

شناسایی علائم ایستای دست با استفاده از ویژگی‌های مبتنی بر نمایش

صدیقه ایلدرآبادی محمد ابراهیمی مهدی یعقوبی حمیدرضا پوررضا
دانشگاه آزاد اسلامی فلاورجان دانشگاه آزاد اسلامی مجلسی دانشگاه آزاد اسلامی مشهد دانشگاه فردوسی مشهد
sedigheh_ildarabadi@yahoo.com hoo_ebrahimi@yahoo.com mmyaghobi@yahoo.com hpourreza@um.ac.ir

چکیده- در این مقاله برای شناسایی علائم از دسته‌بندی‌کننده‌های *MLP*، *SMO*، *BayseNet*، *MultiClassClassifier*، *RBFnetwork* و الگوریتم‌های *KNN*، *FKNN* و *FRNNC* استفاده شد. ویژگی‌های مبتنی بر نمایش و ویژگی‌های آماری از تصاویر استخراج شدند و هر کدام بطور مجزا به دسته‌بندی‌کننده‌ها داده شدند. در مقایسه دیده شد که ویژگی‌های مبتنی بر نمایش نسبت به ویژگی‌های آماری، نتایج بسیار بهتری داشت.

کلید واژه- بینایی ماشین، حالات و اشارات دست، شناسایی علائم ایستای زبان، ویژگی‌های آماری، ویژگی‌های مبتنی بر نمایش.

صورت و زبان، دهان و موقعیت بدن.

۱- مقدمه

در برخی سیستم‌ها، از طریق دستکش‌ها [۱] یا سیستم‌های دریافت حرکت [۲]، داده مستقیماً اندازه‌گیری می‌شود. این گونه روش‌های اندازه‌گیری، اطلاعات مکانی سه‌بعدی دست، انگشتان و دیگر قسمت‌های بدن را بخوبی فراهم می‌کنند. با این حال کاربر بایستی از یک وسیله پوششی استفاده کند که این وسیله حرکت کاربر را محدود می‌کند. همچنین این گونه سیستم‌ها، بدلیل تجهیزاتی که نیاز دارند، مقرون به صرفه نیستند.

در برخی از سیستم‌های مبتنی بر بینایی نیز برای کشف دست‌ها و انگشتان شخص علامت‌دهنده، از دستکش‌های رنگی و لباس‌ها با آستین بلند استفاده شده است. در برخی دیگر نیز محدودیت‌هایی مانند زمینه ایستا یا ثابت بودن موقعیت شخص علامت‌دهنده، در نظر گرفته شده است.

مسئله مهم دیگر این است که بانک‌داده مشترکی برای ارزیابی سیستم‌های شناسایی زبان علائم، وجود ندارد. تقریباً همه بانک‌های داده‌ای که توسط محققان استفاده می‌شود با یکدیگر فرق دارند و در دسترس دیگر محققان نیست. بعلاوه این بانک‌های داده در نوع زبان، اندازه فرهنگ لغت، قیود گرامری و علائم متفاوتند.

علائم زبان به دو دسته؛ علائم ایستا و علائم پیسته تقسیم می‌شوند مروری کلی درباره شناسایی زبان علائم، چگونگی

حالات و اشارات دست، شکلی از ارتباط غیر شفاهی انسان‌ها می‌باشد. اغلب این حالات و اشارات در سیستم‌ها بسته به کاربرد، معانی خاصی دارند. امروزه سیستم‌های قابل توجهی برای شناسایی حالات و اشارات، و زبان علائم، وجود دارد. دو دلیل اصلی برای تحقیق بر روی حالات و اشارات و شناسایی زبان علائم عبارتند از:

