ZESZYTY NAUKOWE POLITECHNIKI ŚLĄSKIEJ

Henryk PALUS WYBRANE ZAGADNIENIA PRZETWARZANIA OBRAZÓW BARWNYCH



POLITECHNIKA ŚLĄSKA ZESZYTY NAUKOWE NR 1742 P. 3342/06

Henryk PALUS

WYBRANE ZAGADNIENIA PRZETWARZANIA OBRAZÓW BARWNYCH

Opiniodawcy Prof. dr. hab. inż. Małgorzata KUJAWIŃSKA Prof. dr hab. inż. Witold MALINA

Kolegium redakcyjne

REDAKTOR NACZELNY	_	Prof. dr hab. inż. Andrzej BUCHACZ
REDAKTOR DZIAŁU	_	Dr inż. Krzysztof SIMEK
SEKRETARZ REDAKCJI	_	Mgr Elżbieta LEŚKO

PRACTWARZANDA ORRAZOW BARWING

Wydano za zgodą Rektora Politechniki Śląskiej

PL ISSN 0434-0760

WYDAWNICTWO POLITECHNIKI ŚLĄSKIEJ ul. Akademicka 5, 44-100 Gliwice; tel./faks (0-32) 237-13-81 http://wydawnictwo.polsl.pl

Sprzedaż i Marketing

tel. (0-32) 237-18-48 wydawnictwo_markp@polsl.pl

 Nakł.
 100+50
 Ark.
 wyd.
 13
 Ark.
 druk.
 13,625
 Papier
 offset.
 70x100,80g

 Oddano do druku 21.12.2006 r.
 Podpisano do druku 21.12.2006 r.
 Podpisano do druku 21.12.2006 r.
 Druk ukończ.
 w styczniu 2007 r.

10

Wydrukowano w Zakładzie Graficznym Politechniki Śląskiej w Gliwicach, ul. Kujawska 1 zam. 2/07

Spis treści

Sp	ois ak	ronimów	7
Sp	ois wa	ażniejszych oznaczeń	9
	Wp	rowadzenie	11
2.	Rep	rezentacja barwy w obrazach cyfrowych	17
	2.1.	Przestrzeń barw RGB	18
	2.2.	Przestrzeń barw CIE XYZ	20
	2.3.	Telewizyjne przestrzenie barw	22
		2.3.1. Przestrzeń barw YUV	23
		2.3.2. Przestrzeń barw YIQ	23
		2.3.3. Przestrzeń barw YC_bC_r	24
	2.4.	Przestrzeń barw opozycyjnych	25
	2.5.	Przestrzeń Ohty	27
	2.6.	Percepcyjne przestrzenie barw	28
	2.7.	Percepcyjnie równomierne przestrzenie barw	35
		2.7.1. Przestrzeń CIELAB	35
		2.7.2. Przestrzeń <i>CIELUV</i>	38
	2.8.	Ocena właściwości przestrzeni barw	39
	Poz	yskiwanie obrazów barwnych	45
	3.1.	Kalibracja systemu wizyjnego	47
	3.2.	Ocena wierności odwzorowania barw w systemie wizyjnym $\ .\ .\ .\ .$	58
	3.3.	Wyznaczanie charakterystyk widmowych kamer $\ . \ . \ . \ . \ . \ .$	67
	3.4.	Ocena barwności pozyskanego obrazu	70

Spis treści 4. Badanie właściwości wybranych przestrzeni barw

4.	Bac	lanie właściwości wybranych przestrzeni barw	87
	4.1.	Przestrzeń barw HSx	88
		4.1.1. Możliwości modyfikacji przestrzeni HSx	90
		4.1.2. Porównanie transformacji z RGB na HSx	90
		4.1.3. Badanie struktury przestrzeni HSx	94
	4.2.	Przestrzeń $K_1 K_2 K_3 \ldots \ldots$	96
		4.2.1. Dekorelacja składowych RGB	98
		4.2.2. Podstawowe właściwości KLT	101
		4.2.3. Zastosowania KLT	105
5.	Seg	mentacja obrazów barwnych	111
	5.1.	Progowanie obrazów barwnych	116
		5.1.1. Zastosowanie wiedzy o tle obrazu	118
		5.1.2. Detekcja odblasków	123
	5.2.	Klasteryzacja w przestrzeni barw	130
	5.3.	Rozrost obszaru w obrazie barwnym	136
		5.3.1. Technika ziarnistego rozrostu obszaru	138
		5.3.2. Technika nieziarnistego rozrostu obszaru	142
		5.3.3. Cienie i odblaski w procesie rozrostu obszaru	147
	5.4.	Przetwarzanie końcowe	149
	5.5.	Ilościowa ocena wyników segmentacji	155
	5.6.	Dalsze przetwarzanie obrazu po segmentacji	159
6.	Prze	etwarzanie wstępne obrazów barwnych	167
	6.1.	Kwantyzacja barwy	167
	6.2.	Filtracja odszumiająca	173
7.	Pod	sumowanie i wnioski	187
Bi	bliog	rafia	193
Do	date	k: Obliczanie różnicy barw według formuły CIEDE2000	209
Sti	eszc	zenie	211

Contents			
Contents			
List of acronyms			7
List of actomynics			
List of symbols			9
1 Introduction			11
1. Introduction			
2. Representations of colour in digita	al images		17
2.1. RGB colour space			18
2.2. CIE XYZ colour space			20
2.3. Television colour spaces			22
2.3.1. YUV colour space			23
2.3.2. YIQ colour space			23
2.3.3. YC_bC_r colour space			24
2.4. Opponent colour spaces			25
2.5. Ohta colour space			27
2.6. Perceptual colour spaces			28
2.7. Perceptually uniform colour space	es		35
2.7.1. CIELAB colour space .			35
2.7.2. CIELUV colour space .			38
2.8. Evaluation of colour spaces prope	erties		39
			15
3. Colour image acquisition			40
3.1. Calibration of vision system	· · · · · · · · · · · · · · ·		41 БQ
3.2. Evaluation of colour reproduction	accuracy in the imag	ge	00
3.3. Determination of spectral response	ses of colour cameras		07
3.4. Evaluation of colourfulness for ac	quired image		70

6

Contents

4. Re	esearch on properties of selected colour spaces	8'
4.1.	HSx colour spaces	8
	4.1.1. Modifications of HSx spaces \ldots \ldots \ldots \ldots	90
	4.1.2. Comparison of RGB to HSx transforms	90
	4.1.3. Research on structure of HSx spaces	
4.2.	$K_1K_2K_3$ colour space	96
	4.2.1. <i>RGB</i> components decorrelation	98
	4.2.2. Fundamental properties of KLT	101
	4.2.3. KLT applications	105
5. Co	lour image segmentation	111
5.1.	Thresholding	116
	5.1.1. Application of knowledge about the image background	118
	5.1.2. Highlight detection	123
5.2.	Clustering in colour space	130
5.3.	Region growing	136
	5.3.1. Seeded region growing technique	138
	5.3.2. Unseeded region growing technique	142
	5.3.3. Shadows and highlights in region growing process	147
5.4.	Postprocessing	149
5.5.	Quantitative evaluation of segmentation results	155
5.6.	Further image processing after segmentation	159
$0. \mathbf{Pre}$	eprocessing of colour images	167
0.1.		167
6.2.	Denoising filtration	173
7. Sur	nmary and conclusions	187
Referen	nces	103
		100
Append	dix: CIEDE2000 colour difference formula	209
Summa	ary	211

Spis akronimów

AGC — Automatic Gain Control

AMF — Arithmetical Mean Filter	${\bf K-N}$ — Kuwahara-Nagao Filter
CCD — Charge Coupled Device	\mathbf{LCFT} — Liquid Crystal Tunable Filter
CCT — Correlated Colour Temperature	LUT — Look-Up Table
CIE — Commission Internationale de l'Eclairage	NCD — Normalized Colour Difference
CIECAM — CIE Colour Appearance Model	NTSC — National Television Standards Committee
CMOS — Complementary Metal Oxide	\mathbf{PAL} — Phase Alternation by Line
CPL Colour Pendering Index	\mathbf{PCA} — Principal Component Analysis
CHI — Control Weighted Vector	PGF — Peer Group Filter
Median Filter	\mathbf{PSNR} — Peak Signal Noise Ratio
DDT — Directional Distance Filter	RAG — Region Adjacency Graph
DPA — Digital Path Approach	ROI — Region Of Interest
DRM — Dichromatic Reflection Model	RMSE — Root Mean Square Error
${f EBU}$ — European Broadcasting Union	\mathbf{SAW} — Self-Avoiding Random Walk
EPS — Edge Preserving Smoothing	\mathbf{SMF} — Scalar Median Filter
FCC — Federal Communications	\mathbf{SNN} — Symmetric Nearest Neighbour Filter
Commission	\mathbf{SNR} — Signal Noise Ratio
FPGA — Field Programmable Gate Array	SRG — Seeded Region Growing
ITU — International Telecommunication Union	${f SVD}$ — Singular Value Decomposition
${f JND}$ — Just Noticeable Colour Difference	\mathbf{VM} — Validity Measure
KI.A — KI. Average	\mathbf{VMF} — Vector Median Filter

KLT — Karhunen-Loeve Transform

8

Spis ważniejszych oznaczeń

$\ \cdot\ $ — norma wektora lub macierzy	e_j — błąd barwy obszaru R_j
arnothing — zbiór pusty	E(I) — skupienie energii w składowych obrazu I
a^* — składowa chromatyczna przestrzeni CIELAB (czerwoność-zieloność)	η_{pq} — znormalizowany moment centralny rzędu $p + q$
$A(R_i)$ — pole powierzchni obszaru R_i b — współrzędna trójchromatyczna dla $Bk^* — alkodowa skremetwane przectrzeni$	${f f}(i,j)$ — binarna funkcja obrazowa Φ_1 — niezmiennik momentowy Hu
 CIELAB (żółtość-niebieskość) 	g — współrzędna trójchromatyczna dla G
B — barwa niebieska	G — barwa zielona
B_f, B_d — długość konturu obszaru (kod łańcuchowy)	$\gamma_R, \gamma_G, \gamma_B$ — współczynniki nieliniowości charakterystyk kamery
c_1, c_2 — wektory barw w przestrzeni BGB	H — odcień w przestrzeniach HSx
C — chroma	${f H}(R_i)$ — predykat jednorodności dla obszaru R_i
C_j — środek klastera K_j	I-intensywność barwy w przestrzeni HSI
C_b — składowa chromatyczna przestrzeni YC_bC_r	I_1, I_2, I_3 — składowe barwy przestrzeni Ohty
C_r — składowa chromatyczna przestrzeni YC_bC_r	$K(R_i)$ — współczynnik kompaktowości obszaru R_i
Ch — chromatyczność barwy	k — liczba klasterów lub współczynnik
Cy — barwa turkusowa	proporcjonalności (rozdz. 4)
d,δ,T — wartości progów	K_1, K_2, K_3 — składowe barwy obrazu po transformacji KLT
d_1, d_2, d_3 — czynniki korekcyjne punktu bieli	K_i — etykieta klastera i
D_{55} — wzorzec bieli CIE, 5500 K	L — jasność w przestrzeni HLS
D_{65} — wzorzec bieli CIE, 6500 K	L^* — jasność w przestrzeniach $CIELAB$ i
ΔE — różnica barw	CIELUV

SPIS WAŻNIEJSZYCH OZNACZEŃ

	~		-3		
$\lambda_1, \lambda_2,$	λ_3	 wartości	własne	macierzy	U

— moment rzędu p + q m_{pq}

 λ — długość fali świetlnej

- M barwność obrazu
- Mqbarwa purpurowa.
- min, μ , max funkcje wyznaczające wartości: minimalna, średnia, maksymalna
- moment centralny rzędu p + q μ_{pq}

n, N— liczba pikseli w obszarze lub na histogramie

 N_j — liczba pikseli należących do klastera K_i

 $O(R_i)$ — obwód obszaru R_i

- P pole segmentu pierścieniowego
- $\Pi(m,n)$ kwaternion barwy piksela o współrzędnych (m, n)
- Q(I) funkcja jakości obrazu I po segmentacji
- $r\,$ współrzędna trójchromatyczna dla R
- R barwa czerwona lub liczba obszarów w obrazie po segmentacii

 R_T — składnik rotacyjny formuły CIEDE2000

 ρ — promień segmentu pierścieniowego

 $\rho_{RG}, \rho_{RB}, \rho_{GB}$ — współczynniki korelacji pomiędzy składowymi RGB

- S nasycenie w przestrzeniach HSx S_L, S_C, S_H — współczynniki skalujące formuly CIE94 σ — odchylenie standardowe T — macierz transformacji Θ — kat między 2 wektorami barw u^* — składowa chromatyczna przestrzeni
- CIELUV
- U(I) liczba barw w obrazie I
- U macierz wariancyjno-kowariancyjna obrazu RGB
- v^* składowa chromatyczna przestrzeni CIELUV
- V wartość barwy w przestrzeni HSV
- Vol objętość segmentu trójwymiarowego
- x współrzędna trójchromatyczna dla X
- X składowa barwy w przestrzeni CIE XYZ
- y współrzędna trójchromatyczna dla Y
- Y składowa barwy w przestrzeni CIE XYZ (luminancja)
- Ye barwa żółta
- z współrzędna trójchromatyczna dla Z
- Z składowa barwy w przestrzeni CIE XYZ

Rozdział 1 Wprowadzenie

Traktując zagadnienie barwy bardzo ogólnie można stwierdzić, że barwa jest wieloaspektowym zjawiskiem występującym w przyrodzie i w kulturze. Podchodząc bardziej precyzyjnie do tego zagadnienia uważa się, że barwa jest wrażeniem powstającym w odpowiedzi na pobudzenie naszego systemu wzrokowego przez światło. Barwa jest więc wielkością psychofizyczna. Percepcja barwy jest oparta na działaniu trzech rodzajów czopków znajdujacych się w siatkówce ludzkiego oka, co skutkuje stosowaniem zwykle trzech składowych do opisu barwy. Barwa nie tylko dodaje piękna temu, co widzimy, ale i stanowi podstawowy środek komunikacji miedzyludzkiej oraz jeden z podstawowych czynników poznania i kształtowania środowiska człowieka. Percepcja barwy odgrywa tak samo istotną rolę w układzie wzrokowym człowieka, jak percepcja kształtu, ruchu i głębi. Barwa, obok kształtu, rozmiaru i tekstury, należy do najważniejszych czynników służących rozpoznawaniu obiektów.

Szczególna rola barwy wynika z faktu, że większość układów wzrokowych (biologicznych systemów wizyjnych) potrafi rozróżniać dużo więcej barw niż poziomów szarości. Dla ludzkiego układu wzrokowego gestość informacji pozyskiwana przy jednakowej rozdzielczości przestrzennej jest kilka rzędów wyższa dla widzenia chromatycznego niż dla widzenia achromatycznego. W klasycznej pracy Judda i Wyszeckiego [123] twierdzi się, że w przypadku porównywania barw parami w idealnych warunkach obserwacji można rozróżnić do 10 milionów barw; możliwości rozróżniania barw w jednym wielobarwnym obrazie sa zdecydowanie mniejsze, np. według [165] wynoszą około 30 tysięcy barw. Stwierdzono również, że w zmiennych warunkach oświetleniowych informacja barwna dostarcza bardziej "odpornych" (ang. robust) cech do rozpoznawania niż informacja jasnościowa [91].

Percepcja wzrokowa sprowadza się do wydzielania "użytecznej informacji" z otaczającego obserwatora świata rzeczywistego, dostarczając mu 70-90% informacji o otoczeniu.

Barwa w procesie percepcji spełnia dwie istotne funkcje [29,91]:

- bierze udział w kształtowaniu konturów obiektów,
- umożliwia postrzeganie kategoryzujące (ang. categorical perception).

Podstawą postrzegania obiektów jest postrzeganie ich kształtu poprzez percepcję konturów. W świecie rzeczywistym najczęściej jednocześnie występują różnice w składowych chromatycznych i achromatycznych (luminancyjnych) barw obiektów znajdujących się w polu widzenia. Dlatego większość konturów w obrazie jest formowana przez ich złożenie. Małe straty informacji konturowej w obrazach czarno-białych świadczą wyraźnie o dominacji konturów achromatycznych. Fakt, że przez długi czas zadowalano się fotografią, filmem i telewizją czarno-białą wskazuje na to, że dużą część informacji zawartej w obrazie barwnym można wyekstrahować z czarno-białej wersji obrazu. Występują jednak sytuacje, gdy kontury chromatyczne stanowią wyłączną podstawę do percepcji konturów. Sytuacja taka powstaje, gdy w polu widzenia znajdują się obiekty izoluminancyjne. Na ogół jednak różnice chromatyczne "dokładają się" do konturów powstałych w wyniku różnic luminancji. Barwa pełni również pomocniczą rolę w stereoskopowej percepcji głębi [122].

W życiu codziennym kojarzymy barwę z obiektami. Oczekujemy, że ten sam obiekt ma określony kolor gdziekolwiek i kiedykolwiek go spotykamy. Wśród zadań postrzegania kategoryzującego, opartych na wykorzystaniu percepcji barwy i ogólnoludzkiej wiedzy o barwie, można przykładowo wymienić:

- ocenę jakości produktów spożywczych,
- ocenę dojrzałości owoców,
- ocenę temperatury rozgrzanego przedmiotu,
- poszukiwanie określonego samochodu na parkingu,
- rozróżnianie graczy przeciwnych drużyn sportowych,
- rozpoznawanie znaków drogowych i stanów sygnalizacji świetlnej.

Istnieją zadania rozwiązywane z wykorzystaniem systemów wizyjnych, których nie da się rozwiązać bez udziału widzenia barwnego lub w których widzenie barwne ułatwia rozwiązanie. W charakterze przykładów można wymienić: rozpoznawanie barwnego kodu paskowego na elementach elektronicznych lub farmaceutykach, zliczanie barwionych komórek w preparatach, rozpoznawanie raka skóry czy rozpoznawanie położenia owoców cytrusowych na drzewie. W tym ostatnim przypadku wydaje się, że owoce mogą być rozpoznawane na podstawie: kształtu, poziomu szarości, czy barwy. Tymczasem kształt owoców ulega zmianie ze względu na zasłanianie ich przez liście. Również poziom szarości jest cechą zawodną, bo zmienia się wraz ze zmianami oświetlenia. Pozostaje więc barwa.

Barwa większości obiektów zależy od materiałów, z których są one zbudowane, a właściwie od tego, jak pochłaniają energię świetlną w określonych przedziałach widma widzialnego. Barwa nie jest jednak cechą bezwarunkowo związaną z obiektem. Zależy ona nie tylko od rozkładu widmowego współczynnika odbicia obiektu, ale również składu widmowego źródła światła i czułości widmowych receptorów oka. Barwa obiektu znajdującego się w terenie ulegnie pewnej zmianie, gdy chmury zasłonią słońce lub obiekt wskutek swego ruchu znajdzie się w cieniu. Układ wzrokowy człowieka radzi sobie z tym problemem za pomocą niezbyt jeszcze dobrze zbadanego mechanizmu stałości barwy (ang. colour constancy). Potrafimy identyfikować obiekty w scenie na podstawie ich barwy mimo zmian oświetlenia. Zjawisko stałości barwy wskazuje na to, że widzenie barwne jest procesem złożonym, a postrzeganie barwy obiektu jest bardziej zależne od charakterystyki widmowej współczynnika odbicia obiektu niż od charakterystyki widmowej źródła światła.

Barwny obraz zapewnia bogatszą informację o obiektach i dlatego jego pozyskanie i przetwarzanie może znacznie uprościć rozpoznawanie i lokalizację tych obiektów. Człowiek całe życie przetwarza obrazy barwne: albo te dostarczane mu na jawie przez układ wzrokowy, albo te wyszukiwane podczas snu w pamięci stanowiącej swoistą obrazową bazę danych. Natomiast cyfrowe przetwarzanie obrazów barwnych ma historię sięgającą jedynie ponad 30 lat wstecz [302]. Rozwój tej dziedziny wymagał kamer kolorowych, specjalnych interfejsów (ang. *framegrabbers*), wydajnych komputerów i wysokiej jakości monitorów kolorowych. Stąd też szybszy był rozwój przetwarzania obrazów binarnych i monochromatycznych (ang. *grey-level images*). Pod koniec lat siedemdziesiątych pojawiły się pierwsze podręczniki przetwarzania obrazów barwnych. W latach osiemdziesiątych zaczęto stosować systemy wizyjne przetwarzające obrazy barwne w przemyśle i biomedycynie [179] oraz w szeroko rozumianej robotyce [182, 185].

Obecnie sytuacja uległa zmianie i "bariera sprzętowa" została przełamana. Łatwy dostęp do urządzeń pozyskiwania obrazów barwnych, takich jak: skanery, cyfrowe aparaty fotograficzne i cyfrowe kamery, spowodował, że to obrazy barwne stały się głównym obiektem przetwarzania i rozpoznawania, a obrazy binarne i monochromatyczne pozosta-

ROZDZIAŁ 1. Wprowadzenie

ły istotne w niektórych zastosowaniach widzenia maszynowego, przetwarzania obrazów medycznych itp. Chociaż przetwarzanie informacji barwnej jest niewątpliwie bardziej złożone, to jednocześnie informacja ta jest bardziej kompletna. W końcu ubiegłego wieku opublikowano dwie pierwsze książki poświęcone w całości cyfrowemu przetwarzaniu obrazów barwnych [231, 248]. Równocześnie publikowane są specjalne numery czasopism naukowych na temat przetwarzania obrazów barwnych [7, 258]. Autor, jako edytor takiego podwójnego numeru *Machine Graphics and Vision* [199], zebrał 14 artykułów, których tematyka rozciąga się od zagadnień teorii retineksu przez filtrację i segmentację obrazów barwnych aż po zastosowania przetwarzania obrazów barwnych w przemyśle spożywczym. Cyfrowe przetwarzanie obrazów barwnych wiele korzysta z osiągnięć nauki o barwie (ang. *colour science*), a szczególnie techniki pomiarowej, jaką jest kolorymetria. Niektórzy [157] uważają nawet, że systemy przetwarzania obrazów barwnych stanowią uzupełnienie tradycyjnych kolorymetrów w zakresie obiektów wielobarwnych, z teksturą, niepłaskich itp. i proponują nazwę kolorymetria obrazowa (ang. *imaging colorimetry*).

Prezentowane w niniejszej rozprawie badania autora koncentrowały się wokół problemów niskopoziomowego przetwarzania obrazu (ang. *low level processing*), począwszy od zagadnienia reprezentacji barwy w obrazach, poprzez pozyskiwanie obrazu barwnego, przetwarzanie wstępne (filtracja, kwantyzacja barwy) aż po segmentację obrazu barwnego i ocenę wyników tej segmentacji. W systemach wizyjnych przetwarzających obrazy barwne na tych właśnie etapach w najwyższym stopniu przejawia się specyfika barwy. Dla wysokopoziomowego przetwarzania obrazu (ang. *high level processing*), polegającego m.in. na ekstrakcji cech, rozpoznawaniu obiektów czy interpretacji sceny barwa obszaru może stanowić jedną z wielu wykorzystywanych cech, obok kształtu, rozmiaru i tekstury. Obserwując problemy, jakie napotykają twórcy systemów przewarzania obrazów barwnych przy rozwiązywaniu zadań technicznych, medycznych i multimedialnych oraz ich dążenia w kierunku umożliwienia pełniejszego korzystania z informacji zawartej w obrazie barwnym, wyraźnie widać konieczność wykorzystania osiągnięć nauki o barwie w metodach i algorytmach przetwarzania obrazów barwnych.

Celem tej rozprawy jest przedstawienie rozwiązań w zakresie wybranych zagadnień cyfrowego przetwarzania obrazów barwnych, które to rozwiązania uzyskano w wyniku zastosowania wiedzy o właściwościach przestrzeni barw, wykorzystania zgodnych z percepcją miar różnicy barw oraz globalnych atrybutów percepcyjnych. W ten sposób potwierdza się inspirujący wpływ nauki o barwie na przetwarzanie obrazów barwnych.

W rozdziale 2 przeprowadzono przegląd kilkunastu najbardziej popularnych w dzie-

dzinie przetwarzania obrazów przestrzeni barw. Wiekszość z tych przestrzni została wykorzystana w dalszych rozdziałach rozprawy. Uwagę skupiono na transformacjach z bazowej przestrzeni *RGB* i ich podstawowych właściwościach takich, jak liniowość transformacji, stabilność obliczeniowa transformacji oraz ich równomierność percepcyjna.

W rozdziale 3 autor przedstawił zagadnienia związane z pozyskiwaniem obrazów barwnych. Badania w tym zakresie wymagały konfigurowania systemów wizyjnych oraz stosowania próbek z atlasów barw i specjalnych wzorników barwy. Zaproponowano i sprawdzono procedurę kalibracji kolorymetrycznej systemu wizyjnego. Pokazano jak standaryzacja przestrzeni *RGB* może poprawić jego wierność odwzorowania barw. Do oceny wierności stosowano najnowsze formuły różnicy barw proponowane przez Międzynarodową Komisję Oświetleniową. Ponieważ każdy pozyskany obraz barwny może zostać opisany za pomocą barwności, to definicję barwności obrazu dołączono do tego "sprzętowego" rozdziału i zbadano jej zależność od lokalnych atrybutów percepcyjnych takich jak odcień, nasycenie i jasność.

Rozdział 4 przybliża właściwości szczególnie ważnej w niniejszej rozprawie klasy przestrzeni barw objętej wspólnym oznaczeniem HSx. Pokazano w nim między innymi, że stosowanie kątowej skali odcieniowej może być źródłem problemów nawet podczas prostych działań arytmetycznych, a z drugiej strony niezmiennicze właściwości odcienia i nasycenia stanowią ważną zaletę tych przestrzeni. Przedstawiono badania doświadczalne, które pokazały przewagę wersji HSI nad 3 innymi wersjami przestrzeni percepcyjnych. Drugą przestrzenią barw omawianą w rozdziale 4 jest przestrzeń $K_1K_2K_3$ będąca wynikiem transformacji Karhunena-Loèvego na obrazie RGB, zapewniająca dekorelację składowych barwy. Wiele uwagi poświęcono efektywnej realizacji tej złożonej obliczeniowo transformacji. Zauważono również, że w zmiennych warunkach oświetleniowych kontrast w obrazie K_1 utrzymuje się znacznie dłużej niż w obrazie jasnościowym Y.

Rozdział 5, najobszerniejszy, poświęcono zagadnieniu segmentacji obrazów barwnych, które ma kluczowe znaczenie w zastosowaniach. Zakres rozpatrywanych w tym rozdziale technik ograniczono do technik pikselowych (progowanie i klasteryzacja) i obszarowych. Wiele miejsca poświęcono progowaniu obrazów barwnych w przestrzeni *HSI*, w szczególności z wykorzystaniem wiedzy o tle oraz w zastosowaniu do detekcji odblasków. Omówiono dwie wersje techniki rozrostu obszaru: ziarnistą i nieziarnistą. Ta ostatnia automatyczna technika segmentacji daje szczególnie dobre wyniki, zależne od stosowanego kryterium jednorodności wraz z progiem oraz od przestrzeni barw. Przetwarzanie końcowe np. usuwanie małych obszarów pozwala dodatkowo poprawić jakość segmentacji. Miarą jakości segmentacji mogą być specjalne funkcje oceny, których przydatność została potwierdzona w pracy.

Następny rozdział, dotyczący przetwarzania wstępnego (kwantyzacja barwy i filtracja odszumiająca), odnosi się co prawda do zagadnień poprzedzających segmentację, jednak autor zdecydował się przedstawić go naruszając kolejność odpowiadającą przetwarzaniu. U podstaw takiej zmiany kolejności był pomysł zastosowania funkcji oceny segmentacji do zbadania wpływu filtracji odszumiającej na wyniki segmentacji. Poza tym przeprowadzono badania 3 metod kwantyzacji barwy: równomiernej w RGB oraz w HSV, a także kwantyzacji metodą k-means. Do oceny wyników kwantyzacji i filtracji stosowano miarę teoriosygnałową PSNR (ang. *Peak Signal Noise Ratio*), różnicę barw ΔE oraz dodatkowo zaproponowano różnicę barwności ΔM . W rozdziale 7 podsumowano najważniejsze wyniki i osiągnięcia rozprawy.

Polska terminologia stosowana w nauce o barwie (ang. *colour science*) i przetwarzaniu obrazów barwnych nie jest jeszcze ustalona. Dlatego, aby uniknąć nieporozumień, w wielu przypadkach obok terminu polskiego umieszczono w nawiasach terminy stosowane w literaturze anglojęzycznej.

*

Autor rozprawy wyraża podziękowanie wszystkim Koleżankom i Kolegom z kraju i zagranicy, z którymi prowadził badania naukowe w dziedzinie przetwarzania obrazów barwnych. Szczególnie inspirująca w tym zakresie była współpraca z Kolegami z Instytutu Automatyki Politechniki Śląskiej, przede wszystkim z dr. hab. Bogdanem Smołką, dr. inż. Damianem Bereską, dr. inż. Arturem Balem i dr. inż. Markiem Szczepańskim.

twierdzona w pracy. Następny rozdział, dotyczący przetwarzania wstępnego (kwantyzacja barwy i filtracja

Rozdział 2

Reprezentacja barwy w obrazach cyfrowych

W przeciwieństwie do obrazów monochromatycznych, gdzie reprezentacja poziomu szarości piksela jest jednoznaczna, w przypadku barwy piksela istnieje poważny problem wyboru reprezentacji. Podstawową formą reprezentacji barwy w obrazach cyfrowych jest przestrzeń barw nazywana niekiedy modelem barw lub układem współrzędnych barw. Ponieważ układ wzrokowy człowieka wykorzystuje w celu percepcji barwy trzy receptory (czopki), to na ogół przestrzeń barw jest trójwymiarowym modelem geometrycznym, gdzie każdy wymiar (każda składowa barwy) reprezentuje jeden z jej atrybutów. Poszczególne barwy mogą być traktowane jako punkty w przestrzeni barw, a zbiór punktów reprezentujących wszystkie możliwe barwy tworzy tzw. bryłę barw. Przestrzeń barw wprowadza porządek w zbiorze kilku milionów (niektórzy twierdzą, że 10 mln [301]) postrzeganych barw. Przestrzeń barw może być pojęciem związanym z konkretnym urządzeniem technicznym (ang. device-dependent colour space) lub być niezależna od urządzenia (ang. device-independent colour space).

Poniżej przedstawiono wybór przestrzeni barw najbardziej popularnych w przetwarzaniu obrazów. Przestrzenie te opisano głównie poprzez przedstawienie równań transformacji do tych przestrzeni z bazowej przestrzeni *RGB*, nie prezentując transformacji odwrotnych. W przypadku liniowych transformacji wzory dla transformacji odwrotnych powstają przez odwrócenie macierzy transformacyjnej, a w przypadku nieliniowych transformacji ich wyprowadzenie jest bardziej złożone. Transformacje wiążą barwy w jednej przestrzeni barw z barwami w innej przestrzeni. W tym przeglądzie autor zestawił przestrzenie barw uznane za przydatne w przetwarzaniu obrazów barwnych. Dlatego opis przestrzeni uzupełniono przykładami ich zastosowań. Szeroki przegląd pozwala na porównanie właściwości tych przestrzeni dokonane w 2.2. Większość przedstawionych przestrzeni znajdzie zastosowanie w dalszych rozdziałach rozprawy.

Przestrzenie barw stosowane w przetwarzaniu obrazów barwnych mogą pochodzić z różnych dziedzin nauki i techniki. Przykładowo, modelowanie widzenia barwnego jest źródłem bazowej przestrzeni *RGB* czy przestrzeni barw opozycyjnych. Z kolei przestrzenie będące standardami CIE (franc. *Commission Internationale de l'Eclairage*) i pochodzące z kolorymetrii zostały opisane w podręcznikach z tej dziedziny techniki [301]. Przestrzenie percepcyjne (*HLS*, *HSV* itp.) są szeroko stosowane w grafice komputerowej [78], a przestrzenie *YUV* i *YIQ* w telewizji [263]. Niektóre przestrzenie zostały zaproponowane specjalnie dla celów przetwarzania obrazów, np. przestrzeń Ohty. Poniższy przegląd jest rozszerzoną wersją publikacji autora poświęconych przestrzeniom barw [180, 193].

2.1. Przestrzeń barw RGB

Idea tej przestrzeni pochodzi z trójchromatycznej teorii widzenia barw. Przestrzeń *RGB* jest przestrzenią barw najczęściej stosowaną w przetwarzaniu obrazów zarówno jako przestrzeń w urządzeniach wejściowych (kamery, skanery itp.) jak i w urządzeniach wyjściowych (monitory). Jednak należy zwrócić uwagę, że są to odrębne wersje przestrzeni *RGB*, różniące się wartościami przyjętych w nich barw podstawowych. Na przykład dla kamer pracujących w standardzie TV możemy mówić o przestrzeniach PAL *RGB* i NTSC *RGB*, a więc o przestrzeniach zdefiniowanych dla konkretnych urządzeń. Różne cyfrowe aparaty fotograficzne mogą wykorzystywać własne niestandardowe przestrzenie *RGB*.

Bryła barw w przestrzeni RGB ma kształt sześcianu (rys. 2.1). Każda barwa, opisana za pomocą składowych RGB, jest reprezentowana przez punkt leżący wewnątrz lub na powierzchni tego sześcianu. Każda barwa jest wynikiem mieszania w odpowiednich proporcjach trzech barw podstawowych RGB. Barwy monochromatyczne znajdują się na przekątnej sześcianu, zaczynając od barwy czarnej (R = G = B = 0) aż po barwę białą (R = G = B = Max). Każda ze składowych RGB może przyjmować jedną z 256 wartości z zakresu [0, 255] lub [0, 1].

Jedną z głównych wad przestrzeni RGB jest w przypadku obrazów naturalnych duża korelacja pomiędzy składowymi barwy. To czyni przestrzeń RGB słabo przydatną w kompresji obrazu. Z drugiej jednak strony to skorelowanie składowych RGB piksela jest wykorzystywane przez konstruktorów kamer z filtrem mozaikowym i procedurami interpolacyjnymi. Inne wady przestrzeni RGB to:

• nieintuicyjność psychologiczna, tj. trudno jest człowiekowi myśleć o barwie w kate-



Rys. 2.1. Reprezentacja barwy w przestrzeni *RGB* Fig. 2.1. *RGB* colour space as a method of colour representation

goriach składowych RGB,

 nierównomierność percepcyjna, tj. odległość pomiędzy barwami w przestrzeni RGB nie wyraża postrzeganej przez obserwatora różnicy barw pomiędzy nimi.

Niezbyt doświadczonemu użytkownikowi przestrzeni *RGB* trudno jest ocenić, jak postrzegana jest barwa opisana przez składowe (100, 160, 200) lub jaka jest różnica pomiędzy dwiema barwami reprezentowanymi przykładowo przez składowe: (150, 30, 30) i (150, 230, 230). Podobnie w sytuacji odwrotnej: dysponując np. próbką barwy "oliwkowej" trudno jest ustalić jej składowe *RGB*. Wymienione wady mogą być przyczyną zniekształceń barwnych, powstających w wyniku przetwarzania obrazów w przestrzeni *RGB*.

Opierając się na składowych *RGB* można zbudować nową przestrzeń poprzez wprowadzenie układu współrzędnych sferycznych:

$$L = \sqrt{R^2 + G^2 + B^2}$$

$$\Theta = \operatorname{arc} \operatorname{tg} \frac{G}{R}$$

$$\varphi = \operatorname{arc} \sin \frac{\sqrt{R^2 + G^2}}{\sqrt{R^2 + G^2 + B^2}}$$
(2.1)

Sferyczna przestrzeń barw, reprezentowana przez trójki (L, Θ, φ) , znalazła m.in. zastosowanie w segmentacji obrazów dermatologicznych [292] i w wykrywaniu barwnych znaczników, ułatwiających nawigację robota mobilnego [126].

2.2. Przestrzeń barw CIE XYZ

widzeniu barw przeciętnego człowieka (tzw. normalny obserwator kolorymetryczny) jest niezależna od urządzenia. Składowa Y reprezentuje luminancję bodźca barwnego. Składowe trójchromatyczne XYZsą związane ze składowymi ${\rm CIE}\,RGB$ za pomocą poniższych równań [301]:

$$X = 0,490R + 0,310G + 0,200B$$

$$Y = 0,177R + 0,812G + 0,011B$$

$$Z = 0,000R + 0,010G + 0,990B$$
(2.3)

W kolorymetrii na bazie składowych XYZ zaproponowano również współrzędne trójchromatyczne xyz [301]:

$$x = \frac{x}{X+Y+Z}$$

$$y = \frac{x}{X+Y+Z}$$

$$z = \frac{x}{X+Y+Z} = 1 - (x+y)$$
(2.4)

Zależności pomiędzy XYZ a CIE RGB wyprowadzono po przyjęciu dla przestrzeni ${\rm CIE}\, RGB$ barw podstawowych o następujących współrzędnych trójch
romatycznych:

$x_{B} = 0,735$	$y_R = 0,265$	(0 F)
$x_G = 0,274$	$y_G=0,717$	(2.5)
$x_B=0,167$	$y_R=0,009$	

łącznie z następującą bielą odniesienia:

$$x_W = 0,333 \qquad \qquad y_W = 0,333 \tag{2.6}$$

W przetwarzaniu obrazów zamiast przestrzeni CIERGBstosowane są inne wersje przestrzeni RGB nazywane niekiedy EBU RGB (Europa) i FCC RGB (USA), pochodzące z techniki telewizyjnej. Te przestrzenie barw są oparte na barwach podstawowych odpowiadających barwom luminoforów stosowanych w telewizji:

• EBU	$egin{aligned} x_R &= 0,640 \ x_G &= 0,290 \ x_B &= 0,150 \end{aligned}$	$y_R = 0,330$ $y_R = 0,600$ $y_R = 0,060$	(2.7)
• FCC	$x_R = 0,670$ $x_G = 0,210$ $x_B = 0,140$	$egin{aligned} y_R &= 0,330 \ y_R &= 0,710 \ y_R &= 0,080 \end{aligned}$	(2.8)

oraz na następujących bielach odniesienia, zdefiniowanych przez CIE:

• D_{65} (dla przestrzeni EBU RGB)

$$x_W = 0,313$$
 $y_W = 0,329$ (2.9)

ROZDZIAŁ 2. Reprezentacja barwy w obrazach cyfrowych

Łatwo zauważyć również, że wartość składowych RGB w obrazie jest proporcionalna do natężenia oświetlenia sceny przedstawianej przez obraz. Aby uniezależnić się od oświetlenia w kolorymetrij wprowadzono tzw. współrzedne chromatyczne (barwy znormalizowane), które wyrażają poniższe wzory [301]:

$$\dot{r} = \frac{R}{R+G+B}$$

$$\dot{r} = \frac{G}{R+G+B}$$

$$\dot{r} = \frac{B}{R+G+B} = 1 - (r+g)$$
(2.2)

Wzory (2.2) odpowiadają przekształceniu trójwymiarowej przestrzeni barw w dwuwymiarową płaszczyznę chromatyczności, tzn. wektory z przestrzeni trójwymiarowej zostaja odwzorowane na punkty na płaszczyźnie. Ponieważ współrzędne rgb są liniowo zależne, to wystarczy stosować 2 pierwsze współrzędne, a trzecia obliczać jedynie w celu weryfikacji ich sumy. Można pokazać, że dopóki nie nastąpi zmiana rozkładu widmowego światła oświetlającego scenę, to zmiany natężenia oświetlenia zasadniczo nie zmieniają wartości współrzednych rgb. Potwierdzono to eksperymentalnie w kilku pracach [2, 5, 25]. Barwy znormalizowane zastosowano m.in. w analizie zdjęć lotniczych [3], detekcji krawedzi [170] i segmentacji obrazu [105]. W czasie segmentacji obrazu współrzedne rab dla ciemnych pikseli (niskie wartości składowych RGB) są obliczane z dużym błędem wywołanym przez szum w obrazie (R + G + B jest bliskie 0). Dlatego takie piksele są często pomijane w czasie segmentacji [99].

Niektóre źródła obrazu barwnego dostarczają składowych RGB celowo poddanych wcześniej nieliniowej transformacji (tzw. korekcja gamma). W wielu zastosowaniach stanowi to problem i takie kamery wymagają linearyzacji charakterystyk [271]. Różnice pomiędzy tak rozumianymi liniowymi i nieliniowymi przestrzeniami RGB omówiono szczegółowo w pracy [231]. Transformacje prowadzące do przestrzeni barw pochodzących z techniki telewizyjnej (patrz rozdz. 2.3.) są przeprowadzane na składowych RGB z korekcją gamma.

Przestrzeń barw CIE XYZ 2.2.

Przestrzeń barw XYZ została zaproponowana przez CIE w 1931 roku w celu przedstawienia wszystkich barw widzianych przez obserwatora za pomocą składowych nieujemnych i od tego czasu jest szeroko stosowana w kolorymetrii. Jako przestrzeń oparta na

22

ROZDZIAŁ 2. Reprezentacja barwy w obrazach cyfrowych

• C (dla przestrzeni FCC RGB)

 $x_W = 0,310$ $y_W = 0,316$ (2.10)

Odpowiednie transformacje $RGB\to XYZ$ są przedstawione w podręcznikach techniki telewizyjnej [263]:

• EBU

$$X = 0,430R + 0,342G + 0,178B$$

$$Y = 0,222R + 0,707G + 0,071B$$

$$Z = 0,020R + 0,130G + 0,939B$$
(2.11)

• FCC

$$\begin{array}{rcl} X &=& 0,607R+0,174G+0,200B \\ Y &=& 0,299R+0,587G+0,114B \\ Z &=& 0,000R+0,066G+1,116B \end{array} \tag{2.12}$$

W przetwarzaniu obrazów barwnych przestrzeń CIE XYZ ma znaczenie pomocnicze i jest często stosowana jako przestrzeń pośrednia w procesie wyznaczania składowych przestrzeni percepcyjnie równomiernej CIELAB lub CIELUV. W przypadku pozyskiwania obrazu barwnego za pomocą niestandardowej kamery lub cyfrowego aparatu fotograficznego często istnieje potrzeba wyznaczenia swoistej macierzy transformującej składowe RGB do przestrzeni CIE XYZ.

2.3. Telewizyjne przestrzenie barw

Przestrzenie barw pochodzące z telewizji (np. YUV, YIQ) są w rzeczywistości przestrzeniami barw opozycyjnych (patrz rozdz. 2.4.), ponieważ definiują składową luminancyjną (Y, niekoniecznie równą CIE Y) oraz dwie składowe chrominancyjne oparte na sygnałach różnicy barw: R - Y i B - Y. Przestrzenie te zostały zaprojektowane w celu minimalizacji pasma zespolonego sygnału wizyjnego, tj. umieszczenia sygnałów różnicowych barwy w paśmie istniejącej telewizji czarno-białej.

Układ wzrokowy człowieka jest znacznie mniej wrażliwy na szczegóły przestrzenne w przypadku składowych chrominancyjnych niż luminancyjnych [263]. Z tego powodu pasmo przeznaczone dla sygnałów chrominancji jest znacznie węższe niż w przypadku sygnału luminancji, np. w przypadku przestrzeni YIQ podział pasma wygląda następująco: Y – 4 MHz, I - 1,5 MHz i Q - 0,5 MHz. Zaletą telewizyjnych przestrzeni barw jest liniowa transformacja z RGB oraz oddzielenie informacji luminancyjnej od chrominancyjnej (przestrzeń luminacja-chrominancje [65]).

2.3. Telewizyjne przestrzenie barw

2.3.1. Przestrzeń barw YUV

W Europie stosuje się przestrzeń barw YUV, która jest podstawą telewizyjnego systemu kodowania sygnału PAL [263]:

$$Y = 0,299R + 0,587G + 0,114B$$

$$U = 0,493(B - Y) = -0,147R - 0,289G + 0,437B$$

$$V = 0.877(R - Y) = 0,615R - 0,515G + 0,100B$$
(2.13)

Przestrzeń YUV może być przekształcona w przestrzeń barw typu HSI (patrz rozdz. 2.6.):

$$H_{YUV} = \arctan \frac{V}{U}$$
(2.14)

$$S_{YUV} = \sqrt{U^2 + V^2}$$
 (2.15)

Niektóre techniki segmentacji są stosowane na płaszczyźnie UV [75]. Przestrzeń YUV jest stosowana głównie do kodowania pojedynczych obrazów i ich sekwencji. Ponieważ chrominancja w sekwencji obrazów zmienia się znacznie mniej niż luminancja, to w procesie kodowania takiej sekwencji można proponować następujący przydział danych: 80% dla składowej Y oraz po 10% dla składowych chrominancyjnych U i V [164]. Specjalne układy scalone przeznaczone do kompresji wideo stosują składowe YUV ze zredukowaną rozdzielczością przestrzenną dla składowych chrominancyjnych $(YUV \ 4 : 2 : 2, YUV \ 4 : 1 : 1).$

2.3.2. Przestrzeń barw YIQ

Przestrzeń barw YIQ opartą na amerykańskim systemie kodowania sygnału telewizyjnego NTSC można wyrazić poprzez składowe FCC RGB [19]:

$$Y = 0,299R + 0,587G + 0,114B$$

$$I = 0,74(R - Y) - 0,27(B - Y) = 0,596R - 0,274G - 0,322B$$

$$Q = 0,48(R - Y) + 0,41(B - Y) = 0,211R - 0,523G + 0,312B$$
(2.16)

Składowa Y (luminancja) jest identyczna z Y w przestrzeni XYZ, a składowa I (ang. *in-phase*) (oś pomarańczowo-turkusowa) wraz ze składową Q (ang. *quadrature*) (oś purpurowo-zielona) wyrażają wspólnie odcień i nasycenie. Po wprowadzeniu układu współrzędnych walcowych przestrzeń YIQ przekształca się w przestrzeń typu HSI (patrz rozdz. 2.6.):

$$H_{YIQ} = \operatorname{arc} \operatorname{tg} \frac{Q}{I}$$

$$S_{YIQ} = \sqrt{I^2 + Q^2}$$
(2.17)

Ponieważ producenci kamer niekiedy wyprowadzają na ich wyjścia dodatkowo sygnały YUV lub YIQ, to w niektórych zastosowaniach wykorzystuje się te przestrzenie. I tak przestrzeń YIQ zastosowano m.in. w:

ROZDZIAŁ 2. Reprezentacja barwy w obrazach cyfrowych

- rozpoznawaniu obiektów barwnych [264],
- segmentacji obrazów barwnych [141],
- detekcji krawedzi [269].
- kodowaniu obrazów barwnych [176],
- detekcji twarzy w obrazach barwnych [55].

2.3.3. Przestrzeń barw YC_bC_r

Ta przestrzeń barw została zdefiniowana specjalnie dla potrzeb standardowej telewizji cyfrowej (625/525 linii) przez ITU (ang. International Telecommunication Union) [240] i jest wyrażana w następującej postacj:

$$Y = 0,299R + 0,587G + 0,114B$$

$$C_b = 0,564(B - Y) = -0,169R - 0,331G + 0,500B$$

$$C_r = 0,713(R - Y) = 0,500R - 0,419G - 0.081B$$
(2.18)

Można zauważyć, że ponieważ składowa luminancyjna jest oparta głównie na składowej zielonej G, to dwie składowe chrominancyjne wykorzystują pozostałe składowe: niebieską Bi czerwoną R (rys. 2.2). Świadczy o tym fakt przyporządkowania tym barwom najjaśniejszych poziomów szarości (rys 2.2(f), (g)). W tej transformacji po wyliczeniu składowej Y można ją zastosować do obliczenia składowych chrominancji. Jeżeli wartości RGBzawarte są w przedziale [0, 255], to wartości C_b i C_r przyjmują wartości z przedziału [-128, 127]. Dlatego niekiedy, aby również te ostatnie składowe utrzymać w zakresie [0,255],dodaje się do prawych stron równań dla C_b i C_r składnik 255.

W porównaniu do wyjściowej przestrzeniRGBkorelacja pomiędzy składowymi w przestrzeni YC_bC_r ulega zm
niejszeniu. Dlatego przestrzeń ta znalazła szerokie zastosowanie w kodowaniu pojedynczych obrazów - formaty JPEG (ang. Joint Photographic Expert Group) i JPEG2000 oraz sekwencji obrazów - format MPEG-2 (ang. Moving Pictures Expert Group) [118]. Duża kompresja jest osiągana m.in. przez redukcję rozdzielczości przestrzennej chrominancji np. używając schematu próbkowania redukcyjnego (ang. subsampling) 4 : 2 : 2 oraz przez zgrubną kwantyzację składowych C_b
i C_r (np. na 4 bitach każda zamiast 8). Liniowe transformacje składowych barwnych oparte na macierzach z elementami niecałkowitymi wprowadzają błędy zaokrągleń, co jest szczególnie istotne w kodowaniu obrazów [67].



- Rys. 2.2. Obraz KLOCKI1 i jego składowe: (a) oryginał, (b) składowa R, (c) składowa G, (d) składowa B, (e) składowa Y, (f) składowa $C_b,$ (g) składowa $C_r.$ Barwna wersja rysunku, patrz strona 43
- Fig. 2.2. Image KLOCKI1 and its colour components: (a) original, (b) component R, (c) component G, (d) component B, (e) component Y, (f) component C_b , (g) component C_r . Colour version of image, see page 43

Przestrzeń barw opozycyjnych 2.4.

Przestrzeń barw opozycyjnych, nazywana niekiedy przestrzenią fizjologiczną [232], powstała w wyniku zainteresowania procesami widzenia barw opozycyjnych (ang. colour opponent processes). W drugiej połowie XIX wieku niemiecki fizjolog Ewald Hering stworzył teorię barw opozycyjnych, która była w stanie wytłumaczyć niektóre zjawiska barwne niewytłumaczalne na gruncie klasycznej teorii trójchromatycznej [106]. Te zjawiska to np., iluzje zwane powidokami (iluzoryczne plamy w barwach dopełniających) i obserwacja, że nie istnieją barwy czerwono-zielona i żółto-niebieska, chociaż istnieją barwy czerwono-niebieska i żółto-zielona. Hering założył więc, że barwa czerwona jest "opozycyjna" w stosunku do zielonej, a barwa niebieska w stosunku do żółtej. W przyjętym przez niego modelu układu wzrokowego występują 3 następujące kanały (procesy) opozycyjne: czerwono-zielony RG, żółto-niebieski YeB oraz kanał achromatyczny (biało-czarny) WhBl.

Wiele lat później potwierdzono istnienie w siatkówce i w innych częściach układu wzrokowego komórek opozycyjnych. Opracowany w latach dziewięćdziesiątych XX wieku trójstopniowy model widzenia barwnego zawiera stopień barw opozycyjnych [293]. Wyni-kiem różnej ilości poszczególnych czopków w siatkówce (najwięcej czerwonych, najmniej niebieskich) jest różna rozdzielczość przestrzenna poszczególnych kanałów opozycyjnych: największa dla kanału achromatycznego a najmniejsza dla kanału żółto-niebieskiego. W modelu tym przyjęto, że barwa nie może być jednocześnie czerwona i zielona albo żółta i niebieska, ponieważ te barwy są opozycyjne względem siebie. Przedstawione poniżej równania i rysunek pokazują sposób transformacji sygnałów pochodzących z czopków *RGB* w barwy opozycyjne (rys. 2.3).

$$RG = R - G$$

$$YeB = 2B - R - G$$

$$WhBl = R + G + B$$

$$(2.19)$$



Rys. 2.3. Schemat stopnia barw opozycyjnych w układzie wzrokowym Fig. 2.3. Schematic representation of opponent colour stage of human visual system

Przestrzeń barw opozycyjnych znalazła liczne zastosowania w przetwarzaniu obrazów barwnych:

• procesor do detekcji krawędzi w obrazach barwnych [156],

- 2.5. Przestrzeń Ohty
 - barwna stereowizja [32],
 - kodowanie sekwencji obrazów barwnych [296],
 - rozpoznawanie obiektów wielobarwnych [31,275,299].

To ostatnie zastosowanie opiera się na wykorzystaniu trójwymiarowych i dwuwymiarowych histogramów barw opozycyjnych do opisu obiektów wielobarwnych. Niekiedy modyfikuje się równania transformacji definiującej przestrzeń. Na przykład w pracy dotyczącej rozpoznawania barwnych liter [304] stosuje się przestrzeń barw opozycyjnych ze zmodyfikowanym drugim równaniem:

$$YeB = 0, 4(R+G) - B$$
(2.20)

Ponieważ odpowiedzi czopków w układzie wzrokowym człowieka są proporcjonalne do logarytmów intensywności bodźców, to opracowano również "logarytmiczną" wersję przestrzeni barw opozycyjnych [77]:

$$RG = \log R - \log G$$

$$YeB = \log B - (\log G + \log R)/2$$

$$WhBl = \log G$$

(2.21)

Takie "logarytmiczne" podejście ma długą tradycję w przetwarzaniu obrazów barwnych [73]. Niekiedy składowe w przestrzeni barw opozycyjnych oznacza się AC_1C_2 , gdzie A to składowa achromatyczna, a 2 pozostałe są składowymi chromatycznymi opozycyjnymi.

2.5. Przestrzeń Ohty

Przestrzeń zaproponowana specjalnie dla przetwarzania obrazów jest wynikiem liniowej transformacji składowych [172]:

$$I_{1} = (R + G + B)/3$$

$$I_{2} = (R - B)/2$$

$$I_{3} = (2G - R - B)/4$$
(2.22)

Składowa I_1 reprezentuje intensywność, a I_2 i I_3 reprezentują informację chromatyczną. Ohta wyprowadził te zależności w wyniku poszukiwania całkowicie niezależnych statystycznie składowych barw. Przestrzeń Ohty jest dobrą aproksymacją transformacji Karhunena-Loevego, która zapewnia pełną dekorelację składowych RGB [235].

ROZDZIAŁ 2. Reprezentacja barwy w obrazach cyfrowych

Przestrzeń Ohty znalazła zastosowanie między innymi w:

- segmentacji obrazów barwnych [141],
- kompresji obrazów barwnych [59],
- znajdowaniu odpowiedniości w parze obrazów stereo [133],
- rozpoznawaniu barw w procesie pomiaru grubości struktur mikroelektronicznych [18],
- analizie obrazów wątroby [272].

2.6. Percepcyjne przestrzenie barw

W układzie wzrokowym człowieka czopki absorbują światło z największą czułością w zielonej, czerwonej i niebieskiej części widma, aby następnie dokonać dalszego przetwarzania informacji barwnej. Obserwatorowi bez specjalnego treningu trudno jest ustalić udział poszczególnych składowych RGB w danej barwie. Ten sam obserwator bez problemu rozpoznaje podstawowe percepcyjne atrybuty barwy: odcień H, nasycenie S oraz intensywność I (jaskrawość B, jasność L). Dlatego przestrzenie barw oparte na tych atrybutach można nazywać percepcyjnymi przestrzeniami barw. Ze względu na ścisły związek z percepcją barw zajmują one szczególne miejsce w przetwarzaniu obrazów barwnych. Pierwszym, który zaproponował system opisu barw oparty na wyżej wymienionych pojęciach percepcyjnych, był amerykański malarz Albert Munsell (1905).



Odcień H, który jest definiowany ostensywnie, odpowiada wrażeniu związanemu z długością fali dominującej bodźca barwnego i na ogół od niego pochodzi nazwa barwy. Składowa H jest kątem z zakresu $[0^{\circ}, 360^{\circ}]$ mierzonym w kierunku przeciwnym do ruchu wskazówek zegara (rys. 2.4).

Na ogół przyjmuje się, że wartość $H = 0^{\circ} = 360^{\circ}$ odpowiada odcieniowi czerwonemu. Nasycenie odpowiada natomiast względnej czystości barwy i w przypadku czystej barwy wynosi ono 100%. Na przykład dla "żywej" czerwieni S = 100%, a dla bladej czerwieni (różowy) S = 50%. Barwy z zerowym nasyceniem to poziomy szarości. Intensywność barwy obiektu, achromatyczna składowa barwy, jest miarą energii światła odbitego przez obiekt w całym zakresie widma widzialnego. Maksymalna intensywność jest postrzegana jako biel, a minimalna jako czerń. Oko ludzkie rozróżnia ponad 100 odcieni, od kilkudziesięciu do ponad 100 poziomów intensywności i około 20 poziomów nasycenia, zależnie od odcienia barwy. Zmiany intensywności mają miejsce wzdłuż osi bryły, zmiany odcienia wzdłuż obwodu przekroju poprzecznego tej bryły, a zmiany nasycenia wzdłuż promienia od osi do powierzchni bryły.

