بسم الله الرحمن الرحيم

دانشگاه آزاد اسلامی

واحد تهران مرکزی

دانشکده فنی و مهندسی

پایان‌نامه برای دریافت درجه کارشناسی‌ارشد (M.Sc)

گرایش: الکترونیک

عنوان :

بهینه سازی شبیه سازی تعقیب بلادرنگ دست مبتنی بر شبکه عصبی بر روی FPGA

استاد راهنما :

دکتر فرداد فرخی

اسلک مشاور:

دکتر کاوه کنگرلو

پژوهشگر :

مهدی حیدریان

تابستان ۱۳۹۲

تقدیم به :

پدر و مادر عزیز و مهربانم

شکر و سپاس خدا را که بزرگترین امید و یاور در لحظه لحظه زندگیست که از روی لطف و کرم، پدر و مادری فداکار نسبیم ساخته تا در سایه درخت پربار وجودشان بیاسایم و از ریشه آنها شاخ و برگ گیرم و از سایه وجودشان در راه کسب علم و دانش تلاش نمایم. والدینی که بودنشان تاج افتخاری است بر سرم و نامشان دلیلی است بر بودنم، چرا که این دو وجود، پس از پروردگار، مایه هستی ام بوده اند دستم را گرفتند و راه رفتن را در این وادی پر فراز و نشیب زندگی آموختند. آموزگارانی که برایم زندگی، بودن و انسان بودن را معنا کردند تا من به جایگاهی که اکنون در آن ایستاده ام برسم.

نه می توانم موهایشان را که در راه عزت من سفید شد، سیاه کنم و نه برای دستهای پینه بسته شان که ثمره تلاش برای افتخار من است، مرهمی دارم . پس توفیقم ده که هر لحظه شکر گزارشان باشم و ثانیه های عمرم را در عصای دست بودنشان بگذرانم.

تشکر و قدردانی :

استاد گرامی جناب آقای دکتر فرخی

دلسوزی ، تلاش و کوشش حضرتعالی در تعلیم و تربیت و انتقال معلومات و تجربیات ارزشمند در کنار برقراری رابطه صمیمی و دوستانه با دانشجویان و ایجاد فضائی دلنشین برای کسب علم و دانش و درک شرایط دانشجویان حقیقتاً قابل ستایش است .اینجانب بر خود وظیفه میدانم در کسوت شاگردی از زحمات و خدمات ارزشمند شما استاد گرانقدر تقدیر و تشکر نمایم.

سپاس بی کران پروردگار یکتا را که هستی مان بخشید و به طریق علم و دانش رهنمونمان شد و به همنشینی رهروان علم و دانش مفتخرمان نمود و خوشه چینی از علم و معرفت را روزیمان ساخت .از خداوند متعال برایتان سلامتی ، موفقیت و همواره یاد دادن را مسئلت دارم.

با تشکر شاگرد همیشگی

مهدی حیدریان

در پایان از خواهر عزیز و مهربانم که صبورانه اینجانب را در تهیه این پایان نامه تحمل کردن که با آرامش روحی و آسایش فکری و حمایت های همه جانبه در محیطی مطلوب، مراتب تحصیلی و همچنین پایان نامه را به اتمام برسانم صمیمانه سپاسگذارم.

تعهد نامه اصالت پایان نامه **کارشناسی ارشد**

اینجانب.....مهدی حیدریان.....**دانش اموخته مقطع کارشناسی ارشد ناپیوسته به شماره دانشجویی**..... ۸۹۰۹۲۹۶۹۳۰۰.....**در رشته** ...برق گرایش الکترونیک.....**که در تاریخ** ۹۲/۶/۱۸ **از پایان نامه خود تحت عنوان** بهینه سازی شبیه سازی تعقیب بلادرنگ دست مبتنی بر شبکه عصبی بر روی FPGA

با کسب نمره ..... ودرجه ..... دفاع نموده ام بدینوسیله متعهد می شوم:

۱- این پایان نامه حاصل تحقیق وپژوهش انجام شده توسط اینجانب بوده ودر مواردی که از دستاوردهای علمی وپژوهشی دیگران (اعم از پایان نامه ، کتاب ، مقاله و...) استفاده نموده ام، مطابق ضوابط و رویه های موجود ،نام منبع مورد استفاده وسایر مشخصات آن را در فهرست ذکر ودرج کرده ام

۲ - این پایان نامه قبلا برای دریافت هیچ مدرک تحصیلی (هم سطح ، پایین تر یا بالاتر) درسایر دانشگاهها و موسسات آموزش عالی ارائه نشده است.

۳ - چنانچه بعد از فراغت از تحصیل ، قصد استفاده وهرگونه بهره برداری اعم از چاپ کتاب ، ثبت اختراع و ....از این پایان نامه داشته باشم ، از حوزه معاونت پژوهشی واحد مجوزهای مربوطه را اخذ نمایم.

٤ - چنانچه در هر مقطع زمانی خلاف موارد فوق ثابت شود ، عواقب ناشی از آن را بپذیرم و واحد دانشگاهی مجاز است با اینجانب مطابق ضوابط ومقررات رفتار نموده ودر صورت ابطال مدرک تحصیلی ام هیچگونه ادعایی نخواهم داشت .

