

VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA INFORMAČNÍCH TECHNOLOGIÍ
ÚSTAV POČÍTAČOVÉ GRAFIKY A MULTIMÉDIÍ

FACULTY OF INFORMATION TECHNOLOGY
DEPARTMENT OF COMPUTER GRAPHICS AND MULTIMEDIA

DETEKCE ODPOVÍDAJÍCÍCH SI BODŮ VE DVOU
FOTOGRAFIÍCH

BAKALÁŘSKÁ PRÁCE
BACHELOR'S THESIS

AUTOR PRÁCE
AUTHOR

MARTIN PTAŠEK

BRNO 2008



VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
BRNO UNIVERSITY OF TECHNOLOGY



FAKULTA INFORMAČNÍCH TECHNOLOGIÍ
ÚSTAV POČÍTAČOVÉ GRAFIKY A MULTIMÉDIÍ
FACULTY OF INFORMATION TECHNOLOGY
DEPARTMENT OF COMPUTER GRAPHICS AND MULTIMEDIA

DETEKCE ODPOVÍDAJÍCÍCH SI BODŮ VE DVOU FOTOGRAFIÍCH

DETECTION OF CORRESPONDING POINTS IN IMAGES

BAKALÁŘSKÁ PRÁCE
BACHELOR'S THESIS

AUTOR PRÁCE
AUTHOR

MARTIN PTAŠEK

VEDOUCÍ PRÁCE
SUPERVISOR

ING. MICHAL ŠPANĚL

BRNO 2008

Detekce odpovídajících si bodů ve dvou fotografiích

Detection of Corresponding Points in Images

Vedoucí:

Španěl Michal, Ing., UPGM FIT VUT

Oponent:

Beran Vítězslav, Ing., UPGM FIT VUT

Zadání:

1. Prostudujte metody zpracování obrazu. Zaměřte se na detekci významných obrazových bodů.
2. Seznamte se s algoritmy pro vyhodnocení korespondence významných bodů ve dvou fotografiích stejného objektu.
3. Navrhněte detektor významných obrazových bodů a postup jejich sesouhlasení.
4. Experimentujte s vaší implementací.
5. Případně navrhněte vlastní modifikace metod.
6. Diskutujte dosažené výsledky a možnosti budoucího vývoje.
7. Vytvořte stručný plakát prezentující vaši bakalářskou práci, její cíle a výsledky.

Kategorie:

Počítačová grafika

Implementační jazyk:

C/C++, C# nebo Python

Operační systém:

MS Windows, Linux (pokud možno přenositelný kód)

Literatura:

- Dle pokynů vedoucího.

Licenční smlouva

Licenční smlouva je uložena v archivu Fakulty informačních technologií Vysokého učení technického v Brně.

Abstrakt

Bakalářská práce se zabývá metodami detekce shodných bodů ve dvou digitálních obrazech. Shrnuje a popisuje metody, které jsou pro téma detekce významných bodů v obraze základní nebo nové a moderní. Popisuje také metody používané k určování korespondencí částí obrazů. Práce je zaměřena na návrh a popis implementace detektoru shodných bodů s důrazem na segmentaci obrazu metodou MSER. Aplikace je vytvořena v jazyce C s využitím grafické knihovny OpenCV v operačním systému Windows XP. Výsledky, které shrnují vlastnosti detektoru, jsou ukázkou možného potenciálu vybraných metod.

Klíčová slova

MSER, obraz, významný bod, detekce, korespondence

Abstract

Bachelor thesis is concerned with detection of corresponding points in two digital images. It summarises and describes basic, new and modern methods of detection of interest points. It also describes methods for determining equivalence of image parts. Work itself includes design of application and describes implementation of corresponding points detector based on segmentation method MSER. Application was written in C programming language and uses Open CV graphic library under Windows XP operating system. Results that summarise characteristics of detector illustrate potential of chosen methods.

Keywords

MSER, image, interest point, detection, equivalence

Citace

Ptašek Martin: Detekce odpovídajících si bodů ve dvou fotografiích. Brno, 2, bakalářská práce, FIT VUT v Brně.

Detekce odpovídajících si bodů ve dvou fotografiích

Prohlášení

Prohlašuji, že jsem tuto bakalářskou práci vypracoval samostatně pod vedením Ing. Michala Španěla. Uvedl jsem všechny literární prameny a publikace, ze kterých jsem čerpal.

.....
Martin Ptašek
8. ledna 2008

Poděkování

Děkuji vedoucímu bakalářské práce Ing. Michalu Španělovi za vstřícnost, konstruktivní kritiku a odborné vedení při zpracování této práce.

© Martin Ptašek, 2008.

Tato práce vznikla jako školní dílo na Vysokém učení technickém v Brně, Fakultě informačních technologií. Práce je chráněna autorským zákonem a její užití bez udělení oprávnění autorem je nezákonné, s výjimkou zákonem definovaných případů..

Obsah

Úvod	2
1 Vymezení základních pojmů.....	3
1.1 Rastrový obraz.....	3
1.2 Významný bod	4
1.3 Robustnost metod.....	5
1.4 Korespondence	6
2 Detekce významných bodů	7
2.1 Detekce rohů	8
2.2 Extrakce příznaků.....	11
2.3 Maximálně stabilní extrémní regiony (MSER)	13
3 Korespondence významných bodů	17
3.1 Minimální vzdálenost vektorů.....	17
3.2 RANSAC.....	18
4 Návrh detektoru	20
5 Implementace detektoru.....	22
5.1 Knihovna OpenCV	22
5.2 Parametry detektoru	22
5.3 Segmentace obrazu.....	22
5.4 Korespondence oblastí	26
6 Dosažené výsledky.....	29
6.1 Vhodná vstupní data.....	29
6.2 Návrh modifikací a rozšíření.....	34
7 Závěr	35
Literatura	36
Seznam zkratk.....	38
Seznam příloh	39
Příloha č.1 - Tabulky hodnot.....	40
Příloha č.2 - Výstupní soubor.....	42

Úvod

Schopnost mozku zpracovávat obrazové informace je pro každého člověka nedocenitelná. S rozvojem počítačové techniky, robotiky a automatizace jsou stále častěji vyžadovány a uplatňovány prostředky, které dokáží, mozku vlastní schopnosti, v mnoha ohledech napodobit. Otázkou jak v obraze identifikovat objekty, rozlišit je od sebe, určit jejich souvislosti, nalézt důležité a specifické oblasti či mnoho dalších otázek ohledně informací, které nám obraz poskytuje, se zabývá mnoho odvětví informatiky již od svého počátku.

V roce 1955 vyslovil Oliver G. Selfridge, člověk stojící u základů umělé inteligence, myšlenku, aby obor počítačového vidění poskytl počítačům oči a uši.

Pro představu, v kolika oborech se na tuto myšlenku navazuje[1]:

- Průmysl - kontrola kvality, geometrická měření
- Robotika - navigace, kooperace, autonomní systémy
- Monitorování - bezpečnostní systémy, inteligentní domácnosti
- Geografie - geo-informační systémy, ekologie, obrana
- Analýza dokumentů - rozpoznávání psaného textu
- Medicína - srovnávání obrazů, modelování tkání
- Virtuální realita - generování obrazu, konstrukce modelu

Ve všech oborech, kde se zpracovává obrazová informace, však musí lidé řešit několik základních problémů. Jak vytvořit robustní metody, které by poskytly vhodná data, jak tyto metody zrychlit při zachování účinnosti a v neposlední řadě, jak zaručit, že jsou dané výsledky správné.

Tato práce se bude zabývat metodami pro detekci významných obrazových bodů či oblastí ve dvou fotografiích a jejich následné sesouhlasení určující vzájemné korespondence. Jejím výsledkem je tento dokument a program implementující detektor shodných bodů na základě vybraných metod.

Vybrané metody, prostudované a popsány v tomto dokumentu jsou v praxi použitelné například, pro spojování navazujících obrazů, tedy tvorbě panoramatických snímků, jsou také používány pro sledování pohybu objektu nebo pro rekonstrukci prostoru.

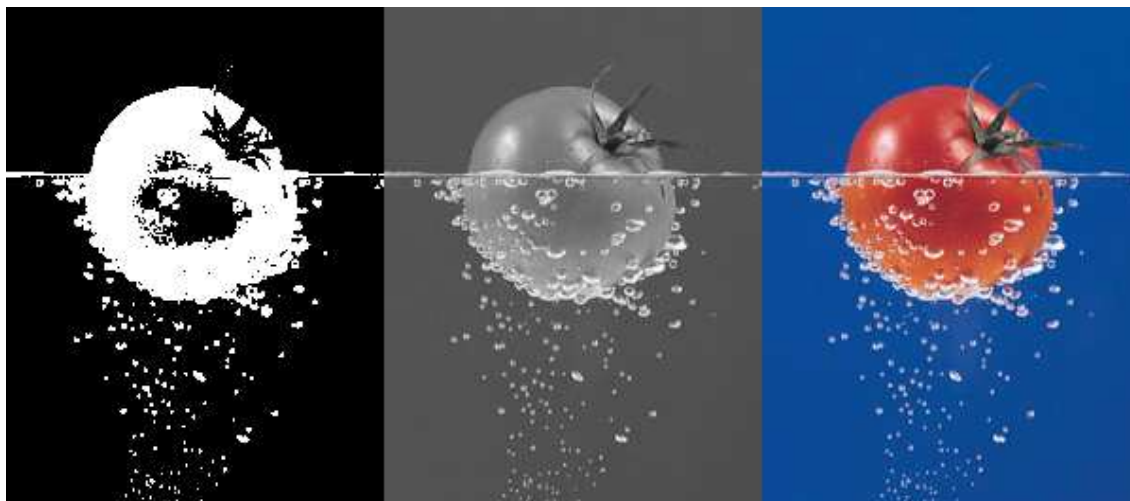
Samotný dokument obsahuje devět kapitol. V kapitole 2 jsou stručně vyloženy základní pojmy k danému tématu. Kapitola 3 je rychlým vhladem a vysvětlením základních metod používaných nebo použitelných v praxi pro detekci významných bodů. Základní možné postupy pro řešení problému korespondence jsou v následující kapitole 4. Kapitola 5 obsahuje návrh detektoru. V kapitole 6 jsou popsány a vysvětleny algoritmy, prostředky a postupy použité k implementaci detektoru. Kapitola 7 je zaměřena na vyhodnocení výsledků, které detektor poskytuje. V kapitole 8 jsou stručně popsány možnosti rozšíření, modifikací a možným vylepšením detektoru do budoucna. Závěr této práce je shrnutím nabytých informací a jejich přínos.

1 Vymezení základních pojmů

1.1 Rastrový obraz

Rastrový obraz (dále jen obraz) je jedním ze způsobů jak můžeme skutečný obraz digitálně zachytit a reprezentovat. Jde o pravidelnou síť nejmenších prvků, neboli pixelů, uspořádaných jako dvourozměrná matice bodů. Každý pixel nese specifické informace, například o jas, barvě, průhlednosti bodu, nebo kombinaci těchto hodnot. Základní dělení obrazu podle počtu bitů použitých pro informaci o barvě je následující:

- **monochromatický:** obsahuje pouze barvu ve stavu zapnuto nebo vypnuto (černo-bílý obraz)
- **odstíny šedé (grayscale):** barva uložena na osmi bitech
- **barevný:** podle zvoleného modelu se mohou nároky lišit, ale nejčastěji se používá osm bitů pro zakódování každé ze tří barevných složek (RGB, CMY, HSV, HLS)



[obr. 1.1] Příklad obrazu s různou informací o barvě. Zleva, monochromatický obraz, obraz v odstínech šedé, barevný obraz.

Pro potřeby této práce však postačuje barevný model s odstíny šedi, tedy 256 hodnot. Jas je v případě barevného RGB obrazu přepočítáván podle vztahu:

$$Y = 0.3 * R + 0.59 * G + 0.11 * B \quad (1.1)$$

Jde o nelineární kódování vztahené ke schopnosti lidského oka vnímat barvy.

Relativní snadnost zpracování takto reprezentovaných informací je vykoupena určitými nedostatky. Jde především o paměťové nároky na uchování obrazu, neboť pixel již není možné dále dělit a detaily je tak nutné reprezentovat na jejich větším počtu. Velikost obrazu navíc zásadním

způsobem ovlivňuje dobu potřebnou k jakémukoli zpracování. V dnešní době však technika dovoluje velké množství metod použít tzv. real-time, neboli za běhu (nejméně 25 snímků za sekundu).

