Univerzitet u Banjoj Luci Elektrotehnički fakultet

Vladimir Risojević

Klasifikacija slika dobijenih tehnikama daljinske detekcije

doktorska disertacija

Banja Luka 2013.

Mojoj porodici.

Sadržaj

1	Uvo	od	1
	1.1	Problem analize snimaka dobijenih	
		daljinskom detekcijom	3
	1.2	Klasifikacija snimaka dobijenih tehnikama	
		daljinske detekcije	4
	1.3	Semantička klasifikacija slika	10
		1.3.1 Deskriptori slika	10
		1.3.2 Klasifikator	11
	1.4	Doprinos disertacije	12
	1.5	Organizacija disertacije	13
	1.6	Publikacije	13
2	Mat	terijal i metodologija	15
	2.1	Kolekcija slika Banja Luka	15
	2.2	Kolekcija slika Window on the UK	17
	2.3	Kolekcija slika UC Merced Land Use	17
	2.4	Kolekcija slika Wuhan	18
	2.5	Aero-snimak donjeg toka rijeke Sane	18
	2.6	Metodologija testiranja	21
3	Gał	oorov deskriptor teksture	22
	3.1	Modeliranje receptivnih polja prostih ćelija pomoću Gaborovih filtara	23
	3.2	Gaborov deskriptor teksture	27
	3.3	Klasifikacija pokrivača i načina korišćenja zemljišta korišćenjem Ga-	
		borovih deskriptora	27
		3.3.1 Kolekcija slika Banja Luka	29
		3.3.2 Window on the UK kolekcija slika	29
4	Pro	šireni Gaborov deskriptor teksture	32
	4.1	Statističke zavisnosti između podopsega Gaborove transformacije	32
	4.2	Prošireni Gaborov deskriptor teksture	33
	4.3	Redukcija dimenzionalnosti deskriptora	37
		4.3.1 Korišćenje koeficijenata iz susjednih podopsega	38
		4.3.2 Analiza glavnih komponenata	38
		4.3.3 Linearna diskriminantna analiza	39
	4.4	Eksperimentalni rezultati	40
		4.4.1 UC Merced Land Use kolekcija slika	40

		4.4.2 Wuhan kolekcija slika	. 46 51	
	4.5	Prošireni Gaborov deskriptor teksture za slike u boji	. 51 . 51	
		4.5.1 Oponentni prošireni Gaborov deskriptor teksture	. 53	
		4.5.2 Kvaternionski prošireni Gaborov deskriptor teksture	. 56	
	4.6	Eksperimentalni rezultati	. 59	
		4.6.1 UC Merced Land Use kolekcija slika	. 59	
		4.6.2 Wuhan kolekcija slika	. 63	
		4.6.3 Diskusija	. 67	
	4.7	Zaključak	. 68	
5	Stru	ıkturalni deskriptor teksture	70	
	5.1	Motivacija	. 70	
	5.2	Strukturalni deskriptor teksture	. 73	
	5.3	Mjera strukturne sličnosti	. 74	
	5.4	Klasifikacija slika	. 75	
	5.5	Eksperimentalni rezultati	. 77	
		5.5.1 UC Merced Land Use kolekcija slika	. 79	
		5.5.2 Wuhan kolekcija slika	. 82	
6	Fuzi	ija globalnih i lokalnih deskriptora	84	
	6.1	Lokalni deskriptori slike	. 84	
	6.2	Hijerarhijska fuzija deskriptora	. 86	
	6.3	Poređenje globalnih i lokalnih deskriptora	. 88	
		6.3.1 UC Merced Land Use kolekcija slika	. 89	
		6.3.2 Wuhan kolekcija slika	. 90	
	6.4	Eksperimentalni rezultati za hijerarhijski klasifikator	. 92	
		6.4.1 UC Merced Land Use kolekcija slika	. 93	
		6.4.2 Wuhan kolekcija slika	. 96	
		6.4.3 Detekcija urbanih područja	. 99	
	6.5	Diskusija i zaključci	. 103	
_				
7	Zak	ljučak	106	
	7.1	Pregled naučnih doprinosa disertacije	. 108	
	7.2	Budući rad	. 110	
Α	Kor	elacije odziva Gaborovih filtara	111	
В	Gist	t deskriptor scene	114	
С	C SIFT deskriptor		116	
D	թոր	olikacije kandidata u vezi sa disertacijom	118	
т.	itoratura			
	teratura 119			

Skraćenice

BoW	Skup riječi (<i>Bag-of-words</i>)
CORINE	Coordination of Information on the Environment
CW-SSIM	Mjera strukturalne sličnosti zasnovana na kompleksnim talasićima (Complex wavelet structural similarity)
DA	Diskriminantna analiza
EGTD	Prošireni Gaborov deskriptor teksture (<i>Extended Gabor Texture</i> Descriptor)
GLCM	Gray Level Co-occurrence Matrix
GTD	Gaborov deskriptor teksture (Gabor Texture Descriptor)
HOG	Histogram orijentisanih gradijenata
LBP	Lokalni binarni uzorak (Local Binary Pattern)
LDA	Linearna diskriminantna analiza
Lidar	Light Detection and Ranging
NN-STSIM	Nearest Neighbor Structural Texture Similarity
ODD	Deskriptor razlika orijentacija (Orientation Difference Descriptor)
PCA	Analiza glavnih komponenata (Principal Component Analysis)
RBF	Radijalna bazna funkcija (Radial Basis Function)
RDA	Regularizovana diskriminatna analiza
RGB	Red-Green-Blue
SAR	Synthetic Aperture Radar
SDD	Deskriptor razlika skala (Scale Difference Descriptor)
SIFT	Scale-Invariant Feature Transform
STD	Strukturalni deskriptor teksture (Structural Texture Descriptor)

SKRAĆENICE

SSIM	Mjera strukturalne sličnosti (Structural Similarity Measure)
STSIM	Strukturalna sličnost tekstura (Structural Texture Similarity)
SVM	Mašina sa vektorima nosačima (Support Vector Machine)

Slike

1.1	Primjer aero-snimka dijela područja Novog Grada i ušća rijeke Sane	2
1.2	Primjer aero-snimka dijela područja Novog Grada i ušća rijeke Sane u Unu i odgovarajuće tematske mape.	5
1.3	Primjer klasifikacije snimka na klase pokrivača i načina korišćenja zomljišta proma programu COBINE	7
1.4	Primjeri slika dobijenih daljinskom detekcijom u različitim rezolucijama.	8
1.5	Primjeri dominantnih vizuelnih obilježja na slikama iz različitih klasa pokrivača i načina korišćenja zemljišta.	11
2.1	Primjeri slika iz kolekcije Banja Luka. Sa lijeva u desno, po kolonama su slike iz klasa: <i>kuće, groblje, industrija, polje, rijeka</i> i <i>drveće.</i>	16
2.2	Raspodjela slika po kategorijama za kolekciju slika Banja Luka.	16
2.3	Primjeri slika iz kolekcije Window on the UK. Sa lijeva u desno, po kolonama su slike iz klasa: <i>brodovi, zgrade, polje, trava, rijeka, put,</i> <i>drueće</i> i <i>vozila</i>	17
2.4	Primjeri slika iz kolekcije UC Merced Land Use: (a) poljoprivreda, (b) avioni, (c) bejzbol teren, (d) plaža, (e) zgrade, (f) prerija, (g) gusto naseljeno rezidencijalno zemljište, (h) šuma, (i) autoput, (j) teren za golf, (k) luka, (l) raskrsnica, (m) rezidencijalno zemljište srednje gustine, (n) kamp prikolice, (o) nadvožnjak, (p) parking, (q) rijeka, (r) aerodromska pista, (s) rijetko naseljeno rezidencijalno zemljište,	11
2.5	 (t) cisterne, (u) teniski tereni. Primjeri slika iz kolekcije Wuhan: (a) aerodrom, (b) plaža, (c) most, (d) poslovne zgrade, (e) pustinja, (f) njive, (g) stadion, (h) šuma, (i) industrija, (j) livada, (k) planine, (l) park, (m) parking, (n) jezero, (o) luka, (p) željeznička stanica, (q) rezidencijalno zemljište, (r) rijeka, (s) vijadukt. 	19 20
3.1 3.2	Primjer 2D Gaborove funkcije	25
3.3 3.4	veće od polovine maksimuma amplitudne karakteristike	26 30
	Gaborov deskriptor teksture	31

4.1	Moduli koeficijenata Gaborove transformacije na različitim skalama	
	(m) i orijentacijama (n).	33
4.2	Dvodimenzionalni uslovni histogrami realnih i imaginarnih dijelova	
	Gaborovih koeficijenata na skalama 2 i 3 i orijentaciji 1	34
4.3	Dvodimenzionalni uslovni histogrami realnih i imaginarnih dijelova	
	Gaborovih koeficijenata na skali 2 i orijentacijama 1 i 2	34
4.4	Centar-rub organizacija receptivnog polja.	35
4.5	Dvodimenzionalni uslovni histogrami realnih i imaginarnih dijelova	
	Gaborovih koeficijenata na skalama 2 i 4 i orijentaciji 1	38
4.6	Dvodimenzionalni uslovni histogrami realnih i imaginarnih dijelova	
	Gaborovih koeficijenata na skali 2 i orijentacijama 1 i 4	39
4.7	Zavisnost tačnosti klasifikacije UC Merced Land Use kolekcije slika	
	od vrijednosti parametra p	42
4.8	Matrica konfuzija za UC Merced Land Use kolekciju slika dobijena	
	za EGTD sa $p = 0, 25. \ldots$	42
4.9	Zavisnost tačnosti klasifikacije UC Merced Land Use kolekcije slika	
	od dimenzionalnosti deskriptora za PCA redukciju dimenzionalnosti	
	primjenjenu na kompletan EGTD	43
4.10	Zavisnost tačnosti klasifikacije UC Merced Land Use kolekcije slika	
	od ukupne dimenzionalnosti deskriptora za PCA redukciju dimenzi-	
	onalnosti primjenjenu na pojedine komponente EGTD	44
4.11	Zavisnost tačnosti klasifikacije postignute na Wuhan kolekciji slika	
	od vrijednosti parametra p	47
4.12	Matrica konfuzija za Wuhan kolekciju slika dobijena za EGTD sa	
	$p = 0, 25. \ldots \ldots$	48
4.13	Zavisnost tačnosti klasifikacije Wuhan kolekcije slika od dimenzional-	
	nosti deskriptora za PCA redukciju dimenzionalnosti primjenjenu na	
	kompletan EGTD	48
4.14	Zavisnost tačnosti klasifikacije Wuhan kolekcije slika u zavisnosti od	
	dimenzionalnosti deskriptora dobijenog primjenog PCA na pojedine	
	komponente EGTD.	49
4.15	Tipovi receptivnih polja ćelija mrežnjače	52
4.16	Moduli odziva Gaborovog filtra za različite spektralne opsege	53
4.17	Dvodimenzionalni uslovni histogrami realnih i imaginarnih dijelova	
	koeficijenata Gaborove transformacije za crveni i zeleni spektralni opseg.	54
4.18	Simplektička dekompozicija RGB slike	57
4.19	Tačnost klasifikacije UC Merced Land Use kolekcije slika u zavisno-	
	sti od parametra p za različite komponente oponentnog proširenog	
	Gaborovog deskriptora teksture	60
4.20	Tačnost klasifikacije UC Merced Land Use kolekcije slika za opo-	
	nentni prosireni Gaborov deskriptor teksture u zavisnosti od ukupne	<u> </u>
1.01	dimenzionalnosti deskriptora.	61
4.21	Tacnost klasifikacije UC Merced Land Use kolekcije slika u zavisnosti	
	od parametra p za razlicite komponente kvaternionskog proširenog	0.2
	Gaborovog deskriptora teksture	62

vi

SLIKE

4.22	Tačnost klasifikacije UC Merced Land Use kolekcije slika za kvater- nionski prošireni Gaborov deskriptor teksture u zavisnosti od ukupne	
4.23	dimenzionalnosti deskriptora	62
4.24	sjednih podopsega	64
4.95	tora teksture.	64
4.20	dimenzionalnosti deskriptora.	65
4.26	Tačnost klasifikacije Wuhan kolekcije slika u zavisnosti od parametra p za različite komponente kvaternionskog proširenog Gaborovog	
4.27	deskriptora teksture	66
4.00	dimenzionalnosti kvaternionskog deskriptora	66
4.28	Matrica konfuzija za Wuhan kolekciju slika dobijena za kvaternionski EGTD uz redukciju dimenzionalnosti na 100-D pomoću PCA	68
5.1	Računanje originalnih STSIM i STSIM-2 mjera sličnosti	71
5.2	Osjetljivost originalne STSIM mjera sličnosti na prostorni raspored	79
5.3	Modifikacija računanja strukturalne sličnosti tekstura	73
5.4	Sličnost slika-klasa. Pojedine sličnosti blok-klasa su naznačene na	-
5.5	strelicama	76
	testne slike i te klase	77
5.6 5.7	Poređenje performansi različitih klasifikatora zasnovanih na STD Matrice konfuzije za skup slike Panje Luke	78
5.7 5.8	Matrica konfuzija za UC Merced Land Use kolekciju slika dobijena	00
	za NN-STSIM 4×4 slučaj iz Tabele 5.2 za RGB slike	81
5.9	Matrica konfuzija za Wuhan kolekciju slika dobijena za NN-STSIM 4×4 slučaj iz Tabele 5.3 za RGB slike	82
6.1	Blok dijagram hijerarhijske šeme za fuziju deskriptora.	86
6.2	Poređenje globalnih i lokalnih deskriptora slike na UC Merced Land Use kolekciji. Moguće je uočiti klase pokrivača i načina korišćenja	0.4
6.3	zemljišta za koje pojedini deskriptori imaju bolje performanse Primjeri slika iz klasa UC Merced Land Use kolekcije slika na kojima se bolji rezultati dobijaju korišćenjem globalnog deskriptora (gornji	91
	red) i lokalnog deskriptora (donji red).	91
6.4	Poređenje globalnih i lokalnih deskriptora slike na Wuhan kolekciji. Moguće je uočiti klase pokrivača i načina korišćenja zemljišta za koje	
	pojedini deskriptori imaju bolje performanse	93

SLIKE

6.5	Primjeri slika iz Wuhan kolekcije koji pripadaju klasama na kojima	
	se bolji rezultati dobijaju korišćenjem globalnog deskriptora (gornji	
	red) i lokalnog deskriptora (donji red).	. 94
6.6	Matrica konfuzija za UC Merced Land Use kolekciju slika dobijena	
	hijerarhijskom kombinacijom globalnih i lokalnih deskriptora	. 97
6.7	Matrica konfuzije za Wuhan kolekciju slika dobijena hijerarhijskom	
	kombinacijom globalnih i lokalnih deskriptora	. 99
6.8	Dio aero-snimka područja Novog Grada sa kojeg se izdvajaju blokovi	
	za obučavanje detektora urbanog područja	. 101
6.9	Dio aero-snimka područja Novog Grada na kojem se testira detektor	
	urbanih područja. Granice područja su određene manuelno i urbano	
	područje je osjenčeno crvenom, a neurbano zelenom bojom	. 101
6.10	Primjeri detekcije urbanog područja na dijelu aero-snimka područja	
	Novog Grada prije i poslije postobrade matematičkom morfologijom.	
	Urbano područje je osjenčeno crvenom, a neurbano zelenom bojom.	. 102
6.11	Primjer detekcije urbanog područja na dijelu aero-snimka područja	
	Novog Grada. Crvenom i zelenom bojom su osjenčena tačno detekto-	
	vana urbana i neurbana područja. Zutom bojom su osjenčena lažno	100
	pozitivna, a plavom lažno negativna urbana područja	. 103
B.1	Ilustracija izračunavanja Gist deskriptora.	. 114
C.1	Primjer izračunavanja SIFT deskriptora.	. 116

viii

Tabele

$3.1 \\ 3.2$	Tačnosti klasifikacije za kolekciju slika Banje Luke	$\begin{array}{c} 29\\ 30\end{array}$
4.1 4.2	Oznake deskriptora korišćene u eksperimentalnim rezultatima Tačnosti klasifikacije (%) postignute na UC Merced Land Use ko- lekciji slika za $p = 2$ i različite kombinacije komponenata EGTD	40
4.3	deskriptora	41
4.4	koeficijenata iz susjednih podopsega, PCA – analiza glavnih kompo- nenata, LDA – linearna diskriminatna analiza, RDA – regularizovana linearna diskriminatna analiza, Komp. – redukcija dimenzionalnosti primjenjena po komponentama EGTD	45
$\begin{array}{c} 4.5\\ 4.6\end{array}$	nentama EGTD	50 63 67
$5.1 \\ 5.2 \\ 5.3$	Tačnosti klasifikacije (%) za skup slika Banje Luke Tačnosti klasifikacije (%) za UC Merced Land Use kolekciju slika Tačnosti klasifikacije (%) za Wuhan kolekciju slika	79 80 82
6.1 6.2	Tačnosti klasifikacije (%) za globalne i lokalne deskriptore na UC Merced Land Use kolekciji slika	89
6.3	kolekciji slika	92
	tora na UC Merced Land Use kolekciji slika.	94

TABELE

6.4	Tačnosti klasifikacije $(\%)$ za različite metaklasifikatore i njihove stac-	
	kingC varijante (označene sa C) za fuziju globalnih deskriptora redu-	
	kovane dimenzionalnosti i lokalnih deskriptora na UC Merced Land	
	Use kolekciji slika.	j
6.5	Poređenje tačnosti klasifikacije za UC Merced Land Use kolekciju slika. 96	3
6.6	Tačnosti klasifikacije (%) za različite metaklasifikatore i njihove stac-	
	kingC varijante (označene sa C) za fuziju globalnih i lokalnih deskrip-	
	tora na Wuhan kolekciji slika	7
6.7	Tačnosti klasifikacije (%) za različite metaklasifikatore za fuziju glo-	
	balnih deskriptora redukovane dimenzionalnosti i lokalnih deskriptora	
	na Wuhan kolekciji slika	3
6.8	Poređenje tačnosti klasifikacije za Wuhan kolekciju slika 99)
6.9	Tačnosti detekcije urbanog područja.	L

Glava 1

Uvod

Zajednički imenilac svih definicija daljinske detekcije (*remote sensing*) [1, 2, 3] je prikupljanje informacija na daljinu, tj. bez fizičkog kontakta sa objektom koji se posmatra. Ipak, termin daljinska detekcija je ponikao i prvenstveno se odnosi na prikupljanje informacija o Zemljinoj površini korišćenjem senzora koji se nalaze iznad nje, najčešće u vazduhu (aero-snimanje) ili svemiru (satelitsko snimanje). Pored toga, uobičajeno je i da se termin daljinska detekcija ograniči na prikupljanje informacija koje se mogu predstaviti u formatu slike, a dobijaju se korišćenjem elektromagnetnog zračenja reflektovanog ili emitovanog sa Zemljine površine. Ova dva ograničenja postavljaju okvir u kojem se termin daljinska detekcija koristi u ovoj disertaciji. Dakle, daljinska detekcija se odnosi na dobijanje informacija iz slika prikupljenih pomoću senzora elektromagnetnog zračenja reflektovanog ili emitovanog sa Zemljine površine, pri čemu se senzori nalaze iznad predmeta posmatranja. Na Slici 1.1 dat je primjer aero-snimka dijela područja Novog Grada i ušća rijeke Sane u Unu.

Daljinska detekcija ima širok spektar primjena, kako vojnih tako i civilnih. Među najznačajnijim primjenama su analiza i planiranje pokrivača i načina korišćenja zemljišta, izrada mapa, praćenje i planiranje urbanog rasta, vremenska prognoza, proučavanje dugoročnih klimatskih promjena, nadzor usjeva, proučavanje erozije, detekcija vodenih tokova, procjena poplavnog rizika, detekcija mreže puteva, detekcija vojnih ciljeva i mnoge druge. Sve ove primjene zahtijevaju razvoj efikasnih načina za akviziciju, obradu, prenos, čuvanje, pretraživanje i analizu snimaka.

Senzori za daljinsku detekciju se nalaze na čitavom nizu aero i svemirskih platformi i koriste se za akviziciju slika u raznim opsezima elektromagnetnog spektra, od vidljivog i infracrvenog dijela spektra do mikrotalasa i radara. Oni mogu biti pasivni i mjeriti zračenje koje prirodno potiče ili se reflektuje od površine Zemlje i/ili njene atmosfere, kao što je npr. Sunčevo zračenje (vidljivi i infracrveni dio spektra) ili Zemljino termalno zračenje (infracrveni dio spektra i mikrotalasi). Druga grupa senzora su aktivni senzori koji mjere zračenje koje potiče iz vještačkog izvora, a reflektuje se od Zemljine površine ili atmosfere. U ovu grupu spadaju radarski (*Synthetic Aperture Radar* – SAR) i laserski senzori (*LIght Detection And Ranging* – Lidar).

Količina podataka dobijena uređajima za daljinsku detekciju se svakodnevno povećava. Povećanje se ogleda ne samo u kontinuiranoj akviziciji podataka korišćenjem



Slika 1.1: Primjer aero-snimka dijela područja Novog Grada i ušća rijeke Sane u Unu.

GLAVA 1. UVOD

postojećih senzora, već i u povećanju broja senzora kojima se podaci prikupljaju. Na primjer, uvećanje obima podataka misije ENVISAT, koju vodi Evropska svemirska agencija, na godišnjem nivou je reda stotina terabajta, a planirano je da će rast obima podataka nove misije, pod nazivom Sentinel, biti za red veličine veći [4]. Ovo povećanje obima podataka predstavlja sve veći problem za postojeće načine njihovog čuvanja, prenosa i analize.

1.1 Problem analize snimaka dobijenih daljinskom detekcijom

Da bi se iz podataka dobijenih tehnikama daljinske detekcije izvukle korisne informacije potrebno ih je analizirati. Tradicionalno, podatke, odnosno snimke, dobijene daljinskom detekcijom, analiziraju posebno obučena lica – analitičari u okviru procesa koji se naziva interpretacija slika. Ljudska bića posjeduju veoma razvijen sistem koji omogućava razumijevanje složenih vizuelnih scena, čak i u uslovima slabog osvjetljenja, prisustva šuma, razlika u položaju, skali i uglu posmatranja, djelimičnog zaklanjanja objekata, te prisustva velikog broja irelevantnih objekata. Ova izuzetna svojstva ljudskog vida dijelom su rezultat obučavanja na svakodnevnom iskustvu u interpretaciji vizuelnih scena. Međutim, u interpretaciji snimaka dobijenih daljinskom detekcijom nije se moguće u potpunosti osloniti na čovjekovo svakodnevno iskustvo. Razlog za ovo su neke osobine snimaka dobijenih daljinskom detekcijom koje se razlikuju od osobina scena koje čovjek interpretira u svakodnevnom životu. Na prvom mjestu je neuobičajen ugao posmatranja dobijen kao posljedica akvizicije slika senzorom koji se nalazi iznad predmeta posmatranja. Druga razlika se odnosi na činjenicu da se u svakodnevnom životu uglavnom srećemo sa slikama formiranim korišćenjem zračenja u vidljivom dijelu spektra, dok slike dobijene daljinskom detekcijom često sadrže spektralne informacije izvan vidljivog dijela spektra. Konačno, geografska pokrivenost Zemljine površine uređajima za daljinsku detekciju je veoma velika i analitičari su često u situaciji da interpretiraju snimke dijelova Zemljine površine koji su izvan njihovog direktnog iskustva. Da bi se prevazišle ove prepreke neophodna je formalna obuka analitičara.

Prvi korisnici snimaka dobijenih daljinskom detekcijom bili su vojske i vlade država. Vremenom su troškovi akvizicije snimaka opadali, a broj senzora se povećavao i, kao rezultat tog procesa, danas je krug korisnika snimaka dobijenih daljinskom detekcijom veoma širok i obuhvata subjekte od političke vlasti na svim nivoima, preko komercijalnih korisnika i organizacija za zaštitu okoline, do pojedinaca. Upotreba računara za analizu snimaka postala je neizbježna, a razvoj Interneta je donio mogućnost jednostavnog pristupa velikoj količini podataka. Zahvaljujući servisima kao što su Google Maps¹ ili Bing Maps² gotovo svi korisnici Interneta su se susreli sa satelitskim snimcima Zemljine površine.

Međutim, povećanje obima raspoloživih podataka dobijenih daljinskom detekcijom i proširenje kruga korisnika ovih podataka naglašava problem njihove analize kako bi se dobile informacije koje se mogu upotrebiti za donošenje odluka. Mnogim

¹http://maps.google.com

²http://www.bing.com/maps/

GLAVA 1. UVOD

zainteresovanim korisnicima obuka analitičara nije ekonomski isplativa, a i sam obim podataka već uveliko prevazilazi mogućnosti manuelne analize, čak i kada se radi o iskusnim analitičarima. S obzirom na sveprisutnost upotrebe računara u gotovo svim fazama akvizicije, obrade, prenosa, čuvanja i pretraživanja snimaka dobijenih daljinskom detekcijom, logična je težnja da se i njihova analiza automatizuje u što je moguće većoj mjeri.

Osnovne aktivnosti u interpretaciji snimaka su: klasifikacija, enumeracija, mjerenje i označavanje [1]. U okviru klasifikacije se objekti, strukture i područja na snimcima svrstavaju u klase iz predefinisanog skupa. Ova aktivnost uključuje i zadatke kao što su detekcija i prepoznavanje objekata. Enumeracija predstavlja formiranje liste, odnosno, utvrđivanje broja diskretnih objekata određenog tipa na snimku. Mjerenje se odnosi na određivanje prostornih dimenzija posmatranih struktura i objekata, čime se bavi *fotogrametrija*, kao i kvantitativno određivanje svjetline snimka, što je predmet *fotometrije* i *radiometrije*. Označavanje uključuje određivanje dijela Zemljine površine koji zauzima određeni prostorni entitet, odnosno, određivanje granica između različitih područja. Predmet istraživanja u ovoj disertaciji je automatizacija klasifikacije i, u izvjesnoj mjeri, označavanja različitih klasa pokrivača i načina korišćenja zemljišta.

U ovoj disertaciji usvojena je sljedeća terminološka konvencija. Termin *snimak* odnosi se na sliku određenog dijela Zemljine površine dobijenu nekom od tehnika daljinske detekcije. Snimak pokriva veći prostorni obuhvat i na njemu se može identifikovati više tipova pokrivača i načina korišćenja zemljišta. Sa druge strane, termin *slika* će se koristiti za manje dijelove snimka, koji se odnose na mali prostorni obuhvat i uglavnom sadrže samo jedan tip pokrivača, odnosno, načina korišćenja zemljišta. U istom kontekstu koristiće se i termin *blok*.

1.2 Klasifikacija snimaka dobijenih tehnikama daljinske detekcije

Jedan od osnovnih zadataka u analizi snimaka dobijenih daljinskom detekcijom je njihova klasifikacija. U ovom kontekstu termin *klasifikacija snimka* se odnosi na pridruživanje klase iz predefinisanog skupa svakom pikselu snimka. Rezultat klasifikacije će biti podjela snimka na skup disjunktnih regiona tako da svi pikseli iz jednog regiona pripadaju istoj klasi, odnosno, zadovoljavaju neki kriterijum homogenosti. Kada se pikselima svake klase dodijeli određena boja dobija se *tematska mapa* koja prikazuje prostornu raspodjelu klasa na posmatranom dijelu Zemljine površine. Aero-snimak, prethodno prikazan na Slici 1.1 i odgovarajuća tematska mapa su prikazani na Slici 1.2

Klasifikacija snimaka dobijenih daljinskom detekcijom se koristi za pripremu podataka za unos u geografske informacione sisteme, planiranje urbanog rasta, nadzor poljoprivrednih dobara, proučavanje erozije, hidrološke analize itd. U nekim od ovih primjena klasifikacija je konačni proizvod analize snimaka dobijenih daljinskom detekcijom, a u nekim je samo međukorak u kojem se pripremaju ulazni podaci za dalje analize. Karakterističan primjer konačnog proizvoda dobijenog klasifikacijom snimaka je klasifikacija pokrivača i načina korišćenja zemljišta (*land cover/land use*



(a) Aero-snimak

(b) Tematska mapa

Slika 1.2: Primjer aero-snimka dijela područja Novog Grada i ušća rijeke Sane u Unu i odgovarajuće tematske mape.

classification). Pokrivač zemljišta čine prirodne ili vještačke strukture koje zauzimaju određeni dio Zemljine površine, kao što su, na primjer, vegetacija, vodena površina, zgrade, put, itd. Klase pokrivača zemljišta je moguće odrediti na osnovu posmatranja snimaka Zemljine površine. Način korišćenja zemljišta je prevashodno ekonomska kategorija i odnosi se na način na koji čovjek koristi određeni dio Zemljine površine, na primjer, poljoprivredno, stambeno ili komercijalno zemljište. Način korišćenja zemljišta nije uvijek moguće odrediti samo na osnovu posmatranja snimaka Zemljine površine. Problem je u tome što region koji odgovara jednoj klasi načina korišćenja zemljišta može da sadrži više tipova pokrivača zemljišta. Na primjer, klasa korišćenja zemljišta "urbano stambeno područje" može da sadrži klase pokrivača zemljišta kao što su: vegetacija, pločnici, ulice i kuće, odnosno, zgrade. Dakle, da bi se tačno odredile klase načina korišćenja zemljišta ponekad je potrebno koristiti i druge izvore informacija. Vidimo da određivanje pokrivača i načina korišćenja zemljišta predstavljaju dva različita zadatka. Ipak, oni su čvrsto povezani i potrebno ih je razmatrati zajedno uz razumijevanje specifičnosti svakog od njih.

Skup klasa u koje se klasifikuje snimak zavisi od potreba korisnika klasifikacije. Ipak, klase ne mogu biti proizvoljne. Njihovi opisi i uputstva za klasifikaciju moraju biti specificirani tako da proces klasifikacije bude ponovljiv, a potrebno je i obezbijediti klasifikaciju svakog dijela posmatranog geografskog područja. Pored toga, često je potrebno upoređivanje rezultata klasifikacije pokrivača i načina korišćenja zemljišta iz dvije epohe ili na različitim parcelama. Ovaj zahtjev se može ispuniti samo ukoliko se koriste isti ili kompatibilni skupovi klasa. Takođe, potrebno je osigurati i kompatibilnost rezultata klasifikacije iste parcele na različitim skalama, sa različitim nivoima detalja. Zbog toga su kreirani različiti sistemi za klasifikaciju snimaka dobijenih daljinskom detekcijom. Anderson i saradnici su predložili

GLAVA 1. UVOD

skup kriterijuma koje mora zadovoljavati sistem za klasifikaciju snimaka dobijenih daljinskom detekcijom [5]:

- 1. Tačnost klasifikacije pokrivača i načina korišćenja zemljišta korišćenjem snimaka dobijenih daljinskom detekcijom treba da bude minimalno 85%.
- 2. Tačnost klasifikacije za različite kategorije treba da bude približno jednaka.
- Rezultati interpretacije treba da budu ponovljivi nezavisno od osobe koja vrši analizu i vremena snimanja.
- 4. Klasifikacioni sistem mora biti primjenljiv na šire geografsko područje.
- 5. Klasifikacioni sistem bi trebalo da omogući da se klase pokrivača zemljišta koriste kao zamjena za aktivnosti koje se odvijaju na posmatranom području.
- 6. Klasifikacioni sistem mora biti primjenljiv na podatke dobijene snimanjima u različitim godišnjim dobima.
- 7. Potrebno je omogućiti efikasno korišćenje podataka dobijenih terenskim snimanjem ili poboljšanim senzorima za daljinsku detekciju.
- 8. Mora biti moguća agregacija kategorija.
- 9. Potrebno je omogućiti poređenje sa budućim podacima.
- 10. Kada je moguće, potrebno je prepoznati višestruke načine korišćenja zemljišta.

Polazeći od ovih kriterijuma predložen je i hijerarhijski sistem za klasifikaciju pokrivača i načina korišćenja zemljišta koji sadrži četiri nivoa hijerarhije. Klase pokrivača i načina korišćenja zemljišta na prva dva nivoa su specificirane, dok klase na višim nivoima mogu specificirati lokalni korisnici u zavisnosti od potreba. Uslov koji pri tome treba da bude zadovoljen je da se tako definisane klase uklapaju u hijerarhiju na nižim nivoima. Ovaj sistem za klasifikaciju je osnova sistema koji koristi Služba za geološka istraživanja Sjedinjenih Američkih Država (United States Geology Survey) [6]. Takođe, Evropska agencija za životnu sredinu (European Environment Agency) je u okviru programa CORINE (Coordination of Information on the Environment) formirala hijerarhijski sistem sa 44 klase pokrivača zemljišta [7]. Primjer klasifikacije snimka na klase pokrivača i načina korišćenja zemljišta prema programu CORINE prikazan je na Slici 1.3.

Algoritam kojim je definisana određena procedura za klasifikaciju snimaka naziva se *klasifikator*. Klasifikator u svom radu primarno koristi vrijednosti piksela na snimku, ali u nekim slučajevima se koriste i pomoćni podaci, kao što su topografski, pedološki ili geološki. Klasifikator koji za klasifikaciju svakog piksela koristi samo njegove vrijednosti u različitim spektralnim opsezima, naziva se *spektralni* klasifikator. Ukoliko klasifikator, pored vrijednosti piksela koji se klasifikuje, koristi i vrijednosti piksela iz njegove okoline radi se o *prostornom* klasifikatoru. Jedna varijanta prostornog klasifikatora koristi lokalnu prostornu informaciju, ali donosi odluku o samo jednom pikselu. Druga varijanta ne klasifikuje svaki piksel pojedinačno, već svim pikselima iz posmatranog regiona pridružuje istu klasu [8]. Posmatrani region,



Slika 1.3: Primjer klasifikacije snimka na klase pokrivača i načina korišćenja zemljišta prema programu CORINE.

u opštem slučaju, može biti proizvoljnog oblika. U literaturi je uobičajeno da se ovakav klasifikator naziva *objektno-orijentisani* klasifikator [1, 9], ukoliko su vizuelno homogeni regioni dobijeni nenadgledanom segmentacijom, a *blokovski* klasifikator, ako se radi o regionima dobijenim fiksnom podjelom snimka.

Kriterijum homogenosti kod spektralnih klasifikatora se odnosi na sličnost vrijednosti piksela u različitim spektralnim opsezima. Prema tome, spektralno homogeni regioni predstavljaju *spektralne klase*. Sa druge strane, *informacione klase* predstavljaju kategorije koje su od interesa za korisnika informacija dobijenih daljinskom detekcijom. Primjeri informacionih klasa su klase pokrivača i načina korišćenja zemljišta, kao, na primjer, one prikazane na Slici 1.3. Kod spektralnih klasifikatora, pripadnost piksela informacionim klasama je moguće odrediti samo posredno, pronalaženjem spektralnih klasa i uspostavljanjem korespondencije između spektralnih i informacionih klasa. Nažalost, ova korespondencija nije jednoznačna jer unutar informacionih klasa mogu postojati spektralne varijacije zbog varijabilnosti prirodnih osobina objekata koji se posmatraju, postojanja pokrivača različitih tipova na površini jednog piksela što rezultuje pojavom mješovitih piksela (*mixed pixels*), atmosferske varijabilnosti, godišnjeg doba, osvjetljenja, itd. Spektralnim varijacijama definisane su spektralne podklase. Zbog toga snimak nije moguće uvijek klasifikovati korišćenjem isključivo spektralnih informacija.

Većina komercijalnih klasifikatora snimaka dobijenih daljinskom detekcijom (ER Mapper³, ENVI⁴, TNTmips⁵), kao i neki akademski (MultiSpec⁶) baziraju se na

³http://www.erdas.com

⁴http://www.exelisvis.com/ProductsServices/ENVI.aspx

⁵http://www.microimages.com/products/tntmips.htm

⁶https://engineering.purdue.edu/~biehl/MultiSpec/



Slika 1.4: Primjeri slika dobijenih daljinskom detekcijom u različitim rezolucijama.

korišćenju isključivo spektralnih informacija za klasifikaciju snimaka. Pošto su spektralni klasifikatori zasnovani na pronalaženju spektralno homogenih regiona na snimku – spektralnih klasa – i njihovom uparivanju sa informacionim klasama, postojanje spektralnih varijacija negativno utiče na kvalitet klasifikacije. Da bi se uticaj spektralnih varijacija umanjio obično se za automatsku klasifikaciju pokrivača i načina korišćenja zemljišta koriste snimci sa niskom prostornom rezolucijom. Na primjer, za klasifikaciju pokrivača i načina korišćenja zemljišta na nivou I Andersonovog sistema prikladni su multispektralni snimci sa veličinama piksela od 20 do 100 metara, dok su rezolucije potrebne za klasifikaciju na nivou II između 5 i 20 metara [3]. Kada se koriste snimci niske rezolucije, detalji na slici se gube. Neke detalje je nemoguće detektovati, a neki se manifestuju u vidu mješovitih piksela. Međutim, ukoliko je potrebno klasifikovati snimak na nivoima III ili IV, ili je rezultat klasifikacije ulazni podatak za dalju analizu snimka (npr. detekciju objekata), neophodno je više detalja, odnosno, korišćenje snimaka visoke rezolucije. Uključivanje detalja povećavanjem rezolucije rezultuje vizuelnim nehomogenostima koje otežavaju spektralnu klasifikaciju pokrivača i načina korišćenja zemljišta. Ova pojava je ilustrovana na Slici 1.4. Vidi se da se, sa porastom rezolucije, povećava nivo raspoloživih detalja na slici, ali i spektralne nehomogenosti postaju sve izraženije. Dok je na Slici 1.4a bilo moguće klasifikovati npr. urbano područje korišćenjem samo spektralne informacije, sa porastom rezolucije taj zadatak postaje sve teži i zahtijeva, na primjer, korišćenje teksture, Slike 1.4b i 1.4c, te pojedinih objekata i njihovih prostornih odnosa, Slika 1.4d. U ovim primjerima, kao i u ostatku disertacije, pojam rezolucija se odnosi na prostornu rezoluciju snimka.

Dakle, moguće je sumarizovati osnovne nedostatke spektralnih klasifikatora:

- Spektralne varijacije onemogućavaju da se uspostavi jednoznačna korespondencija između spektralnih i informacionih klasa;
- Kada se koriste visokorezolucioni snimci, varijacije vrijednosti piksela unutar jedne klase pokrivača ili načina korišćenja zemljišta se povećavaju, onemogućavajući na taj način pronalaženje homogenih regiona;
- Složeni tipovi načina korišćenja zemljišta sadrže heterogene klase pokrivača zemljišta sa različitim spektralnim osobinama pa ih je nemoguće klasifikovati korišćenjem samo spektralnih informacija;
- Spektralna klasifikacija je u najvećoj mjeri bazirana na vrijednostima pojedinačnih piksela dok se čovjek pri analizi vizuelne scene malo oslanja na takvu tačkastu informaciju, a više na korišćenje teksture, oblika, kompozicije, itd.

Svi ovi problemi su posljedica korišćenja isključivo spektralnih informacija za klasifikaciju. Upravo analogija sa onim što je poznato o tome kako obučeni analitičari pristupaju analizi snimaka predstavlja inspiraciju za njihovo rješavanje. U cilju dobijanja što je moguće boljih rezultata poželjno je da interpretatori donose zaključke korišćenjem podataka iz različitih izvora i različitih elemenata interpretacije snimaka jer se tako uspostavlja potpuniji kontekst za analizu snimka. Dakle, osim spektralnih informacija potrebno je koristiti i druge elemente interpretacije snimaka, kao što su: tekstura, oblik, veličina, raspored i prostorni odnosi objekata [1]. Uključivanje ovih elemenata znači i uključivanje prostorne informacije, odnosno, prelazak na korišćenje prostornih klasifikatora. Odluka o klasifikaciji piksela ili regiona sada se donosi na osnovu podataka koji su na raspolaganju u prozoru oko posmatranog piksela, odnosno na osnovu svih piksela u posmatranom regionu. U [8] je ispitivano koja je veličina prozora optimalna da bi obučeni interpretatori uspješno klasifikovali regione izdvojene sa snimka rezolucije od $1, 5m \times 1, 5m$ korišćenjem kvadratnog prozora u kategorije načina korišćenja zemljišta na nivou II Andersonovog sistema za klasifikaciju. Pokazano je da se rezultati poboljšavaju sa povećavanjem prozora i da je za postizanje 75% tačno klasifikovanih regiona u kategorije urbanog načina korišćenja zemljišta potrebna veličina prozora od 40×40 piksela, odnosno, $60m \times 60m$.

Ovaj rezultat predstavlja motivaciju za istraživanje u oblasti objektno-orijentisanih klasifikatora što je i predmet istraživanja u ovoj disertaciji. Problem klasifikacije je postavljen kao problem semantičke klasifikacije slika, pri čemu slike, u stvari, predstavljaju regione visokorezolucionog snimka izdvojene korišćenjem pravougaonog prozora. Zbog specifičnog oblika regiona i načina njegovog izdvajanja, ovaj pristup se, kao što je pomenuto, češće sreće pod nazivom *blokovska klasifikacija*. Semantičke klase u koje se slike klasifikuju predstavljaju klase pokrivača i načina korišćenja zemljišta.

1.3 Semantička klasifikacija slika

Semantička klasifikacija slika opšte namjene je u poslednjih 15 godina predmet intenzivnog proučavanja. Problem semantičke klasifikacije slika se sastoji u dodjeljivanju oznaka klasa iz predefinisanog skupa slikama. Ovaj problem se sreće i pod nazivom označavanje slika. Iako je, u opštem slučaju, moguće da se jednoj slici dodijeli više oznaka klasa, razmatranja u ovoj disertaciji ograničena su na slučaj u kojem se jednoj slici dodjeljuje tačno jedna oznaka klase. Posljedica ovog ograničenja je da se jednom bloku snimka koji se posmatra dodjeljuje tačno jedna oznaka klase pokrivača, odnosno, načina korišćenja zemljišta. Kao što je pomenuto u odjeljku 1.2, ovaj pristup se koristi kod objektno-orijentisane i blokovske klasifikacije snimaka dobijenih daljinskom detekcijom u klase pokrivača i načina korišćenja zemljišta.