نام ونام خانوادگی:

تاریخ و امضاء:

فهرست مطالب

عنوان صفحه

فصل اول : کلیات طرح
۱-۱ چکیده..... ۲
۱-۲ بیان مسئله..... ۲
۱-۲-۱ سوالات و فرضیه تحقیق..... ۵
۱-۲-۲ اهداف، اهمیت و ضرورت تحقیق..... ۷
۱-۲ تحقیقات مرتبط انجام شده..... ۸
فصل دوم: مطالعات نظری
مقدمه
۲-۱ انواع روش‌های ردیابی در تصویر..... ۱۹
۲-۱-۱ ردیابی دست دو بعدی مبتنی بر تصویر..... ۱۹
۲-۱-۲ جریان نوری..... ۲۲
۲-۱-۳ میانگین شیفت..... ۲۵
۲-۲ مروری بر شبکه‌های عصبی..... ۲۷
۲-۲-۱ شبکه عصبی MLP..... ۲۸
۲-۲-۲ کاربرد شبکه‌های عصبی..... ۳۱
۲-۳ بررسی FPGA به‌عنوان بستری مناسب برای پیاده‌سازی شبکه عصبی به همراه نرم‌افزارها و زبان‌های سخت‌افزاری موجود..... ۳٤
۲-۳-۱ ساختار FPGAها..... ۳٦
۲-۳-۲ سلول‌های منطقی FPGA..... ۳۷
۲-۳-۳ مزایای FPGA..... ۳۹
۲-۳-٤ Verilog/VHDL..... ٤۰
۲-۳-۵ نرم‌افزار Xilinx ISE..... ٤۳

۲-۳-۵-۱ ابزارهای سنتز..... ٤٥

فصل سوم :روش شناسایی تحقیق (متدولوژی)

مقدمه

۳-۱ ردیابی دست ٤٨

۳-۱-۱ آشکار سازی پوست (شبکه اول) ٤٨

۳-۱-۲ آشکارسازی حالات مختلف دست (شبکه دوم)..... ٥٤

۳ -۱-۲-۱ انتخاب ویژگی (DCT,DFT,Hu invariant moment)..... ٥٥

۳-۲ پیاده‌سازی انجام شده..... ٥٨

۳-۲-۱ واحد محاسباتی با موازی‌سازی چندگانه..... ٥٩

۳-۲-۱-۱ ساختار واحد محاسباتی..... ٥٩

۳-۲-۱-۲ مدل محاسباتی بین لایه ورودی و لایه میانی..... ٦٠

۳-۲-۱-۳ مدل محاسباتی برای نرون خارجی..... ٦١

۳-۲-۲ پیاده‌سازی سخت‌افزاری تابع تحریک سیگموئید..... ٦٣

۳-۲-۳ پیاده‌سازی لینک..... ٦٥

۳-۳ سنتز شبکه‌های ۲-۸-۷ و ۲-۸-۴۲..... ٦٧

۳-٤ شبیه‌سازی..... ٧١

۳-٥ پیاده‌سازی شبکه‌های ۲-۸-۷ و ۲-۸-۴۲..... ٧٦

فصل چهارم :تجزیه و تحلیل یافته های تحقیق

مقدمه

٤-۱ نتایج شبکه اول (آشکارسازی پوست) و شبکه دوم (آشکارسازی دست) در مطلب..... ٨٤

٤-۲ نتایج شبکه اول (آشکارسازی پوست ۲-۸-۷) و دوم(آشکارسازی دست ۲-۸-٤۲) در FPGA... ٨٨

فصل پنجم: نتیجه گیری

نتیجه‌گیری و پیشنهادات..... ١٠٣

فهرست منابع و ماخذ..... ۱۰۵
پیوست ها و ضمایم..... ۱۱۲
چکیده انگلیسی..... ۱۱۶

فهرست جدول ها

عنوان صفحه

۱- ۱ جدول مقایسه انواع پیاده‌سازی تابع سیگموئید در تحقیق دوم..... ۱۲
۲- ۱ جدول مقایسه زبان VHDL و Verilog..... ۴۳
۳- ۱ جدول سطر اول RAM مربوط به وزن لینک‌های بین لایه ورودی و لایه میانی..... ۶۵
۳- ۲ جدول سطر اول RAM مربوط به مقادیر بایاس لایه میانی..... ۶۵
۳- ۳ جدول RAM مربوط به وزن‌های متناظر با لینک‌های بین نرون‌های لایه میانی و نرون‌های لایه خروجی..... ۶۶
۳- ۴ جدول RAM مربوط به مقادیر بایاس نرون‌های خروجی..... ۶۶
۳- ۵ جدول نتایج نگاشت و مسیریابی شبکه ۲-۸-۴۲ (شبکه آشکارساز دست)..... ۸۲
۴- ۱ جدول نتایج شبکه اول (آشکارسازی پوست) در مطلب..... ۸۵
۴- ۲ جدول نتایج شبکه اول (آشکارسازی پوست) در مطلب..... ۸۶
۴- ۳ جدول تعداد تصویرهایی که از هر حالت، به شبکه آشکارساز دست آموزش داده شده است..... ۸۶
۴- ۴ جدول نتایج حاصل از شبکه دوم (آشکارسازی دست) در مطلب..... ۸۷
۴- ۵ جدول مقایسه نتایج حاصل از شبکه اول (آشکارسازی پوست) در مطلب و FPGA برای ماتریس تست این شبکه در مطلب..... ۹۶
۴- ۶ جدول مقایسه نتایج شبکه دوم (آشکارسازی دست) در مطلب و FPGA برای ۱۲ سطر اول ماتریس تست این شبکه..... ۹۷
۴- ۷ جدول مقایسه نتایج شبکه دوم (آشکارسازی دست) در مطلب و FPGA برای سطر ۱۲ تا ۱۸ ماتریس تست این شبکه..... ۹۸

