



## استخراج نقشه عمق از تصاویر استریو بر مبنای الگوریتم تطبیق بلوکی

زهرا فرجی مهبیاری<sup>۱</sup>، شاهین رفیعی<sup>۲\*</sup>، خدیجه فرجی مهبیاری<sup>۳</sup>

۱- دانشجوی دوره کارشناسی ارشد رشته مهندسی مکانیک بیوسیسستم، دانشگاه تهران- پردیس کشاورزی و منابع طبیعی؛

۲- عضو هیئت علمی، دانشگاه تهران- پردیس کشاورزی و منابع طبیعی؛ shahinrafiee@ut.ac.ir

۳- کارشناس ارشد رشته مهندسی مکانیک ماشین‌های کشاورزی، دانشگاه تهران- پردیس کشاورزی و منابع طبیعی.

### چکیده

با گسترش روزافزون کاربرد تصاویر سه بعدی به منظور مکان‌یابی از حسگرهای مختلفی استفاده می‌شود تا بتوان اطلاعات بصری را با کیفیتی بالا شبیه‌سازی نمود. بینایی استریو، یکی از حسگرهای بینایی است که با داشتن حداقل دو تصویر از یک صحنه و بر مبنای تطابق استریو به استخراج عمق می‌پردازد. در این روش برای یافتن مختصات سه بعدی یک نقطه در فضا با حداقل دو دوربین، باید نقطه مورد نظر را در یک تصویر مشخص کرد و متناظر این نقطه را در تصویر دیگر یافت. توسط این تناظر و خط مبنای بین دو دوربین می‌توان مختصات سه بعدی را به دست آورد. یافتن همه نقاط متناظر هزینه محاسباتی بالایی به سیستم تحمیل می‌کند، اما با داشتن دانش اولیه از هندسه محیط و جسم مورد مطالعه می‌توان پیچیدگی محاسباتی را کاهش داد. در این راستا مهم‌ترین کار، افزایش سرعت اجرا و دقت و نیز کاهش پیچیدگی محاسباتی است. بنابراین از تصاویر استریوی رنگی برای افزایش دقت و از محدودیت خط اپیپولار برای کاهش فضای جستجو استفاده می‌شود. این مقاله به معرفی هندسه سیستم بینایی استریو و روابط حاکم بر آن و نیز بازسازی سه بعدی تصاویر بر مبنای الگوریتم تطابق بلوکی می‌پردازد.

**واژه‌های کلیدی:** الگوریتم تطابق بلوک، بینایی استریو، تطابق استریو، محدودیت خط اپیپولار و نقاط متناظر.

### مقدمه

پیشرفت فنی و تخصصی رشته‌هایی از قبیل رباتیک کشاورزی و کشاورزی دقیق نیازمند آگاهی از اطلاعات فضایی خاصی می‌باشد. بر این اساس باید روشی اتخاذ شود که بتواند هر سه مؤلفه نقطه در مختصات جهانی را به وضوح بیان کند. در دو دهه گذشته، بینایی کامپیوتر و تکنیک‌های پردازش تصویر پیشرفت‌های چشم‌گیری داشته‌اند، اما به طور کلی دید سه بعدی نسبت به تصاویر معمولی دو بعدی نمایش دقیق‌تری از واقعیت ایجاد می‌کند (Rovira-Más *et al.*, 2008). درک خصوصیات سه بعدی یک منظره واقعی و یافتن ویژگی‌های اجسام موجود در آن از مهم‌ترین زمینه‌های تحقیقاتی در بینایی ماشین و کامپیوتر می‌باشد که توسط سیستم بینایی استریو انجام می‌شود. بینایی استریو، به معنای توانایی استخراج عمق بر اساس تحلیل دو تصویر گرفته شده از زوایای مختلف یک صحنه می‌باشد (جوادی، ۱۳۸۵) که از سیستم بینایی انسان الهام گرفته شده است؛ در سیستم بینایی انسان،



وجود دو چشم و همچنین انحراف و عدم تطابق<sup>۱</sup> اشیاء در نمای سه بعدی دوچشمی در مقایسه با دو نمای یک چشمی است که منجر به درک عمق می‌شود (Bhatti, 2012).

امروزه، نیاز به نقشه‌برداری سریع و در عین حال دقیق از مزارع موجب جایگزینی نقشه زمینی ربات کاوشگر (با وضوح بالا) به جای نقشه‌های ماهواره‌ای هواپیماهای بدون سرنشین (با وضوح کم و غیر قابل کنترل از راه دور) شده است (Bhatti, 2011). در این راستا از یک دوربین استریوی دو چشمی مستقر بر یک هلیکوپتر و یا یک هلیکوپتر کنترل از راه دور (Rovira-Más *et al.*, 2005) استفاده شد. با این حال، راه حل جالب‌تر و راحت‌تر قرار دادن موتور نقشه‌برداری استریو بر روی ماشین‌های کشاورزی می‌باشد که به طور هم‌زمان با انجام عملیات زراعی به کشاورزان اجازه نقشه‌برداری نیز می‌دهد (Rovira-Más *et al.*, 2008). همچنین با توجه به اهمیت بازسازی سطوح با کیفیت بالا در کاربردهای مختلف رباتیک، به خصوص هدایت مستقل ربات در محیط‌های ناهموار و طبیعی، از سیستم بینایی استریو با هندسه معین و پردازشگر مستقر بر ربات سیار استفاده شده است (Correal *et al.*, 2013).