Tradicionalno, u oblasti klasifikacije slika opšte namjene razlikuju se dva zadatka: prepoznavanje objekata i prepoznavanje scena. Prepoznavanje objekata podrazumijeva klasifikaciju čitave slike u jednu od klasa u zavisnosti od toga da li se na slici nalazi instanca objekta iz te klase [10]. Pri tome se podrazumijeva da se na slici nalazi objekat iz samo jedne klase i da njegova lokacija na slici nije bitna. Pored objekta, slika može sadržati i određenu pozadinu. Pokazalo se da metodi zasnovani na lokalnim deskriptorima predstavljaju obećavajući smjer istraživanja u ovoj oblasti [11, 12, 13, 14]. Dobri rezultati postignuti su i korišćenjem biološki inspirisanih reprezentacija [15, 16] kao i nenadgledanog učenja reprezentacije [17, 18, 19]. Sa druge strane, prepoznavanje scena podrazumijeva dodjelu oznake klase čitavoj slici koja može sadržati proizvoljan broj objekata i pozadinu u prirodnom rasporedu elemenata. I u ovoj oblasti dobri rezultati su postignuti metodima zasnovanim na lokalnim deskriptorima [12, 20], ali su značajni i holistički pristupi koji koriste globalne deskriptore [21, 22].

Problem semantičke klasifikacije slika dobijenih daljinskom detekcijom predstavlja kombinaciju prepoznavanja objekata i prepoznavanja scena. U nekim slučajevima slika se klasifikuje na osnovu njenog globalnog izgleda, dok je u drugim klasa u koju će slika biti klasifikovana određena prisustvom određenog objekta na njoj. Zbog toga su, za semantičku klasifikaciju slika dobijenih daljinskom detekcijom, potrebni deskriptori koji uz holističku reprezentaciju slike sadrže i informaciju o izraženim strukturama na slici, kao i klasifikatori koji omogućavaju kombinovanje komplementarnih reprezentacija. Tema istraživanja u ovoj disertaciji su upravo ta dva elementa klasifikatora slika dobijenih daljinskom detekcijom – deskriptori i algoritmi za klasifikaciju, odnosno, klasifikatori u užem smislu riječi.

1.3.1 Deskriptori slika

Da bi se neka slika mogla klasifikovati korišćenjem automatskog klasifikatora potrebno je formirati njenu prikladnu reprezentaciju. Naime, vrijednosti piksela slike nisu pogodne kao ulazni podaci za klasifikator zbog njihove velike varijabilnosti na slikama iz jedne semantičke klase, kao i međusobne sličnosti vrijednosti piksela na slikama iz različitih klasa. Pored toga, vrijednosti piksela slike sadrže visok stepen redundanse i zbog toga predstavljaju neefikasan način za kodiranje sadržaja slike. Kako bi se izbjegli pomenuti nedostaci korišćenja vrijednosti piksela, reprezentacija



(a) Trava – dominantna boja.

(b) Urbano područje – pravilni objekti.

(c) Put – izražene ivice.

Slika 1.5: Primjeri dominantnih vizuelnih obilježja na slikama iz različitih klasa pokrivača i načina korišćenja zemljišta.

slike se formira kvantitativnim opisom vizuelnih obilježja na osnovu kojih je moguće odrediti pripadnost slike nekoj od semantičkih klasa. Najčešće korišćenja vizuelna obilježja u klasifikaciji slika su: boja, tekstura, oblik, raspored ivica, prostorni odnosi objekata, itd. Pretpostavlja se da slike koje pripadaju istoj semantičkoj klasi imaju zajednička vizuelna obilježja. Na primjer, na aero-snimcima travnatih površina preovlađuje zelena boja, u urbanim sredinama preovlađuju pravilni, pravougaoni objekti, putevi imaju jasno definisane duge ivice, itd, Slika 1.5.

Vizuelna obilježja se kvantitativno opisuju korišćenjem deskriptora. Deskriptori su, tipično, višedimenzionalni vektori čiji elementi opisuju raspodjelu vizuelnih obilježja na slici. Vrijednosti elemenata deskriptora se izračunavaju tehnikama digitalne obrade slike, najčešće filtriranjem slike i određivanjem statistike vrijednosti na izlazu filtra.

1.3.2 Klasifikator

Deskriptori slika se koriste kao ulazni podaci za algoritam za klasifikaciju, odnosno, sam klasifikator, u užem smislu riječi. Klasifikator može biti implementiran tako da se slika klasifikuje na osnovu skupa pravila koja je definisao ljudski ekspert. Međutim, takva realizacija klasifikatora bi bila izuzetno neefikasna i skupa jer bi za svaku klasu bilo potrebno ručno specificirati skup pravila. Kako bi se ovakav način konstrukcije klasifikatora izbjegao, koristi se algoritmi za njihovo obučavanje.

Pretpostavimo da je potrebno sliku klasifikovati u jednu od klasa iz predefinisanog konačnog skupa $\mathcal{C} = \{c_1, c_2, \ldots, c_K\}$. Klasifikator je funkcija $\hat{c} = f(\mathbf{x}; \theta)$ koja deskriptoru $\mathbf{x} \in \mathbb{R}^d$, date slike, dodjeljuje oznaku klase $\hat{c} \in \mathcal{C}$. Vektorom θ su predstavljeni parametri klasifikatora koji se određuju u procesu njegovog obučavanja.

U procesu obučavanja klasifikatora se koriste podaci, tj. deskriptori slika za koje je poznata pripadnost klasama. Ovaj skup uzoraka naziva se trening skup. Vektor θ parametara klasifikatora se određuje na takav način da se što je moguće veći procenat primjera iz trening skupa tačno klasifikuje. Ovo se postiže tako što se formira neka funkcija vektora parametara čija će vrijednost odražavati tačnost klasifikatora na trening skupu. Ova funkcija se naziva *funkcija cilja*. Obučavanje klasifikatora je proces određivanja vrijednosti vektora parametara za koju funkcija cilja ima ekstremum. Važno je primjetiti da je dostizanje savršene klasifikacije uzoraka iz trening skupa trivijalno, a da je pravi cilj klasifikacije određivanje oznaka klasa za nove uzorke koji ne pripadaju trening skupu. Zbog ovoga je potrebno obezbijediti dobre performanse klasifikatora na novim podacima, odnosno, dobru *generalizaciju* od strane klasifikatora.

1.4 Doprinos disertacije

Osnovna tema ove disertacije je klasifikacija slika dobijenih daljinskom detekcijom. Kada su predmet klasifikacije slike visoke rezolucije, u reprezentaciju slike se, pored spektralnih vrijednosti piksela, moraju uključiti i drugi elementi interpretacije snimaka, kao što su: tekstura, oblik, veličina, raspored i prostorni odnosi objekata. Pošto se, na problemu klasifikacije slika dobijenih daljinskom detekcijom, tekstura pokazala kao vrlo dobro vizuelno obilježje, u ovoj disertaciji je predložen prošireni Gaborov deskriptor teksture koji, pored reprezentacije teksture, uključuje i reprezentaciju istaknutih vizuelnih elemenata slike na različitim skalama i orijentacijama. Kako bi se ovo postiglo, u deskriptor su uključene i mjere korelacije između koeficijenata Gaborove transformacije na različitim skalama i orijentacijama. Eksperimentalno je pokazano da klasifikator slika zasnovan na proširenom Gaborovom deskriptoru teksture daje dobre rezultate na testiranim kolekcijama slika.

Originalno, prošireni Gaborov deskriptor teksture je predložen za panhromatske slike. Međutim, pošto je spektralna informacija tradicionalno osnov klasifikatora slika dobijenih daljinskom detekcijom, predložena su dva pristupa za izračunavanje proširenog Gaborovog deskriptora teksture za RGB slike. Oba pristupa donose dalje poboljšanje performansi klasifikatora. Dobijeni deskriptori su visokodimenzionalni pa je razmatrana i redukcija njihove dimenzionalnosti i pokazano je da se pomoću deskriptora značajno niže dimenzionalnosti mogu dobiti približno iste tačnosti klasifikacije.

Pored proširenog Gaborovog deskriptora teksture koji je globalni deskriptor slika, predložen je i lokalni strukturalni deskriptor teksture koji takođe sadrži mjere kroskorelacija između modula koeficijenata Gaborove transformacije na različitim skalama i orijentacijama. Ovaj deskriptor se računa za blokove pozicionirane oko pravilno raspoređenih tačaka na slici. Za određivanje sličnosti slika predložena je strukturalna sličnost tekstura koja je perceptualno zasnovana mjera sličnosti. Slika se sada posmatra kao neuređen skup deskriptora i za klasifikaciju je predložen klasifikator na principu najbližih susjeda uz računanje sličnosti između testne slike i svake od klasa pokrivača i načina korišćenja zemljišta.

Pošto trenutno ne postoji konsenzus o tome da li su globalni deskriptori teksture ili lokalni deskriptori bolji za klasifikaciju slika dobijenih daljinskom detekcijom analizirane su dobre i loše strane obje grupe deskriptora i pokazano je da su, u stvari, informacije kodirane ovim deskriptorima u određenoj mjeri komplementarne. Zbog toga je moguće njihovom fuzijom postići poboljšanje performansi klasifikatora. Predložen je hijerarhijski metod za kombinovanje klasifikatora koji je detaljno eksperimentalno testiran i pokazano je da se njegovim korišćenjem može značajno poboljšati tačnost klasifikacije slika u klase pokrivača i načina korišćenja zemljišta.

1.5 Organizacija disertacije

Ocjena performansi klasifikatora slika dobijenih daljinskom detekcijom formira se njegovim testiranjem na nekoj kolekciji slika. U **Glavi 2** opisane su kolekcije slika na kojima su testirani deskriptori i klasifikatori predloženi u ovoj disertaciji. Pored samih kolekcija, opisana je i metodologija testiranja algoritma za klasifikaciju.

Polazna tačka za razvoj proširenog Gaborovog deskriptora teksture je poznati Gaborov deskriptor teksture za koji je teorijska osnova data u **Glavi 3**. U ovoj glavi razmatrano je modeliranje receptivnih polja prostih ćelija u vizuelnom korteksu sisara pomoću Gaborovih funkcija i zatim je opisan Gaborov deskriptor teksture. Eksperimentalno je pokazano da Gaborov deskriptor teksture predstavlja dobru osnovu za istraživanje u oblasti klasifikacije snimaka dobijenih daljinskom detekcijom.

U Glavi 4 je ilustrovano postojanje statističkih zavisnosti između koeficijenata Gaborove transformacije na različitim skalama i orijentacijama, kao i njihova veza sa centar-okolina organizacijom receptivnih polja prostih ćelija. Polazeći od ovih zapažanja, predložen je prošireni Gaborov deskriptor teksture koji uključuje i mjere korelacija između koeficijenata Gaborove transformacije na različitim skalama i orijentacijama. Eksperimentalno je pokazano da se na ovaj način poboljšavaju performanse klasifikatora. Pošto je prošireni Gaborov deskriptor teksture visokodimenzionalan, razmatrani su i različiti pristupi za redukciju njegove dimenzionalnosti.

Strukturalni deskriptor teksture koji uključuje mjere korelacija između modula koeficijenata Gaborove transformacije predložen je u **Glavi 5**. Pored samog deskriptora, predložena je i perceptualno zasnovana mjera sličnosti i algoritam za klasifikaciju zasnovan na principu najbližih susjeda i računanju sličnosti između testne slike i svake od klasa.

U **Glavi 6** razmatra se odnos globalnih i lokalnih deskriptora slika dobijenih daljinskom detekcijom. Pošto su informacije koje sadrže ova dva deskriptora komplementarne njihovom fuzijom se može postići poboljšanje performansi klasifikatora. Kao metod za fuziju deskriptora koji omogućava dobru generalizaciju predložena je hijerarhijska klasifikacija. Eksperimentalno je pokazano da se na ovaj način postiže poboljšanje performansi klasifikatora.

1.6 Publikacije

Neki od rezultata prezentovanih u ovoj disertaciji su već objavljeni u drugim publikacijama. U [23] su objavljeni eksperimentalni rezultati iz Glave 3, a ovdje je dodata teorijska osnova Gaborovog deskriptora teksture. Deskriptor razlika orijentacija i njegovo kvaternionsko proširenje na slike u boji su prvobitno objavljeni u [24], a varijanta proširenog Gaborovog deskriptora teksture objavljena je u [25]. Preliminarni eksperimenti sa redukcijom dimenzionalnosti Gaborovog deskriptora teksture i proširenog Gaborovog deskriptora teksture objavljeni su u [26] i [25]. Teorijska osnova i eksperimentalni rezultati za ove deskriptore su prošireni u Glavi 4. Strukturalni deskriptor teksture, predložen u Glavi 5, objavljen je u [27], a njegova implementacija na grafičkom procesoru opšte namjene je objavljena u [28]. Hijerarhijska fuzija globalnih i lokalnih deskriptora, predložena u Glavi 6, objavljena je u

GLAVA 1. UVOD

[25]. Lista publikacija kandidata u vezi sa disertacijom je data u Dodatku D. Pored toga, istraživanja provedena u ovoj disertaciji su dio projekta **Automatska klasi-fikacija pokrivenosti i načina korišćenja zemljišta**, sufinansiranog od strane Ministarstva nauke i tehnologije Republike Srpske.

Glava 2 Materijal i metodologija

Deskriptori i klasifikatori predloženi u ovoj disertaciji testirani su na četiri kolekcije slika dobijenih daljinskom detekcijom: Banja Luka, Window on the UK, UC Merced Land Use i Wuhan, te aero-snimku donjeg toka rijeke Sane. Slike u prve četiri kolekcije su dobijene podjelom aero i satelitskih snimaka visoke rezolucije (do 0,3 m) u blokove koji se jednoznačno mogu klasifikovati u klase pokrivača i načina korišćenja zemljišta. Pošto se radi o snimcima visoke rezolucije za klasifikaciju pokrivača i načina korišćenja zemljišta nije dovoljna spektralna informacija sadržana u pojedinim pikselima, već je potrebno uzeti u obzir i teksturu, oblik i prostorne odnose vizuelnih elemenata na slikama. Ove osobine čine korišćene kolekcije prikladnim za testiranje predloženih pristupa za klasifikaciju. Od navedenih kolekcija dvije, Banja Luka i Window on the UK, sadrže manji broj slika i klasa i korišćene su za inicijalne eksperimente i podešavanja parametara deskriptora i klasifikatora. Kolekcije UC Merced Land Use i Wuhan sadrže veći broj slika raspoređenih u 21, odnosno, 19 klasa načina korišćenja zemljišta i na njima su detaljno testirani svi predloženi deskriptori i klasifikatori. Neke od klasa u ovim kolekcijama su vizuelno homogene i određene teksturom na slici, za druge je, zbog vizuelne nehomogenosti, potrebno uzeti u obzir prostorni raspored tekstura, dok je izvjestan broj klasa određen postojanjem karakterističnih objekata na slici, bez obzira na ostatak slike. Potreba da se odredi kriterijum po kojem se klase razlikuju otežava zadatak koji se postavlja pred testirane klasifikatore. Aero-snimak donjeg toka rijeke Sane je uvršten u materijal kako bi se ispitala mogućnost primjene predloženih deskriptora i klasifikatora na čitav snimak, što je ranije uglavnom bilo rezervisano za spektralne [1] i objektno-orijentisane klasifikatore koji koriste nenadgledanu segmentaciju [9].

2.1 Kolekcija slika Banja Luka

Kolekcija slika Banja Luka nastala je podjelom RGB aero-snimka dijela grada Banje Luke. Čitav snimak je veličine 4500×6000 piksela i sadrži raznolike strukture, kako vještačke, kao što su zgrade, fabrike i skladišta, tako i prirodne, kao što su polja, drveće i rijeke. Prostorna rezolucija ovog snimka je 0,5 m. Dio ovog snimka je podijeljen u blokove veličine 128×128 piksela i u eksperimentima se koristi ukupno 606 ovih blokova za koje će u daljem tekstu biti korišćen termin *slike*. Sve slike su manuelno klasifikovane u šest klasa načina korišćenja zemljišta: *kuće, groblje, indu*-



Slika 2.1: Primjeri slika iz kolekcije Banja Luka. Sa lijeva u desno, po kolonama su slike iz klasa: *kuće, groblje, industrija, polje, rijeka* i *drveće*.



Slika 2.2: Raspodjela slika po kategorijama za kolekciju slika Banja Luka.

strija, polje, rijeka i drveće. Primjeri slika iz različitih klasa su prikazani na Slici 2.1. Raspodjela slika po klasama je prikazana na Slici 2.2. Vidi se da je ova raspodjela izrazito neravnomjerna sa velikim brojem primjera iz klasa *polje* ili *kuće*, a samo 28 primjera iz klase *groblje*. U eksperimentima u kojima se koristi ova kolekcija slika, polovina slika iz svake klase se koristi za treniranje, a druga polovina za testiranje klasifikatora. S obzirom na to da se radi o maloj kolekciji slika, upotrebljivost ove kolekcije je ograničena i korišćena je samo za inicijalne eksperimente sa Gaborovim deskriptorima te za podešavanje parametara strukturalnog deskriptora teksture.



Slika 2.3: Primjeri slika iz kolekcije Window on the UK. Sa lijeva u desno, po kolonama su slike iz klasa: *brodovi, zgrade, polje, trava, rijeka, put, drveće* i *vozila*.

2.2 Kolekcija slika Window on the UK

Kolekcija slika pod nazivom Window on the UK¹ je prvi put korišćena u radu [29]. Ovaj skup se sastoji od 1040 RGB aero-slika dimenzija 64×64 piksela čija prostorna rezolucija nije navedena. Sve slike su manuelno klasifikovane u osam klasa pokrivača i načina korišćenja zemljišta: *zgrađe, put, rijeka, polje, trava, drveće, brodovi* i *vozila*. Primjeri slika iz ove kolekcije dati su na Slici 2.3. U svakoj od kategorija je po 130 slika pa je raspodjela slika po kategorijama u ovom skupu uniformna. Autori su u [29] predložili i podjelu ovog skupa na trening i test skup sa po 520 slika koja je korišćena i u eksperimentima u ovoj disertaciji. U originalnom radu testirani su deskriptori boje, teksture i strukture sa klasifikatorom zasnovanim na linearnoj diskriminantnoj analizi, kao i na mašinama sa vektorima nosačima (*support vector machine, SVM*). Pošto se i u ovom slučaju radi o maloj kolekciji slika, u ovoj disertaciji je upotrebljena samo za inicijalne eksperimente sa Gaborovim deskriptorima.

2.3 Kolekcija slika UC Merced Land Use

Kolekcija slika UC Merced Land Use² se sastoji od slika dimenzija 256×256 piksela u RGB kolor-prostoru izdvojenih iz aero-snimaka sa prostornom rezolucijom od 0,3 m. Slike su manuelno klasifikovane u 21 klasu: poljoprivreda, avioni, bejzbol teren, plaža, zgrade, prerija, gusto naseljeno rezidencijalno zemljište, šuma, autoput, teren za golf, luka, raskrsnica, rezidencijalno zemljište srednje gustine, kamp prikolice, nadvožnjak, parking, rijeka, aerodromska pista, rijetko naseljeno rezidencijalno zemljište, cisterne i teniski tereni. Primjeri slika iz ove kolekcije dati su na Slici 2.4.

¹http://jfauqueur.free.fr/research/GTDB/

²http://vision.ucmerced.edu/datasets/

Svaka klasa sadrži 100 slika što ovaj skup čini najvećim javno dostupnim skupom za klasifikaciju slika dobijenih tehnikama daljinske detekcije. Ova kolekcija slika je prethodno korišćena u [30, 31] za testiranje klasifikatora zasnovanih na SIFT deskriptorima i skupu riječi, te u opsežnoj analizi lokalnih deskriptora u pretraživanju baza slika dobijenih daljinskom detekcijom [32]. Autori su koristili po 80 slika iz svake klase za obučavanje klasifikatora i testirali ga na ostalim slikama. Ovaj ciklus se ponavlja pet puta i navode se srednje vrijednosti dobijenih tačnosti klasifikacije. I u ovoj disertaciji je prihvaćena ista metodologija testiranja s tim što se, kako bi se analizirao uticaj broja trening primjera na tačnost klasifikacije, u eksperimentima dodatno varira veličina trening skupa.

2.4 Kolekcija slika Wuhan

Wuhan kolekcija se sastoji od 950 satelitskih slika rezolucije 0,5 m preuzetih sa servisa Google Earth. Sve slike su u RGB kolor-prostoru i dimenzija 600 × 600 piksela. Slike su podijeljene u 19 klasa načina korišćenja zemljišta: *aerodrom, plaža, most, poslovne zgrade, pustinja, njive, šuma, industrija, livada, planine, park, parking, jezero, luka, rezidencijalno zemljište, rijeka, vijadukt, stadion* i *željeznička stanica.* Primjeri slika iz ove kolekcije dati su na Slici 2.5. U svakoj klasi se nalazi po 50 slika. Ova kolekcija slika je korišćena u [33] za testiranje različitih lokalnih deskriptora sa klasifikatorima na principu najbližih susjeda i mašinama sa vektorima nosačima, a njena verzija sa 12 klasa načina korišćenja zemljišta korišćena je u [34] i [35]. U ovim radovima autori koriste po 25 slika iz svake klase za obučavanje klasifikatora, a ostatak za njegovo testiranje. Ciklus obučavanja i testiranja klasifikatora se ponavlja 10 puta i u rezultatima se navode srednje vrijednosti dobijenih tačnosti klasifikacije. U ovoj disertaciji se, pored opisane metodologije testiranja, analizira i uticaj veličine trening skupa na tačnost klasifikacije pa se eksperimenti ponavljaju sa različitim veličinama trening skupa.

2.5 Aero-snimak donjeg toka rijeke Sane

Pored kolekcija slika formiranih izdvajanjem blokova određene veličine iz visokorezolucionih snimaka dobijenih daljinskom detekcijom, deskriptori i klasifikatori predloženi u ovoj disertaciji su testirani i na kompletnom RGB snimku geografskog područja donjeg toka rijeke Sane do Novog Grada i ušća u rijeku Unu [36]. Ovaj snimak je dobijen aerofoto snimanjem pomenutog područja, a zatim je izvršena njegova aerotriangulacija i formiran je digitalni model terena. Snimak je zatim ortorektifikovan i izrađeni su listovi digitalnog ortofoto plana u razmjeri 1:5000, prostorne rezolucije 0,5 m. Dio dobijenog snimka prikazan je na Slici 1.1.

Na ovom snimku su manuelno označena *urbana područja* i u eksperimentalnom dijelu ove disertacije korišćen je za testiranje automatskog detektora urbanih područja. Detektor urbanih područja je, u stvari, binarni klasifikator kojim se snimak klasifikuje u dvije klase načina korišćenja zemljišta – urbano i neurbano. Klasifikator je obučavan na slučajno izabranim blokovima iz obje klase, a testiran je na dijelu snimka iz kojeg nisu izdvajani blokovi korišćeni za obučavanje.



Slika 2.4: Primjeri slika iz kolekcije UC Merced Land Use: (a) poljoprivreda, (b) avioni, (c) bejzbol teren, (d) plaža, (e) zgrade, (f) prerija, (g) gusto naseljeno rezidencijalno zemljište, (h) šuma, (i) autoput, (j) teren za golf, (k) luka, (l) raskrsnica, (m) rezidencijalno zemljište srednje gustine, (n) kamp prikolice, (o) nadvožnjak, (p) parking, (q) rijeka, (r) aerodromska pista, (s) rijetko naseljeno rezidencijalno zemljište, (t) cisterne, (u) teniski tereni.



Slika 2.5: Primjeri slika iz kolekcije Wuhan: (a) aerodrom, (b) plaža, (c) most, (d) poslovne zgrade, (e) pustinja, (f) njive, (g) stadion, (h) šuma, (i) industrija, (j) livada, (k) planine, (l) park, (m) parking, (n) jezero, (o) luka, (p) željeznička stanica, (q) rezidencijalno zemljište, (r) rijeka, (s) vijadukt.

Ortofoto snimak donjeg toka rijeke Sane na kojem su manuelno označena urbana područja obezbijedila je firma Geo-centar d.o.o. iz Banje Luke u okviru projekta **Automatska analiza pokrivenosti i načina korišćenja zemljišta**.

2.6 Metodologija testiranja

Kao što je pomenuto, slike u svim kolekcijama su u RGB kolor-prostoru, odnosno, sadrže crveni, zeleni i plavi spektralni opseg. U eksperimentima su, pored RGB, korišćene i panhromatske verzije ovih slika. One su generisane polazeći od RGB slika pomoću jednačine

$$I(x,y) = 0,299r(x,y) + 0,587g(x,y) + 0,114b(x,y), \qquad (2.1)$$

gdje su r(x, y), g(x, y) i b(x, y) vrijednosti piksela na koordinatama (x, y) u crvenom, zelenom i plavom spektralnom opsegu, respektivno. U digitalnoj obradi slike, signal određen jednačinom (2.1) se često naziva luminentni signal [37].

U svim eksperimentima korišćena kolekcija slika podijeljena je na dva dijela: trening i test skup. Klasifikatori su obučeni korišćenjem primjera iz trening skupa, a zatim su performanse klasifikatora mjerene procentom tačno klasifikovanih slika iz test skupa, koje nisu korišćene pri obučavanju klasifikatora. Ova procedura je u izvjesnoj mjeri drugačija za klasifikator na principu najbližih susjeda zato što kod njega ne postoji eksplicitna faza obučavanja. Ipak, za njegovo funkcionisanje potreban je skup označenih primjera koji je formalno ekvivalentan trening skupu.

Kako bi se obezbijedila nezavisnost tačnosti klasifikacije od konkretne podjele kolekcije na trening i test skup, ciklus obučavanja i testiranja klasifikatora se ponavlja više puta sa različitim, slučajno izabranim, podjelama. U rezultatima se, zatim, navode srednje vrijednosti i standardne devijacije tačnosti klasifikacije dobijenih za pojedine podjele. Pored tačnosti klasifikacije kao jedinstvene mjere performansi klasifikatora, za klasifikatore koji su pokazali najbolje rezultate date su i matrice konfuzija. Na ovaj način se dobija uvid u tačnosti klasifikacije po klasama, kao i klase za koje se najčešće javljaju greške u klasifikaciji.

Da bi se dobijeni rezultati mogli uporediti sa rezultatima iz literature, u eksperimentima se koriste isti protokoli testiranja kao i u drugim radovima u kojima su korišćene iste kolekcije. Konkretni detalji protokola, veličine trening i test skupova te broj ponavljanja ciklusa obučavanja i testiranja klasifikatora, navedeni su u opisima pojedinih kolekcija slika. Pored opisanih protokola testiranja, izvršeni su i eksperimenti u kojima je varirana veličina trening skupa kako bi se dobila zavisnost tačnosti klasifikacije od broja trening primjera.

Glava 3 Gaborov deskriptor teksture

Vizuelni sistem obuhvata oči i dio centralnog nervnog sistema, a njegov zadatak je akvizicija i obrada vizuelnih informacija. Oči su zadužene za akviziciju podataka čiji je fizički nosilac svjetlost, a njihova obrada počinje u mrežnjači gdje fotoreceptori pobuđeni svjetlošću generišu nervne impulse koji se optičkim nervom prenose u dio moždane kore koji se nalazi u zatiljnom režnju, u zadnjem dijelu mozga, i poznat je pod nazivom vizuelni korteks (vizuelna kora mozga). Hubel i Wiesel [38, 39] su iznijeli pretpostavku da je vizuelni korteks organizovan hijerarhijski, tj. da se receptivna polja¹ ćelija na višem nivou vizuelnog korteksa formiraju tako da primaju signale iz ćelija na nižem nivou. Tako se velika, složena receptivna polja formiraju kombinovanjem malih, jednostavnih receptivnih polja. Vremenom je, kroz mnogobrojne fiziološke i psihofizičke eksperimente, ova pretpostavka i potvrđena [40].

Vizuelni korteks se sastoji od strijatnog korteksa (primarni vizuelni korteks ili oblast V1) i ekstrastrijatnih oblasti (oblasti V2, V3, V4 i V5). Hubel i Wiesel su ćelije u strijatnom korteksu podijelili u tri grupe: proste, kompleksne i hiperkompleksne ćelije. Proste ćelije su detektori ivica i linija određene orijentacije, kompleksne reaguju na pokretne linije određene orijentacije, a hiperkompleksne na linije određene orijentacije i dužine.

Pretpostavku da se receptivna polja prostih ćelija u oblasti V1 mogu modelirati Gaborovim funkcijama prvi je iznio Marčelja [41]. On je pokazao da se rezultati eksperimentalnih mjerenja profila receptivnih polja prostih ćelija mogu opisati jednodimenzionalnim Gaborovim funkcijama. Kada su eksperimentalni podaci o dvodimenzionalnoj strukturi receptivnih polja prostih ćelija postali raspoloživi, Daugman [42] je proveo detaljnu analizu njihovog odnosa sa dvodimenzionalnim Gaborovim funkcijama i pokazao njihovo značajno slaganje. Konačno, Jones i Palmer su u [43] pokazali da receptivna polja nekih prostih ćelija u oblasti V1 kod mačaka odgovaraju dvodimenzionalnim realnim Gaborovim funkcijama. Od izuzetnog značaja je i otkriće Pollena i Ronnera koji su pokazali u [44] da su receptivna polja susjednih prostih ćelija u kvadraturi što je poslužilo kao motivacija da se za njihovo modeliranje koriste kompleksne Gaborove funkcije čiji su realni i imaginarni dio u kvadraturi. Pošto će u nastavku isključivo biti korišćene kompleksne Gaborove funkcije za njih će biti korišćen koristiti termin *Gaborove funkcije* dok će u slučajevima kada se koriste realne Gaborove funkcije ta činjenica biti posebno naglašena.

 $^{^1 \}mathrm{Receptivno}$ polje je dio senzorske nervne ćelije čija stimulacija rezultuje odzivom ćelije.

3.1 Modeliranje receptivnih polja prostih ćelija pomoću Gaborovih filtara

Jednodimenzionalna Gaborova funkcija je sinusoida modulisana Gausovom funkcijom. Inicijalno je predložena za vremensko-frekvencijsku analizu signala [45]. Poznato je da se pri vremensko-frekvencijskoj analizi signala rezolucija u vremenu i rezolucija u frekvenciji ne mogu nezavisno podešavati i njihov proizvod je ograničen odozdo. Ovo ograničenje je ekvivalent Hajzenbergovog principa neodređenosti u kvantnoj mehanici i onemogućava istovremeno postizanje proizvoljno dobrih rezolucija u vremenu i frekvenciji. Kada se za vremensko-frekvencijsku analizu signala koriste Gaborove funkcije postiže se teorijski minimum proizvoda neodređenosti u vremenu i frekvenciji. Dakle, Gaborovim funkcijama se minimizira združena neodređenost u vremenu i frekvenciji što ih čini vrlo pogodnim izborom u oblasti obrade signala i slike. Dvodimenzionalnu Gaborovu funkciju je uveo Daugman [42] upravo kako bi modelirao dvodimenzionalna receptivna polja prostih ćelija i pokazao je da ona takođe minimizira proizvod neodređenosti u prostoru i prostornoj frekvenciji².

Najopštiji oblik funkcije koja minimizira proizvod neodređenosti u prostoru i frekvenciji je

$$e^{-\left(Ax^2+Bxy+Cy^2+Dx+Ey+F\right)}.$$
(3.1)

gdje je $B^2 < 4AC$, a D, E i F su kompleksni brojevi [42]. Kod Gaborove funkcije je $\operatorname{Re}\{D\} = \operatorname{Re}\{E\} = \operatorname{Re}\{F\} = 0$, a pošto su receptivna polja prostih ćelija u kvadraturi onda i realni i imaginarni dio funkcije treba da budu u kvadraturi. Iz ovoga slijedi da možemo usvojiti $\operatorname{Im}\{F\} = 0$. Konačno, jednostavnosti radi, smatraćemo da Gausova i sinusna funkcija imaju istu orijentaciju pa je B = 0. Takođe, bez gubitka opštosti, može se usvojiti da je E = 0. Gaborova funkcija ima isti funkcionalni oblik u prostornom i frekvencijskom domenu.

Saznanje da se receptivna polja prostih ćelija u oblasti V1 mogu modelirati Gaborovim funkcijama iniciralo je upotrebu filtara čiji su impulsni odzivi Gaborove funkcije za analizu slike u biološki inspirisanim sistemima računarskog vida. U nastavku će za filtar čiji je impulsni odziv Gaborova funkcija biti korišćen termin *Gaborov filtar*. Prvi korak u analizi signala je, dakle, njegovo filtriranje Gaborovim filtrom, odnosno, konvolucija sa Gaborovom funkcijom čime se model receptivnog polja lokalno primjenjuje na čitavu sliku.

Iz (3.1) se, uz navedene uslove, dobija oblik Gaborove funkcije koji će biti korišćen u ovom radu [46]

$$g(x,y) = \frac{1}{2\pi\sigma_x\sigma_y} e^{-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2}\right)} e^{j\omega x},$$
(3.2)

gdje je ω frekvencija Gaborove funkcije, a σ_x i σ_y određuju širinu njene Gausove anvelope. Furijeova transformacija ove funkcije je:

$$G\left(\omega_x, \omega_y\right) = e^{-\frac{1}{2} \left[\frac{\left(\omega_x - \omega\right)^2}{\sigma_{\omega_x}^2} + \frac{\omega_y^2}{\sigma_{\omega_y}^2}\right]},\tag{3.3}$$

pri čemu je $\sigma_{\omega_x} = \frac{1}{\sigma_x}$ i $\sigma_{\omega_y} = \frac{1}{\sigma_y}$. Na Slici 3.1 prikazani su primjeri realnog i imaginarnog dijela Gaborove funkcije, kao i modula njene Furijeove transformacije.

 $^{^2\}mathrm{U}$ daljem tekstu će umjesto termina prostorna frekvencija biti korišćen termin frekvencija.

Sa Slike 3.1e može se primjetiti da amplitudni spektar Gaborove funkcije ima oblik Gausove funkcije, što znači da je odgovarajući Gaborov filtar propusnik opsega, Slika 3.1e. Širina propusnog opsega ovog filtra određena je standardnim devijacijama njegove Gausove anvelope.

Pošto su u oblasti V1 pronađene ćelije sa receptivnim poljima koja rezultuju maksimalnim odzivima na različitim skalama i orijentacijama vizuelne pobude, u modelu receptivnih polja prostih ćelija se koristi banka Gaborovih filtara na različitim skalama i orijentacijama. Impulsni odzivi Gaborovih filtara na različitim skalama i orijentacijama su skalirane i rotirane verzije Gaborove funkcije g(x, y)

$$g_{mn}(x,y) = a^{-m}g(x',y'), \qquad (3.4)$$

$$x' = a^{-m} \left(x \cos \theta + y \sin \theta \right), \qquad (3.5)$$

$$y' = a^{-m} \left(-x \sin \theta + y \cos \theta \right), \qquad (3.6)$$

za a > 1 i $\theta = n\pi/K$, pri čemu su $m = 0, \ldots, S - 1$ i $n = 0, \ldots, K - 1$ indeksi skale i orijentacije, respektivno. Skalirajući faktor a^{-m} u jednačini (3.4) služi za normalizaciju energije.

Neka je sada $I(x, y), (x, y) \in \Phi$, slika čija reprezentacija se određuje. Izlaz iz filtra $g_{mn}(x, y)$ dat je konvolucijom

$$w_{mn}(x,y) = I(x,y) * g_{mn}(x,y) = = \iint_{\Phi} I(\xi,\eta) g_{mn}^{*}(x-\xi,y-\eta) d\xi d\eta,$$
(3.7)

gdje je Φ skup tačaka slike.

Pošto se Gaborova filtar banka sastoji od filtara čiji se impulsni odzivi dobijaju skaliranjem i rotiranjem Gaborove funkcije, može se povući paralela sa filtar bankom kojom se realizuje talasna (*wavelet*) transformacija [37]. Zbog toga se za jednačinu (3.7) koristi termin Gaborova transformacija slike I(x, y). Analogno talasnoj transformaciji, Gaborovu funkciju g(x, y) nazivaćemo talasić majka, a izlazni signal iz Gaborove filtar banke, $w_{mn}(x, y)$, predstavlja koeficijente Gaborove transformacije na skali m i orijentaciji n. Takođe, pojedine skale i orijentacije Gaborove transformacije nazivaćemo podopsezi.

U mnogim primjenama, kao npr. u kompresiji signala, se koristi ortogonalna talasna transformacija, pa je onda i odgovarajuća filtar banka takođe ortogonalna, a dobijeni koeficijenti su nekorelisani. Sa druge strane, Gaborova filtar banka nije ortogonalna pa je neophodno na neki način kontrolisati redundansu sadržanu u koeficijentima transformacije. Kako bi se ovo postiglo, parametri a, σ_{ω_x} i σ_{ω_y} , a time i σ_x i σ_y , se biraju tako da se amplitudne karakteristike filtara dodiruju na polovini njihovog maksimuma [46]. Polazeći od tog uslova ovi parametri se izračunavaju


(a) Realni dio 2D Gaborove funkcije.



(c) Uzorak luminanse koji odgovara realnom dijelu 2D Gaborove funkcije.



(e) Amplitudni spektar 2D Gaborove funkcije.



(b) Imaginarni dio 2D Gaborove funkcije.



(d) Uzorak luminanse koji odgovara imaginarnom dijelu 2D Gaborove funkcije.



(f) Uzorak luminanse koji odgovara amplitudnom spektru 2D Gaborove funkcije.

Slika 3.1: Primjer 2D Gaborove funkcije.

korišćenjem sljedećih jednačina:

$$a = \left(\frac{\omega_h}{\omega_l}\right)^{\frac{1}{M-1}},\tag{3.8}$$

$$\sigma_{\omega_x} = \frac{(a-1)\,\omega_h}{(a+1)\,\sqrt{2\ln 2}},\tag{3.9}$$

$$\sigma_{\omega_y} = \operatorname{tg}\left(\frac{\pi}{2K}\right) \left[\omega_h - 2\ln 2\left(\frac{\sigma_{\omega_x}^2}{\omega_h}\right)\right] \left[2\ln 2 - \frac{(2\ln 2)^2 \sigma_{\omega_x}^2}{\omega_h^2}\right]^{-\frac{1}{2}},\qquad(3.10)$$

gdje je frekvencija Gaborovog talasića majke data sa $\omega = \omega_h$. U gornjim jednačinama ω_l i ω_h predstavljaju donju i gornju centralnu frekvenciju od interesa. Kako globalni nivo osvijetljenosti slike ne bi uticao na izlaze iz filtara, realni dio dvodimenzionalne Gaborove funkcije treba da ima nultu srednju vrijednost. Ovo se može postići dodavanjem odgovarajuće konstante ili postavljanjem G(0,0) u jednačini (3.3) na nulu. Na Slici 3.2 prikazani su poluprofili amplitudnih karakteristika filtara iz Gaborove filtar banke. Banka je formirana od filtara na S = 4 skale i K = 6orijentacija, uz digitalne centralne frekvencije $\omega_l = 0, 1\pi$ i $\omega_h = 0, 8\pi$.



Slika 3.2: Osjenčene elipse predstavljaju frekvencijske opsege u kojima amplitudne karakteristike filtara iz Gaborove filtar banke imaju vrijednosti veće od polovine maksimuma amplitudne karakteristike.

Gaborovi filtri se koriste u mnogim algoritmima za analizu i klasifikaciju slika. Neki od najpoznatijih primjera uključuju: identifikaciju osoba pomoću dužice oka [47], prepoznavanje lica [48], klasifikaciju scena [21, 49], prepoznavanje objekata [15], detekciju ciljeva na satelitskim slikama [50], segmentaciju slika [51], detekciju istaknutih regiona na slikama [52], kao i analizu teksture [46, 53].

3.2 Gaborov deskriptor teksture

Klasifikacija tekstura je jedna od oblasti u kojoj su Gaborovi filtri često korišćeni i, pod određenim uslovima, pokazuju dobre rezultate [53, 54]. Kako bi se formirala reprezentacija teksture, slika se prvo filtrira Gaborovom filtar bankom (3.7). Ako pretpostavimo da je tekstura na slici homogena, možemo smatrati da koeficijenti Gaborove transformacije imaju istu raspodjelu vjerovatnoće na čitavoj slici. Prema tome, za reprezentaciju slike se mogu upotrebiti jednostavne statistike koeficijenata Gaborove transformacije, w_{mn} . Gaborov deskriptor teksture sadrži srednje vrijednosti, μ_{mn} , i standardne devijacije, σ_{mn} , modula koeficijenata Gaborove transformacije izračunate usrednjavanjem po pikselima slike [46]

$$\mu_{mn} = \iint_{\Phi} |w_{mn}(x, y)| \, dx dy, \qquad (3.11)$$

$$\sigma_{mn} = \sqrt{\iint_{\Phi} \left(|w_{mn}\left(x,y\right)| - \mu_{mn}\right)^2 dx dy}.$$
(3.12)

U Gaborov deskriptor teksture **TD** se uključuju veličine date jednačinama (3.11) i (3.12) na skalama m = 1, ..., S i orijentacijama n = 1, ..., K

$$\mathbf{TD} = \left[\mu_{00} \ \sigma_{00} \ \mu_{01} \ \sigma_{01} \ \dots \ \mu_{S-1,K-1} \ \sigma_{S-1,K-1}\right]^T.$$
(3.13)

Sličan deskriptor sa bankom filtara na 5 skala i 6 orijentacija usvojen je i u MPEG-7 standardu [55], a u MPEG-7 eksperimentalnom modelu predložena je implementacija Gaborovih filtara u polarnim koordinatama korišćenjem Radonove transformacije [56].

