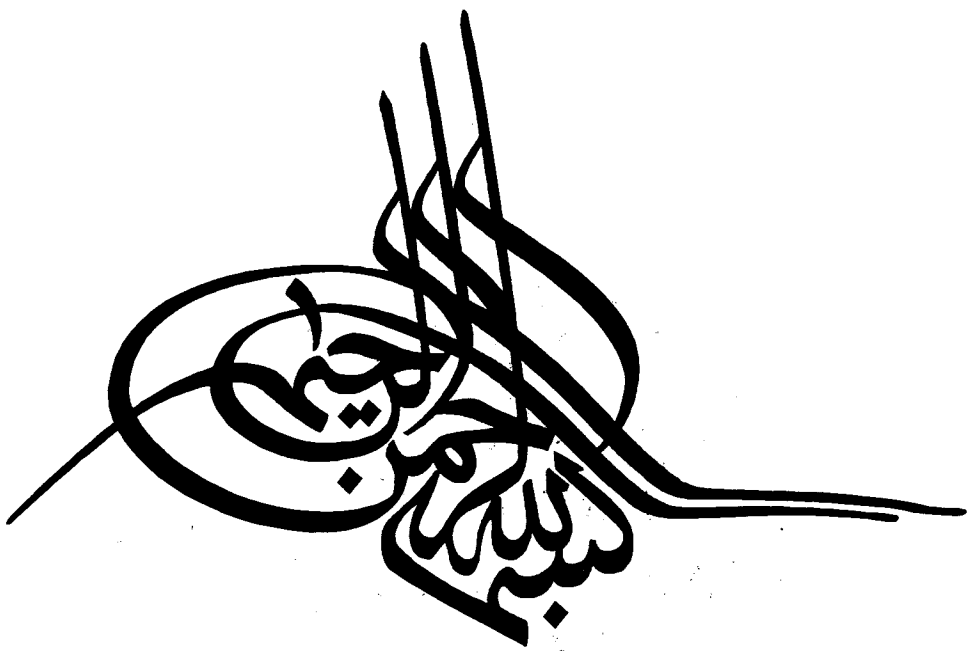

٩٤١١٣

دانشگاه رازی
دانشکده فنی مهندسی
گروه مهندسی برق

پایان‌نامه جهت اخذ درجه کارشناسی ارشد رشته مهندسی برق
گرایش الکترونیک

مدلسازی درک تصویر در مغز انسان: توصیف‌های مبتنی بر بخش جهت اشکال دیداری

استاد راهنما:
دکتر قاسم عازمی

۱۳۸۷/۱۱/۱۰

نگارش:
محمد اکبری فرمانبر

آذر ماه ۱۳۸۶

۹۲۱۱۳

کلیه حقوق مادی مترتب بر نتایج مطالعات، ابتکارات و نـــوآوری های ناشـــی از تحقیق موضوع این پایان‌نامـــه متعلق به دانشگاه رازی است.

دانشگاه رازی
دانشکده فنی مهندسی
گروه مهندسی برق

پایان‌نامه جهت اخذ درجه کارشناسی ارشد رشته‌ی برق گرایش الکترونیک

محمد اکبری فرمانبر

تحت عنوان

مدلسازی درک تصویر در مغز انسان: توصیف‌های مبتنی بر بخش جهت اشکال دیداری

در تاریخ ۸۶/۹/۱۳ توسط هیأت داوران زیر بررسی و با درجه عالی به تصویب نهایی رسید.

۱- استاد راهنمای پایان نامه دکتر **قاسم عازمی** با مرتبه‌ی علمی **استادیار** امضاء

۲- استاد داور داخل گروه دکتر **محسن حیاتی** با مرتبه‌ی علمی **استادیار** امضاء

۳- استاد داور خارج از گروه دکتر **امیر رجب‌زاده** با مرتبه‌ی علمی **استادیار** امضاء

سپاسگزاری

ستایش خدای مهربان، کردگار روزی رسان، یکتا در نام و نشان، خداوندی که ناجُسته یابد و نادریافته شناسد. خدای را شاکرم که همواره در مراحل مختلف این تحقیق، بنده حقیر را مورد عنایات خود قرار داده است.

از دست و زبان که بر آید کز عهده شکرش بدر آید

جا دارد از استاد راهنمای گرانقدرم جناب آقای دکتر عازمی که همواره با نظرات، پیشنهادات و همراهی‌هایشان، بنده را مورد لطف قرار داده و در جهت پیشبرد کارها کمک زیادی نموده‌اند، یاد کنم و بدینوسیله مراتب تشکر و سپاسگزاری خود را از ایشان به عمل آورم.

سایر اساتیدی که شایسته است از آنها قدردانی صورت گیرد، دکتر بابک نجار اعرابی و دکتر فریبا بهرامی از دانشگاه تهران، دکتر احسان‌الله کبیر و مهندس احسان کامرانی از دانشگاه تربیت مدرس، دکتر عباسیان از پژوهشگاه دانش‌های بنیادی (IPM) می‌باشند که هرکدام با راهنمایی‌های خود و ارجاع به مقالات به نحوی در مشخص ساختن مسیر و پیشرفت تحقیق کمک حال بنده بوده‌اند.

در پایان لازم است از سایر اساتید و دوستان بزرگوارم، به ویژه آقای علیرضا نعیمی، و نیز از همه اعضای خانواده‌ام بخصوص پدر و مادر عزیزم، که همواره به صورت‌های مختلف معنوی و مادی، از کمکهای ایشان بهره‌مند بوده‌ام تشکر کنم.