علاوه بر این، کاهش نیروی کار در مناطق روستایی و نیاز به بهبود بهره‌وری، نیز موجب جایگزینی نیروی انسانی با تراکتورهای خودکار شده است. با این حال، تراکتور خودکار نیازمند دستگاه‌هایی برای اطمینان از ایمنی در حین انجام عملیات بدون کنترل و نظارت انسان می‌باشد. شناسایی موانع اطراف دستگاه خودکار و تعیین موقعیت آن‌ها، دو موضوع اساسی در توسعه دستگاه‌های ایمنی است. بنابراین، مشکل ایمنی یک زمینه تحقیقاتی مهم برای دستگاه‌های خودکار کشاورزی است که در مزرعه استفاده می‌شوند (Yang and Noguchi, 2012). در این راستا استفاده از بینایی استریوی تمام جهته به کمک یک دوربین استریوی دوار (Lin *et al.*, 2008) و یا دو دوربین تمام جهته کاتادیوپتريک<sup>۲</sup> (Lima *et al.*, 2001; Andreasson *et al.*, 2007) و نیز استفاده از رادار و بینایی استریو (Wu *et al.*, 2009; Fang *et al.*, 2002) و یا یک دوربین رنگی مبتنی بر تکنیک‌های ثبت تصویر، که در آن مانع از طریق مقایسه تصویر لحظه‌ای با تصویری از پایگاه داده شناسایی می‌شود، (Morales *et al.*, 2011) توسعه یافت.

با وجود مطالعات صورت گرفته در زمینه بینایی استریو در دهه اخیر، عملکرد سیستم‌های بینایی استریو با سیستم بینایی انسان قابل قیاس نیست. عمق استریو با برآورد مطابقت‌های پیکسلی ویژگی‌های مشابه در نمای سه بعدی استریوی حاصل از صحنه سه بعدی به دست می‌آید. با این حال، پیدا کردن نقاط متناظر صحیح هنوز هم به عنوان یک مسئله دشوار مطرح است. در این مقاله به معرفی هندسه سیستم بینایی استریو و روابط حاکم بر آن و نیز بازسازی سه بعدی تصاویر بر مبنای الگوریتم تطابق بلوکی پرداخته شده است.

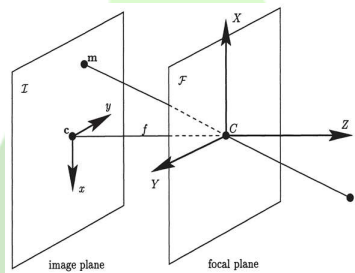
<sup>1</sup> Offset

<sup>2</sup> Catadioptric



## مواد و روش‌ها

در حالت کامل، یک دوربین را می‌توان به صورت خطی با یک حفره کوچک<sup>۳</sup> مدلسازی نمود. که در آن اشعه‌های نوری ساطع یا منعکس شده از هدف با عبور از حفره، تصویر وارونه‌ای را بر روی صفحه تصویر تشکیل می‌دهند. هر نقطه از هدف، نقطه تصویری متناظر آن و حفره، یک خط را تشکیل می‌دهند. این نوع تصویربرداری از فضای سه بعدی بر یک صفحه، تصویربرداری پرسپکتیو<sup>۴</sup> نامیده می‌شود. بنابراین یکی از ابعاد هر نقطه از فضا در هنگام نگاشتن آن بر صفحه دو بعدی از دست می‌رود. برای بازیابی این بعد از دست رفته، تکنیک‌های بازیابی بعد سوم از تصاویر دو بعدی مطرح می‌شوند که یکی از مهم‌ترین آن‌ها روش تصویربرداری استریو است. در این روش، توسط دو دوربین از یک صحنه معین دو تصویر فراهم می‌شود و با استفاده از پارامترهای دوربین‌ها و روابط حاکم بر مدل هندسی تصاویر استریو، هر سه بعد نقاط قابل محاسبه می‌باشد (شریف‌زاده و همکاران، ۱۳۸۶).



