

به نام خالق نیکی‌ها

۱۳۱۷۶۶



دانشگاه شهید بهشتی
دانشکده مهندسی برق و کامپیوتر

تخمین تصویر انحرافی با استفاده از تبدیل ریجالت، به منظور ساخت تصویر میانی
در بینایی استریو

پایان نامه کارشناسی ارشد مهندسی برق
گرایش مخابرات

توسط:

محمد اسلامی

۱۳۸۸ / ۱۲ / ۲

کتابخانه مرکزی دانشگاه شهید بهشتی

استاد راهنما:
فرح ترکمنی آذر

۱۳۸۸

۱۳۱۷۶۶



دانشگاه شهید بهشتی
دانشکده مهندسی برق و کامپیوتر

پایان نامه کارشناسی ارشد مهندسی برق - گرایش مخابرات
تحت عنوان:

تخمین تصویر انحرافی با استفاده از تبدیل ریجالت، به منظور ساخت تصویر میانی در
بینایی استریو

در تاریخ ۸۸/۴/۲۵
پایان نامه دانشجو، (محمد اسلامی)، توسط کمیته تخصصی داوران مورد بررسی و تصویب نهائی
قرار گرفت.

امضاء
امضاء
امضاء
امضاء

دکتر فرح ترکمنی آذر
دکتر محمد عشقی
دکتر جمشید شنبه زاده
دکتر فرح ترکمنی آذر

۱- استاد راهنما اول:
۴- استاد داور (داخلی)
۵- استاد داور (خارجی)
۶- نماینده تحصیلات تکمیلی

کلیه حقوق مادی مترتب بر نتایج مطالعات،

ابتکارات و نوآوریهای ناشی از تحقیق موضوع

این پایان نامه متعلق به دانشگاه شهید بهشتی

می باشد.

به نام خدا


نام و نام خانوادگی: محمد اسلامی

عنوان پایان نامه: تخمین تصویر انحرافی با استفاده از تبدیل ریجالت، به منظور ساخت تصویر میانی در بینایی استریو

استاد راهنما: دکتر فرح ترکمنی آذر

اینجانب محمد اسلامی تهیه کننده پایان نامه کارشناسی ارشد حاضر خود را ملزم به حفظ امانت داری و قردادانی از زحمات سایر محققین و نویسندگان بنا بر قانون Copyright می دانم. بدین وسیله اعلام می نمایم که مسئولیت کلیه مطالب درج شده با اینجانب می باشد و در صورت استفاده از اشکال؛ جداول، و مطالب سایر منابع، بلافاصله مرجع آن ذکر شده و سایر مطالب از کار تحقیقاتی اینجانب استخراج گشته است و امانتداری را به صورت کامل رعایت نموده ام. در صورتی که خلاف این مطلب ثابت شود، مسئولیت کلیه عواقب قانونی با شخص اینجانب می باشد.

نام و نام خانوادگی دانشجو: محمد اسلامی

امضاء و تاریخ:


تقدیم به خانواده و به خصوص برادرم که همواره پشتیبان بوده‌اند

فهرست مطالب

صفحه	
۱	۱-مقدمه -----
۸	۱-۱- ساختار هندسی دوربین پین هول -----
۱۳	۲-۱- خلاصه و جمع بندی -----
۱۴	۲- بینایی استریو -----
۱۵	۱-۲- روابط و تعاریف بنیادی در بینایی استریو -----
۱۵	۲-۱-۱- پیش زمینه -----
۱۶	۲-۱-۲- هندسه اپی پولار -----
۱۷	۲-۱-۳- قید اپی پولار -----
۱۸	۲-۱-۴- ماتریس بنیادی -----
۱۹	۲-۱-۵- یکسوسازی -----
۲۰	۲-۱-۶- بازسازی 3D -----
۲۳	۲-۲- بینایی استریو در انسان و درک عمق -----
۲۵	۳-۲- ساختارهای مختلف دوربین های استریو -----
۳۱	۴-۲- بدست آوردن تصویر میانی -----
۳۵	۵-۲- سنتز نما -----
۳۶	۶-۲- ساختار و حرکت -----
۳۸	۷-۲- مساله تناظر پیکسل ها -----
۳۸	۲-۷-۱- تعریف مساله -----
۴۳	۲-۷-۲- آنالیز استریو استاندارد مبتنی بر ناحیه -----
۴۷	۸-۲- پس پردازش -----
۴۹	۹-۲- جمع بندی -----
۵۰	۳- متد پیشنهادی -----
۵۱	۳-۱- تبدیل رادون و ریجالت پیوسته -----
۵۱	۳-۱-۱- تبدیل رادون پیوسته -----
۵۴	۳-۱-۲- ریجالت پیوسته -----
۵۵	۳-۲- تبدیل رادون و تبدیل ریجالت متناهی -----
۵۵	۳-۲-۱- تبدیل رادون متناهی -----
۵۸	۳-۲-۲- تبدیل ریجالت متناهی -----
۶۰	۳-۳- رابطه و الگوریتم پیشنهادی -----
۶۰	۳-۳-۱- رابطه پیشنهادی -----

