

金



دانشکده مهندسی کامپیوتر و فناوری اطلاعات

گروه هوش مصنوعی

بازشناسی و مکان‌یابی اشیا در صحنه های پیچیده

عباس زهره وند

اساتید راهنما:

دکتر علی اکبر پویان

دکتر علیرضا احمدی فرد

استاد مشاور:

دکتر جواد صدری

پایان نامه جهت اخذ درجه کارشناسی ارشد

بهمن 1392

خیال حوصله بحر می‌پزد بهیات چه باست در سر این قطره محال اندیش

حافظ

تقدیم به:

پدرم، مادرم و، همسرم زهرا

مشکر و قدردانی

پاس خدایی که آدمی را به نعمت تفکر آراست و اساتید فرزانه‌ای چون دکتر علی اکبر پویان و دکتر علیرضا احمدی فرد
رادر مسیر راهم قرار داد تا از اندیشه نابشان بهره‌گیرم و دانش و بینششان راره توشه خویش سازم. پاس می‌دارم اندیشه
بلندتان را و ارج می‌نهم همت و الیایان را.

در انتها مشکر می‌نمایم از پدر و مادرم، همسر و برادر و خواهرانم که وجودشان تکیه‌گاهی برای تمام لحظه‌های سخت من و
دعایانشان تنها سرمایه‌ی بال‌کشودنم بسوی خوشبختی است.

چکیده

در این پایان‌نامه روشی برای بازشناسی و مکان‌یابی اشیا در صحنه‌های پیچیده (شلوغ) ارائه شد. یکی از مهمترین و ابتدایی‌ترین مراحل در بینایی ماشین، بازشناسی اشیا می‌باشد. اگر این فرآیند به درستی و با قابلیت اطمینان انجام نشود مراحل بعدی به درستی صورت نمی‌گیرد. روش پیشنهاد شده در نهایت ترکیبی از دو روش موجود می‌باشد. هر دو روش موجود از یک فرآیند آرام‌سازی احتمالاتی برای تطبیق استفاده می‌کنند. این دو روش ابتدا با استفاده از توصیفگر SIFT نقاط کلیدی را برای صحنه و مدل بدست آورده و سپس با استفاده از تعریف همسایگی، از نقاط بدست آمده یک گراف می‌سازند. هر نود در گراف ساخته یکی از همان نقاط کلیدی است. این گراف سپس با استفاده از دو ویژگی توصیف می‌شود. ویژگی یکانی که برای توصیف هر نود به تنهایی استفاده شده و همان بردار 128 تایی SIFT می‌باشد. ویژگی باینری برای توصیف همسایگی دو نود استفاده می‌شود. برای ویژگی باینری از ویژگی‌های مقیاس و جهت استخراج شده توسط SIFT استفاده می‌شود.

برای بررسی تحمل دو روش تطبیق از یک سو و همچنین تعیین دقیق پارامترهای مسئله تطبیق، یک محیط شبیه‌سازی ساخته شد. در ادامه اتفاقات متداول دنیای واقعی در این محیط شبیه‌سازی شد. سپس تحمل دو روش تطبیق در این محیط شبیه‌سازی مورد ارزیابی قرار گرفت. حاصل این ارزیابی نشان داد که روش اول با استراتژی نود ساختگی پوچ عملکرد بهتری نسبت به روش دوم داشت. از طرفی وقتی نویز به صحنه وارد می‌شد روش دوم عملکردی بهتری داشت. هر دو روش زمانیکه قدرت نویز از یک آستانه فراتر می‌رفت دچار یک آشوب می‌شوند لذا نرخ بازشناسی خوبی بالایی نداشتند. در این حالت روش اول از آنجا که به دلیل وجود نویز هیچ مشاهده قابل اعتمادی وجود نداشت تمامی نودها را به سمت نود پوچ میل می‌داد.

در نهایت چندین سناریوی واقعی ترتیب داده شد و دو روش تطبیق در این سناریوها مورد ارزیابی قرار گرفت. همانطور که انتظار می‌رفت روش اول قابلیت اطمینان (منظور از قابلیت اطمینان تعداد تطبیق‌های اشتباه کمتر می‌باشد) بیشتری نسبت به روش اول داشت. اما روش دوم تعداد تطبیق‌های درست بیشتری را پیدا می‌کرد. برای استفاده از ویژگی‌های مناسب دو روش، یعنی قابلیت اطمینان بالا در کنار نرخ بازشناسی بالا ترکیب دو روش پیشنهاد شد. روش پیشنهادی نه تنها قابلیت اطمینان بالاتری نسبت به دو روش اول و دوم داشت، بلکه نسبت به هر دو روش نرخ بازشناسی بالاتری داشت.

کلید واژه: بازشناسی، مکان‌یابی، شیء، فرآیند آرام‌سازی احتمالاتی، گراف مدل، گراف صحنه،

مدل گراف رابطه‌ایی، ویژگی‌های یکانی، ویژگی‌های باینری

لیست مقالات مستخرج از پایان نامه

1. A.Zohrevand, A.Ahmadyfard, A.Pouyan, Z.Imani,"A SIFT based object recognition using contextual information", Proc. IEEE 12th Iranian conference on intelligent system, ICIS 2014.
2. A.Zohrevand,A.Pouyan,A.Azizian,"A Neuropsychological Based Model for Object Recognition Using Artificial Intelligence Methods",2th Basic and Clinical Neuroscience Congress 2013
3. A.Zohrevand,A.Ahmadyfard,A.Pouyan,"Object Recognition Using Contextual Information and SIFT Descriptor", International Journal of Computer Applications (IJCA), (Accepted)

فهرست

2	فصل اول
2	1-1 مقدمه
3	2-1 تعریف شیء و الفاظ مرتبط با بازشناسی اشیاء
3	3-1 اجزای مهم یک سیستم بازشناسی
5	4-1 ساختار پایان نامه
8	فصل دوم
8	1-2 مقدمه
8	2-2 سیستم های مرسوم بازشناسی اشیاء
13	1-2-2 توصیفگر SIFT
16	2-2-2 تطبیق
17	1-2-2-2 مروری بر روشهای تطبیق گراف
22	3-2 روش های الهام گرفته از بینایی انسان
22	1-3-2 محدودیت های سیستم های مرسوم بازشناسی
23	2-3-2 سیستم بینایی انسان
26	فصل سوم
26	1-3 مقدمه
26	2-3 SIFT به عنوان یک توصیفگر برای استخراج نقاط کلیدی
27	1-2-3 ساخت فضای مقیاس
30	2-2-3 آشکار کردن نقاط کمینه و بیشینه در فضای مقیاس
31	3-2-3 بهبود دقت موقعیت و حذف اکستریم های ناپایدار
34	4-2-3 تخصیص جهت به هر نقطه کلیدی
35	5-2-3 نمایش توصیفگر نقاط کلیدی
37	3-3 تطبیق
38	1-3-3 مروری بر تئوری تکنیک آرامسازی احتمالی
45	4-3 خلاصه و نتیجه گیری
48	فصل چهارم
48	1-4 محیط شبیه سازی شده

49	1-1-4 تعریف محیط شبیه سازی شده
51	1-1-1-4 انسداد
52	2-1-1-4 چرخش
54	3-1-1-4 نرخ شلوغی
56	2-4 مروری کوتاه بر روش اول و دوم
57	3-4 روش اول و در نظر نگرفتن نویز در محیط شبیه سازی
62	4-4 روش دوم بدون در نظر گرفتن نویز در محیط شبیه سازی
69	5-4 جمع بندی دو روش بدون وجود نویز
69	6-4 وجود نویز در صحنه
69	1-6-4 تعیین محدوده قدرت نویز
71	2-6-4 اضافه کردن نویز تنها به ویژگیهای یکانی
76	3-6-4 اضافه کردن نویز تنها به ویژگیهای باینری
78	4-6-4 اضافه کردن نویز به ویژگیهای یکانی و باینری
80	5-6-4 وجود نویز در روش دوم تطبیق
84	7-4 نتیجه گیری
88	فصل پنجم
88	1-5 توصیف تصویر با استفاده از توصیفگر SIFT
89	2-5 ویژگیهای یکانی
90	3-5 ویژگیهای باینری
90	4-5 بازنمایی تصاویر مدل و صحنه در ARG
91	5-5 تعریف مسئله
93	6-5 نتایج
95	1-6-5 چرخش مدل در صحنه
97	2-6-5 انسداد
99	3-6-5 وجود دو شی به صورت همزمان در صحنه
99	4-6-5 تغییر مقیاس
100	5-6-5 تغییرات روشنایی
101	7-5 بهبود تطبیق با ترکیب روش اول و دوم
102	8-5 بررسی تاثیر نویز
104	9-5 جمع بندی

106.....	فصل ششم
106.....	1-6 دیدگاه اول: دیدگاه تکنیکی
107.....	2-6 دیدگاه دوم: دیدگاه تقلیدی از انسان
108.....	پیوست ریاضی
111.....	منابع

فهرست اشکال

- شکل 3-1: ساخت هرم تصویری 28
- شکل 3-2: تصاویر بدست آمده حاصل از کانالو تابع گوسی با تصویر در مقیاسهای مختلف 30
- شکل 3-3: نمایش فضای مقیاس تفاضلی در مقیاسهای مختلف تصویر 30
- شکل 3-4: نحوه مقایسه نقاط برای بدست آوردن نقاط اکسترمم 31
- شکل 3-5: نمایش نقاط ویژگی بدون اعمال هیچ شرطی 32
- شکل 3-6: نقاط ویژگی بعد از اعمال شرط کنتراست 33
- شکل 3-7: نقاط ویژگی بعد از اعمال شرط مربوط به تابع هسین 33
- شکل 3-8: انتخاب یک پنجره در اطراف هر نقطه کلیدی به منظور پارامتر جهت 34
- شکل 3-9: نمایش بردار گرادیان هر یک از نقاط ویژگی 35
- شکل 3-10: نحوه تعیین بردار مشخصه برای هر ویژگی 36
- شکل 4-1: نمای کلی الگوریتم بازشناسی ارائه شده 48
- شکل 4-2: مدل استاندارد 51
- شکل 4-3: شبیه سازی انسداد در مدل شکل 4-1 52
- شکل 4-4: نمایی از یک شیء در دو حالت در اثر چرخش 53
- شکل 4-5: شبیه سازی برخی تغییرات در اثر چرخش 53
- شکل 4-6: شبیه سازی برخی تغییرات در اثر چرخش 54
- شکل 4-7: نمونه ای از شلوغ شدن صحنه، شکل بالا نودهای اضافه شده و پایین ارتباط بین نودهای اضافه شده نودها مدل 55
- شکل 4-8: شکل نهایی گراف صحنه پس از اعمال تغییرات ذکر شده 56
- شکل 4-9: صحنه و مدل ایجاد شده برای آزمایش اول 58

- شکل 4-10: صحنه و مدل ایجاد شده برای آزمایش دوم..... 64
- شکل 4-11: تعیین محدوده مجاز نویز در ویژگیهای یکانی..... 70
- شکل 4-12: تعیین محدوده مجاز نویز در ویژگیهای باینری. شکل بالا ویژگی مقیاس و پایین جهت..... 71
- شکل 4-13: تحمل روش اول تطبیق در برابر نویز اضافه شده تنها به ویژگیهای یکانی..... 72
- شکل 4-14: دومثال دیگر برای تحمل روش اول تطبیق در برابر نویز اضافه شده تنها به ویژگیهای یکانی..... 72
- شکل 4-15: تحمل روش اول در برابر نویز اضافه شده تنها به ویژگیهای یکانی در چندین آزمایش با یک قدرت مشخص..... 73
- شکل 4-16: رفتار روش اول زمانیکه نویز به ویژگیهای یکانی وارد میشود..... 74
- شکل 4-17: احتمالهای پسین برای روش اول در برابر وارد شدن نویز به ویژگیهای یکانی..... 75
- شکل 4-18: توابع حمایت برای روش اول در برابر وارد شدن نویز به ویژگیهای یکانی..... 75
- شکل 4-19: رفتار روش اول در برابر وارد شدن نویز به ویژگیهای باینری..... 76
- شکل 4-20: احتمالهای پسین برای روش اول در برابر وارد شدن نویز به ویژگیهای باینری..... 77
- شکل 4-21: توابع حمایت برای روش اول در برابر وارد شدن نویز به ویژگیهای باینری..... 78
- شکل 4-22: رفتار روش اول در برابر وارد شدن نویز به ویژگیهای یکانی و باینری..... 79
- شکل 4-23: رفتار تابع حمایت روش اول در برابر وارد شدن نویز به ویژگیهای یکانی و باینری..... 80
- شکل 4-24: رفتار روش دوم در برابر وارد شدن نویز به ویژگیهای یکانی..... 81
- شکل 4-25: رفتار تابع حمایت روش دوم در برابر وارد شدن نویز به ویژگیهای یکانی..... 82
- شکل 4-26: رفتار روش دوم در برابر وارد شدن نویز به ویژگیهای باینری..... 83
- شکل 4-27: رفتار روش دوم در برابر وارد شدن نویز به ویژگیهای یکانی و باینری..... 83
- شکل 4-28: رفتار تابع حمایت روش دوم در برابر وارد شدن نویز به ویژگیهای یکانی و باینری..... 84
- شکل 5-1: مراحل کار بر روی یک تصویر واقعی..... 92
- شکل 5-2: نمایی از پایگاه داده COIL..... 94

- شکل 3-5: نمونه ایی از تصاویر ساخته شده برای صحنه 95
- شکل 4-5: تطبیق صحنه و مدل، بالا با تطبیق مستقیم، وسط روش اول و پایین روش دوم..... 96
- شکل 5-5: نتایج حاصل از آزمایش انسداد، بالا روش اول و پایین روش دوم..... 98
- شکل 6-5: حاصل آزمایش برای وجود چند شی در صحنه 99
- شکل 7-5: تغییر مقیاس: سمت چپ با استفاده از روش اول، سمت راست با استفاده از روش دوم..... 100
- شکل 8-5: حاصل آزمایش نبود مدل در صحنه 100
- شکل 9-5: تطبیق: سمت چپ بالا با روش پیشنهادی، بالا سمت راست روش دوم، وسط سمت چپ روش دوم از نمای نزدیک، وسط سمت راست روش اول از نمای نزدیک ، سمت چپ پایین روش پیشنهادی از نمای نزدیک 101
- شکل 10-5: اضافه کردن نویز به سیستم و آزمایش و روش اول، قدرت نویز در تصویر بالا 0.0005، پایین 0.05..... 103
- شکل 11-5: اضافه کردن نویز به سیستم و آزمایش و روش اول، قدرت نویز در تصویر بالا 0.05، پایین 0.5..... 104

فهرست جداول

- جدول 4-1: ماتریس سردرگمی 59
- جدول 4-2: ماتریس شباهت های نودها درمدل و صحنه شکل 4-9 تنها با استفاده از ویژگیهای یکانی 60
- جدول 4-3: احتمالهای پسین برای روش اول 60
- جدول 4-4: تابع حمایت برای صحنه و مدل شکل 4-9 باروش اول 61
- جدول 4-5: احتمالهای پسین برای روش دوم 63
- جدول 4-6: تابع حمایت برای صحنه و مدل شکل 4-9 باروش دوم 63
- جدول 4-7: ماتریس شباهت های نودها درمدل و صحنه شکل 4-10 تنها با استفاده از ویژگیهای یکانی 64
- جدول 4-8: ماتریس سردرگمی برای شکل 4-10 و با استفاده از روش اول 65
- جدول 4-9: احتمالهای پسین برای روش اول برای شکل 4-10 66
- جدول 4-10: تابع حمایت برای صحنه و مدل شکل 4-10 باروش اول 67
- جدول 4-11: ماتریس سردرگمی برای شکل 4-10 و با استفاده از روش دوم 67
- جدول 4-12: احتمالهای پسین برای روش دوم برای شکل 4-10 68
- جدول 4-13: تابع حمایت برای صحنه و مدل شکل 4-10 باروش دوم 68

فرهنگ لغت

English phrase	ترجمه فارسی
Object	شی
Unary measurement	ویژگی یکتایی
Binary measurement	ویژگی باینری
Relaxation theory	نظریه آرام‌سازی
Relaxation labeling	برچسب گذاری به روش آرام‌سازی
Probabilistic Relaxation labeling	برچسب گذاری به روش آرام‌سازی احتمالاتی
Scale Invariant Feature Transform (SIFT)	تبدیل ویژگی‌های مقیاس ثابت
Affine Transform	تبدیل تکراری (ثابت)
Heuristic	غیر قانون مند
Object recognition	بازشناسی اشیا
Object localization	مکان‌یابی اشیا
Scene and Model	صحنه و مدل
Scene analysis	آنالیز صحنه
Descriptor	توصیفگر
Scene graph and Model graph	گراف صحنه و گراف مدل
Discrete and Continues relaxation labeling	فرآیند آرام‌سازی گسسته و پیوسته
Segmentation and region	بخش بندی و ناحیه
Eigen value and Eigen vector	مقدار ویژه و بردار ویژه
Attribute Relational Graph(ARG)	گراف ویژه رابطه‌ایی
Attribute Graph (AG)	گراف رابطه‌ایی
Key-point	نقطه کلیدی - عارضه
Adjacency matrix	ماتریس مجاورتی

به نام خداوند جان خرد کزین برتر اندیشه برنگذرد

فصل اول

فردوسی ۳۱۹-۳۹۷

مقدمه

فصل اول

1-1 مقدمه

سیستم‌های بینایی مصنوعی، بشر را از دوران ماقبل تاریخ شیفته خود کرده است. اولین اشاره به سیستم‌های بینایی به افسانه‌ها بر می‌گردد، جایی که یک غول برونزی به نام Talos توسط خدای باستانی (Hephaestus) ایجاد و به پادشاه جزیره مدیترانه Crete هدیه داده شد [1]. بر اساس این افسانه این روبات برای محافظت از مردم جزیره خدمت می‌کرد و برای محافظت، جزیره را هر روز سه بار دور می‌زد. این شیفتگی و علاقه تا به امروز ادامه داشته و نه تنها در مطالعات محض مرتبط با علوم، بلکه از پتانسیل بالای آن در علوم کاربردی هم استفاده بسیاری شده است [1]. مهمترین قسمت هر سیستم بینایی مصنوعی، بخش بازشناسی اشیاء¹ می‌باشد [1].

یکی از اساسی‌ترین و ابتدایی‌ترین بخش‌های بینایی ماشین، مسئله بازشناسی اشیاء می‌باشد. امروزه جدا از موفقیت‌های آشکار این سیستم‌ها در محیط‌های خاص منظوره، طراحی سیستمی که در شرایط کنترل نشده (مانند تغییرات روشنایی، انسداد²، تغییر زاویه دید³ و غیره) بتواند کلاس یک شیء نسبتاً پیچیده را تشخیص دهد چالش حل نشده‌ای می‌باشد. این در حالی است که توافق کلی روی الگوریتم‌های بازشناسی اشیاء در وضعیت‌های متفاوت وجود ندارد. این پایان نامه تمرکز خود را روی بازشناسی اشیاء در صحنه‌های پیچیده (شلوغ) قرار می‌دهد.

¹Object Recognition

²Occlusion

³View-point

2-1 تعریف شیء و الفاظ مرتبط با بازشناسی اشیاء

اغلب الفاظ و کلماتی در حوزه بازشناسی اشیاء وارد می‌شود که تعریف مناسبی برای آنها وجود ندارد و در

مواردی باعث ابهام¹ می‌شود [1]، مشهورترین این اصطلاحات عبارتند از [1]:

- تشخیص²: آیا یک شیء خاص در تصویر وجود دارد؟
- مکان‌یابی³: تشخیص بعلاوه مشخص کردن محل دقیق یک شیء خاص
- بازشناسی⁴: مکان‌یابی تمامی اشیاء حاضر در یک صحنه خاص
- فهمیدن⁵ یا تحلیل: بازشناسی و بررسی نقش هر شیء با توجه به محیط و اطراف

گاهی اوقات تعریف شیء⁶ ابهام برانگیز است. در واقع یک شیء بر اساس کاربرد تعریف می‌شود. برای مثال

تعریف شیء در بازشناسی کلمات دستنویس با تعریف شیء زمانی که هدف بازشناسی کروموزوم⁷ ها در

تحقیقات پزشکی متفاوت است. به عنوان یک تعریف ساده می‌توان یک شیء را مجموعه‌ای از قالب‌های⁸

معین دانست که در شرایط مشخص می‌تواند حس شود [1].

3-1 اجزای مهم یک سیستم بازشناسی

بازشناسی شامل دو قسمت مهم است: بازنمایی و تطبیق. برای بازنمایی روش‌های بسیاری وجود دارد.

بعضی از روشها یک تصویر را بخش بندی کرده و سپس از این بخش بندی برای توصیف تصویر استفاده

می‌کنند. با توجه به مشکلات بخش‌بندی محققان دیگر از توصیفگرها برای بازنمایی استفاده می‌کنند.

¹Ambiguity

²Detection

³Localization

⁴Recognition

⁵Understanding

⁶Object

⁷Chromosome

⁸Template

توصیف گرها تصویر را با استفاده از یک سری نقاط کلیدی به صورت یکتا بازنمایی می کنند. از انواع توصیفگرها می توان به لبه ها، گوشه ها، تقاطع ها نام برد که هر کدام دارای کاربردهای خاص خود هستند. از یک دیدگاه توصیفگرها به دو قسمت سراسری و محلی تقسیم می شوند. در حالی که توصیفگرهای سراسری سعی دارند که کل تصویر را به صورت یکپارچه توصیف کنند، توصیف کننده های محلی، سعی دارند تا نواحی مختلف تصویر را بصورت محلی توصیف کنند. به دلیل پیچیدگی هایی که تصویر شامل شیء (اشیاء) دارد و همچنین عدم توزیع یکنواخت این پیچیدگی ها در کل تصویر، توصیفگرهای محلی ابزار مناسب تری برای بازنمایی هستند.

تا کنون توصیفگرهای بسیاری طراحی و معرفی شده است که هر کدام دارای معایب و مزایایی هستند. با توجه به تغییراتی که می تواند در یک شیء اتفاق بیفتد، هر توصیفگر باید بتواند در برابر این تغییرات مقاوم باشد. یکی از این توصیفگرهای قدرتمند، توصیفگر SIFT می باشد. این توصیفگر هر تصویر را با استفاده از تعدادی نقاط کلیدی توصیف می کند. هر نقطه کلیدی دارای یک بردار با طول 128 می باشد. از این بردار برای توصیف نقاط کلیدی به صورت مجزا استفاده می شود. این توصیفگر همچنین برای هر نقطه کلیدی یک جهت و یک مقیاس نیز در نظر می گیرد.

طول بردار شامل 128 ویژگی برای هر نقطه کلیدی باعث بالا رفتن پیچیدگی محاسباتی می شود لذا برای بهبود پیچیدگی محاسباتی نسخه های سریع تری ارائه شد. از مهمترین آنها می توان به SURF اشاره کرد که طول بردار در این توصیفگر 64 می باشد.

در این پایان نامه از توصیفگر SIFT برای توصیف و بازنمایی استفاده خواهد شد

پس از بازنمایی یک شیء نوبت به تطبیق می رسد. روش های بسیار زیادی برای تطبیق وجود دارد. یکی از داده ساختارهایی که می تواند بطور مناسب داده ها و ارتباط بین داده ها را نشان دهد گراف است. دو

دسته برای حل مسائل گراف وجود دارد: روش‌های دقیق و روش‌های غیردقیق. در حالی که دسته اول سعی می‌کنند شباهت دو گراف را به صورت دقیق اندازه‌گیری کند، دسته دوم میزانی خطا را می‌پذیرد.

4-1 ساختار پایان‌نامه

در این پایان‌نامه یک روش مبتنی بر گراف برای تطبیق مورد استفاده قرار خواهد گرفت. به این صورت که ابتدا برای هر تصویر (مدل)، توصیفگر SIFT روی آن اعمال شده و نقاط کلیدی استخراج می‌شوند. در ادامه برای هر نقطه کلیدی k نزدیک‌ترین همسایه را به عنوان همسایه‌های نقطه مفروض در نظر گرفته می‌شوند، این ساختار را سپس در گراف ARG برای مدل در نظر گرفته می‌شود. با روندی مشابه همین کار را برای تصویر ورودی تست (صحنه) نیز انجام می‌شود. پس از پایان این مرحله دو گراف ساخته می‌شود. در نهایت مسئله تطبیق یافتن میزان شباهت بین این دو گراف است. از آنجا که بسیاری از مسائل مبتنی بر گراف NP-Complete می‌باشد (به این معنی که برای راه حل الگوریتمی با پیچیدگی زمانی چندجمله‌ای ارائه نمی‌دهند) میزانی خطا پذیرفته می‌شود و اصطلاحاً مسئله آرام¹ می‌شود. برای تطبیق دو گراف صحنه و مدل از روش آرام سازی احتمالاتی استفاده می‌شود.

ساختار این پایان‌نامه به شرح زیر است:

در فصل دوم مهمترین کارهای انجام شده مورد بررسی قرار گرفته، کارهای مرتبط با مسئله معرفی می‌شوند و نقاط ضعف و قوت آنها بررسی می‌شود. در فصل سوم به مبانی نظری پرداخته می‌شود. در مدل ارائه شده برای بازشناسی از توصیفگر SIFT برای بازنمایی و همچنین از فرایند آرام سازی برای تطبیق استفاده می‌شود، لذا در فصل سوم مبانی نظری آنها مورد بحث قرار می‌گیرد.

قسمت ارزیابی و ارائه نتایج شامل دو فصل است.

¹Relax

از آنجا که تعیین محدوده پارامترهای موثر بر سیستم بسیار مهم است در فصل چهار یک محیط شبیه سازی شده پیشنهاد می شود که در آن دنیای واقعی شبیه سازی شده است. در این فصل نتایج شبیه سازی در دو حالت: 1- زمانی که نویزی به گراف وارد نشود و 2- زمانیکه نویز به گراف وارد شود، مورد بررسی قرار می گیرد. در این محیط تحمل دو الگوریتم که از دو نسخه مختلف فرایند آرام سازی برای حل مسئله استفاده می کنند، بررسی می شود و نقاط ضعف و قوت آنها بحث می شود.

در فصل پنجم در دنیای واقعی نتایج در محیط شلوغ و پیچیده ارائه می شود. یک پایگاه داده از تصاویر اشیا است که پس زمینه بسیار ساده دارند¹ Coil 100 برای ساخت مدل، همچنین از تصاویری طبیعی و شلوغ برای صحنه استفاده می شود. در انتها سیستم ارائه شده در این محیط شلوغ مورد ارزیابی قرار می -

گیرد

¹coil-100 @ www.cs.columbia.edu, 2013

هرچه می‌توانید از اشتباهات دیگران چیزهای جدید بیاموزید، زیرا شما وقت کافی برای همه این تجربیات نخواهید داشت

آلفرد شین ولد ۱۹۱۲-۱۹۷۰

فصل دوم

کارهای انجام شده

فصل دوم

1-2 مقدمه

همانطور که گفته شد مهم‌ترین بخش هر سیستم بینایی، بخش بازشناسی اشیاء است. با توجه به اینکه کارهای بسیار زیادی در این حوزه انجام شده یا در حال انجام است، در این فصل منابع به دو دسته تقسیم شده و سپس در هر دسته مهم‌ترین و مرتبط‌ترین کارهای انجام شده مرور مختصری خواهد شد.

این دو دسته عبارتند از:

- سیستم‌های مرسوم بازشناسی اشیاء¹
- سیستم‌های بازشناسی اشیاء که از بینایی انسان الگو می‌گیرند و سعی می‌کنند مانند انسان عمل کنند.²

2-2 سیستم‌های مرسوم بازشناسی اشیاء

در این دسته بازشناسی را می‌توان در حالت کلی در دو مرحله انجام داد [2]:

- بازنمایی اشیاء³
- تطبیق⁴

در مسئله بازنمایی باید بتوان یک شیء مفروض را به بهترین نحو ممکن بازنمایی کرد. به عبارت دیگر باید بتوان یک شیء را به فضایی دیگر نگاشت⁵ داد. این فضا باید در مقابل تغییراتی که اتفاق می‌افتد مقاوم¹ باشد.

¹Traditional object recognition systems

²Cognitive object recognition systems

³Object representation

⁴Matching

⁵Mapping

به صورت کلی در دسته بندی روش‌های بازنمایی توافق جامعی وجود ندارد و منابع مختلف این دسته بندی ها را به صورت مختلفی انجام می‌دهند.

در یک نگاه جامع می‌توان دو دسته برای بازنمایی یک شیء در نظر گرفت [3]. در حالی که دسته اول سعی دارند که یک شیء را با استفاده از اطلاعاتی نظیر لبه، مرزها، گوشه ها و نقاط تقاطع نشان دهند. دسته دوم از روشنایی و رنگ برای بازنمایی استفاده می‌کنند. جدا از دسته بندی بالا از دیدگاه کاربردی تر می‌توان روش‌های متفاوتی برای بازنمایی یک شیء در نظر گرفت، از جمله [3]:

- روش‌های مبتنی بر هیستوگرام
- روش‌های مبتنی بر فضای ویژه²
- روش‌های مبتنی بر گوشه ها و لبه ها
- روش‌های مبتنی بر گراف³

در میان دسته بندی های مبتنی بر هیستوگرام می‌توان به کار Swain [4] اشاره کرد که از رنگ به عنوان یک ویژگی اصلی برای بازشناسی استفاده کرده است. همچنین می‌توان از کار Stricker [5] که از هیستوگرام مرز شیء برای بازشناسی استفاده کرده است، نام برد. از آنجا که روش‌های مبتنی بر هیستوگرام ساده، دارای سرعت مناسب و در برابر تغییرات مقاوم هستند، روش‌های جالبی هستند اما در عین حال ضعف‌هایی دارند [3]. از جمله عیب‌های بزرگ این روش‌ها، ناتوانی آنها در مدل کردن ساختار و شکل شیء و تنها تمرکز بر روی ویژگی‌های رنگ می‌باشد.

¹Robust

²Eigensapce-based

³Graph-based

روش‌های مبتنی بر فضای ویژه، کل تصویر را به صورت یک بردار در نظر می‌گیرند و سعی می‌کنند با استفاده از روش‌های کاهش بعد مانند آنالیز مولفه‌های اصلی¹، بعد آنها را کاهش دهند. از کارهای انجام شده در این دسته می‌توان به کار Turk [6] اشاره کرد. این روش‌ها زمانی کارایی مناسبی دارند که برای مثال تمامی پس زمینه دارای یک توزیع یکنواخت قابل شناسایی باشد [3]. روش‌های موجود در این دسته زمانیکه یک شیء کل تصویر را دربر می‌گیرد بسیار مناسب است [7].

در روش‌های مبتنی بر گراف تصویر به نواحی تشکیل دهنده آن بخش می‌شود و هر ناحیه یک نود از گراف را تشکیل می‌دهد. سپس بین هر ناحیه و نواحی مجاور یک ارتباط² در نظر گرفته می‌شود [2]. از مزیت‌های روش‌های مبتنی بر گراف این است که ویژگی‌های ساختاری را می‌توان به راحتی به صورت گراف در نظر گرفت و سپس از یک الگوریتم تطبیق برای تطبیق زیر گراف‌ها استفاده کرد [8].

جدا از دسته بندی‌های ذکر شده می‌توان به دسته بندی ارائه شده توسط احمدی فرد [2] اشاره کرد. در این مقاله ابتدا عنوان می‌شود که هر سیستم بازشناسی شیء مبتنی بر نقطه دید دارای دو قسمت است: بازنمایی و تطبیق. سپس بیان می‌شود که روش‌های مبتنی بر نقطه دید در دو دسته کلی قرار دارند: روش‌های مبتنی بر ویژگی³ و روش‌های مبتنی ظاهر⁴. روش‌های مبتنی بر ویژگی سعی می‌کنند یک شیء را با استفاده از ویژگی‌های محلی⁵ نمایش دهند. این ویژگی‌ها معمولاً ویژگی‌های ساده‌ای می‌باشند. هدف از این روش‌ها یافتن شباهت بین دو تصویر از یک شیء است زمانیکه در تصاویر تبدیلات هندسی صورت گرفته باشد. این ویژگی‌ها در مقابل تغییراتی مانند انسداد مقاوم هستند. البته باید عنوان کرد، زمانی این

¹ Principle Component Analysis (PCA)

² Edge

³ Feature-based

⁴ Appearance-based

⁵ Local

ویژگی‌ها مفید خواهند بود که با دقت استخراج شوند. تلاش‌های زیادی برای استخراج ویژگی‌های مناسب با تعریف بالا شده است [9][10].

در مقابل روش‌های مبتنی بر ویژگی، روش‌های مبتنی بر ظاهر وجود دارند که سعی می‌کنند یک شیء را با استفاده از تصاویر آن مدل کنند به عنوان مثال در [11] سعی می‌شود که یک شیء با استفاده از تعداد زیادی تصویر که از زوایای مختلف گرفته شده‌اند مدل شود. سپس با استفاده از ارتباط بین نماهای تصویر شیء، مدل فشرده‌ای از آن در فضای ویژه بدست آورد. از معایب این روش می‌توان به حساسیت آنها نسبت به داده‌های پرت¹ اشاره کرد. در تلاش برای ترکیب مزایای روش‌های مبتنی بر ویژگی و روش‌های مبتنی بر ظاهر روش‌های مبتنی بر توصیفگر² ارائه شد [12].

برخلاف روش‌های مبتنی بر ویژگی، توصیفگر محلی یک توصیف یکتا³ از تصویر ارائه می‌دهد. از مزایای مهم روش‌های مبتنی بر توصیفگرهای محلی این است که می‌توان با استفاده تنها از این توصیفگرها به مقایسه دو تصویر از یک شیء پرداخت. علاوه بر آن از آنجا که این توصیفگرها نسبت به تغییرات هندسی مقاوم هستند برخلاف روش‌های مبتنی بر ظاهر به تعداد نماهای کمتری برای بیان یک شیء نیاز دارند [2].

توصیفگرها را می‌توان به دو بخش سراسری⁴ و محلی تقسیم کرد. از آنجا که توصیفگرهای سراسری سعی می‌کنند تمام تصویر را به صورت ساده‌ای توصیف کنند در بسیاری از اوقات نمی‌توانند جزئیات تصویر را مشخص کنند. این جزئیات می‌تواند شامل تغییرات هندسی باشد که در بخشی از تصویر اتفاق می‌افتد [3][13].

¹Outlier data

²Descriptor

³Unique

⁴Global

با توجه به قدرت توصیفگرهای محلی در توصیف کامل تصویر، سعی می‌شود مرور کوتاهی به مهم‌ترین توصیفگرها شود. توصیفگرهای بسیار زیادی تاکنون ارائه شده است که هر کدام ویژگی‌های خاصی را مد نظر قرار می‌دهند. در [14] تقسیم بندی کلی روی توصیفگرهای محلی می‌شود که این تقسیم بندی را می‌توان به صورت زیر خلاصه کرد:

- توصیفگرهای مبتنی بر توزیع¹
- تکنیک‌های مکانی - فرکانسی²
- توصیفگرهای تفاضلی³

توصیفگرهای مبتنی بر توزیع سعی می‌کنند با استفاده از هیستوگرام ویژگی‌های مختلف ظاهر یا شکل یک شیء را بیان کنند [14]. یک توصیفگر ساده در این دسته، همان پیکسل‌های خود تصویر می‌باشد. توصیفگر بعدی که در این خانواده قرار دارد و شهرت بسیار دارد توصیفگر ارائه شده توسط Lowe [15] می‌باشد که در ادامه به طور مفصل بحث خواهد شد. خانواده دوم از توصیفگرها، توصیفگرهای مکانی-فرکانسی است. این تکنیک‌ها سعی می‌کنند که محتوای تصویر را با استفاده از فرکانس‌ها بیان کنند [14]. برای مثال، تبدیل فوریه⁴ محتوای یک تصویر را به توابع اولیه تجزیه می‌کند. اما از آنجا که ارتباط مکانی بین نقاط از بین می‌رود و همچنین تعداد این توابع اولیه بی نهایت هستند برای توصیف محلی مناسب نیستند [14]. تبدیل گابور⁵ [14] مشکل فوق را حل می‌کند اما مشکل دیگری دارد و آن این است که برای یک تغییر کوچک در فرکانس و جهت باید تعداد بسیار زیادی فیلتر محاسبه شود. این

¹Distribution-Based Descriptors

²Spatial-Frequency Techniques

³Differential Descriptors

⁴Fourier Transform

⁵Gabor wavelet

کار باعث رشد فزاینده پیچیدگی محاسباتی می‌شود. فیلتر گابور و تبدیلات موجک [14] برای کلاسه بندی بافت¹ مورد استفاده قرار می‌گیرند.

