

Posudek na diplomovou práci

Autor: Matěj Klaučo

Název: Podpora rozpoznávání štruktúry matematických vzorcov v rámci OCR systému

Podaná diplomová práce se zabývá rozpoznáváním matematických vzorců v obrazové podobě (získané například skenováním tištěných dokumentů) a jejich následným převodem do textové podoby ve formátu TeX. Cílem práce bylo seznámit se s existujícími metodami a implementovat vybrané řešení jako rozšíření existujícího OCR (Optical Character Recognition – optické rozpoznávání znaků) systému.

Rozpoznávání a převod vzorců probíhá následovně: Nejprve se obraz binarizuje pomocí lokálního prahování a odstraní se šum. Potom se rozpoznávají jednotlivé matematické symboly na základě příznaků a kaskády klasifikátorů: (1) jednotlivé souvislé komponenty jsou rozděleny do skupin podle jejich vnějšího obrysu za pomoci Kohonenovy mapy, (2) v rámci skupin jsou symboly klasifikovány podle směru hran v různých oblastech symbolu, (3) chybně klasifikované symboly slouží k trénování SVM (Support Vector Machines) klasifikátorů korigujících tyto chyby. Rozpoznané symboly, jejich pozice a velikost jsou vstupem pro rozpoznávání struktury vzorců na základě definovaných pravidel. Na začátku je vzorec reprezentován grafem nejbližších rozpoznávaných symbolů a tento graf je postupně redukován podle pravidel, která definují jak se mohou jednotlivé podgrafy redukovat na neterminály. (Například integrál se skládá ze symbolu integrálu, ke kterému mohou být přidruženy spodní a horní meze).

V kapitole 5 je experimentálně ověřena úspěšnost klasifikace pro různé nastavení parametrů klasifikátorů. Nejlepší úspěšnost klasifikace symbolů je 93,7%; bezchybně bylo rozpoznáno 61% vzorců a jednu nebo dvě chyby obsahovalo 27% vzorců, z čeho plyne potřeba manuálních korekcí. Je potřeba si však uvědomit, že tento problém je složitější a symbolů je daleko více než při běžném rozpoznávání latinky. Součástí diplomové práce jsou moduly do OCR systému implementující popsané řešení, aplikace pro trénování klasifikátorů (s podporou distribuce na více počítačů) a grafický editor pravidel, ve kterém uživatel může definovat své vlastní matematické struktury.

Celkově práce působí dobře, student nastudoval relevantní metody, implementoval je, otestoval jejich úspěšnost, začlenil je od OCR systému a to včetně podpůrných aplikací. Hlavní přínos vidím v metodě pro rozpoznávání struktury vzorců, která umožňuje definici vlastních složených výrazů, což bývá uživateli TeXu využíváno. Moje výhrady směřují k textu práce, kde bych v práci uvítal podrobnější popis rozšířeného klasifikátoru (sekce 3.4.2 a 3.4.3), který je na hranici pochopitelnosti, a také z popisu experimentu není zřejmé, na kterých datech byly klasifikátory trénovány. Nicméně, toto nesnižuje kvalitu diplomové práce jako celku. Mým dotazem na studenta je, kdy přesně a které symboly jsou použity pro trénink SVM klasifikátorů.

Doporučuji tuto diplomovou práci klasifikovat známkou „výborně“.

V Praze, 15. srpna 2011

Mgr. Pavel Vácha, Ph.D.