمیتی ندارد. برای مثال یک گره‌ی مکعب با طول، عرض و ارتفاع مشخص را می‌توان به هر دو صورت زیر معرفی کرد.

*Cube { width 2 height 4 depth 6 }*

*Cube { height 4 depth 6 width 2 }*

در این بخش تنها گره‌های مورد استفاده در نمایش اشیاء ایستا تشریح می‌شوند.

- Coordinate3 : این گره در دسته‌ی گره‌های خاصیت قرار داشته و در نتیجه عملکرد ترسیم‌ی ندارد. این گره نقاطی را که توسط گره‌های IndexedFaceSet, IndexedLineSet و PointSet می‌بایست ترسیم شوند، مشخص می‌کند. این گره دارای یک فیلد به صورت زیر می‌باشد.

```
Coordinate3 {  
    point 0 0 0 # MFVec3f  
}
```

دقت شود که point از نوع MF است. بنابراین می‌توان چند نقطه را با استفاده از حروف " ، [ ] " توسط این گره معرفی کرد.

- IndexedFaceSet: این گره از نوع گره‌های شکل بوده و یک شیء سه بعدی را با استفاده از چندضلعی‌ها نمایش می‌دهد. چند ضلعی‌ها با استفاده از اندیس نقاط مطرح شده در Coordinate3، در فیلد coordIndex مشخص می‌شوند. هر چندضلعی یا وجه توسط اندیس 1- از سایر وجوه متمایز می‌شود. علاوه بر این امکان انتساب بافت به هر یک از رئوس چندضلعی‌ها نیز با استفاده از فیلد textureCoordIndex وجود دارد. در این فیلد می‌بایست اندیس پیکسل‌های موجود در گره‌ی TextureCoordinate2 قرار گیرند. عملکرد این گره نیز متعاقباً تشریح خواهد شد. باید توجه کرد که تعداد اندیس‌های موجود در فیلد textureCoordIndex می‌بایست با تعداد اندیس‌های موجود در فیلد coordIndex برابر باشد. این الزام جهت افزایش سرعت پیاده‌سازی بافت بر روی چندضلعی‌ها در نظر گرفته شده است. VRML هنگام درون‌یابی نقاط روی چند ضلعی‌ها، مشخصات رنگ رئوس چندضلعی را از فیلد textureCoordIndex استخراج می‌کند. بنابراین در صورت انتساب رنگ‌های مختلف به یک رأس مشترک در چند وجه مشکلی از لحاظ پیاده‌سازی بافت، بوجود نخواهد آمد.

علاوه بر موقعیت فضایی و بافت هر رأس، امکان مشخص کردن بردارهای متعامد و مشخصات نوری رئوس همچون ضرایب بازتابش محیطی، ضرایب بازتابش پخشی و آینه‌ای و میزان شفافیت نیز وجود دارد. این مشخصات به ترتیب توسط فیلدهای `normalIndex` و `materialIndex` بیان می‌شوند. باید توجه کرد که در صورت مشخص کردن بافت وجوه با استفاده از `textureCoordIndex`، مشخصات نوری رئوس تحت تأثیر قرار می‌گیرند. این تأثیر بر اساس تعداد اجزاء تصویر دو بعدی به صورت ضرب و یا جایگزینی می‌باشد. از آنجایی که هدف ما تنها خواندن اطلاعات مربوط به شکل و بافت چهره‌های موجود در پایگاه داده است، در این بخش به جزئیات این مطلب نمی‌پردازیم.

