روشـــی جدید برای ارتباط بین انسان و ماشین

بر پایه پردازش تصویـــر و کاربرد آن در توسعه یک سامانه اطلاع‌رسانی کاربردی

سهیل اخلاقی امیری/ دانشکده مهندسی/ دانشگاه فردوسی مشهد
مشهد/ایران/ s.akhlaghiamiri@gmail.com
مهدی سعادتمند طرزجان/ مرکز پژوهشی مهندسی پزشکی/دانشکده مهندسی
دانشگاه فردوسی مشهد/ مشهد/ ایران/ saadatmand@um.ac.ir

در ایـن مقالـه، یک سیسـتم جدید بـرای ارتباط بین ماشـین و انسـان (HMI) بر مبنای پردازش تصویر ارائه شـده اسـت. در سیسـتم پیشـنهادی، فرآمین مورد نظر کاربر براساس موقعیت قرارگیری وی در مقابل دوربین اجرا می‌شـود. به این ترتیب که یک صفحه شاخص شامل تعدادی دکمه مجازی در مقابل دوربین قرار گرفته و سیسـتم پیشـنهادی با پردازش شـکل دکمه‌های مجـازی، موقعیـت کاربر بـر روی صفحه شـاخص را تشـخیص می‌دهد. همچنین، در این مقاله، برای تشـخیص شـکل دکمه‌ها، یک الگوریتم تشـخیص شـکل جدید بر پایه پـردازش تغییرات زاویه لبه ارائه شده است.
الگوریتم پیشـنهادی در برابر انتقال و تغییرات نور صحنه مقاوم می‌باشـد. از روش HMI پیشـنهادی برای توسـعه یک سـامانه اطلاع‌رسـانی کاربردی در محیط Microsoft Power Point استفاده شـده اسـت؛ به طوری که کاربر با ایسـتادن بـر روی دکمه‌های مجازی مناسب می‌تواند چهار فرمان مختلف شامل نمایش فایل بعدی، نمایش فایل قبلی، نمایش اسـلاید بعدی و نمایش اسلاید قبلی را در محیط مذکور اجرا نماید. نتایج تجربی بیانگر عملکرد مناسـب و کارایی سیستم پیشنهادی در شـرایط نوری مختلف می‌باشد.

کلید واژه: سیستم واسط بین انسان و ماشین (HMI)، تشخیص شکل، بخش‌بندی تصویر، فیلتر هومومورفیک

۱- مقدمه

وظیفه یک سیستم واسط بین انسان و ماشین[1] (HMI) انتقال فرآمین کاربر به ماشـین می‌باشـد. از این منظر، HMI-ها گسـتره وسیعی از سیستم‌ها از صفحه کلید ساده کامپیوتر تا دسته هدایت پرواز در کابین خلبان را شامل می‌شوند.
در سال‌های اخیر، محققین تلاش گسترده‌ای در زمینه توسعه HMI-هـای زیسـتی و فیزیولوژیک داشـته‌اند؛ که از آن جملـه می‌توان به سیستم‌های مبتنی بر پردازش سیگنال EEG [۱] اشاره نمود.
وجه مشترک سیستم‌های اخیر در استفاده حداقلی از تجیزات جانبی برای درک فرامین کاربر می‌باشـد. در این میان، اسـتفاده از پردازش تصویر برای تشـخیص وضعیت بدن [۲، ۳]، وضعیت چشـم[۴] و یا حالت لب [۵] کاربر مورد توجه بوده است.
به طور کلی، ارزیابی حرکات انسـان با اسـتفاده از پـردازش تصاویر ویدیویی می‌تواند منجر به صرفه‌جویی در زمان و هزینه گردد[۶].
در این مقاله، برای برقرار ارتباطی هوشـمندانه بین انسـان و ماشین، سیسـتمی مبتنی بر بینایی کامپیوتری ارائه شده است. قبلا محققین در کارهای دیگری نیز از بینایی ماشـین برای برقرار ارتباط با ماشین سـود جسته‌اند. به عنوان مثال، Fukumoto و همکاران [۲] با استفاده از محاسـبه جهت قرارگیری انگشتان و دست توانسته‌اند رابطی میان انسان و کامپیوتر ایجاد نمایند.
همچنیـن Wren و همـکاران [۳] از مدلی دو بعدی بدن و دسـت‌ها برای این منظور استفاده نموده‌اند. در سیستم پیشنهادی، با تشخیص موقعیت ایستادن کاربر در مقابل دوربین دستور مورد نظر وی تشخیص داده می‌شـود. به این ترتیب که بر روی زمین و در مقابل دوربین یک صفحه شاخص (شکل ۱ را ببینید) نصب می‌گردد.
بر روی صفحه شـاخص تعداد پنج دکمه مجازی برای چهار دسـتور مجزا در نظر گرفته شـده اسـت (یک دکمه بی‌اثر می‌باشـد). با قرار گرفتـن کاربر بر روی هر دکمه سیسـتم به صورت خودکار دسـتور متناظـر با آن دکمه را اجرا خواهد نمود. ما از روش پیشـنهادی برای توسعه یک سـامانه اطلاع‌رسـانی در محیط Microsoft Power Point اسـتفاده کرده‌ایم. در این سیسـتم، کاربر می‌تواند بـا انتخاب دکمه

