

عنوان مقاله:

شاخص گذاری و بازیابی اطلاعات از پایگاه تصاویر اسناد

محل انتشار:

چهاردهمین کنفرانس مهندسی برق ایران (سال: 1385)

تعداد صفحات اصل مقاله: 9

نویسندگان:

محمد اکبری - دانشکده ریاضی و علوم کامپیوتر، دانشگاه صنعتی امیرکبیر

رضا عرمی - دانشکده فنی و مهندسی، دانشگاه الزهرا

خلاصه مقاله:

در این مقاله، یک روش برای بازیابی اطلاعات از تصاویر اسناد فارسی بدون استفاده از بازشناسی نوری حروف معرفی کرده ایم. ابتدا برای ریز کلمات سند یک ویژگی به نام برچسب کانتور بالایی استخراج شده سپس بر مبنای این ویژگی یک دیکشنری تصویری برای ریز کلمات ایجاد شده است. به کمک این روش برای دو مسئله مختلف در بازیابی اسناد با استفاده از جستجوی کلمات کلیدی و بازیابی بر اساس شباهت اسناد راه حل ارائه شد. روشهای ارائه شده بر روی یک پایگاه تصاویر اسناد فارسی ارزیابی شد. نتایج بدست آمده توانایی این روش ها در بازیابی اطلاعات از تصاویر اسناد را تایید کرده است.

کلمات کلیدی:

بازیابی تصاویر اسناد، بردار ویژگی سند، جستجوی کلمات کلیدی، شباهت اسناد، دیکشنری تصویری

لینک ثابت مقاله در پایگاه سیویلیکا:

<https://civilica.com/doc/54838>