Udaljenost dva Gaborova deskriptora tekstur
e \mathbf{TD}_1 i \mathbf{TD}_2 određuje se korišćenjem težinske Menhet
n metrike

$$d\left(\mathbf{TD}_{1},\mathbf{TD}_{2}\right) = \sum_{m=1}^{S} \sum_{n=1}^{K} \left(\left| \frac{\mu_{mn}^{(1)} - \mu_{mn}^{(2)}}{\alpha \left(\mu_{mn}\right)} \right| + \left| \frac{\sigma_{mn}^{(1)} - \sigma_{mn}^{(2)}}{\alpha \left(\sigma_{mn}\right)} \right| \right),$$
(3.14)

gdje su $\alpha(\mu_{mn})$ i $\alpha(\sigma_{mn})$ standardne devijacije μ_{mn} i σ_{mn} određene na trening skupu.

3.3 Klasifikacija pokrivača i načina korišćenja zemljišta korišćenjem Gaborovih deskriptora

U nedavnoj studiji [54] Gaborov deskriptor teksture je pokazao robusnost na varijacije spektralnih karakteristika osvjetljenja te relativno malu računsku složenost, ali nešto lošije performanse u prisustvu afinih i neafinih transformacija. Pošto se klasifikacija pokrivača i načina korišćenja zemljišta uglavnom obavlja na ortofoto snimcima na kojima nema transformacija usljed promjene perspektive, Gaborov deskriptor teksture je dobar izbor za ovaj zadatak. Ipak, translacije i rotacije su uobičajene na snimcima dobijenim daljinskom detekcijom, ali njihov uticaj se može ublažiti uključivanjem transformisanih primjera u trening skup. Dobre performanse Gaborovog deskriptora teksture na zadatku klasifikacije pokrivača i načina korišćenja zemljišta potvrđene su npr. u [23] i [57].

U ovoj disertaciji Gaborov deskriptor teksture se koristi kao polazna tačka za konstruisanje novih deskriptora za klasifikaciju pokrivača i načina korišćenja zemljišta. Takođe, tačnosti klasifikacije dobijene korišćenjem Gaborovog deskriptora teksture će u nastavku poslužiti kao referenca u odnosu na koju se ocjenjuje uspješnost predloženih deskriptora i klasifikatora. Zbog toga su tačnosti klasifikacije dobijene za slike iz UC Merced Land Use i Wuhan kolekcija date u eksperimentalnim rezultatima u Glavi 4, a ovdje su, ilustracije radi, dati samo rezultati dobijeni za manje kolekcije Banja Luka i Window on the UK.

Osim Gaborovog deskriptora teksture u ovim eksperimentima testiran je i Gist deskriptor koji je vrlo popularan referentni deskriptor u klasifikaciji prirodnih scena [21], a takođe je zasnovan na statistikama izlaza iz Gaborove filtar banke. Kratak opis Gist deskriptora dat je u Dodatku B. Iako je često korišćen u klasifikaciji scena i predložena su njegova unapređenja [49, 58], nema mnogo primjera korišćenja Gist deskriptora u klasifikaciji slika dobijenih daljinskom detekcijom. Izuzeci su korišćenje Gist deskriptora za klasterizaciju slika u okviru detekcije drveća [59] i za detekciju ciljeva na satelitskim slikama [50], kao i poboljšani Gist model koji se koristi za detekciju kratera u [60].

U eksperimentima opisanim u ovom odjeljku, Gaborovi deskriptori slika se računaju na 8 skala i 8 orijentacija. Rezultati za filtar banke sa drugačijim brojem skala i orijentacija su slični. Gaborov deskriptor teksture, predložen u [46], se izračunava za panhromatske slike. Pošto su slike u obje kolekcije u RGB kolor-prostoru, izračunavaju se Gaborovi deskriptori za sva tri spektralna opsega slike i konačni deskriptor slike se dobija njihovom konkatenacijom. Na ovaj način, slike su predstavljene $3 \times 8 \times 8 \times 2 = 384$ -dimenzionalnim deskriptorima. Poređenja radi, računaju se i Gaborovi deskriptori za panhromatske verzije slika i oni su $8 \times 8 \times 2 = 128$ dimenzionalni.

Za izračunavanje Gist deskriptora zadržane su nominalne vrijednosti parametara, tj. filtar banka na 4 skale i 8 orijentacija. I za ovaj deskriptor se računaju panhromatska varijanta, koja je $4 \times 8 \times 16 = 512$ -dimenzionalna i RGB varijanta, koja daje $3 \times 4 \times 8 \times 3 = 1536$ -dimenzionalan deskriptor.

Klasifikator za oba deskriptora je mašina sa vektorima nosačima. Pošto se udaljenosti Gaborovih deskriptora računaju pomoću (3.14), na osnovu ove metrike se konstruiše kernel

$$K\left(\mathbf{TD}_{i}, \mathbf{TD}_{j}\right) = \exp\left[-d\left(\mathbf{TD}_{i}, \mathbf{TD}_{j}\right)\right], \qquad (3.15)$$

gdje su \mathbf{TD}_i i \mathbf{TD}_j Gaborovi deskriptori teksture dati sa (3.13), a $d(\mathbf{TD}_i, \mathbf{TD}_j)$, je njihova udaljenost, data sa (3.14). Ova kernel funkcija je suštinski zasnovana na težinskoj L_1 -normi i zadovoljava Merserov uslov [61]. Za Gist deskriptor se, kao u [21], koristi SVM sa kernelom zasnovanim na radijalnim baznim funkcijama.

SVM je binarni klasifikator, ali se može proširiti i koristiti za klasifikaciju u više klasa tako što se obučava onoliko binarnih klasifikatora koliko postoji klasa u koje se slike klasifikuju. Pri tome se, pri obučavanju svakog od klasifikatora, smatra da su primjeri iz jedne klase pozitivni, a primjeri iz svih ostalih klasa negativni (jedan-protiv-svih). Testni primjer se klasifikuje u kategoriju koja odgovara SVM sa maksimalnom izlaznom vrijednošću.

Deskriptor	Panhromatske (%)	RGB (%)
Gabor (pun)	84.5	88.0
Gabor (sr. vr.)	80.7	84.5
Gist	79.5	89.3

Tabela 3.1: Tačnosti klasifikacije za kolekciju slika Banje Luke.

3.3.1 Kolekcija slika Banja Luka

U eksperimentima sa ovom kolekcijom slika, klasifikatori su testirani korišćenjem unakrsne validacije sa 10 različitih slučajnih podjela kolekcije slika. U svakoj podjeli klasifikator se trenira sa po 50% slika iz svake od klasa, a testira na ostatku. Dobijene tačnosti klasifikacije za sve podjele se usrednjavaju. U Tabeli 3.1 su date srednje tačnosti klasifikacije za sve kategorije, pri čemu je sa Gabor (pun) označen Gaborov deskriptor teksture dat sa (3.13), dok je sa Gabor (sr. vr.) označen deskriptor dobijen korišćenjem samo srednjih vrijednosti odziva filtara.

Najbolje performanse se dobijaju korišćenjem Gist deskriptora izračunatog za sve spektralne opsege RGB slike. Cijena koja se pri tome plaća je visoka dimenzionalnost deskriptora. Vrijedi primjetiti da znatno jednostavniji Gaborov deskriptor, čija je dimenzionalnost četiri puta niža, daje slične rezultate. Još značajnija je činjenica da za panhromatske slike Gaborov deskriptor daje bolje rezultate od Gist deskriptora. Iz ovih rezultata jasno je da uključivanje informacija iz različitih spektralnih opsega poboljšava performanse klasifikatora. Kada se radi sa panhromatskim slikama, standardne devijacije odziva Gaborove filtar banke obezbjeđuju bogatiju informaciju o teksturi što rezultuje boljim performansama. Značaj ove informacije može se vidjeti iz pada performansi kada se koriste samo srednje vrijednosti odziva Gaborove filtar banke. Drugi zaključak koji se može izvesti iz ovih rezultata je da, za razliku od scena iz prirođe, prostorni raspored odziva filtar banke ne poboljšava performanse klasifikatora snimaka dobijenih daljinskom detekcijom. Sličan zaključak je izveden i u [30] za slučaj lokalnih deskriptora i kernela sa piramidalnim uparivanjem.

Matrica konfuzija za Gaborov deskriptor je data na Slici 3.3a. Vidi se da su konfuzije najčešće između kategorija koje su teške i za ljude. Najistaknutiji primjeri konfuzija su između klasa *kuće* i groblja, zbog postojanja pravougaonih struktura sa izraženim orijentisanim ivicama, te *rijeka* i *polje*, zato što obje klase karakteriše homogena, fina tekstura bez izraženih ivica. Važno je naglasiti i da su konfuzije između prirodnih (*rijeka, drveće, polje*) i vještačkih (*kuće, groblje, industrija*) klasa rijetke.

Matrica konfuzija za Gist deskriptor je data na Slici 3.3b. I u ovom slučaju mogu se izvesti slični zaključci kao i za matricu konfuzija dobijenu za Gaborov deskriptor teksture.

3.3.2 Window on the UK kolekcija slika

U drugom eksperimentu korišćena je Window on the UK kolekcija slika koja je korišćena i u radu [29] u kojem se više deskriptora boje, teksture i strukture spaja u 310-dimenzionalnu reprezentaciju slike. U istom radu predložena je i podjela ove



(a) Matrica konfuzija kada se koristi Gaborov deskriptor teksture za RGB slike.



(b) Matrica konfuzija kada se koristi Gist deskriptor za RGB slike.

Slika 3.3: Matrice konfuzija za kolekciju slika Banja Luka.

Tabela 3.2: Tačnosti klasifikacije n	a Window on th	e UK kolekciji slika
--------------------------------------	----------------	----------------------

Metod	Tačnost (%)
SVM sa Gaborovim deskriptorom (RGB)	90.8
SVM sa Gist deskriptorom (RGB)	87.1
Algoritam iz [29]	89.4
SVM sa deskriptorima iz [29]	92.3

kolekcije na trening i test skupove sa po 520 slika. U ovoj disertaciji je, kako bi bilo moguće poređenje rezultata, usvojena ista podjela.

Tačnosti klasifikacije za ovu kolekciju slika date su u Tabeli 3.2. I u ovom slučaju, Gaborovim i Gist deskriptorom se postižu slične performanse, ali je ovaj put blaga prednost na strani Gaborovog deskriptora. Ovo je u skladu sa prethodnim zaključcima o diskriminativnoj moći ova dva deskriptora. Pored toga, vidi se da su performanse klasifikatora sa Gaborovim deskriptorom bolje od performansi algoritma predloženog u [29], a samo malo lošije od performansi SVM klasifikatora obučenog korišćenjem deskriptora iz istog rada.

Matrica konfuzija za Gaborov deskriptor je prikazana na Slici 3.4. Vidimo da se najčešće pogrešne klasifikacije ponovo javljaju u slučajevima koji mogu zbuniti i ljude, kao što su *zgrada* nasuprot *vozila* ili *polja* nasuprot *trave*. Važno je primjetiti da se, i u ovom slučaju, rijetko javljaju konfuzije između prirodnih i vještačkih kategorija.

Iz prikazanih rezultata može se zaključiti da Gaborov deskriptor teksture u kombinaciji sa SVM klasifikatorom i kernelom zasnovanom na težinskoj L_1 normi predstavlja dobru osnovu za klasifikaciju slika dobijenih daljinskom detekcijom u klase pokrivača i načina korišćenja zemljišta. Uprkos svojoj jednostavnosti, opisani pristup je nadmašio Gist deskriptor koji ima znatno veću dimenzionalnost i predstavlja referentni deskriptor u klasifikaciji scena. Takođe, iako Gaborov deskriptor pred-



Slika 3.4: Matrica konfuzija za Window on the UK kolekciju slika kada se koristi Gaborov deskriptor teksture.

stavlja samo reprezentaciju teksture na slici, iz Tabele 3.2 se vidi da su dobijeni rezultati, u najgorem slučaju, bliski rezultatima iz [29], dobijenim korišćenjem deskriptora boje, teksture i strukture.

Glava 4

Prošireni Gaborov deskriptor teksture

Dobri rezultati koji se dobijaju korišćenjem Gaborovog deskriptora teksture u klasifikaciji slika dobijenih daljinskom detekcijom u klase pokrivača i načina korišćenja zemljišta inspirisali su dalje istraživanje u ovoj oblasti. Jedan od zanimljivih smjerova za proširenje Gaborovog deskriptora teksture je korišćenje statističkih osobina koeficijenata Gaborove transformacije. Moguće je, na primjer, matematički modelirati funkciju gustine raspodjele koeficijenata i koristiti dobijeni model za klasifikaciju ili pretraživanje baza slika, kao što je, na primjer, urađeno u [62] za slučaj talasne transformacije. U ovoj disertaciji je usvojen nešto jednostavniji pristup koji je zasnovan na postojanju statističkih zavisnosti između koeficijenata Gaborove transformacije koji potiču iz različitih podopsega.

4.1 Statističke zavisnosti između podopsega Gaborove transformacije

Ortogonalnom talasnom transformacijom uklanja se korelacija između koeficijenata u različitim podopsezima. Međutim, iako su koeficijenti dekorelisani, statističke zavisnosti višeg reda između njih i dalje postoje [63]. Ovo se najlakše može uočiti posmatranjem modula talasnih koeficijenata. Zavisnosti postoje između modula koeficijenata na istim relativnim koordinatama u različitim podopsezima, kao i između modula susjednih koeficijenata u istom podopsegu.

Gaborova transformacija nije ortogonalna zbog čega postoje statističke zavisnosti između koeficijenata na različitim skalama, odnosno, orijentacijama. Ove zavisnosti se mogu uočiti na Slici 4.1 na kojoj su prikazane amplitudne karakteristike Gaborovih filtara (drugi red) i odgovarajući moduli koeficijenata Gaborove transformacije na skalama 2 i 3, te orijentacijama 1 i 2 (treći red) za sliku iz prvog reda. Iz ovog primjera vidi se da moduli koeficijenata koji se nalaze na istim relativnim prostornim lokacijama u podopsezima na različitim skalama i orijentacijama imaju slične vrijednosti što ukazuje na postojanje njihovih statističkih zavisnosti.

Uvid u statističku zavisnost koeficijenata može se dobiti posmatranjem njihovih uslovnih histograma. Uslovni histogram neke slučajne promjenljive uslovljene dru-



Slika 4.1: Moduli koeficijenata Gaborove transformacije na različitim skalama (m) i orijentacijama (n).

gom slučajnom promjenljivom je sastavljen od svih histograma te slučajne promjenljive određenih za različite vrijednosti (ili opsege vrijednosti) slučajne promjenljive kojom je uslovljena. Uslovni histogram predstavlja estimaciju uslovne raspodjele vjerovatnoće neke slučajne promjenljive i može se grafički prikazati kao intenzitetska slika na kojoj su intenziteti piksela proporcionalni vrijednostima u ćelijama histograma, pri čemu su kolone skalirane nezavisno jedna od druge kako bi se iskoristio čitav opseg vrijednosti intenziteta.

Na Slici 4.2 su prikazani dvodimenzionalni uslovni histogrami realnih i imaginarnih dijelova Gaborovih koeficijenata na skalama 2 i 3 i orijentaciji 1, a na Slici 4.3, na orijentacijama 1 i 2 i skali 2. U ovim primjerima korišćeni su koeficijenti Gaborove transformacije aero-slike date u prvom redu na Slici 4.1. Iz uslovnih histograma se vidi da vrijednosti realnih i imaginarnih dijelova koeficijenata na različitim skalama, odnosno, orijentacijama nisu nezavisni. Ovo se može zaključiti, na primjer, iz činjenice da varijanse realnog i imaginarnog dijela koeficijenta w_{21} rastu sa porastom apsolutnih vrijednosti realnog i imaginarnog dijela koeficijenta w_{31} . Ovakva zavisnost se može uočiti kako za koeficijente na različitim skalama, Slika 4.2, tako i za koeficijente na različitim orijentacijama, Slika 4.3.

Modeli statističkih zavisnosti između podopsežnih koeficijenata korišćeni su npr. u uklanjanju šuma [64], sintezi teksture [65] i ocjeni vizuelnog kvaliteta slike [66].

4.2 Prošireni Gaborov deskriptor teksture

Mnoge ćelije vizuelnog sistema, kao na primjer ćelije mrežnjače i lateralnog genikulatnog jezgra, imaju receptivna polja podijeljena u dva dijela [40], kao što je prikazano na Slici 4.4a. Kod jedne grupe ćelija, kada je stimulisan (na primjer osvjetljavanjem) centralni kružni dio receptivnog polja, ćelija će biti pobuđena i odziv će biti većeg intenziteta, dok stimulacija rubnog dijela inhibira ćeliju i odziv slabi. Ponašanje druge grupe ćelija je obrnuto, intenzitet odziva je veći prilikom stimulacije



Slika 4.2: Dvodimenzionalni uslovni histogrami realnih i imaginarnih dijelova Gaborovih koeficijenata na skalama 2 i 3 i orijentaciji 1.



Slika 4.3: Dvodimenzionalni uslovni histogrami realnih i imaginarnih dijelova Gaborovih koeficijenata na skali 2 i orijentacijama 1 i 2.

rubnog dijela, a slabi prilikom stimulacije centralnog dijela. Ovakva organizacija receptivnih polja naziva se *centar-rub (center-surround) organizacija*. Na ovaj način ćelija pokazuje prostorni antagonizam.

U [67] je pokazano da, za realne Gaborove filtre, korelacije između koeficijenata na različitim skalama odgovaraju centar-rub organizaciji receptivnih polja u ćelijama vizuelnog sistema sisara. Ovaj rezultat je moguće poopštiti na kompleksne Gaborove filtre kakvi se koriste u ovom radu. Neka su $w_{mn}(x, y), (x, y) \in \Phi$ podopsežni koeficijenti Gaborove transformacije, odnosno, odzivi Gaborovih filtara na skalama $m = 1, \ldots, S$ i orijentacijama $n = 1, \ldots, K$, izračunati pomoću (3.7). Energija podopsežnih koeficijenata je

$$\mu_{mn}^{2} = \iint_{\Phi} |w_{mn}(x,y)|^{2} dx dy.$$
(4.1)

Prostorni antagonizam se može modelirati razlikom normalizovanih odziva Gaborovih filtara na različitim skalama m i m' i orijentaciji n

$$\Delta w_{mm'n}(x,y) = \frac{w_{mn}(x,y)}{\mu_{mn}} - \frac{w_{m'n}(x,y)}{\mu_{m'n}},$$
(4.2)



(a) Centar-rub organizacija receptivnog polja.

(b) Profil receptivnog polja.

Slika 4.4: Centar-rub organizacija receptivnog polja.

Energija razlike odziva je

$$\psi_{mm'n}^2 = \iint_{\Phi} \left| \Delta w_{mm'n} \right|^2 dx dy.$$
(4.3)

Uvrštavanjem (4.2) u (4.3) dobijamo

$$\psi_{mm'n}^{2} = \iint_{\Phi} \left| \frac{w_{mn}(x,y)}{\mu_{mn}} - \frac{w_{m'n}(x,y)}{\mu_{m'n}} \right|^{2} dx dy =$$

$$= \iint_{\Phi} \left(\frac{w_{mn}(x,y)}{\mu_{mn}} - \frac{w_{m'n}(x,y)}{\mu_{m'n}} \right) \left(\frac{w_{mn}(x,y)}{\mu_{mn}} - \frac{w_{m'n}(x,y)}{\mu_{m'n}} \right)^{*} dx dy =$$

$$= \iint_{\Phi} \left(\frac{|w_{mn}(x,y)|^{2}}{\mu_{mn}^{2}} + \frac{|w_{m'n}(x,y)|^{2}}{\mu_{m'n}^{2}} - \frac{w_{mn}(x,y)w_{m'n}^{*}(x,y)}{\mu_{mn}\mu_{m'n}} - \frac{w_{m'n}(x,y)w_{mn}^{*}(x,y)}{\mu_{mn}\mu_{m'n}} \right) dx dy.$$

$$(4.4)$$

Imajući u vidu (4.1) slijedi da je

$$\psi_{mm'n}^2 = 2 - \iint_{\Phi} \left(\frac{w_{mn}\left(x,y\right)w_{m'n}^*\left(x,y\right)}{\mu_{mn}\mu_{m'n}} + \frac{w_{m'n}\left(x,y\right)w_{mn}^*\left(x,y\right)}{\mu_{mn}\mu_{m'n}} \right) dxdy.$$
(4.5)

Gaborovi filtri su propusnici opsega pa je

$$\overline{w}_{mn} = \iint_{\Phi} w_{mn}(x, y) \, dx dy = 0. \tag{4.6}$$

Zbog toga je μ_{mn} jednako standardnoj devijaciji podopsežnih koeficijenata dobijenoj prostornim usrednjavanjem

$$\mu_{mn} = \sqrt{\iint_{\Phi} |w_{mn}(x,y) - \overline{w}_{mn}|^2 dx dy}.$$
(4.7)

Sada je očigledno da u (4.5) sabirci u zagradi predstavljaju koeficijente korelacije podopsežnih koeficijenata w_{mn} i $w_{m'n}$

$$r_{mm'n} = \iint_{\Phi} \frac{[w_{mn}(x,y) - \overline{w}_{mn}] [w_{m'n}(x,y) - \overline{w}_{m'n}]^{*}}{\mu_{mn}\mu_{m'n}} dxdy = = \iint_{\Phi} \frac{w_{mn}(x,y) w_{m'n}^{*}(x,y)}{\mu_{mn}\mu_{m'n}} dxdy = = \iint_{\Phi} \frac{w_{m'n}(x,y) w_{mn}^{*}(x,y)}{\mu_{mn}\mu_{m'n}} dxdy.$$
(4.8)

Druga jednakost u (4.8) slijedi iz činjenice da su kroskorelacije između realnih i imaginarnih dijelova odziva Gaborovih filtara, kako u istom tako i u različitim podopsezima, jednake nuli (vidjeti Dodatak A). Konačno (4.3) postaje

$$\psi_{mm'n}^2 = 2 - 2r_{mm'n}.\tag{4.9}$$

Upoređivanjem jednačina (3.11) i (4.1) vidimo da se one razlikuju samo u nelinearnoj funkciji koja se primjenjuje na module podopsežnih koeficijenata. Moguće je ovu nelinearnost poopštiti tako da definišemo

$$\mu_{mn} = \left[\iint_{\Phi} |w_{mn}(x,y)|^p dxdy\right]^{\frac{1}{p}},\qquad(4.10)$$

pri čemu je $p \in \mathbb{R}$. Pošto je u [46] pokazano da, u slučaju Gaborovog deskriptora teksture, uključivanje standardnih devijacija modula koeficijenata poboljšava performanse klasifikatora u deskriptor uključujemo i

$$\sigma_{mn} = \sqrt{\iint_{\Phi} \left(|w_{mn}(x,y)|^p - \mu_{mn}^p \right)^2 dx dy}.$$
 (4.11)

Deskriptor koji se sastoji od ove dvije komponente je uopštenje deskriptora predloženog u [46] pa se zadržava i njegov originalni naziv – *Gaborov deskriptor teksture* (*Gabor Texture Descriptor, GTD*). Dimenzionalnost ovog deskriptora je $N_1 = 2SK$.

Neka su $w_{mn}(x, y)$ i $w_{m'n}(x, y)$ odzivi Gaborovih filtara na skalama m i m' i orijentaciji n, određeni jednačinom (3.7). Kako bismo u deskriptor uključili mjeru korelacije između različitih podopsega formiramo razlike normalizovanih podopsežnih koeficijenata na različitim skalama $m \neq m'$ i istoj orijentaciji n date jednačinom (4.2). Za ove razlike računamo statistike na isti način kao i za module koeficijenata

$$\psi_{mm'n} = \left[\iint_{\Phi} |\Delta w_{mm'n}(x,y)|^p \, dx dy\right]^{\frac{1}{p}},\tag{4.12}$$

$$\upsilon_{mm'n} = \sqrt{\iint_{\Phi} \left(\left| \Delta w_{mm'n} \left(x, y \right) \right|^p - \psi^p_{mm'n} \right)^2 dx dy}, \tag{4.13}$$

Deskriptor koji se sastoji od ove dvije komponente naziva se deskriptor razlika skala (Scale Difference Descriptor, SDD) i njegova dimenzionalnost je $N_2 = 2K \binom{S}{2}$. Energija razlika normalizovanih koeficijenata na različitim skalama (4.3) dobija se iz (4.12) za p = 2. Kada je $p \neq 2$ relacija (4.9) više ne važi ali se i dalje (4.12) može smatrati mjerom kroskorelacije koeficijenata na različitim skalama.

Analogno, mogu se formirati razlike normalizovanih podopsežnih koeficijenata na različitim orijentacijama $n \neq n'$ i istoj skali m

$$\Delta w_{mnn'}(x,y) = \frac{w_{mn}(x,y)}{\mu_{mn}} - \frac{w_{mn'}(x,y)}{\mu_{mn'}}$$
(4.14)

i izračunati njihove statistike

$$\rho_{mnn'} = \left[\iint_{\Phi} \left|\Delta w_{mnn'}\left(x,y\right)\right|^p dxdy\right]^{\frac{1}{p}},\tag{4.15}$$

$$\nu_{mnn'} = \sqrt{\iint_{\Phi} \left(|\Delta w_{mnn'}(x,y)|^p - \rho_{mnn'}^p \right)^2 dx dy}.$$
 (4.16)

Deskriptor koji se sastoji od ove dvije komponente naziva se deskriptor razlika orijentacija (Orientation Difference Descriptor, ODD) i njegova dimenzionalnost je $N_3 = 2S\binom{K}{2}$. Energija razlika normalizovanih koeficijenata na različitim orijentacijama dobija se iz (4.15) za p = 2. U tom slučaju je moguće povezati (4.15) i kroskorelaciju podopsežnih koeficijenata na skali m i orijentacijama n i n', kao što je u (4.9) pokazano za kroskorelacije podopsežnih koeficijenata na različitim skalama. Kada je $p \neq 2$ ovakve relacije više ne važe, ali se i dalje može smatrati da je (4.15) mjera kroskorelacije podopsežnih koeficijenta na različitim orijentacijama.

Konačno, prošireni Gaborov deskriptor teksture, (Extended Gabor Texture Descriptor, EGTD) se dobija objedinjavanjem veličina datih jednačinama (4.10)–(4.13), (4.15), te (4.16) u $N = N_1 + N_2 + N_3 = SK(S+K)$ -dimenzionalan vektor $\mathbf{f} = [f_1, f_2, \ldots, f_N]^T$.

Za izračunavanje udaljenosti između dva EGTD, $\mathbf{f}^{(a)}$ i $\mathbf{f}^{(b)}$, koristimo metriku zasnovanu na težinskoj L_1 normi

dist
$$\left(\mathbf{f}^{(a)}, \mathbf{f}^{(b)}\right) = \sum_{j=1}^{N} \left| \frac{f_j^{(a)} - f_j^{(b)}}{\alpha(f_j)} \right|,$$
 (4.17)

gdje su sa $\alpha(f_j)$ označene standardne devijacije odgovarajućih elemenata deskriptora izračunate na trening skupu.

4.3 Redukcija dimenzionalnosti deskriptora

Prošireni Gaborov deskriptor teksture može da rezultuje visokodimenzionalnim deskriptorima slike. Na primjer, u slučaju kada se koristi Gaborova filtar banka sa S = 4 skale i K = 6 orijentacija dimenzionalnost deskriptora je N = SK(S + K) =240. Pored potrebe za većim memorijskim prostorom, visokodimenzionalni deskriptori mogu usporiti obučavanje klasifikatora i dovesti do slabe generalizacije. Zbog toga je korisno razmotriti mogućnosti redukcije dimenzionalnosti deskriptora. U nastavku se razmatraju tri načina za redukciju dimenzionalnosti deskriptora: formiranje deskriptora korišćenjem samo koeficijenata iz susjednih podopsega – skala i orijentacija, analiza glavnih komponenata (principal component analysis, PCA) i linearna diskriminantna analiza (linear discriminant analysis, LDA).

4.3.1 Korišćenje koeficijenata iz susjednih podopsega

Na Slikama 4.5 i 4.6 prikazani su uslovni histogrami realnih i imaginarnih dijelova koeficijenata Gaborove transformacije na orijentaciji 1 i skalama 2 i 4, te na skali 2 i orijentacijama 1 i 4, respektivno, određeni za aero-sliku prikazanu u prvom redu na Slici 4.1. U poređenju sa Slikama 4.2 i 4.3, vidimo da su statističke zavisnosti između koeficijenata Gaborove transformacije izraženije na susjednim nego na udaljenim skalama, odnosno, orijentacijama.

Zbog toga se ideja da se u deskriptor teksture uključe samo statistike razlika podopsežnih koeficijenata na susjednim skalama, odnosno, orijentacijama čini intuitivno prihvatljivom. Konkretno, deskriptor razlika skala, dat jednačinama (4.23) i (4.24), izračunava se samo za slučaj kada je |m - m'| = 1, a deskriptor razlika orijentacija, dat jednačinama (4.26) i (4.27), za slučaj kada je |n - n'| = 1. Na taj način se, kada Gaborova filtar banka ima 4 skale i 6 orijentacija, dobija 124-dimenzionalni deskriptor, što je gotovo dva puta niža dimenzionalnost nego u slučaju originalnog deskriptora.



Slika 4.5: Dvodimenzionalni uslovni histogrami realnih i imaginarnih dijelova Gaborovih koeficijenata na skalama 2 i 4 i orijentaciji 1.

4.3.2 Analiza glavnih komponenata

Analiza glavnih komponenata je standardna tehnika za redukciju dimenzionalnosti vektora obilježja u oblasti klasifikacije uzoraka [68]. Osnovna ideja ovog pristupa je da se pronađe linearna transformacija prostora obilježja u prostor niže dimenzionalnosti kojom će se dobiti najbolja reprezentacija vektora obilježja u smislu najmanjih kvadrata. Ovo se postiže tako što se novi vektori obilježja formiraju na takav način da se zadrži maksimalan procenat energije polaznih vektora, odnosno,



Slika 4.6: Dvodimenzionalni uslovni histogrami realnih i imaginarnih dijelova Gaborovih koeficijenata na skali 2 i orijentacijama 1 i 4.

da bazu novog prostora obilježja čine sopstveni vektori kovarijansne matrice polaznih vektora koji odgovaraju najvećim sopstvenim vrijednostima. Na taj način se u novoj reprezentaciji zadržavaju one dimenzije za koje je varijabilnost ulaznih podataka najveća. Kada se koristi PCA, dimenzionalnost rezultujućih deskriptora je ulazni parametar samog algoritma.

U ovoj disertaciji testirana su dva pristupa za redukciju dimenzionalnosti proširenog Gaborovog deskriptora teksture pomoću PCA. Prvi je primjena PCA na kompletan 240-dimenzionalni EGTD, a drugi je primjena PCA na pojedine komponente EGTD, odnosno, na GTD, SDD i ODD pojedinačno, a zatim njihovo ponovno spajanje u jedinstven deskriptor.

4.3.3 Linearna diskriminantna analiza

Nedostatak PCA kao tehnike za redukciju dimenzionalnosti je neuzimanje u obzir informacije o pripadnosti vektora obilježja ciljnim klasama. Naime, PCA pronalazi smjerove u prostoru obilježja u kojima je varijabilnost ulaznih podataka najveća, ali ne garantuje da će nove dimenzije omogućiti bolju klasifikaciju ulaznih vektora obilježja u predefinisani skup klasa. PCA po definiciji i ne može to da postigne zbog toga što se baza novog prostora obilježja izračunava samo na osnovu ulaznih vektora obilježja, zanemarujući njihovu pripadnost klasama. Diskriminantna analiza (DA) [68] predstavlja način da se riješi ovaj problem i da se pronađe novi prostor obilježja u kojem će biti obezbijeđena dobra diskriminabilnost vektora obilježja. U ovoj disertaciji testirana je redukcija dimenzionalnosti pomoću linearne diskriminantne analize (LDA), kao i regularizovane diskriminantne analize (RDA) [69]. Kada se koristi DA dimenzionalnost rezultujućeg deskriptora je za jedan manja od broja klasa u koje se uzorci klasifikuju.

Kao i u slučaju PCA, razmatrana su dva pristupa za redukciju dimenzionalnosti

Oznaka	Deskriptor
GTD	Gaborov deskriptor teksture, (4.10) i (4.11)
SDD	Deskriptor razlika skala, (4.12) i (4.13)
ODD	Deskriptor razlika orijentacija, (4.15) i (4.16)
GTD+SDD	Kombinacija GTD i SDD
GTD+ODD	Kombinacija GTD i ODD
EGTD	Prošireni Gaborov deskriptor teksture (GTD+SDD+ODD)

Tabela 4.1: Oznake deskriptora korišćene u eksperimentalnim rezultatima.

EGTD. Prvi je integralni pristup, kada se DA primjenjuje na kompletan EGTD, a drugi je primjena DA na pojedine komponente EGTD i spajanje dobijenih deskriptora u novi deskriptor niže dimenzionalnosti.

4.4 Eksperimentalni rezultati

Performanse klasifikatora u kojem se za reprezentaciju slika koristi EGTD testirane su na UC Merced Land Use i Wuhan kolekcijama slika, prema protokolima testiranja opisanim u poglavlju 2. Deskriptori slika se izračunavaju filtriranjem Gaborovom filtar bankom sa S = 4 skale i K = 6 orijentacija. Za klasifikaciju se koristi SVM sa kernelom koji predstavlja varijantu kernela sa radijalnim baznim funkcijama (radial basis function, RBF) ali, umjesto Mahalanobisove metrike, koristi metriku (4.17)

$$K\left(\mathbf{f}^{(a)}, \mathbf{f}^{(b)}\right) = e^{-\operatorname{dist}\left(\mathbf{f}^{(a)}, \mathbf{f}^{(b)}\right)}.$$
(4.18)

Klasifikacija u više klasa postignuta je treniranjem binarnih SVM oblika jedanprotiv-svih i klasifikovanjem testnog uzorka u klasu kojoj odgovara najviša vrijednost odziva SVM.

Kao što je opisano, EGTD se sastoji iz više komponenata. Kako bi se ocijenio uticaj uključivanja pojedinih komponenata u deskriptor slike, u eksperimentima su za reprezentaciju slika korišćene pojedine komponente deskriptora kao i njihove kombinacije, čije su oznake sumarizovane u Tabeli 4.1. Dobijeni rezultati su dati u nastavku.

4.4.1 UC Merced Land Use kolekcija slika

Prvo je testiran deskriptor koji se dobije kada se u jednačinama (4.10)-(4.13), (4.15) i (4.16) usvoji vrijednost p = 2. Ovaj deskriptor je proširenje originalnog deskriptora koji uzima u obzir prostorni antagonizam odziva ćelija [67] i sadrži samo veličine (4.10) i (4.12). Protokol testiranja opisan je u odjeljku 2.3, a rezultati su prikazani u Tabeli 4.2. Upoređeni su rezultati dobijeni korišćenjem deskriptora koji sadrže samo srednje vrijednosti koeficijenata i razlika koeficijenata sa rezultatima dobijenim uključivanjem standardnih devijacija koeficijenata i njihovih razlika.

Iz rezultata prikazanih u Tabeli 4.2 može se vidjeti da uključivanje standardnih devijacija u deskriptore dovodi do značajnog poboljšanja performansi klasifikatora. Dakle, kao i za originalni Gaborov deskriptor teksture [46], i za razlike (4.2) i (4.14)

Deskriptor	srednje vrijednosti	lnosti srednje vrijednosti i	
		standardne devijacije	
GTD	$61,71 \pm 2,40$	$71, 19 \pm 1, 46$	
SDD	$56,67 \pm 1,51$	$69, 29 \pm 2, 36$	
ODD	$71, 38 \pm 1, 70$	$75,85 \pm 1,78$	
GTD+SDD	$69,05 \pm 0,82$	$77,90 \pm 1,46$	
GTD+ODD	$75, 38 \pm 2, 34$	$78,86 \pm 1,78$	
EGTD	$79, 10 \pm 1, 69$	$81, 48 \pm 1, 66$	

Tabela 4.2: Tačnosti klasifikacije (%) postignute na UC Merced Land Use kolekciji slika za p = 2 i različite kombinacije komponenata EGTD deskriptora.

uključivanje bogatije statistike koeficijenata u deskriptor poboljšava rezultate klasifikacije. Ovo poboljšanje je veoma izraženo u slučajevima pojedinačnih deskriptora i iznosi od 4% do gotovo 13%, dok je za kombinacije deskriptora nešto manje, ali ipak značajno jer za kompletan EGTD deskriptor iznosi oko 2,5%. Zbog toga će u nastavku biti prezentovani samo rezultati dobijeni korišćenjem deskriptora koji uključuju i srednje vrijednosti i standardne devijacije kako nelinearno transformisanih koeficijenata tako i njihovih normalizovanih razlika.

Kada se međusobno uporede različite komponente proširenog Gaborovog deskriptora može se uočiti da se najslabije performanse dobijaju korišćenjem samo deskriptora razlika skala. Slijede Gaborov deskriptor teksture i deskriptor razlika orijentacija. Vidimo da se proširivanjem Gaborovog deskriptora teksture dodavanjem deskriptora razlika skala, odnosno, orijentacija, performanse klasifikatora poboljšavaju. Razlog za ovo je što se dodavanjem informacije o korelaciji između koeficijenata u različitim podopsezima u stvari dobija informacija o strukturama na slici koje su izražene na određenom skupu skala, odnosno, orijentacija. Na ovaj način se lakše mogu napraviti razlike između slika koje pripadaju različitim klasama.

Zanimljivo je primjetiti da uključivanje razlika orijentacija u deskriptor donosi veće poboljšanje tačnosti klasifikacije od uključenja razlika skala. Ovo suštinski znači da strukture izražene na različitim orijentacijama omogućavaju bolju diskriminaciju različitih klasa pokrivača i načina korišćenja zemljišta od struktura izraženih na različitim skalama. Mogući razlog za ovo je to što sve slike imaju istu rezoluciju te nema značajnih razlika u skali između struktura prikazanih na njima.

Interesantno je ispitati uticaj parametra p, koji figuriše u svim veličinama koje čine EGTD, na tačnost klasifikacije. Na Slici 4.7 prikazane su tačnosti klasifikacije u funkciji vrijednosti p za sve razmatrane kombinacije deskriptora. Vidimo da, do određene granice, tačnost klasifikacije raste sa smanjenjem vrijednosti p. Brzina rasta je najveća do vrijednosti p = 1 i za slučaj EGTD deskriptora opažamo poboljšanje tačnosti od preko 4% pri prelasku sa p = 2 na p = 1. Za vrijednosti p < 1 za sve deskriptore primijećujemo rast tačnosti do vrijednosti p = 0, 5, nakon čega za GTD, SDD i GTD+SDD tačnost počinje da opada. Međutim, u slučajevima ODD, GTD+ODD i EGTD deskriptora rast tačnosti se nastavlja do vrijednosti p = 0, 25i dodatno iznosi nešto više od 2%. Bez obzira na vrijednost parametra p relativni odnos performansi dobijenih različitim deskriptorima ostaje isti.

Na Slici 4.8 prikazana je matrica konfuzija dobijena korišćenjem EGTD sa p=



Slika 4.7: Zavisnost tačnosti klasifikacije UC Merced Land Use kolekcije slika od vrijednosti parametra p.



Slika 4.8: Matrica konfuzija za UC Merced Land Use kolekciju slika dobijena za EGTD sa p = 0, 25.



Slika 4.9: Zavisnost tačnosti klasifikacije UC Merced Land Use kolekcije slika od dimenzionalnosti deskriptora za PCA redukciju dimenzionalnosti primjenjenu na kompletan EGTD.

0,25. U prosjeku, najčešće konfuzije su između gusto i rezidencijalnog zemljišta srednje gustine, cisterni i terena za bejzbol, zgrada i raskrsnica, te rijeka i rijetko naseljenog rezidencijalnog zemljišta.

Redukcija dimenzionalnosti deskriptora

Kao što je opisano u odjeljku 4.3, u ovoj disertaciji se, između ostalog, ispituje mogućnost redukcije dimenzionalnosti proširenog Gaborovog deskriptora teksture pomoću analize glavnih komponenata, diskriminantne analize i formiranjem deskriptora korišćenjem samo koeficijenata iz susjednih podopsega.

Na Slici 4.9 prikazana je zavisnost tačnosti klasifikacije od dimenzionalnosti deskriptora kada se PCA primjenjuje na kompletan EGTD. Vidljivo je da se, za razliku od deskriptora bez redukcije dimenzionalnosti, najbolji rezultati dobijaju za deskriptor kod kojeg je p = 0, 5. Tačnost klasifikacije, kao što je i očekivano, najprije raste sa povećanjem dimenzionalnosti deskriptora do 50 dimenzija kada iznosti 80,48%. Međutim, za više dimenzionalnosti tačnost klasifikacije blago opada zbog toga što je u dimenzijama koje odgovaraju malim sopstvenim vrijednostima odnos signal-šum slab i ove dimenzije nisu pouzdane za klasifikaciju [70]. U najboljem slučaju, dakle, PCA redukcijom dimenzionalnosti kompletnog EGTD dobijaju se oko 7% lošije performanse u odnosu na deskriptor pune dimenzionalnosti.

Na Slici 4.10 prikazana je zavisnost tačnosti klasifikacije od ukupne dimenzionalnosti deskriptora kada se PCA primjenjuje na pojedine komponente EGTD. Ukupna dimenzionalnost deskriptora se dobija kao zbir dimenzionalnosti pojedinih komponenata. Može se primjetiti da se povećanjem dimenzionalnosti deskriptora poboljšava



Slika 4.10: Zavisnost tačnosti klasifikacije UC Merced Land Use kolekcije slika od ukupne dimenzionalnosti deskriptora za PCA redukciju dimenzionalnosti primjenjenu na pojedine komponente EGTD.

tačnost klasifikacije do ukupne dimenzionalnosti 60 kada iznosi 85,43%, a zatim ponovo blago opada. U ovom slučaju dimenzionalnost deskriptora je smanjena 4 puta, a tačnost klasifikacije za samo 2%.

