# ABSTRAKT

# ABSTRACT

# KLÍČOVÁ SLOVA

Deskriptory, redukce šumu, zlepšování obrazu, detekce oblastí zájmu, přiřazování klíčových oblastí, detekce tvarů

# KEY WORDS

Descriptors, noise reduction, image enhancement, detection of reagions of interrest, key regions matching, shape detection

# BIBLIOGRAFICKÁ CITACE

LABUDOVÁ, K. Zpracování obrazu pro ovládání robotické ruky. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, 2017. **x** s. Vedoucí bakalářské práce Ing. Vratislav Harabiš Ph.D.

# PROHLÁŠENÍ

Prohlašuji, že svoji seminární práci na téma Zpracování obrazu pro ovládání robotické ruky jsem vypracovala samostatně a pod vedením vedoucího seminární práce a s použitím odborné literatury a dalších informačních zdrojů, které jsou všechny citovány v práci a uvedeny v seznamu literatury na konci této práce. Jako autorka uvedené seminární práce dále prohlašuji, že v souvislosti s vytvořením této seminární práce jsem neporušila autorská práva třetích osob, zejména jsem nezasáhla nedovoleným způsobem do cizích autorských práv osobnostních a/nebo majetkových a jsem si plně vědoma následků porušení ustanovení S 11 a následujících autorského zákona č. 121/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon), ve znění pozdějších předpisů, včetně možných trestněprávních důsledků vyplývajících z ustanovení části druhé, hlavy VI. díl 4 Trestního zákoníku č. 40/2009 Sb.

V Brně dne 3.ledna 2017 ………………………..

podpis autora

# PODĚKOVÁNÍ

Děkuji vedoucímu bakalářské práce Ing. Vratislavu Harabišovi Ph.D. za účinnou metodickou, pedagogickou a odbornou pomoc a další cenné rady při zpracování mé seminární práce.

V Brně dne 3. ledna 2017 ………………………….

podpis autora

# OBSAH

Obsah

[ÚVOD 6](#_Toc468643995)

[1 PŘEDZPRACOVÁNÍ OBRAZU 7](#_Toc468643996)

[2 DETEKCE OBRAZOVEK 8](#_Toc468643997)

[2.1 Deskriptory 8](#_Toc468643998)

[2.1.1 SIFT 8](#_Toc468643999)

[2.1.2 SURF – Speeded Up Robust Features 12](#_Toc468644000)

[2.1.3 ORB 12](#_Toc468644001)

[2.1.4 BRIEF – Binary Robust Independent Elementary Features 12](#_Toc468644002)

[2.2 Srovnávání deskriptorů 12](#_Toc468644003)

[2.2.1 Klastrování pomocí Houghovy transformace 12](#_Toc468644004)

[2.2.2 Efektivní indexování nejbližšího souseda 12](#_Toc468644005)

[2.2.3 Vyhledávání podobných klíčových bodů (keypoint matching) 12](#_Toc468644006)

[3 SEZNAM LITERATURY 15](#_Toc468644007)

# SEZNAM OBRÁZKŮ

# ÚVOD

Tato práce se zabývá zpracováním obrazu pro účely navádění robotické ruky po dotykovém displeji embedded zařízení tiskáren. Tento systém s robotickou rukou pomáhá ve firmě Y Soft Corporation testovat software SafeQ určený pro tiskový management. Celé oddělení Research and Development funguje na bázi kontinuálního vývoje a proto je několik vybuildovaných verzí SafeQ týdně. Ideálně by se mělo provést regresní testování pro každou verzi a každého výrobce tiskáren, kterých máme v dnešní době 12, pro podchycení chyb vzniklých v průběhu vývoje.Toto je velmi náročné až nemožné splnit, kvůli časové náročnosti a počtu lidí, kterých by bylo potřeba. Navíc buildování se děje večer a nejlepší by bylo zjistit případné problémy už v noci, aby je vývojáři mohli začít opravovat už ráno. Robot je dobrým řešením, protože může pracovat přes noc, neunaví se a může provádět repetitivní úkony aniž by ztrácel pozornost jako lidští testeři.

