بسم الله الرحمن الرحيم

دانشگاه الزهرا (س)

دانشکده فنی و مهندسی

پایان نامه جهت اخذ درجه کارشناسی ارشد

رشته مهندسی کامپیوتر-گرایش هوش مصنوعی

عنوان

ردیابی اشیا

استاد راهنما

دکتر بهروز قلی زاده

استاد مشاور

دکتر رضا عزمی

دانشجو

فاطمه نوروزی هریس

کلیه دستاوردهای این تحقیق متعلق به دانشگاه الزهرا(س)است.

از اساتید و معلمانم برای راهنماییهایشان ،

از پدر عزیز و بزرگوارم برای دل دریایی اش ،

از همسر و دختر مهربانم برای صبوری ، همراهی و بزرگواریشان ،

و ازتمام کسانی که در تمام مراحل حامی و پشتیبانم بوده اند

سپاسگزارم

تقدیم به

همسر و دختر نازنیم که صبوری شان ادامه راه را برایم ممکن ساخت

و

خانواده عزیزم که همواره پشتیبانم بوده اند.

چکیده

در سال‌های گذشته، تعدادی سیستم‌های تجزیه و تحلیل تصویر به‌صورت خودکار، برای تشخیص، تجزیه و تحلیل و ردیابی اشیای متحرک در اماکن و کاربردهای گوناگون، توسعه داده شده‌اند. این موضوع یکی از مسائل مهم و در حال توسعه در پردازش تصویر و بینایی ماشین، است. ردیابی اشیاء، نمایش تغییرات موقعیت یک شیء و دنبال کردن آن در یک دنباله از تصاویر ویدئویی، با هدفی خاص است که باید با دقت مطلوبی انجام شود. اگرچه سابقه ایجاد پدیده ردیابی اشیاء به مسائل نظامی برمی‌گردد ولی امروزه به دلیل کاربردهای بسیار گسترده‌ی ردیابی اشیاء در زمینه‌های مختلف، مثلاً کنترل ترافیک و تشخیص حرکات غیرمعمول، این مقوله و جوانب مختلف آن مورد توجه ویژه‌ای قرار گرفته است. از جمله مسائلی که همواره عملکرد الگوریتم‌های ردیابی را با مشکل مواجه ساخته است، تعامل آنها با روش‌های تشخیص هدف، ظاهر متغیر اهداف و همچنین ردیابی همزمان چند هدف است. در این پژوهش، یک روش با بهره وری بالاتر و دقیق‌تر از روش‌های قبل برای ردیابی اشیاء با استفاده از تصویر برداری با دوربین ثابت معرفی می‌گردد. در این سیستم تصاویر ویدئویی گرفته شده از محل مورد نظر، پردازش شده و اشیای متحرک آن شناسایی و استخراج می‌گردند. الگوریتم پیشنهادی در چندین مرحله عمل می‌کند. ابتدا پس‌زمینه‌ی ثابت را شناسایی و نویز را از آن حذف می‌کند. این پس‌زمینه برای تفریق شدن از اشیای متحرک به‌کار می‌رود. پس از آن طی یک مرحله‌ی فیلتر کردن تصویر، سایه‌ها و نویزهای تصویر فیلمبرداری شده حذف و در نهایت با استفاده از روش مسیریابی حباب [1]شی متحرک یا اشیای متحرک تفکیک، شناسایی و ردیابی می‌شوند. آزمایش این سیستم، بر روی تصاویری از یک فرد در حیاط منزل و جاده‌های شهری انجام شد. نتایج تجربی نشان می‌دهد که مدل پیشنهادی برای تشخیص اشیاء متحرک و ردیابی آنها، به خوبی کار می‌کند و می‌تواند تخمین حرکت و مسیر حرکت اشیاء را از نظر سرعت و دقت تا حد مطلوبی بهبود دهد. این بهبود از نظر دقت نسبت به روش‌های گذشته به 10 درصد نیز می‌رسد.