شکل ۱- نمایی از سیستم HMI پیشنهادی شامل صفحه شاخص و دوربین.

مجازی مناسب فایل مورد نظر خود را در نرم افزار مربوطه بارگذاری نموده و یا اسلایدهای آن را به دلخواه نمایش دهد. کاربری سیستم پیشنهادی ساده و جذاب بوده و نیازمند بکارگیری تجهیزات خاص (مانند صفحه کلید، دسته یا ماوس) نمی‌باشد.

در سیستم پیشنهادی، برای تشخیص دکمه‌های مجازی (رسم شده بر روی صفحه شاخص) به صورت مقاوم در برابر تغییرات نور محیط و جابجایی، از یک الگوریتم جدید بر پایه ویژگی شکل استفاده شده است. به طور کلی، تشخیص شکل یکی از مباحث مهم در حوزه پردازش تصویر می‌باشد. معمولا روش‌های تشخیص شکل از ویژگی‌های سطح پایین مانند روشنایی پیکسل‌ها و یا ویژگی‌های سطح بالاتر از قبیل مرز [۷]، خطوط [۸]، گوشه‌ها [۹] و ناحیه [۱۰] استفاده می‌نمایند. به عنوان مثال، Anuar و همکاران [۱۰] از ممان‌های زرنیک برای توصیف شکل استفاده نموده‌اند. ممان‌های زرنیک نسبت به دوران نامتغیر بوده اما وابسته به تغییر مقیاس و انتقال می‌باشند.

Wei و همکاران [۱۱] با نگاشت تصویر ورودی بر دایره‌ای با شعاع مشخص، روشی برای حل مشکل فوق پیشنهاد نموده‌اند.

در کار دیگری، Wee و همکاران [۱۲] از انتقال تمام نواحی تصویر به فضای قطبی-شعاعی برای بهبود عملکرد ممان‌های زرنیک استفاده نموده‌اند. همچنین، تحقیقات متعددی نیز بر مبنای ویژگی‌های سطح بالا برای توصیف شکل انجام شده است.

به عنوان مثال، Lu و همکاران [۱۳] از اجزای مرز شی مطلوب برای ایجاد بردار ویژگی به منظور توصیف شکل استفاده نموده‌اند. برخی محققین نیز برای توصیف شکل، ارتباط بین پاره‌خط‌های لبه را مورد توجه قرار داده‌اند. به عنوان مثال، Prakash و همکاران [۱۴] از روش فوق برای تشخیص گوش در تصویر نیم‌رخ انسان استفاده نموده‌اند.

در سیستم ارائه شده در این مقاله، ابتدا تصویر شاخص به صورت دستی، به پنج ناحیه تجزیه می‌شود تا موقعیت کلی کلیدهای مجازی در تصویر مشخص گردد. سپس، در هر ناحیه بردار ویژگی پیشنهادی برای توصیف شکل استخراج می‌گردد. به عبارت دقیقتر، برای هر پیکسل لبه (با اندازه گرادیان قابل قبول)، پیکسل‌های لبه در یک همسایگی مشخص انتخاب شده و زاویه پاره‌خط‌های مرتبط کننده آنها به صورت مرتب در یک بردار ویژگی قرار می‌گیرند. سپس، بردار ویژگی نهایی به عنوان توصیفی از شکل دکمه مجازی ذخیره می‌گردد. در حین اجرا، بردار ویژگی فوق به روش مشابه در هر فریم دوباره محاسبه شده و با بردار ویژگی مبنا (ذخیره شده) مقایسه می‌گردد. در صورتی که اختلاف آنها از یک حد مشخص بیشتر باشد، به منزله آن است که کاربر بر روی دکمه مذکور ایستاده و در نتیجه فرمان مربوط به آن دکمه، اجرا خواهد شد.

