

***Žilinská univerzita v Žiline***

***Fakulta riadenia a informatiky***

***Projekt 1***

***Inteligentná analýza obrazu***

***Spájanie dokumentov získaných kamerou***

*Patrik Hrmo*

*5ZU011*

*2016/2017*

[1. Úvod 3](#_Toc475084271)

[2. Analýza súčasného stavu 3](#_Toc475084272)

[3. Voľba vhodnej začiatočnej implementácie 3](#_Toc475084273)

[3.1. OpenCV 3](#_Toc475084274)

[3.2. BoofCV 5](#_Toc475084275)

[4. Dataset 7](#_Toc475084276)

[5. Implementácia 7](#_Toc475084277)

[5.1. Výpočet veľkosti spojeného obrázka na základe homografie 7](#_Toc475084278)

[6. Bibliografia 8](#_Toc475084279)

# Úvod

V tejto práci sa budem venovať spájaniu snímok textových dokumentov. Tieto snímky môžu byť vytvárané mobilným zariadením, takže budú podliehať perspektívnej transformácií. Výsledkom bude teda obrázok dokumentu poskladaný z týchto snímok tak, aby ho bolo možné prečítať či už ľudským okom, alebo nástrojom OCR.

# Analýza súčasného stavu

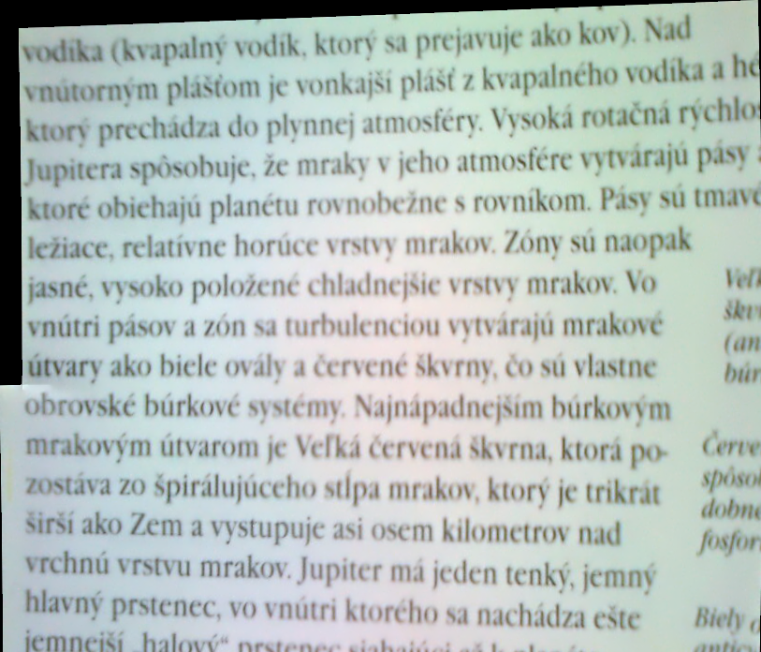
K vyriešeniu tohto problému je napísaných mnoho článkov a kníh. Tieto sa však často venujú spájaniu dokumentov na základe textových vlastností (1)ako napríklad tok textu. V tejto práci sa však sústredím na úpravu algoritmov pre spájanie panorám (2) tak, aby boli funkčné aj na textové dokumenty. Algoritmy na spájanie panorám sú veľmi často využívané, a práve z tohto dôvodu bývajú implementované v rôznych frameworkoch s počítačovým videním.

# Voľba vhodnej začiatočnej implementácie

Ako prvé je teda nájsť vhodný framework pre prácu s obrazom. Najlepšie taký v ktorom je funkčná implementácia spájania panorám.

## OpenCV

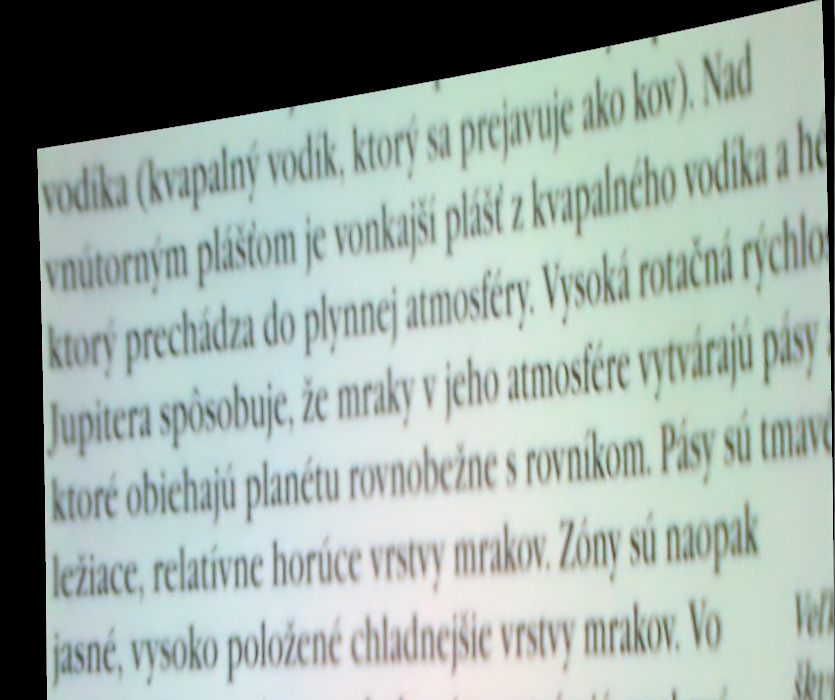
Jeden z najznámejších frameworkov na prácu s počítačovým videním je OpenCV. OpenCV obsahuje triedu Stitcher, ktorá vie spájať obrázky do panorám. Tento algoritmus vie aj dobre spájať snímky textových dokumentov.