Tento robot už je několik let ve fázi vývoje. Robotická ruka je naváděná kamerou, která snímá displej embedded zařízení. Softwarová část funguje na Windows servrech a je připojená k databázi, která je pro tuto práci důležitá. V databázi se nachází referenční snímky obrazovek, zajímavé regiony a text obražený ve všech obrazovkách. Referenční obrazovky jsou uložené ve stromové structuře podle kontextu (verze SafeQ, výrobce a model tiskárny). Dále se tam nachází informace o pozici ovládácích prvků a na jakou obrazovku má aplikace přejít po stisknutí ovládacího prvku. Část pro zpracování obrazu je naimplementovaná, ale neobsahuje žádné předzracování obrazu a není dostatečně rychlá ani spolehlivá. Náplň této práce je vybrat vhodné metody pro předzpracování obrazu a implementovat je. Dále nastudovat algoritmy pro rozpoznávání a klasifikaci obrazů, vybrat ty nejslibnější na základě odborné literatury a poté je implementovat a porovnat. Nejlepší metoda se použije v praxi. Na závěr se zhodnotí rychlost a spolehlivost řešení.

# PŘEDZPRACOVÁNÍ OBRAZU

## Normalizace obrazu

### Detekce rohů displeje

### Normalizace jasu

## Redukce šumu

# DETEKCE OBRAZOVEK

## Detekce zajímavých bodů v obraze

Harissův detektor rohu

* Není invariantní vůči změně měřítka

Harris-Laplace and Hessia-Laplace

* Je invariantní vůči změně měřítka
* Mikolajczyk a Smidt
* Vice stabilní a opakovatelné

Aproximace Laplacianů Gaussianů (LoG) pomocí rozdílu Gaussianů (DoG)

* Lowe
* Rychlejší

Salient region Detector

* Kadir a Brandy

## Deskriptory

1. Detection
2. Description
3. matching

### SIFT

* Jeden z nejvíce používaných, základ těchto metod
* Nevýhoda je vysoká dimenze dekriptoru

Je to deckriptor, který je invariantní vůči otočení a škálování. To znamená, že může srovnávat scény s řadou affiních transformací, změnou úhlu pohledu, osvělení nebo se zašuměním.

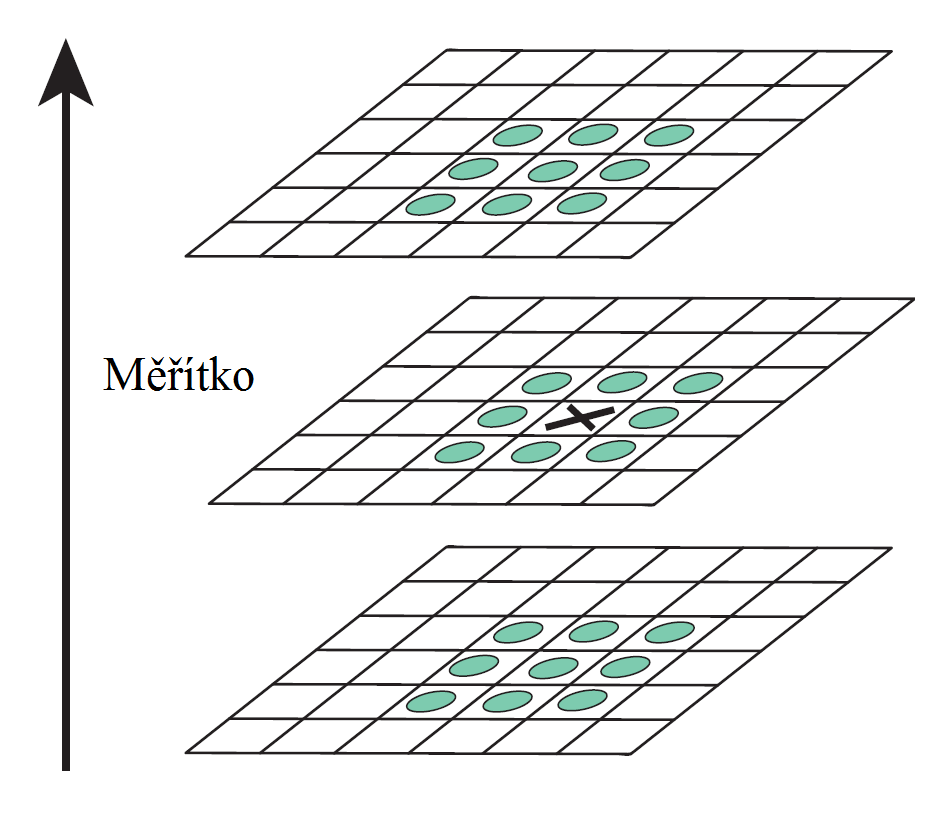
Generování příznaků probíhá v několika fázích:

#### Detekce místních extrémů

Nejdříve se vypočítá rozdíl dvou 2D Gaussových funkcí G(x,y,kσ)[1] s různým měřítkem *k* a získaný rozdíl se konvoluje originálním obrazem I(x,y). Tím vzniká rozdíl měřítkových prostorů původního obrazu D(x,y,σ).

|  |  |  |
| --- | --- | --- |
|  |  | (1) |

Tento postup se provádí pro několik různých měřítek, pak se obraz podvzorkuje a opakuje se stejný postup. Volba měřítka závisí našem uvážení. Menší měřítko zachytí více detailů a generuje vice bodů zájmu, je to ale vykoupené vyšší výpočetní náročností.



Obr. 1 - Okolí bodu [2]

Dalším krokem je srovnávání každého pixelu s okolím 3x3x3 tzn. s 26 sousedícími pixely. Maximum resp. minimum je deketováno, pokud má porovnávaný pixel nejvyšší resp. nejnižší hodnotu v daném okolí. Tyto body jsou dale vyhodnoceny jako spolehlivé příznaky.

#### Lokalizace bodů zájmu

Tyto příznaky se dále selektují pomocí rovnice, kde se počítá kontrast v bode extrému a málo kontrastní body, tzn. se vyřazují.

|  |  |  |
| --- | --- | --- |
|  |  | (2) |

Pro zvýšení stability se vyřazují i body, které pochází z hran, protože jsou nestabilní vůči šumu. Takovéto body se vyznačují charakteristickým zakřivením, kdy mají vysoké zakřivení kolmo na hranu a podél hrany výrazně nižší.

Nejprve se vypočítá determinant a stopa 2x2 Hessiánu , která odpovídá pozici a měřítku bodu zájmu.

|  |  |  |
| --- | --- | --- |
|  |  | *(3)* |

Derivace jsou aproximovány rozdílem sousedících bodů a bodu zájmu. Podle nerovnice *(4)* se vyřadí všechny body zájmu, které mají poměr zakřivení v hlavních osách vyšší než r = 10.

