

# Lokálne príznaky vo farebných obrazoch

Paula Budzáková\*

Školiteľ: Elena Šikudová†

Katedra aplikovanej informatiky, FMFI UK, Mlynská Dolina 842 48 Bratislava

**Abstrakt:** V tomto článku prezentujeme metódu na extrakciu lokálnych príznakov z farebného obrazu. Ako základný model pre metódu sme si vybrali prístup zameraný na ľudský vizuálny systém, s využitím chromatických oponentných kanálov a metódy SIFT. Ideou riešenia je zakomponovanie chromatických oponentných kanálov nahradením šedotónových informácií v metóde SIFT tak, že kľúčové body sú detegované na dvoch separovaných oponentných kanáloch. Pre nájdené zaujímavé body v oboch kanáloch sú vytvorené príznakové vektory, ktoré sú pre následné príznakové párovanie zreťazené.

*Kľúčové slová:* lokálne príznaky, SIFT, SURF, farebný obraz, teória oponentných farieb

## 1 Úvod

Aj keď je farba zvyčajne vnímaná ako nenahradiateľný prvok opisujúci svet okolo nás, techniky na extrakciu lokálnych príznakov sú najčastejšie založené na popise tvaru a úplne ignorujú farebnú informáciu. Farebný obraz vieme opísať pomocou príznakov farby s využitím histogramu jednotlivých farebných kanálov. Ak chceme extrahovať príznaky nejakého objektu v obraze, voľba histogramu nie je veľmi vhodná, pretože histogram opisuje obraz ako celok, nerozlišuje popredie od pozadia. Tým extrahuje hodnotu alebo vektor hodnôt pre celú maticu pixelov obrazu. Lokálne príznaky hľadajú v obraze zaujímavé body, a potom extrahujú príznaky z okolia týchto bodov, čím obraz popíšu takým množstvom hodnôt, koľko je zaujímavých bodov v obraze. Existuje množstvo metód, ktoré už aspoň čiastočne využívajú farebnú informáciu z obrazu. V našom výskume sme sa zamerali na analyzovanie existujúcich modifikácií metód SIFT a SURF. Navrhli sme metódu, ktorá sa pri extrakcii lokálnych príznakov z farebného obrazu zameriava na ľudský vizuálny systém. V závere vyhodnocujeme úspešnosť navrhnutých metód porovnaním výsledkov metódy SIFT pracujúcej na šedotónovom obraze.

## 2 Lokálne príznaky

Lokálne príznaky opisujú body obrazu, ktoré sa bezprostredne odlišujú od svojho okolia. Tie sú najčastejšie spojené so zmenou farby, intenzity alebo textúry. Počet extrahovaných lokálnych príznakov z obrazu závisí od počtu detegovaných zaujímavých bodov v obraze. Extrakcia lokálnych príznakov z obrazu je zložená z dvoch zložiek, a to *detektora*, ktorý vyhľadáva zaujímavé body a *deskriptora*, ktorý následne nájdené zaujímavé body a ich okolie opíše vo forme príznakov. Metódy na extrakciu lokálnych príznakov sú najčastejšie invariantné voči afínnym transformáciám.

### 2.1 Detektor

Detegované body v obraze môžu predstavovať rohy, hrany alebo vzory, ktoré sa s vysokou pravdepodobnosťou opakujú v obraze. Zvyčajne sa pracuje na menších častiach obrazu, ktoré si detektor vytvorí rozdelením obrazu na regióny pixelov, s ktorými sa ďalej zaoberá ako s maticami. Metódy detektorov sú navrhnuté tak, aby našli konkrétne typy príznakov, a preto neexistuje jediná metóda, ktorá by bola vhodná na spracovanie všetkých rôznych typov obrazov. Na základe typu príznaku extrahovaného z obrazu rozlišujeme tri hlavné triedy detektorov [Tuytelaars and Mikolajczyk, 2008]: rohové detektory, detektory škrvny a detektory regiónu.

