



KAPITAŁ LUDZKI  
NARODOWA STRATEGIA SPÓŁNOŚCI



Politechnika Wrocławska

UNIA EUROPEJSKA  
EUROPEJSKI  
FUNDUSZ SPOŁECZNY



## Wstęp do przetwarzania obrazów przemysłowych

**Ewaryst Rafajłowicz i Wojciech Rafajłowicz**

**„Zamawianie kształcenia na kierunkach technicznych, matematycznych, przyrodniczych - pilotaż**



**KAPITAŁ LUDZKI**  
NARODOWA STRATEGIA SPÓŁNOŚCI



Politechnika Wrocławskiego

UNIA EUROPEJSKA  
EUROPEJSKI  
FUNDUSZ SPOŁECZNY



Recenzenci:  
Marek Kurzyński

© Copyright by Oficyna Wydawnicza Politechniki Wrocławskiej



**KAPITAŁ LUDZKI**  
NARODOWA STRATEGIA SPÓŁNOŚCI



Politechnika Wrocławska

UNIA EUROPEJSKA  
EUROPEJSKI  
FUNDUSZ SPOŁECZNY



OFICYNA WYDAWNICZA POLITECHNIKI WROCŁAWSKIEJ  
Wybrzeże Wyspiańskiego 27, 50 – 370 Wrocław

ISBN – 978-83-7493-613-2

# **Wstęp do przetwarzania obrazów przemysłowych**

**Ewaryst Rafajłowicz i Wojciech Rafajłowicz**



# Przedmowa

W niniejszej książce koncentrujemy się na zagadnieniach przetwarzanie obrazów, które dedykowane są zastosowaniom w monitorowaniu jakości produkcji przemysłowej, zarówno ciągłej jak i dyskretnej (jednostkowej). Motywacji do jej napisania dostarczył fakt, że W USA zarejestrowanych jest kilka tysięcy firm zajmujących się dostarczaniem sprzętu i oprogramowania do bieżącej oceny jakości wytwarzania. Również w Polsce liczba takich firm rośnie, a stosowanie systemów wizyjnych stało się dla producentów i odbiorców ich wyrobów wyznacznikiem nowoczesności produkcji. Czynnikami wzrostu tego segmentu rynku są: znaczny spadek cen kamer przemysłowych o wysokiej rozdzielczości i wzrost mocy obliczeniowej komputerów, które są w stanie obsłużyć strumień obrazów w czasie rzeczywistym. Łącznie, czynniki te przekładają się na relatywnie niski koszt inwestycji w systemy wizyjne, który zwraca się bardzo szybko.

W części pierwszej zebrano informacje podstawowe o kamerach, reprezentacji obrazów i ich wstępnym przetwarzaniu. Główny materiał, dotyczący metod i algorytmów wykrywania obiektów i defektów na obrazach, zebrano w części II. W części III zebrano podstawowe metody poprawiania obrazów, jeśli jest ono niezbędne. Zamiast zakończenia, w części IV, proponujemy zapoznanie się z przykładami zastosowań kamer termowizyjnych, które znacznie poszerzają możliwości monitorowania jakości produkcji i eksploatacji.

Poza zakresem tej książki znalazło się wiele ważnych i ciekawych nurtów przetwarzania obrazów. Wymieniamy je – wraz z odsyłaczami do literatury – w uwagach bibliograficznych zawartych w rozdziale wstępnym. Ważnym kryterium doboru materiału była możliwość efektywnej implementacji danej metody w taki sposób, by nadawała się ona do pracy w czasie rzeczywistym, z prędkością powyżej 10 klatek na sekundę, przy zastosowaniu procesora (lub procesorów) o rozsądnej cenie.

Autorzy dziękują Prezesowi firmy OPTOSOFY S.A. z Wrocławia Panu Włodzimierzowi Rohlederowi za wyrażenie zgody na opublikowanie sekwencji zdjęć z wybranych procesów produkcyjnych, dla których firma OPTOSOFT opracowała i wdrożyła wizyjne systemy monitorowania jakości produkcji.



# Spis treści

<b>I Podstawy</b>	<b>17</b>
<b>1 Przykłady i problemy</b>	<b>19</b>
1.1 Przykłady . . . . .	19
1.2 Obszary zastosowań . . . . .	20
1.3 Klasyfikacja zadań i metod . . . . .	25
1.4 Źródła obrazów . . . . .	27
1.5 Oprogramowanie i źródła literaturowe . . . . .	28
1.5.1 Uwagi na temat oprogramowania . . . . .	28
1.5.2 Komentarze bibliograficzne . . . . .	29
<b>2 Kamery</b>	<b>33</b>
2.1 Rodzaje kamer . . . . .	33
2.2 Przetworniki obrazu . . . . .	34
2.3 Budowa kamery . . . . .	35
2.4 Połączenie kamery z komputerem . . . . .	42
<b>3 Odwzorowanie obrazu w pamięci</b>	<b>45</b>
3.1 Obraz jako macierz . . . . .	45
3.2 ROI . . . . .	47
3.3 Uwagi o próbkowaniu i kwantyzacji . . . . .	47
3.4 Przechowywanie obrazów . . . . .	49
<b>4 Transformacje punktowe</b>	<b>53</b>
4.1 Obroty, odbicie lustrzane . . . . .	53
4.2 Jasność . . . . .	54
4.3 Kontrast . . . . .	55
4.4 Gamma . . . . .	55
4.5 LUT . . . . .	56

<b>II Wykrywanie obiektów i defektów</b>	<b>59</b>
<b>5 Proste narzędzia</b>	<b>63</b>
5.1 Histogram . . . . .	63
5.2 Statystyki globalne dla obrazów . . . . .	66
5.3 Lokalne zastosowanie prostych statystyk . . . . .	67
<b>6 Operacje arytmetyczne na obrazach</b>	<b>73</b>
6.1 Obraz wzorcowy i szacowanie tła . . . . .	73
6.2 Operacje na parach obrazów . . . . .	75
6.2.1 Dodawanie i odejmowanie par obrazów . . . . .	75
6.2.2 Dzielenie dwóch obrazów . . . . .	75
6.2.3 Inne operacje na parach obrazów: . . . . .	78
6.2.4 Przykłady operacji na dwóch obrazach . . . . .	78
<b>7 Obiekty i defekty jako grupy pikseli</b>	<b>81</b>
7.1 Wstęp . . . . .	81
7.1.1 Segmentacja . . . . .	81
7.1.2 Przetwarzanie obrazu przed progowaniem . . . . .	82
7.2 Progowanie . . . . .	82
7.2.1 Dobór progu na postawie histogramu . . . . .	83
7.2.2 Inne warianty progowania . . . . .	85
7.3 Automatyczny dobór progu . . . . .	85
7.3.1 Metoda Otsu . . . . .	86
7.3.2 Podstawowa metoda wyboru progu . . . . .	86
7.3.3 Metoda 2-średnich . . . . .	87
7.3.4 Wybór progu jako bayesowski problem decyzyjny . . . . .	88
7.3.5 Przykład – bayesowski wybór progu . . . . .	89
7.3.6 Jeszcze raz metoda Otsu . . . . .	90
7.3.7 Zasada maksymalnej entropii . . . . .	90
7.3.8 Uwagi o wyborze metody progowania . . . . .	92
7.4 Progowanie adaptacyjne . . . . .	93
7.4.1 Progowanie adaptacyjne . . . . .	93
7.4.2 Progowanie wielopoziomowe . . . . .	94
7.4.3 Progowanie dwustronne – przykład . . . . .	94
7.5 Jeszcze o progowaniu . . . . .	95
7.5.1 Proste progowanie – przykład pozytywny . . . . .	95
7.5.2 Dwie proste reguły . . . . .	95
7.5.3 Metoda Trussell'a . . . . .	96
7.5.4 Kolejne cztery proste metody . . . . .	98
7.6 Uwagi dotyczące segmentacji obrazów kolorowych . . . . .	100