در دسته سوم سعی می‌شود تخمینی از همسایه‌های یک نقطه یا پیکسل محاسبه شود [14]. از مهم‌ترین کارهای اخیر در این دسته، الگوهای های باینری محلی² [16] است. این روش در عین حال که ساده است ولی کاربردهای بسیار زیادی مخصوصاً در آنالیز بافت دارد. این توصیفگر یک عملگر مربعی 3×3 را در نظر گرفته و روی تصویر شیء حرکت می‌دهد. این مربع روی هر پیکسلی که قرار می‌گیرد 8 پیکسل همسایه با پیکسل مرکزی مقایسه می‌شوند و اگر مقدار شدت روشنایی آنها از پیکسل همسایه بیشتر بود مقدار 1 و در غیر این صورت مقدار صفر را می‌گیرند. در نهایت مقدار شدت روشنایی پیکسل مرکزی، جمع وزن‌دار مقادیر پیکسل‌های همسایه است. سپس مربع را حرکت داده و روی پیکسل مجاور قرار می‌گیرد. این کار تا انتها ادامه می‌یابد. هیستوگرام مقادیر، نتیجه اعمال الگوی دودویی محلی می‌باشد. از این توصیفگر در کاربردهای بسیار استفاده شده است برای مطالعه بیشتر به [15][17]-[19] مراجعه کنید.

در میان خانواده توصیفگرهای معرفی شده، از SIFT کارایی بسیار مناسبی گزارش شده است [15][14]. به این دلیل کمی مفصل‌تر به ویژگی‌های این توصیفگر پرداخته و آن را با دیگر توصیفگرهای هم خانواده خود، مقایسه می‌شود.

2-2-1 توصیفگر SIFT³

این توصیفگر یکی از پرکاربردترین توصیفگرهایی است که در مقابل تغییراتی مانند چرخش، تغییر مقیاس، کشیدگی تصویر⁴ و نورپردازی پایداری مناسبی دارد. این روش توصیف بر پایه‌ی نقاط کلیدی

¹ Texture

² Local Binary Pattern (LBP)

³ Scale Invariant Feature Transform

⁴ Affine

استخراج شده از تصویر کار می‌کند. استخراج این نقاط کلیدی را می‌توان طی سه قسمت اصلی انجام داد: 1- یافتن نقاط کلیدی¹: که در یک راهکار کلی این مشکل با جستجو کردن در طول تصویر و یافتن نقاط منحصر بفردی که دارای ویژگی‌های خاصی هستند، انجام می‌گیرد. این نقاط را می‌توان با یافتن گوشه‌ها²، حفره‌ها³ و تقاطع‌های تی شکل⁴ در تصویر یافت. 2- توصیف نقاط کلیدی: توصیف این نقاط باید به گونه‌ای باشد که در مقابل نویز محیط، تغییرات هندسی و نورپردازی، نمایش یکسانی را از نقاط کلیدی داشته و متمایز و غیر حساس باشند. 3- مرحله آخر تطبیق این نقاط بین تصاویر مختلف است. عموماً در این مرحله از روش‌های محاسبه‌ی فاصله، همانند اقلیدسی، میان بردارهای ویژگی که در مرحله‌ی قبل بوجود آمده‌اند، استفاده می‌شود.

در مقایسه‌ای که میان روش‌های مختلف توصیف ویژگی‌ها انجام شده است [14]، تبدیل ویژگی غیر حساس به مقیاس (SIFT) [15] بارزترین توصیف را از شیء ارائه داده است. توصیفگر SIFT روی نقاط کلیدی غیر حساسی که از طریق تفاوت‌های گوسی⁵ (DOG) استخراج شده است، تکیه می‌کند. نسخه‌های مختلفی از توصیفگر SIFT ارائه شده است که در هر کدام سعی شده است تا یکی از ویژگی‌های آن تقویت شود. در برخی از موارد هم ویژگی‌های جدید مانند رنگ به آن اضافه شده است. این توصیفگر دارای دقت مناسبی است، اما تعداد نقاط ویژگی استخراج شده در آن بالاست، که محاسبات بیشتری را در پی خواهد داشت. این مساله در تصاویری که دارای پیچیدگی بیشتری هستند، نمود بیشتری دارد. در همین راستا روش PCA-SIFT⁶ [20] جهت کاهش ابعاد بردار ویژگی SIFT پایه از 128 به 36 ارائه شد. این روش باعث کاهش پیچیدگی زمانی می‌شود اما انعطاف‌پذیری زیادی ندارد. در مواقعی که ساختار

¹ Key-point

² Corner

³ Bubbles

⁴ T-junction

⁵ Difference of Gaussian

⁶ Principal Component Analysis SIFT

بافتی مشابه‌ای در تصاویر اتفاق افتاده است، SIFT خصوصیات محتویات شکل را در جهت بالابردن توان تمایز، به بردار ویژگی قبلی اضافه می‌کند [21].

در سال 2009 میچل و یو توصیفگر ASIFT¹ [22] را ارائه دادند. این توصیفگر علاوه بر داشتن تمام ویژگی‌های SIFT قدرت بسیار بالایی در تصاویری که تبدیل کشیدگی² بر روی آن اعمال شده است، داشت. در توصیفگر SIFT-CCH [23]، از ماتریس هم‌اتفاقی³ [24] در جهت شرکت دادن ویژگی رنگی در توصیفگر SIFT استفاده شده است. ماتریس هم‌اتفاقی اولیه بر روی تصاویر سطح خاکستری کاربرد داشت، ولی نسخه‌های دیگری از آن روی تصاویر رنگی هم قابل استفاده می‌باشد. این روش به دلیل ترکیب دو ویژگی در یک بردار ویژگی به SIFT-CCH معروف است.

در سال 2006 یک توصیفگر ویژگی غیر حساس به مقیاس دیگر تحت عنوان SURF⁴ [25] ارائه شد. توصیفگر SURF نیز همانند SIFT نقاط کلیدی را استخراج می‌نماید با این تفاوت که هر نقطه توسط یک بردار ویژگی با 64 مولفه نمایش داده می‌شود. بر اساس ارزیابی‌های انجام شده، به نتیجه رسیده‌اند که SIFT قدرت تشخیص بیشتری نسبت به SURF دارد ولی در مقابل SURF سریع‌تر و منسجم‌تر عمل می‌نماید. نسخه‌های تغییر یافته متفاوتی از توصیفگر SIFT به منظور بهبود ویژگی‌های تطبیق و کاهش پیچیدگی محاسباتی آن ارائه شده است. به منظور مطالعه بیشتر در مورد توصیفگرهای ویژگی غیر حساس به مقیاس می‌توانید به مراجع [14] [25]–[27] مراجعه نمایید.

¹ Affine-SIFT

² Affine transform

³ Co-occurrence matrix

⁴ Speeded Up RobustFeatures

2-2-2 تطبیق¹

همانطور که در ابتدای این فصل بیان شد پس از بیان مناسب صحنه و مدل، فاز بعدی تطبیق مناسب بین این دو می‌باشد. هدف از تطبیق بررسی میزان شباهت² بین دو بردار ویژگی استخراج شده از مدل و صحنه است. وقتی صحبت از میزان شباهت یا معیار شباهت می‌شود خود به خود بحث میزان دوری یا نزدیکی دو بردار تداعی می‌شود. می‌توان گفت که کارایی یک سیستم بازشناسی وابسته به معیار شباهت است [28]. در سال‌های اخیر تلاش‌های بسیار زیادی برای طراحی یک معیار شباهت مناسب مطرح شده است [29]–[32]. برخی از این معیارهای فاصله عبارتند از: فاصله اقلیدسی، هاسدورف [33]، فاصله هدف³ [34]، تقاطع هیستوگرام⁴. در میان این معیارها فاصله اقلیدسی بسیار ساده و در عین حال پر استفاده است. اما اگر شیء، حتی تغییر شکل ناچیزی داشته باشد، این روش کارایی بسیار پایینی خواهد داشت. فاصله هدف در محاسبه ارقام دست‌نویس بسیار کاربرد دارد [28]، این نوع معیار شباهت به تغییر شکل‌های محلی حساس است. [28] فاصله هاسدورف⁵ نه تنها به نویز مقاوم است بلکه این امکان را نیز می‌دهد تا بخشی از یک تصویر را با تصویر دیگر مقایسه شود. البته در تمامی این معیارهای شباهت فرض بر این است که تصویر به درستی توصیف شده است. همانطور که اشاره شده یکی از روش‌های بازنمایی اشیاء، استفاده از روش‌های مبتنی بر گراف است. وقتی صحبت از گراف می‌شود و ساختار گراف مطرح می‌شود برای تطبیق نیز باید سراغ روش‌های تطبیق دو گراف رفت.

¹ Matching

² Similarity

³ Target distance

⁴ Histogram intersection

⁵ Hausdorff distance

الگوریتم‌های متعلق به دسته دوم به دنبال این هستند که میزان سازگاری بین دو گراف را اندازه‌گیری کنند. در مقایسه با دسته اول این الگوریتم‌ها نیازی به جستجو در فضای حالت¹ برای یافتن راه‌حل ندارند. بنابراین الگوریتم‌های موجود در دسته دوم پیچیدگی محاسباتی بسیار کمتری دارند. الگوریتم‌های زیادی که در محدوده بهینه سازی گسسته² قرار گرفته از شبیه سازی تبریدی³[45][46]، آرام سازی گسسته⁴[47] و جستجوی مبتنی بر ژنتیک⁵[48][49] تا تجزیه مقادیر ویژه⁶[50][51]، آرام سازی لاگرانژ⁶[52] و آرام سازی احتمالاتی⁷[53] که در محدوده بهینه سازی پیوسته⁸ قرار می‌گیرند در این دسته (دسته دوم) طبقه بندی می‌شوند.

همانطور که در بالا گفته شد کلاسی از روش‌ها در دسته دوم برای تطبیق دو گراف از تکنیک آرام‌سازی استفاده می‌کنند. در این روش‌ها مسئله تطبیق گراف، به عنوان انتساب برچسب مناسب از پایگاه داده (گراف مدل) به هر نود از گراف ورودی (گراف صحنه) می‌باشد. بسته به مدل‌های تصمیم‌گیری این روش‌ها را می‌توان در دو کلاس طبقه‌بندی کرد:

- پیام محور⁹
- شیء محور¹⁰

¹ State space

² Discrete optimization

³ Simulated Annealing

⁴ Discrete relaxation

⁵ Genetic search

⁶ Lagrangian relaxation

⁷ Probabilistic relaxation

⁸ Continues optimization

⁹ Message oriented

¹⁰ Object oriented

در الگوریتم‌های پیام محور با یک انتساب یکجا¹، طوری برچسب‌ها به همه نودهای گراف ورودی اختصاص داده می‌شود که بیشترین سازگاری² در دو گراف برقرار شود [54]. روش‌های شی محورد در هر زمان بر روی اختصاص برچسب مناسب به یک نود متمرکز می‌شوند [55].

الگوریتم پایه و استاندارد آرام سازی گسسته [47] از یک واژه نامه برای بهینه کردن سازگاری بین دو گراف استفاده می‌کند. ضعف اساسی این روش در نظر نگرفتن شواهد تجربی³ در برقراری سازگاری است. کارهای بعدی این ضعف را مورد توجه قرار دادند برای نمونه [54] مسئله را در یک چارچوب احتمالاتی بیزین⁴ مدل کرد. بطوریکه معیار تطبیق را احتمال برچسب خوردن، با توجه به ویژگی های یکانی تعریف نمود.

در مقابل آرام‌سازی گسسته، تکنیک مبتنی بر آرام‌سازی احتمالاتی وجود دارد، که در خانواده الگوریتم‌های شی محور قرار می‌گیرند. آرام‌سازی گسسته سازگاری در انتصاب برچسب‌ها را اندازه می‌گیرد، در حالی که مدل احتمالاتی با استفاده از به‌روزرسانی احتمال برچسب‌ها، به برچسب زدن نودها می‌پردازد. این تکنیک اولین بار توسط [53] در یک حالت کاملاً هیورستیک⁵ ارائه شد. ضعف معیار مشاهده در روش اصلی توسط [55] برطرف شد. نویسندگان مرجع [55] برای حل مشکل هیورستیک از آرام‌سازی احتمالاتی با استفاده از فرمول سازی آن در چارچوب بیزین استفاده کردند. کاربرد این چارچوب سپس توسط Christmas و همکارانش [56] با استفاده از اطلاعات باینری در فرآیند آرام‌سازی بهبود داده شد. این روش نه تنها فهم مسئله را راحت‌تر می‌کرد، بلکه استراتژی واضح‌تری برای این فرآیند مشخص می‌کرد. کارایی مناسب این روش در کاربردهایی مانند بینایی استریو گزارش شده است. [56]

¹Join assignment

²Consistence

³Evidence

⁴Bayesian Probabilistic Framework

⁵heuristic

پس از اینکه با دو بخش بسیار مهم سیستم های بازشناسی اشیاء و کارهای مهم در هر بخش به صورت مجزا آشنا شدیم. در این قسمت از مهمترین و مرتبطترین سیستم‌هایی که تا کنون انجام شده است چند نمونه معرفی می‌شود. احمدی فردو همکارانش [2] سیستمی را برای تشخیص اشیاء ارائه دادند. تمامی کار آنها را میتوان در سه بخش خلاصه کرد:

- بخش بندی تصویر¹
- استخراج ناحیه‌ها²
- نرمال کردن ناحیه‌ها³

آنها قبل از بخش بندی برای حذف نویز تصویر را هموار کردند. پس از استخراج همه نواحی تصویر، هر ناحیه به یک فضای تغییر ناپذیر با تبدیل افاین نگاشت می‌شود. سپس تمامی ناحیه‌های حاصل از بخش بندی، در یک ARG قرار داده می‌شود. هر نود از این گراف که معرف یک ناحیه است با توجه به ویژگی‌های یکانی⁴ توصیف شده و هر ارتباط بین دو نود توسط ویژگی‌های باینری توصیف می‌شود. برای تطبیق از آرام‌سازی احتمالی استفاده می‌شود. آنها همچنین در روش پیشنهادی خود زمانیکه هیچ نودی از صحنه تناظری با نودی از مدل نداشته باشد راهکاری را در نظر گرفتند. این راهکار، اضافه کردن یک نود ساختگی⁵ به مدل است. یکی از مشکلات سیستم بالا مشکل بخش بندی تصویر است. امروزه پس از سالها کار بر روی بخش بندی تصویر هنوز راهکار کامل و جامعی برای بخش بندی وجود ندارد [57].

در سال 2005 Kostin و همکارانش [8] برای اینکه مشکلات بخش‌بندی را از بین ببرند و از طرفی بتوانند ناحیه‌های تصویر شیء را به راحتی و با قدرت بیشتری از نظر مقاوم بودن در برابر تغییرات استخراج کنند، از

¹Image segmentation

²Region extraction

³Region normalization

⁴Unary measurement

⁵Dummy node

توصیفگر SIFT برای استخراج نواحی استفاده کردند، با این نگاه هر ناحیه معادل با یک نقطه کلیدی استخراج شده توسط SIFT بود. با توجه به اینکه توصیفگر SIFT نسبت به تغییرات مقاوم بود لذا برخلاف کار قبلی نیازی به نگاشت نبود. در ادامه بین هر نقطه کلیدی و نقاط کلیدی همسایه یک همسایگی در نظر گرفتند. آنها سپس این ساختار را در گراف ARG قرار دادند. در نهایت آنها از مقادیر بردار 128 تایی SIFT برای ویژگی های یکانی، و برای ویژگی های باینری از مقادیر بزرگنمایی و جهت استخراج شده توسط SIFT استفاده کردند.

در سال 2012 sanroma و همکارانش [۵۸] روش دیگری ارائه دادند که به دو روش قبل تا حدی نزدیک بود. آنها به جای اینکه پیچیدگی را به نقاط کلیدی منتقل کنند، ابتدا با استفاده از کشف کننده هریس [59] که شناساکننده ساده‌ایی است، نقاطی را استخراج می‌کنند. سپس به جای اینکه تنها با استفاده از توصیفگرها به تطبیق پردازند، با استفاده از تعریف دو ارتباط ساختاری و هندسی این کار را انجام می‌دهند. این روش برخلاف اینکه از مقادیر مطلق¹ پیکسل‌ها استفاده کند از مقادیر هندسی آنها استفاده می‌کند [60][61]. در الگوریتم ارائه شده سعی بر این است که پیوستگی سراسر فرآیند تناظر حفظ شود. و این برخلاف الگوریتم های دیگری است که از آرام‌سازی گسسته برای تناظر استفاده می‌کنند [62]–[64]. در این الگوریتم، آنها با یک مدل ترکیب شده روبرو بودند که برای حل آن از روش حداکثر انتظار² استفاده می‌کنند. در این نگاه هدف، موفقیت در تخصیص احتمال است که با استفاده از انتصاب نرم³ انجام می‌شود. نکته مهم الگوریتم ارائه شده این است که می‌تواند خطا را در حین فرآیند یافته و اصلاح کند.

¹Absolute

²Expectation Maximization (EM)

³Soft-assign

همانطور که در ابتدای فصل گفته شد روش های بسیار زیادی برای حل مسئله بازشناسی ارائه شده است. ارائه تمامی آنها در این فصل و این پایان نامه میسر نیست برای مطالعه بیشتر به [2]، [13]، [65]–[69] مراجعه شود.

2-3 روش های الهام گرفته از بینایی انسان

آلن تورینگ¹ (1912-1954) [70] هوش مصنوعی را ساخت ماشینی دانست که می تواند از انسان تقلید کند. با الهام از این تعریف تورینگ دسته دوم شامل روشهایی است که سعی می کنند بینایی انسان را تقلید کنند.

اگر کمی به بینایی انسان توجه کنیم، مشاهده خواهیم کرد که انسان به راحتی بسیار از اجسام را می تواند تشخیص دهد بدون اینکه زحمت زیادی را متحمل شود. این در حالی است که بازشناسی اشیاء مسئله پیچیده ای می باشد. هدف اصلی تمام کارهایی که در این دسته انجام می شود این است که انسان چگونه اشیاء را تشخیص می دهد؟ [71].

2-3-1 محدودیت های سیستم های مرسوم بازشناسی

روش های مرسوم بازشناسی همیشه به یک پایگاه داده وابسته هستند و واضح است که هیچ پایگاه داده ای کامل نیست لذا این روش ها عمومیت ندارند. مشکل بعدی به این برمی گردد که روش های مرسوم بازشناسی یکبار آموزش می بینند و سپس مورد آزمایش قرار می گیرند که این نیز یک نقص است زیرا وقتی به بینایی انسان باز می گردیم کاملاً متوجه خواهیم شد که انسان مدام در حال یادگیری است [72].

¹Alan Turing

مشکل بعدی سیستم‌های مرسوم این است که وقتی یک سیستم مرسوم پس از آموزش در یک محیط ناآشنا (منظور تعریف راسل از محیط است [70]) قرار داده می‌شود عملاً هیچ کاری نمی‌تواند انجام دهد. در واقع نمی‌تواند خود را با شرایط جدید وفق¹ دهد. اما انسان در یک محیط جدید به راحتی پس از مدت بسیار اندکی با محیط وفق پیدا می‌کند. این مشکلات تنها بخشی از ناتوانی سیستم‌های مرسوم در مقابل سیستم بینایی انسان است. در ادامه این فصل مرور بسیار مختصری بر روش‌هایی که در این دسته قرار می‌گیرند خواهیم کرد.

2-3-2 سیستم بینایی انسان

تمامی کارهای که در این دسته قرار می‌گیرند از یک مقاله [73] سرچشمه می‌گیرند. در این مقاله یک مدل ریاضی برای لایه کورتکس² مغز پیشنهاد شد. همانطور که مشخص شده است عمل بازشناسی اشیاء در لایه از مغز به نام کورتکس انجام می‌شود. ما بدون اینکه بدانیم انرژی بسیار زیادی برای عمل بازشناسی اشیاء مصرف می‌کنیم. این مقاله یک معماری سلسه مراتبی از عمل بازشناسی ارائه می‌دهد این مدل HMAX نام دارد.

مقالات بسیار زیادی در سال‌های اخیر براساس اینمقاله ارائه شد از جمله:

در سال 2012 Jeong و همکارانش [72] روش سلسله مراتبی را ارائه دادند که از همان معماری کورتکس استفاده می‌کرد. این تحقیق ابتدا به مشکل یک بار آموزش دیدن سیستم‌های مرسوم اشاره می‌کند و سپس یک مدل افزایشی برای ارائه ویژگی‌ها ارائه می‌دهد. سپس بیان می‌کند که به جای اضافه کردن داده‌های اضافی از انعطاف پذیری داده‌ها استفاده شود. از طرفی اضافه کردن داده‌های اضافی فراموشی را نیز به سیستم تحمیل نخواهد کرد.

¹ Adopt

² Cortex

در کار دیگر Siagian و همکارانش [74] روشی مبتنی بر کورتکس برای روبات‌های متحرک در محیط‌های باز ارائه می‌دهند. روش‌های ارائه شده تا کنون در محیط‌های کنترل شده کارایی مناسبی داشته‌اند اما زمانی که محیط تغییر می‌کند و وارد محیط‌های باز می‌شویم کارایی سیستم‌های مرسوم بشدت کاهش می‌یابد. دلیل کاهش کارایی محیط بسیار پیچیده ای است که ذات محیط‌های روباز است.

این مقاله به ارائه معماری می‌پردازد که از لحاظ مکانیزم و همچنین پیچیدگی محاسباتی از انسان تقلید می‌کند. در این تحقیق اشاره می‌شود انسان همواره بر روی منابعی که در در میدان دیدش وجود دارد تمرکز می‌کند. در واقع انسان آن قسمت که بیشتر مورد علاقه است را مورد بررسی و جستجو قرار می‌دهد. تجزیه و تحلیل بسیار سریع تصویر جامع در این استراتژی، منجر به یک مجموعه کوچک از مکان‌های کاندید مهم در صحنه می‌شود.

انسان نباید هیچ امری را به عنوان حقیقت قبول کند، مگر آنکه به راستی در نظر او حقیقت باشد.

دکارت ۱۵۹۵-۱۶۵۰

فصل سوم

مباحث نظری

فصل سوم

1-3 مقدمه

در این فصل فرض می‌شود که هر تصویر از یک شی شامل مجموعه ایی از نقاط کلیدی¹ است. تغییراتی که در یک تصویر ممکن است اتفاق بیفتد: شامل چرخش، بزرگنمایی، کشیدگی تصویر² و نورپردازی است. این نقاط کلیدی باید نسبت به تغییرات گفته شده مقاوم باشد.

پس از استخراج نقاط کلیدی از تصویر یک شی سعی خواهد شد از یک ساختار برای نمایش این فضای تغییر یافته استفاده کرد این ساختار گراف ARG نام دارد. هر نود در این گراف نماینده یک نقطه کلیدی است که توسط ویژگی‌های یکانی توصیف خواهد شد. در این گراف یک همسایگی تعریف می‌شود. برای توصیف این همسایگی از ویژگی باینری استفاده می‌شود. از این گراف تحت عنوان گراف صحنه نام برده می‌شود. به طور مشابه برای هر مدل از شی در پایگاه داده یک ARG ساخته می‌شود.

ادامه این فصل به این شرح است: ابتدا مختصری درباره توصیفگر SIFT بحث خواهد شد و روش استاندارد این توصیفگر برای استخراج نقاط کلیدی شرح داده می‌شود. در ادامه به تشریح فرآیند آرام سازی در تطبیق گراف‌ها پرداخته خواهد شد و مبانی نظری این روش تطبیق بررسی می‌شود.

3-2 SIFT به عنوان یک توصیفگر برای استخراج نقاط کلیدی

الگوریتم SIFT به وسیله‌ی David Lowe در سال 1999 ارائه گردید [75]. SIFT یک روش برای آشکارسازی و استخراج ویژگی‌های مستقل و مشخص از تصاویر می‌باشد که می‌تواند برای تناظریابی³ بین تصاویر و یا شی با یک تصویر مورد استفاده قرار گیرد. به طور کلی مراحل توصیف یک تصویر توسط این

¹Key-point

²Affine

³Matching

توصیفگر را می‌توان به چهار قسمت اصلی تقسیم کرد. در ادامه جزئیات مربوط به هر مرحله بیان شده است.

3-2-1 ساخت فضای مقیاس¹

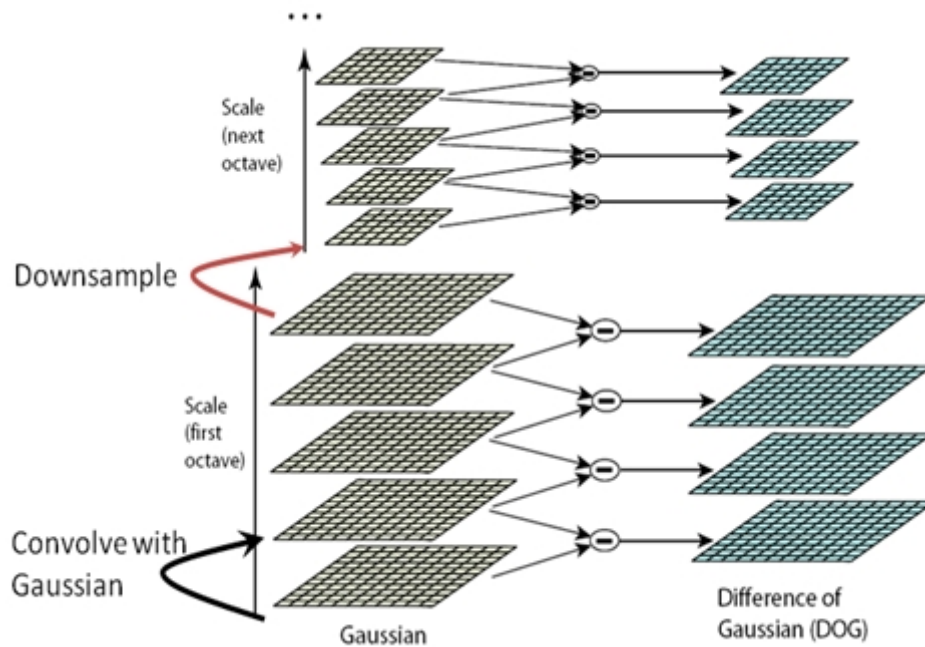
گام نخست برای یافتن نقاط کلیدی، پیدا کردن مکان‌هایی است که نسبت به تغییر مقیاس تصویر مقاوم هستند. این مکان‌ها با جستجو برای ویژگی‌های پایدار میان تمامی مقیاس‌های ممکن، توسط تابع پیوسته‌های مقیاس که به نام فضای مقیاس شناخته شده است و همچنین استفاده از مفاهیم هرم تصویری² یافته می‌شود. به وسیله‌ی هرم تصویری یک تصویر رقمی در قدرت تفکیک‌های متفاوت نمایش داده می‌شود و با توجه به مفهوم فضای مقیاس، تصویر در مقیاس‌های مختلف ساخته می‌شود (شکل 3-1)، در این شکل ابتدا تصویر را در چند مقیاس مختلف تشکیل می‌شود، سپس برای هر مقیاس، پنج تصویر (اکتاو³) با انحراف معیارهای مختلف به وجود می‌آید. از خصوصیات اصلی تصاویر هرمی، کاهش نویز تصویر در اثر نرم کردن در قدرت تفکیک‌های پایین‌تر می‌باشد، برای این کار از فیلتر گوسین⁴ استفاده شد.

¹Scale-Space

²Image Pyramid

³Octave

⁴GaussianFilter



شکل 3-1: ساخت هرم تصویری

میزان تاری¹ با پارامتر انحراف استاندارد σ در تابع گوسین کنترل می‌شود. σ برای اکتاو صفرم برابر $1/6$ و برای اکتاو بعدی دو برابر می‌شود و این روند ادامه می‌یابد. همچنین اولین تصویر از هر اکتاو به غیر از اکتاو صفر در فضای مقیاسی با زیر نمونه‌برداری از تصویر مناسبی از اکتاو قبلی بدست می‌آید. بعد از ایجاد هر اکتاو تصویر گوسی که مقیاس آن دو برابر مقیاس اولیه (2σ) می‌باشد، انتخاب شده و ابعاد آن با نمونه‌برداری مجدد نصف شده و به عنوان تصویر ابتدائی اکتاو بعدی در نظر گرفته می‌شود. این فرآیند تکرار می‌شود.

در مرحله‌ی بعد هر دو تصویر مجاور در داخل دسته‌های پنج‌تایی با اختلاف تفاضلی بر اساس جبر ماتریسی از هم تفریق می‌شوند، تا فضای تفاضلی² (DOG) نشان داده شده در سمت راست شکل (3-1) بدست آید. معادلات مربوط به اجرای این مرحله در روابط (3-1)، (3-2) و (3-3) آورده شده است:

¹ Blurring

² Different of Gaussian

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \quad 1-3$$

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2} \quad 2-3$$

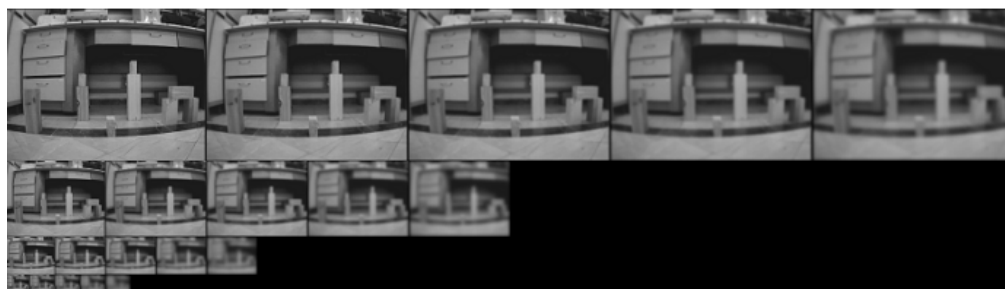
$$D(x, y, \sigma) = (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) = L(x, y, k\sigma) - L(x, y, \sigma) \quad 3-3$$

در معادلات بالا، تصویر اصلی است و G تابع کرنل گوسی می‌باشد. L ، تصویر تولید شده در اثر ضرب تلفیقی¹ تصویر I با کرنل G را است. در هر دسته‌ی پنج تایی، با تفریق دو تصویر مجاور که اختلاف آن‌ها در انحراف معیار σ می‌باشد، تصویر D بر اساس اختلاف تفاضل گوسی تولید می‌شود. پارامتر σ بیانگر مقیاس هر تصویر است و با استفاده از پارامتر ثابت K در یک روند افزایشی در سطوح هر اکتاو زیاد می‌شود. عبارت سمت راست تساوی رابطه (3-3) لاپلاسیان نرمالیزه شده تابع گوسین² است که برای تشخیص بردارهای ویژگی مستقل از مقیاس، مناسب است. محاسبه این معادله از نظر برنامه نویسی راحت‌تر است. زیرا از تفریق دو تصویر قابل حصول است و از آنجا که ضرب $(K-1)$ مقداری ثابت است، در پیدا کردن اکستریم تاثیری ندارد.

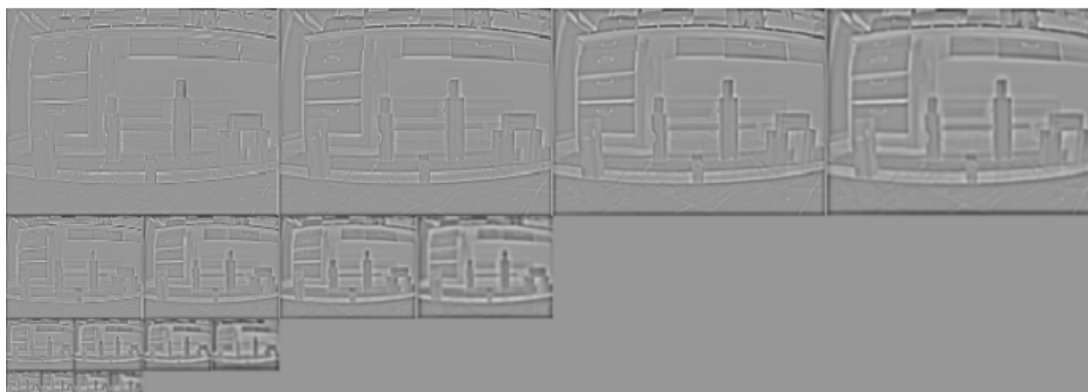
در شکل (2-3) و (3-3) تصاویر بدست آمده در مرحله اول اجرای الگوریتم که مربوط به مخروط‌های تصاویر گوسی و DOGها در مقیاس‌های مختلف است، نشان داده شده است.

¹ Convolution

² Laplacian of Gaussian



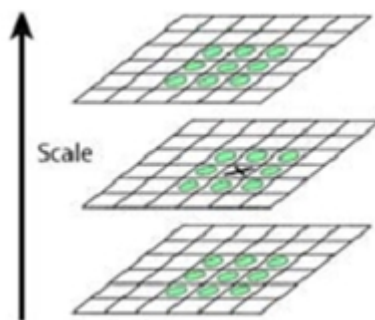
شکل 3-2: تصاویر بدست آمده حاصل از کانالو تابع گوسی با تصویر در مقیاس‌های مختلف



شکل 3-3: نمایش فضای مقیاس تفاضلی در مقیاس‌های مختلف تصویر

3-2-2 آشکار کردن نقاط کمینه و بیشینه در فضای مقیاس

مرحله بعد یافتن نقاط ماکزیمم و مینیمم در هر اکتاو است. این کار را با مقایسه هر پیکسل با همسایه‌های 26 گانه در ناحیه 3×3 تمامی سطوح DOG مجاور در همان اکتاو انجام می‌گیرد (شکل 3-4)). اگر نقطه مورد نظر کوچک‌تر یا بزرگ‌تر از همه همسایگانش باشد به عنوان نقطه کلیدی انتخاب می‌شود. با توجه اینکه در هر اکتاو، در فضای تفاضلی امکان اکستریم‌یابی برای دو تصویر وجود داشته باشد، اگر در هر اکتاو فضای تفاضلی، به S بازه تقسیم گردد باید $S+3$ تصویر در فضای مقیاس غیرتفاضلی تولید شود. برای رابطه $k=2^{1/s}(3-3)$ و در شبیه سازی $S=3$ در نظر گرفته می‌شود.



شکل 3-4: نحوه‌ی مقایسه نقاط برای بدست آوردن نقاط اکسترمم

3-2-3 بهبود دقت موقعیت و حذف اکسترمم‌های ناپایدار

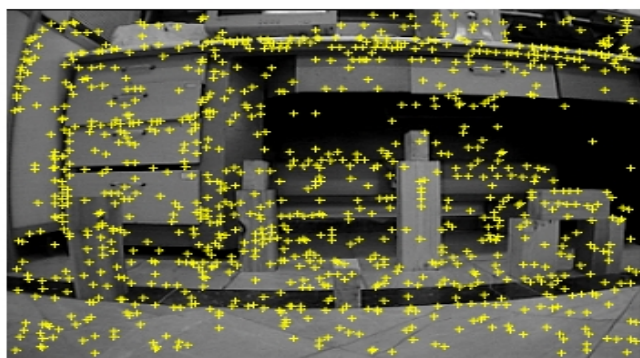
اکسترمم‌های استخراج شده دارای دو مولفه‌ی مختصات (مکان هر اکسترمم) و یک مولفه‌ی مقیاس هستند. مولفه‌ی مقیاس هر نقطه کلیدی براساس فضای مقیاس تفاضلی که از آن استخراج شده‌اند انتخاب می‌شود. بعد از استخراج نقاط کلیدی اولیه در تمامی فضای مقیاس، موقعیت و مقیاس دقیق آن‌ها با استفاده از برازش یک تابع درجه 3 در همسایگی 26 تایی هر عارضه تعیین می‌شود. این تابع با استفاده از بسط سری تیلور مرتبه‌ی دوم برای تابع فضای مقیاس حول نقطه‌ی مورد نظر برآورد می‌شود [75]. به منظور تعیین موقعیت بهبود داده شده، اکسترمم تابع فوق با استفاده از مشتق‌گیری تعیین شده و به عنوان آفست موقعیت هر نقطه به آن اضافه می‌شود. برای برآورد مشتقات از اختلافات میان درجات خاکستری پیکسل‌ها در همسایگی هر نقطه استفاده می‌شود. بدین ترتیب موقعیت مکانی هر نقطه کلیدی در حد زیر پیکسل محاسبه شده و میزان مقیاس دقیق آن نیز برآورده می‌شود.