Obrázok spojenie textových dokumentov v Opencv

Obrázok 2 spojenie textových dokumentov v OpenCV

Problém s touto implementáciou nastáva ak sa pokúsime spojiť snímky ktoré sú pod vplyvom rotácií. V tomto prípade algoritmus zlyhá alebo vráti veľmy zlý výsledok.

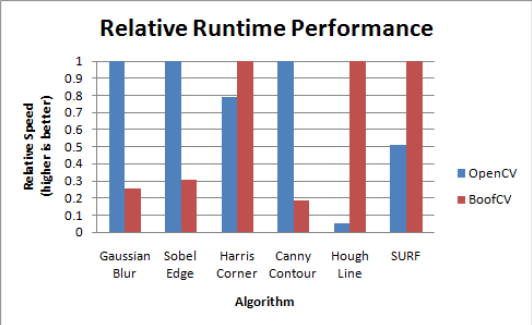


Obrázok 3 zlé spojenie dokumentov v OpenCV

Je teda nutné upraviť tento algoritmus. S týmto sú však spojené problémy ohľadom dokumentácie, ktorá je pri triedach a metódach spájania zanedbaná. Pre lepšiu produktivitu je teda potrebné nájsť iný framework.

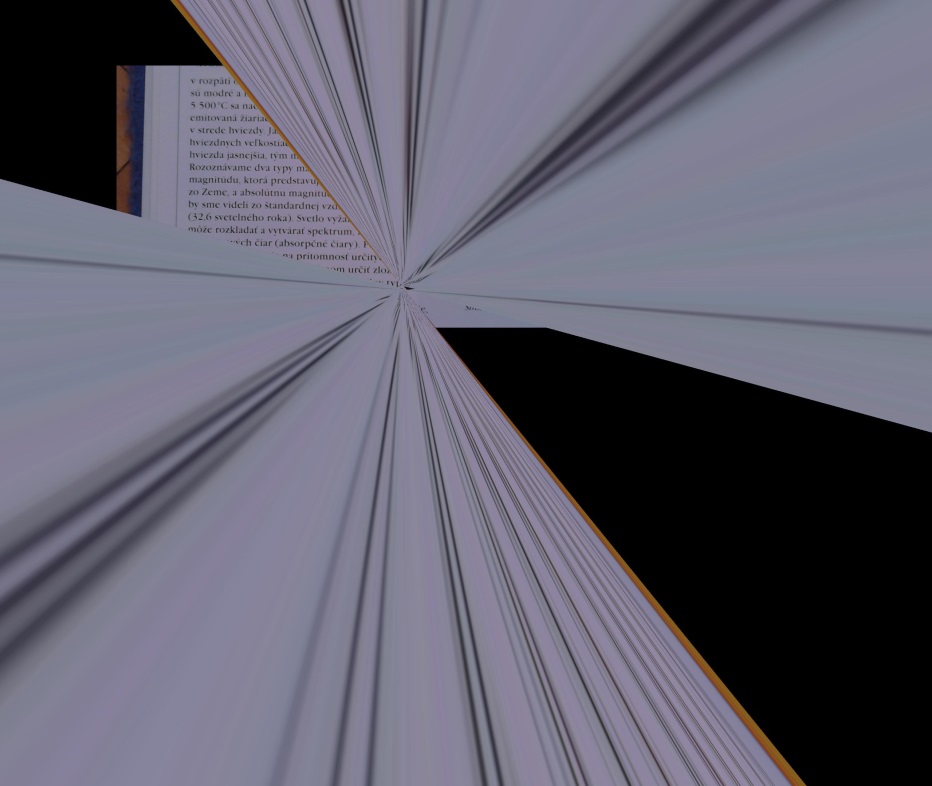
## BoofCV

BoofCV je pomerne nový framework ktorý je napísaný čisto v jave. Z toho samozrejme vyplíva že operácie na nízkej úrovni sú pomalšie ako v OpenCV. Pre tento projekt je však dôležitá hlavne rýchlosť detekovania a popisu *feature points* snímok. Tu je však BoofCV približne 2x rýchlejší ako OpenCV.