فیلدهای موجود در این گره و مقادیر پیش‌فرض آن به صورت زیر می‌باشند.

```
IndexedFaceSet {
    coordIndex      0 # MFLong
    materialIndex   -1 # MFLong
    normalIndex     -1 # MFLong
    textureCoordIndex -1 # MFLong
}
```

- `Separator`: این گره مهمترین و پرکاربردترین گره‌ی گروهی می‌باشد. با استفاده از این گره عملکرد تمام گره‌های مشتق، مستقل از سایر بخش‌های گراف صحنه خواهد بود. به عبارت دیگر هیچ یک از گره‌های خاصیت موجود در خارج این گره، تأثیری بر گره‌های مشتق این گره نخواهند داشت. علاوه بر این گره‌های خاصیت موجود در این گره نیز تأثیری بر گره‌های خارجی نخواهند گذاشت. همانطور که در بخش مربوط به دستگاه‌های مختصات مطرح شد، با استفاده از این گره می‌توان دستگاه‌های مختصات محلی نیز ایجاد کرد.

- `TextureCoordinate2`: این گره شامل تعدادی مختصات دو بعدی است که در فیلد `textureCoordIndex` مربوط به گره‌ی `IndexedFaceSet` استفاده می‌شوند. این مختصات دو بعدی به پیکسل‌های یک تصویر دوبعدی که توسط گره‌ی `Texture2` معرفی می‌شود، اشاره می‌کنند. مشخصات رنگ پیکسل مذکور به عنوان

مشخصات بافت رأس مورد نظر، در نظر گرفته می‌شوند. فیلدهای گرهی TextureCoordinate2 به صورت زیر می‌باشند.

```
TextureCoordinate2 {
    point 0 0 #MFVec2f
}
```

فیلد point که بیانگر پیکسل مورد نظر در تصویر دوبعدی است، می‌بایست مقداری نرمالیزه شده بین 0 و 1 باشد.

- Texture2: این گره مشخص کننده‌ی تصویری است که توسط TextureCoordinate2 به عنوان بافت رؤس چندضلعی‌ها در نظر گرفته می‌شود. معرفی این تصویر به صورت یک فایل تصویری و یا استفاده از فیلد SFImage می‌باشد. در صورت استفاده از فایل تصویری، اطلاعات موجود در فیلد SFImage غیر معتبر خواهند بود. فیلدهای این گره و مقادیر پیش فرض آن به صورت زیر می‌باشند.

```
WRAP ENUM
    REPEAT #Repeats texture outside 0-1 texture coordinate range
    CLAMP #Clamps texture coordinates to lie within 0-1 range
Texture2 {
    filename "" #SFString
    image 0 0 0 #SFImage
    wrapS REPEAT #SFEnum
}
```

فیلد warpS نحوه‌ی برخورد با مقادیری از فیلد point در گرهی TextureCoordinate2 را که خارج از حدود 0 و 1 باشند، نشان می‌دهد.

- Transform: این گره از گره‌های خاصیت بوده و بیانگر تبدیلی هندسی شامل انتقال، دوران و تغییر مقیاس می‌باشد. فیلدهای مربوط به این گره و مقادیر پیش فرض آن به شرح زیر می‌باشند.

```
Transform {
    translation T1 0 0 0 #SFVec3f
    rotation R1 0 0 1 0 #SFRotation
    scaleFactor S 1 1 1 #SFVec3f
    scaleOrientation R2 0 0 1 0 #SFRotation
    center T2 0 0 0 #SFVec3f
}
```

با توجه به نحوه‌ی عملکرد تبدیل فوق و مطالب ارائه شده در فصل ۲، این تبدیل را معادل با تبدیلات متوالی زیر می‌باشد.