*Rohové detektory* (Harrisov detektor rohov [Haralick and Shapiro, 1992], SUSAN detektor [Smith and Brady, 1997]) nemusia nutne detegovať len tie body obrazu, ktoré ľudským okom pôsobia ako roh. Bod detegovaný v 2D obraze ako roh predstavuje bod s vysokou krivosťou, v 3D projekcii rohov to ale nutne platiť nemusí. Roh sa môže nachádzať na vysoko štrukturovanom povrchu textúry, uzavretých hraniciach a podobne. Pre mnoho aplikácií je to postačujúce, či už ide o skutočný roh alebo nie, je považované za irelevantné. *Detektory škrvny* (SURF [Bay et al., 2008], SIFT [Lowe, 2004] Hessian detektor [Tuytelaars and Mikolajczyk, 2008]) sú v istom zmysle komplementárne k rohovým

\*budzakova9@uniba.sk

†sikudova@fmph.uniba.sk

detektorom, v dôsledku čoho sa často používajú spoločne, čím sa dosiahne lepšie pokrytie obrazu. Vo všeobecnosti platí, že štruktúra škvرنy sa v obrazovej rovine lokalizuje horšie ako roh, aj keď ich škála a tvar sú definované lepšie. Na lokalizovanie rohu nám stačí jeden bod, pri škvرنe môže byť lokalizovaná celá jej hranica, ktorá je zvyčajne nepravidelného tvaru a poskytuje dobrý odhad o veľkosti a škále škvرنy. Detegované zaujímavé body *detektorom regiónu* (IBR [Tuytelaars and Van Gool, 2000], MSER [Matas et al., 2004]) zvyčajne predstavujú homogénne oblasti, ktoré sú štruktúrou často podobné ako pri detekcii detektorom škvرنy, čo môže spôsobiť problémy neskôr pri deskripcii a párovaní. Pre odstránenie možných komplikácií je potrebné pri výpočte deskriptora použiť zväčšenú škálu, ktorá pokrýva časť okolitých štruktúr obrazu a taktiež zachytáva tvar hranice oblasti.

Každý detektor z týchto tried síce deteguje zaujímavé body svojím špecifickým spôsobom, no dajú sa kombinovať, čím sa dopĺňajú a fungujú lepšie. Vo všeobecnosti platí, že zaujímavý bod musí byť ľahko nájditel'ný a ideálne rýchlo vypočítaný, čiže by mal mať dobrú pozíciu na opísanie príznakov deskriptorom. Pre zabránenie nožnej falošnej detekcie je množina detegovaných zaujímavých bodov odfiltrovaná od nekvalitných zaujímavých bodov.

## 2.2 Deskriptor

Po detekcii zaujímavých bodov je potrebné vytvoriť príznakový vektor deskriptorom [Krig, 2014]. Môže byť vypočítaný pomocou každého bodu záujmu, čím poskytne viac informácií o regióne pixelov z okolia zaujímavého bodu. Niektoré metódy informácie o regióne nepotrebujú. Využívajú mriežku nemenej veľkosti (kruhového alebo štvorcového tvaru), ktorá určuje, z akého rozsiahleho okolia bodu bude deskriptor vytvorený. Veľké množstvo deskripčných metód odlišujeme spôsobom vytvárania príznakového vektora.

*Lokálne binárne deskriptory* (FREAK, BRISK [Leutenegger et al., 2011]) vytvárajú binárne kódovaný bitový deskriptor, v ktorom každý bit zodpovedá jednému výsledku binárneho testu. *Spektrálne deskriptory* (SIFT [Lowe, 2004], SURF [Bay et al., 2008]) využívajú širokú škálu spektrálnych hodnôt, ako gradienty a priemery oblasti. Najčastejšie využívané spektrum je gradient lokálnej oblasti využitý v SIFT-e. *Bázové priestorové*

*deskriptory* vektor príznakov kódujú do množiny básových funkcií ako Fourierov rad z funkcií sínus, kosínus. *Mnohouhlnkové deskriptory* využívajú tvar objektu zmeraný štatistickými metrikami ako plocha, ťažisko, obvod. Tvary sú následne extrahované pomocou morfológických operácií a algoritmov popsujúcich oblasti.