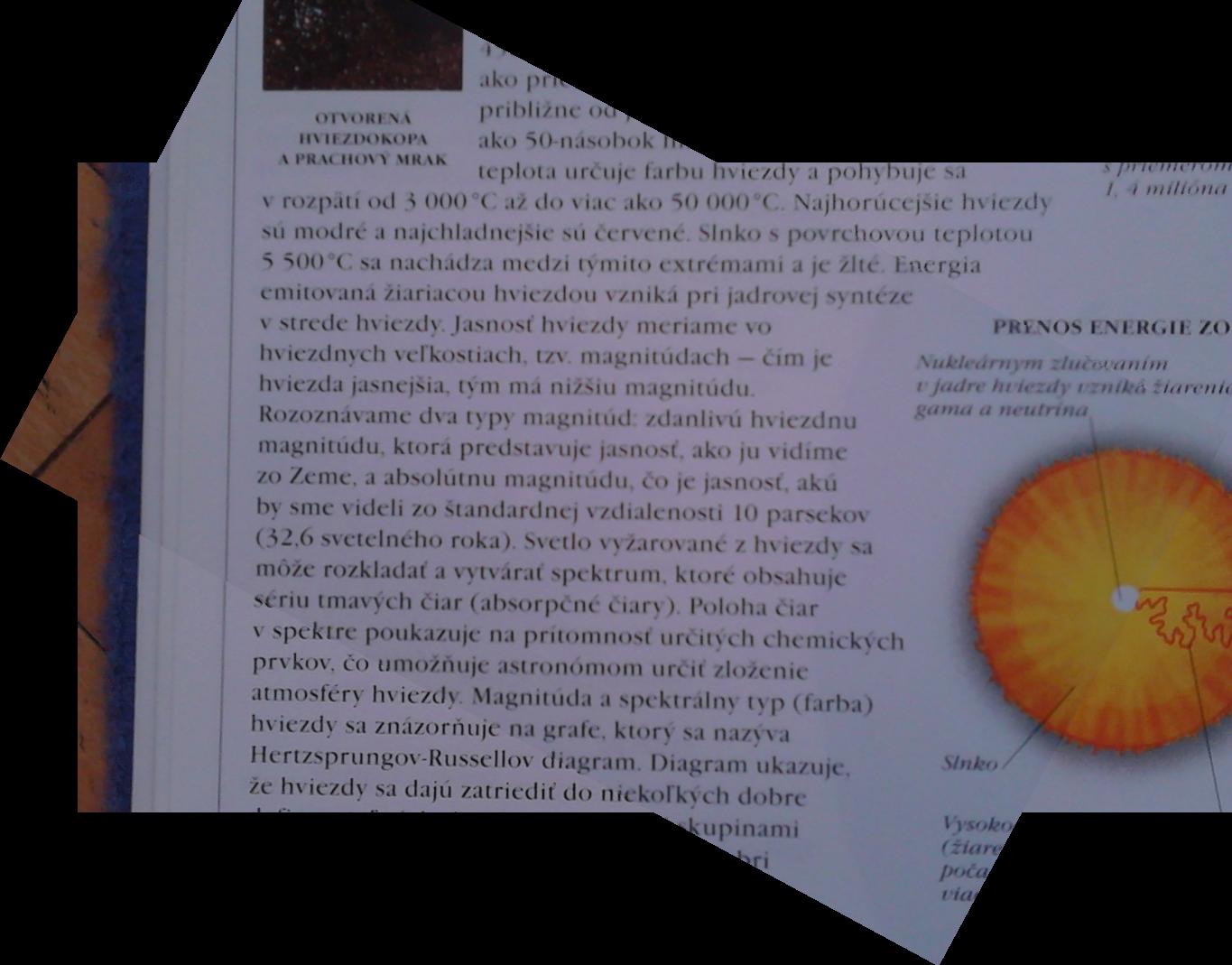
Obrázok 4 porovnanie rýchlostí OpenCV a BoofCV

BoofCV neposkytuje defaultne žiadnu triedu na spájanie obrázkov, avšak medzi príkladmy využitia je aj jednoduchý algoritmus spájania obrázkov. Tento príkladný algoritmus je však nevhodný pre spájanie textových dokumentov.