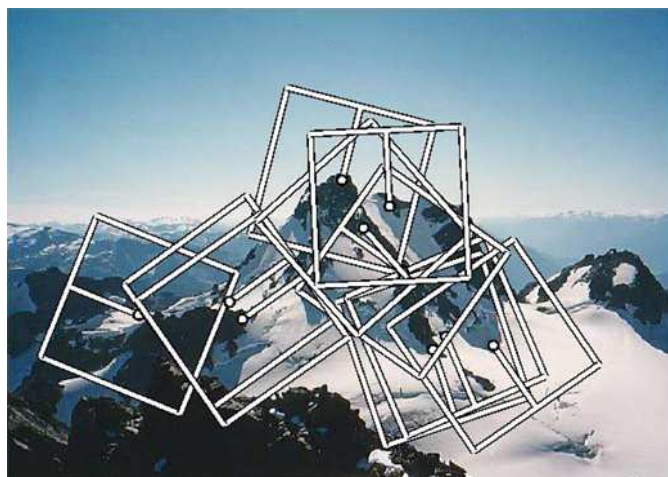
1.2 Významný bod

Významný bod je v obraze charakterizován následovně:

- je, pokud možno, matematicky definován
- má dobře definovanou polohu v prostoru
- okolí bodu obsahuje velké množství informací v kontextu celého obrazu
- je stabilní vůči lokálním i globálním změnám, které zahrnují geometrické transformace (posun, rotace, zvětšení, apod.) i změny barevného prostoru obrazu (světlost, kontrast)



[a]



[b]

[obr. 1.2] Ukázka významných bodů v obraze [a] originální obrázek [b] obrázek s vyznačenými významnými body. Čtverce reprezentují některé vlastnosti detekovaných významných bodů (pozici, orientaci, měřítko). Tento popis se liší podle zvolené metody. [15]

1.3 Robustnost metod

Při kladení požadavku na robustnost metody jsou obvykle vyžadovány takové vlastnosti, aby byly výsledky správné i v případě, že vstupní data nebudou, v určitých ohledech, zcela ideální. V případě práce s fotografiemi a detekcí významných bodů je důležitá robustnost vůči několika základním faktorům. Především robustnost vůči šumu, geometrickým změnám a změně barevného prostoru.

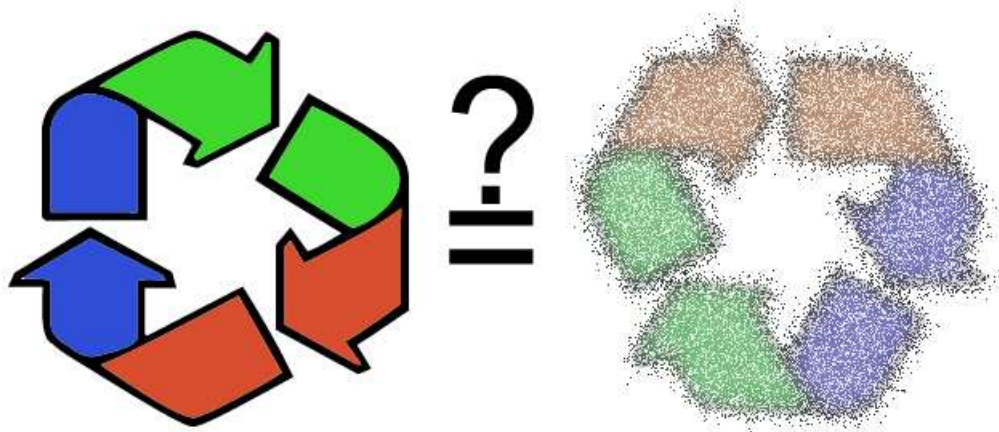
Šum v obraze je chápán jako data bez významu, tedy data škodlivá, která nejsou použitelná pro relevantní zpracování a naopak poškozují původní „čistá“ data. Jedná se většinou o barevné skvrny (body či plochy) na snímku, které snižují ostrost a detaily v obraze.

Příčin pro vznik šumu v obraze existuje poměrně velké množství. Může jít o náhodné uvolňování či přeskokování elektronů přímo na snímači zařízení pořizujícího obraz, jehož důvodem je vnější elektronické rušení, teplotní nestálost křemíku a případně nekvalitní snímač jako takový. Výrazným činitelem je kvalita a přenos analog-digitálního převodníku, který zpracovává výstupní data ze samotného snímače. V AD převodníku dochází i k tzv. zvýšení citlivosti snímače. Proto i kvalita AD převodníku určuje kvalitu výsledné obrazové informace. Dále může jít o granularitu filmového materiálu, relativní pohyb mezi objektem a snímačem, špatné zaostření, zpracování nebo zaokrouhlování hodnot digitální fotografie obrazovým procesorem.

Geometrické zkreslení vzniká především díky perspektivě reálného světa nebo nepřesnosti snímačového zařízení (např. geometrické zkreslení družicového snímku vlivem zakřivení země).

Zkreslení barevného prostoru je možné chápat jako změnu jasové informace obrazu změnou barvy či intenzity osvětlení. Důležitou roli také hraje světlost a kontrast.

U všech těchto faktorů existují meze použitelnosti, při jejichž překročení již není možné očekávat spolehlivý výstup.



[obr. 1.3] Nástin problému robustnosti

1.4 Korespondence

Objekt ve fotografiích, které pořídíme například z různého úhlu nebo při jiném osvětlení je v mozku většinou snadno identifikován jako tentýž. Samotná data však mohou ukazovat na obrazy zcela rozdílné a právě korespondence určuje vzájemnou souhlasnost některých jejich částí.

Ačkoliv korespondence obvykle určuje pouze vztah několika částí, v této práci je chápána jako ukazatel totožnosti právě dvou objektů si navzájem.



[obr. 1.4] Naznačení korespondence objektu ve dvou obrazech

2 Detekce významných bodů

V současné době zaznamenává obor zpracování obrazu a počítačového vidění obrovský zájem. Využití a vývoj níže popsaných metod jde ruku v ruce se současným rychlým vývojem civilizace. Zatím co, v historii můžeme vidět využití počítačového vidění především ve vědecké, akademické či průmyslové sféře, v současnosti se již produkty a metody stojící na tomto oboru dostávají do obecného využití a zájmu širší veřejnosti. Značné uplatnění v praxi je v současnosti také neoddiskutovatelné.

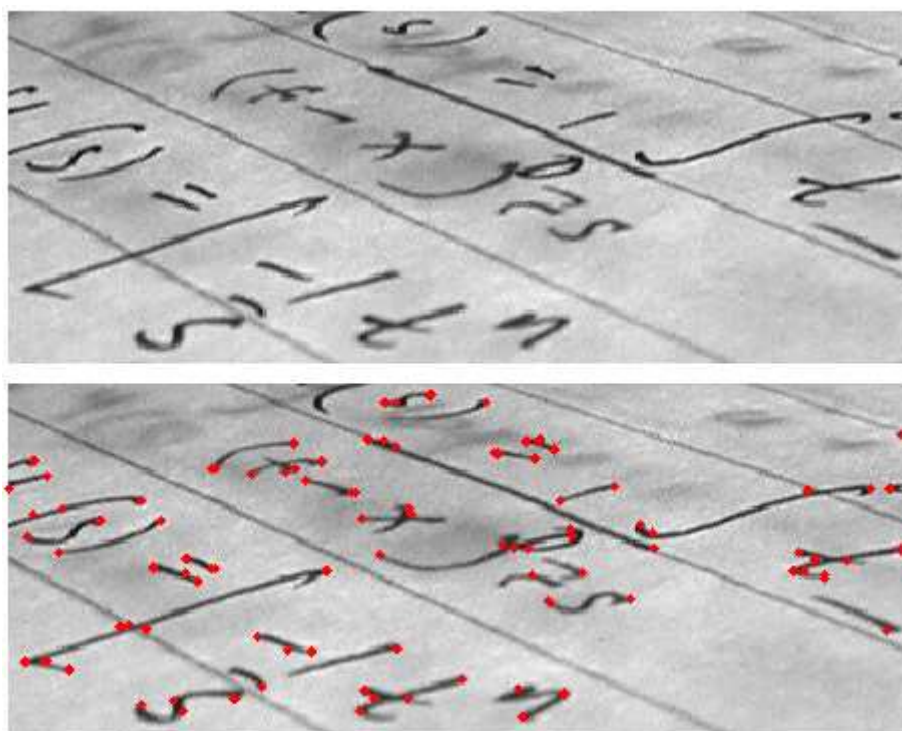
Systémy pro nalezení obličeje nebo srovnání otisků prstů, založené na metodách hledající významné body, jsou již v policejní praxi nebo při využití bezpečnostních systémů, téměř nepostradatelné. Ve vzorku široké veřejnosti by jsme také snadno našli mnoho lidí, kteří běžně používají software schopný spojovat fotografie, na základě korespondence významných bodů, do panoramatických snímků. Pro představu, je tento software možné nalézt i v digitálních fotoaparátech, které jsou součástí mobilních telefonů.

V poslední době se můžeme také setkat s implementací některých níže zmíněných metod pro sledování určených oblastí ve videu, což v daných systémech umožňuje například, velmi pohodlně použití speciálních digitálních efektů, které by bylo dříve v domácích podmínkách nemyslitelné.

Metody detekce významných bodů můžeme rozdělit do několika skupin, podle způsobu samotné detekce a výsledného výstupu, který nám mohou poskytnout. Do první skupiny patří metody zaměřené na získání jednotlivých významných bodů v obraze, jimiž jsou rohy. Následují metody, které zkoumaný bod transformují na měřítkově nezávislou množinu příznaků. Jako poslední způsob detekce významných bodů, je představena metoda hledající maximálně stabilní extrémní regiony, která patří do specifické skupiny blob detekce.

2.1 Detekce rohů

Roh můžeme definovat jako průsečík dvou hran. Dále může být definován jako bod, který má dva hlavní a odlišné směry hrany v blízkém okolí bodu. S přihlédnutím k vymezení významného bodu v kapitole 1.2, může být roh také konec čáry, nebo bod křivky tvořící její lokální maximum zakřivení.



[obr. 2.1] Typický výstup rohových detektorů [12]

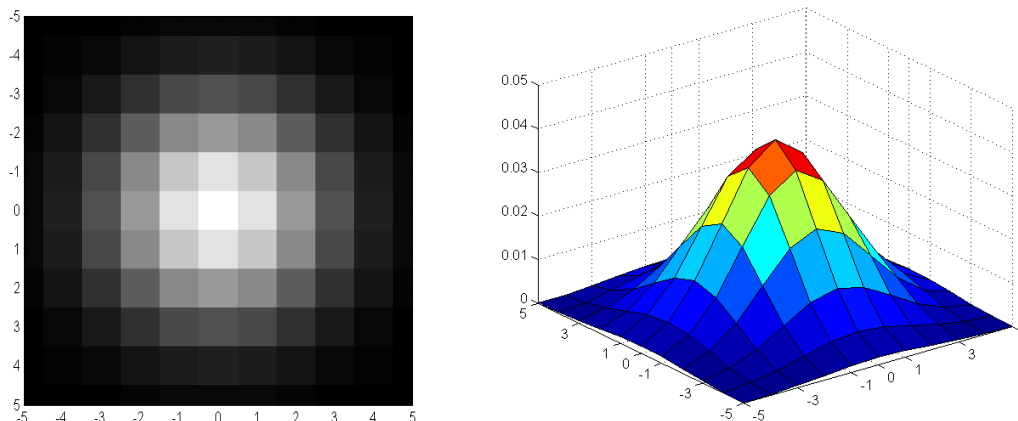
Harris & Stephens / Plessey rohový detektor

Harris Stephens algoritmus se stal od roku 1988 základní metou pro zvládnutí rohové detekce. Navazuje na Moravec algoritmus z roku 1977 a pracuje na principu posunu zkoumané části obrazu v různých směrech a výpočtu gradientu v jednotlivých bodech. Základem metody je výpočet rozdílů čtverců jasů ve všech bodech snímku. Princip funguje pouze tehdy, jestliže porovnávané části mají stejný jas a kontrast. Nedostatkem algoritmu Harris Stephens je tedy citlivost na změny jasu a kontrastu.

Z tohoto algoritmu vychází množství novodobých algoritmů. Jedná se o metody Shi & Tomasi či více stupňový Harrisův detektor.

Wang Brady detektor

Zkoumá zakřivení hran a roh detekuje při náhlé změně směru přímky. U této metody je nezbytné použít Gaussův filtr pro snížení šumu. Jedná se o rozostření obrazu, získaného jako vážený průměr okolí podle konvolučního jádra,



[obr. 2.2] Ukázka konvolučního jádra Gaussovy funkce

které odpovídá Gaussově funkci definované vztahem:

$$G(x, y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \quad (2.1)$$

kde σ je parametr určující velikost jádra.

Vyhlazení pomocí Gaussova filtru ovšem způsobí špatné rozmístění detekovaných rohů. Algoritmus Wang Brady tento problém řeší aplikováním speciálního výrazu na detekované body, kdy po aplikování tohoto výrazu dojde k umístění rohů na odpovídající místa. Mezi nevýhody tohoto algoritmu patří špatné určování významných bodů při velmi malých a velmi velkých úhlech.

SUSAN rohový detektor

SUSAN (Smallest Univalued Segment Assimilating Nucleus) [10] rohový detektor pracuje s algoritmem z roku 1995, který porovnává hodnotu zkoumaného bodu, takzvaného základu (nukleus) s jeho kruhovým okolím, neboli maskou, určené velikosti. Každý pixel obrazu je porovnáván se základem podle funkce:

$$c(\vec{m}) = e^{\frac{(I(\vec{m}) - I(\vec{m}_0))^6}{t}} \quad (2.2)$$

kde \vec{m} jsou pixely masky, \vec{m}_0 tvoří základ, t určuje rozsah podobnosti a exponent byl určen empiricky. USAN je oblast, jejíž všechny body mají podobnou jasovou hodnotu jako zkoumaný základ (jádro oblasti).