Pre správnu extrakciu lokálnych príznakov musí byť metóda detektoru a deskriptoru zvolená tak, aby spolu vedeli dobre spolupracovať a zároveň dobre fungovať na zvolenom type obrazu, ktorý spracovávajú. Atribúty robustnosti ako kontrast, škála a rotácia musia vybraný detektor a deskriptor považovať za zohranú dvojicu. Pre analýzu šedotónového obrazu je prioritnou oblasťou osvetlenie, ktoré je oddelené od farby a farebného priestoru. Pri spracovaní farebného obrazu s analýzou farby je presnosť farby veľmi kritická. Je nutné sa zamerať na farebný priestor, farebné kanály a bitovú hĺbku farby. Extrakciu príznakov môže skomplikovať aj zlá kvalita obrazu (šum, rozostrenie), ktorú je potrebné odstrániť. Časťým problémom je aj rozlíšenie a škála. Na odstránenie tohto problému môže byť potrebný viac ako jeden spôsob deskriptora.

## 2.3 Párovanie

Deskriptor je základňou pre príznakové párovanie, ktoré vedie k porozumeniu obrazu, analyzovaniu scény, či k sledovaniu objektov. Hlavným problémom párovania je, ako určiť príznak, ktorý je odlišný od podobných príznakov a zároveň je súčasťou väčšieho objektu. Vybratie správnej metódy párovania je veľmi náročné z dôvodu opakovanosti, robustnosti, výpočtovej náročnosti alebo veľkosti pamäte. Relatívne páry sú určené funkciou vzdialenosti, kde prichádzajúce množiny príznakových deskriptorov sú porovnávané s množinou už známych deskriptorov.

## 3 Existujúce metódy

V rámci práce sme preskúmali nájdené existujúce metódy na extrakciu lokálnych príznakov z farebného obrazu. Zamerali sme sa na metódy SIFT a SURF, ktoré sú predstavené v tejto kapitole.

Na základe jednotlivých využití farebných priestorov bolo navrhnutých niekoľko modifikácií metódy SIFT a SURF s využitím farebnej informácie. V [Li et al., 2014] autori pridávajú RGB farebné informácie do SIFT deskriptora počítaním

klasickej metódy SIFT na všetkých troch kanáloch RGB zvlášť, čím vznikne deskriptor veľkosti  $3 \times 128$  a zachytí viac informácií o farbe. Rovnaký princíp je využitý s modelom HSV [Bosch et al., 2008] alebo YCbCr [Li et al., 2014]. V [Li et al., 2011] [Jalilvand et al., 2011] na získanie farebnej informácie využívajú lokálny kernel histogram, ktorý následne zakomponujú do deskriptora. Ďalšou variantou je vytvorenie si vlastného farebného modelu v [Cui et al., 2010] s využitím funkcie 3D farebného priestoru. Niektoré metódy pracujú aj so šedotónovým aj s farebným obrazom paralelne a až pri príznakovom párovaní spoja získané informácie. Pri technike najbližšieho suseda je najčastejšie využívaná euklidovská alebo vzdialenosť Bhattacharyya.

## 4 Implementácia

V rámci našej práce sme navrhli metódu, ktorá predstavuje modifikáciu metódy SIFT s využitím farebnej informácie z obrazu, ktorú popisujeme v tejto kapitole. V prvej časti je popísaná úprava farebného obrazu pre detekciu zaujímavých bodov, ďalej samotná detekcia bodov, vytváranie deskriptora a príznakové párovanie.