Obrázok 5 zlé spojenie textového dokumentu v BoofCV

Analýzov som zistil že toto spojenie zlyhalo kôli nedostatočnému počtu feature pointov. Tak isto som zmenil konfiguráciu RANSACU ktorý sa v tomto algoritme použiva na určenie homografie medzi obrázkami.



Obrázok 6 spojenie v BoofCV po drobných zmenách

# Počiatočný dataset

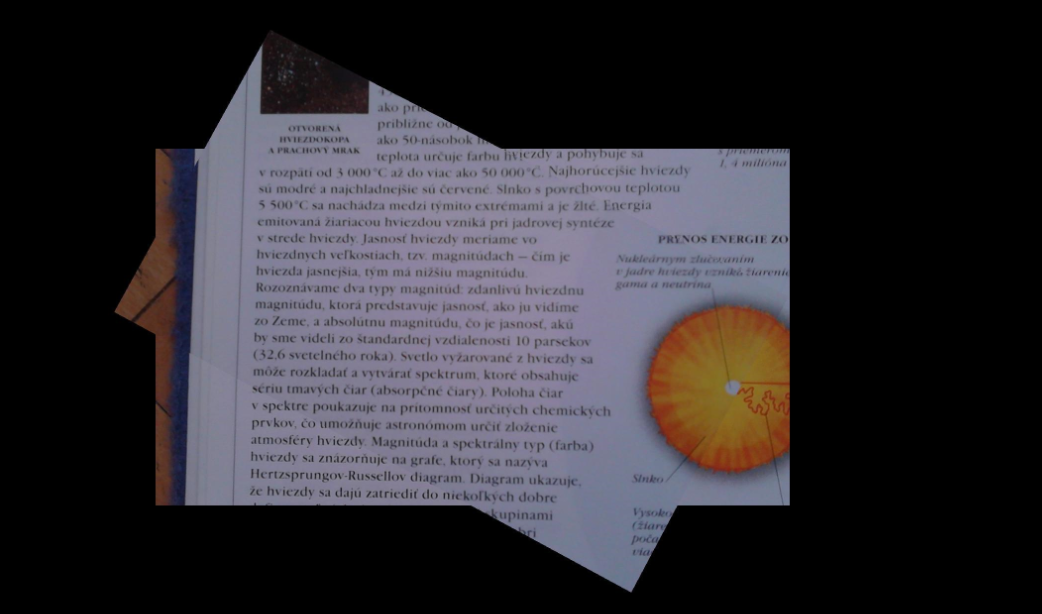
Keď už máme základnú funkčnosť algoritmu implementovanú, je dobré si vytvoriť dataset ktorý bude slúžiť pre budúce testy. Dataset je tvorený priečinkami, kde každý priečinok obsahuje dve snímky – jednu bez transformácie(prípadne len so zanedbateľnou transformáciou) a druhú s transformáciou. Takto môžem ľahko sledovať úspešnosť spájania pri rôznych transformáciách. Bol vytvorený aj testovací program, ktorý prejde a pospája všetky snímky v príslušných priečinkoch. To nám urýchli testovanie nakoľko nebude nutné meniť názvy snímok a priečinkov v kóde.

# Implementácia

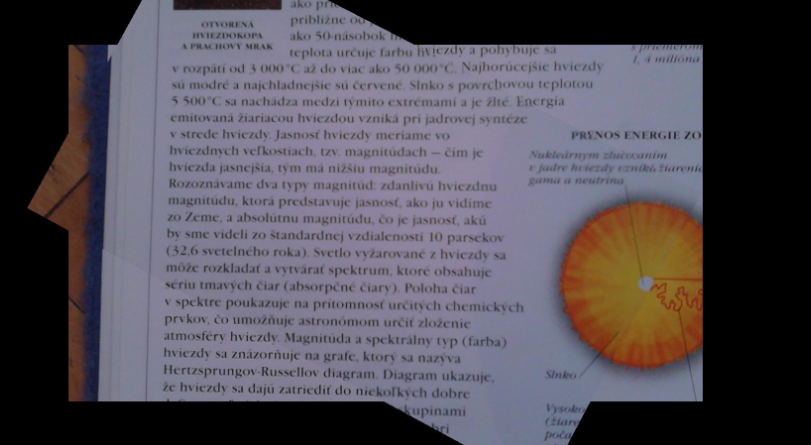
Teraz môžme pristúpiť k pokročelejším úpravám a implementácií chýbajúcich častí algoritmu.

## Výpočet veľkosti spojeného obrázka na základe homografie

Je potrebné upraviť veľkosť obrázka ktorý vznikne po spájaní. Jeho veľkosť je momentálne určená konštantne, čo spôsobuje zbytočné preplnenie pamäte. V horšom prípade by táto vlastnosť mohla zapríčiniť nemožnosť spojenia tohto obrázka s ostatnými, nakoľko sa časti tohot obrázka nemusia nachádzať vo vykreslenej zóne!



