



VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
FAKULTA INFORMAČNÍCH TECHNOLOGIÍ

Ročníkový projekt

2005/2006

Tomáš Loyka

Proděkování

Děkuji Doc. Ing. Františku Zbořilovi, CSc. nejen za odborné vedení, rady a podněty, které mi během práce poskytoval, ale také za jeho vstřícnost při odborné pomoci a poskytování informačních zdrojů.

Prohlášení

Prohlašuji, že jsem tento ročníkový projekt vypracoval samostatně pod vedením Doc. Ing. Františka Zbořila, CSc.

Uvedl jsem všechny literární prameny a publikace, ze kterých jsem čerpal.

.....
Tomáš Loyka

Abstrakt

Projekt se zabývá technikami počítačového vidění a rozpoznávání obrazu. Smyslem je určit polohu robota v laboratoři.

Klíčová slova

počítačové vidění, zpracování obrazu, digitalizace obrazu, předzpracování, segmentace, popis objektů, klasifikace

Abstract

This project is engaged in machine vision and image processing and analysis. The reason is to create a application to determine position of the robot in the laboratory.

Key words

machina vision, image processing, image digitization, pre-processing, segmentation, shape representation and description, image understanding

Obsah

Úvod	5
1 Snímání a digitalizace obrazu.....	5
1.1 Snímání.....	5
1.2 Digitalizace.....	6
1.2.1 Popis barev v obraze.....	8
2 Předzpracování obrazu	9
2.1 Jasové transformace.....	9
2.2 Geometrické transformace.....	11
2.3 Filtrace a ostření	12
3 Segmentace	12
3.1 Prahování	13
3.2 Detekce hran	15
3.3 Srovnávání se vzorem.....	17
4 Popis objektů	17
4.1 Identifikace oblastí	18
4.2 Popis objektů založený na popisu hranic.....	19
4.3 Popis objektů založený na popisu plochy.....	20
5 Klasifikace objektů.....	22
6 Návrh programu.....	22
6.1 Digitalizace, formát obrazu	23
6.2 Předzpracování obrazu	23
6.3 Segmentace obrazu, klasifikace objektů.....	25
7 Závěr, zhodnocení výsledků.....	27
8 Literatura	28
9 Přílohy	28

Úvod

Problematika rozpoznávání objektů je částí disciplíny zvané počítačové vidění. Počítačové vidění je disciplína, která se snaží technickými prostředky napodobit lidské vidění. Pro počítačové vidění je typická snaha porozumět obecné trojrozměrné scéně, např. takové, jakou zahlédnete při pohledu z okna. Postupy počítačového vidění jsou značně složité, s těžištěm v interpretaci obrazových dat, která jsou nejčastěji reprezentována symbolicky. Jádrem pokročilejších postupů jsou znalostní systémy a techniky umělé inteligence. Této části počítačového vidění se říká vyšší úroveň, první částí je nižší úroveň.

Cílem nižší úrovně je analyzovat vstupní dvojrozměrná obrazová data číselného charakteru a najít kvalitativní symbolickou informaci potřebnou pro vyšší úroveň. Pro nižší úroveň se také používá název zpracování obrazu počítačem.

Předmětem zpracování a případné rozpoznání obrazu je obrazová informace o reálném světě, která do počítače vstupuje televizní kamerou, fotoaparátem, skenerem. Počítačové vidění řeší úlohu vytvoření explicitního popisu fyzikálních objektů v obraze.

Vlastní průběh zpracování a rozpoznávání obrazu reálného světa obvykle rozdělujeme do několika základních kroků. Toto rozdělení postupu však není zcela jednoznačné, záleží již na dané aplikaci, zda budou provedeny všechny kroky tak, jak jsou zde uvedeny.

Jednotlivé kroky jsou:

- Snímání a digitalizace obrazu
- Předzpracování
- Segmentace obrazu
- Popis objektů
- Klasifikace

1 Snímání a digitalizace obrazu

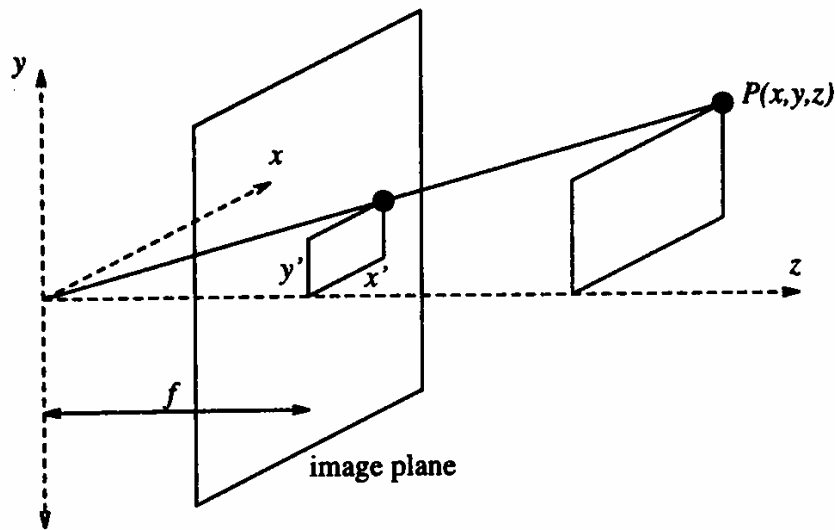
Základním kamenem pro zpracování a rozpoznání obrazu je vlastní získání obrazu reálného světa, jeho převod do digitální podoby vhodné pro uložení a další zpracování v počítači.

1.1 Snímání

Snímání obrazu je převod optické veličiny na elektrický signál, který je spojitý v čase i v úrovni. Vstupní informací pro snímání nemusí být vždy jen jas z kamery či scanneru, ale mohou to být i jiné veličiny, jako jsou tepelné záření, rentgenové záření či ultrazvuk. Snímání dále může probíhat buď v jednom nebo v několika spektrálních pásmech. Pro barevné snímání postačí složky tři: červená, zelená a modrá.

Na výsledný sejmутý obraz má vliv mnoho různých faktorů. Může to být například osvětlení snímaného objektu, jeho vlastnosti, odrazivost, a další.

Snímáním získáme 2D obraz dané scény, který je perspektivní projekcí 3D prostředí které snímáme. Geometrie perspektivního zobrazení je ukázána na obrázku.



obr.: perspektivní zobrazení

Necht' x , y a z jsou souřadnice bodu P ve 3D scéně a f je ohnisková vzdálenost. Potom bod P má po perspektivním zobrazení souřadnice:

$$x' = \frac{xf}{z}, \quad y' = \frac{yf}{z}$$

Tento model je velice realistický, a vlastně odpovídá obrazu který vzniká na sítnici lidského oka. Mnoho informací se však ztratí, a následná rekonstrukce 3D scéně z 2D nasnímaného obrazu je velice obtížná, a možná jen za předpokladu, že o snímané scéně máme dostatečný počet informací. Při snímání a rekonstrukci 3D scéně se však většinou používají jiné metody, které jsou založeny na vyšším počtu snímacích zařízení. Řešení úloh, které mají 2D povahu je tudíž mnohem jednodušší, a ve velké většině případů i dostatečné.

1.2 Digitalizace

Druhým krokem při získávání obrazu vhodného pro další zpracování v počítači je převod spojitého analogového signálu odpovídajícího monochromatickému signálu na signál digitální – digitalizace.

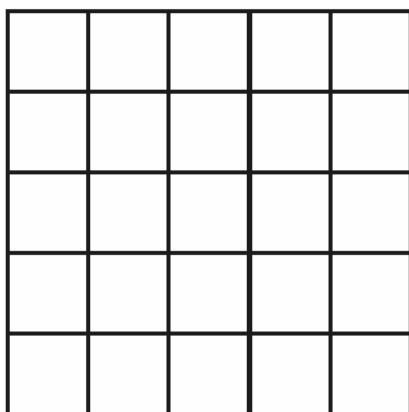
Digitální obraz je ekvivalentem spojitě obrazové funkce $f(i,j)$ kde i a j jsou souřadnice v obraze. Je získán vzorkováním obrazu do matice $M \times N$ bodů a kvantováním do K úrovní jasu. Jednomu prvku matice se říká obrazový element (picture element – pixel). Z hlediska zpracování obrazu jde o dále nedělitelnou jednotku.

Vzorkování se dále řídí známou Shannonovou větou. Z té plyne, že nejmenší detail v digitálním obraze musí být minimálně dvojnásobkem vzorkovacího intervalu. Volba vhodného rozlišení je tedy jedním z nejdůležitějších kroků digitalizace. Při nízkém rozlišení budeme ztrácet informace o detailech v obraze které bychom mohli dále potřebovat, a naopak při vysokém rozlišení nám bude stoupat výpočetní náročnost a tedy i spotřebovaný čas na další zpracování obrazu.