### 4.1 Získanie farebného príznaku

Farby sú veľmi dôležité pre ľudské vnímanie a pozornosť. Ich vnímanie spôsobuje reakcia oka na elektromagnetické vlny určitej dĺžky. V [Jost et al., 2005] pracujú s obrazom v RGB farebnom priestore upravenom podľa ľudského vizuálneho systému na spracovanie farby, ktoré prebieha v hlbších štruktúrach sietnice a mozgovej kóry s charakterom oponentných procesov. Príznakové mapy [Jost et al., 2005], ktorými sme sa inšpirovali, využívajú na spracovanie farby oponentný princíp kódovania farieb [Hurvich and Jameson, 1957]. Sú vytvorené pomocou chromatických oponentných kanálov a intenzity obrazu, ktorá nahrádza achromatický čierno-biely oponentný kanál.

V našej metóde je intenzita obrazu  $I$  získaná z váhovaného súčtu farebných kanálov RGB obrazu s rozdielnymi koeficientmi pre každý kanál, na základe zmeraných intenzít vnímania ľudí bez poruchy trichromatického vnímania farieb.

$$I = 0.2126 * R + 0.7152 * G + 0.0722 * B \quad (1)$$

Už ako môžeme vidieť podľa hodnôt koeficientov v 1, ľudský zrak je najcitlivejší na zelenú farbu a naopak, najmenej citlivý na farbu modrú.

V [Engel et al., 1997] skúmali ľudský vizuálny cortex a jeho farebné ladenie pomocou magnetickej rezonancie. Experimentmi zistili, že najsilnejšia reakcia vo vizuálnej kortikálnej oblasti mozgu je na červeno-zelené podnety a na žlto-modré podnety, pri ktorých, oproti červeno-zeleným podnetom, ale podnet rýchlo klesá.

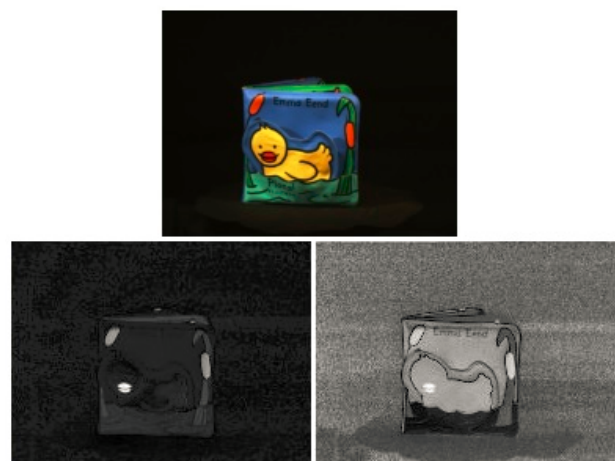
Ako základný model pre úpravu farebného obrazu sme si vybrali prístup založený na experimentoch s využitím chromatických oponentných kanálov a trichromatickej teórie farieb na získanie intenzity. Obraz je kombináciou týchto dvoch teórií rozdelený na červeno-zelenú zložku obrazu (oponentný filter  $R^+G^-$ ) a žlto-modrú zložku obrazu (oponentný filter  $B^+Y^-$ ) ako:

$$R^+G^- = \frac{(R - G)}{I} \quad (2)$$

$$B^+Y^- = \frac{(B - Y)}{I} \quad (3)$$

$$Y = \frac{(R + G)}{2} \quad (4)$$

Detekcia zaujímavých bodov prebieha separovateľne na vyexpertovaných obrazoch zo vstupného obrazu, znázornených na obrázku 1



Obr. 1: Rozdelenie vstupného obrazu na červeno-zelený (vľavo dole) a žlto-modrý oponentný kanál (vpravo dole).