Obrázok 7 ukážka voľného vyplitvaného priestoru



Obrázok 8 ukážka orezania časti pospájaného obrázka

Za pomoci homografie je možné vypočítať presnú veľkosť výsledného obrázka. Ako prvé je potrebné určiť si body reprezentujúce rohy obrázkov ktoré spájame. U základného obrázka je to ľahké nakoľko nebude podliehať žiadnej transformácií (jeho body budú teda [0,0],[0,sirka], ...). U druhého obrázka je to zložitejšie – treba vypočítať ich súradnice po transformácií. Keď už máme vypočítaných všetkých osem bodov, je úloha nájdenia veľkosti výsledného obrázka triviálna(odčítanie y súradnice najvrchnejšieho bodu od najspodnejšieho, analogcky pre najľavejší a najpravjší bod).

# Bibliografia

1. http://www.cfar.umd.edu/~daniel/daniel\_papersfordownload/LiangICPR2006.pdf. [Online]

2. http://matthewalunbrown.com/papers/ijcv2007.pdf. [Online]

# Pokročilý dataset

Pre nové testy je potrebné vytvoriť dataset ktorý bude pokrývať širšiu škálu rôznych vzťahov medzi snímkami. Potrebné je pozorovať ako sa bude bude kvalita spojenia správať pod rôznymi rotáciami a sklonmy kameri.

## Štruktúra datasetu

Štruktúra datasetu je podobná ako pri klasických súboroch – v stromovej štruktúre. Je tu však zopár obmedzení. Samotné snímky ktoré budeme zlučovať sa nachádazjú iba v listoch tejto štruktúry. V každom liste sa teda musia nachádzať snímky ktorých spjenie ideme testovať. Okrem nich sa tu musí nachádzať aj obrázok s názvom grandTruth.jpg, na ktorom je zobrazený celý dokument vo vysokej kvalite.

## Popis datasetov

1 – velke prekrytie ziadna rotacia maly sklon

2 – velke prekrytie velka rotacia maly sklon

3- stredne prekrytie ziadna rotacia ziadny sklon

5 – stredne prekerytie ziadna rotacia maly sklon

6 – stredne prekrytie stredna rotacia ziadny sklon

7 – stredne prekerytie stredna rotacia maly sklon

8 – stredne prekerytie mala rotacia ziadny sklon

9 – malé prekrytie žiadna rotácia žiadny sklon

10 – malé prekrytie žiadna rotácia malý sklon

11 – malé prekrytie stredná rotácia žiadny sklon

12 - malé prekrytie malá rotácia žiadny sklon

## Vyhodnocovanie

Kvlaitu spojenia posudzujeme porovnaním textu, ktorý získame pomocou nástroja OCR. Ako za pravdivý text (grand truth) budeme považovať text zo súboru grandTruth.jpg ktorý sa musí nahádzať v každom datasete. Tento text budeme porovnávať s textom ktorý získame z obrázka ktorý vznikol spojením snímok v tomto datasete.

Jednotlivé texty sú porovnávané na základe upravenej metriky levenštejnovej vzdialenosti. Tú bolo potrebné normalizovať aby jej vyhodnotenie bolo v intervale od 0(texty sa vôbec nezhodujú) až po 1(texty sú rovnaké). Na základe tohto čísla môžeme potom povedať ako úspešné bolo spojenie snímok.

# Testy

Parametrov algoritmu spájania viacerých snímok je 11, a nebude možné nájsť optimálne parametre kombinovaním rôznych hodnôt týchto parametrov.

## Zisťovanie vhodnej kombinácie hlavných parametrov

zameriame len na dve najviac významné parametre spájania – počet iterácií RANSACu a počet feature pointov pre každú snímku.