$$T1.T2.R1.R2.S.R2^{-1}.T2^{-1}$$

در انتهای این بخش، یک فایل wrl مربوط پایگاه داده چهره‌ی سه‌بعدی GavabDB را بررسی می‌کنیم. چهره‌های موجود در این پایگاه داده تنها دارای اطلاعات مربوط به موقعیت فضایی می‌باشند. خلاصه‌ای از این فایل به صورت زیر می‌باشد.

```
#VRML V1.0 ascii
Separator {
  Transform {
    translation 0 0 0
    scaleFactor 0.5 0.5 0.5
  }
  Separator {
    Transform {
      translation 0 0 0
    }
    Coordinate3 {
      point [
        39.39 122.05 -1243.33,
        40.85 122.05 -1243.32,
        42.26 121.90 -1241.73,
        -4.31 118.84 -1223.68,
        ...
      ]
    }
    IndexedFaceSet {
      coordIndex [
        0, 8, 9, 1, -1,
        1, 9, 10, 2, -1,
        2, 10, 11, -1,
        3, 12, 13, 4, -1,
        ...
      ]
    }
  }
}
```

همانطور که مشاهده می‌شود در این پایگاه داده گره‌های Coordinate3 و IndexedFaceSet جهت تعیین و ترسیم چندضلعی‌ها استفاده شده است.

Abstract



One of the most important challenges in biometric recognition area, is finding new recognition methods with more applicability and proper functionality. Because of its importance in human recognition capabilities, face recognition is attended as a convenient case for this purpose.

Although a lot of researches have been done in finding an effective method for face recognition, existing algorithms operate well just in controlled conditions and cannot be used in real environments. With respect to 3D nature of face, 3d face recognition methods have more potential than 2D methods in solving existing problems. In addition, because of their capability in face synthesis and their lack of need to 3D input images, model based 3d face recognition methods, seem to be more exciting and more important.

In this research, after a literature review of face recognition methods, advantage and disadvantages of each of their branches would be studied. Because of wide application of computer graphics in 3D face recognition, this subject has been attended carefully and completely.

Model based 3d face recognition is divided into two main methods. The first one is an approach based on generic models and the second method applies morphable models. This thesis is focused on studying and implementation of morphable model based method. But for creating a morphable model, we need a 3D face database. Because of lack of such a database, we decided to make a 3D face models database from 2D images. So, the photogrammetry algorithm which is based on generic models, has been used. Hence, both main methods of model based 3D face recognition have been studied and implemented in this thesis. Creating a 3D face database from 2D images, in addition to its algorithmic complexity, because of its need to determining quite large number of landmarks, is a time consuming process. In this research two 3D face databases have been created from 2D images.

In first one, 2D face images are extracted from MPI databases, and in second one, face images of different Iranian races have been used.

Because of its remarkable features and advantages, face recognition based on morphable model is studied as the most important 3D face recognition method. Indeed, morphable model is a combination of deformable models and computer graphics techniques for simulation of illumination effects. In this model changes in face pose and illumination is implemented simply. Also this model can estimate carefully pose, shape, texture and illumination of face model with no need of operator's attendances. In addition face synthesis and capability of facial expressions implementation on synthesized model, are of other advantages of this method.

The core step of morphable modeling is finding a dense point to point correspondence between instant models. For implementing this step, in this research, a novel and graph based method have been presented which optimizes image matching results, optimization results of this method, show remarkable improvement in morphable model functionality and more quality in face synthesis. This algorithm is a graph optimization algorithm. So it is independent of matching method and it can be implemented on results of all image matching methods. In addition to optimization of matching results, injectivation is one of the other benefits of this novel method.

After implementation of morphable model, effect of 3D face database on recognition and synthesis quality was studied. So, input images of Iranian faces, was synthesized by two morphable models based on two presented 3D databases. High quality of images synthesized by the morphable model based on Iranian 3D face database, shows effective role of database in morphable modeling. In this research, faces were recognized by rate of 100% by the morphable model algorithm. In spite of low number of input test images, this high recognition rate, shows usefulness and power of this model in face recognition.