## 4.2 Detekcia zaujímavých bodov

Detektor SIFT využíva na vyhľadávanie zaujímavých bodov škálovú pyramídu [Tuytelaars and Mikolajczyk, 2008] s využitím rozdielu Gaussiánov. Obrázky sú konvolované s rozdielnou škálou a opätovne filtrované Gaussiánom. Zaujímavé body reprezentujú lokálne extrémny rozdiel Gaussianov (Difference of Gaussians - DoG) vstupného obrazu v rôznych škálach s využitím farebnej informácie z oponentných filtrov obrazu. V DoG obraze sa následne hľadajú lokálne extrémny v okolí 3x3 aktuálneho obrazu a taktiež v okolí rovnakej veľkosti v nasledujúcom a predchádzajúcom obraze škálovej pyramídy. Nájdene lokálne extrémny (minimá alebo maximá) sú identifikované ako zaujímavé body. Dostatočne kontrastne zaujímavé body sú klasifikované ako kľúčové body, pre ktoré je následne vypočítaný deskriptor. Na obrázku 2 môžeme vidieť detekciu zaujímavých bodov našou metódou a pôvodnou metódou SIFT, ktorá pracuje na šedotónovom obraze.



Obr. 2: Detekcia zaujímavých bodov na červeno-zelenom kanáli (vľavo) a žltomodrom kanáli (vpravo)

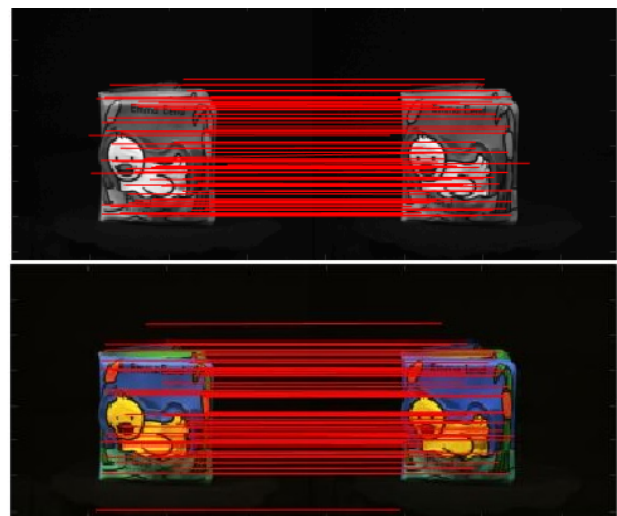
## 4.3 Deskriptor a príznakové párovanie

Pre výpočet deskriptora je najprv zistená dominantná orientácia každého kľúčového bodu, ku ktorej je relatívne vytvorený bodový deskriptor [Krig, 2014]. Následne sa obraz 16 bodmi navzorkuje a pre každú vzorku je vypočítaná veľkosť a orientácia farebného gradientu oponentných kanálov. Vypočítaným veľkostiam a orientáciám sú priradené váhy pomocou Gaussového kruhového okna, z ktorých sú vytvorené histogramy orientácií pre šestnásť regiónov z okolia kľúčového bodu. Pre každý kľúčový bod je vypočítaná hodnota deskriptora. Takto sú vypočítané deskriptory pre body detegované v červeno-zelenom kanáli a body detegované v žltomodrom kanáli zvlášť.

Pre následné príznakové párovanie sú tieto dve množiny deskriptorov zret'azené do jednej, obsahujúcej farebné informácie z oboch kanálov.

Príznakové párovanie je zrealizované pomocou techniky *najbližšieho suseda*. V rámci jednej dvojice deskriptorov (deskriptor originálneho obrazu a deskriptor porovnávaného obrazu) technika najbližšieho suseda nájde najbližšieho suseda a druhého najbližšieho suseda pre deskriptor každého kľúčového bodu (pre každý riadok v poli z dvojice množín deskriptorov). Za najbližšieho suseda je identifikovaný taký riadok v poli deskriptorov, ktorý má najmenšiu euklidovskú vzdialenosť.

