

UNIVERZITA KOMENSKÉHO V BRATISLAVE
FAKULTA MATEMATIKY, FYZIKY A
INFORMATIKY

LOKÁLNE PRÍZNAKY VO FAREBNÝCH
OBRAZOCH

Diplomová práca

UNIVERZITA KOMENSKÉHO V BRATISLAVE
FAKULTA MATEMATIKY, FYZIKY A
INFORMATIKY

LOKÁLNE PRÍZNAKY VO FAREBNÝCH
OBRAZOCH

Diplomová práca

Študijný program: Aplikovaná informatika
Študijný odbor: 2511 aplikovaná informatika
Školiace pracovisko: Katedra aplikovanej informatiky
Školiteľ: RNDr. Elena Šikudová, PhD.



Univerzita Komenského v Bratislave
Fakulta matematiky, fyziky a informatiky

ZADANIE ZÁVEREČNEJ PRÁCE

Meno a priezvisko študenta: Bc. Paula Budzáková
Študijný program: aplikovaná informatika (Jednoodborové štúdium, magisterský II. st., denná forma)
Študijný odbor: 9.2.9. aplikovaná informatika
Typ záverečnej práce: diplomová
Jazyk záverečnej práce: slovenský
Sekundárny jazyk: anglický


Názov: Lokálne príznaky vo farebných obrazoch
Local features in color images


Cieľ: Metódy na vyhľadávanie lokálnych príznakov vo farebných obrazoch:
 1. Naštudovanie
 2. Návrh zlepšenia
 3. Implementácia
 4. Porovnanie výsledkov

Vedúci: RNDr. Elena Šikudová, PhD.
Katedra: FMFI.KAI - Katedra aplikovanej informatiky
Vedúci katedry: doc. PhDr. Ján Rybár, PhD.
Dátum zadania: 20.10.2014

Dátum schválenia: 24.10.2014

prof. RNDr. Roman Ďurikovič, PhD.
garant študijného programu


študent


vedúci práce

Čestné vyhlásenie

Čestne prehlasujem, že som túto diplomovú prácu vypracovala samostatne s použitím citovaných zdrojov.

V Bratislave

.....

Pod'akovanie

Chcela by som sa pod'akovať mojej školiteľke, RNDr. Elene Šikudovej, PhD, za pomoc a rady pri navrhovaní a testovaní metódy a za obetovaný čas pri vylepšovaní celej diplomovej práce.

Abstrakt

Aj keď je farba vnímaná ako nenahraditeľný prvok opisujúci svet okolo nás, techniky na extrakciu lokálnych príznakov sú najčastejšie založené na popise tvaru a úplne ignorujú farebnú informáciu.

V práci navrhujeme metódu na extrakciu lokálnych príznakov z farebného obrazu. Ako základný model 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 zret'azené.

Kľúčové slová: *lokálne príznaky, oponentný proces, trichromatický proces, SIFT*

Abstract

Although the color is perceived as an irreplaceable element describing the world around us, the techniques for extracting of the local features are mostly based on the description of the shape - while the color information is being fully ignored. This paper proposes a method for extracting of the local features of the color image. As a basic model we have chosen the approach to the human visual system using chromatic opponent channels and the SIFT method. The idea of this solution is the incorporation of the opponent chromatic channels by replacing the grayscale information in the SIFT method, so that the key points are detected on two separate opponent channels. For the interesting points found in the two channels, the descriptors are formed which are then successively concatenated.

Keywords: *local features, opponent process theory, trichromatic theory, SIFT*

Obsah

Úvod	1
1 Prehľad problematiky	2
1.1 Lokálne príznaky	3
1.1.1 Lokálne vs. globálne príznaky	3
1.1.2 Lokálne príznaky	3
1.1.3 Detektor	4
1.1.4 Deskriptor	6
2 Prehľad existujúcich prác	10
2.1 Farebné modifikácie metódy SIFT	10
2.2 Farebné modifikácie metódy SURF	18
2.3 Analýza metód	23
3 Špecifikácia	25
3.1 Motivácia a cieľ práce	25
3.2 Softvér	26
3.3 Základný model	26
3.3.1 Získanie farebného príznaku	26
3.3.2 Detekcia zaujímavých bodov	28
3.3.3 Vytváranie deskriptora	29
3.3.4 Navrhnuté riešenie	30
4 Implementácia	31
4.1 Predspracovanie	33
4.2 Detektor a deskriptor	34
4.3 Príznakové párovanie	35
5 Validácia	37
5.1 Databáza	37

<i>OBSAH</i>	IX
5.2 Analýza	38
5.3 Klasifikácia obrázkov	39
5.4 Budúca práca	43
6 Záver	47
Literatúra	51

Zoznam obrázkov

1.1	Taxonómia pre rozmery deskriptora: hustota (density) príznaku, ako globálna, regionálna alebo rozhádzaná lokálna, tvar a vzor pixlov (shape a pattern) použitý na výpočet deskriptora, ako kruh alebo obdĺžnik, spektrum, ktoré obsahuje informácie obsiahnuté v samotných funkciách [18].	7
1.2	Príznakový vektor metódy SIFT [24]	8
2.1	Princíp metód na extrakciu lokálnych príznakov rozdelených na fázu detekcie a fázu deskripcie [26].	12
2.2	Prehľad navrhovanej CH-SIFT metódy. Farebný a šedý variant obrázkov je analyzovaný oddelene a párovanie je vykonávané s rozdielnymi váhami [15].	15
2.3	Párovanie pomocou metódy SIFT, C-SIFT a CH-SIFT. [15]	16
2.4	Nový farebný priestor a PC-SIFT deskriptor [8].	18
2.5	Aproximácie Gaussiánov druhej parciálnej derivácie v smere x,y a xy a vážené aproximácie box filtrov [4].	18
2.6	Určenie orientácie. Sumov zložiek odoziev získame vektor [10].	19
3.1	Rozdelenie vstupného obrazu na chromatické oponentné kanály.	27
3.2	Hľadanie lokálnych extrémov porovnávaním pixla v okolí $3 \times 3 \times 3$ [24].	29
4.1	Vývinový diagram navrhutej metódy.	32
4.2	Normalizované kanály použité ako vstup pre detekciu a deskripciu. Vľavo červeno-zelený kanál, vpravo žltomodrý	33
4.3	Princíp príznakového párovania pomocou metódy najbližšieho suseda.	35
5.1	Ukážka použitej databázy	37
5.2	Detekcia zaujímavých bodov metódou SIFT (vľavo) a navrhovanou metódou (vpravo)	38
5.3	Graf zobrazujúci počty detegovaných bodov a zhôd v obrázkoch kocky so zmenou smeru osvetlenia. X-ová os znázorňuje jednotlivé analyzované obrázky, Y-ová os predstavuje hodnoty detekcie a zhôd.	39

5.4	Matica zámen reprezentovaná teplotnou mapou pre klasifikáciu objektov pomocou navrhutej metódy	41
5.5	Matica zámen reprezentovaná teplotnou mapou pre metódu SIFT	42
5.6	Citlivosť a presnosť klasifikácie navrhnutou metódou na vybranej databáze.	43
5.7	Citlivosť a presnosť klasifikácie metódou SIFT na vybranej databáze. . . .	43
5.8	Výsledky I.	45
5.9	Výsledky II.	46

Zoznam tabuliek

2.1	Analýza farebných modifikácií metód z predchádzajúcich kapitol	24
5.1	Tabuľka vybraných objektov a ich počet obsiahnutý v databáze vybraných 303 obrázkov.	40
5.2	Percentuálne zastúpenie správnej klasifikácie objektu hračka pri jednotlivých prahoch navrhnutou metódou a metódou SIFT.	40
5.3	Percentuálne zastúpenie nesprávnej klasifikácie objektu hračka pri jednotlivých prahoch navrhnutou metódou a metódou SIFT	40

Úvod

Predstavme si, že by sme videli svet len čiernobielo. Dúha by bola len akýsi pás odtieňov šedej, západ slnka by nebol tak romantický, všetko okolo by stratilo akúsi tú iskru. Nechýbala by nám farba?

Farba v našom živote zohráva naozaj veľmi dôležitú rolu. Vďaka vnímaniu farby dokážeme spracovávať informácie s výraznejšími pocitmi a postrehmi. Rovnako je farba dôležitým prvkom aj pre rozlišovanie objektov. Ak by sme videli svet čiernobielo, nevedeli by sme rozlíšiť napríklad červené jablko od zeleného. Objekt jablko je síce rovnaký, no nemali by sme ich zaradiť do rovnakej triedy, aj keď pri šedotónovej reprezentácii by sme to pravdepodobne urobili.

Rovnako je to aj pri extrakcii lokálnych príznakov. Dokážeme zo šedotónového obrazu extrahovať všetky potrebné informácie o obraze? Vedeli by sme bez farebnej informácie správne klasifikovať červené jablko a nie zelené?

Všetky existujúce metódy, ktoré extrahujú lokálne príznaky zo šedotónového obrazu, dosahujú dobré výsledky. Tieto metódy sú využívané v širokom spektre počítačového videnia bez ohľadu na to, že nespracovávajú farbu. Vieme, že šedotónový obraz obsahuje veľké množstvo informácií na spracovanie obrazu. No ale, čo ak by farebná informácia zlepšila proces rozpoznávania alebo klasifikácie? Mnoho autorov navrhlo modifikácie existujúcich metód, ktoré využívajú aj farebnú informáciu z obrazu. Pristupujú k nej rôznymi spôsobmi, či už cez histogram alebo návrh vlastného farebného priestoru, do ktorého potom obraz prevedú. Každý návrh modifikácie je prispôbený aktuálnym potrebám autorov, napríklad obrázkom pod vodnou hladinou. No možno nie je najlepšou cestou zamerať sa na samotný obraz, ale skôr na jeho celkové vnímanie.

Cieľom tejto diplomovej práce je naštudovanie existujúcich metód na extrakciu lokálnych príznakov z farebného obrazu, návrh vlastnej metódy, implementácia a porovnanie dosiahnutých výsledkov s metódou pracujúcou na šedotónovom obraze.

Kapitola 1

Prehľad problematiky

V kapitole prehľadu problematiky sme zhrnuli základnú problematiku potrebnú pre výskum v tejto diplomovej práci. V prvej časti kapitoly predstavujeme príznaky. Ďalej sa venujeme lokálnym príznakom a postupom extrakcie. Pri detektoroch sme sa zamerali na rôzne metódy detekcie zaujímavých bodov rozdelených podľa konkrétneho typu príznakov. Kapitola deskriptoru popisuje rôzne typy deskriptorov na základe vytvárania vektora príznakov.

Pojem príznak v počítačovom videní definuje veľmi široké spektrum informácií a využití pri spracovaní obrazu. Pod pojmom príznak sa rozumie výsledok merania, ktoré kvantifikuje nejakú vlastnosť objektu, čiže extrahuje nejakú informáciu o obraze. Takto získané príznaky sa spájajú do vektora príznakov. Existuje veľa rôznych druhov príznakov. Príznaky v počítačovom videní definujeme do troch tried príznakov [35] na základe úrovne abstrakcie:

- *nízkoúrovňové príznaky* - opisujú základné vlastnosti, ako farbu, tvar, textúru atď.
- *stredoúrovňové príznaky* - vznikajú spojením nízkoúrovňových príznakov, ktorých pomocou vieme detegovať ľudskú tvár.
- *vysokoúrovňové príznaky* - príznaky najvyššej úrovne, opisujú sémantiku zobrazenej scény.

Každú úroveň delíme ešte na viaceré podkategórie príznakov na základe ich konkrétneho zamerania. Lokálne príznaky patria do úrovne nízkoúrovňových príznakov, spolu s globálnymi a bodovými príznakmi.

1.1 Lokálne príznaky

1.1.1 Lokálne vs. globálne príznaky

Pri rozpoznávaní objektov vo všeobecnosti rozlišujeme dve úlohy:

1. **Rozpoznávanie kategórií objektov**
2. **Rozpoznávanie konkrétnych objektov**

Bolo navrhnutých mnoho globálnych príznakov, ktoré opisujú celý obraz alebo len vysegmentované objekty obrazu. Medzi globálne príznaky patria, napríklad: dominantná farba, tvar oblasti, textúra a pod. Farebný obraz vieme opísať pomocou príznakov farby, využitím histogramu jednotlivých farebných kanálov. Tento prístup je veľmi dobrý výber pri obrazoch, ktoré obsahujú výrazné farby. Ak chceme extrahovať príznaky nejakého objektu v obraze, voľba histogramu farieb nie je vhodná. Histogram opisuje obraz ako celok, nerozlišuje popredie od pozadia a extrahuje jednu hodnotu alebo vektor hodnôt pre celú maticu pixlov obrazu. Na rozdiel od globálnych príznakov, lokálne príznaky hľadajú v obraze zaujímavé body, a potom extrahujú príznaky z okolia týchto bodov. Obraz je tak popísaný takým množstvom hodnôt, aký je počet zaujímavých bodov v obraze.

1.1.2 Lokálne príznaky

Lokálne príznaky opisujú body obrazu, ktoré sa bezprostredne odlišujú od svojho okolia. Tieto odlišnosti sa najčastejšie spájajú so zmenou vlastností obrazu, ako zmena intenzity, farby alebo textúry. Počet extrahovaných lokálnych príznakov z obrazu závisí od počtu detegovaných zaujímavých bodov v obraze. Ak máme obraz obyčajnej bielej steny bez zachytenia akýchkoľvek tieňov, nevieme z neho extrahovať žiadne lokálne príznaky, pretože na obraze nie je možné nájsť žiadny lokálny zaujímavý bod.

Lokálne príznaky sú napríklad: body záujmu (*interest points*), významné body (*salient points*), príznaky oblasti (*region-based features*). Proces extrakcie lokálnych príznakov z obrazu sa skladá z dvoch častí:

1. **Detektor** - vyhľadáva zaujímavé body v obraze
2. **Deskriptor** - opisuje zaujímavé body a ich okolie vo forme príznakov

Väčšina metód na získavanie lokálnych príznakov je založená na invariantnosti príznakov vzhľadom na nejakú transformáciu. To znamená, že tieto metódy vedú vytvárať opisné vektory zaujímavých bodov, ktoré zostanú rovnaké, aj keď hľadaný objekt bude posunutý, škálovaný, otočený a podobne. Pri lokálnych príznakoch sa najčastejšie zaoberáme

metódami invariantnými voči afinným transformáciám.

Lokálne invariantné príznaky sú významným a veľmi silným nástrojom, ktorý je úspešne aplikovaný v širokej škále aplikácií a systémov.

1.1.3 Detektor

Detektor slúži na vyhľadávanie zaujímavých bodov v obraze. Môžu to byť rohy, hrany alebo vzory, ktoré sa s vysokou pravdepodobnosťou opakujú v obraze. Vo všeobecnosti platí, že zaujímavý bod musí byť ľahko nájditeľný a ideálne rýchlo vypočítaný, to znamená, že by mal mať dobrú pozíciu na vypočítanie príznakového deskriptora. V skratke, zaujímavý bod definujeme ako kvalifikátor alebo kľúčový bod, z ktorého okolia vieme vypočítať príznaky, vypočítať vektor orientácie alebo vytvoriť deskriptor. Neisté alebo nekvalitné kľúčové body sú pred vytvorením deskriptora odfiltrované, aby sa zabránilo "falošnej detekcii".

V súčasnosti sú najviac využívané detektory, ktoré detegujú zaujímavé body na šedotónových obrazoch viacerými spôsobmi. Existujú rôzne koncepty metód vyhľadávania zaujímavých bodov. Detektor zvyčajne pracuje s menšími oblasťami obrazu. Obraz si rozdelí na regióny pixlov veľkosti, napr. 3x3, s ktorými ďalej zaobchádza ako s maticami. Metódy detektorov sú navrhnuté tak, aby našli konkrétne typy príznakov, a preto neexistuje jedna metóda, ktorá by bola vhodná na spracovanie všetkých typov obrazov. Na základe typu príznaku extrahovaného z obrazu sa detektory klasifikujú do troch hlavných tried [31]:

1. **Rohové detektory** (corner detectors)
2. **Blob detektory** (blob detectors)
3. **Detektory regiónu** (region detectors)

Rohové detektory

Termín roh, v tomto prípade použitia, má veľmi špecifický význam. V 2D obraze detegovaný bod, ako roh, predstavuje bod s vysokou krivosťou. Pri 3D projekcii rohov to ale nutne platiť nemusí. Roh sa môže nachádzať na vysoko štruktúrovanom povrchu textúry, uzavretých hraniciach a podobne. Pre mnoho praktických aplikácií je to postačujúce, či už ide o skutočný roh alebo nie, je považované za irelevantné.

Medzi rohovými detektormi [31][13] patrí napríklad *Harrisov detektor rohov*, ktorý využíva deriváciu šedotónového obrazu. Cieľom Harrisovho detektora rohov je pomocou kovariančnej matice derivátov intenzity obrazu v smere x a y klasifikovať bod ako roh. Pre detekciu rohov sú dôležité vlastné hodnoty kovariančnej matice, ktoré vyjadrujú variáciu v dvoch hlavných smeroch. Ak je variácia v oboch smeroch veľká, bod leží na rohu, ak je veľká len jedna z nich, bod leží na hrane.

SUSAN detektor [29] je ďalší z kategórie rohových detektorov, ktorý používa morfológické

prístupy. Pre každý pixel obrazu posudzuje kruhové okolie pevnej veľkosti susedných pixlov. Pixel, ktorý sa nachádza v centre tohto okolia, sa nazýva jadro. Hodnota intenzity tohto jadra predstavuje referenčnú hodnotu, s ktorou následne porovnáva hodnoty intenzity pixlov z jeho okolia. Pomocou lokálnych miním určí rohy v obraze.

Detektory škvorny

Detektory škvorny [31] sú v istom zmysle komplementárne k rohovým detektorom, v dôsledku čoho sa často používajú spoločne. Pri použití niektorých doplnujúcich funkcií detektorov sa obraz lepšie pokryje. Vo všeobecnosti platí, že štruktúra škvorny sa v obrazovej rovine horšie lokalizuje ako roh, aj keď ich škála a tvar sú definované lepšie. Na lokalizovanie rohu stačí jediný bod, zatiaľ čo pri škvorne môže byť lokalizovaná celá jej hranica. Aj keď je hranica zvyčajne nepravidelného tvaru, poskytuje dobrý odhad o veľkosti a škále škvorny. Detektory škvorny sú využívané v rôznych aplikačných oblastiach, no sú menej vhodné na 3D rekonštrukcie.

Detektor metódy *SURF* môže byť tiež kategorizovaný ako detektor škvorny, ako aj *Hessian detektor*, ktorý sa považuje za doplnkovú metódu k Harrisovmu detektoru rohov. Na miesto derivátov intenzity obrazu využíva Taylorov rozvoj intenzity obrazu, s využitím vyhladených Gausiánov obrazových derivátov druhého rádu. Pomocou Laplacianovho filtra a filtra rozdielov Gausiánov nachádza v obraze lokálne maximá, ktoré sa nachádzajú v blízkosti obrysov alebo hrán. Deriváty druhého rádu sú symetrické filtre, čím dávajú slabé reakcie presne v bode, kde je najvýznamnejšia zmena signálu. Z tohto dôvodu sú lokálne maximá lokalizované na hranici škvorny, ktoré zodpovedajú veľkosti Gaussového jadra podľa veľkosti štruktúry škvorny.