<b>8 Lokalizacja, etykietowanie i pomiar</b>	<b>101</b>
8.1 Poszukiwanie obiektu o znanym kształcie . . . . .	101
8.2 Histogramy brzegowe . . . . .	105
8.3 Histogramy brzegowe – przykład . . . . .	107
8.4 Uwagi o analizie grup pikseli . . . . .	107
8.5 Etykietowanie obiektów . . . . .	111
8.5.1 Spójne obszary . . . . .	112
8.5.2 Algorytm etykietowania . . . . .	115
8.5.3 Rozwiązywanie konfliktów . . . . .	116
8.6 Opisywanie obiektów . . . . .	117
8.6.1 Cechy geometryczne . . . . .	119
8.6.2 Momenty geometryczne . . . . .	120
8.7 Znajdowanie obiektów . . . . .	121
8.7.1 Przykład . . . . .	121
<b>9 Kontury obiektów i defektów</b>	<b>129</b>
9.1 Okontuowywanie . . . . .	129
9.1.1 Maski . . . . .	130
9.1.2 Definicja krawędzi . . . . .	131
9.1.3 Pole gradientu . . . . .	132
9.1.4 Metody gradientowe . . . . .	133
9.1.5 Metody drugiego rzędu . . . . .	136
9.1.6 Inne detektory krawędzi . . . . .	143
9.1.7 Ocena detektorów krawędzi . . . . .	143
9.2 Opisy krawędzi . . . . .	144
9.2.1 Kodowanie łańcuchowe . . . . .	145
9.2.2 Odcinkowo-liniowe łączenie krawędzi . . . . .	146
9.2.3 Fourierowski opis konturu . . . . .	147
<b>10 Wykrywanie obiektów na podstawie konturów</b>	<b>151</b>
10.1 Proste metody wykrywania . . . . .	151
10.1.1 Operacje logiczne na obrazach binarnych . . . . .	151
10.1.2 Wykrywanie fragmentów linii – maski . . . . .	152
10.2 Współrzędne biegunowe – sygnatury obiektów . . . . .	152
10.2.1 Sygnatury . . . . .	152
10.2.2 Centroid . . . . .	154
10.2.3 Dopasowywanie wykresów . . . . .	155
10.3 Transformacja Hougha . . . . .	158
10.3.1 Wykrywanie prostych – wersja dydaktyczna . . . . .	159
10.3.2 Dogodniejsza parametryzacja prostej . . . . .	162
10.3.3 Uwagi na temat transformacji Hougha . . . . .	164

10.3.4 Modyfikacje i uogólnienia . . . . .	165
<b>III Poprawianie obrazów</b>	<b>167</b>
<b>11 Wyrównywanie histogramu</b>	<b>171</b>
11.0.5 Przykład wstępny . . . . .	171
11.0.6 Podstawy teoretyczne . . . . .	171
11.0.7 Wersja empiryczna i przykłady . . . . .	174
<b>12 Proste metody filtracji</b>	<b>177</b>
12.1 Proste filtry . . . . .	179
12.1.1 Filtr ruchomej średniej . . . . .	180
12.1.2 Filtry bazujące na innych średnich . . . . .	181
12.1.3 Filtr ruchomej mediany . . . . .	183
12.2 Filtry z wertykalnymi wagami . . . . .	185
12.2.1 Idea wertykalnie uciętego pudełka . . . . .	186
12.2.2 Ogólniejsza klasa filtrów z wertykalnymi wagami . . . . .	187
<b>13 Operacje morfologiczne</b>	<b>193</b>
13.1 Podstawowe pojęcia . . . . .	193
13.1.1 Definicje i podstawowe własności . . . . .	193
13.1.2 Dylatacja . . . . .	195
13.1.3 Erozja . . . . .	196
13.2 Otwarcie i domknięcie . . . . .	196
13.2.1 Definicje . . . . .	196
13.2.2 Przykłady działania operacji otwarcia i domknięcia . . . . .	198
13.3 Złożone operacje morfologiczne . . . . .	201
13.3.1 Rekonstrukcja . . . . .	201
13.3.2 Hit or miss . . . . .	203
13.3.3 Inne zastosowania . . . . .	203
<b>IV Epilog</b>	<b>205</b>
<b>14 Zamiast zakończenia – zastosowania termowizji</b>	<b>207</b>
14.1 Bliska podczerwień – zastosowania . . . . .	207
14.1.1 Obserwacje obiektu . . . . .	207
14.1.2 Wykrywanie wad . . . . .	211
14.2 Projektowanie układu pomiarowego . . . . .	212
14.3 Ocena strat cieplnych w systemach energetycznych . . . . .	217
14.4 Straty ciepła w budynkach . . . . .	217





# Spis rysunków

1.1	Ostrze nożyczek (po lewej) i fragment ostrza pilnika . . . . .	21
1.2	Ostrze szpilki i fragment tkaniny . . . . .	21
1.3	Fragment orderu i stalówka wiecznego pióra . . . . .	21
1.4	Zadanie polega na policzeniu płyt miedzianych . . . . .	22
1.5	Czy skorodowana, metalowa płyta musi być już wymieniona ?	22
1.6	Ocena defektów w produkcji ciągłej . . . . .	23
1.7	Termowizyjny obraz współczesnego budynku użyteczności publicznej . . . . .	28
2.1	Matryce kamer analogowych i cyfrowych . . . . .	34
2.2	Tradycyjny układ i układ Bayera . . . . .	35
2.3	Budowa kamery . . . . .	36
2.4	Dystorsja beczkowata – uzyskana sztucznie w celu pokazania zjawiska. . . . .	36
2.5	Dystorsja typu rybie oko – uzyskana sztucznie w celu pokazania zjawiska. . . . .	36
2.6	Obiektyw w przekroju (uproszczony) . . . . .	37
2.7	Głębia ostrości . . . . .	38
2.8	Przykład szumu spowodowanego zbyt dużym wzmacnieniem.	39
2.9	Ograniczenia czasu naświetlania, wynikające z zastosowanej prędkości odbioru obrazów z kamery. . . . .	40
2.10	Po lewej rozgrzana płyta, po prawej średnia intensywność na kolejnych klatkach . . . . .	41
2.11	Obraz nieprawidłowy – widoczne są promienie słoneczne. . . . .	42
3.1	Bufor, określenie współrzędnych . . . . .	46
3.2	Obraz funkcji gaussowskiej, która – mimo gładkości – nie ma ograniczonego widma (rysunek po lewej) i funkcji $\sin(r)/r$ , która ma ograniczone widmo ( $r = \sqrt{x^2 + y^2}$ ). . . . .	48
3.3	Przykładowe defekty w wersji 8-bitowej i 4-bitowej. . . . .	49