Redukcijom dimenzionalnosti po komponentama se dobijaju konzistentno bolji rezultati u odnosu na slučaj kada se PCA primjenjuje na kompletan EGTD. Ovo je posljedica činjenice da je EGTD formiran konkatenacijom tri komponente: GTD, SDD i ODD. Korelacije između elemenata unutar jedne komponente su izraženije nego korelacije između elemenata koji pripadaju različitim komponentama. Kada se PCA globalno primjenjuje sve komponente se posmatraju zajedno tako da u niskodimenzionalnom deskriptoru može dominirati informacija koja potiče od samo jedne komponente EGTD. Iako je na takav način očuvana maksimalna energija deskriptora u novom prostoru odbačene su potencijalno diskriminativne informacije iz ostalih komponenata EGTD što rezultuje lošijim performansama klasifikatora. Tako je, za 60-dimenzionalan deskriptor, PCA primjenjena po komponentama bolja za 5% od PCA primjenjene na kompletan deskriptor.

Kada se dimenzionalnost deskriptora redukuje korišćenjem samo koeficijenata iz susjednih podopsega kao rezultat se dobija deskriptor fiksne dimenzionalnosti – 124 dimenzije. Slično, kada se koristi diskriminantna analiza, dimenzionalnost rezultujućeg deskriptora je za jedan manja od broja klasa. U slučaju UC Merced kolekcije slika, ovo znači da se, kada se DA primjenjuje na kompletan EGTD, dobija deskriptor dimenzionalnosti 20, a kada se primjenjuje na pojedine komponente EGTD, deskriptor $3 \times 20 = 60$ -dimenzionalan.

Sa druge strane, dimenzionalnost deskriptora dobijenih pomoću PCA je ulazni

Tabela 4.3: Tačnosti klasifikacije (%) postignute na UC Merced Land Use kolekciji slika za različite metode redukcije dimenzionalnosti i različite dimenzionalnosti deskriptora. EGTD – prošireni Gaborov deskriptor teksture, Susj. podops. – EGTD izračunat korišćenjem samo koeficijenata iz susjednih podopsega, PCA – analiza glavnih komponenata, LDA – linearna diskriminatna analiza, RDA – regularizovana linearna diskriminatna analiza, Komp. – redukcija dimenzionalnosti primjenjena po komponentama EGTD.

Metod	Dimenzionalnost			
	20	60	124	240
EGTD	_	—	—	$87,57 \pm 1,68$
Susj. podops.	_	_	$85, 95 \pm 0, 65$	_
PCA	$78,29\pm1,34$	$80, 38 \pm 1, 72$	$78,71 \pm 2,51$	$77, 76 \pm 2, 13$
Komp. PCA	$82,62\pm0,78$	$85, 43 \pm 0, 97$	$83, 43 \pm 0, 40$	$82,24 \pm 1,66$
LDA	$76, 19 \pm 2, 45$	_	_	_
RDA	$79,43 \pm 1,58$	_	_	_
Komp. LDA	_	$81,57\pm0,62$	_	_
Komp. RDA	_	$83,95\pm0,99$	_	_

parametar ovog algoritma za redukciju dimenzionalnosti i na ovaj način je moguće dobiti deskriptore proizvoljne dimenzionalnosti manje ili jednake dimenzionalnosti originalnih deskriptora. Pošto poređenje performansi klasifikatora obučenih korišćenjem deskriptora različitih dimenzionalnosti nije u potpunosti korektno, u Tabeli 4.3 su dati rezultati poređenja za deskriptore iste dimenzionalnosti dobijene različitim metodama za redukciju dimenzionalnosti. Polazeći od toga da neki metodi daju deskriptore fiksne dimenzionalnosti, prikazani su rezultati za deskriptore dimenzionalnosti 20, 60, 124, kao i za puni 240-dimenzionalni deskriptor. Prazna polja u tabeli odgovaraju dimenzionalnostima deskriptora koje nisu praktično ostvarljive.

U koloni koja odgovara 124-dimenzionalnim deskriptorima dato je poređenje deskriptora dobijenog korišćenjem samo koeficijenata iz susjednih podopsega sa deskriptorima dobijenim korišćenjem PCA primjenjene na kompletan EGTD kao i na pojedine komponente. Dati su najbolji rezultati dobijeni testiranjem za različite vrijednosti parametra p. Interesantno je da se u slučaju korišćenja samo susjednih podopsega najbolji rezultat dobija za vrijednost p = 0,25, dok se za deskriptore dobijene pomoću PCA najbolji rezultat dobija za vrijednost p = 0,5. U slučaju PCA primjenjene po komponentama dat je rezultat dobijen za 120-dimenzionalni deskriptor kod kojeg su sve komponente EGTD jednako procentualno zastupljene.

Iz prikazanih rezultata se vidi da se korišćenjem samo susjednih podopsega pri izračunavanju deskriptora performanse klasifikatora smanjuju za manje od 2% dok je dimenzionalnost deskriptora oko dva puta niža. Primjenom PCA na pojedine komponente EGTD deskriptora se takođe dobija dobar rezultat. Međutim, prednost korišćenja samo susjednih podopsega je manja složenost s obzirom na to da ne postoji eksplicitan korak redukcije dimenzionalnosti deskriptora, a smanjena je i složenost izračunavanja samog deskriptora pošto se ne računaju statistike razlika koeficijenata iz svih parova podopsega.

U koloni koja odgovara 60-dimenzionalnim deskriptorima date su tačnosti klasifi-

kacije za PCA, PCA primjenjenu na komponente EGTD, te LDA i RDA primjenjene na komponente EGTD. U svim slučajevima prikazani su najbolji rezultati koji su dobijeni za p = 0, 5. Ponovo se vidi da se primjenom PCA po komponentama EGTD dobijaju bolji rezultati nego njenom primjenom na kompletan deskriptor. U ovom slučaju najbolje performanse se dobijaju pomoću PCA primjenjene na EGTD po komponentama. Regularizovana diskriminatna analiza je bolja od linearne diskriminantne analize, ali ipak nešto lošija od PCA. Važno je primjetiti da je rezultat dobijen pomoću PCA primjenjene po komponentama u ovom slučaju za 2% bolji od rezultata dobijenog pomoću PCA za ukupnu dimenzionalnost 124, a od tačnosti klasifikacije dobijene korišćenjem susjednih podopsega se razlikuje za manje od jedne standardne devijacije.

Kada je dimenzionalnost deskriptora jednaka 20 prikazane su tačnosti klasifikacije za PCA primjenjenu na kompletan EGTD, odnosno njegove komponente, te LDA i RDA primjenjene na kompletan EGTD. U svim slučajevima prikazani su najbolji rezultati koji su dobijeni za p = 0, 5.

Kada se porede tehnike za redukciju dimenzionalnosti primjenjene na kompetan EGTD najbolje rezultate daje regularizovana diskriminantna analiza. Međutim, PCA primjenjena po komponentama i u ovom slučaju daje tačnost klasifikacije bolju za 3% od RDA primjenjene na kompletan deskriptor. Još je značajnije što je ovaj rezultat oko 3% lošiji od tačnosti klasifikacije dobijene 124-dimenzionalnim deskriptorom, a svega oko 5% lošiji od tačnosti klasifikacije dobijene deskriptorom pune dimenzionalnosti.

Diskriminatna analiza nastoji da pronađe reprezentaciju deskriptora koja omogućava najbolju diskriminaciju uzoraka. Iz Tabele 4.3 se vidi da se primjenom RDA dobijaju bolji rezultati nego primjenom PCA za 20-dimenzionalne deskriptore. Međutim, vidi se i da su, kada se diskriminantna analiza primjenjuje po komponentama, rezultati lošiji nego primjenom PCA. Razlog za ovo je činjenica da raspoloživi trening primjeri nisu reprezentativni za čitav prostor obilježja, kao što je zaključeno i u [71].

Konačno, za punu dimenzionalnost deskriptora (240-D) tačnosti klasifikacije dobijene kada se koristi PCA su znatno lošije od tačnosti dobijene originalnim deskriptorom. Razlog za ovo je, kao što je već pomenuto, uključivanje u deskriptor nepouzdanih dimenzija koje odgovaraju malim sopstvenim vrijednostima kovarijansne matrice.

4.4.2 Wuhan kolekcija slika

Na Slici 4.11 prikazana je zavisnost tačnosti klasifikacije Wuhan kolekcije slika od vrijednosti parametra p za EGTD i njegove pojedine komponente. Vidimo da je ova zavisnost slična kao u slučaju UC Merced kolekcije slika. I u ovom slučaju se maksimumi tačnosti klasifikacije postižu za vrijednosti p = 0,25 i p = 0,5. Za SDD, ODD i GTD+ODD maksimum tačnosti se postiže za p = 0,5, a za GTD, GTD+SDD i EGTD za p = 0,25. Najbolje performanse se, bez obzira na vrijednost p, i u ovom slučaju dobijaju korišćenjem EGTD, a najlošije korišćenjem samo SDD. Zanimljivo je primjetiti da, osim za vrijednost p = 2, u ovom slučaju ODD daje bolje performanse od kombinacije GTD+SDD.



Slika 4.11: Zavisnost tačnosti klasifikacije postignute na Wuhan kolekciji slika od vrijednosti parametra p.

Matrica konfuzija dobijena korišćenjem EGTD sa p = 0,25 i obučavanjem klasifikatora na 50% slika data je na Slici 4.12. U prosjeku, najčešće konfuzije se javljaju između klasa: rezidencijalno zemljište i poslovne zgrade, luka i jezero, parking i luka, kao i park i rijeka. Ove konfuzije su posljedica izražene vizuelne sličnosti slika iz pomenutih kategorija.

Redukcija dimenzionalnosti deskriptora

U eksperimentima sa redukcijom dimenzionalnosti metodologija testiranja je ista kao za klasifikator sa deskriptorima pune dimenzionalnosti. Kao i za UC Merced skup slika, testirani su: deskriptor koji koristi samo koeficijente iz susjednih podopsega, te PCA i LDA primjenjene na kompletan EGTD kao i na njegove komponente.

Pošto PCA omogućava da se varira dimenzionalnost rezultujućeg deskriptora najprije je ispitan njen uticaj na tačnost klasifikacije. Na Slici 4.13 prikazana je zavisnost tačnosti klasifikacije od dimenzionalnosti deskriptora dobijenog primjenom PCA na kompletan EGTD. Različite krive su dobijene za različite vrijednosti parametra p. Globalno, ove krive su vrlo slične odgovarajućim krivama dobijenim za UC Merced skup slika, prikazanim na Slici 4.9. Najveća tačnost klasifikacije se dobija kada je p = 0,5 i deskriptor je 50-dimenzionalan, a iznosi 80,12%. Ovaj rezultat je oko 3% lošiji u odnosu na tačnost postignutu deskriptorom pune dimenzionalnosti. Sa slike se takođe vidi da se, sa porastom dimenzionalnosti, u deskriptor dodaju dimenzije u kojima je odnos signal-šum loš što rezultuje lošijim performansama klasifikatora.

Na Slici 4.14 prikazana je zavisnost tačnosti klasifikacije od ukupne dimenzionalnosti deskriptora dobijenog primjenom PCA na pojedine komponente EGTD i



Slika 4.12: Matrica konfuzija za Wuhan kolekciju slika dobijena za EGTD sa p=0,25.



Slika 4.13: Zavisnost tačnosti klasifikacije Wuhan kolekcije slika od dimenzionalnosti deskriptora za PCA redukciju dimenzionalnosti primjenjenu na kompletan EGTD.



Slika 4.14: Zavisnost tačnosti klasifikacije Wuhan kolekcije slika u zavisnosti od dimenzionalnosti deskriptora dobijenog primjenog PCA na pojedine komponente EGTD.

njihovom konkatenacijom u jedinstven deskriptor. I u ovom slučaju zavisnost je slična kao za UC Merced skup slika. Najbolji rezultati se ponovo dobijaju kada je p = 0, 5, bez obzira na dimenzionalnost deskriptora. Ukupno, najbolja tačnost klasifikacije od 83,47% se postiže korišćenjem 50-dimenzionalnog deskriptora i približno je jednaka tačnosti klasifikacije dobijenoj pomoću EGTD pune dimenzionalnosti. Sa porastom dimenzionalnosti performanse klasifikatora se pogoršavaju, mada u nešto blažoj mjeri nego u slučaju kada se PCA primjenjuje na kompletan EGTD.

Slike u Wuhan kolekciji su klasifikovane u 19 klasa pokrivača i načina korišćenja zemljišta. Iz toga slijedi da će se diskriminantnom analizom dobiti 18-dimenzionalni deskriptori, kada se primjeni na kompletan EGTD, a $3 \times 18 = 54$ -dimenzionalni deskriptori kada se DA primjeni na EGTD po komponentama.

U Tabeli 4.4 uporedno su date tačnosti klasifikacije za sve analizirane metode redukcije dimenzionalnosti. Prikazani su rezultati dobijeni za deskriptore čija je dimenzionalnost redukovana na 18, 54 i 124 dimenzije, kao i za deskriptore pune dimenzionalnosti 240.

Deskriptori dimenzionalnosti 124 dobijaju se korišćenjem koeficijenata iz susjednih podopsega, te primjenom PCA, kako na kompletan EGTD, tako i na pojedine njegove komponente. U slučaju korišćenja samo susjednih podopsega dat je rezultat dobijen za vrijednost p = 0, 25, dok je za redukciju dimenzionalnosti pomoću PCA dat rezultat dobijen za vrijednost p = 0, 5. U slučaju PCA primjenjene po komponentama dat je rezultat dobijen za 120-dimenzionalni deskriptor.

Vidi se da se, za razliku od UC Merced skupa slika, najbolji rezultat u ovom slučaju dobija korišćenjem PCA primjenjene na pojedine komponente EGTD. Ovaj

Tabela 4.4: Tačnosti klasifikacije (%) postignute na Wuhan kolekciji slika za različite metode redukcije dimenzionalnosti i različite dimenzionalnosti deskriptora. EGTD – prošireni Gaborov deskriptor teksture, Susj. podops. – EGTD izračunat korišćenjem samo koeficijenata iz susjednih podopsega, PCA – analiza glavnih komponenata, LDA – linearna diskriminatna analiza, RDA – regularizovana linearna diskriminatna analiza, Komp. – redukcija dimenzionalnosti primjenjena po komponentama EGTD.

Motod	Dimenzionalnost			
Metod	18	54	124	240
EGTD	—	—	_	$83,74\pm0,95$
Susj. podops.	—	_	$79,72 \pm 1,34$	_
PCA	$79,66\pm1,37$	$79,61\pm1,81$	$78, 10 \pm 1, 33$	$62, 47 \pm 1, 89$
Komp. PCA	$80, 58 \pm 1, 07$	$83, 10 \pm 0, 99$	$82, 43 \pm 1, 40$	$77,25 \pm 1,83$
LDA	$64, 88 \pm 2, 38$	_	_	_
RDA	$78,77\pm1,68$	_	_	_
Komp. LDA	—	$79,73 \pm 1,30$	_	_
Komp. RDA	_	$82, 29 \pm 1, 48$	_	_

rezultat se razlikuje od tačnosti dobijene korišćenjem deskriptora pune dimenzionalnosti za manje od jedne standardne devijacije tako da se u praksi može smatrati da su njihove performanse jednake.

Deskriptori dimenzionalnosti 54 dobijaju se redukcijom dimenzionalnosti pomoću PCA, PCA primjenjene na komponente EGTD, te LDA i RDA primjenjene na komponente EGTD. U svim slučajevima prikazani su najbolji rezultati koji su dobijeni za p = 0, 5.

Najbolji rezultati za ovu dimenzionalnost deskriptora se ponovo dobijaju primjenom PCA na pojedine komponente EGTD. Rezultat dobijen regularizovanom diskriminantnom analizom je unutar jedne standardne devijacije u odnosu na rezultat dobijen pomoću PCA primjenjene po komponentama. Iako diskriminantna analiza nastoji da pronađe reprezentaciju koja omogućava najbolju diskriminaciju uzoraka, rezultati dobijeni njenim korišćenjem su, ipak, nešto lošiji u odnosu na rezultate dobijene primjenom PCA. Kao i u slučaju UC Merced Land Use kolekcije, ovo se dešava zbog toga što raspoloživi trening primjeri nisu dovoljno reprezentativni za čitav prostor obilježja. Zanimljivo je da su oba rezultata unutar jedne standardne devijacije u odnosu na najbolji rezultat za 124-dimenzionalni deskriptor, kao i u odnosu na tačnost klasifikacije dobijenu za deskriptor pune dimenzionalnosti.

Deskriptori dimenzionalnosti 18 dobijaju se pomoću PCA, PCA primjenjene na komponente EGTD, te LDA i RDA primjenjene na kompletan EGTD. U svim slučajevima prikazani su najbolji rezultati koji su dobijeni za p = 0, 5.

Kao i za više dimenzionalnosti deskriptora, najbolji rezultat se i ovdje dobija primjenom PCA po komponentama EGTD. Razlika u performansama između deskriptora dobijenog primjenom PCA na kompletan EGTD i na pojedine komponenten EGTD je u ovom slučaju znatno manja nego u slučajevima deskriptora više dimenzionalnosti. Iako je dimenzionalnost deskriptora u ovom slučaju značajno smanjena u odnosu na dimenzionalnost originalnog deskriptora, tačnost klasifikacije je niža za samo oko 3%. Konačno, korišćenjem originalnog 240-dimenzionalnog deskriptora dobija se značajno bolja tačnost klasifikacije nego korišćenjem deskriptora iste dimenzionalnosti ali nakon primjene PCA. Razlog za ovo je nepouzdanost estimacije malih sopstvenih vrijednosti što rezultuje slabim odnosom signal-šum u odgovarajućim dimenzijama novog prostora obilježja [70].

4.4.3 Diskusija

Na osnovu rezultata dobijenih za obje kolekcije slika vidimo da se redukcijom dimenzionalnosti deskriptora na vrijednost između 50 i 60, primjenom PCA na pojedine komponente EGTD, tačnost klasifikacije zanemarljivo smanjuje (maksimalno 2%) u odnosu na korišćenje originalnog 240-dimenzionalnog deskriptora. Međutim, dimenzionalnost deskriptora se smanjuje bar četiri puta. Ukoliko se može tolerisati pad tačnosti od 4-5% moguće je koristiti 124-dimenzionalne deskriptore izračunate samo korišćenjem koeficijenata iz susjednih podopsega, čime se smanjuje složenost izračunavanja deskriptora ili 20-dimenzionalne deskriptore dobijene analizom glavnih komponenata, čime se ostvaruje još veća redukcija dimenzionalnosti.

4.5 Prošireni Gaborov deskriptor teksture za slike u boji

Iako se često analizira odvojeno od boje, tekstura, u opštem slučaju, nastaje kombinovanjem elemenata (tekstona) različite boje i informacija o boji može olakšati njenu analizu. U [54] je pokazano da korišćenje deskriptora teksture u boji formiranih konkatenacijom deskriptora izračunatih za pojedinačne kolor-kanale poboljšava tačnost klasifikacije tekstura.

Uključivanje informacije o boji u prošireni Gaborov deskriptor teksture inspirisano je teorijom oponentnih procesa u ljudskom vidu [40]. U odjeljku 4.2 već je pomenut prostorni antagonizam koji pokazuju fotoreceptori u centralnim i rubnim dijelovima receptivnih polja ćelija mrežnjače i lateralnog genikulatnog jezgra. Tom prilikom nisu postavljena nikakva ograničenja u pogledu spektralnog sastava (hromatskih svojstava) svjetlosti na koju su fotoreceptori osjetljivi. Međutim, fotoreceptori mogu biti osjetljivi na svjetlost različitih talasnih dužina te se, pored prostornog, javlja i hromatski antagonizam. U zavisnosti od spektralne osjetljivosti ćelija razlikuju se tri tipa receptivnih polja prikazana na Slici 4.15 [72].

Neka su A i B dvije klase fotoreceptora, npr. fotoreceptori koji odgovaraju crvenom i zelenom dijelu spektra. Receptivna polja tipa I u centralnom dijelu pobuđuju fotoreceptori klase A, a u rubnom dijelu ih inhibiraju fotoreceptori klase B, kao što je prikazano na Slici 4.15a. Ćelije koje pokazuju ovakvo ponašanje nazivaju se *jednostruko oponentne ćelije* Na Slici 4.15b prikazana je organizacija i profil odziva receptivnog polja tipa II, odnosno, *dvostruko oponentne ćelije*. U ovom slučaju receptivno polje u centralnom dijelu pobuđuju fotoreceptori klase A, a inhibiraju fotoreceptori klase B, dok je u rubnom dijelu situacija obrnuta. Ćelije, razmatrane u odjeljku 4.2, koje pokazuju samo prostorni, a ne i spektralni antagonizam imaju receptivna polja tipa III, Slika 4.15c.



Slika 4.15: Tipovi receptivnih polja ćelija mrežnjače.



Slika 4.16: Moduli odziva Gaborovog filtra za različite spektralne opsege.

Prvi korak u izračunavanju proširenog Gaborovog deskriptora teksture je filtriranje slike u boji bankom Gaborovih filtara. U obradi slika u boji, generalno, postoje dva pristupa: marginalni i vektorski [73]. Marginalni pristup se zasniva na filtriranju svake spektralne komponente slike posebno, dok se kod vektorskog pristupa vrijednosti piksela u različitim spektralnim komponentama posmatraju kao cjelina, odnosno, vektor u skupu \mathbb{R}^c , gdje je c broj spektralnih komponenata. Dobra osobina marginalnog pristupa je njegova jednostavnost, a loše strane su mogućnost pojave kolor artifakata i neuzimanje u obzir združenog efekta različitih spektralnih opsega. Vektorski pristupi za obradu multispektralnih slika su rezultat nastojanja da se prevaziđu ovi problemi. Cijena koja se pri tome plaća je praktično uvijek povećanje složenosti algoritama.

Deskriptor koji se dobija kada se za filtiranje slike u boji koristi marginalni pristup nazvaćemo oponentni prošireni Gaborov deskriptor teksture, zbog njegove sličnosti sa oponentnim kolor deskriptorom predloženim u [67]. Vektorski pristup filtriranju slika u boji realizovaćemo pomoću kvaternionske reprezentacije slika u boji [74] i kvaternionskog Gaborovog filtra [75]. Rezultujući deskriptor ćemo zvati kvaternionski prošireni Gaborov deskriptor teksture.

4.5.1 Oponentni prošireni Gaborov deskriptor teksture

Razmotrimo najprije marginalni pristup filtriranju RGB slika Gaborovom filtar bankom. Neka je data RGB slika $I(x, y), (x, y) \in \Phi$, gdje je Φ skup piksela slike i neka je $I_i(x, y)$ njen *i*-ti spektralni opseg, $i = 1, \ldots, B$. RGB slika se kod ovog pristupa filtrira Gaborovim filtrom tako što se svaki njen opseg filtrira zasebno. Sada je izlaz Gaborovog filtra sa impulsnim odzivom, $g_{mn}(x, y)$, na skali $m = 1, \ldots, S$ i orijentaciji $n = 1, \ldots, K$, za *i*-ti spektralni opseg dat konvolucijom

$$w_{imn}(x,y) = I_i(x,y) * g_{mn}(x,y).$$
(4.19)



Slika 4.17: Dvodimenzionalni uslovni histogrami realnih i imaginarnih dijelova koeficijenata Gaborove transformacije za crveni i zeleni spektralni opseg.

Ovo je, u stvari, ista jednačina kao (3.7) sa izuzetkom oznake indeksa spektralnog opsega.

Na Slici 4.16 prikazani su moduli odziva Gaborovog filtra na skali 2 i orijentaciji 1 za crveni, zeleni i plavi kolor spektralni opseg jedne slike. Vidi se da moduli koeficijenata odziva na istim prostornim lokacijama za različite spektralne opsege imaju slične vrijednosti što ukazuje na njihovu korelisanost.

Postojanje korelacija između koeficijenata Gaborove transformacije u različitim spektralnim opsezima se može potvrditi pomoću dvodimenzionalnih uslovnih histograma realnih i imaginarnih dijelova koeficijenata Gaborove transformacije na skali 2 i orijentaciji 1 za crveni i zeleni spektralni opseg, prikazanih na Slici 4.17. Vidljivo je da srednja vrijednost koeficijenata dobijenih za zeleni spektralni opseg izraženo zavisi od vrijednosti koeficijenata dobijenih za crveni spektralni opseg što ukazuje na korelisanost koeficijenata. Slični histogrami se dobijaju i za druge kombinacije spektralnih opsega pa se može zaključiti da postoje izražene korelacije između podopsežnih koeficijenata za različite spektralne opsege.

Oponentni prošireni Gaborov deskriptor teksture predstavlja nadogradnju proširenog Gaborovog deskriptora teksture, definisanog za panhromatske slike, uključivanjem informacije o boji, odnosno, o oponentnosti boja. Analogno jednačinama (4.10) i (4.11), za svaki spektralni opseg, i = 1, ..., B, definišu se

$$\mu_{imn} = \left[\iint_{\Phi} |w_{imn}\left(x,y\right)|^p dxdy \right]^{\frac{1}{p}}, \qquad (4.20)$$

$$\sigma_{imn} = \sqrt{\iint_{\Phi} \left(|w_{imn}\left(x,y\right)|^p - \mu_{imn}^p\right)^2 dx dy}.$$
(4.21)

Ukupan broj elemenata deskriptora određenih ovim jednačinama je $N_1 = 2BSK$.

Kako bi se hromatski antagonizam koji pokazuju receptivna polja ćelija u ljudskom vizuelnom sistemu, a time i korelacije koeficijenata Gaborove transformacije izračunatih za različite spektralne opsege, uključile u prošireni Gaborov deskriptor teksture, pored razlika normalizovanih odziva Gaborovih filtara na različitim skalama (4.2) i orijentacijama (4.14), formiramo i razlike odziva filtara za različite spektralne opsege.

Neka su dati odzivi Gaborovih filtara za spektralne opsege i i j, skale m i m' te orijentaciju n, određeni jednačinom (4.19). Razlike normalizovanih odziva filtara na različitim skalama su date sa

$$\Delta w_{ijmm'n}(x,y) = \frac{w_{imn}(x,y)}{\mu_{imn}} - \frac{w_{jm'n}(x,y)}{\mu_{jm'n}}.$$
(4.22)

Analogno jednačinama (4.12) i (4.13) u deskriptor uključujemo veličine

$$\psi_{ijmm'n} = \left[\iint_{\Phi} \left|\Delta w_{ijmm'n}\left(x,y\right)\right|^p dxdy\right]^{\frac{1}{p}},\tag{4.23}$$

$$\upsilon_{ijmm'n} = \sqrt{\iint_{\Phi} \left(\left| \Delta w_{ijmm'n} \left(x, y \right) \right|^p - \psi^p_{ijmm'n} \right)^2 dx dy}.$$
(4.24)

Ukoliko je $i \neq j$ i $m \neq m'$ ovim veličinama su modelirani spektralni i prostorni antagonizam koji pokazuju ćelije tipa I na Slici 4.15. Nadalje, ukoliko je i = j i $m \neq m'$ modeliran je samo prostorni antagonizam koji pokazuju ćelije tipa III. U ovom slučaju veličine date jednačinama (4.23) i (4.24) su identične veličinama (4.12) i (4.13) koje se dobijaju za panhromatske slike. Broj elemenata deskriptora koji se dobija za ova dva slučaja je $N_2 = 2KB^2 {S \choose 2}$. Ukoliko je $i \neq j$ i m = m' modeliran je samo spektralni antagonizam koji pokazuju dvostruko oponentne ćelije tipa II. Broj elemenata deskriptora u ovom slučaju je $N_3 = 2KS {B \choose 2}$.

Neka su, dalje, dati odzivi Gaborovih filtara za spektralne opsege i i j, skalu m i orijentacije n i n'. Razlike normalizovanih odziva filtara na različitim orijentacijama date su sa

$$\Delta w_{ijmnn'}(x,y) = \frac{w_{imn}(x,y)}{\mu_{imn}} - \frac{w_{jmn'}(x,y)}{\mu_{jmn'}},$$
(4.25)

Analogno jednačinama (4.15) i (4.16) u oponentni prošireni Gaborov deskriptor teksture uključujemo veličine

$$\rho_{ijmnn'} = \left[\iint_{\Phi} |\Delta w_{ijmnn'}(x,y)|^p \, dx dy\right]^{\frac{1}{p}},\tag{4.26}$$

$$\nu_{ijmnn'} = \sqrt{\iint_{\Phi} \left(|\Delta w_{ijmnn'}(x,y)|^p - \rho_{ijmnn'}^p \right)^2 dx dy}.$$
(4.27)

Slično kao za veličine (4.23) i (4.24) i ovdje razlikujemo tri slučaja. Ukoliko je $i \neq j$ i $n \neq n'$ deskriptor pokazuje i spektralni i antagonizam orijentacija. U slučaju kada je i = j, a $n \neq n'$ deskriptor pokazuje samo antagonizam orijentacija. Ukupan broj elemenata deskriptora za ova dva slučaja je $N_4 = 2SB^2 \binom{K}{2}$. U slučaju kada je $i \neq j$ i n = n' dobijaju se iste vrijednosti kao jednačinama (4.23) i (4.24) za $i \neq j$ i m = m', tako da te vrijednosti neće ponovo biti uključene u deskriptor. Dakle, ukupna dimenzionalnost oponentnog poboljšanog Gaborovog deskriptora teksture je $N = N_1 + N_2 + N_3 + N_4$. Na ovaj način se, u stvari, u deskriptor uključuju koeficijenti korelacije između Gaborovih koeficijenata za različite spektralne opsege. Za pojedine komponente proširenog Gaborovog deskriptora teksture zadržavaju se isti nazivi, s tom razlikom što se u njihovo izračunavanje uključuju različiti spektralni opsezi. Tako se u Gaborov deskriptor teksture, pored srednjih vrijednosti (4.20) i standardnih devijacija (4.21) modula koeficijenata, uključuju i srednje vrijednosti i standardne devijacije modula razlika koeficijenata na istoj skali i orijentaciji, ali iz različitih spektralnih opsega. Ove veličine se dobijaju iz (4.23), (4.24), (4.26) i (4.27) za $i \neq j, m = m'$ i n = n'. Nadalje, deskriptor razlika skala se izračunava pomoću (4.23) i (4.24), a deskriptor razlika orijentacija pomoću (4.26) i (4.27) za $m \neq m'$, odnosno, $n \neq n'$, respektivno. Za poređenje deskriptora se koristi metrika (4.17).

4.5.2 Kvaternionski prošireni Gaborov deskriptor teksture

Alternativa marginalnom pristupu filtriranju slika u boji, u kojem se svaki spektralni opseg filtrira posebno, je korišćenje vektorskog pristupa u kojem se vrijednosti piksela u različitim spektralnim opsezima posmatraju kao elementi jednog vektora [73]. Operacije filtriranja slike se, u ovom slučaju, definišu nad vektorima čime se svi spektralni opsezi istovremeno obrađuju uz uzimanje u obzir veza i interakcija između njih.

Jedna takva, vektorska, reprezentacija RGB slika je kvaternionska reprezentacija koju je uveo Sangwine u [74]. U kvaternionskoj reprezentaciji, RGB slika se može prikazati kao čisti kvaternion

$$I(x,y) = r(x,y)i + g(x,y)j + b(x,y)k,$$
(4.28)

gdje su i, j, k jedinični kvaternioni i vrijedi $i^2 = j^2 = k^2 = -1$ i ijk = -1. Za sliku I(x, y) i dva proizvoljna čista jedinična kvaterniona μ_1 i μ_2 , takva da je $\mu_1 \perp \mu_2$, je moguće odrediti simplektičku dekompoziciju [76]

$$I(x,y) = I^{(1)}(x,y) + I^{(2)}(x,y) \mu_2, \qquad (4.29)$$

pri čemu je

$$I^{(1)}(x,y) = q_1(x,y)\,\mu_1,\tag{4.30}$$

$$I^{(2)}(x,y) = q_2(x,y) + q_3(x,y)\,\mu_1.$$
(4.31)

Kvaternionske slike $I^{(1)}(x, y)$ i $I^{(2)}(x, y)$ se nazivaju *simpleksni* i *perpleksni* dio slike I(x, y), respektivno, i izomorfne su kompleksnim slikama, tj. slikama čiji pikseli imaju vrijednosti iz skupa kompleksnih brojeva.

Simplektička dekompozicija slike je ekvivalentna promjeni baze vektorskog dijela kvaterniona u (4.28) iz (i, j, k) u (μ_1, μ_2, μ_3) , pri čemu vrijedi $\mu_3 = \mu_1 \mu_2$. Slika je tada jednaka

$$I(x,y) = q_1(x,y)\,\mu_1 + q_2(x,y)\,\mu_2 + q_3(x,y)\,\mu_3,\tag{4.32}$$

gdje su $q_1(x, y)$, $q_2(x, y)$ i $q_3(x, y)$ komponente kvaterniona u novoj bazi. Pored pomenutih uslova ortonormalnosti, baza (μ_1, μ_2, μ_3) ne mora da zadovoljava druge posebne uslove.

Kako bi se dobila prikladna reprezentacija slike koja će uzeti u obzir spektralni anatagonizam između kolor kanala, koristi se baza data sa

$$\mu_1 = \frac{i+j+k}{\sqrt{3}},\tag{4.33}$$

$$\mu_2 = \frac{i-j}{\sqrt{2}},\tag{4.34}$$

$$\mu_3 = \frac{i+j-2k}{\sqrt{6}},\tag{4.35}$$

pa μ_1 odgovara sivoj liniji u RGB kolor-prostoru. Prema tome, simpleksni dio $I^{(1)}(x, y)$ u (4.29) odgovara luminantnoj, a perpleksni dio $I^{(2)}(x, y)$ odgovara hrominantnoj komponenti slike, kao što je prikazano na Slici 4.18. Ova reprezentacija je povezana sa oponentnim kolor-prostorom datim sa

$$o_1 = q_2 = \frac{r - g}{\sqrt{2}} \tag{4.36}$$

$$o_2 = q_3 = \frac{r+g-2b}{\sqrt{6}} \tag{4.37}$$

$$p_3 = q_1 = \frac{r+g+b}{\sqrt{3}}.$$
(4.38)



Slika 4.18: Simplektička dekompozicija RGB slike

Filtriranje Gaborovom filtar bankom je ekvivalentno prostorno zavisnoj Furijeovoj transformaciji uz korišćenje Gausove prozorske funkcije. S obzirom na to da se generalizovana kvaternionska Furijeova transformacija može dobiti zamjenom imaginarne jedinice, u definiciji Furijeove transformacije, jediničnim kvaternionom [76], kvaternionska Gaborova funkcija se može definisati kao

$$g(x,y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2}\right) + \mu_1\omega x\right].$$
(4.39)

Impulsni odzivi filtara u kvaternionskoj Gaborovoj filtar banci se sada dobijaju skaliranjem i rotacijom kvaternionske Gaborove funkcije.

Odzivi Gaborovih filtara $w_{mn}(x, y)$, dobijeni filtriranjem slike I(x, y), su puni kvaternioni i njihova simplektička dekompozicija je

$$w_{mn}(x,y) = w_{mn}^{(1)}(x,y) + w_{mn}^{(2)}(x,y)\,\mu_2.$$
(4.40)

Kada se koristi baza povezana sa oponentnim kolor-prostorom, simpleksni dio odziva, $w_{mn}^{(1)}(x, y)$, odgovara odzivu Gaborovog filtra na luminantnu komponentu, a perpleksni dio, $w_{mn}^{(2)}(x, y)$ odgovara odzivu na hrominantnu komponentu. Pošto su ova dva dijela odziva izomorfni kompleksnim signalima koji se dobiju kada se panhromatska slika filtrira Gaborovim filtrom (3.7), moguće je za svaki od njih odrediti veličine koje čine prošireni Gaborov deskriptor teksture za panhromatske slike:

• Kvaternionski Gaborov deskriptor teksture

$$\mu_{mn}^{(i)} = \left[\iint_{\Phi} \left| w_{mn}^{(i)} \left(x, y \right) \right|^p dx dy \right]^{\frac{1}{p}}, \tag{4.41}$$

$$\sigma_{mn}^{(i)} = \sqrt{\iint_{\Phi} \left[\left| w_{mn}^{(i)}(x,y) \right|^p - \left(\mu_{mn}^{(i)} \right)^p \right]^2 dx dy}.$$
 (4.42)

• Kvaternionski deskriptor razlika skala

$$\Delta w_{mnn'}^{(i)}(x,y) = \frac{w_{mn}^{(i)}(x,y)}{\mu_{mn}} - \frac{w_{mn'}^{(i)}(x,y)}{\mu_{mn'}},\tag{4.43}$$

$$\psi_{mm'n}^{(i)} = \left[\iint_{\Phi} \left| \Delta w_{mm'n}^{(i)} \left(x, y \right) \right|^p dx dy \right]^{\frac{1}{p}}, \qquad (4.44)$$

$$v_{mm'n}^{(i)} = \sqrt{\iint_{\Phi} \left[\left| \Delta w_{mm'n}^{(i)}(x,y) \right|^p - \left(\psi_{mm'n}^{(i)} \right)^p \right]^2 dx dy}.$$
 (4.45)

• Kvaternionski deskriptor razlika orijentacija

$$\Delta w_{mnn'}^{(i)}(x,y) = \frac{w_{mn}^{(i)}(x,y)}{\mu_{mn}} - \frac{w_{mn'}^{(i)}(x,y)}{\mu_{mn'}}, \qquad (4.46)$$

$$\rho_{mnn'}^{(i)} = \left[\iint_{\Phi} \left| \Delta w_{mnn'}^{(i)} \left(x, y \right) \right|^p dx dy \right]^{\frac{1}{p}}, \qquad (4.47)$$

$$\nu_{mnn'}^{(i)} = \sqrt{\iint_{\Phi} \left[\left| \Delta w_{mnn'}^{(i)} \left(x, y \right) \right|^p - \left(\rho_{mnn'}^{(i)} \right)^p \right]^2 dx dy}.$$
 (4.48)

U gornjim jednačinama $i \in \{1, 2\}$ označava simpleksni i perpleksni dio kvaterniona, a μ_{mn} je ukupna energija odziva filtra na skali m i orijentaciji n

$$\mu_{mn} = \iint_{\Phi} |w_{mn}(x,y)| \, dxdy. \tag{4.49}$$

Ove veličine se računaju za sve skale i orijentacije uz uslove $m \neq m'$ i $n \neq n'$ za deskriptore razlika skala i orijentacija, respektivno. Dakle, ovaj deskriptor je 2SK(S+K)-dimenzionalni vektor. Za računanje mjere nesličnosti između slika usvajamo metriku (4.17).

4.6 Eksperimentalni rezultati

Oponentni i kvaternionski prošireni Gaborov deskriptor teksture testirani su na UC Merced Land Use i Wuhan kolekcijama slika, prema protokolima testiranja opisanim u poglavlju 2. U oba slučaja klasifikator je SVM sa kernelom (4.18).

Za izračunavanje oponentnog proširenog Gaborovog deskriptora teksture se koristi Gaborova filtar banka sa S = 4 skale i K = 6 orijentacija. Pored kompletnog deskriptora, testira se i kako njegove pojedine komponente, opisane u odjeljku 4.5.1, utiču na tačnost klasifikacije. Pošto je za panhromatske slike pokazano da tačnost klasifikacije zavisi i od parametra p koji figuriše u jednačinama oponentnog deskriptora razmatra se i njegov uticaj na tačnost klasifikacije.

Oponentni prošireni Gaborov deskriptor može biti izuzetno visokodimenzionalan. Na primjer, u razmatranom slučaju, dimenzionalnost deskriptora je N = 2016. Zbog toga će, pored deskriptora pune dimenzionalnosti, biti testirani i deskriptori redukovane dimenzionalnosti. Polazeći od rezultata dobijenih u odjeljku 4.4.1 razmatraće se deskriptor koji se dobija kada u jednačinama (4.22)-(4.27) figurišu samo iste ili susjedne skale, odnosno, orijentacije, tj. kada je $|m - m'| \leq 1$ i $|n - n'| \leq 1$, kao i redukcija dimenzionalnosti pomoću PCA primjenjene na pojedine komponente proširenog Gaborovog deskriptora teksture. PCA primjenjena na kompletan deskriptor i diskriminatna analiza ovdje neće biti razmatrane zato što su u Odjeljku 4.4.1 rezultovale konzistentno nižim tačnostima klasifikacije u odnosu na druge dvije tehnike.

Za izračunavanje kvaternionskog proširenog Gaborovog deskriptora teksture se koristi kvaternionska Gaborova filtar banka sa S = 4 skale i K = 6 orijentacija. Kao i kod oponentnog deskriptora, testiraju se doprinos pojedinih komponenata deskriptora i uticaj parametra p na tačnost klasifikacije. Dimenzionalnost punog kvaternionskog proširenog Gaborovog deskriptora teksture iznosi 480. Iako je ova dimenzionalnost znatno niža od dimenzionalnosti punog oponentnog proširenog Gaborovog deskriptora teksture i dalje se radi o visokoj dimenzionalnosti pa će i u ovom slučaju biti testirane tehnike redukcije dimenzionalnosti deskriptora primjenom PCA na pojedine komponente deskriptora te korišćenjem samo susjednih podopsega Gaborove transformacije, tj. $|m - m'| \leq 1$ u (4.43), odnosno, $|n - n'| \leq 1$ u (4.46).

4.6.1 UC Merced Land Use kolekcija slika

Na Slici 4.19 date su tačnosti klasifikacije za pojedine komponente oponentnog proširenog Gaborovog deskriptora teksture, kao i za njihove kombinacije, u zavisnosti od parametra p. U svim slučajevima, osim za deskriptor razlika skala, najveća tačnost klasifikacije se dobija, kao i za panhromatske slike, za p = 0, 25. Za deskriptor razlika skala tačnost dobijena za p = 0, 5 je neznatno veća. Primjetno je da je Gaborov deskriptor teksture sada komponenta za koju se postiže veća tačnost klasifikacije nego korišćenjem deskriptora razlika skala i deskriptora razlika orijentacija. Ovaj rezultat se razlikuje od rezultata dobijenog za panhromatske slike. Razlog je uključivanje informacije o korelacijama između različitih spektralnih opsega u ovaj deskriptor čime se poboljšava njegova diskriminativna sposobnost.