Oblast USAN je popsána vzorcem:

$$n(M) = \sum_{\vec{m} \in M} c(\vec{m}) \quad (2.3)$$

V případě, že c je rozhodovací funkce, pak n je počet pixelů v masce, jejichž hodnota je v mezích t .

Odpověď operátoru USAN je dána následovně:

$$R(M) = \begin{cases} g - n(M) & \text{pokud } n(M) < g \\ 0 & \text{jindy} \end{cases} \quad (2.4)$$

Operátor USAN nabývá kladných pouze v případě dostatečně malých oblastí. Po odstranění nemaximálních hodnot získáme operátor SUSAN. K samotnému nalezení rohových bodů je potřeba vyhledání těžišť USAN oblastí daleko od jádra. Posledním nezbytným krokem je ujištění se, že všechny body na spojnici od jádra přes těžiště na okraj oblasti masky náležejí do USAN. Tato technika je chráněná ve Velké Británii patentem číslo 2272285.

2.2 Extrakce příznaků

SIFT

SIFT (Scale-invariant feature transform) [3] je algoritmus pro extrakci výrazných rysů nebo také příznaků obrazu. Příznaky získané touto metodou jsou neměnné při změně velikosti obrazu či jeho otočení. Příznaky jsou také částečně neměnné (robustní) vůči změně úhlu pohledu a změně osvětlení. Jméno této metody (Scale-invariant feature transform) bylo zvoleno podle algoritmu, který převádí obrazová data na souřadnice velikostně neměnné a při tom, data relativní vůči lokálním rysům obrazu.

Reprezentace obrazu pomocí SIFT byla vytvořena jako analogie funkce nervových buněk spodního spánkového laloku u biologického řádu Primáti, do kterého patří i člověk.

Tento algoritmus poprvé publikoval David Lowe v roce 1999 a současným držitelem patentu je University of British Columbia.

Vlastní algoritmus [9] je možné zjednodušeně popsat ve čtyřech následujících krocích:

- Je potřeba vytvořit měřítkově nezávislou reprezentaci zkoumaných obrazů. K tomu je potřeba použít sérii Gaussových filtrů. Srovnáním vzniklých obrazů se naleznou extrémy, jenž detekují shodné měřítko obrazů. Extrémy jsou popsány vektorem, který určuje měřítko, orientaci (natočení) a umístění.
- Lokalizace klíčových bodů z nalezených extrémů podle míry jejich stability. Prahováním (thresholding) se odstraní body s nízkou stabilitou a pomocí bilineární interpolace se upřesní jejich umístění v prostoru obrazu.
- Přidělení orientace každému z klíčových bodů. Orientace je počítána v okolí 16x16 a určuje směr dominantního gradientu. Všechny následující operace jsou prováděny nad obrazovými daty, která jsou relativně transformována k přidělené orientaci, měřítku a umístění. To zajišťuje neměnnost vůči právě těmto transformacím.
- Generování deskriptorů bodů. Tímto deskriptorem je pro každý klíčový bod 128-rozměrný vektor vypočtený z gradientů okolí ve vybraném měřítku. Tato reprezentace je důležitá z hlediska stability metody v případě lokální deformace objektu a fotometrických změn.

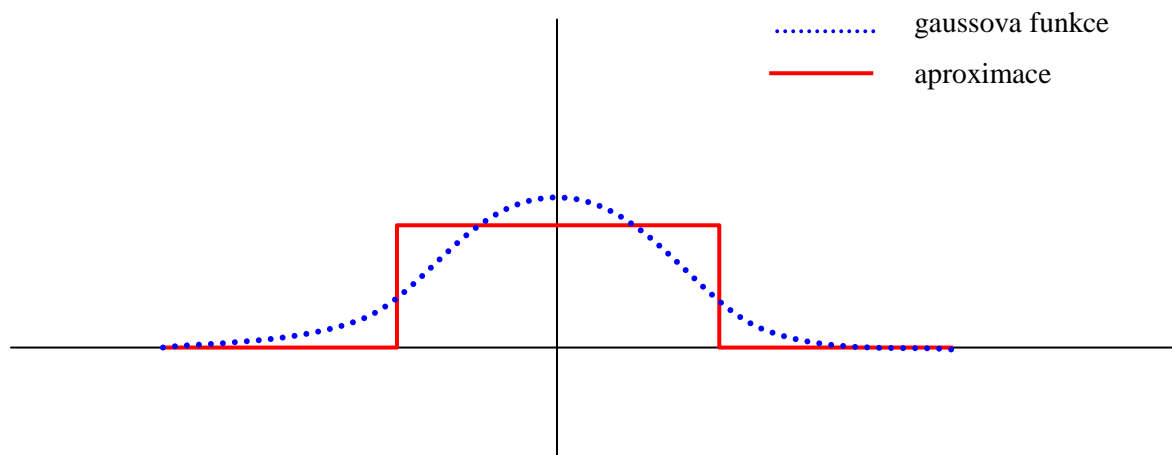
První dva body algoritmu pro výpočet příznaků obrazu jsou shodné s metodami blob detekce s automatickým výběrem měřítka a jsou odvozené z návrhu Lindeberga z roku 1998. Pro aplikace SIFT, používané při spojování a rozpoznávání objektů, Lowe používal algoritmus nejbližšího souseda, následovaný Houghovou transformací pro rozpoznávání objektů.

Detekce metodou SIFT je velice robustní, ale díky snaze o přesnou a co možná nejstabilnější lokalizaci významných bodů není z důvodů výpočetních nároků v praxi využívána. Přesto však tvoří dobrý základ pro vznik metod nových.

SURF

SURF (Speeded Up Robust Features) [8] je nová metoda, která hledá robustní příznaky v obraze. Na rozdíl od SIFT detekce, na jejímž základě stojí, se nesnaží o naprostou přesnost vyhledaných příznaků. Bylo by ovšem chybou myslet si, že tato metoda není přesná. V SURF metodě se pouze autoři snažili aproximovat složité a pomalé prostředky SIFT, jednoduššími a rychlejšími, při zachování původní robustnosti a účinnosti. Dle jejich vyjádření v roce 2006, kdy tuto metodu poprvé publikovali, se jim to povedlo. Je to zároveň důkaz o tom jak dynamické, odvětví počítačového vidění, je.

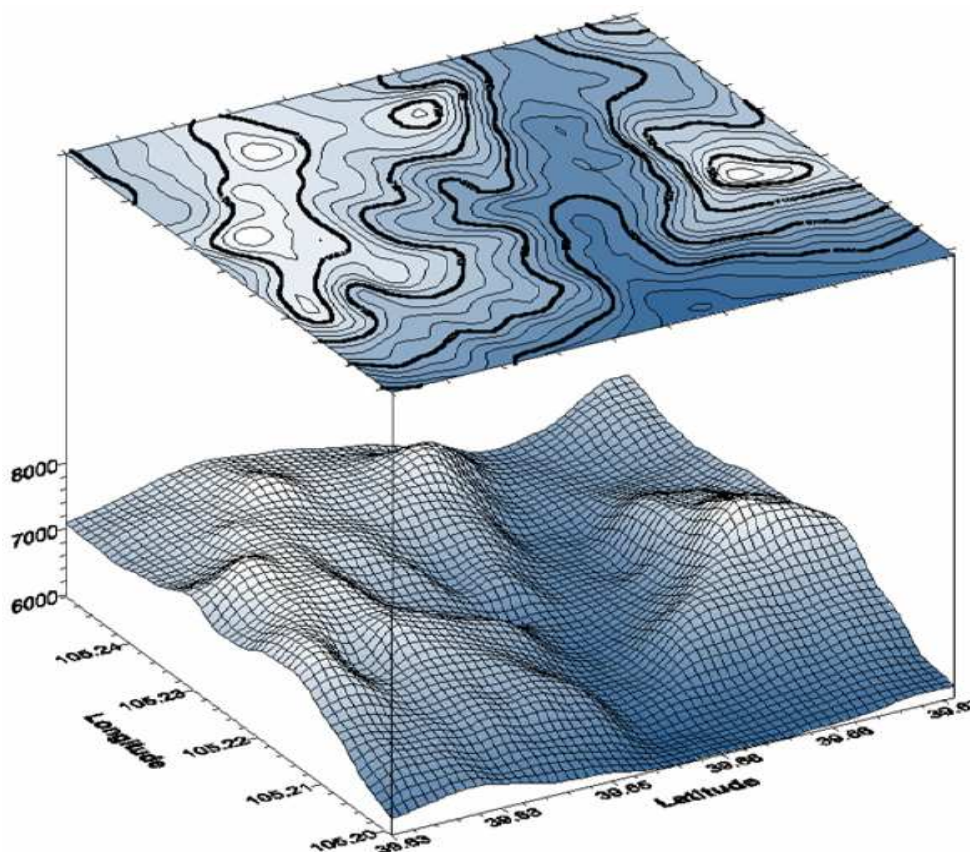
Samotná metoda Speeded Up Robust Features sestává ze shodných kroků jako metoda SIFT, s rozdílem v použitých prostředcích. Mezi hlavní rozdíly, které přispívají k výraznému zrychlení, je nahrazení konvolučního jádra podle derivace gaussovy funkce jádry založenými na funkci obdélníkové. Další zrychlení přineslo určení orientace pouze podle dominantního gradientu a zanedbání všech ostatních. Téměř poslední změnou oproti SIFT je jednodušší generování deskriptorů, které ve výsledku obsahují poloviční množství dat. Následné srovnávání je pak mnohem rychlejší, neboť je potřeba menšího množství operací.



[obr. 2.3] Ukázka aproximace gaussovy funkce.

2.3 Maximálně stabilní extrémní regiony (MSER)

Metoda maximálně stabilních extrémních regionů (MSER) spadá do skupiny blob detektorů. Jde o detektory významných bodů nebo regionů v obraze, které jsou tmavší nebo naopak světlejší než jejich okolí. Jde o algoritmy, které původně vznikaly pro potřeby geografie avšak obraz je také možné chápat jako výškový reliéf krajiny, čehož se využívá v počítačovém vidění.



[obr. 2.4] Digitální model terénu na základě obrazu.

Blob detektory je možné rozdělit do dvou hlavních skupin, které se liší metodami použitými k určování potřebného kritéria hledaného prvku a okolí.

První skupina metod je založena na první derivaci. Porovnává vzájemné intenzity bodů a určujícím faktorem je rozdíl.

Další skupina metod je založena na určování lokálních extrémů intenzit, tedy na hledání lokálních maxim a minim. Matematicky jde o druhou derivaci.

Důvodů pro používání a dalšímu výzkumu blob detektorů existuje hned několik. Zejména jde o poskytnutí doplňujících informací, které nejsou získány detektory jiného typu, detektory hran a rohů.

Popis metody

Metodu MSER představil v roce 2002 Jiří Matas jako novou, rychlou metodu řešící korespondenci dvou obrazů pořízených z rozdílného úhlu. Základní myšlenka [1] je popsána následujícím způsobem:

Obraz ve stupních šedi prahujeme všemi prahy. Hodnotu prahu zvyšujeme resp. snižujeme od nejnižší resp. nejvyšší hodnoty. Představíme-li si, že je obraz prahován nejvyšším prahem, je výsledný obraz zcela černý. Když následně hodnotu prahu snižujeme, objevují se body nejvyšší intenzity původního obrazu jako bílé body, které se postupně zvětšují. Tyto jednotlivé bílé části tvoří samostatné regiony. V místech, kde dojde ke spojení regionů, leží lokální extrém, lokální minimum. Na konci máme obraz zcela bílý a sadu všech spojených regionů, která tvoří nelineární, stromový seznam všech maximálních regionů.

Z předchozího popisu vyplývá následující definice:

- Vnější okraj regionu

$$\partial Q = \{q \in D \setminus Q : \exists p \in Q : qAp\} \quad (2.5)$$

kde okraj ∂Q regionu Q je sada pixelů příbuzných nejméně k jednomu pixelu Q , ale samotná do Q nenáleží, D je celý obraz.

- Extrémní region Q je takový, pro který platí

$$p \in Q, q \in \partial Q : I(p) > I(q) \text{ nebo } I(p) < I(q) \quad (2.6)$$

kde I je vstupní obraz.

- Maximálně Stabilní Extrémní Region (MSER) je Q_i pouze když

$$q(i) = \frac{|Q_{g+\Delta} \setminus Q_{g-\Delta}|}{|Q_g|} \quad (2.7)$$

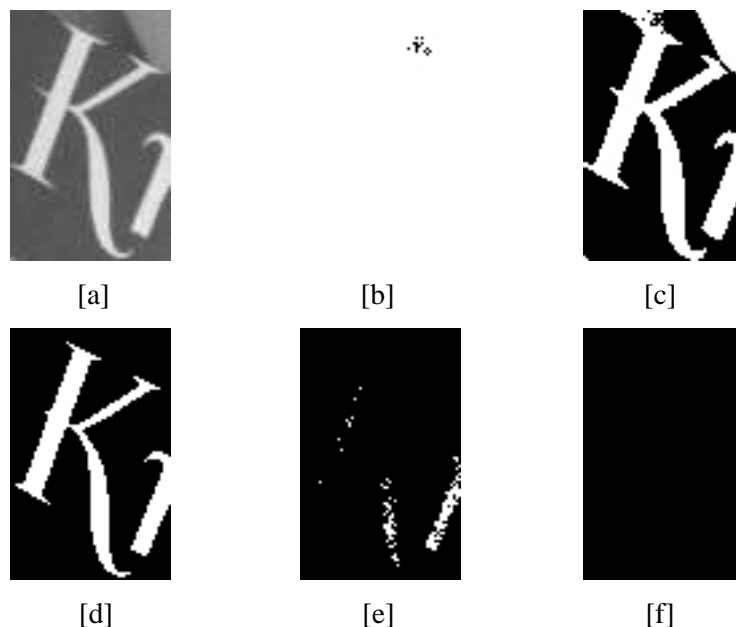
má lokální minimum. $|\cdot|$ určuje kardinalitu regionů a Δ je parametrem metody, g je hodnota prahu.