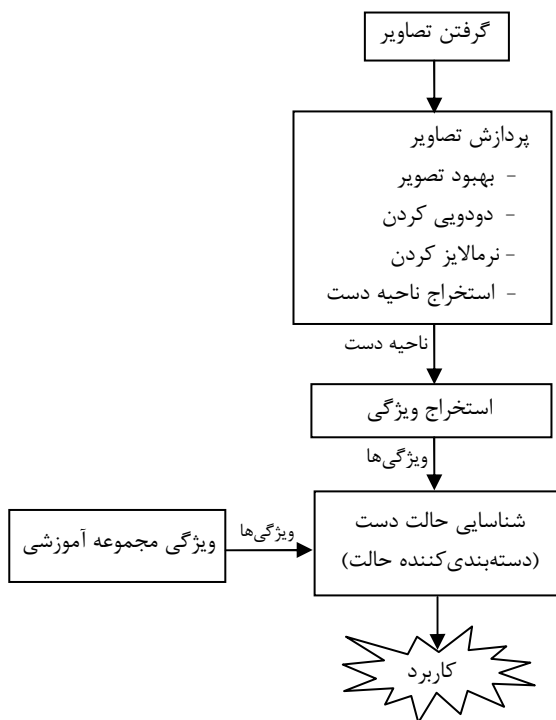
- ایجاد رابطی برای کمک به افراد ناشنوا
- ایجاد رابطی برای تعامل بین انسان و کامپیوتر

زبان علائم، زبانی طبیعی افراد ناشنوا می‌باشد. زبان علائم، زبانی است که در آن از حالات و اشارات، بجای صدا استفاده می‌شود. در این زبان برای رساندن معنا، ترکیبی از شکل دست، جهت و جابجایی دست‌ها، بازوها و بدن، سیمای چهره و شکل لب، استفاده می‌شود. زبان علائم، زبان بین-المللی نیست و حتی در نواحی مختلف از کشورها نیز ممکن است تغییر کند. زبان علائم، زبانی دیداری است و شامل سه جزء اصلی می‌باشد:

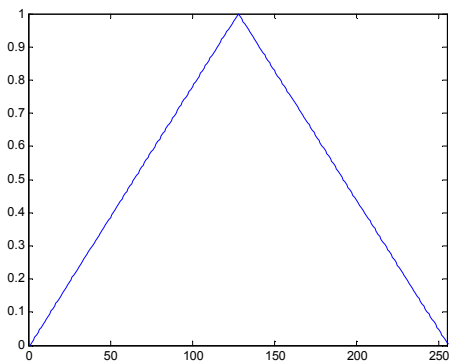
- هجی کردن با انگشتان: برای هجی کردن حرف به حرف کلمه استفاده می‌شود.
- فرهنگ لغتی از علائم: با هدف برقراری ارتباط راحت‌تر بکار می‌رود.
- ویژگی‌هایی که مربوط به دست نیستند: سیمای

تصویر داده شده انتخاب می‌کند. در این روش سعی می‌شود واریانس بین کلاسی سطوح خاکستری تصویر کمینه شود.

اکنون با استفاده از حد آستانه otsu تصویر دودویی می‌شود. زمینه تصویر مقادیر صفر و ناحیه دست محتوی مقادیر یک می‌باشد. براحتی با اعمال and منطقی تصویر دودویی حاصل با تصویر اولیه، ناحیه دست از تصویر اصلی، تقطیع می‌شود. در نهایت مستطیل محاطی دست با حذف نواحی زمینه بدست می‌آید. حال نواحی مربوط به دست را به ابعاد 32×32 نرمالایز می‌کنیم.



شکل ۱. چهار مرحله مهم در ساخت رابط کاربری برای شناسایی حالات و اشارات دست



شکل ۲. تابع عضویت مثلثی جهت تقریب روی نواحی دست (این تابع سطوح خاکستری بین صفر تا ۲۵۵ را به صفر تا یک می‌برد)

استخراج ویژگی‌ها و روش‌های دسته‌بندی در [۳] آمده است.