Očekivano, korišćenjem kombinacija deskriptora tačnost klasifikatora se pobolj-



Slika 4.19: Tačnost klasifikacije UC Merced Land Use kolekcije slika u zavisnosti od parametra p za različite komponente oponentnog proširenog Gaborovog deskriptora teksture.

šava i najveća tačnost od 88,86% se dobija kombinacijom sva tri deskriptora, tj. korišćenjem punog oponentnog proširenog Gaborovog deskriptora teksture. Međutim, u ovom slučaju poboljšanje u odnosu na kombinaciju Gaborovog deskriptora teksture i deskriptora razlika orijentacija je vrlo malo, a za slučaj p = 1 je čak oponentni prošireni Gaborov deskriptor teksture lošiji od kombinacije Gaborovog deskriptora teksture i deskriptora razlika orijentacija. Ovaj pad performansi se može objasniti lošom generalizacijom uzrokovanom veoma visokom dimenzionalnošću deskriptora.

Na Slici 4.20 data je zavisnost tačnosti klasifikacije od ukupne dimenzionalnosti oponentnog deskriptora nakon primjene PCA na njegove pojedine komponente. Deskriptor redukovane dimenzionalnosti se formira tako što se zadržava isti procenat dimenzionalnosti svake komponente. Izuzetno je zanimljivo primjetiti da tačnost klasifikacije maksimalnu vrijednost od 87,71% poprima za relativno niskodimenzionalan deskriptor (252-D), a da sa porastom dimenzionalnosti izraženo opada. Ovo se dešava zbog toga što se, sa porastom dimenzionalnosti, u deskriptor uključuju nepouzdane dimenzije koje odgovaraju malim sopstvenim vrijednostima kovarijansne matrice deskriptora [70]. Zbog toga je generalizacija loša i tačnost klasifikacije opada. Takođe se može primjetiti da se najbolje performanse u ovom slučaju dobijaju za p = 0, 5, slično kao i za panhromatske slike.

Kod oponentnog proširenog Gaborovog deskriptora, interakcije između spektralnih opsega se uključuju kroz eksplicitno izračunavanje razlika Gaborovih koeficijenata određenih za pojedine spektralne opsege. Zbog velikog broja kombinacija spektralnih opsega, skala i orijentacija za koje se koeficijenti izračunavaju, rezultujući deskriptor ima veoma visoku dimenzionalnost. Sa druge strane, kada se koristi


Slika 4.20: Tačnost klasifikacije UC Merced Land Use kolekcije slika za oponentni prošireni Gaborov deskriptor teksture u zavisnosti od ukupne dimenzionalnosti deskriptora.

kvaternionski pristup za obradu slika u boji interakcije između pojedinih spektralnih opsega su implicitno uključene u definicije operatora za obradu slike. Ovo omogućava dobijanje deskriptora niže dimenzionalnosti.

Na Slici 4.21 date su zavisnosti tačnosti klasifikacije od parametra p za pojedine komponente kvaternionskog deskriptora i njihove kombinacije. Međusobni odnos pojedinih komponenata je isti kao i za panhromatske slike – najbolje performanse se dobijaju korišćenjem deskriptora razlika orijentacija, a najlošije za deskriptor razlika skala. U ovom slučaju, međutim, deskriptor razlika orijentacija pokazuje bolje performanse od kombinacije Gaborovog deskriptora teksture i deskriptora razlika skala. Najveća tačnost klasifikacije od 88,05% se dobija korišćenjem punog kvaternionskog proširenog Gaborovog deskriptora. Ipak, u poređenju sa kombinacijom Gaborovog deskriptora teksture i deskriptora razlika orijentacija, poboljšanje tačnosti je manje od jedne standardne devijacije. Najbolji rezultati se uglavnom dobijaju za p = 0,25, osim za deskriptor razlika skala (p = 1) i njegovu kombinaciju sa Gaborovim deskriptorom teksture (p = 0, 5).

Na Slici 4.22 prikazana je tačnost klasifikacije u zavisnosti od ukupne dimenzionalnosti kvaternionskog deskriptora za različite vrijednosti parametra p, nakon primjene PCA na njegove pojedine komponente. Maksimalna tačnost klasifikacije se dobija za p = 0,5 i 100-dimenzionalni deskriptor, a iznosi 86,19%. Sa povećanjem dimenzionalnosti tačnost klasifikacije opada, iako nešto blaže u odnosu na oponentni deskriptor, zato što je ukupna dimenzionalnost u slučaju kvaternionskih deskriptora znatno niža od dimenzionalnosti punog oponentnog proširenog Gaborovog deskriptora teksture.



Slika 4.21: Tačnost klasifikacije UC Merced Land Use kolekcije slika u zavisnosti od parametra p za različite komponente kvaternionskog proširenog Gaborovog deskriptora teksture.



Slika 4.22: Tačnost klasifikacije UC Merced Land Use kolekcije slika za kvaternionski prošireni Gaborov deskriptor teksture u zavisnosti od ukupne dimenzionalnosti deskriptora.

Metod	Dim.	Tačnost
EGTD panhrom.	240	$87,57 \pm 1,68$
EGTD op.	2016	$88,86 \pm 0,74$
EGTD kvater.	480	$88,05 \pm 1,07$
EGTD op. susj. podop.	972	$89,00 \pm 1,03$
EGTD kvater. susj. podop.	248	$86,90 \pm 1,14$
EGTD op. PCA komp.	252	$87,71 \pm 0,95$
EGTD kvater. PCA komp.	100	$86, 19 \pm 0, 71$

Tabela 4.5: Tačnosti klasifikacije(%) za UC Merced Land Use kolekciju slika.

U Tabeli 4.5 date su tačnosti klasifikacije za UC Merced Land Use kolekciju slika kada se koriste različiti deskriptori izračunati za slike u boji. Pored rezultata dobijenih korišćenjem punih deskriptora i najboljih rezultata dobijenih redukcijom dimenzionalnosti pomoću PCA, navedeni su i rezultati dobijeni za prošireni Gaborov deskriptor teksture izračunat korišćenjem samo razlika koeficijenata iz susjednih podopsega. Takođe, poređenja radi, uključena je i tačnost klasifikacije koja se dobija proširenim Gaborovim deskriptorom teksture za panhromatske slike. Vidi se da se korišćenjem deskriptora izračunatog za slike u boji postiže povećanje tačnosti klasifikacije. Cijena koja se pri tome plaća je povećanje dimenzionalnosti deskriptora. Ovo povećanje je nešto niže kada se koristi kvaternionska reprezentacija slike u boji. Interesantno je primjetiti da se za oponentni deskriptor izračunat korišćenjem samo koeficijenata iz susjednih podopsega dobijaju neznatno bolje (mada u okviru jedne standardne devijacije) performanse nego za puni deskriptor. U ostalim slučajevima redukcijom dimenzionalnosti se performanse u maloj mjeri pogoršavaju i postaju približno jednake performansama dobijenim za panhromatske slike.

Na Slici 4.23 prikazana je matrica konfuzija dobijena za oponentni EGTD izračunat korišćenjem samo koeficijenata iz susjednih podopsega. Konfuzije se najčešće javljaju između klasa: zgrade i raskrsnica, gusto rezidencijalno i srednje rezidencijalno, te cisterne i raskrsnica. Uzrok ovih konfuzija je izražena vizuelna sličnost slika iz navedenih klasa.

4.6.2 Wuhan kolekcija slika

Na Slici 4.24 prikazane su tačnosti klasifikacije za Wuhan kolekciju slika u zavisnosti od parametra p za oponentni prošireni Gaborov deskriptor teksture i njegove pojedine komponente. Najbolji rezultati se postižu kada se koristi puni EGTD, a vrlo bliski su i rezultati dobijeni korišćenjem kombinacije GTD+ODD. Za razliku od UC Merced Land Use kolekcije slika, u ovom slučaju su performanse dobijene korišćenjem samo oponentnog ODD približno jednake ili čak bolje od performansi dobijenih kombinacijom GTD+SDD za vrijednosti $p \leq 1$. Pored toga, za ovu kolekciju slika GTD je bolji samo od SDD što je primjećeno i za panhromatske slike iz obje kolekcije. Najbolje performanse se ponovo dobijaju za vrijednosti p od 0,25 i 0,5.

Na Slici 4.25 prikazana je zavisnost tačnosti klasifikacije od ukupne dimenzionalnosti oponentnog deskriptora nakon primjene PCA na njegove pojedine komponente.



Slika 4.23: Matrica konfuzija za UC Merced Land Use kolekciju slika dobijena za oponentni EGTD izračunat korišćenjem samo koeficijenata iz susjednih podopsega.



Slika 4.24: Tačnost klasifikacije Wuhan kolekcije slika u zavisnosti od parametra p za različite komponente oponentnog proširenog Gaborovog deskriptora teksture.



Slika 4.25: Tačnost klasifikacije Wuhan kolekcije slika u zavisnosti od ukupne dimenzionalnosti deskriptora.

Rezultati su prikazani za različite vrijednosti parametra *p*. Pošto je ukupan broj trening slika 475, a dimenzionalnosti SDD i ODD su 648 i 1080, respektivno, onda se klasifikator testira sa maksimalnom dimenzionalnošću ova dva deskriptora od 456. Vidimo da je potrebna relativno mala dimenzionalnost deskriptora (100-D) da bi se postigla maksimalna tačnost klasifikacije.

Na Slici 4.26 prikazana je tačnost klasifikacije za Wuhan kolekciju slika u zavisnosti od vrijednosti parametra p za kvaternionski prošireni Gaborov deskriptor teksture. Uporedo su prikazani rezultati dobijeni za puni kvaternionski deskriptor kao i za njegove pojedine komponente i njihove kombinacije. U većini slučajeva najbolje performanse se dobijaju za p = 0,25, sa izuzetkom SDD i ODD za koje su rezultati dobijeni za p = 0,5 za nijansu bolji. Što se tiče uzajamnog odnosa performansi pojedinih deskriptora vidimo da on gotovo ne zavisi od vrijednosti p, izuzev u slučaju Gaborovog deskriptora teksture i SDD.

Očekivano, najbolja tačnost klasifikacije se dobija korišćenjem punog kvaternionskog proširenog Gaborovog deskriptora teksture. Međutim, kombinacija Gaborovog deskriptora i deskriptora razlika orijentacija je samo neznatno, tj. unutar jedne standardne devijacije, slabija. Zanimljivo je da je čak i sam deskriptor razlika orijentacija svega nekoliko procenata lošiji od punog deskriptora, iako znatno niže dimenzionalnosti.

Kada se za redukciju dimenzionalnosti kvaternionskog deskriptora koristi PCA primjenjena po komponentama deskriptora, sa Slike 4.27 se vidi da se najbolja tačnost klasifikacije dobija za p = 0, 5. Kao što je bio slučaj i u ostalim eksperimentima, i za kvaternionski deskriptor najbolje performanse se dobijaju korišćenjem deskriptora relativno niske dimenzionalnosti u poređenju sa originalnim deskriptorom. Tako



Slika 4.26: Tačnost klasifikacije Wuhan kolekcije slika u zavisnosti od parametra p za različite komponente kvaternionskog proširenog Gaborovog deskriptora teksture.



Slika 4.27: Tačnost klasifikacije Wuhan kolekcije slika u zavisnosti od ukupne dimenzionalnosti kvaternionskog deskriptora.

Metod	Dim.	Tačnost
EGTD panhrom.	240	$83,74 \pm 0,95$
EGTD op.	2016	$85,01 \pm 1,26$
EGTD kvater.	480	$84, 88 \pm 1, 17$
EGTD op. susj. podop.	972	$82,68 \pm 1,23$
EGTD kvater. susj. podop.	248	$83,65 \pm 0,93$
EGTD op. PCA komp.	100	$86, 27 \pm 1, 46$
EGTD kvater. PCA komp.	60	$86,02 \pm 0,98$

Tabela 4.6: Tačnosti klasifikacije (%) za Wuhan kolekciju slika.

se već za 60-dimenzionalni deskriptor dobija tačnost klasifikacije od 86,02%.

Performanse različitih varijanti proširenog Gaborovog deskriptora teksture za slike u boji, na Wuhan kolekciji slika, date su u Tabeli 4.6. Navedene su tačnosti klasifikacije za puni oponentni i kvaternionski deskriptor, za deskriptore snižene dimenzionalnosti koji su formirani samo korišćenjem koeficijenata iz susjednih podopsega, kao i za deskriptore čija je dimenzionalnost redukovana korišćenjem PCA. Poređenja radi, data je i tačnost klasifikacije koja se dobija za panhromatske slike.

Vidimo da se tačnost klasifikacije blago poboljšava izračunavanjem deskriptora za slike u boji. Međutim, to poboljšanje je praćeno značajnim povećanjem dimenzionalnosti deskriptora. Ipak, redukcijom dimenzionalnosti pomoću PCA moguće je dobiti niskodimenzionalne deskriptore koji omogućavaju čak i nešto bolje performanse od deskriptora pune dimenzionalnosti. Ova pojava je posljedica odbacivanja nepouzdanih dimenzija koje odgovaraju malim sopstvenim vrijednostima kovarijansne matrice [70].

Na Slici 4.28 prikazana je matrica konfuzija dobijena za kvaternionski EGTD čija je dimenzionalnost redukovana na 100 pomoću PCA. Konfuzije se najčešće javljaju između klasa: plaža i most, rezidencijalno zemljište i poslovne zgrade, te aerodrom i vijadukt. I u ovom slučaju, uzrok ovih konfuzija je izražena vizuelna sličnost slika iz navedenih klasa.

4.6.3 Diskusija

Iz dobijenih eksperimentalnih rezultata na obje kolekcije slika može se primjetiti da se uključivanjem informacija iz različitih spektralnih opsega tačnost klasifikacije generalno poboljšava u odnosu na slučaj kada se koriste samo panhromatske slike. Performanse dobijene oponentnim i kvaternionskim proširenim Gaborovim deskriptorom teksture su pribiližno jednake, ali kvaternionski deskriptor ima značajno nižu dimenzionalnost što ga čini primamljivijim izborom u ovom slučaju. Razlog za postizanje dobrih rezultata korišćenjem kvaternionskog deskriptora, uprkos nižoj dimenzionalnosti, je uključivanje spektralne informacije u izračunavanje deskriptora od samog početka čime su implicitno uzete u obzir interakcije između spektralnih opsega.

Uključivanje informacije o boji u deskriptor u oba slučaja rezultuje povećanjem njegove dimenzionalnosti. Ovo povećanje je manje u slučaju kvaternionskog proširenog Gaborovog deskriptora teksture, ali je rezultujući deskriptor ipak 480-



Slika 4.28: Matrica konfuzija za Wuhan kolekciju slika dobijena za kvaternionski EGTD uz redukciju dimenzionalnosti na 100-D pomoću PCA.

dimenzionalan. Pokazano je da se redukcijom dimenzionalnosti deskriptora korišćenjem analize glavnih komponenata ili formiranjem deskriptora korišćenjem samo koeficijenata iz susjednih podopsega mogu dobiti deskriptori značajno niže dimenzionalnosti uz, u najgorem slučaju, smanjenje tačnosti klasifikacije za manje od 3%. U nekim slučajevima je čak primjećeno povećanje tačnosti klasifikacije nakon redukcije dimenzionalnosti što je posljedica odbacivanja dimenzija koje sadrže informacije koje nisu dovoljno diskriminativne.

4.7 Zaključak

Prošireni Gaborov deskriptor teksture za panhromatske slike se formira dodavanjem mjere korelisanosti između podopsega Gaborove transformacije na različitim skalama, odnosno, orijentacijama. Korelacije između Gaborovih koeficijenata na različitim skalama povezane su sa centar-rub organizacijom receptivnih polja ćelija vizuelnog sistema. Eksperimentalno je pokazano da se njihovim uključivanjem u deskriptor slike poboljšava tačnost klasifikacije. Sa druge strane, korelacije između Gaborovih koeficijenata na različitim orijentacijama nisu biološki zasnovane, ali iz eksperimentalnih rezultata se vidi da se njihovim korišćenjem dobija još veće povećanje tačnosti klasifikacije.

U oba slučaja korelacije između Gaborovih koeficijenata ukazuju na postojanje struktura na slici koje su izražene na različitim skalama, odnosno, orijentacijama.

Imajući u vidu da su sve slike u jednoj kolekciji iste rezolucije, informacija o postojanju struktura izraženih na različitim orijentacijama je korisnija za klasifikaciju od informacije o postojanju struktura izraženih na različitim skalama.

Predloženi prošireni Gaborov deskriptor teksture može, u zavisnosti od broja skala i orijentacija upotrebljene filtar banke, biti visokodimenzionalan. Testirane su različite tehnike za redukciju dimenzionalnosti deskriptora i pokazano je da se konzistentno najbolji rezultati dobijaju primjenom PCA na pojedine komponente deskriptora. Na ovaj način je moguće značajno smanjiti dimenzionalnost deskriptora uz neznatno smanjenje performansi klasifikatora.

Iz Tabele 4.3 se može primjetiti da se korišćenjem samo susjednih podopsega pri formiranju deskriptora za UC Merced Land Use kolekciju slika dobija tačnost klasifikacije koja je neznatno lošija od tačnosti dobijene korišćenjem deskriptora pune dimenzionalnosti. Međutim, ovo ponašanje nije primjećeno za Wuhan kolekciju slika, Tabela 4.4. Mogući razlog su krupnije strukture na slikama iz Wuhan kolekcije. Ove strukture su izražene na većem broju skala ili orijentacija čime su i korelacije između udaljenih skala, odnosno, orijentacija izraženije.

Testirana su i dva načina za uključivanje boje u izračunavanje proširenog Gaborovog deskriptora teksture: marginalni i vektorski. Rezultujući deskriptori su nazvani oponentni i kvaternionski prošireni Gaborov deskriptor teksture, respektivno. Oba pristupa rezultuju poboljšanim performansama klasifikatora, ali po cijenu povećane dimenzionalnosti deskriptora. Ipak, kvaternionski deskriptor je nešto niže dimenzionalnosti od oponentnog, a dobijene tačnosti klasifikacije su približno jednake.

Problem vrlo visoke dimenzionalnosti deskriptora se, i u ovom slučaju, može riješiti redukcijom dimenzionalnosti primjenom PCA na pojedine komponente deskriptora. Međutim, iz Tabele 4.5 se vidi da se tako za UC Merced Land Use kolekciju slika dobijaju tačnosti klasifikacije približno jednake tačnosti dobijenoj za panhromatske slike. Sa druge strane, za Wuhan kolekciju slika, Tabela 4.6, se, na taj način, performanse klasifikatora čak poboljšavaju. Kada se za formiranje deskriptora koriste samo susjedni podopsezi, sa druge strane, poboljšanje performansi se može uočiti na slučaju UC Merced Land Use kolekcije slika, dok za Wuhan kolekciju slika performanse opadaju što je slično ponašanju primjećenom za panhromatske slike.

Glava 5 Strukturalni deskriptor teksture

U klasifikaciji visokorezolucionih snimaka dobijenih daljinskom detekcijom tekstura igra značajnu ulogu u diskriminaciji klasa pokrivača i načina korišćenja zemljišta. Jedan od najpopularnijih deskriptora teksture u analizi slika dobijenih daljinskom detekcijom je Gaborov deskriptor teksture razmatran u Odjeljku 3.2. Međutim, Gaborov deskriptor teksture se sastoji samo od srednjih vrijednosti i standardnih devijacija podopsežnih koeficijenata. Sa druge strane, izračunavanje Gaborove transformacije se može posmatrati i kao filtriranje slike bankom Gaborovih filtara na različitim skalama i orijentacijama. Moduli odziva Gaborove filtar banke, odnosno, koeficijenti Gaborove transformacije, su, kako je pokazano u Odjeljku 4.1, izraženo korelisani. Korišćenjem koeficijenata korelacije između podopsežnih koeficijenata, dobijenih pomoću banke filtara na različitim skalama i orijentacijama, moguće je formirati statistički model teksture pogodan za njenu sintezu [65]. Imajući ovo u vidu, zanimljivo je ispitati mogućnost korišćenja koeficijenata korelacije između modula koeficijenata iz različitih podopsega Gaborove transformacije kao deskriptora teksture.

5.1 Motivacija

Nedavno je predložena nova mjera sličnosti slika koja uzima u obzir osobine čovjekovog vizuelnog sistema [77, 78] i sastoji se od tri člana

$$S(\mathbf{x}, \mathbf{y}) = l(\mathbf{x}, \mathbf{y}) c(\mathbf{x}, \mathbf{y}) s(\mathbf{x}, \mathbf{y}) = = \left(\frac{2\mu_x \mu_y + K_1}{\mu_x^2 + \mu_y^2 + K_1}\right) \left(\frac{2\sigma_x \sigma_y + K_2}{\sigma_x^2 + \sigma_y^2 + K_2}\right) \left(\frac{\sigma_{xy} + K_3}{\sigma_x \sigma_y + K_3}\right),$$
(5.1)

gdje su K_1, K_2 i K_3 male konstante uvedene kako bi se poboljšala numerička stabilnost mjere sličnosti. Prva dva člana u (5.1) su mjere sličnosti između srednjih vrijednosti i standardnih devijacija vrijednosti piksela na respektivnim slikama, a treći je *strukturalni član*, zasnovan na kroskorelacijama između slika koje se porede. Smatra se da se uključivanjem ovog člana pri poređenju uzima u obzir strukturna informacija na slikama, koja je ljudskim subjektima veoma važna u ocjenjivanju sličnosti slika. Zbog toga je ova mjera sličnosti nazvana *mjera strukturalne sličnosti (Structural Similarity Measure, SSIM)*. Pošto je slika realizacija nestacionarnog



 $STSIM = \sum_{x,y} S(x,y)$

Slika 5.1: Računanje originalnih STSIM i STSIM-2 mjera sličnosti.

slučajnog procesa, pomenute statistike se ne računaju globalno, već lokalno, korišćenjem klizećih prozora koji se simultano pomjeraju po obje slike. Na kraju se dobijene vrijednosti lokalne sličnosti usrednjavaju u ukupnu mjeru sličnosti.

Originalna mjera strukturalne sličnosti neposredno koristi vrijednosti piksela i osjetljiva je na transformacije slike. U [79] je predložena *strukturalna sličnost zasnovana na kompleksnim talasićima (Complex wavelet structural similarity, CW-SSIM)* koja je robusna na male translacije, rotacije i promjene skale. Za izračunavanje mjere sličnosti slika, CW-SSIM, umjesto vrijednosti piksela, koristi vrijednosti podopsežnih koeficijenata dobijenih kompleksnom usmjerenom talasnom dekompozicijom (steerable wavelet decomposition) [80]. Važno je primjetiti da CW-SSIM koristi model obrade koji je sličan biološkom, jer se, kao što je već razmatrano, obrada u vidnoj kori velikog mozga može modelirati orijentisanim filtrima na različitim skalama.

Iako je moguće koristiti CW-SSIM kao mjeru sličnosti u klasifikaciji i pretraživanju baza slika na osnovu sličnosti, kao npr. u [81], strukturalni član je, u suštini, poređenje slika piksel-po-piksel. Međutim, čak i dvije vizuelno slične slike mogu imati značajne razlike u vrijednostima piksela. Da bi se izbjeglo poređenje slika piksel-po-piksel, u [82] je predložena *strukturalna sličnost tekstura (Structural Texture Similarity, STSIM)*. Ova mjera sličnosti je zasnovana na premisi da slične teksture mogu imati značajne razlike u vrijednostima piksela i njena osnovna namjena je poređenje tekstura pri kompresiji slike sa gubicima. U STSIM mjeri sličnosti se ne koristi strukturalni član iz (5.1) već umjesto njega figuriše mjera sličnosti autokorelacija podopsežnih koeficijenata susjednih po horizontalnoj, odnosno, vertikalnoj osi. U [83] STSIM je proširena uvođenjem kroskorelacija između podopsega na različitim skalama, odnosno, orijentacijama i novodobijena mjera sličnosti je nazvana STSIM-2. Autori su u [83] testirali usklađenost STSIM-2 sa ljudskom ocjenom sličnosti tekstura, a u [84] su je koristili za pretraživanje baze tekstura.

Da bi se izračunala strukturalna sličnost dvije slike tekstura korišćenjem STSIM-2, svaka slika se prvo filtrira filtar bankom kompleksnih usmjerenih filtara. Zatim se, kao što je ilustrovano na Slici 5.1, za svaki prozor dimenzija 7×7 piksela koji se pomjera po svakom podopsegu izračunavaju sljedeće statistike podopsežnih koeficijenata: srednja vrijednost, standardna devijacija, koeficijenti autokorelacije za pomak od jednog piksela po horizontalnoj i vertikalnoj osi, kao i koeficijenti kroskorelacije



Slika 5.2: Osjetljivost originalne STSIM mjera sličnosti na prostorni raspored vizuelnih elemenata na slici.

između podopsega. Sličnost dvije teksture se izračunava poređenjem statistika podopsežnih koeficijenata za odgovarajuće prozore, tj. prozore na istim lokacijama na slikama, i integracijom dobijenih lokalnih sličnosti.

Iako usmjereni talasići, koji se koriste u obje varijante STSIM, omogućavaju izvjesan stepen robusnosti na male geometrijske transformacije, ovaj pristup je ipak osjetljiv na prostorni raspored objekata ili regiona na slikama. Ova osjetljivost može imati kao posljedicu netačnu vrijednost sličnosti kada se ovaj pristup koristi za nehomogene teksture, kao što je obično slučaj kod aero-snimaka visoke rezolucije. Ovaj problem je ilustrovan primjerom na Slici 5.2 u kojem se sličnost slika izračunata pomoću STSIM koristi za njihovu klasifikaciju korišćenjem klasifikatora na principu najbližeg susjeda. Slike u gornjem redu su označeni primjeri¹ iz klasa *rijetko naseljeno* i *teniski tereni*. Testna slika (u donjem redu) treba biti klasifikovana kao teniski teren. Korišćenjem originalne STSIM-2 mjere sličnosti testna slika je sličnija označenom primjeru iz klase rijetko naseljeno nego iz (tačne) klase teniski tereni, što bi rezultovalo njenom netačnom klasifikacijom ako se koristi klasifikator na principu najbližeg susjeda.

Osjetljivost na promjene prostornog rasporeda regiona na slikama može biti korisna u klasifikaciji scena iz prirode gdje je prostorni raspored važan za klasifikaciju [21]. Međutim, pokazano je da apsolutni prostorni raspored, uzet u obzir pomoću prostornog piramidalnog uparivanja (spatial pyramid match kernel) [30] ili Gist deskriptora [23], ne poboljšava tačnost klasifikacije za slike dobijene daljinskom detekcijom. Shodno tome, apsolutni prostorni raspored elemenata na slici dobijenoj daljinskom detekcijom se ne može smatrati korisnim za njihovu klasifikaciju. U

¹Kada se koristi klasifikator na principu najbližeg susjeda ne postoji faza obučavanja klasifikatora pa se ne koristi ni termin trening skup. Ipak, da bi klasifikator mogao da se koristi za klasifikaciju novih primjera potrebno je da postoji skup primjera čija je pripadnost klasama poznata. Ovi primjeri se nazivaju *označeni primjeri*.



Slika 5.3: Modifikacija računanja strukturalne sličnosti tekstura.

Odjeljku 5.4 je predložena modifikacija načina izračunavanja strukturalne sličnosti tekstura kojom se ublažava nepovoljan uticaj prostornog rasporeda vizuelnih elemenata na slici. Korišćenjem ove modifikovane mjere sličnosti testna slika u primjeru sa Slike 5.2 će biti tačno klasifikovana.

5.2 Strukturalni deskriptor teksture

Nedostaci o kojima je bilo govora u prethodnom odjeljku potiču od računanja sličnosti korišćenjem klizećih prozora na istim lokacijama na slikama koje se porede. Da bi se izbjegli ovi nedostaci, a ipak iskoristila deskriptivna sposobnost strukturalne reprezentacije teksture, moguće je na osnovu nje konstruisati lokalne deskriptore. Ovi deskriptori se i dalje izračunavaju u lokalnim prozorima, ali se slike posmatraju kao neuređeni skupovi prozora i omogućeno je poređenje prozora na različitim prostornim lokacijama. Ovo je ilustrovano na Slici 5.3. Za svaki prozor sa testne slike, pronalazi se najsličniji prozor na označenoj slici, a ukupna sličnost dvije slike se sada dobija integracijom dobijenih sličnosti za sve prozore sa testne slike.

U poređenju sa [83], u ovoj disertaciji su u algoritam za formiranje reprezentacije slike i računanje sličnosti unesene sljedeće promjene. Umjesto usmjerenih filtara koriste se Gaborovi filtri, izbačene su korelacije između prostorno susjednih Gaborovih koeficijenata, veličina prozora nije ograničena na 7×7 piksela već se tipično koriste veći prozori (64×64 piksela i veći) i slika se posmatra kao neuređeni skup prozora. Gaborovi filtri su upotrebljeni zato što su suštinski slični usmjerenim, a u eksperimentima su rezultovali većom tačnošću klasifikacije od usmjerenih filtara. Korelacije između prostorno susjednih Gaborovih koeficijenata su izbačene zbog osjetljivosti na geometrijske transformacije slike, a veći prozori se koriste kako bi se obezbijedila bolja statistička karakterizacija teksture. Konačno, posmatranjem slike kao neuređenog skupa prozora omogućava se poređenje nehomogenih tekstura.

Izračunavanje lokalnih deskriptora predloženih u ovoj disertaciji počinje filtriranjem ulazne slike Gaborovom filtar bankom na S skala i K orijentacija, kao što je dato jednačinom (3.7). Na ovaj način se dobijaju koeficijenti iz SK podopsega. Kako bi notacija bila jednostavnija u nastavku se, umjesto posebnih indeksa skale i orijentacije, koristi jedinstven indeks podopsega. Neka su $w_k^{\mathbf{x}}$ koeficijenti podopsega $k = 1, \ldots, SK$ u prozoru \mathbf{x} . Za svaki prozor \mathbf{x} se izračunavaju sljedeće statistike vrijednosti koeficijenata:

• Srednje vrijednosti modula podopsežnih koeficijenata

$$\mu_k^{\mathbf{x}} = E\left\{ |w_k^{\mathbf{x}}| \right\} \tag{5.2}$$

• Standardne devijacije modula podopsežnih koeficijenata

$$\sigma_k^{\mathbf{x}} = E\left\{ \left(|w_k^{\mathbf{x}}| - \mu_k^{\mathbf{x}} \right)^2 \right\}$$
(5.3)

• Koeficijenti kroskorelacije između modula koeficijenata u podopsezima k i l, na istoj skali ali različitim orijentacijama, kao i na istoj orijentaciji ali različitim skalama.

$$\rho_{kl}^{\mathbf{x}} = \frac{E\left\{\left(|w_k^{\mathbf{x}}| - \mu_k^{\mathbf{x}}\right)\left(|w_l^{\mathbf{x}}| - \mu_l^{\mathbf{x}}\right)\right\}}{\sigma_k^{\mathbf{x}}\sigma_l^{\mathbf{x}}}, \ k \neq l.$$
(5.4)

U gornjim jednačinama matematička očekivanja se izračunavaju usrednjavanjem po svim koeficijentima u prozoru \mathbf{x} u odgovarajućim podopsezima.

Ove veličine se zatim objedinjavaju u jedan lokalni deskriptor prozora \mathbf{x} . U nastavku će za ovaj deskriptor biti korišćen naziv strukturalni deskriptor teksture (Structural Texture Descriptor, STD). STD je povezan sa Gaborovim deskriptorom teksture [46] i Gist deskriptorom [21], ali se razlikuje od njih po tome što se u deskriptor uključuju i mjere korelacije između podopsežnih koeficijenata i što se radi o lokalnom deskriptoru, nasuprot druga dva koji su globalni. Pored toga, iako STD izračunavamo u pravilno raspoređenim prozorima, sliku posmatramo kao neuređen skup deskriptora ne obraćajući pažnju na njihove prostorne lokacije, dok je Gist deskriptor zasnovan na fiksnom prostornom rasporedu elemenata na slici.

5.3 Mjera strukturne sličnosti

Mjera strukturne sličnosti (Structural Texture SIMilarity, STSIM) dva prozora, **x** i **y**, predstavljena svojim STD, računa se usrednjavanjem sličnosti deskriptora u svim podopsezima,

$$S(\mathbf{x}, \mathbf{y}) = \frac{1}{SK} \sum_{k=1}^{SK} l_k^{\frac{1}{3}}(\mathbf{x}, \mathbf{y}) c_k^{\frac{1}{3}}(\mathbf{x}, \mathbf{y}) r_k^{\frac{1}{3}}(\mathbf{x}, \mathbf{y}), \qquad (5.5)$$

gdje je $l_k(\mathbf{x}, \mathbf{y})$ sličnost srednjih vrijednosti

$$l_{k}(\mathbf{x}, \mathbf{y}) = \frac{2\mu_{k}^{\mathbf{x}}\mu_{k}^{\mathbf{y}} + K_{1}}{(\mu_{k}^{\mathbf{x}})^{2} + (\mu_{k}^{\mathbf{y}})^{2} + K_{1}},$$
(5.6)

a $c_k(\mathbf{x}, \mathbf{y})$ sličnost standardnih devijacija podopsežnih koeficijenata u podopsegu $k = 1, \ldots, SK$

$$c_k\left(\mathbf{x},\mathbf{y}\right) = \frac{2\sigma_k^{\mathbf{x}}\sigma_k^{\mathbf{y}} + K_2}{\left(\sigma_k^{\mathbf{x}}\right)^2 + \left(\sigma_k^{\mathbf{y}}\right)^2 + K_2}.$$
(5.7)

U jednačinama (5.6) i (5.7) K_1 i K_2 su male konstante čija je uloga da poboljšaju numeričku stabilnost mjere sličnosti.

Clan $r_k(\mathbf{x}, \mathbf{y})$ predstavlja doprinos koeficijenata kroskorelacije mjeri sličnosti i dobija se usrednjavanjem sličnosti svih koeficijenata kroskorelacije:

$$r_k(\mathbf{x}, \mathbf{y}) = \frac{1}{N_{\rho}} \left[\sum_{k \neq l} \left(1 - \frac{1}{2} |\rho_{kl}^{\mathbf{x}} - \rho_{kl}^{\mathbf{y}}| \right) \right],$$
(5.8)

gdje je $N_{\rho} = S + K - 2$ broj koeficijenata kroskorelacije za jedan podopsežni prozor.

5.4 Klasifikacija slika

Da bi se klasifikovala slika predstavljena lokalnim deskriptorima potrebno je formirati reprezentaciju slike koja će služiti kao ulazni podatak za sam klasifikator. Jedan od najpopularnijih pristupa za formiranje reprezentacije slike polazeći od lokalnih deskriptora je pristup poznat pod nazivom *skup riječi (Bag-of-words)* [11] koji koristi kvantizaciju lokalnih deskriptora čime se formiraju *vizuelne riječi*. Slika se zatim predstavlja histogramom pojavljivanja pojedinih vizuelnih riječi, a dobijeni deskriptori se koriste sa nekim od klasičnih algoritama za klasifikaciju. Ovaj pristup je inspirisan klasifikacijom tekstualnih dokumenata gdje se dokument predstavlja histogramom pojavljivanja pojedinih riječi [85].

U [13] su identifikovana dva problema vezana sa ovim pristupom. Prvi je gubitak diskriminativnih informacija zbog kvantovanja lokalnih deskriptora, a drugi je činjenica da većina algoritama za klasifikaciju uključuje izračunavanje mjere sličnosti između testne slike i slika iz trening skupa (sličnost slika-slika). Na primjer, klasifikator zasnovan na principu najbližih susjeda uključuje izračunavanje mjere sličnosti između testne slike i označenih slika pa testnu sliku klasifikuje u klasu koja sadrži najsličniju označenu sliku, dok mašine sa vektorima nosačima koriste mjeru sličnosti između testne slike i slika koje odgovaraju vektorima nosačima (*support* vektorima). Sa druge strane, u [13] je predloženo korišćenje lokalnih deskriptora bez kvantovanja, te da se, umjesto sličnosti slika-slika, koristi sličnost slika-klasa između testne slike i svake od klasa. Testna slika se klasifikuje u klasu za koju ova mjera sličnosti ima maksimalnu vrijednost. U nastavku je izložena varijanta ovog pristupa prilagođena korišćenju sa strukturalnim deskriptorima teksture.

Neka je data slika \mathbf{X} koja se može prikazati i kao skup prozora ili blokova $\mathbf{X} = {\mathbf{x}_1, \ldots, \mathbf{x}_n}$. Neka su, dalje, za datu sliku izračunati strukturalni deskriptori teksture za određen skup prozora ili blokova , kao što je opisano. Ako slika \mathbf{X} pripada klasi C, smatra se da svi njeni blokovi \mathbf{x}_i pripadaju istoj klasi što se označava sa $\mathbf{x}_i \in C$. Sličnost blok-klasa se definiše kao maksimalna vrijednost sličnosti datog bloka \mathbf{x} i svih blokova iz posmatrane klase C

$$S(\mathbf{x}, C) = \max_{\mathbf{y} \in C} S(\mathbf{x}, \mathbf{y}).$$
(5.9)

Sličnost testne slike $\mathbf{X} = {\mathbf{x}_1, \dots, \mathbf{x}_n}$ i klase C je data sa

$$S(\mathbf{X}, C) = \prod_{i=1}^{n} S(\mathbf{x}_i, C).$$
(5.10)

Izračunavanje sličnosti slika-klasa je ilustrovano na Slici 5.4. Za svaki blok testne slike pronalazi se najsličniji blok iz skupa označenih slika koje pripadaju istoj klasi. Ukupna sličnost slika-klasa je proizvod pojedinih sličnosti blok-klasa za svaki od blokova testne slike.

Testna slika \mathbf{X} se sada klasifikuje na sljedeći način:

- 1. Izračunati STD za blokove $\mathbf{x}_1, \ldots, \mathbf{x}_n$ testne slike \mathbf{X} ;
- 2. Za svaki blok \mathbf{x}_i , i = 1, ..., n testne slike i za svaku klasu C_j , j = 1, ..., k, gdje je k broj klasa, izračunati sličnost $S(\mathbf{x}_i, C_j)$ korišćenjem (5.10);



Označeni primjeri iz klase *Industrija* Sličnost slika-klasa 0.8226

Slika 5.4: Sličnost slika-klasa. Pojedine sličnosti blok-klasa su naznačene na strelicama.

3. Izračunati sličnosti slika-klasa testne slike i svake od klasa

$$S\left(\mathbf{X}, C_{j}\right) = \prod_{i=1}^{n} S\left(\mathbf{x}_{i}, C_{j}\right), \ j = 1, \dots, k;$$

$$(5.11)$$

4. Klasifikovati sliku ${\bf X}$ u klasu

$$\hat{C} = \arg\max_{C} S\left(\mathbf{X}, C\right). \tag{5.12}$$

U primjeru sa Slike 5.2 se vidi da je, kada se slika posmatra kao neuređen skup prozora i koristi modifikacija mjere sličnosti data jednačinama (5.9) i (5.10), testna slika sličnija označenoj slici iz iste klase (teniski teren) nego označenoj slici iz druge klase (rijetko naseljeno) i biće tačno klasifikovana korišćenjem klasifikatora na principu najbližih susjeda.

Dodatno poboljšanje klasifikatora na principu najbližih susjeda može se postići korišćenjem sličnosti slika-klasa umjesto uobičajene sličnosti slika-slika. Prednost korišćenja sličnosti slika-klasa u poređenju sa sličnošću slika-slika u klasifikaciji slika ilustrovana je na Slici 5.5. Neka su dati označeni primjeri iz dvije klase, *industrija* i *kuće*, te testna slika za koju je pozitivna klasa *industrija*. Za sve slike izračunat je strukturalni deskriptor teksture korišćenjem podjele podopsega mrežom od 4×4 blokova koji se ne preklapaju. Na osnovu vrijednosti slika-slika, najsličnija testnoj slici bila označena slika koja pripada negativnoj klasi. Zbog toga bi, u slučaju korišćenja sličnosti slika-slika, testna slika bila pogrešno klasifikovana. Sa druge strane, kada se koristi sličnost slika-klasa, testna slika će biti tačno klasifikovana, pošto je sličnost slika-klasa za klasu *industrija* veća nego za klasu *kuće*.



Pozitivna klasa Industrija (Sličnost slika-klasa 0.8226)

Negativna klasa Kuće (Sličnost slika-klasa 0.7893)

Slika 5.5: Poređenje sličnosti slika-klasa i sličnosti slika-slika. Vrijednosti sličnosti slika-slika testne slike i označenih slika su naznačene na strelicama, a za svaku klasu je navedena vrijednost sličnosti slika-klasa testne slike i te klase.

5.5 Eksperimentalni rezultati

Strukturalni deskriptor teksture je predložen za reprezentaciju panhromatskih slika. Da bi se deskriptor izračunao prvo se svaka slika filtrira Gaborovom filtar bankom na S = 4 skale i K = 6 orijentacija. Podopsezi se zatim dijele na blokove nekom od mreža dimenzija 1×1 , 2×2 ili 4×4 blokova. U svakom od tako dobijenih blokova izračunavaju se srednje vrijednosti (5.2) i standardne devijacije modula koeficijenata (5.3), kao i koeficijenti korelacije modula koeficijenata iz različitih podopsega (5.4), čime se formira STD panhromatske slike. Za slike u boji deskriptori i sličnosti se za svaki spektralni opseg iz RGB kolor-prostora izračunavaju posebno, a zatim se dobijene sličnosti usrednjavaju. Važno je primjetiti da se kada se deskriptori izračunavaju za čitav podopseg, tj. za podjelu podopsega se koristi mreža dimenzija 1×1 , u stvari računaju i koriste globalne statistike podopsežnih koeficijenata. Sa druge strane, kada se koriste podjele podopsega na veći broj blokova, testiranje ovog deskriptora je ograničeno velikom računskom složenošću algoritama za njegovo izračunavanje i klasifikaciju tekstura na principu najbližih susjeda.