Předchozí definice je založena na zkoumání nárůstu okraje regionu při prahování, přičemž hledáme regiony, jejichž okraj se mění nejméně při co nejvyšším počtu prahů. Jde však o původní definici Matase a v dnešní době již existují, pro maximální stabilitu regionu, i jiná řešení:

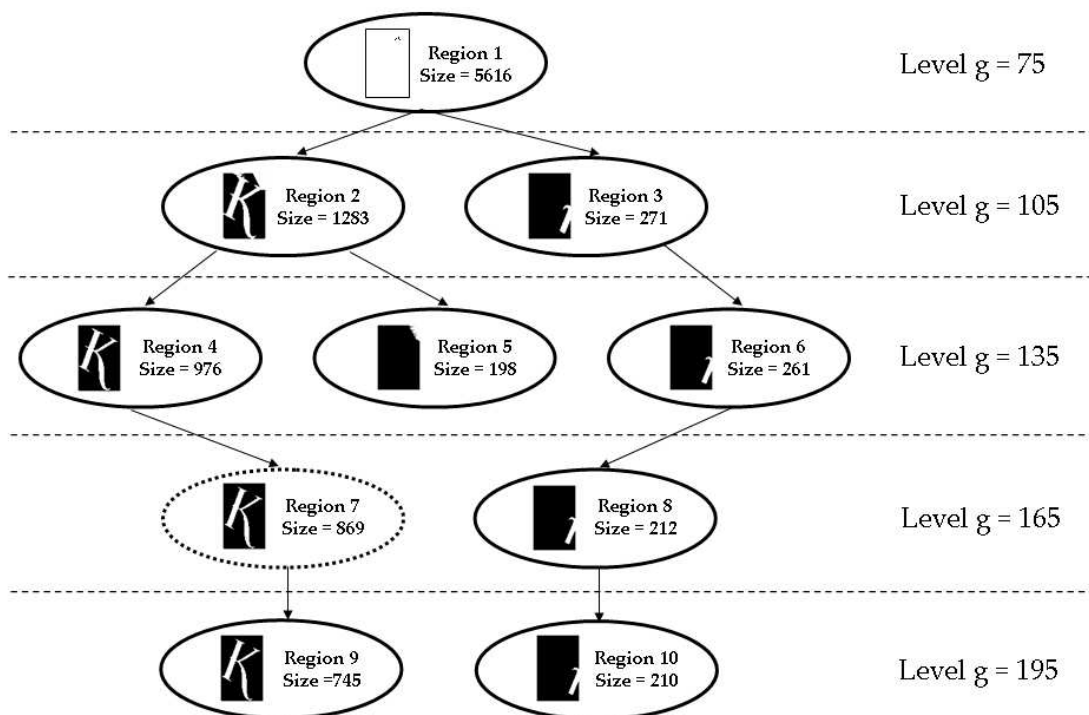
- maximální stabilita je hledána jako nejmenší změna tvaru elipsy, která je vepsána do každého extrémního regionu
- změna pozice těžiště regionu je kritérium, které je kvůli změně afinity potřeba normalizovat. Jako nejlepší prostředek byla experimentálně určena penalizace regionu podle kompaktnosti regionu. Tato kompaktnost je vyjádřena poměrem obsahu a velikosti okraje regionu.
- stabilitu regionu je možné získat kombinací předešlých dvou metod

Algoritmus pro výpočet MSER

Vlastní algoritmus pro výpočet MSER je následující: zkoumaný obraz převedeme na obraz ve stupních šedi. Všechny body vzniklého obrazu seřadíme vzestupně podle intenzity. Tyto body následně procházíme a vkládáme do nového obrazu od nejvyšší intenzity v sestupném nebo vzestupném pořadí. Body, které splňují podmínky, tvoří extrémní regiony. Případné spojení regionů předurčuje výskyt menšího regionu, který je ve vzniklém obsažen. Tento proces vytváří seznam spojených komponent a tvoří základ pro určení maximální stability podle některého z kritérií popsaných v předchozí části.



[obr. 2.5] Prahování [a] vstupní obraz [b] práh $g = 75$ [c] $g = 105$ [d] práh $g = 165$ se výrazně nemění až do $g = 195$ [e] $g = 225$ [f] $g = 255$, obraz je zcela černý [2]



[obr. 2.6] Část vytvořeného stromu regionů vytvořeného podle obrázku 2.5. Region 7 je označen jako MSER. [2]

3 Korespondence významných bodů

Hledání korespondence pro každý pixel zvlášť by bylo výpočetně náročné. Proto se v každém obraze nejprve určí množina klíčových bodů resp. regionů, ze kterých budou vybrány korespondující páry. Tyto klíčové body resp. regiony odpovídají těm, které byly metodami, popsány výše, určeny jako významné. Pro korespondenci je však potřeba najít v různých obrazech odpovídající si body, které lze opakovaně identifikovat nehledě na geometrické a fotometrické transformace obrazu. Jelikož, jak bylo popsáno v předchozích kapitolách, je tento problém dosti složitý a v některých případech stále těžko řešitelný, musí se při zjišťování korespondence vyvíjet a používat další robustní metody.

Některé z metod určování korespondence jsou popsány v následující kapitole.

3.1 Minimální vzdálenost vektorů

Kritériem podobnosti, při volbě této metody, je euklidovská vzdálenost popisů zkoumaných bodů. Proto se pro každý nalezený významný bod obrazu vyhledá, z druhého obrazu bod s nejbližším popisem.

K vyhledání nejbližšího „sousedá“ je možné dojít dvěma způsoby:

- nalezneme vzájemně nejbližší popisy. To znamená, že korespondence je vytvořena jen pokud zvolený bod prvního obrazu je nejbližší ze všech bodů obrazu druhého a zároveň nalezený bod obrazu druhého je nejbližší ze všech bodů obrazu prvního.
- nalezneme ke každému popisu bodu z prvního obrazu nejbližší dva v obrazu druhém. Když je následný poměr vzdáleností prvního/druhého nejbližšího popisu menší než např. 0.8, popisy korespondují. Tuto metodu je například možné použít pro diskriminativní popisy metody SIFT.

Možnosti využití této metody jsou odvislé a zároveň limitované právě popisem hledaných bodů nebo regionů. Je-li v obrazech detekováno malé množství objektů (bodů či regionů) existuje vysoká pravděpodobnost, že vyhledané korespondence budou správné i při relativně slabém popisu daných objektů. Na druhou stranu, obrazy vedoucí, při spojení s některými metodami, k detekci velkého počtu bodů, potřebují generovat popis komplexnější, kvůli výskytu vysokého počtu nesouhlasných párů.

3.2 RANSAC

RANSAC (Random Sample Consensus) [13], poprvé publikovaný Fischlerem Bollesem v roce 1981, je algoritmus pro odhad parametrů matematických modelů ze sady zkoumaných dat, která obsahuje i vzorek dat, vzdálených hledanému modelu (outliers). RANSAC je především vhodný pro data obsahující velké množství dat nenáležících hledanému modelu. V tomto případě je totiž většina z klasických metod nepoužitelná, neboť jsou tyto metody ovlivněny i daty mimo hledaný model.

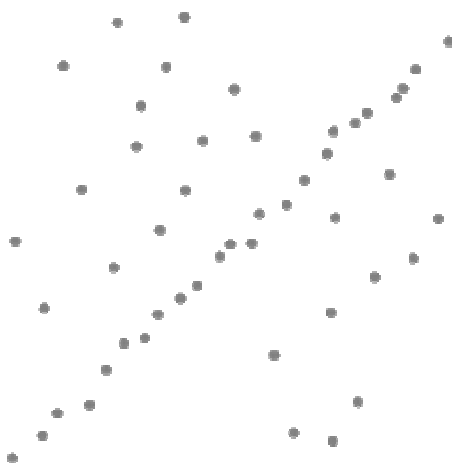
Algoritmus

Algoritmus je popsán následujícím způsobem:

- náhodně vybereme potřebný počet vzorků dat n
- provedeme odhad hledaného modelu z vybraných vzorků (vypočteme parametry), tedy stanovíme hypotézu
- vypočítáme kolik vzorků n odpovídá hledanému modelu, je shodných se stanovenou hypotézou, při zvolené toleranci (např. vzdálenost). Zvolenou hypotézu takto ohodnotíme.
- v případě, že je počet vzorků větší než pro dosud nejlépe ohodnocenou hypotézu, stanovíme novou hypotézu jako nejlepší
- celý algoritmus procházíme zvoleným počtem opakování

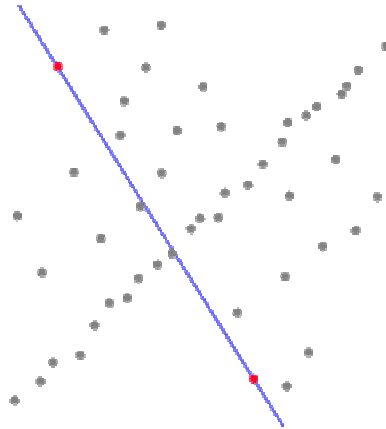
Pro snazší představu následuje příklad.

V tomto příkladu budeme hledat přímku, která nejlépe odpovídá zvolenému vzorku bodů.



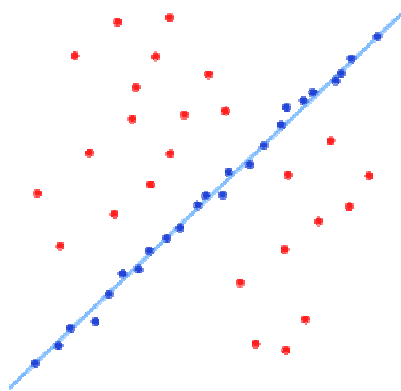
[obr. 3.1] Vzorek bodů, který obsahuje hledanou přímku

Počet bodů, které náhodně vybíráme je určen nejmenším počtem bodů potřebným k výpočtu parametrů hledaného modelu. Přímkou můžeme popsat jako spojnicí, která prochází nejméně dvěma body, tedy vybereme náhodně právě dva a přímkou jimi proložíme.



[obr. 3.2] Náhodně zvolené dva body, které protíná přímka

Způsobů, jak stanovit hypotézu je několik. Můžeme spočítat vzdálenosti všech bodů vzorku od dané přímky a hledaná přímka pak bude mít tyto vzdálenosti minimální. Další možností je stanovit hranice kolem dané přímky a spočítat sumu bodů v této oblasti. Hledaná přímka v tomto případě musí mít tuto hodnotu maximální.



[obr. 3.3] Výsledná nalezená přímka

Hlavní výhoda této metody byla již zmíněna dříve, ovšem nevýhody jsou zde také. Jde o čas poskytnutý této metodě, který se nedá předem určit. V případě omezení času mohou být následující výsledky neoptimální. Další nevýhodou metody RANSAC je nespolehlivost při hledání více shodných modelů na shodné sadě zkoumaných dat.

V praxi je tato metoda používána pro řešení problému korespondence, hledání základních matematicky popsaných objektů a pro odhad základní matice, která se používá pro určení stereometrie obrazů.

4 Návrh detektoru

Tato a následující kapitoly se, na rozdíl od předchozí teorie, zabývají vlastní prací, která dle zadání směřuje k vytvoření vlastního detektoru shodných bodů ve dvou fotografiích.

Funkční aplikace detektoru se musí skládat z několika částí, které v součinnosti poskytnou požadovaný výsledek. Jejím vstupem budou dva obrázky a zvolené parametry detektoru. Výstupem bude obrázek pouze jeden, který však bude obsahovat oba vstupní, a měl by vhodnou formou znázorňovat vybrané významné body nebo regiony, společně s názorným určením jejich vzájemné korespondence v případě, že bude nějaká detekována.

Navržený a implementovaný detektor by měl splňovat základní požadavky, které jsou kladeny na tento druh aplikací. Zároveň však bude patrná snaha vytvořit a modifikovat původní metody, s cílem vylepšení některých vlastností.

