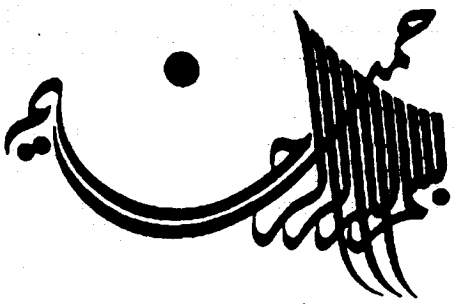
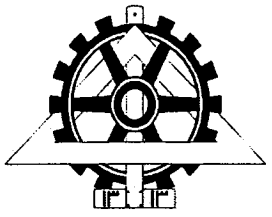
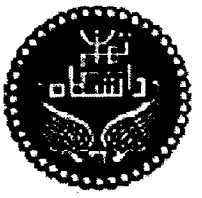


١٠٨٤٩٧٥

۸۷/۱/۱۰۵۵۳۸
۸۷/۱/۲۴





دانشگاه تهران
پردیس دانشکده‌های فنی
دانشکده مهندسی برق و کامپیوتر

یادگیری کنترل توجه در یادگیری مفاهیم در محیط های پیچیده

نگارش:
احسان نوروزی

استاد راهنمای اول: دکتر مجید نیلی احمدآبادی
استاد راهنمای دوم: دکتر بابک نجار اعرابی

پایان‌نامه برای دریافت درجهٔ کارشناسی ارشد
در
رشتهٔ مهندسی برق، گرایش کنترل

مهرماه ۱۳۸۷

۱۳۸۷/۱۱/۰۷

۱۰۸۴۹۵

دانشکده مهندسی برق و کامپیوتر

پردیس دانشکده های فنی

بسمه تعالی

گواهی دفاع از پایان نامه کارشناسی ارشد

هیأت داوران پایان نامه کارشناسی ارشد آقای/خانم **احسان نورزوی** در رشته مهندسی برق و کامپیوتر، گرایش : کنترل

را، با عنوان: **"یادگیری کنترل توجه در یادگیری مفاهیم در محیط های پیچیده"**

	به عدد	به حروف
نمره نهایی پایان نامه:	۱۷	هفده

در تاریخ ۱۳۸۷/۰۷/۱۳

و درجه بسیار خوب ارزیابی نمود.

مشخصات هیأت داوران	نام و نام خانوادگی	مرتبه دانشگاهی	دانشگاه یا موسسه	امضاء
۱-استاد راهنما	دکتر مجید نیلی احمدآبادی	دانشیار	تهران	
استاد راهنمای دوم (حسب مورد)	دکتر بابک نجاراعرابی	دانشیار	تهران	
۲-استاد مشاور	--	--	--	
۳ -استاد مدعو داخلی (یا استاد مشاور دوم)	دکتر مسعود اسدپور	استادیار	تهران	
۴-استاد مدعو خارجی	دکترمنصورجم زاد	دانشیار	صنعتی شریف	
۵-داور و نماینده کمیته تحصیلات تکمیلی دانشکده	دکتر منوچهر مرادی سبزوار	استادیار	تهران	

تذکر:این برگه پس از تکمیل توسط هیأت داوران در نخستین صفحه پایان نامه درج می گردد.

دانشگاه تهران

پردیس دانشکده‌های فنی

دانشکده مهندسی برق و کامپیوتر

پایان‌نامه برای دریافت درجهٔ کارشناسی ارشد در رشتهٔ مهندسی برق، گرایش کنترل

عنوان: یادگیری کنترل توجه در یادگیری مفاهیم در محیط های پیچیده

نگارش: احسان نوروزی

این پایان نامه در تاریخ ۱۳ مهرماه ۱۳۸۷ در مقابل هیات داوران دفاع گردید و مورد تصویب قرار گرفت.