4.1	Obraz orginalny (u góry) i obraz z odwróconymi poziomami szarości (u dołu). . . . .	54
4.2	Obraz orginalny i o powiększonej jasności. . . . .	54
4.3	Obraz orginalny i o zwiększym i zmniejszonym kontraście .	55
4.4	Obraz orginalny i o zwiększym kontraście, w tym przypadku zmiana jest bardzo potrzebna . . . . .	55
4.5	Wykresy transformacji gamma dla $\gamma = 0.3$ , $\gamma = 1.7$ oraz, dla porównania, dla $\gamma = 1$ . . . . .	56
4.6	Obraz orginalny i po transformacji gamma dla $\gamma = 0.3$ i $1.7$ .	56
4.7	Przykładowe transformacje punktowe . . . . .	57
5.1	Nie patrząc na lewy obraz, możemy powiedzieć, że zawiera on tylko cztery odcienie szarości w podobnych proporcjach (ale nie możemy powiedzieć gdzie się znajdują) . . . . .	64
5.2	Prosty przypadek – wzgórze po lewej odpowiada obiekowi, wieża po prawej to poziomy szarości tła . . . . .	65
5.3	Zdarza się też sytuacja idealna – poziomy szarości defektów idealnie odseparowane od tła . . . . .	65
5.4	Teraz odgadnięcie poziomu defektów nie jest tak łatwe – $\Delta = 25$ . . . . .	66
5.5	Wybór $\Delta$ jest bardzo ważny – teraz $\Delta = 5$ . . . . .	67
5.6	Część dobra (po lewej) oraz uszkodzona (po prawej) . . . . .	68
5.7	Obraz testowy i przekroje poziomów szarości – pionowe, co 100 pikseli wzdłuż slabu . . . . .	69
5.8	Obraz testowy i przebieg mediany (u góry) oraz średniej. . . . .	70
5.9	Obraz testowy i dyspersje oraz minima w przekrojach. . . . .	70
6.1	Wynik odejmowania obrazów . . . . .	76
6.2	Wynik dzielenia (przycięte i poprawione przed podzieleniem – dla lepszej widoczności) . . . . .	76
6.3	Wynik odejmowania obrazów (po lewej) i wartości bezwzględnej ich różnicy. . . . .	77
6.4	Wynik mnożenia obrazów (po lewej) i operacji minimum (po prawej) . . . . .	78
7.1	Histogram, pozwalający na łatwe ustalenie progu . . . . .	83
7.2	Obraz po progowaniu – $T = 177$ . . . . .	84
7.3	Histogram obrazu, dla którego znalezienie progu jest trudne. .	84
7.4	Przykład – metoda Otsu . . . . .	86
7.5	Detal i histogram . . . . .	89
7.6	Progowanie metodą Otsu i 2-średnich; $T = 123$ . . . . .	89

7.7	Wynik stosowania progu $T = 166$ , otrzymanego za pomocą podejścia bayesowskiego . . . . .	90
7.8	Histogram i odpowiadające mu $H_{sum}(t)$ . . . . .	91
7.9	Próg $\tilde{T} = 174$ – obliczony metodą maksimum entropii . . . . .	92
7.10	Histogram rdzy . . . . .	93
7.11	Wybór obszarów o różnym stopniu korozji . . . . .	95
7.12	Przykład skuteczności prostego progowania . . . . .	96
7.13	Wykres statystyki t-Studenta ma maksimum dla $T = 110$ (po prawej). Wynik progowania dla $T = 110$ (po lewej) . . . . .	97
7.14	. . . . .	98
7.15	Wynik progowania metodą Metoda Kittler'a-Illingworth'a . . . . .	99
7.16	Możemy szybko wykryć czerwone tabletki, cienie są automatyczne pomijane. . . . .	100
8.1	Obiekt o znanym kształcie . . . . .	102
8.2	Przykład zastosowania metody wyszukiwania wzorca – opis w tekście . . . . .	103
8.3	Przykład zastosowania metody wyszukiwania wzorca do obrazu w odcieniach szarości – opis w tekście . . . . .	103
8.4	Ilustracja działania metody histogramów brzegowych . . . . .	105
8.5	Histogramy brzegowe – sprężynka 3 . . . . .	107
8.6	Histogramy brzegowe – sprężynka 1 . . . . .	108
8.7	Histogramy brzegowe – sprężynka 2 . . . . .	109
8.8	Binarny obraz kapsułek . . . . .	110
8.9	Obraz defektów po przypisaniu kolorowych etykiet. . . . .	111
8.10	Kratery widoczne w części defektów po progowaniu nie powinny być uzupełniane. . . . .	112
8.11	Zebra nie byłaby szczęśliwa, gdyby jej paski traktować jako oddzielne obiekty. . . . .	113
8.12	Dwie kapsułki nie zostały rozróżnione w procesie etykietowania	113
8.13	Wykrywanie wycieku – orginalny obraz . . . . .	122
8.14	Wykrywanie wycieku – przekształcenie od odcieni szarości . . . . .	123
8.15	Wykrywanie wycieku – zwiększenie kontrastu . . . . .	123
8.16	Wykrywanie wycieku – segmentacja . . . . .	124
8.17	Wykrywanie wycieku – poprawa rezultatów segmentacji . . . . .	124
8.18	Wykrywanie wycieku – lokalizacja . . . . .	125
8.19	Wykrywanie wycieku 2 – orginalny obraz . . . . .	125
8.20	Wykrywanie wycieku 2 – przekształcenie od odcieni szarości . . . . .	126
8.21	Wykrywanie wycieku 2 – zwiększenie kontrastu . . . . .	126
8.22	Wykrywanie wycieku 2 – segmentacja . . . . .	127
8.23	Wykrywanie wycieku 2 – poprawa rezultatów segmentacji . . . . .	127

8.24 Wykrywanie wycieku 2 – lokalizacja . . . . .	128
9.1 Powierzchnia obrazu z krawędzią i jej przekrój . . . . .	131
9.2 Krawędzie o coraz mniejszym nachyleniu . . . . .	131
9.3 Krawędzie z zakłóceniami . . . . .	131
9.4 Powierzchnia obrazu z krawędzią i odpowiadające jej pole gradientów . . . . .	132
9.5 Obraz po zastosowaniu maski Prewitta . . . . .	134
9.6 Obraz po drugiej masce Prewitta . . . . .	134
9.7 Kierunek $x$ (po lewej) i $y$ (po prawej). Bez progowania, ale z odwróceniem poziomów szarości . . . . .	135
9.8 Wykrywanie krawędzi detektorem Sobela z różnym poziomem progu $T$ . . . . .	136
9.9 Na wykresach (od lewej) pokazano przekrój krawędzi, pochodną i drugą pochodną . . . . .	136
9.10 Obrazy (macierze $R$ ), otrzymane w wyniku zastosowania masek operatora Laplace'a do obrazu testowego . . . . .	138
9.11 Po lewej – krawędzie wykryte dla $T = 0.005$ (wszystko jest zadowalający). Po prawej – $T = 0.1$ (wynik wciąż nie jest zadowalający) . . . . .	139
9.12 Po lewej – krawędzie dla $T = 0.25$ . Po prawej – $T = 0.12$ . . . . .	139
9.13 Obraz oryginalny (po lewej) i po LoG (po prawej) . . . . .	139
9.14 Jądro gaussowskie i jego zdyskretyzowana wersja w oknie $9 \times 9$ dla $\sigma = 1$ . . . . .	140
9.15 Wynik detekcji LoG . . . . .	141
9.16 Od lewej: obraz oryginalny, obraz po filtracji gaussowskiej, wynik wykrywania krawędzi detektorem LoG . . . . .	141
9.17 Obraz po detekcji LoG – okno o rozmiarze $3 \times 3$ . . . . .	142
9.18 Obraz przed i po LoG – rezultat jest niezadowalający . . . . .	142
9.19 Idea łączenia punktów krawędzi . . . . .	146
9.20 Parametryczny opis krzywej zbliżonej kształtem do elipsy . . . . .	147
9.21 Krzywa $c(t)$ o kształcie zbliżonym do elipsy (po lewej) i wykres modułów współczynników $ C_k $ . . . . .	148
9.22 Rekonstrukcja kształtu na podstawie 14 deskryptorów. . . . .	149
9.23 Krawędzie monety i wykres 295 modułów wartości współczynników Fouriera . . . . .	149
9.24 Wynik rekonstrukcji brzegu monety na podstawie 52 deskryptorów Fouriera . . . . .	150
10.1 Koło i jego wykres we współrzędnych biegunkowych (sygnatura). . . . .	153
10.2 Elipsy i ich sygnatury. Obroty wyłącznie je przesuwają . . . . .	154