کلمات کلیدی: ردیابی اشیاء، اشیاء متحرک، تصاویر فیلتر شده، تفریق زمینه، مسیردهی حباب

[1] مرکز تصویر به صورت حباب در نظر می‌گیرد

فهرست مطالب

فهرست اشکال و فرمولها و جداول.....ح

چکیده.....ک

هدف تحقیق.....2

مروری بر کارهای گذشته.....3

فصل دوم: ارائه‌ی مفاهیم و اصطلاحات پایه.....6

ردیابی اشیاء.....6

2-1 ردیابی نقطه‌ای :.....8

2-1-1 روش‌های قطعی برای ارتباط :.....9

2-1-2 روش‌های آماری برای ایجاد ارتباط :.....11

الف- تخمین حالت شی منفرد :.....11

ب- تخمین حالت و داده‌های انجمنی چند شیئی:.....12

2-2 ردیابی هسته.....12

2-2-1ردیابی با استفاده از الگو و مدل‌های ظاهری مبتنی بر تراکم :.....13

الف- ردیابی اشیای جداگانه :.....13

ب- ردیابی اشیای مشترک :.....13

2-2-2ردیابی با استفاده از مدل‌های ظاهری چنددیدی: 14
2-3 ردیابی سایه 14
2-3-1 تطبیق شکل : 15
2-3-2 ردیابی لبه 15
2-3-2-1 ردیابی توسط مدل‌های فضای حالت 16
2-3-2-2 ردیابی با کمینه‌سازی مستقیم تابع انرژی لبه : 16
الگوریتم‌های متنوع در ردیابی شئ 16
ردیابی اشیاء سه‌بعدی 20
فصل سوم: الگوریتم پیشنهادی 24
3-1 بیان کلی مسئله 25
3-2 تشخیص مناطق متحرک 26
3-2-1 روش تفاضل دو تصویر متوالی 26
3-2-2 روش حذف پس زمینه 27
2-2-4 تفریق پس زمینه 28
3-3 ردیابی شی 30
معادله ردیابی 31
3-4-روش‌های تک‌لایه‌ای 35
روش‌های فضا زمان 36

فصل چهارم: ارزیابی روش پیشنهادی 39
4-1 شبیه سازی 40
4-2 کاربرد آمار در شبیه سازی 42
داده ها و رویدادهای تصادفی 42
4-3 نتایج شبیه سازی برای دنبال کردن اشیاء متحرک 43
4-5 نتایج شبیه سازی برای مقایسه با الگوریتم‌های دیگر 54
فصل 5: نتیجه‌گیری و کارهای آینده 56
منابع 66

فهرست اشکال و فرمولها و جداول

نمودار1-نمودار دسته‌بندی ردیاب‌ها 7
شکل 2-1 7
شکل 2-2 9
شکل2-3 10
شکل 3-1 بلوک دیاگرام مربوط به یک سامانه ردیاب 26
فرمول(1) 27
(2) فرمول 27

فرمول(3)..... 28
شکلهای 3-2 تصویر یک فریم که فاقد شی متحرک است و نتیجه تفریق پس زمینه 30
شکل 3-3 : اضافه کردن حباب به تصویر اصلی 30
(4)فرمول..... 31
فرمول(5)..... 31
شکل3-4 مدل مخفی مارکوف به عنوان نماینده شبکه های بیزی با برش زمان. 33
یکی از نمونه های فرم از این سناریو است. 34
شکل3-5 یک عمل انسانی مانند مشت زدن را درفضای سه‌بعدی نشان می‌دهد. 37
محتویات فایل اصلی (main.m): 43
محتویات فایل ردیابی اشیاء (backtrack.m): 43
شکل 4-1. تصویر پس‌زمینه‌ای که در ابتدا با دوربین گرفته می‌شود. 46
شکل 4-2. تصویر فرد تقریباً در ابتدای حرکت خود در ویدئو 47
شکل 4-3. تصویر برده شده به حیطهٔ سیاه و سفید به همراه حباب مرکز آن و چهارضلعی ردیابی. 48
شکل 4-4. تصویر فرد در اواسط حرکت خود در ویدئو. 49
شکل 4-5. تصویر ردیابی شده فرد در اواسط حرکت خود در ویدئو. 49
شکل 4-6. تصویر فرد تقریباً در انتهای مسیر خود در ویدئو. 51
شکل 4-7. تصویر ردیابی شده فرد در انتهای حرکت خود در ویدئو. 51
شکل 4-8 به‌کارگیری الگوریتم برای ردیابی وسایل نقلیه متحرک در یک جاده. 52