معاون آموزشی و تحصیلات تکمیلی پردیس دانشکده‌های فنی: دکتر جواد فیض

رئیس دانشکدهٔ مهندسی برق و کامپیوتر: دکتر پرویز جبه‌دار مارالانی

معاون پژوهشی و تحصیلات تکمیلی دانشکدهٔ مهندسی برق و کامپیوتر: دکتر سعید

اساتید راهنما: دکتر مجید نیلی احمدآبادی

دکتر بابک نجار اعرابی

عضو هیات داوران: دکتر منصور جم‌زاد

عضو هیات داوران: دکتر مسعود اسدپور

عضو هیات داوران: دکتر منوچهر مرادی سبزوار

تعهد نامهٔ اصالت اثر

اینجانب احسان نوروزی تایید می‌کنم که مطالب مندرج در این پایان نامه حاصل کار پژوهشی این جانب بوده و به دستاوردهای پژوهشی دیگران که در این نوشته از آنها استفاده شده است مطابق مقررات ارجاع گردیده است. این پایان نامه قبلا برای احراز هیچ مدرک هم‌سطح یا بالاتر ارائه نشده است.

کلیهٔ حقوق مادی و معنوی این اثر متعلق به دانشکدهٔ فنی دانشگاه تهران می‌باشد.

نام و نام خانوادگی دانشجو:

امضای دانشجو:

تقدیم به خانواده‌ام

وتقدیم به استاد عزیزم آقای دکتر حسین روان‌بد

چکیده

عامل‌های هوشمند، چه طبیعی و چه مصنوعی، از طریق مجموعه حسگرهای خود از محیط بیرون خود کسب اطلاعات می‌کنند. برای تعامل موثر با محیط بایستی این عامل‌ها توانایی پردازش این اطلاعات در زمان واقعی را داشته باشند. اما به طور معمول حجم عظیم اطلاعات حسگری، محاسبات مربوط به یادگیری ماشین را غیرقابل ردیابی می‌کند. از طرف دیگر در موارد بسیار زیادی بخشی از اطلاعات ورودی گمراه کننده هستند که بایستی این موضوع تشخیص داده شود، به عنوان مثال هنگامی که قسمتی از جسمی که قرار است بازشناسی شود از دید عامل پوشیده شده باشد. یکی از راه‌های مقابله با چنین مسائلی یافتن استراتژیی برای کنترل توجه است. به این معنی که عامل فضای حسگری خود را، بر حسب درجهٔ اهمیتی که در وظیفهٔ جاری عامل ایفا می‌کند، تقسیم‌بندی کرده و این قسمتها را به نوبت و بر حسب درجهٔ اهمیت، پردازش کند. به این ترتیب محاسبات، عملی و قابل ردیابی می‌شود.

از آنجایی که پیاده‌سازی کنترل توجه به شدت وابسته به وظیفه است، در این پایان نامه بازشناخت چهره به عنوان وظیفه انتخاب شده است. در بازشناخت چهره لازم است بردارهایی با بعد زیاد مطالعه شوند. ابتدا تصاویر چهره، بلوک‌بندی شده و سپس یک ساختار یادگیری تقویتی ترتیب اهمیت بلوک‌ها رادر توانایی طبقه‌بندی به دست می‌آورد. حاصل، یک استراتژی

کنترل توجه بود که بر طبق آن می‌توان با نگاه کردن به کم‌ترین تعداد بلوک‌ها وظیفهٔ بازشناسی را به انجام رسانید.

کارکرد مهم کنترل توجه این است که عامل، قسمت‌های مهم‌تر فضای حسگری را مورد بررسی قرار دهد. این امر موجب می‌شود عامل به تدریج یاد بگیرد که قسمت‌هایی از این فضا را که به طور نظام‌مند غیر قابل رویت است در تصمیم‌گیری‌های خود دخالت ندهد. به این ترتیب می‌توان به عملکردهایی بالاتر دست یافت. برای نشان دادنِ چنین نتیجه‌ای از بازشناسی چهره‌های دچار پوشیدگی استفاده شد. به این صورت که در یک آزمایش به عامل تصاویر ناپوشیدهٔ افرادی آموزش داده می‌شود و عامل بایستی چهرهٔ این افراد را در حالتی که پوشیدگی‌هایی عارض شده بازشناسی کند. در آزمایش دیگر عامل با تصاویر پوشیده آموزش دیده و بایستی تصاویر بدون پوشش را بازشناسی کند. این آزمایش بر روی بانک تصاویر چهرهٔ AR صورت پذیرفت و نتایج آن با نتایج روش متداولِ کلیت‌گرایانه مورد مقایسه قرار گرفت. در آزمایش اول روش پیشنهاد شده بهبودی بیش از ۲۸ درصد و در آزمایش دوم بهبودی نزدیک ۵ درصد، در مقایسه با روش کلیت‌گرایانه حاصل شد.