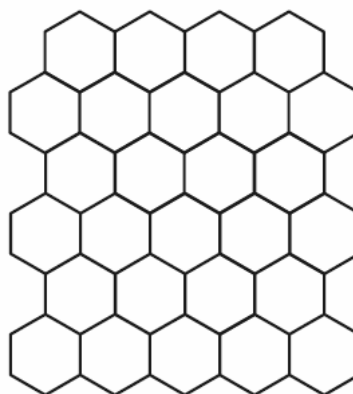
Velikost obrazu se většinou udává v pixelech, rozlišení obrazu je obvykle udáváno v jednotkách body/palec (dots per inch – DPI).

U kvantování úrovní jasu volíme jejich počet tak, aby nedocházelo k falešným obrysům. Tento efekt nastává pokud kvantovaný počet úrovní jasu je nižší než dokáže člověk rozlišit. Toto číslo je závislé na mnoha okolnostech, například průměrná světlost, ale od počtu úrovní 100 a více už k tomuto jevu nedochází.

Další důležitou součástí digitalizace je volba vzorkovací mřížky. Nejčastěji používanými mřížkami je mřížka čtvercová a hexagonální. Jejich struktura je naznačena na obrázku.



obr.: čtvercová mřížka

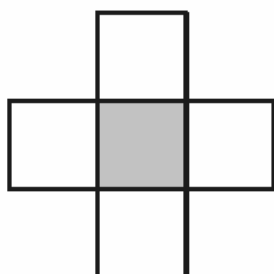


hexagonální mřížka

Čtvercová mřížka vychází z konstrukce většiny snímacích prvků, a ze struktury paměti a je tudíž velice snadno realizovatelná. Její nevýhody se však projeví při měření vzdálenosti mezi jednotlivými body, a také při zjišťování spojitosti objektů. Hexagonální mřížka tyto problémy řeší, avšak je obtížněji realizovatelná a pro některé operace i méně vhodná. Dále budeme pracovat jen se čtvercovou vzorkovací mřížkou. Zavedeme si tedy vzdálenost mezi dvěma body (i,j) a (h,k) v čtvercové mřížce. Obecně je tato vzdálenost rovna Euklidovské vzdálenosti D_E , definované podle vztahu:

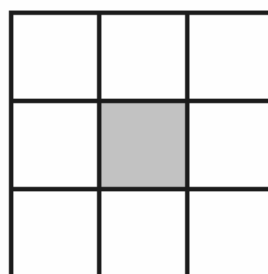
$$D_E = \sqrt{(i-h)^2 + (j-k)^2}$$

Pokud je ale prostor diskretizován, tak definujeme nejdříve sousedství bodu a pak vzdálenost bodů v obraze za předpokladu daného sousedství. Při práci se čtvercovou mřížkou můžeme uvažovat buď 4-sousedství, nebo 8-sousedství:



obr.:

4-sousedství



8-sousedství

Vzdálenost dvou bodů je pak dána podle následujících vztahů:

$$D_4 = |i - h| + |j - k|$$

$$D_8 = \max\{|i - h|, |j - k|\}$$

1.2.1 Popis barev v obraze

Barva obrazu je důležitým nositelem informace. Pro určení přesného odstínu barvy je vytvořeno několik různých barevných modelů. Ty se liší podle toho, které složky jsou při popisu barvy použity.

1.2.1.1 Model RGB

Tento barevný model je využíván v zobrazovacích zařízeních (monitor, TV), kde jeden viditelný bod (pixel) tvoří tři velmi blízko umístěné body barevných složek Red, Green, Blue (červená, zelená, modrá). Jejich sloučením vznikají další barvy různých odstínů.

Tento model je aditivní, to znamená že pro hodnoty RGB (0,0,0) je výsledná barva černá. Se zvyšující hodnotou se jas zvyšuje, až po maximální hodnotu (1,1,1) kdy je výsledná barva bílá. Pokud jsou hodnoty jednotlivých složek stejné, jedná se o odstín šedé, pokud jsou hodnoty složek rozdílné, dostáváme všechny barvy spektra.

V počítačových aplikacích se rozsah jednotlivých hodnot určuje v celých číslech, většinou v rozmezí 0-255. Toto rozmezí se nazývá také barevná hloubka.

Barevná hloubka 8 bitů je tvořena tabulkou 256 definovaných barev. Dále se používá barevná hloubka 24 bitů, tedy 256 hodnot pro každou složku RGB. Tato barevná hloubka se nazývá také TrueColor, protože je možno zobrazit již všechny odstíny barev které je možné vidět ve skutečném světě.

1.2.1.2 Modely CMY a CMYK

Tento barevný model je využíván pro tisk a při výrobě fotografií. Na rozdíl od předchozího modelu se jedná o subtraktivní systém. Barvy se neskládají, ale odečítají od původní bílé. Se zvyšující hodnotou dané složky je tedy celkový jas nižší.

Vztah mezi RGB a CMY (Cyan, Magenta, Yellow – azurová, purpurová, žlutá):

$$\begin{bmatrix} C \\ M \\ Y \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

Barva (1,0,0) v RGB je tedy ekvivalentní barvě (0,1,1) v CMY.

Model CMYK doplňuje původní složky CMY ještě o složku K (black - černá). Tento model se používá často v tiskárnách, jinak by bylo nutné nanést všechny tři složky CMY a výsledná barva by byla spíše šedá.

1.2.1.3 Model HSI

Na rozdíl od předchozích modelů nejsou jeho složky tvořeny jednotlivými základními barvami, ale jejich třemi vlastnosti: Hue, Saturation, Intensity (barva, sytost, jas).

Barva je udávána jako úhel v rozmezí 0-360°. Barvy tvoří uzavřený kruh a jsou lineární. Lze tedy zadat libovolný zlomek úhlu pro určení požadované barvy.

Sytost určuje množství přidané bílé složky.

Jas určuje kolik světla barva odrazí, tedy jak zářivá bude.

Tento model má výhodu v tom že zohledňuje lidské vnímání barev, používá se tedy v některých algoritmech při digitálním zpracování obrazu.

1.2.1.4 Modely YUV, YIQ

Tyto barevné modely se při zpracování obrazu nepoužívají. Jejich využití je v televizním vysílání, kdy složka Y představuje jas nebo šedost, a U a V (I a Q) jsou barevné složky.

1.2.1.5 Šedá škála

Někdy vystačíme s obrazem v šedé škále. To znamená rozmezí bílá, stupně šedé, černá. Většinou se používá 256 odstínů šedi, tedy barevná hloubka 8 bitů.

2 Předzpracování obrazu

Po úspěšném získání obrazu a jeho digitalizaci máme k dispozici digitální obraz zkoumané scény. Tento obraz však může být zkreslen díky způsobu snímání nebo nevhodných podmínkách při jeho získání. Pokud je znám charakter zkreslení, je možné tuto chybu opravit pomocí korekcí, které jsou jednou z metod předzpracování obrazu. Existuje také však velké množství dalších metod, které usnadňují analýzu obrazu, identifikaci objektů, nebo jen zvýrazňují důležité rysy obrazu pro snazší pozorování člověkem.

Předzpracování je společný název pro operace na nízké úrovni abstrakce, kdy jsou vstupem i výstupem matice které představují digitální obrazovou funkci.

Je důležité si uvědomit, že v průběhu předzpracování nezískáme žádnou novou informaci. Můžeme jen nějakou informaci potlačit nebo zvýraznit. Metody předzpracování využívají značné nadbytečnosti údajů v obraze, po provedení transformace se hodnota informace v obraze obsažené většinou sníží.

Základní rozdělení metod předzpracování obrazu:

- jasové transformace
- geometrické transformace
- filtrace a ostření

2.1 Jasové transformace

Jasové transformace lze rozdělit do dvou skupin. Jsou to jasové korekce a modifikace jasové stupnice. U jasových korekcí závisí jas v bodě výstupního obrazu pouze na jasu bodu ve vstupním obraze se stejnými souřadnicemi. Modifikace jasové stupnice transformuje určitou hodnotu jasu ve vstupním obraze na jinou výstupní hodnotu, a to bez ohledu na pozici v obraze.

První skupinou jsou jasové korekce. V ideálním případě by snímací a digitalizační zařízení mělo mít ve všech místech stejnou citlivost bez ohledu na umístění bodu v obraze. Tento požadavek však v praxi není často splněn. Světlo procházející dál od optické osy je více

zeslabováno, také světlocitlivá vrstva snímacího senzoru nemusí mít ve všech místech stejnou citlivost. Zdrojem odchylek je i nerovnoměrné osvětlení snímané scény.

Jsou-li uvedené odchylky systematické, lze je potlačit jasovými korekcemi.