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Ransac 1500  Feature 5000 | | Ransac 1250  Feature 5000 | | Ransac 1000  Feature 5000 | | Ransac 750  Feature 5000 | | Ransac 500  Feature 5000 | | Ransac 250  Feature 5000 | |
| dataset | [%] | Cas  [ms] | [%] | Cas  [ms] | [%] | Cas  [ms] | [%] | Cas  [ms] | [%] | Cas  [ms] | [%] | Cas [ms] |
| celkom | 84.64 | 86909 | 84.38 | 87469 | 84.17 | 86955 | 83.04 | 86940 | 82.54 | 86830 | 81.07 | 87346 |
| 1 | 92.21 | 9157 | 92.21 | 9235 | 92.21 | 9095 | 92.21 | 9017 | 92.21 | 9048 | 92.21 | 8985 |
| 2 | 85.11 | 8377 | 85.11 | 8408 | 85.11 | 8377 | 85.11 | 8424 | 82.76 | 8377 | 82.76 | 8393 |
| 3 | 95.68 | 7722 | 95.68 | 7784 | 94.41 | 7707 | 94.41 | 7722 | 94.41 | 7722 | 95.51 | 7675 |
| 4 | 80.95 | 7925 | 80.95 | 7894 | 80.95 | 7956 | 80.95 | 7863 | 80.95 | 7878 | 81.50 | 7862 |
| 5 | 82.46 | 8471 | 82.46 | 8643 | 81.42 | 8548 | 62.81 | 8549 | 61.63 | 8502 | 75.70 | 8643 |
| 6 | 69.88 | 7878 | 67.08 | 7847 | 67.08 | 7800 | 71.18 | 7831 | 71.18 | 7753 | 71.18 | 7784 |
| 7 | 93.24 | 8144 | 94.15 | 8236 | 94.15 | 8206 | 94.15 | 8206 | 92.15 | 8190 | 92.15 | 8144 |
| 8 | 84.27 | 7114 | 83.28 | 7161 | 83.28 | 7098 | 83.28 | 7145 | 83.28 | 7098 | 84.31 | 7052 |
| 9 | 85.87 | 6864 | 85.87 | 6973 | 85.87 | 6958 | 87.93 | 6942 | 87.93 | 7052 | 84.26 | 6973 |
| 10 | 85.02 | 7769 | 85.02 | 7831 | 85.02 | 7831 | 85.02 | 7769 | 85.02 | 7815 | 68.49 | 8222 |
| 11 | 76.38 | 7488 | 76.38 | 7457 | 76.38 | 7379 | 76.38 | 7472 | 76.38 | 7395 | 63.70 | 7613 |

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Ransac 10 000  Feature 5000 | | Ransac 100 000  Feature 5000 | | Ransac 3000  Feature 5000 | | Ransac 5000  Feature 5000 | | Ransac 1000 000  Feature 5000 | |
| dataset | [%] | Cas  [ms] | [%] | Cas  [ms] | [%] | Cas  [ms] | [%] | Cas  [ms] | [%] | Cas  [ms] |
| celkom | 81.31 | 93324 | 81.28 | 135212 | 84.84 | 90778 | 81.51 | 90708 | 81.16 | 571861 |
| 1 | 92.21 | 9801 | 93.26 | 14733 | 92.21 | 10012 | 92.21 | 9532 | 91.67 | 60854 |
| 2 | 84.45 | 8863 | 82.80 | 12101 | 85.08 | 8880 | 85.08 | 9124 | 84.96 | 41652 |
| 3 | 94.74 | 8414 | 95.37 | 13103 | 96.27 | 8018 | 96.27 | 8127 | 95.03 | 60323 |
| 4 | 81.24 | 8462 | 81.24 | 12570 | 80.92 | 8125 | 82.01 | 8204 | 78.75 | 58547 |
| 5 | 83.24 | 9056 | 82.59 | 12157 | 82.37 | 8796 | 83.63 | 8826 | 82.55 | 44218 |
| 6 | 66.91 | 8428 | 66.91 | 12118 | 69.70 | 8109 | 66.91 | 8103 | 68.23 | 48829 |
| 7 | 93.32 | 8827 | 92.81 | 13924 | 93.88 | 8476 | 93.32 | 8625 | 92.86 | 72497 |
| 8 | 84.47 | 7499 | 84.68 | 11850 | 85.59 | 7315 | 84.24 | 7402 | 85.22 | 54799 |
| 9 | 84.47 | 7399 | 85.35 | 10814 | 85.85 | 7209 | 84.42 | 7257 | 86.78 | 50124 |
| 10 | 53.96 | 7955 | 54.57 | 10804 | 85.02 | 8113 | 52.16 | 7869 | 52.26 | 38493 |
| 11 | 75.36 | 8622 | 74.49 | 11038 | 76.35 | 7725 | 76.35 | 7639 | 74.51 | 41398 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Ransac 3000  Feature 5000  Inlier treshold 2 | | Ransac 3000  Feature 5000  Inlier treshold 3 | |
| dataset | [%] | Cas  [ms] | [%] | Cas  [ms] |
| celkom | 84.71 | 88187 | 81.68 | 88906 |
| 1 | 92.70 | 9251 | 93.08 | 9937 |
| 2 | 82.16 | 8487 | 85.08 | 8518 |
| 3 | 95.65 | 8174 | 95.68 | 7925 |
| 4 | 82.68 | 7971 | 82.68 | 8034 |
| 5 | 82.37 | 8783 | 82.37 | 8674 |
| 6 | 69.70 | 7878 | 69.70 | 7894 |
| 7 | 94.53 | 8269 | 94.53 | 8424 |
| 8 | 84.74 | 7129 | 84.74 | 7301 |
| 9 | 85.02 | 7831 | 84.97 | 6989 |
| 10 | 76.35 | 7441 | 49.29 | 7675 |
| 11 | 75.36 | 8622 | 76.35 | 7535 |

