FAKULTA APLIKOVANÝCH VĚD ZÁPADOČESKÉ UNIVERZITY V PLZNI

STUDENTSKÁ VĚDECKÁ KONFERENCE 2016

Využití dvourozměrné Fourierovy transformace v úloze zarovnání naskenovaného dokumentu

Petr Neduchal[1], Lukáš Bureš[2]

1 Úvod

Papír je stále jedním z nejspolehlivějších prostředků pro uchovávání informací. Přesto může být užitečné informace digitalizovat. Takovým případem je digitalizace dokumentů za účelem snazšího vyhledávání informací v nich obsažených, nebo digitalizace faktur za účelem uložení dat do databáze. Součástí tohoto převodu je proces naskenování dokumentu. Výstup ze scanneru však není vždy dokonalý a může být natočený. To může způsobit problém při optickém rozpoznávání znaků (OCR), proto je důležité dokument zarovnat.

Cílem tohoto článku je analýza využití dvourozměrné Fourierovy transformace v této úloze a zhodnocení její přesnosti a rychlosti v porovnání s metodou Houghovy přímkové transformace, která se v tomto problému často používá.

2 Zarovnání dokumentu

Při skenování dokumentu může dojít ke špatnému srovnání dokumentu obsluhou viz obrázek 1 vlevo. Toto natočení, které může dosahovat hodnot $\pm$ 5 stupňů může zapříčinit nesprávné rozpoznání znaků v dalším kroku. K zarovnání dokumentu tak jak je vidět na obrázku 1

Obrázek 1: Vlevo ukázka zdrojového dokumentu. Vpravo ukázka dokumentu zarovnaného pomocí představené metody.

vpravo se většinou přistupuje pomocí Houghovy přímkové transformace [Rosner et al. (2010)], která v obrázku nalezne všechny přímky se zadanou minimální délkou. Přímky jsou hledány pomocí parametrů ρ a θ v parametrickém prostoru definovaném vztahem

$$\rho = xcos(\theta) + ysin(\theta), \tag{1}$$

kde ρ je vzdálenost od pravého horního rohu dokumentu, kterou tvoří kolmice k nalezené přímce. A θ je úhel, který tato kolmice svírá s osou x. Z nalezených ρ a θ lze rekonstruovat

[1] student doktorského studijního programu Aplikované vědy a informatika, obor Kybernetika, specializace Počítačové vidění, e-mail: neduchal@kky.zcu.cz

[2] student doktorského studijního programu Aplikované vědy a informatika, obor Kybernetika, specializace Počítačové vidění, e-mail: lbures@kky.zcu.cz

nalezené přímky. Vzhledem k omezenému rozsahu úhlů lze z těchto přímek vybrat pouze ty, které by měly odpovídat natočení. Jako celkové natočení je pak brán modus získaných úhlů natočení vybraných přímek.

Jiným přístupem je metoda využívající algoritmus rychlé Fourierovy transformace (FFT) [Šonka et al. (2007)], která má asymptotickou složitost $O\left(nlog(n)\right)$. FFT převede vstupní obraz do frekvenční oblasti. V této reprezentaci se objevují vysoké hodnoty tvořící přímky, které odpovídají kolmicím k přímkám v původním obraze.

Natočením oblasti okolo středu frekvenčního obrazu a nalezením maxima součtu hodnot ve středovém sloupci se následně určí úhel, o který je nutné otočit původní obraz.

3 Experimenty a výsledky

Metody byly porovnávány vzhledem ke dvěma kvalitativními kritérii. Prvním kritériem je kvalita odhadu a s ní spojená robustnost proti šumu v datech. Druhým kritériem byla potom časová náročnost algoritmu. Vše je shrnuto v tabulce 1. Experiment byl proveden na datasetu obsahujícím 125 obrázků o velikosti 2516x3272 pixelů.

	Přesnost ve stupních			Čas [s]		
Metoda	Průměr	Max	Min	Průměr	Min	Max
Fourier	0.60	0.02	1.19	5.60	5.39	6.03
Hough	2.49	0.003	10.77	0.29	0.06	0.81

Tabulka 1: Tabulka výsledků. U času i přesnosti platí, že nižší číslo znamená lepší výsledek.

Z výsledků je patrné, že metoda využívající Fourierovu transformaci je přesnější, ale platí za to vyšší časovou náročností

4 Závěr

Tento článek zkoumal využití Fourierovy transformace v úloze zarovnání naskenovaného dokumentu. Chyba navržené metody dosahovala v průměru mírně přes půl stupně, přičemž se nikdy nedostala nad 1.2 stupně. Bohužel má o mnoho vyšší časové nároky na výpočet oproti Houhově transformaci, kterou je možné

Poděkování

Tento příspěvek byl podpořen grantovým projektem SGS-2016-039. ...

Literatura

Šonka, M., Hlavac, v., Boyle, R., 2007. Image Processing, Analysis, and Machine Vision. *Thomson Learning, Toronto* .

Rosner, D., et al., 2010. Text line processing for high-confidence skew detection in image documents. *Intelligent Computer Communication and Processing (ICCP), 2010 IEEE International Conference.*