10.3 Trójkąt o zaokrąglonych kątach. . . . .	154
10.4 Część krawędzi koła i jej wykres we współrzędnych bieguno- wych z centroidem na $(0,0)$ . . . . .	155
10.5 Sygnatura okręgu ze źle dobranym środkiem. . . . .	156
10.6 Ilustracja potrzeby skalowania – dwie elipsy, których jedna jest dwa razy większa od drugiej . . . . .	156
10.7 Moneta, wykryty brzeg i dopasowanie ich profili. $Q = 0.88$ . . .	158
10.8 Obraz z linią . . . . .	159
10.9 Po lewej – prosta $y = 2 \cdot x + 1$ i punkty $\{0, 1\}, \{1, 3\}, \{2, 5\}$ . Po prawej proste $b = 1$ , $b = -a + 3$ , $b = -2 \cdot a + 5$ , z punktem wspólnym $(2, 1)$ . . . . .	160
10.10 Proste w przestrzeni parametrów i komórki akumulatora . . .	161
10.11 Akumulator w naszym przykładzie. Maksimum w okolicy punk- tu $(2, 1)$ jest wyraźnie widoczne . . . . .	162
10.12 Przestrzeń parametrów: $(\rho, \theta)$ . . . . .	163
10.13 Przykładowy rysunek (po lewej) i wykryte linie (po prawej) . .	163
11.1 Fragment niepoprawnie naświetlonej metalowej płyty i jego histogram . . . . .	172
11.2 Fragment metalowej płyty i jego histogram . . . . .	172
11.3 Ciemny samochód na ciemnym tle (po lewej) i wynik wyrów- nywania histogramu (po prawej) . . . . .	173
11.4 Histogramy zdjęć przed i po wyrównaniu histogramu . . . .	174
11.5 Slab przed i po wyrównaniu histogramu . . . . .	175
11.6 Histogramy slabu . . . . .	175
12.1 Szum typu sól z pieprzem . . . . .	178
12.2 Szum o rozkładzie równomiernym . . . . .	178
12.3 Szum gaussowski . . . . .	179
12.4 Wynik działania filtru gaussowskiego o oknie $11 \times 11$ . . . .	181
12.5 Porównanie zachowania mediany i średniej na krawędzi . . .	184
12.6 Wynik działania filtru medianowego o oknie $9 \times 9$ . . . .	185
12.7 Pudełko nadążające za powierzchnią obrazu . . . . .	187
12.8 Zasada działania filtru zachowującego krawędzie . . . . .	188
12.9 I wiersz Vert. mediana, V-średina $H = 0.1$ , $M = 5$ V-średnia Class-średnia Class-mediana . . . . .	191
13.1 Tło i obiekt . . . . .	194
13.2 Dylatacja $3 \times 3$ SE . . . . .	195
13.3 Dylatacja $13 \times 13$ SE . . . . .	196
13.4 Erozja $7 \times 7$ SE . . . . .	197

13.5 Erozja 15x15 SE . . . . .	197
13.6 Domknięcie SE 7x7 . . . . .	198
13.7 Domknięcie SE 19x19 – bardzo domknęty . . . . .	199
13.8 Otwarcie SE 3x3. . . . .	199
13.9 Otwarcie SE 7x7. . . . .	200
13.10 Otwarcie SE 25x25 – szeroko otwarty. . . . .	200
13.11 Porównanie otwarcia z domknięciem . . . . .	200
13.12 Różnice na brzegu . . . . .	201
13.13 Porównanie erozji z dylatacją (wiersz środkowy). Porównanie otwarcia z domknięciem (wiersz dolny) . . . . .	202
13.14 SE - koło, $r=7$ . Znaleźć koła większe od $r=7$ . Znaleziono prawidłowe obiekty, ale zostały one zmniejszone . . . . .	203
13.15 SE - koło $r = 12$ . Znaleziono koła większe niż $r=12$ , ale również zostały one zmniejszone . . . . .	204
 14.1 Wybrane zdjęcia termowizyjne w bliskiej podczerwieni – klatki od 10 do 110. . . . .	208
14.2 Wybrane zdjęcia termowizyjne w bliskiej podczerwieni – klatki od 130 do 230. . . . .	209
14.3 Wybrane zdjęcia termowizyjne w bliskiej podczerwieni – klatki od 200 do 500. . . . .	210
14.4 Rozkład temperatury bębna nagrzewnicy – ustawiona moc 10 kW. . . . .	214
14.5 Rozkład temperatury bębna nagrzewnicy – ustawiona moc 30 kW. . . . .	215
14.6 Rozkład temperatury bębna nagrzewnicy – ustawiona moc 50 kW. . . . .	216
14.7 Termowizyjny obraz transformatora 1. . . . .	217
14.8 Termowizyjny obraz transformatora 2. . . . .	218
14.9 Termowizyjny obraz metalowego słupa trakcji energetycznej. .	218
14.10 Termowizyjne obrazy budynków . . . . .	219
 15.1 Wykresy $I(\lambda)$ w temperaturze $0^\circ C$ . . . . .	222
15.2 Wykresy $I(\lambda)$ w temperaturze $2000^\circ C$ . . . . .	222
15.3 Schemat ciała doskonale czarnego . . . . .	223
15.4 Schemat punktowego pomiaru temperatury. . . . .	224

# Część I

**Podstawy:**

**Przykłady i problemy**

**Literatura**

**Kamery**

**Obraz w pamięci**

**Transformacje punkt-punkt**



# Rozdział 1

## Przetwarzanie obrazów przemysłowych – przykłady i problemy

Niniejszy rozdział ma charakter poszerzonego wstępu. Postaramy się zarysować zagadnienia poruszane w tej książce a także naszkicować obszary zastosowań przetwarzania obrazów w różnych gałęziach przemysłu.

### 1.1 Przykłady

Rozpoczniemy od przykładów obrazów i pytań, które możemy stawiać. Przy okazji, postaramy się pokazać, że nawet przemysłowe obrazy mogą być – po prostu – ładne.

- Na Rys. 1.1 pokazano ostrze nożyczek i fragment ostrza pilnika. Założymy, że mamy więcej takich przykładowych zdjęć ostrzy poprawnych i zużytych. Czy potrafimy tak dobrą algorytmami przetwarzania obrazów, by móc – z małym prawdopodobieństwem błędu – automatycznie rozstrzygać, czy dane ostrze nadaje się jeszcze do użytku, czy też należy je wymienić? Kładziemy nacisk na słowo automatycznie, które oznacza tu tyle co *bez udziału człowieka w procesie decyzyjnym*. Dopuszczamy udział człowieka na etapie dostarczenia przedmiotu przed kamerą. Oczywiście lepiej mieć system, który także i w tej czynności wyręczy człowieka, ale ten aspekt pozostawiamy poza zakresem tej książki, odsyłając Czytelnika do opracowań z automatyki i mechaniki precyzyjnej.
- Na Rys. 1.2 pokazano ostrze szpilki i fragment tkaniny. Czy szpilka ta

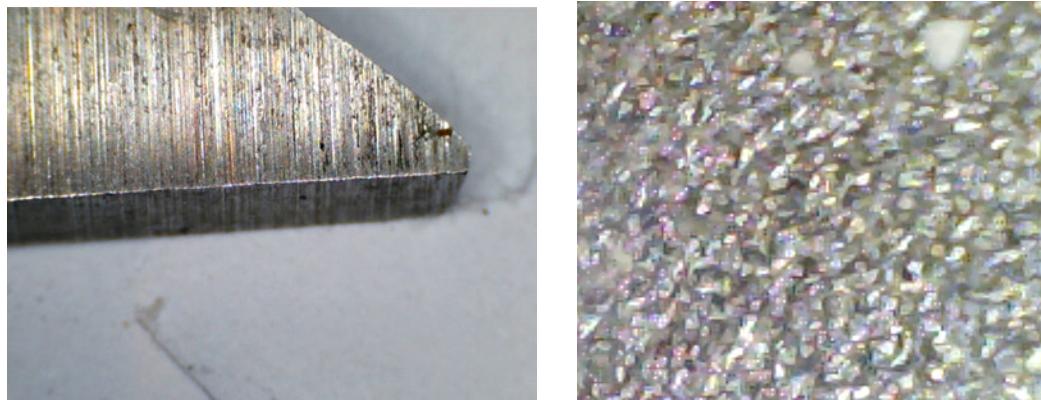
jest dostatecznie ostra ? Czy tkanina jest prawidłowo utkana ? Podobnie jak wyżej, aby dobrać metody przetwarzania tych obrazów, musimy mieć więcej przykładów poprawnych i niepoprawnych.