فهرست شکل ها

عنوان صفحه

۱- ۱ شکل سیستم مجازی چند مدلی برای HCI ۳
۱- ۲ شکل دستکش سایبر (Cyber-Glove) به همراه حسگرهای آن ٤
۱- ۳ شکل مدل ریاضی شبکه عصبی MLP استفاده شده در تحقیق اول ۹
۱- ٤ شکل بردارهای ورودی (پیکسل ها) در تحقیق اول ۹
۱- ٥ شکل تقریب تابع tanh با LUT ۹
۱- ٦ شکل a تصویر اصلی.b تصویر حاصل از شبیه‌سازی و سنتز با $\gamma_e = 199$ .c حاصل از شبیه‌سازی و سنتز با $\gamma_e = 221$ ۱۰
۱- ۷ شکل شبکه MLP در تحقیق دوم(پیاده‌سازی آشکارسازی صورت بر روی FPGA) ۱۱
۱- ۸ شکل معماری توزیع به‌کار رفته در تحقیق سوم ۱٥
۲- ۱ شکل تجزیه و تحلیل حرکت دست با استفاده از الگوریتم جریان نوری ۲٤
۲- ۲ شکل ردیابی دست با استفاده از اثر حرکت و رنگ دست ۲٥
۲- ۳ شکل الگوریتم شیفت میانگین ۲٦
۲- ٤ شکل ردیابی دست با استفاده از الگوریتم CamShift ۲٦
۲- ٥ شکل شبکهٔ پرسپترون چند لایه ۲۸
۲- ٦ شکل مدل نرون McCulloch-Pitts ۲۹
۲- ۷ شکل مدل گراف الگوریم پس انتشار ۳۰
۲- ۸ شکل موازنه بین بسترهای مختلف سخت‌افزاری ۳٥
۲- ۹ شکل ساختار کلی یک FPGA به همراه آرایه‌ای از سلول‌های منطقی و کانال‌های مسیریابی افقی و عمودی ۳۷
۲- ۱۰ شکل چگونگی پیاده‌سازی مدار توسط LUT ۳۸
۲- ۱۱ شکل تحقق یک LUT به‌صورت فیزیکی سلول منطقی مبتنی بر مالتی پلکسر ۳۸

۲- ۱۲ شکل نحوه پیاده‌سازی یک مدار کوچک در سلول منطقی مبتنی بر مالتی پلکسر.....۳۹
۳- ۱ شکل نمونه‌های پوست به‌کار رفته در شبکه اول (آشکارسازی پوست).....٤۹
۳- ۲ شکل فلوچارت آموزش شبکه اول (آشکارسازی پوست).....۵۱
۳- ۳ شکل فلوچارت تست شبکه اول(آشکارسازی پوست).....۵۳
۳- ٤ شکل ساختار شبکه اول (آشکارسازی پوست) به همراه تعداد نرون‌ها و بردار ورودی.....۵٤
۳- ۵ شکل واحد محاسباتی با موازی‌سازی چندگانه.....٦۰
۳- ٦ شکل مدل محاسباتی بین لایه ورودی و لایه میانی.....٦۱
۳- ۷ شکل دیاگرام برای چهار نرون خروجی موازی.....٦۲
۳- ۸ شکل جزئیات مربوط به هر نرون خروجی.....٦۳
۳- ۹ شکل بررسی لینک‌های شبکه با افزایش آدرس حافظه‌ها.....٦۷
۳- ۱۰ شکل یک مدار نمونه جهت شبیه‌سازی.....۷۲
۳- ۱۱ شکل نتایج یک شبیه‌سازی رویدادگرا.....۷۳
۳- ۱۲ شکل نتایج شبیه‌سازی مدار نمونه در Isim.....۷۳
۳- ۱۳ شکل نتایج شبیه‌سازی شبکه ۲-۸-۷.....۷۵
۳- ۱٤ شکل نتایج شبیه‌سازی شبکه ۲-۸-٤۲ (شبکه تشخیص دست).....۷٦
۳- ۱۵ شکل نمایش ویرایشگر PlanAhead جهت تعیین پورت های I/O در قطعه هدف (XCS1000 4fg456) برای شبکه ۲-۸-۷.....۷۸
۳- ۱٦ شکل جانمایی و مسیر یابی حافظه (RAM) در بردارنده وزن‌های بین لایه ورودی و میانی در قطعه هدف(XCS1000 4fg456) برای شبکه ۲-۸-۷.....۷۹
۳- ۱۷ شکل ایجاد یک فایل PROM جهت برنامه‌ریزی قطعه هدف توسط IMPACT.....۸۱
٤- ۱ شکل اعمال شبکه MLP بر روی تصویر ویدئویی.....۸۵
٤- ۲ شکل تشخیص صورت و ردگیری حالت‌های مختلف دست در مطلب.....۸۷
٤- ۳ شکل تصاویر خروجی FPGA در شبکه آشکارسازی پوست (۲-۸-۷).....۱۰۰
٤- ٤ شکل خروجی شبکه دوم،آشکارساز دست (۲-۸-٤۲) در FPGA.....۱۰۰