شکل 4-9 نتیجه نهایی حاصل از به‌کارگیری الگوریتم برای ردیابی وسایل نقلیه. 53

جدول 4-10. مقایسهٔ سطح دقت روش پیشنهادی در مقابل روش‌های دیگر در شدت کنتراست‌های مختلف 55

و سرعت‌های **مختلف** تصاویر ویدئویی 55

شکل 5-1 شبه سازی تصاویر 56

شکل 5-3 تصویری از محیط داخل با اعمال تنظیمات در جدول 5-2 56

شکل 5-2 تصویری از محیط بیرون با اعمال تنظیمات در جدول 5-2 57

جدول 5-1 نشان دهنده تصویر پردازش شده تا دریافت تصویر خروجی 57

جدول5-2 مربوط به نتایج پردازش بر روی تصاویر نمونه 58

فصل اول: مقدمه

ردیابی شی کار مهمی در حوزه‌ی بینایی ماشین محسوب می‌شود . گسترش روزافزون کامپیوترهای قوی، کیفیت بالا و در دسترس بودن دوربین‌های فیلم‌برداری ارزان قیمت و نیاز روزافزون برای تجزیه و تحلیل خودکار فیلم‌ها باعث افزایش کاربرد ردیابی اشیاء گشته است . سه گام کلیدی در تجزیه و تحلیل فیلم‌های ویدیویی وجود دارد :

1- شناسایی شی متحرک مورد نظر
2- ردیابی اشیا فریم به فریم
3- تجزیه و تحلیل حرکت شی برای تشخیص رفتار آن

ردیابی می‌تواند شامل تعریف مسیر و حرکت یک شی در آن مسیر باشد . به عبارت دیگر ردیابی یعنی برچسب زدن به شی مورد نظر در فریم‌های مختلف تصویر. علاوه بر این، ردیابی می‌تواند اطلاعات دیگری را نیز مانند جهت‌گیری ، مساحت یا شکل جسم در اختیار قرار دهد. ردیابی اشیاء می‌تواند به دلایل زیر پیچیده باشد :

1- از دست رفتن بخشی از اطلاعات دنیای واقعی سه بعدی [2]در زمان تبدیل به تصویر دو بعدی
2- نویز [3]در تصاویر
3- حرکات پیچیده‌ی شی
4- ماهیت غیر شبکه‌ای یا بند بند بودن شی
5- شکل پیچیده‌ی شی
6- ناقص یا مسدود بودن جسم

[2] 3D
[3] Noise

7- تغییرات در نور صحنه

8-احتیاج به پردازش بلادرنگ[4]

لذا در ردیابی جسم می‌توان در ابعادی مثل حرکت یا شکل جسم، محدودیت‌هایی قائل شد. تقریباً در تمام الگوریتم‌های ردیابی فرض می‌شود که حرکت جسم، صاف و بدون تغییر مسیر ناگهانی است. دیگر اینکه می‌توان حرکت جسم را با سرعت ثابت یا شتاب ثابت براساس اطلاعات سابق فرض کرد و همچنین می‌توان فرضیه‌های اولیه در مورد تعداد ، اندازه یا ظاهر و شکل شی داشته باشیم . روش‌های متعددی برای ردیابی شی در نظر گرفته شده است. در درجه اول باید به سوالات زیر پاسخ داده شود: کدام شی برای ردیابی مناسب است؟ کدام ویژگی تصویر مناسب است؟ چگونه حرکت، ظاهر و شکل جسم مدل[5] شود؟ پاسخ به این پرسش‌ها به محیط ردیابی شی بستگی دارد. هدف از این بررسی، مطالعه‌ی روش‌های ردیابی و دسته‌بندی آنها و شرح جامع از روش‌ها و رویکردها در هر مجموعه است. علاقمندیم به خوانندگان نشان دهیم که کدام یک از الگوریتم‌های موجود برای کاربرد خاص موردنظرشان مناسب‌تر است. علاوه‌براین هدف ما شناسایی و معرفی روندها و فناوری‌های جدید می‌باشد.

در این پژوهش ما قصد ارائه‌ی رویکردی داریم تا در آن اشیای متحرک را با استفاده از جابه‌جایی مبتنی بر لبه و حذف پس‌زمینه ردیابی کنیم.