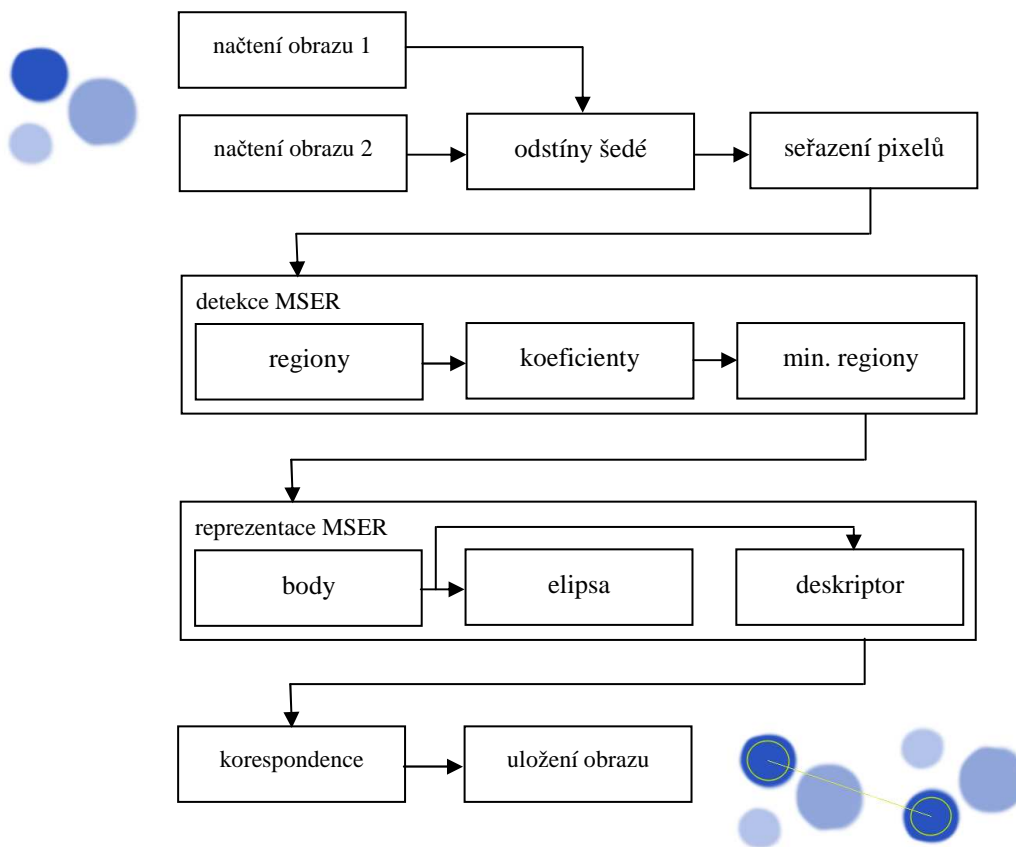
Celkový úkol detekce shodných bodů v obrazech je možné rozdělit na dílčí podproblémy. Nejvýznamnější z nich je část detekce významných bodů s následným určením vzájemné korespondence. Tyto problémy tvoří hlavní kostru celého detektoru a výsledná funkčnost je odvislá právě od volby metod, jež je řeší.

Výsledná aplikace pracuje v několik základních krocích:

- načtení prvního obrázku
- převedení obrázku na odstíny šedé
- detekce významných bodů
- popis a reprezentace nalezených bodů
- opakování všech předchozích bodů pro druhý obrázek
- určení korespondence bodů v obou obrázcích
- vykreslení a uložení názorného výstupu

Uvedené kroky je potřeba rozšířit o specifikaci a zvolené metody. Tedy, obrázky, které detektor v úvodu načítá by měly být běžného rozlišení, co se týče velikosti a barvy. Většinou by mělo jít o obrazy jednoho objektu, ovšem nemusí být totožné a mohou se lišit v různých ohledech.

Výběr vhodné metody pro detekci významných bodů, je stěžejním problémem při návrhu tohoto detektoru. Obyčejně je nutné dobře zvážit výhody a nevýhody různých řešení. V tomto případě bylo nakonec vybíráno ze tří metod, které by měly poskytovat kvalitní výsledky. Z trojice metod Harris rohové detekce, SIFT a MSER byla vybrána právě metoda maximálně stabilních extrémních regionů MSER. Mimo jiné, byl také důvodem pro tuto volbu fakt, že metoda MSER není příliš známá, a přestože by měla poskytovat vynikající výsledky, není její algoritmus popsán či probírán stejně často, jako zbylých dvou metod. Zároveň jde, z této trojice, o jedinou metodu, která nepracuje ve výsledku s body, ale s celými regiony.



[obr. 4.1] Schématický postup práce detektoru

Algoritmus i s postupem výpočtu MSER je rozebrán v kapitole 2.3.

Na základě vlastností metody zvolené pro detekci významných bodů, tedy MSER, byla pro korespondenci vybrána metoda minimální vzdálenosti vektorů, která byla popsána dříve.

Regiony by měly být reprezentovány elipsami, které jsou do nich vloženy. Parametry těchto elips by byly následně použity jako hodnoty tvořící deskriptor regionu.

5 Implementace detektoru

Tato kapitola se zabývá vlastní realizací detektoru shodných bodů ve dvou obrazech na základě návrhu, který byl popsán v předchozí kapitole.

5.1 Knihovna OpenCV

OpenCV (Open Computer Vision library) je volně dostupná grafická knihovna pro analýzu obrazu, která byla původně vyvíjena společností Intel s motivací využití vysokých výkonů nových procesorů. Jejím cílem je podpora komerčního využití počítačového vidění pro interakci člověk-stroj, robotiku, monitorovací a bezpečnostní systémy. Obsahuje velké množství funkcí pro zpracování a analýzu obrazu, ale i tvorbu uživatelských rozhraní a mnohé jiné.

V tomto projektu je využívána právě pro analýzu a zpracování obrazu. Jsou využity funkce pro načítání a výstup obrazu, převod barevného obrazu na odstíny šedé, vložení elips do nalezených regionů, jejich tisk do obrazu a tisk spojnic korespondujících si regionů.

5.2 Parametry detektoru

Program Detektor, který provádí detekci shodných bodů v obraze, se spouští z příkazové řádky, v ideálním případě, s pěti parametry.

- První dva jsou názvy obrazů, které jsou určeny k detekci. Podporovány jsou formáty JPG, PNG a BMP.
- Třetí parametr je název výstupního obrazu, do něhož budou uloženy vstupní obrazy, společně s detekovanými regiony a naznačením jejich souhlasnosti.
- Čtvrtým parametrem je stabilita, s kterou je detekce prováděna. Úloha a význam tohoto parametru bude blíže rozebrán později, v místě implementace.
- Poslední parametr je práh koeficientu maximální stability. Stejně jako u předchozího, jde o parametr, jehož nastavení je kritické, tedy vysokou měrou určuje výsledné vlastnosti detektoru. Bližší vlastnosti a vhodná nastavení budou popsána v kapitole 6.

5.3 Segmentace obrazu

Po načtení prvního obrazu, je tento převeden na odstíny šedi. Pro další zpracování je však nutné uložit jednotlivé pixely obrazu tak, aby je bylo možné seřadit podle velikosti. K tomuto účelu slouží pole struktur obsahující pozici pixelu a jeho hodnotu (intenzitu).

Seřazení pixelů podle velikosti je prováděno standardní funkcí `gsort()`, jejíž rychlost je důležitá především při větších rozměrech obrazů. Quick sort však není jediným řadícím algoritmem,

který je možné použít. Pro tento typ dat, tedy velké množství celočíselných hodnot v malém rozsahu, je zajímavou alternativou algoritmus radix sort. Ten však není standardně implementován a detektor tak využívá výše zmíněné, optimalizované metody quick sort.

Již seřazené pole pixelů je zpracováno funkcí MSER(), která je hlavní metodou použitou k segmentaci obrazu na jednotlivé regiony. Tato funkce pracuje ve dvou cyklech:

- v prvním, který obsahuje celou funkční část segmentace, je procházeno právě pole pixelů seřazených podle intenzity, od nejvyšší po nejnižší. Teoretický rozsah hodnot intenzity je 256, ovšem většina fotografií má tento rozsah mírně zkrácen. Čím je tento rozsah větší, tím je i kontrast obrazu větší, což je pro správnou detekci významných bodů pozitivní.
- v druhém cyklu jsou, pro každý jednotlivý bod z prvního cyklu, zkoumány jednotlivě čtyři sousedé. Při tom je potřeba zjistit, zda soused existuje a následně vyhledat jeho hodnotu, resp. hodnotu kořene regionu, do kterého náleží. Hodnota je funkcí IsInDelta() porovnána s hodnotou původního pixelu.

Vlastním řešením se vytváří strom regionů, jehož uzly obsahují další data potřebná k pozdějšímu zpracování.

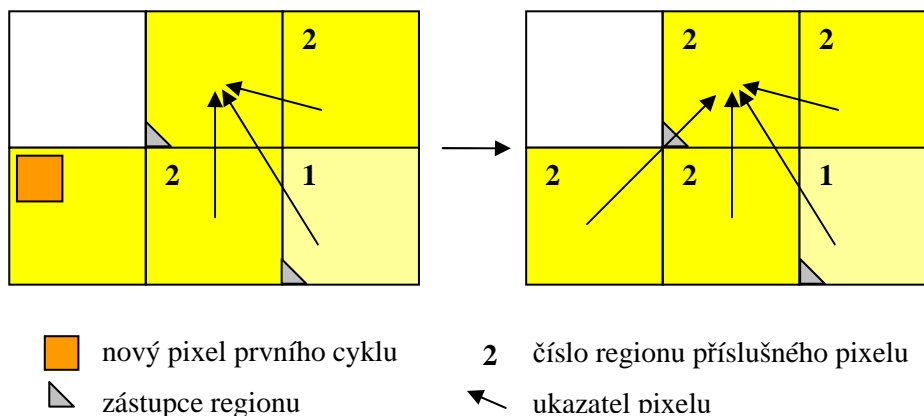
Kvůli snazšímu pochopení i možnosti snazšího popisu následujících stavů, je nejprve vysvětlena samotná struktura uzlů:

- zástupce regionu je pixel, na nějž ukazují všechny pixely regionu, do kterého samotný zástupce patří. Je vytvářen při zakládání nového regionu, z čehož vyplývá, že má v tomto regionu nejvyšší intenzitu. Právě od této intenzity a zvolené stability, je počítán rozdíl, který určuje, zda hledaný pixel ještě náleží do regionu zástupce, nebo nikoli. Stabilita tak určuje rozsah intenzit, které se mohou vyskytovat v jednom regionu.
- předek je indexem rodičovského uzlu ve vytvářeném stromu. Určuje region, do kterého původní region s vyšší intenzitou spadá.
- okraj regionu je tvořen pixely, jenž splňují definici (2.5). V případě, že je kritérium stability uvažováno dle původní definice (2.7), je velikost okraje regionu pro výpočet koeficientu maximální stability nezbytná.
- koeficient maximální stability je vypočten dle definice (2.7).
- příznak maximální stability regionu určuje pouze regiony, které jsou dle daných podmínek maximálně stabilní a extrémní. Pouze regiony s tímto kladným příznakem jsou určeny k dalšímu zpracování.

Ve skutečnosti dochází při zkoumání sousedních pixelů ke čtyřem stavům, které je potřeba řešit.

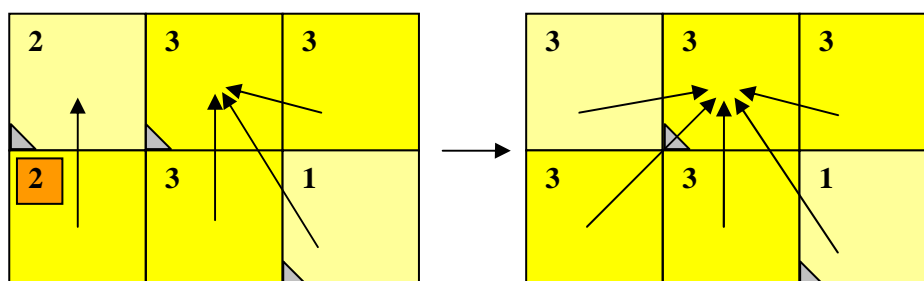
V případě, že původní pixel nemá žádného aktivního souseda v mezích stability, určených funkcí `IsInDelta()` na základě rozdílu intenzit, je nutné původní pixel označit jako nový region. Vytvoří se tak uzel stromu, který tvoří list a není tak předkem žádného jiného regionu. Zároveň je pixel označen jako aktivní.

V ostatních případech jde vždy o aktivního souseda, který již náleží nějakému regionu. Nejjednodušším stavem jenž nastává, je případ nově sledovaného pixelu z prvního cyklu, který ještě není přiřazen žádnému regionu. V případě nalezení aktivního souseda, jenž leží v mezích stability, je původní pixel připojen do regionu souseda, resp. kořene regionu souseda. Ačkoliv jde o snadnou funkci nesmíme vynechat definici (2.5) a v případě, že původní pixel nepřipojujeme přímo k regionu souseda, je původní pixel vnějším okrajem regionu sousedního.