Składowe HSI są obliczane na podstawie wzorów wyrażających w przybliżeniu psychofizyczny sens tych pojęć. Wzory te są przedstawiane w literaturze w różnej formie. Dotyczy to szczególnie odcienia. Wynika to z chęci osiągnięcia kompromisu pomiędzy dokładnością transformacji a jej złożonością obliczeniową. Wzór, wyprowadzony przez Tenenbauma i in. [282], i traktowany jako bazowy [100] jest złożony i dlatego często cytowany z błędami. Jego poprawna postać dana jest przez:

$$H = \arccos \frac{2r - g - b}{\sqrt{6\left[\left(r - \frac{1}{3}\right)^2 + \left(g - \frac{1}{3}\right)^2 + \left(b - \frac{1}{3}\right)^2\right]}}$$
(2.23)

jeżeli b > g, to $H := 360^{\circ} - H$

Poniżej przedstawiono wersję dla składowych RGB:

2.6. Percepcyjne przestrzenie barw

$$H = \arccos \frac{\frac{1}{2} \left[(R - G) + (R - B) \right]}{\sqrt{(R - G)^2 + (R - B) (G - B)}}$$
jeżeli B > G, to H := 360° - H
$$(2.24)$$

Dokładne wyprowadzenie powyższych wzorów, wynikające z transformacji kartezjańskiego układu współrzędnych RGB na układ współrzędnych walcowych HSI, przedstawiono w pracach [95, 136]. Kender zaproponował "szybszą" wersję wzoru dla H, która wymaga mniej mnożeń i nie stosuje pierwiastka kwadratowego [129]: ROZDZIAŁ 2. Reprezentacja barwy w obrazach cyfrowych

if R = G = B then {przypadek achromatyczny} H := niezdefiniowanyelse {przypadek chromatyczny} if R > B and G > B then $\{B = \min(R, G, B)\}$ $H = \frac{\pi}{3} + \operatorname{arctg}\left[\frac{\sqrt{3}\left(G - R\right)}{\left(G - B\right) + \left(R - R\right)}\right]$ else if G > R then (2.25) $\{R = \min(R, G, B)\}$ $H = \pi + \operatorname{arc} \operatorname{tg} \left[\frac{\sqrt{3} \left(B - G \right)}{\left(B - R \right) + \left(G - R \right)} \right]$ else $\{G = \min(R, G, B)\}$ $H = \frac{5\pi}{3} + \arccos\left[\frac{\sqrt{3}(R-B)}{(R-G) + (B-G)}\right]$ end end end

W literaturze można znaleźć prostsze wzory do obliczania odcienia, ale wyniki uzyskane za ich pomocą różnią się od wyników pochodzących ze wzorów Tenenbauma i Kendera. Przykładem jest wersja transformacji zaproponowana w 1985 przez Bajona i in., w której wzory do obliczania odcienia zostały uproszczone i już nie zawierają funkcji trygonometrycznych [12]:

if
$$\min(R, G, B) = B$$
 then
 $H = \frac{G - B}{3(R + G - 2B)}$
elseif $\min(R, G, B) = R$
 $H = \frac{B - R}{3(G + B - 2R)} + \frac{1}{3}$
elseif $\min(R, G, B) = G$
 $H = \frac{R - G}{3(R + B - 2G)} + \frac{2}{3}$
end

Ta transformacja znalazła zastosowanie w widzeniu komputerowym [13, 158].

S

Nasycenie ${\cal S}$ jest obliczane najczęściej według jednego z dwóch równoważnych wzorów:

$$= 1 - 3\min(r, g, b)$$
(2.27)

2.6. Percepcyjne przestrzenie barw

oraz

$$S = 1 - 3\frac{\min(R, G, B)}{(R + G + B)}$$
(2.28)

Istnieje wersja wzoru do obliczania nasycenia, która zmniejsza niestabilność obliczeń wokół punktu czerni S(0,0,0) [96]:

if
$$I \leq \frac{1}{3}I_{max}$$
 then
 $S = I_{max} - 3\frac{\min(I_{max} - R, I_{max} - G, I_{max} - B)}{3 - (R + G + B)}$
else
 $S = I_{max} - 3\frac{\min(R, G, B)}{(R + G + B)}$
end
(2.29)

Intensywność I jest obliczana według następującego wzoru:

$$I = \frac{1}{3} \left(R + G + B \right) \tag{2.30}$$

lub rzadziej jako:

$$I = R + G + B \tag{2.31}$$

Stosowanie wzoru dla I jest znacznym uproszczeniem. Przykładowo barwa żółta jest postrzegana jako 6-krotnie jaśniejsza od barwy niebieskiej, a ze wzoru (2.30) wynika tylko 2-krotnie większa jasność tej barwy.

Dwie inne przestrzenie percepcyjne, zbliżone do HSI, stosowane są w grafice komputerowej i nazywane modelami barw: HSV i HLS [78]. Bryły barw w obu przestrzeniach pochodzą od sześcianu RGB oglądanego z początkowego punktu układu współrzędnych wzdłuż kierunku linii poziomów szarości, z którą pokrywają się ich osie symetrii. Obie bryły są oparte na sześciokącie foremnym z barwami R, Ye, G, Cy, B, Mg w jego wierzchołkach. Zaletą HLS jest jej symetria w odniesieniu do jasności i ciemności (ang. darkness); tej cechy nie posiada przestrzeń HSV. Przestrzenie HSV i HLS różnią się wzorami do obliczeń składowej achromatycznej i nasycenia (rys. 2.5):

$$V := max$$
if $max \neq 0$ then
$$S := \frac{max - min}{max}$$
else
$$S := 0$$
end

(2.32)

2.6. Percepcyjne przestrzenie barw



33

• HLS

 $L := \frac{max + min}{2}$ if max = min then S := 0else if $L \le 0.5$ then $S := \frac{max - min}{max + min}$ else $S := \frac{max - min}{2 - max - min}$ end end

ROZDZIAŁ 2. Reprezentacja barwy w obrazach cyfrowych



- Rys. 2.5. Bryły barw: (a) HSV (ostrosłup sześciokatny), (b) HLS (podwójny ostrosłup sześciokątny)
- Fig. 2.5. Colour solids: (a) HSV (hexagonal pyramid), (b) HLS (double hexagonal pyramid)

Obydwie przestrzenie odnoszą się do składowych znormalizowanych $r,\ g,\ b.$ Tak więc można zapisać dla nich:

$$min := \min(r, g, b)$$

$$max := \max(r, g, b)$$
(2.34)

Odcień oblicza się, jednakowo dla obu przestrzeni *HSV* i *HLS*, według poniższej procedury:



Te dwa modele pokazują, że barwy stają się mniej nasycone, gdy cecha achromatyczna osiąga swój minimalny lub maksymalny poziom, np. bryła barw HLS ulega zwężeniu na obu końcach skali jasności. Jednak nie wszystkie zjawiska związane z percepcją barw zostały uwzględnione w tych modelach. W pełni nasycone barwy o różnych odcieniach mają jednakowe wartości V = 1 w przypadku HSV i te same wartości jasności L =0,5 w przypadku przestrzeni HLS. Natomiast w rzeczywistości w pełni nasycona barwa żółta jest zawsze jaśniejsza od w pełni nasyconej barwy niebieskiej. Obie przestrzenie są stosowane również w widzeniu komputerowym: [144, 236, 281]. Modele barw oparte na składowych percepcyjnych zostały zintegrowane w jeden uogólniony model barw: GLHS[139].

Wśród ważnych zalet przestrzeni HSI można wymienić jej dobrą zgodność z ludzką intuicją, rozdzielenie barw chromatycznych od achromatycznych oraz możliwość stosowania tylko jednej cechy (H) w segmentacji i rozpoznawaniu obiektów. Zamiast przeprowadzania operacji z wykorzystaniem wszystkich 3 składowych RGB można ją przeprowadzić szybciej na jednej składowej H [85, 246]. Przykładem może być detekcja przez system wizyjny robota ogrodniczego żółtych kwiatów pomidora [246]. Odcień jest w wielu przypadkach niezależny od cieni i orientacji obiektu [228]. Składowe H i S są szczególnie przydatne w rozpoznawaniu obiektów, opisując obiekty barwne w sposób niezależny od zmian intensywności.

Jednak przestrzeń HSI posiada też pewne istotne wady (patrz rozdz. 4.1.):

32

(2.35)

- ROZDZIAŁ 2. Reprezentacja barwy w obrazach cyfrowych
- nieusuwalne osobliwości transformacji $RGB \rightarrow HSI$ (*H* dla wszystkich barw achromatycznych oraz w przypadku *S* dla czerni),
- \bullet wrażliwość składowych HSIna niewielkie odchylenia wartości RGBw pobliżu punktów osobliwych,
- percepcyjna nierównomierność przestrzeni mimo percepcyjnego charakteru przestrzeni,
- problemy z operacjami na wartościach kątowych odcienia, np. obliczanie średniego odcienia [52].

Często celem stosowania przestrzeni HSI jest rozdzielenie barw chromatycznych od achromatycznych na podstawie wartości nasycenia S [13, 96]. Składowe przestrzeni HSIsą przydatne w procesach poprawiania obrazów barwnych, np. w filtracji medianowej [312] czy w poprawie kontrastu [130, 284, 305]. Systemy wizyjne wykorzystujące zalety przestrzeni HSI zastosowano do następujących zadań:

- identyfikacja elementów elektronicznych z kodem paskowym [13,96],
- kontrola jakości mieszanek warzywnych [85],
- zrobotyzowane zbieranie owoców [264],
- manipulacje kulturami bakteryjnymi [158],
- sortowanie i klasyfikacja owoców [127],
- rozpoznawanie znaków drogowych [128],
- rozpoznawanie skrzynek z napojami [162].

Z powodu istnienia wielu wariantów transformacji definiujących percepcyjne przestrzenie barw można mówić o przestrzeniach typu HSI, niekiedy oznaczanych jako HSx. Ta klasa przestrzeni zawiera również przestrzenie powstałe w wyniku podwójnej transformacji, np. $RGB \rightarrow YUV \rightarrow HSI$ [25], $RGB \rightarrow YIQ \rightarrow HSI$ [264] lub $RGB \rightarrow I_1I_2I_3 \rightarrow$ HSI [246]. Ciągle powstają nowe wersje przestrzeni percepcyjnych, np. Yagi i in. [303] zaproponowali dwie nowe wersje przestrzeni HSV, tak zmieniając wzory na nasycenie i intensywność, aby powstała przestrzeń barw lepiej opisywała ludzką percepcję barwy. Nowe wersje okazały się lepsze podczas segmentacji obszarowej zdjęć lotniczych niż 3 "klasyczne" wzory transformacyjne.

2.7. Percepcyjnie równomierne przestrzenie barw

2.7. Percepcyjnie równomierne przestrzenie barw

Przestrzeń barw jest percepcyjnie równomierna, gdy stosowana w niej miara podobieństwa barw jest zgodna z ich podobieństwem percepcyjnym. Chociaż percepcyjna przestrzeń *HSI* bardziej intuicyjnie opisuje barwę niż np. przestrzeń *RGB*, to jednak oparta jest ona na uproszczonych założeniach i nie jest przestrzenią percepcyjnie równomierną [243]. Od dawna próbowano zbudować przestrzeń barw, w której jednakowym różnicom barw postrzeganym przez obserwatora odpowiadałyby jednakowe odległości euklidesowe.

Takie przestrzenie zaproponowała CIE w 1976 roku. Jest to tzw. przestrzeń CIELABze składowymi $L^*a^*b^*$ (stosowana głównie dla światła odbitego) i tzw. przestrzeń CIELUVze składowymi $L^*u^*v^*$ (stosowana głównie dla światła emitowanego) [301]. Obydwie przestrzenie należą do klasy przestrzeni niezależnych od urządzeń, w których są stosowane. Percepcyjnie równomierna przestrzeń barw może być również zrealizowana materialnie w postaci zbioru dyskretnych próbek, np. atlas Munsella. CIE, budując przestrzenie LUVi LAB wykorzystała wiedzę o procesach opozycyjnych w percepcji barw i zorganizowała te przestrzenie wokół 3 par barw opozycyjnych: czarny-biały, czerwony-zielony i żółtyniebieski.

W ostatnich latach w przetwarzaniu obrazów częściej stosuje się przestrzeń *CIELAB*. Przestrzeń ta jest wykorzystywana dla porównywania dwóch obrazów barwnych oraz w systemach zarządzania barwą (ang. *colour management system*). Chociaż przestrzeń *CIELAB* jest standardem CIE, to jej równomierność percepcyjna jest przybliżona i dlatego CIE niedawno zaproponowała wyrażanie różnicy barw za pomocą odległości nieeuklidesowych, np. CIE94, CIEDE2000.

2.7.1. Przestrzeń CIELAB

Przestrzeń CIELAB jest zdefiniowana za pomocą następujących zależności:

$$L^{*} = 116 f\left(\frac{Y}{Y_{0}}\right) - 16$$

$$a^{*} = 500 \left[f\left(\frac{X}{X_{0}}\right) - f\left(\frac{Y}{Y_{0}}\right)\right]$$

$$b^{*} = 200 \left[f\left(\frac{Y}{Y_{0}}\right) - f\left(\frac{Z}{Z_{0}}\right)\right],$$

(2.36)

gdzie pomocnicza funkcja f ma nastepującą postać:

$$f(q) = q^{\frac{1}{3}} \text{ jeżeli } q > 0,008856$$

$$f(q) = 7,787q + \frac{16}{116} \text{ jeżeli } q \le 0,008856$$
(2.37)

 L^* oznacza wartość jasności (ang. *lightness*), (rys. 2.6), która uwzględnia nieliniową zależność od luminancji. Składowa a^* oznacza cechę barwną na skali czerwoność-zieloność, a

 $\mathbf{34}$



Rys. 2.6. Układ składowych przestrzeni *CIELAB* Fig. 2.6. Coordinates of *CIELAB* colour space

składowa b^* cechę na skali żółtość-niebieskość (patrz rozdz. 2.4.). Ponieważ obie składowe chromatyczne przyjmują wartości od ujemnych do dodatnich, to punkty w przestrzeni *CIELAB* z wartościami a^* i b^* równymi 0 odpowiadają barwom szarym. Wartości z indeksem 0 oznaczają barwę traktowaną jako biel odniesienia. Przykładowo dla standardowej bieli odniesienia D_{65} , zdefiniowanej przez CIE, odpowiednie składowe wynoszą: $X_0 = 95,05, Y_0 = 100, Z_0 = 108,91$. Najczęściej stosowane zakresy zmienności składowych to od 0 do 100 dla składowej L^* oraz od -128 do 127 dla a^* i b^* .

Dysponując programem transformującym $RGB \rightarrow XYZ \rightarrow CIELAB$ można, dokonując obliczeń składowych CIELAB dla wszystkich 16,7 mln barw z sześcianu RGB, wyznaczyć dokładnie ich wartości graniczne. Wyniki te zależą od przyjętej przez nas macierzy transformującej RGB na XYZ oraz zastosowanego wzorca bieli. Stosując dla RGB standard sRGB (rozdz. 3.2.) oraz iluminant D_{65} jako wzorzec bieli a następnie przekształcając wzór (3.17), otrzymujemy:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0,4124 & 0,3576 & 0,1805 \\ 0,2126 & 0,7152 & 0,0722 \\ 0,0193 & 0,1192 & 0,9505 \end{bmatrix} \begin{bmatrix} R_{sRGB} \\ G_{sRGB} \\ B_{sRGB} \end{bmatrix}$$
(2.38)

Poddając składowe XYZ standardowej transformacji wyznaczamy wartości graniczne CIELAB przedstawione w tab. 2.1.

Zgodnie z przewidywaniami wartości graniczne luminancji L^* dotyczą bieli i czerni w RGB. Z tabeli 2.1 wynika jednak, że maksymalna wartość składowej a^* jest osiągana nie przez barwę czerwoną, lecz purpurową. Drugi wniosek wynikający z tego eksperymentu

2.7. Percepcyjnie równomierne przestrzenie barw

Tabela 2.1. Wartości graniczne składowych CIELAB

Składowa	Wartość	\overline{R}	G	B
L _{min}	0,00	0	0	0
Lmax	99,99	1	1	1
a _{min}	-86,14	0	1	0
amax	98,29	1	0	1
b_{min}	-107,84	0	0	1
bmax	94,50	1	1	0

dotyczy składowej b^* (żółtość-niebieskość). Rozpiętość bryły barw wzdłuż tej osi jest zdecydowanie większa niż rozpiętość wzdłuż osi a^* . Tak więc do opisu bryły barw *CIELAB* należy używać elipsoidy, a nie kuli.

Do wyznaczania różnicy barw stosuje się odległość euklidesową:

$$\Delta E_{ab}^* = \sqrt{(\Delta L^*)^2 + (\Delta a^*)^2 + (\Delta b^*)^2}$$
(2.39)

Przedstawiona różnica barw jest używana do oceny jakości reprodukcji obrazów barwnych, np. po kompresji stratnej [59]. Odpowiada ona ludzkiej ocenie różnicy barw, a jej wartość równa $\Delta E_{ab}^* = 2,3$ reprezentuje średni próg postrzegalności różnicy barw JND, (ang. Just Noticeable colour Difference), [148].

W przestrzeni *CIELAB* można wprowadzić współrzędne walcowe i zdefiniować wielkości odpowiadające odcieniowi i chromie:

$$L^{*} = L^{*}$$

$$H^{*}_{ab} = \operatorname{arc} \operatorname{tg} \frac{b^{*}}{a^{*}}$$

$$C^{*}_{ab} = \sqrt{a^{*2} + b^{*2}}$$
(2.40)

Percepcyjnie równomierne przestrzenie z metryką Euklidesa są szczególnie przydatne w segmentacji obrazów scen naturalnych technikami klasteryzacyjnymi. Dotyczy to zarówno wersji ze współrzędnymi prostokątnymi [151,285] jak również walcowymi [14,36]. Odcień H^* jest stosowany jako podstawowa cecha w wykrywaniu obszarów o określonej barwie [93]. W systemach działających w czasie rzeczywistym i wykorzystujących przestrzeń CIELAB duże znaczenie ma czas realizacji nieliniowych wzorów transformacyjnych: można stosować numeryczne aproksymacje funkcji pierwiastkowej lub tabele pośrednie LUT (ang. Look-Up Table) [48,49]. W pracy [288] zastosowano nierównomierną kwantyzację dla poszczególnych składowych przestrzeni CIELAB i $L^*H^*C^*$.

Wzór na różnicę barw w przestrzeni *CIELAB* jest stosowany przy założeniu, że porównywana jest barwa jednolitych, względnie dużych obszarów barwnych. Jednak w obrazach

2.8. Ocena właściwości przestrzeni barw

ROZDZIAŁ 2. Reprezentacja barwy w obrazach cyfrowych

występują również wysokie częstotliwości przestrzenne (małe obszary barwne). Z drugiej strony oko obserwatora działa jak przestrzenny filtr dolnoprzepustowy, szczególnie dla kanałów chrominancji. Stosowanie dla obrazów miary ΔE_{ab}^* liczonej na zasadzie porównywania barw pomiędzy pikselami może wprowadzać dodatkowy błąd. Dlatego w przetwarzaniu obrazów zaproponowano wersję przestrzenną *CIELAB*, zwaną *S-CIELAB* (ang. *spatial CIELAB*) [310]. Obraz barwny wyrażony przez składowe CIE XYZ jest transformowany do przestrzeni barw opozycyjnych, w której odbywa się filtracja. Dla każdej składowej stosowane są przestrzenne (dwuwymiarowe) filtry Gaussa o różnych parametrach. Następnie wraca się do przestrzeni XYZ, aby przejść do *CIELAB* i stosować standardowy wzór na różnicę barw lub jego nowsze wersje (CIE94, CIEDE2000).

2.7.2. Przestrzeń CIELUV

Przestrzeń *CIELUV* jest rekomendowana przez CIE do zastosowań opartych na addytywnym mieszaniu świateł (np. monitory) i jest definiowana za pomocą następujących wzorów:

$$L^{*} = \begin{cases} 116 \left(\frac{Y}{Y_{0}}\right)^{\frac{1}{3}} - 16 \text{ jeżeli } \frac{Y}{Y_{0}} > 0,008856\\ 903,3 \left(\frac{Y}{Y_{0}}\right) \text{ jeżeli } \frac{Y}{Y_{0}} \leqslant 0,008856\\ u^{*} = 13L^{*} \left(u' - u'_{0}\right)\\ v^{*} = 13L^{*} \left(v' - v'_{0}\right), \end{cases}$$
(2.41)

gdzie:

$$u' = \frac{4X}{X + 15Y + 3Z} = \frac{4x}{-2x + 12y + 3}$$
$$v' = \frac{9Y}{X + 15Y + 3Z} = \frac{9y}{-2x + 12y + 3}$$
(2.42)

Składowa L^* jest wyznaczana podobnie jak w przypadku przestrzeni *CIELAB*. Wartości z indeksem 0 odpowiadają barwie traktowanej jako biel odniesienia. Składowe u^* i v^* różnią się od składowych U i V w przestrzeni YUV. W przestrzeni *CIELUV* można wprowadzić współrzędne walcowe i zdefiniować atrybuty percepcyjne:

$$L^{*} = L^{*}$$

$$H^{*}_{uv} = \operatorname{arc} \operatorname{tg} \frac{v^{*}}{u^{*}}$$

$$C^{*}_{uv} = \sqrt{u^{*2} + v^{*2}}$$
(2.43)

Dodatkowo w tej przestrzeni można zdefiniować nasycenie:

$$S_{uv}^* = \frac{C_{uv}^*}{L^*}$$
(2.44)

Odległość euklidesowa w przestrzeni CIELUV dobrze przybliża percepcyjną różnicę barw:

$$\Delta E_{uv}^* = \sqrt{(\Delta L^*)^2 + (\Delta u^*)^2 + (\Delta v^*)^2}$$
(2.45)

Średni próg postrzegalności różnicy barw JND w przestrzeni CIELUVodpowiada wartości .

Przestrzeń *CIELUV* jest stosowana w segmentacji obrazów barwnych technikami klasteryzacyjnymi [251, 291] oraz do rozpoznawania i lokalizacji obiektów dwuwymiarowych na znanym tle [252]. Ponieważ przestrzeń jest niezależna od urządzeń, to może być stosowana do generacji skal barwnych na wyświetlaczach obrazu [295]. W pracy [271] zaproponowano procedurę kalibracji systemu wizyjnego w przestrzeni *CIELUV*.

2.8. Ocena właściwości przestrzeni barw

Wybór przestrzeni barw może być ważnym czynnikiem wpływającym na wynik przetwarzania obrazu, np. na jakość segmentacji obrazu, na wartość współczynnika kompresji itd. Dobrym przykładem wagi przestrzeni barw jest praca [89] porównująca algorytmy segmentacji obrazu barwnego w różnych przestrzeniach barw (*RGB*, *YIQ*, *HLS* i *CIELAB*). Praca ta pokazuje, że nie istnieje jedna przestrzeń barw, w której otrzymywano by najlepsze wyniki dla wszystkich badanych algorytmów segmentacji i dla wszystkich obrazów. Można stwierdzić bardziej ogólnie, że nie istnieje jedna idealna lub przynajmniej ogólnie akceptowalna przestrzeń barw jednakowo przydatna na różnych etapach przetwarzania obrazów barwnych. Decyzja odnośnie wyboru przestrzeni barw zależy od konkretnego zadania przetwarzania i obrazów. Trudno jest nieraz podjąć decyzję dotyczącą wyboru przestrzeni barw, ale wiedza o właściwościach tych przestrzeni może pomóc w podjęciu odpowiedniej decyzji.

Istnieją również hybrydowe przestrzenie barw, tj. przestrzenie, których składowe pochodzą z różnych przestrzeni barw i zostały celowo zestawione w jedną nową przestrzeń.

Przestrzenie barw różnią się pomiędzy sobą następującymi cechami: liniowością transformacji, stabilnością obliczeń transformacyjnych i równomiernością percepcyjną. W tab. 2.2 zestawiono te cechy dla transformacji z przestrzeni *RGB* do innych przestrzeni omawianych w tym rozdziale. Z tabeli wynika bezpośrednio, że nie istnieje idealna przestrzeń barw, tzn. taka, której transformacja byłaby jednocześnie liniowa, stabilna obliczeniowo i równomierna percepcyjnie.

Omawiając przestrzenie barw skoncentrowano się na ich właściwościach wykorzystywanych w przetwarzaniu obrazów, pomijając inne istotne np. w kolorymetrii czy tele-

2.8. Ocena właściwości przestrzeni barw

Lp.	Przestrzeń	Liniowość	Stabilność obliczeń	Równomierność percepcyjna	
1.	rgb	_	-		
2.	XYZ	+	+		
3.	xyz	_	-		
4.	YUV	+	+	-	
5.	YIQ	+	+	-	
6.	YC_bC_r	+	+	_	
7.	Opozycyjne	+	+	_	
8.	Ohty	+	+	_	
9.	HSx	_	-	_	
10.	CIELAB	-	+	+	
11.	CIELUV	-	+	+	

Tabela 2.2. Właściwości transformacji pomiędzy RGB i innymi przestrzeniami barw

wizji. W przetwarzaniu obrazów barwnych wykorzystuje się tylko niewielką część przestrzeni barw opracowanych przez lata w nauce o barwie i stosowanych w technologii barwnej [21, 134]. W przedstawionym przeglądzie przestrzeni barw świadomie pominięto niektóre rzadziej stosowane w przetwarzaniu barw przestrzenie [180, 193], np. system Munsella, przestrzeń ATD itp., przestrzenie mające znaczenie jedynie historyczne, np. CIE $U^*V^*W^*$ oraz "przestrzenie komercyjne" np. YCC Kodaka, TekHVC Tektronixa czy YES Xeroxa.

Niezadowolenie z właściwości istniejących przestrzeni barw ciągle stymuluje próby konstruowania nowych przestrzeni o lepszych właściwościach. Punktem odniesienia w tych pracach są składowe istniejących przestrzeni barw. Na przykład w pracy [120] złożono razem odcień H zdefiniowany w przestrzeni CIELAB, z nasyceniem S pochodzącym z przestrzeni HSI i luminancją Y z telewizyjnej przestrzeni YIQ. Powstające w taki sposób nowe przestrzenie barw nazwano przestrzeniami hybrydowymi [294]. W cytowanej pracy na podstawie barwy klasyfikowano sylwetki piłkarzy w obrazach z meczu. W celu efektywnej segmentacji obrazu barwnego rozpatrzono wyniki uzyskane przy zastosowaniu 14 różnych przestrzeni barw. Ustalono 3 składowe pochodzące z różnych przestrzeni, dla których segmentacja była najbardziej efektywna i nazwano je składowymi hybrydowej przestrzeni barw. Inna możliwość tworzenia przestrzeni barw to definiowanie nowych transformacji na RGB. W pracy [249] opisano nową przestrzeń barw, zainspirowaną zaletami przestrzeni HSV i CIELAB, a według autorów pozbawioną ich wad. W nauce o barwie, oprócz klasycznej kolorymetrii przydatnej w pomiarach barwy, ale nie wyjaśniającej wielu zjawisk, powstają matematyczne modele percepcji barwy (ang. *Colour Appearance Models*, CAM) [71, 257]. Modele te uwzględniają dodatkowo: różnorodność warunków widzenia, złożone zjawiska percepcyjne (np. adaptacja chromatyczna, kontrast symultaniczny itp.), oświetlenie otoczenia obiektu i jego barwę. Przestrzenie *CIELAB* i *CIELUV* są przestrzeniami barw oraz jednocześnie mają pewne cechy modeli percepcji barwy. Modele zalecane przez CIE są oznaczone jako: model przejściowy CIE-CAM97s [45] i nowszy model CIECAM02 [166]. Modele CAM nie będą tutaj szczegółowo omawiane, ponieważ ich zastosowanie w przetwarzaniu obrazów barwnych jest jeszcze ograniczone i nie będą wykorzystywane w dalszej części pracy.

W celu skrócenia czasu przetwarzania obrazu barwnego transformacje przestrzeni barw realizowane są w sposób sprzętowy. Przykładowo niektóre framegrabbery wyposażone są w układy scalone, dokonujące w czasie rzeczywistym transformacji $RGB \leftrightarrow HSI$ [163]. Inna możliwość szybkiej realizacji tej nieliniowej transformacji to wykorzystanie tabeli pośredniej LUT, do której załadowano wcześniej wartości HSI [158]. W technologii FPGA (ang. Field Programmable Gate Array) realizowane są również konwertery $RGB \leftrightarrow$ YC_bC_r .

Cyfrowy obraz barwny można traktować jako dwuwymiarową tablicę pikseli. Załóżmy, że tablica składa się z M wierszy i N kolumn, a każdy piksel (m, n) jest reprezentowany przez uporządkowaną trójkę liczb:

$$\forall_{m \in [0, M-1]} \forall_{n \in [0, N-1]} P(m, n) = \{ P_1(m, n), P_2(m, n), P_3(m, n) \}, \qquad (2.46)$$

gdzie $P_i \in \Re, i = \{1, 2, 3\}$. Składowe P_i przyjmują najczęściej wartości rzeczywiste z zakresu [0, 1] lub całkowite z zakresu [0, 255]. Do opisu chromatyczności barwy można stosować liczby zespolone [82], np.:

$$Ch(m,n) = S(m,n) e^{H(m,n)i}$$
 (2.47)

gdzie: Ch – chromatyczność, H – odcień, S – nasycenie. Taki opis barwy jest niepełny; barwa jest wielkością trójwymiarową. Uogólnieniem liczb zespolonych są kwaterniony. Od dziesięciu lat [226] rozwijane jest podejście kwaternionowe w przetwarzaniu obrazów barwnych. Barwa piksela może być wyrażona w postaci następującego kwaternionu:

$$\Pi(m,n) = P_0(m,n) + P_1(m,n)i + P_2(m,n)j + P_3(m,n)k, \qquad (2.48)$$

gdzie *i*, *j*, *k* to jednostki urojone. W przypadku reprezentacji barw w przestrzeni RGB elementy kwaternionu przyjmują wartości: $P_0 = 0$, $P_1 = R$, $P_2 = G$ oraz $P_3 = B$.

Takie kwaterniony z zerową częścią rzeczywistą, zwane czystymi kwaternionami, znalazły zastosowanie w różnych transformacjach obrazów barwnych, takich jak: FFT (ang. Fast Fourier Transform) [247], KLT (ang. Karhunen-Loeve Transform) [26] itp.

W tym zdecydowanie przeglądowym rozdziale autor zawarł dwa elementy własne. Pierwszy to przeprowadzone wyznaczanie wartości granicznych dla składowych przestrzeni *CIELAB*, które dostarczają informacji o kształcie bryły barw. Drugi element to idea tabeli 2.2, która może być pomocna przy doborze przestrzeni barw do zadania przetwarzania obrazów.





Rys. 2.2. Obraz KLOCKI1 i jego składowe: (a) oryginał, (b) składowa R, (c) składowa wa G, (d) składowa B, (e) składowa Y, (f) składowa C_b , (g) składowa C_r

Fig. 2.2. Image KLOCKI1 and its colour components: (a) original, (b) component R, (c) component G, (d) component B, (e) component Y, (f) component C_b , (g) component C_r

44

Rozdział 3 Pozyskiwanie obrazów barwnych

Podstawowe urządzenie służące do pozyskiwania obrazów barwnych scen trójwymiarowych to:

- stosowana obecnie tylko w niektórych systemach ze względu na swoją wysoką cenę trójprzetwornikowa kamera kolorowa CCD (ang. *Charge-Coupled Device*) z drogim układem optycznym do widmowego rozdziału światła (ang. *beamsplitter*),
- najczęściej obecnie stosowana jednoprzetwornikowa kamera kolorowa z przetwornikiem wykonanym w technologii CCD albo coraz częściej w technologii CMOS i z nałożonym na niego barwnym filtrem mozaikowym (np. filtr Bayera 2G: R: B lub filtr oparty głównie na barwach dopełniających Cy: Mg: Ye: G),
- kamera kolorowa z liniową strukturą CCD do pozyskiwania obrazów barwnych scen znajdujących się w ruchu prostopadle do osi sensora kamery,
- stosowana od dawna do pozyskiwania obrazów scen nieruchomych kamera monochromatyczna z kołem filtracyjnym R, G, B, opisana np. w pracy [302].

Przykładem dużych możliwości rozwojowych technologii CMOS (ang. Complementary Metal Oxid Semiconductor) są produkowane od kilku lat sensory X3 firmy Foveon (USA). Ich specjalna architektura pozwala pozyskiwać obrazy barwne bez stosowania filtru mozaikowego i specjalnych procedur interpolacyjnych (ang. demosaicking). Foveon wykorzystuje w tych czujnikach zależność głębokości, na której absorbowane jest światło od długości jego fali [308]. I tak niebieskie światło jest absorbowane najbliżej powierzchni sensora, a czerwone najdalej. Na tych różnych poziomach są detekowane odpowiednie składowe sygnału barwy.

Rozwiązanie w postaci kamery z kołem filtracyjnym, które swego czasu odegrało dużą rolę w rozwoju przetwarzania obrazów barwnych, coraz częściej jest zastępowane przez

ROZDZIAŁ 3. Pozyskiwanie obrazów barwnych

kamerę z filtrem ciekłokrystalicznym o regulowanych charakterystykach widmowych (ang. *Liquid Crystal Tunable Filter*, LCTF). Przykładem jest filtr Micro*Color firmy CRI, Inc. (USA) [34]. To rozwiązanie eliminuje części ruchome i związane z nimi drgania, które mogą być powodem zniekształceń w pozyskanym obrazie. Bezpośrednie, tzn. bez procedur interpolacyjnych, wyznaczanie składowych barwy pikseli skutkuje wysoką jakością obrazu barwnego.

Pozyskiwanie obrazów barwnych wymaga stosowania oprócz odpowiedniej kamery z odpowiednim obiektywem również interfejsu wizyjnego (ang. framegrabber), pozwalającego pozyskiwać obrazy z 24 bitami informacji barwnej o każdym pikselu (najczęściej 8 bitów na każdą z barw podstawowych R, G, B) oraz układu oświetleniowego ze źródłami światła o wysokim wskaźniku oddawania barw CRI (ang. Colour Rendering Index). Stosowanie kamer cyfrowych znacznie upraszcza konstrukcję interfejsu wizyjnego, sprowadzając problem do zastosowania interfejsu standardowego, np. IEEE-1394 czy USB.

Konfiguracja systemu wizyjnego może wymagać rozwiązania wielu szczegółowych problemów technicznych, takich jak:

- zapewnienie równomierności oświetlenia na dużej powierzchni [58],
- zapewnienie jakości odwzorowania optycznego,
- zapewnienie automatycznego balansu bieli w kamerze odpowiadającego adaptacji chromatycznej układu wzrokowego,
- zastosowanie chłodzenia sensora w celu osiągnięcia wyższej wartości SNR (ang. *Signal Noise Ratio*) kamery,
- optymalne powiązanie wartości natężenia oświetlenia, apertury obiektywu, czasu naświetlania i czułości sensora.

Często nie docenia się etapu pozyskiwania obrazów uważając, że w dalszym przetwarzaniu obrazów można poprawić jakość pozyskanego obrazu. Nie zawsze jest to jednak możliwe. Przykładowo, jeżeli punkt pracy systemu wizyjnego znajduje się poza jego zakresem pracy dynamicznej, to następuje tzw. obcinanie sygnału (ang. *clipping*), którego nie da się skorygować w dalszym przetwarzaniu.

Szybki rozwój technologii kamer kolorowych powoduje coraz częściej dezaktualizację niektórych związanych z nimi problemów, np. efektu kwitnienia (ang. *blooming*), potrzeby blokowania podczerwieni itp. Dlatego w dalszej części tego rozdziału zostaną przedstawione jedynie te zagadnienia związane z pozyskiwaniem obrazów, które jeszcze długo będą stanowić obszar zainteresowania specjalistów zajmujących się przetwarzaniem obrazów barwnych.

Do takich zagadnień należą przedstawione poniżej: zadanie kalibracji kolorymetrycznej systemu wizyjnego, metodyka oceny wierności odwzorowania barw w systemie wizyjnym czy procedura wyznaczania charakterystyk widmowych kamery. Co prawda, w badaniach w tym zakresie autor wykorzystywał jedynie jednoprzetwornikowe kamery CCD, ale jest przekonany o przydatności zaproponowanych procedur i metodyki dla kamer CMOS i trójprzetwornikowych kamer CCD.

W ostatnim podrozdziale przedstawiono problematykę nie związaną bezpośrednio ze sprzętem wizyjnym. Każdy z pozyskanych obrazów barwnych można ocenić ze względu na jego barwność. To niedawno zdefiniowane pojęcie poddano konfrontacji z pojęciem liczby barw w obrazie oraz badano jego zależność od percepcyjnych atrybutów barwy. Barwność obrazu została wykorzystana w badaniach przedstawionych w (patrz rozdz. 6).

3.1. Kalibracja systemu wizyjnego

Racjonalne stosowanie systemów przetwarzania obrazów barwnych wymaga przed etapem pozyskiwania obrazów ustawienia takich parametrów systemu wizyjnego, aby pracował on w swoim zakresie dynamicznym oraz żeby można było uniknąć negatywnych zjawisk, takich jak: obcinanie, olśnienie czy brak równowagi bieli. Wpływ tych zjawisk trudno jest skorygować po pozyskaniu obrazu. W wielu przypadkach potrzebna jest linearyzacja charakterystyk kamery i niekiedy kalibracja kolorymetryczna (ang. *colorimetric calibration*). Potrzeba takiej kalibracji jest spowodowana tym, że w przypadku większości kamer kolorowych ich charakterystyki widmowe nie są kombinacją liniową charakterystyk widmowych czopków, tj. nie spełniają pochodzącego z początku XX wieku warunku dokładnej reprodukcji barwy Luthera-Ivesa [257]. Kalibracja ta może także ułatwić porównywanie obrazów barwnych tej samej sceny pozyskanych w różnych warunkach oświetleniowych i za pomocą różnych kamer.

Zjawisko obcinania występuje w wyniku ograniczenia sygnału barwy z powodu przekroczenia zakresu pracy dynamicznej danego członu przetwarzającego. W systemach wizyjnych obcinanie nie występuje we wszystkich 3 kanałach równocześnie, co prowadzi do znacznych zniekształceń barwy. Na przykład dla barwy z największą składową G nasycenie i obcinanie najszybciej nastąpi dla tej właśnie składowej. Dalszy, nawet liniowy wzrost dwóch pozostałych składowych będzie powodować zniekształcenia barwne.

ROZDZIAŁ 3. Pozyskiwanie obrazów barwnych

Wzrost wartości sygnału poza granicę obcinania może prowadzić do wystąpienia zjawiska olśnienia. Zjawisko to polega na wypływie nośników ładunku z "przepełnionej", w wyniku oddziaływania dużego strumienia fotonów, studni potencjału elementu matrycy CCD do studni potencjału sąsiednich elementów. Rezultatem jest "rozpłynięcie się" części obrazu odpowiadającej silnie naświetlonemu obszarowi. W nowszych kamerach stosuje się sensory CCD z układami zabezpieczającymi przed olśnieniem. Na występowanie olśnienia może mieć również wpływ obecność w scenie gładkich powierzchni, które są źródłem odblasków (ang. *highlights*).

Brak równowagi bieli jest najbardziej zauważalny w obrazach obiektów achromatycznych. Przyczyną fizyczną tego stanu jest różna czułość elementów matrycy CCD i różna transmisja barwnego filtru mozaikowego dla różnych zakresów widma. Nawet stosowanie kamery z układem automatycznej równowagi bieli nie do końca rozwiązuje ten problem.

O potrzebie kalibracji systemu wizyjnego można łatwo się przekonać obserwując różnice barw w obrazach tej samej sceny pozyskanych za pomocą 2 różnych systemów wizyjnych. W pracy [83] zaproponowano następującą kilkukrokową procedurę kalibracyjną systemu wizyjnego wykorzystującego kamerę analogową z interfejsem wizyjnym:

- Zgrubne ustawienie równowagi bieli oznacza osiągnięcie równości sygnałów wyjściowych R, G, B kamery dla obiektu achromatycznego będącego w jej polu widzenia. Różne mogą być przyczyny braku tej równowagi, np.: nierównomierna czułość sensora, różne wartości współczynnika transmisji dla różnych elementów barwnych filtru mozaikowego itd. Równowagę w trzech kanałach uzyskuje się poprzez ręczną lub automatyczną regulację wzmocnienia. Zmiana układu oświetleniowego wymaga nowego ustawienia tej równowagi.
- 2. Wyznaczenie poziomów czerni w trzech kanałach kamery następuje przez pomiar sygnału wyjściowego (zapis obrazu) przy zamkniętym obiektywie. Zerowa wartość poziomu czerni w jednym z kanałów oznacza, że poziom ten leży poniżej poziomu, któremu przetwornik a/c przypisuje wartość 0. W tej sytuacji należy obniżyć ujemne napięcie odniesienia przetwornika a/c w tym kanale.
- 3. Określenie zakresu pracy dynamicznej jest istotne ze względu na niedopuszczenie do stanu nasycenia systemu. W tym celu po ustawieniu maksymalnego oświetlenia należy otwierać przysłonę obiektywu kamery aż do momentu, gdy wartości sygnałów R, G, B przestaną się zmieniać (stan nasycenia), a następnie nieco przymknąć obiektyw.

- 4. Precyzyjne ustawienie równowagi bieli i czerni można osiągnąć poprzez programową regulację dodatniego i ujemnego napięcia odniesienia przetworników a/c. Wartości wyznaczonych poziomów bieli i czerni wykorzystuje się do korekcji składowych *RGB* w następujący sposób:
 - $R = \frac{R R_m}{R_b R_{cz}}$ $G = \frac{G G_m}{G_b G_{cz}}$ $B = \frac{B B_m}{B_b B_{cz}},$ (3.1)

gdzie:

- R_m, G_m, B_m wartości mierzone składowych RGB,
- R_{cz}, G_{cz}, B_{cz} wartości składowych RGB dla poziomu czerni,
- R_b, G_b, B_b wartości składowych RGB dla poziomu bieli.
- 5. Linearyzacja charakterystyk przenoszenia wynika z nieliniowości "wbudowanej" w kamerę dla potrzeb kompensacji nieliniowości monitora (tzw. korektor gamma). Dla takich kamer współczynnik $\gamma = 0,45$. Również liniowość kamer, w których obwód korekcji gamma można odłączyć ($\gamma = 1$), pozostawia wiele do życzenia. Aby zlinearyzować kamerę, należy wyznaczyć charakterystyki przenoszenia 3 kanałów kamery, co można uzyskać przez pomiar składowych RGB za pomocą systemu wizyjnego dla zestawu próbek achromatycznych o różnych poziomach szarości. Charakterystyki przenoszenia mogą być wyznaczane metodą transmisyjną (próbki w postaci filtrów neutralnych) lub reflektancyjną. Następnie wykorzystując znane a priori wartości składowych Y barw tych próbek można obliczyć wartości współczynników nieliniowości γ_R , γ_G , γ_B :

$$Y = R^{\frac{1}{\gamma_R}}$$
 $Y = G^{\frac{1}{\gamma_G}}$ $Y = B^{\frac{1}{\gamma_H}}$, (3.2)

stąd:

 $\gamma_R = \frac{\log R}{\log Y}$ $\gamma_G = \frac{\log G}{\log Y}$ $\gamma_B = \frac{\log B}{\log Y}$ (3.3)

Na podstawie wyznaczonych charakterystyk przenoszenia i współczynników nieliniowości można ustalić wartości współczynników korekcyjnych, które umieszczone w tabeli pośredniej LUT interfejsu wizyjnego lub wykorzystane w programie korekcyjnym zapewnią liniową pracę kamery.

6. Kalibracja systemu na zbiorze obiektów barwnych o znanych składowych barwy (np. próbki z atlasu barw) polega na wyznaczeniu w danym systemie wizyjnym składowych *RGB* tych obiektów, ich transformacji do wybranej przestrzeni barw (np. CIE *XYZ*) i ustaleniu wartości korekcyjnych.

50

Doświadczalnie potwierdzono [83] przydatność opisanej procedury dla systemów z kamerą jednoprzetwornikową (XC-711P firmy Sony, TK-1070E firmy JVC). Badano zmienność wartości mierzonych współrzędnych trójchromatycznych rg i xy reprezentatywnego zbioru 9 barw, wywołaną przez zmianę natężenia oświetlenia sceny. Te zmiany oświetlenia symulowano zmieniając wartość przysłony obiektywu od 1 : 5,6 do 1 : 11 oraz od 1 : 8 do 1 : 5,6. Współrzędne trójchromatyczne są teoretycznie niewrażliwe na takie zmiany. Uzyskana zmienność współrzędnych, w przypadku gdy w systemie wizyjnym przeprowadzono wcześniej linearyzację według opisanej procedury, była wyraźnie mniejsza niż dla systemu z nieliniowymi charakterystykami przeproszenia.

Niektóre zastosowania wymagają jednak pełnej kalibracji kolorymetrycznej. Jej celem jest redukcja błędu odwzorowania barw i wyrażenie go w postaci ΔE_{uv}^* lub ΔE_{ub}^* . Wśród metod kalibracji kolorymetrycznej opisanych w literaturze [50] istnieją dwie zasadnicze grupy metod: metody spektrofotometryczne [138, 155] i metody próbek rzeczywistych (ang. *real samples metod*) [56, 271]. Metody pierwszej grupy wymagają albo znajomości charakterystyk widmowych oświetlenia, kamery i próbek albo ich pomiaru przy pomocy spektrofotometru. Metody drugiej grupy, ze względu na swoją prostotę i stosunkowo dobre wyniki pomiarowe, zostały zastosowane w telewizji programowej [56] do wyznaczania wierności odwzorowania barw kamer, a następnie także w systemach przetwarzania obrazów barwnych [271].

W raporcie [184] przestawiono wyniki kalibracji kolorymetrycznej systemu wizyjnego, w którego skład wchodziły następujące elementy: układ oświetleniowy złożony z dwóch świetlówek kompaktowych firmy Osram o mocy 36 W i wartości strumienia świetlnego 1900 lm każda, jednoprzetwornikowa kamera kolorowa WV-CL350 firmy Panasonic oraz interfejs wizyjny Microeye1C firmy Digithurst. Temperatura barwowa najbliższa (ang. *Correlated Colour Temperature*, CCT) świetlówek wynosi 5400 K, a wskaźnik CRI jest większy od 90, co pozwala traktować je w przybliżeniu jako źródła światła białego D_{55} o temperaturze barwowej 5500 K. Wykorzystywane w procesie kalibracji próbki barw pochodziły z atlasu barw L. Gericke [90]. System barw tego atlasu oparty jest na kole barw podzielonym na 24 elementy (po 3 elementy z grupy: barw żółtych, pomarańczowych, czerwonych, purpurowych, fioletowych, niebieskich, turkusowych i zielonych) oraz Tabela 3.1. Współrzędne wybranych próbek chromatycznych

Nr próbki	x	y	Y
1.	0,4310	0,4760	66
2.	0,5235	0,3860	33
3.	0,5380	0,3860	11
4.	0,2905	0,2495	5
5.	0,2125	0,2205	11
6.	0,1925	0,2285	15
7.	0,2555	0,4595	31
8.	0,3920	0,4975	53
9.	0,2015	0,3335	17

21-elementowej skali szarości rozciągającej się od bieli (próbka 00) do czerni (próbka 020) (rys. 3.1).

Barwa każdej próbki określona była przez wartości współrzędnych trójchromatycznych CIE x i y oraz składowej trójchromatycznej Y. W systemie barw tego atlasu bielą odniesienia była biel C. W pracy [178] przedstawiono rozkład jasności próbek z koła barw wyznaczonej z obrazów próbek pozyskanych w pewnym systemie wizyjnym. Na rys. 3.1(c) pokazano rozkład jasności tych próbek w sytuacji, gdy obrazy pozyskiwano za pomocą cyfrowego aparatu fotograficznego Olympus C-770. W obu przypadkach przebieg wykresów jest podobny: dla próbek z numerem od 1 do 13 jasność spada, a następnie rośnie aż do ostatniej próbki koła barw. Z wykresu widać, że istnieją barwy, które są nierozróżnialne na podstawie ich jasności tj., nierozróżnialne w obrazach monochromatycznych.

W kalibracji wykorzystano 9 próbek chromatycznych do oceny odwzorowania barw oraz 9 próbek achromatycznych do linearyzacji charakterystyk przenoszenia. Wartości współrzędnych CIE dla wybranych próbek zawarto odpowiednio w tab. 3.1 i tab. 3.2.

W ramach opisanej powyżej procedury kalibracyjnej w kamerze wyłączono obwód automatycznej regulacji wzmocnienia (ang. Automatic Gain Control, AGC) oraz ustawiono minimalną wartość poziomu czerni (ang. Pedestal). Następnie za pomocą najjaśniejszej próbki białej (Y = 85) wyznaczono składowe RGB poziomu bieli:

$$R_b = 229 \ G_b = 226 \ B_b = 233$$

(3.4)







Rys. 3.1. Atlas barw GERICKE: (a) 24 próbki z koła barw, (b) 21 próbek achromatycznych, (c) rozkład jasności próbek z koła barw wyznaczony z obrazu pozyskanego przez system wizyjny

Fig. 3.1. Gericke's colour atlas: (a) 24 chips from colour wheel, (b) 21 achromtic chips,(c) lightness distribution of colour chips build on the base of colour image acquired by the vision system

Nr próbki	x	y	Y
1.	0,3100	0,3160	85
2.	0,3100	0,3160	73
3.	0,3100	0,3160	62
4.	0,3100	0,3160	51
5.	0,3100	0,3160	42
6.	0,3060	0,3120	27
7.	0,3010	0,3050	14
8.	0,2960	0,2995	7
9.	0,2960	0,2995	2

Tabela 3.2. Współrzędne wybranych próbek achromatycznych

Pozyskano obraz przy całkowicie zamkniętym obiektywie i na tej podstawie wyznaczono składowe *RGB* poziomu czerni:

$R_{cz}=23$	
$G_{cz} = 24$	(3.5)
$B_{cz}=20$	

Wyznaczenie wartości poziomów bieli i czerni pozwala stosować dla składowych RGB wzory korekcyjne (3.1). Jak widać, ograniczone możliwości stosunkowo prostego interfejsu wizyjnego nie pozwoliły uzyskać szerszego zakresu pracy dynamicznej i lepiej zrównoważyć poszczególne kanały. Wartości składowych RGB oraz Y^* zmierzone za pomocą systemu wizyjnego przed linearyzacją dla 9 próbek achromatycznych przedstawiono w tab. 3.3, a charakterystyki przenoszenia kamery ilustruje rys. 3.2.

Dla trzech wyznaczonych nieliniowych charakterystyk przenoszenia kamery (rys. 3.2) za pomocą wzorów (3.3) obliczono współczynniki nieliniowości:

$$\begin{aligned}
\gamma_R &= 0,5716 \\
\gamma_G &= 0,5728 \\
\gamma_B &= 0,5573
\end{aligned}$$
(3.6)

Na ich podstawie zbudowano krzywe linearyzujące, które pozwoliły zlinearyzować charakterystyki przenoszenia kamery (tab. 3.4, rys. 3.3).

Ostatni etap kalibracji kolorymetrycznej systemu wymagał wyznaczenia błędu odwzorowania barw. W tym celu pozyskano 9 obrazów z próbkami barwnymi, w każdym pozyskanym obrazie znajdował się tylko jeden barwny obiekt, w celu wyeliminowania błędów związanych z odbiciem światła od innych obiektów sceny. Obrazy zostały pozyskane

Tabela 3.3. Składowe RGB oraz Y dla próbek w obrazie przed linearyzacją

Nr próbki	R	G	В	Y*	Y
1.	254	254	255	100	85
2.	230	232	236	91	77
3.	211	220	230	86	62
4.	183	193	210	75	51
5.	163	172	187	67	42
6.	120	128	148	50	27
7.	69	75	88	29	14
8.	34	34	38	13	7
9.	21	19	22	8	2

Tabela 3.4. Składowe RGB oraz Y dla próbek w obrazie po linearyzacji

Nr próbki	R	G	В	Y*	Y
1.	216	216	216	85	85
2.	192	191	188	75	77
3.	159	161	164	63	62
4.	129	130	130	51	51
5.	107	107	107	42	42
6.	69	69	69	27	27
7.	36	36	36	14	14
8.	16	16	16	6	7
9.	5	4	5	2	2

dwukrotnie: raz w systemie skalibrowanym według opisywanej powyżej procedury, a drugi raz w systemie niekalibrowanym. Uśrednione w ramach każdej próbki składowe *RGB* były przeliczane do przestrzeni CIE *XYZ*, a następnie do przestrzeni *CIELUV*. Wymagało to wyprowadzenia właściwej dla badanego systemu wizyjnego macierzy transformującej składowe *RGB* na składowe *XYZ*.

W procesie wyznaczania transformacji $RGB \rightarrow XYZ$ przyjęto dla uproszczenia, że układ oświetleniowy można przybliżyć za pomocą iluminantu D_{55} . Współrzędne trójchro-



ROZDZIAŁ 3. Pozyskiwanie obrazów barwnych





Rys. 3.3. Charakterystyki przenoszenia kamery po linearyzacji Fig. 3.3. Camera transfer functions after linearization

matyczne dla barw podstawowych w systemie EBU i dla bieli D_{55} wynosza:

$$\begin{array}{ccccccc} x_R = 0, 64 & y_R = 0, 33 & z_R = 0, 03 \\ x_G = 0, 29 & y_G = 0, 60 & z_G = 0, 11 \\ x_B = 0, 15 & y_B = 0, 06 & z_B = 0, 79 \\ & & \\ x_{D_{55}} = 0, 3324 \\ & y_{D_{55}} = 0, 3476 \\ & & \\ z_{D_{55}} = 0, 3200 \end{array}$$
 (3.8)

ROZDZIAŁ 3. Pozyskiwanie obrazów barwnych

Sposób wyznaczenia macierzy transformacji pochodzi z literatury [245]. Przedstawiając zależności pomiędzy składowymi RGB i XYZ w następujący sposób:

$$\begin{bmatrix} X\\Y\\Z \end{bmatrix} = \begin{bmatrix} S_R x_R & S_G x_G & S_B x_B\\S_R y_R & S_G y_G & S_B y_B\\S_R z_R & S_G z_G & S_B z_B \end{bmatrix} \begin{bmatrix} R\\G\\B, \end{bmatrix}$$
(3.9)

gdzie S_R , S_G i S_B są współczynnikami skalującymi oraz przyjmując, że punkt bieli w przestrzeni RGB jest określony jako (1, 1, 1), a w przestrzeni XYZ przez biel odniesienia, a także przyjmując typową normalizację Y = 1, można zapisać:

$$\begin{bmatrix} \frac{x_{D_{55}}}{y_{D_{56}}}\\ 1\\ \frac{z_{D_{55}}}{y_{D_{56}}} \end{bmatrix} = \begin{bmatrix} 0.95644\\ 1\\ 0.92085 \end{bmatrix} = \begin{bmatrix} S_R 0.64 & S_G 0.29 & S_B 0.15\\ S_R 0.33 & S_G 0.60 & S_B 0.06\\ S_R 0.03 & S_G 0.11 & S_B 0.79 \end{bmatrix} = \begin{bmatrix} 1\\ 1\\ 1\\ 1 \end{bmatrix}$$
(3.10)

Po prostych przekształceniach otrzymujemy układ trzech równań z trzema niewiadomymi:

$$S_R 0,64 + S_G 0,29 + S_B 0,15 = 0,95644$$

$$S_R 0,33 + S_G 0,60 + S_B 0,06 = 1,00000$$

$$S_R 0.03 + S_G 0,11 + S_R 0,79 = 0.92085$$

(3.11)

Rozwiązując powyższy układ równań otrzymano:

$$S_R = 0,7389$$

$$S_G = 1,1627$$

$$S_R = 0.9757$$

(3.12)

W rezultacie otrzymano następującą macierz transformacji:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0,4729 & 0,3372 & 0,1464 \\ 0,2439 & 0,6976 & 0,0585 \\ 0,2222 & 0,1279 & 0,7708 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$
(3.13)

Dalsza transformacja z przestrzeni XYZ do przestrzeni CIELUV została opisana w rozdziale 2. W podręcznikach kolorymetrii, np. [97], można znaleźć dane kolorymetryczne iluminantu D_{55} niezbędne do przeprowadzenia tej transformacji:

$$Y_0 = 100
 x_0 = 0,3324
 y_0 = 0,3474,
 (3.14)$$

a na podstawie wzorów (2.42) można wyznaczyć współrzędne punktu bieli D_{55} na płaszczyźnie u^*v^* :

$$u'_0 = 0,2044$$

 $v'_0 = 0,4807$
(3.15)

Dla 9 próbek barwnych wyznaczono błędy odwzorowania barw ΔE_{uv}^* i obliczono ich wartość średnią. Wynosiła ona dla systemu nieskalibrowanego 33,74 oraz 27,51 dla systemu po kalibracji.

Symulując zmianę natężenia oświetlenia poprzez przymknięcie przysłony powtórzono badania uwzględniając obniżenie poziomu bieli:

$$R_b = 180$$

 $G_b = 179$
 $B_b = 181$
(3.16)

i przeprowadzając nową linearyzację charakterystyki kamery. Otrzymane wyniki potwierdziły intuicyjną zależność błędu odwzorowania barwy od natężenia oświetlenia: zmniejszenie natężenia powodowało wzrost błędu. Można stwierdzić, że badany system wizyjny, podobnie jak oko ludzkie, lepiej "postrzega" barwę przy wyższym natężeniu oświetlenia niż przy niższym.

Wyniki badań doświadczalnych pokazują, że zaproponowana procedura kalibracji systemu wizyjnego wyraźnie zmniejsza błąd odwzorowania barw. Uzyskiwane średnie wartości błędu ΔE_{uv}^* , większe od 20, pokazują różnice w jakości pomiędzy kamerami jednoprzetwornikowymi a znacznie lepiej odwzorowywującymi barwę kamerami trójprzetwornikowymi, dla których średnia wartość tego błędu wynosi poniżej 10 [271], a nawet w przypadku kamer profesjonalnych około 5 [56].