Multiplikační koeficient $e(i,j)$ popisuje odchylku každého bodu obrazu od ideální převodní charakteristiky. Pro každý bod původního (neporušeného) obrazu $g(i,j)$ získáme na výstupu zkreslený bod $f(i,j)$:

$$f(i, j) = e(i, j)g(i, j)$$

K určení chybového koeficientu $e(i,j)$ stačí při stálých snímacích podmínkách sejmout obraz $g(i,j)$ o známém průběhu jasové funkce. Nejjednodušší způsob je sejmout obraz o konstantním jasu c , který po digitalizaci označíme $f_c(i,j)$. Potom můžeme systematické chyby v digitalizovaném obrazu jednoduše korigovat podle vztahu

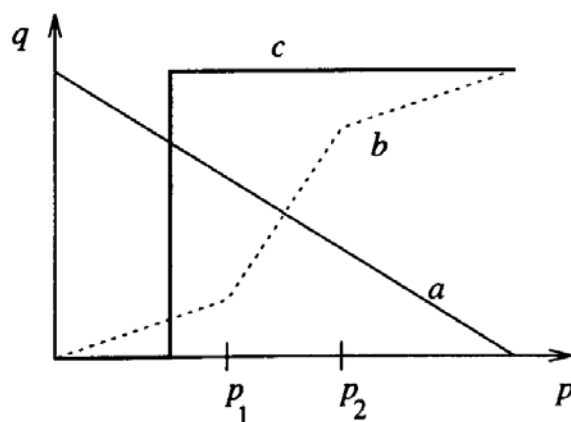
$$g(i, j) = \frac{f(i, j)}{e(i, j)} = \frac{c \cdot f(i, j)}{f_c(i, j)}$$

Tato metoda může být použita jen pokud jsou snímací podmínky stálé. Pokud chceme korigovat tuto chybu při měnících se snímacích podmínkách, musíme čas od času kalibrovat chybovou matici.

Druhou skupinou je modifikace jasové stupnice. Tato transformace nezávisí na poloze bodu v obraze. Transformace T výchozí hodnoty jasu p na novou hodnotu jasu q je dána vztahem

$$q = T(p)$$

Obvyklé transformace stupnice jasu jsou vyznačeny na obrázku. Přímka a označuje negativní transformaci, po částech lineární průběh funkce b zvyšuje kontrast mezi hodnotami jasu p_1 a p_2 . Funkce c se nazývá prahování a jeho výsledkem je černo-bílý obrázek.



obr.: obvyklé transformace stupnice jasu: a) negativ, b) zvýšení kontrastu c) prahování

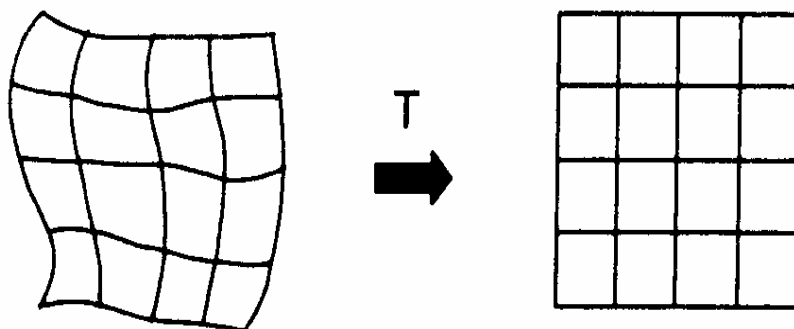
Digitalizovaný obraz má velmi omezené množství úrovní jasu, tudíž transformace jasové stupnice jsou velice snadno realizovatelné. Je vytvořena vyhledávací tabulka (look-up table) o tolika paměťových místech, kolik je jasových úrovní. Nové hodnoty jasu, které jsou výsledkem transformace, jsou pak přímo obsahem paměti. Vyhledávací tabulka může být zařazena do cesty obrazovému signálu, a tudíž je možné pozorovat výsledek transformace v reálném čase.

Stejný princip může být použit i pro barevné obrazy, kdy jsou vytvořeny tři vyhledávací tabulky pro jednotlivé barevné složky R, G, B.

Hodnoty jasové transformace pro zvýšení kontrastu jsou často nalezeny automaticky, podle analýzy histogramu.

2.2 Geometrické transformace

Geometrické transformace najdou uplatnění pokud chceme provést korekci geometrického zkreslení nebo při posunutí, rotaci, například při zpracování dokumentů OCR programem.



obr.: geometrická transformace v rovině

Geometrická transformace plošného obrazu je vektorová funkce T která transformuje bod (x, y) v rovině na novou pozici (x', y') . Transformaci plošného obrazu ilustruje obrázek. T je definována vztahy

$$x' = T_x(x, y) \quad y' = T_y(x, y)$$

Transformační vztahy T_x a T_y mohou být známy buď předem, jako je tomu například v případě, rotace, posunu, zvětšení, nebo je možné hledat transformační vztah na základě znalosti původního i transformovaného obrazu. Při hledání vztahu se obvykle využívá několika bodů které si odpovídají a lze je snadno najít v obou obrazech.

Geometrická transformace se skládá ze dvou kroků. První je plošná transformace bodu, která najde vstupnímu bodu s diskrétními souřadnicemi odpovídající výstupní bod. Souřadnice výstupního bodu musí být počítány jako spojité, protože po poloha výstupního transformaci obvykle nesouhlasí s celočíselnou mřížkou. Druhým krokem je nalezení úrovně jasu v konkrétním bodě mřížky, který se obvykle interpoluje z několika bodů v okolí, které po provedení transformace nemají celočíselné souřadnice. Toto okolí bývá většinou malé, vzhledem k výpočetním náročnosti. Tři nejpoužívanější metody jsou metoda nejbližšího souseda, která je nejjednodušší a nejrychlejší, dále potom lineární interpolace a bi-kubická interpolace.

2.3 Filtrace a ostření

Tyto metody využívají pro výpočet jasů bodu ve výstupním obraze jen lokálního okolí odpovídajícího bodu ve vstupním obraze. Podle cíle zpracování můžeme lokální předzpracování rozdělit do dvou velkých skupin: vyhlazování a gradientní operace.

Vyhlazování obrazu vede k potlačení vyšších frekvencí obrazové funkce. Žádaným výsledkem vyhlazování je tedy potlačení náhodného šumu a drobných odchylek. Bohužel vyhlazování také potlačuje ostatní náhlé změny jasové funkce, dochází tedy k oslabení ostrých čar a hran.

Gradientní operace a ostření obrazu naopak vedou ke zdůraznění vyšších frekvencí. Výsledkem je tedy zvýraznění těch částí obrazu kde se jasová funkce náhle mění, dochází tedy ke zvýraznění hran. Nežádoucím výsledkem je současné zvýraznění šumu v obraze.

Vyhlazovací a gradientní operace mají tedy protichůdné cíle. Řešením tohoto rozporu jsou některé algoritmy, které oba postupy kombinují, a provádějí tedy vyhlazení obrazu a detekci hran současně.

Podle vlastností funkčního vztahu pro výpočet výsledného jasů bodu na základě okolí ve vstupním obraze, dále metody předzpracování dělíme na lineární a nelineární. Lineární metody počítají výsledný jas v bodě $g(i,j)$ jako lineární kombinaci jasů v lokálním okolí bodu $f(i,j)$ vstupního obrazu. Váha bodů v lokálním okolí f je určena koeficientem h :

$$f(i, j) = \sum_{(m,n) \in O} h(i - m, j - n)g(m, n)$$

Na tento vztah lze pohlížet také jako na diskretní konvoluci s konvolučním jádrem h . V mnoha praktických případech se využívá pravoúhlého okolí O . Necht' je okolí O například obdélníkové o M řádcích a N sloupcích. Aby bylo použité okolí symetrické vůči středovému bodu $f(i,j)$ volí se nejčastěji rozměry M, N z množiny přirozených lichých čísel.

Lokální operace předzpracování obvykle jen málo využívají předběžné znalosti o povaze obrazu. Vzhledem k malému okolí zpracovávaného bodu je velice obtížné si tuto informaci odvodit i při vlastním průběhu zpracování. Zvláště vyhlazovací operace je vhodné znát charakter poruch v obraze, což mohou být například statistické parametry šumu, a podle nich přizpůsobit zpracování.

Volba lokální transformace, velikosti a tvaru příslušného okolí zpracovávaného bodu, závisí na velikosti objektů v obraze. Jsou-li objekty poměrně velké, obraz může být vylepšen vyhlazením malých nespojitostí.

Konvoluční lineární operace (filtry) jsou užitečné pro vyhlazování, detekci hran nebo čar. Lze poměrně jednoduše vytvořit specializované technické prostředky pro zrychlení výpočtu těchto operací. Zrychlení je jasně patrné pro speciální tvary filtrů. Příkladem jsou rekurzivní a separabilní filtry.

3 Segmentace

Jedním z nejtěžších kroků zpracování obrazu je segmentace obrazu. Jedná se o analýzu obrazu vedoucí k nalezení objektů v obraze. Za objekty se zde považují části obrazu, které jsou bodem zájmu v průběhu dalšího zpracování. Cílem segmentace je tedy rozdělení obrazu do částí odpovídající předmětům či oblastem z reálného světa. Výsledkem segmentace by měl být soubor oblastí, které odpovídají objektům ve vstupním obraze. Jedná se pak o tzv. kompletní segmentaci.