در ادامه مقاله، در بخش (۲)، الگوریتم پیش‌پردازشی مورد استفاده برای ایجاد تصویر لبه شرح داده خواهد شد. بخش (۳) به شرح الگوریتم تشخیص شکل پیشنهادی اختصاص یافته است. در بخش (۴)، نتایج تجربی ذکر گردیده و در نهایت، بخش (۵) به نتیجه‌گیری می‌پردازد.

۲- پیش‌پردازش

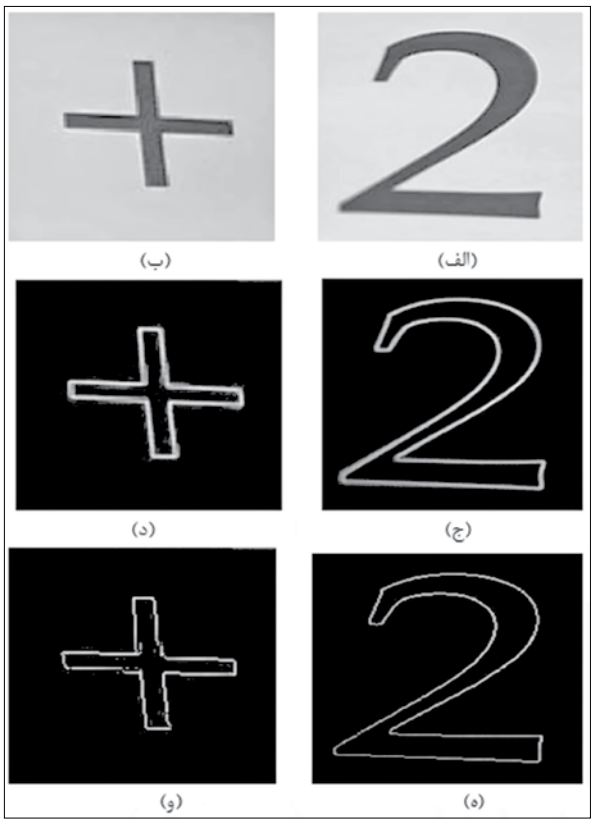
می‌دانیم که لبه‌ها مهمترین ویژگی سطح پایین توصیف کننده شکل در تصویر می‌باشند [۱۵]. لذا، با پردازش لبه‌ها می‌توان به توصیف مناسبی از شکل دست یافت. برای ایجاد تصویر لبه، ابتدا با استفاده از فیلتر گوسی نویز تصویر تضعیف می‌شود.

برای محاسبه گرادیان تصویر از فیلتر Sobel استفاده شده است. به طور کلی، اندازه و زاویه بردار گرادیان در هر پیکسل از تصویر $g(x,y):R2\rightarrow R$، به ترتیب، مطابق معادله‌های زیر محاسبه می‌گردد:

$$\|\nabla g(x,y)\| = \sqrt{\left(\frac{\partial g}{\partial x}\right)^2 + \left(\frac{\partial g}{\partial y}\right)^2} \quad (۱)$$

$$\varphi(x,y) = \tan^{-1}\left(\frac{g_x}{g_y}\right) \quad (۲)$$

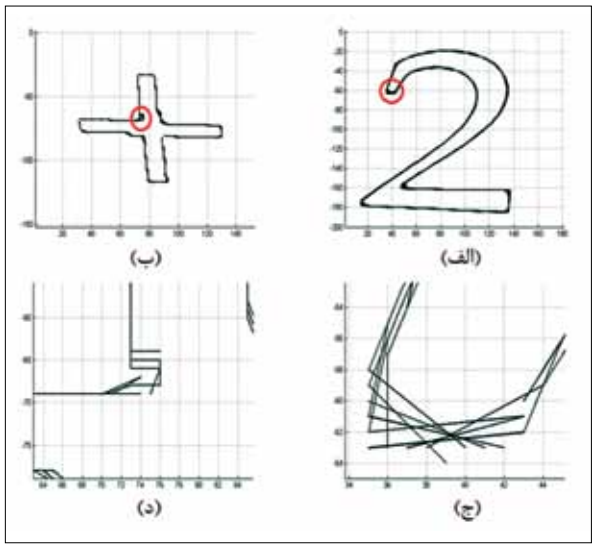
به عنوان مثال، در شکل (۲)، اندازه گرادیان برای تصویری از صفحه شاخص نشان داده شده است. همانطور که مشاهده می‌شود، لبه‌ها



