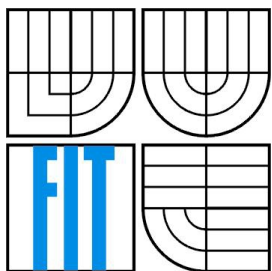




VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
BRNO UNIVERSITY OF TECHNOLOGY



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

ROZPOZNANIE POČASIA Z POHLĎADU VONKAJŠEJ STACIONÁRNEJ KAMERY

RECOGNITION OF WEATHER IN AN OUTDOOR STATIONARY CAMERA VIEW

BAKALÁRSKA PRÁCA
BACHELOR'S THESIS

AUTOR PRÁCE
AUTHOR

MICHAL JENČO

VEDÚCI PRÁCE
SUPERVISOR

DOC. ING. ADAM HEROUT, PH.D.

BRNO 2014

Abstrakt

Tato práce řeší rozpoznávání počasí z pohledu stacionární venkovní kamery se záběrem krajiny, a to konkrétně mlhy, jasného, polojasného a zamračeného počasí. Zvolený problém byl vyřešen výpočtem pěti obrazových příznaků a strojového učení. Podařilo se dosáhnout celkové úspěšnosti rozpoznávání 95% s jenom malými odchýlkami mezi jednotlivými typmi počasí. Hlavným zjištěním této práce je, že za pomoci zvolené sady jednoduchých příznaků je možno úspěšně odlišit zvolené typy počasí. Výsledky této práce umožňují graficky vykreslit průběh počasí počas dne.

Abstract

This thesis deals with classification of weather from stationary outdoor camera images with a landscape view. It classifies fog, clear, partly cloudy and overcast weather in particular. The problem was solved by computing five image flags and using machine learning. Hit rate of 95% was achieved with only small variances between weather types. The main finding of this work is that it is possible to successfully differentiate between the selected weather types with the chosen set of simple image flags. The implementation of this system enables plotting a graph of weather progression during a chosen day.

Klíčová slova

rozpoznávání počasí, klasifikace počasí, stacionární kamera, obrazové příznaky, strojové učení, detekce hran, saturace, kontrast, počítačové vidění, SVM

Keywords

weather recognition, weather classification, stationary camera, image flags, machine learning, edge detection, saturation, contrast, computer vision, SVM

Citácia

Michal Jenčo: Rozpoznanie počasia z pohľadu vonkajšej stacionárnej kamery, bakalárska práca, Brno, FIT VUT v Brne, 2014

Rozpoznanie počasia z pohľadu vonkajšej stacionárnej kamery

Prehlásenie

Prehlasujem, že som túto bakalársku prácu vypracoval samostatne pod vedením Doc. Ing. Adama Herouta, Ph.D.

Uviedol som všetky literárne pramene a publikácie, z ktorých som čerpal.

.....
Michal Jenčo
11. mája 2014

PodĎakovanie

Rád by som sa poďakoval vedúcemu bakalárskej práce Doc. Ing. Adamovi Heroutovi, Ph.D. za odborné a motivačné vedenie.

© Michal Jenčo, 2014

Táto práca vznikla ako školské dielo na Vysokom učení technickom v Brne, Fakulte informačných technológií. Práca je chránená autorským zákonom a jej použitie bez udelenia oprávnenia autorom je nezákonné, s výnimkou zákonom definovaných prípadov.

Obsah

1	Úvod.....	2
2	Súčasný stav v oblasti rozpoznávania počasia z obrazu.....	3
2.1	Rozmanité využívanie kamier na rozpoznávanie počasia.....	4
2.2	Hmla.....	6
2.3	Oblačnosť.....	6
3	Návrh systému rozpoznávania oblačnosti a hmly z obrazu kamery.....	11
3.1	Cieľ práce.....	11
3.2	Špecifikácia podmienok na obraz kamery potrebný pre úspešné rozpoznávanie.....	11
3.3	Voľba obrazových príznakov pre rozpoznávanie.....	12
3.4	Zhrnutie rozpoznávacej funkcie jednotlivých príznakov.....	16
4	Implementácia navrhutej metódy.....	18
4.1	Voľba, zber a kategorizácia pracovnej sady.....	18
4.2	Implementácia výpočtu obrazových príznakov.....	19
4.3	Implementácia strojového učenia.....	20
5	Testovanie a vyhodnotenie navrhutej metódy.....	21
5.1	Pracovná sada.....	21
5.2	Vplyv veľkosti trénovacej sady na úspešnosť rozpoznávania.....	22
5.3	Vplyv jednotlivých príznakov na úspešnosť rozpoznávania.....	23
5.4	Úspešnosť rozpoznávania pri postupnom pridávaní príznakov.....	24
5.5	Znázornenie vlastností klasifikátora pomocou matice zámen.....	25
5.6	Úspešnosť rozpoznávania v jednotlivých miestach.....	28
5.7	Znázornenie priebehu počasia počas celého dňa.....	29
6	Záver.....	30
	Príloha 1 – priebeh počasia, Klínovec.....	31
	Príloha 2 – priebeh počasia, Churáňov.....	32
	Príloha 3 – priebeh počasia, Klatovy.....	33
	Príloha 4 – prezentačný plagát.....	34
	Príloha 5 – obsah CD.....	35
	Literatúra.....	36
	Zoznam príloh.....	38

1 Úvod

Cieľom tejto práce bolo navrhnuť, implementovať, otestovať a užitočne aplikovať systém na rozpoznanie počasia, a to konkrétne hmly, jasného, polooblačného a zamračeného počasia, zo snímkov statickej vonkajšej kamery s pohľadom na scénu krajiny.

Moje riešenie implicitne predpokladá scénu, v ktorej sa horizont nachádza v strede obrazu. Toto obmedzenie sa opiera o fakt, že pre spoľahlivé rozpoznanie miery oblačnosti v mieste umiestnenia kamery je potrebné v scéne vidieť značnú časť oblohy; na druhej strane, pre spoľahlivé rozpoznanie hmly je potrebné vidieť značnú časť zeme. Riešenie poskytuje výsledky pre ľubovoľné obrázky, ale pre čo najúspešnejšiu detekciu je potrebné na toto obmedzenie brať ohľad.

Pre klasifikáciu obrázkov som zvolil nasledujúce príznaky: miera hrán v spodnej polovici obrázku, jas, kontrast, saturácia oblohy a miera hrán v oblohe. Myšlienkový proces ich výberu bol taký, že som uvažoval, na základe čoho rozlišuje počasie pri pohľade na obrázok človek, a zároveň ktoré príznaky budú jednotlivé typy počasia od seba jednoznačne odlišovať.

Práca je štrukturovaná nasledujúcim spôsobom. Prvá časť je venovaná problematike extrakcie informácií o počasí zo snímkov kamier v nedávnych prácach a výskumoch. Druhá časť obsahuje popis návrhu mojej metódy. Tretia časť sa venuje implementácii navrhutej metódy – zberu a klasifikácii snímkov, výpočtu príznakov a strojovému učeniu. Štvrtá časť opisuje spôsob a výsledky testovania, úspešnosť rozpoznávania, hlavné problémy a silné miesta.

2 Súčasný stav v oblasti rozpoznávania počasia z obrazu

V tejto kapitole predstavím a popíšem nedávne práce venujúce sa rozpoznávaniu počasia z obrazu kamery. Popíšem ich motiváciu, postupy a dosiahnuté výsledky.

Nie je jednoduché nájsť oblasť života, v ktorej sa s rozmachom vedy a techniky nezačali nejakým spôsobom používať videokamery. Či už ide o oblasť verejnú, súkromnú, štátnu, technologickú, alebo výrobnú, vo väčšine z nich už videokamery majú stabilné miesto. Medzi týmito nechýba ani sledovanie počasia.

Meteorologické ústavy a stanice disponujú veľmi často okrem rôznych špecializovaných snímačov a detektorov aj jednoduchou kamerou, zaznamenávajúcou scénu buď kontinuálne, alebo v dopredu určených dlhších časových intervaloch. Napríklad Český hydrometeorologický ústav (ČHMÚ) používa videokamery s intervalovým záznamom na 66 rôznych meteorologických staniciach rozmiestnených po celej Českej republike, ktorých snímky sú verejne dostupné na webovej URL ¹. Interval medzi jednotlivými snímkami je najčastejšie 5 minút, snímky sú zvyčajne dostupné pre aktuálny deň a dva predošlé.



Obrázok 2.1: Typický príklad snímku z meteorologickej stanice ². Pohľad ukazuje krajinu s viditeľným horizontom, kamera je umiestnená vodorovne a zem a obloha majú v obraze približne rovnaké zastúpenie.

¹ <http://www.chmi.cz/files/portal/docs/meteo/kam/index.html>

² stanica Maruška, ČHMÚ, prebraté z <http://www.chmi.cz/files/portal/docs/meteo/kam/prohlizec.html?cam=maruska>

Z takýchto snímkov vie človek bez akýchkoľvek ďalších znalostí o danom mieste veľmi jednoducho pohľadom určiť, aký typ počasia je na nich zobrazený. Samozrejme, veľké množstvo údajov o počasi sa z obrázku iba vizuálnou inšpekciou zistiť nedá vôbec, alebo len veľmi obmedzene. Na určenie teploty nutne potrebujeme teplomer, na určenie atmosférického tlaku barometer a podobne. Čo sa však na jednotlivých obrázkoch vizuálne rozpoznať jednoznačne dá, je oblačnosť a prítomnosť hmly. Prirodzene teda existuje motivácia automatizovať tento proces.

2.1 Rozmanité využívanie kamier na rozpoznávanie počasia

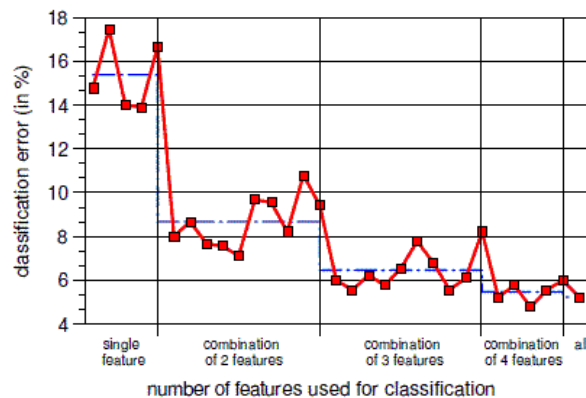
Jednou z oblastí, kde sa v nedávnej dobe začali využívať kamery na rozpoznávanie poveternostných podmienok, sú asistenčné systémy vo vozidlách. Využívajú sa napríklad na sledovanie jazdných pruhov a vybočenia z nich (výskum [9]), detekciu okoloidúcich vozidiel (výskum [10]) a podobne. Tieto a mnohé iné pomocné systémy však často spoľahlivo fungujú iba za ideálnych poveternostných podmienok. Preto sú žiadané systémy, ktoré dokážu poveternostné podmienky z vozidla rozpoznať a poskytnúť informácie o nich ďalším asistenčným subsystémom. Napríklad informácie o daždivom počasi a vode na vozovke, alebo o slabej viditeľnosti a hmle sú veľmi užitočné.

Kurihata a kolektív [11] navrhli a vytvorili systém na rozpoznanie daždivého počasia použitím kamery namontovanej vo vozidle snímajúcej čelné sklo. Ich prístup spočíval v predpoklade, že kvapky na čelnom skle v obraze kamery majú isté spoločné vlastnosti, ktoré je možné automaticky rozpoznať, a to konkrétne obrysové hrany, rozmazané hrany objektov za kvapkami a refrakciu svetla. Vytvorili vzor predstavujúci obraz kvapky na čelnom skle ako sa javí v obraze kamery, a ten postupne posúvali po obraze a hľadali s ním zhodu. Touto metódou sa im podarilo dosiahnuť 90-percentnú úspešnosť odhalenia prítomností kvapiek na čelnom skle.

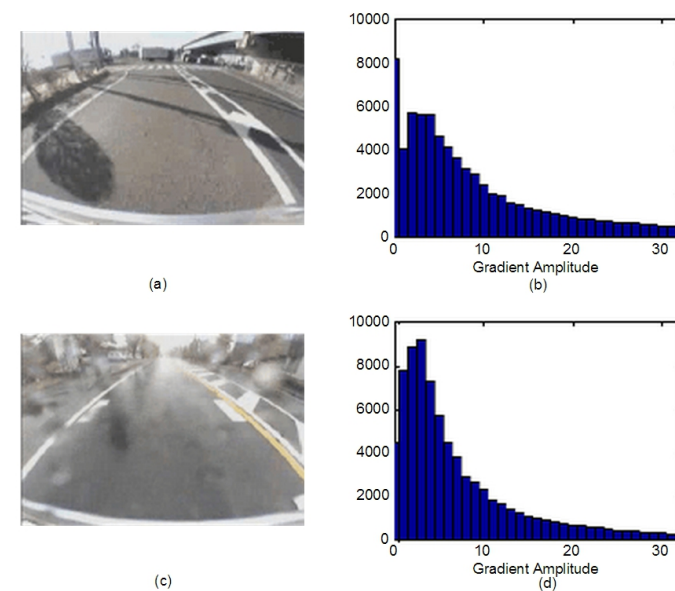
Roser a Moosmann [2] predstavili a implementovali systém schopný rozlíšiť počasie bez zrážok, slabý dážď a silný dážď, taktiež použitím jednej kamery namontovanej v interiéri vozidla. Obraz z kamery rozdelili na záujmové regióny a tie použili na extrakciu obrazových príznakov. Za skúmané príznaky zvolili lokálny kontrast, minimálny jas, ostrosť, odtieň a saturáciu. Pre rozpoznávanie zvolili metódu podporných vektorov (SVM) s lineárnym jadrom. Pracovali s tromi testovacími množinami; prvá obsahovala snímky z diaľnic, v druhej bola polovica snímok z diaľnic a druhá polovica z vidieckych ciest a v tretej k nim pridali ešte snímky z mestských prostredí. Pre prvú testovaciu množinu dosiahli Roser a Moosmann chybovosť iba 2,04%, pre druhú 10,67% a pre tretiu 14,81%. Napriek tomu, väčšina chýb predstavovala omyl medzi príľahlými triedami, v druhej testovacej množine ostala úspešnosť rozlíšenia medzi silným dažďom a suchom na úrovni 99,5%, v tretej, najnáročnejšej, na úrovni 99,1%.