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Ransac 3000  Feature 5000  Scorer 0.1 | | Ransac 1250  Feature 5000  Scorer 0.3 | | Ransac 1000  Feature 5000  Scorer 0.5 | | Ransac 750  Feature 5000  Scorer 0.7 | | Ransac 500  Feature 5000  Scorer 2 | | Ransac 250  Feature 5000  Scorer max | |
| dataset | [%] | Cas  [ms] | [%] | Cas  [ms] | [%] | Cas  [ms] | [%] | Cas  [ms] | [%] | Cas  [ms] | [%] | Cas [ms] |
| celkom | 81.93 | 88684 | 84.84 | 87490 | 84.84 | 87562 | 84.84 | 87171 | 84.84 | 86830 | 84.84 | 87346 |
| 1 | 93.05 | 10030 | 92.21 | 9303 | 92.21 | 9408 | 92.21 | 9266 | 92.21 | 9235 | 92.21 | 8985 |
| 2 | 85.01 | 8486 | 85.08 | 8611 | 85.08 | 8408 | 85.08 | 8392 | 82.76 | 8377 | 82.76 | 8393 |
| 3 | 94.66 | 7800 | 96.27 | 7707 | 96.27 | 7894 | 96.27 | 7768 | 94.41 | 7722 | 95.51 | 7675 |
| 4 | 84.83 | 9094 | 80.92 | 7800 | 80.92 | 7831 | 80.92 | 7863 | 80.95 | 7878 | 81.50 | 7862 |
| 5 | 71.11 | 7831 | 82.37 | 8611 | 82.37 | 8611 | 82.37 | 8596 | 61.63 | 8502 | 75.70 | 8643 |
| 6 | 93.71 | 8253 | 69.70 | 7846 | 69.70 | 7846 | 69.70 | 7706 | 71.18 | 7753 | 71.18 | 7784 |
| 7 | 84.33 | 7223 | 93.88 | 8252 | 93.88 | 8252 | 93.88 | 8221 | 92.15 | 8190 | 92.15 | 8144 |
| 8 | 86.05 | 7036 | 85.59 | 7191 | 85.59 | 7176 | 85.59 | 7114 | 83.28 | 7098 | 84.31 | 7052 |
| 9 | 54.91 | 7456 | 85.85 | 6942 | 85.85 | 6942 | 85.85 | 6973 | 87.93 | 7052 | 84.26 | 6973 |
| 10 | 74.41 | 7581 | 85.02 | 7941 | 85.02 | 7862 | 85.02 | 7893 | 85.02 | 7815 | 68.49 | 8222 |
| 11 | 79.81 | 7485 | 76.35 | 7286 | 76.35 | 7332 | 76.35 | 7379 | 76.38 | 7395 | 63.70 | 7613 |

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Ransac 3000  Feature 5000  Feature int 10 | | Ransac 1250  Feature 5000  Feature int 5 | | Ransac 1000  Feature 5000  Feature int 2 | | Ransac 750  Feature 5000  Scorer 0.7 | | Ransac 500  Feature 5000  Scorer 2 | | Ransac 250  Feature 5000  Scorer max | |
| dataset | [%] | Cas  [ms] | [%] | Cas  [ms] | [%] | Cas  [ms] | [%] | Cas  [ms] | [%] | Cas  [ms] | [%] | Cas [ms] |
| celkom | 82.01 | 83777 | 81.22 | 86544 | 81.71 | 88155 | 84.84 | 87171 | 84.84 | 86830 | 84.84 | 87346 |
| 1 | 94.26 | 8960 | 93.39 | 9180 | 92.93 | 9531 | 92.21 | 9266 | 92.21 | 9235 | 92.21 | 8985 |
| 2 | 84.16 | 8018 | 84.83 | 8325 | 85.87 | 8581 | 85.08 | 8392 | 82.76 | 8377 | 82.76 | 8393 |
| 3 | 94.76 | 8112 | 94.76 | 7820 | 95.07 | 8004 | 96.27 | 7768 | 94.41 | 7722 | 95.51 | 7675 |
| 4 | 79.25 | 7613 | 79.86 | 7771 | 78.45 | 7961 | 80.92 | 7863 | 80.95 | 7878 | 81.50 | 7862 |
| 5 | 82.01 | 7947 | 83.86 | 8365 | 84.50 | 8668 | 82.37 | 8596 | 61.63 | 8502 | 75.70 | 8643 |
| 6 | 70.41 | 7400 | 68.06 | 7692 | 72.42 | 7940 | 69.70 | 7706 | 71.18 | 7753 | 71.18 | 7784 |
| 7 | 94.36 | 7852 | 93.94 | 8117 | 92.25 | 8322 | 93.88 | 8221 | 92.15 | 8190 | 92.15 | 8144 |
| 8 | 85.02 | 6768 | 85.21 | 6982 | 85.24 | 7021 | 85.59 | 7114 | 83.28 | 7098 | 84.31 | 7052 |
| 9 | 86.40 | 6631 | 86.68 | 6778 | 83.48 | 6896 | 85.85 | 6973 | 87.93 | 7052 | 84.26 | 6973 |
| 10 | 53.18 | 7443 | 50.63 | 7653 | 53.47 | 7634 | 85.02 | 7893 | 85.02 | 7815 | 68.49 | 8222 |
| 11 | 77.502 | 7033 | 72.24 | 7801 | 75.10 | 7597 | 76.35 | 7379 | 76.38 | 7395 | 63.70 | 7613 |