۱۳۸۷ /۱/ ۱۸۱

چکیده:

در این پایان‌نامه سعی داریم یک مدل محاسباتی مبتنی بر بخش ازچگونگی درک تصاویر در مغز انسان ارائه دهیم. اهمیت اصلی چنین مدلی آن است که تاکنون چنین مدل محاسباتی یکپارچه‌ای از درک تصویر آدمی ارائه نشده است. روش بخش‌بندی ارائه شده در این مدل مستقل از شکل تصاویر می‌باشد، و این مدل در گستره وسیعی از اشکال به درستی عمل می‌کند.

بینایی انسان شکل اشیاء را بر اساس بخش‌های تشکیل دهنده آن و ارتباط فضایی بین این بخش‌ها درک می‌کند. طبق شواهد تجربی، بخش‌ها بطور سریع و بصورت اولیه در پردازش بصری محاسبه می‌شوند. قاعده مینیما بیان می‌کند که سیستم بینایی انسان از نقاط با مینیمم انحنای منفی (NCM) برای تعریف کرانه‌های بین بخش‌ها بهره می‌برد. سیستم بینایی انسان بیشترین حساسیت خود را نسبت به لبه‌های تصویر (ویژگی‌های خارجی) نشان می‌دهد. بر این اساس در این مدل به کمک ویژگی‌های خارجی، نقاط NCM استخراج می‌شود و با استفاده از این نقاط، محاسبه کرانه بین بخش‌ها (برای محاسبه part-cutها) صورت می‌گیرد. قاعده short-cut برپایه آزمایشات روان-تنی، به اهمیت انتخاب کوتاهترین فواصل بین نقاط NCM جهت ایجاد بهترین part-cutها می‌پردازد. در اینجا با معرفی مفاهیم Neck ، برجستگی و تحدب، به انتخاب بهینه part-cutها بر اساس قاعده short-cut اشاره شده، و بر این اساس ضریب مجاورت تعریف شده است. ضمناً با معرفی مفاهیم Limb و شفافیت به اهمیت امتداد خوب در ساختن part-cutها اشاره شده، و بر این اساس ضریب همواری و کرویت تعریف شده است. ضمناً شرط تقارن محلی نیز در ضرایب کرویت و همواری تعبیه‌شده است. در نهایت به کمک ضرایب همواری، کرویت و مجاورت، بخش‌بندی نهایی صورت می‌گیرد که این بخش‌بندی نسبت به تغییر مقیاس و چرخش شکل مقاوم بوده و در مقابل تغییرات کلی نیز پایدار می‌باشد.

در مرحله آخر با تشخیص و تعبیر الگو روی این بخش‌های پایه‌ای، درک نهایی تصویر صورت خواهد گرفت. در این مدل، استفاده از شبکه عصبی مصنوعی کوهونن با روش آموزش بدون ناظرش که منطبق بر خصوصیات درکی سیستم بینایی می‌باشد جهت دسته‌بندی الگوها، پیشنهاد شده است، و جهت تعبیر و درک نهایی نیز استفاده از یک سیستم خبره بجای روش‌های تعبیر بر اساس منطق‌ـگزاره‌ها و یا شبکه‌های معنایی، پیشنهاد شده است زیرا پایگاه دانش، موتور استنتاج و روش‌های حل تناقض در این سیستم‌ها، خیلی مشابه سیستم درکی انسان می‌باشد، در نتیجه انتظار می‌رود نتایج بهتری حاصل گردد.

کلید واژگان:

بخش‌بندی محاسباتی، درک تصویر، سیستم بینایی انسان، قاعده مینیما، مینیمم‌های منفی انحنا.