فهرست مطالب

مقدمه ۱

۱- کنترل توجه ۵

۱-۱- کنترل توجه، تعریف ، لزوم و مثال ٦

۱-۲- نقشهٔ برجستگی ۸

۱-۳- انواع کنترل توجه ۱۰

۱-۳-۱ روش‌های به دست آوردنِ نقشه برجستگی ۱۰

۱-۳-۲ کنترل توجه پنهان و کنترل توجه آشکار ۱۰

۱-۳-۳ کنترل توجه پایین به بالا و بالا به پایین ۱۱

۱-۴- کنترل توجه به عنوان هرم پردازش داده‌ها ۱۳

۲- بازشناسی چهره ۱۵

۲-۱- مفاهیم بنیادی بازشناسی چهره ۱۵

۲-۲ – روش‌های زیرفضایی برای کاهش بعد ۱۷

۲-۲-۱ - تحلیلِ مؤلفه اصلی (PCA) ۱۹

۲-۲-۲- تحلیل جداپذیری خطی (LDA) ۲۱

۲-۳- بازشناسی ترتیبی چهره ۲۳

۲-۴- بازشناسی چهره در ترکیب با کنترل توجه ۲٤

۳- یادگیری تقویتی ۲۹

۳-۱- فرآیند تصمیم‌گیری مارکوف ۳۰

۳-۲- یادگیری تقویتی ۳۲

۳-۲-۱- مدل‌های رفتار بهینه ۳٤

۳-۲-۲- تابع ارزش ۳٦

۳-۲-۳- راه حل تابع ارزش برای حل فرآیند تصمیم‌گیری مارکوف ۳۷

۳-۲-۴- استراتژی‌های انتخاب عمل ۳۹

۴- ساختار پیشنهاد شده ٤۳

۴-۱- تشریح مساله ٤۳

۴-۲- انتخاب و تفکیک مجموعهٔ داده‌ها ٤٦

۴-۳- بلوک‌بندی تصویر و استخراج ویژگی ٤٦

۴-۴- طبقه‌بندی کنندهٔ ترتیبی ۴۹
۴-۵- یادگیری استراتژی کنترل توجه به صورتِ تقویتی ۵۳
۴-۶- ساختار کلی ۶۰
۵- پیاده‌سازی و نتایج ۶۳
۵-۱- معرفی بانک تصاویر چهرهٔ به کار گرفته شده ۶۴
۵-۱-۱- بانک تصاویر AT & T ۶۴
۵-۱-۲- بانک تصاویر AR ۶۵
۵-۲- اثر کنترل توجه بر کاهش حجم محاسبات ۶۶
۵-۲-۱- اجرای الگوریتم کلیت گرایانه ۶۷
۵-۲-۲- اجرای الگوریتم ترتیبی همراه با کنترل توجه ۷۰
۵-۳- بهبود عملکرد در حالتی که چهره‌ها پوشیدگی دارند ۷۶
۵-۳-۱- پوشیدگی‌های مصنوعی ۷۶
۵-۳-۲- پوشیدگی‌های واقعی ۷۸
۶- جمع‌بندی، تحلیل، نتیجه‌گیری و پیشنهادهایی برای ادامهٔ کار ۸۷
مراجع و مآخذ ۹۱
پیوست ۱- مقالات استخراج شده از پایان نامه ۹۴
پیوست ۲- فهرست واژگان انگلیسی ۹۶