Detektory regiónu

Lokálne príznaky detegované pomocou detektorov regiónu zvyčajne predstavujú homogénne oblasti, štruktúrou často podobné, ako pri detekcii detektorom škvorny. Pri detekcii to nemusí predstavovať žiadny problém, no neskôr pri deskripcii a párovaní môžu vzniknúť problémy, ktoré sa ale dajú prekonať zvýšením meracích oblastí. Čiže použijeme zväčšenú škálu oblasti na výpočet deskriptora tak, aby obsahovala časť okolitých štruktúr obrazu a taktiež zachytávala tvar hranice oblasti.

Detektory regiónu [31] sú v lokalizácii obvykle presné. Veľmi dobre pracujú najmä na obrazoch s dobre štruktúrovanou scénou, ako obrazy obsahujúce objekty s textúrou alebo budovy. Medzi detektory regiónu zaradíme *Intensity-based Region (IBR)* [32], ktorý vychádza z extrémov intenzity, detegovaných v rôznych mierkach. Pri každom extréme sa zaujíma o obraz v jeho okolí, ktoré vytvára radiálnym spôsobom. Vzniknuté oblasti ľubovoľného tvaru nahrádza elipsami, a to tak, že z daného lokálneho extrému intenzity

vyžarujú lúče, pozdĺž ktorých sa skúma funkcia intenzity. Ďalšou metódou je *MSER* (Maximally Stable Extremal Regions) [25], ktorá najlepšie funguje na štrukturovaných obrazoch, ktoré je možné dobre rozdeliť. V najlepšom prípade, ak sú na snímke jednotlivé oblasti oddelené silnými zmenami intenzity. V súčasnosti je považovaný za najúčinnjší afinne invariantný detektor.

Oba spomenuté detektory extrahujú malé oblasti, na ktorých vystupuje intenzita obrazu, s ohľadom na ich bezprostredné okolie. Tento prístup pripomína tradičné segmentovanie obrazu. Avšak segmenty obrazu sú typicky relatívne veľké, príliš veľké na to, aby boli použité ako lokálne príznaky. Zvýšený počet segmentov na reprezentácii nového obrazu je možné získať tam, kde segmenty obrazu sú kompromisom medzi lokálnou a rozlišovacou spôsobilosťou. Takéto zoskupenia pixlov do regiónov označujeme ako *superpixle*, ktoré možno považovať za prirodzenejšiu a vnemovo zmysluplnejšiu alternatívu pre pôvodné pixle obrazu. Hoci *superpixle* zdieľajú niektoré charakteristiky detektorov regiónu, nie sú rovnaké. Urýchľujú analýzu obrazu tým, že namiesto analyzovania všetkých pixlov obrazu sa analyzujú *superpixle*.

Každý detektor z týchto tried deteguje zaujímavé body špecifickým spôsobom. Postupným prechádzaním tried prichádzame k rôznym prepojeniam medzi nimi. Dajú sa kombinovať, čím sa dopĺňajú a fungujú lepšie. Ďalšie metódy, ktoré by sme mohli zaradiť do triedy efektívne implementovaných detektorov, ako SIFT, SURF, ORB, sú viac spomenuté v nasledujúcej kapitole.

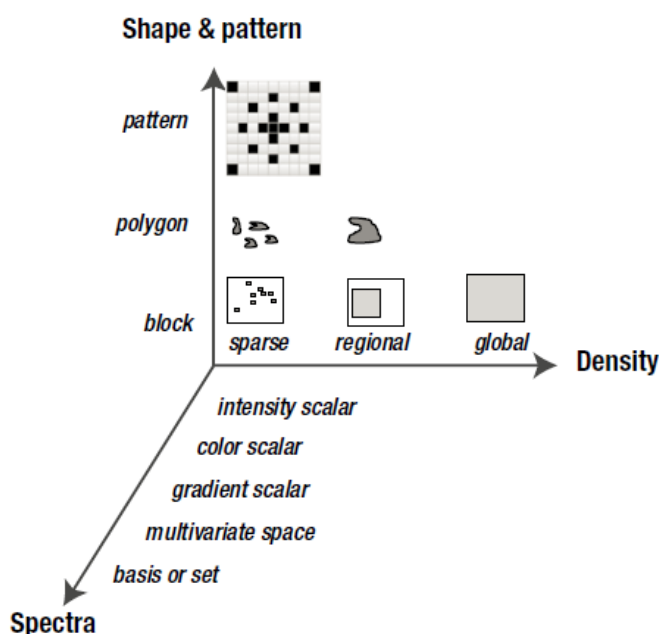
1.1.4 Deskriptor

Príznakový deskriptor môže byť vypočítaný pomocou každého bodu záujmu, čím poskytne viac informácií o regióne pixlov z okolia zaujímavého bodu. Avšak sú metódy, ktoré tieto informácie o regióne pixlov nepotrebujú. Využívajú mriežku nemennej veľkosti, ktorá určuje, z akého rozsiahleho okolia bodu záujmu bude deskriptor vytvorený. Príznakový deskriptor využíva buď nejakú kombináciu farieb z obrazu alebo kanálov intenzity šedej farby. Taktiež môže využívať lokálne informácie o gradientoch a farbách. Deskriptor môže byť viacpremenný (multivariate), viacrozmerný (multidimensional) a multigeometrický, obsahujúci viacero kvantít vypočítaných cez viacero kanálov intenzít, viac geometrických škál a rôznych perspektív. Na obrázku 1.1 vidíme Vision Metrics taxonómiu [18]. Rozmery deskriptora predstavujú tri osi: *tvar a vzor*, *spektrum a hustota*. Skupinu príznakových deskriptorov delíme na:

1. **Lokálne binárne deskriptory** (*local binary descriptors*) - z bodových dvojíc v lokálnej oblasti vytvorí binárne kódovaný bitový vektor. 1 bit = 1 porovnanie

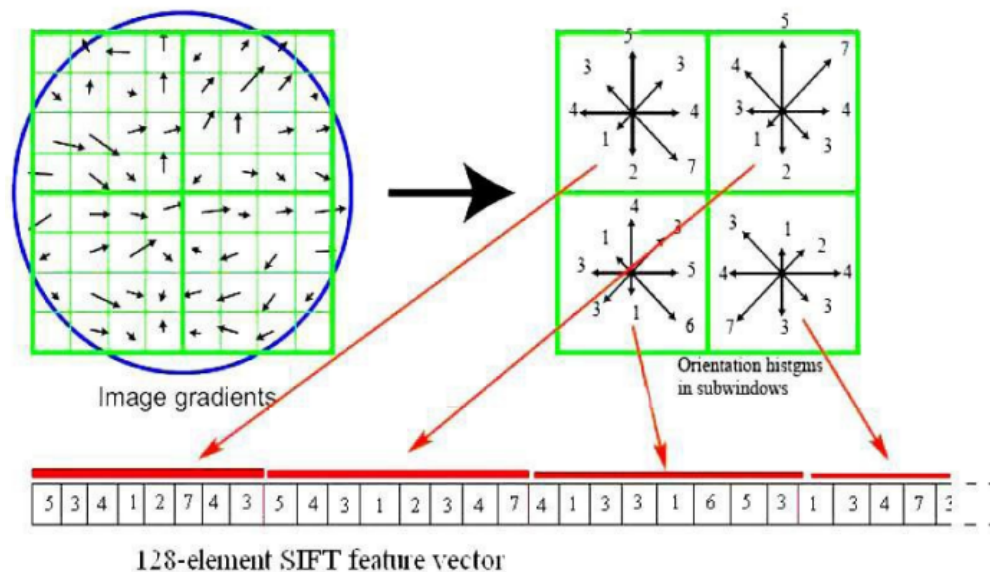
podrobené vyhovujúcej funkcii vzdialenosti. Metódy binárnych deskriptorov [20] sú napr. FREAK, ORB, BRISK.

2. **Spektrálne deskriptory** (*spectra descriptors*) - používajú širokú škálu spektrálnych hodnôt, ako gradienty a priemery oblasti. Najbežnejšie sa využívajúce spektrum v detektoroch je gradient lokálnej oblasti, využitý v SIFT-e.
3. **Bázové priestorové deskriptory** (*basic space descriptors*) - vektor príznakov kódujú do množiny bázových funkcií, ako Fourierov rad z funkcií sínus, kosínus.
4. **Mnohouholníkové deskriptory** (*polygon shape descriptors*) - využívajú tvar objektu zmeraný štatistickými metrikami, ako je plocha, ťažisko a obvod. Tvary sú extrahované pomocou morfológických operácií a algoritmov oblasti.



Obr. 1.1: Taxonómia pre rozmery deskriptora: hustota (density) príznaku, ako globálna, regionálna alebo rozhádzaná lokálna, tvar a vzor pixlov (shape a pattern) použitý na výpočet deskriptora, ako kruh alebo obdĺžnik, spektrum, ktoré obsahuje informácie obsiahnuté v samotných funkciách [18].

Deskriptor vytvárame napríklad pomocou kruhového alebo štvorcového tvaru okolia. Pomocou point-pair pixlov vypočítame alebo porovnáme vybrané hodnoty obrazových bodov, čím získame **príznakový vektor**. Na obrázku 1.2 je znázornený vypočítaný príznakový vektor metódy SIFT. Každý bod záujmu, použitý na vytvorenie deskriptora, obsahuje určité množstvo invariance a robustnosti, ako napríklad informáciu o škále a rotácii. V mnohých prípadoch orientácia deskriptora závisí od bodu záujmu. Kombinácia deskriptora a bodu záujmu poskytuje väčšiu množinu atribútov invariance. Pre správnu



Obr. 1.2: Príznačový vektor metódy SIFT [24]

extrakciu lokálnych príznačkov musí byť metóda detektora a deskriptora starostlivo 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. Robustnosť a presnosť sú kombináciou týchto faktorov:

- *Presnosť bodu záujmu* - pretože mnoho deskriptorov závisí od umiestnenia a orientácie keypointu.
- *Presnosť deskriptora* - pretože každá metóda deskriptora sa líši.
- *Presnosť funkcie klasifikácie a vzdialenosti* - pretože zlá klasifikácia a umiestnenie môže viesť k zlým výsledkom.

Atribúty a kritéria robustnosti zoskupujeme do viacerých obsiahlejších skupín. Svetlo je prioritnou oblasťou pre analýzu šedotónového obrazu. *Osvetlenie* má niekoľko podôb a je spracovávané oddelene od farby a farebného priestoru. Môže byť upravené zmenou zdroja svetla alebo jeho pridaním, pri predspracovaní je zase potrebné jeho odstránenie pre ďalšiu analýzu obrazu a extrakciu príznačkov. Za kľúčové vlastnosti osvetlenia považujeme nerovnomerné osvetlenie, jas, kontrast. Pri spracovaní farebného obrazu s analýzou farby je presnosť farby veľmi kritická. Hlavné aspekty presnosti sú: *farebný priestor, farebné kanály a bitová hĺbka farby*. Každý farebný priestor má presnosť a úžitkové aspekty, ako jednoduchosť a transformácia farieb do farebných priestorov, ako RGB, HSV, YIQ. Farebné kanály poskytujú RGB dáta, z ktorých sú extrahované intenzity šedotónového obrazu. Bitová hĺbka farby disponuje s informáciami o farbách, ktoré sú použité v obraze.

Extrakciu príznačkov môže skomplikovať aj neúplnosť obrazu, ako *šum*, ktorý je potrebné

odstrániť, *rozostrenie* spôsobené napríklad pohybom, *odľahlé hodnoty*, ktoré musia byť ignorované. Častým problémom je aj *vzdialenosť*, *škála* a *rozlíšenie*. V závislosti od aplikácie môže byť potrebný viac ako jeden spôsob deskriptora na zvládnutie rozlíšenia a škály. V praxi sa využíva napríklad škálová pyramída, ktorú využíva metóda SIFT.

Deskriptor je základňou pre príznakový "matching"(párovanie), ktorý vedie k porozumeniu obrazu, analyzovaniu scény, či k sledovaniu objektov. Hlavný problém 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 opakovateľnosti, robustnosti, výpočtovej náročnosti alebo veľkosti pamäte. Relatívne páry sú určené vzdialenostnou funkciou, kde prichádzajúce množiny príznakových deskriptorov sú porovnávané s množinou už známych deskriptorov.

Kapitola 2

Prehľad existujúcich prác

Aj keď je farba zvyčajne vnímaná ako nenahraditeľný prvok opisujúci svet okolo nás, techniky na extrakciu lokálnych príznakov sú najčastejšie založené na popise tvaru a úplne ignorujú farebnú informáciu. Popis farby obmedzuje veľké množstvo variácií, čo spôsobí, že namerané hodnoty farby sa môžu výrazne líšiť.

V tejto kapitole sme zhrnuli množinu vybraných existujúcich modifikácií metód SIFT [24] [23] a SURF [4]. Množstvo prístupov použitých v spracovaní farebnej informácie v SIFT metódach bolo totožných aj pri spracovaní farby metódou SURF. Z tohto dôvodu je obsiahlejšia podkapitola popisujúca modifikácie SIFT metódy. V závere sú popísané metódy analyzované.

2.1 Farebné modifikácie metódy SIFT

Technika SIFT („*Scaleable Invariant Feature Transform*“) vytvorená Lowem [24] [23] predstavuje škálovo a rotačne invariantný detektor a deskriptor zaujímavých bodov. V posledných rokoch je veľmi využívaná v rôznych oblastiach počítačového videnia, ako je rozpoznávanie objektov, klasifikácia obrazu a podobne. Detektor v metóde SIFT využíva na vyhľadávanie zaujímavých bodov tzv. škálovú pyramídu, ktorú vytvára s využitím rozdielu Gaussiánov. Hodnota deskriptora je vypočítaná pre každý zaujímavý bod z histogramu orientácií pre 16 štvorcových oblastí v jeho okolí. Ide o veľmi robustný invariantný deskriptor, rešpektujúci geometrické zmeny.

Tradičný SIFT deskriptor je vypočítaný na šedotónovom obraze, na ktorom dva úplne rozličné objekty môžu byť považované za rovnaké. Z tohto dôsledku boli vyvinuté rôzne druhy deskriptora SIFT, využívajúce farebnú informáciu. Existuje viacero rôznych farebných priestorov rozdelených do kategórií: homogénne priestory farieb, hárdverovo orientované modely, modely farieb v televízii a videotechnike a používateľsky orientované modely.

Na základe jednotlivých využití farebných priestorov bolo navrhnutých niekoľko modifikácií

SIFT deskriptora, s využitím farebnej informácie. Červený, zelený a modrý (*RGB*) farebný priestor, ako aj najznámejší farebný model, poskytuje veľa informácií pre aplikácie počítačového videnia. Qing Li a kol. [22] pridávajú *RGB* farebné informácie do SIFT deskriptora počítaním klasického deskriptora na všetkých troch kanáloch *RGB* zvlášť, čím vznikne deskriptor veľkosti 3×128 a zachytí viac informácií o farbe obrázkov.

Bosh a kol. [5] počítajú SIFT deskriptor cez všetky 3 farebné kanály farebného modelu HSV podobne ako pre *RGB-SIFT*e. Model HSV je užívateľsky orientovaný model, v ktorom hlavnými parametrami sú odtieň farby (H-hue), sýtosť (S-saturation) a jasová hodnota (V-value). Vypočítaním deskriptora cez všetky farebné kanály dostávame deskriptor veľkosti 3×128 , čiže 128 na každý kanál. Kanál H vo farebnom modeli je škálovo a posunovo nemenný vzhľadom na intenzitu svetla. Avšak kombináciou kanálov HSV výsledný deskriptor nemá žiadne vlastnosti invariance.

Farebný priestor *YCbCr* sa používa v *JPEG* a *MPEG* štandardoch. Je to široká oblasť využívaná pri kompresii videa, kvantizácii obrazu alebo analýzy kontúr. *YCbCr* má jednu jasovú zložku (*Y*) a modrú (*Cb*) a červenú (*Cr*) chromatickú zložku. *YCbCr-SIFT* [22] deskriptor je počítaný na všetkých kanáloch farebného priestoru.

V 19.storočí Ewald Hering navrhol oponentný princíp kódovania farieb a vytvoril oponentný farebný priestor, ktorý sa skladá z troch kanálov (O_1, O_2, O_3) [33]. O_3 reprezentuje čierno-biely kanál obrazu, zvyšné dva kanály opisujú oponentnú farbu v obraze. Qing Li a Junzhou Chen [22] na každom farebnom kanáli vyextrahujú jednotlivé deskripty, ktoré následne spolu skombinujú. Výsledný deskriptor používajú ako *Opponent-SIFT*.

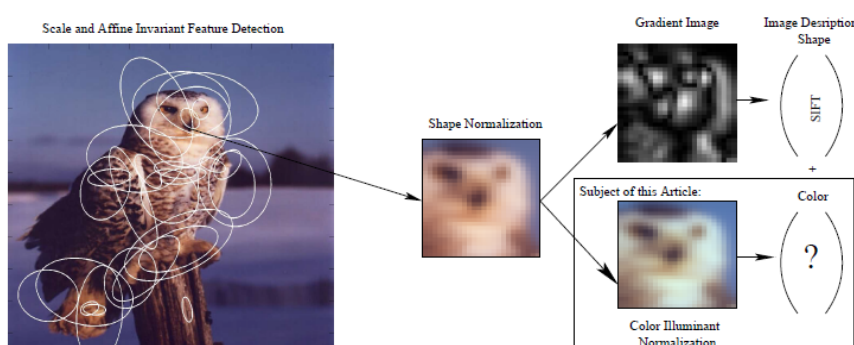
Deskriptor *rg-SIFT* [22] je získaný z farebného priestoru *rg*, konvertovaného z normalizovaného modelu *RGB*. Na popis farebnej informácie v obraze využívajú len *r* a *g* kanál. Ak sú uvedené kanály *r* a *g*, kanál *b* je konštantný.

Rozsiahla štúdia Mykolajczyka a Schmida [26] vyhodnotila deskriptor SIFT ako metódu s najlepším výkonom. Opis lokálnej farby získal relatívne málo pozornosti a ako výsledok metódy používajú svietivosť a ignorujú farebnú informáciu. Weijer a Schmid vo svojej práci [34] rozšírili deskriptor lokálnych príznakov o farebnú informáciu zlúčením farebného deskriptora *K* k deskriptoru tvaru *S* podľa

$$B = (\hat{F}, \lambda \hat{K}) \quad (2.1)$$

kde *B* je kombinovaný farebný a tvarový deskriptor, λ je váhový parameter a $\hat{\cdot}$ indikuje normalizovaný vektor. Pre opis tvaru využívajú deskriptor SIFT. Vzhľadom na kombináciu farebného a tvarového deskriptora, ktorá nemusí obsahovať žiadne priestorové informácie, využívajú lokálne histogramy. Koloritu histogramu dodržia na základe štyroch hlavných kritérií:

- *fotometrická robustnosť* - deskriptor by mal byť robustný na fotometrické invariance, ako tieň, tieňovanie, zmeny svetelného zdroja.
- *geometrická robustnosť* - deskriptor by mal byť robustný voči geometrickým zmenám, ako zmena orientácie objektu, priblíženie.
- *fotometrická stabilita* - deskriptor by mal vedieť spracovať nestability zavedené fotometrickou invariantnou transformáciou.
- *všeobecnosť* - farebný deskriptor by mal byť použiteľný pre širokú škálu aplikácií, ako porovnávanie, klasifikácia, vyhľadávanie. Mal by byť robustný na zmeny kvality obrazu z vysokokvalitných obrazov na komprimované kvality snímok.