Pokud ale oblasti neodpovídají přesně objektům, tak tuto segmentaci nazýváme částečnou. Kompletní segmentace obecně využívá vyšší úroveň zpracování, která je založena na znalostech řešeného problému. Částečná segmentace je založena na principu homogenity obrazových vlastností (např. jas, barva) uvnitř segmentu. Přesto existuje třída úloh, jejichž řešení umožňuje z dobrými výsledky i nižší úroveň zpracování. Častá je situace, kdy je obraz tvořen kontrastními objekty na pozadí neměnného jasu (krevní buňky, psaný text,...). Tam lze využít jednoduché globální postupy a dosáhnout kompletní segmentace obrazu na objekty a pozadí. Takové postupy nezávisí na obsahu, neuvžívá se žádný model zpracovávané oblasti, k řízení procesu segmentace nepřispívají znalosti výsledné oblasti.

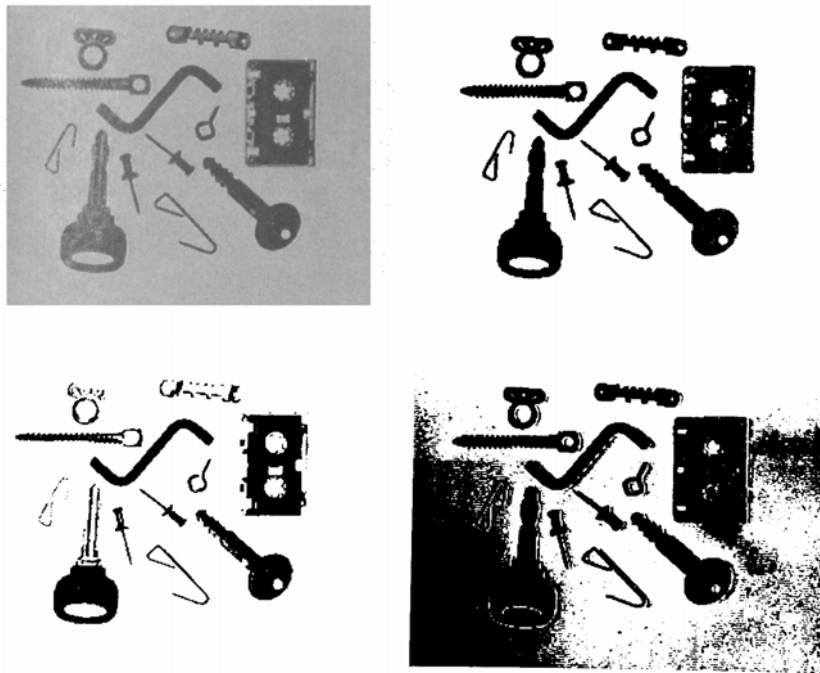
Při částečné segmentaci je výsledkem rozdělení obrazu do oddělených částí, které jsou homogenní vzhledem k zvolené vlastnosti jako je jas, barva, odrazivost, textura, atd. Po zpracování složitěho obrazu, jako je například letecká fotografie, je výsledkem seznam oblastí, které jsou homogenní v jistých rysech. Tyto oblasti se mohou překrývat. Částečně segmentovaný obraz musí být podroben dalšímu zpracování, které pomocí vyšší úrovně umožní získat výslednou segmentaci obrazu.

Úplně správná a kompletní segmentace složitých obrazů není v této fázi zpracování dosažitelná. Rozumným cílem je získat částečnou segmentaci, která je použita pro další zpracování operacemi vyšších úrovní. Okamžitým přínosem segmentace je výrazná redukce objemu zpracovávaných dat.

Jedním z hlavních problémů ovlivňujících segmentaci je nejednoznačnost obrazových dat, často doprovázená též informačním šumem. Segmentační metody lze rozdělit do tří skupin, a to podle dominantní vlastnosti které je pro segmentaci využíváno. V první z nich je využíváno globálních znalostí obrazu reprezentovaných obvykle histogramem určitých vlastností. Druhá využívá detekci hran, třetí detekci oblastí. Druhá a třetí skupina tohoto třídění řeší duální problém, neboť každá oblast je jednoznačně reprezentována svou hranicí a každá uzavřená hranice jednoznačně vypovídá o oblasti kterou určuje. V důsledku odlišného charakteru algoritmů tvorby hranic a oblastí přinášejí obě skupiny metod poněkud rozdílné segmentační výsledky, jinou informaci. Výsledky segmentačních postupů obou skupin lze proto kombinovat a vytvořit jedinou popisnou strukturu. Příkladem může být relační struktura, kde uzlům odpovídají oblasti, hranám relace sousednosti pomocí detekovaných hranic.

3.1 Prahování

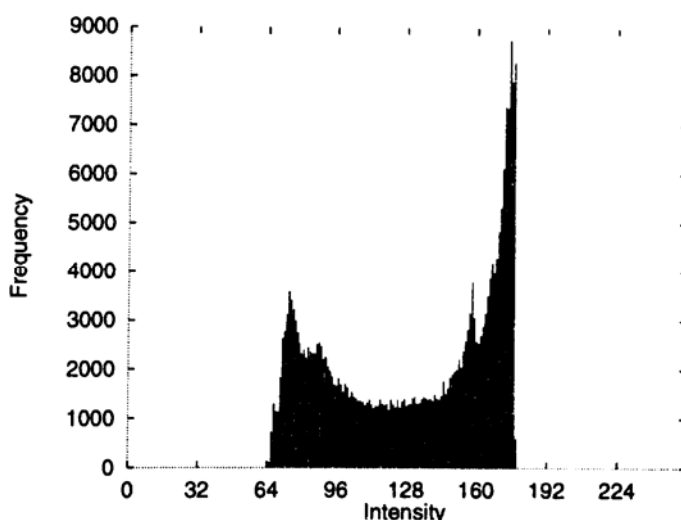
Prahování je nejjednodušším segmentačním postupem. Vychází ze skutečnosti, že mnoho objektů či oblastí obrazu je charakterizováno konstantní odrazivostí či pohltivostí svého povrchu. Pak je možno využít určené jasové konstanty – prahu, k oddělení objektů od pozadí. Prahování je nejstarší segmentační metodou a v jednoduchých případech je i nadále široce užíváno. Vzhledem k výpočetní nenáročnosti a jednoduchosti jej lze provádět v reálném čase.



obr.: prahování obrazu: původní obraz; segmentace prahováním; práh zvolen příliš malý; práh zvolen příliš velký

Nejjednodušší situace nastává, je-li předem známa vlastnost kterou má obraz po segmentaci mít. Práh určujeme tak, abychom tuto vlastnost segmentací splnili. Úlohou daného typu je segmentace stránky tištěného textu na znaky a pozadí, pokud víme, že písmena pokrývají $1/p$ plochy. S využitím znalosti poměru ploch objektů a pozadí je snadné určit (na základě histogramu) takovou hodnotu prahu T , aby právě $1/p$ plochy mělo úroveň jasu menší než T a zbytek mělo hodnotu jasu větší než T . Tato metoda se nazývá procentní prahování. Bohužel většinou informaci o procentním poměru objektů a pozadí nemáme. Tato informace může být někdy nahrazena jinou vlastností, známe například průměrnou šířku čar znaků. Práh můžeme zvolit tak, aby šířka čar v segmentovaném obraze odpovídala naší znalosti.

Složitější skupina metod je založena na analýze tvaru histogramu. Obsahuje-li obraz objekty přibližně téhož jasu které se odlišují od jasu pozadí, je výsledný histogram jasu dvouvrcholový (bi-modální). Jeden vrchol odpovídá jasu objektů, druhý vrchol odpovídá jasu pozadí.



obr.: dvouvrcholový histogram

Z tvaru histogramu vyplívá, že hodnoty jasu ležící mezi oběma vrcholy nejsou v obraze příliš časté, a pravděpodobně odpovídají jasu hraničních bodů mezi objekty a pozadím. Výsledný práh by měl splňovat požadavek minimální segmentační chyby. Je proto logické určit práh pro oddělení objektů od pozadí jako hodnotu jasu, jejíž četnost je minimem ležícím mezi dvěma výše uvedenými maximy. V případě, že histogram je vícevrcholový (multi-modální), může být určeno více hodnot prahu, a to vždy v minimech mezi dvěma maximy. Pro každý práh je pak výsledek segmentace samozřejmě jiný.

3.2 Detekce hran

Velkou skupinu segmentačních metod tvoří metody založené na informaci o hranách obrazu. Toto byl jeden z prvních způsobů segmentace obrazu a stále zůstává jedním z nejdůležitějších. Segmentace na základě detekce hran vychází ze skutečnosti, že hranice oblastí obrazu se sestávají z hran, které jsou v obraze nalezeny aplikací některého z hranových operátorů. Tyto nalezené hrany označují místa obrazu, v nichž dochází k nespojitosti v hodnotě jasu, barvy, nebo v textuře. Obraz, který vznikl aplikací hranového operátoru, je ale jako segmentační výstup nepoužitelný. Proto musí po detekci hran následovat další zpracování obrazu hran, které spojuje hrany do řetězců lépe odpovídající průběhu hranic. Cílem je dosáhnout alespoň částečné segmentace obrazu, tzn. vytvořit z velkého množství hran obraz, v němž se budou vyskytovat jen takové hrany, jimž v původním obraze odpovídají hranice objektů nebo částí obrazu.