- Na Rys. 1.3 widać fragment orderu i stalówkę wiecznego pióra. Czy order ten i stalówka zostały w pełni starannie wykonane ?
- Na zdjęciu Rys. 1.4 widać miedziane płyty. Zadanie polega na policzeniu ile ich jest. Utrudnieniem jest fakt, że płyty mają nierówną grubość w różnych punktach tej samej płyty, co powoduje, że ich krawędzie nie są równoległe. Ponadto, grubość różnych płyt też jest różna.
- Płyta pokazana na Rys. 1.5 uległa procesowi korozji. Stopień korozji jest różny w poszczególnych jej fragmentach. Czy potrafimy ocenić kiedy płytę taka należy wymienić ?
- Na Rys. 1.6 pokazano schemat stanowiska do oceny jakości produkcji ciągłej, na przykład, powierzchni rur. Cztery kamery rozmieszczone co  $90^\circ$  obserwują rurę przesuwającą się z dużą prędkością, prostopadle do płaszczyzny kartki. Przekrój rury zaznaczono na czarno. Ze względu na potrzebną moc obliczeniową, obraz z każdej z kamer przetwarzany jest w niezależnym komputerze, a wyniki oceniane są w komputerze nadzorującym.
- Wiele mechanizmów jest wciąż składanych przez człowieka. Montaż takiego jest narażony na wiele niedokładności. Kontrola poprawności takiego montażu jest wzgleśnie łatwa, gdyż czas przetwarzania pojedynczego obrazu może wynieść kilka sekund. Dodatkowo można ustabilizować warunki oświetleniowe, Detaile są ręcznie układane pod kamerą. Pracownik otrzymuje informację dobrze/źle na monitorze i sortuje elementy do odpowiedniego pudełka.

## 1.2 Obszary zastosowań

Z punktu widzenia oceny jakości, produkcję wyrobów można zgrubnie podzielić na dwa rodzaje.

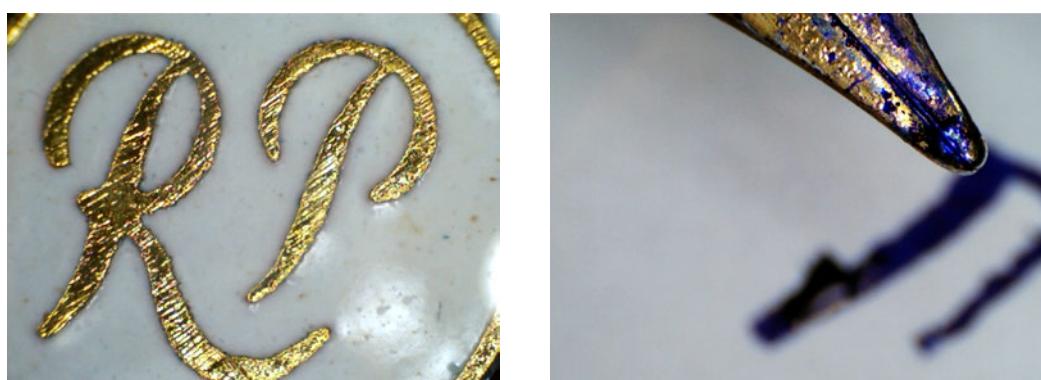
1. Produkcja dyskretna – elementy pojawiają się pod kamerą sztuka po sztuce w odstępach rzędu sekund lub większych (płytki drukowane układów elektronicznych, skompletowane pudełka wyrobów z elementami pomocniczymi lub wymiennymi).



Rysunek 1.1: Ostrze nożyczek (po lewej) i fragment ostrza pilnika



Rysunek 1.2: Ostrze szpilki i fragment tkaniny



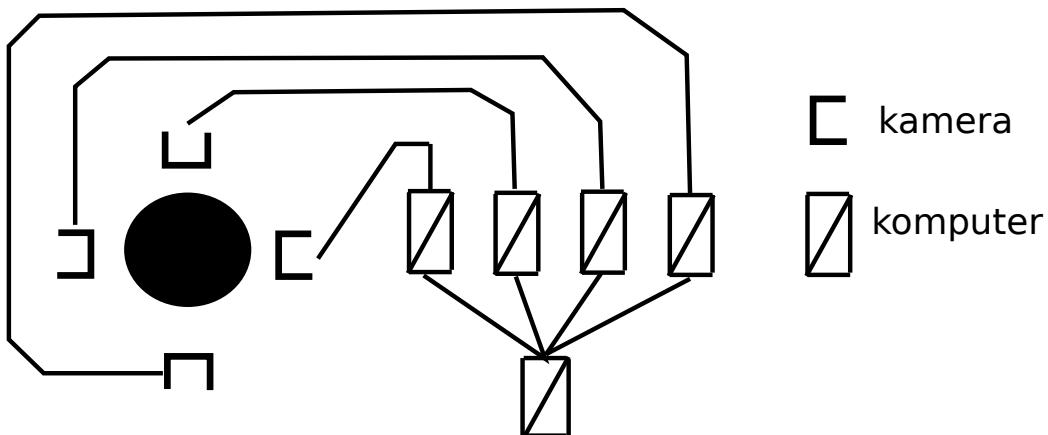
Rysunek 1.3: Fragment orderu i stalówka wiecznego pióra



Rysunek 1.4: Zadanie polega na policzeniu płyt miedzianych



Rysunek 1.5: Czy skorodowana, metalowa płyta musi być już wymieniona ?



Rysunek 1.6: Ocena defektów w produkcji ciągłej

2. Produkcja ciągła – produkt przesuwa się przed kamerą w sposób ciągły i jego powierzchnię należy oceniać na bieżąco (plastikowa rura, odlewane pasmo, walcowana blacha).

Przypadkami pośrednimi są przykłady produkcji, która w zasadzie jest dyskretna, lecz odbywa się z tak dużą prędkością, że wymagane prędkości przetwarzania porównywalne są z tymi, które narzuca szybka produkcja ciągła. Przykładów dostarczają procesy rozlewania i kapslowania, na przykład, mleka lub napojów gazowanych.

Poniżej przedstawiamy przegląd obszarów zastosowań metod wizyjnych w poszczególnych rodzajach przemysłu. Wobec gwałtownie rosnącej różnorodności obszarów zastosowań, przegląd ten daleki jest od kompletności.

### 1. Przemysł samochodowy:

- jakość szyb,
- montaż bezpieczników i całej wiązki przewodów elektrycznych,
- bezpieczeństwo (napinacze pasów),
- malowania, spasowanie elementów.
- stan tarcz hamulcowych,
- tarcze sprzęgła,
- poprawność wykonania filtrów spalin.

### 2. Przemysł farmaceutyczny:

- poprawność napełnienia ampułek,

- wykonanie plastrów,
- liczba i kolor tabletek.

**3. Przemysł włókienniczy:**

- poprawność tkania,
- poprawność malowania tkanin,
- badanie powłok tkanin specjalnych.

**4. Przemysł materiałów budowlanych:**

- ocena wypalenia cegieł i dachówek,
- powierzchnia płytek ceramicznych,
- izolatory.

**5. Przemysł hutniczy:**

- ocena powierzchni blach i rur,
- jakość walcowania drutu.

**6. Przemysł spożywczy:**

- ocena zbiorów,
- ocena przecierów dla dzieci,
- stan owoców, głównie jabłek.
- nadzór chłodni.