شکل ۱- مدل دوربین حفره‌ای و تبدیل پرسپکتیو

واضح است که رابطه بین مختصات دو بعدی تصویر و مختصات فضای سه بعدی به صورت رابطه ۱ نوشته می‌شود که در آن  $f$ ، فاصله کانونی لنز دوربین است:

$$\frac{x}{X} = \frac{y}{Y} = \frac{f}{Z} \quad (1)$$

### هندسه سیستم استریو

شکل ۲ نحوه قرارگیری دوربین‌ها را در سیستم استریو نشان می‌دهد. که در آن  $O$  و  $O_r$  موقعیت دوربین‌های استریو در صحنه،  $b$ ، فاصله بین آن‌ها (خط مبنا<sup>۵</sup>) و  $P_l$  و  $P_r$ ، افکنش<sup>۶</sup>  $P$  (در صحنه سه‌بعدی) روی تصاویر تشکیل شده در دوربین‌ها می‌باشند. از شکل ۲ می‌توان نتیجه گرفت:

$$x_{pl} = \frac{f}{Z} X \quad \text{و} \quad x_{pr} = \frac{f}{Z} (X - b) \quad (2)$$

$$d = x_{pl} - x_{pr} = \frac{f}{Z} b \quad (3)$$

<sup>3</sup> Pinhole

<sup>4</sup> Perspective

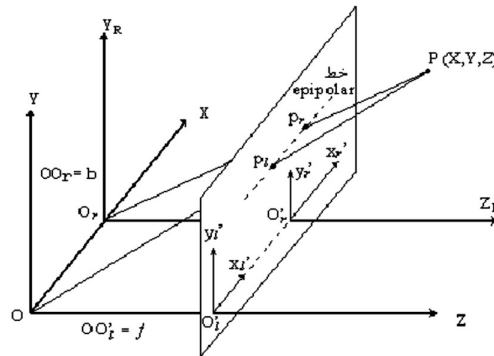
<sup>5</sup> Baseline

<sup>6</sup> Projection

مقدار  $d$ ، اختلاف مکانی<sup>۷</sup> (یعنی اختلاف محل تشکیل تصویر یک نقطه در صفحه تصاویر) نامیده می‌شود. بدین ترتیب عمق هر نقطه با توجه به مقدار اختلاف مکانی آن به دست می‌آید:

$$Z = \frac{f}{d} b \quad (۴)$$

رابطه ۴ حاکی از آن است که عمق نقطه با معکوس اختلاف مکانی متناسب می‌باشد (صادقی و همکاران، ۱۳۸۶).



شکل ۲- هندسه سیستم استریو

## کالیبراسیون دوربین و استریو

کالیبراسیون فرآیند تعیین پارامترهای درونی<sup>۸</sup> و بیرونی<sup>۹</sup> سیستم تصویربرداری است (برادران شکوهی و علیپور، ۱۳۸۵) و از مهم‌ترین روش‌های انجام آن، می‌توان به کالیبراسیون فتومتریک<sup>۱۰</sup> اشاره کرد که در آن از یک جسم سه بعدی با شکل هندسی مشخص که مختصات دقیق آن در دست می‌باشد، استفاده می‌گردد (ابراهیم نژاد صدیق و قاسمیان یزدی، ۱۳۸۴).

برای کالیبراسیون استریو باید حالتی را در نظر گرفت که محورهای نوری دوربین‌ها در بینهایت یکدیگر را قطع کنند، بنابراین با یکسوسازی شرایطی فراهم می‌شود تا دوربین‌ها در آرایش موازی قرار گیرند. برای هر نقطه  $P$  در فضای سه بعدی، هر یک از دوربین‌ها به طور جداگانه کالیبره می‌شوند.

$$P_1 = R_1 P + T_1 \quad \text{و} \quad P_2 = R_2 P + T_2 \quad (۵)$$

$$R_2^T (P_2 - T_2) = R_1^T (P_1 - T_1) \quad (۶)$$

با کالیبراسیون هر دوربین به طور جداگانه ماتریس‌های  $R_1$  و  $R_2$ ،  $T_1$ ،  $T_2$  و سپس ماتریس‌های انتقال و دوران  $R$  و  $T$  به دست می‌آیند.

$$R = R_2 (R_1)^T \quad (۷)$$

$$T = T_2 - R T_1 \quad (۸)$$

<sup>7</sup> Disparity

<sup>8</sup> Intrinsic Parameters

<sup>9</sup> Extrinsic Parameters

<sup>10</sup> Photometric



## استخراج ویژگی و نقاط متناظر در تصاویر استریو

نقاط ویژگی در یک تصویر مانند لبه‌ها و گوشه‌ها از جمله نقاطی هستند که دارای درجه تباین بالایی می‌باشند. به عبارت دیگر گرادین شدت نور یا رنگ در این نقاط از مقدار بالایی برخوردار است. با استفاده از تطبیق استریو می‌توان عمق نقاط ویژگی را با دقت بالایی استخراج نمود. علاوه بر این، در تطبیق نقاط ویژگی نیازی به توازن دقیق شدت نور و رنگ در دوربین‌های مختلف نمی‌باشد.