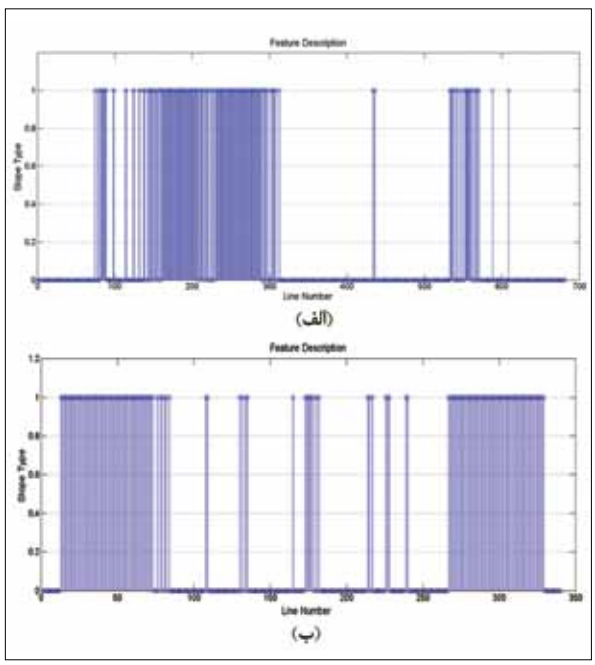
شکل ۲- (الف)، (ب) تصویر دو دکمه مجازی در صفحه شاخص و (ج)، (د) تصاویر اندازه گرادیان حاصل از اعمال فیلتر سوبل بر آنها و (ه)، (و) تصاویر نازکسازی شده متناظر با استفاده از روش NMS.

لزوما در تصویر گرادیان تیز نمی‌باشند (یعنی ضخامت آنها بیش از یک پیکسل است). این افزونگی اطلاعات می‌تواند منجر به کاهش دقت الگوریتم توصیف شکل بر مبنای اطلاعات لبه گردد.

برای نازک‌سازی گرادیان در محل لبه‌ها از روش حذف مقادیر غیرماکزیمم[۲] (NMS) [۱۶] استفاده نموده‌ایم. به این ترتیب که، در تصویر گرادیان، هر پیکسل که اندازه گرادیان آن از هر دو همسایه آن در راستای بردار گرادیان (معادله ۲) بیشتر نباشد، صفر می‌گردد. شکل‌های (۲-ه) و (۲-و) نتیجه اعمال پردازش فوق را، به ترتیب، بر تصاویر گرادیان شکل‌های (۲-ج) و (۲-د) نشان می‌دهند.



شکل ۳- (الف)، (ب) پاره‌خط‌های حاصل از اعمال روش پیشنهادی بر دکمه‌های مجازی نشان داده شده در به ترتیب شکل‌های (۲-الف) و (۲-ب) و (ج)، (د) بزرگنمایی پاره‌خط‌های حاصل در نواحی مشخص شده با دایره در بخش‌های (الف) و (ب).



شکل ۴- (الف)، (ب) بردارهای ویژگی بدست آمده برای دکمه‌های مجازی نشان داده شده در، به ترتیب، شکل‌های (۲-الف) و (۲-ب).

در نهایت، تصویر گرادیان نازک‌سازی با استفاده از روش Otsu [۱۷] آستانه‌گذاری می‌گردد تا تصویر باینری لبه بدست آید.