در این مقاله سیستمی مبتنی بر بینایی ماشین، برای شناسایی علائم ایستا، ارائه شده است. ابتدا نواحی دست با استفاده از آستانه‌یابی otsu و عملیات مورفولوژی استخراج شد، سپس نواحی دست از یک تابع عضویت عبور داده شد. این تابع عضویت سطوح خاکستری بین صفر تا ۲۵۵ را به بازه صفر تا یک می‌برد [۵۴]. ویژگی‌های آماری و ویژگی‌های مبتنی بر نمایش، از نواحی استخراج شده، محاسبه گردید. برای شناسایی علائم از دسته‌بندی‌کننده‌های MLP، MultiClassClassifier، BayseNet، SMO، K Nearest (KNN)، RBFnetwork و الگوریتم‌های (Fuzzy K Nearest Neighbor) FKNNC، (Neighbor Fuzzy Rough K Nearest) FRNNC، و (Classification Neighbor Classification) استفاده شد. در ضمن ویژگی‌های آماری از تصاویر استخراج و به سیستم جهت شناسایی علائم داده شد. در مقایسه دیده شد که ویژگی‌های مبتنی بر نمایش نسبت به ویژگی‌های آماری، نتایج بسیار بهتری داشت.

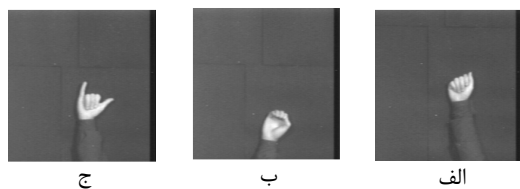
۲- فرایند شناسایی حالات و اشارات دست

بطور کلی چهار مرحله اساسی برای شناسایی حالات و اشارات ایستای دست مبتنی بر بینایی ماشین، وجود دارد. این چهار مرحله همچنان که در شکل (۱) دیده می‌شود؛ عبارتند از: گرفتن تصویر، پردازش تصویر، استخراج ویژگی، شناسایی حالت دست.

۲-۱- پیش‌پردازش تصویر

پیش‌پردازش تصویر شامل سه قسمت بهبود، دودویی کردن و تقطیع ناحیه دست از تصویر می‌باشد. عملیات مورفولوژی از جمله روش‌هایی است که برای حذف نویز از تصویر مورد استفاده قرار می‌گیرد. در اینجا از عملیات مورفولوژی گسترش و فرسایش جهت حذف نویز تصاویر استفاده شده است. برای دودویی کردن تصاویر نیاز به حد آستانه می‌باشد در اینجا از روش آستانه‌گیری otsu برای بدست آوردن حد آستانه‌ی مناسب استفاده شده است. روش دودویی کردن otsu با بیشینه کردن معیار جداکننده یعنی تفکیک پذیری کلاسه‌های ناشی از سطوح خاکستری، آستانه بهینه را برای

دارد. سه نمونه از تصاویر این پایگاه داده در شکل (۳) آمده است. (۳-الف) بیان کننده حرف a، (۳-ب) بیان کننده o و (۳-ج) حرف بیان کننده حرف y می‌باشد.

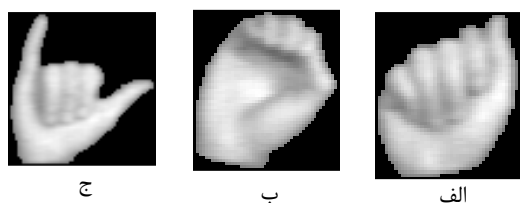


شکل ۳. سه نمونه از تصاویر پایگاه داده الف. حرف a، ب. حرف o، ج. حرف b.