فهرست مراجع و منابع انگلیسی:

- Ackermann, B. and Bunke, H., t. . C. (2000). In *15th International Conference on Pattern Recognition*, pages 809–813.
- Ansari, A. and Abdel-Mottaleb, M. (2003a). 3D face modelling using two orthogonal views and a generic face model. In *International Conference on Multimedia and Expo*, pages 289–292.
- Ansari, A. and Abdel-Mottaleb, M. (2003b). 3D face modelling using two views and a generic face model with application to 3D face recognition. In *IEEE Conference on Advanced Video and Signal Based Surveillance*, pages 37–44.
- Arun, K., Huang, T., and S.D., B. (1987). Least squares fitting of two 3D point sets. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 9(5):698–700.
- Beier, T., and Neely, S. (1992). Feature-based Image Metamorphosis. In *SIGGRAPH 92 Conference Proceedings*, ACM SIGGRAPH, pages 35–42.
- Besl, P., and McKay, N. (1992). A method for registration of 3D shapes. *IEEE Trans. On Pattern Analysis and Machine Intelligence*, 14(2):239–256.
- Beumier, C. and Acheroy, M. (2000). Automatic 3D face authentication. *Image and Vision Computing*, 18(4):315–321.
- Beumier, C. and Acheroy, M. (2001). Face verification from 3D and grey level clues. *Pattern Recognition Letters*, 22:1321–1329.
- Blanz, V. and Vetter, T. (1999). A Morphable Model for the Synthesis of 3D Faces, *Computer Graphics Proc. SIGGRAPH '99*, pages 187-194.

- Blanz, V., Vetter, T. (2003). Face Recognition Based on Fitting a 3D Morphable Model. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(9):1063-1074.
- Bookstein, F.L. (1989). Principal warps: thin-plate splines and the decomposition of Deformations. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 11(6):567–585.
- Bowyer, K. W., [Chang](#), K. I., [Flynn](#), P. J. (2006). A survey of approaches and challenges in 3D and multi-modal 3D + 2D face recognition. [Computer Vision and Image Understanding](#), 101(1):1-15.
- Bregler, C., Covell, M., and Slaney, M. (1997). Video Rewrite: Driving Visual Speech with Audio. In *SIGGRAPH 97 Conference Proceedings*, ACM SIGGRAPH, pages 353–360.
- Bruce, V. (1982). Visual and non-visual coding processes in face recognition. *British Journal of Psychology*, 73:105–116.
- Bruhn, A., Weickert, J., and Schnorr, C. (2002). Combining the Advantages of Local and Global Optic Flow Methods. In *DAGM*, pages 454-462.
- Burt, P. J., and Adelson, E.H. (1985). Merging Images through Pattern Decomposition. *Proc. Applications of Digital Image Processing*, 8(575):173-181.
- Cartoux, J., LaPrete, J., and Richetin, M. (1989). Face authentication or recognition by profile extraction from range images. In *Workshop on Interpretation of 3D Scenes*, pages 194–199.
- Chua, C. and Jarvis, R. (1997). Point signatures - a new representation for 3D object recognition. *International Journal of Computer Vision*, 25(1):63–85.
- Chen, D. T., State, A., and Banks, D. (1995). Interactive Shape Metamorphosis. In *1995 Symposium on Interactive 3D Graphics*, ACM SIGGRAPH, pages 43–44.
- Cyberware Laboratory, Inc, Monterey, California. (1990). *4020/RGB 3D Scanner with Color Digitizer*.
- Duda, R. O., Hart, P. E., and Stork, D. G. (2001). *Pattern Classification*, second ed. John Wiley & Sons.