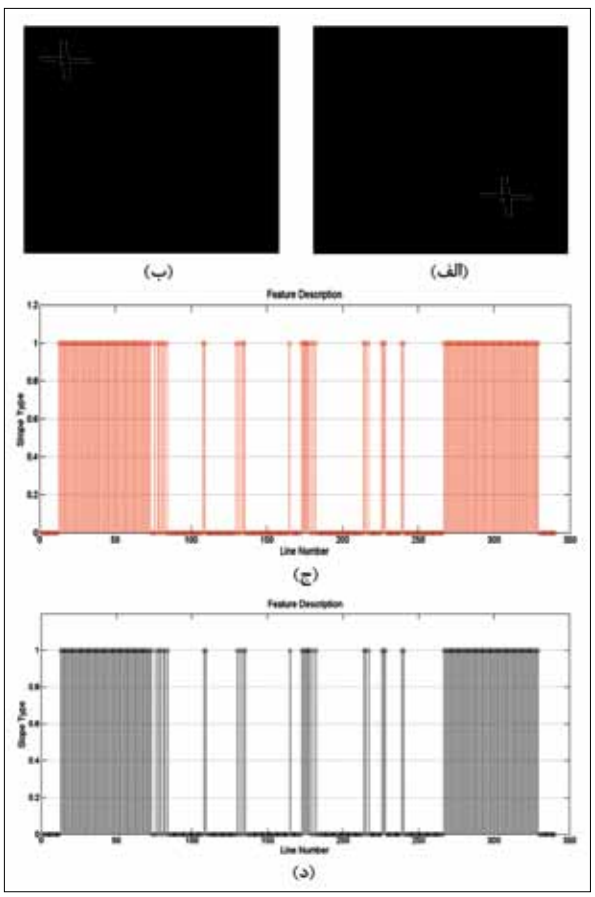
۳- الگوریتم تشخیص شکل پیشنهادی

بردار ویژگی پیشنهادی برای تشخیص شکل، در حقیقت توصیفگر تغییرات شیب لبه‌ها در محدوده شکل مورد نظر می‌باشد. برای این منظور، محدوده کاندیدای شکل (مثلا در اینجا محدوده مشخص شده برای هر دکمه مجازی بر روی صفحه شاخص) در تصویر باینری لبه، پیکسل به پیکسل از بالا به پایین و از چپ به راست اسکن می‌شود تا زوج پیکسل‌های لبه همسایه (با مقدار غیرصفر در تصویر باینری لبه) با فاصله اقلیدوسی r مشخص گردند (در اینجا به صورت تجربی r=5 انتخاب شده است).

برای هر زوج پیکسل لبه، شیب پاره‌خط متصل کننده آنها به یکدیگر محاسبه شده و به ترتیب جستجو در بردار ویژگی پیشنهادی قرار می‌گیرد. توجه شود که برای هر زوج پیکسل همسایه، تنها یک شیب به بردار ویژگی اضافه می‌شود.

برای ساده‌سازی بیشتر، شیب‌های در محدوده ۴۵- تا ۴۵ درجه با مقدار ۱ و سایر شیب‌ها با مقدار صفر مشخص می‌گردند. لذا، بردار ویژگی پیشنهادی تنها شامل مقادیر صفر و یک خواهد بود.

به عنوان مثال، در شکل‌های (۳-الف) و (۳-ب)، پاره‌خط‌هایی حاصل برای دکمه‌های مجازی مربوط به شکل‌های (۲-الف) و (۲-ب) نشان داده شده است. به‌علاوه، بردارهای ویژگی حاصل به ترتیب در



شکل ۵- (الف)، (ب) تصویر دکمه مجازی + در دو موقعیت مختلف و (ج)، (د) بردارهای ویژگی متناظر با آنها.

شکل‌های (۴-الف) و (۴-ب) نشان داده شده‌اند.
پس از محاسبه بردار ویژگی پیشنهادی برای شکل مورد نظر، به منظور تشخیص شکل در تصویر از معیار فاصله مجموع قدر مطلق اختلاف[۳] استفاده می‌شود.
به عبارت دقیقتر، اگر بردار ویژگی مبنا f_r (با طول n_r) و بردار ویژگی مورد ارزیابی f_i (با طول n_i) باشد، اختلاف دو بردار ویژگی مطابق معادله زیر محاسبه می‌گردد:

$$D(f_r, f_i) = \frac{1}{n_r}\left(|n_r - n_i| + \sum_{k=1}^{\min(n_r,n_i)} |f_{r,k} - f_{i,k}|\right) \quad (۳)$$

که در آن، $f_{r,k}$ بیانگر عنصر k-ام بردار ویژگی f_r می‌باشد. همانطور که مشاهده می‌شود، اختلاف طول دو بردار ویژگی به عنوان خطا در تابع هزینه منظور شده است.
بدیهی است که وقتی بردارهای ویژگی f_i و f_r هم‌طول بوده و همه عناصر آنها مشابه باشند، مقدار تابع هزینه (۳) صفر خواهد بود. چنانچه $D(f_r, f_i)$ از مقدار آستانه θ بزرگتر باشد، دو بردار ویژگی متفاوت تلقی شده و در سیستم HMI پیشنهادی فرمان مربوط به کلید مجازی مورد نظر اجرا خواهد شد.