Opracowana w latach dziewięćdziesiątych przez autora i prof. H. Freya kilkukrokowa procedura kalibracji systemu wizyjnego była jedną z pierwszych całościowych propozycji w tym zakresie przedstawionych w literaturze. Nowością procedury jest dbałość o pracę systemu w możliwie szerokim zakresie pracy dynamicznej poprzez regulację przysłony obiektywu i ustawienie parametrów interfejsu wizyjnego. Autor w celu dalszego zwiększenia dokładności procedury, opierając się na równaniach nauki o barwie, wyprowadził i zastosował swoistą macierz transformacji $RGB \rightarrow XYZ$ dla bieli D_{55} . Dodatkowym potwierdzeniem przydatności zaproponowanej procedury kalibracji jest jej wykorzystanie przez innych badaczy w pracach z zakresu przetwarzania obrazów barwnych, np. [24,237]. Mimo upływu czasu zasadnicze idee procedury pozostają aktualne, o czym przekonuje lektura rozdziału 5 nowej monografii z zakresu obrazowania barwnego [257].

ROZDZIAŁ 3. Pozyskiwanie obrazów barwnych

3.2. Ocena wierności odwzorowania barw w systemie wizyjnym

O jakości systemów wizyjnych pozyskujących i przetwarzających obrazy barwne świadczy m.in. wierność odwzorowania barw w pozyskiwanych i przetwarzanych obrazach. W pracach autora [215,217,218] zaproponowano metodykę oceny wierności odwzorowania barw z wykorzystaniem aktualnych wzorów do obliczania różnicy barw opracowanych przez CIE: CIELAB, CIE94, CIEDE2000. Ponieważ wiele urządzeń pozyskujących obrazy barwne wykorzystuje własne przestrzenie barw, to wskazane jest zakodowanie informacji obrazowej w sposób niezależny od urządzenia pozyskującego i platformy komputerowej. Można to osiągnąć stosując standardowe przestrzenie barw RGB (np. sRGB).

Poniżej opisano przestrzeń *sRGB* i proponowane przez CIE formuły do obliczania różnicy barw. Ten aparat pojęciowy został następnie wykorzystany do doświadczalnej oceny wierności odwzorowania barw w obrazie dla dwóch przypadków: reprezentacji barwy poprzez własną przestrzeń *RGB* urządzenia pozyskującego obraz i przestrzeń *sRGB*.

Chociaż początki standaryzacji przestrzeni barw przez CIE sięgają lat trzydziestych XX wieku, to opracowana przez firmy Hewlett-Packard i Microsoft w drugiej połowie lat dziewięćdziesiątych przestrzeń *sRGB* stała się w 1999 roku standardem (de iure [115] i de facto) dla aplikacji działających w systemie operacyjnym Windows. Mimo to w użyciu są również inne przestrzenie *RGB*. W pracy [273] przedstawiono aż 9 różnie zdefiniowanych przestrzeni barw *RGB* propagowanych przez firmy zajmujące się technikami obrazowania i niekiedy przez międzynarodowe instytucje standaryzacyjne. Wszystkie one mają służyć wiernemu przetwarzaniu obrazów barwnych. Ich istnienie w tak dużej liczbie wynika z faktu, że żadna z nich nie jest uniwersalna. Takie niezależne od sprzętu przestrzenie barw są potrzebne do przenoszenia informacji barwnej pomiędzy różnymi urządzeniami.

Przestrzeń *sRGB* jest trójwymiarową przestrzenią, która używa barw luminoforów monitora jako barw podstawowych. Została ona zdefiniowana dla średniej jakości monitorów komputerów klasy PC, pracujących w normalnych warunkach oświetleniowych. Na omawiany standard składają się 2 części: parametry środowiska widzenia (ang. *viewing environment*) oraz transformacja ze standardowej przestrzeni CIE XYZ na sRGB. Parametry środowiska widzenia odpowiadają sytuacji w otoczeniu typowego monitora CRT (ang. *Cathode Ray Tube*), np.:

- poziom luminancji wynosi 80 cd/m²,
- poziom luminancji otoczenia 200 lux,

- 3.2. Ocena wierności odwzorowania barw w systemie wizyjnym
 - punkt bieli monitora D_{65} ,
 - biel otoczenia D_{50} itp.

Są to warunki optymalnej obserwacji obrazów zapisanych w standardzie sRGB na monitorze CRT. Coraz częściej również monitory LCD (ang. *Liquid Crystal Display*), których wadą było słabe odwzorowanie barw, mają możliwość pracy w standardzie sRGB.

Transformacja z przestrzeni CIE XYZ do przestrzeni sRGB odbywa się w 3 następujących etapach: transformacja liniowa z CIE XYZ, korekcja gamma i przeskalowanie składowych. Wychodząc ze współrzędnych trójchromatycznych luminoforów oraz standardowego iluminantu D_{65} wyznaczono macierz transformacyjną pomiędzy sRGB i XYZ:

$$\begin{bmatrix} R_{sRGB} \\ G_{sRGB} \\ B_{sRGB} \end{bmatrix} = \begin{bmatrix} 3,2410 & -1,5374 & -0,4986 \\ -0,9692 & 1,8760 & 0,0416 \\ 0,0556 & -0,2040 & 1,0570 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$
(3.17)

Następnie do składowych sRGB zastosowano korekcję gamma i są one przekształcane na składowe nieliniowe sR'G'B':

Jeżeli $R_{sRGB}, G_{sRGB}, B_{sRGB} \leq 0,00304$, to:

$$R_{sR'G'B'} = 12,92 R_{sRGB}$$

$$G_{sR'G'B'} = 12,92 G_{sRGB}$$

$$B_{sR'G'B'} = 12,92 B_{sRGB},$$
(3.18)

a jeżeli $R_{sRGB}, G_{sRGB}, B_{sRGB} > 0,00304$, to:

$$R_{sR'G'B'} = 1,055 R_{sRGB}^{\frac{10}{2.4}} - 0,055$$

$$G_{sR'G'B'} = 1,055 G_{sRGB}^{\frac{10}{2.4}} - 0,055$$

$$B_{sR'G'B'} = 1,055 B_{sRGB}^{\frac{10}{2.4}} - 0,055$$
(3.19)

Ostatecznie transformację kończy przeskalowanie składowych RGBna skalę 8-bitową:

$$R_{8bit} = 255 R_{sR'G'B'} G_{8bit} = 255 G_{sR'G'B'} B_{8bit} = 255 B_{sR'G'B'}$$
(3.20)

W standardzie sRGB [115] opisano również transformację odwrotną z sRGB na XYZ. Po wprowadzeniu tego standardu w latach dziewięćdziesiątych nastąpiła jego akceptacja i wykorzystanie przez producentów oprogramowania (systemy operacyjne, programy graficzne, formaty plików obrazowych itd.) i sprzętu (monitory, skanery, cyfrowe aparaty fotograficzne itd.). Należy zwrócić uwagę na fakt, że stosowanie sRGB wymaga wcześniejszego określenia wzorów transformujących przestrzeń własną urządzenia (RGB) na standardową przestrzeń CIE XYZ. Dopiero wtedy można zastosować opisaną powyżej transformację z CIE XYZ na sRGB. W pracach [297] i [102] przedstawiono metody takiego charakteryzowania kamer (ang. characterization of camera) i odpowiednio skanerów oparte głównie na regresji liniowej lub regresji trzeciego stopnia.

Problem ilościowej oceny wierności odwzorowania barw sprowadza się do obliczania różnicy barw dla obrazu reprezentatywnego zbioru barw. Naturalną przestrzenią barw, w której można wyznaczać różnice, może być percepcyjnie równomierna przestrzeń CIELAB. W przypadku dwóch obrazów barwnych wartość ΔE_{ab}^* uśredniona po całym obrazie może stanowić miarę percepcyjnej różnicy barw w tych obrazach. Różnica barw pomiędzy 2 barwami czerwonymi równa $\Delta E_{ab}^* = 1$ powinna być postrzegana jako równa różnicy barw pomiędzy 2 barwami o innym odcieniu, wynoszącej $\Delta E_{ab}^* = 1$. Niestety tak nie jest.

Dlatego, opierając się na przestrzeni CIELAB starano się poprawić równomierność percepcyjną poprzez modyfikację stosowanej metryki. I tak w 1995 roku zaproponowano [44] formułę dla wyznaczania różnicy barw oznaczoną jako CIE94. Najpierw dokonuje się przejścia z kartezjańskiego układu współrzędnych przestrzeni CIELAB na walcowe współrzędne przestrzeni CIE $L^*C^*H^*$, gdzie C^*_{ab} oznacza chromę, a H^*_{ab} odcień. Wtedy wyrażenie różnicy barw ma następującą postać:

$$\Delta E_{ab}^* = \sqrt{\Delta L^{*2} + \Delta C_{ab}^{*2} + \Delta H_{ab}^{*2}}$$
(3.21)

Dopiero wprowadzenie odpowiednich współczynników wagi w celu poprawy równomierności percepcyjnej ustanawia nową metrykę CIE94:

$$\Delta E_{94}^* = \sqrt{\left(\frac{\Delta L^*}{k_L S_L}\right)^2 + \left(\frac{\Delta C_{ab}^*}{k_C S_C}\right)^2 + \left(\frac{\Delta H_{ab}^*}{k_H S_H}\right)^2},\tag{3.22}$$

gdzie współczynniki S zależą od chromy porównywanych barw:

$$S_L = 1$$

$$S_C = 1 + 0,045 C_{ab}^*$$

$$S_H = 1 + 0.015 C_{ab}^*$$
(3.23)

a współczynniki k mogą uwzględniać warunki oświetleniowe. Często jednak dla warunków standardowych (D_{65} itp.) przyjmuje się:

$$k_L = k_H = k_C = 1 \tag{3.24}$$

Można zauważyć, że im bardziej nasycone są barwy (większa chroma), tym większa jest różnica pomiędzy miarą ΔE_{ab}^* i ΔE_{94}^* .

W 2001 roku CIE opublikowała dokument przedstawiający jeszcze bardziej dokładną formułę służącą do wyznaczania różnicy barw – CIEDE2000 [46]. Dokładny opis procedury zamieszczono w Dodatku. Ze względu na złożoność formuły CIEDE2000 szczególną uwagę należy zwracać na poprawność jej implementacji. Wykorzystywana dalej implementacja tej formuły została przetestowana na danych zawartych w pracy [259].

W ramach eksperymentu postanowiono ocenić wierność odwzorowania barw stosując omówione powyżej 3 metryki: CIELAB, CIE94 i CIEDE2000. Obiektem obserwowanym przez system wizyjny była tablica wzorców barwnych ColorChecker Chart (GretagMacbeth, USA) oświetlona światłem słonecznym. Tablica ta od ponad ćwierć wieku służy do badań kolorymetrycznych w fotografii, poligrafii i telewizji [159]. Składa się z 24 specjalnie przygotowanych barwnych kwadratów o ściśle określonych współrzędnych kolorymetrycznych zapisanych w systemach opisu barwy takich jak xyY, HVC, nazwa wg ISCC/NBS. Część barw reprezentuje barwy występujące w środowisku naturalnym i biurowym (barwa skóry, roślinności, nieba itp.), a druga część to barwy podstawowe RGB, barwy dopełniające CMY oraz 6 barw achromatycznych. Współrzędne xyY barw pozwalają na obliczenie składowych CIELAB, które mogą stanowić punkt odniesienia dla obiektywnej oceny wierności odwzorowania barw. Tablicę przedstawiono na rys. 3.4.



Rys. 3.4. Wzorce barwne (24 szt.) w tablicy COLORCHECKER CHART Fig. 3.4. Colour chips (24) in COLORCHECKER CHART

Główny element badanego systemu wizyjnego stanowił cyfrowy aparat fotograficzny DiMAGE 7 firmy Minolta. Aparat ten ma inny zakres barw na płaszczyźnie xy niż przestrzeń sRGB [64]. Może on pozyskiwać więcej barw zielonych i niebieskich niż mo-

ROZDZIAŁ 3. Pozyskiwanie obrazów barwnych

że wyświetlić przeciętny monitor pracujący zgodnie z sRGB (rys. 3.5). Aparat zapisuje zdjęcia we własnej przestrzeni RGB, ale jego oprogramowanie pozwala na przekształcenie pozyskanego obrazu na obraz zapisany w jednej ze standardowych przestrzeni RGB (m.in. sRGB, Apple RGB, Adobe RGB, PAL/SECAM RGB). Korzystając z oprogramowania Image Viewer Utility, dołączonego do aparatu DiMAGE 7, dokonano konwersji obrazu z przestrzeni RGB aparatu do standardowej przestrzeni sRGB. Zauważono, że obrazy nie poddawane transformacji na sRGB mają niskie nasycenie barw.





Pozyskiwane obrazy (dla przestrzeni RGB i sRGB) zapisywano w formatach bez kompresji TIFF (ang. Tagged Image File Format) i z kompresją JPEG (współczynnik kompresji 15 : 1). Dalsze przetwarzanie każdego z obrazów wzorca polegało na podziale obrazu na 24 segmenty barwne o wymiarach: 150×145 pikseli. Dla każdego segmentu obliczano wartości średnie składowych RGB oraz sRGB. Następnie dokonywano transformacji do przestrzeni barw CIELAB poprzez XYZ.

Przejście z sRGB na XYZ zostało opisane w standardzie [115], a w przypadku transformacji przestrzeni RGB aparatu na XYZ znaleziono macierz transformacji opierając się na współrzędnych trójchromatycznych wierzchołków trójkąta RGB aparatu (rys. 3.5). Wartości odczytanych współrzędnych xy wynoszą odpowiednio dla punktu R (0,620,0,325), punktu G (0,360,0,695) oraz B (0,038,-0,150). Potrzebne do obliczeń

3.2. Ocena wierności odwzorowania barw w systemie wizyjnym

współrzędne trójchromatyczne z tych punktów znaleziono stosując oczywistą zależność:

$$z = 1 - x - y$$
 (3.25)

Dodatkowo przydatna będzie znajomość współrzędnych punktu bieli D_{65} : (0,3127, 0,3290, 0,3583). Dzieląc te współrzędne przez wartość współrzędnej y otrzymamy składowe XYZ punktu bieli: (0,9505, 1,0, 1,0888). Założenie, że Y = 1, jest często przyjmowane w kolorymetrii dla punktu bieli R = G = B = 1. Wymaga ono również wprowadzenia do równania pewnych czynników korekcyjnych. Po rozwiązaniu poniższego równania wartości czynników korekcyjnych punktu bieli d_1 , d_2 oraz d_3 [Li04]:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} x_R & x_G & x_B \\ y_R & y_G & y_B \\ z_R & z_G & z_B \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \\ d_3 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$
(3.26)

wyniosą odpowiednio: 0,6998, 1,3301 oraz 1,0104. Ostatecznie macierz transformacji T można obliczyć stosując poniższy wzór:

$$T = \begin{bmatrix} x_R & x_G & x_B \\ y_R & y_G & y_B \\ z_R & z_G & z_B \end{bmatrix} \begin{bmatrix} d_1 & 0 & 0 \\ 0 & d_2 & 0 \\ 0 & 0 & d_3 \end{bmatrix}$$
(3.27)

W przypadku rozpatrywanego aparatu fotograficznego można zapisać następującą macierz transformacji $RGB \rightarrow CIE XYZ$:

-	x_R	x_G	x_B	(2.22)
T =	y_R	y_G	y_B	(3.28)
	z_R	z_G	z_B	

Dalsza transformacja z XYZ na CIELAB była jednakowa dla obu wersji (RGB i sRGB) i zgodna ze standardem CIE. Opierając się na przestrzeni CIELAB obliczano średni błąd odwzorowania 24 barw, odnosząc wyniki z eksperymentu do pomiarów spektrofotometrycznych tablicy wzorców ColorChecker Chart, zaczerpniętych ze strony WWW Bruce'a Lindblooma [143]. Obliczenia zostały przeprowadzone w środowisku Matlab. Poniższe wykresy (rys. 3.6 i rys. 3.7) przedstawiają wartość błędu odwzorowania barw dla próbek z tablicy wzorców.

Przyglądając się powyższym wykresom można stwierdzić, że w jednym przypadku (próbka nr 13) błąd odwzorowania barw w przestrzeni sRGB przekracza błąd w RGB. Próbka ta, o barwie ciemnoniebieskiej i nazwie BLUE oraz współrzędnych trójchromatycznych x = 0,1854, y = 0,1258, znajduje się na brzegu zakresu barw dla sRGB (rys. 3.5) i w środku zakresu barw dla RGB, co może tłumaczyć to zjawisko.

ROZDZIAŁ 3. Pozyskiwanie obrazów barwnych







Rys. 3.7. Wyniki eksperymentu dla obrazu w formacie JPEG (metryka *CIELAB*) Fig. 3.7. Results of experiment for the image in JPEG format (*CIELAB* metric)

3.2. Ocena wierności odwzorowania barw w systemie wizyjnym

Na rys. 3.8 i rys. 3.9 przedstawiono wyniki eksperymentu w przypadku stosowania formuły CIEDE2000. Należy zwrócić uwagę, że nie występuje tutaj już problem z niebieską próbką nr 13 (patrz wyżej): błąd w przypadku przestrzeni sRGB jest zawsze nie większy niż w przypadku przestrzeni RGB. Równość błędów występuje tylko w przypadku barwy białej (próbka nr 19). Świadczy to dobrze o nowej formule różnicy barw.



Rys. 3.8. Wyniki eksperymentu dla obrazu w formacie TIFF (formuła CIEDE2000) Fig. 3.8. Results of experiment for the image in TIFF format (CIEDE2000 formula)

Otrzymywane wartości błędu uśredniano po całym polu tablicy wzorców i wyniki w postaci wartości średnich, wartości maksymalnych oaz odchyleń standardowych dla 3 stosowanych formuł różnicy barw zamieszczono w tab. 3.5.

Analizując wyniki można stwierdzić, że w przypadku stosowania standardu *sRGB* otrzymujemy, niezależnie od formatu obrazu, ponad dwukrotnie mniejszy średni błąd odwzorowania barw niż w przypadku stosowania przestrzeni własnej aparatu *RGB*. Nieco większe wartości błędu w przypadku stosowania formatu JPEG można wytłumaczyć stosowaną w tym formacie kompresją stratną.

W przedstawionych badaniach w sposób ilościowy została oceniona wierność odwzorowania barw w obrazie zapisanym przy użyciu składowych przestrzeni *sRGB* i *RGB*. Otrzymane wyniki są doświadczalnym potwierdzeniem tego, że dla niektórych urządzeń pozyskujących obraz przejście od własnej przestrzeni *RGB* do przestrzeni standardowej np. *sRGB* zdecydowanie poprawia wierność odwzorowania barw w systemie wizyjnym. W

64

3.3. Wyznaczanie charakterystyk widmowych kamer



(T) 1

ROZDZIAŁ 3. Pozyskiwanie obrazów barwnych





Tabela 3.5.	Wartości	błędu	odwzorowania	barw	dla	różnych	metrvk i	formatów	ohrazi
-------------	----------	-------	--------------	------	-----	---------	----------	----------	--------

	Średnia		Maksymalna		Odchylenie standardowe			
CIEL (P	Format	TIFF	JPEG	TIFF	JPEG	TIFF	JPEG	
CIELAB	RGB	27,06	31,19	38,51	42,18	7,24	8,04	
	sRGB	13,07	16,92	25,94	34,92	6,23	7,71	
	Format	TIFF	JPEG	TIFF	JPEG	TIFF	JPEG	
CIE94	RGB	24,16	27,88	29,98	34,36	5,74	6,38	
	sRGB	7,73	10,95	12,29	16,59	2,27	3,95	
	Format	TIFF	JPEG	TIFF	JPEG	TIFF	JPEG	
CIEDE2000	RGB	20,46	23,17	27,29	31,40	5,86	6,53	
	sRGB	7,22	9,73	12,39	16,51	2,13	3,33	

opinii autora jest to nowy i istotny wynik, którego uzyskanie wymagało rozwiązania szeregu pomocniczych problemów, jak np. wyznaczenie macierzy transformacji $RGB \rightarrow XYZ$ na bazie znajomości zakresu barw aparatu. Potwierdzona została również przydatność nowej formuły CIEDE2000 dla oceny wierności odwzorowania barw. Co więcej, wyniki uzyskane dla CIEDE2000 są bardziej przekonywujące niż dla starszych formuł CIE. Stosunkowo duża wartość błędu (kilkadziesiąt jednostek) wynika z przyjęcia danych odnie-

sienia powstałych w wyniku precyzyjnych pomiarów spektrofotometrycznych. Zaproponowana metodyka może być wykorzystana do badań porównawczych systemów wizyjnych stosujących różne typy urządzeń wejściowych.

3.3. Wyznaczanie charakterystyk widmowych kamer

W niektórych zastosowaniach kamer kolorowych (detekcja obiektów barwnych, ich kontrola jakości itp.) może być przydatna znajomość charakterystyk widmowych ich kanałów R, G, B. Niekiedy ich wyznaczanie traktuje się jako część procesu kalibracji systemu wizyjnego [224]. Charakterystyki te raczej rzadko są zamieszczane w dokumentacji użytkowej kamer, a mogą być przydatne np. w algorytmach korekcji barwy. Z drugiej strony różnią się one istotnie od charakterystyk standardowego obserwatora kolorymetrycznego CIE oraz są różne dla różnych kamer. Poniżej opisano eksperyment, w którym pokazano, jak można wyznaczać charakterystyki widmowe kamer bez stosowania specjalistycznego sprzętu spektrofotometrycznego.

Bezpośrednie wyznaczanie charakterystyk widmowych kamery lub cyfrowego aparatu fotograficznego wymaga użycia monochromatora, który oświetla specjalną powierzchnię odbijającą lub transmitującą światło. Obraz tej powierzchni wypełnia pole widzenia kamery. Zmieniając długość fali wychodzącej z monochromatora np., co 10 nm w całym zakresie widzialnym (380–780 nm) pozyskuje się obrazy takiej powierzchni. Równocześnie dla każdej stosowanej długości fali mierzy się natężenie światła wysyłanego przez badaną powierzchnię, co pozwala znormalizować uzyskane wyniki. Na danych obrazowych dokonuje się uśredniania po powierzchni np. 64 × 64 piksele. Niekiedy mierzy się charakterystyki tylko w kilku punktach widma i wtedy nie ma potrzeby stosowania monochromatora. Przykładowo w pracy [53] mierzono charakterystyki w 7 punktach widma narzuconych poprzez zastosowane 2 źródła laserowe ($\Delta\lambda < 1$ nm) i 5 filtrów interferencyjnych pracujących z lampą halogenową ($\Delta\lambda < 10$ nm).

Alternatywą dla metody bezpośredniej, wymagającej sprzętu pomiarowego, jest przedstawiona niedawno w literaturze [103] metoda oparta na wykorzystaniu pozyskanych badaną kamerą w znanych warunkach oświetleniowych obrazów wielobarwnego wzorca o znanych charakterystykach odbiciowych. W pracy [23] estymowano tą metodą charakterystyki widmowe jednoprzetwornikowej kamery Cohu 2252 firmy Cohu.

W badaniach korzystano z 2 wielobarwnych wzorców: ColorChecker firmy Gretag-Macbeth (rys. 3.4) oraz IT 8.7/2 firmy Agfa (rys. 3.10). Charakterystyki widmowe współczynnika odbicia próbek barwnych zawartych w wykorzystywanych wzorcach uzyskano z

ROZDZIAŁ 3. Pozyskiwanie obrazów barwnych

literatury [159] oraz witryny internetowej Uniwersytetu Joensuu w Finlandii [154].

W ramach przetwarzania pozyskanych obrazów cyfrowych wyznaczono średnie wartości składowych R, G i B dla każdej z próbek barwnych. Ostatni rząd próbek każdego ze wzorców, zawierający próbki szare o znanych luminancjach, posłużył do wyznaczenia nieliniowych charakterystyk przenoszenia kanałów kamery. Charakterystyki te poddano linearyzacji (patrz rozdz. 3.1.).



- Rys. 3.10. Wzorzec barwny IT 8.7/2 (Agfa)
- Fig. 3.10. Colour chips in IT 8.7/2 (Agfa) chart

Po linearyzacji otrzymano zmodyfikowane wartości składowych R, G i B dla obrazu każdej próbki barwnej obu wzorców. Posłużyły one do estymacji widmowych charakterystyk kanałów R, G i B kamery przy użyciu metody numerycznej, znanej w literaturze jako Principal Eigenvector (PE) [103]. Metoda ta zbliżona jest do metody korzystającej z pseudoodwrotnej macierzy Moore'a-Penrose'a [124]. Estymacja charakterystyk wymaga wyznaczenia tej macierzy dla macierzy zawierającej wartości widmowych współczynników odbicia próbek. Zamiast bezpośredniego wyznaczania macierzy pseudodwrotnej można założyć jej rozkład względem wartości szczególnych (ang. Singular Value Decomposition, SVD) i wziąć pod uwagę jedynie wektory własne odpowiadające najbardziej znaczącym wartościom szczególnym. Uzyskane w ten sposób charakterystyki widmowe $S_R(\lambda)$, $S_G(\lambda)$ i $S_B(\lambda)$ kamery Cohu 2252 przedstawiono na 3.11.

Estymowane charakterystyki można porównać z charakterystykami widmowymi badanej kamery wyznaczonymi na podstawie danych uzyskanych od jej producenta. Jed-

3.3. Wyznaczanie charakterystyk widmowych kamer



Rys. 3.11. Estymowane widmowe charakterystyki czułościowe kamery Cohu 2252 (wzorzec ColorChecker)

Fig. 3.11. Estimated spectral R, G, B chanel responses for Cohu 2252 camera (ColorChecker chart)

noprzetwornikowa kamera Cohu 2252 wyposażona jest w sensor CCD z filtrem mozaikowym Cy Mg Ye G. Dostarczone przez producenta kamery, widmowe charakterystyki kanałów barwnych Cy, Mg, Ye i G przetransformowano na charakterystyki kanałów R, G i B, korzystając z zależności pomiędzy barwami podstawowymi obu przestrzeni barw zaczerpniętymi z dokumentacji kamery. Na rys. 3.12 pokazano uzyskane w ten sposób charakterystyki widmowe $S_R(\lambda)$, $S_G(\lambda)$ i $S_B(\lambda)$. Różnice pomiędzy rys. 3.11 i rys. 3.12 wskazują na przybliżony charakter wyników estymacji charakterystyk i potrzebę bezpośredniego pomiaru charakterystyk w celu oceny błędu estymacji.

W przedstawionym w tym podrozdziale eksperymencie zmodyfikowano metodę numeryczną estymacji charakterystyk widmowych kamery. Modyfikacja polegała na rezygnacji z wyznaczania całej macierzy pseudoodwrotnej i wykorzystaniu jej wektorów własnych do rekonstrukcji charakterystyk widmowych. Uzyskane wyniki są zachęcające, chociaż ten kierunek badań wymaga dalszych działań, a przede wszystkim oceny ilościowej błędu estymacji charakterystyk.
ROZDZIAŁ 3. Pozyskiwanie obrazów barwnych



Rys. 3.12. Widmowe charakterystyki czułościowe kanałów R, G i B kamery Cohu 2252 Fig. 3.12. Spectral R, G, B chanel responses for Cohu 2252 camera

3.4. Ocena barwności pozyskanego obrazu

W dziedzinie nauki o barwie R.W.G. Hunt zaproponował w 1977 roku [114] do opisu wrażenia barwnego nowy atrybut i nazwał go barwnością (ang. *colourfulness*). Barwność, nazywana niekiedy chromatycznością (ang. *chromaticness*) [301], jest atrybutem wrażenia wzrokowego, zgodnie z którym postrzegana barwa pewnej powierzchni wydaje się bardziej lub mniej chromatyczna, tj. posiadająca odcień. Barwność, obok odcienia, jaskrawości, jasności, chromy i nasycenia, stanowi atrybut barwy stosowany w modelach wrażenia barwnego [71]. Definicja barwności jest podobna do definicji chromy i nasycenia, ale te wielkości są wielkościami względnymi. Chroma może być wyrażona jako barwność odniesiona do jaskrawości tak samo oświetlonej bieli odniesienia. Nasycenie może być obliczone jako barwność bodźca odniesiona do jego jaskrawości. Barwność zwykle wzrasta, gdy wzrasta luminancja, np. barwne obrazy scen zewnętrznych, z wyjątkiem sytuacji, gdy jaskrawość jest bardzo duża. Barwność bodźca jest więc miarą "intensywności" jego odcienia i dlatego barwność bodźca achromatycznego jest równa zero.

Powyższa definicja barwności powstała dla pojedynczego bodźca, pojedynczej próbki barwnej (ang. *colour patch*). W ostatnich latach wzrosło zainteresowanie barwnością obrazów (ang. *image colourfulness*), które można traktować jak bodźce wielobarwne. Miara barwności obrazu *M* powinna wyrażać obecność wysokonasyconych barw o różnych odcie-

3.4. Ocena barwności pozyskanego obrazu

niach. Ilościowa ocena barwności obrazu wymaga przede wszystkim transformacji z RGB do przestrzeni bezpośrednio reprezentującej percepcyjne atrybuty barwy bezpośrednio (np. HSI, HSV, HLS itp.) lub pośrednio (np. CIELUV, CIELAB itp.). Obraz barwny można traktować jako zbiór punktów w przestrzeni barw, a barwność obrazu zdefiniować jako funkcję podstawowych parametrów tego zbioru (wartość średnia, rozrzut itp.). W pracy [74] pokazano, że barwność obrazu zależy od 2 czynników: odległości barw od osi barw neutralnych i odległości pomiędzy poszczególnymi barwami. Te czynniki to odpowiednio wartość średnia i odchylenie standardowe chromy, a barwność była obliczana jako ich suma ważona. Yendrikhovskij i in. [306] zaproponowali użycie wskaźnika barwności razem ze wskaźnikiem naturalności obrazu do oceny jakości obrazów naturalnych wyświetlanych na monitorze CRT. Obrazy naturalne są obrazami scen rzeczywistych pozyskanymi za pomocą takich urządzeń jak kamery kolorowe. Autorzy tej pracy zakładali skorelowanie pomiędzy barwnością obrazu i barwami w przestrzeni CIELUV. Zdefiniowano w niej wskaźnik barwności jako sumę wartości średniej nasycenia i jego odchylenia standardowego. Autorzy zignorowali różnice pomiędzy obrazami w zakresie jasności i dlatego mogli zastapić chromę nasyceniem. Winkler w pracy [298] poświęconej ocenie jakości filmów stosował dwa atrybuty: ostrość i barwność. Sprawdzał możliwość stosowania do oceny barwności obrazu wartości nasycenia pochodzących z przestrzeni HSI, HSV oraz HLS, ale ostatecznie zastosował wartości chromy zdefiniowane w przestrzeni CIELUV. Te wartości były obliczane dla całej sekwencji obrazów.

Niedawno zaproponowano nową miarę barwności opartą na statystycznych parametrach rozkładu składowych opozycyjnych R - G i Ye - B [104]. Hasler i Suesstrunk przeprowadzili badania psychofizyczne, w których badani klasyfikowali obrazy (20 osób, 84 obrazów) do 7 kategorii ze względu na ich barwność. Sprawdzili oni różne miary barwności oraz zaproponowali miarę stosunkowo prostą do obliczania i dobrze skorelowaną (95%) z danymi eksperymentalnymi. Zgodnie z [104] barwność obrazu można obliczyć według następującego wzoru:

$$M = \sqrt{\sigma_{rg}^2 + \sigma_{yb}^2} + 0, 3\sqrt{\mu_{rg}^2 + \mu_{yb}^2},$$
(3.29)

gdzie: σ i μ to odpowiednio odchylenie standardowe i wartość średnia współrzędnych opozycyjnych pikseli obrazu. Współrzędne opozycyjne oznaczono zgodnie ze wzorami:

$$RG = R - G \tag{3.30}$$

$$YB = 0.5(R+G) - B \tag{3.31}$$

Jeżeli znany jest obraz oryginalny, to miara barwności może być zastosowana do oceny jakości obrazu po kompresji z wysokim współczynnikiem kompresji (JPEG, JPEG2000), co pokazano w pracy [274]. Barwność obrazu można, obok ostrości i naturalności, zaliczyć do globalnych atrybutów percepcyjnych, które moga służyć do oceny jakości obrazu.

Poniżej przedstawiono wyniki przeprowadzonych przez autora badań doświadczalnych, których celem było poznanie właściwości miary barwności obrazu połaczone z zamiarem jej wykorzystania w przetwarzaniu obrazów.

Po zaimplementowaniu przedstawionej metody obliczania barwności obrazu najpierw sprawdzono poprawność numeryczną programu. Jako pierwszy obraz testowy została zastosowana monochromatyczna wersja obrazu PEPPERS (rys. 3.13). Sprawdzono, że barwność obrazu wynosi zero; obraz monochromatyczny nie jest obrazem barwnym. Badania w zakresie barwności ograniczono do obrazów scen naturalnych, gdyż takie obrazy legły u podstaw definicji barwności obrazu.

Następny test przeprowadzono porównując barwności dwóch obrazów oryginalnych PEPPERS i PEARS z barwnościami ich negatywów (rvs. 3.13 i rvs. 3.14).

$$R_{neg} = 255 - R$$

$$G_{neg} = 255 - G$$

$$B_{neg} = 255 - B$$
(3.32)

Prosta analiza wzorów na barwność obrazów pokazuje, że wprowadzenie do wzorów składowych RGB negatywu zmienia jedynie znak w wyrażeniach rg oraz yb:

$$rg_{neg} = (255 - R) - (255 - G) = -rg \tag{3.33}$$

$$yb_{neg} = 0,5(255 - R + 255 - G) - (255 - B) = B - 0,5(R + G) = -yb$$
(3.34)

Barwność zależy od kwadratów wartości średnich i odchyleń standardowych powyższych wyrażeń. Dlatego barwność negatywu jest równa barwności pozytywu. Barwność obu obrazów PEPPERS wynosi 98,10, a obrazów PEARS wynosi 58,82. Rozmiar pól "chmur" pikseli na płaszczyźnie a^*b^* jest podobna w obu przypadkach (przestrzeń CIELAB). Barwność jest zależna od parametrów statystycznych "chmury" pikseli obrazu zbudowanej na płaszczyźnie wyznaczonej przez osie R - G i Ye - B.

Rysunek 3.14(e) pokazuje położenie na płaszczyźnie a^*b^* punktów odpowiadających 3 barwom podstawowym (R, G, B) i 3 barwom dopełniającym (Cy, Mq, Ye). Brak symetrii

(d)

- Rys. 3.13. Obraz testowy PEPPERS: (a) oryginał, (b) jego negatyw, (c) "chmury" pikseli na płaszczyźnie ab dla obrazu oryginalnego, (d) "chmury" pikseli na płaszczyźnie ab dla negatywu
- Fig. 3.13. Test image PEPPERS: (a) original image, (b) its negative, (c) "clouds" of pixels on the plane ab for original image, (d) "clouds" of pixels on the plane ab for negative image





(a)

(c)



ROZDZIAŁ 3. Pozyskiwanie obrazów barwnych



- Rys. 3.14. Obraz testowy PEARS: (a) oryginał, (b) jego negatyw, (c) "chmury" pikseli na płaszczyźnie *ab* dla obrazu oryginalnego, (d) "chmury" pikseli na płaszczyźnie *ab* dla negatywu, (e) "chmury" pikseli dla wybranych barw
- Fig. 3.14. Test image PEARS: (a) original image, (b) its negative, (c) "clouds" of pixels on the plane *ab* for original image, (d) "clouds" of pixels on the plane *ab* for negative image, (e) "clouds" of pixels on the plane *ab* for selected colours

3.4. Ocena barwności pozyskanego obrazu

w położeniu tych punktów wynika z nieliniowej transformacji z RGB do CIELAB. Barwy dopełniające, powstałe jako negacja barw podstawowych, leżą w ćwiartkach znajdujących się po przeciwnej stronie środka układu współrzędnych. Należy zwrócić uwagę, że wizualizacja na płaszczyźnie a^*b^* jest przybliżeniem, bo barwność M jest tutaj obliczana za pomocą uproszczonych wzorów (3.29)–(3.31), a nie jest wyrażona za pomocą składowych a^* i b^* .

Do dalszych badań [208] wybrano 5 naturalnych obrazów barwnych o różnym poziomie postrzeganej barwności i rozdzielczości 640×480 pikseli (rys. 3.15). W procesie wyboru posiłkowano się liczbą barw w obrazie (ang. *number of unique colours*) U. Podczas wyznaczania U każda trójka (R, G, B) występująca w obrazie jest traktowana jako niezależna barwa i jest jednokrotnie zliczana. Zastosowanie tablic dynamicznych lub struktur listowych oraz odpowiednich procedur sortowania pozwala stosunkowo szybko obliczyć liczbę barw. Niekiedy liczba barw odnoszona jest do liczby pikseli w obrazie i wyrażana w procentach.

Znajomość liczby barw w obrazie może być wykorzystana między innymi:

- w procesie klasyfikacji obrazów na dwie klasy: grafiki i zdjęcia grafiki charakteryzują się znacznie mniejszymi wartościami U [9], czy zdjęcia i obrazy malarskie [54] – te ostatnie mają większą wartość U,
- w klasteryzacji obrazów barwnych punkty odpowiadające nie powtarzającym się barwom w przestrzeni RGB mogą stanowić klastery startowe poddawane łączeniu [289],
- w steganografii zmniejszenie lub zwiększenie liczby barw U w obrazie w porównaniu do obrazu wyjściowego lub grupy obrazów może świadczyć o obecności ukrytej informacji w obrazie [84],
- w procesie kwantyzacji barwy liczba barw U świadczy o efektywności kwantyzacji
 [241] i może być wykorzystana do oceny jakości kwantowanego obrazu [222].

W zbiorze wybranych obrazów o jednakowej rozdzielczości przestrzennej występują obrazy o małej liczbie barw (CHART, DUCK), o średniej liczbie barw (LANDSCAPE) i o wielkiej liczbie barw (CHARACTERS, MOUNTAINS), co pokazano w tab. 3.6.

Należy zwrócić uwagę, że liczba U nie jest właściwą miarą postrzeganej barwności obrazu. Nawet jeżeli liczba barw obrazu jest wysoka, np. w odniesieniu do liczby pikseli przekracza 50%, to większość występujących w nim barw jest percepcyjnie bliska i nie jest

3.4. Ocena barwności pozyskanego obrazu

ROZDZIAŁ 3. Pozyskiwanie obrazów barwnych



- Rys. 3.15. Obrazy testowe: (a) Chart, (b) Duck,(c) Landscape, (d) Characters, (e) Mountains
- Fig. 3.15. Test images: (a) Chart, (b) Duck,(c) Landscape, (d) Characters, (e) Mountains

Tabela 3.6. Liczba barwUw obrazach testowych

Nazwa obrazu	CHART	DUCK	LANDSCAPE	CHARACTERS	MOUNTAINS
Liczba barw	40523	49759	66634	78805	95481

rozróżniana przez obserwatora. Przykładowo liczba U dla obrazu CHART jest mniejsza niż dla obrazu MOUNTAINS, podczas gdy obraz CHART jest postrzegany jako zdecydowanie bardziej barwny! Dlatego dla obrazów testowych wyznaczono barwność według wzoru (3.29). Wyniki zebrano w tab. 3.7.

Tabela 3.7. Barwność M obrazów testowych

Nazwa obrazu	DUCK	LANDSCAPE	MOUNTAINS	CHARACTERS	CHART
Barwność	15,60	35,92	46,04	88,31	96,56

Rysunek 3.16 ilustruje relację pomiędzy barwnością a polem "chmury" pikseli na płaszczyźnie ab. Jeżeli zastosujemy opis słowny skali barwności, to w przypadku badanych obrazów otrzymamy: DUCK – słabo barwny, LANDSCAPE – umiarkowanie barwny, MO-UNTAINS – średniobarwny, CHARACTERS – bardzo barwny oraz CHART – ekstremalnie barwny. Teraz możemy być pewni, że zbiór naszych obrazów barwnych jest reprezentatywny – rozciąga się na całej skali barwności. Przy okazji można zauważyć, że ekstremalnie barwny obraz CHART ma najmniejszą liczbę barw U.

Interesującym zagadnieniem może być odpowiedź na pytanie, jak na barwność obrazu wpływają percepcyjne atrybuty barwy, takie jak odcień, nasycenie i jasność [203]. Wpływ ten można ocenić używając programu umożliwiającego zmianę składowych *HLS* pikseli w obrazie. W tym celu wykorzystano możliwości programu Paint Shop Pro w wersji 9. W przypadku nasycenia i jasności można zmienić ich wartości w obrazie w wybranym stopniu, a w przypadku odcienia dokonać jego przesunięcia o określony kąt na skali kołowej. Najpierw zbadano wpływ nasycenia barwy na barwność obrazu. Wyniki uzyskane dla 5 badanych obrazów przedstawione na 3.17 pokazują jednoznacznie, że wzrost nasycenia barwy pikseli w obrazie powoduje wzrost jego barwności.

Zmiany jasności w obu kierunkach powodują zmniejszanie się barwności obrazu (rys. 3.18), co wynika z zależności pomiędzy jasnością a nasyceniem w przestrzeni *HLS* (bryła barw w postaci 2 stożków złożonych podstawami). Dlatego w przypadku wysokiej jakości obrazów barwnych ich rozjaśnienie lub przyciemnienie powoduje spadek średniego nasycenia w obrazie. Przesunięcie odcienia pikseli obrazu o zadany kąt, zgodnie z oczeki-

3.4. Ocena barwności pozyskanego obrazu

waniami, nie powoduje istotnej zmiany poziomu barwności obrazu (rys. 3.19), ponieważ w wyniku tej operacji nie ulega zmianie ilość odcieni obecnych w obrazie ani ich "intensywność". Za pomocą prostej operacji graficznej możemy zmniejszać ilość odcieni w obrazie usuwając w 4 krokach po 3 odcienie (rys. 3.20) i obserwować spadek barwności obrazu.



Rys. 3.18. Wpływ zmiany jasności pikseli na barwność obrazu Fig. 3.18. Dependence of the colourfulness of the image on the lightness

Podczas przetwarzania obrazu jego barwność może się zmieniać, na przykład wzrastać w wyniku poprawiania obrazu np. poprzez "rozciąganie chromy" [112] lub spadać w wyniku jego kompresji stratnej. Wzrost barwności obrazu poza pewien optymalny poziom obniża jego jakość ocenianą przez obserwatorów, co pokazano np. w pracy [70]. Niekiedy jednak barwność obrazu powinna być niezmienna, np. w procesach takich jak filtracja odszumiająca lub kwantyzacja barwy. Zostało to pokazane poniżej na przykładach, które świadczą o tym, że zmiana barwności obrazu może być przydatna do oceny algorytmów filtracji czy kwantyzacji barwy (patrz rozdz. 4). Należy zauważyć, że barwność jest miarą nie wymagającą, w przeciwieństwie do innych miar np., RMSE (ang. *Root Mean Square Error*), PSNR, znajomości obrazu odniesienia (ang. *no-reference metric*). Obraz odniesienia jest potrzebny do wyznaczenia zmiany barwności. Również rozwój metod wyszukiwania obrazów na podstawie ich zawartości (ang. *content based image retrieval*) spowodował zainteresowanie atrybutami barwnymi obrazów. Do grupy takich atrybutów oprócz barwy dominującej w obrazie czy histogramu barw można zaliczyć także barwność.

Podsumowując przeprowadzone na reprezentatywnym zbiorze obrazów badania barwności obrazu można stwierdzić, że definicja barwności nie preferuje żadnego odcienia z koła





- Rys. 3.16. "Chmury" pikseli na płaszczyźnie *ab* dla wybranych obrazów testowych: (a) DUCK, (b) MOUNTAINS, (c) CHART
- Fig. 3.16. "Clouds" of pixels on the plane *ab* for the selected test images: (a) DUCK, (b) MOUNTAINS, (c) CHART









barw, co prowadzi do równej barwności obrazu i jego negatywu. Pokazano, że mające liczne zastosowania w przetwarzaniu obrazów pojęcie liczby barw w obrazie nie jest związane z barwnością obrazu. Ma ono charakter bardziej "techniczny" niż percepcyjny. Dalsze badania pozwoliły powiązać globalny atrybut barwności obrazu z lokalnymi atrybutami HLS jego pikseli, wskazując na szczególną rolę nasycenia w tym związku. Wskazano również na możliwość śledzenia zmian barwności w procesie przetwarzania obrazu w celu jej wykorzystania do oceny jakości procesu przetwarzania.





(a)



Rys. 3.20. Wpływ liczby odcieni na barwność obrazu: obraz wyjściowy ChartM=96,56: (a) M=84,60, (b) M=67,83, (c) M=51,53, (d) M=33,10

Fig. 3.20. Impact of number of hues on colourfulness of the image: test image CHART: M = 96, 56, (a) M = 84,60, (b) M = 67,83, (c) M = 51,53, (d) M = 33,10





(a)

(b)



- Rys. 3.1. Atlas barw GERICKE: (a) 24 próbki z koła barw, (b) 21 próbek achromatycznych, (c) rozkład jasności próbek z koła barw wyznaczony z obrazu pozyskanego przez system wizyjny
- Fig. 3.1. Gericke's colour atlas: (a) 24 chips from colour wheel, (b) 21 achromtic chips, (c) lightness distribution of colour chips build on the base of colour image acquired by the vision system

ROZDZIAŁ 3. Pozyskiwanie obrazów barwnych



Rys. 3.4. Wzorce barwne (24 szt.) w tablicy COLORCHECKER CHART Fig. 3.4. Colour chips (24) COLORCHECKER CHART



Rys. 3.10. Wzorzec barwny IT 8.7/2 (Agfa) Fig. 3.10. Colour chips in IT 8.7/2 (Agfa) chart

Barwne wersje ilustracji rozdziału 3











Rys. 3.15. Obrazy testowe: (a) CHART, (b) DUCK,(c) LANDSCAPE, (d) CHARAC-TERS, (e) MOUNTAINS

Fig. 3.15. Test images: (a) CHART, (b) DUCK,(c) LANDSCAPE, (d) CHARACTERS, (e) MOUNTAINS

Rozdział 4

Badanie właściwości wybranych przestrzeni barw

Efektywne korzystanie z poszczególnych reprezentacji barw wymaga znajomości ich właściwości. Wiedza o strukturze przestrzeni barw może być pomocna w projektowaniu algorytmów przetwarzania obrazów barwnych, np. segmentacji obrazu. W opinii autora percepcyjne przestrzenie barw HSx ze względu na ich bezpośredni związek z atrybutami percepcji należą do przestrzeni barw wywołujących największe zainteresowanie w przetwarzaniu obrazów. Dlatego w rozdziale 4.1. omówiono relacje pomiędzy składowymi różnych przestrzeni typu HSx, specyfikę odcienia oraz wskazano na niektóre możliwości modyfikacji wzorów transformujących. Wkład własny autora polega na porównaniu różnych wersji tych przestrzeni oraz rozmieszczenia w nich punktów-barw. Porównanie to obejmuje również odrębne zagadnienie wyznaczania obrazu negatywowego w przestrzeniach HSx gdzie autor wyprowadził wzory dla negatywu w przestrzeni HLS.

Drugą wybraną przestrzenią barw jest przestrzeń powstała w wyniku transformacji Karhunena-Loèvego (KLT) na obrazie RGB. Ta przestrzeń wyróżnia się tym spośród innych przestrzeni barw, że dla każdego obrazu oddzielnie trzeba wyznaczyć macierz transformującą. W rozdziale 4.2. przedstawiono szczegóły efektywnej implementacji transformacji KLT dla obrazów barwnych. W rozdziale tym pokazano zasadnicze właściwości transformacji KLT: dekorelację składowych barwnych i skupienie energii obrazu w pierwszej składowej. Najważniejszym uzyskanym przez autora wynikiem jest pokazanie mniejszej wrażliwości składowej K_1 w stosunku do składowej jasnościowej Y ze względu na zmiany luminancji obrazu. Przedstawiono możliwość znalezienia takich współczynników transformujących, które zapewniając dekorelację składowych są jednocześnie stałe dla pewnej klasy obrazów. Pokazano też niepełną dekorelację w przestrzeni Ohty. Dokonano krótkiego przeglądu dotychczasowych zastosowań transformacji KLT w dziedzinie przetwarzania obrazów barwnych.

4.1. Przestrzeń barw HSx

Z opisanych w rozdziale 2.6. nieliniowych wzorów transformacyjnych wynika, że wzory na H i S wykazują nieusuwalne osobliwości (ang. *singularities*), których istnienie może być przyczyną niestabilności obliczeniowej. Łatwo zauważyć, że odcień jest nieokreślony wzdłuż całej osi barw achromatycznych, tzn. dla R = G = B, a nasycenie dla barwy czarnej: R = G = B = 0. Dlatego obiekty achromatyczne mogą być opisywane za pomocą intensywności I (H jest wtedy nieokreślone) i separowane od obiektów chromatycznych za pomocą nasycenia (S = 0). Gdy mianownik we wzorach na odcień (wzory (2.23)–(2.26)) dąży do zera, to odcień staje się niestabilny obliczeniowo. Oznacza to, że zmiana wartości składowych R, G, B o 1 poziom może wywołać niewspółmiernie dużą zmianę wartości H i S. Punkty, odpowiadające poszczególnym barwom, nie są rozłożone równomiernie w przestrzeniach HSx (zagęszczenia i luki), jak to ma miejsce w przestrzeni RGB. Należy pamiętać, że ze względu na dyskretny charakter przestrzeni barw (np. sześcian RGBzawiera 16,7 mln punktów-barw) bryła barw jest w znacznym stopniu pusta.

Podobne problemy występują również dla pikseli o małym nasyceniu. Można to pokazać analitycznie szczególnie łatwo dla wersji HSV badanej przestrzeni. Przyjrzyjmy się pierwszej części wzoru na odcień w tej przestrzeni (2.35) (r - składowa maksymalna) i spróbujmy wyrazić go za pomocą 2 pozostałych składowych:

$$H = \frac{g - b}{max - min} \tag{4.1}$$

Wtedy można zapisać [238]:

$$\frac{\Delta H}{H} = \frac{\Delta \left(g - b\right)}{\left|g - b\right|} + \frac{\Delta V}{V} + \frac{\Delta S}{S}$$
(4.2)

Ostatni wzór pokazuje, że gdy nasycenie jest małe, to nawet nieduża jego odchyłka powoduje dużą odchyłkę odcienia. Stąd wynika celowość nieuwzględniania w dalszym przetwarzaniu odcienia pikseli z nasyceniem poniżej pewnego progu [92]. Odwrotnie: jeżeli jednak interesujące nas obszary w obrazie mają niskie nasycenie, to nieraz rezygnuje się z korzystania z obarczonego dużym błędem odcienia na rzecz składowych S i I [142].

Z kołowej natury skali odcieniowej wynika, że 2 barwy o bardzo zbliżonych odcieniach mogą mieć zupełnie różne wartości H, np. 5° i 355°. Problem dotyczy barwy, dla której

przyjęto wartość $H = 0^{\circ}$; zwykle jest to barwa czerwona. Utrudnia to niektóre operacje, np. wyznaczanie średniego odcienia dla wybranego obszaru, bo zwykłe uśrednianie arytmetyczne wyżej wymienionych wartości H prowadzi do otrzymania barwy o odcieniu $H = 180^{\circ}$, czyli turkusowej (ang. cyan). Stąd w celu wyznaczania średniej arytmetycznej odcieni oraz różnicy dwóch odcieni należy stosować odpowiednie wzory [238]:

4.1 Przestrzeń barw HSz

 \overline{H} :

$$= \begin{cases} \frac{H_1 + H_2}{2} & \text{jeżeli } |H_2 - H_1| \leq 180^{\circ} \\ \frac{H_1 + H_2 - 360^{\circ}}{2} & \text{jeżeli } |H_2 - H_1| > 180^{\circ} \end{cases}$$
(4.3)

$$|\Delta(H_1, H_2)| = \begin{cases} |H_1 + H_2| & \text{jeżeli} \ |H_2 - H_1| \le 180^{\circ} \\ 360^{\circ} - |H_1 + H_2| & \text{jeżeli} \ |H_2 - H_1| > 180^{\circ} \end{cases}$$
(4.4)

Rekurencyjne formuły wyznaczania wartości średniej i wariancji rozkładu odcienia obszaru w obrazie barwnym, słuszne również dla omawianego przypadku, zostały przedstawione w [52]. Inna możliwość obliczania średniego odcienia to konwersja każdego składnika średniej poprzez funkcje trygonometryczne odcienia na 2 składowe w kartezjańskim układzie współrzędnych, następnie niezależne uśrednienie obu składowych i powrót do kątowej skali odcienia dla uzyskanych średnich.

Przestrzeń HSI, mimo swego zorientowania na postrzeganie barw przez człowieka, nie jest percepcyjnie równomierna. Oznacza to, że zmiana odcienia H o np. 10° może wywołać większą percepcyjnie zmianę w niektórych częściach koła barw niż w innych. Przestrzeń ta zawiera również inne uproszczenia w porównaniu z ludzką percepcją barwy. Przykładowo, całkowicie nasycone barwy o różnych odcieniach nie mają w rzeczywistości tej samej wartości intensywności (jasności), jak to ma miejsce w przypadku modeli HSV(V = 1) i HLS (L = 0, 5), np. żółty jest zawsze jaśniejszy niż niebieski.

Szczególne znaczenie dla zastosowań mają następujące właściwości H i S [228]:

• niezmienniczość odcienia ze względu na zmianę natężenia oświetlenia:

$$H(R,G,B) = H(kR,kG,kB)$$
(4.5)

• niezmienniczość nasycenia ze względu na zmianę natężenia oświetlenia:

$$S(R,G,B) = S(kR,kG,kB)$$

$$(4.6)$$

• niezmienniczość odcienia ze względu na zmianę nasycenia:

$$H(R,G,B) = H(R+a,G+a,B+a),$$
(4.7)

ROZDZIAŁ 4. Badanie właściwości wybranych przestrzeni barw

gdzie: $k \ge 0 \land (kR, kG, kB) \in \langle 0, 255 \rangle$ oraz $\forall_a (R + a, G + a, B + a) \in \langle 0, 255 \rangle$.

Zależności (4.5) i (4.6) pokazują, że przestrzeń HSI opisuje barwę w sposób zbliżony do człowieka, również jeżeli chodzi o niewrażliwość na zmiany natężenia oświetlenia. Ostatnia zależność (4.7) odpowiada zjawisku niezmienności odcienia podczas zmniejszania się nasycenia w wyniku dodawania bieli do barwy. Formalną prawdziwość powyższych wzorów można łatwo sprawdzić korzystając ze wzorów na odcień i nasycenie w transformacji $RGB \rightarrow HSI$ (wzory (2.24) i (2.28)). W realizacji praktycznej niezmienniczość jest spełniona dopóty, dopóki wartości RGB pozyskiwane przez system wizyjny leżą w jego zakresie przetwarzania, tj. nie występuje obcinanie sygnału.

4.1.1. Możliwości modyfikacji przestrzeni HSx

W niektórych zastosowaniach nie wykorzystuje się pełnego koła barw. Wtedy wygodnie jest przyjąć dla niewykorzystywanego odcienia, z powodów omówionych powyżej, wartość H = 0. Zwykle przyjmuje się $H(R) = 0^{\circ}$ lub dużo rzadziej $H(B) = 0^{\circ}$. W systemie wizyjnym do klasyfikacji jabłek [127] zastosowano $H(B) = 0^{\circ}$, ponieważ niebieski odcień nie występuje w obrazach jabłek. Podobnie w systemie detekcji cech twarzy [239] wybrano $H(Cy) = 0^{\circ}$. Do takiego ustawienia koła barw szczególnie nadają się przestrzenie HSVi HLS, dla których obrót koła barw o kąt będący wielokrotnością 60° sprowadza się do inkrementacji (dekrementacji) stałych liczbowych modulo 6 we wzorze na odcień (2.72). Tak więc po przejściu od $H(R) = 0^{\circ}$ do $H(Cy) = 0^{\circ}$, któremu odpowiada obrót o 180° $(3 \times 60^{\circ})$ i inkrementacja pierwszego składnika sumy o 3 otrzymamy, [189]:

$$H = \begin{cases} 3 + \frac{G - B}{\max(R, G, B) - \min(R, G, B)} & \text{jeżeli} \max(R, G, B) = R\\ 5 + \frac{B - R}{\max(R, G, B) - \min(R, G, B)} & \text{jeżeli} \max(R, G, B) = G\\ 1 + \frac{R - G}{\max(R, G, B) - \min(R, G, B)} & \text{jeżeli} \max(R, G, B) = B \end{cases}$$
(4.8)

Ciągłe zainteresowanie modyfikacjami i uproszczeniami transformacji $RGB \to HSx$ świadczy o tym, że w tej dziedzinie nie powiedziano jeszcze ostatniego słowa.

4.1.2. Porównanie transformacji z RGB na HSx

Przeprowadzono badania porównujące 4 wersje bezpośrednich transformacji $RGB \rightarrow HSx$ [209, 210]. Te wersje, opisane w rozdziale 2, to wersja klasyczna (wzory: (2.24), (2.28), (2.30)), wersja Bajona (wzory: (2.26), (2.28), (2.30)), HSV (wzory: (2.32), (2.35)) oraz HLS (wzory: (2.33), (2.35)). Kryterium oceny wersji stanowiła zwartość klasterów

4.1. Przestrzeń barw HSx

przedstawiających reprezentatywny zbiór obiektów barwnych w danej przestrzeni HSx. Dla celów segmentacji i klasyfikacji wskazane jest, aby klaster reprezentujący dany obiekt barwny w przestrzeni HSx zajmował możliwie najmniejszą objętość. Rozpatrywano 3 możliwe przypadki odpowiadające różnej liczbie stosowanych składowych:

- klastery w postaci segmentów liniowych (przedziałów) H, gdy wykorzystywany jest tylko odcień,
- klastery w postaci segmentów pierścieniowych na płaszczyźnie HS, gdy wykorzystywany jest odcień i nasycenie (płaszczyzna chromatyczności),
- klastery w postaci segmentów objętościowych w przestrzeni HSx gdy wykorzystywane są wszystkie składowe barwy.