فهرست مطالب

عنوان **صفحه**

فهرست شکل‌ها ..... **ث**

کوتاه نوشت‌ها ..... **خ**

پیشگفتار ..... **د**

فصل اول : مقدمه **۲**

۱-۱- انگیزه و ضرورت ارائه چنین مدلی از درک تصویر ۲

۱-۲- اهمیت توصیف اشکال مبتنی بر بخش‌هایشان ۴

۱-۳- تشریح چندی از رخدادهای دیداری به کمک مفهوم بخش ۶

۱-۳-۱- پدیده شناسی ۶

۱-۳-۲- تشابه ۶

۱-۳-۳- تخصیص شکل- زمینه ۷

۱-۳-۴- شفافیت ۸

۱-۳-۵- تقارن ۸

۱-۳-۶- دفعتاً- پدیدارشدن ۹

۱-۳-۷- حافظه به یادآوری ۹

۱-۳-۸- تمرکز بصری ۹

۱-۴- ساختار و اهداف انجام تحقیق ۹

۱-۵- پیشینه‌ی موضوع ۱۰

۱-۶- مقالات ارائه شده از نتایج این پایان‌نامه ۱۲

فصل دوم : مدل درک تصویر مبتنی بر سیستم بینایی انسان **۱۴**

۲-۱- مقدمه ۱۴

۲-۲- مدل کلی پردازش و درک تصویر بر پایه بخش‌بندی ۱۴

۲-۳- مدل دقیق درک تصویر با الهام از سیستم بینایی انسان ۱۶

۲-۴- جمع بندی ۱۸

فصل سوم : نمایش و توصیف تصویر ۲۱

۳-۱- مقدمه ۲۱

۳-۲- استخراج خصوصیات و نمایش تصویر ۲۱

۳-۳- شناسایی لبه‌های تصویر ۲۲

۳-۳-۱- یافتن لبه ۲۲

۳-۳-۲- اتصال ناپیوستگی‌ها و قطعی‌ها ۲۵

۳-۳-۳- نازک‌سازی لبه‌ها ۲۶

۳-۴- برازش منحنی به کانتور تصویر ۲۷

۳-۴-۱- برازش منحنی به روش cubic B-spline ۲۸

۳-۵- انحنای منحنی و انحنای سطح ۳۰

۳-۵-۱- محاسبه انحنای منحنی ۳۰

۳-۵-۲- محاسبه انحنای سطح ۳۰

۳-۶- جمع‌بندی ۳۱

فصل چهارم : مفاهیم بنیادی بخش‌بندی تصویر بر اساس نقاط با حداقل انحنای منفی ۳۳

۴-۱- مقدمه ۳۳

۴-۲- تعریف کرانه‌های بین بخش‌ها ۳۳

۴-۲-۱- ملاحظات هندسی ۳۳

۴-۲-۲- شواهد روان- تنی ۳۵

۴-۳- استخراج نقاط NCM ۳۶

۴-۴- معرفی Limbها و Neckها ۴۳

۴-۵- برجستگی بخش ۴۵

۴-۶- تئوری برجستگی بخش برای سیلوئت‌ها ۵۵

۴-۷- قاعده Short-Cut ۵۶

۴-۸- تحدب ۵۷

۴-۹- شفافیت ۵۸

۴-۱۰- پردازش پیشـمتوجه ۶۵

۴-۱۱- جمع‌بندی ۶۹

فصل پنجم : بخش‌بندی نهایی تصویر منطبق بر الگوهای درکی سیستم بینایی انسان ۷۱

۵-۱- مقدمه ۷۱

۵-۲- از کرانه‌های بخش تا part-cutها ۷۱

۵-۲-۱- طول cutها ۷۳
۵-۲-۲- شدت کرانه ۷۶
۵-۲-۳- امتداد خوب ۷۹
۵-۲-۴- تقارن محلی ۸۱
۵-۳- بهینه‌سازی انتخاب Part-cutها ۸۶
۵-۳-۱- ضریب کرویت ۸۶
۵-۳-۲- ضریب همواری ۸۸
۵-۳-۳- ضریب مجاورت ۸۸
۵-۴- تقطیع نهایی تصویر به بخش‌هایش ۸۹
۵-۴-۱- انتخاب part-cutهای مناسب با اعمال مقادیر آستانه‌ای ۸۹
۵-۵- جمع‌بندی ۹۴

فصل ششم : شناسایی و تعبیر الگو ۹۶

۶-۱- مقدمه ۹۶
۶-۲- تحلیل تصویر ۹۶
۶-۲-۱- الگوها و دسته‌های الگو ۹۷
۶-۳- روش‌های کلی تشخیص ۹۷
۶-۳-۱- دسته‌بندی‌کننده‌های شبکه‌ی عصبی ۹۸
۶-۳-۲- دسته‌بندی و تشخیص الگو در مغز انسان ۹۹
۶-۴- درک تصویر ۱۰۱
۶-۴-۱- انواع دانش ۱۰۲
۶-۴-۲- استنتاج به کمک حساب گزاره‌ای (استنتاج با اثبات قضیه) ۱۰۳
۶-۴-۳- استنتاج به کمک شبکه‌های معنایی ۱۰۳
۶-۴-۴- استنتاج به کمک سیستم‌های خبره ۱۰۳
۶-۵- جمع‌بندی ۱۰۶

فصل هفتم: جمع‌بندی و پیشنهادات ۱۰۸

۷-۱- خلاصه پایان‌نامه و جمع‌بندی ۱۰۸
۷-۲- مرور نتایج حاصل از این تحقیق ۱۱۰
۷-۳- پیشنهادات ۱۱۱

فهرست منابع و مراجع ۱۱۳

فهرست شکل‌ها

شکل(۱-۱):اهمیت پدیده شناسی در توصیف‌های مبتنی بر بخش اشکال دیداری..... ۶

شکل(۱-۲):اهمیت تشابه در توصیف‌های مبتنی بر بخش اشکال دیداری..... ۷

شکل(۱-۳):اهمیت تخصیص شکل-زمینه در توصیف‌های مبتنی بر بخش اشکال دیداری..... ۷

شکل(۱-۴):اهمیت شفافیت در توصیف‌های مبتنی بر بخش اشکال دیداری..... ۸

شکل(۱-۵):اهمیت تقارن در توصیف‌های مبتنی بر بخش اشکال دیداری..... ۸

شکل(۱-۶):اهمیت دفعتاً پدیدار شدن در توصیف‌های مبتنی بر بخش اشکال دیداری..... ۹

شکل(۲-۱):مراحل کلی پردازش و درک تصویر..... ۱۵

شکل(۲-۲): روندنمای بخش‌های مختلف مدل درک تصویر مبتنی بر "بخش"..... ۱۷

شکل(۳-۱): نمای سه‌بعدی و مقطع عرضی تابع LoG [۴۳]..... ۲۳

شکل(۳-۲): یک پنجره ۵×۵ جهت پیاده‌سازی تابع LoG..... ۲۴

شکل(۳-۳): پنجره ۳×۳ جهت پیاده‌سازی تابع لاپلاسین..... ۲۴

شکل(۳-۴): الگوریتم شرمن جهت نازک‌سازی لبه‌های تصویر..... ۲۶

شکل(۳-۵): یک نمایش بصری از نحوه تقریب یک منحنی پیچیده با چندجمله‌ای‌های تکه‌ای..... ۲۸

شکل(۴-۱): زیر مجموعه اشکال به عنوان بخش‌ها..... ۳۴

شکل(۴-۲):قاعده چین‌های مقعر..... ۳۷

شکل(۴-۳): پلکان شرودر جهت فهمیدن واژگونی درک شده با تغییر شکل و زمینه [۱۶]..... ۳۸