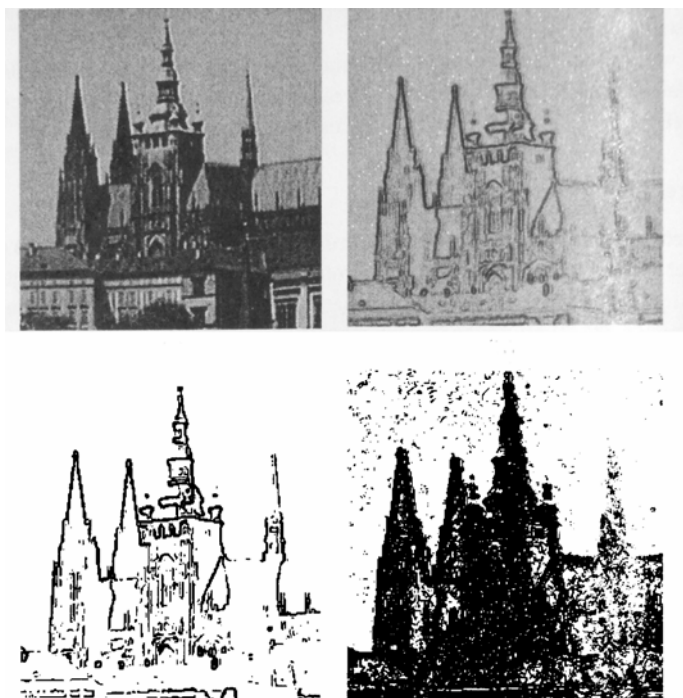
Jednotlivé segmentační metody se liší jednak strategií konstrukce výsledné hranice z hran, jednak množstvím předběžných informací, znalostí o zkoumaném obraze které je možné v konkrétní segmentační úloze využít. Čím více informací je při segmentaci k dispozici, tím lepší segmentační výsledky lze očekávat. Pomocí těchto informací lze také vyhodnotit věrohodnost výsledné získané segmentace. Tyto předběžné znalosti mají také vliv na vlastní průběh segmentace. Pokud je k dispozici velké množství znalostí o požadovaném výsledku, jsou jak

celkový tvar hranice, tak i další vztahy k ostatním objektům v obraze určeny velice přísně a segmentační metoda musí všechny tyto dané podmínky dodržet.

Naopak, pokud je málo informací o výsledné hranici, musí segmentační metoda brát v úvahu lokální vlastnosti obrazu spolu s obecnými znalostmi specifickými pro určitou aplikační oblast. Při malém množství předběžných znalostí také segmentační metody nemohou využívat hodnocení správnosti generované hranice a na jeho výsledku provádět zpětnovazební korekci segmentačního výsledku.

Nejčastějším problémem, se kterým se musí hranové segmentační metody vypořádat, způsobený šumem nebo neužitečnou informací, je výskyt hran v místech bez přítomnosti skutečné hranice a současně absence hran tam, kde hranice ve skutečnosti probíhá. Je jasné, že obě tyto nepřesnosti nepříznivě ovlivňují výsledky segmentace.

Ve vytvořeném obraze hran je většinou jen několik málo míst s nulovou hodnotou velikosti hran. Malé hodnoty hrany odpovídají drobným změnám jasu vlivem kvantizačního šumu a malých nerovností v osvětlení. K odstranění těchto nevýznamných hran se používá jednoduchá metoda prahování obrazu. Metoda vychází z obrazu hran, který je prahován vhodným prahem. Je-li zvolena větší hodnota prahu, pak se na výsledném obraze podílejí pouze významnější hrany. Hodnotu prahu lze určit například metodou procentního prahování. Jsou-li vstupní data dostatečně kontrastní a bez šumu, přináší tato metoda dobré výsledky. Výhodné může být např. vypuštění všech hranic, jejichž délka je menší než jistá hodnota.



obr.: detekce hran: původní obraz; obraz po vyhledání hran; prahování hodnotou 30; prahování hodnotou 10

3.3 Srovnávání se vzorem

Srovnávání se vzorem je další základní metoda segmentace, která se používá k nalezení známých objektů v obraze nebo k hledání specifického vzoru. Srovnávání se může použít například pokud hledáme jeden objekt ve stejné scéně snímané z různých směrů. Jeden obraz může být použit k definování vzorů, a poté můžeme tyto (nebo podobné) vzory vyhledat v dalších obrazech.

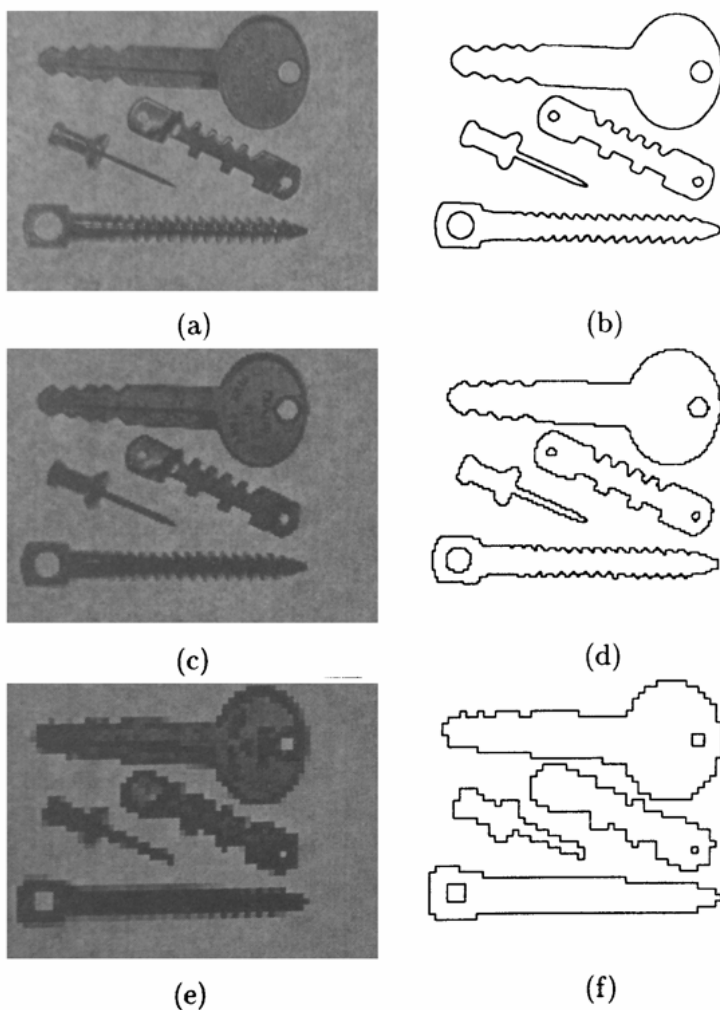
Srovnávací vzor může být velmi malý, nebo to může být celý objekt který hledáme. Srovnávání se vzorem je velice snadné, pokud ve zpracovávaném obraze můžeme očekávat přesnou kopii původního vzoru. Většinou je ale část obrazu poškozená šumem, geometrickou deformací, atd. Proto nehledáme absolutní shodu, ale jen oblast kde je shoda se vzorem největší.

4 Popis objektů

Čtvrtým krokem je popis obrazu nebo též popis nalezených objektů z předešlé segmentace. Existují dva základní způsoby popisu. Jeden je založen na kvantitativním přístupu, což znamená popis objektů pomocí souboru číselných charakteristik. Mohou jimi být například velikost objektu, jeho kompaktnost apod.. Druhou možností je kvalitativní přístup, ve kterém jsou popisovány relace mezi objekty a jejich tvarové vlastnosti. Způsob popisu je zvolen vždy podle toho, k čemu bude dál využit. Ve většině případů je tento popis vstupní informací pro rozpoznávání (klasifikaci) objektů. Výběr popisu je pak závislý na použitém rozpoznávacím algoritmu.

Definovat tvar objektu může být velice obtížné. Tvar je většinou vyjádřen slovně, a lidé používají výrazy jako podlouhlý, kulatý, s ostrými okraji apod. Počítačové zpracování obrazu však potřebuje mít tvar popsany velice přesně. Dosud však není jasné, co je vlastně při popisu tvaru důležité a co ne, a existuje tedy spousta různých metod popisu tvaru objektu.