- Edwards, G. J., Cootes, T. F., and Taylor, C.J. (1998). Face Recognition Using Active Appearance Models. *Proc. Conf. Computer Vision (ECCV '98)*.
- Foley, J.D., Van Dam, A., Feiner, S.K., and Hughes, J.F. (1996). *Computer Graphics: Principles and Practice*, second ed. Addison-Wesley.
- Gokberk, B., Salah, A., and Akarun, L. (2005). Rank-based decision fusion for 3D shapebased face recognition. In *International Conference on Audio- and Video-based Biometric Person Authentication*, pages 1019–1028.
- Gordon, G. (1992). Face recognition based on depth and curvature features. In *IEEE Computer Vision and Pattern Recognition (CVPR)*, pages 808–810.
- Heisele, B., and Blanz, V. (2006). Morphable models for training a component-based face recognition system. In *Face Processing, Advanced Modeling and Methods*, W. Zhao and R. Chellapa (Eds.), Elsevier, pages 439-462.
- Hill, H., Schyns, P., and Akamatsu, S. (1997). Information and viewpoint dependence in face recognition. *Cognition*, 62:201–202.
- Horn, B. K. P. (1984). Extended Gaussian Images. *Proceedings of the IEEE*, pages 1671-1686.
- Horn, B., and Schunck, B. (1981). Determining optical flow. *Artificial Intelligence*, 17:185–203.
- Ip, H. S., and Yin, L. (1996). Constructing a 3D Individualized Head Model from Two Orthogonal Views. *The Visual Computer*, 12:254–266.
- Johnston, A., Hill, H., and Carman, N. (1992). Recognizing faces: effects of lighting direction, inversion and brightness reversal. *Perception*, 21:365–375.
- Jones, M., and Poggio, T. (1996). Model-based matching by linear combination of prototypes. a.i. memo no., Artificial Intelligence Laboratory, Massachusetts Institute of Technology.
- Kim, T. and Kittler, J. (2005). Locally linear discriminant analysis for multimodally distributed classes for face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(3):318–327.

- Kurihara, T., and Arai, K. (1991). A Transformation Method for Modeling and Animation of the Human Face from Photographs. In *Nadia Magnenat Thalmann and Daniel Thalmann, editors, Computer Animation 91, Springer-Verlag, Tokyo*, pages 45–58.
- Lee, J. and Milios, E. (1990). Matching range images of human faces. In *International Conference on Computer Vision (ICCV)*, pages 722–726
- Lee, Y. and Shim, J. (2004). Curvature-based human face recognition using depth-weighted hausdorff distance. In *International Conference on Image Processing*, pages 1429–1432.
- Lee, Y., Song, H., Yang, U., Shin, H., and Sohn, K. (2005). Local feature based 3D face recognition. In *International Conference on Audio- and Video-based Biometric Person Authentication*, pages 909–918.
- Lee, Y., Terzopoulos, D., and Waters, K. (1995). Realistic Modeling for Facial Animation. *Computer Graphics*, 29(4):55-62.
- Lee, S., Wolberd, G., and Shin, S. (1997). Scattered data interpolation with multilevel Bsplines. *IEEE Trans. on Visualization and Computer Graphics*, 3(3):228–244.
- Li, Y., Gong, S., and Lidell, H. (2000). Support vector regression and classification based multiview face detection and recognition. In *International Conference on Face and Gesture Recognition*, pages 300–305.
- Liu, C., Collin, C., Burton, A., and Chaudhuri, A. (1999). Lighting direction affects recognition of untextured faces in photographic positive and negative. *Vision Research*, 39:4003–4009.
- Liu, X., Srivastava, A., and Gallivan, K. (2004) Optimal linear representations of images for object recognition. *IEEE Trans. Pattern Anal. Machine Intell.*, 26(5):662–666.
- Lu, X., Colbry, D., and Jain, A. (2004). Matching 2.5D scans for face recognition. In *International Conference on Pattern Recognition*, pages 362–366.

Lu, X. and Jain, A. (2005a). Deformation analysis for 3D face matching. In *IEEE Workshop on Applications of Computer Vision*, pages 362–366.

Lucas, B., and Kanade, T. (1981). An iterative image registration technique with an application to stereo vision. In *Proc. Seventh International Joint Conf. on Artificial Intelligence*, Vancouver, Canada, 674–679.

Medioni, G. and Waupotitsch, R. (2003). Face recognition and modeling in 3D. In *IEEE International Workshop on Analysis and Modeling of Faces and Gestures*, pages 232–233.