شکل(۴-۴): منحنی کسینوسی که یک منحنی هموار می باشد..... ۳۸

شکل(۴-۵): نرمال‌ها و انحنا روی یک منحنی مسطح..... ۴۰

شکل(۴-۶): (a) تعریف نرمال‌ها و انحنا در اشکال سه‌بعدی..... ۴۱

شکل(۴-۷): کرانه‌های بخش و part-cutها..... ۴۲

شکل(۴-۸):بخش بندی به کمک مفهوم Limb..... ۴۴

شکل(۴-۹): مثال های از بخش‌های مبتنی بر limb برای اشکال مختلف..... ۴۴

شکل(۴-۱۰):بخش‌بندی به کمک مفهوم Neck..... ۴۵

شکل(۴-۱۱):مثال‌های از بخش‌های مبتنی بر neck برای اشکال مختلف..... ۴۵

شکل(۴-۱۲):اشتراک غیرتقاطعی و اشتراک تقاطعی و تعریف زاویه چرخش..... ۴۷

شکل(۴-۱۳):انواع cuspها و نامگذاری آنها برحسب علامت انحنای هر یک از کمانهایشان..... ۴۸

شکل(۴-۱۴):معرفی مفهوم Baseهای بخش..... ۴۹

شکل(۴-۱۵):اثرات اندازه نسبی روی برجستگی یک بخش..... ۵۱

شکل(۴-۱۶):تخصیص شکل-زمینه و نقشه نرمال برای منحنی‌های مسطح..... ۵۲
شکل(۴-۱۷):مفاهیم مرتبط با نقاط cusp..... ۵۳
شکل(۴-۱۸): معرفی locale مربوط به یک کرانه‌ی بخش..... ۵۴
شکل(۴-۱۹):اهمیت تحدب در تعداد و نحوه بخش‌بندی..... ۵۸
شکل(۴-۲۰): اثر شفافیت درک شده در تقطیع تصویر..... ۵۹
شکل(۴-۲۱):توصیف شرایط لومینانس برای شفافیت درکی..... ۶۰
شکل(۴-۲۲):تشریح نقش شرایط شکلی در درک شفافیت..... ۶۱
شکل(۴-۲۳):اشکال بکار رفته در آزمایش اول سینگ و هافمن[۱۹]..... ۶۳
شکل(۴-۲۴):اشکال بکار رفته در آزمایش دوم سینگ و هافمن[۱۹]..... ۶۳
شکل(۴-۲۵):توصیفی دیگر از تعامل بین تقطیع "sideways" یک شکل به بخش‌هایش، و تقطیع عمیق به لایه‌های مختلف شامل شفافیت..... ۶۴
شکل(۴-۲۶):قاعده مینیما و پردازش پیش-متوجه..... ۶۶
شکل(۴-۲۷):نمایش‌های آزمایش زوو و سینگ[۱۵]..... ۶۷
شکل(۴-۲۸):نمایش‌های آزمایش جستجوی شکل تجزیه شده در میان اشکال تجزیه نشده..... ۶۸
شکل(۵-۱):ابهام در مشخص کردن part-cutها..... ۷۲
شکل(۵-۲):استفاده از نقاط کمکی در بدست آوردن part-cutهای طبیعی..... ۷۲
شکل(۵-۳):نقش محور تقارن محلی در مشخص کردن part-cutها..... ۷۳
شکل(۵-۴):نقش طول cut در مشخص کردن part-cutها..... ۷۴
شکل(۵-۵): قاعده short-cut..... ۷۵
شکل(۵-۶):ضعف روش Neck..... ۷۶
شکل(۵-۷): NCMهای تیزتر نسبت به NCMهای ضعیفتر(هموارتر)، part-cutهای قوی‌تری تشکیل می‌دهند..... ۷۷
شکل(۵-۸): وقتی که مینیمم‌های منفی، ضعیف (تقریباً هموار) هستند، سایر ضرایب همچون طول cut ، می‌توانند سبب دور شدن part-cut از NCMها گردد..... ۷۷
شکل(۵-۹): نمایش تعامل بین طول cut و شدت کرانه‌های بخش. ۷۸
شکل(۵-۱۰):یک مثال از اهمیت امتداد خوب در تجزیه اشکال..... ۷۹
شکل(۵-۱۱):مفهوم امتداد خوب در تشکیل part-cutها..... ۸۰
شکل(۵-۱۲):(a) تجزیه تصویر به وسیله تبدیل محور تقارن (SAT) ۸۱
شکل(۵-۱۳):نقش تقارن محلی در مشخص کردن part-cutها..... ۸۲
شکل(۵-۱۴):تشریح مفهوم تقارن محلی..... ۸۳
شکل(۵-۱۵):ضعف تبدیل SAT..... ۸۴
شکل(۵-۱۶):تعریف تقارن محلی هموارشده..... ۸۵
شکل(۵-۱۷):محاسبه میزانی که دو نقطه روی یک شکل از حالت تقارن محلی خمیدگی، منحرف می‌شوند..... ۸۶

شکل(۵-۱۸):Co-circularity بین دو نقطه i و j [۸۷]..... ۸۷
شکل(۵-۱۹):همواری بین نقاط i , j و k ۸۸
شکل(۵-۲۰):تجزیه سیلوئت یک کانگورو به کمک ضرایب کرویت، همواری و مجاورت..... ۹۰
شکل(۵-۲۱):الویت اعمال ضرایب کرویت و همواری نسبت به اعمال ضریب مجاورت..... ۹۲
شکل(۵-۲۲):نسبی بودن بهترین بخش‌بندی جهت بعضی مثال‌های خاص با الویت قائل شدن برای ضریب مجاورت نسبت به ضریب کرویت..... ۹۲
شکل(۵-۲۳):تجزیه‌ی شکل هواپیما که نسبت به تغییر مقیاس و مقیاس-چرخش تغییر ناپذیر است..... ۹۳
شکل(۵-۲۴):تجزیه اشکال مختلف به کمک ضرایب بخش‌بندی مذکور. ۹۴
شکل(۶-۱):بعضی از تصاویر الگوهای درخت و الگوهای غیردرخت جهت نمایش به میمون‌ها..... ۱۰۱
شکل(۶-۲):نمای کلی یک سیستم خبره‌ی نوعی..... ۱۰۴
شکل(۷-۱):روندنمای مدل محاسباتی درک تصویر مبتنی بر بخش و سطوح مختلف پردازشی آن..... ۱۰۹