**7. Przemysł elektroniczny:**

- poprawność trawienia płytEK
- poprawność montażu elementów elektronicznych.

**8. Pakowanie:**

- zamknięcie butelek z mlekiem,
- zakrętki,
- kontrola napełniania butelek,
- kontrola kompletności zestawów,
- zliczanie przy składowaniu.

Opisy metod monitorowania jakości za pomocą kamer, stosowanych w przemyśle włókienniczym i przetwarzania wełny znaleźć można w pracach [1], [75], [83], [92], [84]. Inspekcji w przemyśle spożywczym i pakowania jego wyrobów dotyczą artykuły [6], [26], [7], [26], [22] i monografia [14]. Przemysł elektroniczny i elementów elektrycznych jest bardzo znaczącym odbiorcą wzajemnych systemów [10], [15], [20], [39]. Przegląd metod stosowanych w odlewictwie i przetwórstwie metali znaleźć można w pracach [49], [50], [53], [71], [74], [87].

### 1.3 Przetwarzanie obrazów przemysłowych na tle procesu produkcyjnego

Na przemysłowe przetwarzanie obrazów składa się:

- automatyczna kontrola za pomocą kamer,
- detekcja zmiany w procesie ciągłym lub wykrycie niezgodności w produkcie jednostkowym,
- decyzja czy wykryta zmiana, dotycząca kontrolowanego elementu lub fragmentu procesu ciągłego, jest dyskwalifikująca.

Narzędzi do wykrywania zmian lub niezgodności dostarcza teoria statystycznych metod monitorowania jakości (por. [81], [30], [58]). Podstawowym narzędziem, służącym do wykrywania zmian są karty kontrolne. Czułość kart kontrolnych na zmiany jakości jest zwykle dużo większa niż powszechnie stosowane proste progowanie. W rozważanych tu problemach, danymi wejściowymi kart kontrolnych są cechy wekstrahowane z obrazu w wyniku jego przetwarzania. Uzyskanie takich danych jest głównym tematem tej książki.

Etapem który następuje po wykryciu zmiany lub niezgodności w produkcie jednostkowym jest decyzja o zakwalifikowaniu tej zmiany jako jako istotnej lub jako jeszcze mniejszej się w zakresie dopuszczalnym. Narzędzi ułatwiających podjęcie decyzji dostarczają techniki:

- rozpoznawania obrazów (wzorców) (*pattern recognition*),
- analizy skupień (*cluster analysis*),
- testowania hipotez statystycznych i Bayesowskiej teorii decyzji,
- sztuczne sieci neuronowe.

Problematyką kart kontrolnych i podejmowania decyzji nie będziemy się w tej książce zajmować.

Skupimy się natomiast na podstawowych zadaniach kontroli za pomocą kamer. Należą do nich:

- sprawdzanie poprawności wykonania produktów,
- pomiar parametrów produktów lub produkcji za pomocą przetwarzanie obrazów.

Na podstawie zarejestrowanych zmian parametrów procesu można podejmować również decyzje o prowadzeniu procesu. W realizacji takich decyzji może uczestniczyć człowiek lub zmiany nastaw procesu mogą odbywać się one automatycznie. Jednocześnie warto zbierać dane statystyczne o zmienności procesu w celu podejmowania decyzji na szczeblu zarządzania.

Większość można zaklasyfikować do następujących kategorii lub ich kombinacji:

- Kontrola kształtu i rozmiaru produktu (parametry geometryczne).
- Sprawdzanie obecności bądź nieobecności oraz położenia części, otworów i (lub) ich położenia względem siebie.
- Kontrola powierzchni pod kątem obecności zadanej cechy bądź braku jednolitości.

Pod względem technicznym każde z tych zagadnień można podzielić na następujące etapy:

- pobranie obrazu (klatki),
- poprawa jakości obrazu (zwykle dostosowana do sytuacji),
- wykrycie obiektu lub obiektów
- lokalizacja obiektu,
- pomiary i (lub) porównanie ze wzorcem,
- zebranie cech do rozpoznawania, klasyfikacji itd.,
- ocena obiektów (lub defektów).

Przedstawiona lista jest zbliżona pod względem zawartości do spisu treści tej książki.

Zadania związane z przetwarzaniem obrazów zwykle klasyfikuje się jako:

1. zadania niskiego poziomu przetwarzania (pobieranie, poprawienie obrazów, detekcja obiektów lub defektów),
2. zadania średniego poziomu – wykrywanie obiektów, ich położenia oraz cech,
3. zadania wyższego poziomu przetwarzania – rozpoznawanie i decyzje,
4. próby interpretacji i zrozumienia obrazu.

Próby automatycznego rozumienia obrazów mogą odnosić się do dobrze określonych klas obrazów, gdyż – na przykład – próba zrozumienia obrazu *Picassa Guernica* jest trudna bez znajomości historii.

W książce tej koncentrujemy się na zagadnieniach 1) i 2), czyli na technikach niskiego i średniego poziomu przetwarzania obrazów.

## 1.4 Siegaj gdzie wzrok nie sięga – źródła obrazów

Podstawowym źródłem obrazów przemysłowych są kamery pracujące w świetle widzialnym. Rosnąca rozdzielcość tych kamer i relatywnie niskie ceny powodują gwałtowny wzrost ich zastosowań.

Warto jednak wymienić inne źródła obrazów, które – z racji wyższych cen – stosowane są w sytuacjach wymagających szczególnego traktowania. W tym kontekście można brać pod uwagę:

1. obrazy dostarczane przez źródła promieniowania rentgenowskiego, zarówno płaskie jak i tomograficzne,
2. obrazy powstające w wyniku zastosowania rezonansu magnetycznego.
3. obrazowanie ultradźwiękowe,
4. kamery termowizyjne – pracujące w podczerwieni,
5. kamery pracujące paśmie ultrafioletowym,
6. skanery laserowe,
7. mikroskopy sprzężone z matrycą cyfrową.



Rysunek 1.7: Termowizyjny obraz współczesnego budynku użyteczności publicznej

Zwracamy uwagę na kamery termowizyjne. Znaczący spadek cen tych kamer spowodował, że przestały być one bardzo drogim sprzętem o charakterze laboratoryjnym. Uprościła się także ich obsługa, gdyż nie jest wymagane chłodzenie ciekłym helem. Czynniki te powodują, że rośnie zainteresowanie firm przemysłowym wykorzystaniem termowizji. Najprostsze zastosowanie – do oceny strat ciepła w budynkach – pokazano na Rys. 1.7.

## 1.5 Oprogramowanie i źródła literaturowe

### 1.5.1 Uwagi na temat oprogramowania

Tworzenie oprogramowania do przetwarzania obrazów przemysłowych przebiega zwykle w dwóch zasadniczych – często przeplatających się – etapach.

**Etap tworzenia koncepcji przetwarzania** Zadanie stworzenia oprogramowania do monitorowania danego procesu rzadko udaje się rozwiązać przez zastosowanie gotowych metod. Zwykle jest to proces zestawiania sekwencji metod i badania, czy wynikowa metoda spełnia wymagania. Na tym etapie można i warto posługiwać się gotowymi procedurami elementarnymi. Znajdziemy je w postaci komend systemów takich jak Octave, Matlab i Mathematica (w wersji 7 lub wyższych).

**Etap wytworzenia oprogramowania użytkowego** Po skompletowaniu i przetestowaniu całego ciągu technologicznego przetwarzania obrazów dla danego zastosowania, trzeba napisać efektywne oprogramowanie użytkowe. Nie zawsze warto pisać wszystkie procedury samemu. Dostępna jest biblioteka OpenCV, której licencja pozwala na korzystanie z wielu bardzo dobrych algorytmów zarówno do celów dydaktycznych jak i komercyjnych (na zasadach podanych w licencji). Jako wstęp do zapoznania się z tą biblioteką służyć mogą dwie książki [5] oraz [62], która jest dostępna w Dolnośląskiej Bibliotece Cyfrowej, jako pdf do pobrania.