Pri porovnávaní technikou najbližšieho suseda je dôležitou hodnotou počet zhôd. Za zhodu považujeme ten prípad, keď pre riadok v poli deskriptorov originálneho obrazu platí, že euklidovská vzdialenosť najbližšieho suseda je menšia ako euklidovská vzdialenosť druhého najbližšieho suseda vynásobená koeficientom 0.6 (hodnota koeficientu bola zadaná autorom SIFTu [Lowe, 2004]).

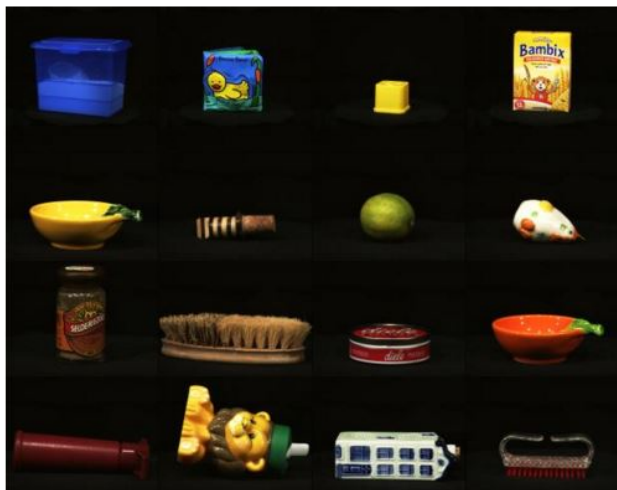


Obr. 3: Príznakové párovanie deskriptorov metódy SIFT (hore) a našej metódy (dole) pre objekty, ktoré majú zmenené teploty farby pomocou osvetlenia.

## 5 Analýza a výsledky

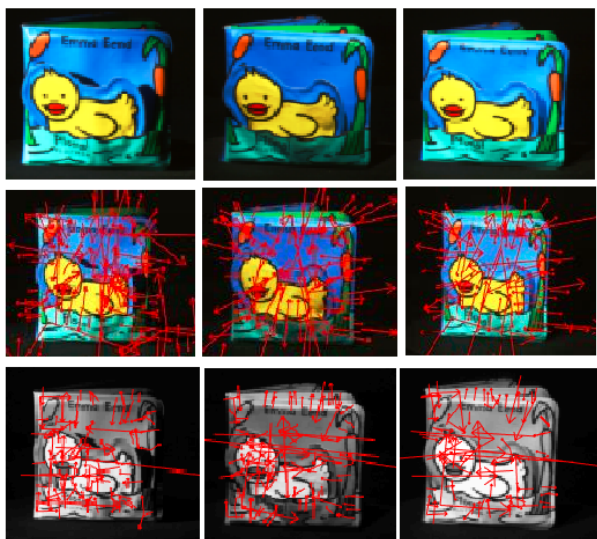
Na analýzu a zhodnotenie navrhnutej metódy bola vybraná databáza ALOI z [Geusebroek et al., 2005] (ukážka databázy je znázornená na obrázku 4), ktorá obsahuje veľké množstvo objektov v rôznych zobrazovacích podmienkach, a to v rôznych smeroch osvetlenia, v rôznej intenzite osvetlenia, zmena tep-

loty farby pomocou osvetlenia a rotovanie objektu. Použili sme 5 sád obrázkov s osvetlením v rôznych smeroch a 5 sád obrázkov so zmenou teploty farby. Obrázky boli vybrané tak, aby boli pokryté všetky farby.



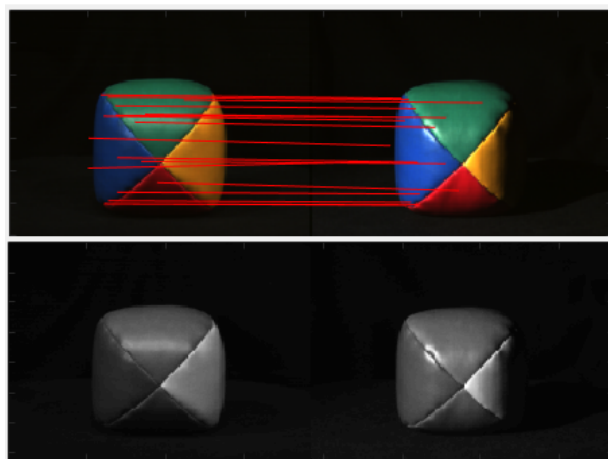
Obr. 4: Ukážka obrázkov z použitej databázy