Za pomocą programu do ręcznej segmentacji obrazu zaznaczano odpowiedni segment prostokątny, obejmujący obraz badanego obiektu barwnego (np. próbki z atlasu) bez punktów brzegowych. Dla każdej z 4 badanych wersji transformacji z RGB na HSx wyznaczano w segmencie obrazu wartości minimalne i maksymalne każdej ze składowych HSx. Następnie obliczano:

• rozmiar segmentu liniowego (przedziału) H:

$$H = H_{max} - H_{min} \tag{4.9}$$

• rozmiar segmentu pierścieniowego P (wycinek pierścienia o kącie rozwarcia $H_{max} - H_{min}$, promieniu zewnętrznym ρ_{max} i promieniu wewnętrznym ρ_{min}):

$$P = \pi \frac{H_{max} - H_{min}}{360^{\circ}} \left(\rho_{max}^{2} - \rho_{min}^{2}\right)$$
(4.10)

- rozmiar segmentu objętościowego Vol (dla przypadku HSI):
 - $Vol = P\left(I_{max} I_{min}\right) \tag{4.11}$

Badania eksperymentalne przeprowadzono przy użyciu systemu wizyjnego opisanego w rozdziale 3.1. i 24 próbek barwnych z atlasu L. Gericke [90], reprezentujących całe koło barw. Program ręcznej segmentacji pozwolił wyznaczyć dla każdej próbki i dla każdej z 4 transformacji rozmiary segmentów liniowych, powierzchniowych i objętościowych w przestrzeni *HSx*. Wartości dla odpowiednich segmentów zsumowano dla 24 próbek. Sumy

4.1. Przestrzeń barw HSx

Tabela 4.1. Zwartość klasterów dla 4 wersji transformacji $RGB \rightarrow HSx$

Transformacja	Segmenty H	Segmenty HS	Segmenty HSx
klasyczna	1,000	1,000	1,000
Bajona	1,070	1,200	1,200
HSV	0,980	1,220	2,120
HLS	0,980	0,970	1,250

te dla przypadku transformacji klasycznej przyjęto jako równe 1. W tab. 4.1 przedstawiono sumaryczne rozmiary segmentów dla pozostałych 3 transformacji w odniesieniu do transformacji klasycznej.

Badania te pokazały, że rozmiar segmentu zależy od barwy obiektu. Niezależnie od wersji transformacji największe segmenty tworzyły obiekty o ciemnych barwach z grupy purpur i fioletów, a najmniejsze – obiekty o jasnych barwach z grupy żółtych. Obiekty reprezentowane w przestrzeni HSx przez niskie wartości nasycenia wykazywały szczególnie duży rozrzut wartości odcienia.

W wyniku badań stwierdzono, że pomimo stosowania różnych wzorów na odcień dla badanych 4 wersji transformacji sumy segmentów liniowych różnią się pomiędzy sobą jedynie o kilka procent. Przejście do segmentów powierzchniowych (wprowadzenie nasycenia S) pokazuje, że transformacje Bajona i HSV tworzą klastery o około 20% większe niż transformacja klasyczna i transformacja HLS. Porównując wyniki uzyskane dla modelu HSV z wynikami dla HLS (ten sam wzór na H) należy stwierdzić, że w modelu HSVzastosowano zbyt uproszczony wzór na nasycenie w stosunku do wersji klasycznej. Wyniki dla segmentów objętościowych pokazują, że cechy intensywnościowe: L i szczególnie V, stosowane w przestrzeniach pochodzących z grafiki komputerowej, prowadzą do powiększenia klasterów w stosunku do przestrzeni korzystających z intensywności I. Wyniki przedstawione w tab. 4.1 świadczą o przewadze klasycznej transformacji $RGB \rightarrow HSI$ nad pozostałymi transformacjami. Wadą klasycznej transformacji jest długi czas obliczeń, stąd wniosek o potrzebie sprawnej implementacji jej wzorów lub zastosowanie nieco gorszej wersji HLS.

Jedną z najprostszych operacji na obrazie jest wyznaczenie jego negatywu (wzory (3.32)). Dysponując obrazem w przestrzeni HSx najlepiej byłoby wyznaczyć jego negatyw bez powrotu do przestrzeni RGB. W tym celu potrzebna jest znajomość zależności pomiędzy składowymi HSx dla negatywu i składowymi HSx obrazu wyjściowego. Poniżej zostały wyprowadzone te zależności dla HLS (transformacja z RGB wg (2.33)–(2.35)).

Aby znaleźć odcień negatywu H_{neg} przyjęto, że max = r i min = b. Wtedy procedura obliczeniowa odcienia (2.35) sprowadza się do wzoru:

$$H = \frac{g-b}{r-b} \tag{4.12}$$

Dla rozpatrywanej barwy po negacji największa składowa to składowa b, a najmniejsza r. Wtedy można na podstawie (2.35) zapisać:

$$H_{neg} = 4 + \frac{r_{neg} - g_{neg}}{max_{neg} - min_{neg}} = 4 + \frac{r_{neg} - g_{neg}}{b_{neg} - r_{neg}}$$
(4.13)

Korzystając ze wzorów definiujących negatyw w przestrzeni *RGB* można dokonać następujących przekształceń:

$$H_{neg} = 4 + \frac{(1-r) - (1-g)}{(1-b) - (1-r)} = 4 + \frac{g-r}{r-b} =$$

$$= \frac{3r - 4b + g}{r-b} = \frac{3r - 3b - b + g}{r-b} = 3 + \frac{g-b}{r-b} = 3 + H$$
(4.14)

Przechodząc z odcieniem na skalę kątową (mnożąc obie strony prze
z $60^\circ)$ otrzymamy:

$$H_{neg} = 180^{\circ} + H \tag{4.15}$$

Jasność jest obliczana według wzoru:

$$L = \frac{max + min}{2} \tag{4.16}$$

Jasność negatywu wynosi:

$$L_{neg} = \frac{max_{neg} + min_{neg}}{2} = \frac{1 - min + 1 - max}{2} = 1 - \frac{max + min}{2} = 1 - L \quad (4.17)$$

Wynik ten przez analogię do obrazów monochromatycznych jest zgodny z intuicją. Aby wyprowadzić zależność dla nasycenia przyjmijmy, jak powyżej, że max = r i min = b oraz dodatkowo, że $L \leq 0.5$. Wtedy wzór na nasycenie (2.33) przyjmuje postać:

$$S = \frac{r-b}{r+b} \tag{4.18}$$

Po negacji obrazu otrzymujemy:

$$max_{neg} = 1 - min = 1 - b$$

$$min_{neg} = 1 - max = 1 - r$$
(4.19)

ROZDZIAŁ 4. Badanie właściwości wybranych przestrzeni barw

Ponieważ jasność negatywu $L_{neg} > 0,5$ to do obliczenia nasycenia barwy negatywu należy zastosować drugą część wzoru (2.33):

$$S_{neg} = \frac{max_{neg} - min_{neg}}{2 - max_{neg} - min_{neg}} = \frac{1 - b + 1 - r}{2 - (1 - b) - (1 - r)} = 1 - \frac{r - b}{r + b} = S$$
(4.20)

Podobną zależność można wyprowadzić również dla innych relacji pomiędzy składowymi r,g i b. Podsumowując ten eksperyment, można stwierdzić, że barwa w obrazie negatywowym ma "dopełniające" odcień i jasność oraz niezmienione nasycenie. Jedynie w przypadku przestrzeni HLS można w tak prosty sposób wyprowadzić zależności pomiędzy odcieniem, jasnością i nasyceniem negatywu i pozytywu. Można łatwo pokazać, że po transformacji z RGB na HSV i z RGB na HSI nie zachodzi relacja $S_{neg} = S$.

4.1.3. Badanie struktury przestrzeni HSx

W pracy [16] wyznaczono liczbę punktów barwnych w poziomych przekrojach bryły barw HSx. Przekroje te można generować zakładając stałą wartość cechy achromatycznej np., I. Przyjmując, że znamy współrzędne środka i promień koła dla maksymalnego przekroju, dla każdego piksela z tego koła wyznaczamy jego współrzędne biegunowe. Odpowiadają one odcieniowi H i nasyceniu S. Dla 3 składowych HSx danego piksela przeprowadzamy transformację odwrotną $HSx \rightarrow RGB$ i wyświetlamy na ekranie monitora piksel o barwie RGB. Jeśli składowe RGB wychodzą poza swój zakres, to piksel przyjmuje barwę neutralnego tła. Tak postępujemy dla wszystkich pikseli koła maksymalnego przekroju. W przypadku przestrzeni HSV i HLS przekroje są sześciokątami (w granicy kołami). Interesujący wynik przynosi generowanie przekrojów poziomych bryły HSI (rys. 4.1).

Wystarczy zauważyć, że wartość nasycenia S zależy od intensywności I. Pokazuje to wyraźnie następująca postać przekształcona wzoru na nasycenie (2.28) w transformacji $RGB \rightarrow HSI$:

$$S = 1 - \frac{\min\left(R, G, B\right)}{I},\tag{4.21}$$

gdzie I = (R + G + B)/3. Dla I = 0 (punkt czerni) wartość nasycenia S jest nieokreślona. Rysunek 4.1 pokazuje, że dla niektórych odcieni (np. R, G, B) wraz ze wzrostem Iprzestaje być osiągalne maksymalne nasycenie. Te 3 barwy podstawowe osiągają wartość S = 1 tylko do poziomu intensywności I = 255/3. Barwy dopełniające (Ye, Cy, Mg) mogą osiągać maksymalne nasycenie do poziomu 2/3 wysokości bryły HSI ($I = 2 \times 255/3$). Powyżej tego poziomu żadna barwa już nie jest maksymalnie nasycena, a nasycenie maleje

4.1. Przestrzeń barw HSx



Rys. 4.1. Wybrane przekroje poziome bryły barwHSI: (a) I=50, (b) I=100, (c) I=150, (d) I=220

Fig. 4.1. Selected horizontal cross-sections of HSI colour solids: (a) I = 50, (b) I = 100, (c) I = 150, (d) I = 220

do punktu bieli. W pewien niedoskonały sposób ilustruje to zjawisko model bryły HSI przedstawiony na rys. 4.2.

Przekroje poziome oprócz informacji o kształcie bryły barw mogą być również źródłem informacji o liczbie i rozmieszczeniu punktów w przestrzeni. Problem zliczania punktów leżących w danym przekroju poziomym bryły barw w przestrzeni HSI można sprowadzić do zadania kombinatorycznego. Podobne podejście dla przekrojów prostopadłych do osi sześcianu RGB zastosowano w pracy [237]. W pracy [16] wyprowadzono wzory pozwalające zliczać punkty w poszczególnych przekrojach poziomych. Cechą wspólną badanych przestrzeni HSx jest to, że dla małych wartości intensywności liczba istniejących punktów-barw jest mała. W przypadku przestrzeni HSI jak i HLS przekroje o największej ilości punktów występują w zakresie średnich jasności, podczas gdy dla przestrzeni





Rys. 4.2. Trójwymiarowe modele brył barw: (a) HSI, (b) HSV [47]Fig. 4.2. Threedimensional models of colour solids: (a) HSI, (b) HSV [47]

HSV w zakresie maksymalnych intensywności, co pokazuje rys. 4.2(b) [47].

Podsumowując powyższe rozważania można stwierdzić, że nieliniowe transformacje z przestrzeni RGB skutkują nierównomiernym rozłożeniem punktów w przestrzeniach typu HSx. Ta nierównomierność jest przyczyną występowania w obszarach o małej gęstości punktów dużych skokowych zmian wartości składowych HSx w odpowiedzi na małe zmiany RGB. Porównanie 4 wariantów przestrzeni HSx na podstawie kryterium zwartości klasterów barwnych w tych przestrzeniach, szczególnie ważne dla rozpoznawania obiektów barwnych, wskazuje na przewagę przestrzeni HSI nad innymi wariantami. Z drugiej strony to wariant HLS pozwala łatwo wyznaczyć negatyw obrazu bez konieczności powrotu do przestrzeni bazowej RGB, co może mieć praktyczne znaczenie w przetwarzaniu obrazów. Autor badając jedynie najpopularniejsze przestrzenie percepcyjne zdaje sobie sprawę z tego, że w tej dziedzinie nie powiedziano jeszcze ostatniego słowa: istniejące modele są wynikiem kompromisu pomiędzy ich złożonością a dokładnością aproksymacji percepcji barwy.

4.2. Przestrzeń $K_1 K_2 K_3$

W pozyskanych obrazach barwnych pomiędzy pikselami występuje redundancja przestrzenna (ang. *spatial redundancy*) a pomiędzy składowymi barwy pikseli redundancja widmowa (ang. *spectral redundancy*), które są wykorzystywane do efektywnej kompresji tych obrazów. Rezultatem redundancji widmowej jest znaczny stopień skorelowania skła-

4.2. Przestrzeń $K_1K_2K_3$

dowych RGB. W celu zilustrowania tego zjawiska wyznaczono współczynniki korelacji składowych RGB pikseli w obrazach z zestawu obrazów scen naturalnych firmy Kodak (krajobrazy, portrety itp.) dostępnych na stronie Center for Image Processing Research (Troy, NY, USA) [132]. Rozdzielczość tych obrazów wynosiła 768 × 512, a barwa była reprezentowana przez 24 bity/piksel tzn., 8 bitów/piksel dla każdej składowej R, G, B. Wyniki zebrano w tab. 4.2.

Tabela 4.2. Współczynniki korelacji składowych RGB dla badanych obrazów barwnych

Lp.	Nazwa obrazu	ρ _{RG}	ρ _{RB}	ρ _{GB}
1.	IMAGE01	0,86	0,84	0,99
2.	IMAGE02	0,51	0,39	0,97
3.	IMAGE03	0,72	0,28	0,55
4.	IMAGE04	0,60	0,68	0,96
5.	IMAGE05	0,90	0,82	0,90
6.	IMAGE06	0,98	0,96	0,99
7.	IMAGE07	0,83	0,76	0,91
8.	IMAGE08	0,97	0,92	0,97
9.	IMAGE09	0,95	0,73	0,85
10.	IMAGE10	0,95	0,89	0,97
11.	IMAGE11	0,82	0,79	0,97
12.	IMAGE12	0,91	0,87	0,97
13.	IMAGE13	0,98	0,95	0,96
14.	IMAGE14	0,86	0,45	0,71
15.	IMAGE15	0,86	0,86	0,99
16.	IMAGE16	0,98	0,89	0,95
17.	IMAGE17	0,99	0,93	0,98
18.	IMAGE18	0,92	0,75	0,84
19.	IMAGE19	0,97	0,82	0,92
20.	IMAGE20	1,00	0,96	0,98
21.	IMAGE21	0,89	0,66	0,91
22.	IMAGE22	0,87	0,80	0,87
23.	IMAGE23	0,61	0,42	0,64
	Wartość średnia	0,74	0,63	0,82
Odch	vlenie standardowe	0,14	0,20	0,12
Juon		, , , , , , , , , , , , , , , , , , , ,		

Jak widać z tab. 4.2 składowe RGB są silnie skorelowane, najsilniej składowe G i B, a najmniejsze wartości korelacji występują pomiędzy składowymi R i B. Podobne zależności uzyskano dla innych zestawów obrazów w pracach [21, 22]. W pracy [20] po-

kazano, że transformacje przestrzeni barw RGB na przestrzenie typu "jedna składowa luminancyjna, dwie składowe chrominancyjne" (np. YUV, YIQ, HSI itp.) częściowo dekorelują składowe barw; w mniejszym stopniu dzieje się to dla przestrzeni powstałych w wyniku transformacji liniowej, a w większym stopniu dla nieliniowej transformacji z RGB na HSI. W przypadku tej przestrzeni największa korelacja występuje pomiędzy nasyceniem S i intensywnością I. Jednym z głównych źródeł korelacji składowych RGBjest nachodzenie na siebie widmowych charakterystyk czułościowych kamery [20], co tłumaczy najmniejszą korelację pomiędzy składowymi R i B, tj. tymi składowymi, którym odpowiadają najbardziej odległe od siebie charakterystyki widmowe.

4.2.1. Dekorelacja składowych RGB

Całkowitej dekorelacji składowych RGB można dokonać za pomocą tzw. transformacji Karhunena-Loèvego (KLT), znanej również w literaturze jako transformacja Hottelinga [95] czy analiza głównych składowych (ang. *Principal Components Analysis, PCA*) [101]. Transformacja ta zastosowana do obrazu barwnego zapewnia otrzymanie 3 kolejnych obrazów składowych o nieskorelowanych wartościach i malejącym znaczeniu statystycznym (coraz mniejsze wartości własne, energia obrazu zawarta głównie w pierwszej składowej). Tak więc KLT minimalizując korelację pomiędzy składowymi jednocześnie maksymalizuje zawartość informacji w pierwszej składowej. Można postrzegać KLT jako obrót układu współrzędnych w przestrzeni barw, pozwalający osi K_1 "ustawić się" w kierunku największej zmienności barwy w obrazie. Wzrost mocy obliczeniowej komputerów, który nastąpił w ostatnich latach, umożliwił nowe zastosowania tej klasycznej metody przetwarzania sygnałów o dużej złożoności obliczeniowej. Transformacja KLT dla obrazu barwnego RGBzdefiniowana jest następująco [233]:

$$\begin{bmatrix} K_1 \\ K_2 \\ K_3 \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33}, \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$
(4.22)

a macierz transformująca sprowadza macierz wariancyjno-kowariancyjną do postaci diagonalnej:

m_{11}	m_{12}	m_{13}	$\int u_{11}$	u_{12}	u_{13}	m_{11}	m_{21}	m ₃₁		λ_1	0	0]	
m_{21}	m_{22}	m_{23}	u_{12}	u_{22}	u_{23}	m_{12}	m_{22}	m_{32}	=	0	λ_2	0	(4.23)
m_{31}	m_{32}	m_{33}	u_{13}	u_{23}	<i>u</i> ₃₃ .	m_{13}	m_{23}	m33 .		0	0	λ_3 ,	. ,

gdzie: λ_1 , λ_2 , λ_3 – wartości własne macierzy wariancyjno-kowariancyjnej obrazu RGB złożonej z elementów u_{ij} , a $[m_{11}, m_{12}, m_{13}]$, $[m_{21}, m_{22}, m_{23}]$, $[m_{31}, m_{32}, m_{33}]$ – wektory własne

tej macierzy. Ponieważ macierz wariancyjno-kowariancyjna jest rzeczywistą macierzą symetryczną, to jej wartości własne są liczbami rzeczywistymi, a odpowiadające im wektory własne są ortogonalne [124].

Podobne równania transformacji KLT można zapisać również dla innych od *RGB* przestrzeni barw. W pracy [39] zastosowano do kodowania obrazów barwnych transformację KLT na obrazie w przestrzeni *CIELAB* a w [300] rozważano transformację KLT na składowych tzw. przestrzeni czopków. Transformacja KLT wyznaczając nowe kierunki w przestrzeni barw nie zmienia jej równomierności percepcyjnej.

Wyznaczenie obrazu po transformacji KLT wymaga następujących działań:

- obliczenia macierzy wariancyjno-kowariancyjnej dla obrazu w przestrzeni wyjściowej,
- obliczenia wartości własnych i wektorów własnych tej macierzy składających się na macierz transformującą,
- pomnożenia macierzy transformacji przez macierz obrazu wyjściowego,
- przeskalowanie otrzymanych wyników do zakresu (0, 255).

W pracy [219] opisano efektywną implementację algorytmu KLT. W algorytmie tym można wyróżnić dwa etapy wymagające złożonych obliczeń numerycznych: wyznaczanie macierzy kowariancji przekształcanego obrazu (najbardziej czasochłonne) oraz obliczanie 3 wartości własnych i 3 wektorów własnych tej macierzy. W powstałym programie zastosowano dwa algorytmy obliczania wartości własnych macierzy o wymiarze 3 × 3 algorytm oparty na metodzie iteracyjnej Kryłowa [80] oraz algorytm oparty na rozwiązywaniu równań 3 stopnia metodą Cardana [69]. Metoda Kryłowa umożliwia poprawne wyznaczenie maksymalnej wartości własnej macierzy w sposób iteracyjny, natomiast pozostałe 2 wartości własne można obliczyć analitycznie rozwiązując równanie 2 stopnia, co jest już proste w realizacji numerycznej. Iteracyjność metody powoduje, że nie zawsze prowadzi ona do rozwiązania, ponieważ jej zbieżność silnie zależy od punktu początkowego. W części programu wykorzystującej metodę Kryłowa zastosowano dwa punkty startowe, tzn. jeżeli metoda nie daje wyników z założoną dokładnością po określonej liczbie iteracji, to obliczenia są rozpoczynane ponownie dla drugiego punktu startowego. Mimo tego zabezpieczenia, które ma znaczący wpływ na czas obliczeń, istnieją obrazy, dla których metoda Kryłowa nie jest w stanie wyznaczyć wartości własnych macierzy kowariancji.

ROZDZIAŁ 4. Badanie właściwości wybranych przestrzeni barw

Metoda Cardana pozwala bezpośrednio rozwiązać równanie 3 stopnia w sposób analityczny, a zatem dokładny, jednak jest bardziej złożona obliczeniowo (m.in. obliczanie wartości funkcji arc cos).

Po wyznaczeniu wartości własnych macierzy kowariancji należy wyznaczyć odpowiadające im wektory własne. Obliczenie 3 wektorów własnych wymaga rozwiazania 3 układów 3 równań z 3 niewiadomymi. Równania te sa jednak równaniami, w których wyraz wolny jest równy 0, dlatego metody numeryczne rozwiazywania takich układów równań daja rozwiązanie trywialne: zerowe. W celu poprawnego rozwiązania tych układów równań została wykorzystana własność wektorów własnych polegająca na tym, iż wektory własne moga być skalowane, a zatem wektor, który nie jest wektorem zerowym, po przeskalowaniu może zawierać składowa równa 1. Wstawiając zatem wartość 1 w miejsce kolejno pierwszej. drugiej i trzeciej niewiadomej możemy redukując rozmiar takiego układu równań znaleźć poprawne rozwiązanie. Gdy układ równań pomimo zastosowania takiej metody nie ma rozwiązania, oznacza to, że wektor własny jest wektorem zerowym. Jeżeli rozwiązań jest wiecej niż jedno, to sa one liniowo zależne i do obliczeń można przyjać dowolne z nich. Dlatego po znalezieniu pierwszego rozwiązania obliczenia są przerywane i pozostałe rozwiązania nie sa już szukane. Znalezione wektory własne sa normalizowane tak, aby ich długości były równe 1. Ze znormalizowanych wektorów własnych zostaje zestawiona macierz transformacji KLT i użyta do dokonania transformacji.

Transformacja KLT obrazu barwnego o wymiarach 768 × 512 za pomocą programu powstałego w wyniku implementacji opisanego algorytmu zajmuje komputerowi klasy PC z mikroprocesorem Pentium (3,4 GHz) 0,6 sekundy. Transformacja KLT może być również zaimplementowana w środowisku Matlab, ponieważ zapewnia ono gotowe funkcje do wyznaczania macierzy kowariancji oraz jej wartości i wektorów własnych. Istotną wadą KLT jest fakt, że każdy nowy obraz ze względu na swoistą macierz wariancyjno-kowariancyjną wymaga obliczenia swojej macierzy transformującej, co jest czasochłonne i mimo efektywnej implementacji ogranicza pole zastosowań. Przykładowy wynik działania programu przedstawia rys. 4.3. Oprócz obrazu oryginalnego pokazano na nim trzy obrazy wynikowe $K_1K_2K_3$ o malejącym znaczeniu statystycznym. Składowa K_1 (najwyższa wartość własna) zawiera najwięcej informacji i stanowi dobrej jakości obraz z poziomami szarości, natomiast w składowej K_3 (najniższa wartość własna) dominuje szum.

4.2. Przestrzeń $K_1K_2K_3$





Rys. 4.3. Rozkład obrazu AIRPLANE na składowe $K_1K_2K_3$: (a) obraz oryginalny, (b) składowa K_1 ($\lambda_1 = 5384,33$), (c) składowa K_2 ($\lambda_2 = 246,76$), (d) składowa K_3 ($\lambda_3 = 73,63$)

Fig. 4.3. Splitting of AIRPLANE image into $K_1K_2K_3$ components: (a) original image, (b) K_1 component ($\lambda_1 = 5384,33$), (c) K2 component ($\lambda_2 = 246,76$), (d) K_3 component ($\lambda_3 = 73,63$)

4.2.2. Podstawowe właściwości KLT

Dysponując programem dokonującym transformacji KLT można przekonać się doświadczalnie o dwóch podstawowych właściwościach transformacji: braku korelacji pomiędzy składowymi $K_1K_2K_3$ oraz o skupieniu energii sygnału (ang. energy compaction), głównie w pierwszej składowej. Niezależnie od obrazu transformacja KLT dokonuje pełnej dekorelacji składowych, tzn. macierz kowariancji dla obrazu w przestrzeni $K_1K_2K_3$ jest ma-

4.2. Przestrzeń $K_1 K_2 K_3$

ROZDZIAŁ 4. Badanie właściwości wybranych przestrzeni barw

cierzą diagonalną i jej wartości własne są równe wariancjom poszczególnych składowych. Do oceny skupienia energii w składowych sygnału zastosowano wariancje tych składowych. Wyznaczono skupienie energii jako udział wariancji pojedynczej składowej w sumie wariancji składowych obrazu, przykładowo dla składowej *R*:

$$E(R) = \frac{\sigma_R^2}{\sigma_R^2 + \sigma_G^2 + \sigma_B^2} \tag{4.24}$$

Wyniki dla 23 obrazów firmy Kodak składowych RGB zebrano w tab. 4.3 Potwierdzają one prawie równomierne rozłożenie energii sygnału obrazowego w trzech składowych RGB oraz zjawisko skupienia energii sygnału głównie w pierwszej składowej KLT. Dla niektórych obrazów składowa K_3 ma zerowy wkład do energii sygnału obrazowego. Jest to typowa sytuacja dla obrazów sztucznych, tj. pochodzących z programów grafiki komputerowej. Również dla obrazu AIRPLANE z rys. 4.3 energia jest skupiona głównie w pierwszej składowej: 95%, 4%, 1%.

W przypadku wielu obrazów barwnych poddanych transformacji KLT obraz K_1 zwraca uwagę swym bardzo dobrym kontrastem. Poniżej porównano obraz K_1 z odpowiednim obrazem jasnościowym Y, uzyskanym jako średnia ważona 3 obrazów składowych R, G, B (2.13). W przetwarzaniu obrazów nie istnieje jedna powszechnie uznana definicja kontrastu obrazu. Obok klasycznych definicji Michelsona czy Webera [227], stosowanych do obliczania kontrastu prostych obrazów, niekiedy używa się wariancji poziomów szarości σ^2 w obrazie jako miary jego kontrastu [95]. Taką miarę zastosowano również w opisywanym eksperymencie. Wykresy na rys. 4.4 zbudowano dla ciągu obrazów powstałych z 5 barwnych obrazów wyjściowych poprzez programowe obniżanie ich jasności; im wyższy numer obrazu, tym niższa jego jasność. Wykresy te pokazują, że w miarę obniżania jasności obrazu, symulującego spadek natężenia oświetlenia sceny, kontrast w obrazach K_1 coraz bardziej przewyższa kontrast w obrazach Y. Z wyników tego eksperymentu można wywnioskować, że obrazy z kamery kolorowej poddane transformacji KLT zapewniają systemowi wizyjnemu pracującemu w warunkach zmiennego oświetlenia dostęp do obrazów monochromatycznych o podwyższonym kontraście.

W wyniku transformacji KLT następuje zmniejszenie redundancji widmowej w obrazie barwnym. Kodowanie takiego obrazu za pomocą bezstratnej techniki powinno więc wymagać mniejszej ilości bitów. Przeprowadzono eksperyment kodując 23 obrazy barwne za pomocą bezstratnej wersji techniki JPEG (tzw. JPEG-LS). Wyniki zamieszczone w tab. 4.4 pokazują, że wcześniejsza transformacja obrazu do przestrzeni $K_1K_2K_3$ średnio zmniejsza liczbę bitów na piksel o ponad 2b/p.

Lp.	Nazwa obrazu	E(R)	E(G)	E(B)	$E(K_1)$	$E(K_2)$	$E(K_3)$
1.	IMAGE01	24%	42%	34%	94%	5%	0%
2.	IMAGE02	45%	28%	27%	73%	27%	0%
3.	IMAGE03	34%	34%	32%	69%	24%	7%
4.	IMAGE04	23%	37%	40%	87%	12%	1%
5.	IMAGE05	34%	38%	28%	92%	6%	2%
6.	IMAGE06	42%	35%	23%	98%	2%	0%
7.	IMAGE07	36%	31%	33%	89%	9%	2%
8.	IMAGE08	35%	37%	28%	97%	3%	0%
9	IMAGE09	39%	36%	25%	91%	8%	1%
10.	IMAGE10	36%	38%	26%	96%	4%	0%
11.	IMAGE11	29%	37%	34%	91%	8%	1%
12.	IMAGE12	26%	40%	34%	95%	4%	1%
13.	IMAGE13	32%	33%	35%	98%	2%	0%
14.	IMAGE14	41%	36%	23%	81%	16%	3%
15.	IMAGE15	25%	41%	34%	95%	5%	0%
16.	IMAGE16	35%	36%	29%	96%	4%	0%
17.	IMAGE17	31%	36%	33%	98%	2%	0%
18.	IMAGE18	49%	31%	20%	91%	7%	2%
19.	IMAGE19	39%	31%	30%	93%	6%	1%
20.	IMAGE20	32%	33%	35%	98%	2%	0%
21.	IMAGE21	29%	33%	38%	89%	11%	0%
22.	IMAGE22	34%	29%	37%	90%	7%	3%
23.	IMAGE23	38%	29%	33%	70%	21%	9%
	Wartość średnia	31%	36%	33%	82%	13%	5%
Odch	vlenie standardowe	6,5%	3,8%	5,1%	8,8%	7,1%	2,4%

Tabela 4.3. Skupienie energii dla składowych obrazów w przestrzeniRGBi $K_1K_2K_3$

Tabela 4.4. Współczynniki kompresji obrazów bezstratną techniką JPEG-LS w przestrzeniRGBi $K_1K_2K_3~({\rm w~bitach~na~piksel})$

Lp.	Nazwa obrazu	RGB	$K_1K_2K_3$
1.	IMAGE01	18,05	15,71
2.	IMAGE02	14,26	12,98
3.	IMAGE03	12,37	9,98
4.	IMAGE04	14,72	13,29
5.	IMAGE05	18,13	13,62
6.	IMAGE06	15,58	13,34
7.	IMAGE07	13,32	11,01
8.	IMAGE08	19,22	16,10
9.	IMAGE09	14,45	13,43
10.	IMAGE10	14,43	11,74
11.	IMAGE11	15,33	12,50
12.	IMAGE12	13,13	10,96
13.	IMAGE13	19,39	14,68
14.	IMAGE14	16,43	13,92
15.	IMAGE15	14,52	12,17
16.	IMAGE16	13,81	12,06
17.	IMAGE17	14,80	11,88
18.	IMAGE18	17,24	14,30
19.	IMAGE19	15,96	12,44
20.	IMAGE20	12,16	11,07
21.	IMAGE21	15,51	13,58
22.	IMAGE22	16,14	14,01
23.	IMAGE23	13,16	11,19
	Wartość średnia	15,61	13,45
Odch	ylenie standardowe	2,04	1,56

4.2. Przestrzeń $K_1 K_2 K_3$

4.2.3. Zastosowania KLT

Złożoność obliczeniowa KLT powoduje, że poszukiwane są liniowe transformacje składowych RGB dające wyniki zbliżone do KLT. Już w 1980 roku zaproponowano aproksymację KLT poprzez stosowanie składowych $I_1I_2I_3$ niezależnych od treści obrazu, nazywanych odtąd w literaturze przestrzenią Ohty [172], która została przedstawiona w rozdziale 2.5. Macierz transformującą ustalono na podstawie wektorów własnych macierzy wariancyjno-kowariancyjnych 8 wybranych obrazów. W pracy [59] zaproponowano 2 inne "KLT-podobne" transformacje. Jedna z nich oznaczona jako KLA (ang. *KL Average*) powstała w wyniku wyznaczenia macierzy transformującej KLT dla metaobrazu złożonego ze wszystkich badanych obrazów, a druga była jej uproszczeniem.

W pracy [21] pokazano za pomocą testów na standardowych obrazach barwnych, że wbrew zapowiedziom zamieszczonym w pracach [59,172] aproksymacje transformacji KLT nie dekorelują całkowicie składowych barwy, a jedynie zmniejszają stopień ich skorelowania do poziomu 0,35–0,45. Pokazano również, że lepszy efekt przynosi zdefiniowanie klasy obrazów podobnych i wyznaczenie dla niej macierzy transformującej KLT w postaci macierzy z uśrednionymi na kilku obrazach z tej klasy wektorami własnymi [21, 22]. Przez obrazy jednolite rozumieć można obrazy pozyskane za pomocą tego samego systemu wizyjnego (kamera, oświetlenie itp.) ze scen o podobnej kolorystyce. Dla określonej tak klasy obrazów pozyskanych dla np. zdeterminowanej sceny robota można znaleźć w ten sposób stałą transformację, tzn. niezależną od obrazu i nie obliczać współczynników transformacji za każdym razem. W tab. 4.5 zestawiono współczynniki korelacji składowych przestrzeni będącej aproksymacją KLT ze współczynnikami dla przestrzeni Ohty. Wyniki te świadczą o przewadze aproksymacji KLT nad przestrzenią Ohty.

Tabela 4.5. Porównanie wyników dekorelacji dla 3 klas obrazów podobnych

	Klasa1			Klasa2			Klasa3		
Proponowana	K_1K_2	K_1K_3	$\overline{K_2}K_3$	K_1K_2	K_1K_3	$\overline{K_2K_3}$	K_1K_2	K_1K_3	K_2K_3
aproksymacja	0,06	0,22	0,10	0,03	0,03	0,01	0,12	0,08	0,07
Przestrzeń	I_1I_2	I_1I_3	I_2I_3	I_1I_2	I_1I_3	I_2I_3	I_1I_2	I_1I_3	I_2I_3
Ohty	0,85	0,75	0,80	0,23	0,41	0,21	0,51	0,32	0,11

Złożoność obliczeniowa KLT spowodowała prawdopodobnie zainteresowanie separowalną transformacją Karhunena-Loevego (SKLT) [35], która jest szybsza i wymagająca





4.2. Przestrzeń $K_1K_2K_3$

mniej pamięci niż KLT. W literaturze można spotkać w ostatnich latach próby wykorzystania transformacji KLT w procesie segmentacji obrazów barwnych. W pracy [253] do segmentacji obrazu skóry na obszar podejrzany o zmiany nowotworowe i jego otoczenie zastosowano z pozytywnym skutkiem metodę złożoną z transformacji KLT obrazu RGB i progowania histogramu najważniejszego obrazu składowego powstałego w wyniku transformacji. W artykule [76] pokazano, że segmentacja barwnego obrazu skóry zastosowana do 2 pierwszych składowych KLT daje lepsze wyniki niż ta sama metoda zastosowana na obrazie w innej przestrzeni barw np., HSI, CIELUV, CIELAB. W artykule [265] segmentacje wododziałowa (ang. watershed segmentation) obrazu barwnego poprzedzono dekorelacja składowych RGB za pomocą transformacji KLT. Wykorzystywany w tej metodzie segmentacji tzw. obraz aktywności (ang. activity image) obliczono jako sumę ważoną obrazów aktywności obliczonych dla obrazów składowych powstałych po transformacji KLT. Współczynniki wagi przy pierwszym i drugim obrazie składowym były wyższe niż przy trzecim. Wynik uzyskany za pomocą tej metody segmentacji obrazu barwnego porównano z odpowiednim wynikiem dla obrazu intensywnościowego, wskazując na znacznie mniejszą nadsegmentację (ang. oversegmentation) w przypadku obrazu barwnego.

Szczególną wartość wśród przeprowadzonych w tym podrozdziale badań stanowi efektywna implementacji transformacji KLT dla obrazów barwnych. Stwierdzona mniejsza wrażliwość składowej K_1 niż składowej jasnościowej Y na zmiany luminancji obrazu ma znaczenie praktyczne. Rozważmy system wizji monochromatycznej pracujący w warunkach zmiennego oświetlenia. Przy coraz niższych cenach kamer kolorowych można zaproponować zastosowanie w nim takiej kamery i dalsze przetwarzanie jedynie składowej K_1 zamiast Y. Pozyskiwane będą w ten sposób obrazy monochromatyczne o wysokim kontraście. Dokonano też krytycznej oceny znanych z literatury aproksymacji KLT, proponując jednocześnie wyznaczanie takich współczynników transformujących, które zapewniając dekorelację składowych są stałe dla pewnej klasy obrazów.

W tym rozdziale na dwóch zupełnie różnych przykładach pokazano, że wiedza o przestrzeniach barw nie sprowadza się jedynie do znajomości transformacji tworzących te przestrzenie i może być bezpośrednio wykorzystana w przetwarzaniu obrazów.

Rys. 4.4. Wykresy zmian kontrastu w obrazach K_1 i Y wywołanych zmianą jasności barwnych obrazów wyjściowych: (a) obraz IMAGE01, (b) obraz IMAGE05, (c) obraz IMAGE07, (d) obraz IMAGE16, (e) obraz IMAGE21

Fig. 4.4. Changes in the contrast of K1 and Y images produced by lightness changes in colour images: (a) IMAGE01, (b) IMAGE05, (c) IMAGE07, (d) IMAGE16, (e) IMAGE21

.

Barwne wersje ilustracji rozdziału 4



(c)



- Rys. 4.1. Wybrane przekroje poziome bryły barw HSI: (a) I = 50, (b) I = 100, (c) I = 150, (d) I = 220
- Fig. 4.1. Selected horizontal cross-sections of HSI colour solids: (a) I = 50, (b) I = 100, (c) I = 150, (d) I = 220



Rys. 4.2. Trójwymiarowe modele brył barw: (a) HSI, (b) HSV[47]Fig. 4.2. Threedimensional models of colour solids: (a) HSI, (b) HSV[47]



(a)



- Rys. 4.3. Rozkład obrazu AIRPLANE na składowe $K_1K_2K_3$: (a) obraz oryginalny, (b) składowa K_1 ($\lambda_1 = 5384,33$), (c) składowa K_2 ($\lambda_2 = 246,76$), (d) składowa K_3 ($\lambda_3 = 73,63$)
- Fig. 4.3. Splitting of AIRPLANE image into $K_1K_2K_3$ components: (a) original image, (b) K_1 component ($\lambda_1 = 5384,33$), (c) K2 component ($\lambda_2 = 246,76$), (d) K_3 component ($\lambda_3 = 73,63$)

Rozdział 5 Segmentacja obrazów barwnych

W różnych zastosowaniach przetwarzania obrazów barwnych podstawową rolę odgrywają techniki automatycznej segmentacji obrazu. Od jakości segmentacji zależą bowiem rezultaty dalszych etapów przetwarzania obrazu: rozpoznawania i śledzenia obiektów, przeszukiwania obrazowych baz danych itd. Celem procesu segmentacji jest podział obrazu na jednorodne i spójne obszary odpowiadające obiektom lub ich częściom znajdującym się w scenie przedstawianej przez obraz. Często segmentacja odbywa się bez stosowania dodatkowej wiedzy o obiektach w obrazie. Jednorodność obszarów w procesach segmentacji obrazów barwnych w naturalny sposób odnosi się do barwy, rzadziej do barwnej tekstury [62]. Niekiedy nie ma potrzeby przeprowadzania całkowitej segmentacji obrazu, wystarczy segmentacja częściowa, polegająca na wydzieleniu z obrazu jedynie interesujących obszarów (ang. *region of interest ROI, salient regions*). Obszary w obrazie po segmentacji posiadają w przeciwieństwie do pojedynczych pikseli wiele interesujących właściwości, takich jak np. kształt, tekstura itp. Można zauważyć, że człowiek za pomocą swojego układu wzrokowego rozpoznaje obiekty w otoczeniu dokonując niejako "po drodze" automatycznej segmentacji obrazu w sposób naturalny i nieświadomy.

Pierwsze artykuły przeglądowe na temat segmentacji obrazów barwnych pochodzą z lat dziewięćdziesiątych [40, 116, 145, 262]. Również w pierwszych podręcznikach przetwarzania obrazów barwnych można znaleźć rozdziały poświęcone tematowi segmentacji obrazów barwnych [231,248]. Wiele technik segmentacji opracowanych wcześniej dla obrazów monochromatycznych [177] znalazło zastosowanie w segmentacji obrazów barwnych. Dobrym przykładem jest technika podziału i łączenia (ang. *split and merge*). Pierwsza publikacja poświęcona tej technice w zastosowaniu do obrazów monochromatycznych [110] pochodzi z 1976 roku, a zastosowanie w obrazach barwnych [51] zostało zaproponowane w 1989 roku. Każde takie rozszerzenie zastosowania techniki segmentacji na obrazy barwne

ROZDZIAŁ 5. Segmentacja obrazów barwnych

jest związane z decyzją o wyborze pewnej przestrzeni barw. Chociaż techniki segmentacji korzystają z bardzo różnorodnego aparatu matematycznego, jednak jak do tej pory nie udało się opracować jednej skutecznej metody pozwalającej na segmentację każdego obrazu.

Nietrudno zauważyć, że wysegmentowanie niektórych obiektów wymaga przetwarzania ich obrazu barwnego. Przykład takiej sytuacji pokazano na rys. 5.1, przedstawiającym wynik segmentacji kwiatu z obrazu barwnego i luminancyjnego KWIAT1 [207]. Te dwa obrazy poddano segmentacji techniką ziarnistego rozrostu obszaru, opisaną w dalszej części tego rozdziału. Celem segmentacji było wysegmentowanie płatków niebieskiego kwiatu umieszczonego centralnie w obrazie. Segmentacja w przypadku obu obrazów (rys. 5.1(a) i rys. 5.1(b)) zaczynała się od ziarna umieszczonego za każdym razem w tym samym punkcie górnego płatka kwiatu. Próby wysegmentowania całego kwiatu z obrazu luminancyjnego związane ze zmianą parametru d (rys. 5.1(c)-(e)) nie dały rezultatów. Segmentacja obrazu barwnego tą samą metodą przyniosła zadowalający rezultat (rys. 5.1(f)), pokazując dobitnie możliwości zawarte w procesach przetwarzania obrazów barwnych.

Jeżeli obraz po segmentacji zawiera dużo małych obszarów odpowiadających jednolitym obiektom w obrazie wyjściowym, to można mówić o nadsegmentacji (ang. oversegmentation) i odwrotnie, gdy zawiera on małą liczbę dużych obszarów, z których każdy odpowiada kilku obiektom, to można mówić o niedosegmentacji (ang. undersegmentation). Na rys. 5.2 [207] przedstawiono dla obrazu PARROTS przykłady nadsegmentacji i niedosegmentacji. Aby lepiej uwidocznić nadsegmentację zastosowano losowy dobór pseudokolorów dla poszczególnych obszarów (rys. 5.2(d)). Błędna segmentacja obrazu (np. nadsegmentacja, niedosegmentacja) pociąga za sobą błędy w procesach analizy i rozpoznawania obrazów. Nadsegmentacja jest jednak wygodniejsza w dalszym przetwarzaniu, ponieważ za pomocą odpowiedniego przetwarzania końcowego można zmniejszyć liczbę obszarów w obrazie. Rysunek 5.2(e) przedstawia względnie dobry wynik segmentacji obrazu PARROTS (62 obszary). Do przedstawienia wyniku segmentacji zastosowano białe krawędzie obszarów naniesione na obraz oryginalny.

Wśród wielu istniejących technik segmentacji obrazów barwnych można wyróżnić cztery zasadnicze klasy: techniki pikselowe, obszarowe, krawędziowe i hybrydowe, które polegają na złożeniu 2 technik z poprzednich klas. Najczęściej składane są techniki pikselowe z obszarowymi [38] oraz techniki obszarowe z krawędziowymi [81]. Niekiedy w tym podziale technik segmentacji wyodrębnia się, ze względu na specyfikę stosowanych narzędzi matematycznych, techniki grafowe, morfologię matematyczną, techniki rozmyte [40, 177]







(c)

(d)



- Rys. 5.1. Wyniki segmentacji obrazu KWIAT1: (a) obraz oryginalny, (b) obraz luminancyjny, (c) posegmentowany obraz luminancyjny z parametrem d = 30, (d) posegmentowany obraz luminancyjny z parametrem d = 50, (e) posegmentowany obraz luminancyjny z parametrem d = 70, (f) posegmentowany obraz luminancyjny z parametrem d = 100
- Fig. 5.1. Example segmentation results: (a) colour image KWIAT1, (b) greyscale image, (c) segmented greyscale image (parameter d = 30), (d) segmented greyscale image (parameter d = 50), (e) segmented greyscale image (parameter d = 70), (f) segmented greyscale image (parameter d = 100)

(a)



3

(c)



Rys. 5.2. Wyniki segmentacji obrazu PARROTS: (a) obraz oryginalny, (b) niedosegmentacja — 48 obszarów, (c) nadsegmentacja — 3000 obszarów, (d) obraz (c) przedstawiony w pseudokolorach, (e) obraz posegmentowany — 62 obszary

(e)

Fig. 5.2. Segmentation of image PARROTS: (a) original image, (b) undersegmentation — 48 regions, (c) oversegmentation — 3000 regions, (d) previous image presented in pseudocolours, (e) segmented image — 62 regions czy techniki wykorzystujące sieci neuronowe [40,145]. Segmentacja obrazów barwnych jest ciągle przedmiotem intensywnych badań.

Metody segmentacji w różnym stopniu korzystają z informacji apriorycznej o scenie lub obrazie tej sceny. Wiedzę tę może stanowić liczba obiektów w scenie, różna od reszty sceny barwa obiektu, który należy zlokalizować, barwa tła, charakterystyki oświetlenia itp. Niekiedy jest to szczegółowa wiedza o rozpoznawanym obiekcie. Dobrym przykładem jest rozpoznawanie twarzy w obrazie barwnym [239]. Korzysta się tutaj z wiedzy o tym, że obszar oczu ma bardzo niskie nasycenie i musi być otoczony przez obszar skóry. Gdy obszar oczu zostanie znaleziony, wykorzystuje się naturalną geometrię twarzy i szuka obszaru ust pod oczami. Jako metodę segmentacji stosowano w tym przypadku progowanie (ang. *thresholding*) w przestrzeni *HSI*.

Posiadanie wiedzy dodatkowej pozwala efektywnie stosować najprostsze algorytmy segmentacji, np. progowanie. Im więcej wiadomo o scenie, tym prostsze, a więc i szybsze algorytmy segmentacji obrazu można stosować. Przyspieszenie procesu segmentacji obrazu barwnego można również osiągnąć nie korzystając z wiedzy a priori, a stosując np. wstępną klasteryzację płaszczyzny chromatyczności [250]. W niektórych zastosowaniach, m.in. w robotyce, wiedza o scenie może być dostępna, a nawet istnieje możliwość kształtowania sceny na etapie projektowania systemu.

Istotnym elementem wiedzy a priori jest wiedza o barwie tła sceny. W literaturze na temat segmentacji obrazów z poziomami szarości są znane prace przedstawiające segmentację obrazu poprzez ekstrakcję tła [244]. Szczególną rolę odgrywa estymacja tła w analizie sekwencji obrazów [242]. Automatyczna segmentacja barwnych obrazów scen naturalnych na obszary tła i pierwszego planu bez żadnej wiedzy a priori o tle wymaga stosowania bardzo złożonych zintegrowanych metod segmentacji obszarowej i krawędziowej [113].

Autor ograniczył swe prace w zakresie segmentacji obrazów barwnych do technik pikselowych i obszarowych. Techniki pikselowe (progowanie w HSI, klasteryzacja k-means) to techniki proste obliczeniowo, szczególnie w przypadku wykorzystania dodatkowej wiedzy o scenie (barwa tła, występowanie odblasków itp.). Sformułowano nową całościową propozycję progowania różnych składowych HSI w zależności od barwy tła. Cenne zastosowanie znalazła tutaj wiedza o właściwościach przestrzeni HSI przedstawiona w poprzednim rozdziale. Badano techniki obszarowe, reprezentowane głównie przez techniki rozrostu obszaru, w których proces rozrostu przebiegał "od ziaren" jak i "bez ziaren". Pokazano, w jakim stopniu przetwarzanie końcowe poprawia wyniki segmentacji i jak ilościowo oceniać te wyniki.

5.1. Progowanie obrazów barwnych

5.1. Progowanie obrazów barwnych

Progowanie obrazów barwnych jest uogólnieniem na 3 kanały zagadnienia progowania obrazów monochromatycznych nazywanego binaryzacją i opisywanego w literaturze, np. [150, 280]. Progowanie w przestrzeni HSI ma długą tradycję i wiele praktycznych zastosowań. Różnią się one doborem składowych do progowania, wymiarowością histogramów i sposobem wyboru progów. Histogramy wyrażają rozkład składowych w całym obrazie barwnym lub w jego wybranym obszarze. Przykładowo w pracy [13] wykorzystano dwuwymiarowe histogramy HS przedstawiające relacje pomiędzy odcieniem i nasyceniem w obrazie oraz jednowymiarowe histogramy obrazów I. W artykule [173] opisano iteracyjne progowanie 3 składowych HSI (w tej kolejności) z progami wyznaczonymi w wyniku analizy histogramów. Inną możliwością jest progowanie obrazu powstałego ze złożenia obrazów H i S [6]. Można ustalić przedziały zmienności składowych RGB lub HSI (tzw. maski chromatyczne, 6 progów) dla barwników stosowanych w preparatach biologicznych i za pomocą progowania wysegmentować odpowiednio zabarwione obiekty [314].

Wartości progów potrzebne do progowania są wyznaczane najczęściej w wyniku analizy histogramów odpowiednich składowych barwy. Do znajdowania progu w jednowymiarowym histogramie można zastosować algorytm Otsu [174] znany ze znajdowania progów dla histogramów poziomów szarości. W przypadku progowania w przestrzeni HSI szczególną rolę odgrywa histogram H. Na rys. 5.3(a) pokazano przypadek obrazu barwnego, który w wyniku automatycznej analizy histogramu obrazu H może zostać posegmentowany na obszary odpowiadające tłu (największy "pik" w histogramie) i 4 zawartym w nim obiektom (mniejsze "piki"). Stąd wynika obecność 5 "pików" (należy pamiętać o kołowej naturze histogramu — początek łączy się z końcem) na histogramie obrazu H. Nie jest problemem umieszczenie progów pomiędzy "pikami". Zamieszczony poniżej histogram obrazu S pokazuje, że piksele achromatyczne i zbliżone do nich ($S \approx 0$) nie odgrywają w obrazie istotnej roli.

Nie zawsze jednak sytuacja wygląda tak prosto. Na rys. 5.3(b) przedstawiono obraz barwny, którego histogram H zawiera dużą ilość wąskich "pików", mimo iż w obrazie występują jedynie 4 obiekty na bardzo ciemnym tle. "Piki" odpowiadają odcieniom, które nie są postrzegane w obrazie. Miejsca, w których występują dodatkowe "piki", to odcienie będące wielokrotnością 30° lub 60°, co można wytłumaczyć trygonometrycznym charakterem wzorów na H. Gdyby tło było idealnie czarne, to jego piksele jako achromatyczne nie byłyby brane pod uwagę w budowie histogramu obrazu H. Z załączonego histogramu obrazu S wynika, że oprócz pikseli achromatycznych w obrazie występuje wiele pikse-







- Rys. 5.3. Wykresy histogramów obrazów H i S dla przykładowych obrazów barwnych:
 (a) obraz wyjściowy, (b) obraz wyjściowy, (c) histogram obrazu H dla (a),
 (d) histogram obrazu H dla (b), (e) histogram obrazu S dla (a), (f) histogram obrazu S dla (b)
- Fig. 5.3. Histograms of H and S images: (a) original image, (b) original image,(c) histogram of H image for (a), (d) histogram of H image for (b), (e) histogram of S image for (a), (f) S histogram of S image for (b)

5.1. Progowanie obrazów barwnych

li do nich zbliżonych ($S \approx 0$). To właśnie te piksele są odpowiedzialne za generowanie dodatkowych "pików" w histogramie obrazu H.

Histogram obrazu H w postaci przedstawionej na rys. 5.3(d) nie nadaje się do analizy w celu wyznaczenia progów. Można dokonać jego korekty poprzez rozszerzenie pojęcia pikseli achromatycznych, nie biorących udziału w konstruowaniu tego histogramu, na piksele do nich zbliżone. W pracy [16] zaproponowano następujący zestaw 9 dodatkowych reguł wiążących ze sobą intensywność i nasycenie dla pikseli achromatycznych w znaczeniu rozszerzonym:

if $I(i,j) \leq 10$ then H(i,j) := niezdefiniowane elseif $I(i,j) \in (10,15)$ and S(i,j) < 0.55 then H(i,j) := niezdefiniowane elseif $I(i,j) \in (15,20)$ and S(i,j) < 0.40 then H(i,j) := niezdefiniowane elseif $I(i,j) \in (20,30)$ and S(i,j) < 0.30 then H(i,j) := niezdefiniowane elseif $I(i,j) \in (30,40)$ and S(i,j) < 0.22 then H(i,j) := niezdefiniowane elseif $I(i,j) \in (40,50)$ and S(i,j) < 0.17 then H(i,j) := niezdefiniowane elseif $I(i,j) \in (50,100)$ and S(i,j) < 0.12 then H(i,j) := niezdefiniowane elseif $I(i,j) \in (100,180)$ and S(i,j) < 0.07 then H(i,j) := niezdefiniowane elseif I(i,j) > 180 and S(i,j) < 0.03 then H(i,j) := niezdefiniowane elseif I(i,j) > 180 and S(i,j) < 0.03 then H(i,j) := niezdefiniowane

(5.1)

gdzie i, j są współrzędnymi piksela w obrazie, a zmienne $I(i,j) \in \langle 0,255 \rangle$, $H(i,j) \in \langle 0^{\circ}, 360^{\circ} \rangle$, $S(i,j) \in \langle 0,1 \rangle$ są składowymi barwy punktu (i,j) w przestrzeni HSI.

W przypadku obrazu barwnego (rys. 5.3(b)) zastosowanie dodatkowych reguł pozwala na wygenerowanie histogramu poddającego się analizie mającej na celu wyznaczanie progów (rys. 5.4(a)). W wyniku takiego progowania otrzymujemy obraz przedstawiony na rys. 5.4(b). Z przeprowadzonych doświadczeń wynika, że histogram obrazu H wymaga wtedy korekty, gdy w obrazie występują obszary postrzegane jako achromatyczne lub chromatyczne o bardzo słabym nasyceniu.

5.1.1. Zastosowanie wiedzy o tle obrazu

Autor sprawdził również możliwości wykorzystania wiedzy o barwie tła w procesie segmentacji obrazów barwnych w przestrzeni HSI, której zalety i wady opisano w [189]. Badana metoda segmentacji pozwala oddzielić obszary odpowiadające obiektom od tła, a następnie przydzielić tym obszarom etykiety i wyznaczyć ich średnią barwę. Wiedza o barwie tła stanowi podstawę wyboru tych składowych HSI, które będą progowane oraz służy do wyznaczenia odpowiednich progów. Nie dysponując wcześniej informacją o barwie tła, można założyć, że zostanie ona pozyskana z pierwszych "brzegowych" pikseli obrazu.



- Rys. 5.4. (a) Skorygowany histogram obrazu H dla rys. 5.3(b), (b) pseudokolorowy wynik progowania obrazu z rys. 5.3(b) po korekcie histogramu obrazu H
- Fig. 5.4. (a) Corrected histogram of H image for Fig. 5.3b, (b) pseudocolourized result of thresholding image Fig. 5.3b after with histogram correction

Przeprowadzono eksperymenty [187, 188] dla rzeczywistych obrazów barwnych scen z różnymi barwami tła. Metodyka postępowania jest tutaj "odwrotna": segmentacja nie koncentruje się na obiektach sceny, lecz na jej tle. Nakłada to ograniczenie na klasę scen, których obrazy są w ten sposób segmentowane: tło powinno być jednorodne lub teksturowe, a obraz kontrastowy. Odnośnie sceny przyjęto, że jest statyczna, równomiernie oświetlona, zawiera obiekty chromatyczne o powierzchniach odbijających dyfuzyjnie (bez odblasków), które się nie przesłaniają. Można rozpatrzeć 5 następujących przypadków tła:

1. Tło achromatyczne, jednorodne, ciemne (rys. 5.5) [212]

W przypadku takiego tła wystarczy progować obraz *I*. Próg wybierano automatycznie w dolinie za pierwszym od lewej "pikiem" histogramu intensywnościowego, odpowiadającym tłu, np. w punkcie I = 70. W ten sposób oddzielono obiekty od tła (rys. 5.5(b)).

2. Tło achromatyczne jednorodne jasne (rys. 5.6) [212]

Gdy tło było białe lub jasne, to stosowano progowanie obrazu S. Nasycenie pikseli tła było bliskie zera, co pozwoliło znaleźć próg na histogramie obrazu S (S = 15%). Histogram obrazu S i wyniki segmentacji przedstawiono na rys. 5.6(b) i rys. 5.6(c). Cienie występujące na powierzchni tła nie są problemem, ponieważ są achromatyczne.





