

Videnie mobilného robota pomocou wifi kamery

Tomáš Ferenczi, Jozef Dorner

Abstrakt

Problémy rozpoznania a sledovania objektov sú aktuálne, aj keď už existuje veľa metód a algoritmov, ale v praxi, kde môže dôjsť k ohrozeniu života sa ešte veľmi nepoužívajú. Máme tým na mysli práve dopravné situácie a rozpoznávanie dopravných značiek samotným automobilom. Tieto systémy sa nevedia dokonale a inteligentne prispôbovať svojmu okoliu, a stále je nutný zásah operátora.

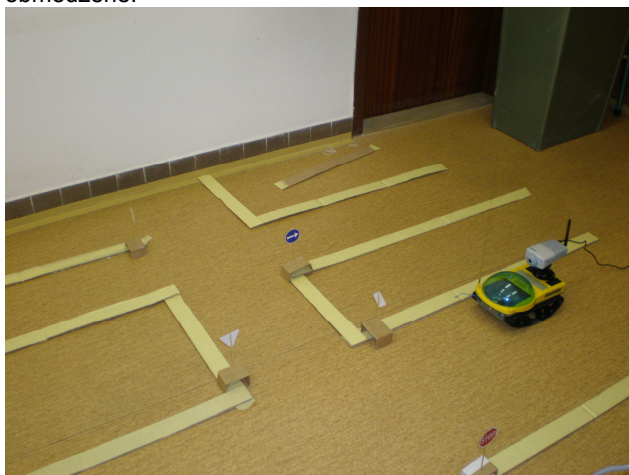
Kľúčové slová: počítačové videnie, spracovanie obrazu, farebné filtre, rozpoznávanie a sledovanie objektov.

Úvod

V tomto článku sa budeme zaoberať videním a sledovaním objektov snímaných z kamery umiestnenej na mobilnom robote. Povieme si o metódach, ktoré budeme využívať na rozpoznávanie dopravných značiek, prípadne iných objektov. S problémom sledovania sa stretávame v každodennom živote, napríklad v doprave, pri rozpoznávaní dopravných značiek osobných automobilov a pod.

1. Popis modelu

Hlavnou časťou simulačného modelu je mobilný robot, ktorý je reprezentovaný pásovým vozidlom. Na vozidle je umiestnená wifi kamera. Budeme využívať iba jednu kameru a nie dve, pretože pri stereo videní rýchlosť sledovania by bola pomalšia a dosiahnutie spracovania v reálnom čase obmedzené.



Obr. 1. Scéna mobilného robota s wifi kamerou
Fig. 1. Scene mobil robot with wifi camera

2. Spracovanie obrazu

Prvou a zároveň najdôležitejšou úlohou je spracovanie obrazu. To, aké kvalitné výstupy dostaneme nám ovplyvní celý ďalší proces. Metóda na spracovanie a vyhľadávanie objektov je v súčasnosti veľa. Niektoré metódy a algoritmy sú však natoľko zložité, že nie je možné ich použiť pri praktickom riešení úloh, pretože majú veľkú výpočtovú zložitosť

a časovú náročnosť výpočtu. Kritériom pre riešenie bude minimalizácia výpočtovej zložitosti a súčasne snaha dosiahnuť čo najlepší výsledok.

2.1. Ladenie farieb

Farby sú často popisované porovnávaním s vytvorenou množinou štandardných farieb a vyhľadávaním najpodobnejšej z nich. Vzniklo veľa takýchto množín so štandardne definovanými farbami a sú využívané vo tlačiarenskom priemysle.

Daná farba so spektrálnou krivkou $S(\lambda)$ sa premietne na jednu časť obrazovky pričom na druhú časť sa premietne kombinácia troch daných testovacích svetiel so spektrálnymi krivkami $A(\lambda)$, $B(\lambda)$, $C(\lambda)$. Pozorovateľ mení intenzitu testovacích svetiel pomocou koeficientov intenzít (a, b, c) až pokým vzniknutá farba je vizuálne rovnaká s danou farbou. T.j. $T(\lambda) = a \cdot A(\lambda) + b \cdot B(\lambda) + c \cdot C(\lambda)$. Pričom spektrálne krivky $S(\lambda)$ a $T(\lambda)$ môžu byť celkom rozdielne. Takto definovaná farba môže byť chápaná ako „súčet“ koeficientov intenzít a, b a c . Definujme algebru nad farbami nasledovne:

Predpokladajme, že spektrálne krivky $S(\lambda)$ a $P(\lambda)$ definujú vizuálne rovnakú farbu. Toto zapisujeme $S = P$. Pridajme tretiu farbu N , so spektrálnou krivkou $N(\lambda)$, k obidvom farbám. Je experimentálne dokázané že aj novovzniknuté farby definujú rovnakú farbu. Symbol $=$ teda definuje vizuálne nerozlišiteľné farby. Pre symbol $+$ platí, že $S + N$ definujú farbu ktorá vznikne spojením spektier $S(\lambda)$ a $N(\lambda)$.

$$\text{Ak } (S = P) \text{ potom } (N + S = N + P) \quad (1)$$

Obdobne dospejeme k záveru, že ak $(S = P)$ potom

$a \cdot S = a \cdot P$, pričom a je skalár. Takto dospejeme k lineárnej kombinácii farieb:

$$T = a \cdot A + b \cdot B. \quad (2)$$

Môžeme zdefinovať vektorovú algebru farieb. Vektory reprezentujú farbu. Vnímanie farieb je trojdimenzionálne z pohľadu vnímania farieb ľudským okom. To znamená, že každá farba môže byť vytvorená kombináciou troch základných farieb. Označme si základné farby R, G a B , potom pre farbu C platí:

$$C = r.R + g.G + b.B \quad (3)$$

Pričom koeficienty r , g a b definujú intenzitu základných farieb obsiahnutých vo farbe C . Dôvodom prečo za základné farby boli zvolené červená (R), zelená (G) a modrá (B) je citlivosť tyčiniek v ľudskom oku práve na tieto tri farby. Za základné farby však môžu byť zvolené ľubovoľné tri farby pre ktoré platí, že žiadna z nich nie je lineárnou kombináciou zvyšných dvoch.

Pomocou základných troch farieb môžeme definovať priestor farieb, pričom každá farba je reprezentovaná bodom (r,g,b) v tomto priestore. Bod $(0,1,0)$ reprezentuje čisto zelenú a $(0.2,0.3,0.5)$ žltú atď. Keď zdvojnásobíme každý koeficient tak získame tú istú farbu ale dvakrát jasnejšiu.

Experimentálne sa pomocou kombinácie troch základných farieb R,G,B hľadali spektrálne čisté farby t.j. farby, ktoré majú dominantnú vlnovú dĺžku a energia je sústredená v oblasti tejto dominantnej vlnovej dĺžky (monochromatické).

$$\text{mono}(\lambda) = r(\lambda).R + g(\lambda).G + b(\lambda).B \quad (4)$$

Napríklad čisto oranžová farba s dominantnou vlnovou dĺžkou $\lambda = 600\text{nm}$, pre pozorovateľov zodpovedá kombinácii $0.37R + 0.08G$. Obvyčajne spektrum oranžového svetla nie je rovnaké so spektrom svetla vzniknutého touto kombináciou (súčtom), ale obidve svetlá vizuálne vyzerajú rovnako. Z obrázku obr.3 vidno, že pre niektoré farby musí byť niektorý s koeficientov záporný aby bola zachovaná platnosť rovnice (4). Tak napríklad pre $\lambda = 520$ je $r(\lambda)$ záporné.

$$C = 0.7R + 0.5G - 0.2B \quad (5)$$

V praxi však nie je možné svetlo, ktoré sa tam nenachádza, odstrániť. Toto vyriešime keď rovnicu prepíšeme od tvaru:

$$C + 0.2B = 0.7R + 0.5G \quad (6)$$

Z toho vyplýva, že farbu C nie je možné vytvoriť lineárnou kombináciou s kladnými koeficientmi pre základné farby. Farbu $C + 0.2B$ však takto vytvoriť vieme. Niektoré farby teda je možné vytvoriť pomocou kladných koeficientov, no niektoré nie. Jednoduchou povedané takýmto skladaním vznikajú menej čisté farby, pričom je nemožné vytvoriť čistejšiu farbu. Je užitočné definovať funkcie $r(\lambda)$, $g(\lambda)$ a $b(\lambda)$ pre ktoré platí:

$$r(\lambda) = r(\lambda) / s(\lambda), \quad g(\lambda) = g(\lambda) / s(\lambda), \quad b(\lambda) = b(\lambda) / s(\lambda)$$