Yan, Luo a Zheng [12] vytvorili metódu schopnú rozpoznať jasné, oblačné a daždivé počasie z obrazu kamery so širokouhlým objektívom umiestnenej na pohybujúcom sa vozidle. Zvolili si tieto obrazové príznaky, na ktorých založili rozpoznávanie: histogram gradientnej amplitúdy (HGA, obrázok 2.3), histogram HSV farebného priestoru a údaj založený na

priemernej šedosti zvoleného výrezu cesty. Na rozpoznávanie použili Yan, Luo a Zheng algoritmus Real AdaBoost, ktorý kombinuje množstvo slabých klasifikátorov do jedného silného. Ďalej využívajú odôvodnený predpoklad, že hoci vo všeobecnej multitriednej klasifikačnej úlohe sú jednotlivé triedy na sebe logicky nezávislé a je potrebné použiť klasifikátory odlišujúce každú triedu od všetkých ostatných, v tejto konkrétnej úlohe, ak bude napríklad obrázok chybné zaradený do kategórie „daždivé“ klasifikátorom „slnечно-daždivo“, bude sa viac pravdepodobne v skutočnosti jednať o zamračený obrázok, než slnečný.



Obrázok 2.2: Vplyv množstva použitých príznakov na úspešnosť detekcie v štúdiu [2]. Je vidieť, že čím viac príznakov sa použilo na detekciu, tým úspešnejšia v priemere bola. Nie je to však všeobecne platný fakt, napríklad Chen a kolektív [1] dosiahli najvyššiu úspešnosť detekcie pri zvolení iba štyroch zo siedmich navrhnutých príznakov.

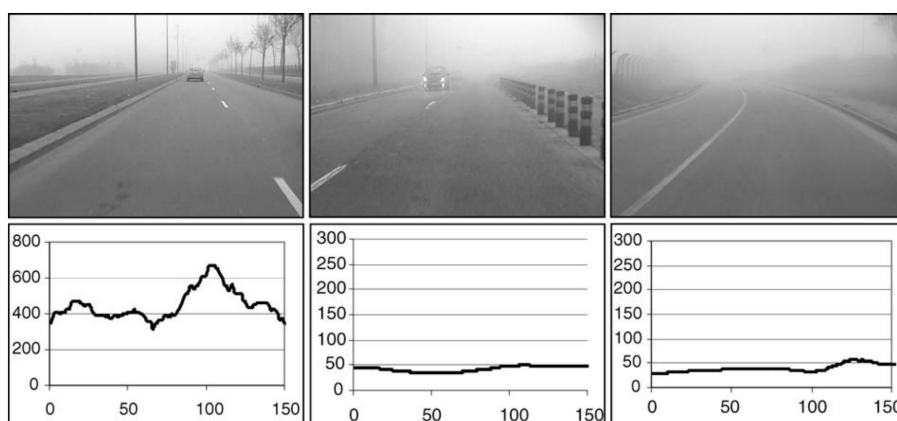


Obrázok 2.3: Znázornenie odlišnosti HGA v štúdiu [12] pri slnečnom (a) a daždivom (b) počasí. Hlavný rozdiel je v zastúpení hodnoty 0 – pri jasnom počasí predstavuje najpočetnejšiu kategóriu, pri daždivom sa maximum posúva do vyšších hodnôt.

Navrhli a využili teda schému, v ktorej prvom kroku obrázok prejde klasifikátorom „slniečno-daždivo“; ak ho ten označí za slnečný, bude ďalej postupovať do klasifikátora „slniečno-zamračené“, inak sa aplikuje klasifikátor „zamračené-daždivo“. Podarilo sa im takto dosiahnuť celkovú úspešnosť detekcie 91,92%.

2.2 Hmla

Hautiére, Tarel, Lavenant a Aubert [13] použili kameru umiestnenú vo vozidle na rozpoznávanie hmly a následný odhad viditeľnosti. Ich motiváciou bol fakt, že podľa výskumu [14] vodiči pri jazde v podmienkach so slabou viditeľnosťou zvyknú skutočnú viditeľnosť preceniť, čo vedie ku vyššej rýchlosti jazdy. Takýto systém by teda mohol vodiča automaticky varovať, že jeho rýchlosť je príliš vysoká na úspešnú reakciu na nový vizuálny podnet, napríklad prekážku na ceste. Za predpoklad si zvolili tvrdenie, že cesta je plochá. To im v kombinácii so známym uhlom kamery voči ceste a známym fyzickým rozmerom svetlocitlivých buniek v snímači kamery poslúžilo ako spôsob výpočtu vzdialenosti bodu v obraze od kamery. Tento výpočet následne kalibrovali pomocou scény so známou vzájomnou vzdialenosťou objektov (prerušovaná deliaca čiara v strede cesty). Následným aplikovaním metódy rozpínania regiónov (region growing) na detekciu plochy cesty, koschmiederovho zákona a ďalších sa im podarilo dosiahnuť „relatívne stabilné výsledky“, hoci presné vyhodnotenie v ich správe nebolo k dispozícii. Uvádzajú, že neistota odhadu viditeľnosti sa pohybovala na úrovni 5,5%.



Obrázok 2.4: Odhad viditeľnosti v troch obrázkových sekvenciách vo výskume [13] (vertikálna os predstavuje vzdialenosť v metroch, horizontálna poradové číslo analyzovaného obrázku). Je vidieť, že ich metóda je citlivá na prítomnosť hmly na ceste.

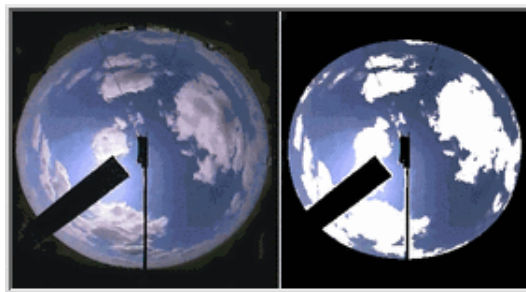
2.3 Oblačnosť

Ďalšou oblasťou súvisiacou s počasím, v ktorej sa využívajú kamerové snímky, je detekcia oblačnosti. Keďže automatické prístroje na určovanie miery oblačnosti už existujú, ale ešte nie sú bežne rozšírené, údaje, ktoré poskytujú meteorologické stanice napríklad na mapách s aktuálnym stavom počasia, sú zväčša určené priamo meteorológmi na meteorologických

staniciach. Miera oblačnosti sa udáva v desatinách (prípadne osminách) pokrytia oblohy, kde 0/10, alebo 0/8, vyjadruje bezoblačnú oblohu, a 10/10, prípadne 8/8 úplne zamračenú oblohu. Táto hodnota sa určuje odhadom a nemusí byť presná; presnosť odhadu skúseného meteorológa na praktické účely postačuje.

2.3.1 Hardwarové zariadenie na automatické určovanie oblačnosti

Hoci miera oblačnosti je meteorologická veličina, ktorá sa ako jedna z mála ešte aj v súčasnosti zvyčajne určuje ľudským pozorovaním, existujú už prístroje, ktoré ju dokážu určiť autonómne. Jedným príkladom je Total Sky Imager od spoločnosti Yankee environmental systems³. Pozostáva z hemisférického zrkadla ošetrovaného vrstvou blokujúcou priame slnečné svetlo a kamery, ktorá je namierená smerom dole na toto zrkadlo. Kamera sníma 360° obraz oblohy odrazený od zrkadla. Prístroj je plne automatický a obsahuje implementáciu algoritmu na odlíšenie oblakov od oblohy, vďaka ktorému dokáže presne vypočítať mieru oblačnosti na danom mieste.



Obrázok 2.5: Ukážka výstupu prístroja Total sky imager. Naľavo je pôvodný obrázok, napravo obrázok po aplikovaní detekcie oblakov. Zaciernený obdĺžnik zakrýva rameno držiace kameru.

Algoritmus zabudovaný v tomto prístroji je založený na prahovaní pomeru červenej a modrej farby každého pixelu – pixel s hodnotou tohto pomeru nižšou, než je prah, je označený za oblohu a naopak, pixel s hodnotou vyššou za oblak. Prah je dynamický a závisí na relatívnej polohe aktuálne skúmaného pixelu a slnka.

2.3.2 Využitie satelitných snímkov na určovanie oblačnosti

Ďalším spôsobom určovania miery oblačnosti, ktorý je už automatizovaný, je použitie satelitných snímkov z meteorologických vesmírnych družíc. Napríklad družica Landsat-7 agentúry NASA využíva algoritmus Automated Cloud-Cover Assessment (ACCA), popísaný v článku [4]. Tento spôsob však nie je aplikovateľný pre potreby meteorologických staníc, keďže tie potrebujú informácie pre rozličné miesta aktualizovať mnohokrát denne. Obežné dráhy satelitov ale nie sú ľubovoľné a toto zvyčajne neumožňujú.

3 Stránka spoločnosti: <http://www.yesinc.com/>

2.3.3 Využitie statickej kamery na účely rozpoznávania počasia podľa oblačnosti

Hoci špecializované prístroje na detekciu oblačnosti predstavujú hardwarové zariadenie, ktoré nie je bez nákupných a prevádzkových nákladov, a satelitné snímky sa na často aktualizované meteorologické účely nedajú použiť, podarilo sa mi nájsť iba jeden výskum [1], ktorý sa zaoberal konkrétne témou rozpoznávania počasia s využitím klasickej stacionárnej kamery s pohľadom na krajinu. Popíšem ho preto detailnejšie. Všetky ďalšie informácie v tejto sekcii budú čerpané z tohoto výskumu.

Chen, Yang, Lindner, Barrenetxea & Vetterli využili panoramatickú kameru umiestnenú na streche svojho ústavu. Tá snímala v pravidelných časových intervaloch 360° obraz scény obsahujúci zem a oblohu. Cieľom ich práce bolo automatizovať zaraďovanie obrázkov z ich kamery do troch ostro vymedzených kategórií – jasno, polooblačno a zamračené podľa tabuľky 2.1. Problém si špecifikovali takto:

- máme k dispozícii N obrázkov automaticky zozbieraných kamerou, z ktorých iba J ($J \ll I$) bude manuálne označených správnym štítkom (jasno, polooblačno, ...)
- všetky ostatné obrázky musia byť označené automaticky s vysokou istotou



(a) Jasno



(b) Polooblačno



(c) Zamračené

Obrázok 2.6: Príklady panoramatických obrázkov, s ktorými pracoval Chen a kolektív.

Počasia	Kritérium
Jasno	Menej než 50% oblakov
Oblačno	Medzi jasným a zamračeným
Zamračené	Nie je viditeľná žiadny modrá obloha

Tabuľka 2.1: popis typov počasia, na ktoré sa zameriaval Chen a kolektív

Keďže údaje o zvolených typoch počasia sa nachádzajú primárne v časti obrázku s oblohou, navrhli metódu extrakcie oblasti oblohy z obrázku za účelom odstránenia rušivých vplyvov popredia (budov, atď). Algoritmus, ktorý navrhli, využíva fakt, že v postupnosti za

sebou nasledujúcich obrázkov sa bude región oblohy meniť oveľa viac, než ostatné regióny. Prahovaním tejto zmeny získali masku oblohy. Týmto postupom určil Chen a kolektív vyše 98% oblohy správne ako oblohu.



Obrázok 2.7: Maska oblohy získaná pomocou zvoleného algoritmu z obrázkov za obdobie 3 dní (30 obrázkov).

Po určení segmentu oblohy rozdelili ďalší postup na extrakciu príznakov a strojové učenie pomocou *metódy podporných vektorov* (Support vector machines, SVM). Na extrakciu príznakov z obrázkov použili metódu „bag of words“ [5], pretože generuje priestorovo nezávislé príznaky, ktoré sa na tento problém hodia, pretože oblačnosť je takisto rozprestrená náhodne, a teda nezávisle. Tabuľka 2.2 ukazuje a popisuje zvolené príznaky.

Príznak	Popis	Typ
H, S, V	odtieň (hue), saturácia a jas (value) podľa farebného modelu HSV	farba
PHOW	Príznak vypočítaný pomocou algoritmu SIFT [6]	tvar
LBP	Lokálne binárne vzory v textúre [7]	textúra
Gradient	Rozsah gradientu vypočítaný pomocou Sobelových operátorov	textúra
Pohyb	Miera pohybu medzi po sebe nasledujúcimi obrázkami	dynamika

Tabuľka 2.2: popis príznakov, s ktorými pracoval Chen a kolektív

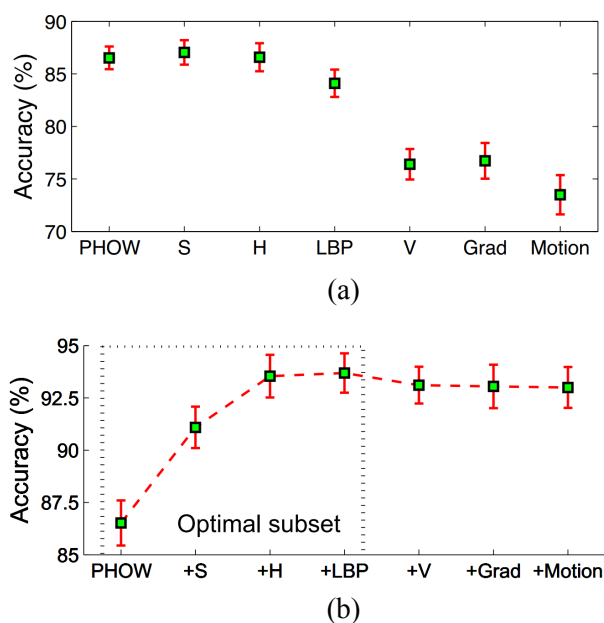
Pre voľbu optimálnej podmnožiny príznakov použili metódu multiple kernel learning (MKL) na zistenie váhy každého príznaku, podľa ktorej ich potom zoradili. Keďže príslušný váhový koeficient každého príznaku priamo predstavuje jeho pomerný prínos ku rozpoznaniu počas, môže byť použitý pre voľbu tejto optimálnej podmnožiny.

Keďže trénovaciu sadu museli označiť ručne, potrebovali ju mať relatívne malú a efektívnu. Zistili však, že pri náhodnom výbere trénovacej sady z plnej množiny obrázkov sa presnosť rozpoznávania medzi iteráciami výrazne líšila. Pre vylepšenie teda zvolili prístup, v ktorom sa trénovacia sada vyberá iteratívne, pričom trénovací algoritmus si počas učenia môže vyžiadať manuálne zaradenie vybraných obrázkov človekom. Tento princíp sa nazýva *aktívne učenie* [8].