Dalším problémem při popisu tvaru objektu je velikost (rozlišení) obrázku, protože změna rozlišení může vyvolat změnu tvaru. Při vysokém rozlišení může dojít na hranici objektu vlivem šumu k změně hran, při nízkém rozlišení zase hrozí ztráta drobných detailů objektu. Vhodné rozlišení obrázku tedy musí být zvoleno s ohledem na charakter zkoumané scény.



obr: a) původní obraz 640x480; b) hranice a; c) původní obraz 160x120; d) hranice c; e) původní obraz 64x48; f) hranice e

4.1 Identifikace oblastí

Identifikace jednotlivých oblastí je nutná pro popis oblastí. Jedna z mnoha metod identifikace oblastí je označit každou oblast jednoznačným (celým) číslem. Nejvyšší číslo oblasti poté obvykle určuje i celkový počet oblastí v obraze. Tato metoda se nazývá barvení oblastí. Jiná metoda je užít menší množství identifikačních čísel (dostačující jsou čtyři), a ujistit se, že žádné dvě sousední oblasti nemají stejné identifikační číslo. Poté se ale při odkazování na tuto oblast se musí uvést i nějaký bod, který do této oblasti náleží.

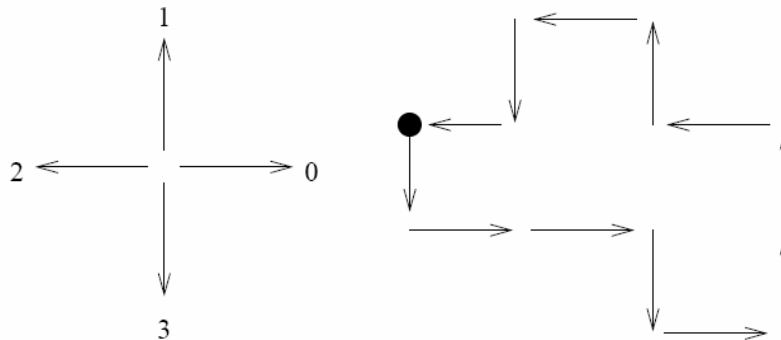
Předpokládejme že segmentovaný obrázek R obsahuje m nezávislých oblastí R_i . Obrázek R většinou rozdělujeme na objekty a pozadí, kde R^c je doplněk, R_b je pozadí, ostatní oblasti jsou jednotlivé objekty.

$$R_b^C = \bigcup_{i=1, i \neq b}^m R_i$$

Vstupní data pro barvicí algoritmus je většinou binární, nebo několikabarevný obrázek, kde pozadí je reprezentováno hodnotou nula a ostatní objekty nenulovými hodnotami. Jako výstup barvicího algoritmu je většinou zvolen několikabarevný obrázek, kde pozadí je reprezentováno nulovou hodnotou a jednotlivé objekty různými hodnotami odpovídající různým barvám.

4.2 Popis objektů založený na popisu hranic

Způsobem popisu objektů založeném na popisu hran jsou řetězové kódy. Řetězové kódy popisují objekt sekvencí jednotkových segmentů čar s danou orientací. Počáteční bod takové sekvence musí být zadán samostatně, musí být určena jeho poloha v obraze. Výsledkem tohoto popisu je seznam čísel, která reprezentují hranu objektu. Řetězový kód může být snadno získán již při algoritmu detekce hran.



obr.: řetězový kód v 4-sousednosti

kód: 3, 0, 0, 3, 0, 1, 1, 2, 1, 2, 3, 2

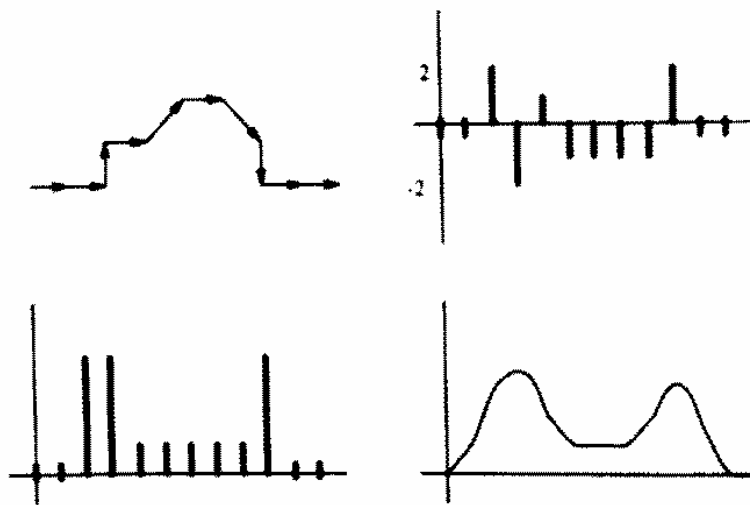
Vlastnosti kterými může být hranice objektu popsána jsou délka hranice, křivost, ohybová energie.

Délka hranice je základní vlastnost objektu, která je jednoduše odvozena z jeho reprezentace řetězovým kódem. Délka každé vertikální a horizontální hrany je jedna, délka diagonální hrany (při uvažovaném 8-sousedství bodu) je $\sqrt{2}$.

Křivost objektu je definována jako poměr mezi celkovou délkou hrany, a počtem bodů ve kterých hrana mění svůj směr. Čím menší je počet těchto bodů, tím přímější hranice je.

Ohybová energie hranice (křivky) může být chápána jako energie potřebná ke zkřivení přímky do požadovaného tvaru. Tato hodnota může být spočítána jako součet čtverců křivosti $c(k)$ na délku hrany L .

$$BE = \frac{1}{L} \sum_{k=1} c^2(k)$$



obr.: ohybová energie; řetězový kód 0, 0, 2, 0, 1, 0, 7, 6, 0, 0; křivost 0, 2, -2, 1, -1, -1, -1, 2, 0; součet čtverců křivosti; vyhlazená křivka

4.3 Popis objektů založený na popisu plochy

Jednoduché objekty lze popsat podle vlastností které má plocha tohoto objektu. Tento popis je jednoduchý, avšak lze jej použít jen na nepříliš složité objekty. Vlastnosti které můžeme u objektu určit jsou: plocha, délka hranice, počet děr, plochy děr, kompaktnost, minimální a maximální vzdálenost hranice od těžiště, průměrná vzdálenost hranice od těžiště, poměr ploch děr k celkové ploše a momenty.

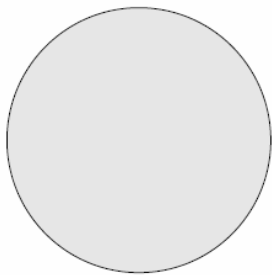
Plocha objektu je nejjednodušší a nezákladnější vlastnost oblasti. Určuje se podle počtu bodů které do dané oblasti patří. Skutečná plocha bodu musí být známa, pokud chceme určit skutečnou plochu oblasti.

Délka hranice je počet hraničních bodů dané oblasti. Pokud se zvyšuje rozlišení obrazu, délka hranice konverguje k nekonečnu.

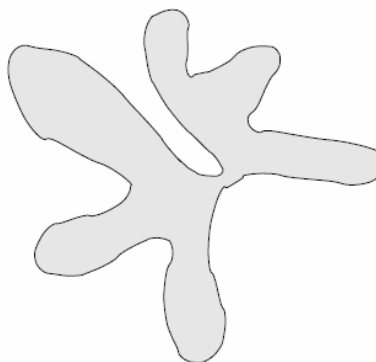
Kompaktnost je další oblíbenou vlastností používanou při popisu ploch.

$$\text{kompaktnost} = \frac{(\text{délka hranice})^2}{\text{plocha}}$$

Nejkompatnější útvar je kruh. Kompaktnost nabývá hodnot $[1, \infty)$.

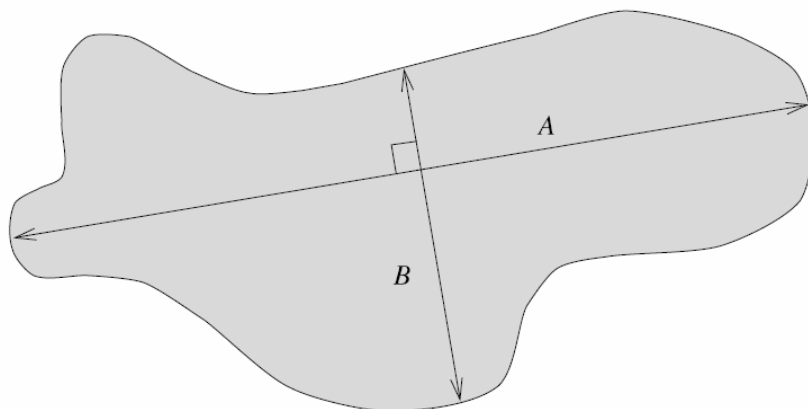


obr.: kompaktní



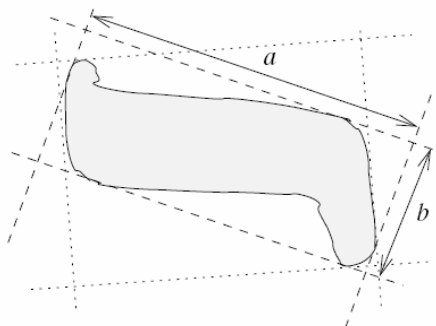
nekompatní

Výstřednost je nejmenší a největší vzdálenost hranice od těžiště.