هدف تحقیق

هدف ردیابی اشیاء، دنبال کردن حرکات در تصاویر ویدئویی از فریمی به فریم دیگر است. این فناوری قدرتمند می‌تواند برای سیستم‌های نظارت تصویری در شهرها و اماکن مهم به کار رود. این فناوری می‌تواند

[4] Straight off

[5] Model

در مقیاس‌های گسترده مورد استفاده قرار گیرد. این روند، ردیابی مسیر یک شی از یک فریم به فریم دیگر در یک ویدیو است. انواع موارد استفاده‌ی ردیابی اشیا به‌صورت زیر است:

1- تشخیص براساس حرکت: مثل شناسایی انسان براساس نوع راه رفتن ، تشخیص خودکار شی و غیره
2- نظارت خودکار : نظارت بر صحنه برای تشخیص فعالیت‌های مشکوک و یا حوادث غیر قابل پیش‌بینی
3- نمایه‌سازی تصویر : حاشیه نویسی اتوماتیک و بازیابی فیلم‌ها در پایگاه داده‌های چندرسانه‌ای
4- تعامل انسان و کامپیوتر : تشخیص چهره [6]، با نگاه کردن فرد به صفحه کامپیوتر و تشخیص حالت فرد برای دادن ورودی به کامپیوتر.
5- نظارت بر ترافیک : جمع‌آوری آمار ترافیک بصورت آنلاین برای هدایت جریان ترافیک
6- ناوبری اشیا: برنامه‌ریزی مسیر و اجتناب از برخوردبا موانع و حرکت در مسیر تعیین شده

لذا در این پژوهش ما قصد ارائه‌ی رویکردی را داریم تا در آن اشیاء متحرک را با استفاده از جابه‌جایی مبتنی بر لبه و حذف پس‌زمینه ردیابی کنیم.

مروری بر کارهای گذشته

در این بخش، به بررسی رویکردهای اخیر مربوط به موضوع پژوهش پرداخته می شود.

ردیابی با استفاده از تشخیص، اخیراً در مقاله‌های [5]- [9] مورد مطالعه قرار گرفته است. این موضوع در واقع، روش‌های ردیابی مبتنی بر تشخیص می‌باشد. چون تشخیص‌گر برون‌خط[7] یاد می‌گیرد می‌تواند به عنوان یک نوع خاص از روش‌های مبتنی بر یادگیری قلمداد گردد.

[6] Face Recognition

در مقاله‌ی [10]، مونتاژ قسمت‌های بدن انسان برای شناسایی و ردیابی تا حدی مورد استفاده قرار می‌گیرد. در مقاله‌ی [4]، تشخیص‌گرهای اندام به شناسایی اندام مردم در طیفی گسترده از شلوغی و جمعیت می‌پردازد. در مقاله‌ی [11]، یک اتصال به بازخورد نوین[8] از تشخیص‌گر شی مطرح شده که (سنجش برحسب میل بصری) برای ردیابی انسان مورد استفاده قرار می‌گیرد. در مقاله‌ی [12]، یک رویکرد دو مرحله‌ای پیشنهاد شده که برای اولین بار یک مدل ظاهری از افراد ایجاد و سپس آنها را با تشخیص این مدل در هر فریم ردیابی می‌کند. این روش‌ها براساس تجزیه و تحلیل قسمت‌های مختلف بدن یا ژست‌های[9] انسان در حالت‌های مختلف، ردیابی انسان را انجام می‌دهند.

با این حال، کیفیت ارائه‌ی تصویر بسیاری از برنامه‌های کاربردی نظارت تصویری، برای این موضوع ممکن است به اندازه‌ی کافی واضح نباشد، در نتیجه حرکت دقیق قسمت‌های بدن به سختی قابل کشف است.