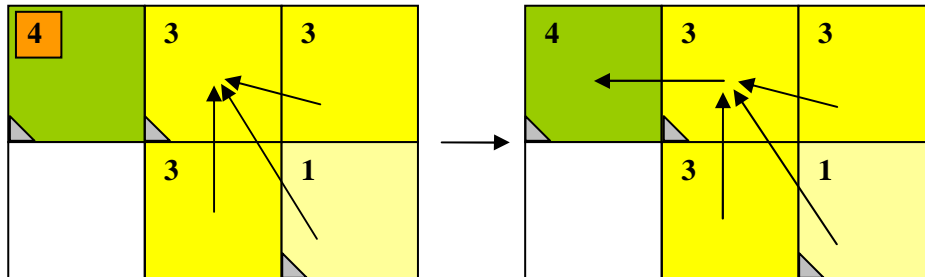
[obr. 5.1] Modelový příklad připojení nového pixelu ke stávajícímu regionu s intenzitou v mezích stability. Uvedené vysvětlivky jsou platné i pro obrázky 5.2 a 5.3

Další stav je odvozen od stavu předchozího, v důsledku je však velice odlišný. Stejně jako v předchozím případě, jde o pixel z prvního cyklu a souseda v mezích stability, ovšem původní pixel je již zpracován, a tedy náleží nějakému regionu. Z toho vyplývá, že jde o bod, kde dochází ke spojení dvou regionů stejné úrovně intenzity. Ve stromu pak dochází ke spojení dvou uzlů. Při tomto spojení je potřeba sloučit všechny pixely obou regionů do jednoho, což zajišťuje funkce UnifyReg(). Tato funkce řeší hned několik problémů, které při spojování regionů stejné stability vznikají. Kvůli úspoře času nejsou rušeny oba původní regiony a vytvářen jeden, který by je nahradil, ale jeden je ponechán a druhý se stane jeho součástí. K tomu je potřeba upravit předky všech regionů, které ukazovaly na druhý region tak, že nyní ukazují na region první a jednotlivé pixely druhého regionu náleží regionu prvnímu.



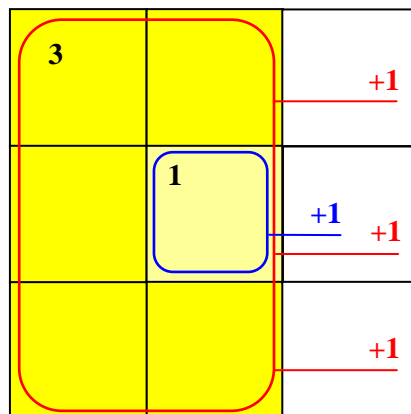
[obr. 5.2] Příklad spojení dvou regionů s intenzitou v mezích stability. Zkoumaný pixel je připojen do regionu 2, ale následně celý region 2 sloučen s regionem 3. Zaniká i zástupce regionu 2.

Poslední stav, ke kterému může v algoritmu MSER dojít je případ, kdy souseď neleží v mezích intenzity. V tomto stavu je potřeba řešit situaci, kdy je původní pixel neaktivní, tedy není ještě přiřazen do žádného regionu a je tak vytvořen region nový, nebo již je součástí nějakého regionu. V každém případě je však region souseďa připojen jako potomek. Připojení potomka nám signalizuje, že souseďní region je již celý zpracován.



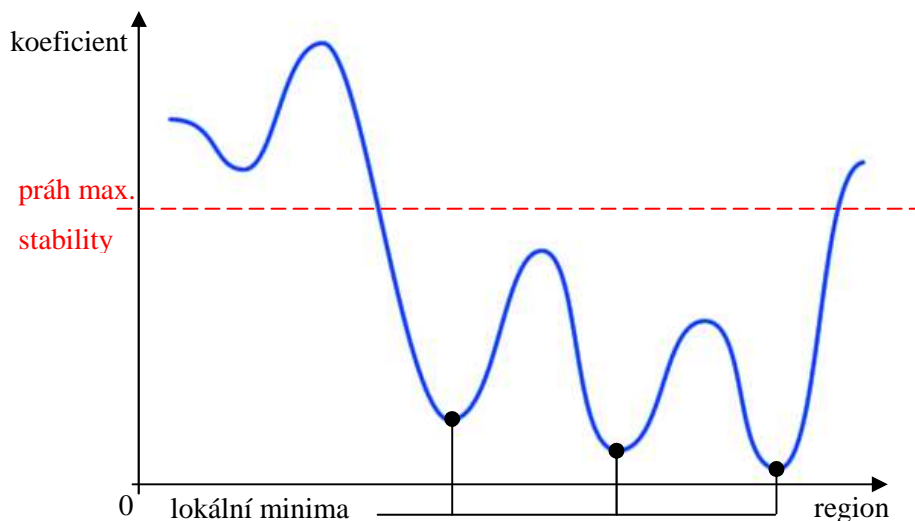
[obr. 5.3] Příklad připojení stávajícího regionu 3 na region 4, jehož intenzita je mimo mez stability.

Zároveň je potřeba správně vytvořit okraje souseďních regionů. Pro lepší představu jak dochází k určování, kterému regionu okraj náleží, je následující obrázek.



[obr. 5.4] Vnější okraj tvoří pixely přímo souseďící s regionem. Souseď tvoří vnější okraj všem regionům, kterých se přímo dotýká. Velikost okraje regionu 1 je 4, regionu 3 je 10.

Následující funkce, zajistí výpočet koeficientu stability pro jednotlivé regiony a zajistí, že budou dále uvažovány pouze regiony s koeficientem větším než je zvolený práh. Následně je potřeba zajistit, aby se u tohoto koeficientu jednalo pouze o lokální minimum. Tím se výsledný počet regionů opět výrazně sníží.



[obr. 5.5] Graf ukazuje, které regiony, dle koeficientu, budou vybrány jako maximálně stabilní.

Koeficient zajišťuje, že vybrané oblasti budou co největší a jejich intenzita se nebude výrazně měnit.

Regiony, které jsou předchozím postupem získány jsou především světlé oblasti s tmavým ohraničením. Ve skutečnosti však reálné obrazy a zejména fotografie, obsahují i oblasti opačné charakteristiky. Jde o tmavé oblasti se světlým ohraničením, které jsou při předchozím postupu zpracovány tak, že nemohou tvořit stabilní regiony.

Z důvodů nejednoznačnosti regionů, které by byly detekovány předchozím postupem i postupem obráceným (detekce tmavých oblastí). Je implementována možnost volit mezi těmito postupy znaménkem parametru stability. Parametr stability se zápornými hodnotami tak určuje právě detekci tmavých oblastí. Samotný postup detekce je analogický předešlému. Je možné si jej představit jako průchod již seřazeného pole pixelů v opačném směru, nebo stejný postup pro obraz, který má hodnoty intenzit jednotlivých pixelů invertované. U obrazu v odstínech šedi jde o doplněk intenzity do 255.

5.4 Korespondence oblastí

Předchozí algoritmus vybral z obrazu významné oblasti. Následující postup již směřuje k jejich sesouhlasení.

Funkce z knihovny OpenCV `cvFitEllipse()` zajistí optimální vložení elipsy každému MSER regionu. K tomu je potřeba nejméně šest bodů, což opět sníží počet výsledných oblastí určených k hledání korespondence.

Po dokončení detektoru, dle návrhu, bylo následným testováním zjištěno nevyhovující chování detektoru. Program vyhodnocoval korespondenci regionů nevyhovujícím způsobem a vykazoval vysokou chybovost právě při určování korespondence.

Důvodem byla nedostatečná reprezentace MSER regionů. Dle návrhu byl použit popis regionu pouze parametry elipsy, která každý reprezentovala. Tento popis je ale velice hrubou aproximací regionu a má význam především pro vizualizaci regionu. Při vyšším počtu nalezených regionů ve složitých obrazech nebo při větší změně afinity obrazu už tato reprezentace nestačila. Z toho důvodu byla reprezentace regionu rozšířena o popis pomocí momentových funkcí.

Momenty jednotlivých regionů, jsou funkce spočtené nad všemi body obsaženými v regionu, jejich intenzitou a polohou. Z důvodu vytvoření co největší robustnosti je využíváno centrálních momentů třech řádů, které mají zaručit neměnnost popisu při změně polohy a rotace obrazu.

Výpočet momentů se řídí vzorcem:

$$M_{ij} = \sum_x \sum_y x^i y^j I(x, y) \quad (5.1)$$

kde, x a y jsou souřadnice bodu v obraze a funkce $I()$ nám dává hodnotu intenzity.

Moment nultého řádu:

$$\mu_{00} = M_{00} = \sum_x \sum_y x^0 y^0 I(x, y) \quad (5.2)$$

Moment prvního řádu:

$$\mu_{11} = M_{11} - \bar{x}M_{01} \quad (5.3)$$

kde M_{11} a M_{01} je analogií M_{00} a těžiště je počítáno podle vztahu:

$$\bar{x} = \frac{M_{10}}{M_{00}} ; \bar{y} = \frac{M_{01}}{M_{00}} \quad (5.4); (5.5)$$

Moment druhého řádu:

$$\mu_{20} = M_{20} - \bar{x}M_{10} \quad (5.6)$$

$$\mu_{02} = M_{02} - \bar{y}M_{01} \quad (5.7)$$

Funkce $Korespond()$ zjišťuje korespondence mezi regiony tak, že hledá nejmenší vzdálenost mezi vektory popisu regionu. Hodnota je vypočítána podle vztahu pro vzdálenost v n -rozměrném euklidovském prostoru:

$$d = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (5.8)$$

V případě hledání nejmenší vzdálenosti však vzniká problém vícečetné korespondence k jednomu regionu, když existuje region, který je nejbližší více regionům. Tento problém je řešen tak, že jako korespondující pár je určen pouze ten, jehož region prvního obrazu má nejbližší region v druhém obrazu a zároveň, pro tento nalezený, je původní region z prvního obrazu také nejbližší ze všech. Pro zachování jednoznačnosti se navíc musí první nejmenší vzdálenost rozcházet s druhou o více než dvojnásobek. V opačném případě je daná korespondence ignorována právě z důvodu možné nejednoznačnosti.

Na výstupu se nakonec vytisknou pouze regiony, jejichž korespondující pár obsahuje elipsy, jejichž velikost je v daných mezích. Dolní mez je určena empiricky na pět pixelů, čímž jsou odstraněny velmi malé regiony a horní mez byla empiricky zvolena jako průměrná velikost elips, čímž jsou odstraněny největší regiony.

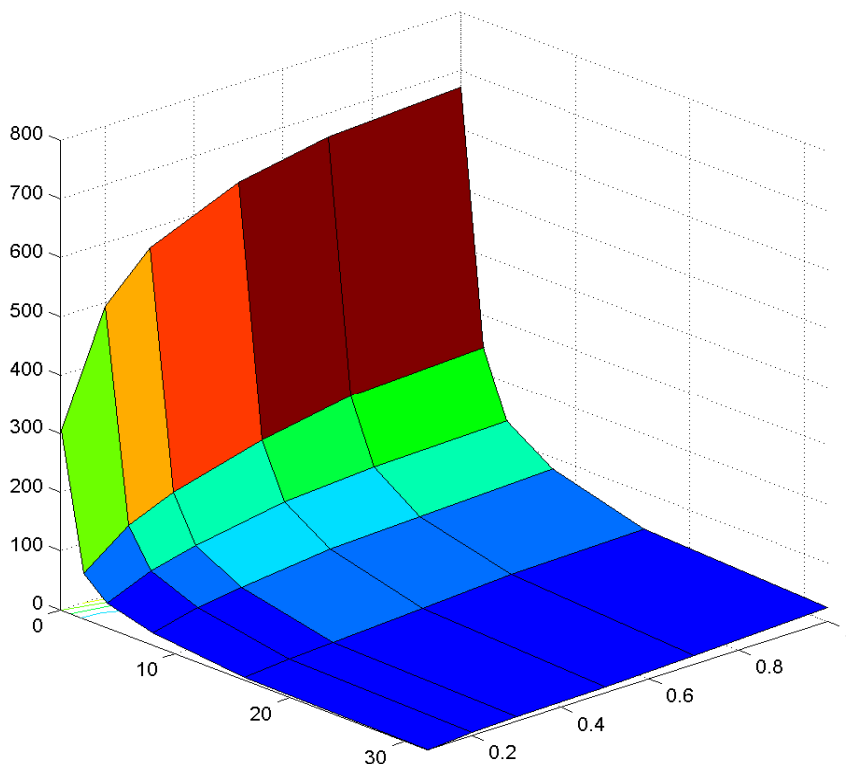
6 Dosažené výsledky

Kapitola by měla stručnou formou shrnout nejen výsledky, kterých bylo dosaženo, ale také funkčnost, případné nedostatky, možnost rozšíření do budoucna, modifikace a využití.

6.1 Vhodná vstupní data

Účinnost detektoru, který je výše popsán, je odvislá především od tří vstupních informací, které jsou mu poskytnuty. Je to parametr stability, který ovlivňuje rozsah hodnot intenzity v jednotlivých regionech. Práh koeficientu stability určuje jaká míra stability je požadována po detekovaných regionech. A informace, která ovlivní detekci nejvíce, je charakter vstupních obrazů. Poslední zmiňovaná informace záleží největší měrou na zkušenosti uživatele, což je aspekt, který pouze těžko popsatelný. Proto bude pouze stručně zmíněn v závěru této kapitoly.

Při zjišťování vhodného nastavení parametrů, které je možno přímo volit, tedy stability a prahu koeficientů, bylo postupováno následujícím způsobem. Detekce byla prováděna nad shodnými obrazy, což nám poskytne nejvyšší možný počet korespondencí, které je možné z daného obrazu získat. Hodnoty obou zmíněných parametrů nabývaly šest volených hodnot. Jejich kombinace a výsledný počet vzniklých MSER regionů je patrný z následujícího grafu.