کوتاه نوشت‌ها

ANN	Artificial Neural Network
FFT	Fast Fourier Transform
HPF	High Pass Filter
IT	Inferior Temporal
LoG	Laplacian of Gaussian
NCM	Negative Curvature Minima
pp-form	Piecewise Polynomial form
SAT	Symmetry Axis Transform
SLS	Smoothed Local Symmetry
SOM	Self-Organizing Map
2D	2 Dimensional
3D	3 Dimensional

پیشگفتار

بیشترین ارتباط انسان با محیط از طریق سیستم بینایی صورت می گیرد و نزدیک به ۴۰ درصد مغز ما در ارتباط با سیستم بینایی عمل می‌کند[۱]. با پیشرفت‌های علمی در زمینه شناسایی و تشریح سیستم پیچیده بینایی انسان، به تدریج سیستم‌های بینایی مصنوعی با الگو گرفتن از نحوه عملکرد سیستم بینایی انسان شکل گرفته‌اند. با وجود روش‌های زیادی که به موازات پیشرفت علوم و تکنولوژی در زمینه پردازش تصویر بکار گرفته شده است، تجربه یک قرن پردازش تصویر ثابت می‌کند که در عمل بهترین و دقیق‌ترین روش‌ها آنهایی هستند که بر نحوه پردازش تصویر بوسیله انسان منطبق‌ترند.

در اینجا قصد داریم به ارائه یک مدل محاسباتی مبتنی بر بخش از نحوه درک تصویر، که بیشترین انطباق را از لحاظ آزمایشات روان–تنی[1] با درک تصویر در مغز انسان دارد بپردازیم، در حقیقت مدلی کارآمد و مقاوم ارائه دهیم. از آنجا که در مدل ارائه شده، بر مبنای اصول اساسی پردازش تصویر در سیستم بینایی انسان پیش می‌رویم، روشی پیشنهاد می‌کنیم که بیانگر حساس‌ترین و تعیین‌کننده‌ترین عوامل شناسایی اجزاء تصویر به کمک سیستم بینایی انسان باشد و نیز سعی نموده‌ایم تا حد امکان این مدل را جهت پردازش‌های خودکار توسط سیستم‌های کامپیوتری سازگار نماییم. این کار از طریق مفاهیم پایه‌ای هندسه تصویری و بهره‌گیری از اصول کلی پردازش تصویر دیجیتال بر مبنای مفاهیم درکی سیستم بینایی تحقق می‌یابد. قابل ذکر است که امروزه در دانش روز دنیا، بخصوص در مبحث پردازش تصویر به پردازش موازی، توجه ویژه‌ای صورت می‌گیرد، با بیان این نکته در اینجا تاکید می‌کنیم که در مدل ارائه شده از الگوهایی بهره می‌بریم که بسیار مناسب جهت پردازش موازی هستند و بطور کلی مدلی سریع و کارآمد جهت فرآیندهای مختلف پردازشی خواهیم داشت.

بطور کلی در پردازش تصویر دو هدف کلی دنبال می‌شود: یکی بهبود و توسعه اطلاعات تصویری برای پردازش و تفسیر بوسیله انسان، ودیگری پردازش اطلاعات برای درک خودکار توسط ماشین. از آنجا که هدف ما در اینجا پردازش خودکار تصویر منطبق با روش‌های پردازش بینایی انسان می‌باشد ، هر دوی این اهداف را در موازات هم دنبال خواهیم نمود.

در اینجا به کلیات محتوایی این پایان‌نامه اشاره‌ای خواهیم داشت؛ در فصل اول پس از تشریح اهمیت و ضرورت ارائه چنین مدلی، با تشریح چندی از رخدادهای دیداری به کمک مفهوم "بخش"، به بیان مقدمه‌ای نسبتاً جامع، راجع به اهمیت بخش‌بندی تصویر، جهت پردازش و درک آن در سیستم بینایی انسان و شکل‌گیری مدل ارائه شده بر مبنای تقطیع تصویر به بخش‌های تشکیل دهنده‌اش، می‌پردازیم، ضمناً در انتهای این فصل به پیشینه موضوع و تحقیقات صورت گرفته پیشین، اشاره‌ای خواهیم داشت. در فصل دوم، به ارائه الگوریتم مدل محاسباتی درک تصویر مبتنی بر بخش می‌پردازیم. در این فصل ابتدا با کمک مفاهیم بخش‌بندیِ ارائه شده در فصل اول، مدلی کلی از سیستم پردازش و درک تصویر طوری که جهت پردازش‌های دیجیتالی تصاویر مناسب باشد، ارائه شده است، سپس مدل دقیق درک تصویر مبتنی بر سیستم بینایی انسان، بر اساس این مدل کلی معرفی شده است. مدلِ دقیق بر مبنای الگوی نقاط با حداقل انحنای منفی و تقطیع تصویر به بخش‌های تشکیل دهنده‌اش به کمک این نقاط، پایه‌گذاری شده است. در فصل سوم به بیان نمایش و توصیف تصویر که گام اول و دوم در مدل ارائه شده می‌باشد پرداخته‌ایم. جهت نمایش و توصیف تصویر ابتدا با یافتن لبه‌های تصویر، کانتور آن را یافته، سپس یک