هر نقطه اکسترمم برای اینکه به عنوان ویژگی انتخاب شود، باید مقدار کنتراست آن بالاتر از یک مقدار مشخص باشد. اندازه قدرمطلق تابع در موقعیت اکسترمم به عنوان میزان کنتراست هر عارضه در نظر گرفته می‌شود. با توجه به این که تصویر اولیه در بازه صفر و یک نرمالیزه می‌شود مقدار کنتراست هر عارضه باید از $0/03$ بیشتر باشد. با این کار نقاط کلیدی که میزان کنتراست آن‌ها از یک حد آستانه کمتر

است، ناپایدار و حساس به نویز بوده و حذف می‌شوند. از طرفی با توجه به این که در نقاطی که برای انتخاب ضعیف هستند، مشتق در یک جهت بزرگ و در جهت عمود به آن بسیار کوچک است. یک معیار برای تشخیص این نقاط استفاده از ماتریس هسین¹ است و چون نسبت بین اندازه مشتقات برابر با نسبت بین مقادیر ویژه ماتریس هسین است می‌توان چنین نقاطی را تمیز داد. برای مناسب بودن طبق رابطه (4-3) باید داشته باشیم:

$$\left(\frac{Tr(H)^2}{Det(H)} < \frac{(r+1)^2}{r}\right), H = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{bmatrix} \quad 4-3$$

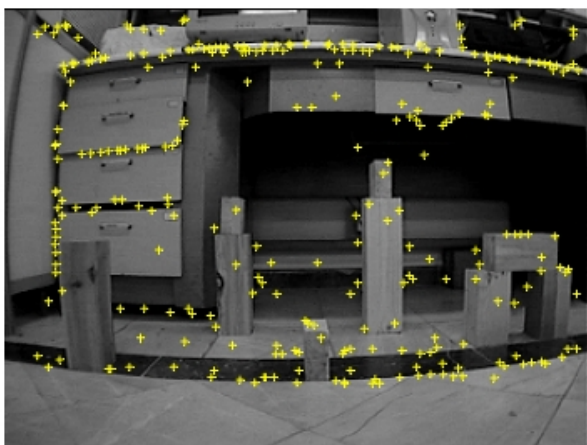
در رابطه 4-3، منظور از $\det(H)$ همان دترمینان ماتریس هسین می‌باشد، $Tr(H)$ تریس یا همان اثر ماتریس مربعی هسین، D_{xx} مشتق دوم نسبت به سطر (x) و D_{xy} مشتق یکبار نسبت به سطر (x) و یکبار نسبت به ستون (y) . مقدار $r = 10$ به عنوان حد آستانه معرفی شده است. برای بدست آوردن تابع هسین که یک ماتریس 2×2 می‌باشد از عملگرهای معروف مشتق گیری و کانوالو آن‌ها با تصویر اصلی استفاده می‌شود. با توجه به روش بیان شده، نقاط ویژگی بدون در نظر گرفتن هیچ شرطی و تنها بر اساس این که نقطه‌ی اکسترمم هستند در شکل (5-3) آمده است.



شکل 5-3: نمایش نقاط ویژگی بدون اعمال هیچ شرطی

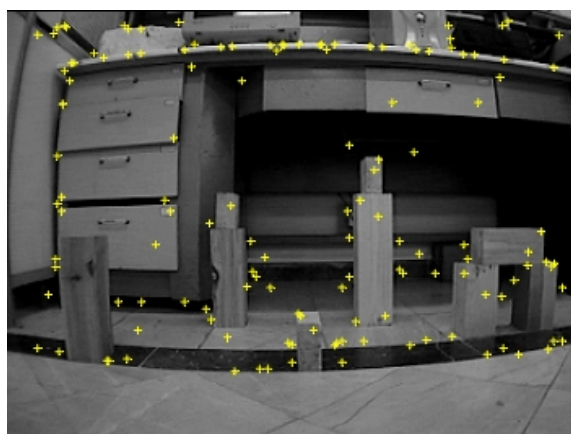
¹Hessian

با اعمال شرط کنتراست داریم:



شکل 3-6: نقاط ویژگی بعد از اعمال شرط کنتراست

همانطور که مشاهده می‌شود نقاطی که در لبه‌های میز قرار دارند، از جمله نقاطی هستند که مشتق آن‌ها در یک جهت، خیلی بزرگ‌تر از مشتق در جهت دیگر است. در نهایت با اعمال شرط مورد نظر و حذف نقاط واقع بر روی لبه‌ها شکل (3-7) را نتیجه خواهد شد که نمایشی از مکان نقاط کلیدی یافت شده در تصویر است.



شکل 3-7: نقاط ویژگی بعد از اعمال شرط مربوط به تابع هسین

3-2-4 تخصیص جهت به هر نقطه کلیدی

برای تناظریابی مستقل از دوران، برای هر یک از نقاط کلیدی استخراج شده از مرحله قبل، پارامتر جهت تعیین می‌شود. ابتدا یک پنجره دایره‌ای شکل در اطراف هر نقطه کلیدی به شعاع سه برابر مقیاس انتخاب می‌شود (شکل (3-8))، سپس مقادیر اندازه و جهت‌گردایان با استفاده از تصاویر DOG توسط روابط (3-5) و (3-6) برای نقاط درون پنجره محاسبه می‌شوند و مقادیر اندازه‌ی آن‌ها با استفاده از یک تابع گوسی با انحراف معیار $1/5$ برابر مقیاس نقطه کلیدی، وزندهی می‌شوند.

$$m(x, y) = \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2} \quad 5-3$$

$$\theta(x, y) = \tan^{-1}((L(x, y+1) - L(x, y-1)) / (L(x+1, y) - L(x-1, y))) \quad 6-3$$

که $m(x, y)$ معادل اندازه و $\theta(x, y)$ معادل جهت‌گردایان تصویر می‌باشند.

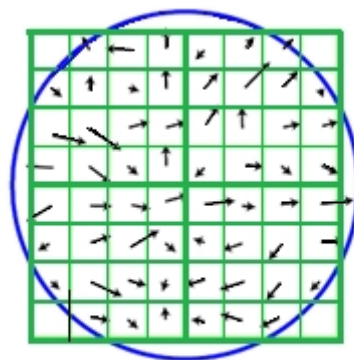
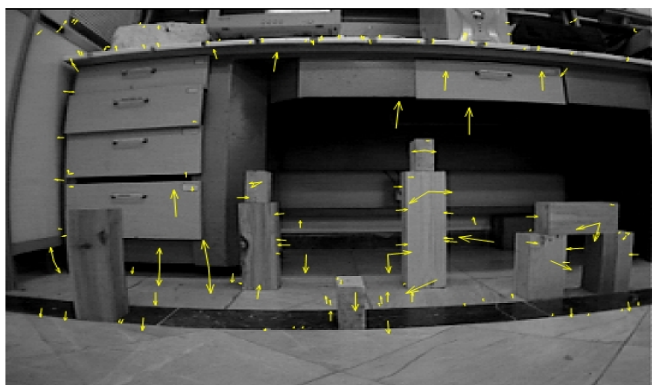


Image gradient

شکل 3-8: انتخاب یک پنجره در اطراف هر نقطه کلیدی به منظور پارامتر جهت

سپس یک هیستوگرام با عنوان هیستوگرام جهت ایجاد می‌شود که محور افقی آن محدوده‌ی 360 درجه را نشان می‌دهد و به ستون‌های 10 درجه‌ای تقسیم می‌شود. ارتفاع هرستون از هیستوگرام با استفاده از مجموع مقادیر گردایان‌های وزندهی شده‌ی پیکسل‌هایی که جهت‌گردایان آن‌ها در محدوده‌ی این ستون قرار دارد، تعیین می‌شود. جهت مربوط به ماکزیمم ستون هیستوگرام به عنوان جهت مربوط به

نقطه کلیدی مورد نظر انتخاب می‌شود. ممکن است برای بعضی از نقاط دو بیشینه در هیستوگرام داشته باشیم که در این صورت ویژگی مربوطه دارای دو جهت‌گیری خواهد بود. این اتفاق برای حدود 15 درصد نقاط کلیدی رخ می‌دهد. ولی در کل باعث بهبود در تطابق بین نقاط ویژگی می‌شود. برای اینکه جهت نقطه کلیدی با دقت بیشتری تعیین شود، جهت‌گیری با یک درون‌یابی بین سه نقطه از نزدیک‌ترین نقاط به نقطه پیک هیستوگرام محاسبه می‌شود. نتیجه نهایی برای تصویر مورد نظر در شکل (3-9) آورده شده است.



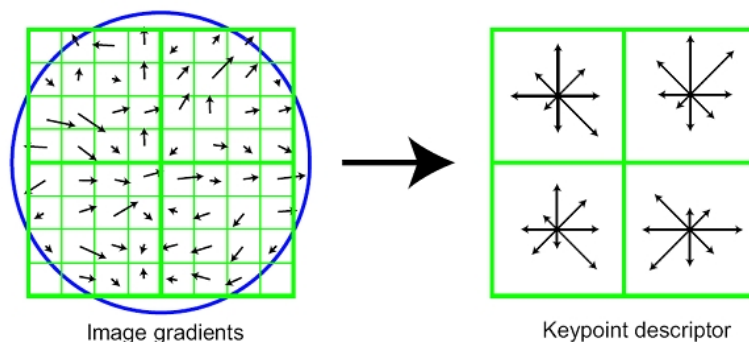
شکل 3-9: نمایش بردار گرادیان هر یک از نقاط ویژگی

اندازه بردارهای مربوطه نشان می‌دهد که مربوط به کدام مقیاس است. بردارهای بزرگ‌تر مربوط به مقیاس‌های بزرگ‌تر هستند. بنابراین هر نقطه کلیدی دارای چهار مولفه، دو مولفه مختصات، یک مولفه مقیاس و یک مولفه جهت خواهد بود و مستقل از تغییرات مقیاس و دوران تصاویر، نقاط کلیدی استخراج خواهد شد.

3-2-5 نمایش توصیفگر نقاط کلیدی

برای اینکه هر ویژگی به طور منحصر به فردی قابل شناسایی باشد، یک بردار مشخصه به آن تخصیص داده می‌شود تا در مراحل بعدی از جمله تطابق ویژگی مورد استفاده قرار گیرد. برای ایجاد توصیفگر ابتدا یک ناحیه به صورت یک شبکه 4×4 در اطراف هر نقطه کلیدی در نظر گرفته می‌شود. ابعاد این ناحیه

مطابق با مقیاس هر نقطه کلیدی و به گونه‌ای انتخاب می‌شود که هر سلول به صورت مربعی با ضلع معادل با سه برابر مقیاس نقطه کلیدی باشد. سپس مختصات شبکه معادل با جهت اصلی عارضه‌ی مورد نظر دوران داده می‌شود. در مرحله بعد مقادیر اندازه و جهت گرادیان پیکسل‌های درون ناحیه دوران داده شده، محاسبه و جهت گرادیان‌ها نیز معادل با جهت اصلی عارضه مورد نظر دوران داده می‌شود. یک تابع گوسی با مقیاس معادل با نصف پهناى شبکه به منظور وزن‌دهی مقادیر گرادیان‌ها استفاده می‌شود. وزن‌دهی با تابع گوسی باعث می‌شود که اولاً بردار مشخصه نسبت به تغییر مکان پنجره مشخصه حساسیت کمتری داشته باشد و ثانیاً اندازه گرادیان نقاط دورتر تاثیر کمتری در بردار مشخصه داشته باشند. در ادامه برای هر سلول در شبکه یک هیستوگرام هشت‌بخشی از جهت گرادیان‌های و وزن‌دهی شده پیکسل‌های درون آن مشابه با هیستوگرام جهت‌ایجاد می‌شود که برآیند هر یک بر روی چهار جهت اصلی و نیمساز آن‌ها ترسیم می‌شود. (مطابق شکل (3-10))



شکل 3-10: نحوه تعیین بردار مشخصه برای هر ویژگی

در این هیستوگرام محدوددهی 360 درجه به 8 قسمت 45 درجه‌ای تقسیم می‌شود. سپس اطلاعات هر یک از این هیستوگرام‌ها به دنبال هم قرار می‌گیرند و از آن‌جا که هر زیر پنجره یک هیستوگرام هشت‌بخشی دارد، بنابراین بردار مشخصه $4 \times 4 \times 8$ تایی خواهد شد. در نهایت بردار ویژگی‌ها برای کاهش اثر تغییرات روشنایی اصلاح می‌شود. در ابتدا بردار به طول واحد نرمالیزه می‌شود.

در مورد تغییر در روشنایی، یک ثابت به هر پیکسل تصویر اضافه می‌شود که تغییری در مقدار گرادیان ایجاد نخواهد کرد، به این دلیل که از روی تفاضل پیکسل‌ها محاسبه می‌شود. بنابراین، توصیفگر در مقابل تغییرات کشیدگی و روشنایی مقاوم است. بنابراین برای کاهش اثر نامطلوب دامنه بزرگ گرادیان، آن‌ها را با آستانه گذاری مناسب محدود به $0/2$ می‌کنیم و سپس به مقدار واحد تبدیل می‌نماییم. این بدین معناست که تطبیق دامنه‌ها برای نقاط گرادیان‌های بزرگ اهمیتی ندارد، در حالی که توصیف جهت‌ها اهمیت بیشتری دارد و مقدار $0/2$ نیز به طور تجربی با استفاده از تصاویری با روشنایی‌های متفاوت برای یک شی سه بعدی بدست آمده است.

3-3 تطبیق

تطبیق، یافتن تناظر¹ بین نقاط اساسی استخراج شده از تصویر صحنه و مدل (مدل‌های) استخراج شده در پایگاه داده² می‌باشد. در بخش قبل روشی برای نمایش هر شی در قالب گراف ARG ارائه شد. بر این اساس گراف شامل نقاط کلیدی استخراج شده از صحنه با عنوان گراف صحنه³ و همچنین با روشی مشابه برای هر شی در پایگاه داده تشکیل می‌شود (گراف مدل⁴)، سپس تمامی مدل‌های شی ساخته شده در قالب یگ گراف با عنوان گراف مدل آماده می‌شود.

با این کار هر شی یک زیر گراف⁵ از گراف ساخته شده از تمامی مدل‌های ساخته شده می‌باشد. با این نگاه تطبیق بین صحنه و مدل (ها) به وسیله گراف‌های این دو انجام می‌شود، حاصل این فرآیند یافتن تناظر بین زیرگراف (ها) صحنه و زیرگراف (ها) از مدل می‌باشد. از روش مدل آرام‌سازی احتمالاتی⁶ برای تطبیق بین دو گراف استفاده می‌کنیم [56]. همانطور که گفته شد از یک گراف برای تمامی مدل‌های موجود در پایگاه

¹Correspondence

²Database

³Graph scene

⁴Model graph

⁵Sub graph

⁶probabilistic relaxation technique

داده استفاده شد. در نتیجه با این نگاه تفاوت بسیار زیادی بین مسئله ما و کاربردهای قبلی از مدل آرام-سازی احتمالاتی وجود دارد [56].

در این بخش برای تطبیق بین دو گراف صحنه و مدل می‌توان دو روش را ارائه داد. اولین روش در نظر گرفتن تمامی نمایش‌های مربوط به هر نود در مدل آرام‌سازی احتمالاتی است. پیچیدگی این روش با افزایش تعداد نودها در مدل (گراف) صحنه رشد فزاینده‌ای دارد. در روش دوم برای کاهش پیچیدگی محاسباتی تطبیق به دو مرحله تقسیم می‌شود:

- در مرحله اول بهترین نمایش برای هر نود بدست می‌آید.
- سپس با استفاده از اطلاعات همسایگی احتمال برچسب زنی به هر نود به روز رسانی می‌شود.

3-3-1 مروری بر تئوری تکنیک آرام‌سازی احتمالی

در تکنیک آرام‌سازی احتمالاتی انتصاب هر برچسب به هر شی با استفاده از یک مقدار احتمالی بیان می‌شود. در ابتدا برچسب احتمالی متناسب با هر شی تنها با استفاده از مشاهدات مربوط به همان نود محاسبه می‌شود. این مقادیر اولیه سپس با استفاده از اطلاعات همسایگی هر نود به روزرسانی می‌شود.

تکنیک آرام‌سازی اولین بار توسط [76] معرفی شد. ضعف اساسی روش اولیه تکنیک آرام‌سازی [53] این بود که استراتژی طراحی آن بسیار هیورستیک بود. کارهای بعدی [55][53] این ضعف را برطرف کردند. در عمل [54] تئوری احتمالاتی بر مبنای چارچوب بیزین برای حل مشکل هیورستیک روش اولیه تکنیک آرام‌سازی ارائه داد. این چارچوب سپس توسط [76] با اعمال اطلاعات باینری گسترش داده شد. معرفی این ابزار (یکتایی و باینری) سراسر فرآیند برچسب زنی توسط تکنیک آرام‌سازی را تحت تاثیر قرار داد تا آنجا که این فرآیند را بسیار کارا تر کرد. نویسندگان [76] نشان دادند که مسئله تطبیق را می‌توان به عنوان مسئله تخصیص یک برچسب مناسب از مجموعه مدل به هر یک از اشیا موجود در صحنه در نظر گرفت

جایی که یک برچسب را به عنوان پوچ در نظر گرفته می‌شود. این برچسب برای اشیایی در نظر گرفته می‌شود که هیچ برچسب مناسبی برای آن در گراف مدل یافت نمی‌شود.

در آن فرمول‌بندی در معیار انتصاب به صراحت از اطلاعات یکانی و باینری با توجه به مشاهدات عنوان شده است:

$$P(\theta_i = \omega_\alpha) = \max_{\omega_\lambda \in \Omega} P(\theta_i = \omega_\alpha | \bar{x}_{j,j \in N_i^+}, \bar{A}_{ij,j \in N_i}) \quad 7-3$$

در تعریف ارائه شده در معادله (7-3) \bar{x}_j به اطلاعات یکانی متناظر با شی \mathbf{o}_j اشاره دارد همچنین \bar{A}_{ij} به اطلاعات باینری متناظر با نودهای \mathbf{o}_i و \mathbf{o}_j اشاره می‌کند. $P(\theta_i = \omega_\alpha)$ احتمال گرفتن برچسب α از مدل توسط نود α از صحنه را بیان می‌کند. با استفاده از قانون بیزین و سپس اعمال تئوری احتمالات معادله 3-7 را می‌توان به صورت زیر بسط داد.

$$P(\theta_i = \omega_\alpha | \bar{x}_{j,j \in N_i^+}, \bar{A}_{ij,j \in N_i}) = \frac{P(\theta_i = \omega_\alpha | \bar{x}_{j,j \in N_i^+}, \bar{A}_{ij,j \in N_i})}{P(\bar{x}_{j,j \in N_i^+}, \bar{A}_{ij,j \in N_i})} \quad 8-3$$

$$= \frac{\sum_{\omega_{\theta_1} \in \Omega} \cdots \sum_{\omega_{\theta_{i-1}} \in \Omega} \cdots \sum_{\omega_{\theta_{i+1}} \in \Omega} \cdots \sum_{\omega_{\theta_N} \in \Omega} P(\theta_1 = \omega_{\theta_1}, \cdots, \theta_i = \omega_\alpha, \cdots, \theta_N = \omega_{\theta_N})}{\sum_{\omega_{\theta_1} \in \Omega} \cdots \sum_{\omega_{\theta_N} \in \Omega} P(\theta_1 = \omega_{\theta_1}, \cdots, \theta_i = \omega_\alpha, \cdots, \theta_N = \omega_{\theta_N}, \bar{x}_{j,j \in N_i^+}, \bar{A}_{ij})}$$

در این بیان از مسئله تطبیق احتمال انتصاب شامل ترکیب تمامی انتصاب‌های ممکن برای هر شی در گراف صحنه می‌باشد. به این معنی که تمامی گراف‌هایی که ممکن است با توجه به شی‌های صحنه ایجاد شود را این تعریف در نظر می‌گیرد.

بدیهی است که پیچیدگی محاسباتی این نگاه بسیار بالاست و با یک حساب سرانگشتی می‌توان به این نتیجه رسید که پیچیدگی محاسباتی آن با افزایش تعداد نودهای صحنه از مرتبه نمایی است (برای دیدن

نحوه این نتیجه به پیوست ریاضی مراجعه شود). برای غلبه بر این مشکل و محدود کردن پیچیدگی محاسباتی معادله 3-8 به چند جمله ای فرضیات زیر مناسب و ضروری می باشد [76]:

- معیار اطلاعات یکانی در مجموعه $\bar{X}_{j,j \in N_i^+}$ کاملاً از هم مستقل هستند.
- ارتباطات باینری نیز از این قاعده مستثنی نبوده و این ارتباطات $\bar{A}_{ij,j \in N_i}$ نیز از یکدیگر مستقل هستند
- نکته دیگر اینکه ارتباط بین ارتباط یکانی و باینری هم باید مشخص شود. وقوع اطلاعات یکانی کاملاً مستقل از ارتباطات باینری در نظر گرفته می شود.
- اشتراک وقوع جفت اتفاق $\theta_i = \omega_{\theta_i}$ و $\theta_j = \omega_{\theta_j}$ صفر در نظر گرفته می شود به عبارت دیگر این دو رویداد نیز مستقل در نظر گرفته می شوند.

$$P(\theta_i = \omega_{\theta_i} \cap \theta_j = \omega_{\theta_j}) = 0 \quad 9-3$$

چگالی احتمال توام که در صورت و مخرج معادله 3-8 ظاهر می شود را می توان به صورت جملات شرطی به صورت زیر نوشت :

$$\begin{aligned} & P(\theta_1 = \omega_{\theta_1}, \dots, \theta_i = \omega_{\theta_i}, \dots, \theta_N = \omega_{\theta_N}, \bar{X}_{j,j \in N_i^+}, \bar{A}_{ij,j \in N_i}) \\ &= P(\bar{X}_{j,j \in N_i^+} | \theta_1 = \omega_{\theta_1}, \dots, \theta_i = \omega_{\theta_i}, \dots, \theta_N = \omega_{\theta_N}, \bar{A}_{ij,j \in N_i}) \\ & \quad \times P(\theta_1 = \omega_{\theta_1}, \dots, \theta_i = \omega_{\theta_i}, \dots, \theta_N = \omega_{\theta_N}, \bar{A}_{ij,j \in N_i}) \end{aligned} \quad 10-3$$

با سومین فرضی که برای سادگی محاسبات در نظر گرفتیم می توان ارتباطات باینری را از لیست متغیرهای شرطی از جمله اول در معادله 10-3 حذف کرد. بنابراین می توان با فاکتورگیری جمله دوم در معادله 10-3 و استفاده از دومین و آخرین فرض می توان معادله 10-3 را به صورت زیر بازنویسی کرد:

$$\begin{aligned}
 & P(\theta_1 = \omega_{\theta_1}, \dots, \theta_i = \omega_{\alpha}, \dots, \theta_N = \omega_{\theta_N}, \bar{A}_{ij}, j \in N_i) \\
 &= P(\bar{A}_{i1} | \bar{A}_{i2}, \dots, \bar{A}_{ii-1}, \bar{A}_{ii+1}, \dots, \bar{A}_{iN}, \theta_1 = \omega_{\theta_1}, \dots, \theta_N = \omega_{\theta_N}) \\
 &\times P(\bar{A}_{i2} | \bar{A}_{i3}, \dots, \bar{A}_{ii-1}, \bar{A}_{ii+1}, \dots, \bar{A}_{iN}, \theta_1 = \omega_{\theta_1}, \dots, \theta_N = \omega_{\theta_N}) \\
 &\times \dots \times P(\bar{A}_{iN} | \theta_1 = \omega_{\theta_1}, \dots, \theta_N = \omega_{\theta_N}) \times P(\theta_1 = \omega_{\theta_1}, \dots, \theta_N = \omega_{\theta_N}) \\
 &= P(\theta_i = \omega_{\theta_i}) \left\{ \prod_{j \in N_i} P(\bar{A}_{ij} | \theta_j = \omega_{\theta_j}, \theta_i = \omega_{\theta_i}) \times P(\theta_j = \omega_{\theta_j}) \right\}
 \end{aligned} \tag{11-3}$$

اگر عبارت ساده شده در معادله 11-3 را در 8-3 جایگزین می توان احتمال پسین را به صورت زیر بازنویسی کرد (برای دیدن نحوه رسیدن به معادله های زیر به پیوست ریاضی مراجعه شود):

$$P(\theta_i = \omega_{\alpha} | \bar{x}_{j, j \in N_i}^+, \bar{A}_{ij}, j \in N_i) = \frac{P(\theta_i = \omega_{\alpha}) \times Q(\theta_i = \omega_{\alpha})}{\sum_{\omega_{\lambda} \in \Omega} P(\theta_i = \omega_{\lambda}) \times Q(\theta_i = \omega_{\lambda})} \tag{12-3}$$

در معادله 12-3 $Q(\theta_i = \omega_{\alpha})$ را می توان به صورت زیر نوشت

$$Q(\theta_i = \omega_{\alpha}) = \prod_{j \in N_i} \left\{ \sum_{\omega_{\beta} \in \Omega} P(\theta_j = \omega_{\beta} | \bar{x}_j) \times P(\bar{A}_{ij} | \theta_i = \omega_{\alpha}, \theta_j = \omega_{\beta}) \right\} \tag{13-3}$$

این تابع، $Q(\theta_i = \omega_{\alpha})$ ، تابع حمایت نامیده می شود. مفهوم تابع حمایت¹ به این شرح است:

¹Support function

• فرض کنید که O_i برچسب ω_α را بگیرد، چقدر همسایه‌های O_i این میزان سازگاری را حمایت می‌کنند. به عبارت دیگر چقدر همسایه‌های O_i این نود را در گرفتن برچسب ω_α حمایت و پشتیبانی می‌کنند.

استفاده از معادله 3-12 می‌توان با توجه به تمامی مشاهدات در دسترس در شبکه گراف می‌توان احتمال اینکه O_i برچسب ω_α را بگیرد را محاسبه کرد. این احتمال به عنوان تابعی از احتمال شرطی برچسب‌ها روی اطلاعات یکانی $(P(\theta_j = \omega_{\theta_j} | \bar{x}_{j, j \in N_i^+}))$ و همچنین توزیع احتمالاتی شرطی از ارتباط باینری $(P(\bar{A}_{ij, j \in N_i} | \theta_i = \omega_{\theta_i}, \theta_j = \omega_{\theta_j}))$ با فرض اینکه دو نود O_i و O_j با هم منتصب شود اشتقاق پیدا می‌کند. با استفاده از قانون بی‌زین احتمال انتصاب یک برچسب به شرط ویژگی‌های یکانی به صورت زیر تعریف می‌شود:

$$P(\theta_j = \omega_{\theta_j} | \bar{x}_j) = \frac{P(\bar{x}_j | \theta_j = \omega_{\theta_j}) \times P(\theta_j = \omega_{\theta_j})}{\sum_{\omega_\alpha \in \Omega} P(\bar{x}_j | \theta_j = \omega_\alpha) \times P(\theta_j = \omega_\alpha)} \quad 14-3$$

در معادله 3-14 $P(\bar{x}_j | \theta_j = \omega_{\theta_j})$ تابع چگالی احتمالی ویژگی‌های یکانی و $P(\theta_j = \omega_{\theta_j})$ احتمال اولیه برای انتصاب می‌باشد. براساس قانون تصمیم‌گیری 7-3 برچسب نود O_i برابر است با مقدار بیشینه احتمال پسین برای این نود.

در معادله 3-14 پیش فرض احتمال اولیه $P(\theta_j = \omega_{\theta_j})$ ، شامل اطلاعات بسیار کمی می‌باشد. با این انگیزه و توجه به اینکه که ویژگی‌های یکانی دارای خطا می‌باشند و همچنین مدل کردن تابع توزیع چگالی برای آنها بسیار پیچیده است، لذا محاسبات با استفاده از معادله 3-14 با خطا همراه می‌باشد. بنابراین به روز کردن احتمال انتصاب‌ها با استفاده از 3-14 کارا و قابل اعتماد برای انتصاب‌ها در قانون 3-7 نیست.

تکنیک آرام‌سازی روشی را پیشنهاد می‌کند تا مدیریت انتصاب با وضوح و مقاومت بیشتری انجام شود و درعین حال خطا را نیز تا حد امکان کم کند.

برای این کار ابتدا احتمال انتصاب ها را با استفاده از قانون 3-14 مقدار دهی اولیه می‌کنیم و سپس با استفاده از معادلات 3-12 و 3-13 به صورت تکراری به روز رسانی می‌کنیم. احتمال انتصاب ها در هر مرحله با استفاده از همسایه ها دوباره به روز رسانی می‌شود. این فرآیند می‌تواند هم به تعداد مشخص تکرار شود و هم تا زمانی که شرط سازگاری برقرار شود و اصطلاحاً فرآیند همگرا شود ادامه پیدا کند. براساس این استراتژی معادلات به روز رسانی 3-12 و 3-13 را می‌توان به صورت زیر به عنوان معادلات نهایی برای تکنیک احتمالاتی ارائه کرد:

$$P^{(n+1)}(\theta_i = \omega_{\theta_i}) = \frac{P^{(n)}(\theta_i = \omega_{\theta_i}) \times Q^{(n)}(\theta_i = \omega_{\theta_i})}{\sum_{\omega_{\lambda} \in \Omega} P^{(n)}(\theta_i = \omega_{\lambda}) \times Q^{(n)}(\theta_i = \omega_{\lambda})} \quad 15-3$$

$$Q^{(n)}(\theta_i = \omega_{\alpha}) = \prod_{j \in N_i} \left\{ \sum_{\omega_{\beta} \in \Omega} P^{(n)}(\theta_i = \omega_{\beta}) \times P(\bar{A}_{ij} | \theta_i = \omega_{\alpha}, \theta_j = \omega_{\beta}) \right\} \quad 16-3$$

در اولین تکرار $P^{(0)}(\theta_i = \omega_{\theta_i})$ مقداردهی اولیه می‌شود و سپس تابع چگالی احتمالاتی برای ویژگی‌های یکانی با استفاده از معادله 3-14 محاسبه می‌شود. بدون هیچ اطلاعات مشاهده احتمال انتصاب ها مساوی در نظر گرفته می‌شود. در [76] غرض بر این است که تابع توزیع احتمالاتی خطا برای ویژگی‌های یکانی دارای توزیع گوسی و به صورت آماری مستقل هستند بر این اساس تابع توزیع ویژگی‌های یکانی به صورت زیر تعریف می‌شود:

$$P(\bar{x}_i | \theta_i = \omega_{\alpha}) = N_{\bar{x}_i}(\bar{x}_i, \Sigma_u) \quad 17-3$$

در معادله 3-17 Σ_u یک ماتریس کورایانس قطری برای بردار \bar{x}_i است و \bar{x}_α ویژگی یکانی متناظر با نود در گراف با برچسب ω_α می‌باشد. در تابع حمایت Q جمله $P(\bar{A}_{ij}|\theta_i = \omega_\alpha, \theta_j = \omega_\beta)$ مانند یک سازگار کننده ضرایب در سایر روش‌های آرام سازی رفتار می‌کند. به عبارت دیگر Q یک تابع توزیع احتمالاتی برای ویژگی باینری \bar{A}_{ij} است زمانی که $\theta_i = \omega_\alpha, \theta_j = \omega_\beta$ تطبیق پیدا می‌کند

$$P(\bar{A}_{ij}|\theta_i = \omega_\alpha, \theta_j = \omega_\beta) = N_{\bar{A}_{ij}}(\bar{A}_{\alpha\beta}, \Sigma_b) \quad 18-3$$

مشابه با ویژگی های یکانی فرض بر این است که انحراف از ویژگی های باینری در نود های گرفته شده از صحنه که در تناظر با نودهای گراف مدل هستند نیز از یک توزیع گوسی به صورت زیر تبعیت می‌کنند در معادله 3-18 Σ_b ماتریس کورایانس برای ویژگی های باینری بردار \bar{A}_{ij} می‌باشد. زمانی که در نمونه خاص حداقل یکی از دو شی در این توزیع احتمالاتی برچسب پوچ را بگیرد، توزیع معرفی شده کارا نخواهد بود. در این وضعیت جایی که هیچ ارتباطی بین نود ساختگی و دیگر نودها برقرار نباشد، استفاده از توزیع یکنواخت به عنوان تابع چگالی توزیع احتمالاتی کارساز است [76].

اگر فرض کنیم که اجزای ویژگی های باینری \bar{A}_{ij} از هم مستقل هستند، می‌توان تابع چگالی توزیع احتمالاتی برای آن را به صورت حاصلضرب توزیع هر یک از اجزای آن دانست. به صورت زیر:

$$P(\bar{A}_{ij}|\theta_i = \omega_\alpha, \theta_j = \omega_\beta) = \prod_{k=1}^{\|\bar{A}_{ij}\|} U(m_k, D_k) \quad 19-3$$

توزیع $U(m_k, D_k)$ متناسب است با k امین جز از بردار \bar{A}_{ij} است که دارای توزیع یکنواخت به مرکزیت m_k و محدوده غیر صفر D_k است. پارامترهای m_k و D_k میانگین و محدوده k امین جز می باشد. فرآیند تکراری زمانی خاتمه خواهد یافت که یکی از شرایط زیر برقرار باشد.

- در تکرار آخر هیچ احتمالی بیشتر از ϵ تغییر نکند.
- تعداد مشخصی تکرار اتفاق بیفتد.

3-4 خلاصه و نتیجه گیری

با توجه به تمامی مطالب گفته شده در این پایان نامه روند ادامه کار به صورت زیر است:

ابتدا برای هر تصویر ورودی توصیفگر SIFT را روی تصویر اعمال می‌کنیم و نقاط کلیدی را استخراج می‌کنیم. در ادامه برای هر نقطه کلیدی k نزدیک‌ترین همسایه را به عنوان همسایگان نقطه مفروض در نظر می‌گیریم و این ساختار را سپس در گراف ویژه رابطه ایی تحت عنوان گراف ویژه رابطه ایی برای صحنه در نظر می‌گیریم. با روندی مشابه برای هر شیء در پایگاه داده، این کار را انجام داده و آنها را در گراف ARG قرار می‌دهیم و در نهایت برای تطبیق از روش آرام سازی احتمالاتی استفاده می‌کنیم.

دلیل استفاده از روش آرام سازی احتمالاتی استراتژی واضح برای فرایند آرام‌سازی و استفاده از ارتباطات باینری در فرآیند تطبیق است.

یک مسئله رایج گاه در آن سطحی که طرح شده نمی توان حل کرد

آلبرت ایشتین ۱۸۷۹-۱۹۷۵

فصل چهارم

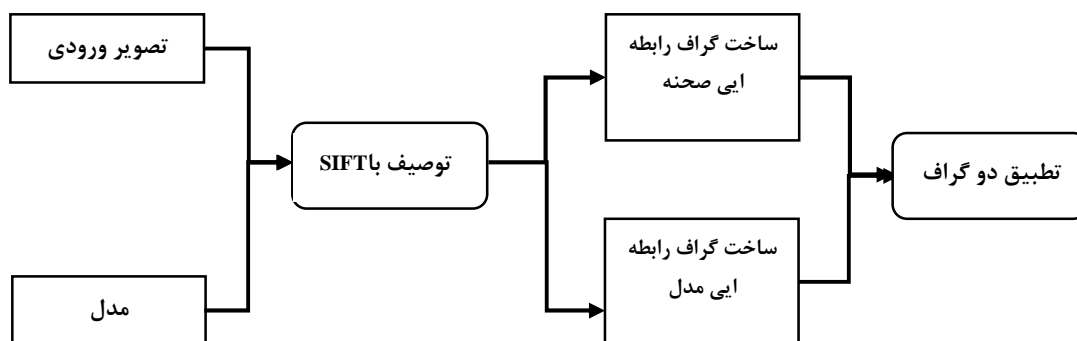
نتایج پیاده سازی محیط شبیه سازی شده

فصل چهارم

4-1 محیط شبیه سازی شده

هدف از ایجاد یک محیط کنترل شده برای بررسی دقیق الگوریتم تطبیق از جمله: تحمل الگوریتم در برابر نویز و همچنین تخمین درست محدوده پارامترهای موثر در الگوریتم است. در دنیای واقعی مشکلات زیادی برای بررسی پارامترها وجود دارد. در دنیای واقعی به جای بررسی ماهیت الگوریتم تطبیق مدت زمان زیادی باید صرف تنظیم همین پارامترها شود. از طرفی از آنجا که در دنیای واقعی نمی توان به صورت کمی بازدهی سیستم را بررسی کرد لذا این کار را باید به صورت چشمی انجام شود. از طرفی با مطالعه ای که در منابع مختلف انجام شد وجود یک محیط کنترل شده برای نیل به اهداف بالا یافته نشد لذا در این فصل سعی خواهد شد این محیط فراهم شود تا تحمل الگوریتم تطبیق در این فضا کنترل شده بررسی شود. در فصل بعد یافته های بدست آمده از این محیط در چند سناریوی واقعی بررسی خواهد شد.

نمای کلی سیستم مورد نظر در این پایان نامه به صورت شکل زیر است



شکل 4-1: نمای کلی الگوریتم بازشناسی ارائه شده

4-1-1 تعریف محیط شبیه سازی شده

این محیط شامل گراف مدل و صحنه است. گراف مدل، مدل یک الگو را ارائه می‌نماید و گراف صحنه محتویات صحنه را ارائه می‌کند. در گراف مدل تعدادی نود مشخص وجود دارد. تعداد نودها در گراف مدل یکی از پارامترهای محیط است و می‌تواند تغییر کند. برای هر نود در گراف مدل یک بردار در نظر گرفته می‌شود که توصیف کننده این نود است (ویژگی یکانی). طول و مقادیر این بردار در شبیه سازی به عنوان پارامتر در نظر گرفته می‌شود.

در نودهای گراف مدل، همسایگی نیز تعریف می‌شود. دو نود از گراف می‌توانند بر اساس معیاری باهم همسایه باشند. برای مثال در محیط واقعی می‌توان k نزدیک‌ترین همسایه‌ی هر نود را (با استفاده از فاصله اقلیدسی) همسایه آن نود در نظر گرفت. یکی از معیارها در محیط شبیه سازی می‌تواند انتخاب تصادفی باشد. از آنجا که در محیط پیشنهادی فضا کاملاً تصادفی است، لذا از یک معیار کاملاً تصادفی برای انتخاب همسایه (همسایه‌ها) برای یک نود استفاده می‌شود. تعداد این همسایه‌ها محدود در یک بازه در نظر گرفته می‌شود.

همانطور که قبلاً بیان شد ارتباط بین دو همسایه با استفاده از ویژگی باینری توصیف می‌شود. برای شبیه سازی این بردار همانند بردار ویژگی یکانی که هر نود را توصیف می‌کند عمل می‌شود. این بردار دارای طول و مقادیر مشخص است و توصیف کننده همسایگی بین دو نود مشخص در مدل می‌باشد.