Príznak	PHOW	S	H	LBP	V	Grad	Motion
Váha	0,308	0,302	0,300	0,282	0,278	0,275	0,268

Tabuľka 2.3: Váhové koeficienty jednotlivých príznakov vypočítaní pomocou MKL, zoradené zostupne podľa prínosu k detekcii.

Jeho základným princípom je, že počas procesu rozpoznávania SVM určuje vzdialenosť w_k medzi neoznačeným obrázkom I_k a deliacou nadrovinou. Keďže SVM počas trénovania hľadá najefektívnejšie nadroviny, w_k sa dá považovať za mieru neistoty zaradenia obrázku I_k . Po zotriedení všetkých nezaradených obrázkov podľa w_k použili tie s nízkymi hodnotami ako trénovaciu sadu v ďalšej iterácii. Pomocou popísaných postupov sa Chenovi a kolektívu podarilo pri manuálnom označení 20% obrázkov dosiahnuť úspešnosť klasifikácie 95%.



Obrázok 2.8: Presnosť rozpoznávania podľa príznakov na testovacej vzorke 500 náhodných obrázkov, bez aktívneho učenia. (a) Presnosť jednotlivých príznakov samostatne. (b) Úspešnosť rozpoznávania počas toho, ako doň boli postupne pridávané ďalšie príznaky.

3 **Návrh systému rozpoznávania oblačnosti a hmly z obrazu kamery**

Nedávne výskumy a štúdie venujúce sa rozpoznávaniu počasia a poveternostných podmienok výlučne z obrazu kamery sa zväčša zameriavali na asistenčné systémy vo vozidlách. Venovali sa detekcii hmly a odhadu viditeľnosti, odlíšeniu suchých podmienok od daždivých na základe rôznych poznatkov o ceste a oblohe dostupných z obrazu kamery – či už rozpoznávaním kvapiek na čelnom skle, alebo charakteru celého obrazu. Iba jedna štúdia, ktorú som našiel (štúdia [1]), sa zaoberala rozpoznávaním počasia z pohľadu statickej kamery. V tej sa však zaoberali iba jediným miestom a výstupom ich systému boli tri ostro ohraničené hodnoty. Hoci ich metódu by bolo istotne možné aplikovať aj na ďalšie kamery, túto možnosť netestovali, preto ich 95% úspešnosť môže byť mierne skreslená; rovnaký typ počasia sa môže na rozličných miestach javiť rozlične, prípadne ho môže byť ťažšie rozpoznať. Taktiež sa nezamerali na ďalší typ počasia, ktorý je vizuálne (aspoň pre človeka) spoľahlivo z obrázku možné rozlíšiť, a to hmlu.

3.1 **Cieľ práce**

Navrhujem preto systém, ktorý bude pomocou viacerých jednoduchých obrazových príznakov schopný na snímkoch kamier umiestnených na rozličných miestach rozpoznať tieto typy počasia: hmlu, bezoblačné počasie, polooblačné počasie a zamračené počasie. Čiastočne budem vychádzať z prístupu, ktorý si zvolil Chen a kolektív [1], no použijem iné príznaky a budem priamo skúmať veľké množstvo kamier. Navyše, moje riešenie nebude poskytovať iba ostro ohraničené predpovede, ale pravdepodobnostné hodnoty príslušnosti každého obrázku do všetkých štyroch typov počasia. Z nich bude samozrejme možné získať ostrú hodnotu počasia (tá s najvyššou pravdepodobnostnou hodnotou), no zároveň bude možné v grafickej podobe vykresliť jeho priebeh počas dňa.

3.2 **Špecifikácia podmienok na obraz kamery potrebný pre úspešné rozpoznávanie**

Chen a kolektív [1] použili metódu extrakcie masky oblohy z obrázka, pretože rozpoznávali iba typy počasia, o ktorých sa všetky informácie nachádzali v oblasti oblohy; štruktúry na zemi (budovy, kopce, atď.) pre nich teda predstavovali skôr zdroj dezinformácií, než užitočných dát.

Pridaním hmly do množiny rozpoznávaných typov sa však situácia mení. Tá sa totiž prejavuje hlavne tak, že zakrýva štruktúry vyskytujúce sa na zemi, preto túto oblasť obrazu nemôžem ignorovať. Zvolil som teda nasledujúcu požiadavku na orientáciu kamery, ktorej čo

najbližšie dodržanie bude predpokladom pre úspešné rozpoznanie všetkých typov počasia: horizont by sa mal nachádzať čo najbližšie stredu obrázku.

Toto obmedzenie je nevyhnutné preto, lebo hoci väčší pomer oblohy v obraze by poskytol viac informácií na presnejšie rozpoznanie jasného, polooblačného a zamračeného počasia, zanechal by veľmi malú oblasť v ktorej by sa na obrázku dala pozorovať prítomnosť hmly. A naopak – hmla by sa dala rozpoznať tým presnejšie, čím viac by sa horizont blížil hornému okraju obrázku, no tým by sa viditeľná oblasť oblohy zmenšila na kúsok nad horizontom, ktorý je fyzicky veľmi vzdialený skutočnému umiestneniu kamery. Nie je nezvyčajné, ak sa počas jasného dňa okolo horizontu nachádza oblačnosť, tá však nepredstavuje oblačnosť v mieste jej pozorovania. Ešte doplním, že oveľa väčšiu nepresnosť spôsobí, ak viac než polovicu obrázku tvorí zem, než ak je to obloha. Pri namierení kamery do zeme totiž býva často obloha tak výrazne preexponovaná, že v nej nie je možné rozoznať stav oblačnosti. Ak však miernu väčšinu snímku zaberá obloha, výrazný problém pre úspešnosť rozpoznávania to nepredstavuje.

Touto odôvodnenou požiadavkou sa zároveň stratila nutnosť počítať masku oblohy. Zároveň sa aj veľmi často zhoduje, alebo blíži, skutočnej orientácii kamier umiestnených na meteorologických staniciach. Preto nepredstavuje obmedzenie, ktoré by významne znížilo praktický význam navrhovanej metódy.

3.3 Voľba obrazových príznakov pre rozpoznávanie

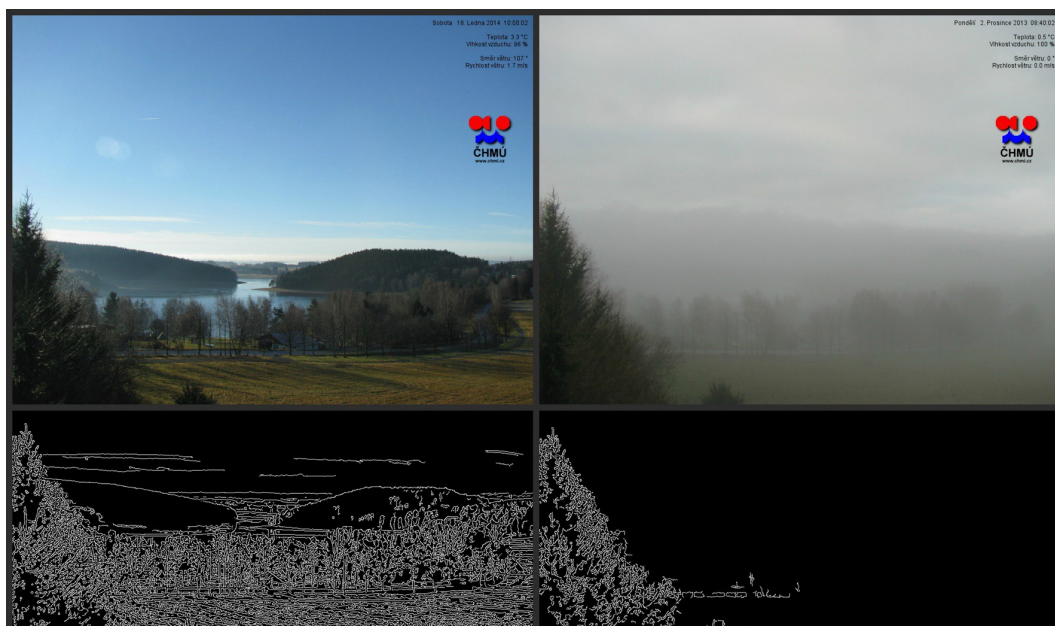
Pri voľbe príznakov pre rozpoznávanie som vychádzal z toho, podľa čoho pri pohľade na obrázok rozpozná počasie človek. Na základe tejto úvahy som zvolil nasledujúcich päť príznakov.

3.3.1 Miera hrán v spodnej polovici obrázku

Keďže najprv som skúmal rozpoznávanie hmly, toto bol prvý príznak, ktorý som zvolil. Na obrázkoch s hmlou je vizuálne charakteristické hlavne to, že čím je táto hustejšia, tým viac zakrýva objekty v blízkosti kamery a rozmazáva, prípadne úplne skrýva ich hrany. Preto som zvolil tento príznak ako promile pixelov, ktoré predstavujú hrany, zo všetkých pixelov v spodnej polovici obrázku. Promile som namiesto percent zvolil preto, aby som dosiahol jemnejší rozptyl hodnôt.

Pôvodne som tento príznak počítal z celého obrázku, pretože na obrázkoch, ktoré som mal dovtedy k dispozícii hmla vždy, keď bola prítomná, zakrývala aj časť obrázku, kde sa mala nachádzať obloha. Neskôr sa však vyskytli aj situácie prízemnej hmly, ktorá bola síce hustá, no oblohu zanechala na obrázku viditeľnú. Preto som využil predpoklad horizontu v strede obrázku, a výpočet promile hrán som aplikoval len na jeho spodnú polovicu.

Očakávam, že tento príznak bude zohrávať kľúčovú úlohu v odlíšení obrázkov s hmlou od všetkých ostatných.



Obrázok 3.1: Ukážka aplikácie detekcie hrán na obrázok s jasným počasím a obrázok s hmlou, oba z rovnakého miesta. Horný rad predstavuje pôvodné obrázky, dolný rad ukazuje hrany rozpoznané v spodnej polovici týchto obrázkov. Obrázok s hmlou obsahuje 36‰ hrán, zatiaľčo obrázok s jasným počasím až 187‰.

3.3.2 Kontrast v obrázku

Ďalším znakom, u ktorého som pri skúmaní obrázkov zistil, že sa dá použiť na odlišenie istých typov počasia, bol kontrast plôch v obrázku. Silný kontrast vznikol hlavne pri zamračenom počasí. Pri ňom totiž obloha predstavovala často súvislú svetlú plochu a zem súvislú tmavú. Expozícia autonómnych kamier, akými sú tie, s obrázkami ktorých som pracoval, je spravidla ponechaná v automatickom režime. Keďže však takéto kamery zvyčajne majú veľmi malé snímače (diagonálny rozmer štvrt' palca až palec), majú relatívne nízky dynamický rozsah. Ten sa prejaví tak, že oblasť mimo oblohy, ktorá je výrazne tmavšia, než zamračená obloha, bude výrazne tmavšia aj na obrázku. To vytvorí silný kontrast, ktorý je možné využiť pri rozpoznávaní počasia. Hoci takýto vysoký kontrast môže teoreticky vzniknúť aj pri jasnom a polooblačnom počasí, prítomnosť hmly celkový kontrast v scéne znižuje, preto by tento príznak mal byť použiteľný hlavne na odlišenie hmly od ostatných typov počasia. Tento príznak bude vypočítaný nasledovne:

- 1) obrázok sa rozdelí na väčšie množstvo rovnako veľkých dlaždíc
- 2) osobitne sa vypočíta priemerný jas každej dlaždice
- 3) priemerný jas všetkých dlaždíc sa sčíta a podelí ich počtom, čím sa získa priemerný jas obrázku
- 4) vypočíta sa rozdiel medzi jasom každej dlaždice a priemerným jasom obrázku
- 5) nakoniec sa sčíta tento rozdiel u všetkých dlaždíc a podelí sa ich počtom
- 6) výsledné číslo bude teda predstavovať priemernú odchýlku jasu dlaždíc od priemerného jasu obrázku

Možný rozsah hodnôt tohto príznaku je 0 (pri obrázku s pixelmi s jedinou hodnotou jasu) až 127 (ak by sa každá dlaždica trafila buď do súvislej bielej alebo čiernej plochy). Predpokladám, že tento príznak posluží hlavne na výraznejšie odlišenie obrázkov so zamračeným počasím, prípadne s hmlou, od ostatných.



Obrázok 3.2: Porovnanie kontrastu v obrázkoch z rovnakého miesta pri jasnom, polojasnom, zamračenom počasí a hmlle. Obrázok s jasným počasím má kontrast 39, obrázok s polojasným počasím 41, obrázok so zamračeným počasím 66 a obrázok s hmlou 31.

3.3.3 Jas obrázku

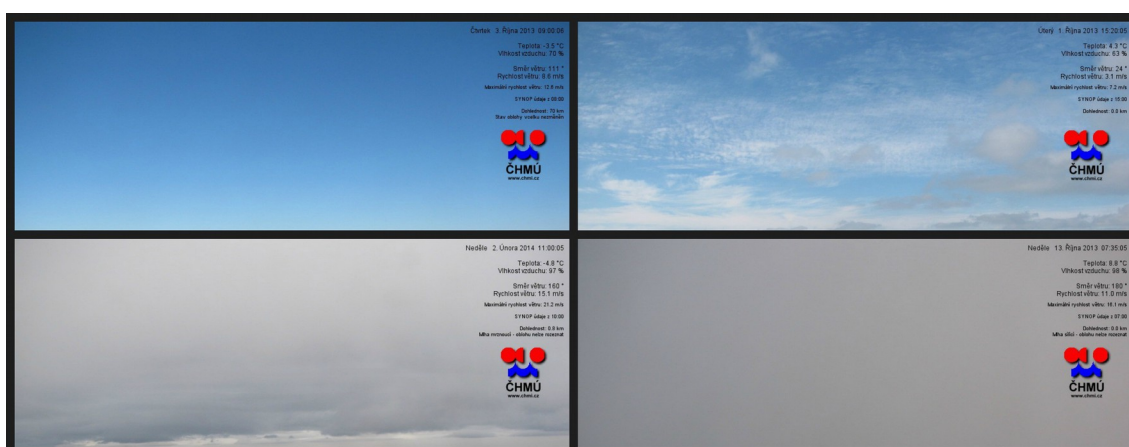
Bude sa počítať ako priemerný jas všetkých pixelov obrázku – súčet hodnôt jasu všetkých pixelov obrázku podelený ich počtom. Tento príznak bude užitočný hlavne na vyradenie príliš tmavých obrázkov z detekcie. Keďže kamery používajú automatickú expozíciu, nepredpokladám, že rôzne typy počasia budú mať odlišný priemerný jas. Rozsah hodnôt jasu vo farebnom modeli RGB je 0 až 255.