۶۱	-----	۲-۳-۳- الگوریتم‌های جستجوی پاره‌ای
۶۲	-----	۳-۳-۳- الگوریتم پیشنهادی
۶۶	-----	۴-۳- خلاصه و جمع‌بندی
۶۷	-----	۴- جمع‌بندی
۸۳	-----	۵- جمع‌بندی و کارهای آتی
۸۴	-----	۵-۱- جمع‌بندی
۸۶	-----	۵-۲- کارهای آتی
۸۸	-----	اختصارات
۸۹	-----	واژه‌نامه
۹۴	-----	مقاله منتشرشده
۱۰۱	-----	مراجع

چکیده

بینایی استریو و بینایی چند سنسوری یکی از مهمترین مسائل و فیلدهای بهروز در زمینه سه بعدی سازی، رندرینگ، هدایت رباتها، ویدئو کنفرانس های استریو و متدهای افزودن واقعیت به شبیه سازها هستند. یکی از مهمترین مراحل در بینایی استریو، یافتن میزان انحراف بین دو تصویر چپ و راست و ساختن تصویری موسوم به تصویر انحراف است. این تصویر در خود اطلاعات مربوط به عمق را دارد. معیارهای محاسبه ی شباهت ها در تخمین انحراف بسیار با معیارهای به کار گرفته شده در تخمین حرکت در ویدئو رابطه دارند و از آنها می توان به MSE, SSD, SAD, NCC و غیره اشاره کرد. هدف این پایان نامه ارائه معیار شباهت متمایزی از SSD و ... با تکیه بر خاصیت لبه یابی تبدیل ریجלט است که نه تنها از معیارهای شباهت فوق الذکر کارایی بهتری داشته باشد بلکه نسبت تغییرات محیطی مانند نویز و .. نیز مقاوم تر باشد. در نهایت با استفاده از ماتریس انحرافی بدست آمده اقدام به بازسازی تصویر میانی شده است.

کلمات کلیدی:

بینایی استریو، تبدیل ریجלט، انحراف، ماتریس انحرافی، ساختار موازی دوربین ها، تصویر میانی

۱- فصل اول: مقدمه

بدون شک یکی از موثرترین ابزار بشر در درک جهان و شناخت آن ، دیدن و یا حس بینایی بوده است. یکی از مهمترین زمینه‌های تحقیقاتی در بینایی ماشین و کامپیوتر، درک خصوصیات سه بعدی یک منظره واقعی و یافتن اطلاعات هندسی اجسام موجود در آن می‌باشد. در ۳۰ سال اخیر این زمینه پژوهشی پیشرفت‌های قابل ملاحظه‌ای داشته است و متدهای بسیاری جهت نیل به این اهداف معرفی شده‌اند. در کاربردهای رباتیک ، استفاده از این تکنیک‌ها سبب می‌شود تا ربات اطلاعاتی از ساختار و شکل هندسی اجسام، مکان آنها (هم نسبت به ربات و هم نسبت به اجسام دیگر) بدست آورد.

از این متدها می‌توان به سیستم‌های پردازش بینایی چند سنسوری و عموماً بینایی استریو اشاره کرد. روش‌های درک خصوصیات سه بعدی منظره به دو گروه تقسیم می‌شوند [1] : بینایی پسیو^۱ و بینایی اکتیو^۲. در بینایی پسیو ، از وسیله‌ای به‌جز دوربین‌ها استفاده نمی‌گردد و سعی می‌شود توسط پردازش تصاویر حاصله و روابط مابین مکان دوربین‌ها اطلاعات عمق و بعدسوم را بدست آورند. در حقیقت این شکل از بینایی همانند بینایی چشم‌های انسان است. تکنیک‌های پسیو شامل مسائل بینایی استریو^۳، بینایی سه‌سنسوری^۱ و بسیاری تکنیک‌های درک شکل تک چشمی^۲ مانند استفاده از بافت^۳، فوکوس^۴، عدم فوکوس^۵، سایه^۶ و دیگر مشخصات تصاویر می‌شود.