5182 N





- Rys. 5.5. Progowanie obrazu przedstawiającego obiekty chromatyczne na jednorodnym ciemnym tle achromatycznym: (a) obraz wyjściowy, (b) obraz po progowaniu, (c) histogram obrazu I
- Fig. 5.5. Thresholding image presented chromatic objects on the achromatic dark uniform background: (a) original image, (b) image after thresholding, (c) histogram of I image

3. Tło achromatyczne, teksturowe (rys. 5.7) [191]

Achromatyczność tła powoduje, że z powodzeniem można stosować progowanie obrazu S (rys. 5.7(c)). Największy "pik" z niskim nasyceniem odpowiada tłu i dlatego lokując próg w pierwszej dolinie za nim np. punkt S = 22% i otrzymamy wynik progowania przedstawiony na rys. 5.7(b). Otrzymany obraz wymaga dalszego przetwarzania w postaci filtracji odszumiającej.

4. Tło chromatyczne, jednorodne (rys. 5.8) [212]

W przypadku tła chromatycznego przyjęto założenie, że scena nie zawiera obiektów w odcieniu tła. Stosowano progowanie przedziałowe obrazu H, tzn. potrzebna była znajomość zakresu odcienia, w którym zawierają się wszystkie piksele tła. Zakres ten powinien być wzięty z pewnym zapasem, ponieważ cienie rzucane przez obiekty na tło sceny rozszerzają nieznacznie zakres odcieni tła na histogramie obrazu H. Dla badanego fioletowego tła przyjęto zakres $H_f = \langle 240^\circ, 310^\circ \rangle$ (rys. 5.8(c)).



- Rys. 5.6. Progowanie obrazu przedstawiającego obiekty chromatyczne na jednorodnym ciemnym tle achromatycznym: (a) obraz wyjściowy, (b) obraz po progowaniu, (c) histogram obrazu S
- Fig. 5.6. Thresholding image presented chromatic objects on the achromatic dark uniform background: (a) original image, (b) image after thresholding, (c) histogram of S image
- 5. Tło chromatyczne, teksturowe (rys. 5.9) [191]

Można sformułować tutaj podobne uwagi jak w p
kt. 4. Zakres odcienia dla turkusowoniebieskiego tła obrazu wynos
i $H_{t-n} = \langle 130^{\circ}, 240^{\circ} \rangle.$

Podsumowując należy stwierdzić, że w przypadku tła achromatycznego można stosować progowanie obrazu I lub S, a w przypadku tła chromatycznego jest przydatne progowanie przedziałowe obrazu H. Niektóre obrazy po progowaniu (rys. 5.7(b), rys. 5.9(b)) wymagają przetwarzania końcowego np. w postaci filtracji morfologicznej [171], które usunie z nich szum.

Tak więc omawiana procedura segmentacji składała się z następujących etapów:

- wyznaczania progów i progowania wybranych składowych HSI,
- przetwarzania końcowego (filtracja morfologiczna otrzymanego obrazu binarnego opcjonalnie),



ROZDZIAŁ 5. Segmentacja obrazów barwnych



- Rys. 5.7. Progowanie obrazu przedstawiającego obiekty chromatyczne na achromatycznym tle teksturowym: (a) obraz wyjściowy, (b) obraz po progowaniu, (c) histogram obrazu ${\cal S}$
- Fig. 5.7. Thresholding image presented chromatic objects on the achromatic textured background: (a) original image, (b) image after thresholding, (c) histogram of S image
- etykietowania obszarów w obrazie binarnym (rys. 5.10(b)) przy zastosowaniu klasycznego algorytmu [279, 280],
- wyznaczania współrzędnych centroidów obszarów w obrazie binarnym,
- wyznaczania średniej barwy obszarów na podstawie barw pikseli w sąsiedztwie centroidów.

Średnia barwa obszaru wraz z dodatkowymi cechami kształtowymi (rozdz. 5.6.) mogą zostać wykorzystane do rozpoznawania jednobarwnych obiektów w scenie [190].

Oprócz stosowania jednowymiarowych histogramów składowych HSI przydatne mogą być histogramy dwuwymiarowe, szczególnie obrazów H-S i S-I. Na rys. 5.11 pokazano obraz sceny złożonej z 2 obiektów: żółtego i niebieskiego. Rzuty klasterów reprezentujących te obiekty w przestrzeni HSI na płaszczyznę H-S (rys. 5.11(b)) są rozłączne, co zdecydowanie ułatwia prawidłową segmentację takiego obrazu. Niewrażliwość odcienia i nasycenia na zmiany natężenia oświetlenia sceny i orientacji powierzchni obiektów zwiększa praktyczną przydatność "histogramu" obrazów H-S [211]. Należy zauważyć, 5.1. Progowanie obrazów barwnych



- Rys. 5.8. Progowanie obrazu z obiektami chromatycznymi na jednorodnym tle chromatycznym: (a) obraz wyjściowy, (b) wynik progowania, (c) histogram obrazu H
- Fig. 5.8. Thresholding image presented chromatic objects on the chromatic uniform background: (a) original image, (b) image after thresholding, (c) histogram of H image

że prezentowane poniżej wykresy kołowe na płaszczyźnie H-Sz formalnego punktu widzenia nie są histogramami.

5.1.2. Detekcja odblasków

Na powierzchni niektórych rzeczywistych obiektów trójwymiarowych występują zjawiska optyczne zakłócające przetwarzanie obrazów takich obiektów. Te zjawiska to: cienie (ang. *shadows*), cienie własne (ang. *shading*), odblaski wywołane przez odbicie zwierciadlane od ich powierzchni (ang. *highlights*) czy interrefleksy (ang. *intereflections*) wywołane przez światło odbite wcześniej od sąsiednich obiektów. Dalsze rozważania zostaną ograniczone do odblasków, zwanych niekiedy refleksami świetlnymi. Są one zjawiskiem tak typowym dla pewnych klas obiektów, że w programach graficznych dodaje się je do obrazów jako środek urealniający wygląd obrazu. Odblaski niosą ze sobą informację, która może mieć znaczenie praktyczne. Odblaski występujące w obrazach monochromatycznych mogą być wykorzystane do wyznaczania rozmiaru i położenia [283] oraz orientacji [11] obiektów.

5.1. Progowanie obrazów barwnych



- Rys. 5.11. Wykresy kołowe na płaszczyźnie H-S: (a) obraz wyjściowy, (b) rzut klastera obiektu żółtego na płaszczyznę H-S, (c) rzut klastera obiektu niebieskiego na płaszczyznę H-S
- Fig. 5.11. H S circular diagrams: (a) original image, (b) projection of cluster of yellow object on H S plane, (c) projection of cluster of blue object on H S plane

Odblaski są jednak często źródłem problemów w systemach wizyjnych. Przeszkadzają w segmentacji obrazu, bowiem są traktowane jako oddzielne obszary o dużej jasności. W związku z tym często zachodzi potrzeba detekcji oraz eliminacji odblasków. Można to zrobić wprowadzając zmiany w systemie wizyjnym np. specjalne oświetlacze dyfuzyjne [146] lub filtry polaryzacyjne [167]. Inna grupa metod to eliminacja odblasków z obrazów przy wykorzystaniu wiedzy o odbiciu światła, niekiedy sformalizowanej w postaci tzw. modeli odbiciowych (ang. *reflection models*). Sposób odbicia światła zależy od mikrostruktury powierzchni obiektu: powierzchnia gładka odbija w jednym kierunku - odbicie kierunkowe (zwierciadlane), a powierzchnia chropowata rozprasza światło w różnych kierunkach – odbicie rozproszone (dyfuzyjne, lambertowskie). Intensywność światła odbitego w sposób rozproszony nie zależy od kąta widzenia oka (kamery). Większość powierzchni obiektów rzeczywistych to powierzchnie hybrydowe, tzn. odbicie od nich jest sumą odbicia kierunkowego i rozproszonego. Ponieważ odblaski są to obszary powstałe w wyniku odbicia kierunkowego, to próby modelowania tego odbicia są najbardziej interesujące.

Jedną z pierwszych prób modelowania odbicia kierunkowego był stosowany w grafice komputerowej model Phonga [229, 307], który używa funkcji trygonometrycznej z parametrem o wartości zależnej od rodzaju odbicia. Współczesna optyka używa dokładnych modeli odbiciowych (np. model Beckmanna – Spizzichino oraz model Torrance'a – Sparrowa), jednak o ograniczonej przydatności dla detekcji odblasków w obrazach.

Jednym z modeli wykazujących największą przydatność zarówno do detekcji jak i eliminacji składowej kierunkowej odbicia jest zaproponowany przez Stevena Shafera Dwuchromatyczny Model Odbiciowy DRM (ang. *Dichromatic Reflection Model*) [255]. W modelu





- Rys. 5.9. Progowanie obrazu przedstawiającego obiekty chromatyczne na chromatycznym tle teksturowym: (a) obraz wyjściowy, (b) obraz po progowaniu, (c) histogram obrazu ${\cal H}$
- Fig. 5.9. Thresholding image presented chromatic objects on the chromatic textured background: (a) original image, (b) image after thresholding, (c) histogram of H image





- Rys. 5.10. Progowanie obrazu przedstawiającego 23 obiekty chromatyczne na achromatycznym ciemnym tle: (a) obraz wyjściowy, (b) obraz po progowaniu i etykietowaniu
- Fig. 5.10. Thresholding image presented 23 chromatic objects on the achromatic dark background: (a) original image, (b) image after thresholding and labeling

Wykryte w obrazie barwnym odblaski mogą być użyte do estymacji barwy źródła światła [137], a nawet zliczania obiektów w scenie [194].

ROZDZIAŁ 5. Segmentacja obrazów barwnych

DRM światło o luminancji *L* odbite przez powierzchnię optycznie niejednorodną, np. z tworzywa sztucznego, można przedstawić jako liniową kombinację światła rozproszonego i światła odbitego kierunkowo:

$$L(\lambda, n, s, v) = L_b(\lambda, n, s, v) + L_s(\lambda, n, s, v)$$
(5.2)

We wzorze L oznacza luminancję, λ – długość fali, n, s, v – kąty: normalnej do powierzchni, kierunku oświetlenia i kierunku widzenia, b – składową odbicia rozproszonego (ang. *body reflection*), s – składową odbicia kierunkowego (ang. *surface reflection*). Model DRM pozwala również na separację w każdym składniku części widmowej c od części geometrycznej m:

$$L(\lambda, n, s, v) = m_b(n, s, v)c_b(\lambda) + m_s(n, s, v)c_s(\lambda)$$
(5.3)

Klaster w przestrzeni *RGB*, odpowiadający pojedynczemu obiektowi z odblaskiem, jest płaski i składa się z 2 wydłużonych ("linearnych") klasterów, z których jeden odpowiada barwie obiektu, a drugi - barwie odblasku. Kierunki i rozmiary obu klasterów są wykorzystywane do detekcji odblasku. Na bazie tego modelu można otrzymać dwa obrazy: obraz bez odblasków i obraz z samymi odblaskami. Przy budowie modelu DRM przyjęto założenie, że źródło światła jest punktowe, chociaż rys. 5.12(d), (f) pokazuje, że model ten działa też dla liniowego źródła światła (świetlówki).

Rysunek 5.12 przedstawia obrazy otrzymane w wyniku implementacji modelu DRM. Rysunek 5.12(a) pokazuje obraz wygenerowany za pomocą programu graficznego, a rys. 5.12(c) obraz pozyskany kamerą w układzie oświetleniowym złożonym z 2 świetlówek, stąd zawarte w nim 2 odblaski liniowe. Wykryte odblaski (rys. 5.12(c), rys. 5.12(d)) są zaznaczone za pomocą czarnej barwy. Detekcję odblasków w obrazie rys. 5.12(a) ograniczono do największego odblasku w centrum obiektu.

W pracach autora [195,196] pokazano, że w pewnych przypadkach, gdy rozwiązywany problem pozwala ograniczyć się do detekcji odblasków bez ich usuwania, można zrezygnować ze stosowania modelu DRM. Eksperymentalnie można sprawdzić, że pole powierzchni i kształt odblasku powstającego na gładkiej powierzchni obiektu zależy od orientacji obiektu i kształtu źródła światła. Badano również histogramy I oraz histogramy H - S dla obszarów z odblaskami. Wraz z pojawieniem się odblasku na powierzchni obiektu wzrasta jego intensywność, a chromatyczność zmienia się w kierunku barwy światła. Gdy chodzi o źródło białego światła, oznacza to zmniejszanie się nasycenia S do zera. Tak więc odbla-



(b)





Rys. 5.12. Detekcja i eliminacja odblasków za pomocą modelu DRM: (a), (b) obrazy barwne z odblaskami, (c), (d) detekcja odblasków, (e), (f) obrazy barwne po eliminacji odblasków

Fig. 5.12. Highlight detection and elimination with the help of DRM model: (a), (b) colour image with highlights, (c), (d) highlight detection, (e), (f) colour images after highlight elimination

5.1. Progowanie obrazów barwnych

ROZDZIAŁ 5. Segmentacja obrazów barwnych

ski można detekować jako obszary o wysokim I i jednocześnie niskim S. Ostatnio coraz częściej stosuje się histogramy dwuwymiarowe S - I do detekcji odblasków [286].

Odblaski mogą służyć jako źródło informacji o obecności obiektów pewnej klasy w scenie. Rozpatrzmy scenę zawierającą pewną liczbę obiektów o gładkich powierzchniach zorientowanych względem układu oświetleniowego i kamery w sposób umożliwiający powstanie na nich odblasków. Rysunek 5.13 przedstawia 3 obrazy takich scen. Obrazy zostały pozyskane przez system wizyjny opisany w [184] i nie były poddawane przetwarzaniu wstępnemu. W pracy [194] zaproponowano zliczanie obiektów w takich scenach na podstawie zliczania odblasków w pojedynczym obrazie barwnym. Proponowana metoda wykorzystuje progowanie w przestrzeni *HSI* i składa się z następujących etapów:

- transformacja $RGB \rightarrow HSI$,
- detekcja odblasków z obrazu barwnego przez progowanie obrazów I oraz S,
- konsolidacja wydzielonych odblasków,
- zliczanie obiektów w obrazie binarnym.

Sceny przedstawione na rys. 5.13(a) i rys. 5.13(d) są oświetlone z prawej strony, natomiast scena przedstawiona na rys. 5.13(g) została oświetlona z obu stron. Dlatego jej tło jest jaśniejsze i na powierzchni obiektów występuje podwójna liczba odblasków. Linearny kształt odblasków jest wynikiem linearnego kształtu źródeł światła (świetlówki). Warunek dla progowania obrazu z odblaskami, które prowadzi do otrzymania obrazu binarnego f(i, j), można zapisać w następujący sposób:

for dla każdego piksela
$$(i, j)$$

if $I(i, j) \in \langle I_{min}, 255 \rangle$ and $S(i, j) \in \langle 0, S_{max} \rangle$ then
 $f(i, j) := 1$
else
 $f(i, j) := 0$
end
end

Zaletą progowania jest prostota i szybkość. Dla obrazów przedstawionych na rys. 5.13 wyznaczono następujące progi: dla intensywności $I_{min} = 150$ oraz dla nasycenia $S_{max} = 0,3$. W wyniku progowania piksele z wartościami I oraz S zawartymi w przedziałach wyznaczonych przez progi są białe w obrazach binarnych (rys. 5.13(b), (e), (h)). Wysegmentowane w ten sposób obszary o wysokich intensywnościach i niskich nasyceniach reprezentują odblaski. Niektóre obszary odpowiadające odblaskom są niespójne. Dlatego zastosowano trzykrotną dylację z elementem strukturującym o rozmiarach 3×3 , który zapewnia



Rys. 5.13. Detekcja odblasków w obrazach barwnych: (a), (d), (g) obrazy z odblaskami, (b), (e), (h) detekcja odblasków poprzez progowanie, (c), (f), (i) obrazy binarne po konsolidacji odblasków

Fig. 5.13. Highlight detection in colour images: (a), (d), (g) image with highlights, (b), (e),(h) highlight detection with the help of thresholding, (c), (f), (i) binary images after highlight consolidation

5.2. Klasteryzacja w przestrzeni barw

ROZDZIAŁ 5. Segmentacja obrazów barwnych

konsolidację obszarów odblasków. Kształt i rozmiar elementu strukturującego powinien odpowiadać kształtowi odblasków. Wyniki konsolidacji przedstawiono na rys. 5.13(c), (f), (i). Zliczanie obiektów w obrazie binarnym nie jest złożonym zadaniem. Jest to faktycznie zadanie etykietowania obszarów w obrazie binarnym, którego algorytmy są opisywane w wielu podręcznikach [279, 280].

Metodę sprawdzono przy zastosowaniu źródła światła białego (świetlówki). Pokazano, że nie wymaga ona precyzyjnej orientacji obiektów, ponieważ zliczane są odblaski a nie ich powierzchnia. Metoda może znaleźć zastosowanie w wizyjnej kontroli jakości (np. kontrola kompletności). Znajomość właściwości przestrzeni HSI pozwoliła autorowi zaproponować prostą i szybką metodę detekcji odblasków, której wersje dla różnych przestrzeni typu HSx są obecnie prezentowane w literaturze, np. [286].

5.2. Klasteryzacja w przestrzeni barw

Jednym z silnie rozwijanych podejść do segmentacji obrazów barwnych jest klasteryzacja (grupowanie danych) pikseli, przedstawianych jako punkty w trójwymiarowej przestrzeni barw, na podstawie podobieństwa ich barwy. Klasteryzacja jest często traktowana jako nienadzorowana klasyfikacja setek tysięcy lub milionów pikseli. W procesie tym nie jest wykorzystywana wiedza a priori o obrazie. Barwy dominujące w obrazie w sposób naturalny tworzą klastery w przestrzeni barw. Rysunek 5.14 pokazuje odpowiadające klasterom trzy "chmury pikseli" w przestrzeni *RGB*.





- Rys. 5.14. Obraz KLOCKII oraz klastery jego pikseli w przestrzeni barw *RGB*: (a) obraz wyjściowy, (b) "chmury" pikseli obrazu
- Fig. 5.14. KLOCK11 image and clusters of its pixels formed in RGB space: (a) original image, (b) "clouds" of pixels

Wiele różnych technik klasteryzacji zaproponowanych w literaturze z zakresu rozpoznawania obrazów [119] może być wykorzystanych do segmentacji obrazów barwnych. Jedną z bardziej popularnych i najszybszych technik klasteryzacyjnych jest technika kmeans.

Technikę k-means zaproponowano w latach 60. [147] i została ona opisana w wielu podręcznikach rozpoznawania obrazów, np. [4]. Pierwszy etap wymaga określenia liczby klasterów k oraz wyboru ich początkowych środków C_i , co jest istotnym ograniczeniem tej metody:

$$C_1, C_2, \dots, C_k$$
 gdzie $C_i = [R_i, G_i, B_i], i = 1, 2, \dots, k$ (5.5)

Podczas klasteryzacji każdy punkt obrazu x jest przyporządkowany do klastera K_j , do którego środka odległość punktu jest najmniejsza. Miarą odległości może być metryka Euklidesa, metryka L1, metryka Mahalanobisa, uwzględniająca korelację między składowymi [270] itp. Warunek przynależności punktu do klastera K_j w czasie *n*-tej iteracji można zapisać następująco:

$$x \in K_j(n) \iff \forall i = 1, 2, \dots, j - 1, j + 1, \dots, k \quad ||x - C_{j(n)}|| < ||x - C_{i(n)}||,$$
(5.6)

gdzie C_j jest środkiem klastera K_j .

Zasadniczą ideą techniki k-means jest modyfikacja położeń środków klasterów przeprowadzana tak długo, aż suma odległości pomiędzy wszystkimi punktami klasterów a ich środkami będzie minimalna. Dla klastera K_j minimalizowany wskaźnik J przyjmuje postać:

$$J_j = \sum_{x \in K_j(n)} \|x - C_j(n+1)\|^2$$
(5.7)

Po alokacji punktów obliczane są składowe barwy nowych środków klasterów. Wychodząc ze wzoru (5.7) można zapisać składowe barwy środka klastera K_j po (n + 1) iteracji jako średnie arytmetyczne składowych barwy punktów należących do tego klastera:

$$C_{jR}(n+1) = \frac{1}{N_j(n)} \sum_{x \in K_j(n)} x_R$$
(5.8)

$$C_{jG}(n+1) = \frac{1}{N_j(n)} \sum_{x \in K_j(n)} x_G$$
(5.9)

$$C_{jB}(n+1) = \frac{1}{N_j(n)} \sum_{x \in K_j(n)} x_B$$
(5.10)

5.2. Klasteryzacja w przestrzeni barw

gdzie $N_j(n)$ – liczba pikseli należących do klastera K_j w *n*-tej iteracji.

Ponieważ uśrednianie według wzorów (5.8)–(5.10) odbywa się dla wszystkich k klasterów, to metodę można nazywać metodą k-średnich. W następnym etapie sprawdzana jest różnica pomiędzy położeniem nowych i starych środków klasterów. Jeżeli ta różnica jest powyżej pewnego progu δ , to w ramach następnej iteracji obliczane są odległości punktów od nowych środków klasterów, zmieniana jest przynależność punktów itd. Jeżeli ta różnica jest poniżej pewnego progu δ , to proces klasteryzacji zostaje zatrzymany.

$$\forall i = 1, 2, \dots, k \quad \left| C_{i(n+1)} - C_{i(n)} \right| < \delta$$
 (5.11)

Innym możliwym kryterium stopu jest wykonanie wcześniej założonej liczby iteracji. W ramach ostatniego etapu działania techniki k-means barwa każdego piksela zostaje zastąpiona barwą odpowiadającą barwie środka klastera, do którego należy dany piksel. Tak więc liczba barw w obrazie posegmentowanym zostaje zredukowana do k barw. Algorytm k-means jest zbieżny, ale jedynie do minimum lokalnego [254].

Na czas klasteryzacji techniką k-means mają wpływ: wybrana na początku klasteryzacji liczba klasterów k oraz rozdzielczość obrazu. Im większe są te parametry, tym dłuższa jest klasteryzacja. W tab. 5.1 pokazano tę zależność w przypadku wykorzystania komputera osobistego z procesorem pracującym z częstotliwością 1,5 GHz i pamięcią operacyjną 752 MB.

Tabela 5.1. Czas klasteryzacji (w sekundach) techniką k-meansa liczba klasterów ki rozdzielczość obrazu

Nazwa obrazu	k=2	k = 5	k = 10	k = 15	k=20	k = 25
Baboon (512×512)	1	1	2	3	3	4
PEPPERS (512×512)	1	1	2	3	3	4
Pills (1236×1236)	8	12	19	27	34	41

Wyniki segmentacji metodą k-means istotnie zależą od położenia początkowych środków klasterów. Te dane wejściowe w wersji półautomatycznej mogą być definiowane przez operatora programu. W wersji automatycznej techniki k-means mogą być one wybierane losowo spośród barw występujących w obrazie. Istnieją również inne możliwości: może to być k pierwszych barw w obrazie lub k szarości z równomiernie podzielonej przekątnej sześcianu. Na rys. 5.15 pokazano dla obrazu OBIEKTY wyniki uzyskane w poszczególnych iteracjach zarówno w dziedzinie obrazu, jak i w dziedzinie przestrzeni barw RGB. Założono istnienie 8 klasterów i położenie początkowych środków klasterów na osi szarości.











(c)



- Rys. 5.15. Kolejne iteracje w procesie segmentacji obrazu OBIEKTY metodą k-means dla k = 8: (a) obraz oryginalny, (b) początkowy rozkład środków klasterów, (c) obraz po 1 iteracji, (d) rozkład środków klasterów po 1 iteracji, (e) obraz po 2 iteracji, (f) rozkład środków klasterów po 2 iteracji
- Fig. 5.15. Iterations in k-means-based segmentation process (OBIEKTY image, k = 8): (a) original image, (b) distribution of initial cluster centers, (c) image after 1st iteration, (d) distribution of cluster centers after 1st iteration, (e) image after 2nd iteration, (f) distribution of cluster centers after 2nd iteration

ROZDZIAŁ 5. Segmentacja obrazów barwnych



- Rys. 5.16. Kolejne iteracje w procesie segmentacji obrazu OBIEKTY metodą k-means dla k = 8: (a) obraz po 3 iteracji, (b) rozkład klasterów po 3 iteracji, (c) obraz po 5 iteracji, (d) rozkład klasterów po 5 iteracji, (e) obraz po 7 iteracji, (f) rozkład klasterów po 7 iteracji
- Fig. 5.16. Iterations in k-means-based segmentation process (OBIEKTY image, k = 8): (a) image after 3rd iteration, (b) distribution of cluster centers after 3rd iteration, (c) image after 5th iteration, (d) distribution of cluster centers after 5th iteration, (e) image after 7th iteration, (f) distribution of cluster centers after 7th iteration

5.2. Klasteryzacja w przestrzeni barw

Przedstawione wyniki segmentacji obrazów techniką k-means otrzymano w przestrzeni RGB. Można dokonać transformacji obrazu z przestrzeni RGB do innej przestrzeni barw i klasteryzacji pikseli w tej przestrzeni. Podobnie istnieje możliwość zwiększenia wymiaru przestrzeni cech w procesie klasteryzacji poprzez wprowadzenie dla każdego piksela dodatkowych parametrów, takich jak: współrzędne geometryczne pikseli w obrazie, gradient barwy, tekstura itd. [79, 256]. Ponieważ algorytmy klasteryzacji zawsze mogą znaleźć w zbiorze pikseli jakieś klastery, to potrzebna jest ich ocena (rozdz. 5.5.).

Liczba iteracji w procedurze k-means zależy od przyjętego wcześniej kryterium stopu, tj. różnicy pomiędzy położeniem nowych i starych środków klasterów. Im mniejsza jest wartość tej różnicy, tym więcej iteracji jest wykonywanych. Liczba iteracji oraz wynik segmentacji zależą też od położenia początkowych środków klasterów. Ilustruje to rys. 5.17. Obrazy na rys. 5.17(b) i rys. 5.17(c) otrzymano stosując losowy wybór początków 20 klasterów. Program zatrzymał się odpowiednio po 26 (rys. 5.17(b)) i 18 (rys. 5.17(c)) iteracjach. Różnica pomiędzy obydwoma posegmentowanymi obrazami jest wyraźnie widoczna w barwnych obrazach otrzymanych po segmentacji: w pierwszym posegmentowanym obrazie występują 2 obszary żółte, których nie ma w drugim obrazie.



Rys. 5.17. Segmentacja k-means k = 20: (a) KWIAT1, (b), (c) po segmentacji

Fig. 5.17. K-means segmentation, k = 20: (a) KWIAT1, (b), (c) after segmentation

W posegmentowanym obrazie piksele przyporządkowane do jednego klastera należą na ogół do wielu różnych obszarów. Im większa jest liczba klasterów k, tym więcej obszarów w obrazie po segmentacji, tym bardziej szczegółowa jest segmentacja. Techniki klasteryzacyjne ze swej natury nie uwzględniają położenia poszczególnych pikseli (sąsiedztwo) w obrazie, czego skutkiem jest ich wrażliwość na szum. Stąd wynika potrzeba przetwarzania końcowego (rozdz. 5.4.), które likwiduje nadsegmentację obrazu. W niniejszym podrozdziale istotnym elementem było pokazanie na tle klasycznej techniki klasteryzacji znaczenia wyboru początkowych środków klasterów i różnych możliwości w tym zakresie (wybór

ROZDZIAŁ 5. Segmentacja obrazów barwnych

losowy z obrazu, pierwsze barwy w obrazie, poziomy szarości z przekątnej sześcianu RGBitp.).

5.3. Rozrost obszaru w obrazie barwnym

W połowie lat 70. S.L. Horowitz i T. Pavlidis [110] zaproponowali formalną definicję segmentacji obszarowej. Segmentacja obszarowa obrazu może być zdefiniowana jako podział obrazu I na obszary R_i , i = 1, ..., N takie, że przy użyciu pewnego predykatu logicznego H, służącego do sprawdzania jednorodności tych obszarów, są spełnione następujące warunki:

$$I = \bigcup_{i=1}^{N} R_i \tag{5.12}$$

 $\forall i \neq j \ R_i \cap R_j = \emptyset \tag{5.13}$

 $\forall i \ H\left(R_i\right) = \text{prawda} \tag{5.14}$

 $\forall i \neq j \ H (R_i \cap R_j) = \text{falsz gdzie } R_i \ i \ R_j \ \text{przylegają do siebie}$ (5.15)

Powyższa formalna definicja stanowi, że każdy piksel należy do jakiegoś obszaru R_i (warunek kompletności (5.12)). Obszary wyznaczone przez technikę segmentacji są rozłączne (warunek rozłączności (5.13)). Dwa pozostałe warunki są warunkami jednorodności. Warunek (5.14) może stanowić kryterium stopu dla procesu podziału, tzn. podział trwa tak długo, jak długo istnieją obszary niejednorodne ze względu na kryterium wyrażone przez predykat H. Podobnie warunek (5.15) jest kryterium stopu dla procesu łączenia, tzn. nie mogą pozostać w obrazie wynikowym po segmentacji obszary przylegające, które po połączeniu spełniałyby kryterium jednorodności.

Techniki obszarowe (ang. *region-based techniques*) polegają na grupowaniu pikseli w obszary jednorodne pod względem barwy. Można wśród nich wyróżnić: rozrost obszaru (ang. *region growing*), podział obszaru (ang. *region splitting*), łączenie obszarów (ang. *region merging*) i inne. Szczególnie technika rozrostu obszaru, zaproponowana po raz pierwszy dość dawno [30] dla obrazów z poziomami szarości, cieszy się ciągle powodzeniem w przetwarzaniu obrazów barwnych.

Rozrost obszaru jest typową techniką "z dołu do góry" (ang. *bottom-up*). Poszczególne sąsiadujące piksele są łączone ze sobą w obszary, jeżeli ich atrybuty (tutaj: barwa) są wystarczająco podobne. Podobieństwo to wyraża predykat, nazywany niekiedy kryterium jednorodności. Jeżeli dla danego piksela kryterium jednorodności jest spełnione, to piksel zostaje dołączony do obszaru i jednocześnie atrybuty obszaru (średnia barwa, po-

5.3. Rozrost obszaru w obrazie barwnym

le powierzchni itp.) zostają zaktualizowane. W klasycznej wersji rozrostu obszaru proces rozrostu zaczyna się od wybranych pikseli-ziaren (ang. *seeds*) i jest kontynuowany do czasu przyporządkowania wszystkich pikseli do obszarów. Techniki rozrostu obszaru różnią się pomiędzy sobą postaciami kryterium jednorodności i sposobem osadzania ziaren. Ich generalną zaletą jest równoczesne uwzględnianie podobieństwa barw pikseli i ich bliskości w obrazie.

Jeżeli barwa w obrazie jest opisywana za pomocą składowych *RGB*, to kryterium jednorodności wyrażające różnicę barw za pomocą metryki L2 ma następującą postać:

$$\sqrt{(R-R^*)^2 + (G-G^*)^2 + (B-B^*)^2} \le d,$$
(5.16)

gdzie: R, G, B – składowe barwy sprawdzanego piksela, R^*, G^*, B^* – składowe średniej barwy powstającego obszaru, a d to parametr, który silnie wpływa na wynik segmentacji obrazu. Kryterium jednorodności w RGB można wyrazić też za pomocą koniunkcji kilku nierówności:

$$R^{*} - T_{1}^{R} \leqslant R \leqslant R^{*} - T_{2}^{R} \quad \wedge \quad G^{*} - T_{1}^{G} \leqslant G \leqslant G^{*} - T_{2}^{G} \quad \wedge \quad B^{*} - T_{1}^{B} \leqslant B \leqslant B^{*} - T_{2}^{B},$$
(5.17)

gdzie T_1^R, \ldots, T_2^B – progi.

Proces rozrostu może być zorganizowany również w innej przestrzeni barw, na przykład w walcowej przestrzeni HSI, HSV lub HLS [193]. Ten rodzaj przestrzeni lepiej oddaje percepcję barw przez człowieka niż przestrzeń RGB. W tym przypadku wzór analogiczny do wzoru (5.16) ma postać nieco bardziej skomplikowaną:

$$\sqrt{(I-I^*)^2 + S^2 + S^{*2} - 2SS^2 \cos\left(H - H^*\right)} \leqslant d,$$
(5.18)

gdzie: H, S, I – odcień, nasycenie i intensywność sprawdzanego piksela, H^*, S^*, I^* – składowe średniej barwy powstającego obszaru. Inne wyrażenia, którymi można wyrazić kryterium jednorodności w tej przestrzeni i ich wpływ na wynik segmentacji, przedstawiono w pracy [192]. Jeżeli wyrażenie nie uwzględnia składowej I, to obraz po segmentacji może nie zawierać cieni i odblasków [213] (rozdz. 5.3.3.). Kryterium jednorodności może zostać również wyrażone poprzez wariancje składowych barw powstającego obszaru: piksel może zostać dołączony do obszaru, jeżeli nie spowoduje to wzrostu wartości wyrażenia opartego na wariancjach. Takie podejście zastosowano w pracy [63] do przestrzeni HSV:

$$\gamma = \frac{A_1}{\sigma_H^2} + \frac{A_2}{\sigma_S^2} + \frac{A_3}{\sigma_V^2},$$
(5.19)

5.3. Rozrost obszaru w obrazie barwnym

ROZDZIAŁ 5. Segmentacja obrazów barwnych

gdzie: A_1 , A_2 , A_3 – stałe współczynniki, σ_H^2 , σ_S^2 , σ_V^2 – wariancje odcienia, nasycenia i wartości w obszarze po dołączeniu piksela-kandydata. Piksel jest dołączany do obszaru, gdy wartość γ po tej operacji wzrośnie.

Niekiedy [86] zwraca się uwagę na to, że odległość euklidesowa powinna wyrażać rzeczywistą różnicę barw pomiędzy pikselem-kandydatem a średnią barwą powstającego obszaru. Wymaga to wykorzystania równomiernej percepcyjnie przestrzeni barw [193], której najpopularniejszym przykładem jest przestrzeń *CIELAB*. Wtedy kryterium jednorodności ma następującą postać:

$$\sqrt{(L-L^*)^2 + (a-a^*)^2 + (b-b^*)^2} \le d,$$
(5.20)

gdzie: L, a, b – odcień, nasycenie i intensywność sprawdzanego piksela, L^* , a^* , b^* – składowe średniej barwy powstającego obszaru.

Jeżeli sprawdzany piksel spełnia kryterium jednorodności, to średnia barwa obszaru może zostać zaktualizowana za pomocą odpowiedniego wzoru rekurencyjnego; poniżej wersja wzoru dla intensywności:

$$I_n^* = \frac{(n-1)I_{n-1}^* + I_{ij}}{n},\tag{5.21}$$

gdzie: I_{ij} – intensywność piksela o współrzędnych ij, I_{n-1}^* – średnia intensywność obszaru z (n-1) pikselami, I_n^* – średnia intensywność obszaru z n pikselami po dołączeniu sprawdzanego piksela. Podobne wzory rekurencyjne można wyprowadzić dla innych charakterystyk obszaru, np. wariancji:

$$\sigma_n^2 = \frac{(n-1)(\sigma_{n-1}^2 + I_{n-1}^{*2}) + I_{ij}^2}{n} - I_n^{*2}, \qquad (5.22)$$

gdzie: σ_{n-1}^2 – wariancja intensywności w obszarze z n-1 pikselami, σ_n^2 – wariancja intensywności w obszarze z n pikselami, po dołączeniu piksela-kandydata. Należy jednak pamiętać, że w czasie procesu segmentacji opisane powyżej statystyki obszaru nie są znane dokładnie, bo nie są znane jeszcze wszystkie piksele, które wejdą w skład obszaru.

W procesie rozrostu obszaru może być stosowana idea 4-spójności lub 8-spójności. Piksel spełniający warunki 4-spójności lub 8-spójności zostaje włączony do powstajacego obszaru, gdy jego barwa spełnia kryterium jednorodności.

5.3.1. Technika ziarnistego rozrostu obszaru

Ziarnisty rozrost obszaru (ang. *seeded region growing, SRG*) jest jedną z klasycznych technik segmentacji. W latach 90. zostały szczegółowo opisane 2 wersje algorytmu SRG dla obrazów z poziomami szarości [1,160]. Ziarna są najczęściej wybierane w obszarach reprezentujących obiekty zainteresowania. Aby zapewnić pełny podział obrazu na N obszarów, musimy wybrać ziarna w każdym obszarze i proces rozrostu obszaru musi zostać przeprowadzony N razy. Dobrym przykładem jest rys. 5.18, który przedstawia napis złożony z 6 liter (8 oddzielnych obszarów) [206].



Rys. 5.18. Ziarnisty rozrost obszaru dla obrazu LITERY: (a) obraz oryginalny z zaznaczonymi ziarnami, (b) obraz posegmentowany -d = 50, (c) obraz posegmentowany -d = 10, (d) obraz posegmentowany -d = 90

Fig. 5.18. Seeded region growing for LITERY colour image: (a) original image with the seeds, (b) segmented image -d = 50, (c) segmented image -d = 10, (d) segmented image -d = 90

Pełna segmentacja obrazu LITERY wymaga więc użycia 8 ziaren i doboru parametru d. Rozmieszczenie ziaren pokazuje rys. 5.18(a), a wynik ziarnistego rozrostu obszaru w przestrzeni RGB z wartościami składowych barwy z zakresu $\langle 0, 255 \rangle$ uzyskany dla d = 50przedstawiono na rys. 5.18(b). Istnieje pewien zakres wartości parametru d, przy których następuje prawidłowa segmentacja. Stosowanie parametru spoza zakresu kończy się niepełnym wysegmentowaniem liter (rys. 5.18(c)) lub powiększeniem liter o piksele tła (rys. 5.18(d)). Podczas procesu segmentacji obszary rozrastają się wokół ziaren. Tło w posegmentowanym obrazie zostało usunięte w celu uproszczenia obrazu. Technika SRG z ręcznym lokowaniem ziaren znajduje zastosowanie w sytuacji, gdy chcemy wysegmentować

5.3. Rozrost obszaru w obrazie barwnym

ROZDZIAŁ 5. Segmentacja obrazów barwnych

z obrazu niezbyt dużą liczbę obiektów, w których osadzamy ziarna [160].

Wyniki segmentacji omawianą techniką zależą jednak od położenia ziaren. Na rys. 5.19 przedstawiono obrazy posegmentowane z wykorzystaniem ziaren umieszczonych w 2 różnych miejscach. Rysunek 5.19(d) dowodzi, że położenie ziarna nie jest obojętne dla wyniku segmentacji. Również użycie 2 ziaren rozmieszczonych na jednym obiekcie (niebieski kwiat) powoduje powstanie 2 obszarów zamiast jednego (rys. 5.19(e)). Dopiero przetwarzanie końcowe może połączyć je w jeden obszar. Badano również wpływ wielkości ziarna (od 1 piksela do 20×20 pikseli) na wyniki segmentacji. Jest on zdecydowanie mniejszy niż wpływ położenia ziaren. Obraz różnicowy pomiędzy wynikami dla małych i dużych ziaren zawierał jedynie kilkadziesiąt pojedynczych pikseli przy rozdzielczości obrazów 640 × 480 pikseli. Wpływ kształtu ziaren był praktycznie żaden.



Rys. 5.19. Ziarnisty rozrost obszaru w obrazie KwiAT2, d = 150: (a) obraz oryginalny z zaznaczonym ziarnem w 2 różnych położeniach, (b) segmentacja od ziarna w środku kwiatu, (c) segmentacja od ziarna na płatku kwiatu, (d) różnica obrazów (b) i (c), (e) segmentacja od 2 ziaren rozmieszczonych tak jak w obrazie (a)

Fig. 5.19. Seeded region growing for image KWIAT2, d = 150: (a) original image with 2 seeds, (b) segmentation from the seed in the middle, (c) segmentation from the seed on the petal, (d) difference of images (b) and (c), (e) segmentation from 2 seeds located as in image (a)

Ziarna mogą być osadzane w obrazie ręcznie lub wybierane losowo. W procesie ręcznego wyboru ziaren operator wykorzystuje swoją wiedzę o obrazie do ich osadzania. Losowy wybór ziaren jest szczególnie ryzykowny, gdy obraz jest zaszumiony, ponieważ ziarno może być ulokowane w nietypowym pikselu, będącym wynikiem szumu. Ziarna można wybierać korzystając z wierzchołków histogramu barwy. Również informacja krawędziowa może być wykorzystana do lokowania ziaren wewnątrz zamkniętych konturów. Sinclair [260] zaproponował umiejscowienie ziaren w wierzchołkach obrazu Woronoja, co wymagało wcześniejszego znalezienia krawędzi w obrazie wyjściowym. Obraz Woronoja to obraz z poziomami szarości powstały na podstawie binarnego obrazu krawędziowego.

Poziom szarości piksela w obrazie Woronoja jest funkcją odległości piksela od najbliższej krawędzi. Dlatego za celowe uznano lokowanie ziaren w najjaśniejszych punktach obrazu Woronoja (rys. 5.20).



Rys. 5.20. Obraz Woronoja powstały na podstawie obrazu barwnego CHART: (a) obraz wyjściowy, (b) obraz Woronoja

Fig. 5.20. Voronoi image build on the base of CHART colour image: (a) original image, (b) Voronoi image

Ouerhani i in. [175] wykorzystali tzw. punkty uwagi pochodzące z modelu uwagi wzrokowej (ang. visual attention) jako naturalnych kandydatów do roli ziaren. W pracy Ikonomakisa [117] ziarna dla obszarów chromatycznych wyznaczano sprawdzając przy użyciu symetrycznych masek wariancję odcienia; gdy ona nie przekraczała progu, to w masce lokowano ziarno, co odpowiadało środkowi obszaru. W konkretnych zastosowaniach do lokowania ziaren wykorzystuje się wiedzę a priori o obrazach. Na przykład gdy robot mobilny wyposażony w system wizyjny ma wykrywać drzwi w pomieszczeniach wewnętrznych, to ziarna są lokowane automatycznie w pewnej odległości od narożników drzwi, znajdowanych za pomocą metody Harrisa [37].


Rys. 5.22. Segmentacja obrazu KWIAT2, zmiana parametru d: (a) obraz oryginalny, (b) obraz posegmentowany z d = 40 (2194 obszary), (c) obraz posegmentowany z d = 70 (413 obszarów), (d) obraz posegmentowany z d = 180 (3 obszary)

Fig. 5.22. Segmentation of image KWIAT2 with different values d: (a) original, (b) image segmented with d = 40 (2194 regions), (c) image segmented with d = 70 (413 regions), (d) image segmented with d = 180 (3 regions)

Przedstawiana tutaj technika jest techniką sekwencyjną, a więc zależy od punktu startu i kolejności przetwarzania pikseli. Dlatego wykazuje ona wrażliwość ze względu na kierunek procesu skanowania obrazu. Dla kilku obrazów przeprowadzono dodatkowo segmentację w odwrotnym kierunku, tzn. od dolnego lewego piksela obrazu do górnego prawego piksela. Uzyskane liczby obszarów nieznacznie różnią się od odpowiednich wartości uzyskanych dla obrazów segmentowanych w kierunku prostym. Porównano obrazy wynikowe i dla każdej pary wyznaczono obraz różnicowy. Rysunek 5.23 przedstawia wyniki takiego eksperymentu dla obrazu KWIAT2. Negatyw obrazu różnicowego (rys. 5.23(c)) nie jest obrazem białym, a więc pokazuje, że proponowana metoda segmentacji jest zależna od kierunku, w którym następuje łączenie pikseli w obszary. Podobne wyniki uzyskano dla standardowych obrazów PEPPERS i PARROTS w pracy [205]. W literaturze [261] zaproponowano również wiele innych, niekiedy nielinearnych, porządków skanowania obrazu. Uniezależnienie wyniku segmentacji od kolejności łączenia pikseli wymaga skomplikowania i zrównoleglenia algorytmu rozrostu. Ideę takiego podejścia dla obrazów z poziomami

5.3.2. Technika nieziarnistego rozrostu obszaru

Poniżej przedstawiono i sprawdzono doświadczalnie propozycję autora w zakresie technik nieziarnistego rozrostu obszaru. Technika ta bywa nazywana też techniką rozrostu obszaru poprzez agregowanie pikseli. Polega ona na łączeniu podobnych pikseli w obszary w czasie skanowania obrazów bez użycia ziaren, typowych dla klasycznego rozrostu obszaru. W proponowanej wersji na początku każdy piksel ma swoją własną etykietę (obszary jednopikselowe). Po każdym włączeniu piksela do obszaru średnia barwa obszaru jest aktualizowana poprzez zastosowanie odpowiedniego wzoru rekurencyjnego. Proces rozrostu, który zaczyna się od lewego górnego piksela i dąży w kierunku prawego dolnego piksela, może być uzupełniony przez dodatkowy proces w kierunku odwrotnym [214]. Taki "podwójny" proces rozrostu prowadzi do obrazów o zmniejszonej liczbie obszarów [205]. Implementacja algorytmu z wykorzystaniem struktur listowych pozwala dokonywać stosunkowo szybkiej segmentacji: komputer osobisty z procesorem pracującym z częstotliwością 3,2 GHz segmentuje obraz o rozdzielczości 512 × 512 w ciągu 0,1 s.

ROZDZIAŁ 5. Segmentacja obrazów barwnych

W przypadku 4-spójności podczas dwóch przebiegów sprawdzanych jest 4 sąsiadów piksela. W przypadku 8-spójności podczas każdego z dwóch przebiegów są sprawdzane dodatkowo dwa sąsiednie piksele (rys. 5.21). Sprawdzanie dwukrotnie większej liczby pikseli wydłuża czas procesu rozrostu. Wynikiem stosowania omawianej techniki segmentacji jest zbiór obszarów scharakteryzowanych przez ich średnią barwę, rozmiar oraz listę pikseli należących do obszaru. Obszary w tym procesie są generowane sekwencyjnie.



Rys. 5.21. Maski stosowane w procesie segmentacji: (a), (b) przypadek 4-spójności, (c), (d) przypadek 8-spójności

Fig. 5.21. Masks used in segmentation process: (a), (b) masks for 4-connectivity, (c), (d) masks for 8-connectivity

W czasie testowania nieziarnistego rozrostu obszaru zaobserwowano, że jeżeli wartość parametru d rośnie, to równocześnie maleje w obrazie posegmentowanym liczba obszarów R. Wyniki dla obrazu KWIAT2 przedstawiono na rys. 5.22. Niska wartość parametru d może być przyczyną nadsegmentacji, a wysoka wartość przyczyną niedosegmentacji.

ROZDZIAŁ 5. Segmentacja obrazów barwnych

szarości pokazano jako algorytm ISRG w pracy [160].



- Rys. 5.23. Segmentacja obrazu KWIAT2, zmiana kierunku skanowania, parametr d = 70: (a) obraz posegmentowany ze skanowaniem prostym (413 obszarów), (b) obraz posegmentowany ze skanowaniem odwrotnym (411 obszarów), (c) negatyw różnicy obrazów (a) i (b)
- Fig. 5.23. Segmentation results on KWIAT2 colour image different scanning directions, parameter d = 70: (a) segmentation based on the direct scanning direction (413 regions), (b) segmentation based on the inverse scanning direction (411 regions), (c) negative image of difference image between (a) and (b)

Badano również wpływ koncepcji spójności (4-spójność, 8-spójność) na wyniki segmentacji. Wprowadzenie 8-spójności, jak należało oczekiwać, zmniejsza liczbę obszarów w posegmentowanym obrazie o kilka lub kilkanaście procent, co powoduje zwiększenie średniego błędu barwy. Stosowanie 8-spójności wydłuża też czas procesu segmentacji o 20%-30% w stosunku do segmentacji opartej na 4-spójności. Dlatego nieraz ze względu na szybkość wybierana jest 4-spójność.

Techniki obszarowe są ze swej natury sekwencyjne i dlatego dla wyniku duże znaczenie ma kolejność, zgodnie z którą przetwarzane są piksele i obszary. Łatwo zauważyć, że technika ziarnistego rozrostu obszaru bardziej nadaje się do wysegmentowania małej liczby obiektów z obrazu, a technika nieziarnistego rozrostu do "pełnej" segmentacji obrazu. Technika ziarnistego rozrostu ma charakter półautomatyczny, gdy ziarna umieszcza w obrazie operator programu, a technika nieziarnistego rozrostu jest techniką automatyczną.

Informacja o krawędziach obszarów uzyskana np. za pomocą gradientu może wspomóc sterowanie procesem rozrostu obszaru. Takie podejście zaproponowano dla obrazów z poziomami szarości w pracy [108]. Powstało wiele technik hybrydowych, gdzie informacja obszarowa i krawędziowa uzupełniają się prowadząc do segmentacji obrazu [72]. Idea rozrostu obszaru znalazła również zastosowanie w segmentacji wododziałowej użytej po raz pierwszy dla obrazów barwnych przez Meyera [161]. Niekiedy rozrost obszaru jest traktowany jako segmentacja początkowa przed głównym procesem segmentacji [225].

5.3. Rozrost obszaru w obrazie barwnym

Nieziarnisty rozrost obszaru stosowany do łączenia małych obszarów (ang. region merging) może wchodzić w skład złożonej techniki podziału i łączenia (ang. split and merge) [110]. Technika ta wykorzystuje hierarchiczną i rekurencyjną strukturę danych, tzw. drzewo czwórkowe (ang. quad tree). Korzeń drzewa odpowiada całemu obrazowi, liście to pojedyncze piksele, a pozostałe wierzchołki drzewa (węzły) reprezentują kwadratowe bloki obrazu. Etapom podziału i łączenia odpowiada poruszanie się po drzewie odpowiednio w dół i w góre. Najpierw cały obraz traktowany jest jako jeden obszar i jest oceniana jego jednorodność przy użyciu wybranego kryterium jednorodności. W przypadku niejednorodności obszar-obraz jest dzielony na 4 jednakowe podobszary. Jednorodność każdego z nich podlega ocenie za pomocą kryterium. Jeżeli podobszar okazuje się jednorodny, to nie podlega on dalszemu podziałowi. Podobszary niejednorodne dalej dzielone są na ćwiartki. Wynikiem takiego podziału obrazu jest oczywiście nadsegmentacja obrazu. Można by kontrolować ten proces podziału tak, aby nie dopuścić do nadsegmentacji. Jednak zamiast kontroli procesu podziału stosuje się drugi etap segmentacji, tj. łączenie obszarów przylegających. Przed połączeniem 2 przylegających obszarów sprawdzana jest jednorodność obszaru, który z takiego połaczenia powstanie. Proces ten jest procesem nieziarnistego rozrostu obszaru, ale w tym przypadku łączone są bloki o różnych kształtach. Etap łączenia jest kontynuowany dopóty, dopóki są możliwe nowe połączenia obszarów.

Miarą jednorodności obszarów w obrazach monochromatycznych może być różnica pomiędzy maksymalną i minimalną intensywnością (poziomem szarości) w obszarze lub wariancja jego intensywności. Jeżeli wartość tej miary jest mniejsza od pewnego zadanego progu, to obszar jest uznawany za jednorodny. Oprócz takich prostych miar statystycznych niekiedy używa się bardziej skomplikowanych kryteriów np. przed połączeniem przylegających obszarów porównuje się za pomocą testu nieparametrycznego Kołmogorowa-Smirnowa podobieństwo ich rozkładów poziomów szarości. Dla techniki podziału i łączenia należy mówić o dwóch kryteriach jednorodności, bo można stosować inne kryterium dla etapu podziału, a inne dla etapu łączenia. Parametry występujące w kryteriach mają duży wpływ na ilość i rozmiary obszarów w obrazie wynikowym.

Wady "wrodzone" techniki podziału i łączenia to: ograniczenie obszaru zastosowań do obrazów kwadratowych i to o wymiarze boku będącym całkowitą potęgą liczby 2 oraz nadmiernie prostokątny (blokowy) kształt powstających obszarów, który zniekształca rzeczywiste krawędzie obiektów i jest źródłem błędów. Klasyczne drzewo czwórkowe, wygodne w procesie podziału, nie ułatwia znajdowania obszarów przylegających do siebie w procesie łączenia. Łączenie utrudnia fakt, że obszary przylegające do siebie w obrazie

5.3. Rozrost obszaru w obrazie barwnym

ROZDZIAŁ 5. Segmentacja obrazów barwnych

mogą leżeć na różnych poziomach drzewa czwórkowego.

W pracy [197] opisano badania właściwości techniki podziału i łączenia dla obrazów barwnych. Na etapie podziału obrazu zastosowano w niej jako kryterium jednorodności obszaru sumę wariancji składowych barwnych np. dla *RGB*:

$$\sigma_R^2 + \sigma_G^2 + \sigma_B^2 \leqslant d_1, \tag{5.23}$$

gdzie:

$$\sigma_R^2 = \frac{1}{N-1} \sum_{i=1}^N \left(R_i - \bar{R} \right)^2, \quad \sigma_G^2 = \frac{1}{N-1} \sum_{i=1}^N \left(G_i - \bar{G} \right)^2, \quad \sigma_B^2 = \frac{1}{N-1} \sum_{i=1}^N \left(B_i - \bar{B} \right)^2, \quad (5.24)$$

a \overline{R} , \overline{G} , \overline{B} to składowe średniej barwy w obszarze, N – liczba pikseli w obszarze. Suma wariancji składowych barwy to, jak nietrudno zauważyć, ślad macierzy kowariancji, proponowany w literaturze jako kryterium jednorodności obszaru [88]. Analogicznie można zdefiniować sumę wariancji składowych barwnych HSI. Na etapie łączenia stosowano natomiast odległość euklidesową pomiędzy średnimi barwami w łączonych obszarach (wersja dla RGB – odpowiednik wzoru (5.16)):

$$\sqrt{(\bar{R}_1 - \bar{R}_2)^2 + (\bar{G}_1 - \bar{G}_2)^2 + (\bar{B}_1 - \bar{B}_2)^2} \le d_2$$
(5.25)

W przypadku przestrzeni HSI uwzględniano walcowość jej składowych (odpowiednik wzoru (5.26)):

$$\sqrt{(\bar{I}_1 - \bar{I}_2)^2 + \bar{S}_1^2 + \bar{S}_2^2} - 2\bar{S}_1\bar{S}_2\cos(\bar{H}_1 - \bar{H}_2) \le d_3$$
(5.26)

Poniżej pokazano działanie techniki podziału i łączenia na obrazie sztucznym KULKA (turkusowa kula oświetlona białym światłem). Obrazy sztuczne, tj. pochodzące z systemów grafiki komputerowej, poprzez swoją prostotę szczególnie dobrze nadają się do testowania algorytmów przetwarzania obrazów. Wyniki segmentacji przedstawiono w postaci obrazów krawędziowych. Analizując te wyniki można stwierdzić, że w procesie podziału im mniejszy jest próg d_1 w kryterium jednorodności (rys. 5.24), tym większa jest w obrazie liczba kwadratowych obszarów.

Podobnie można przedstawić możliwości sterowania procesem segmentacji za pomocą progu d_2 w etapie łączenia obszarów. Rysunek 5.25 pokazuje, jak w wyniku zwiększania się progu d_2 w tym etapie zmniejsza się liczba wysegmentowanych obszarów, co pozwala zmniejszyć nadsegmentację.

Rysunek 5.26 przedstawia wyniki segmentacji w przestrzeni RGB i HSI. Nieliniowa transformacja $RGB \rightarrow HSI$ wydłuża czas działania algorytmu i dlatego jej stosowanie ma



- Rys. 5.24. Monochromatyczna wersja obrazu KULKA i wyniki etapu podziału w procesie segmentacji w przestrzeni RGB (próg $d_2 = 0$): (a) obraz wyjściowy, (b) $d_1 = 100$, (c) $d_1 = 30$, (d) $d_1 = 10$
- Fig. 5.24. Monochrome version of KULKA image and results of splitting in the segmentation process in *RGB* space (threshold $d_2 = 0$): (a) original image, (b) $d_1 = 100$, (c) $d_1 = 30$, (d) $d_1 = 10$



Rys. 5.25. Monochromatyczna wersja obrazu KULKA i całościowe wyniki etapu podziału w procesie segmentacji w przestrzeni RGB (próg $d_1 = 30$): (a) obraz wyjściowy, (b) $d_2 = 10$, (c) $d_2 = 30$, (d) $d_2 = 100$

Fig. 5.25. Monochrome version of KULKA image and results of segmentation process in RGB space (threshold $d_1 = 30$): (a) original image, (b) $d_2 = 10$, (c) $d_2 = 30$, (d) $d_2 = 100$

sens tylko w przypadku przewagi wyników segmentacji w przestrzeni HSI nad wynikami segmentacji w przestrzeni RGB. Rysunek pokazuje istnienie takiej przewagi: segmentacja HSI przy tych samych progach prowadzi do mniejszej liczby obszarów o mniej blokowym kształcie niż segmentacja w RGB. Ten istotny wynik kolejny raz potwierdza sens rozpatrywania w procesie segmentacji obrazu innych niż RGB przestrzeni barw.