دقت شود خاصیت همسایگی دو طرفه است یعنی اگر نود x با نود y همسایه باشد، آنگاه نود y نیز با نود x همسایه است. این از ویژگی‌های محیط واقعی است. در محیط واقعی همچنین یک نود با خودش همسایه نیست. این ویژگی‌ها در محیط شبیه سازی شده کاملاً رعایت شده و گرافی که در نظر گرفته می‌شود، یک گراف کاملاً بی جهت و بدون حلقه است.

با توجه به تمام توضیحات بالا پارامترهای محیط شبیه سازی شده به صورت زیر است:

- تعداد نودهای گراف مدل: این تعداد به عنوان پارامتر مدل قابل تغییر است. در آزمایشها این تعداد برابر 10 در نظر گرفته شده است

- مقادیر و طول بردار یکانی: طول این بردار و مقادیر آن کاملاً قابل تغییر است که در آزمایشات طول بردار 20 و مقادیر هر بردار از یک توزیع نرمال با میانگین صفر و واریانس 1 می باشد.

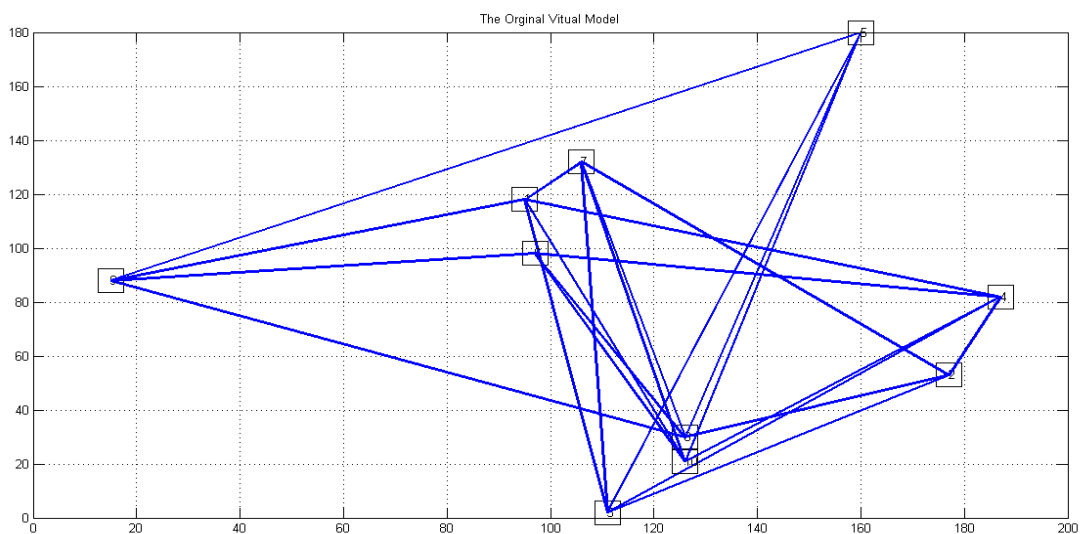
$$x_j \sim \mathcal{N}(0,1) \quad j \in (1 \dots (\text{size unary measurement})) \quad 1-4$$

در تعریف بالا منظور از x_j ، j امین عنصر از بردار یکانی است که دارای یک توزیع نرمال با میانگین صفر و واریانس 1 است.

- مقادیر و طول بردار ویژگی باینری: طول این بردار و مقادیر آن کاملاً قابل تغییر است که در آزمایشات طول بردار 20 و مقادیر هر بردار از یک توزیع نرمال با میانگین صفر و واریانس 1 می باشد. دقت شود نام باینری برای این بردار بخاطر این است که توصیفگر ارتباط بین دو نود گراف است. در حالی که عناصر این بردار مقادیر حقیقی (نه باینری) می باشد.

$$x_{ij} = x_{ji} \sim \mathcal{N}(0,1) \quad j \in (1 \dots \text{size}(Binary\ measurement)) \quad 2-4$$

اگر مجموعه همسایه های نود i را با N_i نشان دهیم نود j جز همسایه های نود i می باشد یعنی $j \in N_i$ در تعریف بالا منظور از x_{ij} ، ارتباط بین i امین و j امین نود در مدل است که دارای یک توزیع نرمال بامیانگین صفر و واریانس 1 است. این ارتباط متقارن است.



شکل 4-2: مدل استاندارد

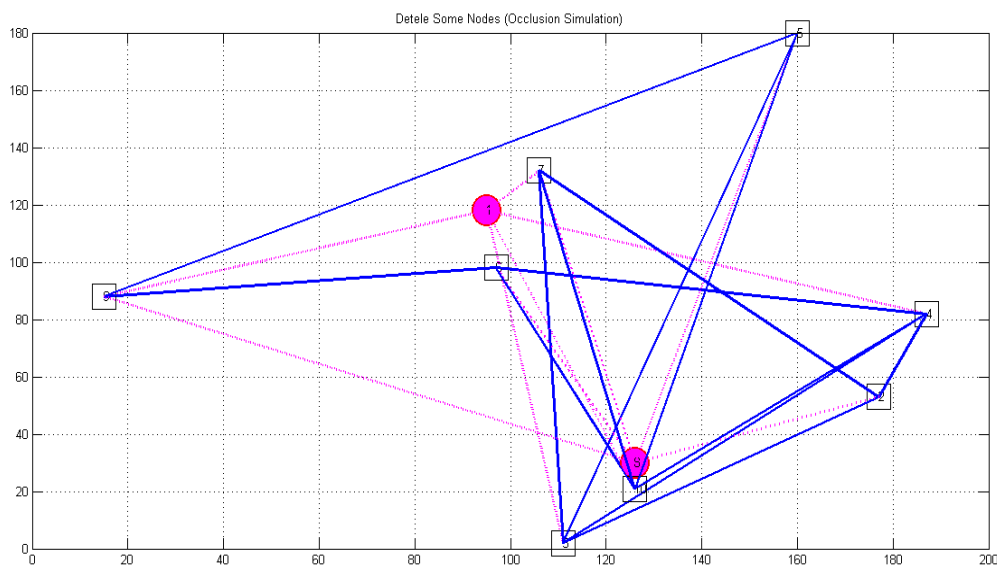
- آخرین پارامتر در این محیط میزان همسایه‌هاست که تعداد آن‌ها متغیر ولی محدود در یک بازه در نظر گرفته می‌شود.

با توجه به توضیحات بالا و با در نظر گرفتن یک مختصات تصادفی برای هر نود، شکل بالا (شکل 4-1) نمونه‌ای از یک مدل ساختگی را نشان می‌دهد. برای شبیه‌سازی گراف صحنه، استراتژی مشابهی را در نظر گرفته می‌شود. در دنیای واقعی صحنه دارای ویژگی‌های جدیدی نیز می‌باشد. در واقع ممکن است از لحاظ نقطه دید ناظر اتفاقاتی در صحنه صورت گیرد که تعدادی از مهم‌ترین آن‌ها به صورت زیر است: انسداد، چرخش و نرخ شلوغی.

4-1-1-1-1-انسداد

این تغییر زمانی در یک محیط واقعی اتفاق می‌افتد که ناظر قسمتی از یک شیء را در اثر شلوغی در صحنه نمی‌بیند. این اتفاق بسیار مرسوم است به این دلیل که نقطه دید برای یک ناظر محدود است. در صحنه ساختگی حذف یک نود معادل حذف تمامی ارتباطات بین آن نود و نودهای همسایه می‌باشد. با این توجیح

انسداد را می‌توان حذف یک یا چند نود شبیه سازی کرد که ارتباطات میان آن نود(نودها) و همسایه‌ها نیز حذف شده باشد. شکل 2-4 نمونه‌ای از عمل انسداد را نشان می‌دهد.



شکل 3-4: شبیه‌سازی انسداد در مدل شکل 1-4

در شکل بالا در اثر انسداد، نودهای 1 و 2 حذف شده‌اند. این نودها در شکل با رنگ قرمز و شکل دایره مشخص شده‌اند.

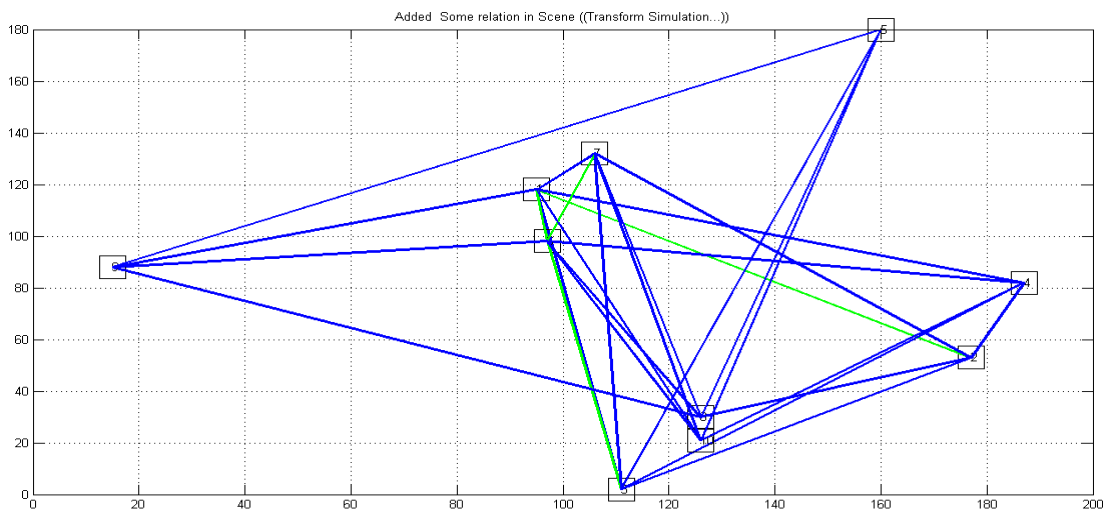
4-1-1-2 چرخش

تغییر دیگری که ممکن است در صحنه اتفاق بیفتد، تغییراتی است که باعث می‌شود شکل تا حدی بچرخد (شکل 3-4). در این شکل، سمت چپ بر اثر چرخش به حالت شکل سمت راست درآمده است. در واقع اگر دقت شود تعدادی از نقاطی که در مدل همسایه بوده‌اند ممکن است در اثر این چرخش دیگر همسایه نباشند. این تغییر را می‌توان با حذف تعدادی از ارتباطات بین نودها در گراف شبیه‌سازی کرد. شکل 4-4 این تغییر را در محیط شبیه‌سازی شده نشان می‌دهد.



شکل 4-4: نمایی از یک شیء در دو حالت در اثر چرخش

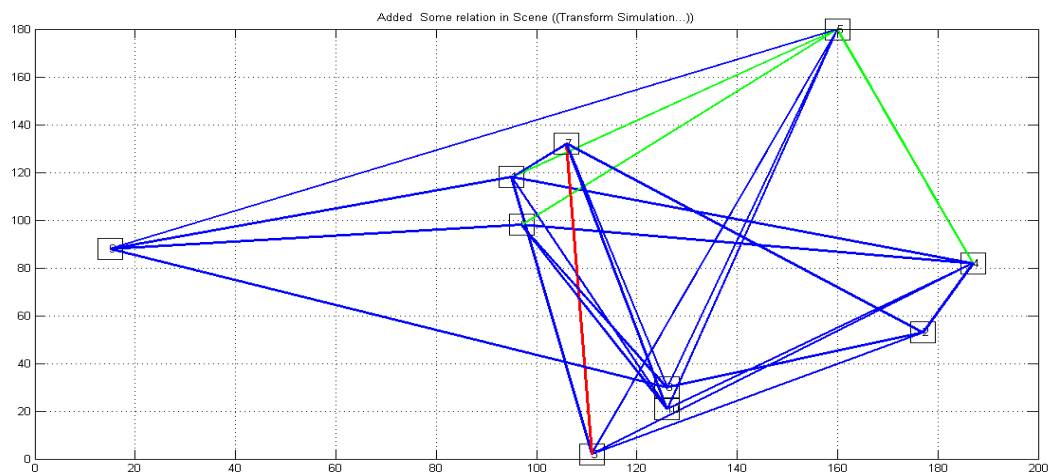
همانطور که در شکل 4-4 مشخص است برخی ارتباطات حذف شده است که این ارتباطات حذف شده با رنگ قرمز مشخص شده است. به طور مشابه، این تغییر ممکن است برعکس هم باشد، به عبارت بهتر تعدادی از نودهایی که همسایه نبوده‌اند در اثر این اتفاق همسایه شوند. این تغییرات را می‌توان با همسایه کردن برخی از نودها به طور تصادفی که همسایه نیستند شبیه سازی کرد.



شکل 4-5: شبیه سازی برخی تغییرات در اثر چرخش

نمونه ایی از این کار را در شکل 4-5 مشاهده می‌کنید. در این شکل ارتباطات سبز رنگ نشان دهنده وجود یک ارتباط در اثر چرخش می‌باشد. کنار هم قرار گرفتن حذف و اضافه شدن تعدادی از ارتباطات به صورت

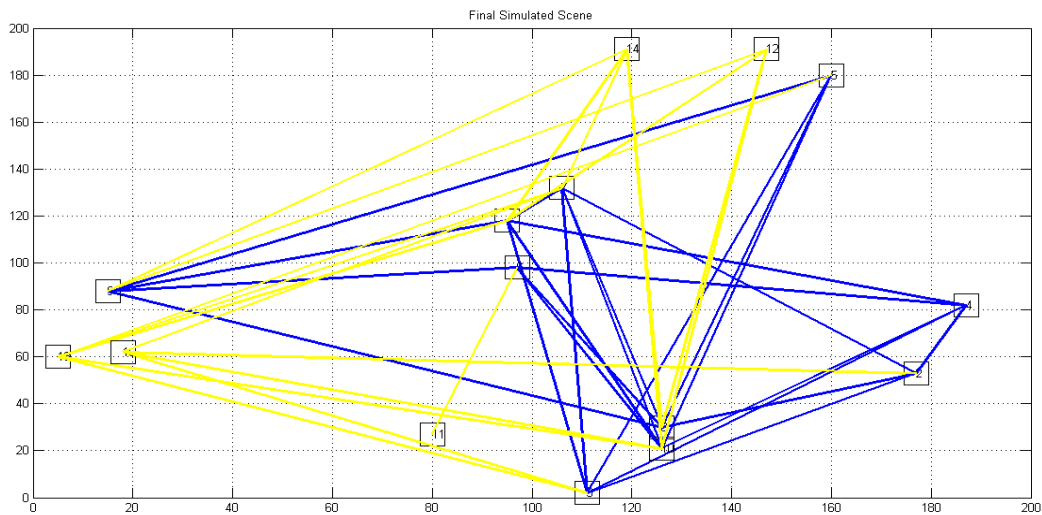
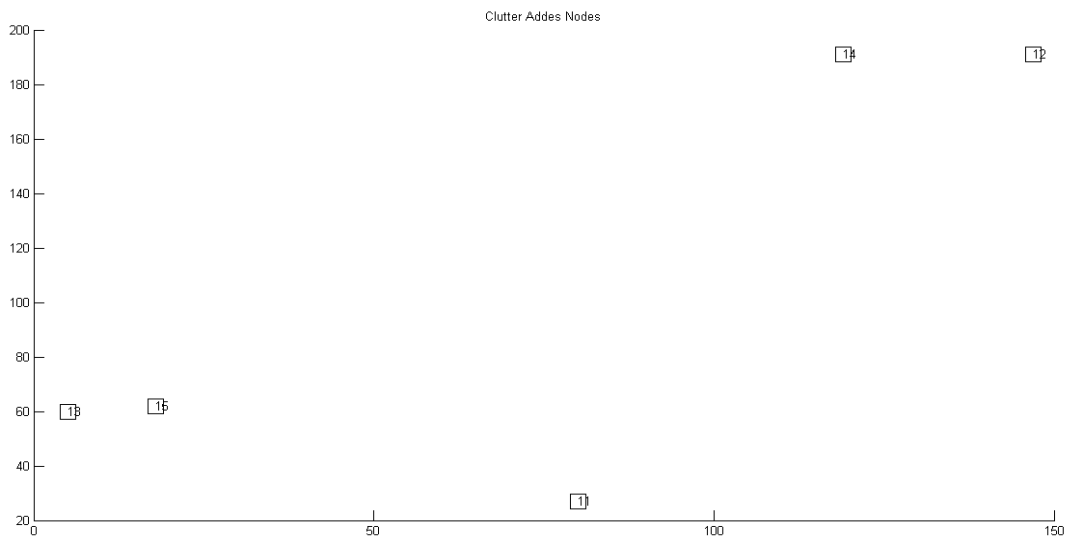
تصادفی، شبیه سازی این تغییرات را کامل می‌کند. در شکل 4-6 نمونه ای از کنار هم قرار گرفتن این تغییرات نشان داده شده است.



شکل 4-6: شبیه سازی برخی تغییرات در اثر چرخش

4-1-1-3 نرخ شلوغی

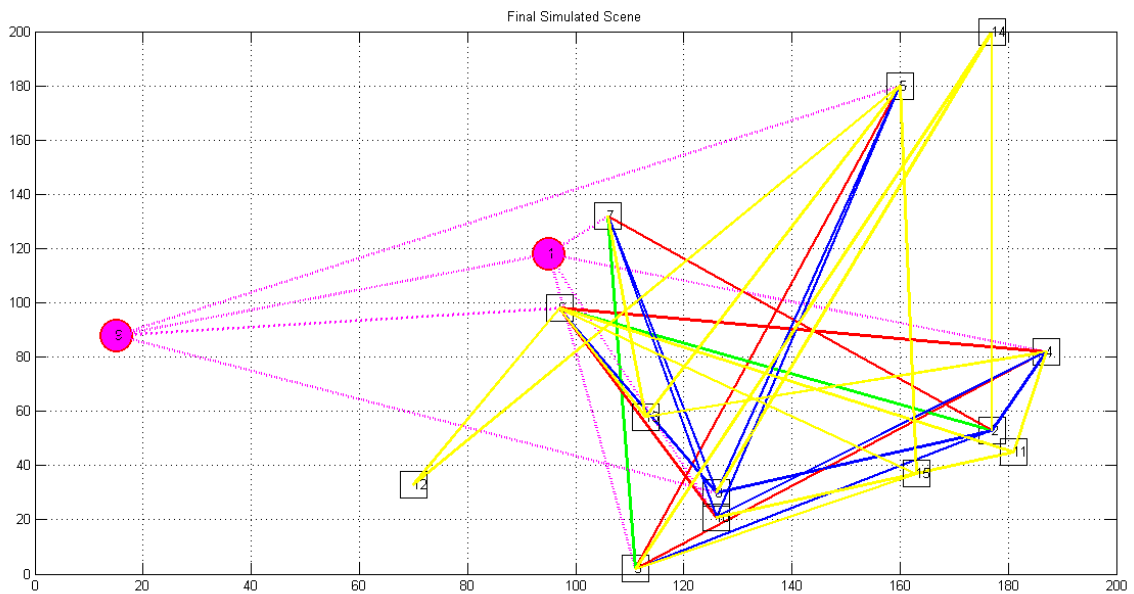
در صحنه‌های واقعی همواره در کنار شیء مورد نظر، سایر اشیا که مورد توجه نیستند وجود دارند. این اتفاق سبب می‌شود که صحنه اصطلاحاً شلوغ باشد. در صحنه‌های دنیای واقعی این اتفاق کاملاً متداول می‌باشد. در نهایت هدف، شبیه سازی دنیای واقعی است. آخرین مرحله قرار دادن مدل در یک صحنه شلوغ است. برای شبیه سازی این فرآیند ابتدا تعدادی نود غیرمرتبط در صحنه ایجاد می‌شود، سپس به طور کاملاً تصادفی ارتباطاتی بین نودهای اضافه شده و همچنین نودهای مدل ایجاد می‌کنیم. نمونه‌ای از این فرآیند را در شکل 4-7 مشاهده می‌شود



شکل 4-7: نمونه ای از شلوغ شدن صحنه، شکل بالا نودهای اضافه شده و پایین ارتباط بین نودهای اضافه شده نودها مدل

پس از تمامی این تغییرات، صحنه‌ای که به دنیای واقعی نزدیک شده را می‌توان ساخت. شکل 4-8 زیر

مجموعه نمونه ای است که حاصل ترکیب تمامی تغییرات ذکر شده در مدل است.



شکل 4-8: شکل نهایی گراف صحنه پس از اعمال تغییرات ذکر شده

پس از ساخت صحنه و مدل در محیط شبیه سازی، برای گزارش نتایج دو حالت را در نظر گرفته می‌شود:

- حالتی که در آن هیچ نویزی به بردارهای ویژگی یکانی و باینری صحنه اضافه نشده باشد
- حالتی که در آن به بردارهای ویژگی یکانی و باینری نویز وارد شود.

در هر حالت تحمل الگوریتم تطبیق را با استفاده دو روش معرفی شده توسط احمدی فرد [2] و Kostin [8] بررسی خواهد شد. روش احمدی فرد را روش اول و روش Kostin را روش دوم نامگذاری می‌شود. قبل از ادامه بحث ابتدا مرور بسیار مختصری بر روش اول و دوم خواهیم داشت .

4-2 مروری کوتاه بر روش اول و دوم

روش اول از فرآیند آرام‌سازی احتمالاتی برای تطبیق استفاده می‌کند. برای هر مدل در پایگاه داده یک گراف ARG ساخته و به طور مشابه همین کار را برای صحنه انجام می‌شود. این روش سپس برای عدم وجود یک نود گراف صحنه در مدل نیز راهکاری ارائه می‌دهد. هرگاه روش اول نتواند برای نودی از صحنه

برچسب مناسب در مدل بیابد، نود مذکور را با برچسب ساختگی پوچ برچسب گذاری می‌کند. این نود متناظر با هیچ جزئی از محیط واقعی مدل نیست. با این استراتژی می‌توان عدم حضور یک نود از مدل در صحنه را نیز توجیه کند. ممکن است از یک مدل چند نمونه در صحنه وجود داشته باشد یا اصلا وجود نداشته باشد. این به معنی غیر متقارن بودن گراف مدل و صحنه است.

روش دوم عنوان می‌کند که روش آرام سازی احتمالاتی کاملا متقارن بوده و لذا وجود نود پوچ به معنی برهم زدن این خاصیت است. با این توجیه ، وجود این نود را منتفی دانسته و آن را حذف می‌کند. استنتاج این روش این است که هر نود صحنه باید با یک و تنها یک نود از مدل متناظر باشد. این دو روش در تابع حمایت نیز با هم تفاوت دارند.

$$Q^{(n)}(\theta_i = \omega_\alpha) = \prod_{j \in N_i} \left\{ \sum_{\omega_\beta \in \Omega} P^{(n)}(\theta_i = \omega_\beta) \times P(\bar{A}_{ij} | \theta_i = \omega_\alpha, \theta_j = \omega_\beta) \right\} \quad 1-4$$

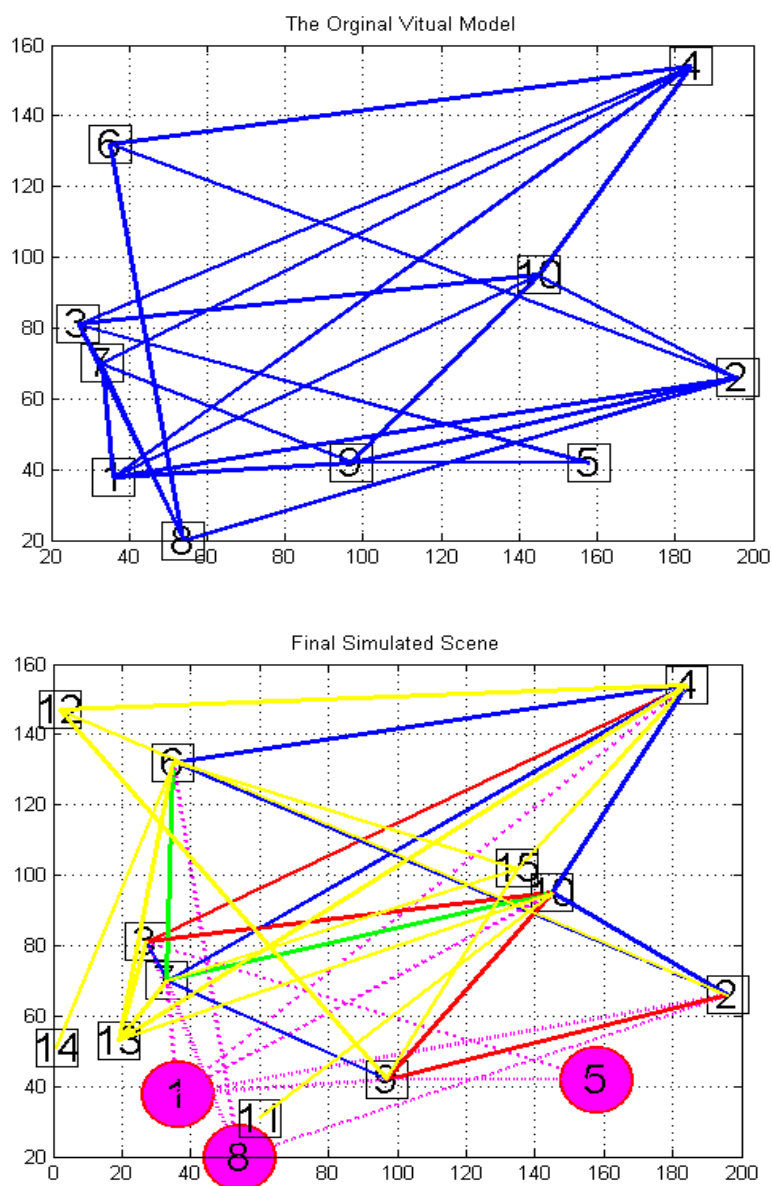
$$Q_{ij} = p_{ij}^n \left(1 - \prod_{(n'_i, n'_k) \in B} (1 - p_{ijk}^{nn}) \right) \quad 2-4$$

معادله 1-4 تابع حمایت برای روش اول و معادله 2-4 تابع حمایت برای روش دوم را نشان می‌دهد. در ادامه این دو روش ، یکبار وقتی هیچ نویزی به شبکه گراف وارد نشده باشد و بار دیگر زمانی که نویزی به گراف وارد شده باشد، بررسی خواهد شد.

3-4 روش اول و در نظر نگرفتن نویز در محیط شبیه سازی

در فصل مباحث نظری ذکر شد که تنها استفاده از بردار ویژگی یکانی برای تطبیق نمی‌تواند کارایی مناسبی داشته باشد. برای اینکه نشان داده شود استفاده از اطلاعات همسایگی قادر است اشتباهاتی که در اثر

شبهات بین ویژگی یکانی نودها اتفاق می افتد را در حین فرآیند تطبیق تصحیح کنند استراتژی زیر در نظر گرفته می شود. به طور تصادفی برخی بردارهای ویژگی یکانی از نودهای گراف مدل را مشابه با تعدادی از نودهای اضافه شده در صحنه در نظر گرفته می شود. سپس آزمایش خواهد شد که روش اول چقدر توانایی تصحیح این اشتباه را دارد. شکل 4-9 صحنه و مدل ایجاد شده را برای اولین آزمایش نشان می دهد.



شکل 4-9: صحنه و مدل ایجاد شده برای آزمایش اول

جدول 4-1: ماتریس سردر گمی

	Label1	Label2	Label3	Label4	Label5	Label6	Label7	Label8	Label9	Label10	NULL
Node1	0	0	0	0	0	0	0	0	0	0	1
Node2	0	1	0	0	0	0	0	0	0	0	0
Node3	0	0	1	0	0	0	0	0	0	0	0
Node4	0	0	0	1	0	0	0	0	0	0	0
Node5	0	0	0	0	0	0	0	0	0	0	1
Node6	0	0	0	0	0	1	0	0	0	0	0
Node7	0	0	0	0	0	0	1	0	0	0	0
Node8	0	0	0	0	0	0	0	0	0	0	1
Node9	0	0	0	0	0	0	0	0	1	0	0
Node10	0	0	0	0	0	0	0	0	0	1	0
Node11	0	0	0	0	0	0	0	0	1	0	0
Node12	0	0	0	0	0	0	0	0	0	0	1
Node13	0	0	0	0	0	0	0	0	0	0	1
Node14	0	0	0	0	0	0	0	0	0	0	1
Node15	0	0	0	0	0	0	0	0	0	0	1

برای آزمایش اول همانطور که در جدول 4-1 مشخص است، روش اول توانسته است که به درستی حذف نود 1، 8 و 5 را گزارش کند. برچسب نودهای 1، 8 و 5 از صحنه، نود پوچ (NULL) از گراف مدل است. از طرفی اگر به جدول 4-2 دقت کنید متوجه خواهید شد در حالی که در بردار ویژگی یکانی اشتباه وجود دارد روش اول توانسته است در حین فرآیند آرام سازی این اشتباه را اصلاح کند، در جدول 4-1 اگر دقت شود نود 11 از صحنه به نود 9 از مدل برچسبی خورده است که این اشتباه می باشد. با توجه به جدول 4-2 بررسی های یکانی در نودهای 11 از صحنه و نود 9 از مدل مشابه است از طرفی وقتی همسایه های نود 11 بررسی شد متوجه شدیم که این نود تنها یک همسایه دارد که این همسایه جزه همسایه های نود 9 نیز می باشد با این توجیه این اشتباه اجتناب ناپذیر است. دقت شود که برای محاسبه مقادیر جدول 4-2 فاصله اقلیدسی بین بردار ویژگی هر نود صحنه با تمامی نودهای مدل محاسبه می شود این مقدار برای نود پوچ 1.5 برابر جمع احتمال های تمامی نودهای در نظر گرفته می شود.

جدول 2-4: ماتریس شباهت‌های نود ها در مدل و صحنه شکل 4-9 تنها با از ویژگی‌های یکانی

	Label1	Label2	Label3	Label4	Label5	Label6	Label7	Label8	Label9	Label10	NULL
Nodel1	1	1,51E-19	7,05E-15	3,62E-18	2,76E-08	1,71E-14	4,3E-13	1,92E-11	1,17E-22	4,31E-19	1,5
Nodel2	1,51E-19	1	5,74E-16	1,51E-18	5,2E-20	6,26E-10	1,16E-18	7,72E-18	2,33E-18	4,54E-18	1,5
Nodel3	7,05E-15	5,74E-16	1	1,37E-15	1,15E-14	3,08E-19	1,98E-12	7,17E-14	1,33E-15	1,1E-07	1,5
Nodel4	3,62E-18	1,51E-18	1,37E-15	1	7,59E-14	2,31E-17	9,08E-15	5,31E-07	4,35E-10	2,47E-17	1,5
Nodel5	2,76E-08	5,2E-20	1,15E-14	7,59E-14	1	8,93E-12	3,67E-09	7,14E-14	5,99E-24	2,75E-16	1,5
Nodel6	1,71E-14	6,26E-10	3,08E-19	2,31E-17	8,93E-12	1	9,1E-12	1,88E-19	9,29E-21	1,41E-20	1,5
Nodel7	4,3E-13	1,16E-18	1,98E-12	9,08E-15	3,67E-09	9,1E-12	1	9,47E-19	2,62E-17	9,18E-15	1,5
Nodel8	1,92E-11	7,72E-18	7,17E-14	5,31E-07	7,14E-14	1,88E-19	9,47E-19	1	6,23E-12	3,14E-16	1,5
Nodel9	1,17E-22	2,33E-18	1,33E-15	4,35E-10	5,99E-24	9,29E-21	2,62E-17	6,23E-12	1	2,62E-17	1,5
Nodel10	4,31E-19	4,54E-18	1,1E-07	2,47E-17	2,75E-16	1,41E-20	9,18E-15	3,14E-16	2,62E-17	1	1,5
Nodel11	1,17E-22	2,33E-18	1,33E-15	4,35E-10	5,99E-24	9,29E-21	2,62E-17	6,23E-12	1	2,62E-17	1,5
Nodel12	3,62E-18	1,51E-18	1,37E-15	1	7,59E-14	2,31E-17	9,08E-15	5,31E-07	4,35E-10	2,47E-17	1,5
Nodel13	4,31E-19	4,54E-18	1,1E-07	2,47E-17	2,75E-16	1,41E-20	9,18E-15	3,14E-16	2,62E-17	1	1,5
Nodel14	4,3E-13	1,16E-18	1,98E-12	9,08E-15	3,67E-09	9,1E-12	1	9,47E-19	2,62E-17	9,18E-15	1,5
Nodel15	2,76E-08	5,2E-20	1,15E-14	7,59E-14	1	8,93E-12	3,67E-09	7,14E-14	5,99E-24	2,75E-16	1,5

احتمال‌های پسین و همچنین تابع حمایت برای روش اول به ترتیب در جداول 3-4 و 4-4 گزارش شده است.

جدول 3-4: احتمال‌های پسین برای روش اول

	Label1	Label2	Label3	Label4	Label5	Label6	Label7	Label8	Label9	Label10	NULL
Nodel1	0	0	0	0	0	0	0	0	0	0	1
Nodel2	0	1	0	0	0	0	0	0	0	0	1,1E-163
Nodel3	0	0	1	0	0	0	0	0	0	0	1,5E-157
Nodel4	0	0	0	1	0	0	0	0	0	0	1,5E-157
Nodel5	0	0	0	0	0	0	0	0	0	0	1
Nodel6	0	0	0	0	0	1	0	0	0	0	8,6E-160
Nodel7	0	0	0	0	0	0	1	0	0	0	2,8E-154
Nodel8	0	0	0	0	0	0	0	0	0	0	1
Nodel9	0	0	0	0	0	0	0	0	1	0	6,2E-149
Nodel10	0	0	0	0	0	0	0	0	0	1	8,3E-153
Nodel11	0	0	0	0	0	0	0	0	1	0	1,63E-07
Nodel12	0	0	0	0	0	0	0	2,2E-236	0	0	1
Nodel13	0	0	8,1E-304	0	0	0	0	0	0	2,96E-26	1
Nodel14	0	0	0	0	0	0	0	0	0	0	1
Nodel15	5,7E-306	0	0	0	0,000213	0	0	0	0	0	0,999787

همانطور که از جدول 3-4 مشخص است، روش اول با استراتژی ایجاد نود ساختگی توانسته است به طور مطلوبی احتمال‌های مناسبی را به هریک از نودهای صحنه تخصیص دهد و در حین فرآیند آرام سازی بتواند این احتمال‌ها را تضعیف یا قوی کند. در جدول 4-4 نیز میزان حمایت هر نود از برچسب‌های مدل برای روش اول نشان داده شده است.

اگر در جدول 4-4 دقت کنید متوجه خواهید شد که در عناصر روی قطر اصلی این جدول در سطرهای یک تا 10 حمایت بسیار مناسبی از نودهای صحنه به درستی شده است و این حمایت در فرآیند آرام سازی توانسته است سبب برچسب خوردن درست نود در صحنه شود.

جدول 4-4: تابع حمایت برای صحنه و مدل شکل 9-4 با روش اول

	Label1	Label2	Label3	Label4	Label5	Label6	Label7	Label8	Label9	Label10
Nodel1	0	0	0	0	0	0	0	0	0	0
Nodel2	8,53E-23	0,021333	1,91E-20	5,58E-24	3,44E-32	3,73E-13	1,83E-23	6,65E-21	5,65E-23	8,32E-22
Nodel3	8,76E-20	1,71E-23	0,016	2,64E-21	5,39E-29	2,51E-40	2,17E-23	4,51E-20	5,64E-22	6,38E-23
Nodel4	8,27E-22	7,57E-22	1,72E-19	0,016003	2,93E-24	9,86E-22	9,62E-20	1,26E-12	6,95E-14	1,02E-21
Nodel5	0	0	0	0	0	0	0	0	0	0
Nodel6	1,02E-17	7,71E-14	8,28E-23	6,37E-22	1,33E-24	0,017778	2,68E-16	2,45E-23	3,48E-24	5,65E-23
Nodel7	5,39E-16	1,65E-22	8,05E-16	1,85E-20	3,08E-15	3,83E-16	0,013714	1,23E-21	2,31E-20	3,27E-19
Nodel8	0	0	0	0	0	0	0	0	0	0
Nodel9	7,27E-26	7,41E-30	5,45E-21	1,67E-14	6,43E-37	2,61E-23	5,2E-23	4,19E-17	0,010671	3,6E-20
Nodel10	6,43E-22	6,9E-24	1,69E-12	9,81E-22	6,69E-23	1,13E-22	2,43E-18	1,85E-19	6,07E-23	0,012803
Nodel11	4,22E-29	4,92E-23	1,26E-21	8,52E-13	3,51E-35	4,53E-40	5,04E-31	1,09E-25	1,39E-05	2,73E-30
Nodel12	1,48E-20	3,73E-25	6,54E-22	2,24E-13	1,87E-17	3,31E-21	1,54E-18	2,1E-10	6,43E-15	9,96E-21
Nodel13	5,2E-24	2,49E-22	9,12E-12	1,04E-21	4,69E-24	1,91E-25	1,75E-18	7,08E-20	3,32E-22	3,04E-06
Nodel14	3,81E-24	7,37E-25	1,63E-26	1,24E-23	3,02E-35	6,03E-27	7,34E-21	6,64E-27	2,04E-32	5,86E-31
Nodel15	8,48E-12	2,52E-27	2,8E-19	2,77E-18	8,49E-06	9,97E-19	1,98E-13	1,94E-18	7,15E-27	1,01E-19

در سطرهای 11 تا 15 که نودهای اضافه شده به مدل اصلی می‌باشد حمایت از نودهای به شدت پایین بوده و میزان احتمال اولیه که به نود پوچ داده شده است سبب شده که به درستی نود پوچ در این حالت‌ها برنده شود. نود 9 به دلیل انسداد از صحنه حذف شده است، لذا این نود هیچ همسایه ای ندارد

که از آن حمایت کند. به عبارت بهتر هویت نود 9 از بین رفته است لذا حمایت برای آن معنا پیدا نمی- کند، به این دلیل تابع حمایت در سطر نهم در تمامی درایه های خود عدد صفر را می بیند.