¹ Passive visual sensing

² Active visual sensing

³ Stereo vision

یکی از مسائل بینایی پسیو، درک عمق و اطلاعات سه‌بعدی تنها از یک تصویر دوبعدی است، که این روش‌ها ذاتاً توانایی خوبی در نیل به این هدف ندارند و اصطلاحاً به آنها مسائل بیمار^۷ گفته می‌شود. بینایی استریو تکنیک دیگری برای درک عمق می‌باشد و نسبت به متد قبلی بسیار از دقت بالاتری برخوردار است. توانایی انسان در درک عمق و فاصله به دلیل حضور دو چشم در انسان است. از مسائل بینایی استریو در بینایی ماشین می‌توان به تنظیمات دوربین-ها^۸ نسبت به هم، محاسبه و تفسیر انحراف^۹ بین دو تصویر حاصله نام برد. مساله محاسبه انحراف تصاویر احتیاج به حل مساله تطابق و یا تناظر^{۱۰} بین نقاط تصاویر دارد که در نوع خود مساله‌ای پیچیده و در نتیجه زمانبر با پردازشی سنگین خواهد بود. یکی دیگر از معایب بینایی استریو پسیو، اینست که به دلیل تهیه تصاویر، می‌بایست همانند چشم انسان، این سیستم‌ها در حداقل نوری قرار داشته‌باشند تا بتوانند تصویری مناسب تهیه و سپس پردازش کنند.

در روش‌های پسیو تنها وسیله نمونه‌برداری دوربین‌ها هستند. همانند شکل (۱-۱) هر نقطه سه‌بعدی در دنیای واقعی به نقطه‌ای در تصویر نگاشت می‌گردد. حال اگر بخواهیم از نقطه‌ی واقع بر تصویر به نقطه‌ی سه‌بعدی برسیم، غیرممکن است. زیرا نقطه سه‌بعدی می‌تواند هر کجای خط واصل در شکل (۱-۱) قرار داشته‌باشد. این خط واصل را خط *backprojection* می‌نامند. ولی اگر دو تصویر از منظره وجود داشته‌باشد، می‌توان توسط برهم‌نهی تصاویر و چگونگی ارتباط آنها، مکان نقطه سه‌بعدی را تخمین زد. شکل (۲-۱) نمایی از این مفهوم را نشان می‌دهد. شکل (۱-۳) تصویر ربات مریخ‌نوردی که از بینایی استریو پسیو در آن استفاده شده است، را نشان می‌دهد. در این گزارش سعی داریم، این نوع بینایی ماشین را بیان و بررسی کنیم.