Obr. 2.1: Princíp metód na extrakciu lokálnych príznakov rozdelených na fázu detekcie a fázu deskripcie [26].

Pri tvorbe farebného deskriptora predpokladajú, že scéna sa skladá len z nehomogénnych materiálov, vrátane papierov a plastov. Homogénne materiály úplne vylúčili. Pri zdroji svetla rozlišujú dve varianty. Ak je zdroj svetla len jeden, modelujú ho lokálne. V prípade väčšieho počtu svetelných zdrojov predpokladajú, že ich kombinácie sa môžu približne určiť ako jediný zdroj svetla. Namerané hodnoty $C \in \{R, G, B\}$ kamery so spektrálnou citlivosťou f^C sú modelované integráciou cez viditeľné spektrum ω

$$C(x) = m^b(x) \int_{\omega} b(\lambda, x) e(\lambda) f^C(\lambda) d\lambda + m^i(x) \int_{\omega} i(\lambda) e(\lambda) f^C(\lambda) d\lambda. \quad (2.2)$$

Geometrickú závislosť odrazu predstavuje m^b (odraz tela) a m^i (odraz rozhrania), ktoré sú závislé na uhle pohľadu, zdroji svetla a orientácii povrchu. x označuje priestorové súradnice. Rovnica nepokrýva difúzne svetlo prichádzajúce z viacerých smerov.

Farebný deskriptor bol odvodený na základe štyroch hlavných kritérií dodržiavania koloritu histogramu. Pri fotometrickej robustnosti posudzovali prípad, pri ktorom nie je prítomný žiadny difúzny osvetľovací zdroj. Vzťah medzi dvoma obrazmi rovnakej scény, C^1 a C^2 , s rôznymi osvetľovacími zdrojmi, je modelovaný skalárnym súčinom, keďže

$$C^2(x) = (m^b(x)b^C(x) + m^i(x))e_2^C = \lambda^C C^1(x) \quad (2.3)$$

kde $\lambda^C = e_2^C/e_1^C$. Farby v oboch scénach prezentuje diagonálna matica $C_2 = \Lambda C_1$. Relácia diagonálnej matice je známa a je dôsledkom predpokladaných delta funkcií citlivosti kamery. Invariance, vzhľadom na osvetlenie, možno získať normalizáciou každého farebného kanálu:

$$C^*(x) = \frac{C(x)}{C(x)} \quad (2.4)$$

Ak je v scéne dominantné difúzne svetlo, rovnica 2.3 nie je postačujúca. Obrazy sú príbuzné cez diagonálnu maticu $C_2 = \Lambda C_1 + a$. Invariance, vzhľadom na dominantný svetelný zdroj, Λ , tak môžu byť získané normalizáciou farebných kanálov s ich priemerným derivátom:

$$C^*(x) = \frac{C(x)}{|C(x)|} \quad (2.5)$$

V prípade matného povrchu ($m^i = 0$) bez difúzneho osvetlenia ($a^C = 0$) považujú normalizované kanály rgb za nemenné, s ohľadom na geometriu osvetlenia a hľadisko (viewpoint), m^b . Vzhľadom k tomu:

$$r = \frac{R}{R + G + B} = \frac{m^b b^R e^R}{m^b (b^R e^R + b^G e^G + b^B e^B)}. \quad (2.6)$$

Podobné rovnice ako 2.6 platia aj pre normalizované g a b . Invariantnosť s ohľadom na geometriu osvetlenia a odrazivosť je získaná z odtieňa:

$$hue = \arctan\left(\frac{O1}{O2}\right) = \arctan\left(\frac{\sqrt{3}(b^R - b^G)}{(b^R + b^G - 2b^B)}\right) \quad (2.7)$$

kde využívajú oponentné farby 2.8, ktoré predstavujú stále varianty pre zmeny geometrického osvetlenia:

$$\begin{aligned} O1 &= 1/\sqrt{2} (R - G) = 1/\sqrt{2} (m^b e (b^R - b^G) + m^i e - m^i e) \\ O2 &= 1/\sqrt{6} (R + G - 2B) = 1/\sqrt{6} (m^b e (b^R + b^G - 2b^B) + 2m^i e - 2m^i e) \end{aligned} \quad (2.8)$$

Geometrickú robustnosť čiastočne získali afinnou invariantnou detekciou zaujímavých bodov. Osobitnú pozornosť venovali práci s derivátom, na základe invarianty, kde navrhli dve nové invarianty nazývajúce farebné uhly. Pri fotometrických invariantoch je známe, že majú inherentnú nestabilitu. Aby nijako neovplyvnili konečné histogramy, upravili váhy hodnôt farby v histograme podľa jeho istoty. Využili *rgb-histogram* a *hue-histogram*.

Na detekciu zaujímavých bodov využili afinne invariantný Harris-Laplace detektor. V kroku tvarovej normalizácie obraz zredukovali na 20x20 susedných pixlov a následne vypočítali SIFT. Pri farebnom deskriptore najprv znormalizovali farbu, potom vypočítali rgb a hue histogramy s upravenými váhami. Deskriptor SIFT mal dĺžku 128, jednodimenzionálny hue deskriptor 37 a dvojdimenzionálny deskriptor rgb 121.

Colored SIFT (ďalej len C-SIFT) je geometricky a farebne invariantný deskriptor navrhnutý Alaa E. Abdel-Hakim [1], ktorý využíva farebný invariantný model závislý na dichromatickom modeli odrazivosti (Dichromatic Reflectance Model), navrhnutý M. Geusebroekom [12]. Farebná invariancia je závislá na teórii Kubelka-Munk [19], ktorá modeluje fotometrickú odrazivosť podľa:

$$E(\lambda, \vec{x}) = e(\lambda, \vec{x})(1 - \rho_f(\vec{x}))^2 R_\infty(\lambda, \vec{x}) + e(\lambda, \vec{x})\rho_f(\vec{x}) \quad (2.9)$$

kde λ prezentuje vlnovú dĺžku a \vec{x} 2D vektor polohy obrazu. $e(\lambda, \vec{x})$ označuje osvetlenie spektra, $\rho_f(\vec{x})$ označuje Fresnelovu odrazivosť na \vec{x} . $R_\infty(\lambda, \vec{x})$ označuje odrazivosť materiálu a $E(\lambda, \vec{x})$ predstavuje odrazené svetlo v smere pohľadu.

Na výpočet farebnej invariantnosti zo známeho RGB priestoru používajú Gaussian farebný model, ako všeobecný model pre reprezentáciu spektrálnych informácií a lokálnej štruktúry obrazu. Lineárne transformácie z RGB priestoru sú použité na získanie *spektrálnych diferenciálnych kvocientov* ($\hat{E}, \hat{E}_\lambda, \hat{E}_{\lambda\lambda}$). *Priestorové diferenciálne kvocienty* ($\hat{E}_x, \hat{E}_{\lambda,x}, \hat{E}_{\lambda\lambda,x}$) sú získané z konvolúcie s filtrom Gausiánov. Vypočítaním prvých troch komponentov ($\hat{E}, \hat{E}_\lambda, \hat{E}_{\lambda\lambda}$) z Gaussian farebného modelu získali dobrú aproximáciu pre ľudský vizuálny systém a pre CIE 1964 XYZ. Požadovanú implementáciu Gaussian farebného modelu v podmienkach RGB získali, ako je znázornené v rovnici 2.10, lineárnou transformáciou z RGB do XYZ a z XYZ do Gaussian farebného modelu. Mieru farebnej invariantnosti získali substitúciou $E, E_\lambda, E_{\lambda\lambda}$ za $\hat{E}, \hat{E}_\lambda, \hat{E}_{\lambda\lambda}$ a pri danom σ_x .

$$\begin{pmatrix} \hat{E} \\ \hat{E}_\lambda \\ \hat{E}_{\lambda\lambda} \end{pmatrix} = \begin{pmatrix} .06 & .63 & .27 \\ .3 & .04 & -.35 \\ .34 & -.6 & .17 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix} \quad (2.10)$$

Zaujímavé body detegovali na extrémoch rozdielov Gausiánov (DoG) vstupného obrazu, s využitím farebnej invariantnosti popísanej vyššie. Pre Gaussian farebný model použili $\sigma_x = 2$. Vzhľadom k tomu, pre filtrovanie Gausiánom použili štandardnú odchýlku $\sigma = 1.4$.

Rovnako ako pri SIFT-e, ako kľúčové vstupy deskriptora použili lokálny histogram orientácií, pre každý zaujímavý bod a jeho okolie. Namiesto šedých gradientov použili gradienty farebných invariantov. Použitie farebných invariantov namiesto použitia hodnoty im zaistili robustnosť vzhľadom k fotometrickým zmenám.

Dandan Li a kol. [21] navrhli *COLOR-SIFT*, ako kombináciu SIFT prístupu s *lokálnym farebným kernel histogramom* [27]. Konečný rozsah histogramu zvýši celkovú odlišnosť a preskúma informácie farebného priestoru spôsobom, ktorý je jednoduchý a účinný. Farebný histogram počítajú v susedných regiónoch jednotlivých zaujímavých bodov.

Na jednotlivé susedné regióny zaujímavých bodov veľkosti 8×8 aplikujú uniformnú farebnú kvantifikáciu a kvantizáciu. Z kvantovaných oblastí vybrali N_C farieb, ktoré spojili s

N_C zložkami histogramu. Pre možné nejasnosti, medzi obrázkom a histogramom, použili Gaussovo jadro, ktoré vo farebnom obraze môže zabezpečiť robustnosť proti šumu. Hlavnou nevýhodou histogramu je, že farby, ktoré sú síce rozdielne, ale veľmi blízko pri sebe, histogram zaradí do dvoch odlišných zložiek. Využitie Gaussovho jadra tento problém zmierňuje tým, že berie do úvahy podobnosť farieb. Pixel rozdelí na niekoľko zložiek podľa Gaussovej váhy G^C . V YUV priestore a pre danú h_j , zložku reprezentujúcu farbu (U_j, V_j) v chrominančnom histograme, príspevok pixlu $S_k(x_k, y_k, Y_k, U_k, V_k)$ na h_j zložku je:

$$\begin{aligned} d_U &= U_k - U_j \\ d_V &= V_k - V_j \\ G_j^c(U_k, V_k) &= \frac{K_j^C}{2\pi\sigma} \exp\left(-\frac{d_u^2 - d_v^2}{2\sigma_c^2}\right) \end{aligned} \quad (2.11)$$

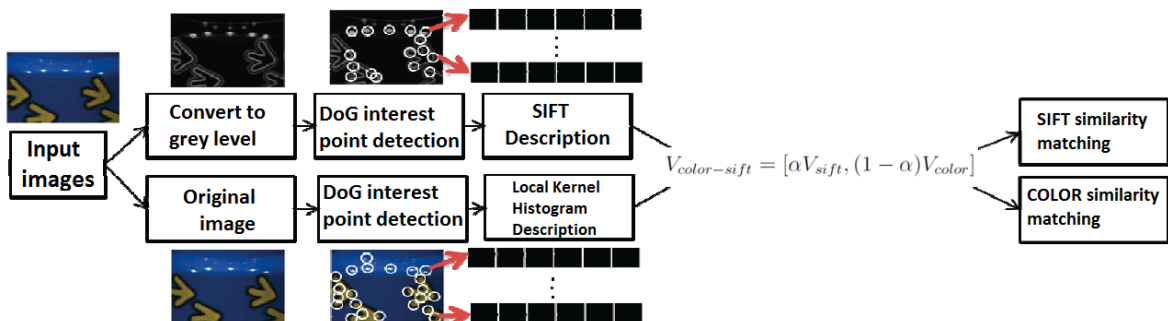
K_j^c je normalizovaný koeficient určený pre farbu (U_j, V_j) medzi n_r farbami v redukovanom priestore $\sum_{i=1}^{n_r} G_j^c(U_i, V_i) = 1$. Štandardná odchýlka σ_c je odhadovaná šumom z kamery. V lokálnej oblasti 1, hodnota h_j zložky histogramu zodpovedajúcej vybranej farbe ($1 \leq j \leq n_c$) je:

$$h_j = \sum_{k=1}^{n^2} G_j^c(S_k) \quad (2.12)$$

COLOR-SIFT deskriptor je dvojprvkový vektor, ktorý vznikne zret'azením SIFT-u a vektorov deskriptorov farebného histogramu.

$$V_{color-sift} = [V_{sift}, V_{color}] \quad (2.13)$$

kde V_{color} je vektor farebného kernel histogramu v rámci regiónu okolo zaujímavého bodu. Pre párovanie COLOR-SIFT deskriptora používajú rôzne metrické vzdialenosti pre porovnanie podobností medzi oboma časťami deskriptora. Pre časť deskriptora SIFT používajú *euklidovskú vzdialenosť*. Pri výpočte podobnosti lokálneho farebného kernel histogramu využívajú jeho *prienik*.

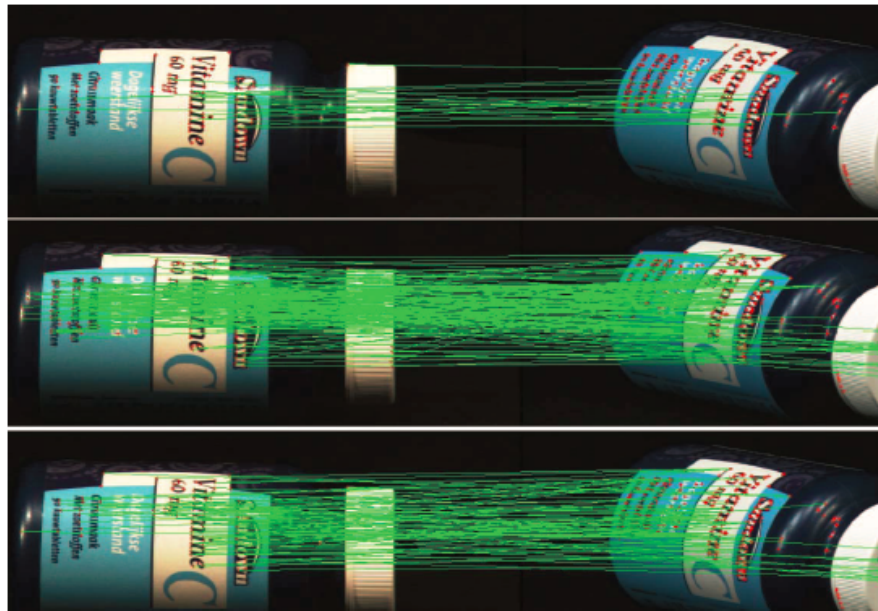


Obr. 2.2: Prehľad navrhovanej CH-SIFT metódy. Farebný a šedý variant obrázkov je analyzovaný oddelene a párovanie je vykonávané s rozdielnymi váhami [15].

Ali Jalilvand a kol. [15] vo svojom *CH-SIFT*-e využili taktiež lokálny farebný kernel histogram s kombináciou s klasickým SIFT-om, ako je znázornené na obrázku 2.2. Jedinou zmenou oproti *COLOR-SIFT*-u je, že pri zret'azení vektorov farebného histogramu a SIFT-u berú do úvahy, že ich hodnoty nemusia byť rovnaké. Na rozdiel od 2.13 využívajú rôzne váhy pre každý vektor deskriptora. Zret'azenie vykonajú na základe 2.14

$$V_{color-sift} = [\alpha V_{sift}, (1 - \alpha)V_{color}] \quad (2.14)$$

α testujú v hodnotách v rozpätí $\langle 0.1, 1 \rangle$. Za najvhodnejšiu bola zvolená hodnota $\alpha = 0.3$. Pri párovaní využívajú pre V_{sift} *euklidovskú vzdialenosť* a pre V_{color} *Bhattacharyaovu vzdialenosť*. Na obrázku 2.3 vidíme párovanie pomocou troch metód, a to SIFT, C-SIFT a CH-SIFT. Metóda CH-SIFT oproti ostatným dvom dosiahla najlepšie výsledky.



Obr. 2.3: Párovanie pomocou metódy SIFT, C-SIFT a CH-SIFT. [15]

Metóda *SIFT-CCH* navrhnutá C. Ancuti a Philippe Bekaertom [2], podobne ako predošlé metódy, využíva kombináciu metódy SIFT s histogramom farebných výskytov, s cieľom zvýšiť odlišnosť SIFT deskriptora. Metóda pracuje v rgb farebnom priestore, ktorý je reprezentovaný ako:

$$r = \frac{R}{R + G + B} \quad g = \frac{G}{R + G + B} \quad b = \frac{B}{R + G + B} \quad (2.15)$$

Keďže histogramy nezohľadňujú priestorovú informáciu, a preto aj veľmi odlišné obrazy môžu mať podobnú farebnú distribúciu, mohli obrázok rozdeliť na niekoľko častí a jednotlivito vypočítať histogram. Takýto prístup vo veľkej miere ovplyvňuje pamäť a výpočtový čas. Pre odstránenie tohto nedostatku sa rozhodli použiť maticu farebných

výskytov (*CCH*), ktorá predstavuje vylepšenú verziu matice výskytov šedej úrovne. Finálny deskriptor je, podobne ako pri metóde *CH-SIFT*, zret'azením *SIFT*-u a *CCH* na základe 2.16

$$V_{SIFT-CCH} = [\alpha V_{SIFT}, (1 - \alpha) V_{CCH}] \quad (2.16)$$

Perception-based color SIFT (ďalej len *PC-SIFT*) navrhnutý Yan Cui a kol. [8] kombinuje štandardnú metódu *SIFT* s farebným modelom, založenom na vnímaní (*perception-based color space*), ktorý poskytuje invariáciu na osvetlenie.

Podmienky vnímania farebného modelu [7] formalizujú pomocou prekladania RGB farebného priestoru do priestoru XYZ.

$$\begin{bmatrix} \hat{X} \\ \hat{Y} \\ \hat{Z} \end{bmatrix} = \frac{1}{0.177} \begin{bmatrix} 0.49 & 0.31 & 0.20 \\ 0.177 & 0.812 & 0.011 \\ 0.00 & 0.01 & 0.99 \end{bmatrix} \begin{bmatrix} G(R) \\ G(G) \\ G(B) \end{bmatrix} \quad (2.17)$$

kde G je funkcia gamma korekcie s $\gamma = 2.0$. Priestor XYZ prekladajú na vnímavý farebný priestor UVW. V nadväznosti na [7] predpokladajú, že materiály a osvetlenie v použitých scénach sú také, že efekt opätovného osvetlenia je dobre aproximovateľný vynásobením každej trichromatickej hodnoty faktorom škály, ktorá je nezávislá na pozorovaných materiáloch. Efekt opätovného osvetlenia zapisujú ako:

$$\vec{x} \mapsto F(\vec{x}) = B^{-1} D B \vec{x} \quad (2.18)$$

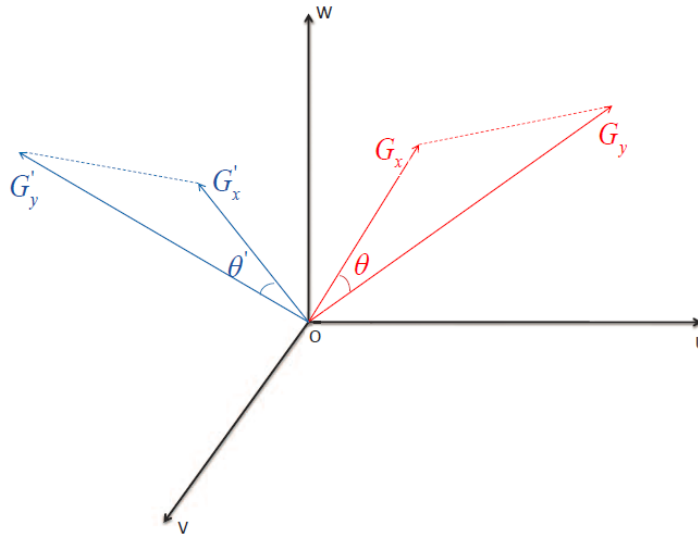
kde \vec{x} je trichromatická hodnota senzora zastúpená v XYZ súradniciach. F zastupuje 3D farebný priestor parametrizácie, na ktorom to riešia. Pevnú zmenu reprezentujú maticou B s vhodným základom a D je diagonálna matica v závislosti na osvetľovacích prostriedkoch. Ako je ukázané v [7], nelineárna funkcia F z rovnice pred musí mať formu:

$$F(\vec{x}) = A(\hat{\ln}(B\vec{x})) \quad (2.19)$$

kde A a B sú inverzné 3×3 matice. Matica B transformuje súradnice farebného základu, v ktorých je opätovné osvetlenie najlepšie zodpovedané násobením diagonálnou maticou, zatiaľ čo matica A obsahuje stupne voľnosti, ktoré môžu byť použité tak, aby zodpovedali vnemovej vzdialenosti. Matice A a B boli experimentálne odhadnuté na databáze podobných farieb.