4-4 روش دوم بدون در نظر گرفتن نويز در محیط شبیه سازی

در ادامه مشابه همین کار را برای روش دوم انجام خواهیم داد و ماتریس احتمال پسین و تابع حمایت بدست آمده از این روش برای صحنه و مدل شکل 4-9 نشان خواهیم داد. اگر به دو جدول 4-5 و 4-6 دقت شود ضعف عدم وجود نود پوچ را در مدیریت برچسب، زمانیکه از روش دوم برای برچسب زنی استفاده می شود ظاهر خواهد شد. اساس روش دوم به این صورت است که در نهایت به هر نود از صحنه یک برچسب از مدل را نسبت می دهد و راهکاری برای عدم وجود یک نود از صحنه در مدل را در نظر نمی گیرد. لذا زمانیکه چند نود ساختگی به صورت تصادفی به صحنه اضافه می شود که دارای ویژگی های یکانی مشابهی است هرچند که در فرآیند آرام سازی این نودها حمایت کمی می گیرند و همانطور که از جدول 4-5 مشخص است احتمال های پسین برای این نودها پایین بوده و همچنین به دنبال آن طبق جدول 4-6 حمایت از آنها نیز کاهش یافته اما به دلیل نبود نود پوچ در نهایت جذب یکی از برچسب های مدل می شوند. این برگترین ضعف روش دوم در برابر روش است.

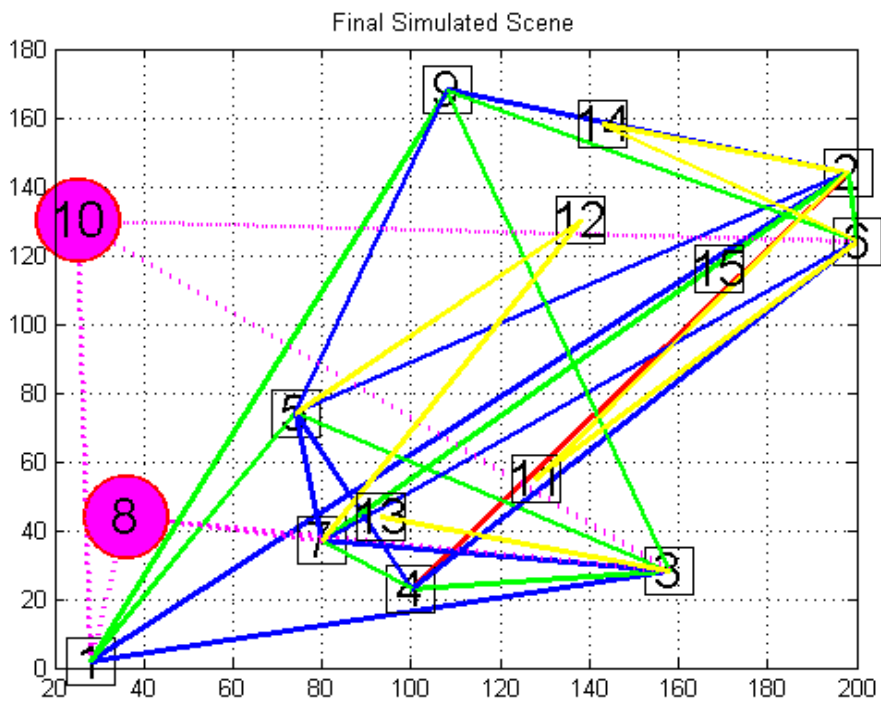
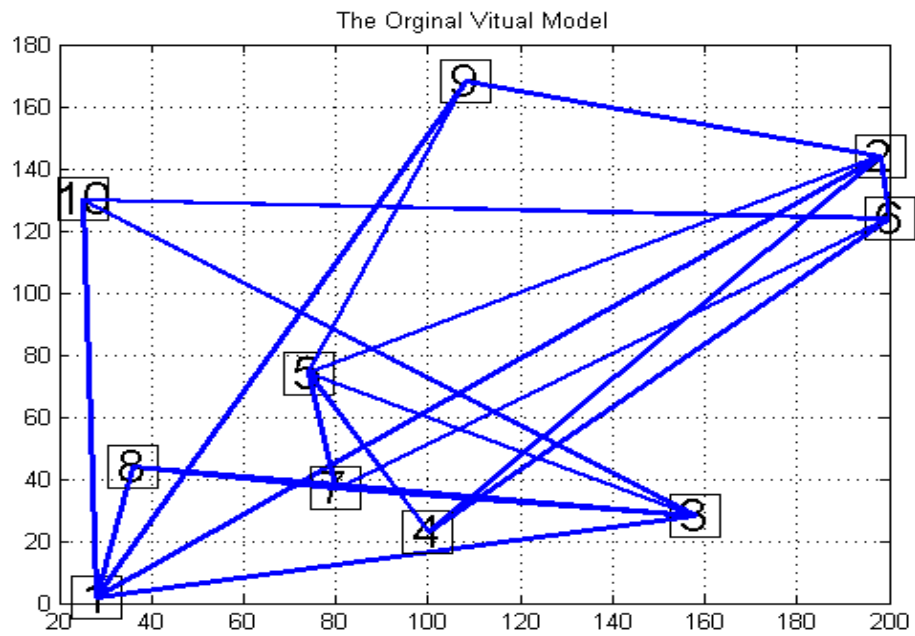
جدول 4-5: احتمال‌های پسین برای روش دوم

	Label1	Label2	Label3	Label4	Label5	Label6	Label7	Label8	Label9	Label10
Node11	0	0	0	0	0	0	0	0	0	0
Node12	0	1	0	0	0	0	0	0	0	0
Node13	0	0	1	0	0	0	0	0	0	0
Node14	0	0	0	1	0	0	0	0	0	0
Node15	0	0	0	0	0	0	0	0	0	0
Node16	0	0	0	0	0	1	0	0	0	0
Node17	0	0	0	0	0	0	1	0	0	0
Node18	0	0	0	0	0	0	0	0	0	0
Node19	0	0	0	0	0	0	0	0	0,5	0
Node110	0	0	0	0	0	0	0	0	0	0,5
Node111	0	0	0	0	0	0	0	0	0,5	0
Node112	0	0	0	0	0	0	0	1	0	0
Node113	0	0	0	0	0	0	0	0	0	0,5
Node114	2E-138	0	0	0	0	0	0	0	0	0
Node115	1	0	0	0	1	0	0	0	0	0

جدول 4-6: تابع حمایت برای صحنه و مدل شکل 4-9 با روش دوم

	Label1	Label2	Label3	Label4	Label5	Label6	Label7	Label8	Label9	Label10
Node11	0	0	0	0	0	0	0	0	0	0
Node12	8E-21	1	1,79E-18	5,23E-22	1,29E-30	2,1E-11	1,72E-21	4,99E-19	5,3E-21	7,8E-20
Node13	5,47E-18	1,07E-21	1	1,65E-19	1,35E-27	0	1,36E-21	2,25E-18	3,53E-20	3,97E-21
Node14	1,55E-19	1,42E-19	3,22E-17	0,999999	2,2E-22	1,11E-19	1,8E-17	1,89E-10	1,3E-11	1,91E-19
Node15	0	0	0	0	0	0	0	0	0	0
Node16	1,88E-15	1,45E-11	1,53E-20	1,19E-19	1E-22	1	5,03E-14	3,66E-21	6,41E-22	1,06E-20
Node17	1,09E-13	3,57E-20	1,75E-13	4,05E-18	2,69E-13	5,03E-14	1	2,14E-19	5,06E-18	7,16E-17
Node18	0	0	0	0	0	0	0	0	0	0
Node19	6,82E-24	6,95E-28	5,11E-19	1,57E-12	2,41E-35	1,47E-21	4,88E-21	3,14E-15	1	3,37E-18
Node110	9,84E-20	1,08E-21	2,64E-10	1,53E-19	4,18E-21	1,06E-20	3,77E-16	2,27E-17	9,49E-21	1
Node111	1,32E-27	1,54E-21	3,94E-20	2,66E-11	4,39E-34	0	1,57E-29	2,73E-24	0,000433	8,52E-29
Node112	1,39E-18	3,5E-23	6,13E-20	2,1E-11	7,02E-16	1,86E-19	1,44E-16	1,58E-08	6,03E-13	9,28E-19
Node113	8,12E-22	3,89E-20	1,42E-09	1,63E-19	2,93E-22	1,8E-23	2,73E-16	8,79E-18	5,18E-20	0,000475
Node114	1,19E-22	2,3E-23	5,08E-25	3,86E-22	0	1,13E-25	0	1,66E-25	6,32E-31	1,83E-29
Node115	1,06E-09	3,15E-25	3,5E-17	3,46E-16	0,000424	7,47E-17	2,48E-11	1,94E-16	8,93E-25	1,26E-17

در ادامه گراف مدل و صحنه دیگری را برای بررسی بیشتر در نظر می‌گیریم.



شکل 4-10: صحنه و مدل ایجاد شده برای آزمایش دوم

جدول 7-4: ماتریس شباهت‌های نودها در مدل و صحنه شکل 4-10 تنها با استفاده از ویژگی‌های یکانی

	Label1	Label2	Label3	Label4	Label5	Label6	Label7	Label8	Label9	Label10	NULL
Nodel1	1	2,48E-11	5,85E-18	6,6E-18	9,6E-12	1,84E-09	2,48E-09	3,77E-14	8,27E-08	9,27E-11	1,5
Nodel2	2,48E-11	1	3,17E-16	7,8E-18	8,01E-08	2,37E-13	3,28E-14	1,06E-14	1,41E-07	3,34E-10	1,5
Nodel3	5,85E-18	3,17E-16	1	6,54E-16	3,85E-15	1,1E-19	2,93E-15	9,48E-09	1,5E-12	1,3E-15	1,5
Nodel4	6,6E-18	7,8E-18	6,54E-16	1	3,57E-14	7,64E-11	1,28E-18	9,87E-20	3,42E-15	3,29E-20	1,5
Nodel5	9,6E-12	8,01E-08	3,85E-15	3,57E-14	1	4,97E-12	1,61E-11	1,66E-13	1,37E-10	2,3E-11	1,5
Nodel6	1,84E-09	2,37E-13	1,1E-19	7,64E-11	4,97E-12	1	1,26E-17	4,04E-15	2,57E-13	5,03E-12	1,5
Nodel7	2,48E-09	3,28E-14	2,93E-15	1,28E-18	1,61E-11	1,26E-17	1	2,52E-19	2,71E-11	5,65E-13	1,5
Nodel8	3,77E-14	1,06E-14	9,48E-09	9,87E-20	1,66E-13	4,04E-15	2,52E-19	1	6,41E-11	1,19E-17	1,5
Nodel9	8,27E-08	1,41E-07	1,5E-12	3,42E-15	1,37E-10	2,57E-13	2,71E-11	6,41E-11	1	1,54E-11	1,5
Nodel10	9,27E-11	3,34E-10	1,3E-15	3,29E-20	2,3E-11	5,03E-12	5,65E-13	1,19E-17	1,54E-11	1	1,5
Nodel11	1,02E-16	4,79E-14	3,69E-21	1,59E-15	2,66E-23	9,02E-14	6,7E-26	1,54E-19	5,06E-17	3,7E-20	0,590729
Nodel12	8,27E-08	1,41E-07	1,5E-12	3,42E-15	1,37E-10	2,57E-13	2,71E-11	6,41E-11	1	1,54E-11	1,5
Nodel13	1,42E-09	3,47E-14	7,94E-16	5,89E-19	1,1E-20	1,18E-15	8,01E-19	1E-14	2,24E-08	1,46E-14	0,376452
Nodel14	2,48E-11	<u>1</u>	3,17E-16	7,8E-18	8,01E-08	2,37E-13	3,28E-14	1,06E-14	1,41E-07	3,34E-10	1,5
Nodel15	6,06E-14	3,97E-09	3,13E-17	2,25E-19	1,34E-10	8,81E-13	7,73E-12	2,15E-17	5,85E-09	3,82E-13	1,068679

جدول 8-4: ماتریس سردرگمی برای شکل 4-10 و با استفاده از روش اول

	Label1	Label2	Label3	Label4	Label5	Label6	Label7	Label8	Label9	Label10	NULL
Nodel1	1	0	0	0	0	0	0	0	0	0	0
Nodel2	0	1	0	0	0	0	0	0	0	0	0
Nodel3	0	0	1	0	0	0	0	0	0	0	0
Nodel4	0	0	0	1	0	0	0	0	0	0	0
Nodel5	0	0	0	0	1	0	0	0	0	0	0
Nodel6	0	0	0	0	0	1	0	0	0	0	0
Nodel7	0	0	0	0	0	0	1	0	0	0	0
Nodel8	0	0	0	0	0	0	0	0	0	0	1
Nodel9	0	0	0	0	0	0	0	0	1	0	0
Nodel10	0	0	0	0	0	0	0	0	0	0	1
Nodel11	0	0	0	0	0	0	0	0	0	0	1
Nodel12	0	0	0	0	0	0	0	0	0	0	1
Nodel13	0	0	0	0	0	0	0	0	0	0	1
Nodel14	0	0	0	0	0	0	0	0	0	0	1
Nodel15	0	0	0	0	0	0	0	0	0	0	1

همانطور که در جدول 9-4 مشخص است روش اول به درستی وظیفه برچسب زنی را انجام داده است شکل 9-4 احتمال‌های پسین برای روش اول نشان می‌دهد. واضح است، برای نودهای درست احتمال بسیار نزدیک به یک بوده و بقیه احتمال‌ها به سمت صفر میل کرده اند. اگر به ماتریس تابع حمایت در جدول 10-4 دقت کنید سطرهای 8 و 10 هیچ حمایتی را ندارند دلیل این اتفاق نیز واضح است، این نودها حذف شده‌اند. وقتی یک نود حذف می‌شود تمامی ارتباطات آن نود با بقیه نودها حذف می‌شود به همین دلیل درایه‌های این دو نود هیچ مقداری نمی‌گیرند.

جدول 9-4: احتمال‌های پسین برای روش اول برای شکل 10-4

	Label1	Label2	Label3	Label4	Label5	Label6	Label7	Label8	Label9	Label10	NULL
Nodel1	1	0	0	0	0	0	0	0	0	0	7,8E-158
Nodel2	0	1	0	0	0	0	0	0	0	0	2,8E-154
Nodel3	0	0	1	0	0	0	0	0	0	0	6,3E-149
Nodel4	0	0	0	1	0	0	0	0	0	0	2E-168
Nodel5	0	0	0	0	1	0	0	0	0	0	2,1E-160
Nodel6	0	0	0	0	0	1	0	0	0	0	1,1E-153
Nodel7	0	0	0	0	0	0	1	0	0	0	2,7E-162
Nodel8	0	0	0	0	0	0	0	0	0	0	1
Nodel9	0	0	0	0	0	0	0	0	1	0	5,9E-164
Nodel10	0	0	0	0	0	0	0	0	0	0	1
Nodel11	0	0	0	0	0	0	0	0	0	0	1
Nodel12	0	0	0	0	0	0	0	0	3,85E-43	0	1
Nodel13	0	0	0	0	0	0	0	0	0	0	1
Nodel14	0	1,56E-09	0	0	0	0	0	0	0	0	1
Nodel15	0	0	0	0	0	0	0	0	0	0	1

جدول 10-4: تابع حمایت برای صحنه و مدل شکل 10-4 باروش اول

	Label1	Label2	Label3	Label4	Label5	Label6	Label7	Label8	Label9	Label10
Nodel1	0,016214	6,01E-15	5,29E-23	2,67E-21	9,6E-16	1,48E-12	3,09E-13	5,64E-19	2,19E-12	8,88E-15
Nodel2	5,68E-16	0,013714	2,62E-21	1,43E-21	1,08E-10	4,65E-21	1,48E-18	1,31E-18	1,41E-12	1,42E-14
Nodel3	4,04E-24	7,42E-21	0,010667	4,78E-21	3,65E-19	1,21E-22	5,96E-23	1,45E-14	1,13E-18	2,16E-20
Nodel4	2,1E-23	2,23E-23	9,95E-20	0,026667	3,15E-17	5,48E-17	7,12E-22	1,26E-24	3,37E-18	3,63E-24
Nodel5	5,58E-16	1,13E-10	3E-18	7,24E-17	0,018286	1,89E-16	8,09E-18	7,73E-17	4,04E-14	9,8E-16
Nodel6	3,94E-14	1,21E-17	6,83E-24	2,4E-14	1,92E-15	0,013336	6,47E-28	2,5E-19	4,03E-16	3,57E-25
Nodel7	1,17E-12	5,31E-17	7,59E-22	6,36E-22	2,08E-16	1,76E-21	0,02	7,57E-28	3,06E-15	4,93E-16
Nodel8	0	0	0	0	0	0	0	0	0	0
Nodel9	3,36E-11	1,44E-10	3,01E-18	3,34E-18	3,39E-14	9,04E-18	5,07E-15	1,68E-14	0,021619	7,35E-16
Nodel10	0	0	0	0	0	0	0	0	0	0
Nodel11	1,17E-22	3,32E-21	1,22E-33	3,76E-18	1,33E-28	6,96E-19	1,12E-29	3,09E-31	1,55E-22	3,71E-22
Nodel12	1,28E-20	1,01E-15	1,39E-17	2,67E-18	2,66E-15	6,48E-22	1,07E-16	1,47E-17	1,37E-06	9,05E-25
Nodel13	1,67E-13	4,22E-30	4,85E-27	2,08E-36	1,54E-23	6E-32	1,94E-23	1,08E-17	1,9E-26	4,62E-17
Nodel14	1,58E-16	6,67E-06	2,13E-28	7,46E-21	2,77E-12	1,2E-20	1,55E-20	4,29E-29	3,88E-13	8,88E-19
Nodel15	0	0	0	0	0	0	0	0	0	0

جدول 11-4: ماتریس سردرگمی برای شکل 10-4 و با استفاده از روش دوم

	Label1	Label2	Label3	Label4	Label5	Label6	Label7	Label8	Label9	Label10
Nodel1	1	0	0	0	0	0	0	0	0	0
Nodel2	0	1	0	0	0	0	0	0	0	0
Nodel3	0	0	1	0	0	0	0	0	0	0
Nodel4	0	0	0	1	0	0	0	0	0	0
Nodel5	0	0	0	0	1	0	0	0	0	0
Nodel6	0	0	0	0	0	1	0	0	0	0
Nodel7	0	0	0	0	0	0	1	0	0	0
Nodel8	0	0	0	0	0	0	0	0	0	0
Nodel9	0	0	0	0	0	0	0	0	1	0
Nodel10	0	0	0	0	0	0	0	0	0	0
Nodel11	0	0	0	1	0	0	0	0	0	0
Nodel12	0	0	0	0	0	0	0	0	1	0
Nodel13	0	0	0	0	0	0	0	0	0	1
Nodel14	0	1	0	0	0	0	0	0	0	0
Nodel15	0	0	0	0	0	0	0	0	0	0

جدول 12-4: احتمال‌های پسین برای روش دوم برای شکل 10-4

	Label1	Label2	Label3	Label4	Label5	Label6	Label7	Label8	Label9	Label10
Nodel1	0,500102	0	0	0	0	0	0	0	0	0
Nodel2	0	0,5	0	0	0	0	0	0	0	0
Nodel3	0	0	1	0	0	0	0	0	0	0
Nodel4	0	0	0	0,5	0	0	0	0	0	0
Nodel5	0	0	0	0	1	0	0	0	0	0
Nodel6	0	0	0	0	0	1	0	0	0	0
Nodel7	0	0	0	0	0	0	1	0	0	0
Nodel8	0	0	0	0	0	0	0	0	0	0
Nodel9	0	0	0	0	0	0	0	0	0,5	0
Nodel10	0	0	0	0	0	0	0	0	0	0
Nodel11	0	0	0	0,5	0	1,76E-61	0	0	0	1,19E-45
Nodel12	0	0	0	0	0	0	0	0	0,5	0
Nodel13	0,499898	0	0	0	0	0	0	1	0	1
Nodel14	0	0,5	0	0	0	0	0	0	0	0
Nodel15	0	0	0	0	0	0	0	0	0	0

جدول 13-4: تابع حمایت برای صحنه و مدل شکل 10-4 باروش دوم

	Label1	Label2	Label3	Label4	Label5	Label6	Label7	Label8	Label9	Label10
Nodel1	1	7,52E-13	6,61E-21	2E-19	1,2E-13	1,48E-10	3,09E-11	4,23E-17	1,64E-10	6,66E-13
Nodel2	1,24E-13	1	5,91E-19	2,43E-19	2,35E-08	8,2E-18	2,59E-16	1,72E-16	1,86E-10	1,87E-12
Nodel3	2,12E-19	1,5E-18	1	5,38E-19	7,01E-17	1,79E-20	8,95E-21	1,64E-12	3,55E-15	1,55E-17
Nodel4	2,62E-21	2,78E-21	1,24E-17	1	3,86E-15	5,48E-15	7,12E-20	9,44E-23	2,53E-16	2,72E-22
Nodel5	1,22E-13	2,43E-08	6,54E-16	9,51E-15	1	3,3E-14	1,42E-15	1,01E-14	5,3E-12	1,29E-13
Nodel6	7,39E-12	2,26E-15	1,28E-21	2,7E-12	3,58E-13	1	1,61E-20	2,81E-17	4,54E-14	2,88E-13
Nodel7	2,19E-10	9,87E-15	1,42E-19	7,15E-20	3,9E-14	2,63E-19	1	8,51E-26	3,45E-13	5,54E-14
Nodel8	0	0	0	0	0	0	0	0	0	0
Nodel9	5,24E-09	2,25E-08	4,7E-16	3,09E-16	5,29E-12	1,13E-15	6,31E-13	1,58E-12	1	6,88E-14
Nodel10	0	0	0	0	0	0	0	0	0	0
Nodel11	1,23E-08	3,5E-07	5,13E-20	0,000238	1,41E-14	5,88E-05	9,46E-16	1,96E-17	9,85E-09	2,35E-08
Nodel12	6,68E-19	6,31E-14	8,7E-16	1E-16	1,67E-13	3,24E-20	5,35E-15	5,53E-16	5,15E-05	2,82E-23
Nodel13	3,31E-05	8,09E-22	9,57E-19	0	3,04E-15	1,1E-23	3,06E-15	1,27E-09	0	5,47E-09
Nodel14	9,88E-15	0,000417	1,26E-26	2,78E-19	1,73E-10	6,02E-19	7,77E-19	1,61E-27	1,45E-11	3,33E-17
Nodel15	0	0	0	0	0	0	0	0	0	0

4-5 جمع بندی دو روش بدون وجود نویز

زمانیکه نویزی به شبکه وارد نمی‌شود، روش اول به راحتی می‌تواند نودهای سرگردان در صحنه را مدیریت کند. در مقابل روش دوم این توانایی را ندارد. تابع حمایت برای هر دو روش زمانیکه نودی از صحنه به دلیل انسداد حذف شده است، مقدار صفر را برای تمامی درایه‌های آن سطر در نظر می‌گیرد.

4-6 وجود نویز در صحنه

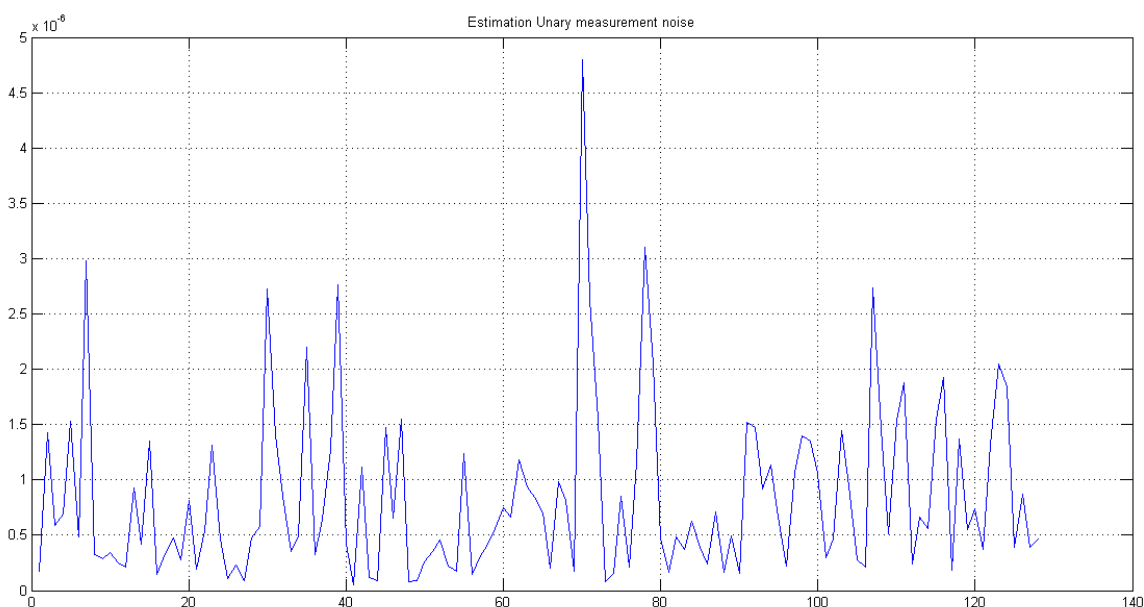
در مرحله قبل فرض بر آن بود که به نودهای صحنه هیچ نویزی وارد نشود. در این قسمت کمی بیشتر محیط شبیه سازی را به واقعیت نزدیک‌تر می‌کنیم. در این مرحله فرض می‌کنیم که هرکدام از عناصر بردار ویژگی یکانی و بردار ویژگی باینری از گراف صحنه توسط نویز نرمال با مقدار انحراف معیار مشخص آلوده شود. با این شرایط جدید سعی می‌کنیم تحمل الگوریتم تطبیق روش اول و روش دوم را در برابر نویز بررسی کنیم.

4-6-1 تعیین محدوده قدرت نویز

محیط شبیه سازی شده باید دارای ویژگی‌های محیط واقعی باشد، لذا محدود نویزی که باید به ویژگی‌های یکانی و همچنین ویژگی‌های باینری اضافه شود، باید به درستی انتخاب شود. برای اینکه محدوده درست نویز در ویژگی‌های یکانی و همچنین در ویژگی‌های باینری مشخص شود، مطالعه‌ای رو تصاویر واقعی انجام شد تا تقریب درستی از این محدوده بدست آید.

در این مطالعه تعدادی تصویر انتخاب شد و توصیفگر SIFT را روی آنها اعمال نمودیم. با استفاده از مقایسه بین بردارهای 128 تایی بین هر نقطه کلیدی از مدل و متناظر آن در صحنه، نزدیک ترین نقطه از صحنه نسبت به یک نقطه مفروض بدست آمد. فرض می‌شود این تطبیق، تطبیق درستی باشد. برای

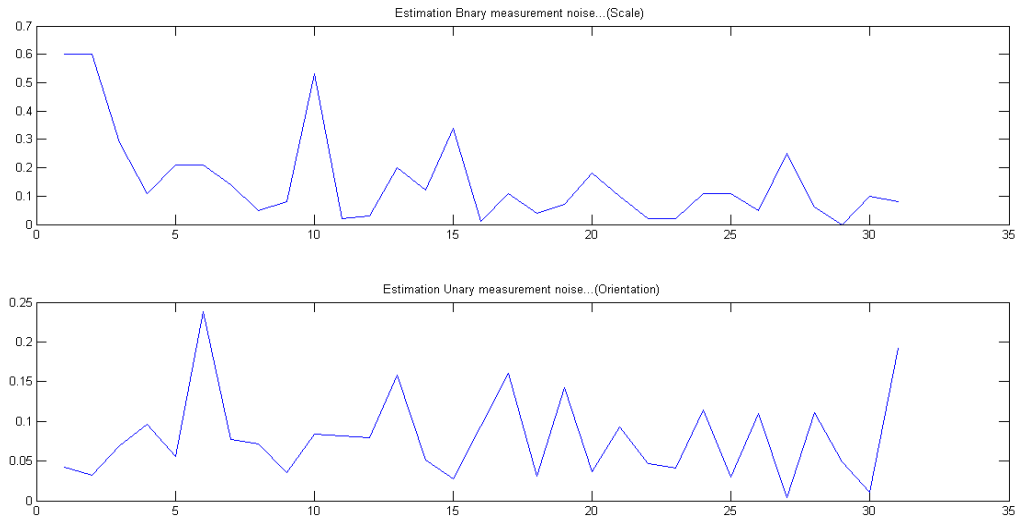
محاسبه میزان نویز ، برای هر دو نقطه تطبیق داده شده از صحنه و مدل، مقدار مطلق تفاضل این دو بردار 128 تایی را در نظر گرفته شد. سپس این مقدار با تقسیم بر اندازه بردار نرمال می‌شود. در نهایت از ماتریس بدست آمده روی تک تک سطرهایش انحراف معیار گرفته شد. مقدار این انحراف معیار، به میانگین عناصر آن سطر تقسیم شد تا مقدار بدست آمده مستقل از محدوده اعداد باشد. حاصل این کار برای دو تصویر مشخص در شکل زیر (شکل 4-11) آمده است.



شکل 4-11: تعیین محدوده مجاز نویز در ویژگی‌های یکانی

در شکل 4-11 محور افقی عناصر یک نقطه کلیدی در توصیفگر SIFT نشان می‌دهد و محور عمودی توان نویز را نشان می‌دهد. اگر بردار 128 تایی را به عنوان ویژگی‌های یکانی در نظر بگیریم، شکل 4-11 نشان می‌دهد که توان نویز در محدوده صفر تا 5×10^{-4} می‌باشد. برای ویژگی باینری هم مشابه این کار انجام می‌شود. در آزمایش‌های محیط واقعی از ویژگی‌های مقیاس و جهت برای ویژگی‌های باینری استفاده شده است، لذا برای تعیین میزان نویز نیز از همین دو ویژگی مشابه روش بالا استفاده کرده و نویز را

تخمین می‌شود. شکل 4-12 برای دو نقطه کلیدی تطبیق داده شده است. در هر تطبیق تفاضل دو جهت برای دو نقطه تطبیق داده شده در نظر گرفته شده است. در نهایت حاصل این تفاضل نرمال شده است.



شکل 4-12: تعیین محدوده مجاز نویز در ویژگی‌های باینری. شکل بالا ویژگی مقیاس و پایین جهت

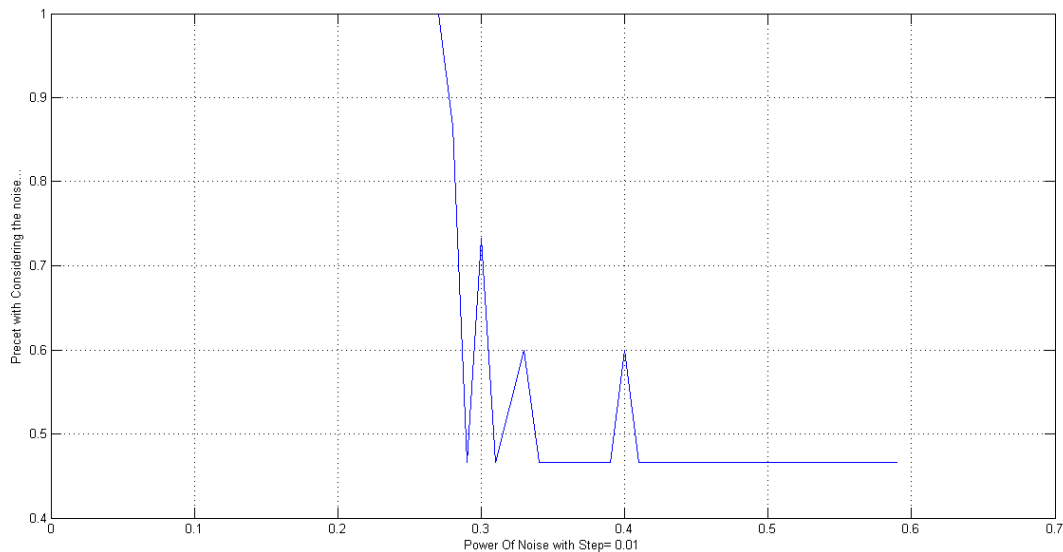
پس از مطالعه محدوده مناسب نویز اثر نویز را در سه حالت زیر بررسی کنیم

- اضافه کردن نویز تنها به بردار ویژگی یکانی
- اضافه کردن نویز تنها به بردار ویژگی باینری
- اضافه کردن نویز هم به بردار ویژگی یکانی و هم به بردار ویژگی باینری

4-6-2 اضافه کردن نویز تنها به ویژگی‌های یکانی

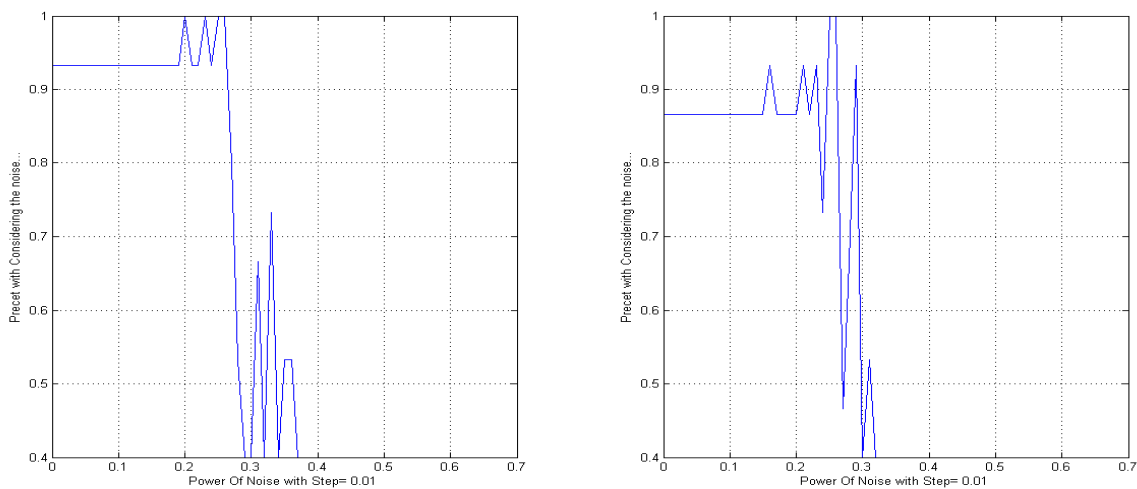
در حالت اول نویز تنها به ویژگی‌های یکانی اضافه می‌شود. سپس رفتار سیستم در یک دوره مورد بررسی قرار می‌گیرد. در هر مرحله کمی قدرت نویز افزایش داده می‌شود. حاصل این آزمایش به صورت شکل 4-13 قابل مشاهده است. همانطور که از این شکل قابل برداشت است تا زمانی که نویز به یک حد آستانه

می‌رسد سیستم قابل اتکا است. زمانی که نویز از آن حد آستانه عبور می‌کند سیستم دچار یک آشوب می‌شود، در این حالت سیستم دیگر قابل اعتماد نیست.



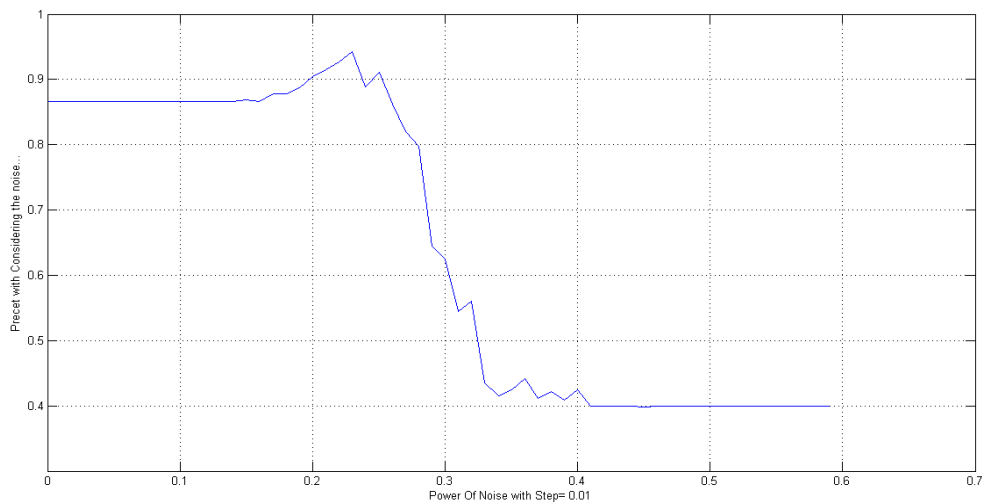
شکل 4-13: تحمل روش اول تطبیق در برابر نویز اضافه شده تنها به ویژگی‌های یکانی

مدل و صحنه‌های دیگری را نیز بررسی کردیم که در زیر نتایج آمده است.