---

1. Psychophysically

منحنی پارامتری به کانتور تصویر برازش می‌نماییم. روش بکار گرفته شده برای برازش منحنی استفاده از روش spline می‌باشد. فصل چهارم به مفاهیم بنیادی بخش‌بندی تصویر بر اساس نقاط با حداقل انحنای منفی می‌پردازد. در این فصل با تعاریف و اهمیت این مفاهیم پایه‌ای جهت بهره‌گیری از آنها در فصول آینده آشنا می‌شویم. برخی از مفاهیم شامل برجستگی، تحدب، شفافیت، قاعده short-cut و پردازش پیش‌ـمتوجه می‌باشد. در فصل پنجم به بخش‌بندی نهایی تصویر منطبق بر الگوهای درکی سیستم بینایی انسان می‌پردازیم. در این فصل به نحوه یافتن بهترین part-cutها که بیشترین انطباق با سیستم بینایی را دارند می‌پردازیم. مفاهیم طول cut، شدت کرانه، تقارن محلی و امتداد خوب به عنوان چهار ضریب هندسی موثر در بخش‌بندی در سیستم بینایی تشریح شده‌اند و بر پایه‌ی این ضرایب هندسی ضرایب کرویت، همواری و مجاورت جهت بخش‌بندی نهایی شکل گرفته‌اند. در فصل ششم به تشریح آخرین مرحله از مراحل پردازشی مدل خواهیم پرداخت. در این فصل ابتدا به شناسایی الگوهای بدست آمده از مراحل پیشین و دسته‌بندی آنها پرداخته و شبکه عصبی کوهونن را به عنوان یک دسته‌بندی کننده مناسب جهت اهداف خود پیشنهاد کرده، سپس به تعبیر الگو و درک نهایی تصویر ورودی می‌پردازیم. پیشنهاد ما جهت تعبیر الگو، یک سیستم خبره نوعی می‌باشد و اشاره کرده‌ایم که چنین سیستمی می‌تواند به خوبی بر پایه خصوصیات درکی انسان عمل کند. و نهایتاً در فصل هفتم به نتیجه‌گیری و جمع‌بندی مراحل مختلف تشریح شده می‌پردازیم. در پایان این فصل به زمینه‌های تحقیقاتی که در خلال انجام این تحقیق، جهت تحقیقات آینده و کنکاش بیشتر مناسب به نظر آمده است، اشاره خواهیم نمود. قابل ذکر است در پایان فصول مختلف، با توجه به مفاهیم ارائه شده، تعدادی از مراجع جهت مطالعات پیشرفته‌تر و تحقیقات در حال انجام، ذکر گردیده است که می‌تواند کمک شایانی جهت مطالعات و تحقیقات آینده باشد.

فصل اول

مقـــدمـــه

در این پایان‌نامه به ارائه مدلی محاسباتی از چگونگی درک تصویر، منطبق بر سیستم درکی بینایی انسان می‌پردازیم. منظور از مدل محاسباتی، آن است که تا جایی که ممکن است به کمک روابط ریاضی و بهره‌گیری از ملاحظات هندسی و بکارگیری تئوری‌های پردازش تصاویر دیجیتالی، نحوه عملکرد و پردازش‌های مختلف صورت گرفته توسط مغز را مدلسازی کنیم. در این فصل قصد داریم به اهمیت بخش‌بندی تصویر جهت درک آن توسط انسان بپردازیم، بنابراین پس از بیان اهمیت و لزوم ارائه چنین مدلی، به اهمیت توصیف اشکال بر پایه بخش‌هایشان خواهیم پرداخت و در همین راستا به تشریح چندی از رخدادهای دیداری[1] به کمک مفهوم "بخش" خواهیم پرداخت. بیان دقیق و مشروح این رخدادها در فصل‌های چهار و پنج خواهد آمد. در پایان این فصل به پیشینه موضوع و تحقیقات صورت گرفته در زمینه‌های مشابه و مقالات ارائه شده از نتایج این پایان‌نامه پرداخته می‌شود.

البته خاطر نشان می‌کنیم که بحث در ارتباط با عناصر دریافت بصری مانند چشم، شبکیه، سلسله اعصاب بینایی[2] و نحوه تشکیل تصویر در چشم و چگونگی تطابق و جداسازی روشنایی و غیره خارج از حوزه مطالعاتی این تحقیق می‌باشد و در اینجا فرض بر آن است که آرایه‌ای گسسته از شدت روشنایی در ورودی مدل دریافت می‌شود و اعمال پردازشی روی آن صورت می‌گیرد. ضمناً روی پردازش تصاویر سطح خاکستری (غیر رنگی) تاکید می‌کنیم و کاری به درک رنگ جهت پردازش نداریم. شایان ذکر است تئوری‌های اصلی ارائه شده، به گونه‌ای است که اغلب ابتدا بر اساس تصاویر سه بعدی تعریف شده‌اند سپس به حوزه دو بعدی تعمیم یافته‌اند. البته تئوری‌های نهایی بکار گرفته شده در این تحقیق، تنها جهت پردازش تصاویر دو بعدی بکار گرفته شده‌اند ولی تئوری‌های پایه به گونه‌ای هستند که مدل مذکور قابل تعمیم به حوزه‌ی سه‌بعدی نیز می‌باشد.