[obr.6.1] Graf - počet výsledných korespondencí v závislosti na zvolených parametrech



[obr. 6.2] Obraz, z kterého byly získány hodnoty pro předchozí graf 6.1

Sérií testů byly zjišťovány nejlepší hodnoty parametrů a jako nejvýhodnější bylo určeno rozmezí:

- stabilita 2 - 15
- práh koeficientů 0.1 - 1

toto rozmezí není možné obecně zúžit, protože je výsledek také výrazně ovlivněn charakterem obrazu. Navíc je určení, zda jde o vhodnou detekci do značné míry subjektivní.

Charakter obrazu, který je vhodný pro tento detektor, je možné popsat podle způsobu práce se samotným obrazem následovně:

- podle frekvenční charakteristiky by měl obraz obsahovat převážně nízké frekvence (jemné přechody intenzity) a menší množství míst s vyšší frekvencí, které tvoří detaily a výrazné hranice.
- z pohledu intenzity bodů, jsou nejvhodnější obrazy obsahující intenzity celého možného rozsahu hodnot
- obecně jsou pro detekci nejvhodnější obrazy obsahující jeden výrazný prvek, jehož tvar je co možná nejkompaktnější, v ideálním případě kruhový nebo oválný.

Odolnost vůči geometrickým změnám

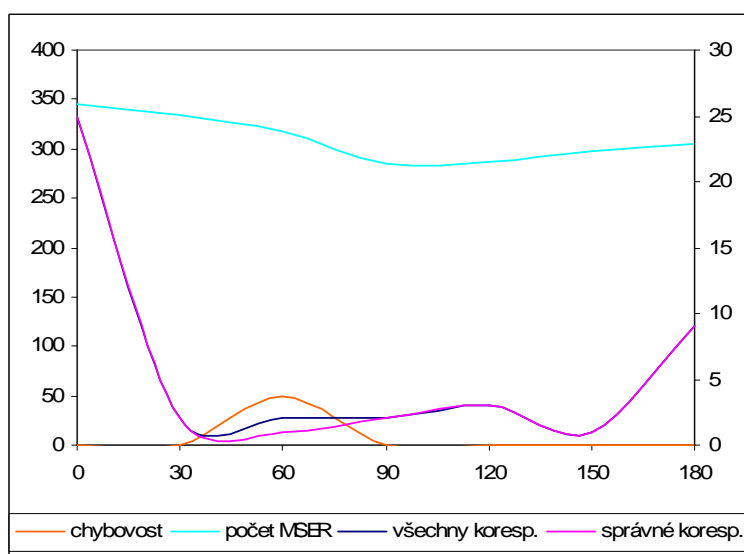
Výsledky ukazují možnosti detektoru při vhodných vstupních parametrech a odpovídají konkrétním obrazům.

Následující tři testy byly provedeny na obraze 6.3, při nastavení hodnot stability -8 a prahu koeficientů na 1,0.



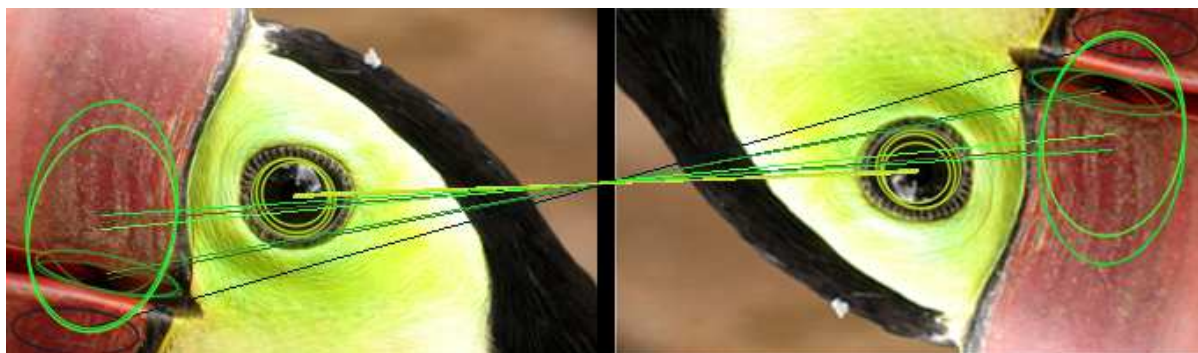
[obr. 6.3] Obraz použitý pro následující tři detekce

Rotace obrazu:



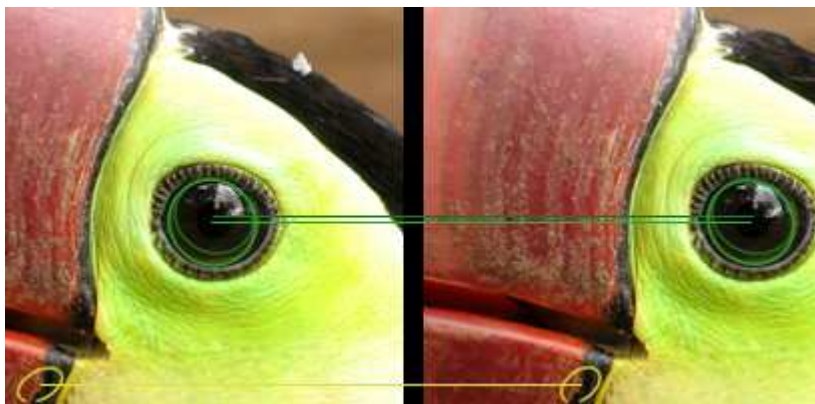
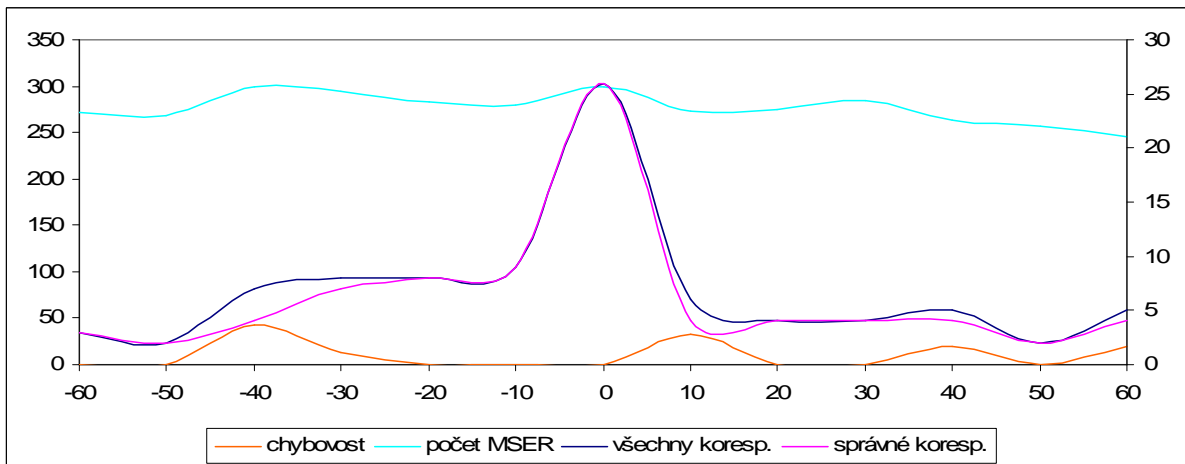
Korespondence je v rozmezí 1 až 25.

Za vynikající výsledek je možné považovat nalezené korespondence při rotaci obrazu o 180°.



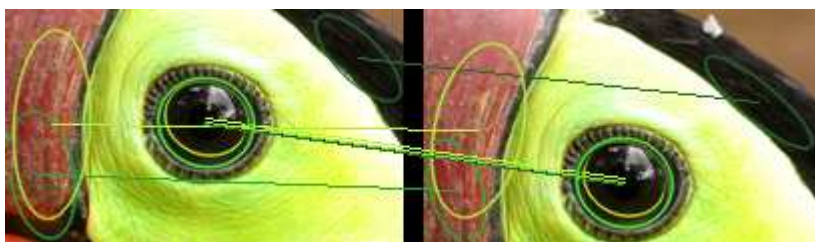
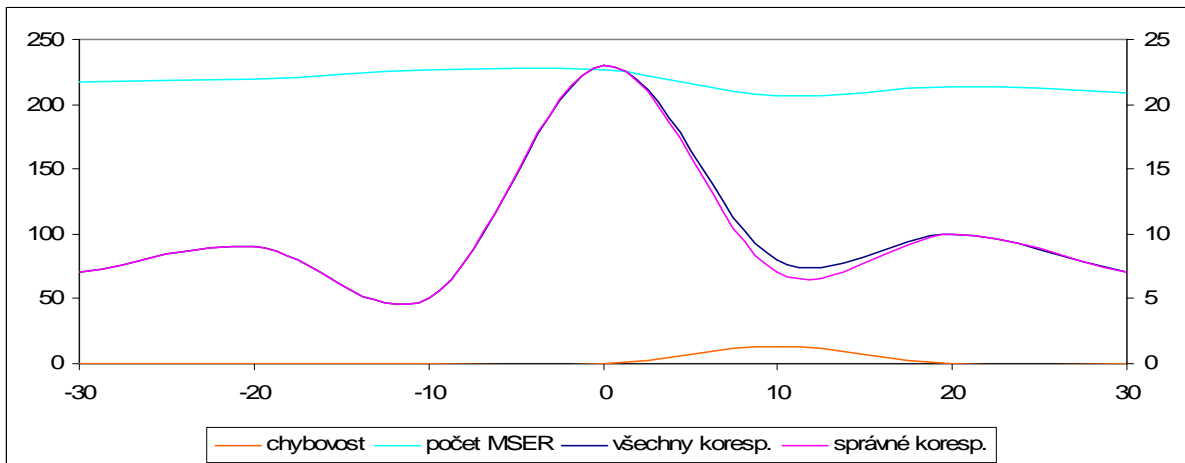
[obr. 6.4] Graf ukazuje počet výsledných korespondencí při změně rotace obrazu. Obrázek je ukázkou samotného výstupu detektoru.

Horizontální posun obrazu:



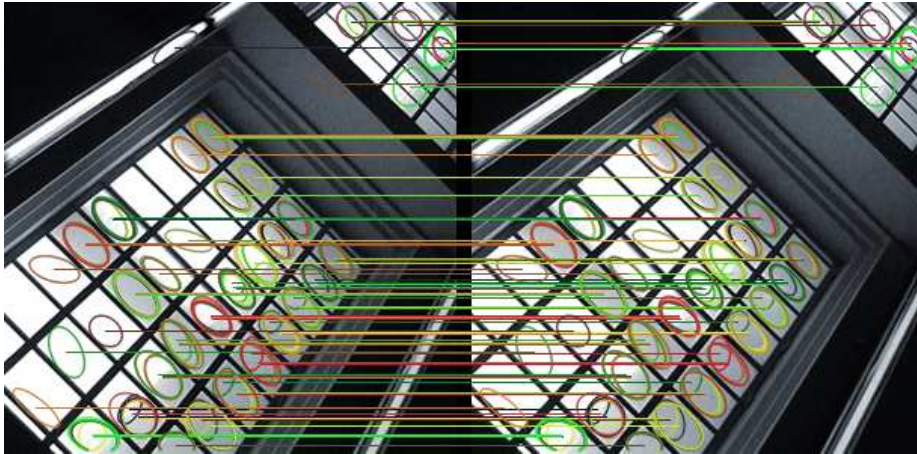
[obr. 6.5] Graf hodnot při posunech obrazu o 10 pixelů ve vodorovném směru. Obrázek korespondencí při posunu o 60 pixelů.

Vertikální posun obrazu:

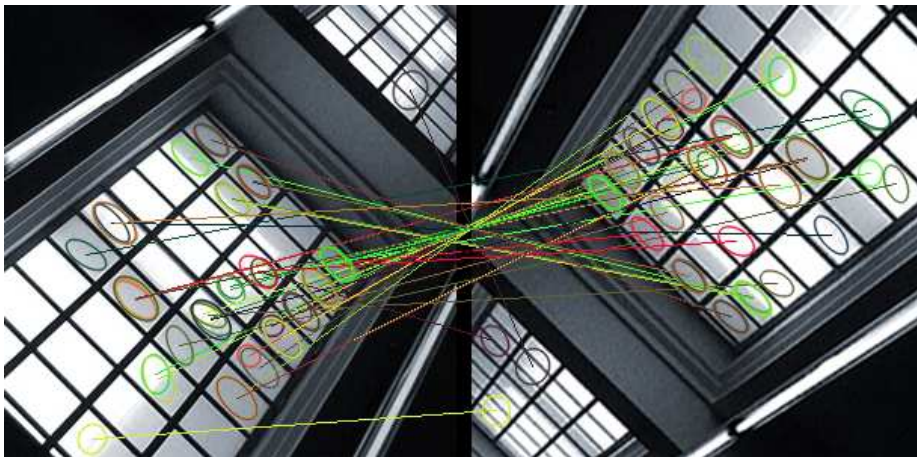


[obr. 6.6] Graf hodnot při posunu o 10 pixelů ve svislém směru. Obrázek korespondencí při posunu o 30 pixelů.