3.3.4 Saturácia oblohy

Saturácia je intenzita odtieňa farby, miera jeho odlišnosti od šedého tónu. Tento príznak bude využívať predpoklad, že pri jasnom počasí bude farba oblohy výrazne saturovaná a s pribúdajúcou oblačnosťou bude priemerná saturácia pixelov klesať. Pri počítaní tohoto príznaku sa bude pracovať iba s hornou polovicou obrázka. Saturácia pixelu vo farebnom modeli RGB sa počíta podľa nasledujúceho vzorca:

$$S = \begin{cases} 0 & R=G=B \\ 255 \frac{(M-m)}{(M+m)} & L < 128 \\ 255 \frac{(M-m)}{(511-(M+m))} & L \geq 128 \end{cases} \quad (3.1)$$

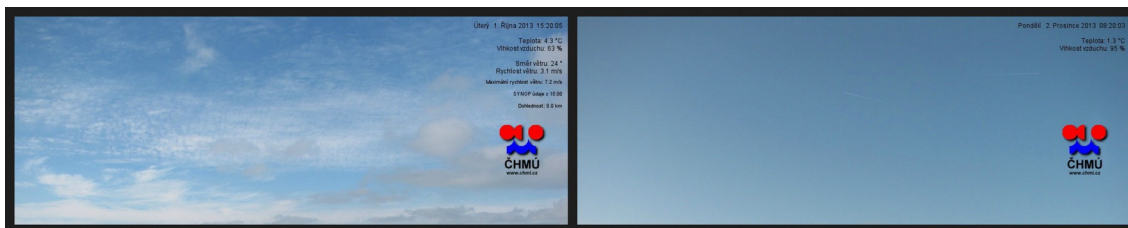
R , G a B predstavujú hodnoty červeného, zeleného a modrého subpixelu. M predstavuje maximálnu hodnotu spomedzi týchto subpixelov, m predstavuje minimálnu hodnotu. L predstavuje jas pixelu. Rozsah hodnôt saturácie vo farebnom modeli RGB je 0 až 255. Očakávam, že tento príznak bude užitočný hlavne v odlíšení obrázkov s jasným a polooblačným počasím od tých so zamračeným počasím a hmlou.



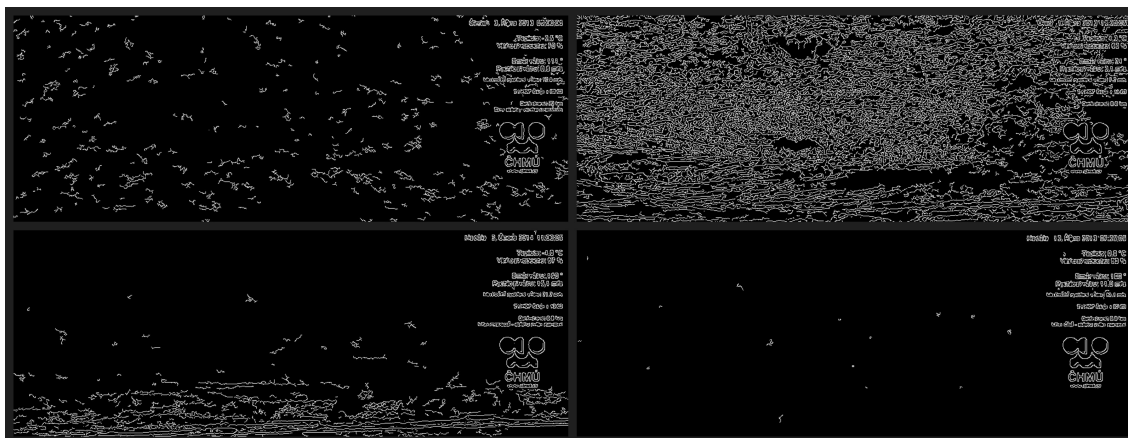
Obrázok 3.3: Porovnanie saturácie oblohy na obrázkoch z rovnakého miesta pri jasnom, polojasnom, zamračenom počasí a hmle. Obloha na obrázku s jasným počasím má priemernú saturáciu pixelov 135, obrázok s polooblačným počasím 90, obrázok so zamračeným počasím 12 a obrázok s hmlou iba 8.

3.3.5 Miera hrán v oblohe

Tento posledný príznak bude predstavovať promile pixelov z hornej polovice obrázku, ktoré predstavujú hrany. Jeho účelom bude výraznejšie odlíšiť polooblačné počasie od všetkých ostatných. Kým saturácia výraznejšie odliší jasné a polooblačné počasie od zamračeného a hmly, nebude pravdepodobne na jej základe spoľahlivo odlíšiť jasné a polooblačné počasie, pretože aj pri jasnej oblohe nemusí táto byť tak silno saturovaná, ako polooblačné počasie nasnímané v inom čase. Preto využijem predpoklad, že polooblačná obloha bude obsahovať výrazne viac hrán, než typická obloha pri ostatných rozpoznávaných typoch počasia.



Obrázok 3.4: Ukážka prípadu, kedy má úplne jasná obloha nižšiu saturáciu, než polooblačná. Obrázok vľavo má priemernú saturáciu 90, obrázok napravo iba 78.



Obrázok 3.5: Porovnanie miery hrán v oblasti oblohy na obrázkoch z rovnakého miesta pri (zľava doprava, zhora dolu) jasnom, polooblačnom, zamračenom a hmlistom počasí. Obloha na obrázku s jasným počasím má obsahuje 18% hrán, obloha s polooblačným počasím 198%, obloha so zamračeným počasím 39% a obloha s hmlou 13%.

3.4 Zhrnutie rozpoznávacej funkcie jednotlivých príznakov

V tejto podkapitole zhrniem funkciu jednotlivých príznakov, aby bolo jasné, kvôli čomu som ich zvolil a čo od nich očakávam.

Príznak	Očakávaný účel
Hrany v spodnej časti obrázku	Výrazné odlišenie obrázkov s hmlou od všetkých ostatných
Kontrast	Pomoc v odlišení hmly od všetkých ostatných obrázkov, mierne odlišenie zamračeného počasia od všetkých ostatných
Jas	Hlavne na vylúčenie príliš tmavých obrázkov (napríklad z noci), potenciálny mierny vplyv aj na samotnú detekciu
Saturácia oblohy	Výrazné odlišenie jasného počasia od zamračeného a hmlistého, menej výrazné, ale dôležité odlišenie polooblačného počasia od zamračeného a hmlistého
Hrany v oblohe	Výrazné odlišenie polooblačného počasia od všetkých ostatných

Tabuľka 3.1: Zhrnutie významu jednotlivých príznakov.

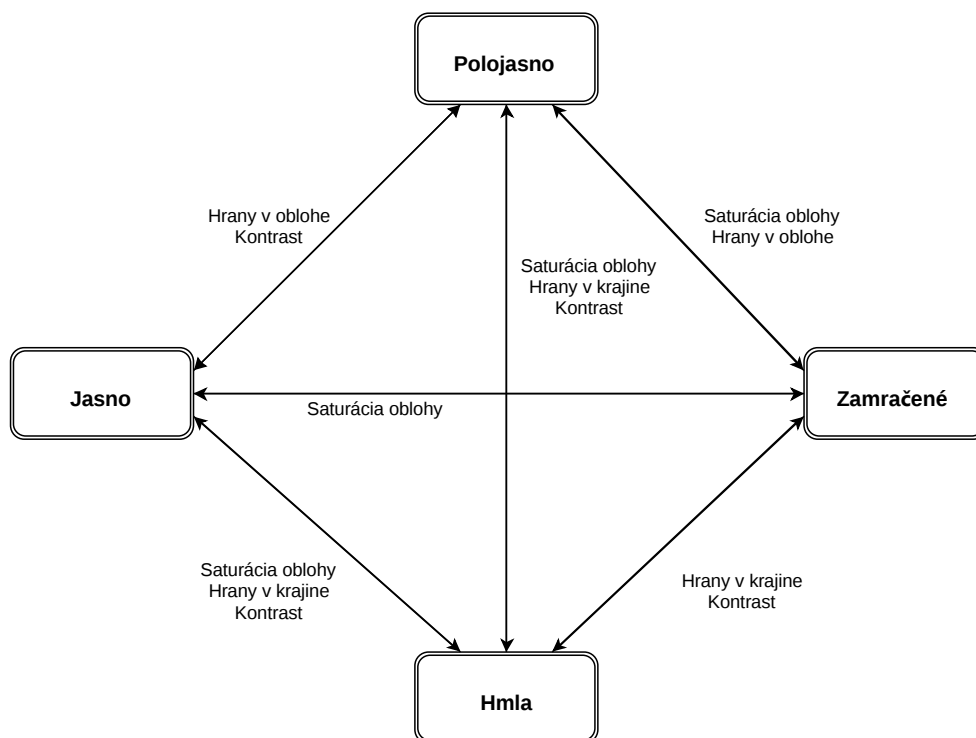


Schéma 3.1: Grafické znázornenie spôsobu odlišenia jednotlivých typov počasia. Popis šípok medzi dvoma typmi počasia vyjadruje príznaky, u ktorých je predpoklad, že budú tieto dva typy najviac odlišovať.

4 Implementácia navrhutej metódy

V tejto kapitole popíšem spôsob, ktorým som navrhnutú metódu rozpoznávania počasia implementoval. Budem sa venovať výberu a zberu pracovnej sady obrázkov, ich kategorizácii, výberu programovacích jazykov a knižníc a ukázkam spustenia, vstupov a výstupov jednotlivých častí.

4.1 Voľba, zber a kategorizácia pracovnej sady

Pre účely vývoja a testovania systému som používal verejne dostupné obrázky z webových kamier Českého hydrometeorologického ústavu, dostupné na URL ⁴. Vybral som 28 z nich, ktoré najviac spĺňali podmienku na polohu horizontu a zároveň spoľahlivo a pravidelne aktualizovali obrázky. Sú to kamery z týchto miest: Bělátn, Brdy, Brno, Doksany, Dyákovice, Holešov, Hradec Králové, Cheb, Churáňov, Klatovy, Klínovec, Kopisty, Kostelní Myšlová, Labská Bouda, Ostrava, Luční Bouda, Lysá Hora, Maruška, Olomouc, Ostrava – Poruba, Praha – Libuř, Příbram, Příbyslav, Přímá, Seč, Šerák, Vlkonice a Vsetín.

Pre potreby automatického zberu týchto obrázkov som vytvoril krátky skript v programovacom jazyku Python. Pôvodne som nastavil jeho pravidelné spúšťanie v časových úsekoch korešpondujúcich s intervalom aktualizácie online obrázkov, ktorý je pre väčšinu kamier 5 minút. Neskôr som však využil fakt, že obrázky z aktuálneho a dvoch predošlých dní sú spätne dostupné pomocou URL obsahujúceho dátum aj čas. Skript som teda zmenil na spätný dávkový zber všetkých obrázkov z dňa určeného parametrom.

Príklad spustenia skriptu:

```
py downloadDataFast.py -o=ibp/imgs -d=20140428 -t=0719 -f=15
```

Skript po spustení s týmito parametrami stiahne a uloží do zložky ibp/imgs všetky obrázky z 28. apríla 2014 zachytené v čase od 7:00 do 18:45 s intervalom 15 minút.

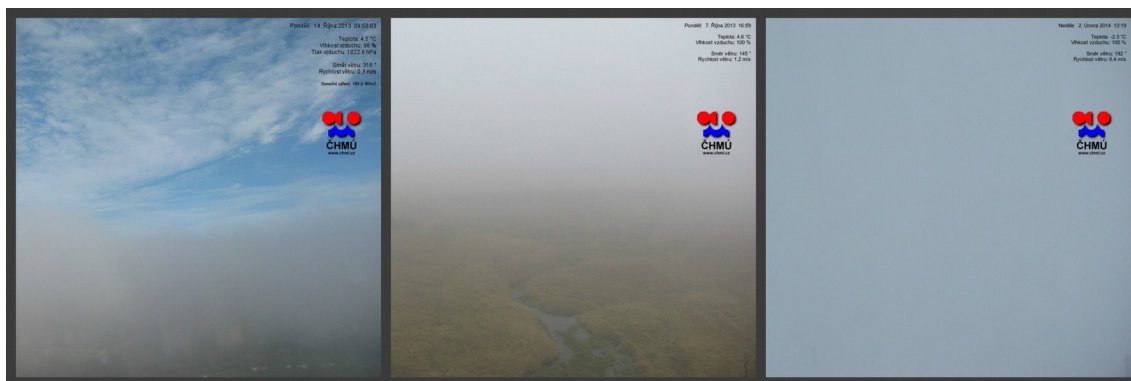
Parameter	-o	-d	-t	-f
Popis	Zložka na umiestnenie sťahovaných obrázkov	Dátum v tvare YYYYMMDD	Hodiny začiatku a konca zberu v tvare HHHH	Interval medzi snímkami v minútach

Tabuľka 4.1: Popis parametrov skriptu na zber testovacích obrázkov.

4 <http://www.chmi.cz/files/portal/docs/meteo/kam/index.html>

Všetky zozbierané obrázky bolo samozrejme potrebné zaradiť do kategórií podľa počasia, aby ich bolo možné využiť pre strojové učenie. Na začiatku som využíval údaje o aktuálnom počasi z URL ⁵, ktoré skript na zber obrázkov automaticky sťahoval ku každému obrázku a ukladal ich do textového súboru. Veľmi skoro som však zistil, že aktuálnosť a presnosť týchto údajov je absolútne nedostatočná pre úspešné použitie na automatické vytvorenie učiacej a testovacej sady. Informácie sa často aktualizovali až po niekoľkých hodinách a pre viacero miest ani neboli dostupné. Preto som zvolil ručnú klasifikáciu zozbieraných obrázkov. Bolo to síce časovo náročné, keďže veľkosť testovacej sady sa pohybovala nad 20 000 obrázkami, no zaručil som si tak potrebnú presnosť. Obrázky som zaraďoval do štyroch zvolených kategórií počasia výlučne.