اخیراً مقالات [1]-[3]، [13]-[15] بر روی روش مبتنی بر آموزش تمرکز کرده‌اند. ردیاب بردار پشتیبان در مقاله‌ی [13]، از یک ماشین بردار پشتیبان یادگیر برون‌خطا استفاده می‌کند که آن را کلاس‌بندی و در یک ردیاب نوری مبتنی بر جریان تعبیه کرده است. در مقاله‌ی [2]، قاب فعلی با استفاده از یک طبقه‌بندی‌کننده آموخته شده در قاب قبلی طبقه‌بندی می‌شود. نسبت واریانس[10]، برای اندازه‌گیری ویژگی قابلیت تشخیص و انتخاب بهترین ویژگی از یک مجموعه ویژگی، برای ردیابی استفاده شده است. در مقاله‌ی [1]، یک گروه از کلاس‌بندی‌کننده‌های ضعیف به‌صورت آنلاین آموزش دیده‌اند و پیکسل‌ها به عنوان هدف یا پس‌زمینه برچسب‌دهی شده‌اند. از آنجا که تنها ساختارهای فضایی محلی استخراج شده‌اند، قدرت تفکیک‌پذیری در

---

[7] Offline

[8] Novel

[9] Gestures

[10] Variance

صحنه‌های شلوغ تنزل پیدا می‌کند. در مقاله‌ی [16]، هدف در یک زیرفضا با ابعاد کوچک که به‌صورت تطبیقی به روز شده، با استفاده از ردیابی تصاویر در فریم‌های قبلی نشان داده شده است. در مقاله‌ی [15]، یک فیلتر ذرات آبشاری با مجموعه‌ای از ویژگی‌های طول عمر مختلف مطرح شده است، در حالی که مدل احتمالی پر سر و صدا و دارای قله‌های زیادی بدون انتشار گاوسی در مرحله نمونه‌برداری است. در مقاله‌ی [17]، یک روش نیمه نظارتی پیشنهاد شده است که در آن نمونه‌های نشاندار، فقط از اولین قاب بار می‌آیند، و پس از آن نمونه‌های آموزش بدون نشانه در سمت چپ آورده می‌شوند. مقاله‌ی [18]، یک طبقه‌بندی تبعیضی به‌صورت آنلاین براساس یادگیری چند نمونه‌ای را آموزش داده است. همه این روش‌های استخراج شده، تنها ساختارهای فضایی را نشان می‌دهند اما در مورد اطلاعات حرکتی غفلت شده است.

برای رسیدن به تعامل‌طبیعی و همه‌جانبه بین انسان و کامپیوتر، دست انسان می‌تواند به عنوان یک دستگاه رابط به کار رود [19,20,21]. قدرت استفاده از حرکات دست از انسانی به انسان دیگر برای کانال‌های برقراری ارتباط فرق می‌کند که بخش عمده‌ای از انتقال اطلاعات در زندگی روزمره ما را انجام می‌دهند. حرکات دست می‌توانند که یک راه آسان برای استفاده و تعامل طبیعی بین انسان وکامپیوتر برقرار کنند. برای مثال، به عنوان وسیله‌ای ارتباطی ورود به سیستم از زبان به طور گسترده در میان مردم استفاده می‌شود. افرادی که توانایی صحبت کردن دارند از بسیاری از حرکات دست برای کمک به ارتباطات خود استفاده می‌کنند. با این حال، حرکات دست به طور کامل برای تعامل بین کامپیوتر و انسان کشف نشده است. در مقایسه با دستگاه‌های سنتی HCI، حرکات دست برای کاربران جهت ارتباط برقرار کردن با کامپیوتر و کشف دنیای مجازی سه‌بعدی راحت‌تر است[22].

فصل دوم: ارائه‌ی مفاهیم و اصطلاحات پایه

ردیابي اشياء[11]