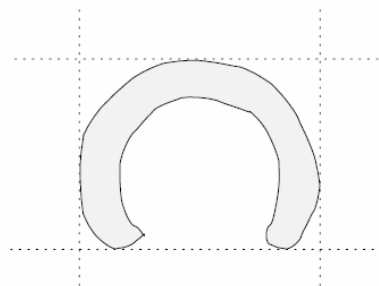


obr.: výstřednost

Protáhlost je definována jako poměr mezi délkou a šířkou nejmenšího obdélníku který ohraničuje námi zkoumaný tvar. Toto kritérium je použitelné jen u objektů, které nejsou příliš zakřivené.



obr.: protáhlost: použitelné



nepoužitelné

5 Klasifikace objektů

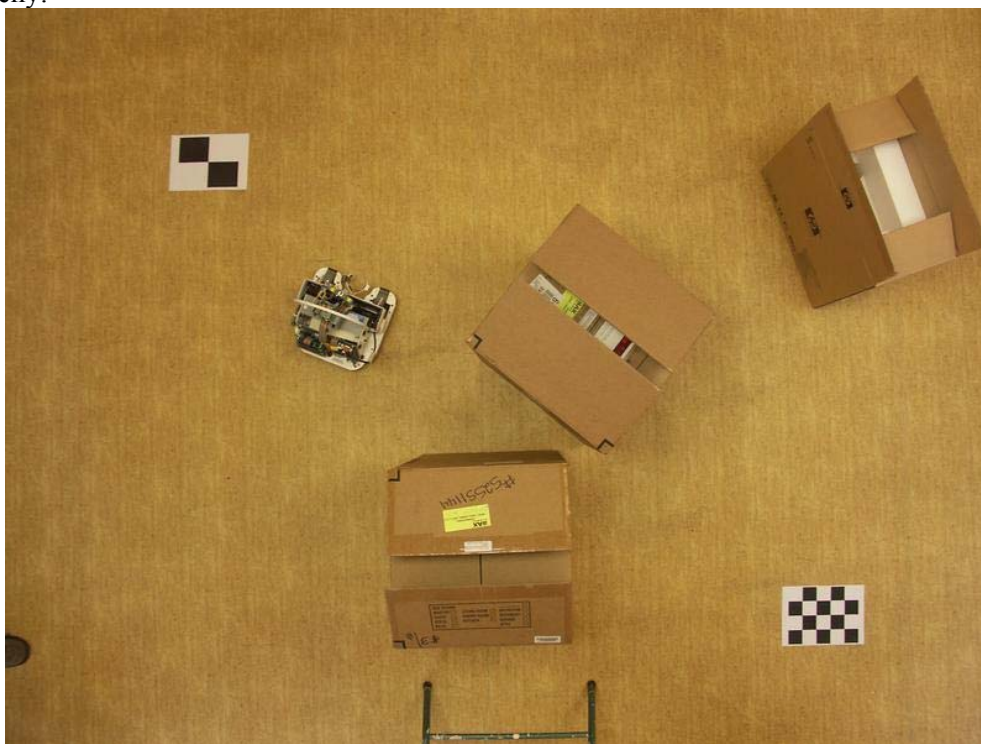
Finálním krokem při zpracování obrazu je klasifikace (rozpoznání obrazu). Ve většině případů se jedná o zařazení objektů nalezených v obraze do skupiny předem známých tříd. Metody klasifikace objektů se dělí do dvou základních skupin, které jsou úzce spjaty se způsobem popisu objektů. Jedná se o příznakové rozpoznávání a strukturální rozpoznávání.

Příznakové metody jsou založeny na principu využití příznaků, což je skupina číselných charakteristik objektu. Učení vlastního klasifikátoru zde může být s trénováním množinou i bez ní, na principu shlukové analýzy.

Strukturální rozpoznávání využívá jako vstupu kvalitativní popis objektů. Objekty jsou zde popsány primitivy. Dále je definována abeceda, jazyk popisu a gramatiky jednotlivých tříd. Vlastní rozpoznávání je pak založeno na principu rozboru slova a kontroly správnosti syntaxe pro jednotlivé třídy.

6 Návrh programu

Pro návrh vlastního programu který by rozpoznával umístění a orientaci robota v laboratoři jsem nejprve nafotil sérii dvaceti fotografií na kterých byl robot umístěn v laboratoři, dále pomocné značky ohraničující prohledávanou oblast a dále nějaké další objekty, převážně prázdné krabice, mezi kterými probíhalo vyhledávání robota. Fotografie byly pořizovány z ptáčích perspektivy a z přibližně stejného místa, aby se minimalizovalo 3D zkreslení dané scény.



obr.: *prohledávaná scéna: robot, značky ohraničující prohledávanou oblast a další objekty*

Na každé z těchto fotografií byly všechny objekty jinak umístěné, s rozdílným natočením, případně se některé objekty dotýkaly.

Jako nástroj pro vytvoření vlastního programu na hledání robota v obraze jsem zvolil jazyk Borland Delphi, a to ve verzi 7. Toto vývojové prostředí poskytuje všechny prostředky pro navrhování, kompilaci a odlaďování aplikací pod Windows. Dále poskytuje některé základní prostředky pro práci s bitmapovou grafikou, čehož jsem dále využil.

6.1 Digitalizace, formát obrazu

Jazyk Delphi poskytuje základní funkce pro práci s bitmapovou grafikou. Je to především prostřednictvím komponenty TImage. Tato komponenta umožňuje náhrávání grafiky ze souboru, její zobrazení a především lze snadno přistupovat přímo k jednotlivým bodům obrazu.

Jeden z podporovaných grafických formátů je BMP. Jako velikost vstupního souboru se kterým bude program pracovat jsem zvolil rozměr 800 x 600 bodů. Tento rozměr jsem zvolil především proto, že jednak poskytuje dostatečné množství detailů v obraze, a jednak je práce s takto malým rozlišením rychlá a nenáročná na paměť.

Po otevření vstupního souboru, dojde tedy k jeho načtení do paměti, a rozdělení na jeho tři barevné složky, Red, Green a Blue, které se uloží do tří polí, se kterými můžeme poté dále samostatně pracovat.

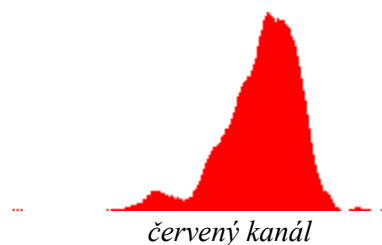
```
c := Image1.Canvas.Pixels[x,y];  
rmatrix[x,y] := (c and $0000FF);  
gmatrix[x,y] := (c and $00FF00) shr 8;  
bmatrix[x,y] := (c and $FF0000) shr 16;
```

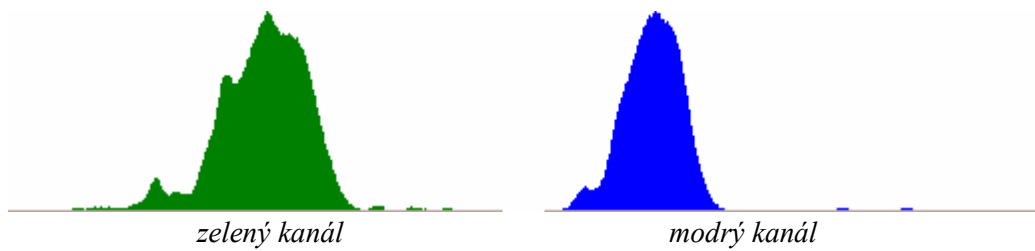
Nyní tedy máme tři pole, každé o rozměru 800 x 600 bodů, které v každém bodě obsahuje číslo 0-255, která udává hodnotu jasu dané barevné složky.

6.2 Předzpracování obrazu

Pro předzpracování obrazu jsem se rozhodl využít metodu jasových transformací obrazu, a to prahování. Prahování je jednoduchá metoda, která poskytuje dobré výsledky.

Pro určení správné hodnoty prahu, bylo nejdříve nutné vytvořit histogramy obrazu. Tyto histogramy bylo nutné vytvořit pro celkový složený obraz, a také pro každou barevnou složku obrazu zvlášť. Celkem tedy vznikly čtyři různé histogramy.





