

Rozpoznawanie języka z użyciem metryki ulic

Mirosław GAJER

Akademia Górniczo-Hutnicza, Katedra Automatyki
al. Mickiewicza 30, 30-059 Kraków
e-mail: mgajer@ia.agh.edu.pl

Otrzymano 17 października 2001 roku

Streszczenie. W artykule opisano system służący do automatycznego rozpoznawania języka tekstów. System taki może stanowić pierwsze ogniwo większego systemu służącego do automatycznego tłumaczenia wielojęzycznych tekstów, gdzie zachodzi konieczność ustalenia języka tłumaczonego dokumentu. W artykule opisano w sposób szczegółowy zaproponowaną metodę rozpoznawania, która opiera się na pomiarze częstości występowania w tekście poszczególnych znaków oraz na pomiarze minimalnej odległości do klasy przynależności. Zamieszczono liczne wyniki eksperymentów, które potwierdzają poprawną pracę opisanego systemu.

Słowa kluczowe: techniki przetwarzania języków naturalnych, rozpoznawanie języka tekstów, metoda minimalnoodległościowa