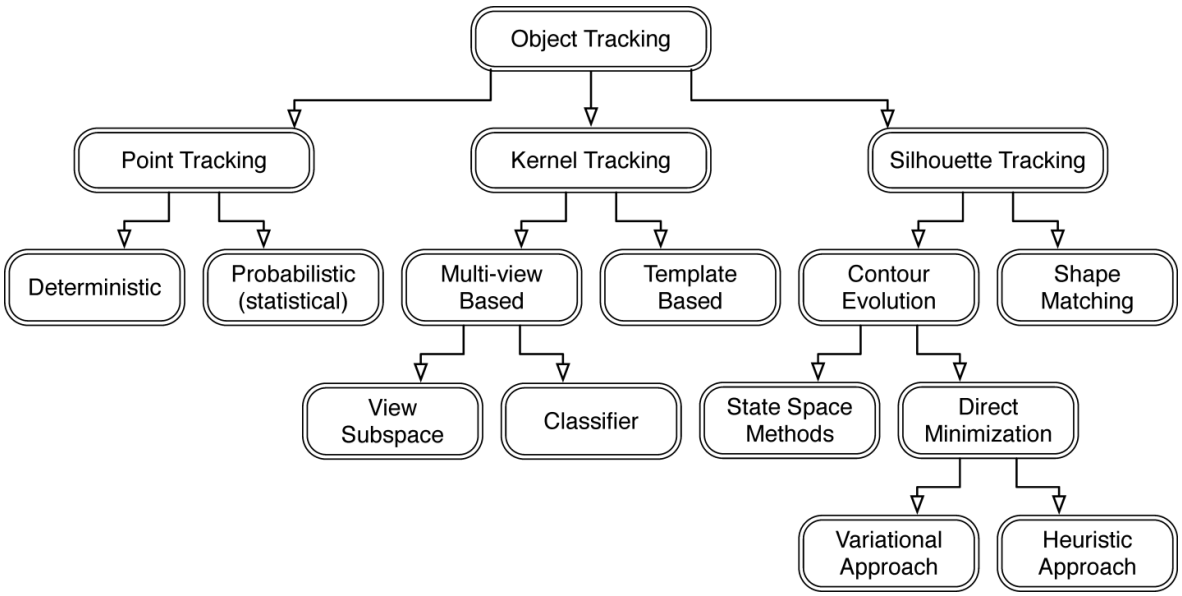
منظور از ردیابی شی، ایجاد مسیری از جابجایی شی در طول زمان در فریم‌های متوالی است و ممکن است ردیاب در هر لحظه از زمان، منطقه‌ی کمی از تصویر را زیر نظر داشته باشد[3] . تشخیص شی و برقراری ارتباط میان شی نمونه در فریم‌های مختلف ممکن است به‌صورت جداگانه یا مشترک صورت پذیرد. در حالت جداگانه، محدودهٔ شی در هر فریم به کمک الگوریتم تشخیص شی پیدا و سپس ردیاب در سراسر فریم‌ها، شی مربوطه را ردیابی می‌کند.

در حالت دوم (مشترک) با استفاده از اطلاعات بدست آمده از فریم‌های قبلی و به روزرسانی اطلاعات مرتبط با شی در هر لحظه، در هر فریم شی در منطقهٔ جدید تشخیص و ردیابی می‌شود.

در هر دو روش، ردیابی اشیا با استفاده از شکل یا ظاهر تشخیص داده می‌شود. سپس براساس شکل جسم و بسته به نوع حرکت و یا تغییر شکل آن، مدلی در نظر گرفته می‌شود. مثلاً اگر شی به شکل یک نقطه نشان داده شود تنها می‌توان از مدل انتقالی برای آن استفاده کرد یا اگر از شکل هندسی مانند بیضی برای نمایش استفاده شود از مدل حرکتی affine یا تغییرات تصویری مناسب می‌توان بهره برد.

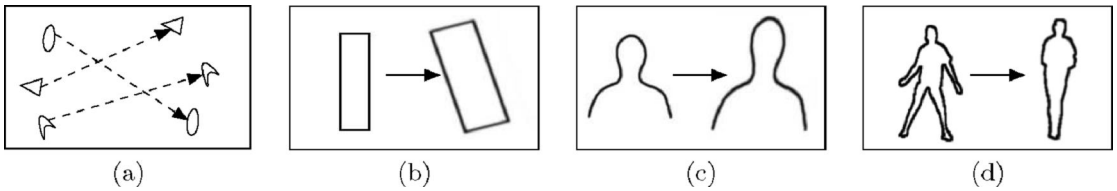
روش‌های مختلف ردیابی در نمودار1 زیر به تصویر کشیده شده‌اند.