### 1.5.2 Komentarze bibliograficzne

Literatura na temat przetwarzania obrazów liczy sobie tysiące artykułów i dziesiątki książek. Nie sposób wymienić tutaj wszystkich i dlatego skupimy się głównie na wątku związanym z przetwarzaniem obrazów przemysłowych pod kątem monitorowania jakości produkcji. Również ten wątek liczy sobie setki artykułów naukowych, natomiast liczba książek jest wyraźnie mniejsza.

Wśród książek wyróżnia się monografia E. R. Daviesa [13]. Zebrano w niej ogromne doświadczenie praktyczne jej autora oraz ważniejsze oryginalne metody, które rozwija i stosuje. W trakcie pisania niniejszej książki korzystaliśmy z monografii [13] wielokrotnie. W szczególności, rozdziały 7.1 i 7.2 bazują na rozwiniętej przez Davies'a technice histogramów brzegowych.

Wiele cennych informacji znaleźćmy w książce [18]. Zawiera ona wiele podpowiedzi o charakterze praktycznym.

O ile autorom wiadomo, na polskim rynku księgarskim jedynie monografia prof. K. J. Kurzydłowskiego i Jego zespołu [?] poświęcona jest przetwarzaniu obrazów przemysłowych. Jej autorzy dzielą się z Czytelnikiem bogatą wiedzą na temat zastosowań w materiałoznastwie, w tym na temat współczesnych technik mikroskopowych i związanych z nimi metod przetwarzania obrazów. Znajdziemy w niej również oceny skuteczności poszczególnych algorytmów przetwarzania obrazów.

W przetwarzaniu obrazów przemysłowych korzysta się szeroko z wybranych, zwykle prostych, szybkich i skutecznych, podejść wypracowanych z myślą o szerszych zastosowaniach przetwarzania obrazów. Z tego względu, w wielu miejscach tej książki korzystaliśmy z monografii Jaina [33], Pratta [54], Gonzalesa i Woodsa [25] oraz [31], [28], [68].

Ogólnym algorytmem przetwarzania obrazów poświęconych jest kilka nadal aktualnych – mimo upływu lat – i bardzo użytecznych książek w języku polskim. Należy do nich – już dziś klasyczna – książka profesora Tadeusiewicza i Korohody [78] oraz książki profesorów Maliny i Smiatacza[42], Nie-

niewskiego [45] oraz Chorasia [8].

Lista czasopism, które publikują prace na temat przetwarzania obrazów przemysłowych obejmuje:

- *Computers in Industry,*
- *Computer Vision, Graphics, and Image Processing,*
- *Image and Vision Computing,*
- *Pattern Recognition,*
- *Mathematics and Computers in Simulation,*
- *Computer Graphics and Image Processing,*
- *Real-Time Imaging,*
- *Pattern Recognition Letters,*
- *Journal of X-Ray Science and Technology,*
- *Int. J. Computer Vision,*
- *IEEE Trans. Pattern Analysis and Machine Intelligence,*
- *IEEE Trans. Image Processing,*
- *International Journal of Production Research,*
- *Vision Research,*

Powyzsza lista nie jest z pewnością kompletna. Artykuły na temat monitorowania jakości produkcji za pomocą kamer publikowane są w czasopismach z danej dziedziny wytwarzania. Wiele czasopism z obszaru matematyki stosowanej również publikuje wiele prac na temat przetwarzania obrazów. Należą do nich, między innymi, *Nonlinear Analysis*, *J. Math. Imaging and Vision*, *SIAM J. Numerical Analysis*, *International Journal on Applied Mathematics and Computer Science*, który niedawno opublikował numer specjalny [36].

Wiele cennych informacji zawierają artykuły przeglądowe: [23], [32], [41], [47], [2], [52], [51] i wiele przeglądów o nieco węższej tematyce.

### Co musieliszmy pominąć

Jak już wspominalismy w Przedmowie, poza zakresem tej książki znalazło się wiele ważnych i ciekawych nurtów przetwarzania obrazów. Istotnym kryterium doboru materiału była możliwość efektywnej implementacji danej metody. Ważnym czynnikiem doboru była też względna łatwość dostrojenia danej metody do konkretnej klasy zadań – metody, które dostarczają użytkownikowi wielu parametrów, które trzeba wybrać, są wprawdzie elastyczne, ale korzystanie z nich jest utrudnione. Kładlismy też nacisk na prostotę koncepcji wybranych metod, gdyż ułatwia to dostrojenie algorytmu.

Poniżej wymieniamy część pominiętych zagadnień wraz z nielicznymi odсылaczami do literatury, które pozwolą na samodzielne rozpoczęcie ich studiowania. Ważnym narzędziem przetwarzania obrazów są sztuczne sieci neuronowe (por. [22], [79]). Kluczowym problemem dla jakości rozpoznawania obrazów jest dobór cech [24], [80], [83], [21], [37]. Nie mogliśmy poświęcić także zbyt wiele miejsca zagadnieniom próbkowania obrazów [16], [64]. Sporo miejsca zajęły zagadnienia wykrywania krawędzi i filtracji zachowującej krawędzie i segmentacji, ale i tak wiele wątków tego ważnego tematu pozostało poza zakresem tej książki [69], [11], edgeflow, [59], [66], [4], [46]. Ze względu na ograniczenia objętości książki musieliszmy także pominąć ważny nurt badań jakości produktów, charakteryzujących się fakturą powierzchni (ang. *texture*) [85], [86], [89].

Nowych narzędzi do filtracji obrazów dostarczają badania nad wykorzystaniem równań o pochodnych cząstkowych jako filtrów zachowujących krawędzie por. [77].

Kamery coraz częściej stanowią element pętli sprzężenia zwrotnego, które służy do automatycznej korekty jakości wyrobów [90], [35], [61].

Przetwarzanie obrazów medycznych jest w wielu aspektach bliskie przetwarzaniu obrazów przemysłowych. W obu dziedzinach mamy często do czynienia z obiektami, które różnią się od tła zaledwie o kilka poziomów szarości. Wysoki jest również poziom zakłóceń.



# Rozdział 2

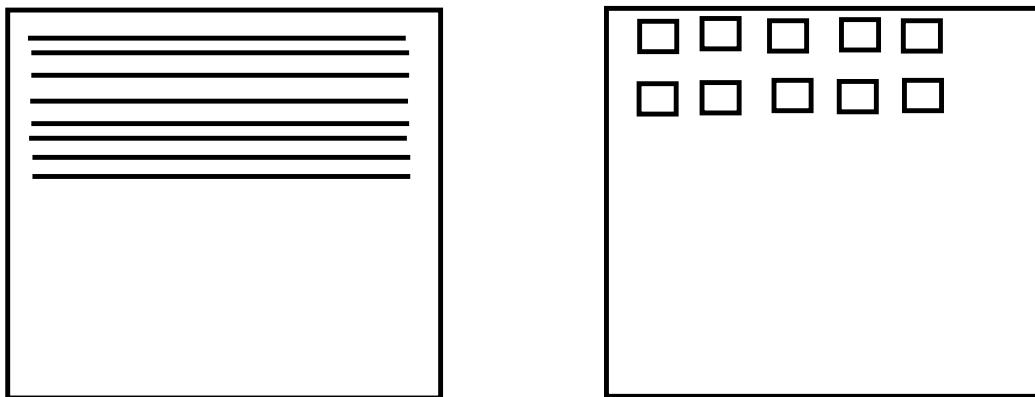
## Kamery

### 2.1 Rodzaje kamer

Kamery można podzielić na trzy grupy:

- analogowe – wprowadzenie obrazu do komputera wymaga specjalnej karty przetworników, zwanej *frame grabber* (termin ten funkcjonuje dość powszechnie także w języku polskim),
- cyfrowe:
  - matryce w technologii CCD (charged coupled devices) – to starszy typ kamer cyfrowych o dobrze opanowanej technologii,
  - matryce w technologii CMOS – nowsza technologia, w porównaniu z kamerami CCD są tańsze przy tej samej rodzinie, nadal doskonalone.
- inne typy kamer – wspominalismy już wcześniej o kamerach pracujących w paśmie podczerwoni i UV,
- warto wymienić także skanery laserowe, które nie są formalnie kamerami, ale powstały przy ich użyciu obrazy można i warto stosować do monitorowania jakości produkcji.