٤- ٥ شکل تصاویر خروجی شبکه دوم(٢-٨-٤٢ ، آشکارساز دست) در FPGA به صورت باینری..... ١٠١
الف- ١ شکل ماژول نرون لایه میانی..... ١٠٥
الف- ٢ شکل ماژول مربوط به نرون خروجی..... ١٠٥
الف- ٣ شکل RAM مربوط به وزن‌های متناظر با لینک‌های بین لایه ورودی و لایه میانی..... ١٠٦
الف- ٤ شکل RAM مربوط به وزن‌های متناظر با لینک‌های بین لایه میانی و لایه خروجی..... ١٠٦
الف- ٥ شکل RAM مربوط به بایاس متناظر با نرون‌های لایه میانی..... ١٠٧
الف- ٦ شکل RAM مربوط به بایاس متناظر با نرون‌های لایه خروجی..... ١٠٧
الف- ٧ شکل ماژول مربوط به واحد کنترل در سه نمای مختلف..... ١٠٧
الف- ٨ شکل شماتیک شبکه ٢-٨-٧ به همراه ارتباط بین اجزای آن..... ١٠٨
الف- ٩ شکل الف و ب شماتیک شبکه ٢-٨-٤٢.ج.ارتباط بین اجزای آن..... ١٠٩

فصل اول : کلیات طرح

1-1 چکیده

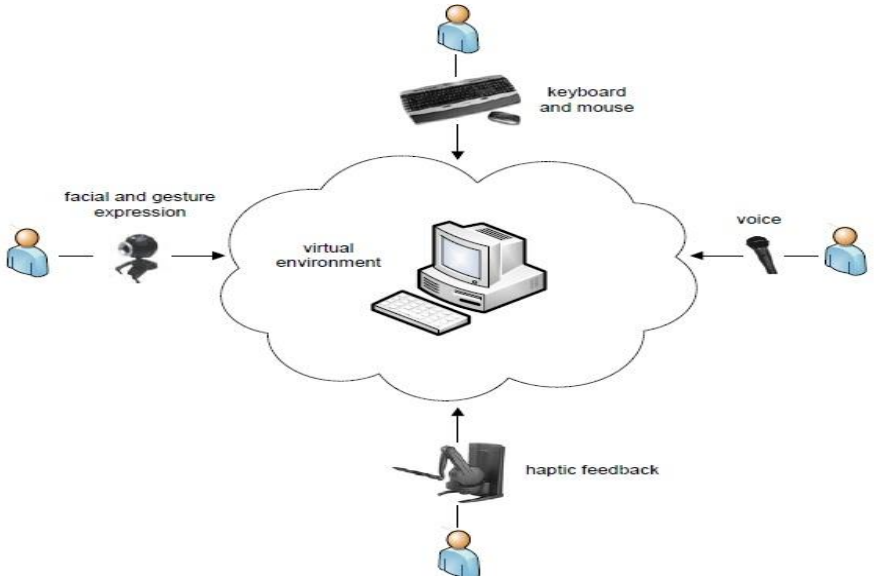
در این پایان‌نامه از دو شبکه عصبی پرسپترون چند لایه (MLP)[1] که به صورت با سرپرست آموزش می‌بینند جهت تشخیص و ردیابی دست استفاده شده است. شبکه نخست جهت تشخیص پوست آموزش می‌بیند و شبکه دوم علاوه بر تشخیص دست، حالات مختلف دست[2] را نیز تشخیص می دهد. از شبکه اول یک تصویر باینری که ناحیه پوست را نشان می دهد، بدست می‌آید و این تصویر به شبکه دوم داده می‌شود، هر دو شبکه دارای دو کلاس است . در شبکه اول، کلاس اول پوست و کلاس دوم غیر پوست. در شبکه دوم برای تشخیص دست از صورت نیز از دو کلاس استفاده شده ؛ که کلاس اول دست و کلاس دوم صورت است. در پایان این دو شبکه بر روی FPGA[3] پیاده‌سازی می‌شوند. ویژگی‌ها در شبکه دوم به صورتی انتخاب شده است که برای ردگیری، وابستگی به نزدیک کردن دست به دوربین نداشته باشد، تا بتوان از هر فاصله و با هر ابعادی دست و حالات آن را ردگیری کرد. همچنین این ویژگی‌ها در عین حال که از نظر تعداد (جهت بالا بردن سرعت پردازش) مناسب است، از دقت قابل قبولی نیز برخوردار است. در واقع در این پایان نامه سعی شده است، علاوه بر سرعت، دقت نیز مورد توجه قرار گیرد که بتوان یک سیستم کاربردی برای HCI[4] بوجود آورد.