۱-۱- انگیزه و ضرورت ارائه چنین مدلی از درک تصویر

در اینجا مناسب است نکاتی چند پیرامون اهمیت، ارزش وکاربرد نتایج این پایان‌نامه را از نظر بگذرانیم؛ به جرأت می‌توان گفت در هر سیستم مصنوعی و دست‌ساز انسان که از الگوها و نمونه‌های مشابه طبیعی الهام

1. Visual phenomena
2. Visual nerves

گرفته شده است، شاهد بهینه شدن سیستم بوده‌ایم. از طرفی علیرغم وجود رشد فزاینده علوم، هنوز این الگوبرداری‌ها وشبیه‌سازی‌ها راهی بس طولانی تا رسیدن به غایت خود دارند. بنابراین اهمیت و ضرورت اصلی این تحقیق را می‌توان کامل‌تر کردن سیستم‌های بینایی مصنوعی که از نحوه‌ی عملکرد سیستم بینایی انسان الگو می‌گیرند، ذکر کرد.

توجه به این نکته جالب است که با توجه به اینکه هنوز در علم پردازش تصویر و بینایی ماشین، یک مدل محاسباتی مبتنی بر بخش، بصورت یکپارچه جهت درک تصویر منطبق با خصوصیات درکی انسان ارائه نشده است، ارائه چنین مدل‌هایی را می‌توان حتی نقطه عطفی در این علوم دانست که کمک شایانی به ارائه الگوریتم‌های جدید علمی در حوزه‌های گسترده‌ای از این علوم می‌نماید. (شایان ذکر است تاکنون مدل‌هایی مانند مدل ارائه شده در [۲]، معرفی شده‌اند که از ویژگی‌های سیستم بینایی الهام گرفته‌اند ، ولی هر یک از آنها بر اساس ویژگی‌های بخصوص دیگری پایه‌گذاری شده‌اند و دارای کلیت لازمه نیز نمی‌باشند.)

در ذیل به بعضی انگیزه‌های ارائه و کاربردهایی که چنین مدلی می‌تواند داشته باشد اشاره می‌کنیم:

- کاربرد در سیستم های هوش مصنوعی نظیر تشخیص دستنوشته‌ها و سایر الگوهای دینامیک.
- پی بردن به ارتباط بین سیستم بینایی انسان و سایر فعالیت‌های ماهیچه‌ای مانند موقعیت سر، موقعیت چشمان، حرکت بدن و همگرایی در چشم.
- ایجاد یک مدل سه بعدی از یک تصویر دو بعدی جهت کاربردهای مختلف نظامی، نجومی، صنعتی و ... (مرجع [۳] را ببینید.)
- کمک به پیشرفت علم ربات‌های انسان‌نما[1] با بکارگیری این مدل جهت بینایی این ربات‌ها و روش درک چنین مدلی جهت سایر زمینه‌های درکی آنها.
- کمک به پیشرفت علومی چون پزشکی و علوم عصبی[2] جهت شناخت بیشتر قشر بینایی مغز[3] و سایر بخش‌های سیستم بینایی و پیچیدگی‌های سیستم درکی.
- کاربرد در علومی چون فشرده‌سازی و بازیابی تصاویر در علم پردازش تصویر و بینایی کامپیوتر جهت ارائه الگوریتم‌های بهینه. (آدمی میلیونها تصویر را در ذهن خویش ثبت و ضبط می‌کند و در مواقع لزوم حتی با وجود تغییرات اساسی در آن، در کمترین زمانی قابلیت به یادآوری و شناسایی مجدد را داراست.)

از دیگر کاربردهایی که برای چنین مدلی می‌توان برشمرد آن است که در بسیاری از کاربردها و مقالات علمی، در بررسی نتایج خروجی و مقایسه‌ها از انسان و در حقیقت سیستم درکی وی جهت ارزیابی شهودی استفاده می‌شود، حال اگر بتوان مدلی ریاضی برای آن ارائه نمود، این مقایسه‌ها و نتیجه‌گیری‌ها بسیار ساخت

---

1. Humanoid Robotics
2. Neuroscience
3. Visual Cortex

یافته‌تر شده و بررسی نتایج خروجی را می‌توان به رایانه‌ها محول نمود که در بعضی موارد با سرعت بیشتر و خطای کمتر همراه خواهد بود و در غالب موارد بسیار کم هزینه‌تر خواهد بود.

۱-۲- اهمیت توصیف اشکال مبتنی بر بخش‌هایشان

جهان دیداری ما پر از ساختارهایی به صورت اشیاء و رویه‌ها[1] می‌باشد. هر چند در ورودی سیستم‌های بینایی و پردازش تصویر، در واقع هیچ شیء یا رویه‌ای نداریم بلکه فقط آرایه‌هایی گسسته از مقادیر شدت روشنایی را داریم که می‌توانند به صورت قابل توجهی با تغییرات کوچک در شرایط دید تغییر کنند. بنابراین اشیاء و رویه‌هایی که ما می‌بینیم لازم است ابتدا توسط سیستم بینایی‌مان از این ورودی‌ها بازسازی شود. یک حالت مهم و بحرانی از این بازسازی، تجزیه آرایه‌های ورودی به نواحی است که به نظر از اشیاء مجزا ایجاد می‌شوند؛ به عبارت دیگر منظور ما همان "واحدهای درکی[2]" می‌باشد.