فهرست جدول‌ها

جدول ۵-۱- پارامترهای مورد استفاده در شبیه سازی ۶۷
جدول ۵-۲- نتایج حاصل شده از شبیه‌سازی الگوریتم کلیت گرایانه (LDA) ۶۹
جدول ۵-۳- عملکرد نهاییِ بازشناسیِ ترتیبی با استفاده از مقدارهای متفاوت واریانسِ حفظ شده ۷۰
جدول ۵-۴- پارامترهای الگوریتم یادگیری تقویتی ۷۱
جدول ۵-۵- مقایسهٔ روش‌های کلیت گرایانه و انتخابی ۷۵
جدول ۵-۶- بهترین درصد طبقه‌بندیِ صحیح در حالات مختلف پوشیدگی چهره ۷۸
جدول ۵-۷- نتایج الگوریتم کلیت‌گرایانه در حالتی که سیستم با تصاویر ناپوشیده آموزش داده شده است. ۸۰
جدول ۵-۸- درصد طبقه‌بندی صحیح در مقابل نحوهٔ بلوک‌بندی ۸۱
جدول ۵-۹- درصد طبقه‌بندی صحیح در مقابل درصد همپوشانی بلوک‌ها ۸۲
جدول ۵-۱۰- نتایج الگوریتم ترتیبی در دو حالت تصادفی و انتخابی، آموزش با تصاویر ناپوشیده ۸۳
جدول ۵-۱۱- نتایج الگوریتم کلیت‌گرایانه در حالتی که سیستم با تصاویر پوشیده آموزش داده شده است ۸۳
جدول ۵-۱۲- نتایج الگوریتم ترتیبی در دو حالت تصادفی و انتخابی، آموزش با تصاویر پوشیده ۸۴

فهرست شکل‌ها

شکل ۱-۱- تصویر شبکیهٔ چشم ۸

شکل ۱-۲- روندنمای محاسبهٔ نقشهٔ برجستگی (مطابق پیشنهادِ [Itti et al., 1998]) ۸

شکل ۱-۳- نمونهٔ به دست آوردن یک نقشهٔ برجستگی ([Itti et al., 1998]) ۹

شکل ۱-۴- آزمایش یاربوس برای نشان دادن تغییر الگوی چرخش چشم افراد مختلف بر روی تصویر مشابه در شرایطی که میخواهند سوالات گوناگونی را پاسخ دهند [Yarbus,1967]..... ۱۲

شکل ۱-۵- هرم پردازش داده‌های ورودی با دخالت کنترل توجهِ پایین به بالا و بالا به پایین ۱۴

شکل ۲-۱- شماتیک منیفولد چهره‌ها در فضای کلّ تصاویر ۱۹

شکل ۲-۲- نحوهٔ آشنا شدن تدریجیِ افراد با چهره‌های ناآشنا، (الف) افزایش نرخ بازشناسی، (ب) کاهش تعداد نگاه‌های لازم برای بازشناسی [Heisz and Shore, 2008] ۲۶

شکل ۳-۱- شماتیک تعامل عاملِ یادگیرندهٔ تقویتی با محیط ۳۲

شکل ۴-۱- نمودار بلوکی تصمیم گیری ترتیبی..... ۴۵

شکل ۴-۲- نحوهٔ بلوک‌بندی تصاویر. (الف) همپوشانی بین بلوک‌های مجاور، (ب) بلوک‌بندی کل تصویر..... ۴۷

شکل ۴-۳- نحوهٔ محاسبهٔ فاصلهٔ یک نمونه از یک کلاس با استفاده از k -نزدیک‌ترین همسایه ۴۹

شکل ۴-۴- نحوهٔ رشد المان‌های بردار باور بر اثر افزایش تعداد بلوک‌های پردازش شده ۵۳

شکل ۴-۵- نحوهٔ عملکرد تابع پاداش در بازشناسی یک تصویر..... ۵۶

شکل ۴-۶- نمودارِ نحوهٔ تخصیص اعتبارِ تخفیف یافته (R_f پاداش نهایی و $0 < \alpha <= 1$ ضریب تخفیف) ۵۸

شکل ۴-۷- روندنمای یادگیری استراتژی کنترل توجه..... ۶۱

شکل ۵-۱- نمونهٔ یکی از تصاویر موجود در بانک تصاویر AT & T ۶۵

شکل ۵-۲- نمونهٔ یکی از تصاویر موجود در بانک تصاویر AR ۶۶

شکل ۵-۳- نمودار درصد طبقه بندی صحیح بر حسب تعداد فیشرچهره‌ها ۶۸

شکل ۵-۴- نمودار قابلیت اعتماد در تصمیم گیری بر حسب تعداد فیشرچهره‌ها..... ۶۹

شکل ۵-۵- افزایش تدریجی درصد طبقه‌بندی صحیح و کاهش تدریجی تعداد بلوک‌های لازم برای پردازش هنگامی که تنبیه حدس غلط برابر یک باشد ۷۱