1-2 بیان مسئله

رابطهای انسان و کامپیوتر از رابطهای مبتنی بر متن از طریق رابطهای مبتنی بر گرافیک دو بعدی، رابطهای پشتیبانی چند رسانه ای، محیط مجازی VE[5] سه بعدی مبتنی بر چند مدلی که به طور کامل تکامل یافته است، توسعه یافته است.در حالی که یک پارادایم(paradigm) جدید سطح بالا برای برقراری ارتباط، یادگیری، آموزش و سرگرمی ارائه می‌شود، محیطهای مجازی نیز چالش‌های جدیدی برای تعامل بین انسان و کامپیوتر ارائه می‌دهند. دستگاه‌های سنتی، HCI دو بعدی مانند صفحه کلید و موس‌ها برای آخرین برنامه‌های کاربردی VE کافی نیستند. در عوض، سیستم VE فرصتی برای استفاده از روش‌های مختلف ارتباطی و تکنولوژی حسی برای ارائه تجربه کاربردی همه جانبه محیا می کند (Pavlovic& Huang& SHrma,1996,30-35),(Turk,2001,100) همانطور که در شکل ۱-۱ نشان داده شده است ،

[1] Multilayer perceptron
[2] Hand Gestures
[3] Field Programmable Gate Array
[4] Human Computer Intraction
[5] Virtual environment

وسایلی که موقعیت و جهت‌گیری بدن، گفتار و صدا، بیان چهره و حالات صورت ، بازخورد لمسی و دیگر جنبه‌های رفتار یا حالت انسان را می‌تواند تشخیص دهد برای تعامل قوی‌تر و مؤثرتر بین انسان و کامپیوتر مورد استفاده قرار می‌گیرد.برای رسیدن به تعامل طبیعی و همه‌جانبه انسان و کامپیوتر ، دست انسان می‌تواند به عنوان یک دستگاه رابط استفاده شود ,(Baudel& Beaudouin, 1993,28-35) (Wexelblatt., 1995,179-200) (Kirishima& ,Sato& Chihara, 2005,351-364)حالات مختلف دست کانال ارتباطی قدرتمند انسان به انسان است که بخش عمده‌ای از انتقال اطلاعات در زندگی روزمره ما می‌باشد. حرکات دست یک راه آسان و طبیعی برای تعامل است. به عنوان مثال، زبان‌های اشاره به طور گسترده در میان افراد معلول از نظر گفتاری استفاده می‌شود. همچنین افرادی که می‌توانند صحبت کنند هم از بسیاری از انواع حرکات برای کمک به ارتباطات خود استفاده می‌کنند.



شکل ۱-۱ سیستم مجازی چند مدلی برای HCI

با این حال، معنی حرکات دست به طور کامل برای تعامل انسان و کامپیوتر کشف نشده است . در مقایسه با دستگاه‌های سنتی HCI، استفاده از حالات دست دشوار است و تعامل برای کاربران با کامپیوتر و کشف دنیای مجازی سه بعدی راحت‌تر است (Wu& Huang, 2001,51-60). حالات دست می‌تواند در طیف گسترده ای از برنامه‌های کاربردی مرسوم ، از طریق حرکات (زبان‌های اشاره گوناگون) و بکارگیری اشیاء در VE به عنوان تعامل انسان و کامپیوتر مورد استفاده قرار گیرد.

برای استفاده از دست انسان به عنوان یک وسیله طبیعی HCI، دستکش داده‌ها مانند دستکش سایبر (Cyber-Glove) (Cyberglove II Datasheet) که در شکل ۱-۲ نشان داده شده برای ثبت حرکات دست استفاده می‌شود، بسنسورهای متصل شده، زوایای مشترک و موقعیت‌های مکانی از دست را می‌توان به طور مستقیم از طریق دستکش اندازه‌گیری کرد. با این حال، دستکش داده‌ها و سیم‌های وابسته به آن هنوز هم برای پوشیدن کاربران مزاحمت ایجاد می کند.علاوه بر این، هزینه‌های دستکش داده اغلب برای کاربران عادی بیش از حد بالاست)(Chen& Etal ,2005,110-115) (Fels& Hinton,1998,205-212) , Yang& Xu& Chen, 1994,1747-1752)

شکل ۱-۲ دستکش سایبر (Cyber-Glove) به همراه حسگرهای آن.

سیستم‌های تشخیص حالات دست مبتنی بر تصویر می‌توانند حرکات مختلف دست را از ورودی‌های ویدئویی شناسایی کنند و از آنها به عنوان دستورات مجازی قابل درک برای رایانه‌ها که به آن پاسخ می‌دهند، استفاده کنند(Geer, 2004,20-23). با دوربین ویدئویی به عنوان دستگاه ورودی و تصویر کامپیوتر و تکنیک‌های پردازش تصویر، سیستم‌های تشخیص حرکت دست مبتنی بر تصویر به دلیل سخت‌افزار ساده و قابل حمل ، راحت‌تر و کارآمدتر هستند . سیستم‌های تشخیص حرکت دست مبتنی بر تصویر می‌تواند یک کانال ارتباطی بصری برای تعامل انسان و کامپیوتر محیا کند . به عنوان مثال، کاربر می‌تواند یک محیط مجازی سه بعدی را هدایت کند و اشیاء مجازی را با مجموعه‌ای از دستورات حرکت، بدون کمک از صفحه کلید، موس یا جوی استیک کنترل کند -Cruz-Neira& Sandin& DeFanti , 1993,135) 142) (Shafer& Etal.,1998.127-130)(Pavlovic& Sharma& Huang, 1996,30-35).

