

Bonbónek 2

Vytváření strukturních obrazů

V prostém textovém módu programově realizujte úlohu detekce primitiv v rastrové reprezentaci písmen typu „OCR – Handprint Numeric“ metodou porovnávání se vzory primitiv (template matching) a klasifikace těchto primitiv jednoduchou strukturní metodou.

Předpokládejte rozměr reprezentujícího rastru 14 x 21 obrazových bodů. Druhy i velikost reprezentace primitiv volte podle vlastního uvážení tak, abyste jimi postihli všechny tvarové charakteristiky rozpoznávaných písmen a přitom dosáhli uspokojivé algoritmické složitosti vyhledávání. V závěru programového chodu zřetězte nalezená (detekovaná) primitiva do jednoznačných strukturních popisů (obrazů) – zřetězení proveďte algoritmicky a paměťově nejméně náročným způsobem (řetězení lze provádět i v průběhu detekce jednotlivých primitiv). Výstupem Vašeho programu bude informace o klasifikaci testovaného znaku, tj. o přiřazení vytvořeného strukturního obrazu k příslušné třídě.

Data pro trénování a testování jsou umístěna na adrese

<http://www.kiv.zcu.cz/studies/predmety/uir/ocr/PismenaOCR.zip>

V archivu naleznete celkem šest adresářů se jmény „abc?“ – v každém adresáři pak 26 souborů s obrazy písmen (soubory a.trs až z.trs). Symbol „?“ ve jménu adresáře je nahrazen číslicí, která udává, do jaké míry jsou písmena v adresáři poškozena (v počtu pixelů). Data z adresáře abc0 použijte pro nastavení klasifikátoru, data z ostatních adresářů pro testování.

[15 bodů]

ABCDEFGHIJKLMN
MNOPQRSTUVWXYZ