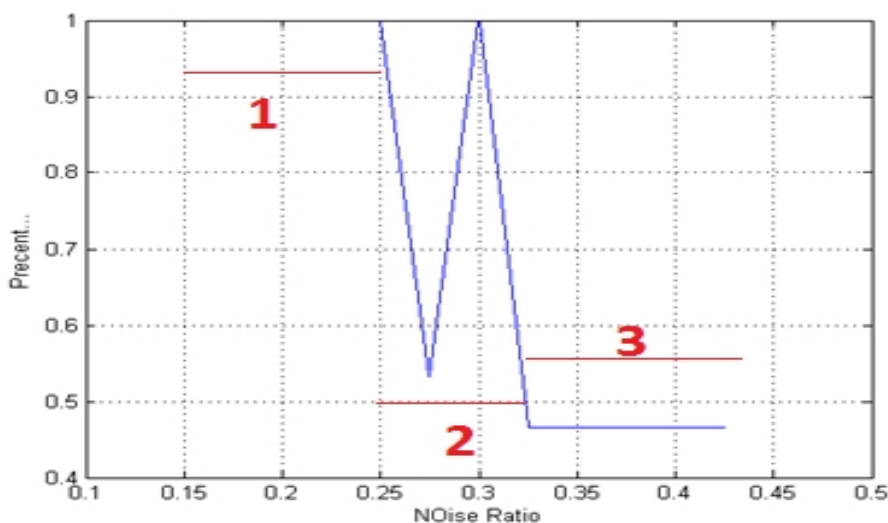
شکل 4-14: دومتال دیگر برای تحمل روش اول تطبیق در برابر نویز اضافه شده تنها به ویژگی‌های یکانی

همانطور که از شکل‌های (4-13 و 4-14) مشخص است زمانی که سیستم به یک آشوب می‌رسد ممکن است که اضافه شدن نویز کمک کند که دقت سیستم افزایش یابد، اما این افزایش قابل اعتماد نیست. نویز تصادفی می‌باشد و هر بار تغییر می‌کند اما قدرت نویز ثابت است. به تعداد معین ولی با قدرت نویز مشخص و ثابت سیستم آزمایش می‌شود.



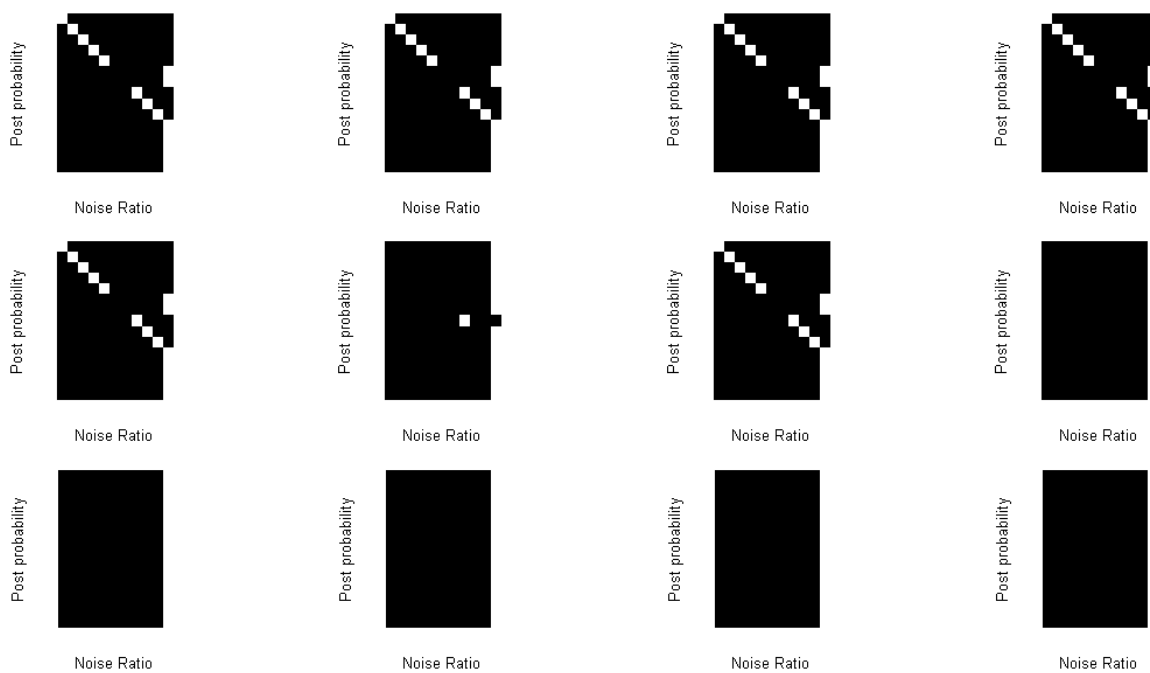
شکل 4-15: تحمل روش اول در برابر نویز اضافه شده تنها به ویژگی‌های یکانی در چندین آزمایش با یک قدرت مشخص

برای اینکه این بحث روشن‌تر شود برای یک مدل خاص نویز را یکبار تنها به ویژگی‌های یکانی و بار دیگر تنها به ویژگی‌های باینری و در انتها به هر دو آن‌ها اضافه شده و نتایج تحلیل می‌شود. پس از اینکه نویز تنها به ویژگی‌های یکانی صحنه به صورت مرحله به مرحله اضافه کردیم، حاصل به صورت زیر شد.



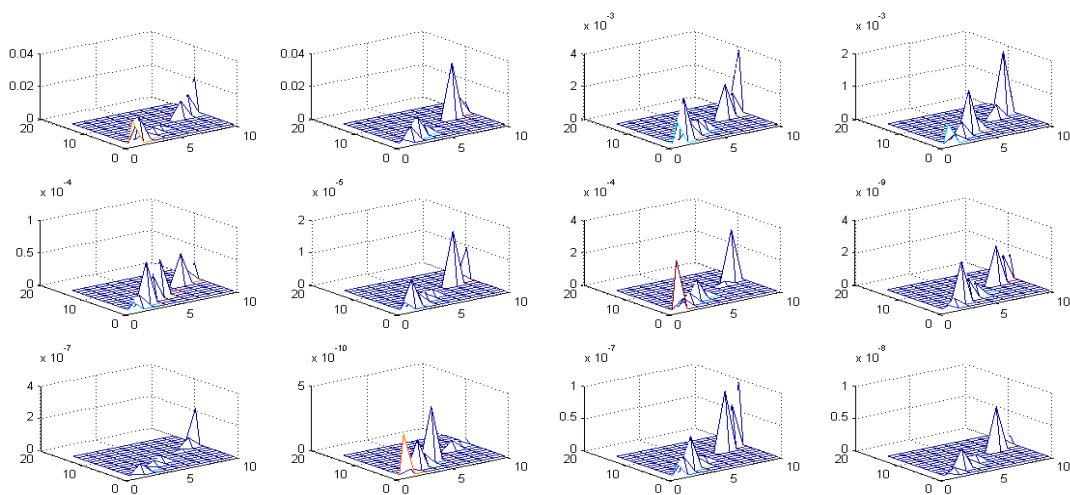
شکل 4-16: رفتار روش اول زمانیکه نویز به ویژگی‌های یکانی وارد می‌شود.

اگر به شکل 4-15 دقت شود، رفتار آن در سه مرحله قابل توضیح است. قبل از یک آستانه سیستم در حالت یک قرار دارد در این حالت سیستم کاملاً در مقابل شرایط نویزی مقاوم است. پس از یک آستانه سیستم دچار یک آشوب می‌شود. این آشوب در بعضی موارد سبب می‌شود که نرخ بازشناسی حتی بالاتر رود (مرحله 2). پس از یک آشوب دوباره سیستم به یک حالت پایسته می‌رسد (مرحله 3). برای تحلیل اینکه چرا در مرحله 2 سیستم در بعضی موارد افزایش بازشناسی را دارد به احتمال‌های پسین دقت شد. با توجه به شکل 4-16 در تحلیل این پدیده باید گفت زمانی که نرخ بازشناسی در فاز آشوب بالا می‌رود دو اتفاق امکان افتادن دارد: اولین اتفاق این است که یک نود یک برچسب درست نگرفته بوده حالا در اثر محیط تصادفی برچسب پوچ بگیرد که یک برچسب درست باشد و یا اینکه یک نود که برچسب پوچ داشته و نادرست بوده در اثر همان محیط تصادفی برچسب درست بگیرد.



شکل 4-17: احتمال‌های پسین برای روش اول در برابر وارد شدن نویز به ویژگی‌های یکانی

پدیده دیگری که اتفاق می‌افتد و در شکل 4-16 مشخص است، این است که بعد از یک مدت تمامی نودها برچسب پوچ می‌گیرند برای تحلیل این حالت، تابع حمایت مورد تحلیل قرار گرفت.



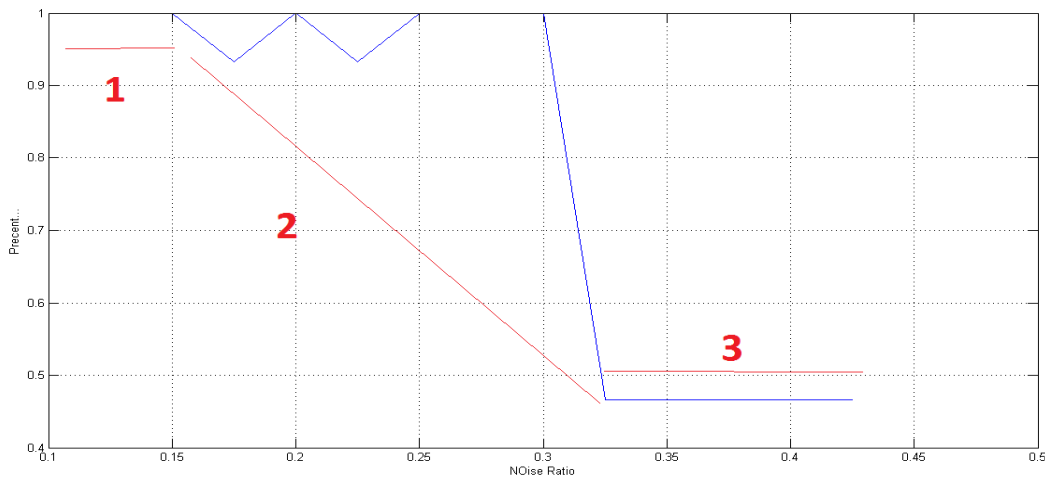
شکل 4-18: توابع حمایت برای روش اول در برابر وارد شدن نویز به ویژگی‌های یکانی

در توجیه این پدیده با استفاده از شکل 4-17 می‌توان گفت هرچقدر توان نویز افزایش می‌یابد در واقع توان تابع حمایت کمتر می‌شود و این بدان معناست که حمایتی از همسایه‌ها برای یک نود خاص صورت نمی‌گیرد. در چنین شرایطی نود پوچ تمامی نودها را به سوی خود جذب می‌کند. برای اینکه در این آزمایش‌ها بتوان قدرت تحمل سیستم را بررسی کرد، از معیار نسبت قدرت سیگنال به نویز استفاده می‌شود. این معیار برای یک سیگنال به صورت زیر تعریف می‌شود.

$$PSNR = 10 \times \log_{10} \left(\frac{MAX^2}{MSE} \right) \quad 1-4$$

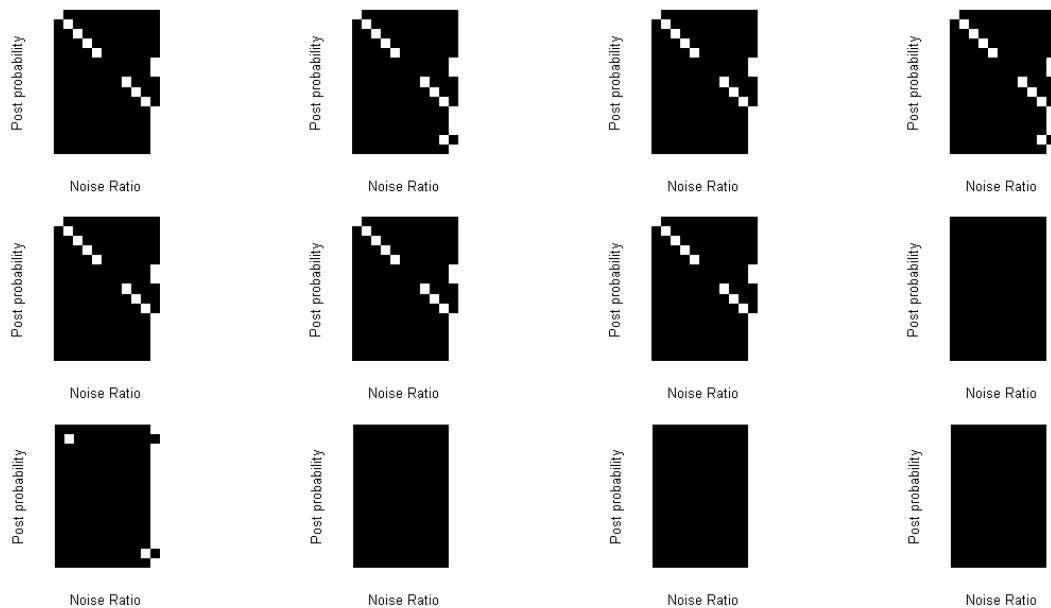
در تمامی آزمایش‌ها، میزان صورت (قدرت سیگنال) در تعریف بالا برابر یک می‌باشد و اگر به شکل 4-15 دقت شود، زمانی که سیستم به فاز آشوب (فاز 2) می‌رود میزان توان نویز برابر 0/25 می‌باشد. لذا میزان PSNR برای آزمایش بالا برابر 12 دسی بل می‌باشد که با توجه به مطالعه محدوده نویز در محیط واقعی، نشان دهنده مقاوم بودن تطبیق مبتنی بر آرام سازی در برابر نویز است.

4-6-3 اضافه کردن نویز تنها به ویژگی‌های باینری

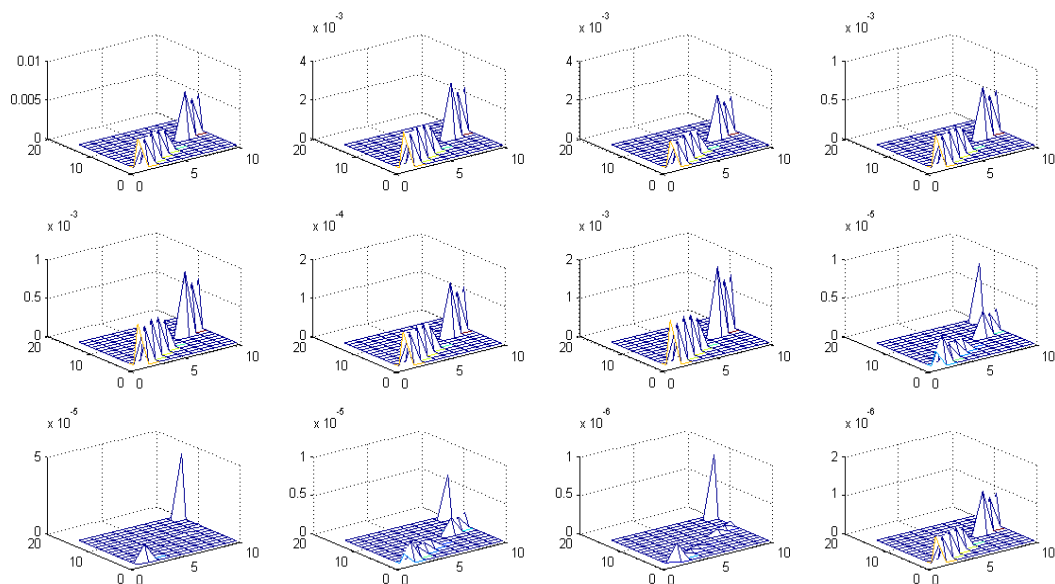


شکل 4-19: رفتار روش اول در برابر وارد شدن نویز به ویژگی‌های باینری

به صورت مشابه این بار تأثیر اضافه کردن نویز به ویژگی‌های باینری مورد آزمایش قرار می‌گیرد. شکل بالا را نیز مانند مرحله قبل می‌توان به سه فاز تقسیم کرد: فاز اول فاز قبل از آشوب، پس از یک آستانه سیستم وارد یک آشوب می‌شود و در فاز سوم دوباره سیستم به یک حالت تعادل می‌رسد. همانطور که در فاز دوم مشخص است در بعضی حالات بازشناسی سیستم حتی افزایش هم یافته که برای این کار مطالعه‌ای روی ماتریس احتمال‌های پسین انجام شد. وقتی به ماتریس شکل 4-19 دقت شود همان تحلیل مشابه در حالت قبل را می‌توان برای این حالت نیز نوشت. و در نهایت اینکه چرا تمامی نودها از یک مرحله به بعد جذب نود پوچ می‌شوند؟ برای تحلیل این حالت نیز مطالعه‌ای بر روی ماتریس توابع حمایت شد که حاصل در شکل 4-20 آمده است.



شکل 4-20: احتمال‌های پسین برای روش اول در برابر وارد شدن نویز به ویژگی‌های باینری

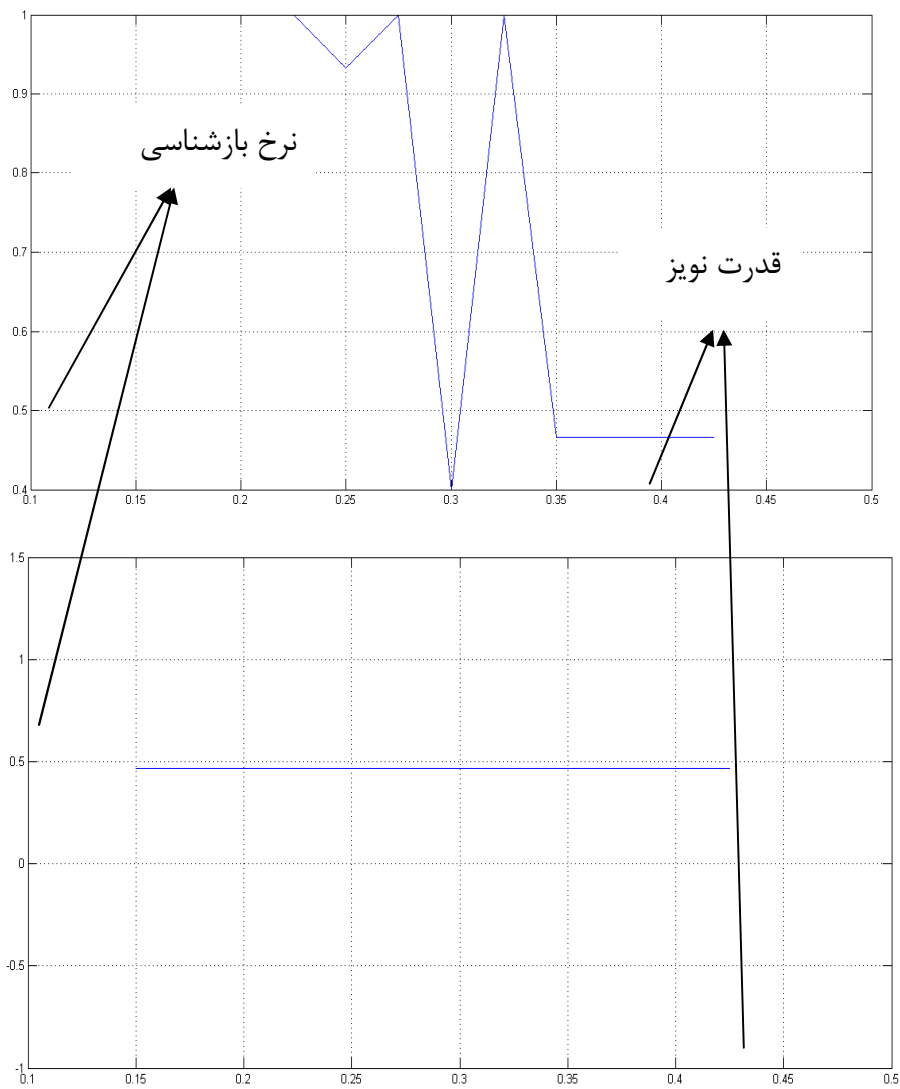


شکل 4-21: نواح حمایت برای روش اول در برابر شدن نویز به ویژگی‌های باینری

همانطور که از شکل 4-20 کاملاً مشخص است میزان حمایت همسایه‌ها از یک نود با افزایش نویز کاهش می‌یابد و این یعنی کاهش اثر مشاهدات در تطبیق. این دلیل جذب شدن نودها به نود پوچ می‌باشد. به طور مشابه می‌توان مقدار PSNR را برای این حالت محاسبه کرد.

4-6-4 اضافه کردن نویز به ویژگی‌های یکانی و باینری

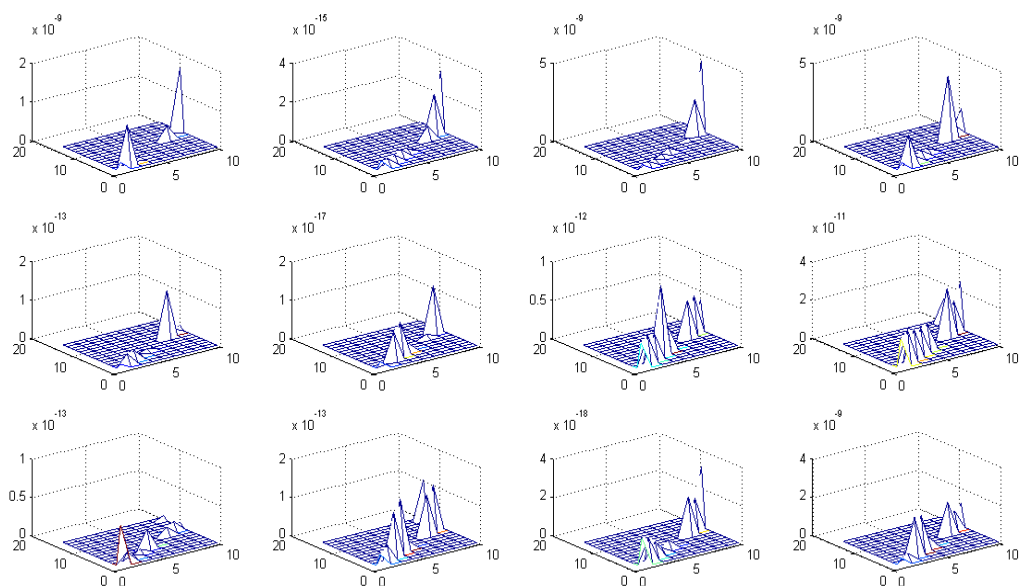
در حالت طبیعی سیستم به کمک ویژگی‌های یکانی و باینری به بازشناسی می‌پردازد. لذا باید حالتی را که هر دو ویژگی دارای نویز باشد را مورد بررسی قرار داد. هرگاه نویز بتواند به صورت توأم به هر دو ویژگی وارد شود، شبکه چقدر توان دارد و چگونه در مقابل این پدیده واکنش نشان می‌دهد؟ در همان مدل و صحنه‌ای که دو آزمایش بالا انجام داده شد این آزمایش نیز انجام می‌شود.



شکل 4-22: رفتار روش اول در برابر وارد شدن نویز به ویژگی‌های یکانی و باینری

در تحلیل شکل بالا (شکل 4-21) باید گفت مادامی که نویز به صورت توأم به شبکه تزریق می‌شود با افزایش نویز انتظار می‌رود تحمل شبکه کاهش یابد که این حالت اتفاق افتاد. همانطور که از شکل 4-21 مشخص است قسمت بالا این شکل زمانی را نشان می‌دهد که نویز وارد شده به ویژگی‌های یکانی و باینری در کمترین میزان خود است و قسمت پایین شکل حالتی را نشان می‌دهد که نویز اضافه شده به هر دو در بالاترین میزان خود است. در این شرایط، سیستم کاملاً تحت تأثیر نویز قرار گرفته است. بیشترین توان نویز در این آزمایش برای هر دو ویژگی برابر $0/45$ می‌باشد. با توجه به اینکه هر دو ویژگی در محدوده

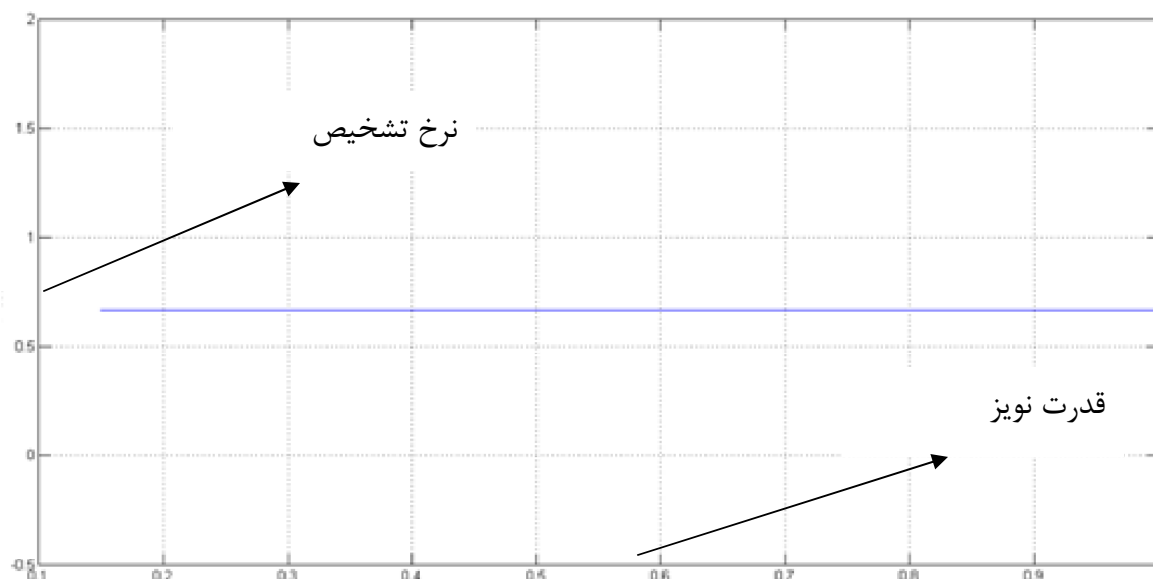
صفر و یک هستند، دارای توان ماکزیمم یک می‌باشند، لذا با این مقدار نویز سیستم تحمل بالایی دارد. مقادیر احتمال های پسین و همچنین تابع حمایت نیز مورد بررسی قرار گرفت. هرگاه قدرت نویز بالا می‌رود مقادیر حمایت کاهش پیدا می‌یابد و تمامی نودها به سمت پوچ تمایل پیدا می‌کنند. در شکل 4-22 برای بیشترین مقدار نویز به صورت توأم، مقادیر تابع حمایت آمده است.



شکل 4-23: رفتار تابع حمایت روش اول در برابر وارد شدن نویز به ویژگی‌های یکانی و باینری

4-6-5 وجود نویز در روش دوم تطبیق

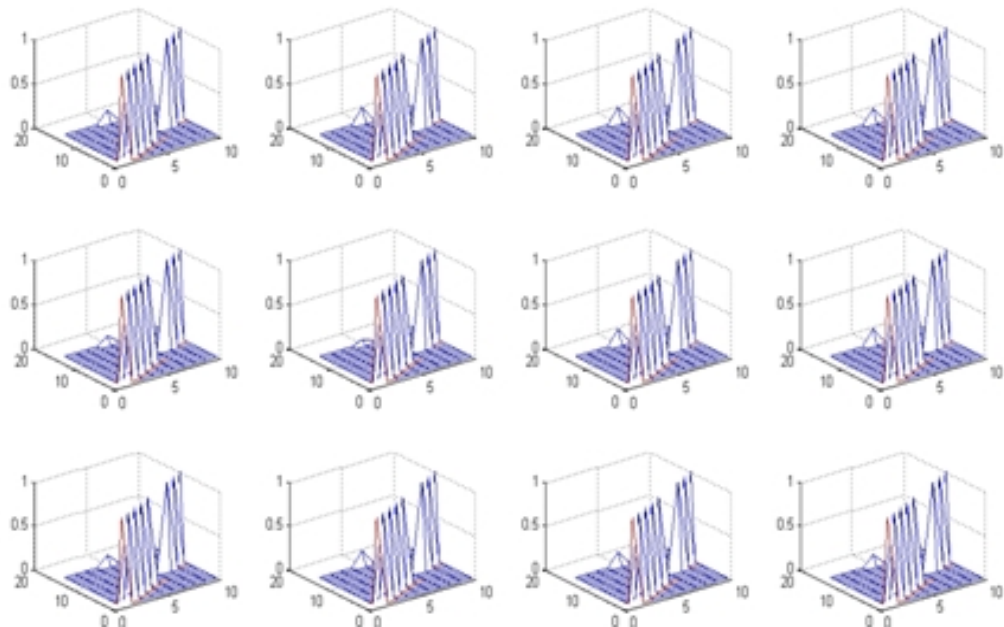
همانند روش اول، در این روش نیز نویز را به صورت مرحله به مرحله به سیستم اضافه می‌شود. سپس سیستم در حالتی که نویز تنها به ویژگی‌های یکانی، تنها ویژگی‌های باینری و در انتها هر دو ویژگی اضافه شده باشد، مورد بررسی قرار می‌گیرد. در اولین حالت نویز را به ویژگی یکانی اضافه کرده و تحمل شبکه گراف را بررسی می‌کنیم. شکل 4-24 اعمال نویز را به ویژگی یکانی نشان می‌دهد.



شکل 4-24: رفتار روش دوم در برابر وارد شدن نویز به ویژگی‌های یکانی

شکل بالا مقاومت بالای روش دوم در برابر نویز را نشان می‌دهد. در مقایسه با روش اول، نقطه ضعف روش دوم نرخ بازشناسی پایین‌تر این روش است. این نرخ پایین به دلیل نبود نود پوچ است. در واقع این روش نمی‌تواند نودهایی که در صحنه معادلی در مدل ندارند مدیریت کند و این یعنی بالا رفتن نرخ خطا¹ که عامل بسیار مهمی در تصمیم‌گیری است. در شکل زیر مشخص است که توابع حمایت توانسته‌اند به درستی نود صحیح را مورد حمایت قرار دهند و در برابر نویز دچار تزلزل نشده‌اند. وقتی این توابع حمایت را با توابع حمایت شکل های 4-17 مقایسه می‌شود، در آن توابع وقتی نویز افزایش می‌یابد محدوده حمایت بسیار کاهش می‌یابد. نویز سبب می‌شود که حمایت در آن روش دچار تزلزل شود. این در حالی است که در روش دوم توابع حمایت در برابر افزایش به درستی عمل می‌کنند. دلیل توانمندی بیشتر روش دوم در برابر نویز هم قدرت بالای تابع حمایت در این روش است.

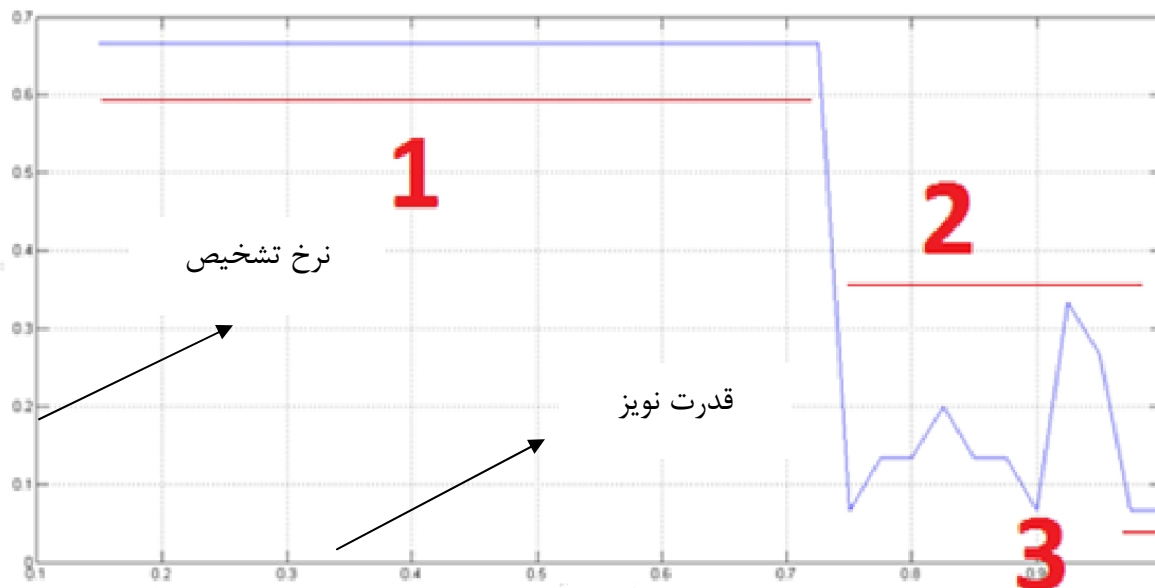
¹False positive



شکل 4-25: رفتار تابع حمایت روش دوم در برابر وارد شدن نویز به ویژگی‌های یکانی

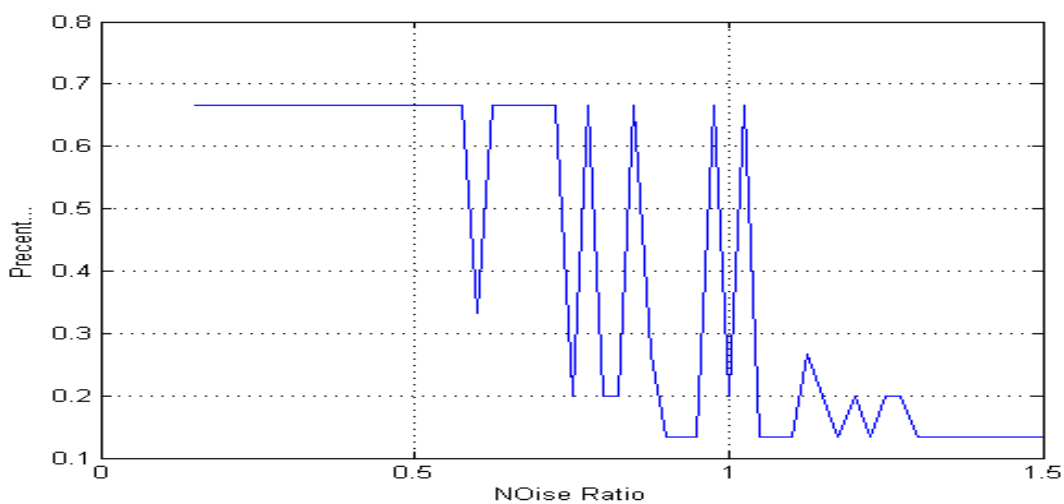
در آزمایش بعدی نویز را به ویژگی‌های باینری اضافه می‌کنیم تا توان الگوریتم را در برابر این نویز آزمایش کنیم. زمانی که نویز به ویژگی‌های باینری اعمال می‌شود نسبت به زمانی که نویز به ویژگی‌های یکانی وارد می‌شد سیستم زودتر متزلزل می‌شود. این پدیده را می‌توان اینگونه تحلیل کرد: روش دوم نود ساختگی پوچ را در نظر نمی‌گیرد، وقتی نود پوچ وجود داشته باشد همواره یک کشمکش بین این نود و بقیه نودهای دیگر برای برچسب زدن وجود دارد. زمانی که قدرت نویز وارد شده شبکه افزایش می‌یابد، از آنجا که روی مشاهدات تأثیر می‌گذارد، نود ساختگی پوچ در این کشمکش پیروز می‌شود لذا با بالا رفتن نویز تمامی نودها به سمت نود پوچ می‌روند. در روش دوم این نود وجود ندارد، پس هیچ کشمکشی هم وجود ندارد، اما از آنجا که اساساً روش دوم و هر روش دیگری که مبتنی بر گراف است به ویژگی‌های باینری وابسته است زمانی که نویز اضافه می‌شود این شبکه در مقایسه با زمانی که نویز به ویژگی یکانی اضافه می‌شود، زودتر به تلاطم می‌افتد. اما به دلیل نبود نود ساختگی پوچ دیرتر از روش اول به تلاطم می‌افتد.

اینجا دوباره باید تاکید کرد که روش دوم در این حالت نیز نسبت به روش اول دارای نرخ بازشناسی کمتری است.