5.3.3. Cienie i odblaski w procesie rozrostu obszaru

Często wymaga się niewrażliwości algorytmu segmentacji na cienie i odblaski występujące w obrazie. Problematyka cieni i odblasków w procesie segmentacji zostanie przedstawiona

5.4. Przetwarzanie końcowe

ROZDZIAŁ 5. Segmentacja obrazów barwnych



- Rys. 5.26. Monochromatyczna wersja obrazu KULKA i wyniki procesu segmentacji: (a) obraz wyjściowy, (b) segmentacja w przestrzeni RGB ($d_1 = 30, d_2 = 30$), (c) segmentacja w przestrzeni HSI ($d_1 = 30, d_2 = 30$)
- Fig. 5.26. Monochrome version of KULKA image and results of segmentation process: (a) original image, (b) segmentation in RGB space $(d_1 = 30, d_2 = 30)$, (c) segmentation in HSI space $(d_1 = 30, d_2 = 30)$

na przykładzie techniki nieziarnistego rozrostu obszarów. Zostaną sprawdzone możliwości tej techniki powstałe w wyniki zmiany jej parametrów z wykorzystaniem wiedzy o właściwościach przestrzeni barw. Cienie występujące w obrazach mogą stanowić na tyle duże obszary, że nie można ich usunąć w ramach przetwarzania końcowego, bo równocześnie zostałyby usunięte inne istotne obszary w obrazie posegmentowanym. Taka sytuacja występuje w obrazie KLOCKI2 (rys. 5.24(a)). Zastosowanie metryki euklidesowej w przestrzeni RGB prowadzi do wysegmentowania obszarów cienia o dużych powierzchniach (rys. 5.24(b)). Dobór odpowiedniej przestrzeni barw oraz metryki stosowanej w kryterium jednorodności wraz z parametrami techniki segmentacji (d, A) pozwala usunąć obszary cienia z posegmentowanego obrazu. Usunięcie składowej luminancyjnej ze wzorów (5.26) i (5.20) pozwala wysegmentować obiekty bez cieni (rys. 5.24(c), (d)). Taka operacja wymaga jednak wcześniejszej nieliniowej transformacji składowych RGB na składowe HSI lub CIELAB. Aby uniknąć czasochłonnej transformacji można przeprowadzić rozrost obszaru w przestrzeni RGB stosując zamiast metryki euklidesowej kąt pomiędzy wektorami barw jako miarę ich podobieństwa [68].

Załóżmy, że średnia barwa powstającego obszaru i barwa piksela, którego przynależność do obszaru rozpatrujemy, są reprezentowane przez następujące wektory:

$$c_1 = [R_1, G_1, B_1], \qquad c_2 = [R_2, G_2, B_2]$$
(5.27)

Kąt pomiędzy wektorami c_1 i c_2 , oznaczony przez θ , można wyrazić za pomocą wzoru:

$$\cos\theta = \frac{c_1^T \circ c_2}{\|c_1\| \|c_2\|},\tag{5.28}$$

korzystając z definicji iloczynu skalarnego wektorów. Im mniejszy kąt θ , tym barwa piksela jest bliższa średniej barwy obszaru. Dlatego jako miara podobieństwa barw zostanie użyta funkcja sinus i na jej bazie można stosować następujące kryterium jednorodności:

$$255 \cdot \sqrt{1 - \left(\frac{c_1^T \circ c_2}{\|c_1\| \|c_2\|}\right)^2} \leqslant d \tag{5.29}$$

Zastosowanie wzoru (5.25) w procesie rozrostu obszaru prowadzi do segmentacji odpornej na cienie, co widać na rys. 5.27(e). Należy jednak pamiętać, że dane kątowe, takie jak kąt pomiędzy wektorami barwy θ , odcień H i inne wymagają m.in. ze względu na swoją okresowość stosowania specjalnych metod analizy statystycznej, tzw. statystyk kierunkowych [152]. Zintegrowanie obszarów cienia z tłem obrazu powoduje wzrost niebieskawego zabarwienia powstałego w obrazie tła (rys. 5.27(c)–(e)).

Również odblaski, wywołane gładkością powierzchni obiektów w scenie i odbiciem zwierciadlanym, mogą stanowić problem w procesie segmentacji (rys. 5.28(a)). Na ogół odblaski tworzą w obrazie oddzielne obszary o barwie zbliżonej do barwy źródła światła np. białej (rys. 5.28(b)). Obszary odblasków można usunąć stosując przetwarzanie końcowe (rys. 5.28(c)). Jest to możliwe, ponieważ rozmiar odblasku jest zdecydowanie mniejszy od rozmiaru obiektu.

5.4. Przetwarzanie końcowe

Bardzo często proces segmentacji kończy się nadsegmentacją, tj. w obrazie po segmentacji występuje bardzo dużo małych obszarów. Przyczyną tego zjawiska może być zaszumienie obrazu poddawanego segmentacji. Nawet, jeżeli obraz jest wysokiej jakości, to na krawędziach obiektów występuje duża ilość małych obszarów. Można je usunąć z obrazu stosując po segmentacji przetwarzanie końcowe (ang. *postprocessing*). Poniżej pokazano na tle metod literaturowych zastosowanie procedury przetwarzania końcowego do nieziarnistego rozrostu obszaru [206], chociaż opisywana procedura może być dołączana do innych technik segmentacji. Doświadczalnie badano wpływ parametru procedury na wynik przetwarzania końcowego.

Jedną z efektywnych metod takiego przetwarzania jest usuwanie małych obszarów z obrazu poprzez dołączanie ich do obszarów sąsiednich najbardziej zbliżonych pod względem barwy. Nie jest to zadanie trudne, ponieważ po segmentacji obszarowej dysponujemy listą obszarów w obrazie, która może być posortowana według pól powierzchni obszarów. Progowa wartość powierzchni małego obszaru A zależy od obrazu. Dla jednego obrazu ten

ROZDZIAŁ 5. Segmentacja obrazów barwnych



- Rys. 5.27. Segmentacja obrazu KLOCKI2: (a) oryginał, (b) segmentacja z metryką euklidesową w RGB (d = 45, A = 710), (c) segmentacja z metryką euklidesową dla HS z przestrzeni HSI (d = 50, A = 1000), (d) segmentacja z metryką euklidesową dla a^*b^* z przestrzeni CIELAB (d = 12, A = 1800), (e) segmentacja z metryką "kąt między wektorami" w RGB (d = 21, A = 2000)
- Fig. 5.27. Segmentation of image KLOCK12: (a) original, (b) segmentation with Euclidean metric in RGB (d = 45, A = 710), (c) segmentation with Euclidean metric for HS components from HSI (d = 50, A = 1000), (d) segmentation with Euclidean metric for a^*b^* from CIELAB (d = 12, A = 1800), (e) segmentation with "angle between vectors" metric in RGB (d = 21, A = 2000)

sam próg pozwala usuwać niepotrzebne artefakty, np. odblaski czy szumy, a dla innego obrazu usuwane są wtedy potrzebne szczegóły. Po połączeniu obszarów zostaje obliczona średnia barwa nowego obszaru o większej powierzchni i zostają zmienione etykiety pikseli świadczące o przynależności do obszaru. W wyniku takiego przetwarzania końcowego znacznie maleje liczba obszarów w obrazie posegmentowanym.

Na rys. 5.29(b) pokazano powiększony fragment posegmentowanego obrazu. Zastosowanie pseudokolorów pozwala łatwo zauważyć małe obszary występujące na krawędziach obiektów, w naszym przypadku na brzegu płatka kwiatu. Na rys. 5.29(b) jest takich małych obszarów 26. Zastosowanie przetwarzania końcowego, polegającego na usuwaniu obszarów o powierzchni mniejszej niż 100 pikseli, zmniejsza liczbę wszystkich obszarów w badanym fragmencie do 4, w tym 1 mały obszar leżący wzdłuż krawędzi (rys. 5.29(c)).

5.4. Przetwarzanie końcowe



- Rys. 5.28. Segmentacja obrazu CANDIES: (a) oryginał, (b) obraz posegmentowany przy użyciu metryki euklidesowej w RGB z progiem d = 60, (c) obraz posegmentowany przy użyciu metryki euklidesowej w RGB z parametrami d = 60 i A = 5000
- Fig. 5.28. Segmentation of image CANDIES: (a) original, (b) segmentation with Euclidean metric in RGB and parameter d = 60, (c) segmentation with Euclidean metric in RGB and parameters d = 60, A = 5000

Rozszerzenie definicji małego obszaru do 1000 pikseli powoduje, że znikają wszystkie małe obszary i zostaje tylko obszar płatka oraz dwa obszary liści (rys. 5.29(d)).

Poniżej zamieszczono wyniki przetwarzania końcowego dla całego badanego obrazu KWIAT2 (rys. 5.30). Można zauważyć, że procedura przetwarzania końcowego może służyć nie tylko do usuwania małych obszarów, ale po ustawieniu dużej wartości parametru A także do łączenia przylegających obszarów. Na rys. 5.30(d), gdy przyjęto wartość A = 10000, można zobaczyć efekt takiego połączenia obszarów - jednolite tło obrazu.

Programy implementujące techniki obszarowe można wykorzystać również do segmentacji obrazów monochromatycznych, które są traktowane jako obrazy barwne z pikselami, dla których R = G = B. Takie podejście zastosowano z powodzeniem do używanych w badaniach medycznych tzw. obrazów kometkowych (ang. *comet assay images*) [268]. Na rys. 5.31 pokazano wynik segmentacji obrazu kometkowego techniką nieziarnistego rozrostu obszaru dla parametrów d = 25, A = 2400.

Łączenie małych obszarów z dużymi obszarami sąsiednimi o podobnej barwie nie musi się opierać na zdefiniowaniu progowej wartości pola dla małego obszaru. Można przyjąć założenie, że np. 90% wszystkich pikseli obrazu stanowią obszary istotne, posortować obszary w obrazie według rozmiaru, wybrać duże obszary zawierające 90% pikseli, a obszary zajmujące pozostałe 10% powierzchni obrazu traktować jako małe [140]. Yagi i in. zaproponowali [303] porównywanie barw obszarów kandydujących do połączenia za pomoca odległości w przestrzeni HSV i zmiennej wartości progu, zależnej odwrotnie pro-



- Rys. 5.29. Przetwarzanie końcowe posegmentowanego obrazu KWIAT2 (parametr d = 70): (a) obraz posegmentowany z zaznaczonym oknem, (b) powiększony obraz krawędziowy dla fragmentu obrazu z okna przed przetwarzaniem końcowym, (c) wynik przetwarzania końcowego z parametrem A = 100, (d) wynik przetwarzania końcowego z parametrem A = 1000
- Fig. 5.29. Postprocessing of segmented image KWIAT2 (d = 70): (a) segmented image with selected window, (b) zoomed and pseudocolorized part of the image before postprocessing,(c) postprocessing with parameter A = 100, (d) postprocessing with parameter A = 1000

porcjonalnie od pola powierzchni obszaru. Obok małej powierzchni obszaru dodatkowym wskaźnikiem dla dołączenia danego obszaru może być niska wariancja barwy (wyrażona np. przez małą wartość śladu macierzy kowariancji) lub położenie obszaru przy granicy obrazu.

Podobieństwo barwy sąsiednich obszarów można badać systematycznie, niezależnie od ich pola. Miarą podobieństwa obszarów może być różnica ich średnich barw wyznaczana za pomocą wzorów (5.16)–(5.20) lub bardziej złożone obliczeniowo cechy, takie jak np.:

- histogramy barw obszarów [276] z wyrażeniem podobieństwa barw poprzez przecięcie histogramów (ang. histogram intersection technique) [121],
- wyrażenie zależne m.in. od średniego gradientu barwy obliczonego dla pikseli należących do tych obszarów [169],

5.4. Przetwarzanie końcowe



Rys. 5.30. Przetwarzanie końcowe posegmentowanego obrazu KWIAT2 (parametr d = 70): (a) obraz posegmentowany przed przetwarzaniem końcowym (413 obszarów), (b) wynik przetwarzania końcowego z parametrem A = 100 (47 obszary), (c) wynik przetwarzania końcowego z parametrem A = 1000 (12 obszarów), (d) wynik przetwarzania końcowego z parametrem A = 10000 (3 obszary)

- Fig. 5.30. Results of postprocessing step on segmented image KWIAT2 (parameter d = 70): (a) segmented image before postprocessing (413 regions), (b) result of postprocessing – parameter A = 100 (47 regions), (c) result of postprocessing – parameter A = 1000 (12 regions), (d) result of postprocessing – parameter A = 10000(3 regions)
- odległość Fishera pomiędzy przylegającymi obszarami dla jednej składowej barwy [251,313]:

$$FD_{12} = \frac{\sqrt{(n_1 + n_2)|\hat{\mu}_1 - \hat{\mu}_2|}}{\sqrt{n_1 {\sigma_1}^2 + n_2 {\sigma_2}^2}}$$
(5.30)

gdzie: n_1 , n_2 , $\hat{\mu_1}$, $\hat{\mu_2}$, $\hat{\sigma_1}^2$, $\hat{\sigma_2}^2$ oznaczają odpowiednio: liczbę pikseli, estymaty średniej wartości i wariancji składowej barwy jednego i drugiego obszaru. Jako ostateczną miarę podobieństwa 2 obszarów barwnych można wziąć maksymalną wartość z wartości obliczonych dla trzech składowych.

Od czasu pierwszych publikacji na temat rozrostu obszaru [30] zwracano uwagę na długość części wspólnej konturów sąsiadujących obszarów. Często dodatkowym warun-

ROZDZIAŁ 5. Segmentacja obrazów barwnych



- Rys. 5.31. Segmentacja obrazu kometkowego: (a) obraz wyjściowy, (b) obraz po segmentacji na 3 obszary (jądro komety, ogon komety, tło)
- Fig. 5.31. Comet assay image segmentation: (a) original image, (b) image after segmentation into 3 regions (nucleus of comet, tail of comet, background)

kiem połączenia sąsiednich obszarów może być przekroczenie pewnej minimalnej wartości przez długość części wspólnej ich konturów [98]:

$$\frac{B_{ij}}{B_i + B_j} > T,\tag{5.31}$$

gdzie:

- B_i, B_j długość konturów obszarów R_i, R_j ,
- B_{ij} długość części wspólnej konturów obszarów R_i , R_j ,
- T próg np. 0,05.

Do ustalenia relacji pomiędzy obszarami i kolejności łączenia obszarów wykorzystuje się metody grafowe [87,149] na czele z grafem przylegania obszarów RAG (ang. region adjacency graph) [287], zdefiniowanym w wielu podręcznikach [225]. Węzły w grafie RAG reprezentują obszary, a łuki grafu łączą dwa węzły, jeżeli odpowiadające im obszary przylegają do siebie. Dla każdego łuku można określić koszt połączenia obszarów. Łączenie obszarów z użyciem RAG polega na usuwaniu łuków o najniższym koszcie i łączeniu odpowiadających im węzłów.

Kolejność łączenia obszarów ma wpływ na końcowy wynik segmentacji [223]. Metodyczne podejście do działań na liście obszarów z zastosowaniem metryki (5.13) jako różnicy barw pomiędzy dwoma obszarami opisano w pracy [41]. Ostatecznym celem łączenia obszarów jest otrzymanie obrazu zawierającego możliwie największe obszary jednorodne.

5.5. Ilościowa ocena wyników segmentacji

5.5. Ilościowa ocena wyników segmentacji

Chociaż istnieje wiele technik segmentacji obrazów barwnych, to jednak dużo mniej osiągnięto w dziedzinie oceny wyników procesu segmentacji. Brak jest zarówno powszechnie akceptowanych kryteriów, jak i jednolitych procedur oceny wyników segmentacji. Wiele algorytmów przedstawianych w literaturze jest testowanych na bardzo małej liczbie obrazów.

Najprostszym rodzajem oceny wyniku segmentacyjnego jest subiektywna ocena dokonywana przez eksperta lub zespół ekspertów. Niektórzy uważają nawet, że człowiek jest najlepszym sędzią w tej sprawie [177]. Jednak w niektórych zastosowaniach segmentacji, takich jak np. rozpoznawanie obiektów "weryfikatorem" konkretnego algorytmu segmentacji, jest osiągnięty współczynnik prawidłowych rozpoznań (ang. *recognition rate*) niezależnie od opinii eksperta o obrazie posegmentowanym. Ilościowe metody oceny wyników segmentacji w klasycznej pracy Zhanga [311] zostały podzielone na analityczne i eksperymentalne. Ponieważ jednak nie istnieje ogólna teoria segmentacji obrazu, to metody analityczne są słabo rozwinięte.

Wśród metod eksperymentalnych dominują dwa podejścia: "dobrociowe" (ang. empirical goodness) oraz "rozbieżnościowe" (ang. empirical discrepancy). Pierwsze podejście nie wymaga obrazu referencyjnego (wzorcowo posegmentowanego), a ocena wyników segmentacji następuje w odniesieniu do obrazu oryginalnego. Przykładem miary dobroci segmentacji może być w przypadku segmentacji obszarowej np. jednorodność obszarów lub kontrast międzyobszarowy. W przypadku podejścia "rozbieżnościowego" wyznacza się miarę rozbieżności pomiędzy obrazem posegmentowanym przez badany algorytm a obrazem referencyjnym, który może być obrazem syntetycznym lub obrazem ręcznie posegmentowanym przez eksperta. Oczywiście otwarte pozostaje pytanie, kiedy można a kiedy nie można "zbudować" obraz referencyjny. Miarą rozbieżności może być liczba źle posegmentowanych pikseli, położenie źle posegmentowanych pikseli itp.

W dziedzinie klasteryzacji danych wytworzono własne miary do oceny jakości wyników. Prostą miarą jest odległość wewnątrzklasterowa (ang. *intra-cluster distance*), która może być liczona jako suma odległości punktów klastera od jego środka i świadczyć o zwartości klastera. Miarą komplementarną do odległości wewnątrzklasterowej może być odległość międzyklasterowa (ang. *inter-cluster distance*), która może być liczona jako suma odległości wszystkich środków klasterów od siebie i świadczyć o wzajemnej separacji klasterów. Obie proponowane miary są ujemnie skorelowane.

5.5. Ilościowa ocena wyników segmentacji

R. Turi [290] zaproponowała miarę prawidłowości klasterów VM (ang. Validity Measure) zdefiniowaną jako stosunek dwóch, wspomnianych powyżej, miar charakteryzujących klastery.

$$VM = y(k) \frac{intra}{inter}$$
 (5.32)
 $y(k) = cN_{2,1}(k) + 1,$ (5.33)

gdzie: *intra* – miara zwartości klasterów, *inter* – miara odległości między klasterami, y(k) – ciągła funkcja malejąca wraz ze wzrostem liczby klasterów k, oparta na funkcji Gaussa. Stała c jest liczbą naturalną z zakresu (15, 25). Dla liczby klasterów większej niż 4 funkcja y(k) ma wartość równą 1 i nie ma wpływu na miarę VM.

Celem segmentacji przez klasteryzację jest minimalizacja miary VM, czyli minimalizacja odległości pomiędzy elementami klasterów i ich środkami, zapewniająca "zwartość" klasterów oraz maksymalizacja odległości między klasterami, zapewniająca dobrą separację klasterów. Do wzoru (5.32) podstawiamy wartość intra, obliczoną na podstawie wzoru:

$$intra = \frac{1}{MN} \sum_{i=1}^{k} \sum_{x \in K_i} |x - C_i|^2, \qquad (5.34)$$

gdzie: $M \times N$ – rozdzielczość obrazu w pikselach, k – liczba klasterów, x – piksel obrazu należący do klastera K_i , C_i to barwa środka klastera K_i .

Miarę odległości między środkami klasterów zdefiniowano następująco:

$$i = 1, 2, \dots, k - 1, \quad \forall j = i + 1, \dots, k, \quad inter = \min\left(\|C_i - C_j\|^2\right)$$
(5.35)

Minimalizacja miary VM może być wykorzystana do wyznaczania optymalnej wartości liczby klasterów k.

Empiryczną funkcję oceny segmentacji Q(I) zaproponowali Borsotti i in. w pracy [28] dla klasteryzacyjnych technik segmentacji:

$$Q(I) = \frac{\sqrt{R}}{10000(N \times M)} \cdot \sum_{i=1}^{R} \left[\frac{e_i^2}{1 + \log A_i} + \left(\frac{R(A_i)}{A_i} \right)^2 \right],$$
(5.36)

gdzie: I – obraz po segmentacji, $M \times N$ – rozdzielczość obrazu w pikselach, R – liczba obszarów w posegmentowanym obrazie, A_i – pole obszaru z indeksem i, a e_i – błąd barwy obszaru z indeksem i. Błąd barwy w przestrzeni RGB jest liczony jako suma odległości euklidesowych pomiędzy składowymi barwy pikseli obszaru i składowymi średniej barwy, która jest atrybutem obszaru w posegmentowanym obrazie. Błędy barwy w różnych przestrzeniach barw trudno ze sobą porównywać i dlatego są one transformowane z powrotem do przestrzeni RGB.

Pierwszy czynnik wyrażenia (5.36) jest czynnikiem normalizującym ocenę ze względu na rozmiar obrazu, drugi czynnik "karze" wyniki zawierające zbyt wiele obszarów (nadsegmentacja), a trzeci czynnik "karze" obrazy zawierające obszary niejednorodne. Ponieważ błąd barwy jest większy dla dużych obszarów, to ostatni czynnik jest dzielony przez powierzchnię obszaru. Im mniejsza jest wartość funkcji oceny Q(I), tym lepszy jest wynik segmentacji. Przydatność funkcji Q(I) do oceny nieziarnistego rozrostu obszaru pokazano w pracach autora [200, 221]. Jest to istotny wynik, ponieważ do tej pory stosowano tę metode oceny jedynie do pikselowych technik segmentacji.

Oba opisane wskaźniki oceny VM oraz Q(I) pozwalają dla danej klasy obrazów dobrać wartości parametrów segmentacji umożliwiające uniknięcie zarówno nadsegmentacji, jak i niedosegmentacji. Ze względu na różny charakter wskaźników one nie zawsze jednoznacznie wskazują jeden najlepiej posegmentowany obraz. Pokazano to w pracy [220], gdzie stosowano te wskaźniki do 5 różnych technik klasteryzacji (k-means, mean shift, ISODATA, technika łączenia klasterów oraz technika podziału klasterów).

Funkcję Q(I) wybrano dla porównania ze sobą wyników segmentacji otrzymywanych metodami klasteryzacyjnymi (k-means) i obszarowymi (nieziarnisty rozrost obszaru). Porównanie przeprowadzono dla prezentowanych w tym rozdziale obrazów oprócz 2 obrazów zawierających cienie (LITERY, KLOCKI2) i wymagających stosowania innej metryki. Dla każdego obrazu znaleziono metodą prób i błędów wartości parametrów segmentacji, odpowiednio liczbę klasterów k oraz próg d, które minimalizują Q(I). Tabela 5.2 prezentuje liczbowe wartości parametrów i ocen Q(I). Widać, że dla każdego obrazu technika obszarowa pozwala uzyskiwać mniejszą wartość Q(I), a więc lepszy wynik segmentacyjny w sensie tego kryterium niż metoda k-means. Autorowi nie jest znane z literatury inne ilościowo uzasadnione porównanie technik pikselowych i obszarowych, które podkreśla przewagę tych ostatnich.

Przetwarzanie obrazów znalazło w ostatnich latach zastosowanie w przetwarzaniu dokumentów barwnych. Coraz częściej są one automatycznie skanowane i odczytywana jest ich treść. Programy optycznego rozpoznawania znaków (ang. OCR) na ogół bez problemów odczytują pismo maszynowe. Jednak wielobarwne czeki wypełnione ręcznie barwą atramentu zbliżoną do barwy dokumentu mogą sprawić kłopot programowi rozpoznającemu. Poniżej przedstawiono przykład [198] takiego dokumentu w odcieniach barwy niebieskiej, wypełnionego niebieskim i czarnym atramentem (rys. 5.32(a)). Automatyczny odczyt dokumentu w poziomach szarości jest praktycznie niemożliwy. Histogram intensywności (rys. 5.32(b)) posiada jedno wyraźne minimum (I = 217), łatwe do znalezienia

A

ROZDZIAŁ 5. Segmentacja obrazów barwnych 5.6. Dalsze przetwarzanie obrazu po segmentacji

Tabela 5.2. Parametry i oceny wyników segmentacji obrazów

Nazwa	k	Q_{min}	d	Q _{min}
KWIAT1	24	3953	34	3675
Parrots	19	2420	29	1019
Klocki1	24	188	25	144
OBIEKTY	8	456	81	266
KWIAT2	17	3023	26	2747
CHART	25	4191	7	94
CANDIES	15	12740	40	3799

przez program analizy histogramu. Progowanie względem tego progu pozwala jednak wyodrębnić jedynie część bardziej kontrastowych znaków zapisanych na czeku (rys. 5.32(c)).



Rys. 5.32. Segmentacja wielobarwnego czeku: (a) oryginał, (b) histogram poziomów szarości, (c) wynik progowania obrazu monochromatycznego, (d) wygładzony i poprawiony barwny obraz czeku, (e) ostateczny wynik segmentacji

Fig. 5.32. Segmentation of multicoloured cheque: (a) original, (b) grey-level histogram,
(c) result of grey-level image thresholding, (d) smoothed and enhanced colour image of cheque, (e) final result of segmentation

Z tego powodu do obrazu barwnego zastosowano technikę nieziarnistego rozrostu obszaru z usuwaniem małych obszarów, co zdecydowanie pozwoliło wygładzić i odszumić obraz (rys. 5.32(d)). Następnie przeprowadzono progowanie obrazu H wraz z klasteryzacją w przestrzeni *CIELUV*. Tak przygotowany obraz binarny (rys. 5.32(e)) może stanowić właściwe dane wejściowe dla programu OCR. W tym przykładzie pokazano, że niekiedy można stosować sekwencyjnie 2 techniki segmentacji, z których pierwsza (rozrost obszaru) jest właściwie presegmentacja.

5.6. Dalsze przetwarzanie obrazu po segmentacji

W systemach widzenia komputerowego po przeprowadzeniu procesu segmentacji obrazu dla każdego z powstałych obszarów wyznaczane są jego cechy służące do identyfikacji obiektu przedstawianego przez obszar. Istnieje szeroka literatura oceniająca przydatność poszczególnych cech kształtowych, których opisano do tej pory kilkaset, do rozpoznawania obiektów [107,150,153]. Wśród cech geometrycznych ważną rolę odgrywają współczynniki kształtu ze współczynnikiem kompaktowości na czele [186,279]:

$$K = \frac{4\pi A}{O^2}$$
, (5.37)

gdzie: *O* – obwód obszaru, *A* – pole obszaru. Wyznaczenie pola obszaru jako sumy pikseli należących do obszaru jest zadaniem prostym, większym problemem jest wyznaczenie dokładnej wartości obwodu. Oprócz zwykłego i niezbyt dokładnego wyrażenia obwodu jako liczby pikseli konturowych możliwe są różne aproksymacje obwodu oparte na idei kodu łańcuchowego, np. klasyczna formuła Freemana:

$$B_f = 1,000n_e + 1,414n_o \tag{5.38}$$

lub Dorsta-Smeuldersa:

$$B_d = 0.980n_e + 1.406n_o - 0.091n_c, \tag{5.39}$$

gdzie: n_e – liczba elementów poziomych i pionowych w kodzie łańcuchowym, n_o – liczba elementów skośnych w kodzie łańcuchowym, n_c – liczba przejść od elementu z kodem parzystym do elementu z kodem nieparzystym i odwrotnie. W pracy [186] pokazano, że zastosowanie formuły Dorsta-Smeuldersa do obliczania obwodu pozwala znacznie zwiększyć dokładność wyznaczania współczynnika kompaktowości (błąd względny 5,7%).

Zaletą współczynnika kompaktowości jest jego niezmienniczość w odniesieniu do translacji i rotacji obiektu w scenie i zmiany skali w systemie wizyjnym. Wada tej cechy to

5.6. Dalsze przetwarzanie obrazu po segmentacji

ROZDZIAŁ 5. Segmentacja obrazów barwnych

możliwość uzyskania przez nią jednakowych wartości dla obiektów o różnych kształtach. Podobne właściwości mają inne współczynniki kształtu. Niemniej istotne znaczenie mają również niezmienniki momentowe zaproponowane w początkowym etapie rozwoju przetwarzania obrazów [111]. Wzór na momenty (ang. spatial moments) rzędu p + q binarnej funkcji obrazowej f(i, j) opisującej obraz o rozdzielczości $M \times N$ można zapisać w następującej postaci:

$$m_{pq} = \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} i^p j^q f(i,j), \qquad (5.40)$$

gdzie: $p, q \in \{0, 1, 2, 3\}$. Na podstawie tych momentów można zdefiniować momenty centralne rzędu p + q:

$$\mu_{pq} = \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} \left(i - \bar{i}\right)^p \left(j - \bar{j}\right)^q f(i, j), \tag{5.41}$$

gdzie: $\overline{i} = \frac{m_{10}}{m_{00}}, \ \overline{j} = \frac{m_{01}}{m_{00}}$. Momenty centralne μ_{pq} można poddać normalizacji:

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{pq}^{\gamma}},\tag{5.42}$$

gdzie: $\gamma = \frac{p+q}{2} + 1$ dla $p + q = 2, 3, \dots$

Znormalizowane momenty centralne η_{pq} różnych rzędów wchodzą w skład wyrażeń na 7 niezmienników momentowych Hu. Najprostszy pierwszy niezmiennik ma postać:

$$\Phi_1 = \eta_{20} + \eta_{02} \tag{5.43}$$

Sprawdzono możliwość rozróżniania za pomocą niezmiennika momentowego Phi_1 stosunkowo prostych obiektów przedstawionych na rys. 5.33. Dla każdego typu obiektu pozyskano kilka obrazów. Uzyskane przedziały zmienności niezmiennika momentowego zawarto w tab. 5.3.



Rys. 5.33. Objects of prostych kształtach geometrycznych Fig. 5.33. Objects with simple geometrical shapes Tabela 5.3. Przedziały zmienności dla niezmiennika momentowego Φ_1

Obiekt	A B		С	D	E	
$\Delta \Phi_1$	0,198-0,202	0,238-0,255	0,220-0,235	0,259–0,290	0,190-0,193	

Łatwo zauważyć, że wszystkie przedziały zmienności są rozłączne, co jest podstawowym warunkiem rozpoznawania obiektów.

Wieloletni rozwój problematyki przetwarzania obrazów binarnych zaowocował zastosowaniem do opisu obszarów również trudniejszych do wyznaczenia cech topologicznych, takich jak: liczba otworów obszaru i ich charakterystyki kształtowe, liczba wierzchołków obszaru, wklęsłość/wypukłość obszaru itp. [278]. Cechy topologiczne, ze swej natury niezależne od translacji i rotacji obiektu w scenie i zmiany skali w systemie wizyjnym, mogą być elementami składowymi modeli obiektów. Bardzo ważny jest właściwy wybór metody opisu kształtu obiektów, aby zapewnić pewne i szybkie ich rozpoznawanie. Cechy topologiczne są pojęciowo proste, jednak wyznaczenie ich wartości wymaga stosowania często złożonych algorytmów. Dobrym przykładem jest aproksymacja konturu obszaru za pomocą wielokąta i znajdowanie jego wierzchołków. Idee będące podstawą dwóch algorytmów znajdowania wierzchołków opisano w pracy [278].

Rzadziej w dziedzinie identyfikacji obiektów stosuje się metody sztucznej inteligencji, takie jak reguły IF-THEN [181]. W pracy tej zaproponowano system oparty na 6 regułach IF-THEN, służący do rozpoznawania prostych figur geometrycznych. Baza reguł przedstawiona w pracy ma możliwość dalszego rozwoju poprzez dodawanie następnych reguł.

Obrazy barwne po przeprowadzeniu segmentacji mogą zostać wykorzystane również w procesie pasowania (ang. *matching*) stanowiącym podstawę stereowizyjnej percepcji głębi [183]. Zgrubna segmentacja (ang. *coarse segmentation*) obrazów barwnych może stanowić proces przygotowania obrazowej bazy danych do jej przeszukiwania na podstawie zawartości obrazów (ang. *content-based image retrieval*).

Uniwersalna technika segmentacji obrazu barwnego prawdopodobnie nie istnieje. Tym bardziej, że cele segmentacji wyraźnie zależą od zadania, którego elementem jest proces segmentacji. Rosnące możliwości obliczeniowe sprzętu sprawiają, że rosną szanse zastosowań metod bardziej złożonych niż prezentowane w rozdziale, szczególnie hybrydowych, metod z wykorzystaniem tekstury barwnej itp.

ROZDZIAŁ 5. Segmentacja obrazów barwnych

Niniejszy rozdział zawiera kilka wyników własnych otrzymanych w zakresie pikselowych i obszarowych technik segmentacji obrazów barwnych. Wiekszość z nich wykorzystuje znajomość właściwości przestrzeni barw typu HSx omówionych w poprzednich rozdziałach rozprawy. Progowanie w HSI w przypadku posiadania lub pozyskania dodatkowej wiedzy o tle sceny może ograniczyć się do wybranych składowych. Odblaski pochodzace od źródeł światła białego można łatwo detekować poprzez progowanie składowych S oraz I. Zaproponowana automatyczna dwuprzebiegowa technika nieziarnistego rozrostu obszaru uzupełniona procedura przetwarzania końcowego pokazała duże możliwości adaptacji do segmentacji różnego typu obrazów. Poprzez odpowiedni wybór przestrzeni barw np. HSI, zmianę formy kryterium jednorodności, zmianę koncepcji spójności czy zmianę wartości progu możliwe jest usuwanie cieni, odblasków, właściwa segmentacja obrazu monochromatycznego czy wygładzanie obrazu barwnego. Rozszerzono zastosowanie funkcji oceny segmentacji Q(I) na techniki obszarowe segmentacji i korzystając z tego pokazano przewagę segmentacji obszarowej nad pikselowa.

Prezentowane w tym rozdziale i poddawane segmentacji obrazy były bardzo dobrej jakości. Gdy poddajemy segmentacji zaszumiony obraz, to kilkakrotnie wzrasta liczba obszarów w posegmentowanym obrazie, rośnie średni błąd barwy oraz jednocześnie znacznie rośnie wartość funkcji oceny Q(I) [205]. Dlatego przed segmentacja w ramach przetwarzania wstępnego warto poddać obraz filtracji. Odpowiedni filtr powinien wygładzać obraz zachowując krawędzie (patrz rozdz. 6).



Barwne wersje ilustracji rozdziału 5





- Rys. 5.2. Segmentacja obrazu PARROTS: (a) oryginał, (b) niedosegmentacja, 48 obsz.; (c) nadsegmentacja, 3000 obsz.; (d) obraz (c) w pseudokolorach, (e) segmentacja, 62 obsz.
- Fig. 5.2. Segmentation of image PARROTS: (a) original, (b) undersegmentation, 48 reg.; (c) oversegmentation, 3000 reg.; (d) image (c) in pseudocolours, (e) segmentation, 62 reg.







(a)

Rys. 5.17. Segmentacja k-meansz parametrem k = 20: (a) obraz KWIAT1, (b) segmentacja 1, (c) segmentacja 2

Fig. 5.17. Segmentation by k-meanstechnique with parameter k = 20: (a) image KWIAT1, (b) segmentation 1, (c) segmentation 2

ROZDZIAŁ 5. Segmentacja obrazów barwnych

(c)



164

- (d) (e)
- Rys. 5.19. Ziarnisty rozrost obszaru, d = 150: (a) obraz KWIAT2 z ziarnem w 2 położeniach, (b) segmentacja (ziarno centralne), (c) segmentacja (ziarno boczne), (d) różnica obrazów (b) i (c), (e) segmentacja od 2 ziaren
- Fig. 5.19. Seeded region growing, d = 150: (a) image KWIAT2 with 2 seeds, (b) segmentation (seed in the middle), (c) segmentation (seed on the petal), (d) difference of images (b) and (c), (e) segmentation from 2 seeds



- Rys. 5.22. Segmentacja obrazu KWIAT2, zmiana parametru d: (a) oryginał, (b) segmentacja d = 40, 2194 obsz.; (c) segmentacja d = 70, 413 obsz.; (d) segmentacja d = 180, 3 obsz.
- Fig. 5.22. Segmentation of image KWIAT2 with parameter d: (a) original, (b) segmentation d = 40, 2194 reg.; (c) segmentation d = 70, 413 reg.; (d) segmentation d = 180, 3 reg.

Barwne wersje ilustracji rozdziału 5





- Rys. 5.27. Segmentacja obrazu KLOCKI2, zmiana metryki: (a) oryginał, (b) metryka euklidesowa w RGB (d = 45, A = 710), (c) metryka euklidesowa dla HS (d = 50, A = 1000), (d) metryka euklidesowa dla a^*b^* (d = 12, A = 1800), (e) metryka "kąt między wektorami" w RGB (d = 21, A = 2000)
- Fig. 5.27. Segmentation of image KLOCK12, change of metric: (a), (b) the Euclidean metric in RGB (d = 45, A = 710), (c) the Euclidean metric for HS components (d = 50, A = 1000), (d) the Euclidean metric for a^*b^* (d = 12, A = 1800), (e) the "angle between vectors" metric in RGB (d = 21, A = 2000)



- Rys. 5.28. Segmentacja obrazu CANDIES: (a) oryginał, (b) segmentacja z metryką euklidesową w $RGB \ d = 60$, (c) segmentacja z metryką euklidesową w $RGB \ d = 60$ i A = 5000
- Fig. 5.28. Segmentation of image CANDIES: (a) original, (b) segmentation with Euclidean metric in $RGB \ d = 60$, (c) segmentation with Euclidean metric in $RGB \ d = 60$, A = 5000



- Rys. 5.29. Przetwarzanie końcowe posegmentowanego obrazu KWIAT2 (parametr d = 70): (a) obraz posegmentowany z zaznaczonym oknem, (b) powiększony obraz pseudokolorowy z zaznaczonego okna przed przetwarzaniem końcowym, (c) przetwarzanie końcowe, A = 100, (d) przetwarzanie końcowe, A = 1000
- Fig. 5.29. Postprocessing of segmented image KWIAT2 (d = 70): (a) segmented image with selected window, (b) zoomed and pseudocolorized part of the image before postprocessing,(c) postprocessing, A = 100, (d) postprocessing, A = 1000



- Rys. 5.32. Segmentacja wielobarwnego czeku: (a) oryginał, (b) histogram poziomów szarości, (c) wynik progowania obrazu monochromatycznego, (d) wygładzony i poprawiony barwny obraz czeku, (e) wynik segmentacji
- Fig. 5.32. Segmentation of multicoloured cheque: (a) original, (b) grey-level histogram, (c) result of grey-level image thresholding, (d) smoothed and enhanced colour image of cheque, (e) final segmentation

Rozdział 6 Przetwarzanie wstępne obrazów

Procesy przetwarzania wstępnego (kwantyzacja barwy, filtracja odszumiająca itp.) mają na celu poprawę obrazu przed dalszymi etapami przetwarzania, przede wszystkim przed jego segmentacją. Urządzenia pozyskiwania i przesyłania obrazów barwnych są jednocześnie źródłami szumu, dlatego ważne jest, aby stosować odpowiednie techniki redukcji szumu, które jednocześnie z tłumieniem szumu zachowują krawędzie w obrazie (ang. *edge preserving filters*). Niekiedy również po przeprowadzeniu transformacji przestrzeni barw (np. z *RGB* na *HLS*) w celu wygładzenia obrazu przed segmentacją w przestrzeni *HLS* dokonuje się filtracji odszumiającej [60].

Poniżej przedstawiono różne techniki służące do przetwarzania wstępnego oraz oceniono ich przydatność w przetwarzaniu obrazów barwnych. Między innymi pokazano, że różnica barwności może być przydatna, obok tradycyjnych miar, takich jak PSNR i ΔE , do oceny algorytmów zarówno kwantyzacji barwy, jak i filtracji odszumiającej.

6.1. Kwantyzacja barwy

barwnych

Typowym zadaniem z zakresu przetwarzania obrazów barwnych jest kwantyzacja barwy [33,66]. Kwantyzacja obrazu barwnego jest procesem transformacji tzw. prawdziwego obrazu barwnego (ang. *true colour image*) opisanego za pomocą 24 bitów/piksel na obraz barwny, zawierający mniejszą liczbę specjalnie wybranych barw (paleta barw). Jest to więc proces redukcji liczby barw w obrazie. Ponieważ obrazy scen naturalnych zawierają dużą liczbę barw, to wierne przedstawienie ich za pomocą palety z niedużą liczbą barw nie jest zadaniem prostym. Proces kwantyzacji barwy jest często stosowany jako pomocnicza operacja w widzeniu i grafice komputerowej. Dobrym przykładem jest użycie kwantyzacji

ROZDZIAŁ 6. Przetwarzanie wstępne obrazów barwnych

barwy jako presegmentacji, która redukuje złożoność problemu segmentacji obrazu. Niektóre złożone techniki segmentacji mają wbudowane algorytmy kwantyzacji barwy [62]. Nowe barwy w obrazie wynikowym po kwantyzacji są wybierane tak, aby zminimalizować różnicę barw pomiędzy obrazem oryginalnym a wynikowym. Mała różnica barwy (błąd kwantyzacji) wymaga na ogół zastosowania palety zaprojektowanej indywidualnie dla danego obrazu. Błąd kwantyzacji zależy od: liczby barw w palecie, np. 256, 64, 16 barw (rys. 6.1), metody doboru barw do palety oraz techniki klasyfikacji pikseli. Kwantyzacja barwy jest więc rodzajem stratnej techniki kompresji obrazu. Jeżeli liczba barw w obrazje jest mniejsza od 256, to taki obraz można zapisać stosując jednobajtową głębie barw, co oznacza 3-krotną kompresję obrazu wyjściowego. Ponieważ obrazy cyfrowe mają już za sobą jeden proces kwantyzacji podczas ich pozyskiwania, to kwantyzacja barwy jest de facto procesem rekwantyzacji.



(a)



- Rys. 6.1. Wyniki kwantyzacji barwy dla obrazu MOUNTAINS: (a) 256 barw, (b) 64 barw, (c) 16 barw
- Fig. 6.1. Colour quantization results for image MOUNTAINS: (a) 256 colours, (b) 64 colours, (c) 16 colours

6.1. Kwantyzacja barwy

Proces kwantyzacji barwy składa się z 2 etapów: projektowania palety barw i odwzorowania pikseli obrazu oryginalnego na barwy z zaprojektowanej palety. Paleta barw może być stała (uniwersalna) lub adaptacyjna (dostosowywana do każdego obrazu). Poniżej przedstawiono wyniki przeprowadzonych przez autora badań 3 metod kwantyzacji barwy ocenione za pomocą dwóch zwyczajowo stosowanych miar oraz jednej nowo wprowadzonej miary.

Pierwsza metoda jest najprostszą i najszybsza metoda kwantyzacji barwy ze stała paletą i polega na równomiernej kwantyzacji barwy w przestrzeni RGB. Główną zasadą stosowaną w tej metodzie jest zasada kwantyzacji skalarnej, tzn. kwantyzacji każdej składowej RGB osobno. Skale z 256 poziomami wartości każdej składowej RGB są dzielone na 8 (w przypadku składowych R i G) lub 4 (dla składowej B) równe przedziały. Tak więc sześcian RGB jest dzielony na $8 \times 8 \times 4 = 256$ jednakowych, prostopadłościennych bloków. Każda z barw oryginału zostaje przypisana do jednego z tych bloków. Zróżnicowane podejście do składowych RGB można wyjaśnić najmniejszą czułością oka na składową niebieska.

Druga metoda kwantyzacji wykorzystuje percepcyjną przestrzeń barw HSV. Ponieważ odcień jest kątem, to składowa H może być zapisywana za pomocą 360 poziomów wartości, a składowe S i V - 255 poziomów. Oprócz barw opisywanych składowymi HSV paleta może zawierać także tzw. barwy achromatyczne, czyli takie, dla których odcień jest niezdefiniowany, a nasycenie równe zeru. Są one definiowane wyłącznie przez składową jasności V = R = G = B. Metoda wymaga od użytkownika podania liczby poziomów dla każdej składowej. Przyjmuje ona automatycznie, że liczba barw achromatycznych będzie większa o 1 od liczby poziomów jasności V. Przykładowo 256-elementowa paleta barw może zawierać 10 odcieni, 5 poziomów nasycenia, 5 poziomów jasności oraz 6 barw achromatycznych (zapis skrócony: $10 \times 5 \times 5 + 6$). Skala każdej ze składowych HSV oraz skala barw achromatycznych (skala szarości) zostaje podzielona na odpowiednią ilość równych przedziałów. Taka równomierna kwantyzacja jest uproszczeniem. Kołowy charakter skali odcieniowej ($H = 0^{\circ} = 360^{\circ}$) stwarza dodatkowy problem podczas kwantyzacji barw o odcieniach czerwonych. Przydzielenie większej liczby poziomów kwantowania odcieniowi niż pozostałym składowym wynika z faktu większej wrażliwości układu wzrokowego na zmiany odcienia. Podobną kwantyzację stosuje się w systemach wyszukiwania obrazów, takich jak np. VisualSEEK ($18 \times 3 \times 3 + 4$) [266], czy FOCUS ($64 \times 10 \times 10$) [57]. W przestrzeni HSV niekiedy stosuje się kwantyzację nierównomierną [309].

6.1. Kwantyzacja barwy

ROZDZIAŁ 6. Przetwarzanie wstępne obrazów barwnych

Trzecią rozpatrywaną tutaj metodą kwantyzacji barwy jest metoda k-means, która została opisana w poprzednim rozdziale jako technika segmentacji. Z punktu widzenia kwantyzacji barwy metoda ta jest iteracyjną metodą poprawiania wybranej na wstępie k-elementowej palety barw i zaliczana jest w literaturze do metod kwantyzacji wektorowej [33].

Duży problem stanowi ocena jakości skwantowanych obrazów. Nie istnieje tutaj jedna miara jakości. Obok ogólnych miar teoriosygnałowych, takich jak błąd średniokwadratowy RMSE czy stosunek szczytowego sygnału do szumu PSNR, stosowane są miary oparte na odległości pomiędzy barwami oryginału i skwantowanego obrazu w percepcyjnie równomiernej przestrzeni barw, np. *CIELAB*. Najczęściej stosuje się następujące wzory oparte na różnicy barw w przestrzeni *RGB* i odpowiednio w przestrzeni *CIELAB*:

$$RMSE = \sqrt{\frac{1}{3MN} \sum_{i=1}^{M} \sum_{j=1}^{N} (R_{ij} - R_{ij}^{*})^{2} + (G_{ij} - G_{ij}^{*})^{2} + (B_{ij} - B_{ij}^{*})^{2}}$$
(6.1)

$$\Delta E = \frac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} \sqrt{(L_{ij} - L_{ij}^*)^2 + (a_{ij} - a_{ij}^*)^2 + (b_{ij} - b_{ij}^*)^2}, \tag{6.2}$$

gdzie: R_{ij} , G_{ij} , B_{ij} są składowymi barwy obrazu oryginalnego, a R_{ij}^* , G_{ij}^* , B_{ij}^* są składowymi barwy obrazu powstałego w wyniku kwantyzacji. Podobne oznaczenia odnoszące się do składowych przestrzeni *CIELAB* zastosowano we wzorze (6.2). Obydwa wzory wyrażają różnicę barw pomiędzy pikselami porównywanych obrazów odniesioną do liczby pikseli. Porównywanie obrazów na zasadzie piksel do piksela nie uwzględnia jednak wpływu sąsiedztwa piksela na postrzeganą barwę. Jest to wada wszystkich tych miar.

Na bazie błędu RMSE często buduje się miarę logarytmiczną PSNR i wyraża ją w decybelach. Dla obrazów barwnych nie istnieje jedna przyjęta formuła PSNR. Zastosowano tutaj następującą wersję tej miary:

$$PSNR = 20 \ log_{10} \frac{255}{RMSE}$$
(6.3)

Łatwo zauważyć, że barwność obrazu w wyniku jego kwantyzacji spada tym szybciej, im mniejsza jest liczba barw w palecie. Wyniki dla 5 obrazów przedstawionych na rys. 3.12, potwierdzające to stwierdzenie, pokazano w tabelach: tab. 6.1, tab. 6.2. Analizując zawartość tych tabel można zauważyć, że w większości przypadków mniejszy spadek barwności występuje w przypadku kwantyzacji przeprowadzonej w przestrzeni *CIELAB*. Podobne wyniki odnośnie do przestrzeni *CIELAB* stosowanej w innej metodzie kwantyzacji barwy otrzymano niedawno [10]. Jeżeli założymy, że barwność obrazu wynikowego po kwantyzacji powinna być równa barwności obrazu oryginalnego [202,208], to do oceny wyników kwantyzacji barwy można zastosować następującą różnicę barwności:

$$\Delta M = |M_{oryg} - M_{kwant}| \tag{6.4}$$

Tabela 6.1. Barwność oryginałów i obrazów po kwantyzacji technik
ąk-means wRGB

RGB	CHART	CHART DUCK LANDSCAPE C		CHARACTERS	Mountains	
Oryginał	96,56	15,60	35,92	88,31	46,04	
256 barw	96,49	15,20	35,75	87,90	45,77	
16 barw	90,60	13,61	34,09	84,21	42,94	

W pracy [202] przedstawiono wyniki badań, w których porównywano ze sobą trzy opisane metody kwantyzacji, przy czym metodę kwantyzacji w HSV rozpatrywano dla 2 wariantów podziału przestrzeni (HSV_1 : $10 \times 5 \times 5 + 6$ i HSV_2 : $12 \times 4 \times 5 + 6$), a metodę k-means dla kwantyzacji w RGB i w CIELAB. W tab. 6.3 zestawiono wartości powyższych miar jakości skwantowanego obrazu (256 barw) uśrednione dla 5 naturalnych obrazów barwnych przedstawionych na rys. 3.15.

Należy zwrócić uwagę, że im wyższa wartość PSNR i im niższe wartości miar ΔE i ΔM , tym wyższa jest jakość skwantowanego obrazu. Uzyskane wyniki są kolejnym potwierdzeniem znanej prawdy, że adaptacyjne metody kwantyzacji (tutaj k-means) są wolniejsze, ale pozwalają otrzymać lepszą jakość obrazów skwantowanych. Generalnie nowa miara jakości ΔM zachowuje się podobnie jak dwie znane miary PSNR i ΔE . Jednak w przypadku metody k-means jak i kwantyzacji w HSV zachodzi rozbieżność, np. PSNR uzyskany dla obrazów kwantowanych metodą k-means w przestrzeni RGB jest wyższy niż w przypadku kwantowania w przestrzeni CIELAB. Błąd ΔE obliczony dla obrazu kwantowanego w przestrzeni RGB jest jednak wtedy większy niż dla obrazu kwantowanego w przestrzeni CIELAB. W takim przypadku dodatkowa miara ΔM może rozstrzygać o tym, która metoda kwantyzacji daje obrazy o lepszej jakości. Łatwo zauważyć, że im mniejsza jest liczba barw palety, tym mniejsza jest barwność obrazu po kwantyzacji, a więc strata barwności ΔM jest większa. Barwność obrazu może być przydatna w procesie oceny algorytmów kwantyzacji barwy oprócz takich tradycyjnych miar, jak: RMSE, PSNR oraz ΔE .

Wiele programów graficznych pozwala użytkownikowi na przeprowadzanie kwantyzacji barwy. Dlatego autor porównał najlepszą z rozpatrywanych metod kwantyzacji, tj. me-

F

ROZDZIAŁ 6. Przetwarzanie wstępne obrazów barwnych

Tabela 6.2. Barwność oryginałów i obrazów po kwantyzacji techniką k-means wCIELAB

CIELAB	CHART	DUCK	LANDSCAPE	CHARACTERS	MOUNTAINS
Oryginał	96,56	15,60	35,92	88,31	46.04
256 barw	96,47	15,36	35,77	87,93	45,84
16 barw	93,47	14,33	35,38	80,45	44.59

Tabela 6.3. Wyniki eksperymentu porównującego metody kwantyzacji barwy

Metoda kwantyzacji	PSNR (dB)	ΔE	ΔM
Równomierna RGB	19,94	21,05	8.78
Równomierna HSV_1	25,17	8,10	2,32
Równomierna HSV_2	25,23	8,40	2,77
K-means RGB	40,35	1,81	0,26
K-means CIELAB	37,90	1,47	0,21

todę k-means z kwantyzacją w popularnych programach Corel PhotoPaint w.11 i Adobe Photoshop w.7. Do porównania wybrano następujące algorytmy kwantyzacji: *Optimized* (PhotoPaint), *Adaptive* (PhotoPaint) oraz *Adaptive* (Photoshop) [216].

Wyniki tego porównania przedstawiono w tab. 6.4 i tab. 6.5.

Tabela 6.4. Wyniki kwantyzacji obrazów na 256 barw

RMSE	CHART	DUCK	LANDSCAPE	CHARACTERS	MOUNTAINS
k-means	2,37	2,27	2,84	4,23	3.34
Paint_optim.	2,54	2,31	2,96	3,92	3.55
Paint_adapt.	4,65	4,14	4,40	5,38	4 29
Photo_adapt.	4,05	2,68	4,33	9,91	4.87
ΔE	CHART	Duck	LANDSCAPE	CHARACTERS	Mountains
k-means	1,13	1,11	1,17	3,83	1.83
Pamt_optim.	1,21	1,16	1,23	5.61	1.95
Paint_adapt.	2,15	2,03	1,81	9.74	2 38
Photo_adapt.	1,98	1,30	1,76	12,81	2,52

Tabela 6.5. Wyniki kwantyzacji obrazów na 16 barw

RMSE	Chart	DUCK	LANDSCAPE	CHARACTERS	Mountains
k-means	21,61	7,22	8,68	14,71	11,33
Paint_optim.	22,15	7,38	8,62	13,60	13,20
Paint_adapt.	32,15	10,97	10,53	20,53	13,25
Photo_adapt.	24,86	6,81	10,21	17,74	11,91
ΔE	CHART	DUCK	LANDSCAPE	CHARACTERS	MOUNTAINS
k-means	8,21	2,52	3,15	11,49	5,03
Paint_optim.	7,00	2,63	2,62	12,58	5,01
Paint_adapt.	11,06	3,23	3,49	16,82	5,52
Photo_adapt.	7,44	2,39	3,00	16,73	4,82

W przypadku kwantyzacji obrazu na 256 barw metoda k-means wykazuje mniejsze błędy kwantyzacji niż porównywane z nią metody stosowane w programach komercyjnych. W przypadku kwantyzacji na 16 barw przewaga k-means nie jest już tak wyraźna. Główną wadą k-means w porównaniu z programami komercyjnymi jest dłuższy czas obliczeń.

6.2. Filtracja odszumiająca

Wpływ filtracji odszumiającej stosowanej przed segmentacją na wyniki segmentacji można oceniać w różny sposób, nie pomijając oceny wzrokowej. Bardziej obiektywną metodą oceny jakości samego filtru jest obliczanie różnych wskaźników jakości, np. PSNR, NCD (ang. *Normalized Colour Difference*) [231] itp. Najczęściej stosowane filtry odszumiające wygładzają obraz, ale równocześnie rozmywają krawędzie.

Poniżej przedstawiono wyniki badań przeprowadzonych z udziałem autora, mających na celu porównanie kilku nieliniowych filtrów redukujących szum w obrazie, a jednocześnie zachowujących krawędzie [109]. Filtry takie oparte są na uogólnieniu koncepcji znanych z zakresu filtracji obrazów monochromatycznych. Stosowanie takich filtrów może znacząco poprawić wyniki segmentacji. Jako punkt odniesienia stosowany był liniowy filtr uśredniający AMF (ang. *Arithmetical Mean Filter*). Co prawda szum addytywny może być skutecznie usuwany przez filtry liniowe (np. filtr Gaussa, AMF) stosowane oddzielnie do każdej składowej barwy, to jednak niestety generują one barwy, których nie było w obrazie oryginalnym. W omawianej pracy badaniu poddano następujące filtry:

- EPS (ang. *Edge Preserving Smoothing*), przedstawiony w wersji monochromatycznej w [168],
- SNN (ang. Symmetric Nearest Neighbour Filter) opisany w [230],
- K-N (ang. *Kuwahara-Nagao Filter*) zaproponowany w [135] dla obrazów monochromatycznych,
- VMF (ang. Vector Median Filter) przedstawiony w [8].

Algorytm filtru EPS stosuje dla każdego piksela 9 różnych masek 5×5 ; trzy z nich zostały pokazane na rys. 6.2. Na podstawie symetrii dla każdej z 2 pierwszych masek (rys. 6.2(a) i rys. 6.2(b)) można wygenerować dodatkowo 3 inne maski 7-elementowe. W ten sposób dysponujemy 8 maskami, a dziewiąta maska 9-elementowa została przedstawiona jako ostatnia na rys. 6.2(c). W wersji monochromatycznej dla każdej maski obliczamy wariancję, a nowy poziom szarości piksela centralnego to średnia dla maski o najmniejszej wariancji. Aby rozszerzyć ten algorytm na dziedzinę obrazów barwnych, obliczamy dla każdej maski macierz kowariancji i jej normę Frobeniusa [124]. Stosując dla macierzy wariancyjno-kowariancyjnej te same oznaczenia co we wzorze (4.23), można zapisać wyrażenie dla normy Frobeniusa tej macierzy:

$$||U||_F = \sqrt{\sum_{i=1}^{3} \sum_{j=1}^{3} |u_{ij}|^2}$$
(6.5)

Nowa barwa piksela centralnego jest średnią barw obliczoną dla maski z macierzą kowariancji o najmniejszej wartości normy Frobeniusa.



Rys. 6.2. Wybrane maski filtru EPS Fig. 6.2. Selected masks of EPS filter

Idea filtru SNN jest zbliżona do filtrów uśredniających i medianowych, lecz pozwala lepiej zachowywać krawędzie. Sąsiedzi piksela centralnego w oknie maski są pogrupowani w pary pikseli symetrycznych: N-S, W-E, NW-SE oraz NE-SW (rys. 6.3(a)). Dla każdej pary jest wybierany piksel bardziej zbliżony ze względu na barwę do piksela centralnego. Barwy tak wybranych 4 pikseli są uśredniane i ta średnia barwa stanowi nową barwę piksela centralnego.