تحقیقات اولیه در مورد حالات دست مبتنی بر تصویر معمولاً به مساعدت نشانگر و یا دستکش‌های رنگی انجام می‌شود(Joslin&Etal.,2005,1706-1711) (Keskin&Erkan&Akarun, 2003,359-366). استفاده از نشانگر برای بالا بردن دقت و سرعت جهت بلادرنگ بودن ردیابی است . اما در حال حاضر در زمینه ردیابی دست مبتنی بر تصویر و تکنیک‌های تشخیص حرکت، تحقیقات بیشتر روی ردیابی دست بدون کمک از نشانگرها و دستکش متمرکز شده است.

۱-۲-۱ سوالات و فرضیه تحقیق

- چرا در تصویر های دارای نویز محیطی برای پردازش سریع و دقیق و بلادرنگ تصویر جهت تشخیص و ردیابی دست باید از شبکه عصبی به جای فیلتر های دیگر استفاده کرد؟
- بهترین مدل شبکه عصبی جهت تشخیص و ردیابی دست کدام است؟
- چرا برای پیاده سازی شبکه عصبی بروی سخت افزار باید از پروسسور هایی مانند FPGA استفاده کرد؟

در بسیاری از کارهای انجام شده وجود تغییرات نور و یا پیچیده شدن پس زمینه، باعث بروز مشکلاتی می‌شود که دقت کار را به شدت کاهش می‌دهد. از این رو استفاده از یک روش و الگوریتم مانند شبکه عصبی که قادر به یاد گیری است تا حدودی باعث افزایش دقت می‌شود. استفاده از شبکه عصبی در تصاویر با اهداف خاص مانند بخشبندی تصویر [1]، فیلتر کردن، شناسایی الگو[2] و یا شناسایی شیئ خاص و غیره . امروزه جای روش‌های سنتی در تشخیص الگو را گرفته است (Egmont& Ridderb& Handelsc ,2001,2279-2301). اگر بخواهیم این روش را با دیگر الگوریتم‌ها نظیر فیلتر کالمن مقایسه کنیم، باید گفت که یافتن محل و تعقیب یک شیء در یک تصویر ویدئویی با نویز محیطی کم ، بوسیله فیلتر کالمن و شبکه MLP تقریباً نتایج مساوی از نظر دقت و سرعت محل یابی شیء به دنبال دارد.اما هنگامی که به تصویر نویز اضافه می‌شود، به دلیل این‌که شبکه عصبی قادر به یادگیری است ، با دقت بیشتری می‌تواند محل شیء را تشخیص و آن را دنبال کند. تحقیقات نشان داده است که خطای شبکه عصبی MLP در مقابل فیلتر کالمن با الگوریتم PVA[3] در یافتن محل یک شیء، ۵ در مقابل ۱۰

[1] Image segmentation
[2] Pattern Recognition
[3] Position,Velocity,Accelaration

است، از نظر فضای اشغال شده اگر بخواهی م شبکه MLP را بر روی یک سخت افزار پیاده‌سازی کنیم حافظه بسیار کمتری مورد نیاز است . به عبارت دیگر می توان گفت که شبکه MLP کمترین فضا را در مقابل فیلتر کالمن اشغال می‌کند. ۲۸۰ بایت در مقابل ۱۳٤٤ بایت (Shareef&Etal.,2007,323-330).

شبکه MLP به خاطر سادگی و سرعت مناسب در پردازش و یادگیری (Aeinfar& Etal.,2009,2075-2080) تبدیل به یکی از شبکه‌های پرکاربرد در زمینه شناسایی الگو است. یادگیری یک شبکه عصبی یکی از مزایای کاربردی در حل مسائل پیچیده است و این شبکه را قادر ساخته که در مقابل تغییرات ناگهانی شرایط از خود انعطاف نشان دهد . همین موضوع باعث شده که در مسائلی مانند تشخیص یک جسم خاص مانند دست انسان، اگر شرایط محیطی مانند نویز به محیط اضافه شود بازهم شبکه بتواند با دقت قابل قبولی عمل کند . تحقیقات نشان داده است که دقت شبکه عصبی در حالتی که شرایط محیط و داده‌های اولیه تغییر کرده است در حدود ۹۵٪ است(Maung, 2009,466-470). در صورتی که بدون استفاده از شبکه عصبی برای حذف نویز باید از فیلترهای گوناگون نظیر فیلتر های مورفولوژیکی[1] جهت تعقیب بلادرنگ و کاهش خطا استفاده شود (Krips& Lammert& Kummert,2002, 313-317). در بیشتر تحقیقات قبلی تمرکز بیشتر بر روی انتخاب ویژگی‌ها و دقت شبکه است و کمتر به کارگیری شبکه بر روی تصویر به صورت بلادرنگ توجه شده است . اما پیاده‌سازی این مهم مستلزم به‌کارگیری بهترین ویژگی‌ها و در عین حال محدود جهت بالا بردن سرعت است. اما ابتدا باید دست را در تصویر آشکار کرد تا بتوان آن را ردیابی کرد. استفاده از مدل مارکوف [2] و بوجود آوردن فضای رنگ متفاوت (Chenaoua&Bouridane, 2006, 2673-2676) می‌تواند پوست را آشکارسازی کند ، اما این الگوریتم با ابعاد بزرگتر دچار مشکل می‌شود و اگر دست را نزدیک به دور بین کنیم ممکن است خطا ای جاد شود(Araban&Farokhi&Kangarloo, 2011,547-552) و از دقت بالایی نیز برخوردار نیست. اما می‌توان از یک شبکه عصبی به همراه یک پایگاه داده و ویژگی‌های مناسب مانند انتخاب فضاهای رنگ مناسب برای تشخیص پوست استفاده کرد.