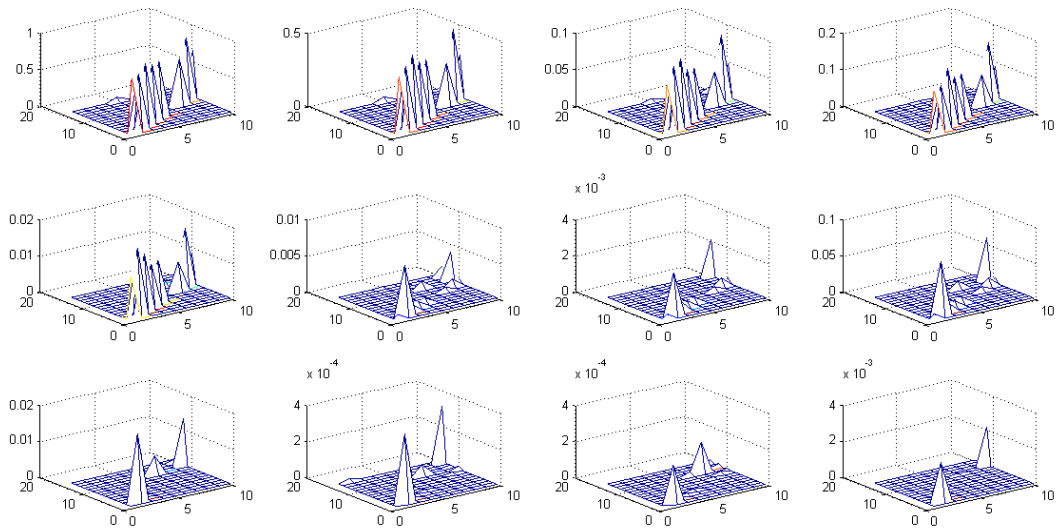


شکل 4-26: رفتار روش دوم در برابر وارد شدن نویز به ویژگی‌های باینری

در نهایت حالتی که نویز به صورت توأم به شبکه اضافه شود (شکل 4-26).



شکل 4-27: رفتار روش دوم در برابر وارد شدن نویز به ویژگی‌های یکانی و باینری



شکل 4-28: رفتار تابع حمایت روش دوم در برابر وارد شدن نویز به ویژگی‌های یکانی و باینری

وقتی که نویز به سیستم وارد می‌شود به این معنا است که اطلاعات درست سیستم کاهش می‌یابد. کاهش یافتن اطلاعات به این معنی است که همسایه‌ها به درستی از نودها حمایت نمی‌کنند. دلیل این موضوع واضح است، ویژگی‌های استفاده شده به دلیل نویز قابل اتکا نیستند. این موضوع را می‌توان از رسم روند تابع های حمایت نیز ملاحظه کرد (شکل 4-27).

4-7 نتیجه گیری

در این فصل یک محیط شبیه سازی شده ساخته شد تا تحمل دو الگوریتم مبتنی بر فرآیند آرام سازی احتمالاتی بررسی گردد. در محیط واقعی بسیاری اوقات تعیین دقیق محدوده پارامترها بسیار مشکل است. از طرفی محدوده درست بسیار در تعیین صحت و کارایی الگوریتم تاثیر دارد. هدف از ایجاد محیط شبیه سازی شده این بود تا بتوان محدوده پارامترهای مهم به درستی محاسبه گردد. امروزه توصیفگرها محلی برای توصیف تصویر بسیار کارا هستند، همچنین توصیفگرها به بلوغ کافی دست پیدا کرده اند. از این رو

بیشترین تلاش این پایان نامه بهبود و تحلیل الگوریتم تطبیق است. یکی از چالش‌هایی که تاثیر بسیار زیادی روی عملکرد الگوریتم تطبیق دارد، وجود نویز است.

رفتار دو الگوریتم احمدی فرد (روش اول) و Kostin (روش دوم) در این فصل در برابر نویز و بدون نویز مورد بررسی قرار گرفت. در شرایطی که نویز در سیستم وجود نداشته باشد، روش اول با استراتژی نود ساختگی پوچ کارایی مناسبی و نرخ بازشناسی بالایی نیز دارد، این در حالی است که روش دوم در برابر نودهایی که در صحنه به عنوان نود پیش زمینه¹ اضافه شدند کارایی مناسبی نداشت و همین مسئله کارایی روش دوم را کاهش می‌دهد. روش دوم در برابر نویز مقاومت بسیار زیادی دارد.

این مقاومت زمانی است که نویز به ویژگی‌های یکانی وارد می‌شود. زمانیکه نویز به ویژگی‌های باینری اضافه شد روش دوم در مقایسه با زمانیکه نویز به ویژگی‌های یکانی وارد می‌شود مقاومت خوبی از خود نشان نداد هرچند بازهم مقاومت مناسبی بود. برای اینکه مقاومت روش اول و دوم نشان داده شود مطالعه ای رو نویز سیستم های واقعی انجام شد. در مقایسه با محدوده نویزی که به صورت طبیعی ایجاد می‌شود، این الگوریتم‌ها دارای مقاوم بالایی هستند. دو الگوریتم زمانیکه نویز از یک حد آستانه بیشتر می‌شود وارد یک حالت آشوب گونه می‌شوند، در این حالت شبکه قابل اتکا نیست. در مرحله آشوب حتی ممکن است نرخ بازشناسی افزایش یابد. در روش دوم افزایش نویز به ویژگی‌های یکانی تاثیر زیادی در شبکه ندارد ولی افزایش نویز به ویژگی‌های باینری سبب می‌شود قدرت گراف که بیشتر در اطلاعات باینری می‌باشد دچار تزلزل شود. روش دوم در مقایسه با روش اول در برابر نویز بسیار مقاوم تر است هرچند که روش اول نیز کارایی خوبی در برابر نویز دارد. در مقابل روش دوم در مقایسه با روش اول نرخ بازشناسی کمتری دارد

¹Background

دلیل پایین بودن نرخ بازشناسی در روش دوم نبود نود پوچ برای مدیریت نودهای سرگردان در صحنه است.

یک مسئله رایج گاه در آن سطحی که طرح شده نمی توان حل کرد

آلبرت له شتین ۱۸۷۹-۱۹۷۵

فصل پنجم

نتایج پیاده سازی سنایومی واقعی

فصل پنجم

در این فصل دو روش احمدی فرد و Kostin را در سناریوی واقعی مورد بررسی قرار می‌دهیم. مراحل کار در سناریوی واقعی به شرح زیر است:

5-1 توصیف تصویر با استفاده از توصیفگر SIFT

زمانیکه توصیفگر SIFT روی تصویر ورودی اعمال شد تعدادی نقاط کلیدی استخراج می‌شوند. هر نقطه کلیدی دارای یک بردار است که با استفاده از آن می‌توان نقطه کلیدی مفروض را توصیف کرد. حال اگر قرار باشد صحنه و مدل را تطبیق داده شود، ساده‌ترین راه‌حل مقایسه نقاط کلیدی به تنهایی می‌باشد. این روش را تطبیق مستقیم¹ می‌گویند. اگر با روش مستقیم و با استفاده از بردار توصیفگر SIFT تطبیق انجام شود نیاز است که توصیفگرها خاصیت بسیار متمایز کننده‌ای² داشته باشند.

در عمل این روش به دو دلیل امکان پذیر نیست

- تصاویر با اشیای مختلف در پایگاه داده‌های بسیار بزرگ ممکن است دارای شباهت‌هایی به صورت محلی³ باشند.
- درصد شلوغ⁴ بودن صحنه این وضعیت را بدتر هم می‌کند.

دلایل بالا توجیه می‌کند که استفاده از تطبیق مستقیم برای مقایسه دو تصویر مناسب به نظر نمی‌رسد. وقتی دو تصویر باهم مقایسه می‌شوند وجود خطا یا اجتناب ناپذیر است. میزان تحمل خطا بستگی به کاربرد دارد.

¹Direct matching

²Discriminative

³Local similarity

⁴Clutter

هدف اصلی در این پایان‌نامه بهبود انطباق بین تصاویر با استفاده از تعریف همسایگی برای هر نقطه کلیدی می‌باشد. برای هر نقطه کلیدی تعدادی همسایه در نظر گرفته می‌شود. تعداد این همسایه ها k نزدیکترین همسایه است که در نزدیکی نقطه کلیدی وجود دارد. برای سنجیدن فاصله بین هر دو نقطه کلیدی از فاصله اقلیدسی استفاده می‌کنیم.

به دو دلیل از تنها k نزدیک ترین همسایه استفاده می‌کنیم:

- اولاً برای کاهش پیچیدگی محاسباتی هم در استخراج ویژگی و هم تطبیق.
- ثانیاً استفاده از تعداد بالای همسایگی باعث می‌شود که خاصیت متمایز کنندگی که مد نظر است (یا به عبارت دیگر قابلیت اطمینان) کاهش یابد.

دو نوع ویژگی برای نقاط کلیدی استخراج شده تصویر در نظر گرفته می‌شود:

- ویژگی‌های یکانی
- ویژگی‌های باینری

با استفاده از ویژگی نوع اول ما هر نقطه کلیدی را به تنهایی توصیف می‌کنیم. این در حالی است که در نوع دوم هر جفت¹ نقاط کلیدی را توصیف می‌کنیم. در بخش بعد هر کدام از این ویژگی‌ها را بررسی می‌کنیم.

5-2 ویژگی‌های یکانی

زمانیکه بر روی هر تصویر توصیفگر SIFT را اعمال می‌شود حاصل، تعدادی نقطه کلیدی است. در تئوری بیان شد که هر نقطه کلیدی در روش SIFT توسط یک بردار با طول 128 توصیف می‌شود.

¹Pair

از این بردار به عنوان ویژگی یکانی استفاده می شود

3-5 ویژگی های باینری

برای توصیف ارتباط بین هر دو نقطه کلیدی همسایه به این صورت عمل می شود :

- در اعمال توصیفگر SIFT برای هر نقطه کلیدی یک جهت استخراج می شود (Orientation_{N_i} جهت برای i امین نقطه کلیدی).
- همچنین برای هر نقطه کلیدی یک مقیاس هم استخراج می شود (scale_{N_i} بزرگنمایی برای i امین نقطه کلیدی).

برای ویژگی های باینری برای هر دو نقطه کلیدی N_i و همسایه N_j دو ویژگی در نظر می گیریم. این توصیفگرها بین دو نود متناظر از صحنه و مدل با هم برابر می باشند.

$$\text{Binary measurement } (N_i, N_j) = \begin{cases} \text{Orientation}_{N_i} - \text{Orientation}_{N_j} & 1-5 \\ \text{Scale}_{N_i} / \text{Scale}_{N_j} \end{cases}$$

4-5 بازنمایی تصاویر مدل و صحنه در ARG

تمامی نقاط کلیدی استخراج شده از مدل شی را در یک گراف نشان می دهیم. این گراف با چندتایی $\bar{G} = \{\bar{O}, \bar{X}, \bar{A}\}$ مشخص می شود. در این بیان $\bar{O} = \{\bar{o}_1, \bar{o}_2, \dots, \bar{o}_m\}$ نشان دهنده مجموعه نودهای گراف است. \bar{X} و \bar{A} به ترتیب ویژگی های باینری و یکانی هستند. هر نود در این گراف یک نقطه کلیدی استخراج شده توسط SIFT است. هر نود \bar{o}_i را با استفاده از یک بردار ویژگی یکانی \bar{x}_i توصیف می شود. \bar{X} نشان دهنده مجموعه بردارهای یکانی برای نقاط استخراج شده از تصویر مدل شی در گراف \bar{G} است.

علاوه بر این برای هر نود \bar{o}_i و یکی از k نزدیکترین همسایه آن \bar{o}_j ، یک لبه در نظر گرفته می-شود. این ارتباط با استفاده از یک بردار ویژگی باینری \bar{A}_{ij} توصیف می-شود. از معیار فاصله اقلیدسی برای یافتن نزدیکترین فاصله استفاده می-شود. \bar{A} به مجموعه تمامی بردارهای باینری در گراف \bar{G} می-باشد.

مشابه مدل، برای صحنه نیز به همین شکل عمل نموده و با استفاده از چندتایی $\mathcal{G} = \{\mathbf{O}, \mathbf{X}, \mathbf{A}\}$ مشخص گراف صحنه را مشخص می-کنیم.

5-5 تعریف مسئله

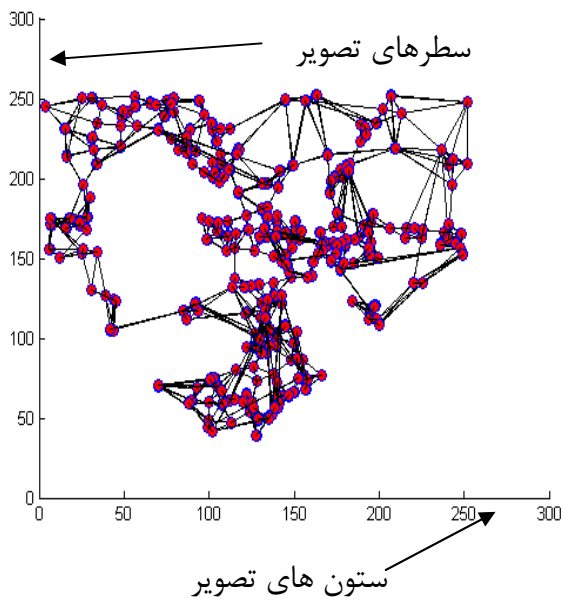
همانطور که در بخش 4-5 اشاره شد از چندگانه $\mathcal{G} = \{\mathbf{O}, \mathbf{X}, \mathbf{A}\}$ و $\bar{\mathcal{G}} = \{\bar{\mathbf{O}}, \bar{\mathbf{X}}, \bar{\mathbf{A}}\}$ به ترتیب برای معرفی گراف صحنه و مدل استفاده می-شود. در این نمایش $\mathbf{O} = \{\mathbf{o}_1, \mathbf{o}_2, \dots, \mathbf{o}_N\}$ و $\bar{\mathbf{O}} = \{\bar{\mathbf{o}}_1, \bar{\mathbf{o}}_2, \dots, \bar{\mathbf{o}}_M\}$ به ترتیب مجموعه نودهای صحنه و مدل می-باشد. هر نود در گراف مدل با استفاده از یک بردار یکانی توصیف می-شود. در این پایان نامه از $\bar{\mathbf{X}}$ برای اشاره به تمامی بردارهای یکانی در $\bar{\mathbf{G}}$ استفاده می-شود. بعلاوه در گراف مدل ارتباط بین هر جفت نود که در همسایگی هم هستند با استفاده از یک بردار باینری تعریف می-شود. برای مشخص کردن مجموعه ای شامل تمامی این بردارها از $\bar{\mathbf{A}}$ استفاده می-شود. مشابه این تعریف در گراف مدل از \mathbf{X} و \mathbf{A} به ترتیب برای توصیف ویژگی های یکانی و باینری در صحنه استفاده می-کنیم.



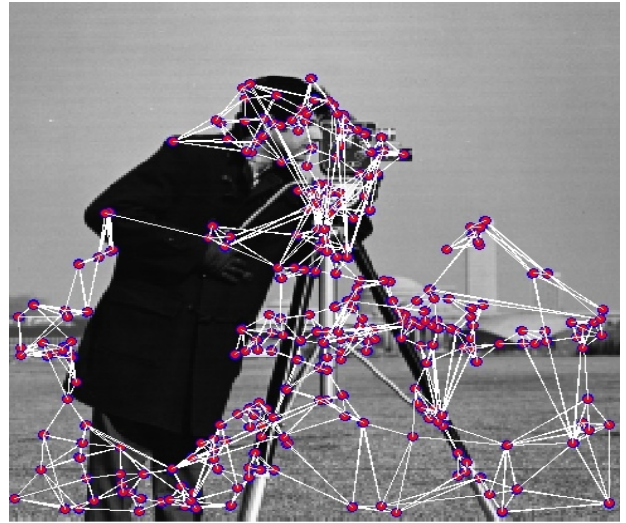
الف) تصویر اصلی



ب) اعمال SIFT روی تصویر الف



ج) رسم گراف با توجه به شکل ج



ج) رسم گراف با توجه به شکل ب

شکل 5-1: مراحل کار بر روی یک تصویر واقعی

به هر نود در دو گراف یک برچسب از مجموعه $\theta = \{\theta_1, \theta_2, \dots, \theta_N\}$ اختصاص داده می‌شود، به این صورت که θ_i برچسب اختصاص یافته شده برای نود \mathbf{o}_i به صورت مشابه از مجموعه $\Omega = \{\omega_0, \omega_2, \dots, \omega_M\}$ به عنوان مجموعه برچسب‌ها برای نودهای گراف مدل استفاده می‌کنیم. ممکن است برای یک یا چند نود از نودهای صحنه نتوان یک برچسب واقعی در مدل (Ω) استخراج کرد که مناسب آن نود باشد. استراتژی‌های روش اول و دوم برای حل این مسئله متفاوت است.

هدف از برچسب زدن انتصاب هر نود صحنه به یک برچسب مناسب از مجموعه $\Omega = \{\omega_0, \omega_2, \dots, \omega_M\}$ است. زمانیکه شی \mathbf{o}_i برچسب ω_α به صورت $\theta_i = \omega_\alpha$ مشخص می‌شود. در این ادبیات θ_i نشان‌دهنده برچسب برای شی \mathbf{o}_i می‌باشد. در مرحله تطبیق برای هر انتصاب به صورت $\theta_i = \omega_\alpha$ میزان سازگاری برچسب همسایه‌ها شی را اندازه گرفته می‌شود. برای ارزیابی این سازگاری بین دو نود $(\mathbf{o}_i, \bar{\mathbf{o}}_\alpha)$ از \mathbf{k} نزدیکترین همسایه در این دو استفاده می‌شود. پارامتر \mathbf{k} در این روش به عنوان کسری از تعداد متوسط نقاط کلیدی تمامی مدل‌ها در نظر گرفته می‌شود. فاصله بین هر دو نود همسایه در این گراف با استفاده از فاصله اقلیدسی محاسبه می‌شود.

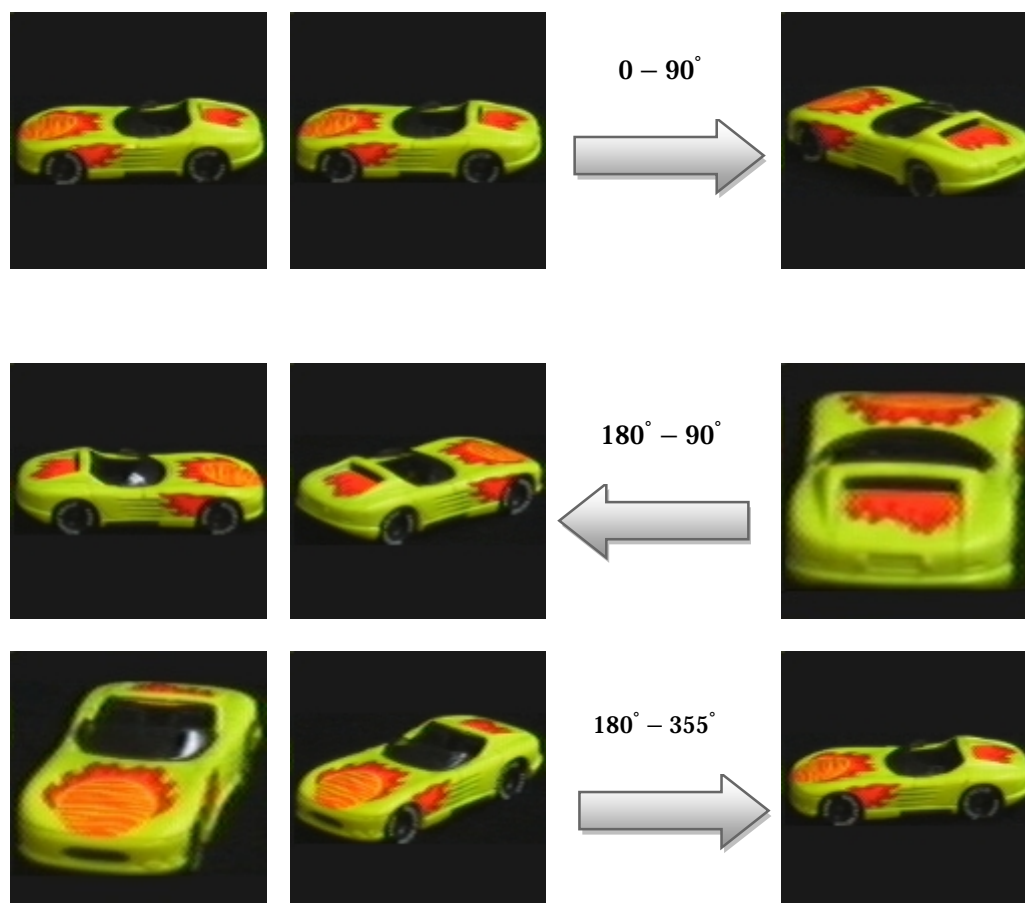
5-6 نتایج

نتایج روی سناریوی واقعی در دو حالت زیر مورد بررسی قرار می‌گیرد:

- روش اول: روش احمدی فرد
- روش دوم: روش Kostin

در هر دو روش برای انتخاب تصاویر مدل از پایگاه داده Coil برای آزمایش استفاده می‌شود. این پایگاه داده از یک شی در زوایای مختلف تصویر برداری می‌کند. پیش‌زمینه در تصاویر این پایگاه داده سیاه می‌-

باشد. از هر شی هر تصویر 71 تصویر وجود دارد که هر تصویر با تصویر قبلی 5 درجه اختلاف دارد. شکل زیر یک شی از این پایگاه داده را در زاویه‌های مختلف بین 0 تا 355 درجه نشان می‌دهد.



شکل 5-2: نمایی از پایگاه داده Coil

برای انتخاب صحنه با استفاده از یک دوربین 2 مگاپیکسل تصاویر از یک میز شلوغ جمع آوری شد و به این ترتیب صحنه ساخته شد سپس از تصاویر مدل انتخاب شده و پس از حذف پس زمینه، تصویر مدل در صحنه آماده شده قرار داده شد. در شکل زیر نمونه ایی از ساخت یک صحنه را مشاهده می‌کنید.



شکل 5-3: نمونه ایی از تصاویر ساخته شده برای صحنه

همانطور در فصل چهارم عنوان شد تغییراتی که در صحنه می تواند اتفاق بیفتد شامل:

- چرخش مدل در صحنه
- انسداد
- وجود دو شی مدل همزمان در صحنه
- تغییر مقیاس
- تغییرات روشنایی

بر این اساس سناریو ها را بر طبق تقسیم بندی بالا ادامه داده و در اولین سناریو را که همان چرخش مدل در صحنه باشد مورد بررسی قرار می گیرد.

5-6-1 چرخش مدل در صحنه

در اولین آزمایش یک مدل را در صحنه بالا قرار داده و ابتدا تنها با استفاده از ویژگی های یکانی عمل تطبیق را انجام شده و سپس با استفاده از دو روش اول و دوم به تطبیق مدل و صحنه پرداخته می شود.



شکل 4-5: تطبیق صحنه و مدل، بالا با تطبیق مستقیم، وسط روش اول و پایین روش دوم

شکل (4-5) نتایج این آزمایش را نشان می‌دهد. همانطور که در فصل چهارم گفته شد استفاده از تنها ویژگی‌های یکانی برای تطبیق اگرچه سریعتر است اما کارایی مناسبی ندارد. روش اول با در نظر گرفتن نود پوچ توانسته است کارایی بهتری داشته باشد. از دیدگاه مهندسی سیستمی کارا تر است که دارای قابلیت اطمینان بیشتری باشد. در کار ما قابلیت اطمینان یعنی تعداد تطبیق‌های غلط کمتر و روش اول با در نظر گرفتن نود ساختگی پوچ توانسته است در مورد نودهایی که نمی‌تواند تصمیم‌گیری کند برچسب اشتباه را انتخاب نکرده و برچسب پوچ را انتخاب کند. از این دیدگاه که دیدگاه بسیار کاربردی می‌باشد روش اول کارایی بالایی داشته و در حقیقت قابل اعتماد است.

برای تحلیل روش دوم اگر به شکل 4-5 قسمت پایین دقت شود تصویر به سه قسمت تقسیم شده است. ناحیه اول ناحیه ای است که تطبیق به درستی انجام شده است و نواحی یک و دو ناحیه‌هایی هستند که تطبیق به درستی در آنها انجام نشده است. این نقطه ضعف روش دوم به نبود نود ساختگی پوچ بر می‌گردد. در مقایسه با روش اول، این روش تعداد تطبیق‌های بیشتری را یافته است اما تعداد اشتباه‌های بیشتری نیز نسبت به روش اول تطبیق دارد.

5-6-2 انسداد

از آنجا که در مرحله قبل اثبات شد که تنها ویژگی‌های یکانی برای تطبیق کافی نمی‌باشند در این آزمایش تطبیق با استفاده از تنها ویژگی‌های یکانی در گرفته نشد. نتایج آزمایش با استفاده از روش اول و دوم روی یک سناریوی واقعی در شکل 5-5 نشان داده شده است. همانطور که مشاهده می‌شود، روش اول تطبیق توانسته است آن قسمتی از مدل که در صحنه وجود دارد را بیابد. نکته مهم تر اینکه برای آن قسمتی از مدل که در صحنه وجود ندارد اشتباه نکرده بلکه برای عدم وجود با استفاده از استراتژی نود پوچ توانسته مشکل را حل کند.

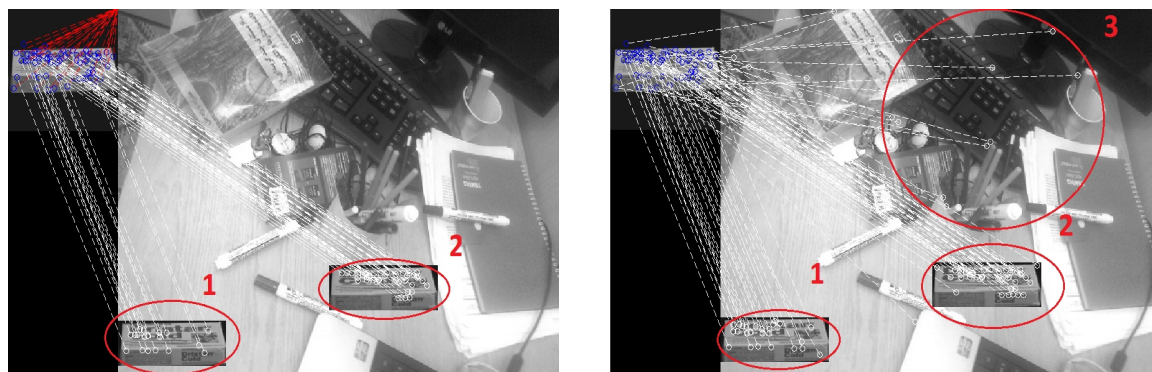


شکل 5-5: نتایج حاصل از آزمایش انسداد، بالا روش اول و پایین روش دوم

اما روش دوم در این حالت نیز توانسته آن قسمت از مدل که در صحنه وجود دارد را به درستی بیابد، اما برای قسمت حذف شده استراتژی مشخصی وجود ندارد. در فصل چهارم نیز به این مشکل اشاره شد. آنجا نیز وقتی نودی از مدل در صحنه وجود نداشت کل درایه های آن سطر برای تابع حمایت صفر در نظر گرفته می شد. هرچند آنجا می شد تحلیل کرد که چون هویت نود از صحنه حذف شده لذا این روش هیچ تناطری نیافته پس همه مقادیر آن سطر را صفر در نظر گرفته ولی در این سناریو نمی توان هیچ تحلیل برای روش دوم آورد.

3-6-5 وجود دو شی به صورت همزمان در صحنه

در این آزمایش نیز همان مشکل قبلی در روش دوم وجود دارد. در حالیکه روش دوم تعداد تطبیق‌های زیادتری یافته است. این مسئله در شکل زیر با مقایسه ناحیه‌های یک و دو برای دو تصویر راست که توسط روش دوم و چپ که توسط روش اول مشخص شده قابل مشاهده می‌باشد. ناحیه 3 که در قسمت سمت راست شکل 6-5 آمده است بیانگر همان نبود نود پوچ برای روش دوم است.

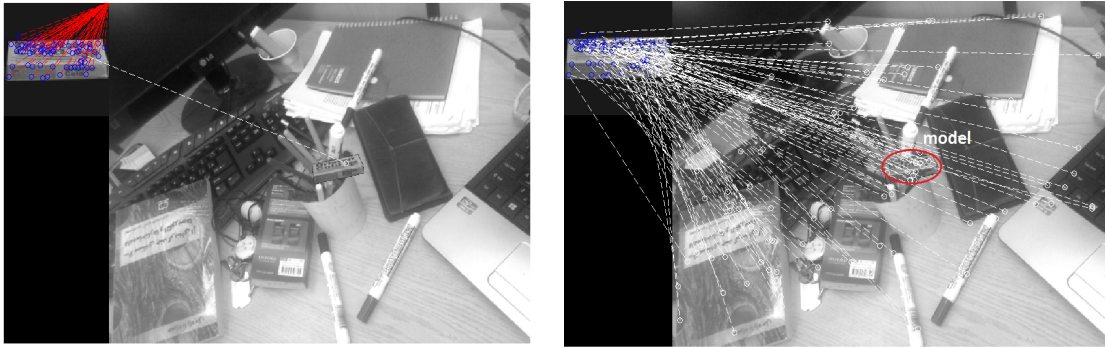


شکل 6-5: حاصل آزمایش برای وجود چند شی در صحنه

4-6-5 تغییر مقیاس

این آزمایش نیز برای هر دو روش انجام شد و البته علاوه بر اینکه شکل نصف شد چرخش نیز داده شد.

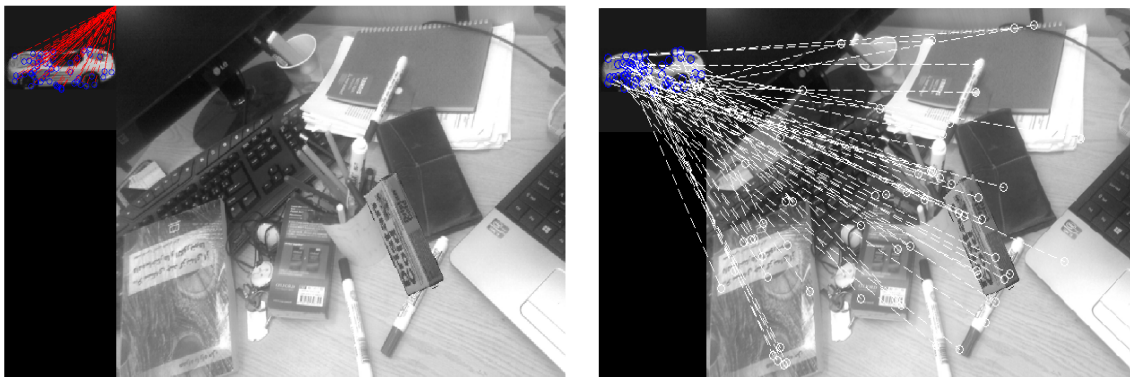
اگر به تصاویر زیر دقت کنید برای انسان نیز بسیار مشکل است که در نگاه اول بتواند مدل را با این وضعیت در صحنه پیدا کند.



شکل 5-7: تغییر مقیاس: سمت چپ با استفاده از روش اول، سمت راست با استفاده از روش دوم

5-6-5 تغییرات روشنایی

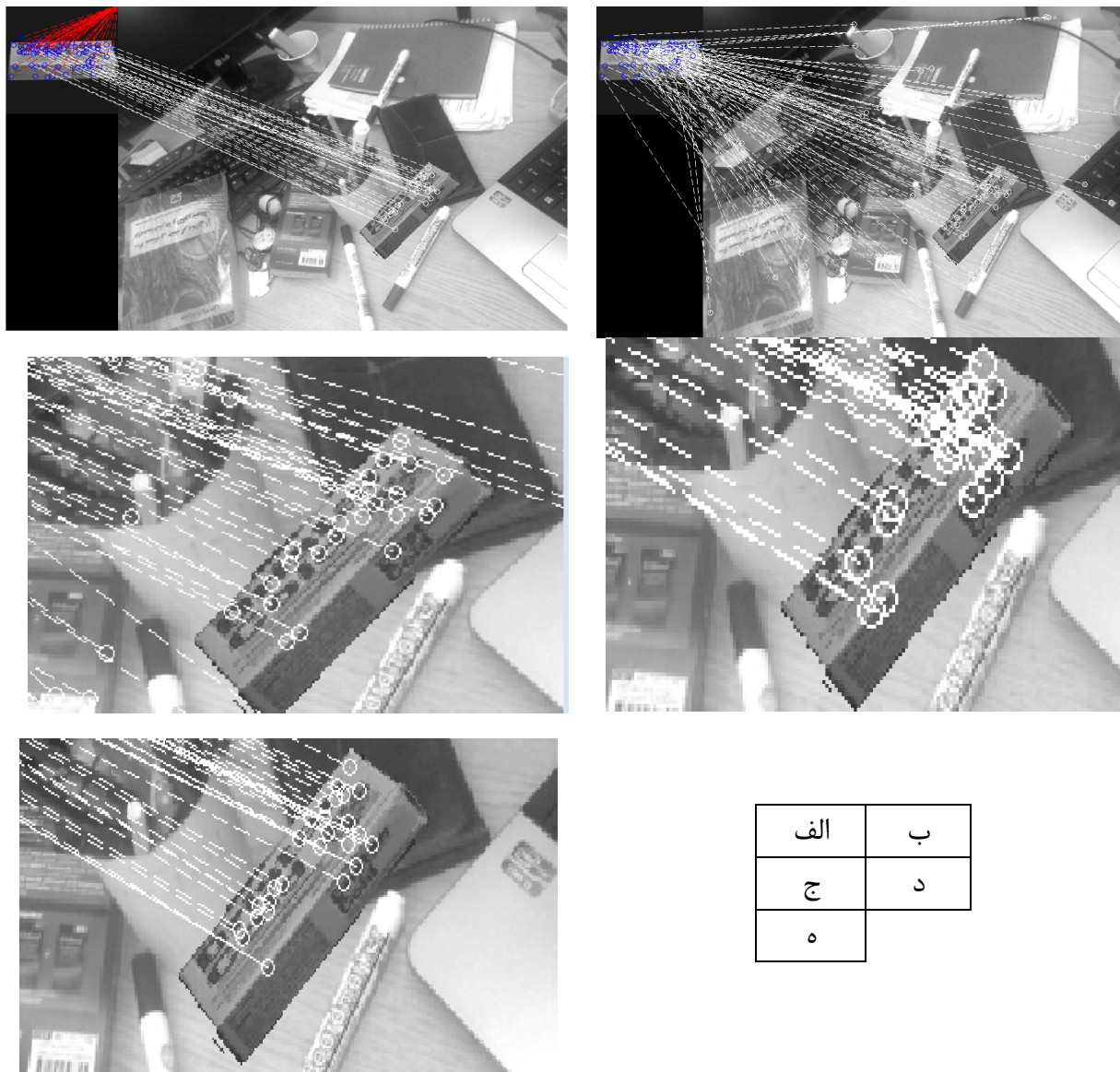
به دلیل اینکه توصیفگر SIFT تا حد بسیار بالایی نسبت به تغییرات روشنایی مقاوم است این تغییر در نظر گرفته نشد. هرچند می‌توان با استفاده از فیلتر همومورفیک¹[77] تا حدودی نور را از صحنه حذف کرد. حالت انتهایی که بررسی می‌شود حالت کلی انسداد است. زمانیکه مدل به کلی در صحنه وجود نداشته باشد. این آزمایش دقیقاً قابلیت اطمینان روش اول را نسبت به روش دوم نشان می‌دهد. همانطور که در شکل 5-8 نشان داده شده است مدل، یک ماشین است که در صحنه وجود ندارد، اما روش دوم در این آزمایش نتوانسته به درستی نبود را گزارش کند. در مقابل روش اول کاملاً این نبود را گزارش کرده است. یعنی تمام نودهای صحنه به نود پوچ متناظر شده‌اند.



شکل 5-8: حاصل آزمایش نبود مدل در صحنه

¹Homomorphic filter

7-5 بهبود تطبیق با ترکیب روش اول و دوم



الف	ب
ج	د
ه	

شکل 5-9: تطبیق: سمت چپ بالا با روش پیشنهادی، بالا سمت راست روش دوم، وسط سمت چپ روش دوم از نمای

نزدیک، وسط سمت راست روش اول از نمای نزدیک، سمت چپ پایین روش پیشنهادی از نمای نزدیک

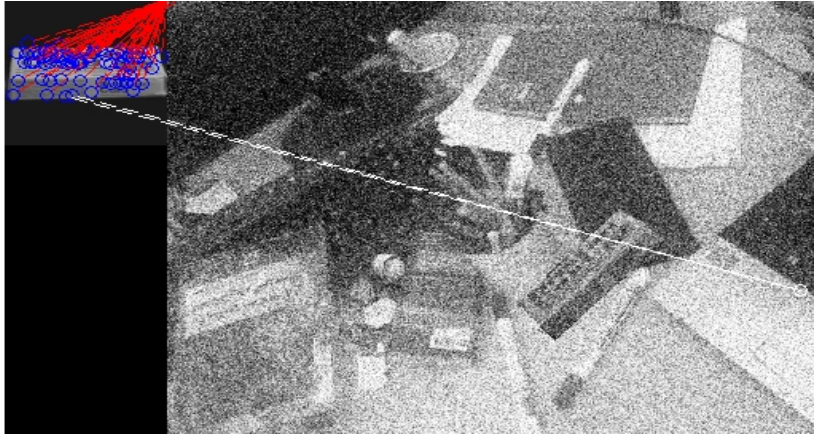
همانطور که در تمامی آزمایشهای بالا ذکر شد روش اول قابل اعتماد تر بود اما در بسیاری از موارد، روش دوم تطبیق‌های درست بیشتری را یافته است. برای اینکه مزایای دو روش را بتوان همزمان داشت سعی

شد در روش اول از قانون به روز رسانی روش دوم استفاده شود. نتایج به صورت بالا بدست آمد. اگر به شکل 5-9 دقت شود مشخص است که زمانیکه از ترکیب دو روش برای تطبیق استفاده شده است الگوریتم حاصل مزیت روش اول را داشته به این معنی که از ویژگی نود ساختگی استفاده مناسبی شده است و زمانیکه به تطبیقها دقت می شود واضح است که تا حد بسیار کمی نیز نسبت به روش اول بهبود داشته است. این بهبود با تغییر قانون برای نود پوچ متناسب با فضای الگوریتم تطبیق قابلیت بهبود بیشتر را داراست. قابل ذکر است چون ویژگی نود پوچ در این الگوریتم ترکیبی وجود دارد لذا در تغییراتی مانند انسداد و نبود مدل در صحنه این الگوریتم مقاوم است.