از معروف‌ترین روش‌های تشخیص لبه می‌توان به پنجره‌های آشکارساز سوبل<sup>۱۱</sup>، کنی<sup>۱۲</sup> و لاپلاسیان<sup>۱۳</sup> اشاره نمود. در دهه اخیر برای تشخیص گوشه روش فضای مقیاس انحنا<sup>۱۴</sup> ارائه شده است. در این روش، مکان کانتورهای مرزی با حداکثر قدر مطلق انحنا، مشخص و به منظور بهبود محلی‌سازی با مقیاس‌های چندگانه ردیابی می‌شوند. لازم به ذکر است که روش نام‌برده در برابر نویز تصویر، مقاوم می‌باشد.

پس از استخراج نقاط ویژگی در تصاویر دوربین‌ها، فرایند تطبیق برای نقاط ویژگی صورت می‌گیرد. هدف از این فرآیند، استخراج مختصات دقیق سه بعدی نقاط ویژگی می‌باشد. به منظور کاهش فضای جستجو و افزایش سرعت در فرآیند تطبیق، از محدودیت خط اپیپولار<sup>۱۵</sup> استفاده می‌شود به این معنی که با یکسوسازی تصاویر چپ و راست می‌توان جستجو در تصویر سمت راستی را از دو بعد به یک بعد کاهش داد و در واقع برای یافتن متناظر یک نقطه ویژه، فقط در خط اپیپولار متناظر با آن در تصویر سمت راستی جستجو از جام داد. ضلع استفاده از تصاویر رنگی به جای تصاویر سطوح خاکستری در این فرآیند، تأثیر نویز و ابهام را کاهش و اطلاعات مفیدتری را نسبت به تصاویر سطوح خاکستری ارائه می‌دهد (صادقی و همکاران، ۱۳۸۶).

هندسه اپیپولار بر هر سیستم دوربینی حاکم است. شکل ۳ وضعیت دوربین‌ها را نشان می‌دهد که در آن  $I$  و  $I'$ ، صفحه تصویر دوربین‌های چپ و راست و  $C$  و  $C'$ ، مرکز نوری آن‌ها می‌باشد. همچنین خط  $CC'$ ، خط مبناست که مرکز دوربین‌ها را به هم وصل می‌کند. محل تقاطع خط  $CC'$  با صفحات  $I$  و  $I'$ ، نقاط  $e$  و  $e'$  می‌باشد که به ترتیب اپیپول<sup>۱۶</sup> چپ و راست نامیده می‌شوند. برای هر نقطه دلخواه  $M$  در فضا، یک صفحه اپیپولار متناظر (II) تعریف می‌گردد که از آن نقطه و مرکز دوربین‌ها عبور می‌کند. افکنش  $M$  به دوربین‌ها با  $m$  و  $m'$  نشان داده می‌شود. همچنین محل تقاطع صفحه اپیپولار با  $I$  و  $I'$ ، خطوط اپیپولار متناظر با  $M$  ( $l_m$  و  $l'_m$ ) می‌باشند که از نقاط  $e$  و  $e'$  نیز عبور می‌کنند (Xu G. and Zhang Z., 1996).

<sup>11</sup> Sobel

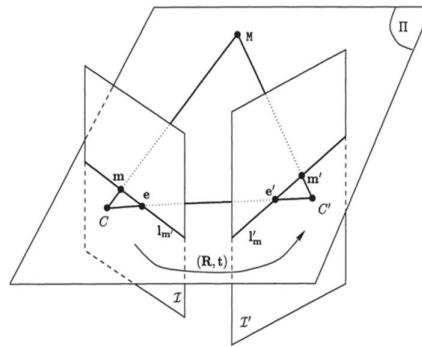
<sup>12</sup> Canny

<sup>13</sup> Laplasian

<sup>14</sup> Curvature Scale Space

<sup>15</sup> Epipolar

<sup>16</sup> Epipole



شکل ۳- هندسه اپیولار

موقعیت دوربین راست نسبت به دوربین چپ توسط ماتریس  $R$  و  $T$  تعیین می‌شود. بنابراین هر نقطه  $(X, Y, Z)$  در سیستم مختصات دوربین چپ با نقطه  $(X', Y', Z')$  در سیستم مختصات دوربین راست متناسب است به طوری که:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = R \begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} + T \quad (9)$$

$$R = \begin{bmatrix} \cos \phi_y \cos \phi_z & \cos \phi_x \sin \phi_z + \sin \phi_x \sin \phi_y \cos \phi_z & \sin \phi_x \sin \phi_z - \cos \phi_x \sin \phi_y \cos \phi_z \\ -\cos \phi_y \sin \phi_z & \cos \phi_x \cos \phi_z + \sin \phi_x \sin \phi_y \sin \phi_z & \sin \phi_x \cos \phi_z + \cos \phi_x \sin \phi_y \sin \phi_z \\ \sin \phi_y & -\sin \phi_x \cos \phi_y & \cos \phi_x \cos \phi_y \end{bmatrix} \quad (10)$$

$$T = \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix} \quad (11)$$

( $\phi_x$  و  $\phi_y$  و  $\phi_z$  زوایای چرخش اولر می‌باشند).