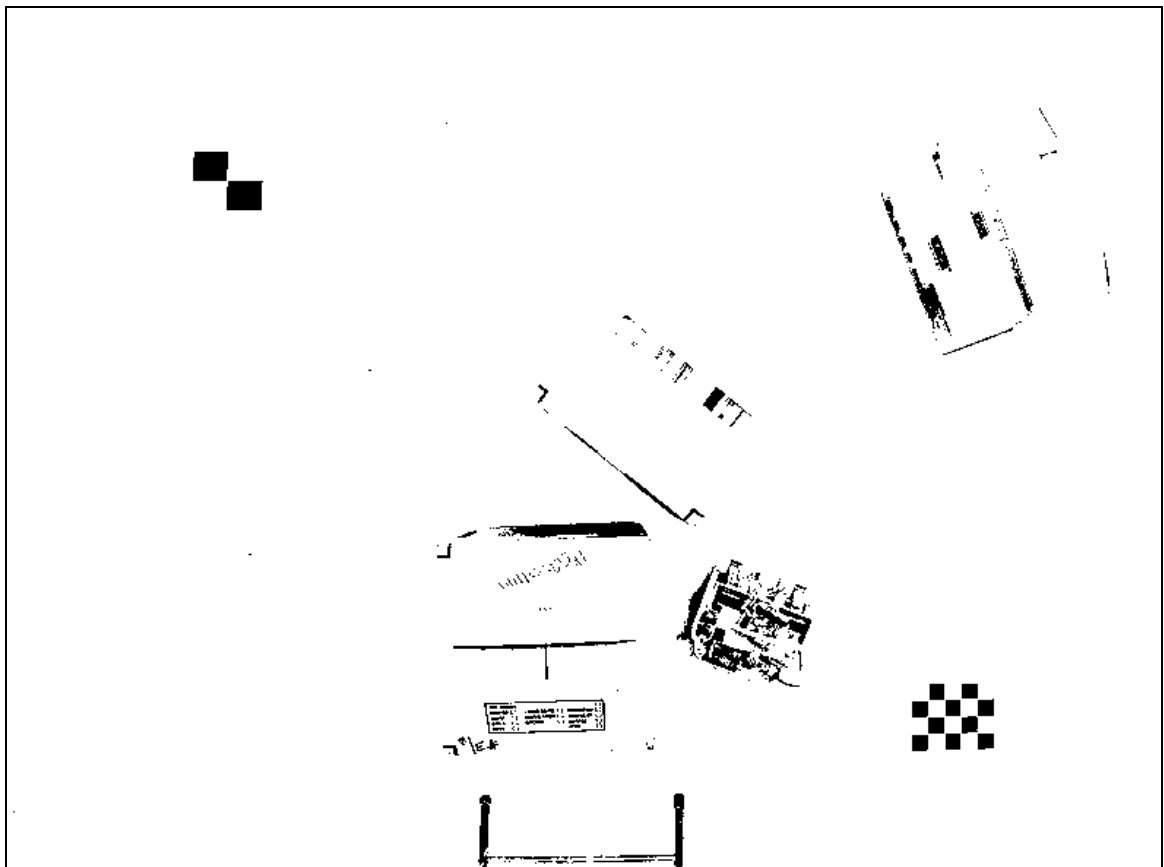
obr.: histogramy

Po zhodnocení histogramů je možno přejít k vlastnímu prahování. Toto prahování probíhá pro každou barevnou složku zvlášť (případně pro složený obraz), a se zvolenou hodnotou prahu, která je vhodná pro právě vyhledávaný objekt v obraze.

```

svetlost := trunc(rmatrix[x,y]*RPodil+gmatrix[x,y]*GPodil+bmatrix[x,y]*BPodil);
if svetlost > prah then cmatrix[x,y] := clWhite
else cmatrix[x,y] := clBlack

```



obr.: prahovaný obraz

Další prováděná úprava obrazu při této fázi je vyhlazení. Tato úprava má za úkol vyhladit prahovaný obraz od jednotlivých osamocených bodů, které vznikly působením šumu. Funguje

tak, že na čtverci 8x8 bodů odstraňuje jednotlivé osamocené body, které tam pravděpodobně nepatří.

```
for l := j-1 to j+1 do
begin
for m := k-1 to k+1 do
begin
if cmatrix[l,m]=clwhite then countw:=countw+1;
if cmatrix[l,m]=clblack then countb:=countb+1;
end;
end;
if countw<=2 then begin
for l := j-1 to j+1 do
begin
for m := k-1 to k+1 do cmatrix[l,m]:=clblack;
end;
end;
```

6.3 Segmentace obrazu, klasifikace objektů

Pro segmentaci obrazu a klasifikaci objektů je využívána opět metoda prahování, dále je užito metody srovnávání se vzorem a dále metody vyhledávání požadovaných vlastností v obraze.

Pro vyhledání levé horní a pravé spodní souřadnicové značky je užito metody srovnávání se vzorem v prahovaném obraze.



obr.: levá horní značka



obr.: pravá spodní značka

K vyhledání těchto dvou značek je celý obraz prahován a poté prohledáván s cílem najít tyto šachovnicové vzory.

Poté co jsou vzory nalezeny, je prohledáván už jen prostor těmito souřadnicemi vymezený a to s cílem najít robota. Jako nejvhodnější metoda se ukázala být metoda vyhledávání místa největšího kontrastu v obraze. V barevném obraze je tedy hledáno místo, o velikosti která přibližně odpovídá velikosti robota, kde se mezi jednotlivými body vyskytuje největší rozdíl v hodnotách všech tří barevných složek.



obr.: vyhledávaný robot

```
for i := x to (x+z) do
begin
for j := y to (y+z) do
begin
result:=result+abs(rmatrix[i,j]-lastr)+abs(gmatrix[i,j]-lastg)+abs(bmatrix[i,j]-lastb);
lastr := rmatrix[i,j];
lastg := gmatrix[i,j];
lastb := bmatrix[i,j];
end;
end;
```

Tato metoda se ukázala být neúčinnější, protože spolehlivě najde robota, který je jinak jen velice obtížně rozpoznatelný, například metodou prahování, a nenechá se zmást ostatními objekty v obraze.

Poté co je nalezen celý objekt robota, je dále pracováno jen s tímto objektem ve snaze učít jeho natočení v prostoru. K tomuto účelu se jako nejlepší metoda ukázala být opět metoda prahování, a to s cílem najít na robotovi nejprve jeho držadlo, které je na všech fotografiích poměrně snadno rozpoznatelné, a poté určit, na které straně od držadla se nachází oblast s větším rozdílem jasu, která odpovídá jeho přední části.

Při hledání držadla, je v obraze prahovaném poměrně vysokým prahem, hledána delší, rovná, nebo mírně zahnutá linie. To je nutné z toho důvodu, že při poloze robota někde při okraji snímané scény, dochází ke zkreslení tohoto držadla, které pak má na fotografii tvar mírného oblouku.



obr.: robot, držadlo tvoří mírný oblouk



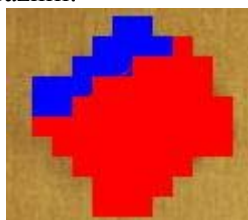
obr.: prahovaný obrázek robota, hledáme držadlo

Poté co máme nalezeno držadlo a tedy robota rozděleného na dvě části, provedeme poslední prahování, po kterém zjišťujeme hodnotu jasu na obou částech. Držadlo je umístěno uprostřed robota, a nemůžeme tedy určit orientaci zjištěním vzdálenosti k okraji. Dále se vlivem perspektivního zkreslení při polohách robota u okraje snímané scény mění poměr části „před“ a „za“ držadlem, takže poté co zjistíme celkovou hodnotu jasu v obou částech, musíme její hodnotu upravit podle viditelného poměru obou částí.



obr.: prahovaný obrázek robota

Potom už tedy máme vše potřebné, abychom robota v obrázku přesně lokalizovali, včetně natočení, a zvýraznili.



*obr.: nalezený a vybarvený robot,
přední část je vybarvená červeně, zadní modře*

7 Závěr, zhodnocení výsledků

Na základě požadavků jsem navrhl program pro lokalizaci robota v laboratoři. Tento program dosahuje poměrně spolehlivých a přesných výsledků, ve všech vstupních obrazech dokáže jednoznačně určit polohu robota. S jeho orientací je to již složitější, ve většině obrazech ji najde správně, ale v několika případech dojde k chybným výsledkům, např. k přehození směru.

Další rozvoj tohoto projektu je směrem k rozpoznávání a jednoznačné identifikaci i v případech, kdy je na snímané scéně umístěno více robotů, kteří by byli opatřeni např. identifikačními čísly pro jejich jednoznačné určení.

Toto téma jsem si vybral, protože se rád zabývám zpracováním digitálního obrazu. Při jeho řešení jsem se seznámil se spoustou nových a zajímavých postupů.

8 Literatura

- [1] Sonka, M., Hlavac, V., Boyle, R.: Image Processing, Analysis, and Machine Vision
Brooks/Cole publishing Company 1999
ISBN 0-534-95393-X

- [2] Lischner R.: Delphi v kostce, pohotová referenční příručka
Computer Press 2000
ISBN 80-7226-361-7

- [3] Swan T.: Mistrovství v Delphi 4
Computer Press 1999
ISBN 80-7226-173-8
..

9 Přílohy

CD-ROM se zdrojovými kódy projektu
Zdrojové obrazy a jejich výsledky po zpracování