شکل ۵-۶- افزایش تدریجی درصد طبقه‌بندی صحیح و کاهش تدریجی تعداد بلوک‌های لازم برای پردازش هنگامی که تنبیه حدس غلط برابر پنج باشد ۷۲

شکل ۵-۷- نقشهٔ برجستگیِ به دست آمده از الگوریتم کنترل توجه بر روی چهرهٔ انسان..... ۷۳

شکل ۵-۸- مقایسهٔ سرعت افزایش درصد طبقه‌بندی صحیح در دوحالتِ تصادفی و انتخابی (تنبیه کم)..... ۷۴

شکل ۵-۹- مقایسهٔ سرعت افزایش درصد طبقه‌بندی صحیح در دوحالتِ تصادفی و انتخابی(تنبیه زیاد) ۷٤
شکل ۵-۱۰- انواع پوشیدگی‌های به کار رفته در شبیه‌سازی‌ها..... ۷٦
شکل ۵-۱۱- نمودار درصد طبقه بندی صحیح بر حسب تعداد فیشرچهره‌ها، تصاویر نیم پوشیده هستند ۷۷
شکل ۵-۱۲- نمودار درصد طبقه بندی صحیح بر حسب تعداد فیشرچهره‌ها، تصاویر ربع پوشیده هستند ۷۷
شکل ۵-۱۳- عملکرد بازشناسیِ ترتیبی، بر حسب تعداد بلوک‌های درگیر، تصاویر نیمه پوشیده هستند..... ۷۸
شکل ۵-۱۴- عملکرد بازشناسیِ ترتیبی، بر حسب تعداد بلوک‌های درگیر، تصاویر ربع پوشیده هستند..... ۷۹
شکل ۵-۱۵- نمودار درصد طبقه بندی صحیح بر حسب تعداد فیشرچهره‌ها، آموزش با تصاویر ناپوشیده ۸۰
شکل ۵-۱۶- مقایسهٔ سرعت افزایش درصد طبقه‌بندی صحیح در دوحالتِ تصادفی و انتخابی، آموزش با تصاویر ناپوشیده..... ۸۲
شکل ۵-۱۷- نمودار درصد طبقه بندی صحیح بر حسب تعداد فیشرچهره‌ها، آموزش با تصاویر پوشیده ۸۳
شکل ۵-۱۸- مقایسهٔ سرعت افزایش درصد طبقه‌بندی صحیح در دوحالتِ تصادفی و انتخابی، آموزش با تصاویر پوشیده ۸٤

مقدمه

یک عامل برای انجام دادن وظیفه‌ای که به وی محول شده است بایستی با محیط خود تعامل برقرار کند. لازمهٔ عمل کردن موفق در هر محیط، درک و به دنبال آن شناسایی محیط است. این ادراک عموما از طریق حسگرهایی که ادراکات حسی عامل را به صورت عددی و کمّی تبدیل می‌کنند حاصل می‌شوند.

محیط‌های پیچیده محیط‌هایی هستند که برای درک و توصیفشان نیاز به استفاده از تعداد زیادی حسگر است. با توجه به این تعریف به راحتی دریافته می‌شود که به جز بعضی از محیط‌های مصنوعی و انتزاعی، اغلب محیط‌های طبیعی و حتی مصنوعی را بایستی به چشم محیط‌های پیچیده نگریست.

پس از ادراک محیط، عامل بایستی با پردازش داده‌های خام حسی آنها را به صورت مفاهیم درآورده و این مفاهیم، پایه‌ای برای عملکرد عامل می‌شوند. اما یادگیری مفهوم در محیط‌های پیچیده به راحتی صورت نگرفته و عامل‌ها معمولا با مشکل اساسی کاستی در توان پردازشی در مقایسه با حجم داده‌ها مواجه هستند.