$$RR^T = I \quad \text{و} \quad \det(R) = 1 \quad (12)$$

میان افکنش یک نقطه دلخواه سه بعدی در تصاویر دو دوربین، یک رابطه ماتریسی برقرار است:

$$\tilde{m}^T F \tilde{m}' = 0 \quad (13)$$

رابطه (۱۳) به ازای یک نقطه ورودی  $m$ ، معادله یک خط خواهد بود و به این مفهوم است که تصویر نقطه  $m$  در صفحه تصویر سمت راستی،  $m'$ ، روی این خط (خط اپیولار) قرار خواهد داشت و محل قرار گرفتن آن بستگی به عمق نقطه‌ای دارد که این تصویر را ایجاد کرده است که در آن ماتریس  $F$ ، ماتریس اساسی<sup>۱۷</sup> نامیده می‌شود و مختصات پیکسلی نقاط در یک صفحه تصویر را به دیگری مرتبط می‌کند:

$$F = A^{-T} E A'^{-1} \quad (14)$$

$A$  و  $A'$ ، ماتریس داخلی<sup>۱۸</sup> دوربین‌ها می‌باشند:

<sup>17</sup> Fundamental Matrix

<sup>18</sup> Intrinsic Matrix



$$A = \begin{bmatrix} fk_u & fk_u \cot \theta & u_0 \\ 0 & fk_v / \sin \theta & v_0 \\ 0 & 0 & 1 \end{bmatrix} \quad (16)$$

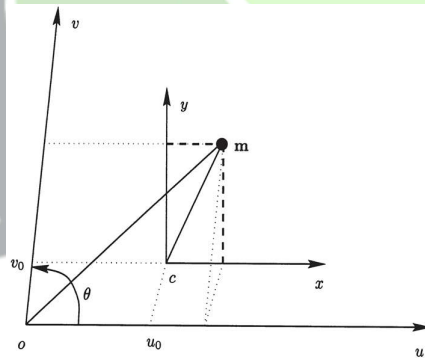
که در آن  $k_u$  و  $k_v$  به ترتیب نسبت واحد محورهای  $u$  و  $v$  در سیستم مختصات پیکسلی تصویر،  $(O, u, v)$ ، به محورهای  $x$  و  $y$  در سیستم مختصات تصویر،  $(C, x, y)$ ، می‌باشند (در اکثر دوربین‌های CCD، پیکسل‌ها به صورت مستطیل با نسبت  $\frac{k_v}{k_u}$  تا  $1/1$  می‌باشند). محورهای  $u$  و  $v$  ممکن است عمود بر هم نباشند و زاویه بین آن‌ها با  $\theta$  مشخص شود.  $[u_0, v_0]^T$ ، مختصات نقطه  $c$  در  $(O, u, v)$ ، است. این پنج پارامتر که به موقعیت و جهت دوربین وابسته نیستند، پارامترهای درونی دوربین نامیده می‌شوند.  $E$  (در رابطه ۱۴)، ماتریس ضروری<sup>۱۹</sup> نام دارد که از دوران و انتقال بین دوربین‌ها به دست می‌آید و مختصات نقطه واقعی را در صفحه چپ و راست به دیگری مرتبط می‌کند:

$$E = [t]_x R \quad (16)$$

(نماد  $[t]_x$  برای ضرب خارجی دو بردار، با نمایش ماتریسی، به کار می‌رود.)

برای حالت خاصی که چرخش فقط حول محور نوری دوربین‌ها انجام می‌شود:

$$E = \begin{bmatrix} 0 & 0 & t_y \\ 0 & 0 & -t_x \\ -t_y \cos \theta - t_x \sin \theta & -t_y \sin \theta + t_x \cos \theta & 0 \end{bmatrix} \quad (17)$$



شکل ۴- پارامترهای داخلی دوربین

با توجه به روابط بیان شده می‌توان رابطه بین مختصات سه بعدی یک نقطه و افکنش آن در تصویر را چنین نوشت:

$$s\tilde{m} = A[R \quad t] \begin{bmatrix} M \\ 1 \end{bmatrix} = A(RA'^{-1}\tilde{m}' + t) \quad (18)$$

$$s'\tilde{m}' = A'[I \quad 0] \begin{bmatrix} M \\ 1 \end{bmatrix} = A'(AR)^{-1}(\tilde{m} - At) \quad (19)$$