|  |  |  |
| --- | --- | --- |
|  |  | *(4)* |

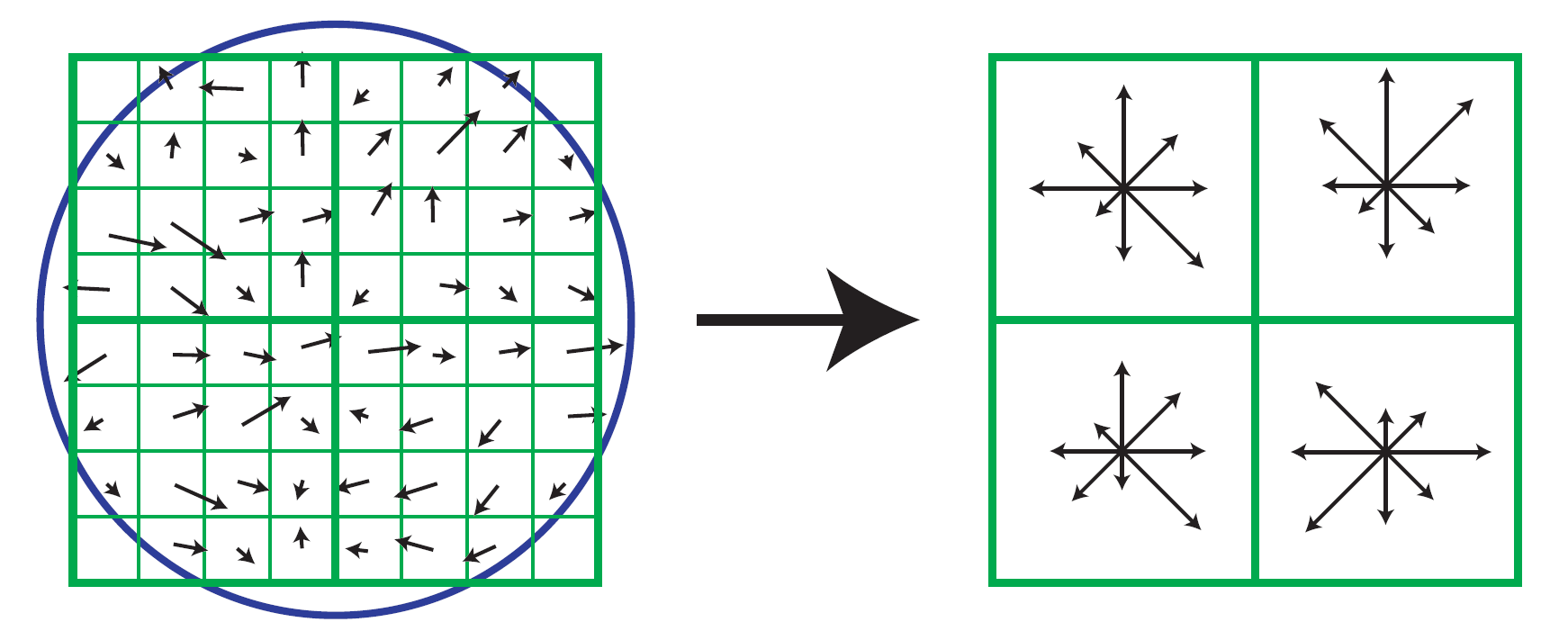
#### Přiřazení orientace

Dalším krokem je výpočet orientace, to zajišťuje invariantnost vůči otočení. Nevýhodou je vyřazení potenciálně hodnotných bodů zájmu, pokud jsou nekonzistentně otočené vůči zbytku obrazu. Pro každý pixel je vypočítána intenzita gradientu m(x,y) a jeho orientace podle následujících rovnic.

|  |  |  |
| --- | --- | --- |
|  |  | (5) |
|  |  | *(6)* |

Z orientací gradientů važených jejich intenzitou a kruhovým oknem váženým Gaussovskou funkcí je vypočítán histogram o 36 sloupcích, které odpovídají 360°. Píky v histogramu označují dominantní směry lokálních gradientů. Směr, který odpovídá maximu histogramu je označen jako hlavní směr bodu zájmu. Pokud se v histogramu nachází další píky, které jsou vysoké alespoň jako 80% maxima v histogramu, vytvoří se další body zájmu s danými orientacemi. Nakonec se na proloží 3 hodnoty v histogramu, které jsou nejblíže píkům k určení pozice píku pro lepší přesnost.