5-8 بررسی تاثیر نویز

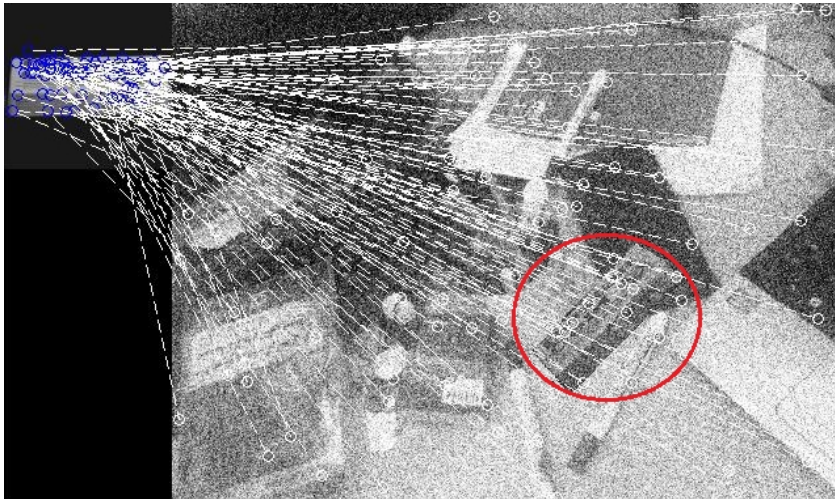
یکی از معمول ترین مشکلات در یک صحنه، وجود نویز است. برای بررسی تاثیر نویز و اینکه شبکه تا چه هنگام قادر است نویز را تحمل کند نویز نرمال با واریانس متفاوت به صحنه و مدل اضافه کرده و تحمل هر دو روش مورد ارزیابی قرار می گیرد. در اولین آزمایش نویزی نرمال با قدرت های مختلف به سیستم وارد کرده و تحمل روش اول را مورد ارزیابی قرار می دهیم.

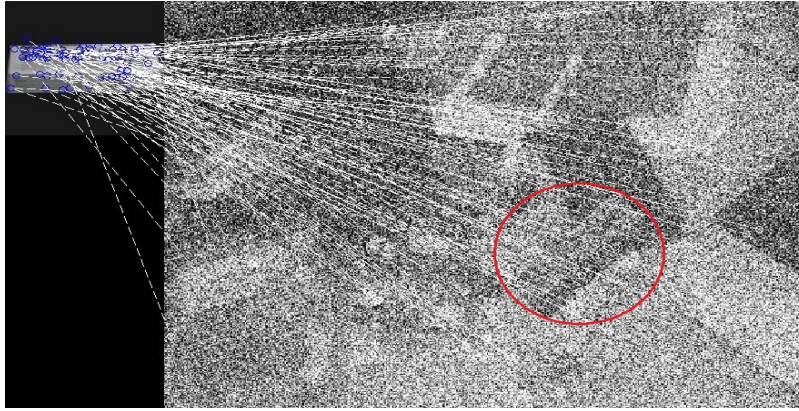




شکل 5-10: اضافه کردن نویز به سیستم و آزمایش و روش اول، قدرت نویز در تصویر بالا 0.0005، پایین 0.05

در فصل چهارم نشان داده شد زمانیکه نویز به شبکه وارد می‌شود پس از یک آستانه سیستم دچار یک تلاطم شده و به شدت نرخ بازشناسی کاهش می‌یافت. در سناریو واقعی نیز تا حدی این وضعیت وجود دارد و سیستم پس از یک آستانه (0.0005) دچار تلاطم شد. روش اول همانطور که از شکل 5-10 مشخص است توانسته در نویز با قدرت 0.05 بسیاری از نودها را به سمت نود ساختگی بکشد. برای روش دوم نیز به صورت مشابه با روش اول نویز با قدرت مختلف به سیستم وارد شد و تحمل این روش مورد ارزیابی قرار گرفت. در فصل چهارم عنوان شد که روش دوم در برابر نویز نسبت به روش اول مقاوم تر است این گفته در سناریوی واقعی هم اتفاق افتاد.





شکل 5-11: اضافه کردن نویز به سیستم و آزمایش و روش اول، قدرت نویز در تصویر بالا 0.05، پایین 0.5

5-9 جمع بندی

در این فصل نتایج روی چندین سناریو مورد ارزیابی قرار گرفت. در هر سناریو دو روش اول و دوم مورد آزمایش قرار گرفت. در هر سناریو مشخص شد که روش اول قابلیت اطمینان بیشتری نسبت به روش دوم از لحاظ تصمیم درست گرفتن دارد. در مقابل روش دوم تعداد تطبیق های درست بیشتری پیدا می کند. برای داشتن مزیت های دو روش، روشی را پیشنهاد داده شد که از تلفیق روش اول و دوم حاصل می گردد. بطوریکه از قانون به روز رسانی روش دوم برای روش اول استفاده شد. نتایج حاکی از آن بود که روش ترکیبی قابلیت اطمینان مناسبی دارد و تا حدودی نسبت به دو روش نرخ بازشناسی بهتری دارد. برای بررسی نویز نیز در محیط واقعی مطالعه ای صورت گرفت و گفته های فصل چهارم در مورد نویز در محیط مجازی اینجا نیز تایید شد. در حضور نویز روش دوم آستانه مقاومت بالاتری دارد هرچند که در روش اول این شرایط نیز قابل اعتماد تر است.

آنجا که قوانین کلاسیک ریاضیات به طبیعت اشاره می‌کنند مطمئن نیستند و آنجا که مطمئن هستند نمی‌توانند به طبیعت اشاره ای داشته باشند

آلبرت اینشتین ۱۸۷۹-۱۹۵۵

فصل ششم

کارهای آتی

فصل ششم

در فصل دوم کارهای انجام شده در دو دسته: 1- روش‌های مرسوم بازشناسی اشیاء و 2- روش‌هایی که برای بازشناسی از بینایی انسان تقلید می‌کنند، دسته بندی شد.

در این فصل هم می‌توان دو دیدگاه را برای ادامه کار پیشنهاد داد. اولین دیدگاه، دیدگاه تکنیکی است به این معنی که برای بهبود کار به تکنیک‌های استفاده شده دقت شود، سپس سعی شود این تکنیک‌ها تا حد امکان بهبود یابند تا بتوان در نهایت کارایی سیستم را بهبود داد. اما دیدگاه دوم تغییر نوع نگرش است.

6-1 دیدگاه اول: دیدگاه تکنیکی

در روش‌های مرسوم، عمل بازشناسی در دو مرحله انجام می‌شود: 1- بازنمایی، 2- تطبیق. در این پایان نامه از توصیفگر محلی SIFT برای بازنمایی تصویر مدل و صحنه استفاده شد. این توصیفگر، توصیفگر قدرتمندی است ولی پیچیدگی محاسباتی نسبتاً بالایی دارد. برای بهبود این توصیفگر از این لحاظ کارهای بسیار زیادی ارائه شده که برای مطالعه بیشتر می‌توانید به فصل کارهای انجام شده مراجعه کنید. یکی از راه‌های تکنیکی برای بهبود سرعت سیستم، استفاده از توصیفگرهایی است که سریعتر هستند. یکی دیگر از پیشنهادات که می‌توان در حوزه بازنمایی در نظر گرفت انجام یک پس پردازش بعد از اعمال توصیفگر SIFT است. تا توصیفگرهای نزدیک به هم را در یک خوشه قرار دهد و از مراکز خوشه‌های بدست آمده برای ساخت گراف استفاده نماید. به عبارت بهتر این پس پردازش یک بازنمایی تُنک از مدل و صحنه ارائه می‌دهد. بهبود عملکرد توصیفگرها با توجه به تحقیقات فراوان صورت گرفته تاکنون چشمگیر نخواهد بود. بهبود دیگر در فاز تطبیق می‌باشد. در این پایان نامه از یک فرایند آرام‌سازی احتمالاتی برای تطبیق استفاده شد. برای مقایسه از دو روش که اساس آنها همین فرآیند است ولی در تابع حمایت باهم متفاوت هستند استفاده شد. همچنین در فصل چهارم مزایا و معایب دو روش تحلیل و بررسی شد. برای

بهبود روش تطبیق از ترکیب این دو روش استفاده نمودیم تا از مزایای هر دو روش بتوان استفاده کرد. از دیگر بهبودهایی که می‌توان اشاره کرد اضافه کردن ویژگی رنگ به عنوان یک ویژگی تاثیر گذار به توصیفگر اشاره کرد

2-6 دیدگاه دوم: دیدگاه تقلیدی از انسان

مشکلات روش‌های مرسوم در فصل کارهای انجام شده آمده است. در کنار مشکلات ذکر شده یکی از مشکلات روش‌های مرسوم بازشناسی ناتوانی در مدل کردن برخی تغییرات (برای مثال الاستیسیته) در تصویر شی (اشیاء) می‌باشد. این تغییر باعث ایجاد یک خاصیت کشسانی در تصویر می‌شود. توصیفگر SIFT با توجه به در نظر گرفتن جهت و مقیاس برای هر نقطه کلیدی تا حدی بسیار کمی می‌تواند این تغییر را مدل کند. اما در تغییرات بیشتر ناتوان است. روش‌های دیگر مرسوم بازشناسی نیز در مدل کردن این تغییر ناتوان هستند. در چنین حالاتی بهبود تکنیک راه گشا نیست بلکه باید تغییر نگرش داد. به عبارتی تکنیک‌های مرسوم توانایی مدل کردن این سطح از پیچیدگی از تصویر را ندارند. یادآور می‌شود که انسان به راحتی در برابر تغییراتی از این دست مقاوم است. به این معنی که انسان در این شرایط می‌تواند عمل بازشناسی را انجام دهد.

یکی از بهبودها، تقلید از توانایی انسان در بازشناسی اشیاء می‌باشد. برای اینکه این اتفاق حاصل شود باید به صورت دقیق بررسی شود که انسان چگونه عمل بازشناسی را انجام می‌دهد. این عمل به صورت سلسله مراتبی در لایه کورتکس مغز انجام می‌شود. انسان ابتدا ویژگی‌های رنگ را مورد بررسی قرار می‌دهد و سپس از ویژگی‌یافت استفاده می‌کند. بهبود شامل استفاده از این معماری برای بازشناسی است برای فهم دقیق تر این موضوع به مرجع [71] مراجعه شود.

پیوست ریاضی

همانطور که در فصل مبانی تئوری گفته شد. در نهایت فرآیند آرم سازی احتمالاتی را میتوان مانند فرمول زیر در نظر گرفت

$$P(\theta_i = \omega_\alpha) = \max_{\omega_\lambda \in \Omega} P(\theta_i = \omega_\alpha | \bar{x}_{j,j \in N_i^+}, \bar{A}_{ij,j \in N_i}) \quad 1$$

همانطور که روشن است اکثر الگوریتم‌های گراف، از مسائل NP می‌باشد. لذا سعی می‌شود تا مسئله به صورت آرام شده مورد نظر گرفته شود. از یک چارچوب بیزین برای این کار استفاده شده و با این توجیح فرمول 1 را می‌توان به صورت زیر در چارچوب بیزین در نظر گرفت:

$$P(\theta_i = \omega_\alpha | \bar{x}_{j,j \in N_i^+}, \bar{A}_{ij,j \in N_i}) = \frac{P(\theta_i = \omega_\alpha | \bar{x}_{j,j \in N_i^+}, \bar{A}_{ij,j \in N_i})}{P(\bar{x}_{j,j \in N_i^+}, \bar{A}_{ij,j \in N_i})} \quad 2$$

فرمول 2 را می‌توان روی تک تک نودهای گراف بسط داد و به این صورت می‌توان فرمول 2 را به صورت زیر نوشت

$$\frac{\sum_{\omega_{\theta_1} \in \Omega} \cdots \sum_{\omega_{\theta_{i-1}} \in \Omega} \cdots \sum_{\omega_{\theta_{i+1}} \in \Omega} \cdots \sum_{\omega_{\theta_N} \in \Omega} P(\theta_1 = \omega_{\theta_1}, \dots, \theta_i = \omega_\alpha, \dots, \theta_N = \omega_{\theta_N}, \bar{x}_{j,j \in N_i^+}, \bar{A})}{\sum_{\omega_{\theta_1} \in \Omega} \cdots \sum_{\omega_{\theta_N} \in \Omega} P(\theta_1 = \omega_{\theta_1}, \dots, \theta_i = \omega_\alpha, \dots, \theta_N = \omega_{\theta_N}, \bar{x}_{j,j \in N_i^+}, \bar{A}_{ij,j \in N_i})} \quad 3$$

این بسط تمامی حالات ممکن برای برچسب زدن نودهای صحنه را در نظر می‌گیرد.

طبق فرمول سوم هرگاه هر نود معادل یک شی در صحنه باشد. مرتبه زمانی الگوریتم ارائه شده برای فرمول 3 از مرتبه زمانی توانی متناسب با تعداد اشیا در گراف مدل صحنه می‌باشد. این بدان معناست که با افزایش تعداد نودها (تعداد اشیا) پیچیدگی الگوریتم به صورت توانی افزایش می‌یابد. برای اینکه این موضوع را نشان دهیم فرض کنید که احتمال انتصاب $P(\theta_1 = \omega_{\theta_1}, \dots, \theta_i = \omega_\alpha, \dots, \theta_N = \omega_{\theta_N})$

($\omega_{\theta_N}, \bar{x}_{j,j \in N_i^+}, \bar{A}_{ij,j \in N_i}$) در t واحد محاسبه انجام شود، اگر N نود در گراف صحنه وجود داشته باشد. برای در نظر گرفتن تمامی حالات طبق اصل ضرب $t \times t \times \dots \times t$ ، باز طول بکشد. با این تفاسیر الگوریتم برای در نظر گرفتن تمامی حالات ممکن دارای پیچیدگی زمانی t^N می‌باشد. این به معنای پیچیدگی از مرتبه توانی می‌باشد. یعنی با افزایش تعدا نودها مرتبه الگوریتم با رشد توانی افزایش می‌یابد. برای اینکه مشکلات توانی بودن مرتبه فرمول 3، حل شود فرض های زیر در نظر گرفته می‌شود:

- معیار اطلاعات یکانی در مجموعه $\bar{x}_{j,j \in N_i^+}$ کاملاً از هم مستقل هستند.
- ارتباطات باینری نیز از این قاعده مستثنی نبوده و این ارتباطات $\bar{A}_{ij,j \in N_i}$ نیز از یکدیگر مستقل هستند
- نکته دیگر که باید مورد توجه قرار گیرد این است که ارتباط بین ارتباط یکانی و باینری هم باید مشخص شود. وقوع اطلاعات یکانی کاملاً مستقل از ارتباطات باینری در نظری گرفته می‌شود.
- اشتراک وقوع جفت اتفاق $\theta_i = \omega_{\theta_i}$ و $\theta_j = \omega_{\theta_j}$ صفر در نظر گرفته می‌شود به عبارت دیگر این دو رویداد نیز مستقل در نظر گرفته می‌شود.

چگالی احتمال توام که در صورت و مخرج معادله 3-8 ظاهر می‌شود را می‌توان به صورت جملات شرطی به صورت زیر نوشت :

$$\begin{aligned}
 & P\left(\theta_1 = \omega_{\theta_1}, \dots, \theta_i = \omega_{\theta_i}, \dots, \theta_N = \omega_{\theta_N}, \bar{x}_{j,j \in N_i^+}, \bar{A}_{ij,j \in N_i}\right) & 4 \\
 & = P\left(\bar{x}_{j,j \in N_i^+} \mid \theta_1 = \omega_{\theta_1}, \dots, \theta_i = \omega_{\theta_i}, \dots, \theta_N = \omega_{\theta_N}, \bar{A}_{ij,j \in N_i}\right) \\
 & \quad \times P\left(\theta_1 = \omega_{\theta_1}, \dots, \theta_i = \omega_{\theta_i}, \dots, \theta_N = \omega_{\theta_N}, \bar{A}_{ij,j \in N_i}\right)
 \end{aligned}$$

با سومین فرض برای سادگی محاسبات می توان ارتباطات باینری را از لیست متغیرهای شرطی از معادله 4 حذف کرد. با فاکتورگیری جمله دوم در معادله 4 می توان معادله 4 را به صورت زیر بازنویسی کرد:

$$\begin{aligned}
 & P(\theta_1 = \omega_{\theta_1}, \dots, \theta_i = \omega_{\theta_i}, \dots, \theta_N = \omega_{\theta_N}, \bar{A}_{ij, j \in N_i}) \quad 5 \\
 & = P(\bar{A}_{i1} | \bar{A}_{i2} \dots, \bar{A}_{ii-1}, \bar{A}_{ii+1}, \dots, \bar{A}_{iN}, \theta_1 = \omega_{\theta_1}, \dots, \theta_N = \omega_{\theta_N}) \\
 & \times P(\bar{A}_{i2} | \bar{A}_{i3} \dots, \bar{A}_{ii-1}, \bar{A}_{ii+1}, \dots, \bar{A}_{iN}, \theta_1 = \omega_{\theta_1}, \dots, \theta_N = \omega_{\theta_N}) \\
 & \times \dots \times P(\bar{A}_{iN} | \theta_1 = \omega_{\theta_1}, \dots, \theta_N = \omega_{\theta_N}) \times P(\theta_1 = \omega_{\theta_1}, \dots, \theta_N = \omega_{\theta_N}) \\
 & = P(\theta_i = \omega_{\theta_i}) \left\{ \prod_{j \in N_i} P(\bar{A}_{ij} | \theta_j = \omega_{\theta_j}, \theta_i = \omega_{\theta_i}) \times P(\theta_j = \omega_{\theta_j}) \right\}
 \end{aligned}$$

حالا ابتدا معادله 2 را کمی بسط داده و در ادامه رابطه 5 را در معادله بسط داده شده جایگزین می کنیم

$$P(\theta_i = \omega_{\theta_i} | \bar{x}_{j, j \in N_i^+}, \bar{A}_{ij, j \in N_i}) = \frac{P(\theta_i = \omega_{\theta_i} | \bar{x}_{j, j \in N_i^+}, \bar{A}_{ij, j \in N_i})}{P(\bar{x}_{j, j \in N_i^+}, \bar{A}_{ij, j \in N_i})} = \frac{P(\theta_i = \omega_{\theta_i} | \bar{x}_{j, j \in N_i^+}) \times P(\bar{x}_{j, j \in N_i^+}, \bar{A}_{ij, j \in N_i})}{P(\bar{x}_{j, j \in N_i^+}, \bar{A}_{ij, j \in N_i})} \quad 6$$

در رابطه 2 $P(\bar{x}_{j, j \in N_i^+})$ ویژگی یکانی را معرفی می کند. در رابطه 6 اگر صورت و مخرج را مانند رابطه 5 بسط دهیم $P(\theta_i = \omega_{\theta_i})$ از صورت و مخرج ساده شده و در انتها روابط زیر حاصل می شود.

$$\begin{aligned}
 & P(\theta_i = \omega_{\theta_i} | \bar{x}_{j, j \in N_i^+}, \bar{A}_{ij, j \in N_i}) = \frac{P(\theta_i = \omega_{\theta_i}) \times Q(\theta_i = \omega_{\theta_i})}{\sum_{\omega_{\lambda} \in \Omega} P(\theta_i = \omega_{\lambda}) \times Q(\theta_i = \omega_{\lambda})} \quad 7 \\
 & Q(\theta_i = \omega_{\theta_i}) = \prod_{j \in N_i} \left\{ \sum_{\omega_{\beta} \in \Omega} P(\theta_j = \omega_{\beta} | \bar{x}_j) \times P(\bar{A}_{ij} | \theta_i = \omega_{\theta_i}, \theta_j = \omega_{\beta}) \right\}
 \end{aligned}$$

- [1] A. Andreopoulos and J. K. Tsotsos, “50 Years of Object Recognition: Directions Forward,” *Comput. Vis. Image Underst.*, 2013.
- [2] A. R. Ahmadyfard and J. Kittler, “Using relaxation technique for region-based object recognition,” *Image Vis. Comput.*, vol. 20, no. 11, pp. 769–781, 2002.
- [3] S. K. Naik and C. A. Murthy, “Distinct multicolored region descriptors for object recognition,” *Pattern Anal. Mach. Intell. IEEE Trans.*, vol. 29, no. 7, pp. 1291–1296, 2007.
- [4] M. J. Swain and D. H. Ballard, “Color indexing,” *Int. J. Comput. Vis.*, vol. 7, no. 1, pp. 11–32, 1991.
- [5] M. A. Stricker, “Color and geometry as cues for indexing,” *Tec. Rep. CS92*, vol. 22, 1992.
- [6] M. A. Turk and A. P. Pentland, “Face recognition using eigenfaces,” in *Computer Vision and Pattern Recognition, Proceedings CVPR’91., IEEE Computer Society Conference on*, pp. 586–591.199.
- [7] P. Tu, R. Hartley, and T. Saxena, “Recognizing objects using color-annotated adjacency graphs,” in *Shape, contour and grouping in computer vision*, Springer, pp. 246–263, 1999.
- [8] A. Kostin, J. Kittler, and W. Christmas, “Object recognition by symmetrised graph matching using relaxation labelling with an inhibitory mechanism,” *Pattern Recognit. Lett.*, vol. 26, no. 3, pp. 381–393, 2005.
- [9] H. J. Wolfson, “Model-based object recognition by geometric hashing,” in *Computer Vision—ECCV 90*, Springer, pp. 526–536, 1990.
- [10] S. Procter and J. Illingworth, “Foresight: Fast object recognition using geometric hashing with edge-triple features,” in *Image Processing. Proceedings., International Conference on*, vol. 1, pp. 889–892.1997, 1997.
- [11] H. Murase and S. K. Nayar, “Visual learning and recognition of 3-D objects from appearance,” *Int. J. Comput. Vis.*, vol. 14, no. 1, pp. 5–24, 1995.
- [12] A. R. Ahmadyfard and J. Kittler, “Using relaxation technique for region-based object recognition,” *Image Vis. Comput.*, vol. 20, no. 11, pp. 769–781, 2002.

- [13] B. McFee, C. Galleguillos, and G. Lanckriet, “Contextual object localization with multiple kernel nearest neighbor,” *Image Process. IEEE Trans.*, vol. 20, no. 2, pp. 570–585, 2011.
- [14] K. Mikolajczyk and C. Schmid, “A performance evaluation of local descriptors,” *Pattern Anal. Mach. Intell. IEEE Trans.*, vol. 27, no. 10, pp. 1615–1630, 2005.
- [15] D. G. Lowe, “Distinctive image features from scale-invariant keypoints,” *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, 2004.
- [16] M. Pietikäinen, T. Ojala, and Z. Xu, “Rotation-invariant texture classification using feature distributions,” *Pattern Recognit.*, vol. 33, no. 1, pp. 43–52, 2000.
- [17] J. Matas, O. Chum, M. Urban, and T. Pajdla, “Robust wide-baseline stereo from maximally stable extremal regions,” *Image Vis. Comput.*, vol. 22, no. 10, pp. 761–767, 2004.
- [18] K. Mikolajczyk and C. Schmid, “Indexing based on scale invariant interest points,” in *Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on*, vol. 1, pp. 525–531, 2001.
- [19] K. Mikolajczyk and C. Schmid, “An affine invariant interest point detector,” in *Computer Vision—ECCV 2002*, Springer, pp. 128–142, 2002.
- [20] Y. Ke and R. Sukthankar, “PCA-SIFT: A more distinctive representation for local image descriptors,” in *Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on*, vol. 2, pp. II–506, 2004.
- [21] S. Belongie, J. Malik, and J. Puzicha, “Shape matching and object recognition using shape contexts,” *Pattern Anal. Mach. Intell. IEEE Trans.*, vol. 24, no. 4, pp. 509–522, 2002.
- [22] J.-M. Morel and G. Yu, “ASIFT: A new framework for fully affine invariant image comparison,” *SIAM J. Imaging Sci.*, vol. 2, no. 2, pp. 438–469, 2009.
- [23] C. Ancuti and P. Bekaert, “SIFT-CCH: Increasing the SIFT distinctness by Color Co-occurrence Histograms,” in *Image and Signal Processing and Analysis, 2007. ISPA 2007. 5th International Symposium on*, pp. 130–135, 2007.
- [24] S.-O. Shim and T.-S. Choi, “Image indexing by modified color cooccurrence matrix,” in *Acoustics, Speech, and Signal Processing, 2003. Proceedings.(ICASSP’03). 2003 IEEE International Conference on*, vol. 3, pp. III–577, 2003.

- [25] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, “Speeded-up robust features (SURF),” *Comput. Vis. image Underst.*, vol. 110, no. 3, pp. 346–359, 2008.
- [26] J. Bauer, N. Sunderhauf, and P. Protzel, “Comparing several implementations of two recently published feature detectors,” in *Proc. of the International Conference on Intelligent and Autonomous Systems, IAV, Toulouse, France*, pp. 224–227, 2007.
- [27] P. Moreno, A. Bernardino, and J. Santos-Victor, “Improving the SIFT descriptor with smooth derivative filters,” *Pattern Recognit. Lett.*, vol. 30, no. 1, pp. 18–26, 2009.
- [28] H. Li, Y. Wei, L. Li, and Y. Yuan, “Similarity learning for object recognition based on derived kernel,” *Neurocomputing*, vol. 83, pp. 110–120, 2012.
- [29] X. Tan, S. Chen, Z.-H. Zhou, and J. Liu, “Face recognition under occlusions and variant expressions with partial similarity,” *Inf. Forensics Secur. IEEE Trans.*, vol. 4, no. 2, pp. 217–230, 2009.
- [30] Y. Tang, L. Li, and X. Li, “Learning similarity with multikernel method,” *Syst. Man, Cybern. Part B Cybern. IEEE Trans.*, vol. 41, no. 1, pp. 131–138, 2011.
- [31] L. Jia and L. Kitchen, “Object-based image similarity computation using inductive learning of contour-segment relations,” *Image Process. IEEE Trans.*, vol. 9, no. 1, pp. 80–87, 2000.
- [32] C.-Y. Yen and K. J. Cios, “Image recognition system based on novel measures of image similarity and cluster validity,” *Neurocomputing*, vol. 72, no. 1, pp. 401–412, 2008.
- [33] D. P. Huttenlocher, G. A. Klanderman, and W. J. Rucklidge, “Comparing images using the Hausdorff distance,” *Pattern Anal. Mach. Intell. IEEE Trans.*, vol. 15, no. 9, pp. 850–863, 1993.
- [34] P. Simard, Y. LeCun, and J. S. Denker, “Efficient pattern recognition using a new transformation distance,” in *Advances in Neural Information Processing Systems 5*, pp. 50–58, 1992.
- [35] H. G. Barrow and R. J. Popplestone, “Relational descriptions in picture processing,” *Mach. Intell.*, vol. 6, no. 377–396, pp. 2–3, 1971.
- [36] A. Pearce, T. Caelli, and W. F. Bischof, “Rulegraphs for graph matching in pattern recognition,” *Pattern Recognit.*, vol. 27, no. 9, pp. 1231–1247, 1994.
- [37] J. Lee, “Feature Correspondence and Deformable Object Matching via Agglomerative Correspondence Clustering.”

- [38] S. W. Lu, Y. Ren, and C. Y. Suen, "Hierarchical attributed graph representation and recognition of handwritten Chinese characters," *Pattern Recognit.*, vol. 24, no. 7, pp. 617–632, 1991.
- [39] R. C. Read and D. G. Corneil, "The graph isomorphism disease," *J. Graph Theory*, vol. 1, no. 4, pp. 339–363, 1977.
- [40] J. R. Ullmann, "An algorithm for subgraph isomorphism," *J. ACM*, vol. 23, no. 1, pp. 31–42, 1976.
- [41] H. Bunke and K. Shearer, "A graph distance metric based on the maximal common subgraph," *Pattern Recognit. Lett.*, vol. 19, no. 3, pp. 255–259, 1998.
- [42] M. A. Eshera and K.-S. Fu, "An image understanding system using attributed symbolic representation and inexact graph-matching," *Pattern Anal. Mach. Intell. IEEE Trans.*, no. 5, pp. 604–618, 1986.
- [43] L. G. Shapiro and R. M. Haralick, "A metric for comparing relational descriptions," *Pattern Anal. Mach. Intell. IEEE Trans.*, no. 1, pp. 90–94, 1985.
- [44] L. G. Shapiro and R. M. Haralick, "Structural descriptions and inexact matching," *Pattern Anal. Mach. Intell. IEEE Trans.*, no. 5, pp. 504–519, 1981.
- [45] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *Pattern Anal. Mach. Intell. IEEE Trans.*, no. 6, pp. 721–741, 1984.
- [46] L. Héroult, R. Horaud, F. Veillon, and J.-J. Niez, "Symbolic image matching by simulated annealing," in *4th British Machine Vision Conference (BMVC'90)*, 1990, pp. 319–324.
- [47] D. L. Waltz, "Generating semantic description from drawings of scenes with shadows," 1972.
- [48] D. B. Fogel, "An introduction to simulated evolutionary optimization," *Neural Networks, IEEE Trans.*, vol. 5, no. 1, pp. 3–14, 1994.
- [49] C. Wang and K. Abe, "Region correspondence by inexact attributed planar graph matching," in *Computer Vision, 1995. Proceedings., Fifth International Conference on*, pp. 440–447, 1995.
- [50] L. S. Shapiro and J. Michael Brady, "Feature-based correspondence: an eigenvector approach," *Image Vis. Comput.*, vol. 10, no. 5, pp. 283–288, 1992.

- [51] S. Umeyama, “An eigendecomposition approach to weighted graph matching problems,” *Pattern Anal. Mach. Intell. IEEE Trans.*, vol. 10, no. 5, pp. 695–703, 1988.
- [52] A. Rangarajan and E. D. Mjolsness, “A Lagrangian relaxation network for graph matching,” *Neural Networks, IEEE Trans.*, vol. 7, no. 6, pp. 1365–1381, 1996.
- [53] A. Rosenfeld and S. W. Zucker, “Scene Labeling,” no. 6, pp. 420–433, 1976.
- [54] E. R. Hancock and J. Kittler, “Discrete relaxation,” *Pattern Recognit.*, vol. 23, no. 7, pp. 711–733, 1990.
- [55] J. Kittler and E. R. Hancock, “Combining evidence in probabilistic relaxation,” *Int. J. Pattern Recognit. Artif. Intell.*, vol. 3, no. 01, pp. 29–51, 1989.
- [56] W. J. Christmas, J. Kittler, and M. Petrou, “Structural matching in computer vision using probabilistic relaxation,” *Pattern Anal. Mach. Intell. IEEE Trans.*, vol. 17, no. 8, pp. 749–764, 1995.
- [57] Y.-J. Zhang, *Advances in image and video segmentation*. IGI Global, 2006.
- [58] G. Sanromà, R. Alquézar, and F. Serratosa, “A new graph matching method for point-set correspondence using the EM algorithm and Softassign,” *Comput. Vis. Image Underst.*, vol. 116, no. 2, pp. 292–304, 2012.
- [59] C. Harris and M. Stephens, “A combined corner and edge detector.,” in *Alvey vision conference*, vol. 15, p. 50, 1988.
- [60] B. Luo and E. R. Hancock, “A unified framework for alignment and correspondence,” *Comput. Vis. Image Underst.*, vol. 92, no. 1, pp. 26–55, 2003.
- [61] M. J. Black and A. Rangarajan, “On the unification of line processes, outlier rejection, and robust statistics with applications in early vision,” *Int. J. Comput. Vis.*, vol. 19, no. 1, pp. 57–91, 1996.
- [62] R. C. Wilson and E. R. Hancock, “Structural matching by discrete relaxation,” *Pattern Anal. Mach. Intell. IEEE Trans.*, vol. 19, no. 6, pp. 634–648, 1997.
- [63] A. D. J. Cross and E. R. Hancock, “Graph matching with a dual-step EM algorithm,” *Pattern Anal. Mach. Intell. IEEE Trans.*, vol. 20, no. 11, pp. 1236–1253, 1998.
- [64] B. Luo and E. R. Hancock, “Structural graph matching using the EM algorithm and singular value decomposition,” *Pattern Anal. Mach. Intell. IEEE Trans.*, vol. 23, no. 10, pp. 1120–1136, 2001.

- [65] J. Zou, C.-C. Liu, Y. Zhang, and G.-F. Lu, "Object recognition using Gabor co-occurrence similarity," *Pattern Recognit.*, 2012.
- [66] M. Calonder, V. Lepetit, M. Ozuysal, T. Trzcinski, C. Strecha, and P. Fua, "BRIEF: Computing a local binary descriptor very fast," *Pattern Anal. Mach. Intell. IEEE Trans.*, vol. 34, no. 7, pp. 1281–1298, 2012.
- [67] C. Strecha, A. M. Bronstein, M. M. Bronstein, and P. Fua, "LDAHash: Improved matching with smaller descriptors," *Pattern Anal. Mach. Intell. IEEE Trans.*, vol. 34, no. 1, pp. 66–78, 2012.
- [68] M. Harandi, J. Taheri, and B. C. Lovell, "Ensemble Learning for Object Recognition and Tracking," in *Pattern Recognition, Machine Intelligence and Biometrics*, Springer, 2011, pp. 261–278.
- [69] Y.-H. Liu, A. J. T. Lee, and F. Chang, "Object recognition using discriminative parts," *Comput. Vis. Image Underst.*, vol. 116, no. 7, pp. 854–867, 2012.
- [70] S. J. Russell, P. Norvig, J. F. Canny, J. M. Malik, and D. D. Edwards, *Artificial intelligence: a modern approach*, vol. 74. Prentice hall Englewood Cliffs, 1995.
- [71] J. J. DiCarlo, D. Zoccolan, and N. C. Rust, "How does the brain solve visual object recognition?," *Neuron*, vol. 73, no. 3, pp. 415–434, 2012.
- [72] S. Jeong and M. Lee, "Adaptive object recognition model using incremental feature representation and hierarchical classification," *Neural Networks*, vol. 25, pp. 130–140, 2012.
- [73] M. Riesenhuber and T. Poggio, "Hierarchical models of object recognition in cortex," *Nat. Neurosci.*, vol. 2, no. 11, pp. 1019–1025, 1999.
- [74] C. Siagian and L. Itti, "Rapid biologically-inspired scene classification using features shared with visual attention," *Pattern Anal. Mach. Intell. IEEE Trans.*, vol. 29, no. 2, pp. 300–312, 2007.
- [75] D. G. Lowe, "Object recognition from local scale-invariant features," *Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on*, vol. 2, pp. 1150–1157 vol.2, 1999.
- [76] W. J. Christmas, J. Kittler, and M. Petrou, "Using Probabilistic Relaxation," vol. 17, no. 8, 1995.
- [77] R. C. Gonzalez and E. Richard, "Woods, digital image processing," ed Prentice Hall Press. ISBN 0-201-18075-8, 2002.

Abstract

Objective in this thesis is to develop a method for establishing an object recognition system based on the matching of image primitives. A primitive is extracted from image by the Scale Invariant Feature Transform(SIFT) descriptors. This method can be applied to a number of computer vision applications such as object recognition (in general) and image retrieval. The motivation for using descriptors as image primitives is that they can be invariantly to a group of affine transforms and stable under scaling and rotation. Each object in database modeled by a single frontal image. The recognition task is to determine the presence of object(s) of interest in scene images. Attributed Relational Graph(ARG) proposed for representation image primitives. Each node in ARG described by two properties: Unary and Binary measurements. The unary measurements use to describe any node in graph individually. In order to describe relation with two nodes binary measurements used. The probabilistic relaxation labeling applied for graph matching. Experimental result divided into two stage: Virtual Environment(VE) and real scenario. Since in the real environment is difficult to control control system parameters, Virtual Environment(VE) constructed. Then structural noise added in this VE for evaluating two algorithm based on probabilistic relaxation labeling: Ahmadyfard and Kostin matching. Later two algorithm in real scenario tested. The first algorithm was reliable than second algorithm. Against the second find the more match than first algorithm.

Finally A new hybrid algorithm based on two algorithm proposed and applied to a real scenario. The proposed algorithm performs from both reliability and recognition rates point to view.

Keywords: object recognition, unary measurement, binary measurement, graph matching, relaxation labeling



Shahrood University of Technology

Faculty of Computer Engineering

Localization and recognition of object in complex scenes

Thesis

Submitted in Partial Fulfillment of the Requirements for the Degree of
Master of Science (M.Sc)

Abbas Zohrevand

Supervisors

Dr. Ali A.Pouyan

Dr. Alireza Ahmadyfard

January 2014