¹ Trinocular vision

² Monocular shape-from-X techniques

³ Texture

⁴ Focus

⁵ Defocus

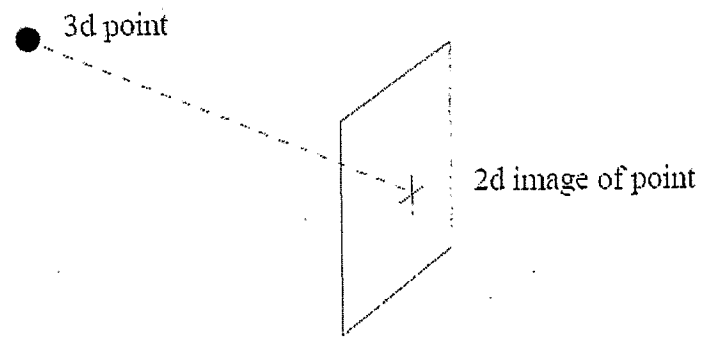
⁶ Shadows & Shadings

⁷ Ill-posed problem

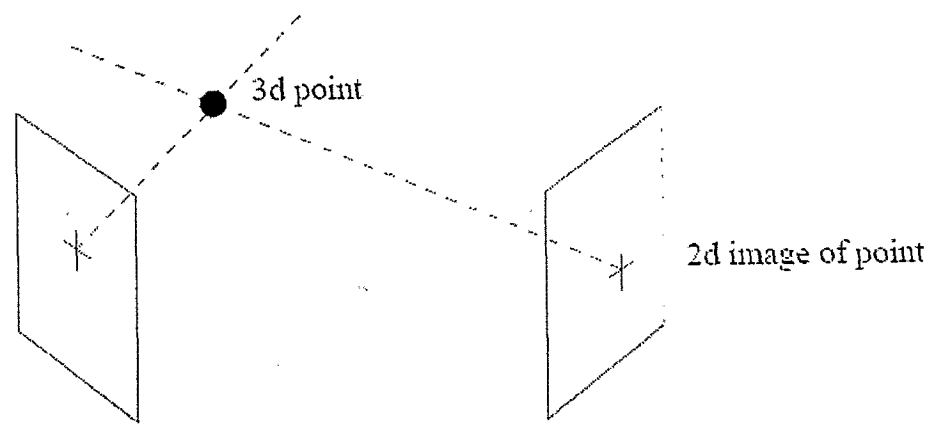
⁸ Camera setup

⁹ Disparity between images

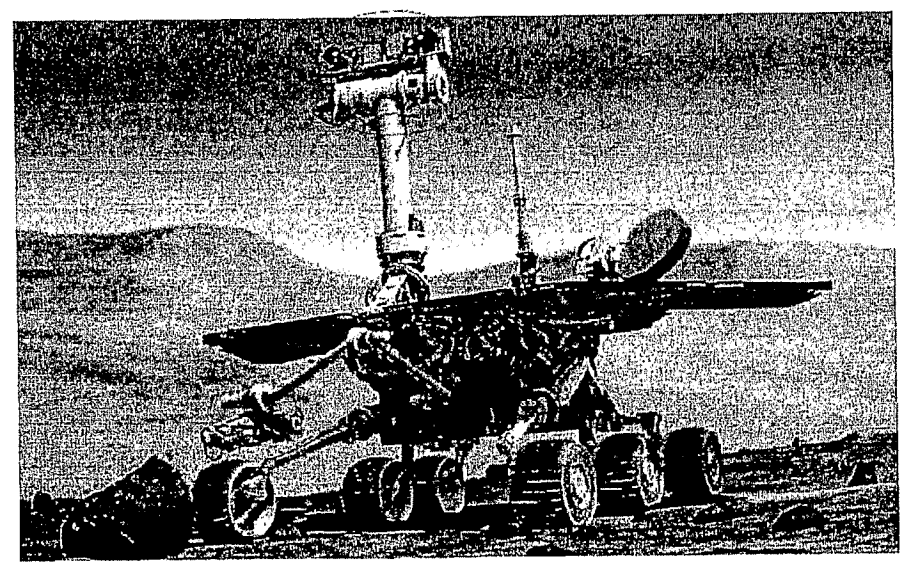
¹⁰ Corresponding matching



شکل ۱-۱: نگاشت نقطه سه‌بعدی به صفحه‌ی تصویر



شکل ۱-۲: چگونگی درک عمق توسط دو تصویر



شکل ۱-۳: مریخ نورد با بینایی استریو پسیو

نتیجه آنکه، یک تصویر منفرد نمی‌تواند مشخصات بعد سوم را ارائه دهد. یکی دیگر از روش‌های حل این مشکل استفاده از ابزار اضافی جهت تخمین عمق می‌باشد. در دید پسیو علاوه بر محدودیت وجود نور حداقل، اگر اختلاف دو تصویر چه از نظر جابه‌جایی و چه از نظر چرخشی کم باشد می‌توان به جواب مطلوبی رسید ولی اگر این اختلاف زیاد باشد عمدتاً حل مساله تطابق دچار مشکل می‌گردد. بنابراین ساخت منظره سه‌بعدی دچار مشکل می‌گردد. به منظور غلبه بر این موضوع، در سال‌های اخیر متدهای بینایی اکتیو عرضه شده‌است. در روش‌های بینایی اکتیو عموماً مساله تطبیق وجود ندارد و می‌توانند با دقت بالایی عمق را تخمین بزنند.

در بینایی اکتیو معمولاً، یک منبع خارجی جهت ارسال و انتشار پرتوهای نوری استفاده می‌گردد. این پرتوها پس از برخورد به سطح موردنظر و انعکاس توسط دوربین‌هایی ثبت می‌شوند. از این پرتوها می‌توان به لیزر و LCD اشاره کرد. در واقع تفاوت اصلی این روش‌ها با روش پسیو، در اینست که در تصویربرداری سیستم‌های پسیو از نور محیط استفاده می‌شود و در سیستم‌های اکتیو از نور مصنوعی ساطع‌شده از منبع شناخته‌شده‌ای استفاده می‌شود.

در روش‌های اکتیو مانند متدهای مبتنی بر لیزر^۱، بینایی و یا تصویر تهیه‌شده به تشعشعات پس‌زمینه و بافت سطوح^۲ بستگی ندارد، در حالیکه یکی از معایب بینایی پسیو را می‌توان همین مورد بیان کرد.

در عمل نوع مناسبی از نور و یا فرم دیگری از انرژی در مسیرهایی به سمت جسم^۳ ساطع می‌شوند، که توسط سطح منعکس و در نهایت توسط گیرنده‌های حسگر^۴ دریافت می‌شوند. فاصله از سطح، توسط روش‌های سه‌گوشه-سازی^۵ و یا زمان پرواز^۶ محاسبه می‌گردد. شکل (۱-۴) نمونه‌ای از اسکن تک نقطه‌ای و روابط حاکم بر این ساختار را نشان می‌دهد.