شکل ۴. تصاویر دودویی بدست آمده با استفاده از روش آستانه‌گیری otsu. الف. حرف a، ب. حرف o، ج. حرف b

در ابتدا با استفاده عملیات گسترش و فرسایش از تصویر ورودی حذف نویز صورت می‌گیرد، سپس با استفاده از روش آستانه‌گیری otsu، حد آستانه مناسب بدست آمده و تصویر دودویی می‌شود. شکل (۴) تصاویر دودویی بدست آمده با استفاده از روش آستانه‌گیری otsu را نشان می‌دهد. پس از دودویی کردن تصویر عملیات تقطیع دست از بقیه تصویر صورت می‌گیرد. ویژگی‌های بیان شده در بخش ۲-۲ از ناحیه دست استخراج می‌شود. شکل (۵) سه نمونه از تصویر تقطیع شده را نشان می‌دهد.



شکل ۵. تقطیع دست از تصاویر اولیه الف. حرف a، ب. حرف o، ج. حرف b

در شکل (۶) سه نمونه از نواحی دست که تقریب خورده‌اند دیده می‌شود.

علاوه بر الگوریتم‌های KNN، FKNNC و FRNNC که با نرم‌افزار MATLAB پیاده‌سازی شده‌اند از نرم‌افزار WEKA نیز جهت دسته‌بندی استفاده شده است. ویژگی‌ها به نرم-

برای زدن تقریب فازی بر روی نواحی دست از تابع عضویت مثلثی استفاده شده است. در شکل (۲) تابع عضویت مثلثی استفاده شده دیده می‌شود.

۲-۲- استخراج ویژگی

ویژگی‌های استخراج شده از تصویر عبارتند از: ممان هندسی مرتبه دو یا واریانس بعنوان معیار بیان کننده کنتراست؛ ممان هندسی مرتبه سه بعنوان بیان کننده میزان کجی (skewness) یا میزان تقارن هیستوگرام حول مقدار متوسط؛ ممان هندسی مرتبه چهار بیان کننده میزان مسطح بودن هیستوگرام؛ انرژی؛ آنترپی. همچنین از ویژگی‌های مرکز ثقل و کشیدگی که در مقابل تغییرات شکل دست مقاومتر هستند و می‌توانند تخمین ثابتی را بسازند، نیز استفاده شده است. این ویژگی‌ها تنها از ناحیه دست که در مرحله پیش‌پردازش از تصاویر استخراج و نرمالایز شده بود، استخراج می‌شود. بنابراین برای هر تصویر موجود در پایگاه داده، بردار ویژگی حاوی مقادیر بدست آمده برای ۷ ویژگی یاد شده بدست می‌آید.

همچنین ویژگی‌های مبتنی بر نمایش نیز از تصاویر استخراج و بار دیگر به دسته‌بندی‌کننده‌ها داده شد. آزمایشات نشان داد که استفاده از ویژگی‌های مبتنی بر نمایش نسبت به ویژگی‌های آماری (ویژگی‌های سطح پایین) نتایج بهتری در برداشتند.

۳-۲- شناسایی حالات و اشارات دست

در مقالات روش‌های مختلفی برای شناسایی و دسته‌بندی حالات و اشارات دست بکار گرفته شده است. در این مقاله ویژگی‌ها به الگوریتم‌های مختلفی داده شده و نتایج با یکدیگر مقایسه شده‌اند. این دسته‌بندی‌کننده‌ها عبارتند از: FRNNC و FKNNC، KNN (K nearest neighbors)

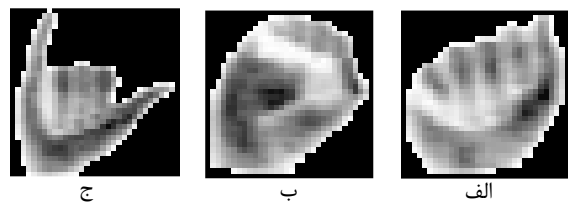
۳- آزمایش‌ها و نتایج

تصاویر پایگاه داده [۶] در سطح خاکستری و با فرمت TIF و به ابعاد ۲۴۸×۲۵۶ می‌باشد. تعداد کل تصاویر ۲۰۸۰ می‌باشد که تصاویر بیان کننده حروف A, B, ..., F و ae از هر کدام ۴۰ نمونه و برای بقیه حروف هر کدام ۱۰۰ نمونه وجود دارد. ضمناً پایگاه داده فاقد حروف J, Z می‌باشد. در مجموع ۲۵ کلاس (حروف الفبا) در این پایگاه داده وجود

افزار WEKA جهت دسته‌بندی داده شد که نتایج بدست آمده برای هر روش آمده است.