Hoci pri kategóriách jasno, polojasno a zamračené je výlučné zaradenie prirodzené, pri hmle je to inak. Je samozrejme možné, aby sa hmla vyskytovala súčasne s jasnou, polooblačnou, alebo zamračenou oblohou, no moja výlučná kategorizácia to pri tvorení učiacej sady neumožňuje zohľadniť; obrázky s hmlou som vždy zaradil do kategórie hmla, bez ohľadu na stav oblačnosti. Prakticky je toto obmedzenie dobre odôvodnené – keďže som všetky obrázky zaraďoval ručne, videl som, že v drvivej väčšine prípadov prítomnosti hmly v obrázku nie je obloha vôbec viditeľná. Opačné prípady sa však objavili. Keďže však môj klasifikátor bude poskytovať pre každý testovaný obrázok pravdepodobnosť jeho príslušnosti do každej zo 4 tried, predpokladám, že napriek zaradeniu týchto prípadov výlučne do kategórie „hmla“ sa prítomnosť jasnej alebo polojasnej oblohy v obrázku s hmlou prejaví na rozložení tejto pravdepodobnosti medzi hmlu a príslušný stav oblohy.



Obrázok 4.1: Príklad obrázku, kde napriek prítomnosti hmly vidno aj oblohu (vľavo). Tento prípad je však veľmi zriedkavý, oveľa častejšie vyzerá obrázok s hmlou ako obrázok v strede (hmla obklopuje kameru aj zhora) alebo napravo (nie sú viditeľné absolútne žiadne prvky krajiny).

4.2 Implementácia výpočtu obrazových príznakov

Keďže som pri implementácii a testovaní svojho systému chcel stráviť čo najmenej času častým opakovaným prekladom mierne zmeneného zdrojového textu, zvolil som si pri výpočte obrazových príznakov modulárny prístup. Každý príznak počíta osobitný program v jazyku C++ s využitím knižnice OpenCV. Návratová hodnota každého tohoto programu priamo predstavuje

⁵ <http://www.wunderground.com/>

hodnotu príznaku, ktorý počíta. Tieto návratové hodnoty preberá a celú réžiu má na starosti skript `computeFlags.py` v programovacom jazyku Python.

Ten na vstupe dostane cestu ku zložke, v ktorej sa nachádzajú obrázky, pre ktoré je potrebné vypočítať príznaky. Táto zložka musí obsahovať podzložky pomenované názvom miesta, ktorého obrázky obsahuje. Tieto podzložky môžu obsahovať buď priamo súbory (ak ide o nezaradené obrázky), alebo podzložky pomenované „fog“, „clear“, „partly cloudy“ alebo „overcast“ obsahujúce už samotné, ručne zaradené, obrázky. Jeho výstupom je súbor `flags.txt`, umiestnený v zložke s obrázkami, ktorý obsahuje pre každý obrázok jeden riadok v nasledujúcom formáte:

```
cesta_k_súboru:: hrany jas kontrast saturácia_oblohy hrany_v_oblohe
```

Čo sa týka konkrétnej implementácie samotných programov na výpočet príznakov, popíšem iba výpočet hrán, pretože konkrétna implementácia ostatných neobsahuje žiadne špecifické úseky.

Program na výpočet hrán, či už v spodnej časti obrázku, alebo v oblohe, využíva Cannyho hranový detektor z knižnice OpenCV so spodným prahom 4 pre hrany v oblohe, respektíve 8 pre hrany v spodnej časti obrázku. Tieto hodnoty som zvolil experimentálne, keď obe predstavujú optimálnu hodnotu pre odlišenie oblačného počasia, respektíve hmly, od ostatných typov počasia. Pri výpočte pomeru hrán sa najprv obrázok prevedie na šedotónový, následne sa naň aplikuje redukcia šumu s maskou o veľkosti 3×3 a nakoniec sa vykoná samotná detekcia hrán. Tá vytvorí maticu so zhodnými rozmermi ako mal pôvodný obrázok, kde biele pixely predstavujú hranu. Následne sa ich počet podelí celkovým rozmerom matice a výsledné desatinné číslo sa vynásobí 1000. Toto nové číslo predstavuje promile hrán v obrázku a zaokrúhlené na jednotky predstavuje návratovú hodnotu programu.

4.3 Implementácia strojového učenia

Pre natrénovanie rozpoznávania počasia som zvolil metódu strojového učenia s učiteľom, a to konkrétne metódu SVM z knižnice `scikit-learn` pre Python [3] s lineárnym jadrom.

Klasická klasifikácia počíta s binárnym problémom – vzorka buď patrí do istej triedy, alebo do nej nepatrí. Mnou riešený problém však obsahuje viac, než dve triedy, preto som použil multitriednu klasifikáciu. Trieda SVM knižnice *sklearn* implementuje túto situáciu pomocou prístupu „každý proti každému“; ak N je počet tried, potom sa interne vytvorí $(N \times (N-1))/2$ binárnych klasifikátorov – pre každú dvojicu tried jeden. Tie sa potom pri požiadavke na určenie triedy nového prvku spustia všetky a ich výsledok sa agreguje, takže navonok sa klasifikátor správa ako jednoliaty.

Keďže medzi jednotlivými miestami sa charakter obrázkov s rovnakým typom počasia často značne líši, vytvoril som namiesto jedného veľkého klasifikátora podporné vektory pre každé miesto zvlášť.

5 Testovanie a vyhodnotenie navrhnutej metódy

V tejto kapitole predstavím a popíšem experimentálne výsledky testovania navrhnutého systému, ktorý som teoreticky popísal v kapitole 3 a detaily jeho implementácie predstavil v kapitole 4. Pri vyhodnocovaní som sa zamerlal hlavne na úspešnosť rozpoznávania a na dosiahnutie čo najmenšieho pomeru výrazne nesprávnej klasifikácie – rozpoznať obrázok s jasným počasím ako polooblačný je menej závažná chyba, než rozpoznať ho ako zamračený. Rýchlosť rozpoznávania nebola prioritou, keďže metóda nie je určená na použitie v reálnom čase a stav počasia sa mení najrýchlejšie v rádoch minút.

5.1 Pracovná sada

Pre zrozumiteľnosť uvádzam, že v ďalšom texte budem pojmom *pracovná sada* označovať všetky obrázky, ktoré som zozbieral, pojmom *trénovacia sada* tie obrázky z pracovnej sady, ktoré budú použité na natrénovanie klasifikátora, a pojmom *testovacia sada* tie, na ktorých budem overovať úspešnosť. Prienik trénovacej a testovacej sady bude pri každom teste prázdna množina.

Obrázky do pracovnej sady som zbieral priebežne od septembra 2013 do marca 2014. Ako je možné vidieť z tabuľky 5.1, sada nie je vyvážená; jednotlivé typy počasia nemajú rovnaké početné zastúpenie. Dôvodom je, samozrejme, fakt, že nie všetky typy počasia sa v období zberu obrázkov vyskytovali rovnako často. Žiaden typ som nepreferoval, vždy som roztriedil všetky snímky z jedného dňa. Toto je však v strojovom učení častá komplikácia, pretože pri výraznej početnej prevahe vzoriek jednej triedy nad druhou (ostatnými) má klasifikátor sklon dávať najväčšiu váhu pri rozpoznávaní tej najpočetnejšej a najmenej početnú zanedbáva. To je, samozrejme, nevyhovujúce, pretože početnosť snímok z jednotlivých počasi nijak neurčuje ich dôležitosť; rozpoznanie všetkých je rovnako dôležité. Problém nastáva aj pri určovaní celkovej úspešnosti rozpoznávania – ak napríklad trénujeme binárny klasifikátor na rozpoznávanie jabĺk od hrušiek na trénovacej sade s 95 obrázkami jabĺk a 5 obrázkami hrušiek, klasifikátor môže dosiahnuť 95% úspešnosť aj bez toho, aby vedel rozpoznať čo i len jedinú hrušku od jablka.

Tento problém sa však výrazne prejavuje hlavne pri silne nevyváženej trénovacej sade, hlavne ak je počet vzoriek niektorej triedy vyslovene nedostatočne veľký na úspešné natrénovanie. V mojom prípade však aj najmenej početná trieda „jasno“ obsahuje viac, než 2000 vzoriek, čo by malo samo o sebe stačiť na úspešné natrénovanie rozpoznávania tejto triedy. Navyše, implementácia metódy SVM, ktorú používam, umožňuje nastaviť automatickú váhu vzorkám z jednotlivých tried inverzne proporcionálnu k početnosti triedy, čo tento problém efektívne eliminuje. Toto nastavenie využívam.

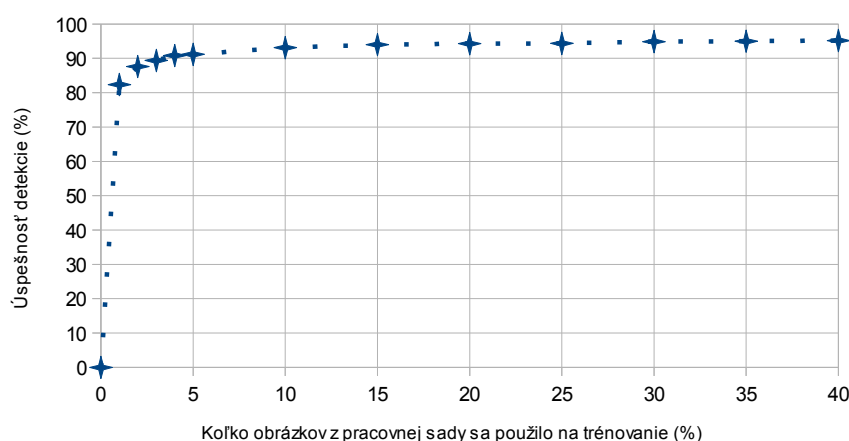
Aj tak však budem okrem celkovej úspešnosti zvlášť uvádzať aj úspešnosť pre každý typ počasia, aby nedošlo k jej skresleniu úspešnosťou rozpoznania najpočetnejšej triedy.

Zamračené	Hmla	Polojasno	Jasno	Dohromady
11916	5481	5392	2054	24843
47,97%	22,06%	21,70%	8,27%	100%

Tabuľka 5.1: Počet a pomer obrázkov s jednotlivými typmi počasia v pracovnej sade.

5.2 Vplyv veľkosti trénovacej sady na úspešnosť rozpoznávania

Keďže moja pracovná sada pozostáva z veľkého množstva obrázkov, potreboval som vedieť, koľko z nich je potrebných na natrénovanie klasifikátora na najvyššiu možnú úspešnosť, po dosiahnutí ktorej už pridávanie ďalších nebude mať na ňu vplyv. Preto som v tomto experimente spúšťal klasifikátor s postupne sa zvyšujúcim počtom trénovacích snímok. Ten je v grafe zobrazený ako percento snímok použitých na trénovanie klasifikátora z celej pracovnej sady.



Graf 5.1: Úspešnosť klasifikácie pri postupnom zvyšovaní počtu trénovacích obrázkov. Každý dátový bod predstavuje priemer vypočítaný z desiatich spustení testovacieho skriptu.

Z grafu je vidieť, že už pri použití 4% (~1000) obrázkov z pracovnej sady na učenie presiahne celková úspešnosť klasifikácie 90%. Postupným pridávaním sa dostane až na hodnotu 95% pri použití 35% obrázkov na učenie, nad ktorú už ďalej nestúpa, preto som ďalšie dátové body do grafu ani nevykresľoval. Hoci na zvýšenie úspešnosti o posledných 5% je potrebný sedemnásobný počet trénovacích obrázkov, oplatí sa to, pretože čas potrebný na ich natrénovanie je aj tak menší, než 2 sekundy.

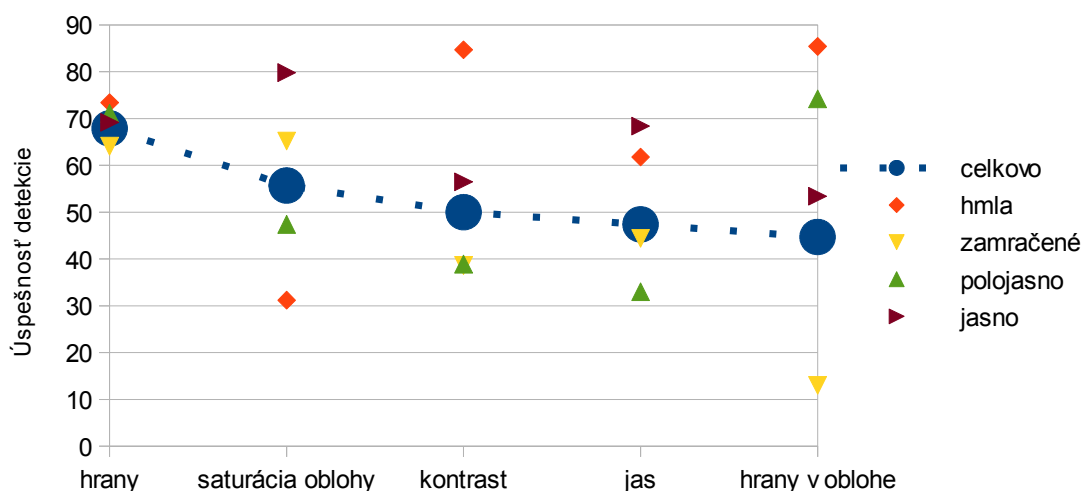
Je potrebné dodať, že do trénovacej sady sa pre každé miesto dostalo náhodným výberom toľko obrázkov z každého typu počasia, aký pomer z celkového počtu obrázkov z daného miesta predstavovali. Uvediem príklad: pracovná sada pre mesto Brno obsahuje 1306 obrázkov, z toho 119 (9,1%) bezoblačných, 266 (20,4%) polooblačných, 905 (69,3%) zamračených a iba 16 (1,2%) s hmlou. To je spôsobené tým, že na rozličných miestach sa v období zberu snímok jednotlivé typy počasia vyskytovali s výrazne odlišnou frekvenciou. To sa prejaví tak, že pri použití 35% obrázkov na trénovanie sa napríklad pri Brne použije priemerne okolo 315

obrázkov so zamračeným počasím, no iba 6 s hmlou. Keďže však mnou použitá implementácia SVM zohľadňuje tieto nepomery medzi triedami, na výsledkoch sa to negatívne neprejavuje.

Pri každom spustení skriptu overujúceho úspešnosť sa teda pracovná sada náhodne rozdelí na trénovaciu a testovaciu sadu v pomere 35:65, pretože zvyšovaním pomeru trénovacej sady nad 35% sa už priemerná úspešnosť rozpoznávania nezvyšuje ani o desatiny percent a zároveň citeľne stúpa čas potrebný na učenie. Jeho znižovaním by sa zase dalo ušetriť isté množstvo času pri učení, no rozpoznávanie by nedosiahlo úplný vrchol úspešnosti.