Cilj prvog skupa eksperimenata je da se ustanovi uticaj koji imaju projektanstki izbori učinjeni pri konstrukciji strukturalnog deskriptora teksture, kao i modifiko-



Slika 5.6: Poređenje performansi različitih klasifikatora zasnovanih na STD.

vane strukturalne mjere sličnosti, na performanse klasifikatora zasnovanog na principu najbližih susjeda koji koristi strukturalnu sličnost tekstura (*Nearest Neighbor Structural Texture SIMilarity – NN-STSIM*. Upoređeni su NN-STSIM klasifikatori zasnovani na Gaborovim filtrima sa: (a) NN-STSIM klasifikatorima zasnovanim na kompleksnim usmjerenim filtrima, kao i NN-STSIM klasifikatorima koji koriste: (b) metriku indukovanu L_1 normom sa STD i sličnošću slika-klasa, (c) sličnost slika-slika sa STSIM i (d) STSIM izračunatu samo sa srednjim vrijednostima i standardnim devijacijama podopsežnih blokova.

Eksperimenti su izvršeni na kolekciji slika Banje Luke, opisanoj u odjeljku 2.1. Za ovu evaluaciju korišćene su samo panhromatske verzije slika. Strukturalni deskriptori teksture ovih slika se izračunavaju na prethodno opisani način korišćenjem Gaborovih, odnosno, kompleksnih usmjerenih filtara. Iz svake klase polovina slika se koristi kao označene slike, a druga polovina kao testne slike. Eksperiment se ponavlja 10 puta sa različitim slučajnim podjelama kolekcije slika i dobijeni rezultati se usrednjavaju.

Poređenja su izvršena za različit broj podopsežnih blokova $(1 \times 1, 2 \times 2 \text{ i } 4 \times 4)$, a performanse su prikazane na Slici 5.6. Može se primjetiti da klasifikatori koji koriste STD zasnovane na Gaborovim filtrima (označeni sa Gabor na Slici 5.6) u svim slučajevima rezultuju boljom tačnošću klasifikacije od klasifikatora zasnovanih na kompleksnim usmjerenim filtrima (označeni sa steerable). Zbog toga će u narednim eksperimentima biti korišćena samo Gaborova filtar banka.

U poređenju sa drugim varijantama klasifikatora, može se primjetiti da klasifikatori koji koriste STSIM imaju bolje performanse od onih koji koriste metriku

Klasifikator	Panhromatske (%)	RGB (%)
NN-STSIM 1×1	84.0	86.8
NN-STSIM 2×2	86.9	89.6
NN-STSIM 4×4	87.3	88.3
SVM Gabor [23]	84.5	88.0
SVM Gist [23]	79.5	89.3

Tabela 5.1: Tačnosti klasifikacije (%) za skup slika Banje Luke

indukovanu L_1 normom što je posljedica perceptualne zasnovanosti STSIM. Nadalje, za klasifikator je povoljna bogatija statistika obilježja pa je klasifikator sa punim deskriptorima uvijek bolji od klasifikatora koji koristi samo srednje vrijednosti i standardne devijacije podopsežnih blokova. Najzad, vidi se da klasifikator sa sličnošću slika-klasa ima bolje performanse od klasifikatora sa sličnošću slika-slika, što je u skladu sa primjerom na Slici 5.5 i rezultatima u [13]. Na osnovu ovih rezultata u narednim eksperimentima će, pored Gaborove filtar banke, biti korišćeni STD koji uključuju koeficijente korelacije između različitih podopsega, perceptualno zasnovana STSIM mjera sličnosti i klasifikator zasnovan na principu najbližih susjeda koji koristi sličnost slika-klasa.

Upoređena je i tačnosti klasifikacije dobijena strukturalnim deskriptorima teksture sa rezultatima dobijenim korišćenjem Gaborovih i Gist deskriptora i SVM klasifikatora iz [23]. Rezultati su dati u Tabeli 5.1. U ovoj Tabeli, uz NN-STSIM klasifikatore naznačene su podjele podopsega. U poslednja dva reda su rezultati dobijeni korišćenjem SVM sa Gaborovim i Gist deskriptorima.

Vidi se da za panhromatske slike NN-STSIM 4×4 klasifikator daje najbolje performanse i čak je bolji od pristupa zasnovanih na SVM. Za slike u boji, najbolje performanse se dobijaju korišćenjem NN-STSIM 2×2 klasifikatora koji je, ponovo, bolji od SVM.

Matrica konfuzije za NN-STSIM 2×2 klasifikator primjenjen na kolor slike je data na Slici 5.7. Može se primjetiti da se konfuzije uglavnom javljaju između klasa koje su teške i za ljudske posmatrače. Najznačajniji primjeri konfuzija su između *kuća* i groblja, kao i između *industrije* i groblja, te između *rijeka* i polja. Konfuzije u prva dva slučaja su posljedica toga što na slikama iz sve tri klase postoje pravougaone strukture sa jakim orijentisanim ivicama. U trećem slučaju konfuzije su posljedica postojanja homogene, glatke teksture bez izraženih ivica. Važno je primjetiti i da nema mnogo konfuzija između prirodnih (*rijeka, drveće, polja*) i vještačkih (*kuće, groblje, industrija*) klasa.

5.5.1 UC Merced Land Use kolekcija slika

Pošto zbog korišćenja klasifikatora na principu najbližeg susjeda ne postoji faza obučavanja klasifikatora, da bi se rezultati mogli uporediti sa ostalim rezultatima dobijenim korišćenjem drugih klasifikatora, u ovom eksperimentu će 80% slika iz svake klase biti označeni primjeri, a preostalih 20% testni primjeri. Tačnosti klasi-



Slika 5.7: Matrica konfuzija za skup slika Banje Luke.

Tabela 5.2: Tačnosti klasifikacije (%) za UC Merced Land Use kolekciju slika.

Klasifikator	Panhromatske (%)	RGB (%)
NN-STSIM 1×1	$76,90 \pm 2,18$	$80, 19 \pm 0, 97$
NN-STSIM 2×2	$81, 10 \pm 2, 36$	$83, 10 \pm 1, 39$
NN-STSIM 4×4	$82, 86 \pm 1, 57$	$85,00\pm0,67$
SVM Gabor	$76,05 \pm 1,46$	$79,52 \pm 1,82$

fikacije dobijene korišćenjem klasifikatora na principu najbližeg susjeda uz STSIM mjeru sličnosti slika i korišćenje sličnosti slika-klasa date su u Tabeli 5.2. Korišćene su tri prostorne podjele podopsega, 1×1 , 2×2 i 4×4 blokova. Dobijeni rezultati su označeni sa NN-STSIM i odgovarajućom prostornom podjelom podopsega. Poređenja radi dati su i rezultati dobijeni korišćenjem SVM klasifikatora i Gaborovog deskriptora teksture (označeni sa SVM Gabor). Svi rezultati su dati za panhromatske i za RGB slike.

Može se primjetiti da se korišćenjem predloženog NN-STSIM klasifikatora dobijaju konzistentno bolji rezultati nego korišćenjem SVM klasifikatora i pored toga što je predloženi klasifikator zasnovan na principu najbližeg susjeda i nije eksplicitno obučavan. Razlozi za ovo leže u:

- 1. korišćenju lokalnih deskriptora prikladnijih za opis nehomogene teksture;
- 2. uključivanju koeficijenata korelacije između koeficijenata u različitim podopsezima u strukturalni deskriptor teksture;
- 3. korišćenju perceptualno zasnovane STSIM mjere sličnosti i
- 4. korišćenju sličnosti slika-klasa kojom se ublažava problem varijabilnosti unutar klasa.



Slika 5.8: Matrica konfuzija za UC Merced Land Use kolekciju slika dobijena za NN-STSIM 4×4 slučaj iz Tabele 5.2 za RGB slike.

Zanimljivo je primjetiti da se već sa prostornom podjelom 1×1 blok, što odgovara globalnim deskriptorima slike, dobijaju približno isti rezultati kao pri korišćenju SVM klasifikatora. Kada se koriste podjele podopsega na 2×2 i 4×4 bloka tačnost klasifikacije dalje raste. Međutim, iako se korišćenjem finijih prostornih podjela dobijaju bolji rezultati, ovu zavisnost nije bilo moguće detaljno ispitati zbog velike složenosti izračunavanja sličnosti slika-klasa. Ipak, ovo ostaje zanimljiva tema za dalje istraživanje.

Uključivanjem informacije o boji u deskriptor, performanse klasifikatora se u izvjesnoj mjeri poboljšavaju. Ovo poboljšanje se kreće u opsegu od 2% do 4%. Dobijeno poboljšanje je posljedica bolje reprezentacije teksture koja se dobija kada se koriste različiti spektralni opsezi. Međutim, s obzirom na to da je najčešće moguće uspješno razlikovati teksture korišćenjem samo intenzitetske informacije, dosta dobra tačnost klasifikacije se dobija već za panhromatske slike, a poboljšanje postignuto izračunavanjem deskriptora teksture u više spektralnih opsega nije veliko. Ipak, intuitivno je jasno da će uključivanjem spektralne informacije u reprezentaciju slike tačnost klasifikacije biti poboljšana. Međutim, pošto je namjena predloženog deskriptora opis teksture, on ne predstavlja adekvatan opis spektralne informacije na slici.

Na Slici 5.8 prikazana je matrica konfuzija za UC Merced Land Use kolekciju slika dobijena korišćenjem NN-STSIM klasifikatora sa podjelom podopsega na 4 × 4 blokova. Deskriptori se računaju za RGB slike. Najčešće konfuzije se javljaju između klasa gusto i srednje naseljeno rezidencijalno zemljište, autoput i nadvožnjak, zgrade i raskrsnica, raskrsnica i gusto naseljeno rezidencijalno zemljište, te šuma i

Klasifikator	Panhromatske (%)	RGB $(\%)$
NN-STSIM 1×1	$65, 48 \pm 1, 68$	$66, 35 \pm 1, 64$
NN-STSIM 2×2	$72, 10 \pm 0, 74$	$72,94\pm0,91$
NN-STSIM 4×4	$78, 24 \pm 1, 44$	$79,04 \pm 1,80$
SVM Gabor	$65, 24 \pm 1, 44$	$66,03 \pm 1,74$

Tabela 5.3: Tačnosti klasifikacije (%) za Wuhan kolekciju slika.



Slika 5.9: Matrica konfuzija za Wuhan kolekciju slika dobijena za NN-STSIM 4×4 slučaj iz Tabele 5.3 za RGB slike.

rijeka. U većini ovih slučajeva konfuzije se javljaju na primjerima koji imaju izražena zajednička vizuelna obilježja i čije je diskriminacija teška i za ljudskog posmatrača.

5.5.2 Wuhan kolekcija slika

Kako bi se omogućilo poređenje rezultata sa literaturom, u ovom eksperimentu se polovina slika iz svake klase koristi kao označeni primjeri, a klasifikator se testira na drugoj polovini. Tačnosti klasifikacije za Wuhan kolekciju slika date su u Tabeli 5.3. Navedeni su rezultati dobijeni korišćenjem klasifikatora na principu najbližeg susjeda uz korišćenje STSIM mjere sličnosti i sličnosti slika-klasa za sve tri prostorne podjele podopsega: 1×1 , 2×2 i 4×4 blokova. Rezultati dobijeni predloženim klasifikatorima označeni su u tabeli sa NN-STSIM i oznakom prostorne podjele podopsega. Poređenja radi, data je i tačnost SVM klasifikatora koji koristi Gaborov deskriptor teksture (označen sa SVM Gabor). Vidimo da su performanse NN-STSIM klasifikatora u najgorem slučaju (prostorna podjela 1×1) jednake performansama SVM klasifikatora, iako se radi o jednostavnijem klasifikatoru u kojem ne postoji eksplicitna faza obučavanja. Razlozi za ovakve rezultate su isti kao i u slučaju UC Merced Land Use kolekcije slika.

U Tabeli 5.3 su dati rezultati kako za panhromatske tako i za RGB slike. Iako su tačnosti klasifikacije dobijene za RGB slike konzistentno bolje od tačnosti dobijenih za panhromatske slike ipak postignuta poboljšanja nisu značajna, tj. razlike u odnosu na tačnosti dobijene za panhromatske slike su unutar jedne standardne devijacije. Ova poboljšanja su još manja nego poboljšanja dobijena za UC Merced Land Use kolekciju slika čime je još više naglašena činjenica da predloženi deskriptor nije prikladan za izdvajanje diskriminativne spektralne informacije iz slika. Ipak, kao deskriptor teksture, on daje vrlo dobre rezultate.

Na Slici 5.9 prikazana je matrica konfuzija za Wuhan kolekciju slika dobijena korišćenjem NN-STSIM klasifikatora sa podjelom podopsega na 4×4 blokova. Deskriptori se računaju za RGB slike. Najčešće konfuzije se javljaju između klasa *rezidencijalno zemljište* i *poslovne zgrade, most* i *jezero, industrija* i *poslovne zgrade, rijeka* i *park*, te *parking* i *luka*. Ove konfuzije su posljedica izražene vizuelne sličnosti primjera iz različitih klasa i u većini slučajeva su izazov i za ljudske posmatrače.

Glava 6

Fuzija globalnih i lokalnih deskriptora

Prošireni Gaborov deskriptor teksture, predložen u ovoj disertaciji, je globalni deskriptor teksture. Dobri rezultati koji se njegovom upotrebom postižu u klasifikaciji snimaka dobijenih daljinskom detekcijom su posljedica velikog značaja teksture za diskriminaciju različitih klasa pokrivača i načina korišćenja zemljišta. Međutim, postoje klase načina korišćenja zemljišta koje su u potpunosti određene pojedinim objektima prisutnim na slikama, kao što su npr. cisterne, sportski tereni, raskrsnice, itd. Pošto se globalnim deskriptorima slika opisuje kao cjelina, ta reprezentacija je osjetljiva na postojanje velikog broja irelevantnih detalja na slici (*clutter*), a često i na geometrijske transformacije slike što rezultuje slabijim tačnostima klasifikacije. U ovim slučajevima bolji rezultati bi se mogli postići korišćenjem lokalnih deskriptora koji se uobičajeno koriste u prepoznavanju objekata opšte namjene.

Globalni i lokalni deskriptori sadrže komplementarne informacije o slici i njihova fuzija ima potencijal da poboljša performanse klasifikatora. Predložena je hijerarhijska šema za fuziju deskriptora u kojoj se klasifikatori na prvom nivou obučavaju korišćenjem globalnih i lokalnih deskriptora slika, a zatim se, na osnovu njihovih ocjena pouzdanosti, formira međureprezentacija slika koja se koristi kao ulazni podatak za klasifikator na drugom nivou.

6.1 Lokalni deskriptori slike

Korišćenjem lokalnih deskriptora slike [86, 87, 88, 89] u prethodnih 15 godina ostvareni su značajni pomaci u rješavanju problema prepoznavanja objekata opšte namjene u prisustvu geometrijskih transformacija slike, promjena osvjetljenja i sjene, kao i postojanja velikog broja irelevantnih detalja na slici. Lokalni deskriptor opisuje raspodjelu prostornih obilježja lokalno, u malom regionu na slici. Ova lokalna informacija je u manjoj mjeri osjetljiva na geometrijske transformacije i postojanje irelevantnih detalja na slici.

Trenutno, korišćenjem lokalnih deskriptora se postižu najbolji rezultati u prepoznavanju objekata i klasifikaciji scena uopšte [90]. Najpopularniji lokalni deskriptori u ovom trenutku su *Scale-Invariant Feature Transform (SIFT)* [88] i *histogram orijentisanih gradijenata (HOG)* [89]. Oba deskriptora su, u suštini, histogrami orijentacija gradijenata u okolini interesne tačke. Korišćenjem gradijenata za opis regiona u okolini interesne tačke postiže se smanjena osjetljivost na varijacije osvjetljenja. Većina algoritama za prepoznavanje objekata i klasifikaciju scena, koji danas dominiraju u literaturi, zasnovani su na varijantama ovih deskriptora.

SIFT deskriptor je inicijalno predložen za pronalaženje objekata na slikama direktnim uparivanjem deskriptora određenih u regionima oko interesnih tačaka [88] i ukratko je opisan u Dodatku C. Ovaj pristup zahtijeva da se čuvaju deskriptori određeni u svim interesnim tačkama na slici. Sa druge strane, ukoliko se čuvaju svi deskriptori reprezentacija slike postaje neefikasna. Zbog toga je u [91] predloženo da se gustina raspodjele vjerovatnoće SIFT deskriptora na slici estimira vektorskom kvantizacijom deskriptora i formiranjem histograma broja pojavljivanja kodnih riječi na slici. Ovaj pristup je sličan paradigmi pod nazivom *skup riječi (bag-of-words)* koja se koristi u pretraživanju baza punog teksta i klasifikaciji tekstualnih dokumenata [85]. Zbog toga je ovaj pristup i nazvan *skup vizuelnih riječi*. Pokazalo se da se kvantovanjem SIFT deskriptora u određenoj mjeri postiže i robusnost na šum i varijacije u izgledu objekata.

U oblasti klasifikacije slika opšte namjene ova paradigma je prvi put upotrebljena u [11]. Ubrzo je primjećeno da se u klasifikaciji bolji rezultati dobijaju ako se ne koristi detektor interesnih tačaka već se deskriptori izračunavaju u regionima oko pravilno raspoređenih tačaka [92]. Budući da je zasnovan na histogramu pojavljivanja kodnih riječi, skup vizuelnih riječi u potpunosti ignoriše prostorni raspored lokalnih deskriptora. U [12] je pokazano da se formiranjem prostorne piramide, tj. uključivanjem grube informacije o lokacijama SIFT deskriptora mogu značajno poboljšati performanse klasifikatora slika opšte namjene. Lokalni deskriptori pokazali su se dobrim kako u prepoznavanju objekata i scena [90], tako i u klasifikaciji tekstura [10].

Kada je riječ o klasifikaciji snimaka dobijenih daljinskom detekcijom rezultati u literaturi ne omogućavaju donošenje konačne odluke. U [93] i [33], na zadacima klasifikacije slika u 6, odnosno, 19 klasa pokrivača i načina korišćenja zemljišta, deskriptori zasnovani na SIFT-u postižu rezultate bolje od ostalih deskriptora. Međutim, u [57] i [30], na skupu slika koji sadrži 21 klasu pokrivača i načina korišćenja zemljišta, rezultati dobijeni pomoću SIFT deskriptora su u rangu ili lošiji od rezultata dobijenih Gaborovim deskriptorima teksture ili čak kolor-histogramom u HLS kolor-prostoru.

Lokalni deskriptori ne moraju nužno biti zasnovani na gradijentu slike. Tako se u [94] lokalni deskriptori formiraju korišćenjem srednjih vrijednosti i standardnih devijacija piksela u tri kolor-kanala, kao i matrica zajedničkih pojavljivanja nivoa sivog (gray level co-occurrence matrix, GLCM). Pokazano je da su performanse klasifikatora zasnovanog na takvim deskriptorima bolje od performansi klasifikatora zasnovanog na SIFT deskriptoru. U [95] su predloženi lokalni deskriptori dobijeni na osnovu PCA transformacije vrijednosti piksela u lokalnom prozoru i testirani na problemu detekcije urbanih regiona. Pokazano je da ovi deskriptori daju bolje rezultate od GLCM deskriptora. Potrebno je, dakle, razmotriti u kojim slučajevima lokalni deskriptori daju dobre rezultate u klasifikaciji slika dobijenih daljinskom detekcijom i kako je moguće postići dalje poboljšanje njihovih performansi.

Pokazalo se i da se uključivanjem prostorne informacije pomoću prostorne pira-



Slika 6.1: Blok dijagram hijerarhijske šeme za fuziju deskriptora.

mide, za razliku od slika opšte namjene, u klasifikaciji slika dobijenih daljinskom detekcijom ne dobija značajno poboljšanje performansi [30]. Ovaj rezultat je sličan rezultatu dobijenom za Gist deskriptor u odjeljku 3.3. Međutim, značaj prostornih odnosa između vizuelnih obilježja za klasifikaciju snimaka dobijenih daljinskom detekcijom se ne može zanemariti [1], pa je u [30] predloženo uključivanje prostorne informacije formiranjem histograma zajedničkih pojavljivanja kodnih riječi. Na ovaj način je ostvareno određeno poboljšanje rezultata, ali ovaj smjer istraživanja još uvijek nudi zanimljive mogućnosti.

Pošto trenutno ne postoji konsenzus o tome da li su globalni deskriptori teksture ili lokalni deskriptori bolji za klasifikaciju slika dobijenih daljinskom detekcijom u ovoj disertaciji eksperimentalno su upoređeni prošireni Gaborov deskriptor teksture i SIFT deskriptor upotrebljen za formiranje reprezentacije pomoću skupa vizuelnih riječi. Pokazano je da odluka o tome koji će deskriptor dati bolje performanse zavisi od klase pokrivača i načina korišćenja zemljišta. U nekim slučajevima bolji su globalni, a u drugim lokalni deskriptori. Dakle, globalni i lokalni deskriptori sadrže komplementarne informacije o slici i njihova fuzija ima potencijal za popravku performansi klasifikatora.

6.2 Hijerarhijska fuzija deskriptora

Predloženi pristup fuziji deskriptora je hijerarhijski i sastoji se iz dva nivoa, prikazana na Slici 6.1. Na prvom nivou (nivo 0) slika se najprije predstavlja korišćenjem različitih deskriptora. Ovi deskriptori mogu biti proizvoljno izabrani, bez ikakvih ograničenja. Jedini zahtjev koji se intuitivno nameće je da bi deskriptori trebalo da sadrže komplementarne informacije o slici kako bi se njihovom kombinacijom poboljšale performanse klasifikatora. Takođe, broj deskriptora na nivou 0 teorijski nije ograničen.

Zatim se za svaki od deskriptora obučava klasifikator baziran na tom deskriptoru

kao reprezentaciji slike. I ovi klasifikatori se mogu proizvoljno izabrati, uz zahtjev da na izlazu svakog klasifikatora postoje ocjene pouzdanosti sa kojim bi posmatrani primjer bio klasifikovan u svaku od klasa. Broj klasifikatora na nivou 0 jednak je broju deskriptora kojima je slika predstavljena i oni se obučavaju nezavisno jedan od drugog. Polazeći od ocjena pouzdanosti klasifikatora sa ovog nivoa, formiraju se međudeskriptori slika konkatenacijom ocjena pouzdanosti dobijenih od svih klasifikatora na nivou 0.

Kao što se vidi sa Slike 6.1, na drugom nivou (nivo 1) postoji samo jedan klasifikator koji se naziva *metaklasifikator*. Ulazni podaci za metaklasifikator su međudeskriptori dobijeni konkatenacijom izlaza klasifikatora na nivou 0. Na izlazu metaklasifikatora dobija se konačna oznaka klase u koju je klasifikovan testni primjer.

Nedavno su slične šeme za klasifikaciju, koje koriste mašine sa vektorima nosačima, predložene za klasifikaciju slika, kako opšte namjene [96], tako i slika dobijenih daljinskom detekcijom [34]. Ovi klasifikatori u stvari koriste *hijerarhijsko dodavanje* klasifikatora (*stacking*) [97] za kombinovanje klasifikatora pomoću SVM. Međutim, u literaturi iz oblasti mašinskog učenja ustanovljeno je da se dobri rezultati mogu postići i korišćenjem vrlo jednostavnih klasifikatora, kao što su linearna [98] ili regularizovana linearna regresija [99]. U nastavku je dat formalan opis predložene tehnike za formiranje ansambla različitih klasifikatora. Oznake nivoa prate [97].

Neka je dat skup od C klasa u koje treba klasifikovati slike i neka je na raspolaganju trening skup sa označenim slikama. Za svaku sliku se izračunava M deskriptora, $\mathbf{f}_i = [f_{i1}, f_{i2}, \ldots, f_{id_i}]^T$, gdje je $i = 1, \ldots, M$, a d_i je dimenzionalnost *i*-tog deskriptora.

Na nivou 0 se obučava M klasifikatora, po jedan za svaki od deskriptora, a njihovi izlazi,

$$\mathbf{p}'_{i} = [p'_{i1}, p'_{i2}, \dots, p'_{iC}]^{T} = cl(\mathbf{f}_{i}), \ i = 1, \dots, M,$$
(6.1)

su ocjene pouzdanosti da ulaznu sliku treba klasifikovati u svaku odCklasa. Ove pouzdanosti se transformišu u aposteriorne vjerovatnoće korišćenjem softmax funkcije

$$p_{ij} = \frac{e^{\beta p'_{ij}}}{\sum_{k=1}^{C} e^{\beta p'_{ik}}}, i = 1, \dots, M, j = 1, \dots, C,$$
(6.2)

gdje se parametrom β kontroliše njena "strmina". Sada je za svaki od deskriptora na raspolaganju vektor posteriornih vjerovatnoća \mathbf{p}_i , $i = 1, \ldots, M$ da će slika biti klasifikovana u svaku od C klasa. Posteriorne vjerovatnoće se zatim nadovezuju u međudeskriptore $\mathbf{P} = \begin{bmatrix} \mathbf{p}_1^T, \ldots, \mathbf{p}_M^T \end{bmatrix}^T$ čija je dimenzionalnost jednaka MC. Konačno, klasifikator na nivou 1 se obučava korišćenjem ovako dobijenih međudeskriptora. Na njegovom izlazu se dobija posteriorna predikcija klase u koju će biti klasifikovana data slika

$$\hat{c} = ml\left(\mathbf{P}\right). \tag{6.3}$$

U fazi klasifikacije, za testnu sliku se takođe izračunava M različitih deskriptora i dobijena reprezentacija se dovodi na ulaze klasifikatora na prvom nivou. Ovi klasifikatori na svojim izlazima (6.1) daju vektore ocjena pouzdanosti da će slika biti klasifikovana u svaku od datih C klasa. Softmax funkcijom (6.2) se dobijaju posteriorne vjerovatnoće klasifikacije slike, a zatim se njihovom konkatenacijom formira međureprezentacija testne slike. Ova međureprezentacija je ulazni podatak za metaklasifikator (6.3) koji daje konačnu predikciju klase kojoj slika pripada.

Zanimljiva varijanta opisane šeme je predložena u [100]. Umjesto korišćenja svih ocjena pouzdanosti za treniranje jednog klasifikatora na nivou 1, možemo trenirati C klasifikatora, po jedan za svaku klasu, korišćenjem samo ocjena pouzdanosti za tu klasu, dobijenih od svih M klasifikatora na nivou 0

$$z_c = m l_c \left(p_{1c}, p_{2c}, \dots, p_{Mc} \right), \ c = 1, \dots, C.$$
(6.4)

Testna slika se klasifikuje u klasu

$$\hat{c} = \arg \max_{c=1,\dots,C} z_c. \tag{6.5}$$

Ovaj pristup se naziva *stackingC*.

Da bi se obučio metaklasifikator potrebno je prvo obučiti klasifikatore na nivou 0 i formirati međureprezentaciju trening skupa. Međutim, pošto se međureprezentacija formira konkatenacijom izlaza klasifikatora na nivou 0, korišćenje ocjena pouzdanosti dobijenih za primjere iz trening skupa bi moglo dovesti do loše generalizacije klasifikatora. Da bi se obezbijedila dobra generalizacija, u fazi obučavanja međudeskriptori se ne smiju formirati korišćenjem istih ulaznih podataka koji su korišćeni za obučavanje klasifikatora na prvom nivou. Kako bi se ovo postiglo međureprezentacija se formira korišćenjem procedure slične unakrsnoj validaciji [97]. Trening skup se slučajno podijeli na P dijelova. Jedan od dijelova je validacioni skup, a klasifikatori na nivou 0 se obučavaju korišćenjem ostatka trening skupa. Međureprezentacija se zatim formira za primjere iz validacioni skup, a trenira se na ostatku trening skupa. Na taj način se formira međureprezentacija za sve primjere iz trening skupa, a metaklasifikator se obučava na cijelom trening skupu.

6.3 Poređenje globalnih i lokalnih deskriptora

U svim eksperimentima u ovom poglavlju je, kao predstavnik globalnih deskriptora slike, korišćen prošireni Gaborov deskriptor teksture izračunat pomoću Gaborove filtar banke sa 4 skale i 6 orijentacija. Dobijeni deskriptor je 240-dimenzionalan. Kao predstavnik lokalnih deskriptora slike korišćen je SIFT deskriptor, uz reprezentaciju slike kao skupa vizuelnih riječi. SIFT deskriptori su računati u prozorima veličina 8×8 i 16×16 piksela sa korakom od 8 piksela, a zatim vektorski kvantovani u 1000 kodnih riječi korišćenjem k-means algoritma. Slika je predstavljena histogramom pojavljivanja kodnih riječi. Dobijeni histogrami su 1000-dimenzionalni i rijetko popunjeni. Oba deskriptora su izračunata za panhromatske slike.

Za klasifikaciju na nivou 0 su testirane SVM sa linearnim, χ^2 , RBF i generalizovanim RBF kernelom uz korišćenje metrike (4.17). Za obje kolekcije slika eksperimenti su ponovljeni uz korišćenje 10, 50 i 90% slika iz svake klase za obučavanje. Ostale slike su korišćenje za testiranje. Ciklus obučavanja i testiranja klasifikatora se ponavlja pet puta za UC Merced Land Use kolekciju slika, a 10 puta za Wuhan

		Broj trening slika po klasi		
Deskr.	Kernel	10	50	90
	linearni	$66,78 \pm 2,02$	$75,94 \pm 1,50$	$80,67 \pm 2,06$
FCTD	χ^2	$60, 13 \pm 2, 28$	$78,86 \pm 1,75$	$83, 81 \pm 1, 30$
EGID	RBF	$62,04 \pm 2,02$	$79,56 \pm 2,48$	$85, 14 \pm 1, 04$
	(4.18)	$66,78 \pm 2,02$	$84,65\pm2,48$	$89,24 \pm 1,04$
	linear	$59,06 \pm 2,13$	$75,50 \pm 1,04$	$81,62 \pm 2,83$
BoW	χ^2	$59,27 \pm 1,56$	$77,09\pm0,61$	$82, 10 \pm 1, 64$
DOW	RBF	$59,09 \pm 2,03$	$76,78 \pm 1,04$	$82, 19 \pm 1, 03$
	(4.18)	$61, 61 \pm 2, 03$	$80,04 \pm 1,04$	$87,05 \pm 1,03$

Tabela 6.1: Tačnosti klasifikacije (%) za globalne i lokalne deskriptore na UC Merced Land Use kolekciji slika.

kolekciju slika i u rezultatima su date srednje vrijednosti i standardne devijacije tačnosti klasifikacije.

Kako bi se stekao bolji uvid u performanse globalnih i lokalnih deskriptora kada je u pitanju klasifikacija snimaka dobijenih daljinskom detekcijom uporedno su testirani klasifikatori zasnovani na predstavnicima obje grupe deskriptora. Prezentovane su ukupne tačnosti klasifikacije dobijene kako za globalne tako i za lokalne deskriptore ali i tačnosti klasifikacije koje se postižu za različite klase pokrivača i načina korišćenja zemljišta. Ideja je da se na ovaj način ispita da li su razmatrane reprezentacije slike komplementarne. Ukoliko je to slučaj onda bi se njihovom fuzijom poboljšala tačnost klasifikacije.

Da bi se bolje ocijenile diskriminativne sposobnosti deskriptora na pojedinim klasama, klasifikatori se obučavaju korišćenjem 10% uzoraka po klasi. Tako mali broj trening uzoraka naglašava potrebu za dobrim reprezentacijama slike pošto u slikama mogu postojati značajne varijacije. Sa druge strane, kada bi se koristio veći broj trening uzoraka, rezultati za neke klase bi odražavali otklon korišćene kolekcije slika, a ne diskriminativne sposobnosti deskriptora.

6.3.1 UC Merced Land Use kolekcija slika

Tačnosti klasifikacije na UC Merced Land Use kolekciji slika, za globalne i lokalne deskriptore date su u Tabeli 6.1, za 10, 50 i 90 trening slika po klasi. Iz prezentovanih rezultata, vidi se da je za oba deskriptora generalizovani RBF kernel (4.18) prikladan izbor. Kada je u pitanju reprezentacija slike kao skupa vizuelnih riječi, ovaj rezultat je interesantan zato što je u literaturi [10] prednost data χ^2 kernelu. Međutim, generalizovani RBF kernel sa L_1 metrikom daje dobre rezultate zbog toga što su deskriptori u stvari histogrami pojavljivanja vizuelnih riječi, a L_1 metrika za normalizovane histograme odgovara presjeku histograma.

Poređenjem klasifikatora zasnovanih na EGTD i skupu riječi vidimo da klasifikator zasnovan na EGTD ima nešto bolje performanse za sve testirane veličine trening skupa. Dobre performanse EGTD na čitavoj kolekciji su posljedica prikladnosti deskriptora teksture za klasifikaciju slika dobijenih daljinskom detekcijom i njihove sposobnosti da prikažu obilježja izražena na različitim skalama i/ili orijentacijama. Sa povećanjem broja trening primjera razlika između performansi ova dva deskriptora se smanjuje. Zanimljivo je da je tačnost klasifikacije od 80,04% dobijena za SIFT skup riječi i 50 trening slika bolja od rezultata iz [30, 31] gdje je za 80 trening slika postignuta tačnost od 76,81%. Razlog za ovo poboljšanje je računanje SIFT deskriptora u tačkama pravilno raspoređenim po cijeloj slici umjesto korišćenja detektora interesnih tačaka. Ovaj rezultat je vrlo blizu i rezultatu iz [101] gdje se, takođe sa 80 trening slika, dobija tačnost klasifikacije od 81,67%.

Na Slici 6.2 uporedo su date tačnosti klasifikacije na UC Merced Land Use kolekciji slika po klasama za globalni i lokalni deskriptor. U ovom slučaju klasifikator je obučavan na po 10 slika iz svake klase i testiran na ostatku. Na lijevoj strani grafikona su klase na kojima globalni deskriptor daje bolje performanse. Idući slijeva u desno performanse klasifikatora koji koristi lokalne deskriptore se poboljšavaju i, najzad, na desnoj strani su klase za koje je lokalni deskriptor bolji. Vidimo da je globalni deskriptor bolji na klasama kao što su *teniski tereni, rijetko naseljeno* rezidencijalno i gusto naseljeno rezidencijalno, koje sadrže obilježja koja globalno dominiraju slikom i uglavnom su teksturalno orijentisana. Primjeri slika iz ovih klasa prikazani su u gornjem redu na Slici 6.3. Sa druge strane, lokalni deskriptor je bolji na klasama koje sadrže karakteristične strukture čije prisustvo se koristi za klasifikaciju slika, npr. *bejzbol igralište, zgrade* i *raskrsnica*, što se može vidjeti iz donjeg reda na Slici 6.3. Zanimljiv je primjer klase rijetko naseljeno rezidencijalno koju karakteriše postojanje manjeg broja kuća na slici što ukazuje na to da bi lokalni deskriptori mogli dati dobre rezultate. U razmatranom primjeru bolje performanse se, ipak, dobijaju korišćenjem globalnih deskriptora zbog toga što je objekat, koji postoji na slici, mali i nije adekvatno predstavljen lokalnim deskriptorima, dok je, sa druge strane, tekstura u okolini objekta izražena i dovoljno diskriminativna za klasifikaciju. Ipak, ukupan procenat tačnosti klasifikacije za ovu klasu, postignut korišćenjem globalnog deskriptora, je mali, oko 62%.

6.3.2 Wuhan kolekcija slika

Rezultati poređenja lokalnih i globalnih klasifikatora na Wuhan kolekciji slika dati su u Tabeli 6.2. I u ovom slučaju, za prošireni Gaborov deskriptor teksture najbolji rezultati se dobijaju korišćenjem generalizovanog RBF kernela. Ukoliko je, pak, slika predstavljena lokalnim deskriptorima, najbolje tačnosti klasifikacije se dobijaju korišćenjem generalizovanog RBF kernela, kada se za obučavanje koristi 5, odnosno, 45 slika po klasi, a korišćenjem χ^2 kernela kada se za obučavanje koristi 25 slika po klasi. Ipak, i u poslednjem slučaju, razlika između dobijenih tačnosti klasifikacije iznosi manje od jedne standardne devijacije tako da će u ostalim eksperimentima biti korišćen generalizovani RBF kernel.

Odnos lokalnih i globalnih deskriptora je isti kao za UC Merced Land Use kolekciju slika i očigledne su nešto bolje performanse globalnih deskriptora s tim da se razlika smanjuje sa povećanjem broja trening primjera. Rezultati dobijeni za SIFT skup riječi reprezentaciju i obučavanje klasifikatora sa 5, odnosno, 25 trening slika po klasi odgovaraju rezultatima iz literature [33].

Kako bi se uporedile tačnosti klasifikacije za pojedine klase, klasifikatori su obučavani sa po 5 slika iz svake klase. Na Slici 6.4 uporedno su prikazane tačnosti



Slika 6.2: Poređenje globalnih i lokalnih deskriptora slike na UC Merced Land Use kolekciji. Moguće je uočiti klase pokrivača i načina korišćenja zemljišta za koje pojedini deskriptori imaju bolje performanse.



Slika 6.3: Primjeri slika iz klasa UC Merced Land Use kolekcije slika na kojima se bolji rezultati dobijaju korišćenjem globalnog deskriptora (gornji red) i lokalnog deskriptora (donji red).

		Broj trening slika po klasi		
Deskr.	Kernel	5	25	45
	linearni	$62,75 \pm 2,91$	$78, 36 \pm 1, 65$	$81,27 \pm 2,61$
FCTD	χ^2	$62, 54 \pm 2, 38$	$79,90 \pm 1,04$	$83, 48 \pm 3, 47$
EGID	RBF	$58, 56 \pm 1, 81$	$78,92\pm0,89$	$84,99 \pm 2,39$
	(4.18)	$67, 80 \pm 1, 83$	$83,74\pm0,95$	$87,71 \pm 2,40$
	linear	$55, 50 \pm 1, 68$	$73,98\pm1,71$	$79,86 \pm 2,28$
BoW	χ^2	$58,99 \pm 1,33$	$79, 35 \pm 1, 27$	$85,57 \pm 2,19$
DOW	RBF	$57, 48 \pm 1, 36$	$76,77 \pm 1,51$	$82, 19 \pm 2, 03$
	(4.18)	$59, 56 \pm 1, 39$	$78,85 \pm 1,62$	$ 86,43\pm1,44 $

Tabela 6.2: Tačnosti klasifikacije (%) za globalne i lokalne deskriptore na Wuhan kolekciji slika.

klasifikacije za pojedine klase iz Wuhan kolekcije slika. Na lijevoj strani grafika su klase na kojima je globalni deskriptor bolji, dok se idući u desno performanse lokalnih deskriptora postepeno poboljšavaju. Ipak, za razliku od UC Merced Land Use kolekcije slika, ovdje postoji veći broj klasa na kojima globalni deskriptori daju bolje rezultate.

Najbolje performanse globalni deskriptor postiže za klase *pustinja, aerodrom* i *livada* čiji su primjeri prikazani u gornjem redu na Slici 6.5. Iz ovih primjera se vidi da se radi o slikama sa izraženom teksturom, kao što su pustinja i livada, kao i sa dominantnim obilježjima na nivou čitave slike, kao što je npr. aerodrom. Sa druge strane, najbolje performanse korišćenjem lokalnih deskriptora postignute su za klase *park, stadion* i *parking.* Iz primjera u donjem redu na Slici 6.5 se vidi da se radi o vizuelno nehomogenim kategorijama sa karakterističnim objektima koji ukazuju na način korišćenja zemljišta.

6.4 Eksperimentalni rezultati za hijerarhijski klasifikator

Da bi se konstruisao klasifikator koji koristi kombinaciju globalnih i lokalnih deskriptora prvo se formira međureprezentacija slika, kao što je opisano u odjeljku 6.2. Na osnovu rezultata iz prethodnog odjeljka, kao klasifikatori na nivou 0 koriste se SVM sa generalizovanim RBF kernelom. Polazeći od ocjena pouzdanosti, dobijenih na izlazima SVM, posteriorne vjerovatnoće se izračunavaju pomoću (6.2) sa $\beta = 2$. Kako bi se izbjegli potencijalni problemi sa generalizacijom pri formiranju međureprezentacije trening skup se dijeli na P = 5 dijelova, kao što je opisano u odjeljku 6.2. Konačno, međureprezentacija slike se dobija konkatenacijom vektora posteriornih vjerovatnoća dobijenih na izlazima klasifikatora na nivou 0. Pošto se na nivou 0 koriste dva deskriptora, dimenzionalnost međureprezentacije za klasično hijerarhijsko dodavanje klasifikatora je u ovom slučaju 2C, gdje je C broj klasa. Ukoliko se koriste stackingC varijante dimenzionalnost međureprezentacije je dva.

Tako dobijena međureprezentacija se dovodi na ulaz klasifikatora na nivou 1. Evaluirani su sljedeći klasifikatori: linearna regresija, regularizovana linearna regre-



Slika 6.4: Poređenje globalnih i lokalnih deskriptora slike na Wuhan kolekciji. Moguće je uočiti klase pokrivača i načina korišćenja zemljišta za koje pojedini deskriptori imaju bolje performanse.

sija, linearna SVM, RBF SVM i regularizovana regresija sa RBF kernelom, kao i njihove stackingC varijante. Lasso regresija [102] nije donijela poboljšanja pa je izostavljena. Klasifikacija u više klasa na nivou 1 se postiže obučavanjem klasifikatora tipa jedan-protiv-svih za svaku od klasa i svrstavanjem testne slike u klasu koja odgovara klasifikatoru sa najvišom ocjenom pouzdanosti.