Hlavným kritériom metódy bola fotometrická robustnosť deskriptora, čiže invariancia na zmeny svetelného zdroja, tieňa, tieňovania a farby. Rozhodli sme sa metódu porovnávať s pôvodnou metódou SIFT pracujúcou na šedotónovom obraze.



Obr. 5: Detegované body záujmu pre konkrétny objekt pod iným smerom a intenzitou osvetlenia. Hore sú pôvodné obrázky, v strede je detekcia našou metódou (zaujímavé body z oboch kanálov) a dole metódou SIFT na šedotónovom obraze.

Obrázok 5 ukazuje detegované zaujímavé body metódou SIFT a nami navrhovanou metódou. Analýzou sme zistili, že metóda dostatočne spĺňa vlastnosti fotometrickej robustnosti. Ako je vidieť, počet nájdených bodov v oponentných kanáloch je väčší ako v šedotónových obrazoch. Je známe, že zvyšovaním počtu detegovaných bodov sa zvyšuje aj výkonnosť rozpoznávacieho procesu, čím môžeme povedať, že z prvého pohľadu na obrázok 5 dosahuje naša metóda lepšie výsledky pokiaľ ide o počet detegovaných zaujímavých bodov. V príznakovom párovaní podáva naša metóda v porovnaní s metódou SIFT lepšie výsledky. Síce nedosahujeme až taký počet zhôd, ktorý by sme očakávali, môžeme povedať, že počet zhôd pre použitie rozpoznávania alebo klasifikácie je dostačujúci. Nižší počet zhôd môže byť spôsobený filtrovaním pixelov pod určitou kontrastnou hranicou pred deskripciou, čím získavame menší počet príznakových deskriptorov pre následné párovanie. Lepšie výsledky dosahujeme pri obrazoch obsahujúcich zmenu teploty farby, znázornené na obrázku 3. Ako vidíme na obrázku 6, pri obrazoch so zmenou smeru osvetlenia, je počet zhôd nižší, no v porovnaní s metódou SIFT dosahuje lepšie výsledky. Celkovo môžeme zhodnotiť, že navrhnutá metóda spĺňa očakávané kritériá a je dostatočne vhodná pre rozpoznávanie objektov z obrazu.



Obr. 6: Príznakové párovanie pre objekty so zmenou smerom osvetlenia. Ako je vidieť, navrhnutá metóda pri tomto type obrázkov dosahuje lepšie výsledky, ako metóda SIFT, ktorá nenašla žiadnu zhodu.



## 6 Záver

Úlohou našej práce bol prehľad a analýza vybraných existujúcich prác, na základe ktorej sme si špecifikovali oblasť záujmu. Zvolili sme si cestu ľudského vizuálneho systému, ktorý sme použili ako základný model pre spracovanie farebného obrazu v metóde SIFT.

V práci boli vysvetlené potrebné teoretické vedomosti na extrakciu lokálnych príznakov. Bez týchto vedomostí by nebolo možné zvoliť si a navrhnúť korektnú metódu. Navrhovaná metóda spracuje obraz na základe trichromatickej teórie farieb a oponentných procesov a ich kombináciou extrahuje obraz, ktorý je následne ďalej spracovávaný metódou SIFT. Na vybranej databáze obrázkov sme ukázali, že metóda dosahuje oproti metóde SIFT lepší výkon, čo nás oprávňuje k ďalšiemu testovaniu a následnému rozvíjaniu navrhutej metódy.