5.3 Vplyv jednotlivých príznakov na úspešnosť rozpoznávania

Podobne ako Chen a kolektív [1] som sa rozhodol preskúmať vplyv jednotlivých príznakov na úspešnosť detekcie. Spúšťal som ju teda postupne s každým príznakom zvlášť a do grafu som vykreslil celkovú úspešnosť i samostatnú úspešnosť rozpoznania každého z nich. Zoradením príznakov podľa úspešnosti rozpoznávania pri ich výlučnom použití sa ukáže, ktorý z nich detekcii pomáha najviac a ktorý najmenej.



Graf 5.2: Úspešnosť rozpoznávania pri použití každého príznaku samostatne. Príznyky sú zoradené zostupne podľa celkovej úspešnosti. Každý dátový bod predstavuje priemer vypočítaný z desiatich spustení testovacieho skriptu.

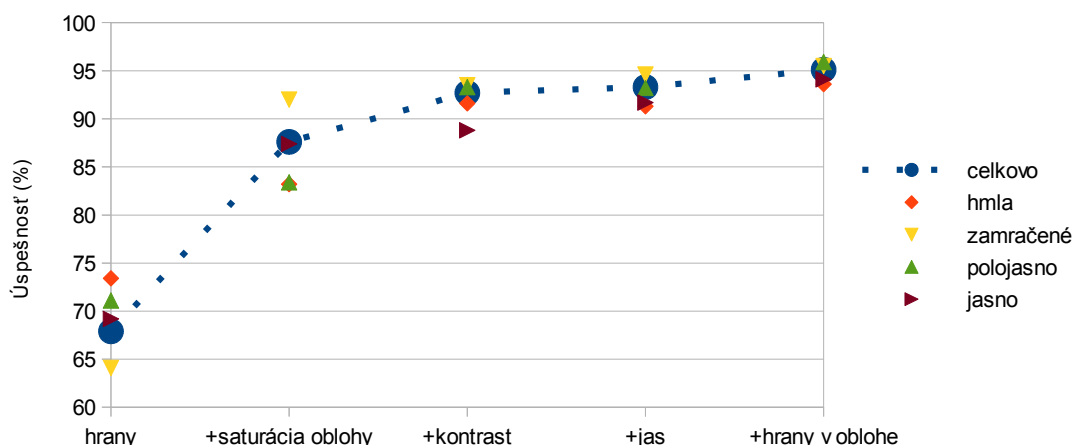
Z grafu je vidieť, ako sa niektoré predpoklady uvedené v tabuľke a schéme 3.1 ukázali ako pravdivé a niektoré nie. Napríklad predpoklad, že hmlu bude najvýraznejšie odlišovať miera hrán v spodnej časti obrázku nebol správny; mierne viac ju odlišuje kontrast a miera hrán v oblohe. Prekvapilo ma tiež, že saturácia oblohy má na odlíšenie hmly taký malý vplyv. Vysvetľujem si to tým, že mnohé obrázky s hmlou majú nesprávne nastavené vyváženie bielej farby – sú zafarbené domodra. Zároveň vďaka faktu, že sa v miere hrán v oblohe až tak výrazne neodlišujú od bezoblačného počasia, a tomu, že ich najväčšou „konkurenciou“ sú obrázky so zamračenou oblohou, ktoré predstavujú zároveň najpočetnejšiu triedu v pracovnej sade, saturácia oblohy až tak nezavážila. Na odlíšenie zamračeného počasia má prekvapivo najvyšší vplyv saturácia oblohy, nie kontrast, ako som predpokladal. Tiež ma prekvapilo, že jas sa

rozpoznávacou schopnosťou vôbec výrazne nelíši od ostatných príznakov a nie je dokonca ani najslabším z nich.

Na druhej strane sa veľmi správnym ukázal predpoklad, že bezoblačné počasie bude najvýraznejšie odlišovať saturácia oblohy. Podobne sa opodstatneným ukázalo očakávanie, že polojasné počasie bude najviac od ostatných odlišovať miera hrán v oblohe.

5.4 Úspešnosť rozpoznávania pri postupnom pridávaní príznakov

Nasledujúci experiment slúžil na zistenie, ako sa bude klasifikátor správať s postupným pridávaním príznakov. Motiváciou bolo zistiť, či sú pre najväčšiu presnosť potrebné všetky, či prípadne pridávanie niektorých bude mať taký malý vplyv, že bude lepšie ušetriť čas ich počítania a vôbec ich do systému nezahrnúť, alebo či sa dokonca bude úspešnosť po istom vrchole s ich postupným pridávaním znižovať, ako sa to stalo v štúdiu [1]. Príznačky som postupne pridával podľa ich priemerného vplyvu na rozpoznávanie, ktorý som zistil v predchádzajúcom experimente.



Graf 5.3: Úspešnosť klasifikácie pri postupnom pridávaní príznakov. Každý dátový bod predstavuje priemer vypočítaný z desiatich spustení testovacieho skriptu.

Z grafu je jasne viditeľné, že na čo najlepšiu klasifikáciu je potrebné použiť všetky navrhnuté príznaky, čím sa dosiahne *celková úspešnosť tesne nad 95%*. Úspešnosť rozpoznania každého počasia jednotlivo aj celková úspešnosť sa pridaním každého príznaku zvýšila, s výnimkou úspešnosti rozpoznania hmly po pridaní jasnu. Aj tá sa ale dostala na svoje maximum až po pridaní posledného príznaku. Ukázalo sa teda, že každý príznak má v tomto systéme svoje miesto a spolu dosahujú veľmi vysokú úspešnosť.

5.5 Znázornenie vlastností klasifikátora pomocou matice zámen

V tejto podkapitole predstavím a popíšem maticu zámen zobrazujúcu vlastnosti klasifikátora a budem sa venovať aj konkrétnym prípadom nesprávnej klasifikácie.

Matica, prípadne tabuľka zámen (v angličtine confusion matrix) sa často používa na znázornenie vlastností multitriednych klasifikátorov, pretože jednoduché percentuálne vyjadrenie úspešnosti o nej ani zďaleka neposkytuje celý obraz. Stĺpce tejto tabuľky označujú rozpoznanú triedu, zatiaľčo riadky tú skutočnú. Tým sa pre každú triedu znázorňujú prípady nesprávnej klasifikácie snímkov z nej do všetkých ostatných tried. V tejto tabuľke sa ľahko dajú nájsť chyby detekcie, pretože nimi sú všetky čísla mimo diagonály, ktorá predstavuje správne rozpoznané obrázky.

		Rozpoznaná trieda			
		Hmla	Zamračené	Polojasno	Jasno
Skutočná trieda	Hmla	3032 (93,4%)	150 (4,6%)	47 (1,4%)	15 (0,1%)
	Zamračené	148 (2,0%)	7191 (95,5%)	89 (1,2%)	103 (1,4%)
	Polojasno	40 (1,2%)	35 (1,0%)	3211 (95,7%)	71 (2,1%)
	Jasno	1 (<0,1%)	41 (3,2%)	43 (3,4%)	1179 (93,3%)

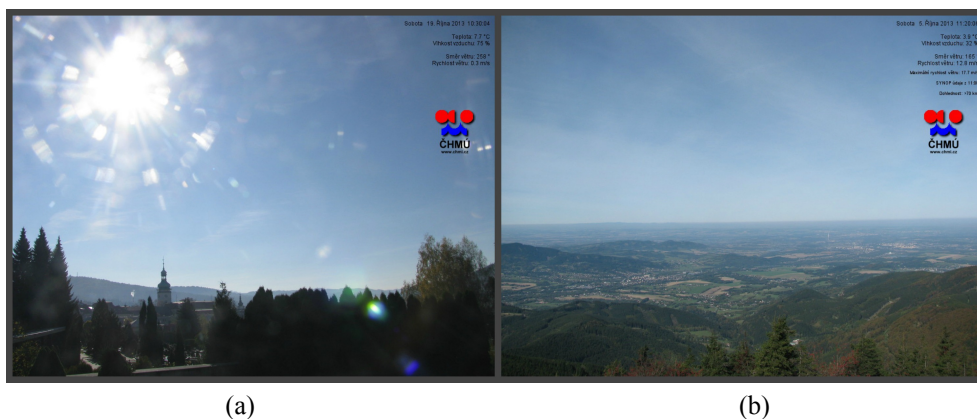
Tabuľka 5.2: Matica zámen implementovaného klasifikátora.

Zelenou farbou zvýraznená diagonála teda predstavuje snímky, ktoré boli klasifikátorom zaradené správne. Je z nej jasne vidieť, že klasifikátor dosahuje vysokú úspešnosť správnej klasifikácie pri nízkom pomere nesprávneho zaradenia. Navyše, aj napriek výrazne nevyváženej trénovacej sade sa úspešnosti rozpoznania jednotlivých tried líšia len veľmi mierne. To dokazuje, že použitá implementácia SVM sa s nevyváženosťou trénovacej sady vyrovnala veľmi dobre. Môže to však byť aj jednoduchým dôsledkom toho, že veľkosť trénovacej sady bola dostatočná na to, aby sa do nej aj z najmenej početnej triedy „jasno“ dostal dostatok snímkov na úspešné natrénovanie. Tým sa zrejme na druhej strane do nej dostal až nepotrebné vysoký počet snímkov z najpočetnejšej triedy „zamračené“, no keďže tréning trvala necelé dve sekundy a je ho na praktické účely (narozdiel od testovacích) potrebné vykonať iba raz, nepredstavuje to záležitosť, ktorú by bolo potrebné odstrániť.

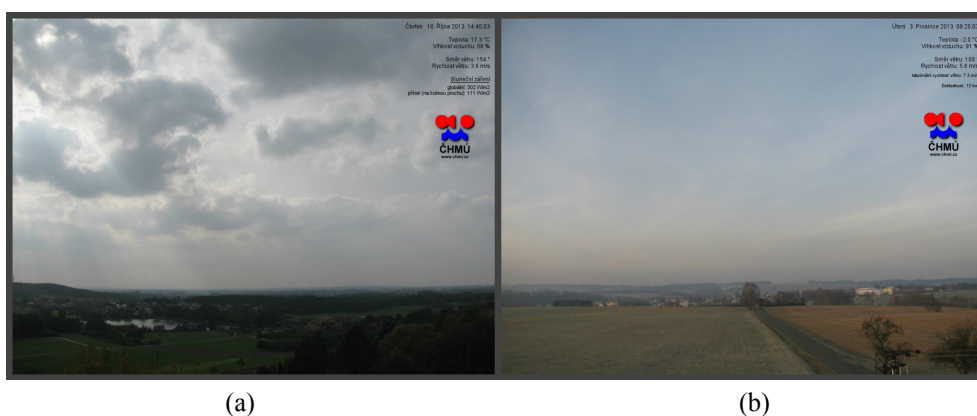
Ako som však už v predošlom texte uviedol, môj klasifikátor nezaraďuje snímky len do jednej triedy, práve naopak, uvádza príslušnosť každého snímku do každej triedy. To je nevyhnutnou vlastnosťou z dôvodu, že počasie sa nemení skokovito. Napríklad pri príchode oblačnosti na bezoblačnú oblohu sa najprv bude spojito znižovať miera jasného počasia na úkor polooblačného a neskôr miera polooblačného na úkor zamračeného. Pre účely vyhodnocovania úspešnosti klasifikácie však bolo potrebné považovať jednotlivé triedy za diskkrétne. Neviem si totiž predstaviť, ako by sa vyhodnocovala úspešnosť rozpoznania spojitely zmeny počasia na testovacej sade, ktorej už ručné diskkrétne roztriedenie bolo časovo veľmi náročné. Pre

zhotovenie tabuľky zámen som teda použil u každého snímku tú triedu, do ktorej mal najvyššie percento príslušnosti. Tým pádom sa medzi jednotlivými typmi vytvorila ostrá hranica, ktorá v skutočnosti neexistuje, čo určite spôsobilo isté nepresnosti vo vyhodnotení.

Oranžovou farbou zvýraznené bunky predstavujú snímky, ktoré boli síce klasifikované nesprávne, ale do podobných tried, napríklad bezoblačné počasie ako polooblačné, a preto nepredstavujú najvýraznejšie omyly (obrázky 5.1 a 5.2). Niektoré z týchto prípadov sú jednoducho dôsledkom toho, že niektoré obrázky boli na hranici dvoch susedných typov počasia a ja som ich tesne zaradil do jednej triedy, zatiaľčo klasifikátor do tej druhej. Iné však predstavujú typ snímkov, ktoré klasifikátor skutočne nevie pomocou navrhnutých príznakov spoľahlivo rozpoznať. Jedným z takýchto prípadov je situácia, ak sa na bezoblačnej oblohe nachádza v zábere slnko – to často vytvorí kvôli nedokonalostiam objektívu svetelné artefakty, ktoré predstavujú hrany mylne považované za oblaky. Tieto oranžovou farbou označené chyby sa vyskytujú len u približne 2% testovaných snímkov.



Obrázok 5.1: Obrázok s jasným počasím (a) nesprávne klasifikovaný ako polooblačný kvôli výskytu artefaktov spôsobených slnkom svietiacim do objektívu. Obrázok s polooblačným počasím (b) nesprávne klasifikovaný ako bezoblačný kvôli prítomnosti jemnej vrstvy oblačnosti bez výrazných hrán a dostatočnej saturácii.