[1] Morphological Filtering
[2] Markov Model

۱-۲-۲ اهداف، اهمیت و ضرورت تحقیق

شبکه های عصبی از دو ویژگی اساسی یادگیری یا نگاشت پذیری بر اساس ارائه داده های تجربی قدرت و توانایی تعمیم پذیری و ساختارپذیری موازی برخوردار می باشند.شبکه های عصبی مصنوعی ابزاری مناسب برای حل مسائل پیچیده در شبیه سازی های نرم افزاری می باشند. برای رسیدن به کارایی مطلوب کاربردهای واقعی، بخصوص مواردی که مستلزم آموزش پیوسته در زمان واقعی هستند نیاز است که طراحان به پیاده سازی سخت افزاری شبکه های عصبی روی آورند تا کنون بیشتر توجه معطوف به شبیه سازی نرم افزاری بوده در حالیکه توجه کمتری به آثار پیاده سازی سخت افزاری شده است.برای رسیدن به کارایی و تراکم بالا جنبه های گوناگونی از شرایط سخت افزاری باید مورد تحلیل و بررسی قرار گیرد. امروزه بسیاری از مسائل دنیای واقعی به شبکه های عصبی پردازشگر سریع برای رسیدن به یک راه حل در زمان واقعی نیاز دارند. در این راستا تشخیص دست به عنوا ن یکی از عنوان های تحقیق کاربردی مورد علاقه بسیاری از محققان و صنعت گران از دیرباز بوده است . از لحاظ عملی می توان کاربردهای فراوانی را برای آن ذکر کرد از جمله : تکنولوژی های برتر دنیا که motion capture نام دارند که در آن یک کاراکتر انیمیشنی قادر است حرکات دست انسان را تقلید کند. امروزه این سیستم جهت ساخت فیلم ها و بازی های کامپیوتری مورد استفاده قرار می گیرد . تشخیص زبان اشاره ایرلندی از ویدیوی زمان حقیقی . زبان اشاره زبانی است که برای انتقال مفهوم به جای انتقال الگوهای صوتی، از الگوهای اشاره استفاده می کند. زبان‌های اشاره، زبان‌هایی مشخص در نوع خودشان و با لغات و گرامر خودشان هستند . یک سیستم تشخیص اشارات دست، مبتنی بر رایانه، دارای کاربرد وسیعی برای جامعه ناشنوایان می باشد .یکی از مشکلات در چنین سیستم هایی، این است که دست یک شیء بند بند ، با بیش از ۲۸ درجه آزادی است. این سیستم‌ها باید برای کاربر های مختلف، با اندازه های متفاوت دست، رنگ‌های پوست گوناگون، تحت شرایط متغیر روشنایی، و در حالی که اشارات به صورت دقیق انجام نمی شوند، قادر به تشخیص اشارات دست باشند .سیستم تشخیص حرکات دست برای جایگزینی موس. کنترل بازی‌ه ای سه بعدی رایانه‌ای با اشاره های دست.کنترل تلویزیون با اشاره ی دست، به‌منظور جایگزینی ریموت کنترل . اما شاید یکی از بهترین کاربرد ها ساخت یک ربات ویلچر که با حرکات دست کار می کند جهت استفاده معلولان باشد.

با بررسی تاریخچه تشخیص و ردیابی دست انسان و تحقیقات گسترده انجام شده در این راستا و به کارگیری انواع روشها و الگوریتم ها و فیلتر های مختلف و پیاده سازی این الگوریتم ها روی پردازنده های مختلف که همه و همه در راستای بهینه سازی نتایج, خطای کم,دقت و سرعت بالا هستند این نتیجه حاصل می شود که به کار گیری دو روش پیشنهاد شده در قسمت فرضیه می تواند نقش بسزایی در بهبود نتایج و خطای خروجی داشته باشد . بنابراین انتظار می رود که در این پایان نامه بتوان دست را در هر محیطی که با نویز های مختلف آمیخته است تشخیص داد و با سرعت قابل قبولی که با انتخاب مناسب تراشه FPGA بدست می آید آن را ردیابی کرد.