#### Výpočet deskriptorů



Obr. 2 - Tvorba deskriptoru z navzorkovaného pole gradientů v bodě zájmu, délka šipek označuje velikost lokálního gradientu, modrý kruh napravo označuje váhování Gaussovým kruhovým oknem [2]

Intenzity a orientace gradientů v bodech zájmu se navzorkují a rozmažou Gaussovskou funkcí o stejném měřítku jako je měřítko bodu zájmu. Pro zachování invariance vůči rotaci, jsou souřadnice deskriptoru a směr gradientu otočeny relativně ke směru bodu zájmu. Tyto hodnoty jsou pro větší efektivitu počítány dopředu pro všechny úrovně pyramidy. Dále jsou váhované Gaussovským kruhovým oknem se σ, která se rovná 1,5násobkem šířky deskriptoru. Tímto zabráníme velkých změn deskriptoru při malé změně pozice, protože dáme větší důraz na gradient ve středu okna a menší důraz na hodnoty dále od centra deskriptoru. Podle [2] je nejlepší vzorkovat polem 4x4 v bodu zájmu a dále postupovat histogramem o 8 sloupcích. Vektor příznaků je tvořen hodnotami ve sloupcích všech histogramů, které korespondují s délkou šipek v  Obr. 2, tzn., že jde o vektor 4x4x8 = 128 hodnot pro každý bod zájmu. Nakonec je vektor normalizován, což zajistí invariantnost vůči osvětlení, protože při změně osvětlení bude přičtena nebo odečtena konstanta pro všechny hodnot vektoru, které nabudou stejných hodnot po normalizaci.

---------------------------------------------------------------------------------------------------------------

Výběr nejlepší shody s obrazy v databázi je proveden na základě Euklidovské vzdálenosti. Kde se srovnávají nejbližší a druzí nejbližší sousedi.

Příznaky se srovnávají s velkou databází již vypočítaných příznaků. Příznaky jsou jedinečné v tom smyslu, že se správně a s vysokou pravděpodobností přiřadí již vypočítanými příznaky obrázků v databázi. Toto srovnávání se provádí výpočtem nejbližšího souseda, následovaný Houghovou transformací k identifikaci klastrů, které patří k jednomu objektu a nakonec verifikací pomocí metody nejmenších čtverců kvůli konzistenci pozičních parametrů. Tento přístup k rozpoznávání objektů je robustní a dosahuje rychlosti blížící se zpracování v reálném čase. Příznaky se nachází jak v místní, tak ve frekvenční doméně, tím se dosahuje robustnosti při velkém počtu objektů v obraze a take při částečném zakrytí objektu jinými. Extrahování příznaků je náročné na výpočet, proto se provádí kaskádové filtrování a nejnáročnější operace jsou počítány jen v regionech, které projdou inicializačním testem.

2) Lokalizace bodů zájmu - vybírají se na základě stability a je určena lokace a měřítko

3) Přiřazení orientace - Na základě místního gradientu se přiřadí orientace.

4) Deskriptor bodu zájmu - Kolem bodů zájmu je změřen gradient v určeném měřítku a dále je tranformovaný do reprezentace, která je invariantní vůči geometrické distorzi a změně osvětlení

To znamená, že transformuje data obrazu do souřadnic, které jsou invariantní vůči měřítku a souvisí s lokálními příznaky. Tento přístup generuje mnoho příznaků, které hustě pokrývají obraz přes velkou řadu měřítek a lokací.

Pro přiřazování obrazů se vypočítají vektor příznaků všech obrazů v databázi a pak se vypočítá Euklidovská vzdálenost k přiřazovanému obrazu pro každý z nich. K identifikaci správné a falešné shody se využívá hashovací tabulky Houghovy transformace. Každý klastr tří nebo více příznaků, které náleží jednomu objektu a pozici se pak dále verifikuje. Dále se vyhodnocuje chyba a odlehlé hodnoty se neberou v potaz. Nakonec se vypočítá pravděpobnost shody a pravděpodobnost falešně pozitivního výsledku.