Zaujímavé body detegujú rovnako ako pri *SIFT*-e s rozdielom, že používajú navrhnutý farebný model založený na vnímaní, čím detegujú body na troch nových farebných kanáloch. Pri vytváraní deskriptora nevyužívajú šedotónové gradienty, ale využívajú nastavenú funkciu 3D farebného priestoru založeného na vnímaní. Ako je znázornené na obrázku 2.4, pre daný pixel v blízkosti zaujímavého bodu označia G_x trojdimenzionálneho farebného gradientu v smere x a G_y v trojdimenzionálneho farebného gradientu v smere y . Uhol θ prezentuje vzdialenosť od vektora G_x k vektoru G_y gradientu orientácií v rozsahu od 0 po 2π .

PC-SIFT je invariantný voči osvetleniu alebo farebnej transformácii.

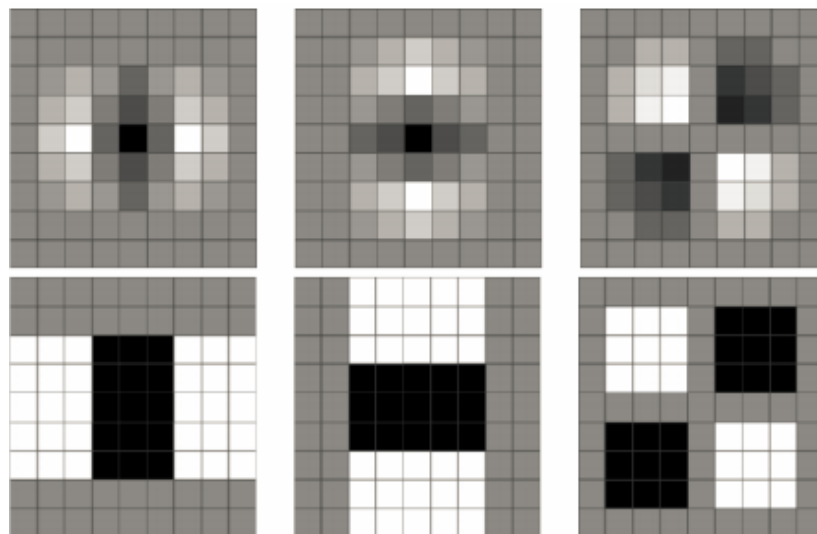


Obr. 2.4: Nový farebný priestor a PC-SIFT deskriptor [8].

2.2 Farebné modifikácie metódy SURF

Ďalšou technikou, ktorá obsahuje detektor aj deskriptor, je technika SURF („Speeded Up Robust Features“), ktorú autori Bay a kol.[4] považujú za vylepšenie SIFT-u a iných podobných techník.

Podobne ako pri metóde SIFT, aj SURF využíva na vyhľadávanie zaujímavých bodov škálovú pyramídu, na ktorú využíva aproximáciu Gausiánov druhej parciálnej derivácie v smere x, y a xy pomocou box filtrov, znázornených na obrázku 2.5. Škálový priestor nevytvára zmenšovaním obrazu a následnou konvolúciou, ale naopak. Využíva konvolúciu pôvodného obrazu s postupne sa zväčšujúcimi filtermi. Následne je na obraz aplikované potlačenie nemaximálnych hodnôt v okolí $3 \times 3 \times 3$, spomínané pri metóde SIFT.

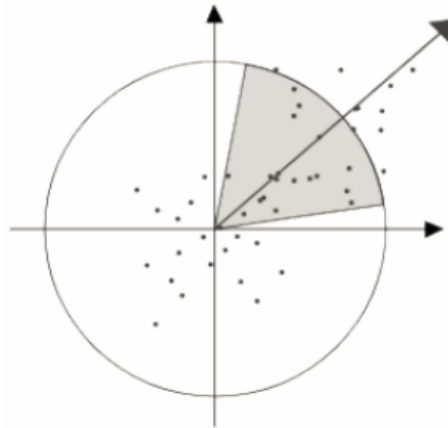


Obr. 2.5: Aproximácie Gausiánov druhej parciálnej derivácie v smere x, y a xy a vážené aproximácie box filtrov [4].

Pre vytvorenie deskriptora v metóde SURF je potrebné zistiť jeho orientáciu a vytvoriť štvorcovú oblasť zarovnanú podľa zistenej orientácie, z ktorej sa bude deskriptor extrahovať. Deskriptor zaujímavého bodu obsahuje 4 hodnoty pre každú zo 16 oblastí, čiže spolu 64 hodnôt. Podobne ako pri metóde SIFT, aj metóda SURF funguje na šedotónovom obraze. Preto boli navrhnuté jej rôzne farebné modifikácie.

Peng Fan a kol. [10] navrhli metódu *Color-SURF*. Podobne ako pri metóde Color-SIFT aj metóda Color-SURF predstavuje použitie metódy *SURF* kombinovanou s *lokálnym farebným kernel histogramom*. Vo svojej metóde využívajú integrálne obrazy, ako urýchľovač času výpočtu. Na detekciu zaujímavých bodov využívajú SURF detektor, založený na parciálnych deriváciách druhého rádu Hesseho matice. Color-SURF detektor rozdelili do dvoch vzdialených úloh: úloha orientácie a komponenty deskriptora.

Na extrakciu zložiek deskriptora a aj na dosiahnutie invariencie rotácie je dôležité priradiť orientáciu pre každý zaujímavý bod. Na určenie orientácie používajú Peng Fan a kol. štandardný SURF detektor s Haar wavelet odozvami vo vertikálnom a horizontálnom smere okolia detegovaného bodu. Dominantná orientácia zaujímavého bodu je určená pomocou rotačnej časti kruhu pokrývajúcej uhol $\frac{\pi}{3}$ v okolí počiatku, zobrazená na obrázku 2.6. V každej pozícii sú horizontálne odozvy x a vertikálne odozvy y , v rámci každého segmentu, zosumované a použité na vytvorenie nového vektora. Zaujímavý bod má dominantnú orientáciu najdlhšieho vytvoreného vektora.



Obr. 2.6: Určenie orientácie. Sumov zložiek odoziev získame vektor [10].

Prvým krokom na extrakciu deskriptora je skonštruovanie štvorcového okna v okolí zaujímavého bodu, ktoré je zarovnané podľa orientácie. Okno si rozdelili na podregióny veľkosti 4×4 . V každom podregióne počítajú 4 príznaky, vid' 2.20, na 25 pravidelne rozmiestnených bodoch.

$$V_{SURF} = \left(\sum d_x, \sum d_y, \sum |d_x|, \sum |d_y| \right) \quad (2.20)$$

kde:

- $\sum d_x$ - suma odoziev Haarovho wavletu v horizontálnom smere
- $\sum d_y$ - suma odoziev Haarovho wavletu vo vertikálnom smere
- $\sum |d_x|$ - suma absolútnych hodnôt odoziev Haarovho wavletu v horizontálnom smere
- $\sum |d_y|$ - suma absolútnych hodnôt odoziev Haarovho wavletu vo vertikálnom smere

Výsledný deskriptor zaujímavého bodu dĺžky 64, zatiaľ neobsahuje žiadnu informáciu o farbe. Farebnú informáciu spracovávajú pomocou lokálneho farebného kernel histogramu, vypočítaného vo vnútri štvorcového okna v okolí zaujímavého bodu. Farebný priestor je kvantovaný podľa najreprezentatívnejších farieb extrahovaných zo scény, čím výrazne znížili množstvo dát bez straty dôležitých informácií. Podobne ako aj pri metóde *Color-SIFT*, aj tu autori používajú Gaussovo jadro na zabezpečenie robustnosti proti šumu a na odstránenie problému farieb veľmi blízkych pri sebe.

Finálny Color-SURF deskriptor 2.21 je dvojprvkový vektor, ktorý vznikne zret'azením deskriptora SURF a deskriptormi vektorov z farebného histogramu.

$$V_{Color-SURF} = [\alpha V_{SURF}, (1 - \alpha) V_{Color}] \quad (2.21)$$

Pri párovaní, pre porovnanie podobností medzi oboma časťami deskriptora, kombinujú dve metrické vzdialenosti. Najprv počítajú *euklidovskú vzdialenosť* medzi všetkými vektormi deskriptora SURF. Medzi farebnými histogramami zaujímavých bodov, ktoré nie sú asociované so SURF vzdialenosťou, počítajú *Bhattacharyaovu vzdialenosť*.

Ďalšou farebnou modifikáciou metódy SURF je metóda *CW-SURF* navrhnutá Ali Jalivandom a kol. [16], ktorá predstavuje nový farebne lokálne invariantný deskriptor založený na metóde SURF a farebne invariantného modelu C_w . Metóda je veľmi podobná metóde C-SIFT spomenutej v predchádzajúcej kapitole až s rozdielom, že autori namiesto SIFT-u využívajú metódu SURF. Rovnako ako v [1] využívajú farebný invariantný model závislý na dichromatickom modeli odrazivosti (Dichromatic Reflectance Model), navrhnutý J. M. Geusebroekom [12]. J.M. Geusebroek vo svojej ďalšej publikácii [6] porovnáva rôzne farebné invarianty a prichádza k záveru, že najviac príznačný farebný invariant bude C-color, ktorý môže byť vyjadrený ako:

$$\vec{C}_w = \sqrt{C_{\lambda x}^2 + C_{\lambda y}^2 + C_{\lambda \lambda x}^2 + C_{\lambda \lambda y}^2} \quad (2.22)$$

Farebná invariancia je závislá na teórii Kubelka-Munk [19], ktorá modeluje fotometrickú odrazivosť podľa 2.9. Ako všeobecný model používajú Gaussian farebný model pre reprezentáciu spektrálnych informácií a lokálnej štruktúry obrazu. Na získanie spektrálnych

a priestorových diferenciálnych koeficientov využívajú lineárne transformácie z RGB priestoru, znázornené na 2.10 a konvolúciu s filtrom Gaussiánov. Z oponentných kanálov $E, E_\lambda, E_{\lambda\lambda}$ z 2.10 môžeme získať $C_\lambda, C_w, C_{\lambda\lambda}$ a H predstavujúce Hesseho maticu v SURF detektore, ktorým detegujú zaujímavé body. Pri vytváraní CW-SURF deskriptora využívajú namiesto šedých gradientov gradienty farebných invariantov, čo výrazne znižuje výpočtovú náročnosť a odolnosť voči rôznym geometrickým transformáciám. Autori metódy pri párovaní testovali niekoľko rôznych vzdialenostných metrík. Po rôznych pokusoch za najvhodnejšiu vybrali euklidovskú vzdialenosť použitú pre V_{SURF} v 2.23:

$$d_{(SURF)}(x, y) = \sqrt{\sum (V_{SURF}^x - V_{SURF}^y)^2} \quad (2.23)$$

Autori C.J.Prabhakar a Praveen Kumar P.U. v [28] skúmali snímky zhotovené pod vodnou hladinou, ktoré obsahujú veľmi vysokú zmenu farby z dôsledku šírenia svetla. Na základe tohto výskumu navrhli *LBP-SURF* deskriptor, ktorý je veľmi účinný na získanie farebne invariantných a textúrnych príznakov zo snímok pod vodnou hladinou. Farebné hodnoty v obraze, ktoré sú zmenené rôznymi radiometrickými faktormi, vytváranými pod vodnou hladinou, upravujú použitím komplexnej normalizácie farebného obrazu. Pri pozorovaní povrchu pod určitým svetlo, ktorý ma dvojnásobnú intenzitu, očakávali zdvojnásobenie zaznamenávaných RGB hodnôt:

$$\begin{pmatrix} R_i \\ G_i \\ B_i \end{pmatrix} = \begin{pmatrix} \rho_i R_i \\ \rho_i G_i \\ \rho_i B_i \end{pmatrix} \quad (2.24)$$

kde ρ_i je jednoduchý skalár.

Zmena intenzity svetla modelovaná 2.24, môže byť zrušená chromatickou normalizáciou:

$$\begin{pmatrix} r_i \\ g_i \\ b_i \end{pmatrix} = \begin{pmatrix} \frac{\rho_i R_i}{\rho_i R_i + \rho_i G_i + \rho_i B_i} \\ \frac{\rho_i G_i}{\rho_i R_i + \rho_i G_i + \rho_i B_i} \\ \frac{\rho_i B_i}{\rho_i R_i + \rho_i G_i + \rho_i B_i} \end{pmatrix} \quad (2.25)$$

kde nové súradnice r, g, b sú nezávislé na ρ_i . Na odstránenie vplyvu geometrie osvetlenia a osvetlenia farby použili komplexnú normalizáciu, s ktorou odstránili obe závislosti, definované ako:

1. $I_0 = I$ inicializácia
2. $I_{i+1} = G(C(I_i))$ krok iterácie
3. $I_{i+1} = I_i$ stav ukončenia

kde $G()$ a $C()$ sú funkcie pre „gray-world“ a chromatickú normalizáciu, ktoré sa aplikujú postupne a opakovane na snímku, až kým výsledný obraz konverguje k fixnému bodu.

Na takto upravenej snímke detegujú zaujímavé body klasickou metódou SURF detektora. Textúrne deskriptor pre každý zaujímavý bod vytvárajú s použitím „*Center-Symmetric Local Binary Patterns*“ (ďalej len CS-LBP) techniky. CS-LBP je modifikovanou verziou „*Local Binary patterns*“ (LBP) deskriptora, ktorý namiesto porovnania každého pixla s centrálnym pixlom porovnáva páry pixlov symetrické centra. Robustnosť v regiónoch snímky je získavaná prahovaním šedoúrovňového rozdielu s malou hodnotou T. Matematická reprezentácia CS-LPB:

$$CS - LBP_{R,N,T}(x, y) = \sum_{i=0}^{(N/2)-1} s(n_i - n_{i+(N/2)})2^i \quad (2.26)$$

$$s(x) = \begin{cases} 1, & x > T \\ 0, & x \leq T \end{cases}$$

kde n_i a $n_{i+(N/2)}$ predstavujú hodnotu šedej v centre symetrických párov pixlov z N rovnomerne rozložených pixlov na kružnice s polomerom R. Na dosiahnutie priestorovej informácie je vstupná oblasť snímky rozdelená karteziánskou mriežkou na bunky, v ktorých je zostavený CS-LBP histogram pre každú bunku. Výsledný deskriptor je 3D CS-LBP histogram pozície a hodnoty. Finálny deskriptor predstavuje zret'azenie príznakových histogramov, vypočítaných v každej bunke, čím sa vytvorí $M \times M \times 2^{N/2}$ dimenzionálny vektor, kde M a N predstavuje veľkosť karteziánskej mriežky.

Pri párovaní snímok autori využívajú princíp navrhnutý Lowem [24], ktorý navrhol párovanie NN prístupom s pomocou euklidovskej vzdialenosti medzi deskriptormi ľavej a pravej snímky, berúc do úvahy vhodnú prahovú hodnotu T. Využívajú NNDR prístup, nazývaný aj ako nejednoznačná vzdialenosť, ktorá predstavuje Loweho NN prístupu s lepším výkonom. NNDR porovnáva vzdialenosť medzi najbližším a druhým najbližším deskriptorom. Vzdialenostný pomer je prahovaný s cieľom zabrániť párovaniu pri veľkej podobnosti najbližších susedov, pixlov. Najbližší deskriptor $d_{L,1}$ a druhý najbližší deskriptor $d_{L,2}$ sú pomocou NNDR definované:

$$NNDR = \frac{\sqrt{\sum_{i=1}^{128} (d_R(i) - d_{L,1}(i))^2}}{\sqrt{\sum_{i=1}^{128} (d_R(i) - d_{L,2}(i))^2}} < T \quad (2.27)$$

Podobnosť medzi deskriptormi počítajú s použitím prístupu založenom na NNDR, kde prahovú hodnotu nastavujú empiricky.

2.3 Analýza metód

Všetky metódy, ktoré sme spomenuli v prechádzajúcich podkapitolách, sa nejakým spôsobom odlišovali. Niektorí autori zvolili prístup zahrnutia farebnej informácie až pri vytváraní deskriptora, iní ju využívali už pri detegovaní zaujímavých bodov. V tabuľke predstavujeme jednoduchú analýzu všetkých metód, ich dôležité odlišnosti, ako využitie rôznych farebných prispôbení formou vytvárania nových farebných modelov, alebo využitie rozdielnych alebo existujúcich metód, až po odlišné použitie vzdialenostných metrík pri párovaní obrazov.

Prvou podstatnou odlišnosťou je rozdielny vstup vo forme obrázka, ktorý je buď šedotónový alebo farebný. Väčšina prístupov v metódach využívajúcich šedotónový obraz spracováva farebnú informáciu len pomocou zvoleného druhu histogramu. Získané výsledky zakomponujú do pôvodného deskriptora klasickým zret'azením a vyhodnotia. Všetky metódy s týmto prístupom mali experimentálne výsledky na veľmi dobrej úrovni.

Pri farebnom vstupe si autori zvyčajne volia náročnejší prístup. Pre dobré detegovanie zaujímavých bodov a následné párovanie priamo farebného obrazu je veľmi dôležitá úprava farby obrazu. Existuje viacero jednoduchých farebných modelov, do ktorých je možné obraz previesť s jemnými modifikáciami a spustiť na ňom detekciu. Môže sa nám ale stať, že výsledky detekcie alebo párovania nebudú postačujúce.

Analyzované metódy vo svojich hodnoteniach poukazujú na výsledky, ktoré nie sú excelentné, ale v celku postačujúce. Každá metóda splnila autormi dané požiadavky, s využitím čiastočne odlišných metód. Autori v C-SIFT metóde sa zamerali na detekciu zaujímavých bodov. Ich výsledkom je detektor, ktorý oproti pôvodnej metóde SIFT dáva lepšie výsledky v detekcii bodov. Veľmi dobré výsledky pri párovaní zaujímavých bodov zaznamenala metóda CH-SIFT, ktorá ako jediná dávala najlepšie výsledky oproti metódam SIFT a C-SIFT.