$s$  و  $s'$  ضرایب عددی هستند. لازم به ذکر است که در یک مجموعه استریو، نقطه اپیپول در تصویر چپ ( $e$ ) از افکنش مرکز دوربین راست در دوربین چپ به دست می‌آید. به همین ترتیب، نقطه  $e'$  نیز افکنش مرکز دوربین چپ در دوربین راست می‌باشد. با فرض این که سیستم مختصات جهانی بر سیستم مختصات دوربین راست منطبق است؛ می‌توان برای این نقاط نیز نوشت:

<sup>19</sup> Essential Matrix





$$s_e \tilde{e} = A[R \quad t] \begin{bmatrix} 0 \\ 1 \end{bmatrix} = At \quad (20)$$

$$s'_e \tilde{e}' = e' [I \quad 0] \begin{bmatrix} C \\ 1 \end{bmatrix} = -A'R^{-1}t \quad (21)$$

با توجه به این که هر خط به وسیله معادله  $ax + by + c = 0$  توصیف می‌شود؛ هر خط را می‌توان با یک بردار،  $l = [a, b, c]^T$  و نقطه‌ای روی آن،  $\tilde{x} = [x, y, 1]^T$ ، تعریف نمود. پس

$$l_{m'} = \tilde{e} \times \tilde{m} = F\tilde{m}' \quad (22)$$

$$l'_m = \tilde{e}' \times \tilde{m}' = F^T \tilde{m} \quad (23)$$

برای یافتن نقطه متناظر  $m$  در صفحه دوربین  $C'$  کافیت عمل جستجو بر روی خط اپیپولار انجام شود.

### الگوریتم تطابق بلوک

الگوریتم تطابق بلوکی روشی برای انجام عمل جستجو روی خط اپیپولار می‌باشد. یک روش ترکیبی متراکم مبتنی بر ظرفیت پنجره که به جای جستجوی شباهت بین پیکسل‌های فردی، تصویر را به پنجره‌ها و یا بلوک‌هایی از پیکسل‌ها تقسیم و آزمون تطابق را اجرا می‌کند. این الگوریتم شامل سه مرحله نرمالیزه کردن شدت تصاویر و بهبود بافت (پس پردازش)، جستجو برای یافتن نقاط متناظر در طول خطوط اپیپولار و حذف تناظرهای اشتباه (پردازش نهایی) می‌باشد. از ویژگی‌های منحصر به فرد این الگوریتم سرعت بالای آن برای تناظریابی و بازسازی سه بعدی تصاویر استریو می‌باشد.

### نتایج و بحث

تصاویر مورد استفاده در این مطالعه که در شکل ۵ نشان داده شده است از وب سایت (<http://www.middlebury.edu/stereo>) به دست آمدند. این تصاویر دیجیتال به صورت تصاویر رنگی با وضوح حداکثر  $350 \times 240$  پیکسل از RGB (قرمز، سبز و آبی) به خاکستری تبدیل شدند. از این تصاویر به عنوان مثالی برای توصیف عملکرد الگوریتم تطابق بلوکی استفاده شده است.



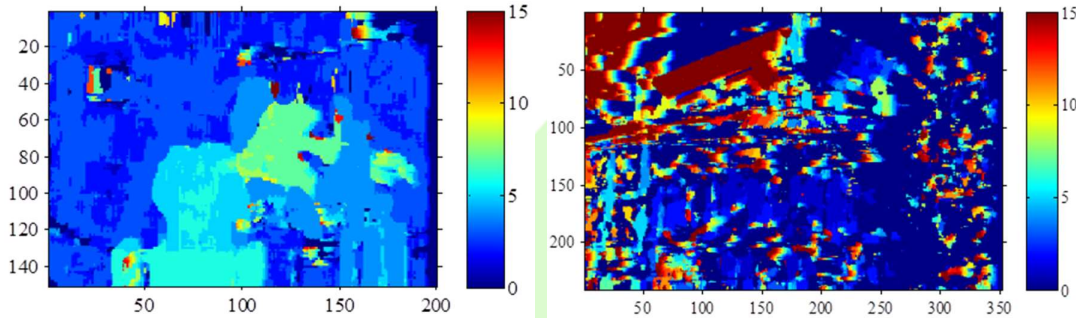
شکل ۵ - جفت تصاویر استریو

به منظور اجرای تطابق بلوکی در هر جفت تصویر، برای هر پیکسل در تصویر سمت راست، بلوک پیکسلی  $7 \times 7$  اطراف آن را استخراج کرده و در امتداد همان سطر در تصویر سمت چپ، بهترین بلوک منطبق بر آن جستجو می‌شود. خروجی محاسبه استریو، نقشه اختلاف مکانی است که فاصله هر نقطه در صحنه فیزیکی را نسبت به دوربین مشخص می‌کند.