6.4.1 UC Merced Land Use kolekcija slika

Pošto UC Merced Land Use kolekcija slika ima 21 klasu načina korišćenja zemljišta, dobijena međureprezentacija je 42-dimenzionalna za klasično hijerarhijsko dodavanje klasifikatora, a dvodimenzionalna za stackingC. Dobijeni rezultati su dati u Tabeli 6.3. Tačnosti postignute konkatenacijom deskriptora slika su date u prvom redu. To je najjednostavniji način fuzije deskriptora i u tom slučaju metaklasifikator se ne koristi. Ostali rezultati prezentovani u tabeli dobijeni su upotrebom različitih metaklasifikatora. Generalno, korišćenjem metaklasifikatora se dobijaju bolji rezultati u odnosu na konkatenaciju deskriptora. Međutim, za 10 trening slika po klasi ovo nije u potpunosti tačno. Mogući razlog za ovo je obučavanje klasifikatora na nivou 0 na malom broj trening slika. Ipak, metoda koja kao metaklasifikator koristi regularizovanu linearnu regresiju C u ovom slučaju daje najbolji rezultat. Za 50 i 90 trening slika najbolji rezultat daje RBF regularizovana regresija, ali tačnosti klasifikacije za regularizovanu linearnu regresiju C su unutar jedne standardne devijacije. Štaviše, zanimljivo je primjetiti da su tačnosti klasifikacije postignute svim



Slika 6.5: Primjeri slika iz Wuhan kolekcije koji pripadaju klasama na kojima se bolji rezultati dobijaju korišćenjem globalnog deskriptora (gornji red) i lokalnog deskriptora (donji red).

Tabela 6.3: Tačnosti klasifikacije (%) za različite metaklasifikatore i njihove stackingC varijante (označene sa C) za fuziju globalnih i lokalnih deskriptora na UC Merced Land Use kolekciji slika.

	Broj trening slika po klasi		
Metaklasifikator	10	50	90
Konkatenacija	$70, 31 \pm 1, 72$	$87,28 \pm 1,22$	$89,71 \pm 1,74$
Linearna SVM	$68,94 \pm 2,71$	$88,51 \pm 1,31$	$92,38\pm0,95$
Linearna SVM C	$69, 64 \pm 2, 63$	$88,50 \pm 1,38$	$92,48\pm0,92$
Linearna regr.	$69,05 \pm 2,52$	$87, 89 \pm 2, 03$	$92,00 \pm 1,03$
Linearna regr. C	$69, 31 \pm 2, 68$	$87, 89 \pm 2, 03$	$92,00\pm1,03$
Reg. lin. regr.	$70, 10 \pm 2, 85$	$88,40 \pm 1,77$	$92, 10 \pm 1, 20$
Reg. lin. regr. C	$72, 12 \pm 2, 08$	$88, 32 \pm 1, 68$	$92,57 \pm 1,15$
RBF SVM	$69, 13 \pm 2, 77$	$88,50 \pm 1,33$	$92,76\pm0,98$
RBF SVM C	$68, 24 \pm 2, 13$	$87,50 \pm 1,39$	$92,86\pm1,47$
RBF reg. regr.	$70, 18 \pm 2, 91$	$88, 82 \pm 1, 27$	$92, 95 \pm 0, 62$
RBF reg. regr. C	$70,57 \pm 2,58$	$88,25 \pm 1,80$	$92,86 \pm 0,75$

testiranim metaklasifikatorima vrlo slične.

Rezultati dati u Tabeli 4.3 pokazuju da se redukcijom dimenzionalnosti proširenog Gaborovog deskriptora teksture sa 240 na 60 performanse klasifikatora smanjuju svega za oko 2%. Ovaj rezultat dobijen je obučavanjem klasifikatora sa 80 trening slika po klasi i testiranjem na preostalim. Potrebno je provjeriti da li ovaj odnos performansi ostaje isti za različite veličine trening skupova, kao i na koji način redukcija dimenzionalnosti proširenog Gaborovog deskriptora teksture utiče na tačnost klasifikacije hijerarhijskog klasifikatora.

	Broj trening slika po klasi		
Metaklasifikator	10	50	90
EGTD komp. PCA	$61, 89 \pm 1, 86$	$81,07 \pm 1,20$	$85,90 \pm 1,29$
Konkatenacija	$65, 50 \pm 1, 76$	$84,48\pm0,87$	$88,67 \pm 1,66$
Linearna SVM	$66, 36 \pm 2, 42$	$87,24 \pm 0,89$	$90,67 \pm 1,24$
Linearna SVM C	$66, 62 \pm 3, 23$	$87,52 \pm 0,84$	$90,86 \pm 0,71$
Linearna regr.	$65,99 \pm 2,81$	$86,93 \pm 1,11$	$90,95 \pm 0,75$
Linearna regr. C	$66, 25 \pm 2, 84$	$86,93\pm1,11$	$90,95 \pm 0,75$
Reg. lin. regr.	$66, 86 \pm 3, 34$	$87,28 \pm 1,17$	$91,05 \pm 0,62$
Reg. lin. regr. C	$68, 10 \pm 2, 67$	$87,45 \pm 0,96$	$90,95\pm0,34$
RBF SVM	$66, 50 \pm 2, 61$	$87, 12 \pm 0, 77$	$90,48 \pm 0,89$
RBF SVM C	$65,47 \pm 2,80$	$86,93 \pm 1,16$	$90,76 \pm 1,41$
RBF reg. regr.	$67, 50 \pm 2, 71$	$87, 90 \pm 0, 23$	$91,05 \pm 0,85$
RBF reg. regr. C	$67,80\pm 2,59$	$87, 31 \pm 1, 03$	$ \hspace{0.1cm} 91 \hspace{0.1cm}, 14 \pm 0 \hspace{0.1cm}, 64 $

Tabela 6.4: Tačnosti klasifikacije (%) za različite metaklasifikatore i njihove stackingC varijante (označene sa C) za fuziju globalnih deskriptora redukovane dimenzionalnosti i lokalnih deskriptora na UC Merced Land Use kolekciji slika.

Rezultati dobijeni fuzijom proširenog Gaborovog deskriptora teksture redukovane dimenzionalnosti i lokalnog deskriptora, za UC Merced Land Use kolekciju slika, dati su u Tabeli 6.4. Može se primjetiti da su postignute tačnosti slične, bez obzira na korišćeni metaklasifikator. Tačnost klasifikacije postignuta korišćenjem samo EGTD redukovane dimenzionalnosti se snižava za 3-5%. Ova razlika zavisi od broja trening slika i smanjuje se sa njegovim povećanjem. Nakon fuzije deskriptora razlika između tačnosti klasifikacije postignutih deskriptorima pune, odnosno, redukovane dimenzionalnosti se smanjuje na 1-4%, ponovo u zavisnosti od broja trening slika. U ovom slučaju se vidi prednost korišćenja metaklasifikatora nad jednostavnom konkatenacijom deskriptora. Kada u deskriptoru dobijenom konkatenacijom figuriše EGTD redukovane dimenzionalnosti, razlike tačnosti klasifikacije u odnosu na slučaj kada se koristi EGTD pune dimenzionalnosti su veće nego kada se koristi metaklasifikator.

Poređenja radi, u Tabeli 6.5 dati su rezultati dobijeni korišćenjem globalnih, odnosno, lokalnih deskriptora, njihovom hijerarhijskom kombinacijom pomoću RBF regularizovane regresije C, te rezultati iz literature postignuti na ovoj kolekciji. Pošto su rezultati u literaturi postignuti za panhromatske slike i rezultati dati u ovoj tabeli su dobijeni pod istim uslovima. Ovi rezultati su postignuti obučavanjem klasifikatora na po 80 trening slika iz svake klase i testiranjem na ostalim i ponavljanjem ciklusa obučavanja i testiranja pet puta¹.

Iz ovih rezultata se vidi da se, za slučaj korišćenja samo jednog deskriptora za opis slike, najbolji rezultati dobijaju kada se koristi prošireni Gaborov deskriptor teksture. Takođe, rezultat dobijen pomoću SIFT skupa riječi je bolji od ekvivalentnih pristupa u [30] i [101]. Razlozi su računanje SIFT deskriptora u pravilno raspoređenim tačkama, a ne u interesnim tačkama određenim detektorom kao u [30] te u korišćenju prozora veličina 8×8 i 16×16 piksela, dok se u [101] SIFT deskriptori

¹Za rezultate iz literature nisu uvijek dostupne standardne devijacije tačnosti klasifikacije

Metod	Tačnost (%)
EGTD	$87,57 \pm 1,68$
SIFT skup riječi	$84, 19 \pm 1, 53$
NN-STSIM 4×4	$82, 86 \pm 1, 57$
Hijerarhijski klasifikator	$ 92,48\pm0,36 $
SIFT skup riječi [30]	76, 81
BOW+SCK $[30]$	77,71
SPCK++[31]	77, 38
SIFT skup riječi [101]	$81,67 \pm 1,23$
BRSP [103]	77, 8
BPP [104]	86, 1

Tabela 6.5: Poređenje tačnosti klasifikacije za UC Merced Land Use kolekciju slika.

računaju samo u prozorima veličine 16×16 piksela. Strukturalnim deskriptorom teksture se dobija slična tačnost klasifikacije kao za SIFT deskriptor. Konačno, hijerarhijskim kombinovanjem klasifikatora se rezultati dodatno poboljšavaju i to je trenutno najbolja tačnost klasifikacije postignuta na ovoj kolekciji.

Na Slici 6.6 prikazana je matrica konfuzija za UC Merced Land Use kolekciju slika dobijena hijerarhijskom kombinacijom globalnih i lokalnih deskriptora pomoću regularizovane regresije C. Tačnosti klasifikacije dobijene za ostale metaklasifikatore su vrlo bliske. Klasifikator je obučavan korišćenjem po 80 slika iz svake klase. Iz matrice konfuzija se vidi da se broj konfuzija značajno smanjio u odnosu na korišćenje samo jednog deskriptora. Najizraženije konfuzije su između klasa *cisterne* i *tereni za bejzbol.* Za sve ostale parove klasa konfuzije se javljaju u manje od 10% slučajeva.

6.4.2 Wuhan kolekcija slika

Slike u Wuhan kolekciji su klasifikovane u 19 klasa načina korišćenja zemljišta tako da se u klasičnom stackingu hijerarhijskom dodavanju klasfikatora dobija 38dimenzionalna međureprezentacija, dok stackingC ponovo koristi dvodimenzionalnu međureprezentaciju. Dobijeni rezultati, kada se koristi hijerarhijska klasifikacija, dati su u Tabeli 6.6. U prvom redu su ponovo tačnosti klasifikacije dobijene konkatenacijom deskriptora, a u ostalim su rezultati dobijeni korišćenjem metaklasifikatora. Ovi rezultati su vrlo slični rezultatima dobijenim za UC Merced Land Use kolekciju slika. Za pet trening slika po klasi najbolje performanse opet pokazuje regularizovana linearna regresija C, a ostali metaklasifikatori su za nijansu lošiji od konkatenacije deskriptora. Kada je broj trening slika veći, svi metaklasifikatori nadmašuju konkatenaciju deskriptora. Za 25 trening slika najbolja je regularizovana linearna regresija sa RBF kernelom, a za 45 trening slika RBF SVM. Takođe, tačnosti klasifikacije dobijene korišćenjem različitih metaklasifikatora se, ponovo, vrlo malo razlikuju.

Rezultati dati u Tabeli 4.4 ukazuju na to da se, redukcijom dimenzionalnosti proširenog Gaborovog deskriptora teksture, primjenom PCA na pojedine komponente deskriptora, ne smanjuje tačnost klasifikacije Wuhan kolekcije slika. Kada


Slika 6.6: Matrica konfuzija za UC Merced Land Use kolekciju slika dobijena hijerarhijskom kombinacijom globalnih i lokalnih deskriptora.

Tabela 6.6: Tačnosti klasifikacije (%) za različite metaklasifikatore i njihove stackingC varijante (označene sa C) za fuziju globalnih i lokalnih deskriptora na Wuhan kolekciji slika.

	Broj trening slika po klasi		
Metaklasifikator	5	25	45
Konkatenacija	$70, 10 \pm 3, 23$	$85,68 \pm 1,75$	$90,20 \pm 1,45$
Linearna SVM	$68,40 \pm 3,28$	$88,01 \pm 1,72$	$92,28 \pm 1,52$
Linearna SVM C	$69,57 \pm 2,33$	$87,91 \pm 1,09$	$92, 36 \pm 1, 72$
Linearna regr.	$67, 44 \pm 2, 26$	$86, 12 \pm 1, 58$	$90,78 \pm 2,00$
Linearna regr. C	$69,67 \pm 1,69$	$86, 16 \pm 1, 62$	$90,78 \pm 2,00$
Reg. lin. regr.	$67, 53 \pm 2, 57$	$87,96 \pm 1,44$	$92, 22 \pm 2, 37$
Reg. lin. regr. C	$74, 01 \pm 2, 23$	$87,89\pm1,37$	$92,24 \pm 2,18$
RBF SVM	$68,02 \pm 3,42$	$88,41 \pm 1,85$	$ \hspace{0.1em} 92, 62 \pm 1, 64 $
RBF SVM C	$68, 10 \pm 3, 39$	$86,87 \pm 1,89$	$92, 31 \pm 1, 94$
RBF reg. regr.	$69,54 \pm 2,57$	$88, 52 \pm 1, 47$	$91,94 \pm 1,67$
RBF reg. regr. C	$71,94 \pm 2,69$	$87,97 \pm 1,70$	$91,62 \pm 2,16$

se deskriptor redukovane dimenzionalnosti koristi u prvom nivou hijerarhijske fuzije sa lokalnim deskriptorom dobijaju se tačnosti klasifikacije date u Tabeli 6.7. Iz rezultata se vidi da tačnosti klasifikacije relativno malo variraju za različite metaklasifikatore, te ne opadaju ili malo opadaju u odnosu na tačnosti dobijene korišćenjem punog proširenog Gaborovog deskriptora teksture. Najveća dobijena razlika tačno-

	Broj trening slika po klasi		
Metaklasifikator	5	25	45
EGTD komp. PCA	$66, 33 \pm 2, 69$	$84,47 \pm 1,07$	$87,50 \pm 2,59$
Konkatenacija	$63, 20 \pm 1, 50$	$83,59\pm1,39$	$89,04 \pm 1,88$
Linearna SVM	$67, 39 \pm 3, 55$	$88,29\pm1,08$	$90,79 \pm 3,10$
Linearna SVM C	$66, 45 \pm 2, 98$	$88,26 \pm 1,10$	$91, 20 \pm 2, 33$
Linearna regr.	$64,70 \pm 3,12$	$86,77\pm0,97$	$89,62 \pm 2,34$
Linearna regr. C	$66, 69 \pm 3, 12$	$86,76 \pm 0,98$	$89,62 \pm 2,34$
Reg. lin. regr.	$66,47 \pm 3,17$	$88,37 \pm 1,28$	$90,72 \pm 2,71$
Reg. lin. regr. C	$72, 38 \pm 2, 42$	$88,42\pm0,95$	$90,81 \pm 2,77$
RBF SVM	$67, 76 \pm 3, 61$	$88,21 \pm 1,11$	$91, 11 \pm 3, 09$
RBF SVM C	$65, 28 \pm 3, 74$	$87,48 \pm 0,82$	$89,97 \pm 2,21$
RBF reg. regr.	$68,92 \pm 3,17$	$88, 61 \pm 1, 09$	$91, 37 \pm 2, 99$
RBF reg. regr. C	$70,93 \pm 2,31$	$88,06 \pm 0,98$	$90,93 \pm 3,10$

Tabela 6.7: Tačnosti klasifikacije (%) za različite metaklasifikatore za fuziju globalnih deskriptora redukovane dimenzionalnosti i lokalnih deskriptora na Wuhan kolekciji slika.

sti klasifikacije od 2% se ponovo dobija za male trening skupove sa samo 5 slika po kategoriji. Sa povećanjem broja trening slika razlika tačnosti klasifikacije nestaje. Međutim, čak i najveća razlika tačnosti klasifikacije je manja od jedne standardne devijacije tačnosti. Sa druge strane, tačnosti klasifikacije dobijene konkatenacijom deskriptora nakon redukcije dimenzionalnosti su više umanjene u odnosu na tačnosti dobijenih konkatenacijom deskriptora pune dimenzionalnosti, što ukazuje na prednost hijerarhijskog kombinovanja klasifikatora.

U Tabeli 6.8 uporedo su prikazani rezultati dobijeni za Wuhan kolekciju slika korišćenjem različitih algoritama razmatranih u ovoj disertaciji i rezultati iz literature. U svim slučajevima deskriptori su računati za panhromatske slike, a klasifikatori su obučavani na po 25 slika iz svake klase i testirani na ostatku. Ciklus obučavanja i testiranja je ponovljen 10 puta.

Iz rezultata se vidi da u ovom slučaju korišćenje prozora različitih veličina ne poboljšava tačnost klasifikacije. Vjerovatan razlog za ovo leži u tome što su dimenzije slika iz Wuhan kolekcije 600×600 piksela pa se manjim prozorima, veličine 8×8 piksela, ne može dobiti dovoljno diskriminativna informacija. Slična tačnost klasifikacije se dobija i strukturalnim deskriptorom teksture. Najbolji rezultat se i u ovom slučaju dobija hijerarhijskom kombinacijom klasifikatora i to je trenutno najbolja tačnost klasifikacije postignuta na ovoj kolekciji slika. Dobijeni rezultat je bolji od rezultata dobijenog konkatenacijom SIFT deskriptora, LBP deskriptora teksture i kolor-histograma, iako je u ovom drugom sadržana i informacija o boji. Razlog za ovo je u slabijoj generalizaciji visokodimenzionalnog deskriptora dobijenog konkatenacijom.

Na Slici 6.7 prikazana je matrica konfuzija za Wuhan kolekciju slika dobijena hijerarhijskom kombinacijom globalnih i lokalnih deskriptora. Kao metaklasifikator iskorišćena je SVM sa RBF kernelom. Za obučavanje je korišćeno po 25 slika iz svake klase. Broj konfuzija je značajno smanjen u odnosu na slučajeve u kojima se

Metod	Tačnost (%)
EGTD	$83,74 \pm 0,95$
SIFT skup riječi	$79,35 \pm 1,27$
NN-STSIM 4×4	$78,24 \pm 1,44$
Hijerarhijski klasifikator	$88, 52 \pm 1, 47$
SIFT skup riječi [33]	79, 1
Konkatenacija (SIFT+LBP+kolor-hist.) [33]	84,8

Tabela 6.8: Poređenje tačnosti klasifikacije za Wuhan kolekciju slika.



Slika 6.7: Matrica konfuzije za Wuhan kolekciju slika dobijena hijerarhijskom kombinacijom globalnih i lokalnih deskriptora.

koristi samo jedan deskriptor. Najizraženije su konfuzije između klasa *rezidencijalno zemljište* i *poslovne zgrade*. Sve ostale konfuzije se javljaju u manje od 10% slučajeva.

6.4.3 Detekcija urbanih područja

Eksperimenti opisani do sada su orijentisani na klasifikaciju blokova izdvojenih sa aero-snimaka pri čemu se ne uzima u obzir njihova lokacija na visokorezolucionom snimku, a sami blokovi se mogu jednoznačno svrstati u semantičke klase. U ovom eksperimentu se razmatra mogućnost korišćenja hijerarhijske fuzije globalnih i lokalnih deskriptora za detekciju urbanih područja na čitavom visokorezolucionom snimku. Detekcija urbanih područja je značajan zadatak u analizi snimaka dobijenih daljinskom detekcijom jer predstavlja osnovu za ažuriranje geografskih informacionih sistema čime se omogućava praćenje i planiranje urbanog razvoja. Kao i ostali zadaci u analizi snimaka dobijenih daljinskom detekcijom i detekcija urbanih područja predstavlja izazov zbog obima i dinamičke prirode podataka koje treba analizirati. Zbog toga su metodi za automatsku detekciju urbanih područja predmet aktivnog istraživanja u prethodnih desetak godina.

Algoritmi predloženi u literaturi se mogu podijeliti u dvije grupe. U prvoj grupi su algoritmi zasnovani na nadgledanoj klasifikaciji, kao na primjer, [105] i [95]. Algoritmi iz druge grupe ne zahtijevaju obučavanje, već su zasnovani na vizuelnim obilježjima karakterističnim za urbana područja, kao što su lokalne interesne tačke [106] ili spojevi ivica [107]. U ovoj disertaciji predmet proučavanja je nadgledana klasifikacija slika dobijenih daljinskom detekcijom u klase pokrivača i načina korišćenja zemljišta pa će ona biti iskorišćena i za detekciju urbanih područja.

Pošto su klasifikatori predloženi u ovoj disertaciji orijentisani na klasifikaciju blokova, polazni visokorezolucioni snimak se dijeli na blokove određene veličine koji se izdvajaju sa pravilno raspoređenih lokacija. Svaki od blokova se posebno klasifikuje kao blok urbanog ili neurbanog područja, što znači da se svi pikseli koji pripadaju tom bloku klasifikuju na isti način. Pošto se blokovi, u opštem slučaju, mogu preklapati, isti piksel se klasifikuje više puta zbog pripadnosti različitim blokovima. Konačna odluka o klasifikaciji se donosi glasanjem – piksel se klasifikuje u klasu u koju je klasifikovano najviše blokova kojima taj piksel pripada. Nakon ove faze, dobijena tematska mapa ima "blokovsku" strukturu pa se dodatno obrađuje operacijama matematičke morfologije kako bi se dobile "glatke" granice regiona.

Ovaj algoritam je testiran na aero-snimku područja Novog Grada i ušća rijeke Sane u Unu opisanom u odjeljku 2.5. Urbanim područjem se, prema [105], u ovom eksperimentu smatra kompaktna grupa građevinskih objekata, što obuhvata klase načina korišćenja zemljišta naseljeno mjesto, okućnice i industrijski kompleks. Klasifikator je obučavan korišćenjem blokova izdvojenih iz dijela snimka prikazanog na Slici 6.8, a testiran na dijelu snimka prikazanom na Slici 6.9, zajedno sa manuelno označenim urbanim i neurbanim područjima. Testni dio snimka je dimenzija 2000×4500 piksela. Veličina bloka koja se koristi je 128×128 piksela što odgovara prostornoj veličini od $64m \times 64m$. Ova veličina je izabrana u skladu sa rezultatima iz [8] gdje je pokazano da je, za klasifikaciju blokova izdvojenih iz aero-snimka u klase načina korišćenja zemljišta na nivou II Andersonovog sistema za klasifikaciju, potrebno da blokovi budu veličine bar $60m \times 60m$.

Svaki blok se predstavlja globalnim i lokalnim deskriptorima. Globalna reprezentacija blokova se dobija korišćenjem oponentnog proširenog Gaborovog deskriptora teksture, kao što je opisano u odjeljku 4.5.1. Kao lokalni deskriptori upotrebljeni su SIFT deskriptori izračunati za svaki spektralni opseg posebno i spojeni u jedinstven deskriptor. Tako dobijeni deskriptori se vektorski kvantizuju i slika se predstavlja kao skup vizuelnih riječi. Na nivou 0 klasifikatori su SVM sa kernelom (4.18), a kao metaklasifikator se koristi regularizovana linearna regresija. Za obučavanje klasifikatora se koristi po 100 slučajno izabranih blokova iz urbanog i neurbanog područja. Pri klasifikaciji blokovi se izdvajaju sa korakom od 32 piksela. Postobrada se sastoji od zatvaranja i otvaranja strukturnim elementom oblika diska i radijusa 32 piksela.

U Tabeli 6.9 date su procentualne vrijednosti broja tačno klasifikovanih i lažno pozitivnih piksela urbanog i neurbanog područja. Prosječan procenat tačno klasifikovanih piksela na čitavoj slici iznosi 89,50%. Dobijeni rezultati odgovaraju aktuelnom stanju istraživanja u ovoj oblasti [95, 105, 106, 107].



Slika 6.8: Dio aero-snimka područja Novog Grada sa kojeg se izdvajaju blokovi za obučavanje detektora urbanog područja.



Slika 6.9: Dio aero-snimka područja Novog Grada na kojem se testira detektor urbanih područja. Granice područja su određene manuelno i urbano područje je osjenčeno crvenom, a neurbano zelenom bojom.

Klasa	Tačnost (%)	Lažno pozitivni (%)
Urbano područje	91,25	12,25
Neurbano područje	87,75	8,75

Tabela 6.9: Tačnosti detekcije urbanog područja.



(a) Primjer detekcije urbanih područja prije postobrade.



(b) Primjer detekcije urbanih područja poslije postobrade.

Slika 6.10: Primjeri detekcije urbanog područja na dijelu aero-snimka područja Novog Grada prije i poslije postobrade matematičkom morfologijom. Urbano područje je osjenčeno crvenom, a neurbano zelenom bojom.

Primjeri prije i poslije postobrade korišćenjem matematičke morfologije su prikazani na Slici 6.10. Uticaj postobrade na tačnost klasifikacije je zanemarljiv, ali je uključena zbog toga što su dobijene granice regiona bliže ljudskom poimanju. Na Slici 6.11 označeni su regioni u kojima su pikseli pogrešno klasifikovani. Sa jedne strane, predloženim algoritmom nisu detektovani neki rijetko naseljeni regioni i rubovi naseljenih regiona u kojima se nalazi mali broj piksela koji pripadaju građevinskim objektima, a sa druge strane, regioni uz puteve su često pogrešno klasifikovani kao urbano područje. Ipak, može se zaključiti da predloženi algoritam daje dobre rezultate u ovom eksperimentu iako nisu korišćena rješenja prilagođena za rješavanje problema detekcije urbanih područja



Slika 6.11: Primjer detekcije urbanog područja na dijelu aero-snimka područja Novog Grada. Crvenom i zelenom bojom su osjenčena tačno detektovana urbana i neurbana područja. Žutom bojom su osjenčena lažno pozitivna, a plavom lažno negativna urbana područja.

6.5 Diskusija i zaključci

Poređenjem tačnosti klasifikacije za pojedine klase, datih na Slikama 6.2 i 6.4 za obje kolekcije slika, može se doći do zaključka da se razmatranim globalnim i lokalnim deskriptorima opisuju različite informacije sadržane na slikama. Iz ovih rezultata se vidi da, u obje kolekcije slika, postoje klase za koje jedan od deskriptora rezultuje izraženo boljim performansama. Ova činjenica ukazuje na to da bi se fuzijom deskriptora mogle poboljšati performanse klasifikatora.

Najjednostavniji način fuzije deskriptora je njihova konkatenacija. Na ovaj način se dobija jedinstven deskriptor koji opisuje sliku i predstavlja ulaz u algoritam za klasifikaciju. Međutim, jednostavna konkatenacija deskriptora bi rezultovala novim deskriptorom visoke dimenzionalnosti što može nepovoljno uticati na brzinu obučavanja i sposobnost generalizacije klasifikatora. Takođe, konkatenacija deskriptora nije izvodljiva u slučajevima kada se za deskriptore koriste različite mjere sličnosti. Zbog toga je predloženo formiranje ansambla klasifikatora njihovim hijerarhijskim kombinovanjem.

Kao globalni deskriptor slika korišćen je prošireni Gaborov deskriptor teksture, a kao lokalni SIFT skup riječi. Klasifikatori na prvom nivou za oba deskriptora su SVM sa generalizovanim RBF kernelom (4.18). Kao metaklasifikatori na drugom nivou testirani su linearna regresija, regularizovana linearna regresija, linearna SVM, RBF regularizovana regresija, RBF SVM, te njihove stackingC varijante.

Na osnovu prikazanih eksperimentalnih rezultata može se primjetiti sljedeće:

 Nema potrebe da se u metaklasifikatorima koriste kerneli. Linearni klasifikatori imaju podjednako dobre performanse i manju složenost. Ovaj zaključak je intuitivno prihvatljiv zato što se međureprezentacija sastoji od ocjena pouzdanosti klasifikatora nivoa 0. Linearni metaklasifikatori kombinuju ove ocjene pozdanosti dajući veće težine izlazima onih klasifikatora nivoa 0 koji sa većom pouzdanošću predviđaju tačnu klasu. Sa druge strane, klasifikatori sa RBF kernelima zavise od Mahalanobisove udaljenosti između uzoraka. Međutim, nije jasno kakva bi bila interpretacija Mahalanobisove udaljenosti u prostoru ocjena pouzdanosti.

- Pošto je dimenzionalnost međudeskriptora relativno niska, nema potrebe da se kao metaklasifikatori koriste SVM. U ovom konkretnom slučaju, dimenzionalnost međudeskriptora je 2C = 42 i vidimo da klasifikatori zasnovani na regularizovanoj linearnoj regresiji imaju vrlo dobre performanse.
- Rezultati dobijeni za stackingC sugerišu da se metaklasifikatori mogu obučavati korišćenjem međudeskriptora čija je dimenzionalnost jednaka broju tipova različitih deskriptora, što je u ovom slučaju dva. Na primjer, klasifikatori zasnovani na regularizovanoj linearnoj regresiji sa dvodimenzionalnim deskriptorima imaju vrlo nisku složenost.
- Kada se koristi redukcija dimenzionalnosti proširenog Gaborovog deskriptora teksture tačnost klasifikacije hijerarhijskog klasifikatora se smanjuje za nekoliko procenata. Smanjenje je izraženije kada se klasifikatori obučavaju na malim trening skupovima, dok se sa porastom broja trening primjera tačnost približava tačnosti klasifikatora obučenog korišćenjem deskriptora pune dimenzionalnosti.

Iako se u opisanom eksperimentu konkatenacijom lokalnih i globalnih deskriptora dobijaju tačnosti klasifikacije uporedive sa tačnostima dobijenim korišćenjem metaklasifikatora, prednosti dodavanja klasifikatora su niža dimenzionalnost međudeskriptora i mogućnost kombinovanja klasifikatora koji koriste različite mjere sličnosti. Naime, kao što je pomenuto, dimenzionalnost međudeskriptora je jednaka MC, odnosno, M ukoliko se koristi stackingC, gdje je M broj deskriptora na nivou 0, a Cbroj klasa. Ove dimenzionalnosti su tipično znatno manje od zbira dimenzionalnosti deskriptora korišćenih na nivou 0. Primjera radi, u opisanim eksperimentima deskriptor dobijen konkatenacijom pojedinačnih deskriptora pune dimenzionalnosti je 1240-dimenzionalan, dok je dimenzionalnost međudeskriptora 42 za UC Merced Land Use, a 38 za Wuhan kolekciju slika, odnosno, dva ako se koristi stackingC. Niža dimenzionalnost deskriptora rezultuje jednostavnijim klasifikatorom i omogućava bolju generalizaciju.

Pored toga, konkatenacijom se formira novi deskriptor u kojem se sve komponente (osnovni deskriptori) tretiraju na isti način. Ovo znači da se za određivanje sličnosti slika u tom slučaju koristi samo jedna mjera sličnosti koja se primjenjuje na novi deskriptor kao cjelinu. Ovaj pristup može biti nepovoljan u slučaju kada se za izračunavanje sličnosti slika predstavljenih osnovnim deskriptorima koriste različite mjere sličnosti. Opisani problem ne postoji kod nadovezivanja klasifikatora zato što se na prvom nivou za svaki deskriptor koristi odgovarajuća mjera sličnosti i njihovi rezultati se kombinuju na drugom nivou.

Za fuziju različitih tipova deskriptora predložen je pristup zasnovan na formiranju međureprezentacije slika nadovezivanjem ocjena pouzdanosti klasifikatora zasnovanih na pojedinim tipovima deskriptora. Pokazano je da se, korišćenjem jednostavnih metaklasifikatora kao što je stackingC sa linearnom ili regularizovanom linearnom regresijom, mogu dobiti dobri rezultati, kako sa stanovišta tačnosti klasifikacije tako i sa stanovišta složenosti klasifikatora. Iako je ovaj pristup ilustrovan na primjeru fuzije globalnih i lokalnih deskriptora, važno je napomenuti da je on generalan i da nije ograničen na deskriptore i klasifikatore razmatrane u ovom poglavlju. Opisani pristup je primjenljiv na različite deskriptore i klasifikatore pod uslovom da su raspoložive ocjene pouzdanosti po klasama.

Glava 7 Zaključak

U ovoj disertaciji prezentovani su rezultati istraživanja u oblasti semantičke klasifikacije slika koje predstavljaju blokove visokorezolucionih snimaka dobijenih daljinskom detekcijom. Ciljne semantičke klase predstavljaju klase pokrivača i načina korišćenja zemljišta. Ovaj zadatak je izazovan zato što na visokorezolucionim snimcima, zbog izraženih detalja, spektralne vrijednosti piksela iz pojedinih klasa pokrivača i načina korišćenja zemljišta više nisu homogene, kao što je to slučaj kod niskorezolucionih snimaka. Zbog toga za diskriminaciju klasa pokrivača i načina korišćenja zemljišta na visokorezolucionim snimcima više nije dovoljna samo spektralna informacija koja predstavlja osnov klasifikacije niskorezolucionih snimaka. Kako bi se ovaj problem riješio u proces klasifikacije je potrebno uključiti i druge vizuelne elemente kao što su tekstura, oblik, struktura, prostorni odnosi između objekata i regiona, itd. Kao posljedica, kriterijumi za klasifikaciju slika zavise od klasa u koje slike pripadaju. Na primjer, neke klase je moguće razlikovati samo na osnovu teksture, za neke je potrebno uključiti i prostorni raspored tekstura, a u nekim slučajevima klase su definisane prisustvom određenih objekata na slici.

Pošto je tekstura jedno od najdiskriminativnijih vizuelnih obilježja visokorezolucionih snimaka, u ovoj disertaciji je predložen prošireni Gaborov deskriptor teksture koji koristi korelacije između podopsega Gaborove transformacije kako bi se dobila bolja reprezentacija struktura na slici izraženih na različitim skalama i orijentacijama. Pored deskriptora za panhromatske slike, predložene su i dvije verzije deskriptora za slike u boji. Eksperimentalno je pokazano je da se klasifikatorom koji koristi reprezentaciju slike proširenim Gaborovim deskriptorom teksture postižu rezultati koji nadmašuju najbolje rezultate iz literature.

Pošto su predloženi deskriptori visokodimenzionalni, razmatrani su i različiti metodi za redukciju njihovih dimenzionalnosti. Pokazano je da je moguće značajno smanjiti dimenzionalnost deskriptora pri čemu se tačnost klasifikacije smanjuje vrlo malo. Deskriptor niže dimenzionalnosti je moguće dobiti korišćenjem samo koeficijenata Gaborove transformacije na susjednim skalama, odnosno, orijentacijama, čime se postiže redukcija dimenzionalnosti sa vrlo malom računskom složenosti. Sa druge strane, primjenom analize glavnih komponenata moguće je dobiti deskriptor vrlo niske dimenzionalnosti čime se značajno smanjuju memorijski zahtjevi klasifikatora.

Srodan deskriptor proširenom Gaborovom deskriptoru teksture je strukturalni deskriptor teksture koji takođe za opis slike koristi korelacije između modula koeficijenata Gaborove transformacije iz različitih podopsega. Predloženo je da se ovaj deskriptor koristi kao lokalni deskriptor i slika je predstavljena kao neuređeni skup deskriptora. Klasifikator koji je predložen u ovom slučaju zasnovan je na principu najbližih susjeda i računanju sličnosti između testne slike i svake od klasa. Pokazano je da predloženi deskriptor i klasifikator postižu vrlo dobre rezultate u klasifikaciji slika dobijenih daljinskom detekcijom iako klasifikator ne sadrži fazu obučavanja. U poređenju sa rezultatima iz literature, dobijene tačnosti klasifikacije su, u mno-gim slučajevima, više od tačnosti postignutih metodama koje uključuju obučavanje klasifikatora.

Konačno, upoređene su performanse klasifikatora koji koriste globalne, odnosno, lokalne deskriptore. Kao globalni deskriptor upotrebljen je prošireni Gaborov deskriptor teksture, a kao lokalni, SIFT deskriptor. Pokazano je da svaki od deskriptora daje bolje rezultate na određenim klasama. Tako, na primjer, globalni deskriptor daje bolje rezultate kada je slika vizuelno uniformna, odnosno, kada slika sadrži globalne vizuelne elemente čija je veličina uporediva sa njenom veličinom. Lokalni deskriptori, sa druge strane, imaju bolje performanse na vizuelno nehomogenim i objektnim kategorijama. Zbog komplementarnosti deskriptora, moguće je njihovom fuzijom konstruisati klasifikator sa poboljšanim performansama. U ovoj disertaciji predložena je fuzija deskriptora hijerarhijskim kombinovanjem klasfikatora. Eksperimentalno je pokazano da se na ovaj način dobijaju tačnosti klasifikacije koje nadmašuju sve dosadašnje objavljene rezultate. Štaviše, tačnost klasifikacije je ovdje premašila 90% čime se približava tačnosti obučenih analitičara [108]. Pored dobrih performansi, predloženi metod za fuziju deskriptora omogućava korišćenje jednostavnih metaklasifikatora bez smanjenja tačnosti klasifikacije. Važno je napomenuti i da je predloženi metod za fuziju deskriptora opšti i njegova primjena nije ograničena na deskriptore i klasifikatore razmatrane u ovoj disertaciji.

U ovoj disertaciji je pokazano da je, uz korišćenje različitih elemenata vizuelne interpretacije slika, moguće poboljšati tačnost klasifikacije visokorezolucionih snimaka dobijenih daljinskom detekcijom. Kako bi se ovo postiglo korišćeni su deskriptori teksture, lokalni deskriptori, kao i njihova fuzija. Rezultujući klasifikator postiže vrlo dobre tačnosti klasifikacije na dvije kolekcije slika koje su korišćene i u drugim radovima iz ove oblasti.

Korišćenje visokorezolucionih snimaka omogućava njihovu klasifikaciju u klase pokrivača i načina korišćenja zemljišta na višim nivoima hijerarhije programa CO-RINE i Andersonovog sistema za klasifikaciju čime je moguće dobiti detaljnije tematske mape na kojima figurišu klase načina korišćenja zemljišta koje nije moguće modelirati korišćenjem isključivo spektralne informacije. Ilustrativan primjer predstavljaju klase načina korišćenja zemljišta karakteristične za urbane sredine. Na snimcima urbanih sredina tipično postoje složene kombinacije različitih klasa pokrivača zemljišta pa je za određivanje klase načina korišćenja zemljišta u nekom regionu potrebno iskoristiti prostornu informaciju iz čitavog bloka. Kako bi se ovo postiglo, u ovoj disertaciji je predloženo korišćenje hijerahijske fuzije globalnih i lokalnih deskriptora. Predloženi pristup je testiran na zadatku detekcije urbanih područja na visokorezolucionom aero-snimku. Dobijeni rezultati su u rangu rezultata iz literature iako nisu korišćena rješenja prilagođena za rješavanje problema detekcije urbanih područja.

7.1 Pregled naučnih doprinosa disertacije

Naučni doprinosi ove disertacije ogledaju se u sljedećem:

- Predložen je prošireni Gaborov deskriptor teksture za panhromatske slike koji, pored statistika koeficijenata Gaborove transformacije, sadrži i mere korelacije između koeficijenata na različitim skalama, odnosno, orijentacijama. Ove mere korelacije su u deskriptor inkorporirane kao statistike razlika normalizovanih koeficijenata. Analiziran je uticaj parametara deskriptora na tačnost klasifikacije slika dobijenih daljinskom detekcijom u klase pokrivača i načina korišćenja zemljišta. Prošireni Gaborov deskriptor teksture je evaluiran eksperimentalno i analizirana je njegova zavisnost od nelinearne funkcije kojom se transformišu koeficijenti. Pokazano je da se korišćenjem proširenog Gaborovog deskriptora teksture postižu tačnosti klasifikacije od 87,57% za UC Merced Land Use i 83,74% za Wuhan kolekciju slika. Ove tačnosti su za 10% više od tačnosti postignutih originalnim Gaborovim deskriptora da opiše strukture na slici koje su izražene na različitim skalama, odnosno, orijentacijama.
- Predložena su dva načina za uključivanje spektralne informacije, odnosno, u ovom slučaju informacije o boji u prošireni Gaborov deskriptor teksture. Oba pristupa su bazirana na postojanju interakcija između kolor-kanala pri percepciji slike. Prvi pristup eksplicitno uzima u obzir ove interakcije kroz izračunavanje mera korelacije između koeficijenata Gaborove transformacije izračunatih za različite kolor-kanale. Drugi pristup je zasnovan na reprezentaciji slike u boji pomoću kvaterniona ili hiperkompleksnih brojeva. U ovom slučaju, interakcije između kolor-kanala se uzimaju u obzir implicitno, kroz definicije kvaternionskih operatora za obradu slike. Eksperimentalno je pokazano da oba načina za uključivanje informacije o boji daju deskriptore koji rezultuju tačnostima klasifikacije od preko 88% za UC Merced Land Use i 85% za Wuhan kolekciju slika. Kada se međusobno uporede, oba pristupa rezultuju sličnim performansama, ali je dimenzionalnost kvaternionskog deskriptora četiri puta manja o dimenzionalnosti oponentnog proširenog Gaborovog deskriptora teksture.
- Oba prethodno predložena deskriptora su visokodimenzionalni. Pošto je, pored deskriptivnosti i diskriminativnosti, za reprezentaciju slike značajno da bude i kompaktna, razmatrana je mogućnost redukcije dimenzionalnosti deskriptora. Sa tim ciljem, analizirana je mogućnost da se deskriptor formira korišćenjem samo koeficijenata iz susjednih podopsega, te mogućnosti primene analize glavnih komponenata i linearne diskriminatne analize. Eksperimentalno je pokazano da se najbolji rezultati dobijaju primenom analize glavnih komponenata primenjene na pojedine elemente deskriptora. Tačnosti klasifikacije dobijene nakon redukcije dimenzionalnosti deskriptora su, u najgorem slučaju, niže za samo 1-2% od tačnosti klasifikacije dobijenih korišćenjem deskriptora pune dimenzionalnosti, dok je dimenzionalnost deskriptora smanjena i do 10 puta u slučaju oponentnog proširenog Gaborovog deskriptora teksture.