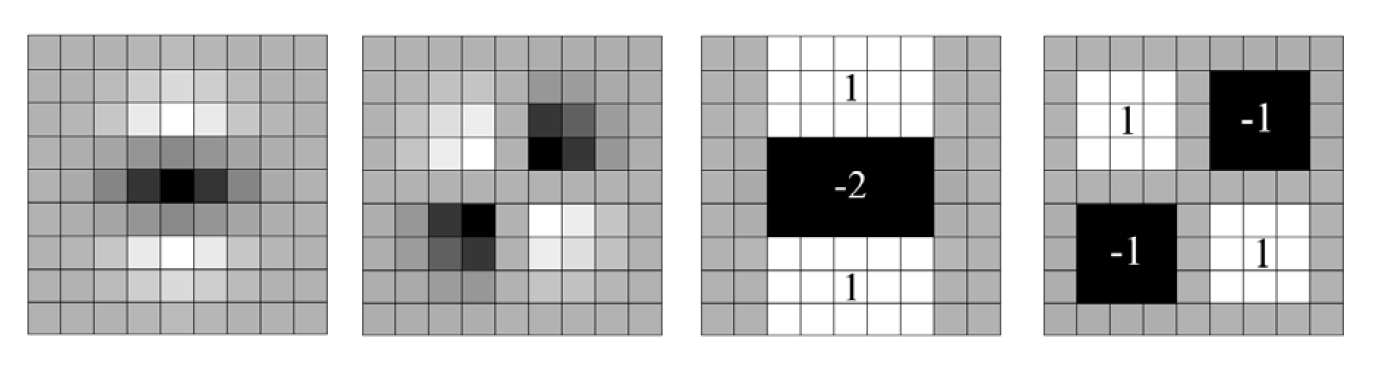
---------------------------------------------------------------------------------------------------------------

### SURF – Speeded Up Robust Features[3]

* Počítá se z integrálních obrázků. Detekce je podobná jako u SIFT, počítá se z DoG a nazývají ho Rychlý Hessian
* Deskriptor popisuje distribuci Haarových vlnek v okolí bodu zájmu
* Používá se 64 dimenzí – to redukuje čas pro srovnávání a hledání podobných deskriptorů
* Oproti SIFT je robustnější, protože používá jiné indexování

#### Rychlý Hessiánský detektor

* Myšlenka je konvolovat druhou parciální derivaci 2D Gaussiálnské funkce s původním obrazem. Dále tyto parciální derivace poskládat do Hessiánské matice [4] a z té vypočítat determinant. To by bylo ale výpočetně náročné a objevoval se aliasing vlivem diskretizace a ořezání Gaussovy funkce. Proto se v praxi využívají krabicové filtry, které aproximují druhé parciální derivace Gaussovy funkce. Jsou mnohem jednodušší a rychlejší na výpočet a neobjevují se u nich výše zmíněné chyby. Měřítkové prostory jsou implementovány jako obrazová pyramida. Obrazy jsou opakovaně vyhlazeny Gaussovským filterem a podvzorkovány. Také se může zvolit jiný přístup – místo podvzorkování se využije větší masky filtru. (začíná se na 9x9 a postupně zvyšuje na 15x15, 21x21, 27x27 atd.). Stejně se zvětšováním masky se pak snižuj frekvence vzorkování. K lokalizování bodů zájmu se používá potlačení bodů, v kterých není maximum v okolí 3x3x3. Maxima determinantů Hessiánské matice jsou pak interpolována v měřítku a v obrazovém prostoru. (Brown)



Obr. 3 – diskrétní a ořezaná druhá parciální derivace Gaussovy funkce Dyy, dále Dxy, krabicový filtr aproximující Dyy, krabicový filtr aproximující Dxy

#### Přiřazení orientace

Nejdříve se vypočítají odezvy na Haarovy vlnky[5] v ose x a y kolem bodů zájmu v okolí o průměru 6*s*, kde *s* je měřítko, v kterém byl bod zájmu detekován. Krok se kterým se vzorkuje okolí je s, stejně tak odezvy Haarových vlnek jsou ve stejném měřítku. Délka jedné Haarovy vlnky je 4*s.* Dále je výsledek vážený Gaussovým kruhovým oknem. Odezva je reprezentována dvěma vektory, jeden je síla odezvy v horizontální rovině, druhý je síla odezvy ve vertikální rovině (přes souřadnice). Dominantní orientace se počítá jako suma všech orientaci kolem bodu zájmu v posouvajícím se okně o úhlu π/3. Horizontální a vektikální vektor jsou sečteny a vzniká nový vektor. Takovýto nejdelší vektor určuje orientaci bodu zájmu.