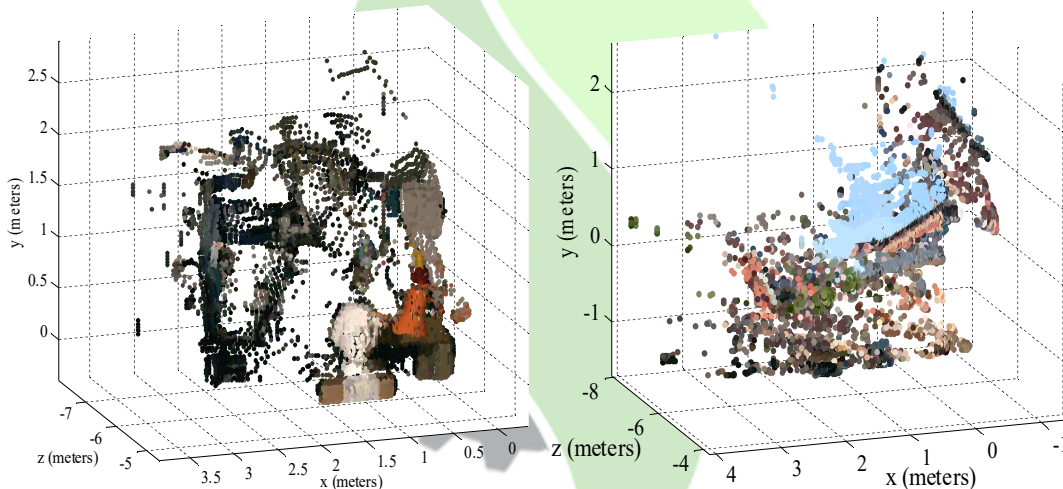




شکل ۶ و ۷ نتایج حاصل از تطابق بلوک می‌باشد که نشان می‌دهد تطابق بلوک به خوبی انجام شده و نقشه عمق به عنوان شکل صحیح صحنه استریو بهبود یافته است. با این حال، تکه‌های نویز و برآورد نادرست عمق در همه جا، به خصوص در مرزها وجود دارد. این‌ها زمانی ایجاد می‌شوند که هیچ ویژگی تصویری قوی در مقایسه پنجره‌های پیکسلی  $7 \times 7$  وجود نداشته باشد. سپس فرایند تطابق در معرض نویز اختلاف مکانی هر پیکسل مستقل از پیکسل‌های دیگر انتخاب می‌شود.



شکل ۶- نقشه عمق حاصل از تطابق بلوک



شکل ۷- بازسازی سه بعدی نقاط

با توجه به نتایج به دست آمده از بازسازی سه بعدی، الگوریتم تطابق بلوک برای تناظریابی در نزدیکی مرزها مناسب نیست؛ چون بخشی از بلوک پیکسلی روی پیش‌زمینه و بخشی روی پس‌زمینه قرار می‌گیرد.

## نتیجه‌گیری کلی

بنیایی استریو یکی از مهم‌ترین و مستقیم‌ترین راه‌ها برای رسیدن به بینایی سه بعدی می‌باشد که از اختلاف دید دو دوربین برای به دست آوردن اطلاعات بعد سوم استفاده می‌کند. از مسائل مهم در این راستا یافتن دقیق نقاط متناظر در تصاویر استریوست. بر این اساس هندسه دوربین و اپیپولار بررسی و معادلات مربوط به خطوط اپیپولار استخراج گردید و عمل جستجو روی آن‌ها با استفاده



از الگوریتم تطبیق بلوکی انجام شد. این الگوریتم استاندارد برای بینایی استریو در سیستم‌های سخت افزاری دارای سرعت بالا می‌باشد. نتایج بازسازی سه بعدی نشان داد این الگوریتم می‌تواند نقاطی را که قویاً مطابق هستند به عنوان نقاط متناظر معرفی کند اما برای تناظریابی در نزدیکی مرزها مناسب نیست.