## Pod'akovanie

Veľká vďaka patrí mojej školiteľke, RNDr. Elene Šikudovej, PhD., za pomoc, nasmerovanie, neoceniteľné nápady a veľkú trpezlivosť.

## Literatúra

- [Bay et al., 2008] Bay, H., Ess, A., Tuytelaars, T., and Van Gool, L. (2008). Speeded-up robust features (SURF). *Computer vision and image understanding*, 110(3):346–359.
- [Bosch et al., 2008] Bosch, A., Zisserman, A., and Muoz, X. (2008). Scene classification using a hybrid generative/discriminative approach. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 30(4):712–727.
- [Cui et al., 2010] Cui, Y., Pagani, A., and Stricker, D. (2010). Sift in perception-based color space. In *Image Processing (ICIP), 2010 17th IEEE International Conference on*, pages 3909–3912. IEEE.
- [Engel et al., 1997] Engel, S., Zhang, X., and Wandell, B. (1997). Colour tuning in human visual cortex measured with functional magnetic resonance imaging. *Nature*, 388(6637):68–71.
- [Geusebroek et al., 2005] Geusebroek, J.-M., Burghouts, G. J., and Smeulders, A. W. (2005). The amsterdam library of object images. *International Journal of Computer Vision*, 61(1):103–112.
- [Haralick and Shapiro, 1992] Haralick, R. M. and Shapiro, L. G. (1992). *Computer and Robot Vision*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1st edition.
- [Hurvich and Jameson, 1957] Hurvich, L. M. and Jameson, D. (1957). An opponent-process theory of color vision. *Psychological review*, 64(6p1):384.
- [Jalilvand et al., 2011] Jalilvand, A., Boroujeni, H. S., and Charkari, M. M. (2011). CH-SIFT: A local kernel color histogram sift based descriptor. In *Multimedia Technology (ICMT), 2011 International Conference on*, pages 6269–6272. IEEE.
- [Jost et al., 2005] Jost, T., Ouerhani, N., Von Wartburg, R., Müri, R., and Hügli, H. (2005). Assessing the contribution of color in visual attention. *Computer Vision and Image Understanding*, 100(1):107–123.
- [Krig, 2014] Krig, S. (2014). *Computer Vision Metrics: Survey, Taxonomy, and Analysis*. Apress.
- [Leutenegger et al., 2011] Leutenegger, S., Chli, M., and Siegwart, R. Y. (2011). Brisk: Binary robust invariant scalable keypoints. In *Computer Vision (ICCV), 2011 IEEE International Conference on*, pages 2548–2555. IEEE.
- [Li et al., 2011] Li, D., Ke, Y., and Zhang, G. (2011). A SIFT descriptor with local kernel color histograms. In *Mechanic Automation and Control Engineering (MACE), 2011 Second International Conference on*, pages 992–995. IEEE.
- [Li et al., 2014] Li, Q., Chen, J., Peng, Q., and Wu, X. (2014). Application of localized soft-assignment coding and CSIFT in image classification. In *Proceedings of International Conference on Internet Multimedia Computing and Service*, page 246. ACM.
- [Lowe, 2004] Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60(2):91–110.
- [Matas et al., 2004] Matas, J., Chum, O., Urban, M., and Pajdla, T. (2004). Robust wide-baseline stereo from maximally stable extremal regions. *Image and vision computing*, 22(10):761–767.
- [Smith and Brady, 1997] Smith, S. M. and Brady, J. M. (1997). Susan—a new approach to low level image processing. *International journal of computer vision*, 23(1):45–78.
- [Tuytelaars and Mikolajczyk, 2008] Tuytelaars, T. and Mikolajczyk, K. (2008). Local invariant feature detectors: a survey. *Foundations and Trends® in Computer Graphics and Vision*, 3(3):177–280.
- [Tuytelaars and Van Gool, 2000] Tuytelaars, T. and Van Gool, L. J. (2000). Wide baseline stereo matching based on local, affinity invariant regions. In *BMVC*, volume 1, page 4.