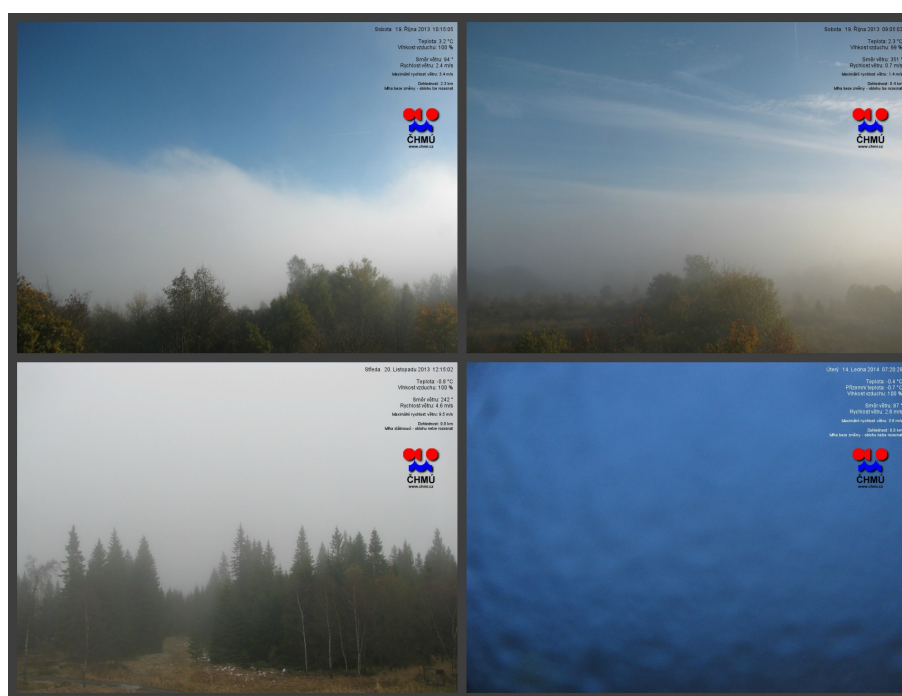


Obrázok 5.2: Obrázok so zamračeným počasím (a) nesprávne klasifikovaný ako polooblačný kvôli prítomnosti súvislej aj rozkúskovanej oblačnosti, ktorej hrany zmiatli klasifikátor. Obrázok s polooblačným počasím (b) nesprávne klasifikovaný ako zamračený kvôli nízkej saturácii oblohy a neprítomnosti výrazných hrán v oblohe.

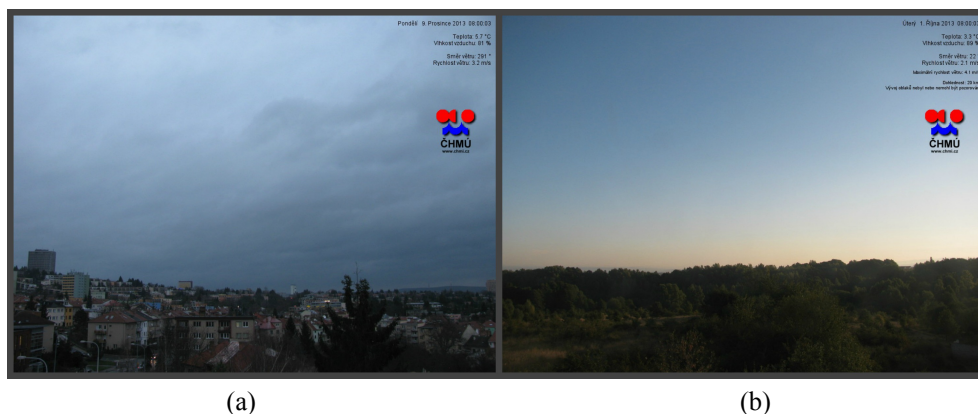
Bunky zvýraznené červenou farbou predstavujú snímky, ktoré boli zaradené úplne nesprávne – konkrétne zámenu medzi bezoblačným a zamračeným počasím (obrázok 5.4). Hoci tieto prípady sa vyskytujú len pre okolo 1,5% obrázkov s bezoblačným a zamračeným počasím, často sú to obrázky, ktoré na prvý pohľad nepredstavujú problém a podobné bývajú bežne zaradené správne. Tento problém sa mi nepodarilo odstrániť ani pridaním viacerých podobných správne zaradených obrázkov do trénovacej sady.

Mnohé z nich však majú známu príčinu. Napríklad množstvo obrázkov so zamračeným počasím zhotovených okolo východu alebo západu slnka, má nesprávne nastavené vyváženie bielej, takže má celá obloha modrú farbu. Ak je zároveň takáto oblačnosť súvislá, žiaden z použitých príznakov ju nedokáže odlišiť od bezoblačnej oblohy, keďže jej saturácia je často vyššia, než pri úplne bezoblačnej oblohe.

Bunky, ktoré predstavujú falošné zásahy detekcie hmly (obrázok 5.3), alebo jej nerozpoznanie som neoznačoval farbou podľa závažnosti chyby. Ako som už spomínal, hmla sa nevylučuje so žiadnym stavom oblačnosti, a keďže som obrázky s hmlou vždy zaradil do kategórie hmla, isté percento omylov je nevyhnutné. Ak klasifikátor zaradil obrázok s hmlou do iného počasia, bolo to buď podľa stavu oblohy (čiže prakticky správne), alebo bol obrázok tak veľmi nekvalitný, či už kvôli úplne chybnému vyváženiu bielej, prítomnosti kvapiek, alebo snehu na objektíve a podobne, že sa správne zaradenie ani nedalo očakávať. Tieto prípady teda nepredstavujú chybu klasifikácie, ale chybu vyhodnotenia úspešnosti. Navyše, pri zaradení obrázku s hmlou a zároveň viditeľnou oblohou isté percento príslušnosti pripadlo aj hmle, čo však vyhodnotenie maticou zámen nedokáže zohľadniť.



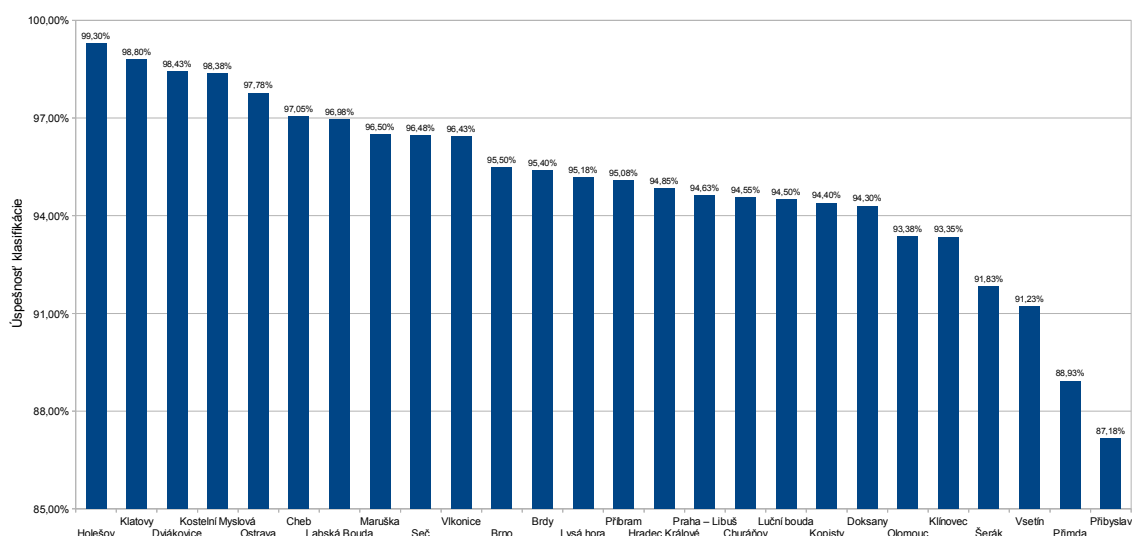
Obrázok 5.3: Príklady „nesprávnej“ klasifikácie obrázkov s hmlou podľa oblačnosti. Obrázok vľavo hore bol určený za bezoblačný, vpravo hore za polooblačný, vľavo dole za zamračený. Obrázok vpravo dole bol identifikovaný ako polooblačný.



Obrázok 5.4: Obrázok so zamračeným počasím (a) nesprávne klasifikovaný ako bezoblačný kvôli nesprávne nastavenému vyváženiu bielej. Obrázok s bezoblačným počasím (b) nesprávne klasifikovaný ako zamračený kvôli nízkej saturácii oblohy a neprítomnosti hrán v nej.

5.6 Úspešnosť rozpoznávania v jednotlivých miestach

Keďže som pre každé miesto vytváral vlastný klasifikátor, vyhodnotil som úspešnosť rozpoznávania aj v každom mieste zvlášť. Výsledky znázorňuje graf 5.4.



Graf 5.4: Úspešnosť rozpoznávania pre jednotlivé miesta. Každý stĺpec predstavuje priemernú hodnotu vypočítanú z desiatich spustení testovacieho skriptu.

Ako je vidieť, úspešnosť rozpoznávania klesla pod 90% iba na dvoch miestach, a to v Přimide a Příbyslave. Nie som si však celkom istý, prečo je ich úspešnosť tak výrazne nižšia, než u ostatných miest. Snímky z nich vyzerajú na prvý pohľad rovnako dobre rozoznateľne, ako tie z ostatných miest, a hoci u oboch je horizont pomerne výrazne posunutý pod polovicu snímku, podobne je to aj u viacerých ostatných miest, kde sa to negatívne neprejavilo.

5.7 Znázornenie priebehu počasia počas celého dňa

S využitím vlastností implementovaného klasifikátora, ktorá umožňuje určiť pre každý snímok pravdepodobnosť príslušnosti do každej z tried, som zostrojil grafy, ktoré ukazujú priebeh počasia počas jedného celého dňa. Pre každý bod na vodorovnej osi sú vykreslené štyri dátové body. Tieto predstavujú pravdepodobnosť príslušnosti snímku z daného času do každej zo štyroch rozpoznávaných tried počasia. Keďže tieto grafy si vyžadujú ručné overenie správnosti, ich celkovú presnosť som nedokázal vyhodnotiť. Mnoho z nich som ale preskúmal a výsledky boli vo väčšine prípadov uspokojivé. Znázornenie prechodu jedného typu počasia do druhého zväčša zodpovedalo skutočnej zmene na analyzovaných snímkoch. V grafoch bolo takisto dobre vidieť spojitosť zaradenia – často bola pravdepodobnosť počasia na snímku rozdelená rovnomerne medzi dve príslušné triedy (napríklad zamračené a polooblačno), čo zodpovedalo skutočnej situácii na snímku.

Príklady tohto výstupu som kvôli veľkosti obrázkov umiestnil do prílohy 1, 2 a 3. Je na nich dobre vidieť, ako veľmi je klasifikátor citlivý napríklad na zmiznutie hmly (v prílohe 1, 2), i na jej náhle objavenie (v prílohe 3). V prílohe 2 je dobre vidieť rozpoznanie chvíľkového rozjasnenia oblohy na konci dňa. V prílohe 3 je pekne znázornený postupný prechod z polooblačného do zamračeného počasia. Veľké množstvo ďalších grafov som umiestnil na priložené CD do zložky `graphs`.

6 Záver

V tejto práci som sa zaoberal rozpoznávaním počasia zo snímok statických kamier. Preštudoval som doterajšie výskumy v tejto oblasti, navrhol vlastný systém založený na detekcii hrán, kontraste, jase a saturácii, implementoval som ho a túto implementáciu som otestoval. Dosiahol som celkovú úspešnosť klasifikácie na úrovni 95%, pričom rozpoznanie jednotlivých typov počasia dosiahlo úspešnosť 93,3% pre bezoblačné počasie, 95,7% pre polooblačné počasie, 95,5% pre zamračené počasie a 93,4% pre hmlu.

Pri návrhu systému som postupoval podľa toho, na základe čoho vie počasie na obrázku rozoznať človek; ktoré znaky si všíma. Týmto uvažovaním som vytvoril skupinu piatich obrazových príznakov, z ktorých každý od seba výrazne odlišuje aspoň jednu dvojicu typov počasia. Ich spojením som dosiahol odlišenie každého počasia od každého. Princíp fungovania implementovaného systému je rozdelený do troch hlavných častí – zber a kategorizácia snímok, výpočet obrazových príznakov a strojové učenie.

Vytvoril som aj nástroj, ktorý dokáže vykresliť priebeh počasia na zadanom mieste v zadanom dni vo forme grafu. Výsledky tohto nástroja si je možné prezrieť na priloženom CD vo forme samotných grafov vizuálnym porovnávaním s analyzovanými obrázkami, alebo aj vo forme krátkeho videa. Taktiež som vytvoril prezentačný plagát, ktorého ukážka je uvedená v prílohe 4.

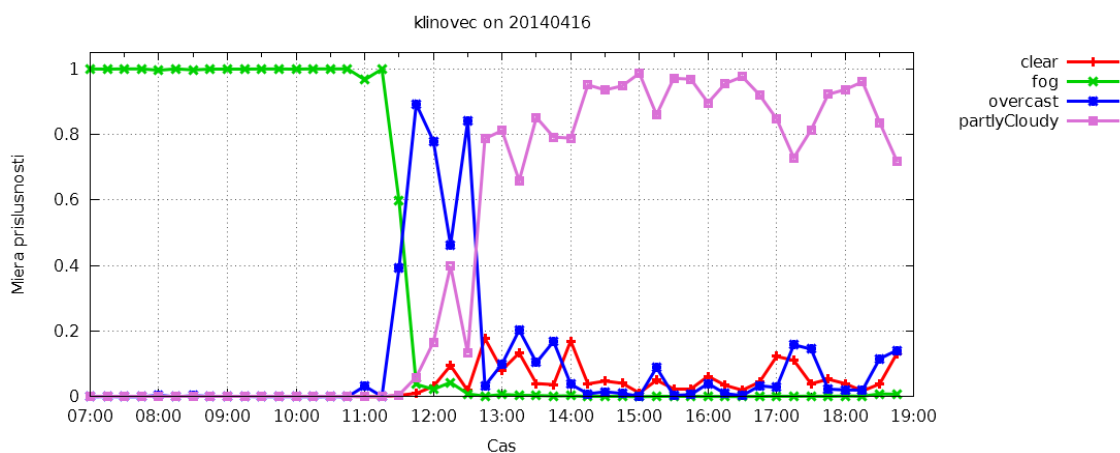
Za najväčší úspech svojej práce považujem fakt, že aj pri nevyváženej pracovnej sade a použití pomerne jednoduchých, na výpočet nenáročných príznakov, sa mi podarilo dosiahnuť veľmi vysokú úspešnosť. Som spokojný aj s tým, že vysokú úspešnosť som dosiahol pri všetkých štyroch skúmaných typoch počasia.

Najväčším prínosom vytvoreného systému je oproti doterajšiemu stavu v tejto oblasti to, že umožňuje vykresliť graf s priebehom počasia počas dňa, v ktorom je pre každý snímok uvedená pravdepodobnosť jeho príslušnosti do každého zo štyroch skúmaných typov počasia. Vďaka tomu je možné zobrazit' prirodzenú plynulú zmenu počasia. Táto možnosť sa podľa mojich znalostí v doterajších prácach v tejto oblasti nenachádza.

Čo sa týka ďalšieho vývoja projektu, rád by som do množiny rozpoznávaných charakteristík počasia pridal dážď a sneženie, prípadne prítomnosť snehovej pokrývky. Hoci samotné padanie dažďa sa mi nepodarilo vizuálne rozpoznať ani na jednom skúmanom obrázku, často sa prejavuje kvapkami na objektíve, ktoré by rozpoznať šlo. Ďalej by sa pri analýze videa namiesto statických obrázkov dal rozpoznávať napríklad vietor, a to podľa pohybu stromov, listov, prípadne iných objektov.