Rys. 6.3. Maski badanych filtrów: (a) SNN, (b) K-N

Fig. 6.3. Masks for the tested filters: (a) SNN, (b) K-N

Maska filtru K-N, pokazana na 6.3(b), jest dzielona na 4 nachodzące na siebie okna o wymiarach 2 × 2, mające wspólną część, którą jest piksel centralny. Dla każdego z tych okien w przypadku obrazu monochromatycznego oblicza się wariancję. Jako nowy poziom szarości piksela centralnego jest stosowana wartość średnia obliczona dla pikseli zawartych w oknie z minimalną wariancją (najbardziej jednorodny obszar). Monochromatyczna wersja filtru K-N została uogólniona na przypadek masek okrągłych [15] oraz udowodniono, że filtr ten jest złożeniem liniowej dyfuzji i wyostrzania morfologicznego (ang. *morphological sharpening*) [27]. Podobnie jak w przypadku filtru EPS, rozszerzając działanie filtru K-N na obrazy barwne, obliczano macierze kowariancji i wybierano tę z najmniejszą normą Frobeniusa. Czas obliczeń dla filtru K-N jest większy niż w przypadku filtru SNN.

Filtr VMF jest zaproponowanym przez J. Astolę i in. [8] uogólnieniem pojęcia filtru medianowego na obszar obrazów barwnych. Barwy zawarte w prostokątnym oknie filtru VMF są traktowane jako wektory i wynikiem VMF w danym oknie jest wektor, którego skumulowana odległość od pozostałych wektorów (barw) w oknie jest najmniejsza. Funkcja odległości korzysta z metryki Euklidesa w przestrzeni barw *RGB*. W pracy [267] w sposób formalny opisano filtr VMF i jego różne warianty.

Obrazy do testowania pobierano z bazy obrazów i zaszumiano je odpowiednim rodzajem i poziomem szumu. W przypadku szumu impulsowego był to szum typu "sól i pieprz". Do tak zaszumionych obrazów stosowano poszczególne filtry, a następnie porównywano uzyskane obrazy z obrazami oryginalnymi (bez szumu), stosując 2 różne miary różnicowe: uniwersalną - stosunek sygnał/szum (SNR) oraz miarę percepcyjną - zgodną z CIE94

ROZDZIAŁ 6. Przetwarzanie wstępne obrazów barwnych

(patrz rozdz. 3) i oznaczoną jako ΔE_{CH} [257]. Im większa wartość SNR oraz im mniejsza różnica ΔE_{CH} , tym bardziej efektywna jest filtracja.

W wyniku tych testów stwierdzono, że w przypadku występowania w obrazie szumu impulsowego i stosowania miary ΔE_{CH} filtr medianowy VMF daje lepsze wyniki niż pozostałe filtry. Podobnie wygląda sytuacja w przypadku, gdy miarą jest SNR i jednocześnie wartość SNR dla obrazu wejściowego jest niska. Filtry liniowe redukują szum gaussowski lepiej niż filtry nieliniowe: filtr AMF uzyskał najmniejszą wartość ΔE_{CH} w przypadku szumu gaussowskiego. Biorąc pod uwagę obie miary należy stwierdzić, że najlepsze wyniki pozwala uzyskać filtr VMF.

Inna możliwość oceny działania filtrów jest oparta na wynikach segmentacji. Ona właśnie jest oceną ostateczną filtru, bo ocenia łącznie filtrację i segmentację a nie tylko walory wizualne obrazu. Do segmentacji zastosowano technikę nieziarnistego rozrostu obszaru w wersji opisanej w pracy [214]. Technika ta stosuje w procesie rozrostu uproszczoną wersję wzoru (5.26), nie uwzględniającą składowej intensywnościowej:

$$\sqrt{S^2 + \bar{S}^2 - 2S\bar{S}\cos(H - \bar{H})} \leqslant d \tag{6.6}$$

Latwo zauważyć, analizując wyniki segmentacji poprzedzonej filtracją, że liczba obszarów w obrazie po segmentacji ulega zmniejszeniu. Tak więc filtracja odszumiająca zapobiega nadsegmentacji obrazu. Przykładowo, gdy standardowy obraz PEPPERS zostanie zaszumiony szumem gaussowskim z parametrem $\delta = 30$ i dodatkowo 4% szumem impulsowym, to bezpośrednie poddanie go segmentacji za pomocą opisanej techniki prowadzi do otrzymania obrazu z 1605 obszarami. Liczbę tę można obniżyć stosując przetwarzanie końcowe w postaci np. usuwania małych obszarów (patrz rozdz. 5.4.) lub filtrację przed segmentacją. W tabeli 6.6 zestawiono wyniki segmentacji (liczba obszarów) obrazu PEP-PERS, uzyskane w sytuacji gdy przed segmentacją stosowano jeden z następujących filtrów: skalarny filtr medianowy SMF (ang. *Scalar Median Filtering*), wektorowy filtr medianowy VMF oraz filtr SAW (ang. *Self-Avoiding random Walk*) wykonywany jednokrotnie (SAW1) lub dwukrotnie (SAW2).

Tabela 6.6. Liczba obszarów w obrazie wynikowym dla różnych filtrów odszumiających

Liczba obszarów	SMF	VMF	SAW ₁	SAW ₂
Po segmentacji	224	209	85	33
Po przetwarzaniu końcowym	73	55	43	23

Obydwa filtry medianowe są szczególnie przydatne do usuwania szumu impulsowego. Filtracja SMF polega na oddzielnym zastosowaniu do każdej składowej barwy filtru medianowego, znanego z dziedziny przetwarzania obrazów monochromatycznych. W literaturze pokazano przewagę wyników uzyskiwanych za pomocą VMF nad SMF [17]. Filtr SAW jest oparty na idei cząstki błądzącej po ścieżce, która nie przechodzi nigdy dwa razy przez ten sam punkt [214].

Jak widać z danych zawartych w tab. 6.6, liczba obszarów uzyskana po zastosowaniu filtru SAW jest znacznie mniejsza niż w przypadku stosowania filtrów medianowych.

W pracach [201,204] autor zaproponował oryginalną ideę: pośrednią ocenę przydatności filtracji odszumiającej można przeprowadzić po segmentacji obrazu, stosując do niej funkcję znaną z procesu oceny samej segmentacji, np. funkcję Q(I) wg wzoru (5.36). Efektywność filtracji można ocenić porównując wartość tej funkcji uzyskaną po segmentacji poprzedzonej filtracją z wartością uzyskaną dla segmentacji bez wcześniejszej filtracji. Badania przeprowadzono dla obrazów oryginalnych i zaszumianych.

Do badań wybrano 5 nastepujących nieliniowych filtrów: SNN, K-N, VMF, PGF oraz DPA. Używano typowych wersji filtrów działających z maskami 3×3 i każdy z nich był stosowany jednokrotnie. Trzy pierwsze filtry zostały opisane powyżej.

Filtr PGF (ang. *Peer Group Filtering*) został zaproponowany przez Denga w pracy [61]. Dla każdego piksela filtr ten znajduje grupę sąsiadów (ang. *peer group*), opierając się na podobieństwie ich barwy i zastępuje barwę piksela średnią barwą sąsiadów. Piksele należące do szumu nie zostają zakwalifikowane do grupy sąsiadów i nie biorą udziału w obliczaniu średniej barwy. Wykrywanie pikseli szumu i odsunięcie ich od wpływu na wynik ma powodować wygładzanie obrazu bez rozmycia.

Filtr DPA (ang. *Digital Paths Approach*) został przedstawiony w pracy M. Szczepańskiego i in. [277]. Wykorzystuje on ideę ścieżek cyfrowych w oknie filtracyjnym i obliczania kosztów połączeń wzdłuż tych ścieżek. W postaci ścieżek rozumianych jako sekwencje pikseli modeluje się losowe błądzenie wirtualnej cząstki na dwuwymiarowej kracie obrazu (rys. 6.4). Koszt połączenia jest miarą zróżnicowania barwy pomiędzy pikselami ścieżki. Miara ta jest zbudowana na podstawie funkcji eksponencjalnej z parametrem wygładzającym. Barwa wynikowa piksela centralnego jest obliczana jako średnia ważona barw w oknie bez udziału piksela centralnego. Wagi są wyznaczane na podstawie ścieżek cyfrowych wychodzących z piksela centralnego do jego sąsiadów.

Do przeprowadzania segmentacji wybrano 2 techniki: jedną pikselową i jedną obszarową. Jako technikę pikselową wybrano technikę nazywaną mean shift [42]. W technice

bieżącym, będzie mniejsze od pewnej ustalonej wartości, np. 0,1. Ostateczny środek sfery jest środkiem klastera. Wtedy spośród punktów nie należących do sfery wybieramy jeden jako środek nowej sfery i powtarzamy proces tak długo, aż wszystkie punkty obrazu będą należeć do sfer-klasterów.

Technika mean shift należy do gęstościowych technik klasteryzacji (ang. density-based clustering). Formalny opis algorytmu mean shift odwołuje się do estymaty gradientu funkcji gęstości prawdopodobieństwa wyrażonej za pomocą funkcji jądrowej np. Epanechnikowa [42]. Wyznaczane w algorytmie wektory przesunięcia średniej (ang. mean shift) służą do przesuwania sferycznych okien w kierunkach lokalnych modów funkcji gęstości prawdopodobieństwa (środków klasterów).

Druga technika segmentacji używana w badaniach to technika obszarowa: nieziarnisty rozrost obszaru opisany w rozdz. 5.3.2..

Badania przeprowadzono na 10 standardowych obrazach barwnych (AIRPLANE, BA-BOON, GIRL, HATS, LENA, LIGHTHOUSE, MOTOCROSS, PARROTS, PEPPERS, SAILBO-ATS) o wysokiej jakości (niski poziom szumu). Dla każdego obrazu po segmentacji wyznaczono wartość funkcji jakości Q(I), a następnie wyznaczono wartość średnią funkcji dla 10 obrazów. Dla każdej z technik segmentacji średnie wartości Q(I) znormalizowano, aby ułatwić porównywanie. Oznacza to, że wartość Q(I) dla obrazów segmentowanych bez filtracji wynosi 100%.

Uzyskane rezultaty pokazują, że filtracja w przypadku segmentacji obszarowej ma bardzo ograniczony wpływ na wyniki segmentacji. W najlepszym przypadku filtr DPA redukuje wartość Q(I) do poziomu 70% (rys. 6.5). Gdy stosowana jest technika mean shift, znaczenie filtracji jest większe. W tym przypadku najlepsze filtry (K-N, DPA oraz VMF) są w stanie obniżyć wartość Q(I) do poziomu 15%.

W następnym eksperymencie zastosowano zaszumione obrazy barwne. Badane obrazy zaszumiano dodając do nich szum gaussowski ($\sigma = 7,5$) i impulsowy (p = 0,1 oraz $p_1 = p_2 = p_3 = 0,02$) [231]. Wyniki przedstawione na rys. 6.6 pokazują, że w tym przypadku wszystkie filtry są bardzo przydatne, ponieważ poprawiają wyniki segmentacji uzyskane przez obie techniki. Największą poprawę można zauważyć, gdy stosuje się filtr DPA lub VMF. Należy zwrócić uwagę na to, że poziom 100% odnosi się do segmentacji obrazów zaszumionych i nie odpowiada poziomowi 100% z rys. 6.5.

W większości badanych przypadków wartości funkcji Q(I) były mniejsze dla segmentacji obszarowej. Oznacza to, że ta technika jest lepsza niż technika mean shift. Dodatkowo można jeszcze zastosować przetwarzanie końcowe. Rysunek 6.7 przedstawia wyniki po



(b)

Fig. 6.4. Example digital paths connected pixels in the window 3×3 : (a) paths of length 2, (b) paths of length 3

tej wykorzystuje się średnią wartość składowych barwy dla punktów obrazu, znajdujących się wewnątrz sfery o zadanym promieniu jako oszacowanie gradientu funkcji gęstości prawdopodobieństwa. Poprzez przesuwanie sfery, uzyskane dzięki zmianie jej środka, w kierunku obszaru charakteryzującego się maksymalną gęstością punktów obrazu uzyskiwane są reprezentatywne środki klasterów. Jedynym parametrem techniki mean shift jest promień sfery.

Na początku ustala się promień sfer, do których w kolejnych iteracjach będą przyporządkowywane punkty obrazu. Wybiera się dowolny punkt obrazu jako początkowy środek pierwszej sfery. Przyporządkowuje się punkty obrazu do sfery, jeżeli ich odległości od środka sfery są mniejsze od ustalonego promienia. Następnie oblicza się średnie wartości składowych R, G i B punktów obrazu, które zostały przyporządkowane do sfery. Te składowe wyznaczają nowy przesunięty środek sfery. Przyporządkowuje się punkty obrazu do nowej sfery o przesuniętym środku. Wyznacza się nowy środek sfery. Proces ten przebiega dopóki przesunięcie środka sfery, czyli różnica między środkiem poprzednim sfery a

ROZDZIAŁ 6. Przetwarzanie wstępne obrazów barwnych



dołączeniu przetwarzania końcowego. Nastąpiła znaczna zmiana: wyniki w stosunku do przypadku segmentacji bez filtracji na ogół nie są zadowalające. Ponieważ filtracja wymaga czasu a jej efektywność nie jest wysoka, to w przypadku stosowania przetwarzania końcowego przetwarzanie wstępne jest bezcelowe. Wniosek ten dotyczy głównie segmentacji obszarowej. Poziom 100% odnosi się do segmentacji obrazów zaszumionych zakończonej procedurą przetwarzania końcowego i nie odpowiada poziomowi 100% z rys. 6.5 i rys. 6.6.



Rys. 6.7. Średnie wartości funkcji Q(I)dla badanych obrazów dodatkowo zaszumianych i poddanych przetwarzaniu końcowemu

Fig. 6.7. Average values of Q(I) function for postprocessed tested images with an additional noise

Oprócz obrazów naturalnych poddano testom dwa obrazy sztuczne o niskiej rozdzielczości (320 × 200 pikseli). Obrazy zawierały, oprócz tła, odpowiednio 5 i 3 obszary. Do obrazów domieszano ten sam rodzaj szumu gaussowsko-impulsowego. W przypadku niestosowania filtracji, szum dominuje w posegmentowanych obrazach: liczba obszarów dla obydwu technik wynosi około 1900. W przypadku zastosowania filtru, najmniejszą liczbę obszarów zbliżoną do odpowiednio 6 i 4 obszarów, otrzymano stosując filtry VMF i DPA. Najgorszy wynik otrzymano stosując filtr SNN.

Badane filtry nieliniowe okazały się szczególnie efektywne w przypadku zaszumionych obrazów. Porównanie filtrów pokazuje, że efektywność filtrów DPA i VMF przewyższa inne testowane filtry. W przypadku stosowania przetwarzania końcowego w postaci usuwania





Rys. 6.5. Średnie wartości Q(I) dla badanych obrazów Fig. 6.5. Average values of Q(I) function for tested images





ROZDZIAŁ 6. Przetwarzanie wstępne obrazów barwnych

małych obszarów korzystanie z przetwarzania wstępnego może być bezcelowe. Efektywność przetwarzania wstępnego zależy od techniki segmentacji obrazu - generalnie przetwarzanie wstępne jest bardziej efektywne w przypadku prostszej segmentacji pikselowej niż segmentacji obszarowej. Podsumowując można stwierdzić, że funkcja jakości Q(I) może być nie tylko stosowana do porównywania technik segmentacji między sobą, ale również do oceny przetwarzania wstępnego przed segmentacją obrazu barwnego.

Poniżej zaproponowano nową ideę do oceny filtracji odszumiającej: zastosowanie miary w postaci zmiany barwności obrazu. Barwność obrazu, podobnie jak inne globalne atrybuty percepcyjne (ostrość, naturalność itp.), nie powinna ulegać zmianie w wyniku filtracji odszumiającej. Dodanie szumu do obrazu barwnego może jednak spowodować zmianę jego początkowej barwności. Wybrano 3 obrazy testowe: słabo barwny obraz DUCK, średnio barwny obraz MOUNTAINS oraz ekstremalnie barwny obraz CHART [203]. Obrazy te zaszumiano 2 rodzajami szumu impulsowego (chromatycznym i achromatycznym), zmieniając stopień zaszumienia od 0% do 18%. Zmienność barwności obrazu, spowodowana szumem impulsowym, została przedstawiona na rys. 6.8. Z wykresów tych widać, że chromatyczny szum impulsowy powoduje wzrost barwności obrazu, a achromatyczny szum impulsowy -spadek barwności obrazu. Wzrost barwności obrazu, a achromatyczny dla obrazów słabo barwnych (np. DUCK), a jej spadek dla obrazów ekstremalnie barwnych (np. CHART). Podobne badania pokazały mniejszy wpływ szumu gaussowskiego na barwność obrazu (rys. 6.9).

Do oceny jakości obrazu po filtracji odszumiającej stosuje się te same miary co do oceny obrazu po kwantyzacji barwy. Wyrażają je wzory (6.1)–(6.3). Jeżeli założymy, że barwność obrazu wynikowego po filtracji powinna być równa barwności obrazu oryginalnego, to do oceny wyników filtracji obrazu można stosować zmianę barwności wyrażoną wzorem (6.4). Aby przekonać się o przydatności ΔM w tym kontekście dla 3 badanych powyżej obrazów, przeprowadzono porównanie 2 medianowych technik filtracji: SMF i VMF.

W tabelach 6.7–6.9 [203] zaprezentowano wyniki uzyskane podczas tego porównania. Do oceny stosowano 3 kryteria: PSNR, ΔE i ΔM , z których tylko pierwsze wymaga maksymalizacji. Dla ułatwienia interpretacji wyników "pogrubiono" te, które wskazują na przewagę filtru SMF. Z poniższych tabel widać, że PSNR daje wyniki przeciwstawne do ΔE , preferując filtr SMF. W takiej sytuacji ΔM może odgrywać rolę trzeciego kryterium, które pozwala podjąć ostateczną decyzję. W większości badanych sytuacji filtr VMF lepiej zachowuje barwność obrazu oryginalnego, szczególnie w sytuacji dużego zaszumienia szumem impulsowym (15% oraz 18%). Wynik porównania dodatkowo potwierdza



Rys. 6.8. Wpływ szumu impulsowego na barwność obrazu Fig. 6.8. Impact of impulsive noise on the image colourfulness



Rys. 6.9. Wpływ szumu gaussowskiego na barwność obrazu Fig. 6.9. Impact of Gaussian noise on the image colourfulness

182

ROZDZIAŁ 6. Przetwarzanie wstępne obrazów barwnych

przydatność VMF, pomimo większego czasu obliczeń, którego on wymaga. Badania te pokazują w sposób pośredni, że zmiana barwności jest bliższa błędowi barwy ΔE niż teoriosygnałowemu wskaźnikowi PSNR.

Tabela 6.7. Wartości PSNR dla obrazów poddanych filtracji medianowej

Szum impulsowy	[%]	3	6	9	12	15	18
Duck	SMF	35,3	34,9	34,5	34,2	33,6	33,1
Book	VMF	34,9	34,4	33,9	33,5	32,7	32,1
MOUNTAINS	SMF	30,2	30,1	30,0	29,9	29,7	29,5
	VMF	30,0	29,9	29,7	29,5	29,3	28,9
Снарт	SMF	42,7	41,6	39,9	37,6	36,5	35,1
	VMF	42,9	41,9	39,8	37,4	35,8	33,7

Tabela 6.8. Wartości ΔE dla obrazów poddanych filtracji medianowej

Szum impulsowy [%]		3	6	9	12	15	18
Duck	SMF	0,99	1,07	1,16	1,23	1,31	1,40
	VMF	0,92	0,97	1,02	1,08	1,16	1,24
MOUNTAINS	SMF	2,54	2,63	2,74	2,85	2,96	3,09
	VMF	2,42	2,47	2,53	2,59	2,67	2,75
Снарт	SMF	1,00	1,03	1,08	1,15	1,21	1,29
	VMF	0,93	0,96	1,01	1,07	1,15	1,27

Dodatkowo porównano 2 filtry znane z literatury: DDT (ang. Directional Distance Filter) [125] oraz CWVMF (ang. Central Weighted Vector Median Filter) [131]. Dla obrazu Mountains zaszumianego 15% szumem impulsowym otrzymano następujące wartości miar: dla filtru DDT (PSNR=28,54 dB, $\Delta E = 3,01$) oraz dla filtru CWVMF (PSNR=28,68 dB, $\Delta E = 3,13$). Obie miary dają sprzeczne wyniki i nie wskazują jednoznacznie, który filtr jest lepszy. Wyznaczone dodatkowo wartości ΔM : 0,22 dla filtru DDT oraz 0,16 dla filtru CWVMF wskazują na przewagę tego ostatniego filtru.

W powyższych badaniach pokazano wpływ szumu w obrazie na jego barwność. Wpływ ten jest szczególnie widoczny dla obrazów o niskiej barwności (szum chromatyczny zwiękTabela 6.9. Wartości ΔM dla obrazów poddanych filtracji medianowej

Szum impulsowy [%]		3	6	9	12	15	18
DUCK	SMF	0,15	0,47	0,79	1,12	1,60	1,99
	VMF	0,10	0,06	0,03	0,11	0,55	1,06
	SMF	0,05	0,02	0,09	0,18	0,35	0,51
MOUNTAINS	VMF	0,08	0,07	0,03	0,01	0,17	0,29
CHART	SMF	0,05	0,03	0,00	0,05	0,16	0,20
	VMF	0,06	0,07	0,06	0,07	0,03	0,08

sza barwność) i dla obrazów o bardzo wysokiej barwności (szum achromatyczny zmniejsza barwność). Filtracja odszumiająca nie powinna zmieniać barwności obrazu. Przykład medianowych technik filtracji pokazuje, że zmiana barwności może być przydatna do oceny jakości obrazu po filtracji, szczególnie, gdy tradycyjne miary (PSNR, ΔE itp.) dają sprzeczne wyniki. Należy podkreślić, że przedstawione tutaj badania wstępne wymagają rozszerzenia swojego obszaru na zdecydowanie bardziej liczny zbiór reprezentatywnych obrazów barwnych.

Podsumowując uzyskane w tym rozdziale wyniki można stwierdzić, że filtry odszumiające mogą być oceniane w sposób pośredni poprzez funkcję oceny segmentacji, np. Q(I), następującej po filtracji. Jest to podejście alternatywne w stosunku do oceny obrazu po filtracji za pomocą miar typu PSNR, NCD itp. Pomimo badań literaturowych autorowi nie są znane publikacje opisujące podobne podejście. Drugi istotny wynik przedstawiony w tym rozdziale to potwierdzenie przydatności różnicy barwności do oceny kwantyzacji i filtracji odszumiającej. Miara ta może stanowić rozstrzygające kryterium oceny, gdy inne miary dają sprzeczne wyniki.

184

Rozdział 7

Podsumowanie i wnioski

Niniejsza monografia stanowi wynik pracy autora prowadzonej od wielu lat w zakresie cyfrowego przetwarzania obrazów barwnych. Stosunkowo dużo uwagi poświęcono w niej zagadnieniu reprezentacji barwy w obrazach cyfrowych, chociaż ograniczono się do najbardziej popularnych przestrzeni barw oraz tych, które zostały wykorzystane w dalszych rozdziałach rozprawy. Była to decyzja wynikająca z przekonania o zasadniczym znaczeniu właściwej reprezentacji barwy dla uzyskiwanych wyników przetwarzania obrazów. Część uzyskanej wiedzy o przestrzeniach barw, związana z transformacjami przestrzeni *RGB*, została wyrażona w postaci tabeli klasyfikacyjnej z 3 kryteriami. Ta tabela pokazuje, że nie istnieje jedna idealna przestrzeń barw, ale sama tabela może być pomocna przy doborze przestrzeni barw do zadania przetwarzania obrazów.

Znajomość transformacji z przestrzeni RGB do danej przestrzeni pozwala wyznaczyć kształt bryły barw. Pokazano to w pracy wyznaczając wartości graniczne składowych przestrzeni CIELAB, która nie jest kulą, jak nieraz bywa przedstawiana w literaturze. Wiedza o przestrzeniach barw nie sprowadza się jedynie do znajomości definiujących je transformacji. Dlatego w rozprawie przeprowadzono bardziej szczegółowe badania dla grupy popularnych w przetwarzaniu obrazów przestrzeni percepcyjnych HSx oraz przestrzeni $K_1K_2K_3$, będącej wynikiem transformacji Karhunena-Loevego. Przestrzenie typu HSx związane złożonymi nieliniowymi transformacjami z przestrzenią RGB w różnym stopniu aproksymują percepcyjne atrybuty barwy i ich relacje. W badaniach doświadczalnych czterech wersji przestrzeni typu HSx pokazano szczególną zwartość klasterów barw w przypadku przestrzeni HSI. Z kolei tylko dla wersji HLS udało się stosunkowo łatwo wyprowadzić wzory dla wyznaczania negatywu w przestrzeni percepcyjnej. Wynik ten może stanowić o przydatności HLS w niektórych zadaniach przetwarzania obrazów. Problemem związanym z przestrzenią $K_1K_2K_3$ jest konieczność wyznaczania dla każdego obrazu ma-

ROZDZIAŁ 7. Podsumowanie i wnioski

cierzy transformującej z RGB. Dlatego implementacja transformacji KLT powinna być efektywna lub zastąpiona aproksymacją za pomocą macierzy ze stałymi współczynnikami. W rozprawie pokazano obydwa podejścia: efektywną implementację KLT i wyznaczanie takich współczynników transformujących, które zapewniając dekorelację składowych są stałe dla pewnej klasy obrazów. Badając właściwości składowych barwy po transformacji KLT zaobserwowano m.in. dużo mniejszą wrażliwość składowej K_1 niż luminancji Y na zmiany jasności w obrazie barwnym RGB. Opierając się na tym wyniku zaproponowano stosowanie, w warunkach zmiennego poziomu oświetlenia sceny i wymaganego wysokiego kontrastu obrazu, monochromatycznych systemów wizyjnych pracujących z kamerą kolorową.

Do zagadnień związanych z pozyskiwaniem obrazów barwnych za pomocą systemów wizyjnych należą problem kalibracji kolorymetrycznej i oceny wierności odwzorowania. barw w systemie wizyjnym oraz pośrednio wyznaczanie charakterystyk widmowych kamery. Zaproponowana i sprawdzona dla systemów wizyjnych z kamerami jednoprzetwornikowymi CCD kilkukrokowa procedura kalibracji kolorymetrycznej pozwala wyraźnie zmniejszyć błąd odwzorowania barwy. W podobnych badaniach dla cyfrowego aparatu fotograficznego pokazano, że przejście od własnej przestrzeni RGB aparatu do przestrzeni standardowej sRGB zdecydowanie poprawia wierność odwzorowania barw. Do jej obliczenia zastosowano klasyczną oraz dwie nowe formuły obliczania różnicy barw proponowane przez CIE. Szczególnie dobre wyniki uzyskano dla najnowszej formuły CIEDE2000, Znajomość zakresu barw aparatu pozwoliła wyznaczyć dla niego macierz transformacji z jego przestrzeni RGB do CIE XYZ. W pracy w wyniku eksperymentu pokazano, jak można wyznaczać charakterystyki widmowe kamer na podstawie znanych charakterystyk odbiciowych sceny i charakterystyk oświetlenia nie stosując specjalistycznego sprzetu spektrofotometrycznego. W tym celu zmodyfikowano metodę numeryczna estymacji charakterystyk widmowych kamery, ograniczając ją do stosowania wektorów własnych macierzy pseudoodwrotnej. Chociaż wyniki sa zachecające, to ten kierunek badań wymaga dalszych działań.

Używane w pracy obrazy barwne były na ogół dobrej jakości. Gdy w obrazie występuje zauważalny poziom szumu (impulsowego, gaussowskiego itp.), to przed jego segmentacją w ramach przetwarzania wstępnego warto poddać obraz filtracji. Gdy nie stosuje się filtru i dokonuje się segmentacji obrazu zaszumionego, to kilkakrotnie wzrasta liczba obszarów w posegmentowanym obrazie, rośnie średni błąd barwy między oryginałem i obrazem posegmentowanym. Odpowiedni filtr powinien wygładzać obraz zachowując krawędzie. Szukając takiego filtru uogólniono koncepcje filtrów nieliniowych stosowane dla obrazów monochromatycznych. Przykładem filtru dającego dobre wyniki jest wektorowy filtr medianowy VMF, będący realizacją idei przetwarzania wektorowego stosowanej również z powodzeniem dla innych problemów przetwarzania obrazów barwnych.

Ciągle otwartym problemem pozostaje jednak problem oceny efektywności tych filtrów. Można korzystać ze znanych wskaźników jakości: teoriosygnałowego PSNR, kolorymetrycznego ΔE i innych, które porównują obrazy metodą "piksel do piksela". W pracy zaproponowano i przetestowano wskaźnik nazwany różnicą barwności obrazów ΔM , który porównuje ze sobą barwności obrazów i może służyć jako dodatkowe kryterium oceny fitracji czy kwantyzacji. Zdefiniowane niedawno w literaturze pojęcie barwności obrazu poddano w pracy konfrontacji z pojęciem liczby barw w obrazie oraz zbadano jego zależność od lokalnych percepcyjnych atrybutów barwy. Celem tych badań było wykorzystanie pojęcia barwności w przetwarzaniu obrazów, między innymi poprzez śledzenie zmian barwności i ocenę jakości przetwarzania na ich podstawie. Tak więc barwność obrazu, będąca jego globalnym atrybutem percepcyjnym, może stanowić podstawę do konstrukcji nowego wskaźnika jakości.

Oceniając filtrację odszumiającą należy pamiętać, że jej ostatecznym celem w tym przypadku ma być poprawa segmentacji. Stwarza to możliwość oceny filtrów za pomocą funkcji oceny Q(I) używanej normalnie do oceny obrazów po segmentacji. Ten oryginalny pomysł pozwolił pokazać, że filtracja odszumiająca odgrywa większą rolę, gdy obraz jest segmentowany za pomocą techniki pikselowej niż w przypadku techniki obszarowej. Przetwarzania wstępnego nie należy sprowadzać jedynie do redukcji szumów. W czasie pozyskiwania cyfrowych obrazów barwnych następuje kwantyzacja sygnału wizyjnego. W niektórych zastosowaniach występuje potrzeba redukcji liczby barw w obrazie (kwantyzacja barwy, a właściwie jej rekwantyzacja), czego skutkiem jest pewnego rodzaju presegmentacja obrazu. Stąd m.in. wynika ciągłe poszukiwanie nowych technik kwantyzacji barwy. Przeprowadzone badania pokazały, że klasyczna technika k-means stosowana w przestrzeniach barw RGB i CIELAB daje bardzo dobre wyniki w porównaniu z programami komercyjnymi, chociaż nie jest to technika szybka. Udowodniono, że zmiana barwności obrazu może być miarą jakości również w procesie kwantyzacji.

W rozprawie szczegółowo przedstawiono wyniki badań stosunkowo prostych technik segmentacji, takich jak progowanie, technika klasteryzacyjna *k*-means, technika rozrostu obszaru, które dobrze nadają się do praktycznych zastosowań. Najprostszą technikę segmentacji, tj. progowanie, szczególnie łatwo jest przeprowadzić w przestrzeni percep-

cyjnej, gdy dysponuje się wiedzą o tle obrazu lub o barwach najważniejszych obiektów w scenie. Opierając się na znajomości właściwości przestrzeni typu HSx zaproponowano segmentację obrazu przez progowanie różnych składowych HSI w zależności od barwy tła. Szczególnym przypadkiem takiego podejścia jest szybka detekcja odblasków w obrazie na podstawie progowania składowych S i I. W przypadku techniki k-means zwrócono uwagę na to, jak wpływa na wyniki segmentacji położenie początkowych środków klasterów oraz kryterium stopu. Szczegółowo zostały opisane dwie wersje techniki rozrostu obszaru: półautomatyczna ziarnistego rozrostu i oryginalna automatyczna nieziarnistego rozrostu. Zaproponowana przez autora dwuprzebiegowa technika nieziarnistego rozrostu obszaru uzupełniona procedurą przetwarzania końcowego pokazała dużą przydatność w segmentacji różnego typu obrazów. Duży wpływ na wygląd obrazu posegmentowanego ma wybrana przestrzeń barw, kierunek agregacji pikseli, forma kryterium jednorodności i przyjęta koncepcja spójności. Pokazano, że odporność algorytmu segmentacji na cienie w obrazie można otrzymać stosując specjalną metrykę kątową. Przetwarzanie końcowe ma duże znaczenie w likwidowaniu nadsegmentacji w obrazie, usuwaniu odblasków i łaczeniu obszarów o podobnej barwie.

Ponieważ nie istnieje jedna, idealna przestrzeń barw, to również podczas rozwiązywania zadania segmentacji obrazu barwnego należy wybrać taką przestrzeń, która daje najlepsze wyniki dla danego zastosowania lub klasy obrazów. Wyniki segmentacji uzyskane za pomocą techniki pikselowej i obszarowej porównano ze sobą stosując ilościowe kryterium Q(I). W literaturze kryterium to stosowano do technik pikselowych, a w niniejszej pracy zastosowano je również do techniki obszarowej. Doświadczalne porównanie jednoznacznie wskazało na przewagę techniki obszarowej nad techniką pikselową. Na licznych przykładach w niniejszej pracy pokazano, że kryterium jednorodności, jego parametry (np. progi) oraz przestrzeń barw, którą wybrano dla reprezentowania obrazu, mają kluczowe znaczenie dla wyników segmentacji.

Uniwersalna technika segmentacji obrazu barwnego prawdopodobnie nie istnieje. Tym bardziej, że cele segmentacji wyraźnie zależą od zadania, którego elementem jest proces segmentacji. Rosnące możliwości obliczeniowe sprzętu komputerowego sprawiają, że wzrastają możliwości zastosowań technik bardziej złożonych niż te prezentowane w pracy, szczególnie technik hybrydowych, technik z wykorzystaniem tekstury barwnej itp.

Realizacja badań przedstawionych w pracy pokazuje, że chociaż tematyka przetwarzania obrazów barwnych należy do zagadnień szeroko rozumianego przetwarzania informacji, to coraz szersze zastosowanie znajdują w niej pojęcia i twierdzenia nauki o barwie. Przykładowo w niniejszej pracy stosowano pojęcie barwności obrazu (globalny atrybut percepcyjny) pochodzące od barwności bodźca oraz najnowszą miarę różnicy barw, zdefiniowaną przez CIE – CIEDE2000.

Dla wielu z przedstawionych w rozprawie zagadnień przetwarzania obrazów barwnych istnieje realna perspektywa dalszych badań. Przykładowo w dziedzinie reprezentacji barw można formułować pewne pytania domagające się rozstrzygnięcia. Na przykład czy zaproponowana specjalnie dla obrazów wersja przestrzeni CIELAB nazwana przez autorów S-CIELAB wprowadza jakieś zalety w rozwiązywaniu zagadnień przetwarzania obrazów barwnych? Jakie zastosowanie mogą znaleźć w przetwarzaniu obrazów barwnych modele percepcji barw typu CAM opracowane w nauce o barwie? Co wnosi do reprezentacji barw podejście spektralne, które uogólnia obrazy barwne na postać wielokanałową (tzw. obrazy wielospektralne)? Chcąc oceniać obrazy barwne trzeba szukać odpowiedzi na pytanie, na ile uprawnione jest porównywanie dwóch obrazów barwnych poprzez porównywanie odpowiadających sobie pikseli. W zakresie przetwarzania wstępnego należy sprawdzić, czy zastosowanie kwantyzacji barwy przed segmentacją poprawia wynik tej ostatniej i w jakim stopniu. Szczególne zainteresowanie autora budzą możliwości konstrukcji adaptacyjnych algorytmów segmentacji poprzez uzależnienie wartości progów d od zawartości obrazu. W ten sposób można wprowadzić do kryteriów jednorodności lokalne statystyki obrazu, np. wariancję. Ogromny obszar dalszych badań stanowi zagadnienie oceny wyniku segmentacji. W pracy stosowano głównie jedno z możliwych podejść – funkcję oceny Q(I). W tej dziedzinie istnieje wiele różnych podejść i aktualnie opublikowano pierwszą monografię na temat oceny wyników segmentacji. W dalszej perspektywie celowe staje się śledzenie nowych trendów w nauce o barwie i sprawdzanie ich przydatności do przetwarzania obrazów barwnych.

Cyfrowe przetwarzanie obrazów barwnych wraz z nauką o barwie współtworzą dziedzinę nauki i techniki nazywaną cyfrowym obrazowaniem barwnym (ang. *digital colour imaging*). Można przypuszczać, że dalszy rozwój obrazowania barwnego będzie wynikiem zarówno postępu w przetwarzaniu obrazów, jak i coraz lepszego rozumienia zjawiska barwy.

190

Bibliografia

- [1] R. Adams and L. Bischof. Seeded region growing. *IEEE Transactions on Pattern Analysis* and Machine Intelligence, 16(6):641-647, 1994.
- [2] H. Ailisto and T. Piironen. Evaluation of color representation methods in a practical vision system. In: Proceedings of the 5th Scandinavian Conference on Image Analysis, vol. 1, Stockholm, Sweden, 1987.
- [3] M. Ali, W. Martin and J. Aggarwal. Color-based computer analysis of aerial photographs. Computer Graphics and Image Processing, 9(3):282-283, 1979.
- [4] M. Anderberg. Cluster Analysis for Applications. Academic Press, New York, USA, 1973.
- [5] I. Andreadis, M. Browne and J. Swift. Image pixel classification by chromaticity analysis. Pattern Recognition Letters, 11(2):51-58, 1990.
- [6] I. Andreadis and P. Tsalides. Coloured object recognition using invariant spectral features. Journal of Intelligent and Robotic Systems, 13(1):93-106, 1995.
- [7] I. Andreadis, (ed.). Special issue on color imaging. Pattern Recognition, 35(8):1641–1806, 2002.
- [8] J. Astola, P. Haavisto and Y. Neuvo. Vector median filters. Proceedings of the IEEE, 78(4):678-689, 1990.
- [9] V. Athitsos, M. Swain and C. Frankel. Distinguishing photographs and graphics on the World Wide Web. In: Proceedings of Workshop on Content-Based Access of Image and Video Libraries (CBAIVL '97), 10-17, San Juan, Puerto Rico 1997.
- [10] A. Atsalakis, N. Papamarkos, N. Kroupis, D. Soudris and A. Thanailakis. Colour quantization technique based on image decomposition and its embedded system implementation. *IEE Proc.-Vis. Image Signal Process.*, 151(6):511–524, 2004.
- [11] M. Babu, C.-H. Lee and A. Rosenfeld. Determining plane orientation from specular reflectance. Pattern Recognition, 18(1):53-62, 1985.
- [12] J. Bajon, M. Cattoen and S. Kim. Real-time colorimetric transformations using in robot vision system. In: *Proceedings of MICAD*, 76-86, Paris, France 1985.
- [13] J. Bajon, M. Cattoen and L. Liang. Identification of multicoloured objects using a vision module. In: Proceedings of the 6 th RoViSeC, 21-30, Paris, France 1986.
- [14] D. Baker, S. Hwang and J. Aggarwal. Detection and segmentation of man-made objects in outdoor scenes: concrete bridges. *Journal of Optical Society of America A*, 6(6):938–950, 1989.
- [15] P. Bakker, L. van Fliet and P. Verbeek. Edge preserving orientation adaptive filtering. In: Proceedings of 5th Annual Conference of the Advanced School for Computing and Imaging, 207-213, Heijen, The Netherlands 1999.
- [16] A. Bal, H. Palus, P. Wołczyk. Selected properties of perceptual colour spaces. In: M. Kurzyński, E. Puchała, M. Woźniak, (eds.), *Computer Recognition Systems*, 125–130. Wrocław University of Technology, Wrocław 2001.

- [17] A. Bardos and S. Sangwine. Selective vector median filtering of colour images. In: Proceedings of 6th Int. Conf. on Image Processing and its Applications, 708–711, Dublin, Ireland 1997.
- [18] M. Barth, S. Parthasarathy, J. Wang and et al. A color vision system for microelectronics: applications to oxide thickness measurements. In: *Proceedings of IEEE Int. Conf. on Robotics and Automation*, vol. 2, 1242–1247, San Francisco, USA 1986.
- [19] K. B. Benson. Television Engineering Handbook. McGraw-Hill, New York 1986.
- [20] D. Bereska. Badania nad korelacją pomiędzy składowymi wektora barwy w cyfrowych obrazach barwnych. Praca doktorska, Instytut Automatyki Politechniki Śląskiej, Gliwice 2002.
- [21] D. Bereska, H. Palus. Korelacja składowych sygnału wyjściowego jednoprzetwornikowej kamery kolorowej CCD. W: Materiały VI Konferencji Naukowej "Czujniki Optoelektroniczne i Elektroniczne" COE 2000, 487–492, Gliwice 2000.
- [22] D. Bereska and H. Palus. Correlation of colour components of camera output signal and decorrelation methods. In: Optoelectronic and Electronic Sensors IV, Proceedings of SPIE, vol. 4516, 299–306, Gliwice 2001.
- [23] D. Bereska, H. Palus. Wyznaczanie charakterystyk widmowych kamer kolorowych CCD.
 W: Materiały VII Konferencji Naukowej "Czujniki Optoelektroniczne i Elektroniczne" COE 2002, 231–236, Rzeszów 2002.
- [24] W. S. Berris and S. Sangwine. Automatic quantitative analysis of healing skin wounds using colour digital image processing. World Wide Wounds, The Electronic Journal of Wound Management Practice, SMTL, Brigend, Wales, Sept. 1997.
- [25] D. Berry. Colour recognition using spectral signatures. Pattern Recognition Letters, 6(1):69-75, 1987.
- [26] N. L. Bihan and S. Sangwine. Quaternion principal component analysis of color images. In: Proceedings of IEEE International Conference on Image Processing, vol. 1, 809–812, Barcelona, Spain 2003.
- [27] R. V. den Boomgaard. Decomposition of the Kuwahara-Nagao operator in terms of linear smoothing and morphological sharpening. In: Proceedings of the 6th International Symposium on Mathematical Morphology, 283-292, Sydney, Australia 2002.
- [28] M. Borsotti, P. Campadelli and R. Schettini. Quantitative evaluation of color image segmentation results. *Pattern Recognition Letters*, 19(8):741–747, 1998.
- [29] R. Boynton. Human color perception. In: K. Leibovic, (ed.), Science of Vision, 211–253. Springer-Verlag, Berlin, Germany 1990.
- [30] C. Brice and C. Fennema. Scene analysis using regions. Artificial Intelligence, 1(3):205-226, 1970.
- [31] S. Brock-Gunn and T. Ellis. Using colour templates for target identification and tracking. In: Proceedings of BMVC'92, 207–216, Leeds, UK 1992.
- [32] D. Brockelbank and Y. Yang. An experimental investigation in the use of color in computational stereopsis. *IEEE Trans. on Systems, Man and Cybernetics*, 19(6):1365–1383, 1989.
- [33] L. Brun and A. Tremeau. Color quantization. In: G. Sharma, (ed.), Digital Color Imaging Handbook, 589–637. CRC Press, Boca Raton, USA 2003.
- [34] Cambridge Research & Instrumentation, CRI, Inc. Woburn, MA, USA, http://www.cri-inc.com.
- [35] D. Cavagnino and A. Werbrouck. Color image compression by means of separable Karhunen-Loeve transform and vector quantization. In: Proceedings of the 5th Int. Workshop on Systems, Signals and Image Processing, 235–238, Zagreb, Croatia 1998.
- [36] M. Celenk. Analysis of color images of natural scenes. Journal of Electronic Imaging, 4(4):382–396, 1995.

- [37] J. Chamorro-Martinez, D. Sanchez and B. Prados-Suarez. A fuzzy color image segmentation applied to robot vision. In: J. M. Benitez, O. Cordon, F. Hoffman and R. Roy, (eds.), Advances in Soft Computing, Engineering, Design and Manufacturing, 129–138, Springer-Verlag, Berlin, Germany 2003.
- [38] T. Chen and Y. Lu. Color image segmentation an innovative approach. Pattern Recognition, 35(2):395–405, 2002.
- [39] Y. Chen, P. Hao and A. Dang. Optimal transform in perceptually uniform color space and its application in image coding. In: *Proceedings of ICIAR 2004*, 269–276, Porto, Portugal 2004.
- [40] H. Cheng, X. Jiang, Y. Sun and J. Wang. Color image segmentation: advances and prospects. *Pattern Recognition*, 34(12):2259–2281, 2001.
- [41] H.-D. Cheng. A hierarchical approach to color image segmentation using homogeneity. IEEE Transactions on Image Processing, 9(12):2071-2082, 2000.
- [42] Y. Cheng. Mean shift, mode seeking and clustering. IEEE Transactions on Pattern Analysis and Machine Intelligence, 17(8):790-799, 1995.
- [43] CIE Supplement No.2 to CIE Publication No.15, Recommendations on uniform colour spaces, colour-difference equations, psychometric colour terms, Bureau Central de la CIE, Paris 1978.
- [44] CIE Publication 116-1995, Industrial colour-difference evaluation, CIE Central Bureau, Vienna 1995.
- [45] CIE TC1-34 Final Report, The CIE 1997 Interim Colour Appearance Model (Simple Version) CIECAM97s, 1998.
- [46] CIE Publication 141-2001, Improvement to industrial colour difference evaluation, CIE Central Bureau, Vienna 2001.
- [47] Ph. Colantoni. Couleur.org, Personal webpage about colour spaces, 2003, http://www.couleur.org.
- [48] C. Connolly. The relationship between colour metrics and the appearance of threedimensional coloured objects. Color Research and Application, 21(5):331-337, 1996.
- [49] C. Connolly and T. Fliess. A study of efficiency and accuracy in the transformation from RGB to CIELAB color space. *IEEE Transactions on Image Processing*, 6(7):1046–1048, 1997.
- [50] C. Connolly and H. Palus. Practical system considerations. In: S. Sangwine and R. Horne, (eds.), *The Colour Image Processing Handbook*, 129–146. Chapman and Hall, London, UK 1998.
- [51] V. Coutance, T. Baron and M. Briot. Segmentation d'images couleur in robotique. In: Proceedings of 7th Congres AFCET, 1115–1122, Paris, France 1989.
- [52] D. Crevier. Computing statistical properties of hue distributions for image analysis. In: Intelligent Robots and Computer Vision XII, Proceedings of SPIE, vol. 2055, 613–623, 1993.
- [53] W. Cudny and L. Chmielewski. Light wave length measurement with a colour camera. Machine Graphics and Vision, 2(3):251–260, 1993.
- [54] F. Cutzu, R. Hammoud and A. Leykin. Estimating the photorealism of images: Distinguishing paintings from photographs. In: Proceedings of International Conference on Computer Vision and Pattern Recognition (CVPR), vol. 2, 305-312, Madison, USA 2003.
- [55] Y. Dai and Y. Nakano. Face-texture model based on SGLD and its application in face detection in a color space. *Pattern Recognition*, 29(6):1007–1017, 1996.
- [56] C. Dalton. The measurement of the colorimetric fidelity of television cameras. Journal of the Institution of Electronic and Radio Engineers, 58(4):181–186, 1988.

- **196**
- [57] M. Das, E. Riseman and B. Draper. Focus: Searching for multi-colored objects in a diverse image database. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition '97, 756-761, San Juan, Puerto Rico 1997.
- [58] E. Davies. Principles and design graphs for obtaining uniform illumination in automated visual inspection. In: Proceedings of 6th IEE Conference on Image Processing and its Applications, 161–165, Dublin, Ireland 1997.
- [59] B. Deknuydt, J. Smolders, L. V. Eycken and A. Oosterlinck. Color space choice for nearly reversible image compression. In: Visual Communications and Image Processing, Proceedings of SPIE, vol. 1818, 1300-1311, 1992.
- [60] C.-H. Demarty and S. Beucher. Color segmentation using an HLS transformation. In: Proceedings of ISMM'98, 231–238, Amsterdam, The Netherlands 1998.
- [61] Y. Deng, C. Kenney, M. Moore and B. Manjunath. Peer group filtering and perceptual color image quantization. In: Proceedings of IEEE International Symposium on Circuits and Systems (ISCAS), vol. IV, 21–24, Orlando, USA 1999.
- [62] Y. Deng and B. Manjunath. Unsupervised segmentation of color-texture regions in images and video. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(8):800– 810, 2001.
- [63] A. De Rosa, A. M. Bonacchi, V. Cappellini and M. Barni. Image segmentation and region filling for virtual restoration of art-works. In: *Proceedings of IEEE Int. Conference on Image Processing (ICIP'01)*, vol. I, 562–565, Thessaloniki, Greece 2001.
- [64] Digital camera product reviews/previews, Minolta, Dimage, 2002, http://www.dpreview.com/reviews/minoltadimage7.
- [65] M. Domański. Zaawansowane techniki kompresji obrazów i sekwencji wizyjnych. Wydawnictwa Politechniki Poznańskiej, Poznań 2000.
- [66] M. Domański and M. Bartkowiak. Compression. In: S. Sangwine and R. Horne, (eds.), The Colour Image Processing Handbook, 242–304. Chapman and Hall, London, UK 1998.
- [67] M. Domański and K. Rakowski. Lossless and near-lossless image compression with color transformations. In: Proceedings of International Conference on Image Processing, vol. 3, 454–457, Thessaloniki, Greece 2001.
- [68] R. Dony and S. Wesolkowski. Edge detection on color images using RGB vector angle. In: Proc. IEEE Canadian Conference on Electrical and Computer Engineering (CCECE), 687–692, Edmonton, Canada 1999.
- [69] A. Empacher, Z. Sęp, A. Zakowska, W. Żakowski. Mały słownik matematyczny. Wiedza Powszechna, Warszawa 1970.
- [70] P. Engeldrum. Extending image quality models. In: Proceedings of PICS2002: IS&T PICS Conference, 65-69, Portland, USA 2002.
- [71] M. Fairchild. Color Appearance Models. Addison-Wesley, Reading, USA 1998.
- [72] J. Fan, D. Yau, A. Elmagarmid and W. Aref. Automatic image segmentation by integrating color-edge extraction and seeded region growing. *IEEE Transactions on Image Processing*, 10(10):1454-1466, 2001.
- [73] O. Faugeras. Digital color image processing within the framework of human visual model. IEEE Transactions on ASSP, 27(4):380-383, 1979.
- [74] E. Fedorovskaya, A. de Ridder and F. Blommaert. Chroma variations and perceived quality of color images of natural scenes. *Color Research and Applications*, 22(2):96–110, 1997.
- [75] F. Ferri and E. Vidal. Colour image segmentation and labeling through multiedit condensing. Pattern Recognition Letters, 13(8):561-568, 1992.

- [76] S. Fischer, P. Schmid and J. Guillod. Analysis of skin lesions with pigmented networks. In: Proceedings of 3rd IEEE International Conference on Image Processing, vol. 1, 323–326, Lausanne, Switzerland 1996.
- [77] M. Fleck, D. Forsyth and C. Bregler. Finding nacked people. In: Proceedings of 4th European Conference on Computer Vision, vol. 2, 592-602, Cambridge, UK 1996.
- [78] J. Foley and A. V. Dam. Fundamentals of Interactive Computer Graphics. Addison-Wesley, Reading, USA 1982.
- [79] D. Forsyth and J. Ponce. Computer Vision. Prentice Hall, Upper Saddle River, USA 2003.
- [80] Z. Fortuna, B. Macukow, J. Wąsowski. Metody numeryczne. WN-T, Warszawa 1998.
- [81] J. Freixenet, X. Munot, D. Raba, J. Marti and X. Cuti. Yet another survey on image segmentation: region and boundary information integration. In: A. Heyden et al., (eds.), ECCV 2002, pages 408-422. Springer-Verlag, Berlin, Germany 2002.
- [82] H. Frey. Digitale Bildverarbeitung in Farbräumen, Dr.-Ing. Dissertation. TU München, München, Germany 1988.
- [83] H. Frey and H. Palus. Sensor calibration for video-colorimetry. In: Proceedings of Workshop on Design Methodologies for Microelectronics and Signal Processing, 109–113, Gliwice-Krakow 1993.
- [84] J. Fridrich, R. Du and L. Meng. Steganalysis of LSB encoding in color images. In: Proceedings of the IEEE Int. Conf. on Multimedia and Expo, vol. 3, 1279–1282, New York, USA 2000.
- [85] G. Gagliardi, G. Hatch and N. Sarkar. Machine vision applications in the food industry. In: Proceedings of Vision'85 Conference, 524–538, Detroit, USA 1985.
- [86] H. Gao, W.-C. Siu and C.-H. Hou. Improved techniques for automatic image segmentation. IEEE Transactions on Circuits and Systems for Video Technology, 11(12):1273-1280, 2001.
- [87] C. Garcia and G. Tziritas. Face detection using quantized skin color regions merging and wavelet packet analysis. *IEEE Transactions on Multimedia*, 1(3):264–277, 1999.
- [88] J. Gauch. Segmentation and edge detection. In: S. Sangwine and R. Horne, (eds.), The Colour Image Processing Handbook, 163–187. Chapman and Hall, London, UK 1998.
- [89] J. Gauch and C. Hsia. A comparison of three segmentation algorithms in four color spaces. In: Proceedings of SPIE, vol. 1818, 1168–1175, 1992.
- [90] L. Gericke, R. Schumitz, O. Richter and K. Schöne. Farbenkatalog für die Gestaltung, I. Ergänzungsteil. Zentralinstitut für die Gestaltung des DAMW, Berlin, Germany 1978.
- [91] R. Gershon. Aspects of perception and computation in color vision. CVGIP, 32(2):244-277, 1985.
- [92] Y. Gong and M. Sakauchi. Detection of regions matching specified chromatic features. Computer Vision and Image Understanding, 61(2):263-269, 1995.
- [93] Y. Gong, H. Zhang and H. Chua. An image database system with content capturing and fast image indexing abilities. In: Proceedings of the IEEE Int. Conf. on Multimedia Computing and Systems (ICMCS'94), pages 121–130, Boston, USA 1994.
- [94] R. Gonzalez. Digital Image Processing. Addison-Wesley, Reading, USA 1977.
- [95] R. Gonzalez and R. Woods. Digital Image Processing. Addison-Wesley, Reading, USA 1992.
- [96] J. Gordillo. Colour representations for a vision machine. In: Proceedings of 2nd Int. Conf. on Machine Intelligence, 375–385, London, UK 1985.
- [97] P. Green and L. MacDonald. Colour Engineering, Achieving Device Independent Colour. John Wiley & Sons, Chichester, UK 2002.

- [98] I. Grinias, Y. Mavrikakis and G. Tziritas. Region growing colour image segmentation applied to face detection. In: Proc. of International Workshop Low Level Video Coding, Athens, Greece 2001.
- [99] A. Gunzinger, S. Mathis and W. Guggenbuehl. Real time color classification. In: V. Cappellini, (ed.), *Time-Varying Image Processing and Moving Object Recognition*, 82–87. Elsevier, Amsterdam, The Netherlands 1990.
- [100] R. Haralick and L. Shapiro. Glossary of computer vision terms. Pattern Recognition, 24(1):69-93, 1991.
- [101] R. Haralick and L. Shapiro. Computer and Robot Vision. Addison-Wesley, Reading, USA 1993.
- [102] J. Hardeberg. Desktop scanning to sRGB. In: Color Imaging: Device Independent Color, Color Hardcopy, and Graphic Arts V, Proceedings of SPIE, vol. 3963, 47–57, 2000.
- [103] J. Hardeberg, H. Brettel and F. Schmitt. Spectral characterisation of electronic cameras. In: Electronic Imaging: Processing, Printing and Publishing in Color, Proceedings of SPIE, vol. 3409, 100-109, 1998.
- [104] D. Hasler and S. Süsstrunk. Measuring colourfulness for natural images. In: Electronic Imaging 2003: Human Vision and Electronic Imaging VIII, Proceedings of SPIE, vol. 5007, 87-95, 2003.
- [105] G. Healey. Segmenting images using normalized color. IEEE Transactions on Systems, Man and Cybernetics, 22(1):64-73, 1992.
- [106] E. Hering. Outlines of a Theory of the Light Sense. Harvard University Press, Cambridge, USA 1964.
- [107] D. Hogg. Shape in machine vision. Image and Vision Computing, 11(6):309-316, 1993.
- [108] S. Hojjatoleslami and J. Kittler. Region growing: A new approach. IEEE Transactions on Image Processing, 7(7):1079-1084, 1998.
- [109] V. Hong, H. Palus and D. Paulus. Edge preserving filters on color images. Lecture Notes in Computer Science, 3039:35-42, 2004.
- [110] S. Horowitz and T. Pavlidis. Picture segmentation by a tree traversal algorithm. Journal of Association for Computing Machinery, 23(2):368-388, 1976.
- [111] M. Hu. Visual pattern recognition by moment invariants. IRE Transactions on Information Theory, 8(2):179–187, 1962.
- [112] K. Huang, Q. Wang and Z. Wu. Color image enhancement and evaluation algorithm based on human visual system. In: *Proceedings of ICASSP*'2004, vol. 3, 721–724, Montreal, Canada 2004.
- [113] Q. Huang, B. Dom, D. Steele, J. Ashley and W. Niblack. Foreground/background segmentation of color images by integration of multiple cues. In: Proceedings of IEEE Int. Conf. on Image Processing, vol. 1, 246-249, Washington, D.C., USA 1995.
- [114] R. Hunt. The specification of colour appearance. Color Research and Applications, 2:55–69 and 109–120, 1977.
- [115] IEC Publ. 61966-2-1: Multimedia systems and equipment Color measurement and management -- Part 2-1: Default RGB colour space — sRGB, 1999.
- [116] N. Ikonomakis, K. Plataniotis and A. Venetsanopoulos. Color image segmentation for multimedia applications. Journal of Intelligent and Robotic Systems, 28(1-2):5-20, 2000.
- [117] N. Ikonomakis, K. Plataniotis and A. Venetsanopoulos. Unsupervised seed determination for a region-based color image segmentation scheme. In: Proceedings of IEEE Int. Conference on Image Processing (ICIP'00), vol. I, 537–540, Vancouver, Canada 2000.
- [118] ISO/IEC 13818-2/ITU-T Rec.262, Generic coding of moving pictures and associated audio, Pt.2: Video, 2000.