Pri farebných modifikáciách metódy SURF za najlepšiu sme zvolili metódu CW-SURF. Aj keď metóda nedáva oveľa lepšie výsledky, no oproti ostatným farebným SURF metódam dáva najkratší čas spracovania pri párovaní bodov.

<i>metóda</i>	<i>vstup</i>	<i>farebný model</i>	<i>reprezentant farby</i>	<i>metrika</i>
C-SIFT	farebný	Gaussian farebný model	gradienty farebných invariantov	–
COLOR-SIFT	šedotónový	RGB	lokálny kernel histogram	euklidovská vzdialenosť, prienik
CH-SIFT	šedotónový	RGB	lokálny kernel histogram	euklidovská vzdialenosť, Bhattacharyaova vzdialenosť
SIFT-CCH	šedotónový	RGB	matica farebných výskytov	euklidovská vzdialenosť
PC-SIFT	farebný	farebný model založený na vnímaní	3D farebný priestor	–
COLOR-SURF	šedotónový	RGB	lokálny kernel histogram	euklidovská vzdialenosť, Bhattacharyaova vzdialenosť
CW-SURF	farebný	Cw model Gaussian farebný model	gradienty farebných invariantov	euklidovská vzdialenosť
LBP-SURF	farebný	komplexná normalizácia	–	euklidovská vzdialenosť, NNDR prístup

Tabuľka 2.1: Analýza farebných modifikácií metód z predchádzajúcich kapitol

Kapitola 3

Špecifikácia

V tretej kapitole tejto práce je popísaná špecifikácia navrhnutej metódy. Najprv je predstavená motivácia a cieľ práce. Ďalej je popísaný softvér, ktorý používame pre implementáciu, základný model, ktorého inšpirácia viedla k návrhu metódy a nakoniec je predstavené navrhnuté riešenie.

3.1 Motivácia a cieľ práce

Hlavným záchytným bodom pri lokálnych príznakoch je šedotónový obraz, ktorý vyhovoval dlhé desaťročia. V poslednej dobe sa metódy lokálnych príznakov začali spájať viac s farbou a farebným obrazom. Využitie farby v rôznych metódach je ale ešte stále v začiatkoch.

Jedným z cieľov tejto diplomovej práce je naštudovanie si problematiky a prebádanie už existujúcich metód. To sme predstavili v predchádzajúcich dvoch kapitolách. Existuje veľké množstvo metód, ktoré sa zaoberajú lokálnymi príznakmi. Ak sú metódy doplnené o farebnú informáciu rôznymi prístupmi, množina metód sa ešte rozrastie. Z tohto dôvodu sme si vyčlenili oblasť nášho záujmu. Pre analýzu existujúcich metód sme sa zamerali na metódy obsahujúce detektor aj deskriptor, presnejšie metódy SIFT a SURF. V druhej kapitole sme predstavili preskúmané modifikácie. V konečnom dôsledku sme sa rozhodli zamerať na metódu SIFT a navrhnuť jej modifikáciu s využitím farebnej informácie, ktorá sa spracovaním farby odlišuje od predstavených metód v druhej kapitole.

Ďalším z cieľov práce je návrh a implementácia nami navrhnutej farebnej metódy a na záver porovnanie výsledkov a zhodnotenie jej úspešnosti.

3.2 Softvér

Na návrh metódy sme použili softvér MATLAB, verziu MATLAB R2015a, ktorý používame spoločne aj s *Image Processing Toolbox*. Tento softvér sme zvolili hlavne preto, že obsahuje množstvo implementovaných, pre nás využiteľných funkcií. Taktiež aj preto, že mnoho nových metód v spracovaní obrazu má práve implementáciu v MATLAB-e.

MATLAB je prostredie, ktoré je veľmi dobre využiteľné pri analýze a navrhovaní systému, na rozvoj sietí a podobne. Je optimalizovaný pre riešenie inžinierskych a vedeckých problémov, modelovanie, simulácie, analýzu údajov alebo spracovania signálov.

Prostredie MATLAB (*matrix laboratory*) bolo prvotne navrhnuté pre jednoduchý prístup k matematickým knižniciam LINPACK (*Linear System Package*) a EISPACK (*Eigen System Package*). Je to interaktívny systém, ktorý sa špecializuje hlavne na maticovú implementáciu rôznych technických výpočtov. Veľkou výhodou je možný import a export programov vytvorených v jazyku C alebo aj knižnice vytvorené pre jazyk Java, ktoré umožňujú dobré prepojenie s týmto jazykom. Veľkou výhodou softvéru MATLAB je jeho doplnenie množinou aplikačno-špecifických riešení, nazývaných *toolbox-y*. Pre riešenia spracovania obrazu využívame množinu funkcií zahrnutých v *Image Processing Toolbox*, ktoré rozširujú prostredie MATLABu.

Jednou z alternatív pre tvorbu systému bolo použitie knižnice v OpenCV (Open Source Computer Vision). Je to knižnica pre real-time počítačové videnie v jazykoch C,C++ alebo Python. Obsahuje množstvo optimalizovaných algoritmov pre strojové učenie, segmentáciu a mnoho ďalších.

3.3 Základný model

V našom výskume sme sa rozhodli upriamiť pozornosť na metódu *SIFT*. Navrhli sme metódu, ktorá predstavuje modifikáciu metódy *SIFT*, s využitím farebnej informácie z obrazu. Mnoho prebádaných existujúcich metód zakomponovalo farbu pri počítaní deskriptora. My sme si zvolili prístup využitia farebnej informácie obrazu už pri detekcii zaujímavých bodov. Jednou z hlavných inšpirácií boli príznakové mapy pre biologicky inšpirovaný Ittiho model navrhnutý v [17].

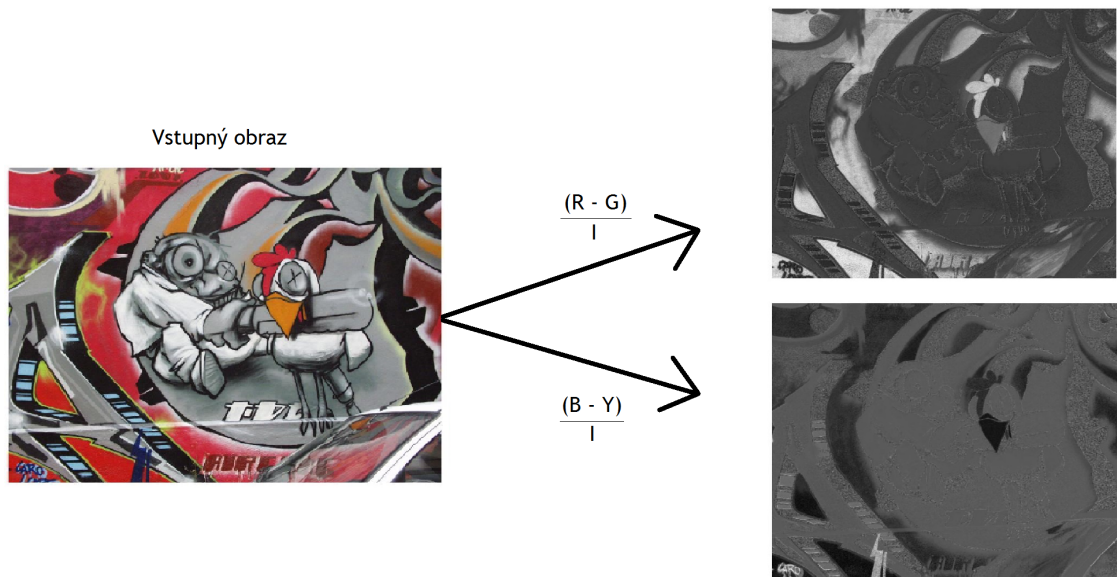
3.3.1 Získanie farebného príznaku

Človek disponuje piatimi zmyslami, z ktorých je zrak považovaný za najdôležitejší zo všetkých zmyslov. Pomocou zmyslového orgánu zraku, čiže okom, dokáže človek vnímať okolitý svet vďaka vizuálnym obrazom a registrovať až 80% informácií zo svojho okolia. Farby sú veľmi dôležité pre ľudské vnímanie a pozornosť. Ich vnímanie spôsobuje reakcia

oka pomocou farebných receptorov - čapíkov na rozličné vlnové dĺžky svetla, a to červenú, zelenú a modrú. Citlivosť jednotlivých čapíkov sa ale prekrýva, čím sa dokážu aktivovať nielen pri primárnych farbách, ale aj pri ich kombináciách. Tento proces skladania farby z čapíkov sa nazýva *trichromatický proces*.

V [17] autori pracujú s obrazom v RGB farebnom priestore upravenom podľa ľudského vizuálneho systému a spracovania farby, ktoré prebieha v hlbších štruktúrach sietnice a mozgovej kôry s charakterom oponentných procesov a trichromatickej teórie farieb.

Teóriu oponentných farieb [35][14] vytvoril Ewald Hering tým, že zrealizoval fyziologické výskumy na základe teórie J.W.Goetheho. Týmito výskumami potvrdil Goetheho tvrdenie, že niektoré farby sa k sebe veľmi dobre hodia a iné zase vôbec, čo ho viedlo k navrhnutiu oponentného princípu kódovania farieb prebiehajúceho na sietnici. V gangliových bunkách na sietnici sa formulujú tri oponentné kanály, jeden *achromatický čierno-biely* a dva *chromatické červeno-zelený* a *žlto-modrý*.



Obr. 3.1: Rozdelenie vstupného obrazu na chromatické oponentné kanály.

Príznakové mapy vytvárané v [17], ktorými sme sa inšpirovali, využívali práve tieto dva chromatické kanály oponentných procesov. Intenzita obrazu I , ktorá je potrebná pre vytvorenie chromatických kanálov, je získaná z váhovaného súčtu farebných kanálov RGB obrazu, s rozdielnymi koeficientmi pre každý kanál 3.1 [3], na základe meraných intenzít vnímania ľudí bez poruchy trichromatického vnímania farieb:

$$I = 0.2126 * R + 0.7152 * G + 0.0722 * B \quad (3.1)$$

Ako vidieť podľa hodnôt koeficientov v 3.1, ľudský zrak je najcitlivejší na zelenú farbu a naopak, najmenej citlivý na farbu modrú.

V našom výskume rozdeľujeme vstupný farebný obraz v RGB na dva chromatické oponentné kanály, vypočítané pomocou kombinácie jednotlivých farebných kanálov RGB, predelených získanou intenzitou obrazu I , vid'. obrázok 3.1. Červeno-zelený oponentný kanál R^+G^- , ako vidíme v 3.2, je vytváraný s použitím červeného a zeleného kanála RGB, žlto-modrý oponentný kanál B^+Y^- , vid'. 3.3, je vytváraný z modrého kanála RGB a dopočítaného žltého kanála, vid'. 3.4.

$$R^+G^- = \frac{(R - G)}{I} \quad (3.2)$$

$$B^+Y^- = \frac{(B - Y)}{I} \quad (3.3)$$

$$Y = \frac{(R + G)}{2} \quad (3.4)$$

3.3.2 Detekcia zaujímavých bodov

Detekcia zaujímavých bodov prebieha na dvoch chromatických kanáloch zvlášť. Pre oba prípady je na detekciu použitý detektor SIFT [24] [23].

SIFT využíva na vyhľadávanie zaujímavých bodov škálovú pyramídu (scale space), v ktorej sú obrázky konvolované Gaussian filtrom s rozdielnou škálou a rozdielnym Gaussianom po sebe idúcich obrázkov. Vytvorí sa oktávy obrazu, ktoré predstavujú podvzorkovaný obrázok do k veľkostí. Zaujímavé body reprezentujú lokálne extrémny rozdielu Gaussiánov (Difference of Gaussians – DoG) vstupného obrazu v rôznych škálach, s využitím farebnej informácie z príznakových máp oponentných filtrov. Škálová pyramída obrazu je definovaná ako funkcia $L(x, y, \sigma)$ (3.5), ktorá je vytvorená z konvolúcie vstupného obrazu $I(x, y)$ s Gaussianom $G(x, y, k\sigma)$ v škále σ .

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \quad (3.5)$$

kde $*$ je operácia konvolúcie v x a v y a Gaussova funkcia je:

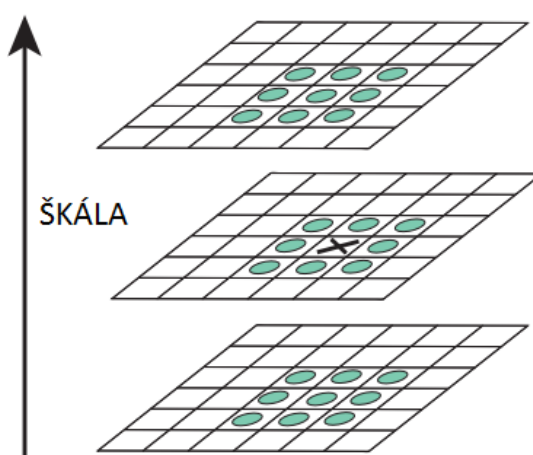
$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2} \quad (3.6)$$

DoG obraz pre hľadanie lokálnych extrémov je definovaný ako:

$$\begin{aligned} D(x, y, \sigma) &= L(x, y, k\sigma) - L(x, y, \sigma) \\ &= (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) \end{aligned} \quad (3.7)$$

V DoG obrázku sa následne hľadajú lokálne extrémny v okolí $3x3x3$ (obrázok 3.2), čo znamená, že sa hľadajú v okolí $3x3$ aktuálneho obrazu, ale taktiež v rovnakom okolí, aj v nasledujúcom a predchádzajúcom obrázku v rámci škálovej pyramídy. Nájdené lokálne

extrémy (minimá alebo maximá) sú identifikované ako zaujímavé body, ktoré sú zároveň aj kandidátmi pre kľúčové body, ktorými sa stanú len dostatočne kontrastné zaujímavé body. Oktáva, v ktorej bol zaujímavý bod identifikovaný, predstavuje škálu daného zaujímavého bodu a zároveň určuje veľkosť jeho okolia, ktoré sa bude používať pri extrakcii deskriptora.



Obr. 3.2: Hľadanie lokálnych extrémov porovnávaním pixla v okolí $3 \times 3 \times 3$ [24].

3.3.3 Vytváranie deskriptora

Hodnota deskriptora je počítaná pre každý kľúčový bod z histogramov orientácií oblasti veľkosti 4×4 jeho okolia. Najprv je zistená dominantná orientácia každého kľúčového bodu, ku ktorej je relatívne vytvorený deskriptor bodu. Následne sa obraz navzorkuje 16 bodmi. Pre každú vzorku obrazu $L(x, y)$ je vypočítaná veľkosť $m(x, y)$ a orientácia $\theta(x, y)$ farebného gradientu oponentných kanálov:

$$m(x, y) = \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2} \quad (3.8)$$

$$\theta(x, y) = \arctan((L(x, y+1) - L(x, y-1)) / (L(x+1, y) - L(x-1, y))) \quad (3.9)$$

Vypočítaným veľkostiam a orientáciám farebných gradientov oponentných kanálov sú priradené váhy pomocou Gaussového kruhové okna, vid' obr. 1.2, z ktorých sú vytvorené histogramy orientácií pre šestnásť (4×4) regiónov z okolia kľúčového bodu. Dĺžka šípky v histograme predstavuje sumu veľkostí farebných gradientov v blízkosti daného smeru oblasti. Ako vidíme na obrázku 1.2, v každom histograme je osem smerov, čo pri oblasti veľkosti 4×4 predstavuje pre jeden kľúčový bod deskriptor obsahujúci 128 hodnôt.

3.3.4 Navrhnuté riešenie

Stephen Engel a kol. v [9] skúmali ľudský vizuálny cortex a jeho farebné ladenie pomocou magnetickej rezonancie. Ľudský vizuálny cortex je súčasťou ľudského vizuálneho systému, ktorý predstavuje časť centrálného nervového systému. Ten pomocou spracovania signálov z fotoreceptorov na sietnici oka, čiže čapíkov, interpretuje informácie o farbe. Vo svojich experimentoch [30] zistili, že najsilnejšia reakcia vo vizuálnej kortikálnej oblasti mozgu je na červeno-zelené podnety. Pri žlto-modrých podnetoch je reakcia taktiež silná, no oproti červeno-zeleným podnetom rýchlo klesá.

Kombinácia trichromatického procesu a procesu oponentného kódovania farieb bola donedávna považovaná za nemožnú. Trichromatický proces predsa hovorí o skladaní farieb z viacerých čapíkov, proces oponentných farieb naopak o zisťovaní farby z ich rozdielu. Zistilo sa však, že oko pracuje na oveľa komplikovanejšej úrovni, a práve tieto dva procesy dokonale kombinuje. Ako základný model pre našu metódu sme si preto vybrali prístup založený na experimentoch v [9], s využitím chromatických oponentných kanálov, navrhnutých v [17] a trichromatickej teórie farieb, na spracovanie farebného obrazu a metódu SIFT.

Ideou riešenia je zakomponovanie farebnej informácie z obrazu do metódy SIFT nahradením šedotónových informácií tak, že kľúčové body sú detegované na dvoch separovaných chromatických kanáloch zvlášť. Achromatický bielo-čierny kanál pri detekcii a deskripcii nie je využívaný. Vytváranie deskriptorov prebieha tiež oddelene. Až pred príznakovým párovaním sú zret'azené pre zjednotenie informácií získaných z oboch kanálov.