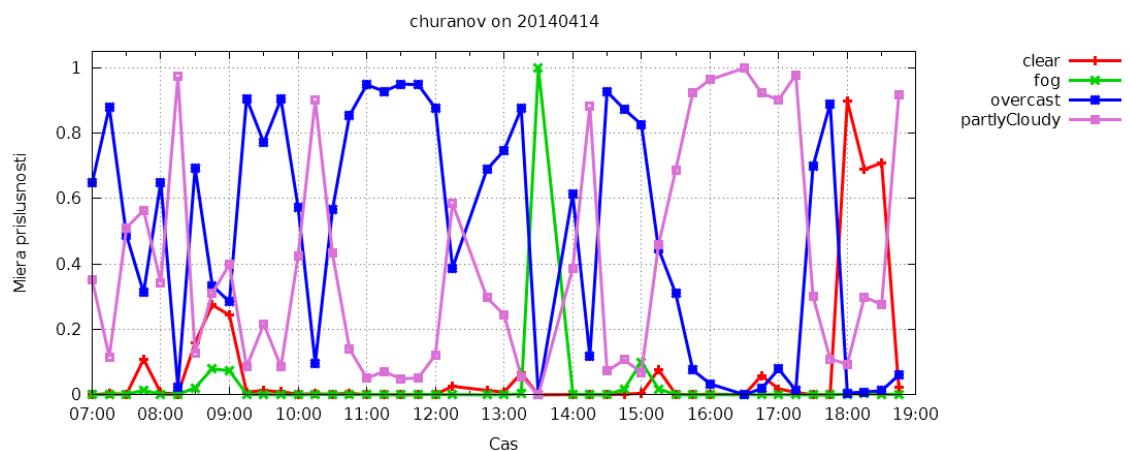
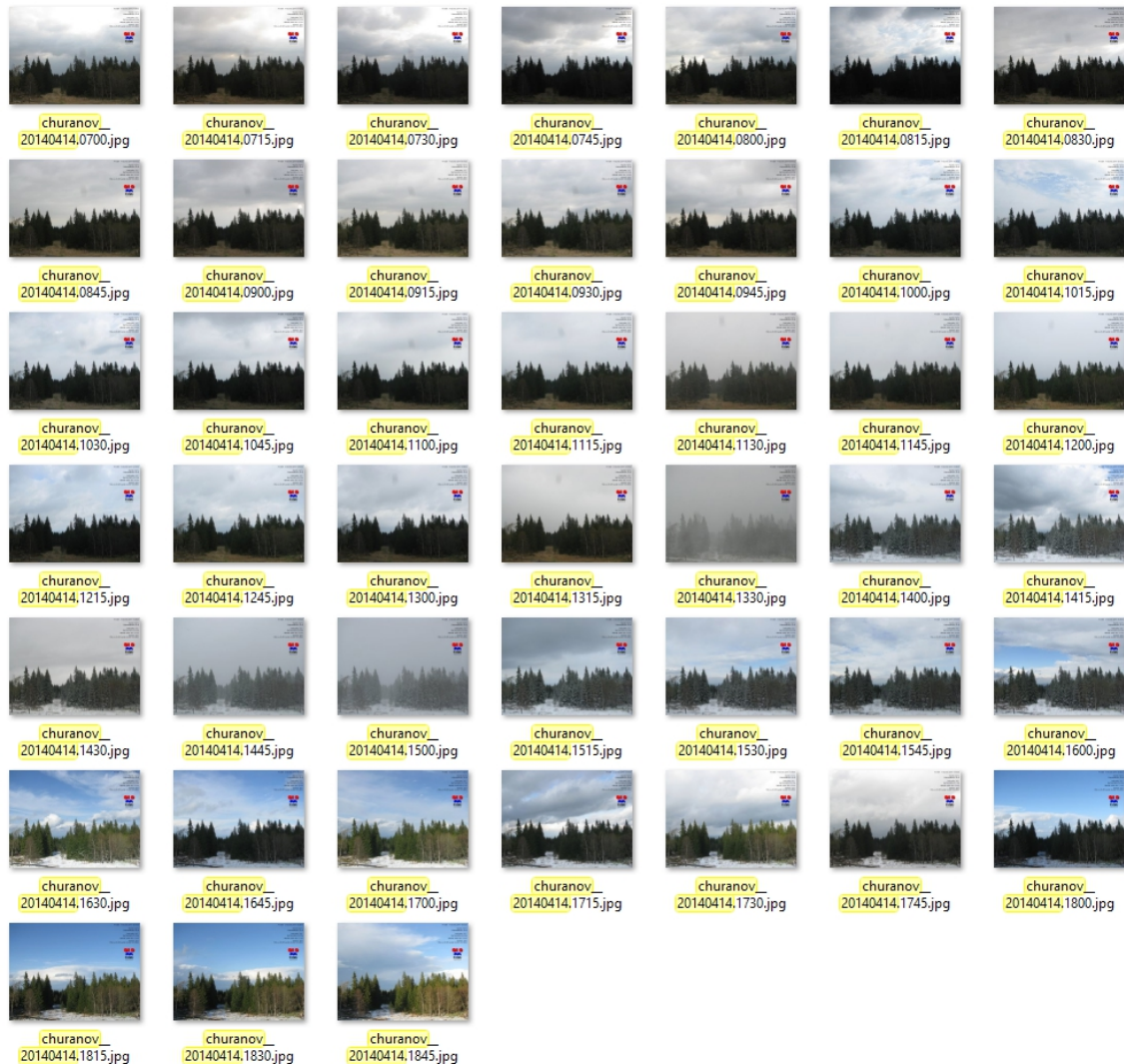
Príloha 1 – priebeh počasia, Klínovec

Znáznornenie priebehu počasia v mieste Klínovec dňa 16. apríla 2014. Intervaly medzi snímkami sú 15 minút.



Príloha 2 – priebeh počasia, Churáňov

Znázornenie priebehu počasia v mieste Churáňov dňa 14. apríla 2014. Intervaly medzi snímkami sú 15 minút.

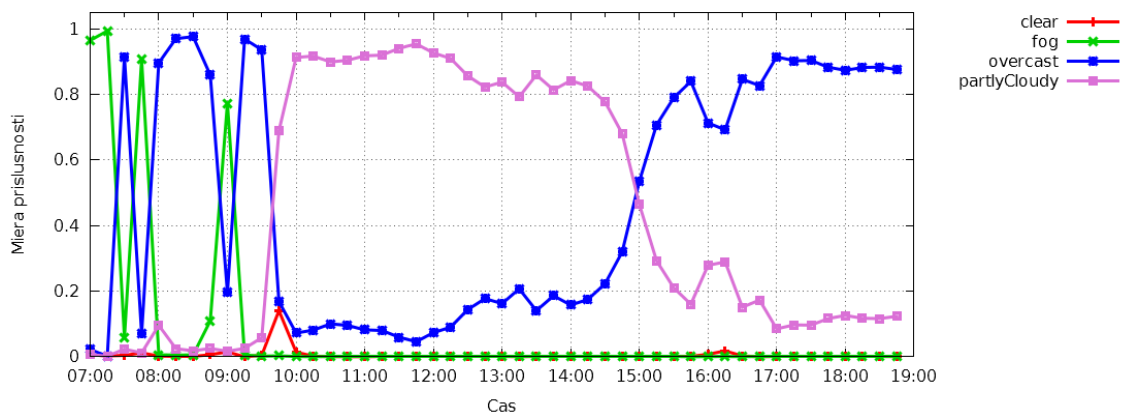


Príloha 3 – priebeh počasia, Klatovy

Znázornenie priebehu počasia v mieste Klatovy dňa 26. apríla 2014. Intervaly medzi snímkami sú 15 minút.



klatovy on 20140426



Príloha 4 – prezentačný plagát

Rozpoznávanie počasia z pohľadu vonkajšej kamery

Autor: Michal Jenčo

Vedúci: Doc. Ing. Adam Herout, Ph.D.

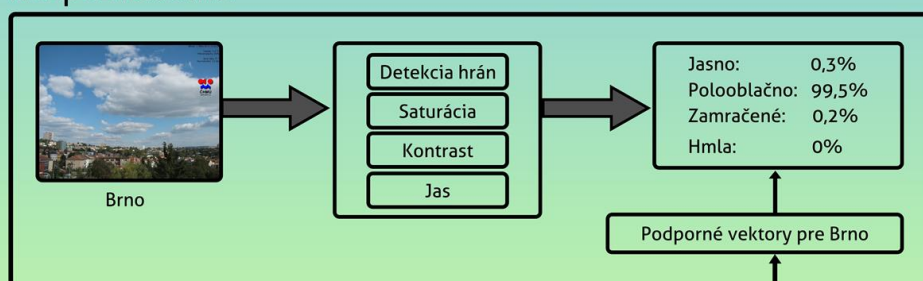
Cieľ

Rozpoznať v snímkoch z veľkého množstva kamier hmlu, jasné, polooblačné a zamračené počasie

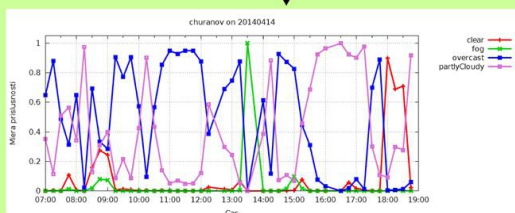
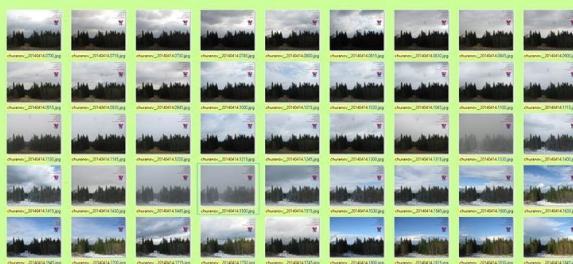
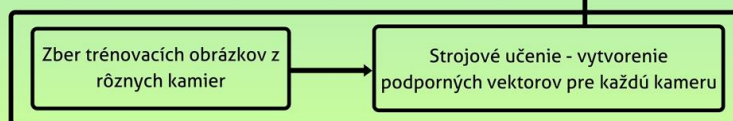
Úspešnosť

- Hmla: 93,4%
- Jasno: 95,5%
- Polooblačno: 95,7%
- Zamračené: 93,3%

Rozpoznávanie



Trénovanie



- rozpoznávanie natrénované pre 28 miest
- spojitý zariadenie - pre každý snímok sa počíta pravdepodobnosť príslušnosti ku každému počasiu
- vďaka tomu je možné vykresliť graf priebehu počasia počas dňa
- ten znázorňuje zmenu (napríklad príchod oblačnosti) postupne, nie náhlým skokom

Príloha 5 – obsah CD

<i>Source/</i>	- adresár so zdrojovými textami
<i>Images/</i>	- adresár so snímkami použitými pri učení a testovaní
<i>Graphs/</i>	- adresár s grafmi priebehov počasia
<i>Video/</i>	- adresár s videom prezentujúcim graf priebehu počasia

Literatúra

- [1] Chen Z., Yang F., Lindner A., Barrenetxea G., Vetterli M.: How is the weather: Automatic inference from images [online]. 2012. Dostupné z: <http://infoscience.epfl.ch/record/176274/files/weather.pdf>.
- [2] Roser M., Moosmann F.: Classification of weather situations on single color images [online]. *Intelligent Vehicles Symposium, 2008 IEEE*, s.798,803, 4-6. jún 2008. ISBN 978-1-4244-2568-6. Dostupné z <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=4621205>.
- [3] Pedregosa a kolektív: Scikit-learn: Machine Learning in Python, *JMLR* 12, s. 2825-2830, 2011.
- [4] Irish R., Barker J., Goward S., Arvidson T.: Characterization of the Landsat-7 ETM + Automated Cloud-Cover Assessment (ACCA) Algorithm [online]. 2006. Dostupné z: http://info.asprs.org/publications/pers/2006journal/october/2006_oct_1179-1188.pdf.
- [5] Csurka G., Dance C., Fan L., Williamowski J., Bray C.: Visual categorization with bags of keypoints, Workshop on statistical learning in computer vision, European Conference on Computer Vision, 2004, vol. 1, s. 22.
- [6] Bosch A., Zisserman A., Muoz X.: Image classification using random forests and ferns, *Proceedings of the 11th International Conference on Computer Vision (ICCV)*, 2007, s. 1-8 .
- [7] Ojala T.: A comparative study of texture measures with classification based on feature distributions, *Pattern recognition*, vol. 29, no. 1, s. 51-59, 1996.
- [8] Ojala T.: Active learning literature survey, *Computer Sciences Technical Report 1648*, University of Wisconsin–Madison, 2009.
- [9] McCall J.C., Trivedi M.M.: Video-based lane estimation and tracking for driver assistance: survey, system, and evaluation, *Intelligent Transportation Systems, IEEE Transactions on*, vol.7, no.1, s.20,37, marec 2006.
- [10] Bertozzi M., Broggi A., Fascioli A., Nichele S.: Stereo vision-based vehicle detection, *IEEE Intelligent Vehicles Symposium*, október 2000.
- [11] Kurihata H., Takahashi T., Ide I., Mekada Y., Murase H., Tamatsu Y., Miyahara T.: Rainy weather recognition from in-vehicle camera images for driver assistance, *Intelligent Vehicles Symposium, 2005. Proceedings. IEEE*, s.205,210, 6-8 jún 2005.
- [12] Xunshi Yan, Yupin Luo, Xiaoming Zheng: Weather recognition based on images captured by vision system in vehicle, In proceeding of: *Advances in Neural Networks*, 6th International Symposium on Neural Networks, 2009.

[13] Hautière N., Tarel J.-P., Lavenant J., Aubert D.: Automatic fog detection and estimation of visibility distance through use of an onboard camera [online], Machine Vision and Applications, 2006, s. 8-20. DOI 10.1007/s00138-006-0011-9. Dostupné z: http://download.springer.com/static/pdf/226/art%253A10.1007%252Fs00138-005-0011-1.pdf?auth66=1397506692_7a1011d7636746312f41eef1d4f4a76b&ext=.pdf

[14] Cavallo V., Colomb M., Dor J.: Distance perception of vehicle rear lights in fog, Hum Factors 43, s. 442-451, 2001.