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Ransac 3000  Feature 5000 | | Ransac 3000  Feature 4000 | | Ransac 1000  Feature 3000 | | Ransac 750  Feature 2000 | | Ransac 3000  Feature 1000 | | Ransac 250  Feature 500 | |
| dataset | [%] | Cas  [ms] | [%] | Cas  [ms] | [%] | Cas  [ms] | [%] | Cas  [ms] | [%] | Cas  [ms] | [%] | Cas [ms] |
| celkom | 84.84 | 88977 | 80.89 | 72464 | 80.31 | 54649 | 80.82 | 36816 | 80.86 | 24024 | 78.21 | 19093 |
| 1 | 92.21 | 10328 | 92.86 | 7785 | 93.23 | 5663 | 93.23 | 4009 | 92.00 | 2902 | 92.68 | 2386 |
| 2 | 85.08 | 8767 | 83.79 | 7083 | 82.33 | 5335 | 82.74 | 3557 | 84.58 | 2309 | 84.90 | 1731 |
| 3 | 96.27 | 8234 | 94.70 | 6506 | 95.85 | 5086 | 95.17 | 3260 | 94.80 | 2075 | 95.82 | 1451 |
| 4 | 80.92 | 7815 | 81.05 | 6603 | 80.39 | 4774 | 78.72 | 3120 | 80.37 | 1965 | 79.25 | 1498 |
| 5 | 82.37 | 8424 | 81.33 | 7145 | 82.36 | 5336 | 82.41 | 3526 | 81.62 | 2309 | 80.65 | 1794 |
| 6 | 69.70 | 7676 | 67.48 | 6552 | 70.09 | 4930 | 66.87 | 3229 | 69.59 | 2059 | 69.25 | 1544 |
| 7 | 93.88 | 8206 | 94.19 | 6646 | 92.66 | 4914 | 92.97 | 3230 | 93.87 | 2012 | 92.99 | 1529 |
| 8 | 85.59 | 7160 | 85.72 | 5772 | 83.62 | 4431 | 84.04 | 3042 | 86.11 | 1903 | 82.95 | 1482 |
| 9 | 85.85 | 6989 | 86.65 | 5818 | 86.48 | 4415 | 85.50 | 2964 | 85.66 | 1887 | 85.38 | 1450 |
| 10 | 85.02 | 8015 | 49.82 | 6318 | 51.99 | 4836 | 51.87 | 3432 | 50.24 | 2403 | 27.04 | 2512 |
| 11 | 76.35 | 7363 | 72.15 | 6225 | 64.38 | 4929 | 75.52 | 3447 | 70.62 | 2200 | 69.40 | 1716 |

Doterajšie výslekdy boli pozorované na po spojení obrázkov tak, aby na výslednom obrázku bola veľkosť písma v pixloch 12. Velkost pismen v datasete je 20 pixelov.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Ransac 3000  Feature 500  Velkost 20 | | Ransac 3000  Feature 1000  Velkost 20 | |
| dataset | [%] | Cas  [ms] | [%] | Cas  [ms] |
| celkom | 89.67 | 39246 | 92.82 | 53140 |
| 1 | 97.73 | 4014 | 98.17 | 5196 |
| 2 | 94.15 | 3698 | 95.75 | 5282 |
| 3 | 98.56 | 3221 | 98.58 | 4437 |
| 4 | 88.83 | 3330 | 94.73 | 4412 |
| 5 | 93.85 | 3964 | 93.16 | 5807 |
| 6 | 87.95 | 3590 | 94.87 | 4610 |
| 7 | 96.28 | 3393 | 97.21 | 4325 |
| 8 | 97.76 | 3721 | 94.5 | 4212 |
| 9 | 94.08 | 3158 | 71.35 | 5245 |
| 10 | 66.36 | 2766 | 91.17 | 5141 |
| 11 | 76.35 | 3205 | 72.15 | 6225 |