Moreno, A., Sanchez, A., Velez, J., and Diaz, F. (2003). Face recognition using 3D surfaceextracted descriptors. In *Irish Machine Vision and Image Processing Conference*.

[MPI] <http://faces.kyb.tuebingen.mpg.de/>

Nagamine, T., Uemura, T., and Masuda, I. (1992). 3D facial image analysis for human identification. In *International Conference on Pattern Recognition*, pages 324–327.

Pan, G., Wu, Z., and Pan, Y. (2003). Automatic 3D face verification from range data. Pages 193–196.

Papatheodorou, T. and Rueckert, D. (2004). Evaluation of automatic 3D face recognition using surface and texture registration. In *International Conference on Automated Face and Gesture Recognition*, pages 321–326.

Parke, F. I. (1972) *Computer Generated Animation of Faces*. Master's thesis, University of Utah, Salt Lake City, UT.

Parke, F. I. (1974) *A Parameterized Model for Human Face*. PhD thesis, University of Utah, Salt Lake City, UT.

Pighin, F., Hecher, J., Lischinski, D., Szeliski, R., and Salesin, D. H. (1998). Synthesizing Realistic Facial Expressions from Photographs. In *SIG-GRSPH 98 Conference Proceedings*, pages 75-84.

Platt, S. M. (1980) *A System for Computer Simulation of the Human Face*. Master's thesis, The Moore School, University of Pennsylvania, Philadelphia.

- Platt, S. M. (1980). A System for Computer Simulation of the Human Face. Master's thesis, The Moore School, University of Pennsylvania, Philadelphia.
- Russ, T., Koch, M., and Little, C. (2005). A 2D range hausdorff approach for 3D face recognition. In *Computer Vision and Pattern Recognition*, pages 1429–1432.
- Sederberg, T., and Parry, S. (1986). Free-form deformation of solid geometric models. In Proc. of the 13th Annual Conf. on Computer Graphics and Interactive Techniques, pages 151–160
- Srivastava, A., Liu, X., and Heshner, C. (2003). Face recognition using optimal linear components of face images. *Journal of Image and Vision Computing*, 24(3):291–299.
- Tanaka, H., Ikeda, M., and Chiaki, H. (1998). Curvature-based face surface recognition using spherical correlation principal directions for curved object recognition. In *International Conference on Automated Face and Gesture Recognition*, pages 372–377.
- Tarr, M. and Bulthoff, H. (1995). Is human object recognition better described by geon structural descriptions or by multiple views. *Journal of Experimental Psychology*, 21:71–86.
- Terzopoulos, D. and Waters, K. (1990). Physically-based Facial Modeling, Analysis, and Animation. *Journal of Visualization and Computer Animation*, 1(4):73-80.
- Tsalakanidou, F., Malassiotis, S., and Strintzis, M. (2004). Integration of 2D and 3D images for enhanced face authentication. In *International Conference on Automated Face and Gesture Recognition*, pages 266–271.
- Tsalakanidou, F., Tzocaras, D., and Strintzis, M. (2003). Use of depth and colour eigenfaces for face recognition. *Pattern Recognition Letters*, 24:1427–1435.
- Turk, M., and Pentland, A. (1991). Eigenfaces for recognition. *Journal for Cognitive Neuroscience*, 3(1):71–86.
- Vetter, T. and Blanz, V. (1998). Estimating Coloured 3D Face Models from Single Images: An Example Based Approach. Proc. Conf. Computer Vision (ECCV '98), vol. 2.

Vetter, T., Jones, M. J., and Poggio, T. (1997). A bootstrapping algorithm for learning linear models of object classes. in IEEE Conference on Computer Vision and Pattern Recognition, CVPR'97, (Puerto Rico, USA), IEEE Computer Society Press.