اختلاف حجم داده‌های حسگری و توان پردازشی، در مورد داده‌هایی که پهنای باند بسیار زیادی دارند، مثلا داده‌های تصویری، خود را با شدت بیشتری نشان می‌دهد. وظیفه‌ای

که در این پژوهش مورد بررسی قرار گرفته، یعنی بازشناسی چهره، از دستهٔ وظایفی است که با اطلاعات تصویری درگیر است. الگوریتم‌های بسیار زیادی در حوزهٔ بازشناسی چهره پیشنهاد شده است که به موفقیت‌هایی نیز رسیده‌اند، اما ویژگی مشترک همهٔ آنها پیچیدگی محاسباتی زیاد است.

پیچیدگی محاسباتی بازشناسی چهره باعث می‌شود وظایفی که در آنها نیاز به واکنش سریع وجود دارد دچار اخلال شوند. از طرف دیگر اگر حجم بانک اطلاعاتی تصاویر زیاد باشد یا تعداد چهره‌هایی که بایستی بازشناسی شوند زیاد باشد مدت زمان لازم برای بازیابی اطلاعات بسیار زیاد می‌شود.

برای کاستن از هزینهٔ محاسباتی یاد شده یک راه، حذف کردن بخش‌هایی از تصویر است که در وظیفهٔ بازشناسی نقشی اندک دارند. به این ترتیب می‌توان توان پردازشی عامل را بر روی سایر نواحی تصویر متمرکز کرد. پیش‌فرض این پیشنهاد این است که در چهرهٔ انسان‌ها مناطقی وجود دارند که در تفکیک چهره‌ها اطلاعات بیشتری فراهم می‌کنند. با توجه به این پیش فرض در این پژوهش تلاش می‌شود با توسعهٔ روشی مبتنی بر یادگیری، عامل به تدریج این مناطق را تشخیص داده و در عمل بازشناسیِ چهره، توجه خود را زودتر به این مناطق معطوف کند. این روش الهام گرفته از راه حلی است تحت عنوان «کنترل توجه» که موجودات زنده برای بازشناسی بصری اشیا به کار می‌برند، به این صورت که به طور ارادی توجه خود را به نواحی مهم‌تر یک صحنه انتقال می‌دهند.

رویکردی که در این پژوهش مورد استفاده قرار می‌گیرد عبارت است از بلوک‌بندیِ تصاویر چهره و مرتب کردن بلوک‌ها به ترتیب اهمیتی که در طبقه‌بندی دارند. این مرتب سازی به تدریج در طول زندگی عامل و در یک فرایند یادگیری تقویتی صورت می‌گیرد. به این ترتیب در پایان نقشه‌ای به دست می‌آید که اهمیت نواحی مختلف بر روی آن مشخص شده است. با داشتن این نقشه استراتژیی برای انتقال توجه عامل به دست می‌آید. به این صورت که می‌توان با استفاده از نقشه بلوک‌ها را به صورت ترتیبی و با توجه به اهمیتشان

انتخاب کرده، و عمل بازشناسی را بر روی تک بلوکِ انتخاب شده صورت داده و نتیجهٔ حاصل را با نتایج حاصل از بلوک‌های پیشین تجمیع کرده و پس از رسیدن به سطحی از کفایت، که می‌تواند پردازش تعداد مشخصی بلوک یا رسیدن به سطحی مناسب از اطمینان باشد، تصمیم‌گیری نهایی را با استفاده از اطلاعات انباشته شده صورت دهد.

از آنجایی که هدف، به دست آوردن استراتژیی برای انتقال توجه است که متناسب با مغز عامل باشد از روش بدون سرپرستیِ یادگیری تقویتی استفاده می کنیم. به این ترتیب می‌توان اطمینان داشت که هیچ گونه پیش فرضی که از اهمیت بلوک‌ها در ذهن طراح انسانی هست در تصمیم گیری عامل دخالت داده نمی‌شود.

