

عنوان مقاله:

بازشناسی برخط کلمات دست‌نویس فارسی با واژگانی گسترده

محل انتشار:

پنجمین کنفرانس ماشین بینایی و پردازش تصویر (سال: 1387)

تعداد صفحات اصل مقاله: 7

نویسندگان:

سیدمحمد رضوی - دانشگاه بیرجند، گروه الکترونیک و مخابرات

احسان اله کبیر - دانشگاه تربیت مدرس، بخش مهندسی برق

خلاصه مقاله:

در این مقاله روشی برای بازشناسی برخط کلمات فارسی ارائه می‌شود. ابتدا زیر-کلمات موجود در کلمه نوشته شده استخراج می‌شوند. علائم هر زیر-کلمه و مکان نسبی آنها تشخیص داده می‌شوند و با توجه به آن گروهی از کلمات، که از نظر تعداد زیر-کلمات و علامتهای موجود در هر زیر-کلمه با کلمه ورودی یکسان هستند، از واژگانی با 30000 کلمه، انتخاب می‌شوند. هر کدام از زیر-کلمات موجود در کلمه ورودی، با توجه به گروه انتخاب شده، با تعدادی از زیر-کلمات مقایسه می‌شود و براساس فاصله ای که با آن زیر-کلمات دارد امتیازی به هرکدام از آنان داده می‌شود. برای هر یک از کلمات گروه انتخابی امتیازی بر اساس امتیازات زیر-کلمات آن محاسبه می‌شود و کلمه ای که بیشترین امتیاز را آورده باشد به کلمه ورودی نسبت داده می‌شود. روش پیشنهادی روی یک متن انتخابی، که توسط یک نویسنده نوشته شده است، آزمایش شد. میزان بازشناسی درست 94,9% برای متن نمونه بدست آمده است.

کلمات کلیدی:

فارسی، بازشناسی، دست‌نویس برخط، کمترین فاصله، واژگان گسترده

لینک ثابت مقاله در پایگاه سیویلیکا:

<https://civilica.com/doc/52078>