در متدهای مبتنی بر لیزر^۷، فاصله مستقیماً توسط اندازه‌گیری میزان زمان تاخیری نور برگشتی از نور ساطع‌شده و دانستن سرعت نور بدست می‌آید. این روش عمدتاً دارای دقت بالایی در تخمین فاصله و عمق است. اگر تعداد نمونه-های لیزر ساطع‌شده در هر بار تخمین را بیشتر کنند، به دقت بالاتری دست می‌یابند. تخمین نهایی توسط متوسط-

¹ Laser-based

² Surface Textures

³ Object

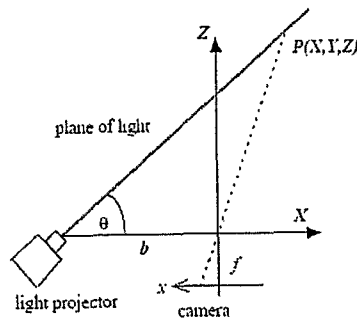
⁴ Sensors

⁵ Triangulation-based methods

⁶ Time of flight

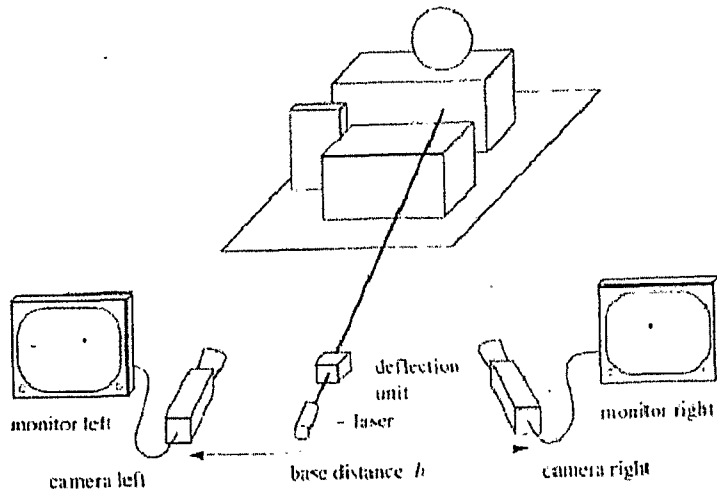
⁷ Laser-based range finder

گیری، از فواصل تخمین زده شده توسط هر نمونه پرتو محاسبه می‌گردد. اساساً روش‌های مبتنی بر لیزر، از دو شکل موج استفاده می‌کنند: موج پالسی و موج پیوسته. در موج پیوسته می‌توان از انواع سیگنال‌های مدولاسیون‌های مختلف مانند مدولاسیون‌های دامنه^۱ و فرکانس^۲ بهره گرفت. از نظر چگونگی جاروب^۳ جسم می‌توان چندین روش به کار برد. به‌عنوان مثال روش جاروب نقطه‌ای در شکل (۵-۱) دیده می‌شود و در آن تنها از یک پرتو لیزر استفاده می‌شود.



$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \frac{b}{f \cot \theta - x} \begin{bmatrix} x \\ y \\ f \end{bmatrix}$$