انتظار می‌رود که این کار موجب کاهش حجم محاسبات و در نتیجه افزایش سرعت بازشناسی شود. از طرف دیگر انتخاب کردن بلوک‌های مهم منجر به حذف شدن بلوک‌های گمراه کننده می‌شود و این امر می‌تواند کیفیت عملکرد را از نظر درصد طبقه بندی صحیح بالا ببرد. برای تایید این فرضیه، عمل بازشناسی را بر روی تصاویر چهره‌ای که دارای پوشیدگی‌هایی هستند انجام می‌دهیم.

ساختار پایان نامه به این صورت است، در فصل اول مفاهیم بنیادین کنترل توجه مورد بررسی قرار گرفته و آشنایی مختصری با ادبیات آن صورت می‌گیرد. همچنین توجیه‌هایی ارائه می‌شود که کنترل توجه وابسته به وظیفه است و به صورت مجرد از یک وظیفه قابل بررسی نیست.

فصل دوم برای آشنایی با ادبیات بازشناسی چهره و تکنیک‌های مورد استفاده در آن آورده شده است. همچنین دو دیدگاه کلیت گرایانه و ترتیبی با هم مقایسه شده و ترکیب کنترل توجه با بازشناسی ترتیبی چهره معرفی می‌شود.

فصل سوم به موضوع یادگیری تقویتی می‌پردازد. یادگیری تقویتی، یک روش یادگیری بدون سرپرستی است که برای به دست آوردن درجهٔ اهمیت بلوک‌های تصویر مورد استفاده قرار گرفته است.

در فصل چهارم تمامی مفاهیم یاد شده در سه فصل قبل در یک ساختار واحد ترکیب شده و ساختار پیشنهاد شده به صورت یک کلیت معرفی می‌شود. فصل پنجم نیز نتایج ناشی از اجرای ساختار پیشنهاد شده را بر روی دو پایگاه تصویریِ معروف ارائه داده می‌شود.

۱- کنترل توجه

پیشرفت‌های پدید آمده در هوش مصنوعی و رُوباتیک منجر به ساخته شدن عامل‌هایی هوشمند شده که می‌توانند تعاملات گسترده‌ای با محیط اطراف خود برقرار کرده و از این طریق وظایفی را به انجام برسانند. عامل‌های هوشمند، خواه طبیعی و خواه مصنوعی، دارای محدودیت‌هایی ساختاری هم در سطح حسگری و هم در سطح پردازشی هستند که درک آنها از محیط را با نایقینی-هایی روبرو می‌کنند. یکی از راه‌هایی که برای کاستن از این نایقینی‌ها به کار می‌رود زیاد کردن حسگرهایی است که عامل توسط آنها محیط را درک می‌کند. این امر منجر به تولید حجم بسیار زیادی دادهٔ خام می‌شود که عامل، برای رسیدن به دانشی سطح بالاتر از محیط، بایستی این حجم زیاد را پردازش کند.

اما پردازش این حجم از داده‌ها خالی از اشکال نیست. سه مشکلِ اساسی که در این میان ممکن است رخ دهد این است که:

- عامل از نظر پردازنده و حافظه به نحوی دچار کاستی باشد که به طور ساختاری توانایی پردازش داده‌هایی که به قالب بردارهایی با طول زیاد هستند را نداشته باشد.

- محاسباتی که در آنها همهٔ داده‌ها مد نظر قرار می‌گیرند حجیم و زمان‌بر بوده و امکان عملکرد عامل در زمان واقعی را می‌گیرد؛ در حالی که در آن وظیفهٔ خاص ممکن است عملکرد بلادرنگ[1] مورد نیاز باشد.
- قسمت‌هایی از فضای حسی، اطلاعاتی گمراه کننده را در خود دارند که بهتر است حذف شوند.

برای حل این مشکلات لازم است سازوکارهایی طراحی شود که داده‌های خام را به نحوی فیلتر کند که عامل در هر لحظه نیاز به پردازش حجم اندکی داده داشته باشد و این حجم اندک حاویِ بیشترین محتوای اطلاعاتیِ صحیحِ کل داده‌ها باشد. فلسفهٔ این پیشنهاد در این است که معمولا عامل برای انجام وظایف خود نیاز به پردازش همزمان تمام داده‌ها ندارد و در هر لحظه تنها قسمت کوچکی از داده‌ها در تصمیم‌گیریِ وی ارزش حیاتی دارند. بدین ترتیب مشکل اطلاعات گمراه کننده نیز حل می‌شود؛ به این صورت که با متمرکز شدن بر نواحی‌ای که اطلاعات متقن‌تر هستند خود به خود نواحی شامل اطلاعات گمراه کننده نادیده گرفته می‌شوند. چنین ساز و کارهایی تحت عنوانِ کلّیِ «کنترل توجه» مورد بررسی قرار می‌گیرد.