- [119] A. Jain and R. Dubes. Algorithms for Clustering Data. Prentice Hall, Englewood Cliffs, USA 1988.
- [120] M. Jamzad, B. Sadjad, V. Mirrokni, M. Kazemi and et al. A fast vision system for middle size robots in Robocup. Lecture Notes in Computer Science, 2377:71-80, 2002.
- [121] X. Jie and S. Peng-Fei. Natural color image segmentation. In: Proc. of IEEE Int. Conf. on Image Processing, (ICIP'03), vol. I, 973-976, Barcelona, Spain 2003.
- [122] J. R. Jordan III, W. Geisler and A. Bovik. Color as a source of information in the stereo correspondence process. *Vision Research*, 30(12):1955–1970, 1990.
- [123] D. Judd and G. Wyszecki. Color in Business, Science, and Industry, 3rd ed. John Wiley & Sons, New York, USA 1975.
- [124] T. Kaczorek. Wektory i macierze w automatyce i elektrotechnice. WN-T, Warszawa 1998.
- [125] D. Karakos and P. Trahanias. Generalized multichannel image filtering structures. IEEE Transactions on Image Processing, 6(7):1038-1045, 1995.
- [126] A. Kasiński, R. Bączyk. Robust landmark recognition with application to navigation. In: M. Kurzyński, E. Puchała, M. Woźniak, (eds.), Computer Recognition Systems, 401–407. Wrocław University of Technology, Wrocław 2001.
- [127] G. Kay and G. D. Jager. A versatile colour system capable of fruit sorting and accurate object classification. In: Proceedings of the COMSIG'92, 145-148, Capetown, RSA 1992.
- [128] N. Kehtarnavaz, N. Griswold and D. Kang. Stop-sign recognition based on color/shape processing. Machine Vision and Applications, 6(4):206-208, 1993.
- [129] J. Kender. Saturation, hue, and normalized color: calculation, digitization effects, and use, Technical Report. Carnegie-Mellon University, Pittsburgh, USA 1976.
- [130] J.-Y. Kim, J.-C. Shim and Y.-H. Ha. Color image enhancement based on modified IHS coordinate system. In: Intelligent Robots and Computer Vision XI, Proceedings of SPIE, vol. 1825, 366-377, 1992.
- [131] S. Ko and Y. Lee. Center weighted median filters and their application to image enhancement. *IEEE Transactions on Circuits and Systems*, 38(9):984-993, 1991.
- [132] Kodak images, Resources of Center for Image Processing Research, Rensselaer Polytechnic Institute, Troy, NY, USA 2002, http://www.cipr.rpi.edu/resource/stills/kodak.html.
- [133] A. Koschan. Dense stereo correspondence using polychromatic block matching. In: Proc. of 5th International Conference on Computer Analysis of Images and Patterns, 538-542, Budapest, Hungary 1993.
- [134] R. Kuehni. Color Space and Its Divisions, Color Order from Antiquity to the Present. John Wiley & Sons, Hoboken, USA 2003.
- [135] M. Kuwahara, K. Hachimura, S. Eiho and M. Kinoshita. Processing of riangiocardiographic images. In: K. Preston and M. Onoe, (eds.), Digital Processing of Biomedical Images, 187–202. Plenum Press, New York, USA 1976.
- [136] S. Ledley, M. Buas and T. Golab. Fundamentals of true-color image processing. In: Proceedings of 10th Int. Conf. on Pattern Recognition, 791–795, Atlantic City, USA 1990.
- [137] H. C. Lee. Method for computing the scene-illuminant chromaticity from specular highlights. Journal of Optical Society of America A, vol.3, no.10, 1694-1699, 1986.
- [138] R. Lee. Colorimetric calibration of a video digitizing system: algorithm and applications. Color Research and Applications, 13(3):180–186, 1988.
- [139] H. Levkowitz and G. T Herman. GLHS: a generalized lightness, hue and saturation color model. CVGIP: Graphical Model and Image Processing, vol. 55, no.4, 271-285, 1993.

- [140] M. Li, I. Sethi, D. Li and N. Dimitrova. Region growing using online learning. In: Proceedings of International Conference on Imaging Science, Systems, and Technology, (CISST'03), vol. 1, 73-76, Las Vegas, USA 2003.
- [141] Y. Lim and S. Lee. On the color image segmentation algorithm based on the thresholding and the fuzzy c-means techniques. *Pattern Recognition*, 23(9):935–952, 1990.
- [142] X. Lin and S. Chen. Color image segmentation using a modified HSI system for road following. In: Proceedings of the IEEE International Conference on Robotics and Automation, vol. 3, 1998–2003, Sacramento, USA 1991.
- [143] B. J. Lindbloom. Personal webpage, 2001, http://www.brucelindbloom.com.
- [144] J. Liu and Y.-H. Yang. Multiresolution color image segmentation IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 16, no.7, 689–700, 1994.
- [145] L. Lucchese and S. Mitra. Advances in color image segmentation. In: Proceedings of Globecom'99, vol. IV, 2038–2044, Rio de Janeiro, Brazil 1999.
- [146] Machine vision illumination products, Siemens NERLITE, Weare, NH, USA, http://www.nerlite.com.
- [147] J. Mac Queen. Some methods for classification and analysis of multivariate observations. In: Proceedings of the 5th Berkeley Symposium on Mathematics, Statistics, and Probabilities, vol. I, 281-297, Berkeley and Los Angeles, USA 1967.
- [148] M. Mahy, L. V. Eyckden and A. Oosterlinck. Evaluation of uniform color spaces developed after the adoption of CIELAB and CIELUV. *Color Research and Application*, 19(2):105– 121, 1994.
- [149] S. Makrogiannis, G. Economou and S. Fotopoulos. A graph theory approach for automatic segmentation of color images. In: Proc. of International Workshop Low Level Video Coding, 162-166, Athens, Greece 2001.
- [150] W. Malina and M. Smiatacz. Metody cyfrowego przetwarzania obrazów. Akademicka Oficyna Wydawnicza EXIT, Warszawa 2005.
- [151] G. Marcu and S. Abe. Blue-print document analysis for color classification. In: Proceedings of 9th Scandinavian Conference on Image Analysis, 569–574, Uppsala, Sweden 1995.
- [152] K. Mardia and P. Jupp. Directional statistics. John Wiley & Sons, Chichester, UK 2000.
- [153] S. Marshall. Review of shape coding techniques. Image and Vision Computing, 7(4):281– 294, 1989.
- [154] E. Marszalec. Database Agfa IT8.7/2 set. University of Joensuu, Finlandia, http://cs.joensuu.fi/~spectral/databases/download/agfait872.htm.
- [155] E. Marszalec and M. Pietikäinen. On-line color camera calibration. In: Proceedings of 12th International Conference of Pattern Recognition, vol. 1, 232–237, Jerusalem, Israel 1994.
- [156] I. Masaki. Real-time multi-spectral visual processor. In: Proc. 1988 IEEE Int. Conf. on Robotics and Automation, vol. 3, 1554–1559, Philadelphia, USA 1988.
- [157] R. Massen. Color monitoring in the production lines: More than alternative to classical non-imaging colorimetry. In: *Polarization and Color Techniques in Industrial Inspection*, *Proceedings of SPIE*, vol. 3826, 36–43, 1999.
- [158] R. Massen, P. Böttcher and U. Leisinger. Real-time grey level and colour image preprocessing for a vision guided biotechnology robot. In: *Proceedings of 7th RoViSeC*, 115–122, Zurich, Switzerland 1988.
- [159] C. McCamy, H. Marcus and J. Davidson. A color rendition chart. Journal of Applied Photographic Engineering, 2(3):95–99, 1976.
- [160] A. Mehnert and P. Jackway. An improved seeded region growing algorithm. Pattern Recognition Letters, 18(10):1065–1071, 1997.

- [161] F. Meyer. Color image segmentation. In: Proc. IEE Int. Conf. Image Processing and its Applications, 303-306, Maastricht, The Netherlands 1992.
- [162] O. Milvang and B. Olafsdottir. Discriminating crates from color images. In: Proceedings of 8th Scandinavian Conference on Image Analysis, vol. 1, 659–669, Tromso, Norway 1993.
- [163] D. Mital, G. Lee and T. Khwang. Color vision for industrial applications. In: Proceedings of 16th Annual Conference of IEEE Industrial Electronics Society, 548–551, Pacific Grove, USA 1990.
- [164] D. Monro and J. Nichols. Low bit rate color fractal video. In: Proceedings of IEEE Int. Conf. on Image Processing, vol. 3, 264–267, Washington, D.C., USA 1995.
- [165] R. Moorhead and Z. Zhu. Signal processing aspects of scientific visualization. IEEE Signal Processing Magazine, 12(5):20–41, 1995.
- [166] N. Moroney, M. Fairchild, R. Hunt, C. Li, M. Luo and T. Newman. The CIECAM02 color appearance model. In: *Proceedings of IS&T/SID 10th Color Imaging Conference*, 23–27, Scottsdale, USA 2002.
- [167] V. Müller. Elimination of specular surface-reflectance using polarized and unpolarized light. In: Proceedings of the 4th European Conference on Computer Vision, vol. 2, 625– 635, Cambridge, UK 1996.
- [168] M. Nagao and T. Matsuyama. Edge preserving smoothing. Computer Graphics and Image Processing, 9(4):374–407, 1979.
- [169] E. Navon, O. Miller and A. Averbuch. Color image segmentation based on adaptive local thresholds. *Image and Vision Computing*, 23(1):69-85, 2005.
- [170] R. Nevatia. A color edge detector and its use in scene segmentation. IEEE Trans. Systems, Man, and Cybernetics, 7(11):820-826, 1977.
- [171] M. Nieniewski. Morfologia matematyczna w przetwarzaniu obrazów. Akademicka Oficyna Wydawnicza PLJ, Warszawa 1998.
- [172] Y. Ohta, T. Kanade and T. Sakai. Color information for region segmentation. Computer Graphics and Image Processing, 13(3):222-241, 1980.
- [173] S. Ong and C. Hew. Segmentation of colour images based on iterative thresholding and merging. In: Proceedings of 2nd Int. Conf. on Image Processing, 721-725, Singapore 1992.
- [174] N. Otsu. A threshold selection method from gray-level histogram. IEEE Transactions on Systems, Man, and Cybernetics, 9(1):62-66, 1979.
- [175] N. Ouerhani, N. Archip, H. Huegli and P. Erard. Visual attention guided seed selection for color image segmentation. In: Proceedings of the 9th International Conference on Computer Analysis of Images and Patterns (CAIP'01), 630–637, Warsaw, Poland 2001.
- [176] L. Overturf, M. Comer and E. Delp. Color image coding using morphological pyramid decomposition. *IEEE Trans. on Image Processing*, 4(2):177–185, 1995.
- [177] N. Pal and S. Pal. Unsupervised segmentation of color-texture regions in images and video. IEEE Transactions on Pattern Analysis and Machine Intelligence, 26(9):1277–1293, 1993.
- [178] H. Palus. Model i badania barwoczułego systemu sensorycznego. Zeszyty Naukowe Politechniki Śląskiej, s. Automatyka, 101:193–204, Gliwice, 1990.
- [179] H. Palus. Cvetnoje mashinnoje zrenie. In: J. Kulikowski and J. Juravlev, (eds.), Lecture Notes of ICB Seminar, 205–215. International Center of Biocybernetics, Warsaw 1991.
- [180] H. Palus. Colour spaces in computer vision. Machine Graphics and Vision, 1(3):543-554, 1992.
- [181] H. Palus. Regułowe systemy wizyjne. Zeszyty Naukowe Politechniki Śląskiej, s. Automatyka, 110:279–289, Gliwice 1992.
- [182] H. Palus. Systemy barwnej wizji komputerowej dla potrzeb robotyki. Prace Naukowe ICT Pol. Wrocl., 94:132–139, 1993.

- BIBLIOGRAFIA
- [183] H. Palus. Barwne stereowidzenie: przegląd metod i analiza możliwości. Prace IPI PAN, 747:1–19, 1994.
- [184] H. Palus. Kalibracja systemu barwnej wizji komputerowej, Raport z pracy BW-1/Rau1/94
 t.15. Instytut Automatyki Politechniki Śląskiej, Gliwice 1994. (praca niepublikowana).
- [185] H. Palus. Systemy wizji komputerowej w górnictwie? Mechanizacja i Automatyzacja Górnictwa, 287–288(5–6):114–116, 1994.
- [186] H. Palus. Wyznaczanie współczynnika kompaktowości w systemach wizyjnej identyfikacji obiektów. Zeszyty Naukowe Politechniki Śląskiej, s. Automatyka, 113:259–270, Gliwice 1994.
- [187] H. Palus. Barwa w rozpoznawaniu obiektów przez system wizyjny robota przemysłowego. Zeszyty Naukowe Politechniki Śląskiej, s. Automatyka, 116:117-127, 1995.
- [188] H. Palus. Farbe in der Objekterkennung: Versuch der Systematisierung und ein Beispiel. Fachberichte Informatik, Universität Koblenz-Landau, (15):17-20, 1995.
- [189] H. Palus. Der IHS Farbraum: Eigenschaften und Modifikationen. Schriftenreihe des ZBS, Ilmenau, (1):73-77, 1996.
- [190] H. Palus. Przestrzeń barw IHS w zastosowaniu do rozpoznawania obiektów. Zeszyty Naukowe Politechniki Śląskiej, s. Automatyka, 119:181-191, Gliwice 1996.
- [191] H. Palus. Nutzung des Wissens über die Farbe des Bildhintergrundes zur Segmentierung von Farbbildern. In: D. Paulus and T. Wagner, (eds.), Dritter Workshop Farbbildverarbeitung, 39–43. IRB-Verlag, Stuttgart, Germany 1997.
- [192] H. Palus. Homogeneity criteria for region-growing image segmentation in IHS colour space. In: Proceedings of 5th Int. Workshop on Systems, Signals and Image Processing, 227–230, Zagreb, Croatia 1998.
- [193] H. Palus. Representations of colour images in different colour spaces. In: S. J. Sangwine and R. E. N. Horne, (eds.), *The Colour Image Processing Handbook*, 67–90. Chapman and Hall, London, UK 1998.
- [194] H. Palus. Wykorzystanie odblasków w obrazach barwnych do zliczania obiektów w scenie. Zeszyty Naukowe Politechniki Śląskiej, s. Automatyka, 125:191–199, 1998.
- [195] H. Palus. Counting of colored objects using highlights. In: Polarization and Color Techniques in Industrial Inspection, Proceedings of SPIE, vol. 3826, 44-51, 1999.
- [196] H. Palus. Remarks on the using colour in computer vision. In: Proceedings of AIC Midterm Meeting, 327–334, Warsaw 1999.
- [197] H. Palus. Segmentacja obrazów barwnych techniką podziału i łączenia. W: Materiały International Conference on Management Systems, 51-59, Bielsko-Biała 1999.
- [198] H. Palus. Barwa w obrazach cyfrowych. W: Materiały Seminarium "Edukacyjne Systemy Internetowe", 99–105, Bielsko-Biała 2001.
- [199] H. Palus. Latest results in color image processing and its applications. Machine Graphics and Vision, 11(2/3):135–137, 2002.
- [200] H. Palus. Region-based colour image segmentation: control parameters and evaluation functions. In: Proceedings of the First European Conference on Color in Graphics, Imaging and Vision, (CGIV'02), 259–262, Poitiers, France 2002.
- [201] H. Palus. Estimating the usefulness of preprocessing in colour image segmentation. In: Proc. of 2nd European Conference on Colour in Graphics, Imaging, and Vision (CGIV2004), 197–200, Aachen, Germany 2004.
- [202] H. Palus. Application of colourfulness of the image in colour image quantization. In: Computer Methods and Systems, Proceedings of Conference, vol. 2, 205–210, Kraków 2005.

- [203] H. Palus. Colourfulness of the image and its application in image filtering. In: Proceedings of 5th IEEE International Symposium on Signal Processing and Information Technology (ISSPIT'2005), 884–889, Athens, Greece 2005.
- [204] H. Palus. Performance evaluation of preprocessing in colour image segmentation. Journal of Imaging Science and Technology, 49(6):583-587, 2005.
- [205] H. Palus. Region-based colour image segmentation technique and its properties. Przegląd Elektrotechniczny, (3):41-44, 2005.
- [206] H. Palus. Region growing technique for colour image segmentation. In: Computer Methods and Systems, Proceedings of Conference, vol. 2, 199-204, Kraków 2005.
- [207] H. Palus. Color image segmentation: selected techniques. In: R. Lukac and K. Plataniotis, (eds.), Color Image Processing: Methods and Applications, 103–128. CRC Press, Boca Raton, USA 2006.
- [208] H. Palus. Colorfulness of the image: definition, computation and properties. In: Proceedings of SPIE, vol. 6158, 05–1–05–6, 2006.
- [209] H. Palus and D. Bereska. The comparison between transformations from RGB colour space to IHS colour space, used for object recognition. In: Proceedings of the 5th Int. Conf. on Image Processing and Its Applications, 825–827, Edinburgh, UK 1995.
- [210] H. Palus, D. Bereska. Segmentacja obrazów w przestrzeni IHS: problem przenoszenia informacji kształtowej. Prace IPI PAN, 772:1-19, 1995.
- [211] H. Palus and D. Bereska. IHS colour space for use in object recognition. In: Proceedings of the 4th International Symposium on Methods and Models in Automation and Robotics, vol. 3, 979–984, Międzyzdroje 1997.
- [212] H. Palus and D. Bereska. Single-coloured object recognition in scenes with a known background. In: Proceedings of the 4th Int. Workshop on Systems, Signals and Image Processing, 145-148, Poznań 1997.
- [213] H. Palus and D. Bereska. Colour image segmentation using region growing in IHS colour space. In: Proceedings of the 6th Int. Workshop on Systems, Signals and Image Processing, 53-57, Bratislava, Slovakia 1999.
- [214] H. Palus and D. Bereska. Region-based colour image segmentation. In: Proc. of 5th Workshop "Farbbildverarbeitung", 67–74, Ilmenau, Germany 1999.
- [215] H. Palus, D. Bereska. Wierna prezentacja barwnych obrazów cyfrowych w Internecie. W: L. Kiełtyka, (ed.), Multimedia w biznesie, 353–364. Kantor Wydawniczy Zakamycze, Kraków 2003.
- [216] H. Palus and D. Bereska. On color image quantization by the k-means algorithm. In: D. Droege and D. Paulus, (eds.), 10. Workshop Farbbildverarbeitung, 58-65. Der Andere Verlag, Toenning, Germany 2004.
- [217] H. Palus and D. Bereska. Wierność odwzorowania barw w systemach wizyjnych. In: R. Gessing, T. Szkodny, (eds.), Automatyzacja procesów dyskretnych, Optymalizacja dyskretna, Robotyka i sterowniki programowalne, 285–294. WN-T, Warszawa 2004.
- [218] H. Palus and D. Bereska. Colour reproduction accuracy in vision systems. In: K. Wojciechowski, B. Smołka, H. Palus and et al., (eds.), Computer Vision and Graphics, 279–286. Springer, Dordrecht, The Netherlands 2006.
- [219] H. Palus, D. Bereska, S. Grela. Transformacja Karhunena-Loevego dla obrazów barwnych. Zeszyty Naukowe Politechniki Śląskiej, s. Automatyka, 131:185–195, Gliwice 2000.
- [220] H. Palus and M. Bogdański. Clustering techniques in colour image segmentation. In: Proceedings of AIMETH International Conference, 103–104, Gliwice 2003.
- [221] H. Palus and T. Kotyczka. Evaluation of colour image segmentation results. Arbeitsberichte des Instituts fuer Informatik Friedrich-Alexander-Universitaet Erlangen-Nuernberg, 34(15):41-46, 2001.

- [222] N. Papamarkos, A. Atsalakis and C. Strouthopoulos. Adaptive color reduction. IEEE Trans. on SMC - Part B: Cybernetics, 32(1):44-56, 2002.
- [223] H. Park and J. Ra. Homogeneous region merging approach for image segmentation preserving semantic object contours. In: Proc. International Workshop on Very Low Bitrate Video Coding, 149–152, Chicago, USA 1998.
- [224] D. Paulus, J. Hornegger and L. Csink. Linear approximation of sensitivity curve calibration. Schriftenreihe des ZBS Ilmenau, (1):3-10, 2002.
- [225] T. Pavlidis. Structural Pattern Recognition. Springer, New York, USA 1977.
- [226] S.-C. Pei and C.-M. Cheng. A novel block truncation coding of color images by using quaternion-moment-preserving principle. In: Proceedings of IEEE Int. Symp. on Circuits and Systems, vol. 2, 684–687, Atlanta, USA 1996.
- [227] E. Peli. Contrast in complex images. Journal of Optical Society of America A, 7(10):2032– 2040, 1990.
- [228] F. Perez and C. Koch. Toward color image segmentation in analog VLSI: algorithm and hardware. International Journal of Computer Vision, 12(1):17-42, 1994.
- [229] B. Phong. Illumination for computer generated pictures. Comm. ACM, 18(6):311-317, 1975.
- [230] M. Pietikainen and D. Harwood. Segmentation of color images using edge-preserving filters. In: V. Cappellini and R. Marconi, (eds.), Advances in Image Processing and Pattern Recognition, 94–99. Elsevier, Amsterdam, The Netherlands 1986.
- [231] K. Plataniotis and A. Venetsanopoulos. Color Image Processing and Applications. Springer-Verlag, Berlin, Germany 2000.
- [232] T. Pomierski and H.-M. Gross. Biological neural architectures for chromatic adaptation resulting in constant color sensations. In: *Proceedings of IEEE Int. Conf. on Neural Networks ICNN'96*, vol. 2, 734–739, Washington, D.C., USA 1996.
- [233] W. Pratt. Spatial transform coding of color images. IEEE Transactions on Communication Technology, 19(6):980-992, 1971.
- [234] W. Pratt. Digital Image Processing. John Wiley & Sons, New York, USA 1978.
- [235] W. Pratt. Digital Image Processing, 2nd ed. John Wiley & Sons, New York, USA 1991.
- [236] L. Priese and V. Rehrmann. A fast hybrid color segmentation method. In: S. Poppl and H. Handels, (Hrsg.), *Mustererkennung 1993*, 297–304. Springer-Verlag, Berlin, Germany 1993.
- [237] A. Pritchard. Object characterisation and image segmentation using a modified HSI colour space, Ph.D. Thesis. University of Reading, Reading, UK 1995.
- [238] P. Pujas and M.-J. Aldon. Robust colour image segmentation. In: Proceedings of the 7th Int. Conf. on Advanced Robotics (ICAR), vol. 1, 145–155, San Feliu de Guixols, Spain 1995.
- [239] P. Ranefall, B. Nordin and E. Bengtsson. Finding facial features using an HLS colour space. In: C. Braccini, L. DeFloriani and G. Vernazza, (eds.), *Image Analysis and Processing*, 191–196. Springer-Verlag, Berlin, Germany 1995.
- [240] Recommendation BT.601, Encoding parameters of digital television for studios, ITU, Geneva 1994.
- [241] P. Reitan. Hybrid approaches to color image quantization, Ph.D. Thesis. University of Maryland, Baltimore, USA 1999.
- [242] C. Ridder, O. Munkelt and H. Kirchner. Adaptive background estimation and foreground detection using Kalman-filtering. In: *Proceedings of ICRAM'95*, 193–195, Istanbul, Turkey 1995.
- [243] P. Robertson. Visualizing color gamuts: A user interface for the effective use of perceptual color spaces in data displays. *IEEE Comp. Graph. and Appl.*, 8(5):50-64, 1988.

- [244] A. Rodriguez and O. Mitchel. Image segmentation by succesive background extraction. Pattern Recognition, 24(5):409-420, 1991.
- [245] D. Rogers and R. Earnshaw. Techniques for Computer Graphics. Springer-Verlag, Berlin, Germany 1986.
- [246] G. Sandini, F. Buemi, M. Massa and M. Zucchini. Visually guided operations in greenhouses. In: Proceedings of IEEE International Workshop on Intelligent Robots and Systems, IROS'90, vol. 1, 279-285, New York, USA 1990.
- [247] S. Sangwine. Fourier transforms of colour images using quaternions, or hypercomplex numbers. *Electronics Letters*, 32(21):1979–1980, 1996.
- [248] S. Sangwine and R. Horne. The Colour Image Processing Handbook. Chapman and Hall, London, UK 1998.
- [249] M. Sarifuddin and R. Missaoui. A new perceptually uniform color space with associated color similarity measure for content-based image and video retrieval. In: Proceedings of ACM SIGIR Workshop on Multimedia Information Retrieval, Salvador, Brazil 2005.
- [250] C. Scheering and A. Knoll. Fast colour image segmentation using pre-clustered chromaticity-plane. In: Proceedings of ICASSP-97, vol. 4, 3145–3147, Munich, Germany 1997.
- [251] R. Schettini. A segmentation algorithm for color images. Pattern Recognition Letters, 14(6):499-506, 1993.
- [252] R. Schettini. Multicolored object recognition and location. Pattern Recognition Letters, 15(11):1089-1097, 1994.
- [253] T. Schindewolf, R. Albert, W. Stolz, W. Abmayr and H. Harms. Klassifikation melanozytärer Hautveränderungen anhand makroskopischer Farbaufnahmen. In: S. J. Poppl, H. Handels, (Hrsg.), Mustererkennung 1993, 436-443, Springer-Verlag, Berlin 1993.
- [254] S. Selim and M. Ismail. K-means-type algorithms. *IEEE Transactions on Pattern Analysis* and Machine Intelligence, 6(1):81-87, 1984.
- [255] S. Shafer. Using color to separate reflection components. Color Research and Application, 10(4):210-218, 1985.
- [256] L. Shapiro and G. Stockman. Computer Vision. Prentice Hall, Upper Saddle River, USA 2003.
- [257] G. Sharma, (ed.). Digital Color Imaging Handbook, CRC Press, Boca Raton, FL, USA 2003.
- [258] G. Sharma and H. J. Trussell, (eds.). Special issue on colour imaging. IEEE Transactions on Image Processing, 6(7):901-932, 1997.
- [259] G. Sharma, W. Wu and E. Dalal. The CIEDE2000 color-difference formula: implementation notes, supplementary test data, and mathematical observations. *Color Research and Application*, 30(1):21–30, 2005.
- [260] D. Sinclair. Voronoi seeded colour image segmentation. Technical Report TR99-4, AT&T Laboratories, Cambridge, UK 1999.
- [261] W. Skarbek. Metody reprezentacji obrazów cyfrowych. Akademicka Oficyna Wydawnicza PLJ, Warszawa 1993.
- [262] W. Skarbek and A. Koschan. Colour image segmentation a survey. Technical Report 94–32, Tech. Univ. of Berlin, Berlin, Germany, October 1994.
- [263] J. Slater. Modern Televisions Systems. Pitman, London, UK 1991.
- [264] D. Slaughter and R. Harrell. Color vision in robotic fruit harvesting. Trans. of the ASAE, 30(4):1144-1148, 1987.

- [265] P. D. Smet, R. Pires and D. D. Vleeschauwer. The activity image in image enhancement and segmentation. In: Proceedings of IEEE Benelux Signal Processing Symposium, 79–82, Leuven, Belgium 1998.
- [266] J. Smith and S. Chang. Visualseek: a fully automated content-based image query system. In: Proceedings of ACM Multimedia Conference, 87–98, Boston, USA 1996.
- [267] B. Smołka. Nonlinear Techniques of Noise Reduction in Digital Color Images. Wydawnictwo Politechniki Śląskiej, Gliwice 2004.
- [268] B. Smołka, M. Szczepański, A. Świerniak, H. Palus and K. Plataniotis. On the segmentation of the comet assay images. Journal of Medical Informatics and Technologies, 2:29–39, 2001.
- [269] J. Solinsky. The use of color in machine edge detection. In: Proceedings of Vision'85, 4-34-4-52, Detroit, USA 1985.
- [270] K. Stąpor. Automatyczna klasyfikacja obiektów. Akademicka Oficyna Wydawnicza EXIT, Warszawa 2005.
- [271] N. Strachan, P. Nesvadba and A. Allen. Calibration of a video camera digitising system in the CIE L*U*V* colour space. *Pattern Recognition Letters*, 11(11):771-777, 1990.
- [272] Y.-N. Sun, C.-S. Wu, X.-Z. Lin and N.-H. Chou. Color image analysis for liver tissue classification. Optical Engineering, 32(7):1609–1615, 1993.
- [273] S. Süsstrunk, R. Buckley and S. Swen. Standard RGB color spaces. In: Proceedings of 7th Color Imaging Conference: Color Science, Systems, and Applications, 127–134, Scotts Valley, USA 1999.
- [274] S. Süsstrunk and S. Winkler. Color image quality on the Internet. In: Electronic Imaging 2004, Internet Imaging V, Proceedings of SPIE, vol. 5304, 118–131, 2004.
- [275] M. Swain. Color Indexing, Ph.D. Thesis. University of Rochester, Rochester, USA 1990.
- [276] M. Swain and D. Ballard. Color indexing. International Journal of Computer Vision, 7(1):11-32, 1991.
- [277] M. Szczepański, B. Smołka, D. Ślusarczyk, K. Plataniotis and A. Venetsanopoulos. Geodesic paths approach to color image enhancement. *Electronic Notes in Theoretical Computer Science*, 46, 2001.
- [278] J. Śmieja, H. Palus. Zastosowanie cech topologicznych w rozpoznawaniu obiektów. Zeszyty Naukowe Politechniki Śląskiej, s. Automatyka, 115:125–135, Gliwice 1994.
- [279] R. Tadeusiewicz. Systemy wizyjne robotów przemyslowych. WN-T, Warszawa 1992.
- [280] R. Tadeusiewicz, P. Korohoda. Komputerowa analiza i przetwarzanie obrazów. Wydawnictwo Fundacji Postępu Telekomunikacji, Kraków 1997.
- [281] R. Taylor and P. Lewis. Colour image segmentation using boundary relaxation. In: Proceedings of 11th IAPR Int. Conf. on Pattern Recognition, vol. 3, 721–724, The Hague, The Netherlands 1992.
- [282] J. M. Tenenbaum, T. D. Garvey, S. Weyl, and H. C. Wolf. An interactive facility for scene analysis research, Technical Report TN 87, SRI International, Menlo Park, California, USA 1974.
- [283] P. Thrift and C. Lee. Using highlights to constrain object size and location. IEEE Transactions on Systems, Man, and Cybernetics, 13(3):426-443, 1983.
- [284] A. Toet. Multiscale colour image enhancement. Pattern Recognition Letters, 13(3):164– 174, 1992.
- [285] S. Tominaga. Color classification of natural color images. Color Research and Application, 17(4):230–239, 1992.
- [286] F. Torres, J. Angulo, and F. Ortiz. Automatic detection of specular reflectance in colour images using the MS diagram. Lecture Notes in Computer Science, 2756:132–139, 2003.

- [287] A. Tremeau and P. Colantoni. Region adjacency graph applied to color image segmentation,. IEEE Transactions on Image Processing, 9(4):735-744, 2000.
- [288] A. Tremeau, V. Lozano, and B. Lager. How to optimize the use of the L*H*C* color space in color image analysis processes. Acta Stereol., 14(2):223-228, 1995.
- [289] R. Turi and S. Ray. Clustering-based colour image segmentation. In: Proceedings of Australia Pattern Recognition Society Student Conference, 76–88, Melbourne, Australia 1996.
- [290] R. Turi and S. Ray. An application of clustering in colour image segmentation. In: Proceedings of 6th International Conference on Control, Automation, Robotics and Vision (ICARCV'00), Singapore 2000.
- [291] T. Uchiyama and M. Arbib. Color image segmentation using competitive learning. IEEE Transactions on Pattern Analysis and Machine Intelligence, 16(12):1197-1206, 1994.
- [292] S. Umbaugh, R. Moss and W. Stoecker. Automatic color segmentation of image with application to detection of variegated coloring in skin tumors. *IEEE Engineering in Medicine* and Biology Magazine, 8(4):43-52, 1989.
- [293] R. D. Valois and K. D. Valois. A multi-stage color model. Vision Research, 33(8):1053– 1065, 1993.
- [294] N. Vandenbroucke, L. Macaire and L.-G. Postaire. Color pixels classification in an hybrid color space. In: Proceedings of International Conference on Image Processing, vol. 1, 176–180, Chicago, USA 1998.
- [295] A. D. Ventura and R. Schettini. Computer-aided color coding for data display. In: Proceedings of 11th IAPR Int. Conf. on Pattern Recognition, vol. 3, 29–32, The Hague, The Netherlands 1992.
- [296] A. Watson and C. Tiana. Color motion video coded by perceptual components. SID Digest of Technical Papers, (23):314–317, 1992.
- [297] S. Westland and C. Ripamonti. Computational Colour Science Using Matlab. John Wiley & Sons, Chichester, UK 2004.
- [298] S. Winkler. Visual fidelity and perceived quality: towards comprehensive metrics. In: Human Vision and Electronic Imaging VI, Proceedings of SPIE, vol. 4299, 114–125, 2001.
- [299] L. Wixson and D. Ballard. Real-time detection of multicolored objects. In: Proceedings of SPIE, vol. 1198, 435–446, 1989.
- [300] S. Wolf, R. Ginosar and Y. Zeevi. Spatio-chromatic image enhancement based on a model of human visual information processing. *Journal of Visual Communication and Image Representation*, 9(1):25–37, 1998.
- [301] G. Wyszecki and W. Stiles. Color Science: Concepts and Methods, Quantitative Data and Formulae. John Wiley & Sons, New York, USA 1982.
- [302] M. Yachida and S. Tsuji. Application of colour information to visual perception. Pattern Recognition, 3(3):307-323, 1971.
- [303] D. Yagi, K. Abe and H. Nakatani. Segmentation of color aerial photographs using HSV color models. In: Proc. IAPR Workshop on Machine Vision Applications, (MVA'92), 367-370, Tokyo, Japan 1992.
- [304] K. Yamaba and Y. Miyake. Color character recognition method based on human perception. Optical Engineering, 32(1):33-40, 1993.
- [305] C. Yang and J. Rodriguez. Saturation clipping in the LHS and YIQ color spaces. In: Proceedings of IS&T/SPIE Int. Symp. on Electronic Imaging, 297–307, San JOse, USA 1996.
- [306] S. Yendrikhovskij, F. Blommaert and H. de Ridder. Optimizing colour reproduction of natural images. In: Proceedings of IS&T/SID 6th Color Imaging Conference, 140–145, Scottsdale, USA 1998.

[307] J. Zabrodzki. Grafika komputerowa, Metody i narzędzia. WNT, Warszawa 1994.

[308] E. Zarakov. Direct image sensor tackles color concerns. Photonics Spectra, 37(11):99–100, 2003.

- [309] L. Zhang, F. Lin and B. Zhang. A CBIR method based on color-spatial feature. In: Proceedings of the IEEE Region 10th Ann. Int. Conf. TENCON'99, 166–169, Cheju, Korea 1999.
- [310] X. Zhang and B. Wandell. A spatial extension to CIELAB for digital color image reproduction. Journal of the Society for Information Display, 5(1):61-63, 1997.
- [311] Y. Zhang. A survey on evaluation methods for image segmentation. Pattern Recognition, 29(8):1335-1346, 1996.
- [312] J. Zheng, K. Valavanis and J. Gauch. Noise removal from color images. Journal of Intelligent and Robotic Systems, 7(3):257-285, 1993.
- [313] S. Zhu and A. Yuille. Region competition: unifying snakes, region growing and Bayes/MDL for multiband image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(9):884-900, 1996.
- [314] K. Zieliński, M. Strzelecki. Komputerowa analiza obrazu medycznego, Wstęp do morfometrii i patologii ilościowej. PWN, Warszawa 2001.

DODATEK

 a'_i C'_i

Procedura obliczania różnicy barw według formuły CIEDE2000

Proces wyznaczania różnicy barw składa się tutaj z 3 etapów:

1. Obliczanie zmodyfikowanej chromy i odcienia C'_i , h'_i :

$$C_{i,ab}^{*} = \sqrt{a_{i}^{*2} + b_{i}^{*2}} \qquad \text{dla } i = 1, 2 \qquad (1)$$

$$\bar{C}_{ab}^{*} = \frac{C_{1,ab}^{*} + C_{2,ab}^{*}}{2} \qquad (2)$$

$$G = 0.5 \left(1 - \sqrt{\frac{\bar{C}_{ab}}{\bar{C}_{ab}^{*}}^{7} + 25^{7}} \right)$$
(3)

$$= (1+G)a_i^* \qquad \qquad \text{dla } i = 1,2 \quad (4)$$

$$=\sqrt{a_i'^2+b_i^{*2}}$$
 dla $i=1,2$ (5)

 $h'_{i} = \begin{cases} 0 & \text{gdy } b^{*}_{i} = a'_{i} = 0\\ \operatorname{arctg}(b^{*}_{i}, a'_{i}) & \text{w przeciwnym przypadku} \end{cases} \quad \text{dla } i = 1, 2 \quad (6)$

2. Obliczanie różnic $\Delta L'$, $\Delta C'$, $\Delta H'$; w przeciwieństwie do poprzednich wzorów różnice wyznaczane tutaj są z dokładnością co do znaku:

$$\Delta L' = L_2^* - L_1^* \tag{7}$$

$$\Delta C' = C'_2 - C'_1 \tag{8}$$

$$\Delta h' = \begin{cases} 0^{\circ} & \text{dla } C'_1 C'_2 = 0\\ h'_2 - h'_1 & \text{dla } C'_1 C'_2 \neq 0 \text{ i } |h'_2 - h'_1| \leqslant 180^{\circ}\\ (h'_2 - h'_1) - 360^{\circ} & \text{dla } C'_1 C'_2 \neq 0 \text{ i } (h'_2 - h'_1) > 180^{\circ}\\ (h'_2 - h'_1) + 360^{\circ} & \text{dla } C'_1 C'_2 \neq 0 \text{ i } (h'_2 - h'_1) < -180^{\circ} \end{cases}$$

$$\Delta H' = 2\sqrt{C'_1 C'_h} \sin \frac{\Delta h'}{2} \qquad (10)$$

 $\mathbf{208}$

 $\mathbf{210}$

DODATEK Procedura obliczania różnicy barw według formuły CIEDE2000

3. Obliczanie różnicy barw zaczyna się od uśrednienia wartości L, C i H:

$$\begin{split} \vec{L}' &= \frac{L_1^* + L_2^*}{2} & (11) \\ \vec{C}' &= \frac{C_1' + C_2'}{2} & (12) \\ \vec{h}' &= \begin{cases} \frac{h_1' + h_2'}{2} & \text{dla } C_1' C_2' = 0 \\ \frac{h_1' + h_2'}{2} & \text{dla } C_1' C_2' \neq 0 \text{ i } |h_1' - h_2'| \leq 180^{\circ} \\ \frac{h_1' + h_2' + 360^{\circ}}{2} & \text{dla } C_1' C_2' \neq 0 \text{ i } |h_1' - h_2'| > 180^{\circ} \text{ oraz } h_1' + h_2' < 360^{\circ} \\ \frac{h_1' + h_2' - 360^{\circ}}{2} & \text{dla } C_1' C_2' \neq 0 \text{ i } |h_1' - h_2'| > 180^{\circ} \text{ oraz } h_1' + h_2' > 360^{\circ} \\ \frac{h_1' + h_2' - 360^{\circ}}{2} & \text{dla } C_1' C_2' \neq 0 \text{ i } |h_1' - h_2'| > 180^{\circ} \text{ oraz } h_1' + h_2' > 360^{\circ} \\ 13) \\ T &= 1 - 0, 17 \cos \left(\bar{h}' - 30^{\circ}\right) + 0, 24 \cos \left(2\bar{h}'\right) + \\ + 0, 32 \cos \left(3\bar{h}' + 6^{\circ}\right) - 0, 20 \cos \left(4\bar{h}' - 63^{\circ}\right) & (14) \\ \Delta\Theta &= 30 \exp \left[-\left(\frac{\bar{h}' - 275^{\circ}}{25}\right)^2 \right] & (15) \\ R_C &= 2\sqrt{\frac{\bar{C}''}{\bar{C}'' + 25^{7}}} & (16) \\ S_L &= 1 + \frac{0,015 \left(\bar{L}' - 50\right)^2}{\sqrt{20 + \left(\bar{L}' - 50\right)^2}} & (17) \\ S_C &= 1 + 0,045 \bar{C}' & (18) \\ S_H &= 1 + 0,015 \bar{C}''T & (19) \\ R_T &= -\sin \left(2\Delta\Theta\right) R_C & (20) \end{split}$$

Ostatecznie otrzymujemy wzór z udziałem tzw. składnika rotacyjnego R_T :

$$\Delta E_{00}^{12} = \Delta E_{00} \left(L_1^*, a_1^*, b_1^*, L_2^*, a_2^*, b_2^* \right) = \\ = \sqrt{\left(\frac{\Delta L'}{k_L S_L} \right)^2 + \left(\frac{\Delta C'}{k_C S_C} \right)^2 + \left(\frac{\Delta H'}{k_H S_H} \right)^2 + R_T \frac{\Delta C'}{k_C S_C} \frac{\Delta H'}{k_H S_H}}$$
(21)

WYBRANE ZAGADNIENIA PRZETWARZANIA OBRAZÓW BARWNYCH

Streszczenie

Barwny obraz zapewnia bogatszą informację o obiektach w scenie i dlatego jego pozyskanie i przetwarzanie może znacznie uprościć rozpoznawanie i lokalizację tych obiektów. Występujące od niedawna: łatwy dostęp do urządzeń pozyskiwania obrazów barwnych, takich jak: skanery, cyfrowe aparaty fotograficzne i cyfrowe kamery oraz wzrost mocy obliczeniowej komputerów spowodowały, że to obrazy barwne stały się głównym obiektem przetwarzania i rozpoznawania, a obrazy binarne i monochromatyczne pozostały istotne w niektórych zastosowaniach widzenia maszynowego, przetwarzania obrazów medycznych itp. Chociaż przetwarzanie informacji barwnej jest niewątpliwie bardziej złożone, to jednocześnie informacja ta jest bardziej pełna. Cyfrowe przetwarzanie obrazów barwnych, oprócz stosowania metod znanych z przetwarzania obrazów monochromatycznych, korzysta z osiągnięć nauki o barwie, a szczególnie z techniki pomiarowej, jaką jest kolorymetria.

Prezentowane w niniejszej rozprawie badania koncentrowały się wokół problemów niskopoziomowego przetwarzania obrazu (ang. *low level processing*), począwszy od zagadnienia reprezentacji barwy w obrazach, poprzez pozyskiwanie obrazu barwnego, przetwarzanie wstępne (filtracja, kwantyzacja barwy), aż po segmentację obrazu barwnego i ocenę wyników tej segmentacji. W systemach wizyjnych przetwarzających obrazy barwne na tych właśnie etapach w najwyższym stopniu przejawia się specyfika barwy. Dla wysokopoziomowego przetwarzania obrazu (ang. *high level processing*), polegającego m.in. na ekstrakcji cech, rozpoznawaniu obiektów czy interpretacji sceny barwa obszaru może stanowić jedną z wielu wykorzystywanych cech, obok kształtu, rozmiaru i tekstury.

W rozdziale 2 przeprowadzono przegląd kilkunastu najbardziej popularnych w dziedzinie przetwarzania obrazów przestrzeni barw. Uwagę skupiono na transformacjach z

STRESZCZENIE

bazowej przestrzeni *RGB* i podstawowych właściwościach definiowanych w ten sposób przestrzeni. Szczególną uwagę zwrócono na liniowość transformacji, stabilność obliczeniową transformacji oraz ich równomierność percepcyjną. Wskazano również na możliwość sprzętowej implementacji takich transformacji oraz znajdujące coraz więcej zastosowań podejście kwaternionowe do reprezentacji barwy w obrazach.

W rozdziale 3 przedstawiono cztery zagadnienia związane z pozyskiwaniem obrazów barwnych. Są one silnie związane z najważniejszymi urządzeniami wchodzącymi w skład systemu wizyjnego (kamery, framegrabbery, układy oświetleniowe itp.). Badania w tym zakresie wymagały konfigurowania systemów wizyjnych oraz stosowania próbek z atlasów barw i specjalnych wzorników barwy. Zaproponowano i sprawdzono procedurę kalibracji kolorymetrycznej systemu wizyjnego. Na przykładzie cyfrowego aparatu fotograficznego pokazano, jak standaryzacja przestrzeni *RGB* może poprawić jego wierność odwzorowania barw. Do oceny wierności stosowano najnowsze formuły różnicy barw proponowane przez Międzynarodową Komisję Oświetleniową. Dla kanałów jednoprzetwornikowej kamery CCD wyznaczono widmowe charakterystyki czułości, nie używając specjalizowanego sprzętu pomiarowego. Globalnym atrybutem pozyskiwanego obrazu barwnego jest jego barwność. Dlatego definicję tego ważnego pojęcia dołączono do tego "sprzętowego" rozdziału, a wśród jego badanych właściwości znalazły się zależności od lokalnych atrybutów percepcyjnych takich, jak odcień, nasycenie i jasność.

Znajomość właściwości przestrzeni barw pozwala na odpowiedni dobór przestrzeni do określonego zadania przetwarzania obrazów. Rozdział 4 przybliża właściwości ważnej i często stosowanej klasy przestrzeni barw objętej wspólnym oznaczeniem HSx. Pokazano, że stosowanie kątowej skali odcieniowej może być źródłem problemów nawet podczas prostych działań arytmetycznych. Z drugiej strony niezmiennicze właściwości odcienia i nasycenia stanowią ważną zaletę tych przestrzeni. Przedstawiono badania doświadczalne, które pokazały przewagę wersji HSI nad 3 innymi wersjami przestrzeni percepcyjnych. Stosując stosunkowo proste przekształcenia algebraiczne wyprowadzono wzory na składowe HLS negatywu, co pozwala dokonywać tej operacji bez konieczności powrotu do przestrzeni RGB. Nierównomierny rozkład punktów-barw w bryłach barw przestrzeni HSx dopełnia właściwości tych przestrzeni. Drugą przestrzenią barw omawianą w rozdziałe 4 jest przestrzeń $K_1K_2K_3$ będąca wynikiem transformacji Karhunena-Loèvego na obrazie RGB. Transformacja KLT zapewnia dekorelację składowych barwy. Wiele uwagi poświęcono efektywnej realizacji tej złożonej obliczeniowo transformacji. Przetwarzając zbiór reprezentatywnych obrazów pokazano zjawisko skupienia energii w pierwszej składo-

wej przestrzeni $K_1 K_2 K_3$. Zauważono również, że w zmiennych warunkach oświetleniowych kontrast w obrazie K_1 utrzymuje się znacznie dłużej niż w obrazie luminancyjnym Y.

Rozdział 5, najobszerniejszy, poświęcono zagadnieniu segmentacji obrazów barwnych, które ma kluczowe znaczenie w zastosowaniach. Ogólny sens segmentacji obrazu to działanie na skończonym zbiorze pikseli w celu jego podziału na podzbiory (np. obszary) zawierające podobne elementy. Zakres rozpatrywanych w tym rozdziale technik ograniczono do technik pikselowych (progowanie i klasteryzacja) i obszarowych. Wiele miejsca poświecono progowaniu obrazów barwnych w przestrzeni HSI, w szczególności z wykorzystaniem wiedzy o tle oraz w zastosowaniu do detekcji odblasków. Podobnie szeroko badano technike klasteryzacyjna k-means zwracając uwagę na dobór parametrów. Omówiono dwie wersje techniki rozrostu obszaru: ziarnistego i nieziarnistego rozrostu. Ta ostatnia automatyczna technika segmentacji daje szczególnie dobre wyniki, zależne od stosowanego kryterium jednorodności wraz z progiem oraz od przestrzeni barw. Odpowiedni dobór kryterium, progu i przestrzeni pozwala ignorować cienie i odblaski w obrazie. Przetwarzanie końcowe, np. usuwanie małych obszarów, pozwala dodatkowo poprawić jakość segmentacji. Mjara jakości segmentacji moga być specjalne funkcje oceny, których przydatność została potwierdzona w pracy. Wynik segmentacji w postaci obrazu binarnego może być punktem wyjścia do detekcji cech kształtowych (współczynniki kształtu, momenty itp.) i topologicznych (np. liczba Eulera), opisujących obiekty przedstawione w obrazie.

Następny rozdział odnosi się do zagadnień poprzedzających segmentację, tzn. przetwarzania wstępnego, które ograniczono do kwantyzacji barwy i filtracji odszumiającej. Przeprowadzono badania 3 metod kwantyzacji: równomiernej w RGB oraz w HSV, a także kwantyzacji metodą k-means. Do oceny wyników kwantyzacji stosowano w rozprawie miarę teoriosygnałową PSNR, różnicę barw ΔE oraz dodatkowo zaproponowano różnicę barwności ΔM . W ramach badań nad filtracją poddano porównaniu kilka nieliniowych filtrów zachowujących krawędzie. Zaproponowano zastosowanie funkcji oceny segmentacji do zbadania wpływu filtracji odszumiającej na wyniki segmentacji. Wynikiem szczególnego zainteresowania autora barwnością obrazu było zastosowanie tego globalnego atrybutu percepcyjnego do porównania skalarnego i wektorowego filtru medianowego. W rozdziale 7 podsumowano najważniejsze wyniki i osiągnięcia rozprawy.
SELECTED PROBLEMS OF COLOUR IMAGE PROCESSING

Summary

Colour images are the sources of rich information on objects in the scene. Therefore, its acquisition and processing can significantly simplify both object recognition and location processes. Nowadays an easy access to the equipment for colour image acquisition e.g. scanners, digital still cameras etc. and still grow-up of computational power are the reasons of colour images use instead of the grey-level images. The binary and monochrome images are still important in some applications of the machine vision, medical image processing, etc. Although, the more complicated colour image processing the fuller colour information is. Apart from the use of the methods popular for the monochrome image processing, the digital colour image processing uses the results of colour science, especially the colour measurement technique (colorimetry).

The research, presented in this monograph, was concentrated around the low level processing problems starting from the representation of colour in digital images, colour image acquisition, preprocessing (filtering, colour quantization), to continue with the problems of colour image segmentation and its evaluation. In the vision systems processing colour images, peculiarity of colour at these low level stages is revealed. In the high level processing (feature extraction, object recognition, scene interpretation etc.) the colour of region can be one of many used features as shape, size or texture.

In Chapter 2 a review of the most popular colour spaces in the field of image processing has been presented. The attention has been given to on the transforms from the RGB colour space, defining these spaces and their fundamental properties. The colour spaces were compared on the basis of linearity of transform, their numerical stability and their perceptual uniformity. The possibility of hardware implementation of transforms is suggested. More and more applications of quaternion approach to the representation of colour in images have been noticed.

In Chapter 3 four problems from the field of colour image acquisition were discussed. These problems are strongly related to the most important devices in the vision system e.g. cameras, framegrabbers, lighting systems etc. Research in this topic required the vision system configurating and applying special colour charts and colour atlases. The procedure of colorimetric calibration of vision system has been proposed and checked. The example of digital still camera has shown that standardization of *RGB* space can improve colour reproduction accuracy. New colour difference formulas, currently proposed by CIE, have been applied to evaluation of accuracy. The spectral sensitivity curves for channels of one-chip CCD camera have been determined without use of any special equipment. The global attribute of acquired colour image is its colourfulness. Therefore, the definition of this important concept has been attached to this "hardware-oriented" Chapter and the impact of such local perceptual attributes as the hue, saturation and lightness on image colourfulness has been studied.

The knowledge of the colour space properties is very useful for the process of space selection in defined task of the colour image processing. Chapter 4 brings the properties of the important and frequently used class of colour spaces closer, which is named HSx. The use of angular scale for hue can be a source of problems even during the simple arithmetic operations. From the other hand the invariant of hue and saturation is important advantage of these spaces. Experimental research presented in this work confirmed that the HSIcolour space has an advantage over the three other versions of perceptual colour spaces. Using relatively simple algebraic transformations the formulae for HLS components of negative image were derived, that makes it possible to do this operation without necessity to go back to the RGB space. Non-uniform distribution of points-colours in the HSxcolour solids is also an important property of these spaces. The second colour space described in Chapter 4 is $K_1K_2K_3$ space that is the result of Karhunen-Loève transformation on the RGB image. This transform decorrelates the colour components. Much attention was devoted to the effective realization of this computationally complex transform. As the result of the representative images processing an effect of the energy compaction for the first component in the $K_1K_2K_3$ space was shown. It was also observed that in the changeable lighting conditions the contrast in the image K_1 remains significantly higher than the contrast in the luminance image Y.

Chapter 5, the longest Chapter of this monograph, is devoted to the problem of colour image segmentation that is very important in many applications. In general sense the image segmentation is an action on limited set of pixels with the aim of dividing into subsets (e.g. regions) with similar elements. The range of the considered segmentation techniques has been limited to the pixel-based (thresholding, clustering) and region-based techniques. A lot of place is devoted to colour image thresholding in HSI colour space, in particular the application of knowledge on the background and application to highlight detection. Also the k-means clustering technique with its parameters has been thoroughly tested. Two versions of region growing technique have been presented: seeded and unseeded versions. The latter automatic technique generates good results. Proper selection of homogeneity criterion, threshold and colour space allows to ignore shadows and highlights in the image. Postprocessing (e.g. small region removing) additionally improves the quality of the segmentation. Special evaluation functions, which usefulness has been confirmed in the research, can be used for the measurement of quality of the segmentation. The segmentation result, in the form of binary image, can be a good output to the detection of shape (shape factors, spatial moments etc.) and topological (e.g. the Euler number) features, describing objects in the image.

The next Chapter concerns the stage preceding the image segmentation i.e., the preprocessing stage, in which the research has been restricted to the problems of colour quantization and denoising filtering. Three colour quantization methods: uniform quantization in RGB space, uniform quantization in HSV space and k-means technique, have been investigated. The following measures for evaluation of the quantization results were used: the fundamental signal processing measure PSNR, the colour difference ΔE and additionally proposed difference of colourfulness ΔM . Within the framework of the filter testing a few nonlinear edge preserving filters have been compared. In this Chapter the use of evaluation function, originally proposed for the image segmentation, for denoising filtering, is also proposed. As a result of special interest in the image colourfulness was an application of this global perceptual attribute in comparison of scalar and vector median filters. In the last Chapter the most important results and achievements of this monograph have been summarized.

216

and the second second second second

Książki Wydawnictwa można nabyć w księgarniach

GLIWICE

- Punkt Sprzedaży Wydawnictwa na Wydziale Górnictwa i Geologii ul. Akademicka 2 (237-17-87)
- "FORMAT" Akademicka 5 na Wydziale Budownictwa
- "LAMBDA" ul. Akademicka 2 (237-21-40)
- "MERCURIUS" ul. Prymasa S. Wyszyńskiego 14 b (032) 230-47-22
- "ŻAK" ul. Kaszubska (budynek Biblioteki)

BIAŁYSTOK

- Dom Książki (Księgarnia 84) ul. Wiejska 45 c
- EKOPRESS Księgarnia Wysyłkowa ul. Brukowa 28 (085) 746-04-95

GDAŃSK

EKO-BIS – ul. Dyrekcyjna 6 (058) 305-28-53

KATOWICE

- Punkt Sprzedaży na Wydziale Transportu ul. Krasińskiego 8
- Hurtownia "DIK" ul. Dulęby 7 (032) 204-82-30
- Hurtownia "JERZY" ul. Słoneczna 24 (032) 258-99-58

KRAKÓW

- Techniczna ul. Podwale 4 (012) 422-48-09
- Punkt Sprzedaży WND AGH, Al. Mickiewicza 30 (012) 634-46-40
- "Zielony Smok" ul. Fabryczna 6 (012) 414-00-25

LUBLIN

Składnica Medyczna LSM – ul. Obywatelska 9 (081) 745-52-03

ŁÓDŹ

- "POLITECHNIKA 100" ul. Żeromskiego 116 PŁ.
- Hurtownia "BIBLIOFIL" ul. Jędrowizna 9a (042) 679-26-77

POZNAŃ

- Księgarnia "POLITECHNIK" ul. Piotrowo 3 (061) 665-23-24
- Księgarnia Techniczna ul. Półwiejska 28 (061) 659-00-38

RYBNIK

- "ORBITA" ul. Rynek 12
- "NEMEZIS" ul. gen. Hallera 26

TYCHY

• "I JA TOURS" - ul. J. Piłsudskiego 10 (217-00-91 w.130)

USTROŃ

• "KRISTECH" - ul. Porzeczkowa 12 (600-55-99-83)

WARSZAWA

- Studencka Pl. Politechniki 1 (022) 628-77-58
- Techniczna ul. Kaliskiego 15 (022) 666-98-02
- Techniczna ul. Świętokrzyska 14
- MDM ul. Piękna 31

WROCŁAW

• "TECH" - ul. Wybrzeże Wyspiańskiego 27

ZABRZE

• Punkt Sprzedaży na Wydziale Organizacji i Zarządzania-ul. Roosevelta 26

BIBLIOTEKA GŁÓWNA Politechniki Śląskiej Drukernia Gliwice



Wydawnictwo Politechniki Śląskiej 44-100 Gliwice, ul. Akademicka 5, tel./faks (0-32) 237-13-81 http://wydawnictwo.polsl.pl Dział Sprzedaży i Reklamy tel. (0-32) 237-18-48 e-mail: wydawnictwo_mark@polsl.pl