#### Výpočet deskriptoru

První krok výpočtu deskriptoru je konstrukce čtverce, v jehož středu je bod zájmu a je stejně jako bod zájmu orientovaný. Velikost tohoto okna je 20*s*. Tento region je rozdělený do menších regionů 4x4. Tohle rozdělení zachovává důležité informace o pozici. Pro každý subregion se vypočítají příznaky a to následujícím způsobem. Vypočítají se odezvy na Haarovy vlnky v x a y rovině, které jsou relativně orientovány vůči bodu zájmu. Tyto odezvy jsou sesumovány v každém regionu zvlášť a tím jsou vytvořeny 2x4 příznaků. Dále sesumujeme všechny absolutní hodnoty odezev v každém subregionu, to je dalších 2x4 příznaků, dohromady tedy 64 příznakový vektor. Vlnkové odezvy jsou invariantní vůči osvětlení. Pro invariantnost vůči měnícímu se kontrastu musíme vektor převést na jednotkový.

#### Modifikace tohoto deskriptoru

U-SURF = Upright SURF

* Který není invariantní vůči otočení, tudíž se dá použít s výhodou jen na některé aplikace. Například na tu naši, kdy k otočení nemůže nikdy dojít, protože je obraz vždy normalizovaný a natočený stejným směrem. Díky tomu, že se úplně vynechává informace o orientaci bodu zájmu, je tento deskriptor rychlejší a robustnější, protože nemá falešně pozitivní výsledky, které vznikly hledáním a srovnávání deskriptorů ve všech směrech.

SURF-128

* Je přesnější, ale taky pomalejší na výpočet a následné srovnávání se snímky v databázi, protože má vyšší dimenzi. Dosahuje se tím, že místo výpočtu sumy absolutních hodnot odezev na Haarovy vlnky se sesumuje d < 0 a d ≥ 0.

Můj návrh

Zkombinovala bych přístup U-SURF a SURF-128.

### ORB

### BRIEF – Binary Robust Independent Elementary Features

BRIEF je efektivní deskritor vlastností bodů.

## Srovnávání deskriptorů

### Klastrování pomocí Houghovy transformace

### Efektivní indexování nejbližšího souseda

### Vyhledávání podobných klíčových bodů (keypoint matching)

# AKTUÁLNÍ ŘEŠENÍ

Samotné rozpoznávání obrazovek se děje ve třech fázích. Nejdříve se ověří, jestli se po stisku tlačítka objevila požadovaná obrazovka. Pokud je výsledek nepřesvědčivý , nastoupí analýza pomocí descriptorů. Pokud i ty selžou, provádí se analýza textu v obraze pomocí OCR a výsledek se srovnává s databází.

# VÝSLEDKY

# ZÁVĚR

# SEZNAM LITERATURY

[1] H. Guo, “A simple algorithm for fitting a gaussian function,” *IEEE Signal Process. Mag.*, vol. 28, no. 5, pp. 134–137, 2011.

[2] D. G. Lowe, “Distinctive image features from scale-invariant keypoints,” *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, Nov. 2004.

[3] H. Bay, T. Tuytelaars, and L. Van Gool, “SURF: Speeded up robust features,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 3951 LNCS, pp. 404–417, 2006.

[4] R. A. Horn and C. R. Johnson, *Matrix Analysis:*, 2nd ed. Cambridge: Cambridge University Press, 2012.

[5] T. Edwards, “Discrete Wavelet Transforms : Theory and Implementation 1 Introduction 2 A Brief Discussion of Wavelets,” *Computer (Long. Beach. Calif).*, no. September, pp. 1–27, 1991.

# SEZNAM ZKRATEK A PŘÍLOH

## Seznam zkratek

LoG – Laplacian of Gaussian

DoG – Difference of Gaussian

SIFT

SURF - Speeded-Up Robust Features

## Seznam příloh