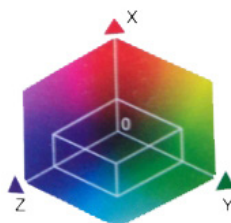
$$\text{pričom: } s(\lambda) = (r(\lambda) + g(\lambda) + b(\lambda))$$

$$\text{a } r(\lambda) + g(\lambda) + b(\lambda) = 1. \quad (7)$$

Tieto funkcie dávajú hodnoty koeficientov pre každú základnú farbu potrebné na vytvorenie svetla s dominantnou vlnovou dĺžkou λ a jednotkovým jasom. Farbu s jednotkovým jasom preto môžeme definovať dvojicou $(r(\lambda), g(\lambda))$, pretože $b(\lambda)$ si dokážeme dopočítať ($b(\lambda) = 1 - r(\lambda) - g(\lambda)$). Zobrazením takto vzniknutých bodov dostávame krivku definujúcu farby s čistým spektrom, pričom z normalizácie vyplýva že táto krivka leží v jednej rovine. Všimnime si, že kvôli zápornosti niektorých koeficientov pre niektoré farby je krivka čiastočne mimo kladného oktantu v priestore. Pre niektoré štandardy, ktoré ďalej popíšeme ostáva spomínaná krivka v kladnom oktante.

2.2. RGB model

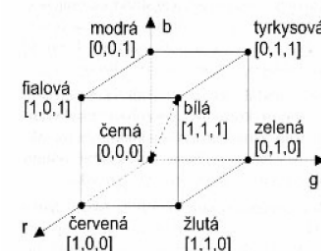
Nakoľko samotný obraz dostávame vo veľkosti 640×480 pixlov, bude obraz celkovo kvalitný a bude vyhovovať požiadavkám na spracovávanie. Každý pixel sa skladá z troch zložiek farieb a to RGB (red, green, blue). Budeme teda spracovávať 921600 pixlov. Práve táto informácia bude pre nás najdôležitejšia. Každá jedna zložka farby RGB môže nadobúdať hodnoty od 0 do 255 .



Obr. 2. RGB kocka

Fig. 2. RGB cube

Akýkoľvek farebný odtieň je zložený z rôznych sýtostí troch základných farieb - červenej, zelenej a modrej. Tieto farby sú pre tento systém základné. Ostatné farby teda dosiahneme miešaním týchto troch farieb. Dané tri základné farby sú rovnocenné, a preto najlepší matematický model RGB je tzv. RGB kocka obr. 3. Táto kocka leží v kladných kvadrantoch osí X,Y,Z .

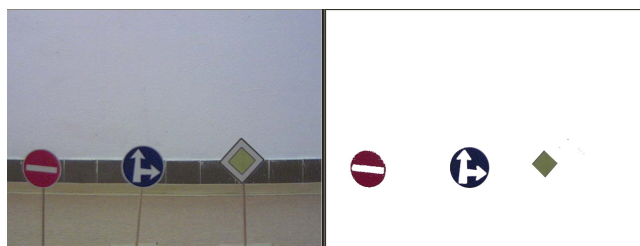


Obr. 3. RGB model

Fig. 3. RGB model

Model RGB je aditívny model, kedy výsledná farba závisí od sčítania farebných hodnôt jednotlivých zložiek. Znamená to, že ak všetky tri farebné zložky nadobúdajú hodnotu nula, výsledná farba je čierna (nič nesvieti). Ale ak všetky tri farebné zložky sú stopercentné, výsledná farba je biela. Jednotlivé osy teda zobrazujú "čisté" odtiene základných farieb: farba na osi X (vpravo) sú rôzne odtiene zelenej, na osi Y (hore) červenej a na osi Z (dopredu) modrej. Dôležité sú pre nás hodnoty v rohoch kocky, teda matematicky body $(1, 1, 0)$, $(1, 0, 1)$, $(0, 0, 1)$, ak chápeme kocku ako jednotkovú.