افزار WEKA جهت دسته‌بندی داده شد که نتایج بدست آمده برای هر روش آمده است.

۱۴۷۰ نمونه از تصاویر جهت آموزش و ۶۱۰ نمونه جهت تست در نظر گرفته شد. جدول (۱) نتایج بدست آمده را نشان می‌دهد.



۴- نتیجه‌گیری

ترکیب شناسایی حالات و اشارات ایستا با شناسایی خط سیر دست در طول چند فریم و استفاده از روش‌هایی مانند HMM، جهت تشخیص حالات و اشارات پویا پیشنهاد می‌شود. هدف نهایی، تشخیص online حالات و اشارات دست در هنگام بیان از طریق ناشنویان می‌باشد.

شکل ۶. تقریب نواحی تقطیع شده از تصاویر الف. حرف a. حرف b. حرف o. ج. حرف b

جدول ۱. نتایج بدست آمده از دسته‌بندی‌کننده‌های MLP ، SMO ، MultiClassClassifier ، BayseNet ، RBFnetwork و الگوریتم‌های KNN (با $k=3$ و $k=7$) ، FKNNC (با $k=3$ و $k=7$) ، و FRNNC (با $k=3$ و $k=7$)

مراجع

- [1] G. Fang, W. Gao: A SRN/HMM System for Signer-Independent Continuous Sign Language Recognition. In 5th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2002), pp. 312–317, Washington, DC, USA, May 2002.
- [2] C. Vogler, D. Metaxas: A Framework for Recognizing the Simultaneous Aspects of American Sign Language. Computer Vision and Image Understanding, Vol. 81, No. 3, pp. 358–384, March 2001.
- [3] S.C. Ong, S. Ranganath: Automatic Sign Language Analysis: A Survey and the Future beyond Lexical Meaning. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 27, No. 6, pp. 873–891, June 2005.
- [4] AboulElla Hassanién, Fuzzy rough sets hybrid scheme for breast cancer detection, Image and Vision Computing, v.25 n.2, p.172-183, February, 2007
- [5] Bhatt and Gopal, On fuzzy-rough sets approach to feature selection. Pattern Recognition Lett. v26. 965-975, 2005.
- [6] http://www.vision.auc.dk/~tbm/Gestures/Gestur_e_images/database.tar

الگوریتم	درصد درستی		
	ویژگی‌های آماری	ویژگی‌های مبتنی بر نمایش	
MLP	۹۳/۹	۹۲/۱	
SMO	۶۰	۹۴/۴	
BayseNet	۷۳/۳	۸۷/۲	
MultiClassClassifier	۴۵/۸	۸۶/۴	
RBFnetwork	۸۶/۲	۹۲/۳	
KNN	K=3	۴۰/۹	۸۸/۲
	K=7	۴۱/۳	۸۶/۵
FKNNC	K=3	۴۴/۷	۹۶/۲
	K=7	۴۴/۹	۹۶/۲
FRNNC	K=3	۴۸/۱	۹۷
	K=7	۴۸/۸	۹۶/۷

بردارهای ویژگی بدست آمده از مرحله قبل (ویژگی‌های مبتنی بر نمایش) به دسته‌بندی‌کننده‌های MLP ، SMO ، MultiClassClassifier ، BayseNet ، RBFnetwork و الگوریتم‌های KNN (با $k=3$ و $k=7$) ، FKNNC (با $k=3$ و $k=7$)