۳-۱- عدم تغییر با انتقال

با توجه به اینکه بردار ویژگی شامل زاویه شیب زوج پیکسل‌های همسایه می‌باشد، می‌توان انتظار داشت که با جابجایی شی در تصویر، بردار ویژگی پیشنهادی تغییر چندانی نداشته باشد. این خصوصیت در شکل (۵) با ذکر مثال نشان داده شده است. همانطور که مشاهده می‌شود، تصویربرداری از دکمه مجازی + در دو موقعیت مختلف، منجر به بردارهای ویژگی یکسان شده است.

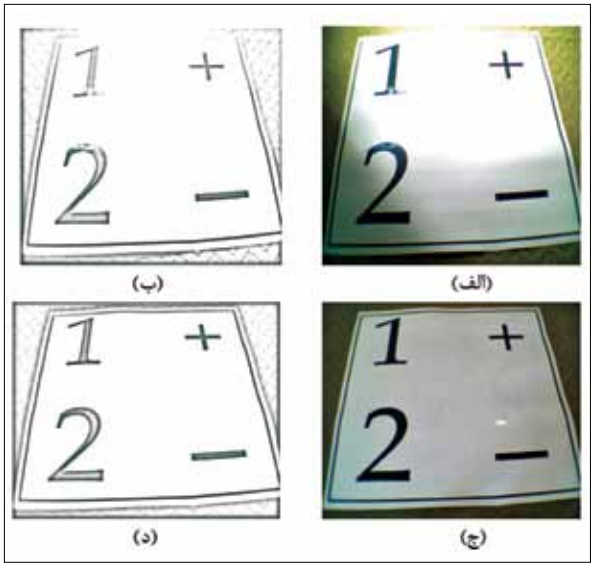
۳-۲- عدم تغییر با تغییر نور صحنه

با توجه به اینکه بردار ویژگی پیشنهادی در حقیقت توصیفگر تغییرات زاویه شیب محلی لبه‌های شکل می‌باشد، می‌توان انتظار داشت که در برابر تغییرات نور محیط مقاومت خوبی داشته باشد. بویژه آنکه تغییرات نور صحنه معمولا منجر به تغییر شکل لبه‌ها نمی‌شوند. با وجود این، ما برای بهبود عملکرد الگوریتم پیشنهادی در برابر تغییرات روشنایی صحنه (بویژه اثر سایه)، از فیلتر هومومورفیک[۴] [۱۸] استفاده نموده‌ایم. بر این اساس، مدل روشنایی-انعکاس برای هر تصویر $f(x,y)$ به صورت زیر تعریف می‌شود:

$$f(x,y) = i(x,y)r(x,y) \quad (۴)$$

که در آن، $i(x,y)$ نور تابشی از منبع و $r(x,y)$ ضریب انعکاس صحنه است. با توجه به اینکه تغییرات نور منبع در محدوده دید دوربین بسیار کم می‌باشد لذا، $i(x,y)$ تابعی با باند فرکانسی بسیار محدود می‌باشد. فیلتر هومومورفیک برای حذف تغییرات روشنایی منبع، ابتدا لگاریتم تصویر را محاسبه نموده تا ترکیب مولفه‌های مربوط به تابش منبع و بازتابش صحنه به صورت خطی درآید:

$$\ln f(x,y) = \ln i(x,y) + \ln r(x,y) \quad (۵)$$



شکل ۶- (الف)، (ج) دو تصویر با شرایط نوری مختلف از صفحه شاخص و (ب)، (د) نتیجه اعمال فیلتر هومومورفیک بر آنها.

سپس، با حذف اطلاعات باند پایین تصویر لگاریتمی عملا $i(x,y)$ به شدت تضعیف می‌گردد. ما برای این منظور، از فیلتر DoG[۵] استفاده نموده‌ایم:

$$g(x,y) = \ln f(x,y) - G_\sigma(x,y) * \ln f(x,y) \quad (۶)$$

که در آن $G\sigma$ کرنل گوسی با انحراف معیار σ است. اکنون، با اعمال تابع نمایی، تصویر از حالت لگاریتمی خارج می‌شود:

$$\hat{f}(x,y) = \exp(g(x,y)) \quad (۷)$$

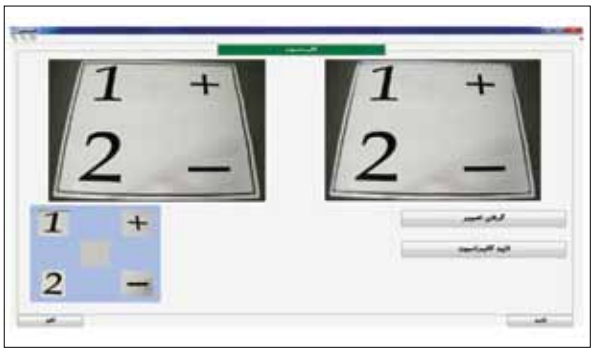
اکنون برای محاسبه بردار ویژگی به جای تصویر f از تصویر بهبود یافته $\hat{f}$ استفاده می‌شود.
در شکل (۶) نتیجه اعمال پردازش فوق بر دو تصویر شاخص با شرایط نوری مختلف نشان داده شده است. همانطور که مشاهده می‌شود اثر سایه در تصاویر حاصل از فیلتر هومومورفیک حذف شده است.

۴- نتایج تجربی

در این مقاله، کلیه نتایج تجربی با استفاده از یک کامپیوتر شخصی با پردازنده AMD Athlon(tm) 64 Processor 3800 2.4GHz و حافظه 3GB انجام شده است.
الگوریتم تشخیص شکل پیشنهادی در عین سادگی بسیار کارا نیز می‌باشد؛ به طوری که زمان پردازشی در محیط برنامه‌نویسی MAT-LAB برای هر فریم ۱٫۰۷ ثانیه بوده است. از سیستم HMI پیشنهادی برای توسعه یک سامانه اطلاع‌رسانی کاربردی استفاده شده است.
در شکل (۷)، صفحه اصلی نرم‌افزار تهیه شده برای این سامانه نشان داده شده است.

شکل ۷- صفحه اصلی نرم افزار تجاری توسعه داده شده

پس از اجرای نرم‌افزار و قرارگیری صفحه شاخص و دوربین در موقعیت مورد نظر، کاربر باید با استفاده از پنجره کالیبراسیون، موقعیت صفحه شاخص را برای نرم‌افزار مشخص نماید (شکل ۸).



شکل ۸- پنجره کالیبراسیون

نرم‌افزار به صورت خودکار موقعیت دکمه‌های مجازی را تشخیص داده و بردار ویژگی مبنا برای هر یک را محاسبه و ذخیره‌سازی می‌کند. سپس، کاربر باید فرآمین مربوط به هر یک از دکمه‌های مجازی را مشخص کند (شکل ۹).

شکل ۹- صفحه تنظیم فرامین مربوط به دکمه‌های مجازی

این فرآمین شامل نمایش فایل بعدی، نمایش فایل قبلی، نمایش اسلاید بعدی در فایل فعلی و نمایش اسلاید قبلی در فایل فعلی در محیط Microsoft Power Point می‌باشند. تنظیمات فوق، تنها یکبار انجام شده و تا زمانی که موقعیت صفحه شاخص و دوربین بیش از حد مجاز تغییر نکرده باشد، معتبر خواهند بود.

اکنون، با اجرای پنجره نمایش انتخابگر و فشردن دکمه شروع محیط Microsoft Power Point با استفاده از فن‌آوری ActiveX اجرا شده و اولین فایل در آن بارگذاری می‌شود. در این حالت، با قرارگیری کاربر بر روی هر کدام از دکمه‌های مجازی صفحه شاخص، فرمان مربوط به آن دکمه در محیط Microsoft Power Point اجرا خواهد گردید. نتایج تجربی بیانگر عملکرد بسیار مناسب سامانه اطلاع‌رسانی مذکور در شرایط نوری مختلف می‌باشد. به علاوه، کار با سیستم برای کاربران بسیار جذاب بوده و ارزیابی‌ها بیانگر رضایت نسبی آنها است.

۵- نتیجه‌گیری

در این مقاله سیستمی برای ایجاد ارتباط هوشمند میان انسان و کامپیوتر بر مبنای پردازش تصویر ارایه گردیده است. سیستم پیشنهادی قادر است با تشخیص موقعیت قرارگیری کاربر در یک صفحه شاخص، فرآمین مورد نظر وی را اجرا نماید. برای تشخیص کلیدهای مجازی موجود بر روی صفحه شاخص نیز یک الگوریتم تشخیص شکل جدید ارائه گردید. روش پیشنهادی علیرغم سادگی نسبت به تغییرات نور و انتقال مقاوم می‌باشد. اگرچه از سیستم HMI پیشنهادی در این مقاله، برای توسعه یک سامانه اطلاع‌رسانی استفاده شده، اما دامنه کاربرد آن می‌تواند بسیار گسترده‌تر باشد و فرآمین بسیار متنوع‌تری را پوشش دهد.