Ďalej budeme spracovávať a analyzovať tieto zložky za účelom rozpoznania dopravnej značky na základe farebného modelu. Po analýze každého pixla dostaneme upravenú informáciu o obraze.



Obr. 4. Spracovanie obrazu a hľadanie dopr. značky

Fig. 4. Image processing and searching traffic signs

Na obrázku obr.4 je vidieť, že po spracovaní obrazu nám zostali iba dopravné značky. Veľkou výhodou tejto metódy je, že dopravná značka môže byť deformovaná napríklad pri bočnom pohľade. Vtedy nevidíme kruh ale elipsu. Samozrejme túto deformáciu musí mať natrénovanú aj neurónová sieť, ktorej úlohu si popíšeme neskôr. Samotný systém však ešte vôbec nevie či ide skutočne o dopravné značky. Preto musíme obraz ďalej spracovať.

3. Odstránenie šumu

Najväčší problém, ktorý vzniká pri hľadaní farebných plôch bolo osvetlenie a tiež aj zašumenie obrazu. Cieľom filtrácie je odstrániť šum, ktorý vznikol pri spracovaní a prenose obrazu, ostré prechody alebo zvýrazniť niektoré črty obrazu.

V prípade zašumeného obrazu, môžeme poškodené pixle nahradiť priemerovaním susedných bodov. Tu využívame malé zmeny susedných bodov a tým aj podobnú hodnotu jas.

Pri spracovávaní frekvencií obrazovej funkcie môžeme rozlišovať dve metódy: vyhladenie a ostrenie obrazu. Pri vyhladzovaní obrazu potláčame vysoké frekvencie a tým aj náhodný šum. Vtedy však dochádza k rozmazávaniu hrán a ostrých prechodov. Naopak ostrením obrazu dochádza k zvýrazneniu vyšších frekvencií a tým aj k zvýrazneniu hrán, ale aj k zvýrazneniu šumu.

Pri modifikácii obrazu sa používajú dva hlavné prístupy: priestorové a frekvenčné. Priestorové pracujú priamo s obrazom alebo jeho časťou $f(x,y)$, ktorú transformujú na výstupný obraz $g(x,y)$ podľa vzorca

$$g(x,y) = T[g(x,y)] \quad (8)$$

Frekvenčné prístupy používajú konvolúciu obrazu. Konvolúcia dvoch funkcií $f(x)$ a $g(x)$ (označ. $f(x)*g(y)$) je definovaná pomocou integrálu:

$$f(x)*g(y) = \int_{-\infty}^{\infty} f(\alpha) \cdot g(x-\alpha) d\alpha \quad (9)$$

kde je α integrálna premenná.

Výsledný obraz $g(x,y)$ je tvorený konvolúciou obrazu $f(x,y)$ a $k(x,y)$ a platí:

$$G(x,y) = K(x,y) * F(x,y) \quad (10)$$

kde G , K a F sú Fourierove transformácie funkcií g , k a f . Výsledný obraz potom vypočítame nasledovne:

$$g(x,y) = F(K(x,y) * F(x,y)) \quad (11)$$

3.1. Vyhladenie obrazu

Vyhladenie obrazu slúži na odstraňovanie šumu. Metódy na odstraňovanie sú: priemerovanie, filtrácia metódou mediánu a nízko frekvenčné filtrovanie

3.1.1. Priemerovanie

Pri filtrácii obrazu filtrujeme lokálnym priemerovaním okolia bodu (x, y) :

$$g(x,y) = \frac{1}{M} \sum_{(i,j) \in O} f(i,j) \quad (12)$$

kde M je počet bodov okolia O .

3.3. Ostrenie obrazu

Hranu v obraze chápeme ako množinu bodov obrazu, v ktorých susedstve sa veľmi mení jas. Ostrením obrazu sa snažíme zvýrazniť tieto hrany.

3.4. Kostra množiny (skelet)

Pod týmto pojmom môžeme rozumieť útvary čiar, ktoré vystihujú danú množinu a zachovávajú jej geometrickú podstatu. Algoritmov na získanie skeletu je veľké množstvo, pričom

ich výstupy sa rôznia, pretože neexistuje presná definícia skeletu množiny v rastrovej grafike. Všeobecný postup spočíva v stenčovaní zadanej množiny.