شکل ۱-۴: مثالی از بینایی اکتیو



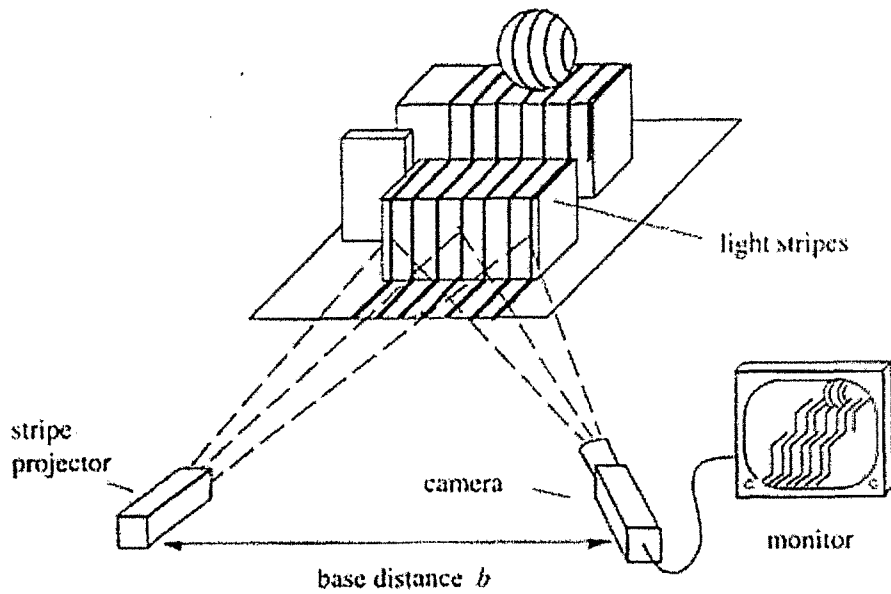
شکل ۱-۵: روش اسکن تک نقطه‌ای اکتیو

¹ AM

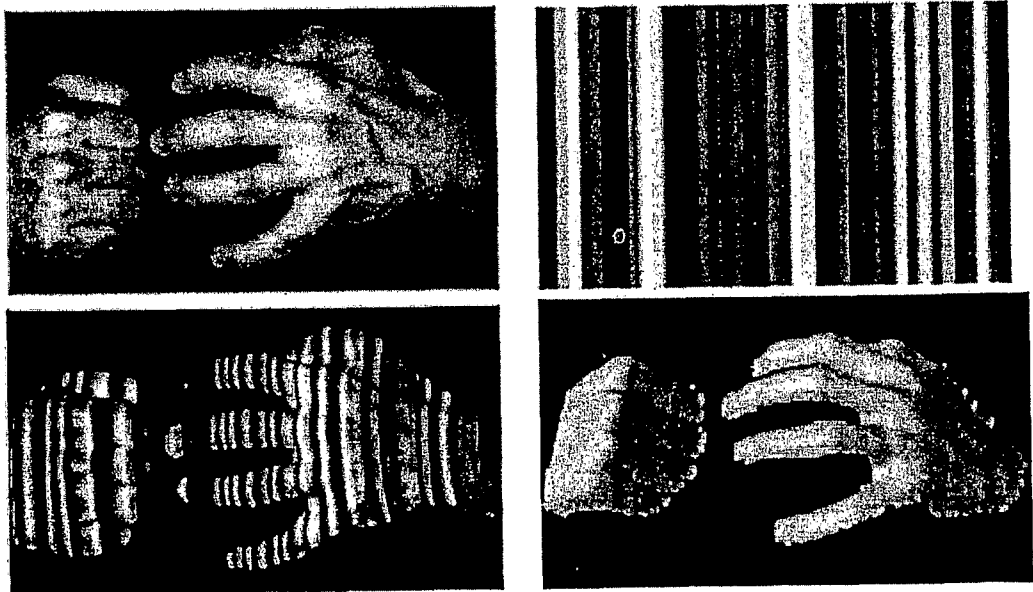
² FM

³ Scan

روش‌های جاروب پیچیده‌تری نیز مانند استفاده از پرتوهای باریکه‌ای کدشده وجود دارند. اشکال (۶-۱) و (۱-۱) نمونه‌ای از این روش‌ها را نشان می‌دهند. هر باریکه، دارای رنگ خاص و در نتیجه سرعت موج متفاوت از باریکه دیگر است. استفاده از این روش‌ها سبب تسهیل امر محاسبه عمق و سرعت بخشیدن به این فرآیند می‌گردد. شکل (۷-۱) مثالی عملی از پوشش باریکه‌ای را نشان می‌دهد. در ادامه تعریف مختصری از ساختار دوربین‌ها و بدست آوردن ماتریس نگاشت ارائه می‌گردد.



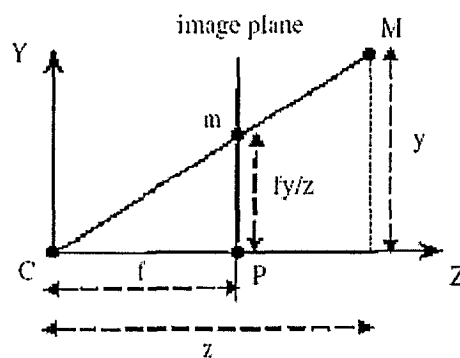
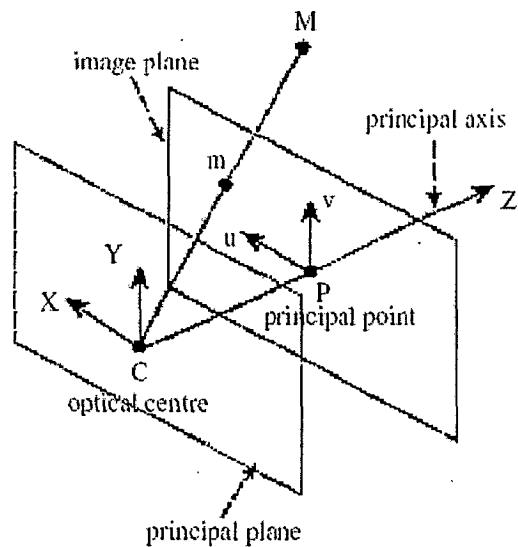
شکل ۶-۱: مثالی از روش بینایی اکتیو یا اسکن باریکه‌ای