Kapitola 4

Implementácia

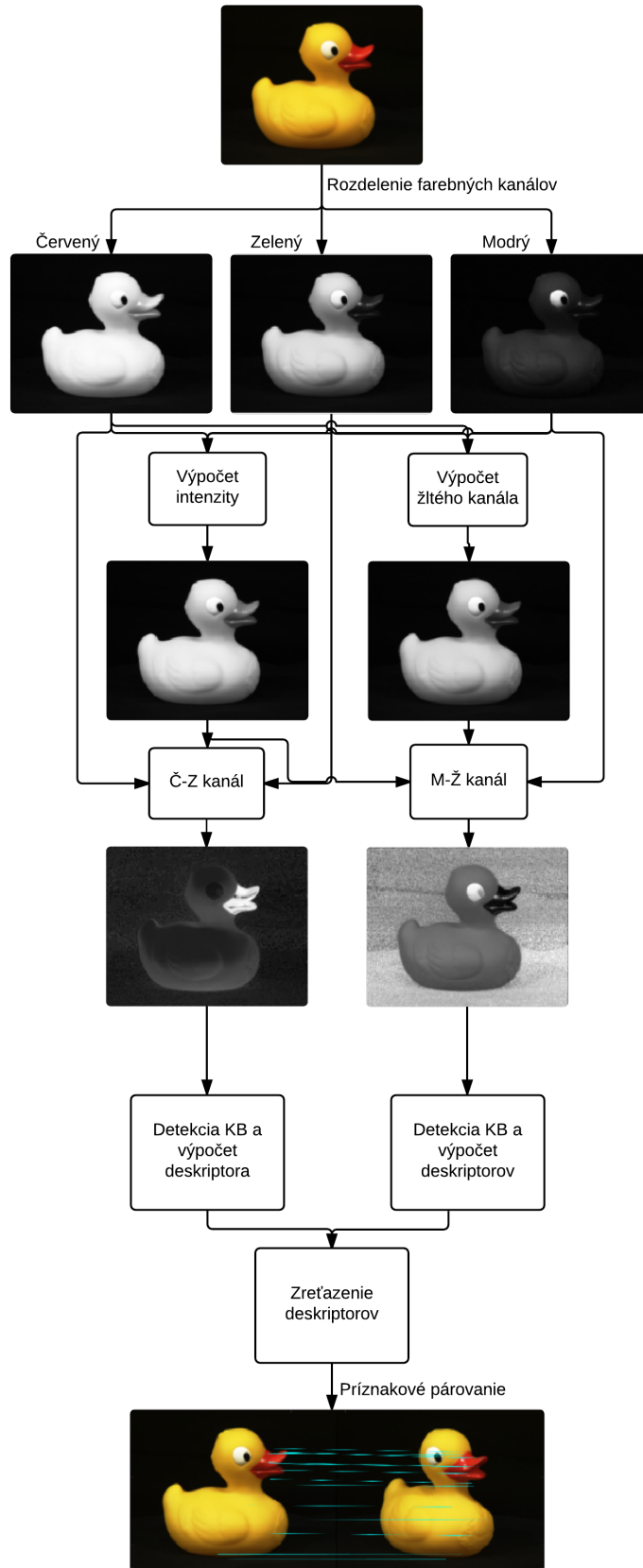
V tejto kapitole je popísaná implementácia navrhutej metódy v prostredí MATLAB. Naprogramovanie metódy spočíva v modifikácii metódy SIFT navrhnutou Lowem [24] [23] o spracovanie farebnej informácie obrazu na základe ľudského vizuálneho systému. Metóda je rozdelená na tri hlavné fázy:

1. predspracovanie
2. detekcia zaujímavých bodov a vytváranie deskriptora
3. príznakové párovanie

Vybraná implementácia metódy SIFT poskytuje základ pre implementáciu návrhu metódy na spracovanie farebného obrazu. Najhlavnejšou časťou metódy je predspracovanie obrazu, ktoré nám pre ďalšie dve podstatné fázy poskytuje potrebné a ďalej spracovávané farebné informácie obrazu na extrakciu lokálnych príznakov. Pre nájdenie farebného príznaku spracovávame obraz v RGB farebnom priestore. Pomocou kanálov RGB vytvárame oponentné kanály obrazu, na ktorých je následne spustená detekcia zaujímavých bodov a deskripcia. Zvolený prístup spúšťa detekciu a deskripciu na oponentných kanáloch zvlášť, čím sú detegované zaujímavé body a vypočítané deskriptory rozdelené podľa kanálov. Pre následné párovanie sú deskriptory zret'azené.

Pri príznakovom párovaní boli okrem metódy SIFT použité dva rôzne prístupy metódy SIFT, s použitím farebných priestorov RGB a HSV, ktoré sú implementované pre potreby validácie a sú viac spomenuté v 5. kapitole tejto práce.

Na obrázku 4.1 je znázornený vývojový diagram navrhutej implementácie metódy



Obr. 4.1: Vývinový diagram navrhutej metódy.

4.1 Predspracovanie

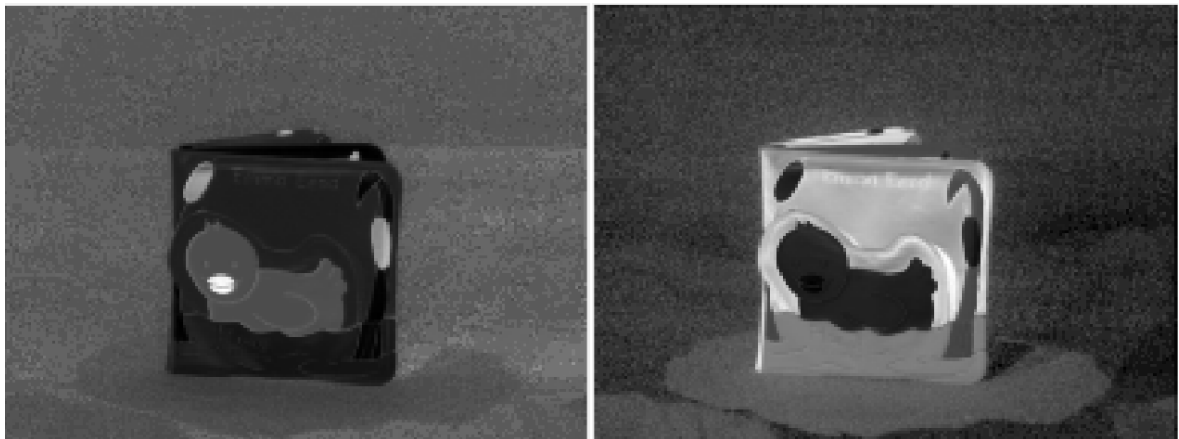
Hlavnou úlohou fázy predspracovania je načítanie obrázka, jeho normalizácia a konverzia farebného RGB obrazu na oponentné kanály. Všetky tieto súčasti sú zahrnuté na začiatku funkcie OppSift.

Na vstupe funkcie sa nachádza načítanie obrázka. V prvom kroku je potrebné zistiť, či je obrázok farebný, presnejšie, či obsahuje 3 farebné kanály RGB. Tie sa oddelia, aby bolo možné s nimi ďalej samostatne pracovať. Pre vytvorenie oponentných modelov je vypočítaná intenzita obrazu I [3] pomocou súčtu preváňovaných kanálov RGB, spomenutých v kapitole 3.3.1 v (3.1).

V ďalšom kroku sú pomocou intenzity I vypočítané oponentné kanály obrazu. Najprv je pomocou červeného a zeleného kanála vypočítaný žltý kanál (3.4). Červeno-zelený oponentný kanál a žlto-modrý kanál je vypočítaný podľa (3.2) a (3.3) v kapitole 3.3.1. Z dôvodu predchádzania vytvárania duplicitných hrán v obraze, boli tieto dva kanály preškálované na rozsah $\langle 0, 255 \rangle$ pomocou Min-max normalizácie:

$$v' = \frac{v - \min A}{\max A - \min A} * (newMax - newMin) + newMin, \quad (4.1)$$

kde A je skúmaný kanál a $\langle newMin, newMax \rangle = \langle 0, 255 \rangle$. Takto predspracovaný obraz, vid' obrázok 4.2, je vstupom pre detekciu zaujímavých bodov a deskripciu.



Obr. 4.2: Normalizované kanály použité ako vstup pre detekciu a deskripciu. Vľavo červeno-zelený kanál, vpravo žlto-modrý

4.2 Detektor a deskriptor

Vo fáze detekcie a deskripcie vo funkcii *OpponentSift* bola použitá SIFT C++ implementácia na detegovanie zaujímavých bodov a na extrakciu SIFT deskriptorov kl'účových bodov z obrázka vytvorená tvorcom metódy Lowem¹[24]. Vstupom do tejto funkcie je predspracovaný obrázok, popísaný v časti predspracovania v tejto kapitole. Výstupom je matica detegovaných zaujímavých bodov na oboch kanáloch a pole deskriptorov kl'účových bodov.

Pre správnu extrakciu je obrázok prevedený do formátu PGM, ktorý je čitateľný pre funkciu *keypoints*,² ktorej výstupom je súbor **názov.key** (ďalej len KEY). Štruktúra súboru KEY je nasledovná:

2 x celé číslo - Prvé číslo **num** predstavuje počet nájdených kl'účových bodov, druhé číslo **len** počet hodnôt deskriptora pre jeden kl'účový bod.

num * (4 x racionálne číslo) - Štyri hodnoty určujú polohu kl'účového bodu, a to v poradí: riadok, stĺpec, škála a orientácia. Takéto hodnoty kl'účových bodov v súbore nájdeme num krát.

num * (len x celé číslo) - Hodnotový deskriptor dĺžky 128 pre každý kl'účový bod.

Súbor KEY je zanalyzovaný kvôli prípadným chybám. Prvým krokom je pokus o otvorenie súboru. Následne sú načítané hodnoty num a len z prvého riadka súboru a je testované, či sa hodnota len rovná 128, teda, či majú deskriptory v súbore správnu dĺžku.

V ďalšom kroku sú vytvorené dve matice, ktoré sú naplnené rozdelením súboru KEY na detegované kl'účové body a vypočítané deskriptory. Celý popísaný postup je aplikovaný pre oba kanály vzniknuté v predspracovaní obrazu zvlášť, čím dostávame dve matice kl'účových bodov a dve matice deskriptorov podľa kanálu.

Pred príznakových párovaním je zistené, či matice kl'účových bodov neobsahujú identické body, aby sa zabránilo duplicitě deskriptorov, čo by mohlo spôsobiť nepresnú analýzu a zlé výsledky validácie. Pod identickými bodmi sa rozumie taký bod detegovaný v červeno-zelenom kanáli, ktorý má rovnakú hodnotu riadka, stĺpca, škály aj orientácie, ako bod detegovaný v modro-žltom kanáli. Kl'účové body z prvého kanála sú porovnávané s hodnotami bodov detegovaných v druhom kanáli.

Pri nájdení bodu obsiahnutého v oboch kanáloch sú identifikovaný bod a jeho deskriptor z matice kl'účových bodov detegovaných na jednom z kanálov a z matice vypočítaných

¹Pre funkciu implementovanú autorom metódy sme sa rozhodli z dôvodu, že funkcia je dostatočne efektívna a rýchla. Nová implementácia tejto metódy by bola zbytočná.

²Funkcia nie je samostatne nájditel'ná. Je obsahom SIFT spustitel'nej implementácie 'siftWin32.exe'.

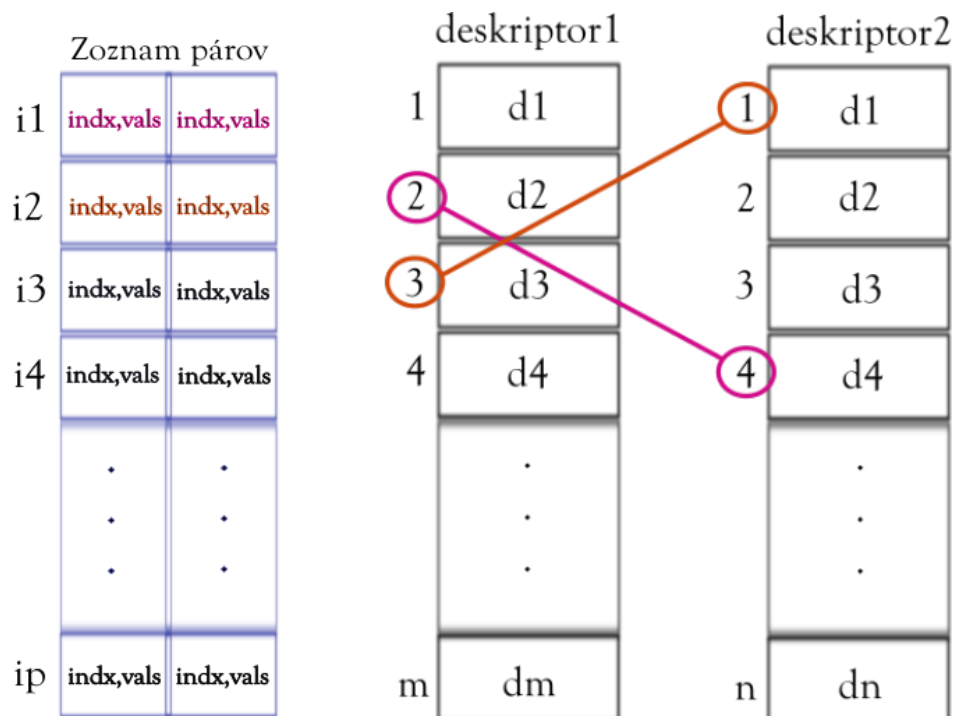
deskriptorov odstránené. V matici kľúčových bodov z druhého kanála a v matici ich deskriptorov je ale zachovaný.

Po filtrácii sú matice deskriptorov a kľúčových bodov, pre následné príznakové párovanie, zret'azené³, čím sú vytvorené matice, obsahujúce popísanú farebnú informáciu z oboch oponentných kanálov. Tieto matice sú spolu s načítaným obrazom výstupom celej funkcie *OppSift*.

4.3 Príznakové párovanie

Podstatou príznakového párovania je postupné zisťovanie počtu zhôd, čiže párovanie, medzi hlavným obrázkom a množinou obrázkov, s ktorými sa porovnáva postupne.

Párovanie je zrealizované pomocou metódy *najbližšieho suseda* (nearest neighbour), ktorá vyhl'adáva najbližšieho suseda v aktuálnej dvojici deskriptorov porovnávaných obrazov, čiže medzi deskriptormi hlavného obrazu a deskriptormi obrazu, s ktorým sa páruje. Pre jednu dvojicu deskriptorov technika najbližšieho suseda nájde pre každý deskriptor hlavného obrazu (teda pre každý riadok v poli deskriptorov) **najbližšieho suseda** a **druhého najbližšieho suseda** v poli deskriptorov druhého obrazu. Za najbližšieho suseda je zvolený taký riadok v poli deskriptorov, ktorého euklidovská vzdialenosť je najmenšia.



Obr. 4.3: Princíp príznakového párovania pomocou metódy najbližšieho suseda.

³Pod zret'azením myslíme zjednotenie množín jednotlivých deskriptorov z kanálov do jednej množiny bez ďalších úprav.

Na obrázku 4.3 vidíme princíp párovania pomocou metódy najbližšieho suseda. V aktuálnej dvojici deskriptorov vyhl'adá technika najbližšieho a druhého najbližšieho suseda. Do zoznamu párov je uložený index (`indx (1)`) pol'a deskriptorov a euklidovská vzdialenosť (`vals (1)`) najbližšieho suseda a index (`indx (2)`) a euklidovská vzdialenosť (`vals (2)`) druhého najbližšieho suseda.

Dôležitou hodnotou pri porovnávaní technikou najbližšieho suseda je počet zhôd. Jednotlivé páry v zozname párov ale nepredstavujú automaticky aj zhodu. Za zhodu považujeme prípad, keď platí, že euklidovská vzdialenosť najbližšieho suseda je menšia ako euklidovská vzdialenosť druhého najbližšieho suseda vynásobená `distRatio`, ktoré bolo autorom metódy SIFT [24] [23] zadané ako 0.6.

Ako sme spomenuli na začiatku kapitoly, metóda bola implementovaná v prostredí MATLAB bez vytvárania GUI prostredia. Funguje pomocou dvoch funkcií:

- `OppSift.m`
- `Match.m`

Funkcia `OppSift.m` predspracuje obraz, deteguje zaujímavé body a vypočíta deskriptory. Funkcia `Match.m` vyhl'adáva počet zhôd v porovnávaných obrazoch pomocou príznakového párovania.

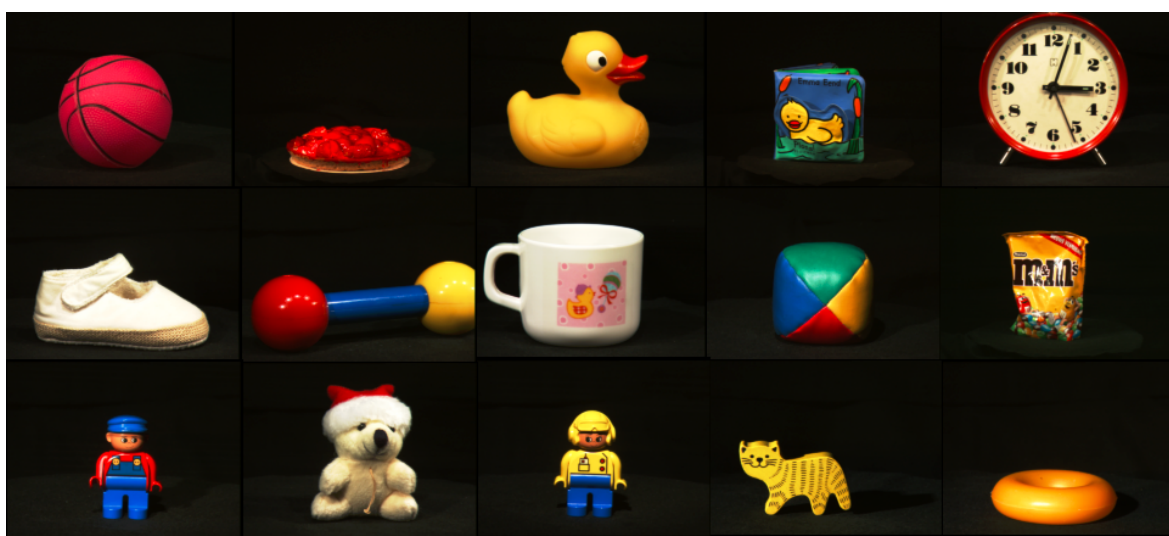
Kapitola 5

Validácia

V piatej kapitole tejto práce sa zaoberáme overením výsledkov navrhnutej metódy, porovnaní s metódou SIFT pracujúcou so šedotónovým obrazom, RGB-SIFT a HSV-SIFT a vyhodnotením najlepších výsledkov. Metódy boli testované na vybranej databáze, ktorá je predstavená v ďalšej časti kapitoly.

5.1 Databáza

Na overenie a zhodnotenie výsledkov sme vybrali databázu ALOI [11], ktorej ukážka je znázornená na obrázku 5.1. Obsahuje veľké množstvo objektov v rôznych zobrazovacích podmienkach. Hlavne sme sa zamerali na obrázky so zmenou smeru osvetlenia, s rôznou intenzitou osvetlenia a zmenou teploty farby pomocou osvetlenia. Vybrané objekty použité na analýzu boli vyberané tak, aby bola obsiahnutá najširšia škála farieb.

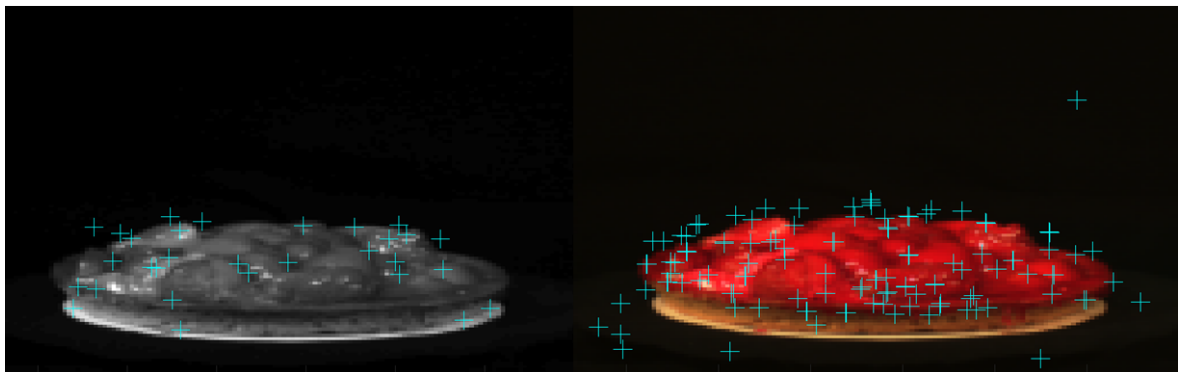


Obr. 5.1: Ukážka použitej databázy

5.2 Analýza

Hlavným kritériom metódy bola fotometrická robustnosť deskriptora. Analyzovali sme, či je deskriptor invariantný na zmeny svetelného zdroja, tieňa, tieňovania a farby. Analýza bola vykonávaná porovnaním s metódou SIFT.