4. Osvetlenie objektu

Problém s ktorým sa pri spracovaní obrazu musíme zaoberať je aj osvetlenie. Ako vieme osvetlenie nám mení hodnoty jednotlivých zložiek farby RGB. Preto musíme definovať množinu, ktorá bude obsahovať hodnoty zložiek RGB pri rôznych typoch osvetlenia.

Využijeme na to vzorec (13). Kde hodnoty $(r1,b1,g1)$ získame z nami zadanej množiny a $(r2,b2,g2)$ sú hodnoty ktoré získavame z každého jedného pixla ktorý aktuálne spracovávame.

$$c = \sqrt{(((b1-b2)*(b1-b2)) + ((g1-g2)*(g1-g2)) + ((r1-r2)*(r1-r2)))} \quad (13)$$

Výsledná premenná c charakterizuje hodnotu, podľa ktorej vieme určiť či sa nájdená farba zhoduje s hľadanou farbou. Čím je c menšie tým je pravdepodobnosť toho, že sme našli správny bod väčšia.

5. Ohraničenie objektov

Ďalším krokom je ohraničenie nájdených objektov.



Obr. 5. Ohraničenie nájdených objektov
Fig. 5. Area boundary find objects

Segmentáciou a ohraničením objektov získame ich vektorový tvar, ktorý nám bude slúžiť ako informácia pre neurónovú sieť. Na obr. 5 v ľavej časti je záber z kamery v reálnom prostredí a v pravej časti obr. 5 je už objekt ohraničený na základe farebného modelu. Pre rozpoznanie objektu a jeho pochopenie sa objekt popíše vektorovými súradnicami a odošle ako výrez z originálu neurónovej siete, pretože má menej chýb a je jasnejší. Neurónovej sieti poskytneme aj informáciu o farebnom zložení daného objektu, aby mohla analýza prebehnúť čo najrýchlejšie, pretože neurónová sieť bude podľa informácie o farbách vedieť, v ktorej množine objektov sa môže nami nájdený objekt nachádzať. Neurónová sieť nám následne rozhodne o type dopravnej značky, ak teda išlo o dopravnú značku alebo pošle informáciu že nájdený objekt nie je dopravná značka. Po vyhodnotení pošle neurónová sieť informáciu ďalšej časti programu, ktorá bude riadiť pohybový systém mobilného robota. Program na rozpoznanie dopravných značiek sa dá ďalej modifikovať podľa požiadaviek. Napríklad z dopravnej značky, respektíve z jej polohy, veľkosti a umiestnenia, vieme určiť, či značku obchádzame vľavo alebo vpravo. Rovnako je program modifikovateľný aj na iné objekty nielen značky. Podľa zadefinovanej množiny farieb vieme nájsť ľubovoľný objekt. Je však potrebné na tento objekt natrénovať aj neurónovú sieť.

Záver

Videnie patrí medzi najdôležitejšie zmysly človeka, preto sa snažíme tento typ vnemu skúmať a získavať čo najviac

možných informácií. Počítačové videnie má pre spoločnosť veľký prínos. Či je to už rozvíjajúci sa automobilový priemysel, výskumné projekty, alebo oblasť medicíny, vždy si nájde široké uplatnenie.

Abstract

Application algorithm process representation knowledge object and application in computer vision. With the development of information technologies in last years also increases the level of artificial intelligence. One of the solving problems is computer vision. The effort of this project will be creating of system and building software which will solve problematic of computer vision and searching traffic signs.

Literatúra

- [1] <http://fractal.dam.fmph.uniba.sk/~ftacnik/>
- [2] Milan Sonka, Václav Hlavac, Roger Boyle: Image Processing Analysis and Machine Vision, the University of Iowa, 1999

PodĎakovanie

Článok vznikol za podpory grantovej agentúry MŠ VEGA č.p. 1/3120/06

Bc. Tomáš Ferenczi

Slovenská Technická Univerzita
Fakulta elektrotechniky a informatiky
Ústav riadenia a priemyselnej informatiky
Ilkovičova 3
81219 Bratislava
tomas.ferenczi@gmail.com

Ing. Jozef Dorner

Slovenská technická univerzita
Fakulta elektrotechniky a informatiky
Ústav riadenia a priemyselnej informatiky
Ilkovičova 3
81219 Bratislava
jozef.dorner@stuba.sk