شکل ۷-۱: مثالی عملی از روش بینایی اکتیو یا اسکن باریکه‌ای

۱-۱- ساختار هندسی دوربین پین هول^۱

این مدل یکی از پرکاربردترین مدل‌های تصویربرداری است. شکل (۸-۱) نمایی از این سیستم را نشان می‌دهد. این مدل توسط مرکز نوری C^2 (که به آن مرکز پروژکتور دوربین^۳ نیز گفته می‌شود) و صفحه تصویر^۴ بیان می‌شود. فاصله بین صفحه تصویر و مرکز C را فاصله کانونی f می‌نامند. صفحه موازی، صفحه تصویر که شامل مرکز C هست را صفحه اصلی^۵ یا صفحه کانونی^۶ دوربین می‌نامند. خطی که مرکز دوربین C را به مرکز صفحه‌ی تصویر متصل می‌کند را راستای اصلی^۷ یا راستای نوری^۸ می‌نامند. چگونگی رابطه بین مختصات سه‌بعدی یک منظره و مختصات قرارگرفته در تصویر توسط سیستم پرسپکتیو^۹ بیان می‌گردد.



شکل ۸-۱: نمایی از ساختار تصویربرداری پین هول

¹ Pinhole geometry

² Optical centre

³ Camera projection centre

⁴ Image Plane

⁵ Principal Plane

⁶ Focal plane

⁷ Principal axis

⁸ Optical axis

⁹ Central or Perspective projection

براساس قضیه تشابه مثلثات [2] و شکل (۸-۱) هر نقطه سه بعدی $(x, y, z)^T$ به نقطه $(\frac{f \cdot x}{z}, \frac{f \cdot y}{z}, f)^T$ نگاشت^۱

می‌شود. این رابطه را می‌توان به شکل ضرب ماتریسی زیر بیان کرد:

$$\begin{pmatrix} f \cdot x \\ f \cdot y \\ z \end{pmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (۱-۱)$$

رابطه (۱-۱) را به شکل زیر و براساس ماتریس پروجکشن و یا نگاشت دوربین^۲ P بیان می‌کنند:

$$zm = PM \quad (۲-۱)$$

که در آن $M = (x, y, z, 1)^T$ مختصات همگن شده^۳ی نقطه 3D و $m = (\frac{f \cdot x}{z}, \frac{f \cdot y}{z}, 1)^T$ مختصات همگن شده^۴ی نقطه متناظر در تصویر است.

توجه: در مختصات نرمال^۴، هر پیکسل در صفحه‌ی تصویر توسط دو مولفه‌ی عمودی y و افقی x ، مانند $m = (x, y)^T$ و هر نقطه در فضا با سه مولفه مانند $(x, y, z)^T$ بیان می‌شود. ولی در مختصات همگن شده هر پیکسل در تصویر را توسط سه مولفه مانند $m = (a, b, w)^T$ و هر نقطه در فضا را با ۴ مولفه مانند، $M = (a, b, c, w)^T$ بیان می‌کنند. برای تبدیل مختصات همگن به نرمال روابط زیر وجود دارد: $x = \frac{a}{w}$ ، $y = \frac{b}{w}$ ، $z = \frac{c}{w}$. بیان نقاط به فرم همگن شده مزایایی دارد. از جمله توانایی بیان نقاط ناپدیدشونده^۵. مانند $[1, 0, 0]^T$ و $[0, 1, 0]^T$ (در فرم همگن نوشته شده‌اند) که به ترتیب نقاطی هستند در فضا که بی‌نهایت از محور x و y فاصله دارند.

ماتریس نگاشتی که در معادله (۱-۱) بیان شد، ساده‌ترین حالت از ماتریس نگاشت است و تنها شامل مشخصات فاصله کانونی است. در حالت کلی، ماتریس P یک ماتریس تمام مرتبه^۶ با اندازه 3×4 است و به فرم‌های دیگری که هر کدام ویژگی جدیدی را اضافه می‌کنند تعریف می‌شوند [3]:

¹ Mapping

² Camera projection matrix

³ Homogeneous coordinate

⁴ Normal coordinate

⁵ Vanishing points

⁶ Full rank

۱-۱-۱- تنظیم رابطه ماتریس نگاشت برای در برگرفتن شیفت مرکز

صفحه‌ی تصویر

در رابطه (۱-۱) فرض شده‌بود که مبدا صفحه ، مرکز تصویر است. ولی در مدل‌های دیگر می‌توان مبدا را نقطه پایین سمت چپ تصویر یا بالا سمت چپ تصویر و در حالت کلی هر جای تصویر انتخاب کرد. در این حالت هر نقطه $M = (x, y, z, 1)^T$ به نقطه‌ی $m = (\frac{f \cdot x}{z} + p_x, \frac{f \cdot y}{z} + p_y, 1)^T$ در صفحه تصویر منتقل می‌گردد. به فرم ماتریسی می‌توان به شکل رابطه (۳-۱) نوشت:

$$\begin{bmatrix} f \cdot x + p_x \cdot z \\ f \cdot y + p_y \cdot z \\ z \end{bmatrix} = \begin{bmatrix} f & 0 & p_x & 0 \\ 0 & f & p_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & p_x \\ 0 & f & p_y \\ 0 & 0 & 1 \end{bmatrix} [I | 0] \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (3-1)$$

می‌توان ماتریس P در رابطه (۳-۱) را به شکل زیر نوشت:

$$P = K [I | 0] \quad (4-1)$$

که در آن K ماتریس کالیبراسیون دوربین^۱ است و بیانگر چگونگی ارتباط مختصات دوربین و مختصات پیکسل‌ها در تصویر است. این ماتریس نقشی حیاتی در مسائل سه‌بعدی سازی بازی می‌کند و اگر این ماتریس دوربین را بدانیم می‌گوییم دوربین کالیبره^۲ است و در غیر اینصورت کالیبره نیست^۳.

۱-۲-۱- تنظیم رابطه ماتریس نگاشت برای درنظر گرفتن تغییرات

مختصات^۴

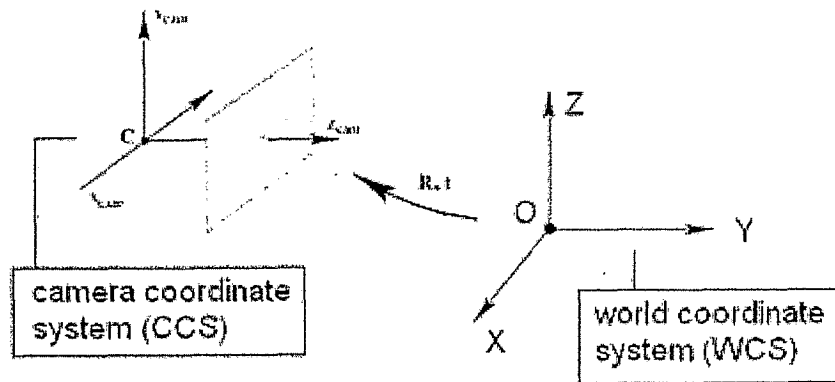
در روابط قبلی که برای ماتریس نگاشت نوشتیم، فرض شده بود که دستگاه مختصات دوربین و دنیای واقعی یکی است. حال اگر مانند تصویر زیر دوربین دارای چرخش یا شیفتی نسبت به مختصات واقعی باشد، آنگاه رابطه ماتریس نگاشت به شکل رابطه (۵-۱) تعمیم می‌یابد:

¹ Camera calibration matrix

² Calibrated

³ Uncalibrated

⁴ Coordinate change



شکل ۹-۱: نمایش چرخش و شیفت صفحه‌ی دوربین

(۵-۱)

مقادیری که در ماتریس K قرار دارند مربوط به پارامترهای ذاتی^۱ دوربین هستند ولی مقادیر موجود در ماتریس-های جابه‌جایی t و ماتریس چرخش R مربوط به پارامترهای برونی^۲ دوربین است. ماتریس K و R دارای سایز ۳×۳ و ماتریس t دارای سایز ۳×۱ است.

۳-۱-۱- تنظیم رابطه ماتریس نگاشت برای درنظر گرفتن شمار پیکسل‌ها^۳

در دوربین‌های دیجیتالی جدید، عموماً تعداد پیکسل‌ها در راستای x و y با هم برابر نیست. این مساله سبب می‌گردد که تعداد پیکسل‌های موجود در یک واحد فاصله در راستای افقی و عمودی با هم برابر نباشند. برای درنظر گرفتن این موضوع باید ماتریس کالیبراسیون (۴-۱) که دربرگیرنده مشخصات ذاتی دوربین است را مطابق رابطه (۶-۱) تغییر داد.

$$k = \begin{bmatrix} \alpha_x & 0 & x_0 \\ 0 & \alpha_y & y_0 \\ 0 & 0 & 1 \end{bmatrix} \tag{۶-۱}$$

در رابطه (۶-۱)، α_x و α_y فواصل کانونی دوربین در راستاهای x و y هستند. x_0 و y_0 مرکز صفحه تصویر با شمارش تعداد پیکسل‌هاست.

¹ Intrinsic

² Extrinsic

³ Pixel units