Pre zistenie fotometrickej robustnosti deskriptora boli zvolené dve množiny obrázkov vybraných z databázy. Prvá množina obsahovala päť sád obrázkov so zmenou teploty farby. V druhej množine boli obrázky so zmenou smeru osvetlenia z ôsmich strán, ktoré boli taktiež rozdelené do piatich sád podľa zastúpenia objektov. Na obrázku 5.2 vidíme detekciu zaujímavých bodov na objekte koláč metódou SIFT a nami navrhovanou metódou.

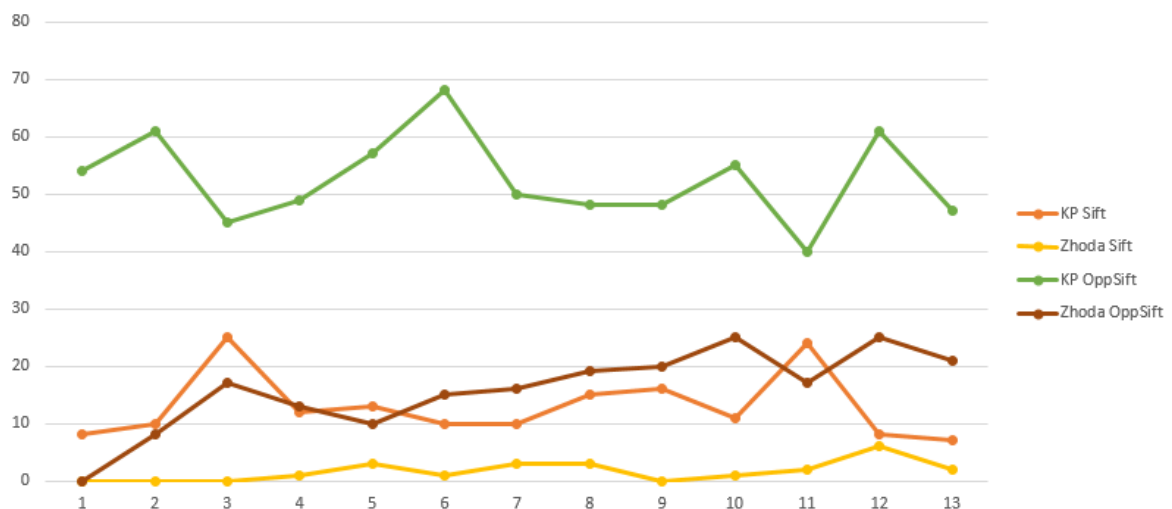


Obr. 5.2: Detekcia zaujímavých bodov metódou SIFT (vľavo) a navrhovanou metódou (vpravo)

Ako vidíme na obrázku 5.2, počet detegovaných zaujímavých bodov je výrazne vyšší oproti metóde SIFT. Výrazne lepšie výsledky sme dosiahli pri obrázkoch so zmenou teploty farby aj so zmenou smeru osvetlenia. Môžeme povedať, že počet nájdených bodov v oponentných kanáloch je väčší ako pri šedotónových obrazoch. Je známe, že zvyšovaním počtu detegovaných bodov sa zvyšuje výkonnosť rozpoznávacieho procesu. Tak môžeme zhodnotiť, že navrhovaná metóda dosahuje lepšie výsledky oproti metóde SIFT, pokiaľ ide o počet detegovaných zaujímavých bodov.

Pre jednoznačné určenie invariantnosti deskriptora nás zaujímal počet zhôd pri jednotlivých typoch obrázkov pri príznakovom párovaní. Pri oboch množinách obrázkov dosahujeme výsledky na úrovni, ktorú sme predpokladali. Pri obrázkoch so zmenou teploty farby síce dosahuje mierne vyššie počty zhôd v porovnaní s metódou SIFT, no predpokladané počty zhôd boli vyššie. Nižší počet zhôd môže byť spôsobený filtrovaním pixlov pod určitou kontrastnou hranicou pred vytváraním deskriptora, čím získavame menší počet príznakových deskriptorov pre následné príznakové párovanie. Aj keď sme nedosiahli predpokladaný počet zhôd, celkovo pracuje navrhovaná metóda porovnateľne, až mierne lepšie, ako metóda SIFT.

Graf 5.3 znázorňuje počet detegovaných kľúčových bodov a počet zhôd v metóde SIFT a v našej metóde na obrazoch so zmenou smeru osvetlenia. Ako vidíme, navrhovaná metóda dosahuje výrazne lepšie výsledky pri detekcii kľúčových bodov, no hlavne dosahuje výrazne lepšie výsledky pri počte zhôd. Metóda SIFT ani v jednom prípade neprekročila hranicu desiatich zhôd pri príznakovom párovaní. Navrhovaná metóda pri obrázkoch so zmenou smeru osvetlenia dosahuje výrazne vyšší počet zhôd v príznakovom párovaní v porovnaní s metódou SIFT. Môžeme teda povedať, že farebná informácia zahrnutá v oponentných kanáloch navrhnutej metódy výrazne zlepšila metódu, splnila očakávané kritériá a je dostatočne vhodná pre rozpoznávanie a klasifikáciu objektov z obrazu.



Obr. 5.3: Graf zobrazujúci počty detegovaných bodov a zhôd v obrázkoch kocky so zmenou smeru osvetlenia. X-ová os znázorňuje jednotlivé analyzované obrázky, Y-ová os predstavuje hodnoty detekcie a zhôd.

5.3 Klasifikácia obrázkov

Ako sme zhodnotili v predchádzajúcej časti kapitoly, metóda splnila očakávané kritériá. Pre overenia nášho tvrdenia sme testovali databázu, ktorá obsahuje 303 obrázkov s rôznymi objektami so zmenou teploty farby, zmenou smeru osvetlenia a taktiež obsahuje aj určité množstvo stereo obrázkov pre odhalenie prípadných nedostatkov. Ako vidíme v tabuľke 5.1, z objektov obsiahnutých v databáze sme vybrali 10 objektov, ktoré pomocou navrhnutej metódy v množine obrázkov rozpoznávame a klasifikujeme. Databáza obrázkov obsahuje aj objekty, ktoré nie sú zastúpené vo vybraných desiatich hlavných objektov. Z tohto dôvodu bol pri rozbere výsledkov zavedený prah počtu zhôd, od ktorého je objekt na obraze klasifikovaný, ako hlavný objekt na porovnanom obraze. Prah bol určený z výsledkov tak, aby maximalizoval počet správne klasifikovaných objektov.

	Objekt	počet obrázkov s objektom v databáze
1	budík	16
2	kačička	16
3	kačka	34
4	kocka	35
5	koláč	35
6	hračka	10
7	m&m	21
8	plyšák	10
9	teniska	12
10	lopta	40
11	iné objekty v databáze	74

Tabuľka 5.1: Tabuľka vybraných objektov a ich počet obsiahnutý v databáze vybraných 303 obrázkov.

Tabuľka 5.2 obsahuje percentuálne vyjadrený počet správne klasifikovaných obrázkov objektu hračka navrhnutou metódou a metódou SIFT pri rôzne nastavenom prahu. Počet percent predstavuje správnu klasifikáciu obrázkov objektu hračka z obrázkov klasifikovaných ako hračka pri zvolenom prahu počtu zhôd. Z tabuľky 5.2 je jasné, že navrhnutá metóda dosahuje lepšie výsledky pri všetkých zvolených prahoch.

hodnota prahu =	3	4	5	6
OppSIFT	28%	53%	77%	100%
SIFT	16%	27%	11%	20%

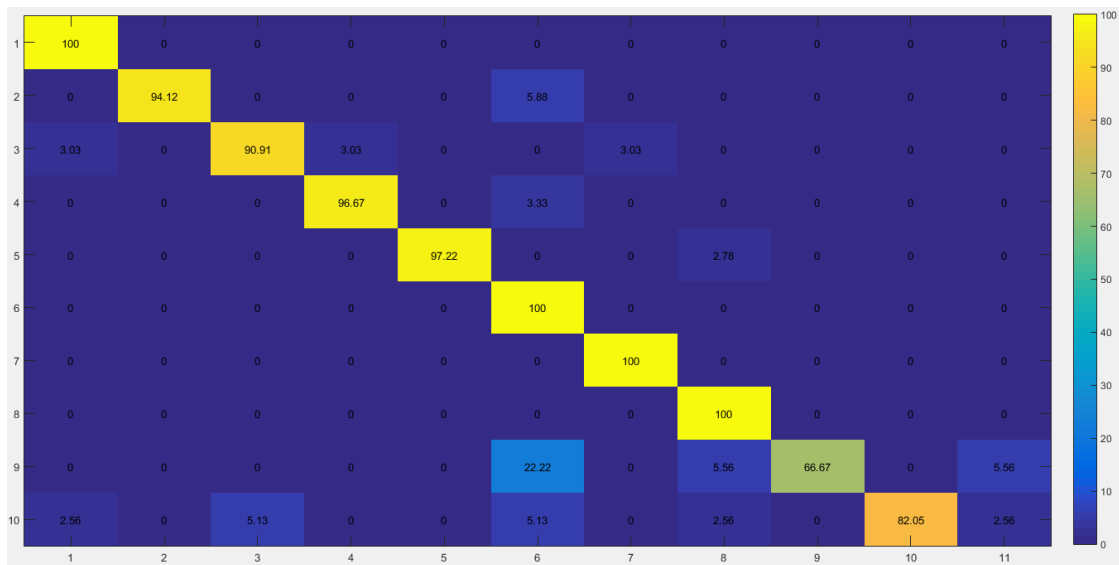
Tabuľka 5.2: Percentuálne zastúpenie správnej klasifikácie objektu hračka pri jednotlivých prahoch navrhnutou metódou a metódou SIFT.

V tabuľke 5.3 môžeme vidieť percentuálne zastúpenie zle klasifikovaných objektov, klasifikovaných ako objekt hračka. Ako vidíme, navrhovaná metóda postupne znižuje percento zlej klasifikácie, no pri metóde SIFT, pri zvolenom objekte, percento zlej klasifikácie nie je jednoznačne klesajúce a pri každom zvolenom prahu je dosť vysoké.

hodnota prahu =	3	4	5	6
nesprávny objekt OppSift	72%	47%	23%	0%
nesprávny objekt SIFT	84%	73%	89%	80%

Tabuľka 5.3: Percentuálne zastúpenie nesprávnej klasifikácie objektu hračka pri jednotlivých prahoch navrhnutou metódou a metódou SIFT

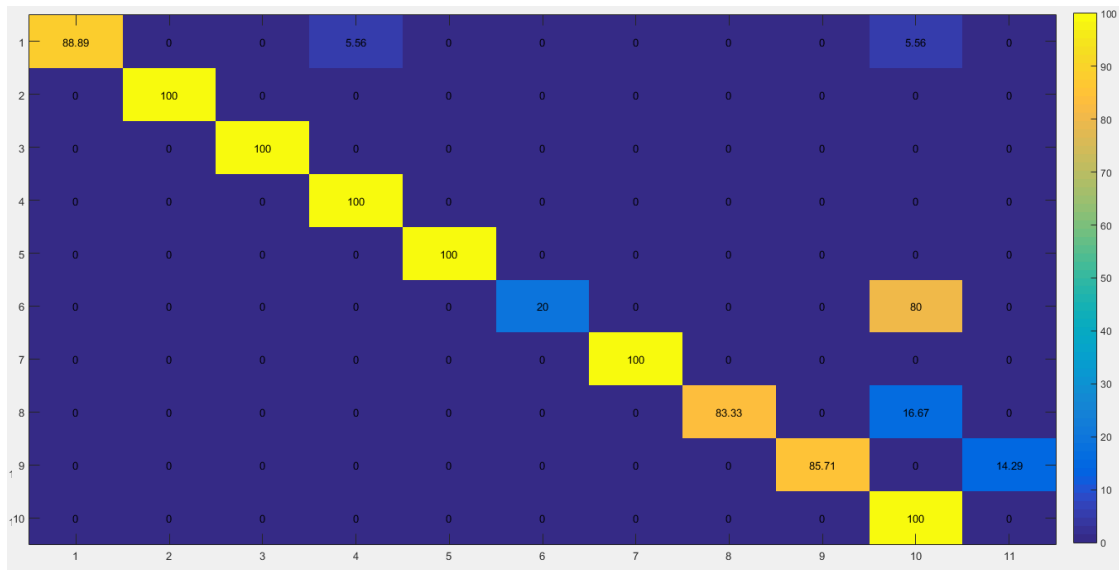
Okrem vhodnosti metód na klasifikáciu a rozpoznávanie objektov z množiny obrázkov nám validačné techniky môžu ponúknuť informácie o klasifikácii vzhľadom na jednotlivé triedy. V práci bolo využitých 10 tried, jedna pre každý obraz objektu vo vybraných hlavných objektoch (pozri tabuľku 5.1) a jedna trieda pre objekty, ktoré sa nevyskytujú v databáze hlavných objektov. Prvým krokom je vytvorenie matice zámen. Keďže počet vzoriek v jednotlivých triedach nie je konštantný, dávame do matice zámen namiesto absolútnych počtov klasifikovaných objektov ich percentuálny podiel v jednotlivých triedach. Pre lepšiu vizualizáciu matice zámen je použitá teplotná mapa.



Obr. 5.4: Matica zámen reprezentovaná teplotnou mapou pre klasifikáciu objektov pomocou navrhutej metódy

Matica zámen je kontingenčnou tabuľkou, do ktorej sa zapisujú počty vzoriek klasifikovaných do jednotlivých tried. Je vytvorená pre navrhovanú metódu a metódu SIFT s prahom rovným šesť. Riadky reprezentujú triedy¹, do ktorých objekty na obrázkoch patria. Stĺpce reprezentujú triedy, kam boli pomocou počtu zhôd z príznakového párovania metódou zaradené. Teda, ak sa obraz nachádza v prvom riadku a v prvom stĺpci, znamená to, že bol obrázok s objektom budík pomocou počtu zhôd klasifikovaný správne. Obraz, ktorý sa nachádza v treťom riadku a prvom stĺpci, je obraz s objektom kačka klasifikovaný ako objekt budík, čiže bol klasifikovaný nesprávne. Matica obsahuje špeciálny stĺpec s číslom 11 (iné objekty v databáze), do ktorých môže metóda pri zlej klasifikácii zaradiť objekty z tried. Obrázok 5.4 reprezentuje maticu zámen vyjadrenú teplotnou mapou pre navrhovanú metódu a obrázok 5.5 pre metódu SIFT.

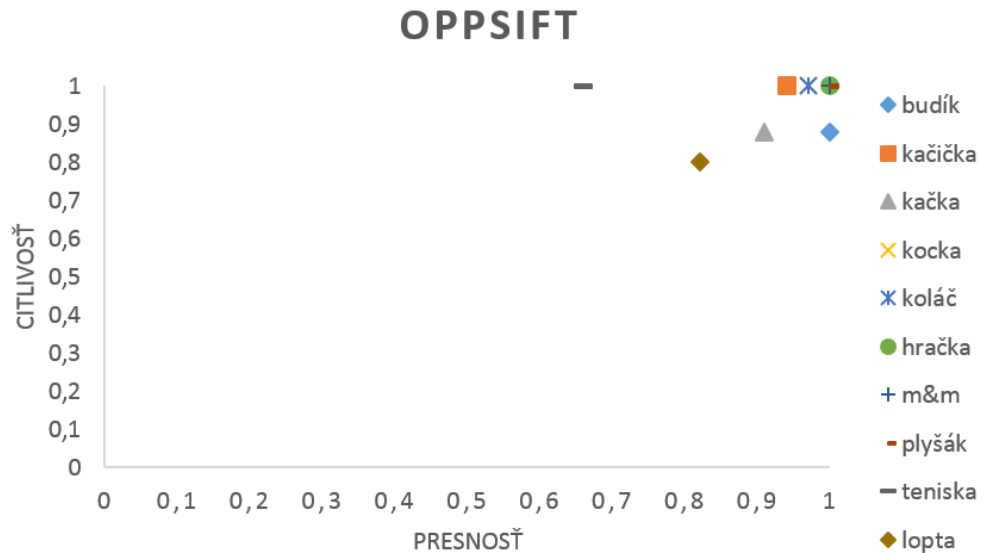
¹Objekty, ktoré zastupujú jednotlivé triedy, sú znázornené v tabuľke 5.1



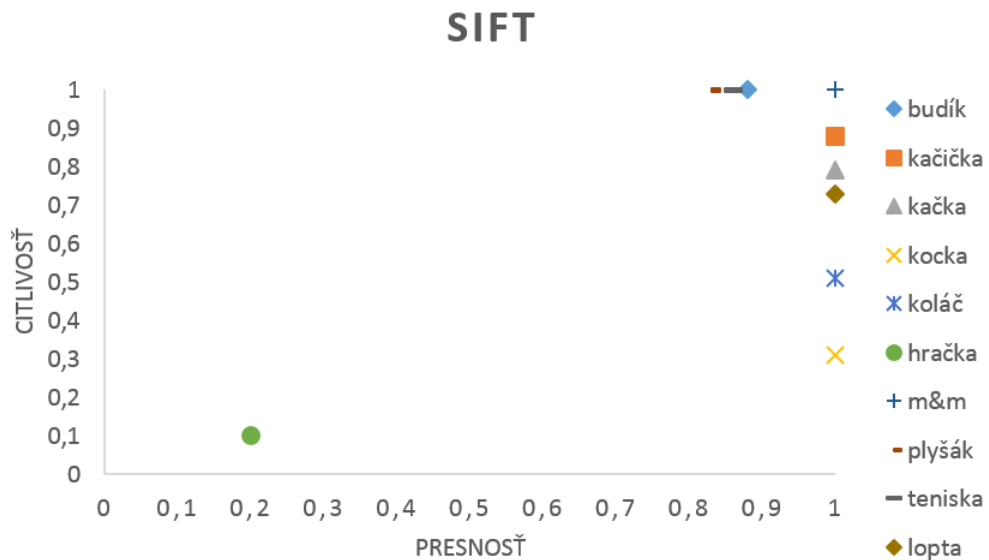
Obr. 5.5: Matica zámen reprezentovaná teplotnou mapou pre metódu SIFT

Po skonštruovaní matíc zámen môžeme odčítať hodnoty citlivosti (recall) a presnosti (precision) klasifikátora pre každú triedu. Citlivosť klasifikátora predstavuje správne klasifikované objekty, predelené počtom všetkých objektov v danej triede. Táto hodnota je ovplyvňovaná počtom obrázkov, na ktorých sa objekt vyskytuje, ale metóda ho nenájde (NN - nesprávne negatívne). Čím je vyšší počet nesprávne negatívnych výsledkov, tým je hodnota citlivosti nižšia, čiže klasifikátor nevie odhaliť všetky výskyty objektu. Ideálny klasifikátor vie odhaliť všetky obrázky s objektom a citlivosť klasifikátora je rovná jednej. Presnosť klasifikátora vyjadruje počet správne klasifikovaných objektov, predelený počtom všetkých objektov klasifikovaných ako daná trieda. Vyjadruje pravdepodobnosť, že na obrázkoch vybraných klasifikátorom sa daný objekt naozaj vyskytuje, čiže je obrázok správne pozitívny. Čím menej nesprávne pozitívnych obrázkov, tým viac sa hodnota presnosti približuje k jednotke. Obrázky 5.7 a 5.6 ukazujú úspešnosť klasifikácie metód vzhľadom na jednotlivé triedy.