۱-۱- کنترل توجه، تعریف، لزوم و مثال

بر مبنای تعریفی که ویلیام جیمز[2]، روانشناس آمریکایی، ارائه کرده است [کنترلِ] توجه عبارت است از سازوکاری که در طی آن قسمت کوچکی از ویژگی[3]های ورودی‌های حسگری پردازش می‌شوند و سایر جنبه[4]های آنها نادیده گرفته می‌شوند [James,1890].

از آنجایی که در جریان کنترل توجه، مقداری از داده‌ها حذف می‌شود مطمئنا قابلیت اعتماد تصمیم نسبت به حالتی که همهٔ داده‌ها مورد پردازش قرار گیرند کمتر می‌شود. بنابراین مهم‌ترین

---

[1] Real time
[2] William James
[3] Feature
[4] Aspect

مزیتی که کنترل توجه [بصری] ایجاد می‌کند هدایت کردن سریع نگاه به سوی شیء مورد علاقه در حوزهٔ بینایی است [Itti & Koch, 2001].

قسمت بسیار زیادی از دستاوردها و پیشرفت‌هایی که در روش‌های کنترل توجه صورت گرفته ریشه در مطالعات سازوکارهای مشابه در سیستم‌های زیستی دارد. به طور خاص، مهم‌ترین مثالی که در این زمینه مورد بررسی قرار گرفته است مسالهٔ بینایی و بازشناخت بصری اشیاء است. اطلاعات تصویری، پهنای باند زیادی داشته و شامل اطلاعات زیادی است و حل کردن مسایل نام‌برده دارای پیچیدگیِ NP-complete هستند [Tsotsos et al.,1995]. در نتیجهٔ این پیچیدگی و حجم زیاد اطلاعات، مغز از نظر ساختاری توانایی حل چنین مسایلی را در زمانی مناسب ندارد. اما مشاهده می‌شود که افراد به راحتی صحنه‌هایی را که می‌بینند تحلیل کرده، محیط خود را درک کرده، اشیا را مورد بازشناسی قرار داده و در زمانی کوتاه واکنش نشان می‌دهند. این امر به این علت امکان‌پذیر شده است که اطلاعات در مراحل و سطوح مختلف فیلتر شده، اطلاعاتی که در آن لحظهٔ خاص اهمیت کمی داشته حذف شده و تنها قسمت بسیار کوچکی از آنها برای پردازش جزیی‌تر به بخش‌های شناختی[5] مغز می‌رسد.

به عنوان مثال ساختار شبکیهٔ چشم به نحویست که در مرکز آن ناحیهٔ نسبتا کوچکی وجود دارد که نحوهٔ چینش سلول‌ها و تراکم آنها موجب شده است که از قدرت تفکیک و حساسیت زیادی برخوردار باشد. در حالی که نواحی اطراف شبکیه دارای تراکم کمتری بوده و تصاویری با قدرت تفکیک کمتری برای مغز فراهم می‌آورند [Koch & Ullman, 1985]. شکل (۱-۱) تصویری واقعی از شبکیه را نشان می‌دهد که ناحیهٔ یاد شده به صورت یک لکهٔ تیره (به علت وجود رگ‌های خون رسانِ بیشتر) بر روی آن مشخص است.

چنین ساختاری باعث می‌شود که در همان سطح حسگری قسمت زیادی از اطلاعات تصویری که در بیرون از مرکز دید قرار می‌گیرد حذف شود و تنها قسمتِ مرکزیِ تصویر با دقت زیاد دیده می‌شود. حال بر حسب اینکه عامل در حال انجام چه وظیفه‌ایست با چرخاندن چشمش می‌تواند

[5] Cognitive