## منابع

۱. ابراهیم نژاد صدیق ح. و قاسمیان یزدی ح.، ۱۳۸۴. استخراج مدل سه بعدی جسم با ادغام ویژگی‌های مستخرج از تصاویر استریو، تخمین حرکت و تصاویر نیم‌رخ. نشریه مهندسی برق و مهندسی کامپیوتر ایران، سال سوم، شماره ۱، نیمسال اول ۱۳۸۴، ص ۲۸-۱۸.
۲. بردران شکوهی ش. و علیپور م.، ۱۳۸۵. روش جدید تناظریابی اتوماتیک در بینایی استریو با استفاده از شبکه‌های عصبی و همبستگی. چهارمین کنفرانس ماشین بینایی و پردازش تصویر ایران، دانشگاه فردوسی مشهد.
۳. جوادی م.، ۱۳۸۵. سیستم‌های پردازش تصویر سه بعدی به روش Triangulation و بررسی روش‌های موفق موجود. نهمین کنفرانس دانشجویی مهندسی برق ایران، دانشگاه تهران.
۴. شریف‌زاده س.، اردشیر غ. و میارنجمی ح.، ۱۳۸۶. تعیین موقعیت اشیای متحرک با استفاده از تصویربرداری استریو بدون نقاط با مختصات معلوم در صحنه. سیزدهمین کنفرانس ملی انجمن کامپیوتر ایران، جزیره کیش، خلیج فارس، ایران.
۵. صادقی ه.، منجمی ا. و معلم پ.، ۱۳۸۶. روشی مقاوم و سریع در تطبیق استریوی تصاویر رنگی به کمک زنجیره‌های پیوسته. سومین کنفرانس بین‌المللی فناوری اطلاعات و دانش، دانشگاه فردوسی مشهد.
6. Andreasson, H., A. Treptow and T. Duckett. 2007. Self-localization in non-stationary environments using omni-directional vision. *Robotics and Autonomous Systems* 55:541-551.
7. Bhatti A. 2012. Current Advancements in Stereo Vision. PP 11-22 in A. Bhatti ed. Published by InTech. Printed in Croatia.
8. Bhatti A. 2011. Advances in Theory and Applications of Stereo Vision. PP 239-245 in A. Bhatti ed. Published by InTech. Printed in India.
9. Correal R., G. Pajares and J.J. Ruz. 2013. Automatic expert system for 3D terrain reconstruction based on stereovision and histogram matching. *Expert Systems with Applications*.
10. Fang, Y., I. Masaki and B. Horn. 2002. Depth-based target segmentation for intelligent vehicles: fusion of radar and binocular stereo. *IEEE Transactions on Intelligent Transportation Systems* 3:196-202.
11. Lima P., A. Bonarini, C. Machado, F. Marchese, C. Marques, F. Ribeiro and D. Sorrenti. 2001. Omni-directional catadioptric vision for soccer robots. *Robotics and Autonomous Systems*. 36:87-102.
12. Lin T.-T., Y.-K. Hsiung, G.-L. Hong, H.-K. Chang and F.-M. Lu. 2008. Development of a virtual reality GIS using stereo vision. *Computers and Electronics in Agriculture* 63:38-48.
13. Morales, N., J.T. Toledo, L. Acosta and R. Armay. 2011. Real-time adaptive obstacle detection based on an image database. *Computer Vision and Image Understanding* 115:1273-1287.
14. Rovira-Mas F., Zhang Q. and Reid J. F., 2008. Stereo Vision three-dimensional terrain maps for precision agriculture. *Computers and electronics in agriculture* 60:133-143.
15. Rovira-Mas, F., Q. Zhang and J.F. Reid. 2005. Creation of three-dimensional crop maps based on aerial stereo images. *Biosystems Engineering* 90 (3):251-259.
16. Wu, S., S. Decker, P. Chang, T. Camus and J. Eledath. 2009. Collision sensing by stereo vision and radar sensor fusion. *IEEE Transactions on Intelligent Transportation Systems* 10:606-614.
17. Yang L. and N. Noguchi. 2012. Human detection for a robot tractor using omni-directional stereo vision. *Computers and Electronics in Agriculture* 89:116-125.
18. Xu G. and Z. Zhang. 1996. Epipolar geometry in stereo, motion and object recognition a unified approach. PP 3-7. O.D. Faugeras, J. J. Koenderink, S. M. Pizer, S. Tsuji and S. W. Zucker eds. Published by Kluwer Academic Publishers.

## Extraction of the depth map of stereo images based on Block Matching

Zahra Faraji Mahyari<sup>1</sup>, Shahin Rafiee<sup>2\*</sup>, Khadijeh Faraji Mahyari<sup>3</sup>

1- M.Sc. Student, Department of Agricultural Machinery Engineering, University of Tehran

2- Professor, Department of Agricultural Machinery Engineering, University of Tehran,

shahinrafiee@ut.ac.ir

3- M.Sc., Department of Agricultural Machinery Engineering, University of Tehran

### Abstract

With the increasing use of three-dimensional images to locate; the sensors are used in order to simulate the visual information with high quality. Stereo vision, a vision sensor that can measure the depth using at least two images of a scene based on stereo matching. This method to find the coordinates of a point in three-dimensional space, point of interest must be specified in one of the images and then find the corresponding point in the other image. We can estimate the three-dimensional coordinates using the corresponding points and the baseline between two cameras. Find all the points corresponding to the system imposes a high computational cost, but having a basic knowledge of geometry and environment and object can reduce the computational complexity. In this regard, the most important task is to increase the speed and accuracy as well as reducing the computational complexity. Therefore be used to increase the accuracy of color stereo images and to reduce the search space of epipolar line constraint. This paper introduces the geometry of stereo vision system and the relationship its and the reconstruction of three-dimensional images based on block matching algorithm.

**Keywords:** Block matching algorithm, Corresponding points, Epipolar line constraint, stereo matching, and Stereo vision.