[11]**Object Traking**



نمودار ۱-نمودار دسته‌بندی ردیاب‌ها

- ردیابی نقطه‌ای[5] : شناسایی اشیا در فریم‌های متوالی با نقطه نمایش داده می‌شود و نقاط درفریم‌های گوناگون به هم وصل می‌شوند. این روش برای تشخیص اشیا، در هر فریم نیاز به مکانیسم‌های خارجی دارد. (شکل 1)



شکل ۲-۱

- ردیابی هسته‌ای[12][4] : هستهٔ جسم به شکل و ظاهر آن اشاره دارد. هسته می‌تواند الگوی مستطیلی یا بیضی شکل همراه با هیستوگرام باشد. اشیا توسط محاسبهٔ حرکت هسته در فریم‌های متوالی ردیابی می‌شوند. این حرکت معمولاً در قالب انتقال پارامتری مثل جابجایی، چرخش یا Affine صورت می‌گیرد .(شکل2-1)

- ردیابی سایه [7]: این روش ردیابی با برآورد منطقهٔ شی در هر فریم انجام می‌شود. در این روش ردیابی اطلاعات در داخل منطقهٔ شی کدگذاری می‌شوند . این اطلاعات می‌تواند شکل ظاهری یا تراکم لبه‌ها باشد. با توجه به مدل شی، سایه می‌تواند ردیابی لبهٔ شکل یا حالت تکاملی آن باشد.

(شکل۲-۳)
هر کدام این روش‌ها به‌صورت قطعه قطعه [13]و با استفاده از اطلاعات تولید شده در فریم قبلی عمل می‌نمایند.

۲-۱ ردیابی نقطه‌ای :

در این نوع ردیابی، اطلاعات بدست آمده از نقاط در فریم‌های مختلف به منظور تشخیص شی به هم وصل می‌شوند. ایجاد ارتباط بین نقاط در صورت وجود اتصال بین اشیا و تشخیص اشیا ورودی و خروجی، کار پیچیده و مشکلی خواهد بود[5] .

به طور کلی این روش را می‌توان به دو دسته کلی تقسیم نمود:

- روش قطعی
- روش آماری

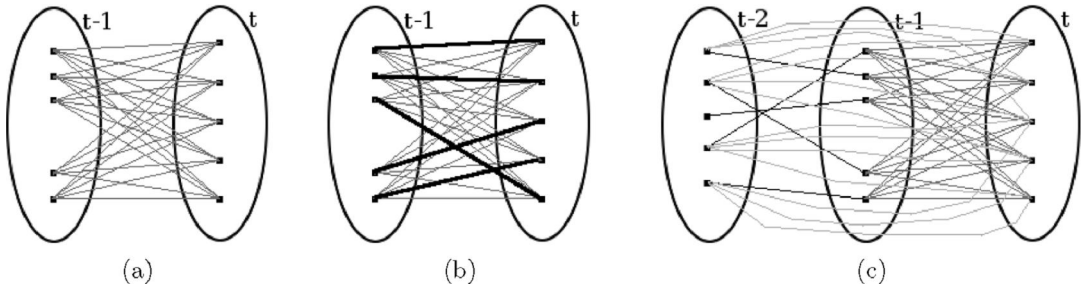
[12]Kernel
[13] Piece by piece

در روش قطعی با استفاده از فناوری‌های حرکت هوشمند کیفی، مشکل ارتباطات حل می‌شود. در روش آماری از اندازه‌گیری مستقیم جسم و محاسبات غیر قطعی برای ایجاد ارتباط استفاده می‌شود.

۲-۱-۱ روش‌هاي قطعي براي ارتباط :

در روش‌های قطعی برای ایجاد ارتباط، بررسی هر شی در t-1 فریم صورت می‌پذیرد[6] . سپس یک هزینه اختصاصی برای این بررسی تعریف می‌شود. لذا در فریم t مجموعه‌ای از محدودیت‌های حرکتی (با توجه به t-1 فریم قبلی) وجود خواهد داشت.

به حداقل‌رسانی این هزینه‌ها یکی از مشکلات بهینه‌سازی این روش است. یکی از روش‌های حل این مشکل، ارتباط یک به یک (متناظر[14]) است[6] . (شکل 2-2)



شکل ۲-۲

این ارتباطات با روش‌های تخصیص بهینه به دست می‌آید. مثل الگوریتم هانگاریان[15]یا روش‌های جستجوی حریصانه[16].

هزینهٔ ارتباطات با استفاده از ترکیب محدودیت‌های زیر به دست می‌آید[8] :

[14] Corresponding

[15]Hangarian

[16] Harisana