قدردانی

بدین وسیله نویسندگان مقاله از شرکت بینایی ماشین و پردازش تصویر بیناتوس (سهامی خاص) مستقر در دانشگاه فردوسی مشهد به خاطر حمایت از این پروژه تشکر می‌نمایند.

***

پی نوشت‌ها

1. Human-machine interface (HMI)
2. Non-maximum suppression (NMS)
3. Sum absolute difference
4. Homomorphic filter
5. Difference of Gaussian (DoG)

مراجع

[1] G. E. Fabiani, D. J. McFarland, J. R. Wolpaw, and G. Pfurtscheller, "Conversion of EEG activity into cursor movement by a brain-computer interface (BCI)," IEEE Trans. Neural Systems and Rehabilitation Engineering, vol. 12, no. 3, pp 331-338, 2004.
[2] M. Fukumoto, Y. Suenaga, and K. Mase, "Finger-pointer: pointing interface by image processing," Computers & Graphics, vol. 18, no. 5, pp 633–642, 1994.
[3] C. R. Wren, A. Azarbayejani, T. Darrell; and A. P. Pentland, "Pfinder: real-time tracking of the human body," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 19, no 7, pp. 780-785, 1997.
[4] T. W. Victor, J. L. Harbluk, and J. A. Engström, "Sensitivity of eye-movement measures to in-vehicle task difficulty," Transportation Research Part F: Traffic Psychology and Behaviour, vol. 8, no. 2, pp. 167-190, 2005.
[5] R. Navarathna, P. Lucey, D. Dean, C. Fookes, and S. Sridharan, "Lip detection for audio-visual speech recognition in-car environment," Int'l Conf. Information Science, Signal Processing and their Applications, pp. 598-601, 2010.
[6] C. Wu and Y. Liu, "Development and evaluation of an ergonomic software package for predicting multiple-task human performance and mental workload in human–machine interface design and evaluation," Computers & Industrial Engineering, vol. 56, no. 1, pp. 323-333, 2009.
[7] N. Paragios and R. Deriche, "Geodesic active contours and level sets for the detection and tracking of moving objects," IEEE Trans. PAMI, vol. 22, no. 3, pp. 266-280, 2000.
[8] A. Mitchie and S. Seida, "Interpretation of structure and motion from line correspondence," IEEE Conf. Pattern Recognition, 1986.
[9] K. Skifstad and A. Jain, "Illumination independent change detection for real world sequence," Computer Vision, Graphics, and Image Processing, vol. 46, no. 3, pp. 387-399, 1989.
[10] F. M. Anuar, R. Setchi, and Y.-K. Lai, "Trademark image retrieval using an integrated shape descriptor," Expert Systems with Applications, vol. 40, no. 1, pp.105–121, 2013.
[11] C.-H. Wei, Y. Li, W.-Y. Chau, and C.-T. Li, "Trademark image retrieval using synthetic features for describing global shape and interior structure," Pattern Recognition, vol. 42, no. 3, pp.386–394, 2009.
[12] C.-Y. Wee and R. Paramesran, "On the computational aspects of Zernike moments," Image and Vision Computing, vol. 25, no. 6, pp. 967–980, 2007.
[13] C. E. Lu, N. Adluru, H. Ling, G. Zhu, and L. J. Latecki, "Contour based object detection using part bundles," Computer Vision and Image Understanding, vol. 114, no. 7, pp. 827–834, 2010.
[14] S. Prakash and P. Gupta, "A rotation and scale invariant technique for ear detection in 3D," Pattern Recognition Letters, vol. 33, no. 14, pp.1924-1931, 2012.
[15] M. Sonka, Image Processing, Analysis, and Machine Vision. Thompson Learning, 3rd ed., Toronto, 2008.
[16] C. Sun and P. Vallotton, "Fast linear feature detection using multiple directional non-maximum suppression," J. Microscopy, vol. 234, no. 2, pp. 147–157, 2009.
[17] N. Otsu, "A threshold selection method from gray-level histograms," IEEE Tran. Systems, Man, and Cybernetics, vol. 9, no. 1, pp. 62-66, 1979.
[18] R. C. Gonzalez and R. E. Woods, Digital Image Processing. Prentice-Hall, 2nd ed., 2003.