2-1 تحقیقات مرتبط انجام شده

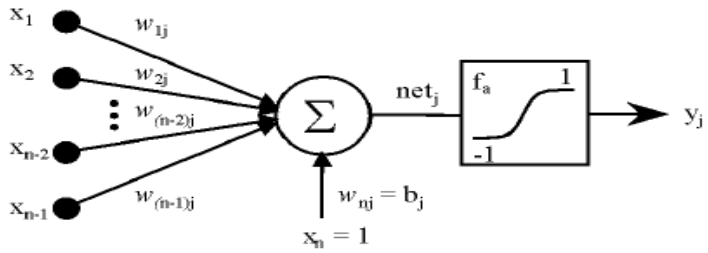
- پیاده‌سازی شبکه عصبی بر روی FPGA جهت ردیابی بلادرنگ دست

FPGA Implementation of a Neural Network for a Real-Time Hand Tracking System

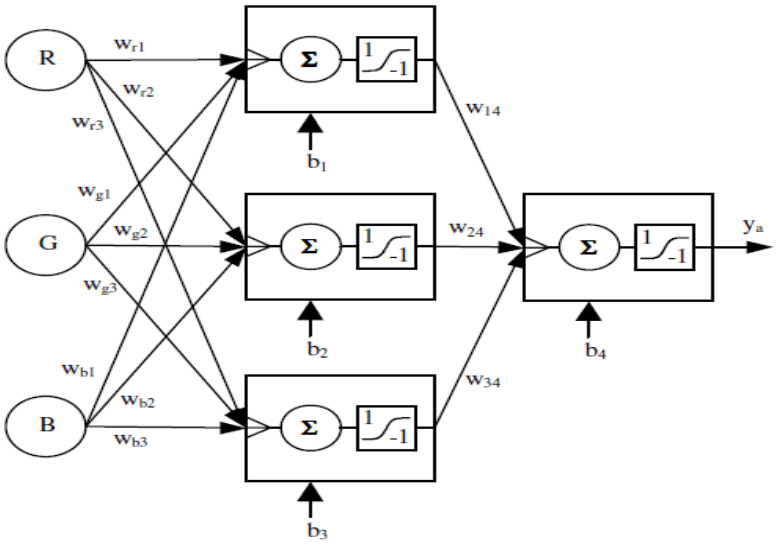
در این مقاله (Lbid, 313-317) همانطور که از شکل ۱-۳ مشخص است از یک شبکه عصبی پرسپترون چند لایه بر اساس الگوریتم پس انتشار[1] جهت تفکیک دست از زمینه استفاده شده است. به این صورت که اگر خروجی یک شد دست و اگر صفر شد زمینه است . در شکل ۱-٤ مشاهده می‌شود که بردارهای ورودی که همان رنگ پیکسل ها هستند (در واقع الگوریتم تشخیص بر اساس تک پیکسل است) به شبکه عصبی داده می شود. مقادیر b_1, b_2 و b_3 مقادیر بایاس هستند که همان یک در نظر گرفته می شود. در مدل ریاضی، نرون j از رابطه $net_{\mathrm{j}=} \sum_{i=1}^{n} x_i w_{ij}$ به دست می آید و $y_i = f_a(net_i)$ به طوری که x_iها بر دار ورودی و w_{ij}ها وزن‌هایی هستند که از نرون i به j می‌روند. f_a تابع تحریک سیستم‌است که یک تابع غیر خطی هایپر بولیک است و در ادامه خواهیم دید که توسط LUT[2] پیاده‌سازی می‌شود. y_i نیز خروجی یک نرون است . برای پیاده‌سازی بر روی FPGA از زبان توصیف سخت‌افزار یا همان VHDL استفاده شده است. البته کد ابتدایی در مطلب نوشته شده است سپس توسط خود مطلب به VHDL تبدیل شده است. وزن‌ها و بایاس‌ها به متغیر Integer تبدیل شده‌اند

[1] Back Propagation
[2] Lookup Table

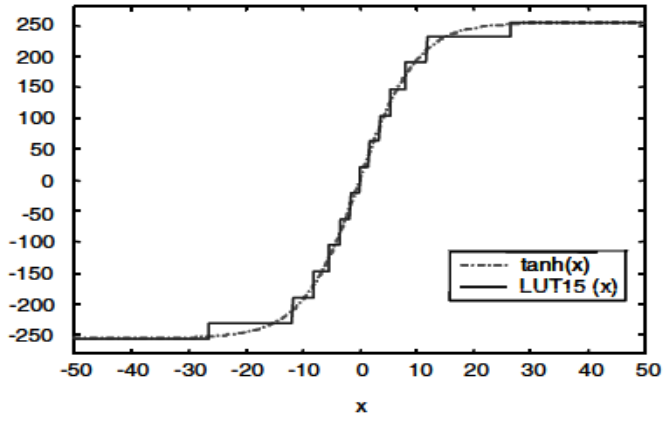
و به طور مستقیم در VHDL به کار برده شده اند. برای تابع تحریک همانطور که گفته شد از تابع tanh استفاده شده که به دلیل غیرخطی بودن تقریب تابع LUT به کار رفته است. سیگنال ورودی به یک بایت تبدیل می‌شود و به تابع داده می‌شود. همانطور که در شکل ۱-۵ مشاهده می‌شود با ۱۵ مرحله تقریب تابع tanh ساخته می‌شود.



شکل ۱-۳ مدل ریاضی شبکه عصبی MLP استفاده شده در تحقیق اول.



شکل ۱-٤ بردارهای ورودی (پیکسل ها) در تحقیق اول



شکل ۱-۵ تقریب تابع tanh با LUT