

عنوان مقاله:

بازشناسی دست نوشته های فارسی با استفاده از کاهش فرهنگ لغت مبتنی بر نقاط برجسته

محل انتشار:

کنفرانس ملی فناوری های نوین در مهندسی برق و کامپیوتر (سال: 1396)

تعداد صفحات اصل مقاله: 19

نویسندگان:

مریم فتاحی وانانی - دانشجوی کارشناسی ارشد برق، گرایش الکترونیک، دانشگاه آزاد اسلامی واحد مبارکه، گروه مهندسی برق و کامپیوتر
مبارکه، اصفهان، ایران

مجید ایرانپور مبارکه - دانشکده مهندسی کامپیوتر و فناوری اطلاعات، دانشگاه پیام نور

خلاصه مقاله:

در این تحقیق به بررسی، مطالعه و طراحی سیستمی به منظور کاهش فرهنگ لغت در بازشناسی دست نوشته های فارسی به منظور بهبود این فرایند پرداخته شده است. به این منظور نخست از روش کاهش واژگان برای بالابردن سرعت سیستم و افزایش دقت استفاده شده سپس از روش ماشینبرداری پشتیبان به دو صورت چند کلاسه و دو کلاسه برای بازشناسی استفاده شده است. روش کاهش واژگان دارای دو مرحله به صورت همزمان میباشد که در یک مرحله از ویژگی سراسری و در مرحله دیگر از ویژگی محلی استفاده شده است. کاهش واژگان براساس ویژگی های سراسری برپایه ویژگی های سراسری LOCI انجام شده است که این ویژگی ها از کل تصویر دستنوشته استخراج میشود و سپس برای خوشه بندی از روش K-means استفاده شده است و توسط این روش خوشه های واژگان استخراج می شود. این مرحله به صورت برون خط صورت میگیرد و از خوشه های استخراج شده در مرحله درون خط استفاده میشود. در مرحله کاهش واژگان بر پایه ویژگی محلی، ابتدا نواحی مرتبط به اجزاء متصل استخراج شده و به هر ناحیه برجستگی اختصاص داده می شود. در گام بعدی خصیصه های مرتبط به موقعیت نسبی محاسبه می شود. در نهایت با استفاده از این خصیصه ها با کمک خوشه های استخراج شده و معیار شباهت، عمل کاهش واژگان انجام می گیرد. این الگوریتم بر روی پایگاه داده ایران شهر مورد ارزیابی قرار گرفته است در مرحله کاهش ویژگی برپایه ویژگی سراسری 200 خوشه نزدیک به تصویر کلمه مورد آزمون ورودی انتخاب شده و در گام بعدی 5 خوشه نزدیک به تصویر ورودی انتخاب میشود که احتمال وجود کلمه آزمون در این خوشه ها، 98/38% و میزان کاهش کلمات کاندید مورد بررسی برای بازشناسی تصویر ورودی 81/54% و درجه تاثیر الگوریتم 80/21% است

کلمات کلیدی:

کاهش فرهنگ لغت، بازشناسی دست نوشته، ویژگی های سراسری LOCI خوشه بندی K-means ماشین بردار

لینک ثابت مقاله در پایگاه سیویلیکا:

<https://civilica.com/doc/758765>