- Predložen je lokalni strukturalni deskriptor teksture za reprezentaciju slika dobijenih daljinskom detekcijom. I ovaj deskriptor uzima u obzir korelacije između koeficijenata Gaborove transformacije tako što se u deskriptor uključuju koeficijenti korelacije između modula koeficijenata na različitim skalama, odnosno, orijentacijama. Pored strukturalnog deskriptora teksture, predloženo je da se kao mera sličnosti dva lokalna deskriptora, odnosno, dva prozora na slikama, koristi perceptualno zasnovana strukturalna mera sličnosti tekstura. Polazeći od predložene reprezentacije i mere sličnosti, predložen je klasifikator slika dobijenih daljinskom detekcijom zasnovan na principu najbližeg suseda i korišćenju sličnosti slika-klasa. Uprkos jednostavnosti i odsustvu faze obučavanja klasifikatora, eksperimentalni rezultati su pokazali da se njegovim korišćenjem dobijaju tačnosti klasifikacije od 85% na UC Merced Land Use kolekciji slika, odnosno, 79% na Wuhan kolekciji slika. Ove tačnosti su u rangu rezultata dobijenih korišćenjem reprezentacije slike kao skupa riječi i složenijih klasifikatora kao što su mašine sa vektorima nosačima.
- S obzirom na to da rezultati u literaturi ne daju mogućnost da se da nedvosmislen odgovor na pitanje o tome da li su lokalni deskriptori bolji od globalnih u zadacima klasifikacije slika dobijenih daljinskom detekcijom u klase pokrivača i načina korišćenja zemljišta, u podnesenoj disertaciji eksperimentalno je izvršena uporedna analiza performansi ova dva tipa deskriptora. Pokazano je da rezultati zavise od klase, odnosno, od sadržaja slike. Zaključeno je da globalni i lokalni deskriptori slike sadrže komplementarne informacije i da se njihovom fuzijom može poboljšati tačnost klasifikacije.
- Predložen je hijerarhijski način za fuziju globalnih i lokalnih deskriptora koji sadrži dva nivoa. Na prvom nivou se koriste deskriptori slika različitog tipa i odgovarajući klasifikatori koji na svojim izlazima daju verovatnoće sa kojim se slika klasifikuje u svaku od klasa. Ove verovatnoće se kombinuju u međudeskriptor koji predstavlja ulaz u metaklasifikator koji donosi konačnu odluku o klasifikaciji slike. Eksperimentalno je testiran veliki broj metaklasifikatora i rezultati su upoređeni sa jednostavnim načinom fuzije konkatenacijom deskriptora. Pokazano je da se hijerarhijskom fuzijom dobijaju više tačnosti klasifikacije nego konkatenacijom deskriptora. Takođe je pokazano da i vrlo jednostavni metaklasifikatori, kao što je regularizovana linearna regresija daju vrlo dobre rezultate što omogućava realizaciju jednostavnijeg hijerarhijskog klasifikatora. Na ovaj način su na dvije testirane kolekcije slika postignute tačnosti klasifikacije od oko 90%.
- Predloženi hijerarhijski klasifikator testiran je i na zadatku detekcije urbanih područja na aero-snimku šireg geografskog područja, kako bi se ocenio kvalitet označavanja regiona dobijenih predloženim blokovskim klasifikatorima. Pokazano je da predloženi pristup daje dobre rezultate u ovom eksperimentu postižući tačnost klasifikacije od 91,25% uz 12,25% lažno pozitivnih detekcija. Pri ovome nisu korišćena rješenja specijalizovana za detekciju urbanih područja.

7.2 Budući rad

Postoji nekoliko zanimljivih smjerova za dalje istraživanje u ovoj oblasti. Prvi je uključivanje spektralne informacije u reprezentaciju slike. Iako su u ovoj disertaciji razmatrani deskriptori teksture za slike u boji, ostaje otvoreno pitanje reprezentacije same boje. Ovo pitanje je važno zato što je boja jedno od najizraženijih vizuelnih obilježja slike, a u analizi slika dobijenih daljinskom detekcijom odgovara spektralnoj informaciji koja tradicionalno predstavlja osnovno obilježje za klasifikaciju snimaka u klase pokrivača i načina korišćenja zemljišta. Takođe, u reprezentaciju slike je moguće uključiti i informacije iz spektralnih opsega koji nisu vidljivi za ljudsko oko. Iako je korišćenje multispektralnih slika uobičajeno u analizi niskorezolucionih snimaka, prostorni klasifikatori visokorezolucionih snimaka su još uvijek rijetki u literaturi. Uključenje novih obilježja u predloženi klasifikator je lako izvodljivo zbog njegovog dizajna koji omogućava njihovo lako dodavanje

Kako bi se poboljšala reprezentacija slike, potrebno je razmotriti i fuziju sa drugim izvorima informacija, kao što su, na primjer, prostorni odnosi između elemenata na slici ili digitalni model terena. Prostorni odnosi između elemenata na slici mogu biti presudni za diskriminaciju nekih klasa načina korišćenja zemljišta, ali njihovo uključivanje treba da bude robusno na varijacije rasporeda prostornih elemenata na slici. Kada je riječ o slikama opšte namjene, jednostavni geometrijski modeli sa fiksnom prostornom podjelom slike, kao što su globalni Gist deskriptor ili piramidalno uparivanje za lokalne deskriptore, značajno poboljšavaju tačnost klasifikacije. Međutim, jednostavni geometrijski modeli daju dobre rezultate u slučaju kada kompozicija vizuelnih elemenata na slici odgovara modelu. Kada su slike dobijene daljinskom detekcijom u pitanju, raspored prostornih elemenata unutar jedne klase pokrivača i načina korišćenja zemljišta može značajno varirati tako da jednostavni geometrijski modeli ne donose očekivano poboljšanje. Sa druge strane, složeniji geometrijski modeli su vrlo računski zahtjevni i njihovo uključivanje u klasifikator snimaka dobijenih daljinskom detekcijom bi bilo nepraktično. Dakle, rješenje ovog problema je potrebno tražiti u klasi fleksibilnih i računski jednostavnih geometrijskih modela.

Većina eksperimenata u ovoj disertaciji je izvedena na ortofoto aero-snimcima. Jedan od proizvoda ortorektifikacije aero-snimka je i digitalni model terena. Digitalni model terena sadrži podatke o elevaciji terena koji je predmet snimanja. Različitim klasama pokrivača i načina korišćenja zemljišta odgovaraju različiti uzorci elevacije terena. Na primjer, poljoprivredno zemljište i vodene površine mogu imati približno konstantnu ili sporopromjenljivu elevaciju, dok elevacija u različitim klasama urbanih sredina može značajno varirati. Zbog toga se može očekivati da će se uključenjem ovog podatka u klasifikaciju tačnost klasifikacije dodatno poboljšati. Fuzija deskriptora predložena u ovoj disertaciji predstavlja jedan od mogućih načina za uključenje digitalnog modela terena u reprezentaciju slike.

Obim podataka uključenih u analizu snimaka dobijenih daljinskom detekcijom je ogroman. Kako bi se realizovali praktično upotrebljivi sistemi za klasifikaciju snimaka potrebno je obratiti pažnju na računsku složenost algoritama, kao i na njihovu realizaciju na računarskim arhitekturama visokih performansi kao što su računarski klasteri i grafički procesori opšte namjene.

Dodatak A Korelacije odziva Gaborovih filtara

Kako bi izlaganje bilo potpunije ovdje je ponovljena definicija Gaborove filtar banke prethodno data u odjeljku 3.2. Impulsni odziv Gaborovog filtra dat je Gaborovom funkcijom (3.2),

$$g(x,y) = \frac{1}{2\pi\sigma_x\sigma_y} e^{-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2}\right)} e^{j\omega x},$$
(A.1)

gdje je ω frekvencija Gaborove funkcije, a σ_x i σ_y određuju širinu Gausove anvelope. Furijeova transformacija ove funkcije je

$$G(\omega_x, \omega_y) = e^{-\frac{1}{2} \left[\frac{(\omega_x - \omega)^2}{\sigma_{\omega_x}^2} + \frac{\omega_y^2}{\sigma_{\omega_y}^2} \right]},$$
(A.2)

pri čemu je $\sigma_{\omega_x} = \frac{1}{\sigma_x}$ i $\sigma_{\omega_y} = \frac{1}{\sigma_y}$. Dakle, Gaborov filtar je propusnik opsega kod kojeg je širina propusnog opsega određena standardnim devijacijama Gausove anvelope.

Gaborovi talasići se dobijaju dilatacijama (skaliranjem) i rotacijama g(x, y):

$$g_{mn}(x,y) = a^{-m}g(x',y'),$$
 (A.3)

pri čemu su $m = 0, \ldots, S-1$ i $n = 0, \ldots, K-1$ indeksi skale i orijentacije talasića, respektivno, a

$$x' = a^{-m} \left(x \cos \theta + y \sin \theta \right) \tag{A.4}$$

$$y' = a^{-m} \left(-x\sin\theta + y\cos\theta \right), \tag{A.5}$$

gdje je a > 1, a $\theta = n\pi/N$. Skalirajući faktor a^{-m} u (3.4) služi za normalizaciju energije. U skladu sa ustaljenom terminologijom [37] svaki par skale i orijentacije određuje jedan podopseg.

Realni i imaginarni dio Gaborovog filtra (A.1) su približno u kvadraturi [45] tako da se Gaborov filtar može koristiti za izdvajanje komponenata signala iz uskog frekvencijskog opsega i aproksimaciju odgovarajućeg analitičkog signala. U ovom slučaju nedostatak definicije (A.1) je što dobijeni realni dio Gaborovog filtra propušta jednosmjernu komponentu signala. Zbog toga se u ovom radu realni dio Gaborovog filtra modifikuje jednostavnim oduzimanjem njegove srednje vrijednosti. Signal na izlazu Gaborovog filtra $w_{mn}(x, y) = c_{mn}(x, y) + js_{mn}(x, y)$ je analitički tako da su njegov realni i imaginarni dio nekorelisani [109]

$$E\{c_{mn}(x,y)s_{mn}(x,y)\} = 0.$$
 (A.6)

Pošto Gaborovi filtri nisu ortogonalni postoje korelacije između različitih podopsega. Radi preglednije notacije u nastavku neće biti eksplicitno označene skale i orijentacije Gaborovih filtara nego će se koristiti jedinstveni indeks podopsega. Dakle, neka su $w_k(x,y)$ i $w_l(x,y)$ odzivi Gaborovih filtara u podopsezima k i l, respektivno, $k \neq l$. Kroskorelacija između podopsega jednaka je

$$\gamma_{w_k w_l} (\xi, \eta) = E \{ w_k (x, y) w_l^* (x - \xi, y - \eta) \} = = E \{ [c_k (x, y) c_l (x - \xi, y - \eta) + s_k (x, y) s_l (x - \xi, y - \eta)] \} + + jE \{ [s_k (x, y) c_l (x - \xi, y - \eta) - c_k (x, y) s_l (x - \xi, y - \eta)] \} = = \gamma_{c_k c_l} (\xi, \eta) + \gamma_{s_k s_l} (\xi, \eta) + j [\gamma_{s_k c_l} (\xi, \eta) - \gamma_{c_k s_l} (\xi, \eta)].$$
(A.7)

Pošto su signali $w_k(x, y)$ i $w_l(x, y)$ analitički vrijedi

$$\gamma_{c_k c_l}\left(\xi,\eta\right) = \gamma_{s_k s_l}\left(\xi,\eta\right),\tag{A.8}$$

Odredimo sada vrijednosti kroskorelacija koje figurišu u imagninarnom dijelu (A.7). Neka su $G_k^{(real)}(u,v)$ i $G_l^{(real)}(u,v)$ funkcije prenosa realnih dijelova Gaborovih filtara u podopsezima k i l, respektivno. Pošto su imaginarni dijelovi ovih Gaborovih filtara približno Hilbertove transformacije realnih dijelova mogu se napisati u obliku

$$G_{k}^{(imag)}\left(u,v\right) = -jG_{k}^{(real)}\left(u,v\right)\operatorname{sgn}\left(u\right)$$
(A.9)

i

$$G_{l}^{(imag)}(u,v) = -jG_{l}^{(real)}(u,v)\operatorname{sgn}(u).$$
 (A.10)

Ovdje se primjenjuje 1D Hilbertova transformacija po x-osi zato što su i oscilacije Gaborove funkcije u smjeru x-ose (3.2).

Imajući u vidu da je realni dio impulsnog odziva Gaborovog filtra parna funkcija, a time je i njegova frekvencijska karakteristika realna, kros-spektralna gustina snage signala $s_k(x, y)$ i $c_l(x, y)$ je [109]

$$\Gamma_{s_k c_l}(u, v) = j G_k^{(real)}(u, v) G_l^{(real)}(u, v) \operatorname{sgn}(u) \Gamma_I(u, v), \qquad (A.11)$$

gdje je $\Gamma_I(u, v)$ spektralna gustina snage ulazne slike. Sa druge strane je

$$\Gamma_{c_l s_k}\left(u, v\right) = -j G_l^{(real)}\left(u, v\right) \operatorname{sgn}\left(u\right) G_k^{(real)}\left(u, v\right) \Gamma_I\left(u, v\right).$$
(A.12)

Sada je

$$\Gamma_{s_k c_l}\left(u, v\right) = -\Gamma_{c_l s_k}\left(u, v\right),\tag{A.13}$$

pa za odgovarajuće kovarijanse važi

$$\gamma_{s_k c_l}\left(\xi,\eta\right) = -\gamma_{c_l s_k}\left(\xi,\eta\right). \tag{A.14}$$

Važan specijalni slučaj se dobija za $\xi = \eta = 0$. Pošto su signali $s_k(x,y)$ i $c_l(x,y)$ realni, vrijedi

$$\gamma_{s_k c_l} (0,0) = \gamma_{c_l s_k} (0,0) , \qquad (A.15)$$

odakle imamo

$$\gamma_{s_k c_l}(0,0) = \gamma_{c_l s_k}(0,0) = 0. \tag{A.16}$$

Dakle, kroskorelacije između realnih i imaginarnih dijelova odziva Gaborovih filtara na istim prostornim lokacijama u različitim podopsezima su jednake nuli.

Konačno, imamo

$$\gamma_{w_k w_l}(\xi, \eta) = 2\gamma_{c_k c_l}(\xi, \eta) + j 2\gamma_{s_k c_l}(\xi, \eta) .$$
(A.17)

Za $\xi = \eta = 0$ dobija se

$$\gamma_{w_k w_l}(0,0) = 2\gamma_{c_k c_l}(0,0).$$
(A.18)

Prema tome, kroskorelacija između odziva u različlitim podopsezima je realna veličina i jednaka je dvostrukoj kroskorelaciji odgovarajućih realnih dijelova odziva.

Dodatak B Gist deskriptor scene

Gist deskriptor je inicijalno predložen u [21] kako bi se dobila globalna niskodimenzionalna reprezentacija scene koja treba da omogući njenu klasifikaciju u neku od predefinisanih semantičkih klasa. Motivacija za njegovo uvođenje bili su eksperimentalni rezultati u oblasti ljudske percepcije scena koji su ukazali na činjenicu da ljudska bića mogu prepoznati scenu korišćenjem samo globalnih obilježja, ne koristeći lokalne informacije o detaljima prisutnim na sceni. To znači da postoji skup vizuelnih svojstava koja su zajednička za scene koje pripadaju istoj semantičkoj klasi. Ovaj skup holističkih osobina scene koje se mogu koristiti za određivanje semantičke klase kojoj scena pripada, bez potrebe da se prepoznaju objekti na sceni, nazvan je *prostorna ovojnica* scene. Eksperimentalno je identifikovano pet osobina prostorne anvelope: stepen prirodnosti, stepen otvorenosti, stepen gruboće, stepen ekspanzije i stepen neravnosti.

Osobine prostorne ovojnica se mogu odrediti korišćenjem spektralnih i grubo lokalizovanih informacija sadržanih u slici. Originalno, za izračunavanje Gist deskriptora korišćena je Furijeova transformacija, odnosno, spektrogram [21], a u kasnijim radovima se koriste usmjereni [110] i Gaborovi filtri [111]. U ovoj disertaciji, Gist deskriptor slike se izračunava tako što se slika prvo filtrira Gaborovom filtar bankom, a zatim se izračuna srednja vrijednost odziva filtara u svakom bloku dobijenom podjelom slike pravilnom mrežom dimenzija 4×4 , kao što je prikazano na Slici B.1.



Slika B.1: Ilustracija izračunavanja Gist deskriptora.

Poređenjem ovog deskriptora sa Gaborovim deskriptorom teksture, može se uočiti da Gist deskriptor u suštini opisuje prostorni raspored tekstura. Takođe, Gist deskriptor ne uključuje standardne devijacije odziva Gaborovih filtara. Pored sličnosti sa Gaborovim deskriptorom teksture, Gist deskriptor se može posmatrati i kao globalna verzija SIFT deskriptora, opisanog u Dodatku C zato što lokalizovane energije odziva usmjerenih filtara čine reprezentaciju slike sličnu lokalizovanim histogramima orijentacija gradijenata koji se formiraju kada se izračunava SIFT deskriptor.

Uprkos jednostavnosti ovaj deskriptor je pokazao dobre rezultate u zadacima klasifikacije scena. Iako je često korišćen u klasifikaciji scena i predložena su njegova unapređenja [49, 58], nema mnogo primjera korišćenja Gist deskriptora u klasifikaciji slika dobijenih daljinskom detekcijom. Izuzeci su korišćenje Gist deskriptora za klasterizaciju slika u okviru detekcije drveća [59] i za detekciju ciljeva na satelitskim slikama [50], kao i poboljšani Gist model koji se koristi za detekciju kratera u [60].

Dodatak C SIFT deskriptor

Jedan od najpopularnijih lokalnih deskriptora koji se koristi u klasifikaciji slika je Scale Invariant Feature Transform (SIFT) deskriptor predložen u [88]. SIFT deskriptor je inicijalno predložen za pronalaženje objekata na slikama direktnim uparivanjem deskriptora određenih u regionima oko interesnih tačaka. Zbog ovoga procedura izračunavanja SIFT deskriptora uključuje i korak u kojem se detektuju interesne tačke. Detektor interesnih tačaka funkcioniše na principu pronalaženja lokalnih ekstremuma na slikama filtriranim razlikom Gausovih filtara, kako po prostornim dimenzijama, tako i po skalama. Na ovaj način se pronalaze interesne tačke koje tipično odgovaraju uglovima na slici i stabilne su na varijacije lokacije, skale i orijentacije. Regioni kvadratnog oblika oko detektovanih interesnih tačaka opisuju se SIFT deskriptorima. Za dati region, SIFT deskriptor se izračunava tako što se najprije izračunavaju vrijednosti modula i orijentacije gradijenta slike na skali na kojoj je detektovana odgovarajuća interesna tačka. Region se pravilnom mrežom dimenzija 4×4 dijeli na blokove, kao što je prikazano na Slici C.1. Svaki blok sadrži 4×4 lokacije u kojima je gradijent izračunat i za njega se formira histogram orijentacija gradijenta, pri čemu je doprinos svakog gradijenta ponderisan njegovim modulom i Gausovom težinskom funkcijom sa ishodištem u interesnoj tački. Prilikom formiranja histograma orijentacije se kvantuju na osam nivoa, tako da je ukupna dimenzionalnost SIFT deskriptora $4 \times 4 \times 8 = 128$.



Slika C.1: Primjer izračunavanja SIFT deskriptora.

DODATAK C. SIFT DESKRIPTOR

Kako bi se formirala kompaktna reprezentacija slike i postigla robusnost na šum i varijacije u izgledu objekta, funkcija gustine raspodjele vjerovatnoće SIFT deskriptora na slici se formira njihovom vektorskom kvantizacijom i formiranjem histograma broja pojavljivanja kodnih riječi. Ova reprezentacija se naziva *skup vizuelnih riječi* i njene varijacije trenutno predstavljaju najpopularnije reprezentacije slike u pretraživanju baza slika, prepoznavanju objekata i klasifikaciji scena. Kada se koristi reprezentacija slike pomoću skupa vizuelnih riječi, eksperimentalno je pokazano [92, 112] da se bolji rezultati dobijaju kada se ne koristi detektor interesnih tačaka, nego se deskriptori računaju u pravilno raspoređenim tačkama na određenom skupu skala.

Dodatak D

Publikacije kandidata u vezi sa disertacijom

- V. Risojević and Z. Babić. Fusion of global and local descriptors for remote sensing image classification. *IEEE Geoscience and Remote Sensing Letters*, 10(4):836–840, 2013
- [2] R. Češnovar, V. Risojević, Z. Babić, T. Dobravec, and P. Bulić. A gpu implementation of a structural-similarity-based aerial-image classification. *The Journal of Supercomputing*, 65(2):978–996, 2013
- [3] V. Risojević, S. Momić, and Z. Babić. Gabor descriptors for aerial image classification. In *Proceedings of 10th International Conference on Adaptive and Natural Computing Algorithms (ICANNGA), Part II*, volume 6594 of *LNCS*, pages 51–60. Springer Berlin / Heidelberg, 2011
- [4] A. Avramović and V. Risojević. Descriptor dimensionality reduction for aerial image classification. In Proceedings of 18th International Conference on Systems, Signals and Image Processing (IWSSIP), pages 1–4, Sarajevo, Bosnia and Herzegovina, June 2011
- [5] V. Risojević and Z. Babić. Aerial image classification using structural texture similarity. In A. Elmaghraby and D. N. Serpanos, editors, *Proceedings of IEEE International Symposium on Signal Processing and Information Technology* (*ISSPIT*), pages 190–195, Bilbao, Spain, December 2011. IEEE
- [6] V. Risojević and Z. Babić. Orientation difference descriptor for aerial image classification. In Proceedings of 19th International Conference on Systems, Signals and Image Processing (IWSSIP), pages 156–159, Vienna, Austria, April 2012

Literatura

- [1] J. B. Campbell. Introduction to remote sensing. Guilford Press, 2006.
- [2] R. A. Schowengerdt. *Remote sensing: models and methods for image processing.* Academic Press, 2007.
- [3] T. M. Lillesand, R. W. Kiefer, and J. W. Chipman. *Remote sensing and image interpretation*. John Wiley & Sons, 2008.
- [4] P. G. Marchetti, G. Rivolta, S. D'Elia, J. Farres, N. Gobron, and G. Mason. A model for the scientific exploitation of earth observation missions: the ESA research and service support. *IEEE Geoscience and Remote Sensing Society Newsletter*, (162), 2012.
- [5] J. R. Anderson, E. E. Hardy, J. T. Roach, and R. E. Witmer. A land use and land cover classification system for use with remote sensor data. U. S. Geological Survey Professional Paper 964, 1976.
- [6] C. H. Homer, J. A. Fry, and C. A. Barnes. The national land cover database. U.S. Geological Survey Fact Sheet 2012-3020, 2012.
- [7] M. Bossard, J. Feranec, and J. Otahel. CORINE land cover technical guide addendum 2000. Technical report, European Environment Agency, 2000.
- [8] M. Hodgson. What size window for image classification: A cognitive perspective. Photogrammetric Engineering & Remote Sensing, 64(8):797–807, August 1998.
- [9] T. Blaschke. Object based image analysis for remote sensing. *ISPRS journal* of photogrammetry and remote sensing, 65(1):2–16, 2010.
- [10] J. Zhang, M. Marszalek, S. Lazebnik, and C. Schmid. Local features and kernels for classification of texture and object categories: A comprehensive study. *International Journal of Computer Vision*, 73:213–238, June 2007.
- [11] G. Csurka, C. Dance, L. Fan, J. Willamowski, and C. Bray. Visual categorization with bags of keypoints. In *Proceedings of ECCV. Workshop on statistical learning in computer vision*, 2004.
- [12] S. Lazebnik, C. Schmid, and J. Ponce. Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories. In *Proceedings* of *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 2, pages 2169–2178, 2006.

- [13] O. Boiman, E. Shechtman, and M. Irani. In defense of nearest-neighbor based image classification. In Proceedings of IEEE Computer Vision and Pattern Recognition Conference (CVPR), pages 1–8, 2008.
- [14] J. Yang, K. Yu, Y. Gong, and T. Huang. Linear spatial pyramid matching using sparse coding for image classification. In *Proceedings of IEEE Conference* on Computer Vision and Pattern Recognition, pages 1794–1801. IEEE, 2009.
- [15] T. Serre, L. Wolf, S. Bileschi, M. Riesenhuber, and T. Poggio. Robust object recognition with cortex-like mechanisms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(3):411–426, 2007.
- [16] J. Mutch and D. G. Lowe. Object class recognition and localization using sparse features with limited receptive fields. *International Journal of Compu*ter Vision, 80(1):45–57, 2008.
- [17] M. Ranzato, F. J. Huang, Y.-L. Boureau, and Y. Lecun. Unsupervised learning of invariant feature hierarchies with applications to object recognition. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition* (CVPR), pages 1–8. IEEE, 2007.
- [18] Y. LeCun, K. Kavukcuoglu, and C. Farabet. Convolutional networks and applications in vision. In *Proceedings of IEEE International Symposium on Circuits and Systems (ISCAS)*, pages 253–256. IEEE, 2010.
- [19] A. Coates, A. Y. Ng, and H. Lee. An analysis of single-layer networks in unsupervised feature learning. In *International Conference on Artificial Intelligence and Statistics*, pages 215–223, 2011.
- [20] A. Bosch, A. Zisserman, and X. Muñoz. Scene classification using a hybrid generative/discriminative approach. *IEEE Transactions on Pattern Analysis* and Machine Intelligence, 30(4):712–727, 2008.
- [21] A. Oliva and A. Torralba. Modeling the shape of the scene: a holistic representation of the spatial envelope. *International Journal of Computer Vision*, 42(3):145–175, 2001.
- [22] J. Wu and J. M. Rehg. Centrist: A visual descriptor for scene categorization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(8):1489– 1501, 2011.
- [23] V. Risojević, S. Momić, and Z. Babić. Gabor descriptors for aerial image classification. In Proceedings of 10th International Conference on Adaptive and Natural Computing Algorithms (ICANNGA), Part II, volume 6594 of LNCS, pages 51–60. Springer Berlin / Heidelberg, 2011.
- [24] V. Risojević and Z. Babić. Orientation difference descriptor for aerial image classification. In Proceedings of 19th International Conference on Systems, Signals and Image Processing (IWSSIP), pages 156–159, Vienna, Austria, April 2012.

- [25] V. Risojević and Z. Babić. Fusion of global and local descriptors for remote sensing image classification. *IEEE Geoscience and Remote Sensing Letters*, 10(4):836–840, 2013.
- [26] A. Avramović and V. Risojević. Descriptor dimensionality reduction for aerial image classification. In *Proceedings of 18th International Conference on Systems, Signals and Image Processing (IWSSIP)*, pages 1–4, Sarajevo, Bosnia and Herzegovina, June 2011.
- [27] V. Risojević and Z. Babić. Aerial image classification using structural texture similarity. In A. Elmaghraby and D. N. Serpanos, editors, *Proceedings of IEEE International Symposium on Signal Processing and Information Technology* (ISSPIT), pages 190–195, Bilbao, Spain, December 2011. IEEE.
- [28] R. Češnovar, V. Risojević, Z. Babić, T. Dobravec, and P. Bulić. A gpu implementation of a structural-similarity-based aerial-image classification. *The Journal of Supercomputing*, 65(2):978–996, 2013.
- [29] J. Fauqueur, N. G. Kingsbury, and R. Anderson. Semantic discriminant mapping for classification and browsing of remote sensing textures and objects. In *Proceedings of IEEE International Conference on Image Processing (ICIP)*, pages 846–849, 2005.
- [30] Y. Yang and S. Newsam. Bag-of-visual-words and spatial extensions for landuse classification. In Proceedings of 18th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems (ACM SIGSPA-TIAL GIS), pages 270–279, 2010.
- [31] Y. Yang and S. Newsam. Spatial pyramid co-occurrence for image classification. In Proceedings of International Conference on Computer Vision (ICCV), pages 1465–1472, 2011.
- [32] Y. Yang and S. Newsam. Geographic image retrieval using local invariant features. *IEEE Transactions on Geoscience and Remote Sensing*, 51(2):818– 832, 2013.
- [33] L. Chen, W. Yang, K. Xu, and T. Xu. Evaluation of local features for scene classification using VHR satellite images. In *Proceedings of Joint Urban Remote Sensing Event (JURSE)*, pages 385–388, Munich, Germany, April 2011.
- [34] G. Sheng, W. Yang, L. Chen, and H. Sun. Satellite image classification using sparse codes of multiple features. In *Proceedings of 10th IEEE International Conference on Signal Processing (ICSP)*, pages 952–955, Oct. 2010.
- [35] D. Dai and W. Yang. Satellite image classification via two-layer sparse coding with biased image representation. *IEEE Geoscience and Remote Sensing Letters*, 8(1):173–176, 2011.
- [36] Geo-centar. Izrada aerofotogrametrijskog snimanja donjeg toka rijeke Sane. Technical report, 2012.

- [37] R. C. Gonzalez and R. E. Woods. *Digital Image Processing*. Prentice Hall, 2nd edition, 2002.
- [38] D. H. Hubel and T. N. Wiesel. Receptive fields of cells in striate cortex of very young, visually inexperienced kittens. *Journal of Neurophysiology*, 26:994–1002, 1963.
- [39] D. H. Hubel and T. N. Wiesel. Receptive fields and functional architecture of monkey striate cortex. *Journal of Physiology*, 195(1):215–243, 1968.
- [40] S. E. Palmer. Vision Science Photons to Phenomenology. The MIT Press, 1999.
- [41] S. Marčelja. Mathematical description of the responses of simple cortical cells. Journal of the Optical Society of America, 70(11):1297–1300, 1980.
- [42] J. G. Daugman. Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. *Journal* of the Optical Society of America A, 2(7):1160–1169, July 1985.
- [43] J. P. Jones and L. A. Palmer. An evaluation of the two-dimensional gabor filter model of simple receptive fields in cat striate cortex. *Journal of Neurop*hysiology, 58(6):1233–1258, 1987.
- [44] D. A. Pollen and S. F. Ronner. Phase relationships between adjacent simple cells in the visual cortex. *Science*, 212(4501):1409–1411, 1981.
- [45] D. Gabor. Theory of communication. part 1: The analysis of information. Journal of the Institution of Electrical Engineers-Part III: Radio and Communication Engineering, 93(26):429–441, 1946.
- [46] B. S. Manjunath and W.-Y. Ma. Texture features for browsing and retrieval of image data. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(8):837–842, August 1996.
- [47] J. Daugman. How iris recognition works. IEEE Transactions on Circuits and Systems for Video Technology, 14(1):21–30, 2004.
- [48] C. Liu and H. Wechsler. Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition. *IEEE Transactions on Image Processing*, 11(4):467–476, 2002.
- [49] D. Song and D. Tao. Biologically inspired feature manifold for scene classification. *IEEE Transactions on Image Processing*, 19(1):174–184, 2010.
- [50] Z. Li and L. Itti. Saliency and gist features for target detection in satellite images. *IEEE Transactions on Image Processing*, 20(7):2017–2029, July 2011.
- [51] W. Y. Ma and B. S. Manjunath. Edge flow: a framework for boundary detection and image segmentation. In *Proceedings of IEEE International Conference* on Computer Vision and Pattern Recognition (CVPR), pages 744–749, June 1997.

- [52] L. Itti, C. Koch, and E. Niebur. A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(11):1254–1259, November 1998.
- [53] S. E. Grigorescu, N. Petkov, and P. Kruizinga. Comparison of texture features based on gabor filters. *IEEE Transactions on Image Processing*, 11(10):1160– 1167, 2002.
- [54] U. Kandaswamy, S. Schuckers, and D. Adjeroh. Comparison of texture analysis schemes under nonideal conditions. *IEEE Transactions on Image Processing*, 20(8):2260–2275, August 2011.
- [55] B. S. Manjunath, J.-R. Ohm, V. V. Vasudevan, and A. Yamada. Color and texture descriptors. *IEEE Transactions on Circuits and Systems for Video Technology*, 11(6):703–715, June 2001.
- [56] Y. M. Ro, M. Kim, H. K. Kang, B. S. Manjunath, and J. Kim. MPEG-7 homogeneous texture descriptor. *ETRI Journal*, 23(2):41–51, June 2001.
- [57] Y. Yang and S. Newsam. Comparing SIFT descriptors and Gabor texture features for classification of remote sensed imagery. In *Proceedings of International Conference on Image Processing (ICIP)*, pages 1852–1855, October 2008.
- [58] Y. Han and G. Liu. A hierarchical gist model embedding multiple biological feasibilities for scene classification. In *Proceedings of International Conference* on Pattern Recognition (ICPR), pages 3109–3112. IEEE, 2010.
- [59] L. Yang, X. Wu, E. Praun, and X. Ma. Tree detection from aerial imagery. In Proceedings of the 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems (ACM SIGSPATIAL GIS), pages 131–137, New York, NY, USA, 2009.
- [60] Y. Mu, W. Ding, D. Tao, and T. F. Stepinski. Biologically inspired model for crater detection. In *Proceedings of International Joint Conference on Neural Networks (IJCNN)*, pages 2487–2494. IEEE, 2011.
- [61] V. Vapnik. Statistical Learning Theory. John Wiley, 1998.
- [62] M. N. Do and M. Vetterli. Wavelet-based texture retrieval using generalized gaussian density and kullback-leibler distance. *IEEE Transactions on Image Processing*, 11(2):146–158, 2002.
- [63] E. P. Simoncelli. Statistical modeling of photographic images. In A. C. Bovik, editor, *Handbook of Image and Video Processing*, pages 431–441. Academic Press, 2005.
- [64] S. Lyu and E. P. Simoncelli. Nonlinear image representation using divisive normalization. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1–8, 2008.

- [65] J. Portilla and E. P. Simoncelli. A parametric texture model based on joint statistics of complex wavelet coefficients. *International Journal of Computer* Vision, 40(1):49–71, October 2000.
- [66] Z. Wang and E. P. Simoncelli. Reduced-reference image quality assessment using a wavelet-domain natural image statistic model. In *Proceedings of SPIE Human Vision and Electronic Imaging*, volume 5666, pages 149–159, 2005.
- [67] A. Jain and G. Healey. A multiscale representation including opponent color features for texture recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 7(1):124–128, January 1998.
- [68] R. O. Duda, P. E. Hart, and D. G. Stork. *Pattern Classification*. John Wiley & Sons, Inc., 2001.
- [69] J. H. Friedman. Regularized discriminant analysis. Journal of the American Statistical Association, 84(405):165–175, 1989.
- [70] X. Jiang. Linear subspace learning-based dimensionality reduction. IEEE Signal Processing Magazine, 28(2):16–26, 2011.
- [71] A. M. Martinez and A. C. Kak. PCA versus LDA. IEEE Transactions on Pattern Analysis and Machine Intelligence, 23(2):228–233, 2001.
- [72] T. N. Wiesel and D. H. Hubel. Spatial and chromatic interactions in the lateral geniculate body of the rhesus monkey. *Journal of Neurophysiology*, 29:1115–1156, 1966.
- [73] K. N. Plataniotis and A. N. Venetsanopoulos. Color image processing and applications. Springer, 2000.
- [74] S. Sangwine. Fourier transforms of colour images using quaternion or hypercomplex, numbers. *Electronic Letters*, 32(21):1979–1980, 1996.
- [75] C. Jones and A. L. Abbott. Color face recognition by hypercomplex gabor analysis. In Proceedings of International Conference on Automatic Face and Gesture Recognition, 2006.
- [76] T. A. Ell and S. J. Sangwine. Hypercomplex fourier transforms of color images. IEEE Transactions on Image Processing, 16(1):22–35, January 2007.
- [77] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4):600–612, 2004.
- [78] Z. Wang and A. C. Bovik. Mean squared error: Love it or leave it. *IEEE Signal Processing Magazine*, 26(1):98–117, January 2009.
- [79] M. P. Sampat, Z. Wang, S. Gupta, A. C. Bovik, and M. K. Markey. Complex wavelet structural similarity: A new image similarity index. *IEEE Transacti*ons on Image Processing, 18(11):2385–2401, November 2009.

- [80] W. T. Freeman and E. H. Adelson. The design and use of steerable filters. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(9):891–906, 1991.
- [81] Y. Gao, A. Rehman, and Z. Wang. CW-SSIM based image classification. In Proceedings of IEEE International Conference on Image Processing (ICIP), pages 1249–1252, Brussels, Belgium, September 2011.
- [82] X. Zhao, M. Reyes, T. Pappas, and D. Neuhoff. Structural texture similarity metrics for retrieval applications. In *Proceedings of IEEE International Conference on Image Processing (ICIP)*, pages 1196–1199, San Diego, CA, USA, October 2008.
- [83] J. Zujovic, T. Pappas, and D. Neuhoff. Structural similarity metrics for texture analysis and retrieval. In *Proceedings of International Conference on Image Processing (ICIP)*, pages 2225–2228, 2009.
- [84] J. Zujovic, T. Pappas, and D. Neuhoff. Perceptual similarity metrics for retrieval of natural textures. In *IEEE International Workshop on Multimedia Signal Processing (MMSP)*, pages 1–5, Rio de Janeiro, Brazil, October 2009.
- [85] F. Sebastiani. Machine learning in automated text categorization. ACM Computing Surveys, 34(1):1–47, March 2002.
- [86] C. Schmid and R. Mohr. Local Grayvalue Invariants for Image Retrieval. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(5):530– 534, 1997.
- [87] T. Leung and J. Malik. Representing and recognizing the visual appearance of materials using three-dimensional textons. *International Journal of Computer* Vision, 43(1):29–44, 2001.
- [88] D. Lowe. Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision, 60(2):91–110, November 2004.
- [89] N. Dalal and W. Triggs. Histograms of oriented gradients for human detection. In Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 886–893, San Diego, CA, USA, June 2005. IEEE.
- [90] T. Tuytelaars, M. Fritz, K. Saenko, and T. Darrell. The NBNN kernel. In Proceedings of International Conference on Computer Vision (ICCV), pages 1824–1831, 2011.
- [91] J. Sivic and A. Zisserman. Video Google: A text retrieval approach to object matching in videos. In *Proceedings of International Conference on Computer Vision (ICCV)*, volume 2, pages 1470–1477, October 2003.
- [92] E. Nowak, F. Jurie, and B. Triggs. Sampling strategies for bag-of-features image classification. In A. Leonardis, H. Bischof, and A. Pinz, editors, *Pro*ceedings of European Conference on Computer Vision (ECCV), volume 3954

of *Lecture Notes in Computer Science (LNCS)*, pages 490–503, Graz, Austria, 2006. Springer.

- [93] B. Ozdemir and S. Aksoy. Image classification using subgraph histogram representation. In Proceedings of IAPR International Conference on Patter n Recognition (ICPR), Istanbul, Turkey, August 2010.
- [94] S. Xu, T. Fang, D. Li, and S. Wang. Object classification of aerial images with bag-of-visual words. *IEEE Geoscience and Remote Sensing Letters*, 7(2):366– 370, April 2010.
- [95] L. Weizman and J. Goldberger. Urban-area segmentation using visual words. IEEE Geoscience and Remote Sensing Letters, 6(3):388–392, July 2009.
- [96] A. Abdullah, R. C. Veltkamp, and M. A. Wiering. Spatial pyramids and two-layer stacking SVM classifiers for image categorization: A comparative study. In *Proceedings of International Joint Conference on Neural Networks* (IJCNN), pages 5–12, Jun. 2009.
- [97] D. H. Wolpert. Stacked generalization. Neural Networks, 5:241–259, 1992.
- [98] K. M. Ting and I. H. Witten. Issues in stacked generalization. Journal of Artificial Intelligence Research, 10:271–289, 1999.
- [99] S. Reid and G. Grudic. Regularized linear models in stacked generalization. In Proceedings of 8th International Workshop on Multiple Classifier Systems, MCS '09, pages 112–121, Berlin, Heidelberg, 2009. Springer-Verlag.
- [100] A. K. Seewald. How to make stacking better and faster while also taking care of an unknown weakness. In *Proceedings of 19th International Conference on Machine Learning (ICML)*, pages 554–561, Sydney, Australia, Jul. 2002.
- [101] A. Cheriyadat. Unsupervised feature learning for aerial scene classification. IEEE Transactions on Geoscience and Remote Sensing, PP(99):1–13, 2013.
- [102] R. Tibshirani. Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society. Series B (Methodological), pages 267–288, 1996.
- [103] Y. Jiang, J. Yuan, and G. Yu. Randomized spatial partition for scene recognition. In Proceedings of European Conference on Computer Vision (ECCV), pages 730–743. Springer, 2012.
- [104] J. Ren, X. Jiang, and J. Yuan. Learning binarized pixel-difference pattern for scene recognition. In Proceedings of IEEE International Conference on Image Processing. 2013.
- [105] P. Zhong and R. Wang. Using combination of statistical models and multilevel structural information for detecting urban areas from a single gray-level image. *IEEE Transactions on Geoscience and Remote Sensing*, 45(5):1469– 1482, 2007.

- [106] B. Sirmacek and C. Unsalan. Urban area detection using local feature points and spatial voting. *IEEE Geoscience and Remote Sensing Letters*, 7(1):146– 150, 2010.
- [107] G. Liu, G.-S. Xia, X. Huang, W. Yang, and L. Zhang. A perception-inspired building index for automatic built-up area detection in high-resolution satellite images. In *Proceedings of IEEE International Geoscience and Remote Sensing* Symposium (IGARSS), 2013.
- [108] C. E. Olson. Is 80% accuracy good enough. In Proceedings of the ASPRS 17th Pecora Conference, 2009.
- [109] B. Picinbono. Random Signals and Systems. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1993.
- [110] A. Torralba, K. P. Murphy, W. T. Freeman, and M. A. Rubin. Context-based vision system for place and object recognition. In *Proceedings of International Conference on Computer Vision (ICCV)*, pages 273–280, 2003.
- [111] C. Siagian and L. Itti. Rapid biologically-inspired scene classification using features shared with visual attention. *IEEE Transactions on Pattern Analysis* and Machine Intelligence, 29(2):300–312, 2007.
- [112] L. Fei-Fei and P. Perona. A bayesian hierarchical model for learning natural scene categories. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 2, pages 524–531, 2005.