اگرچه اشیاء به صورت کلی، فقط واحدهای درکی نیستند. اشیایی که ما می‌بینیم همگی غیرساختاری[3] نیستند و در واقع بیشترشان دارای ساختار بخش[4] می‌باشند. برای مثال وقتی یک میز را درک می‌کنیم، آن را به عنوان یک شی درکی منسجم تشخیص می‌دهیم؛ همچنین این درک به عنوان یک چینش[5] فضایی از بخش‌هایی که به شکل واضحی آنها را تعریف کرده‌ایم نیز می‌تواند باشد: یک صفحه روی میز و چهار عدد پایه. بنابراین "واحدهای درکی" در سطوح مختلفی وجود دارند: در سطح اشکال کلی، در سطح بخش‌ها، و در سطح بخش‌های کوچکتر که به صورت تودرتو درون یک بخش بزرگتر قرار می‌گیرند[۴, ۵].

تئوری‌های زیادی بیان می‌دارند که توصیف کننده‌های اشکال در سیستم بینایی انسان مبتنی بر بخش نیستند و علت آن را نیز حساسیت سیستم بینایی در شناسایی شکل به زاویه دید[6]، ذکر می‌کنند. هر چند این اعتقاد بر پایه این فرض است که یک رهیافت مبتنی بر بخش، معادل با یک رهیافت حجمی و متمرکز شده روی شی، جهت توصیف آن در مجموعه‌ای از بخش‌های سه بعدی پایه‌ای است که به شکل اولیه مشخص شده‌اند[۵, ۶]. اما در اینجا بیان می‌کنیم که این ادعا صرفاً صحیح نیست، البته خاطر نشان می‌کنیم که تئوری‌های زیادی تابحال در این زمینه به بررسی سطوح مختلف پردازش در سیستم بینایی انسان جهت درک اشیاء پرداخته‌اند و پس از این نیز خواهند پرداخت و شاید واقعاً به تنهایی نتوان گفت یکی از آنها کاملاً صحیح است و در همه موارد به شکل غالب عمل می‌کند (برای مثال [۷, ۸] را جهت نمونه‌هایی از بینایی

---

1. Surfaces
2. Perceptual units
3. Unstructured
4. Part
5. Arrangement
6. Viewpoint

سطح میانی[1] ملاحظه کنید که می‌توان گفت نحوه درک مبتنی بر بخش که بصورت اولیه و در سطح پایینی از پردازش قرار دارد را نقض می‌کند، ولی می‌توان گفت سیستم بینایی‌مان در سطوح مختلف پردازشی عمل می‌کند و نه یک سطح خاص.)، ولی هدف این تحقیق، بررسی نحوه درک اشیاء در پایین‌ترین سطح پردازش و ارائه مدل منطبق بر آن می‌باشد که ثابت خواهیم کرد که در اغلب موارد نحوه فعالیت سیستم بینایی انسان مبتنی بر بخش می‌باشد، علیرغم آنکه در بعضی روندهای پردازشی مانند پردازش حرکت، پردازش استریو و جستجوی تصاویر طبق [۷, ۸] همیشه خوب عمل نمی‌کند. بخش‌ها جهت عملیات پردازشی 2D (اشیاء تصویری دوبعدی)، $2-\frac{1}{2}$ D (اشیاء رویه ای[2]) و 3D (اشیاء حجمی) کاربرد دارند و بکار بردن بخش‌ها سبب نمی‌شود که به مجموعه‌ای ثابت از اشکال پایه‌ای محدود شویم[۹].

در مدلی که در این پایان‌نامه بدان می‌پردازیم برپایه معیارهای محاسباتی[3] و تجربی[4] نشان می‌دهیم که توصیف اشکال بصری توسط سیستم بینایی به طور غالب مبتنی بر بخش می‌باشد، طوری که اشیاء به بخش‌های مولفه‌ای‌شان تجزیه شده و به کمک این بخش‌ها و ارتباط فضایی بین آنها، دسته‌بندی می‌شوند و مورد تشخیص الگو و درک قرار می‌گیرند.

از دیدگاه محاسباتی، بخش‌ها به دلایل زیادی مفید هستند. اول آنکه اشیاء زیادی به صورت مفصلی[5] هستند: اشیاء مفصلی، شامل بخش‌هایی هستندکه می‌توانند بدون محدودیت خاصی در ارتباط با سایر بخش‌ها حرکت کنند. برای مثال انگشتان دستها بطور مستقل می‌توانند حرکت کنند. یک توصیف مبتنی بر بخش، اجازه جداسازی اشکال از بخش‌ها را بر اساس ارتباط فضایی بین این بخش‌ها می‌دهد ــ بنابراین یک روش طبیعی جهت توصیف و تشخیص اشیاء مفصلی ارائه می‌دهد. مزیت دیگر آنکه همیشه ناظر، تمام شی را از یک زاویه دید نمی‌بیند؛ قسمت پشت اشیاء توسط قسمت جلوی آنها پوشیده شده است و یا حتی قسمت جلوی شی ممکن است از یک زاویه دید، توسط شی دیگری پوشیده شده باشد. توصیف اشیاء با بخش‌هایشان اجازه می‌دهد که عمل شناسایی، با آن بخش‌هایی که قابل رویت هستند ادامه یابد.

از دیدگاه تجربی، شواهد اخیر علمی نشان می‌دهند که بخش‌ها، خیلی سریع، بصورت خودکار، و با پردازش موازی در محیط دیداری، قابل محاسبه هستند[۱۰-۱۵]. همچنین در [۱۶] نشان داده شده است که بخش‌ها روی تخصیص شکل و زمینه نیز تاثیر می‌گذارند.

در واقع نشان می‌دهیم که بخش‌ها، تنوع قابل توجهی را در رخدادهای بینایی نشان می‌دهند که در اینجا به اختصار به بعضی از آنها اشاره می‌کنیم و در فصول بعدی به شرح جزئیات بیشتر و دقیقتری از هر کدام خواهیم پرداخت؛

---

1. Mid-level vision
2. Surface-based
3. Computational
4. Empirical
5. Articulated