Obraz s velkým počtem podobných regionů:



[a]



[b]

[obr. 6.7] [a] detekované korespondence ve shodném obraze při parametrech, stabilita 5, práh koeficientů 0.5 [b] korespondence při otočení obrazu [a] o 180°.

Velké množství podobných oblastí obrazu 6.6 způsobí, že některé korespondence jsou určeny chybně. Jde o problém, jenž je automaticky, velice obtížně řešitelný. Aby byl odstraněn, bylo by potřeba detektor rozšířit o mnoho složitých metod zaměřených přímo na řešení tohoto problému, což by značně přesahovalo rozsah bakalářské práce.

6.2 Návrh modifikací a rozšíření

Detektor navržený a implementovaný, způsobem popsaným v této práci, má značný počet nevýhod, díky kterým by byl v praxi použitelný jen s obtížemi. Návrh rozšíření, či modifikací, které by některé nevýhody odstranily je obsahem této kapitoly.

Nevýhoda: Vysoká časová náročnost

Trend vývoje aplikací v poslední době vyžaduje v hojné míře zpracování v reálném čase a právě čas je jedním z klíčových nedostatků tohoto detektoru. Vysoká časová náročnost se projevuje především při zpracování větších obrazů. Nejvyšší měrou ji způsobuje segmentace obrazu do jednotlivých regionů. Časově nejnáročnější je konkrétně funkce `UnifyReg()`, a to při spojování velkých regionů. Ve složitých obrazech navíc dochází k tomuto spojování velice často.

Dobu potřebnou pro zpracování jednoho obrazu také, více než dvojnásobně, prodlužuje dvojitě zpracování, hledající regiony i v invertovaném obraze. Empiricky bylo zjištěno, že obrazy, které obsahují regiony nenáročné k segmentaci, mají často v invertované verzi problém s vysokým počtem spojovaných regionů (je nutné použít funkci `UnifyReg()`), nebo regionů, které nesplňují podmínky MSER.

Řešení: můžeme najít v optimalizaci samotné funkce `UnifyReg()`, případně by bylo vhodné rozšířit detektor o uživatelské rozhraní, které by umožnilo vybrat v obraze, resp. v obrazech menší oblast pro zpracování. Další rozšíření by mohlo spočívat v přidání funkce, která by určila, zda jde o obraz, ve kterém je potřeba hledat světlé body nebo tmavé body a zda segmentovat obraz původní, nebo invertovaný, popřípadě oba.

Nevýhoda: Účinnost detekce

I ty nejsofistikovanější detektory jsou svou účinností závislé na charakteru vstupních dat.

Řešení: je možné nalézt například ve vhodné úpravě vstupních dat před zpracováním, v použití účinnějších metod pro hledání korespondencí (RANSAC) nebo ve spolupráci s uživatelem, který může upravit, či upřesnit některé parametry tak, aby mu výstup detektoru zcela vyhovoval.

Během práce na této bakalářské práci bylo také experimentováno s vlastní implementací. V zásadě ale nikdy výsledek nepředčil řešení popsané tvůrci vybraných metod. Algoritmus MSER funguje podle předpokladů, vlastní implementace ale za originální metodou značně časově zaostává. Výsledná účinnost je snížena v hlavní míře, procesem hledání korespondencí. Samostatný výsledek, jenž poskytuje metoda MSER je naopak velice dobrý.

7 Závěr

Práce je zaměřena na detekci významných bodů v obrazech a určení jejich korespondence. Teoretický základ slouží ke stručnému seznámení se základními, ale i novými či modernějšími metodami, které se v oboru počítačového vidění pro daný problém používají. Stěžejní část je především směřována k detailnímu prostudování metody MSER, která je používána v kategorii blob detektorů pro detekci významných regionů. Na této metodě je také z velké části postaven návrh samotné aplikace detektoru. Zdrojové kódy nejsou běžně dostupné, proto bylo nutné metodu nově implementovat.

Implementace byla vytvořena na základě původního návrhu metody [1]. Jedná se hlavně o samotný algoritmus a výběr kritéria stability. Detaily, které původní návrh neobsahuje byly vytvořeny empiricky, na základě sérií testování.

S popisem regionů, deskriptorem, který je potřebný k určování korespondencí, bylo experimentováno, přičemž nejlepších výsledků bylo dosaženo při použití momentových funkcí obrazu nad jednotlivými regiony. Výsledné korespondence určuje nejmenší vzdálenost vektorů popisu.

Body zadání se podařilo splnit. Výsledkem práce je aplikace, která splňuje základní požadavky. I přes nedostatky, hlavně ve velké časové náročnosti a nutnosti individuálního hledání vhodných parametrů pro každou dvojici obrazů, může tento detektor poskytovat velmi dobré výsledky při samotné detekci shodných bodů v obrazech, s důrazem na počet a kvalitu nalezených korespondencí.

Samotná aplikace by byla v praxi využitelná jen obtížně. S odstraněním některých nedostatků a rozšířením o nové funkce (např. automatické určení vhodných parametrů, určení sledované oblasti) je ale možné použít tento detektor jako základ aplikací pro spojování panoramatických snímků, sledování pohyblivých objektů nebo rekonstrukci prostoru.

Literatura

[1] MATAS, J., aj.: Robust Wide Baseline Stereo from Maximally Stable Extremal Regions, Prague: CTU Center for Machine Perception Department of Cybernetics, 19. července 2002.

Dostupný na URL: <http://cmp.felk.cvut.cz/~matas/papers/matas-bmvc02.pdf> (leden 2008)

[2] DONOSER, M., BISCHOF, H.: Efficient Maximally Stable Extremal Region (MSER) Tracking, Graz: Institute for Computer Graphics and Vision Graz University of Technology, 30. března 2006.

Dostupný na URL: <http://www.icg.tugraz.at/pub/pdf/docvpr2006.pdf> (leden 2008)

[3] LOWE, D.W.: Distinctive image features from scale-invariant keypoints, Vancouver: Computer Science Department University of British Columbia, 5. ledna 2004.

Dostupný na URL: <http://www.cs.ubc.ca/~lowe/papers/ijcv04.pdf> (leden 2008)

[4] ŠPANĚL, M., BERAN, V.: Obrazové segmentační techniky – Přehled existujících metod, Brno: VUT Fakulta informačních technologií, 12. října 2005.

Dostupný na URL: <http://www.fit.vutbr.cz/~spanel/segmentace/en> (leden 2008)

[5] ZEMČÍK, P., ŠPANĚL, M.: Zpracování obrazu, [studijní opora], Brno: VUT Fakulta informačních technologií, listopad 2006.

[6] NEUMANN, B.: Computer Vision, Hamburg: Universität Hamburg Fakultät für Informatik, 2003.

Dostupný na URL: <http://lki-www.informatik.uni-hamburg.de/~neumann/BV-SS-2003/BV-Woche1.pdf> (leden 2008)

[7] FROLOVA, D., SIMAKOV, D.: Matching with Invariant Features, Rehovot: The Weizmann Institute of Science, 18. března 2004.

Dostupný na URL:

http://www.wisdom.weizmann.ac.il/~deniss/vision_spring04/files/InvariantFeatures.ppt (leden 2008)

[8] BAY, H., TUYTELAARS, T., GOOL, L.V.: SURF: Speeded Up Robust Features, Zurich: ETH Zurich, květen 2006.

Dostupný na URL: <http://www.vision.ee.ethz.ch/~surf/eccv06.pdf> (leden 2008)

- [9] PERŔOCH, M.: Invariantní detekce stabilních oblastí ve 2D obrazech, [DP], Czech Technical University in Prague, Faculty of Electrical Engineering, květen 2004.
Dostupný na URL: <http://cmp.felk.cvut.cz/~perdom1/dp/final.pdf> (leden 2008)
- [10] HÝNA, P.: Detekce rohů v obraze, [BP], Brno: VUT Fakulta informačních technologií, 12. května 2007.
Dostupný na URL: <http://www.fit.vutbr.cz/study/DP/rpfile.php?id=4930> (leden 2008)
- [11] Wikipedie: The Free Encyclopedia: Interest point detection [online].
Dostupný na URL: http://en.wikipedia.org/wiki/Interest_point_detection (leden 2008)
- [12] Wikipedie: The Free Encyclopedia: Corner detection [online].
Dostupný na URL: http://en.wikipedia.org/wiki/Corner_detection (leden 2008)
- [13] Wikipedie: The Free Encyclopedia: Scale-invariant feature transform [online].
Dostupný na URL: http://en.wikipedia.org/wiki/Scale-invariant_feature_transform (leden 2008)
- [14] Wikipedie: The Free Encyclopedia: RANSAC [online].
Dostupný na URL: <http://en.wikipedia.org/wiki/RANSAC> (leden 2008)
- [15] BROWN, M., SZELISKI, R., WINDER, S.: Multi-Image Matching using Multi-Scale Oriented Patches, Redmond: Microsoft Corporation, prosinec 2004.
Dostupný na URL:
<http://vision.bc.edu/~dmartin/teaching/cs342/2006f/readings/BrownSzeliskiWinder2004-MOPs.pdf>
(leden 2008)

Seznam zkratek

RGB Red Green Blue

CMY Cyan Magenta Yellow

HSV Hue Saturation Value

HLS Hue Lightness Saturation

AD Analogově Digitální

SUSAN Smallest Univalve Segment Assimilating Nucleus

USAN Univalve Segment Assimilating Nucleus

SIFT Scale-Invariant Feature Transform

SURF Speeded Up Robust Features

MSER Maximally Stable Extremal Region

Seznam příloh

Příloha 1. Tabulky hodnot

Příloha 2. Výstupní soubor

Příloha 3. CD/DVD

Příloha č.1 - Tabulky hodnot

Tabulka hodnot použitých v grafu 6.1

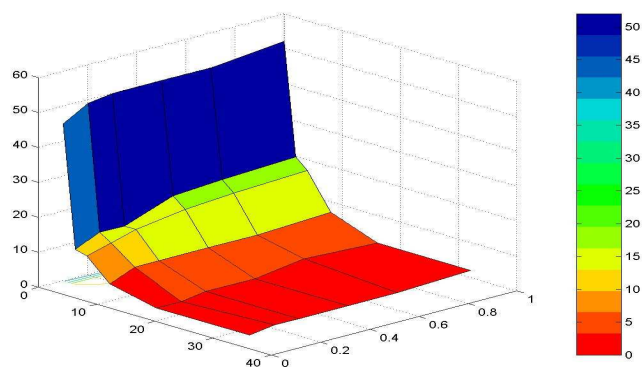
k/s	0	2	4	8	16	32
1	305	493	569	631	660	672
0,7	77	135	167	207	234	245
0,5	42	72	91	116	127	133
0,3	21	40	49	66	72	82
0,2	5	11	15	25	34	39
0,1	1	6	8	11	16	24

k = koeficient stability <0,1; 1>

s = stabilita <0; 32>

Tabulka a graf výsledného počtu korespondencí v obrázku 6.2

k/s	0	2	4	8	16	32
1	45	49	50	50	50	52
0,7	10	13	13	18	19	20
0,5	9	11	13	15	16	17
0,3	3	6	6	6	6	7
0,2	0	0	1	1	3	2
0,1	0	1	1	1	1	2



Skokové hodnoty počtu korespondencí jsou způsobeny funkcí, která se snaží vynechat možné nejednoznačné korespondence.

Tabulka grafu 6.4 - rotace

0	25	25	0	346
30	2	2	0	335
60	2	1	50	317
90	2	2	0	285
120	3	3	0	287
150	1	1	0	298
180	9	9	0	305

Tabulka grafu 6.5 - horizontální posun

posun	počet	správné	chybovost	celkem
-------	-------	---------	-----------	--------

[pixel]	koresp.	koresp.	[%]	MSER
-60	3	3	0	271
-50	2	2	0	269
-40	7	4	42,8	300
-30	8	7	12,5	295
-20	8	8	0	283
-10	9	9	0	279
0	26	26	0	300
10	6	4	33,3	273
20	4	4	0	274
30	4	4	0	284
40	5	4	20	264
50	2	2	0	257
60	5	4	20	246

Tabulka grafu 6.6 - vertikální posun

posun [pixel]	počet koresp.	správné koresp.	chybovost [%]	celkem MSER
-30	7	7	0	217
-20	9	9	0	220
-10	5	5	0	227
0	23	23	0	227
10	8	7	12,5	207
20	10	10	0	214
30	7	7	0	209

Příloha č.2 - Výstupní soubor

Ukázka kompletního výpisu souboru vystup.txt.

```
reg ; okraj ; koe  
57 ; 18 ; 0.111111  
335 ; 63 ; 0.126984  
409 ; 44 ; 0.136364  
452 ; 7 ; 0.142857  
508 ; 8 ; 0.125000  
1159 ; 91 ; 0.087912
```

pocet MSER:6

```
reg ; okraj ; koe  
64 ; 21 ; 0.047619  
210 ; 14 ; 0.142857  
267 ; 18 ; 0.111111  
395 ; 17 ; 0.058824  
478 ; 90 ; 0.177778  
541 ; 24 ; 0.041667  
1300 ; 15 ; 0.133333  
1433 ; 315 ; 0.028571  
1550 ; 38 ; 0.026316  
1890 ; 38 ; 0.131579  
2089 ; 31 ; 0.161290
```

pocet MSER:11

stabilita:32 ; msi:0.200000 ; n-korespond:0