[VRML]: [www.web3d.org/x3d/specifications/vrml/VRML1.0](http://www.web3d.org/x3d/specifications/vrml/VRML1.0)

Waters, K. (1987) A Muscle Model for Animating Three-Dimensional Facial Expressions. *Computer Graphics (ACM SIGGRAPH'87)*, 21(4):17-24.

Wiskott, L., Fellous, J.M., Kruger, N., and Von Der Malsburg, C. (1997). Face Recognition by Elastic Bunch Graph Matching. *IEEE Trans. On Pattern Analysis and Machine Intelligence*, 19(7):775–779.

Wong, H., Chueng, K., and Ip, H. (2004). 3D head model classification by evolutionary optimization of the extended gaussian image representation. *Pattern Recognition*, 37(12):2307–2322.

Wu, Y., Pan, G., and Wu, Z. (2003). Face authentication based on multiple profiles extracted from range data. In *Audio- and Video-Based Biometric Person Authentication*, pages 515– 522.

Xu, C., Wang, Y., Tan, T., and Quan, L. (2004). Automatic 3D face recognition combining global geometric features with local shape variation information. In *International Conference on Automated Face and Gesture Recognition*, pages 308–313.

## **Abstract**

One of the most important challenges in biometric recognition area is finding new recognition methods with more applicability and proper functionality. Because of its importance in human recognition capabilities, face recognition is attended as a convenient case for this



purpose. With respect to 3D nature of face, 3D face recognition methods have more potential than 2D methods in solving existing problems. In addition, because of their capability in face synthesis and their lack of need to 3D input images, model based 3D face recognition methods, seem to be more exciting and more important.

In this research, after a literature review of face recognition methods, advantage and disadvantages of each of their branches would be studied. Because of wide application of computer graphics in 3D face recognition, this subject has been attended carefully and completely.

Model based 3D face recognition is divided into two main methods. The first one is an approach based on generic models and the second method applies morphable models. For creating a morphable model, we need a 3D face database. Because of lack of such a database, we decided to make a 3D face models database from 2D images. So, the photogrammetry algorithm which is based on generic models, has been used. Hence, both main methods of model based 3D face recognition have been studied and implemented in this thesis. In this research two 3D face databases have been created from 2D images. In first one, 2D face images are extracted from MPI databases, and in second one, face images of different Iranian races have been used. Because of its remarkable features and advantages, face recognition based on morphable model is studied as the most important 3D face recognition method. Indeed, morphable model is a combination of deformable models and computer graphics techniques for simulation of illumination effects. In this model changes in face pose and illumination is implemented simply. Also this model can estimate pose, shape, texture and illumination of face model carefully and with no need of operator's attendances. In addition face synthesis and capability of facial expressions implementation on synthesized model, are of other advantages of this method.

The core step of morphable modeling is finding a dense point to point correspondence between instant models. For implementing this step, in this research, a novel and graph based method have been presented which optimizes image matching results, optimization results of this method, show remarkable improvement in morphable model functionality and more quality in face synthesis. This algorithm is a graph optimization algorithm. So it is independent of matching method and it can be implemented on results of all image matching methods. In addition to optimization of matching results, injectivation is one of the other benefits of this novel method.

After implementation of morphable model, effect of 3D face database on recognition and synthesis quality was studied. So, input images of Iranian faces, was synthesized by two morphable models based on the two presented 3D databases. High quality of images synthesized by the morphable model based on Iranian 3D face database, shows effective role of database in morphable modeling. In this research, faces were recognized by rate of 100% by the morphable model algorithm. In spite of low number of input test images, this high recognition rate, shows usefulness and power of this model in face recognition.

Keywords: 3D face recognition, 3D face modeling, photogrammetry, morphable model, image matching



Shahrood University of Technology

Electrical and Robotics Faculty

## **Model Based 3D face Recognition**

Mehdi Tajiani

Supervisor:  
Dr. Ali Solaymani

September 2008