Na grafoch hodnôt citlivosti a presnosti metód vidíme hodnoty citlivosti a presnosti klasifikácie vzhľadom na jednotlivé triedy. Vzhľadom na hodnoty za kvalitnejší klasifikátor môžeme skôr považovať navrhovanú metódu. Na obrázkoch 5.8 a 5.9 sú zobrazené ukážky príznakových párovaní objektov navrhovanou metódou a metódou SIFT, na ktorom bola založená klasifikácia objektov. Pri objektoch hračka, kocka a koláč ide o obrázky so zmenou smeru osvetlenia. Vidíme, že navrhovaná metóda výrazne zvýšila počet zhôd v párovaní oproti metóde SIFT, čím dosahujeme oveľa lepšie a presnejšie výsledky pri rozpoznávaní a klasifikácii tohto druhu obrázkov. Obrázok objektu teniska reprezentuje obrázky so zmenou teploty farby. Ako vidíme, aj pri tomto druhu obrázkov metóda navýšila počet zhôd. Celkovo môžeme povedať, že farebná informácia v navrhovanej metóde výrazne zvýšila počet zhôd v príznakovom párovaní a dosahuje lepšie výsledky pri rozpoznávaní a klasifikácii.



Obr. 5.6: Citlivosť a presnosť klasifikácie navrhnutou metódou na vybranej databáze.

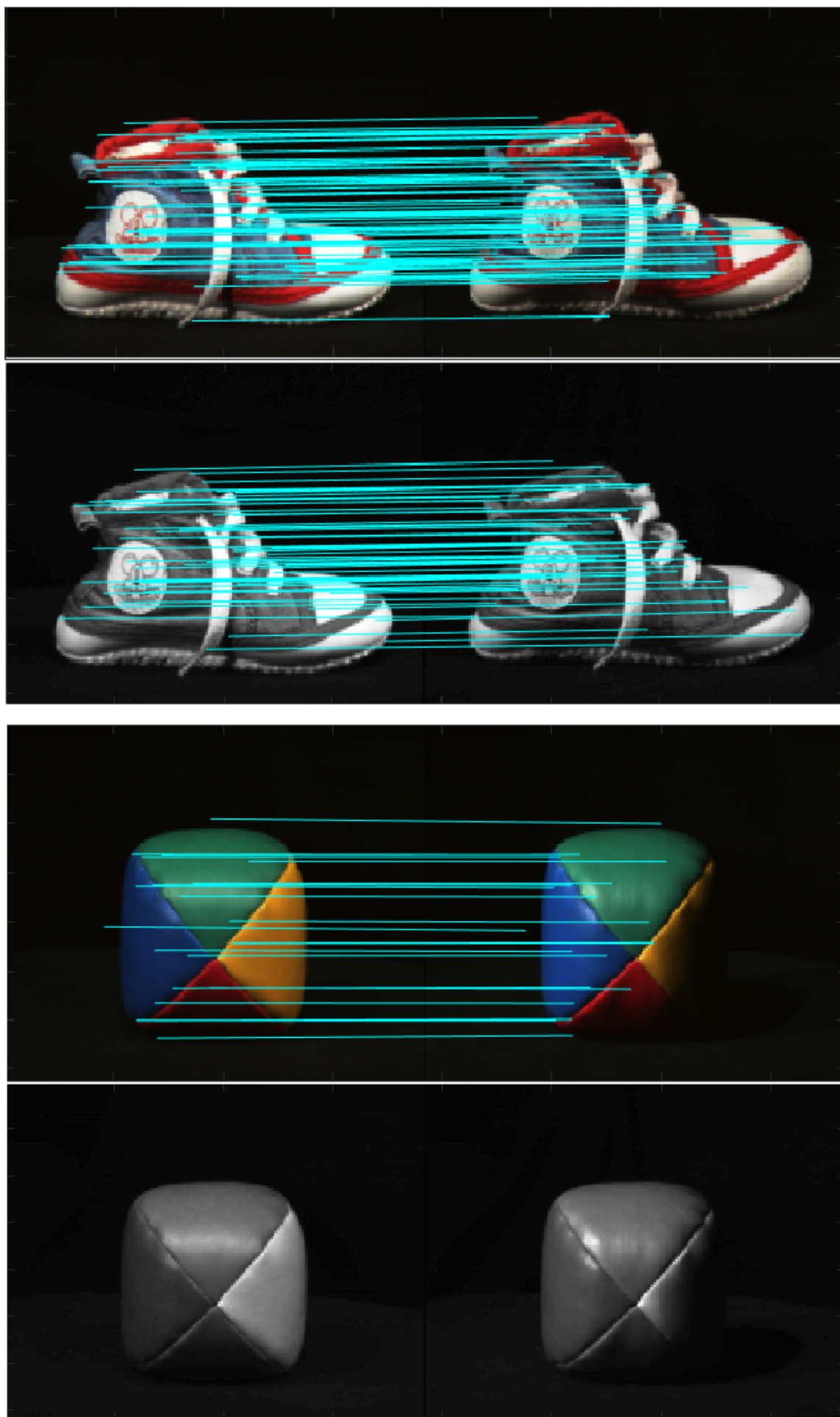


Obr. 5.7: Citlivosť a presnosť klasifikácie metódou SIFT na vybranej databáze.

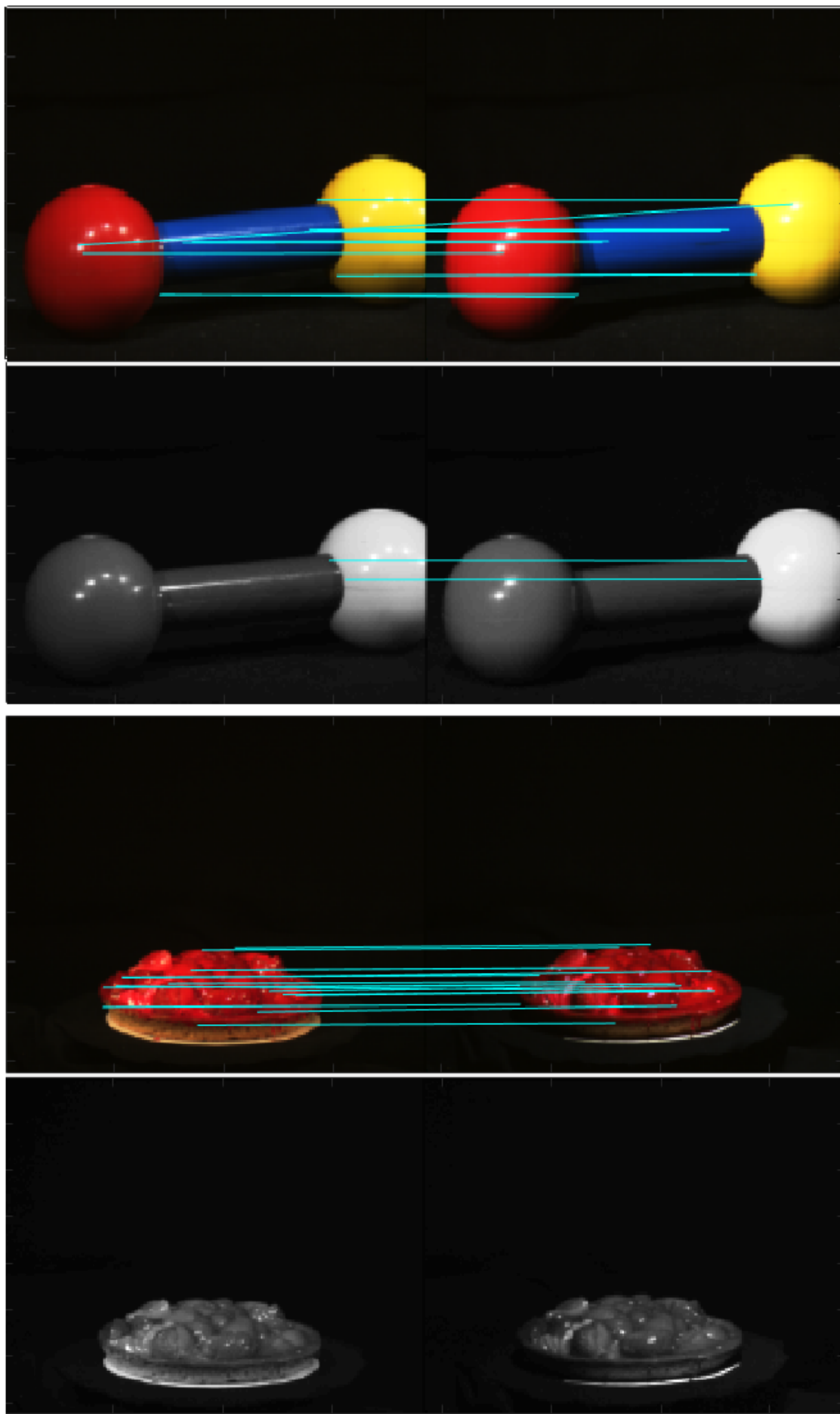
5.4 Budúca práca

V navrhutej metóde na extrakciu lokálnych príznakov spracovávame chromatické oponentné kanály. Naším cieľom bolo navrhnuť metódu, ktorá bude pracovať na farebnom obraze. Z tohto dôvodu neprebíha detekcia a deskripcia na šedotónovej časti obrazu, na ktorom pracuje pôvodná metóda SIFT. Ako sme ukázali v predchádzajúcich podkapitolách, výsledky metódy dosahujú lepšie hodnoty aj bez použitia achromatického čierno-bieleho kanálu. No intenzita obrazu obsahuje dôležité informácie o obraze, ktoré by mohli tieto dosiahnuté výsledky ešte vylepšiť a skvalitniť. Z tohto dôvodu by sme v budúcej práci chceli skúmať detekciu a deskripciu aj na zanedbávanom achromatickom oponentnom kanáli obrazu, spolu s informáciou z chromatických kanálov.

Pred príznakovým párovaním vypočítané deskriptory z oboch chromatických kanálov zret'azíme. To znamená, že ich spojíme do jednej množiny. Nič sa nezmení na tom, že deskriptory z červeno-zeleného kanála sú vypočítané pre nájdený kľúčový bod len z červeno-zeleného okolia. Rovnako to platí aj pri deskriptoroch kľúčových bodov nájdených zo žltomodrého kanála. V budúcnosti by sme pri výpočtoch deskriptorov nájdených kľúčových bodov chceli využiť informácie okolia z každého kanála. Tým by sme dostali deskriptor veľkosti 3×128 , ktorý by mal popísané okolie v rámci každého kanála.



Obr. 5.8: Výsledky I.



Obr. 5.9: Výsledky II.

Kapitola 6

Záver

Diplomová práca splnila požadovaný cieľ. V práci sa nám podarilo navrhnúť metódu, ktorá extrahuje lokálne príznaky z farebného obrazu. Na základe analýzy spracovania farebnej informácie vybranými existujúcimi metódami sme zvolili cestu ľudskeho vizuálneho systému. Farebný obraz sme spracovali na základe princípu vnímania farieb ľudským okom pomocou farebných receptorov - čapíkov, ktoré vieme takmer dokonale vyjadriť kombináciou trichromatického procesu a procesu oponentného kódovania farieb. Na detekciu zaujímavých bodov a výpočet deskriptorov sme použili metódu SIFT.

V analýze sme ukázali, že deskriptor spĺňa požadovanú vlastnosť - fotometrickú robustnosť, čo znamená, že je dostatočne invariantný na zmeny osvetlenia, tieňovanie. Oproti metóde SIFT je metóda výrazne vyššia v detekcii zaujímavých bodov, čím môžeme povedať, že dosahujeme lepšiu výkonnosť rozpoznávacieho procesu oproti metóde SIFT. V porovnaní so šedotónovým obrazom pri príznakovom párovaní použitá farebná informácia v metóde výrazne zlepšila výsledky a je dostatočne vhodná pre rozpoznávanie a klasifikáciu objektov z obrazu.

V práci sme sa hlavne zamerali na farebnú informáciu obsiahnutú v obraze, ktorú sme nekombinovali s informáciami o obraze obsiahnutými v čierno-bielom achromatickom kanáli. V budúcej práci by sme chceli metódu vylepšiť použitím informácií práve z tohto kanála, čím by sme zvýšili počet detegovaných zaujímavých bodov o počet detekcií práve v šedotónovej časti obrazu.

Literatúra

- [1] Alaa E Abdel-Hakim and Aly A Farag. CSIFT: A SIFT descriptor with color invariant characteristics. In *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on*, volume 2, pages 1978–1983. IEEE, 2006.
- [2] Cosmin Ancuti and Philippe Bekaert. SIFT-CCH: Increasing the SIFT distinctness by color co-occurrence histograms. In *Image and Signal Processing and Analysis, 2007. ISPA 2007. 5th International Symposium on*, pages 130–135. IEEE, 2007.
- [3] Matthew Anderson, Ricardo Motta, Srinivasan Chandrasekar, and Michael Stokes. Proposal for a standard default color space for the internet—srgb. In *Color and imaging conference*, volume 1996, pages 238–245. Society for Imaging Science and Technology, 1996.
- [4] Herbert Bay, Andreas Ess, Tinne Tuytelaars, and Luc Van Gool. Speeded-up robust features (SURF). *Computer vision and image understanding*, 110(3):346–359, 2008.
- [5] Anna Bosch, Andrew Zisserman, and Xavier Muoz. Scene classification using a hybrid generative/discriminative approach. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 30(4):712–727, 2008.
- [6] Gertjan J Burghouts and Jan-Mark Geusebroek. Performance evaluation of local colour invariants. *Computer Vision and Image Understanding*, 113(1):48–62, 2009.
- [7] Hamilton Y Chong, Steven J Gortler, and Todd Zickler. A perception-based color space for illumination-invariant image processing. In *ACM Transactions on Graphics (TOG)*, volume 27, page 61. ACM, 2008.
- [8] Yan Cui, Alain Pagani, and Didier Stricker. Sift in perception-based color space. In *Image Processing (ICIP), 2010 17th IEEE International Conference on*, pages 3909–3912. IEEE, 2010.
- [9] Stephen Engel, Xuemei Zhang, and Brian Wandell. Colour tuning in human visual cortex measured with functional magnetic resonance imaging. *Nature*, 388(6637):68–71, 1997.

- [10] Peng Fan, Aidong Men, Mengyang Chen, and Bo Yang. Color-SURF: A surf descriptor with local kernel color histograms. In *Network Infrastructure and Digital Content, 2009. IC-NIDC 2009. IEEE International Conference on*, pages 726–730. IEEE, 2009.
- [11] Jan-Mark Geusebroek, Gertjan J Burghouts, and Arnold WM Smeulders. The amsterdam library of object images. *International Journal of Computer Vision*, 61(1):103–112, 2005.
- [12] Jan-Mark Geusebroek, Rein Van den Boomgaard, Arnold WM Smeulders, and Hugo Geerts. Color invariance. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 23(12):1338–1350, 2001.
- [13] Robert M. Haralick and Linda G. Shapiro. *Computer and Robot Vision*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1st edition, 1992.
- [14] Leo M Hurvich and Dorothea Jameson. An opponent-process theory of color vision. *Psychological review*, 64(6p1):384, 1957.
- [15] Ali Jalilvand, Hamidreza Shayegh Boroujeni, and Masrollah Moghadam Charkari. CH-SIFT: A local kernel color histogram sift based descriptor. In *Multimedia Technology (ICMT), 2011 International Conference on*, pages 6269–6272. IEEE, 2011.
- [16] Ali Jalilvand, Hamidreza Shayegh Boroujeni, and Nasrollah Moghadam Charkari. CWSURF: A novel coloured local invariant descriptor based on SURF. In *Computer and Knowledge Engineering (ICCKE), 2011 1st International eConference on*, pages 214–219. IEEE, 2011.
- [17] Timothée Jost, Nabil Ouerhani, Roman Von Wartburg, René Müri, and Heinz Hügli. Assessing the contribution of color in visual attention. *Computer Vision and Image Understanding*, 100(1):107–123, 2005.
- [18] Scott Krig. *Computer Vision Metrics: Survey, Taxonomy, and Analysis*. Apress, 2014.
- [19] Paul Kubelka. New contributions to the optics of intensely light-scattering materials. part i. *JOSA*, 38(5):448–448, 1948.
- [20] Stefan Leutenegger, Margarita Chli, and Roland Y Siegwart. Brisk: Binary robust invariant scalable keypoints. In *Computer Vision (ICCV), 2011 IEEE International Conference on*, pages 2548–2555. IEEE, 2011.
- [21] Dandan Li, Yongzhen Ke, and Guiling Zhang. A SIFT descriptor with local kernel color histograms. In *Mechanic Automation and Control Engineering (MACE), 2011 Second International Conference on*, pages 992–995. IEEE, 2011.

- [22] Qing Li, Junzhou Chen, Qiang Peng, and Xiao Wu. 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, 2014.
- [23] David G Lowe. Object recognition from local scale-invariant features. In *Computer vision, 1999. The proceedings of the seventh IEEE international conference on*, volume 2, pages 1150–1157. Ieee, 1999.
- [24] David G Lowe. Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60(2):91–110, 2004.
- [25] Jiri Matas, Ondrej Chum, Martin Urban, and Tomas Pajdla. Robust wide-baseline stereo from maximally stable extremal regions. *Image and vision computing*, 22(10):761–767, 2004.
- [26] Krystian Mikolajczyk, Tinne Tuytelaars, Cordelia Schmid, Andrew Zisserman, Jiri Matas, Frederik Schaffalitzky, Timor Kadir, and Luc Van Gool. A comparison of affine region detectors. *International journal of computer vision*, 65(1-2):43–72, 2005.
- [27] Philippe Noriega, Benedicte Bascle, and Olivier Bernier. Local kernel color histograms for background subtraction. In *VISAPP (1)*, pages 213–219, 2006.
- [28] CJ Prabhakar and PU Kumar. LBP-SURF descriptor with color invariant and texture based features for underwater images. In *Proceedings of the eighth Indian conference on computer vision, graphics and image processing*, page 23. ACM, 2012.
- [29] Stephen M Smith and J Michael Brady. Susan—a new approach to low level image processing. *International journal of computer vision*, 23(1):45–78, 1997.
- [30] Jay E Thornton and Edward N Pugh. Red/green color opponency at detection threshold. *Science*, 219(4581):191–193, 1983.
- [31] Tinne Tuytelaars and Krystian Mikolajczyk. Local invariant feature detectors: a survey. *Foundations and Trends® in Computer Graphics and Vision*, 3(3):177–280, 2008.
- [32] Tinne Tuytelaars and Luc J Van Gool. Wide baseline stereo matching based on local, affinely invariant regions. In *BMVC*, volume 1, page 4, 2000.
- [33] Koen EA Van De Sande, Theo Gevers, and Cees GM Snoek. Evaluating color descriptors for object and scene recognition. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 32(9):1582–1596, 2010.

- [34] Joost Van De Weijer and Cordelia Schmid. Coloring local feature extraction. In *Computer Vision–ECCV 2006*, pages 334–348. Springer, 2006.
- [35] Černeková Z. a kol. *Počítačové videnie. Detekcia a rozpoznávanie objektov*. Wikina, Livornská 445, 109 00 Praha 10, 2014.

Prílohy

CD obsahujúce:

- Elektronická verzia diplomovej práce
- Zdrojové kódy navrhutej metódy
- Databáza obrázkov