Część optyczna kamer analogowych i cyfrowych jest zasadniczo podobna. Różnice dotyczą przetwornika obrazu.



Rysunek 2.1: Matryce kamer analogowych i cyfrowych

## 2.2 Przetworniki obrazu

### Przetworniki analogowe

Przetwornik kamery analogowej to element światłoczuły, który jest w różny sposób przemiatany (np. wiązką elektronów jak w widikonie) tak, aby odczytywać kolejne linie z obrazu. Teoretycznie, w jednej z osi (zwykle poziomej) rozdzielczość jest nieskończoność duża. W praktyce, ograniczeniem szczegółowości kamer analogowych jest pasmo wzmacniaczy wizyjnych.

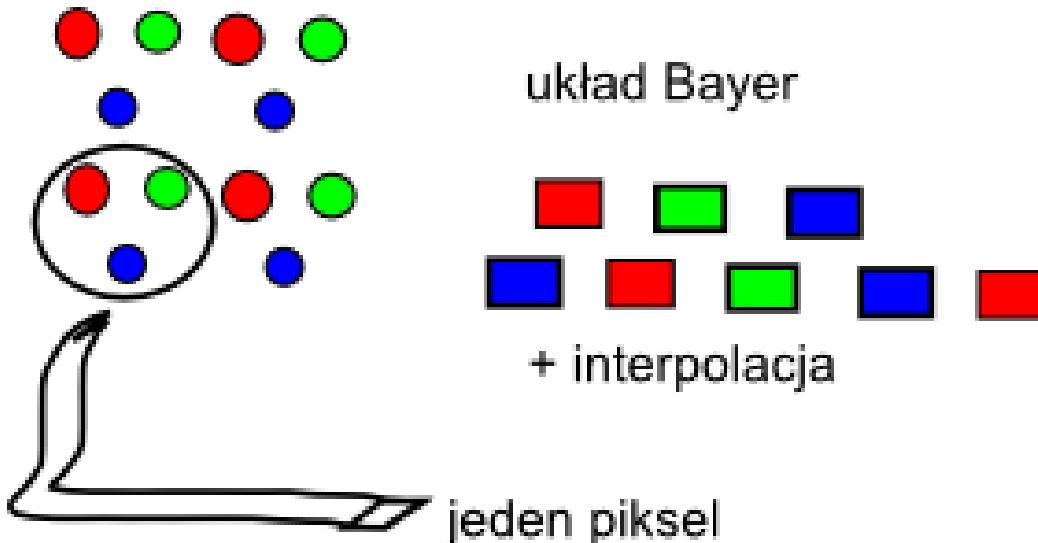
Kamery analogowe nadal są używane, głównie w systemach dozoruienia, lecz ich rola w monitorowaniu jakości produkcji znacznie maleje.

### Przetworniki cyfrowe

Matryca cyfrowa składa się z dużej ilości małych elementów światłoczułych, odczytywanych punkt po punkcie. Rozdzielczość kamery zależy od ilości tych punktów. W celu uzyskania obrazu cyfrowego z kamery analogowej stosuje się tzw. frame-grabbery, które są przetwornikami A/C. Za dyskretyzację w „ciągłe” osi odpowiada czas próbkowania przetwornika.

### Uzyskiwanie koloru

Elementy matrycy nie są czułe na kolory. Kolor uzyskuje się stosując filtry RGB (*red, green, blue*). Trzy bliskie elementy są napełniane różnymi filtrami. Stosuje się też format Bayer – przetworniki poszczególnych kolorów układane są w odpowiednią mozaikę i wartości kolorów w węzłach prostokątnej siatki uzyskiwane są w wyniku interpolacji.



Rysunek 2.2: Tradycyjny układ i układ Bayera

## 2.3 Budowa kamery

### Obiektyw

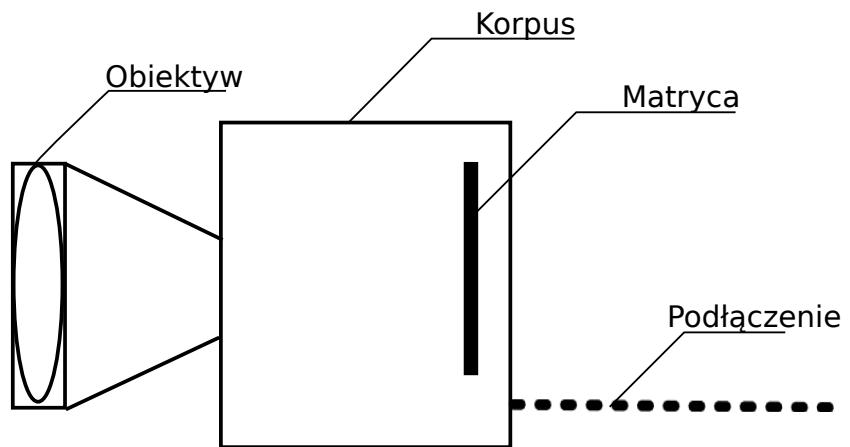
Obiektyw jest układem soczewek, skonstruowanym tak, aby zachowywać się jak pojedyncza soczewka o określonej ogniskowej. Obiektyw stosowany do kontroli jakości musi być dobrej jakości – powinien mieć małe zniekształcenia.

Warto ocenić zniekształcenia wnoszone przez kamerę, fotografując obraz testowy np. szachownicę. Na Rys. 2.4 i Rys 2.5 pokazano, w przerysowanej formie, przykłady zniekształceń obrazu. Występują również zniekształcenia chromatyczne (koloru).

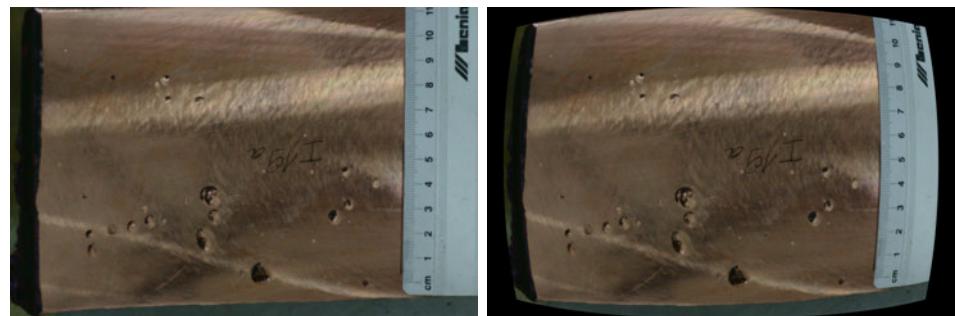
Z punktu widzenia kontroli jakości, dokuczliwe bywa winetowanie – zniekształcenie polegające na mniejszej jasności obrazu w rogach i na obrzeżu. Jego wpływ można zmniejszyć dobierając ogniskową obiektywu i/lub odległość badanego przedmiotu tak, by znajdował się on w centralnej części obrazu. Odbywa się to kosztem pewnego zmniejszenia liczby milimetrów przypadających na jeden piksel, lecz – przy przy osiąganych dzisiaj rozdzielczościach kamer – wadę tę można wyeliminować zwiększąc rozdzielcość kamery.

### Ogniskowa obiektywu

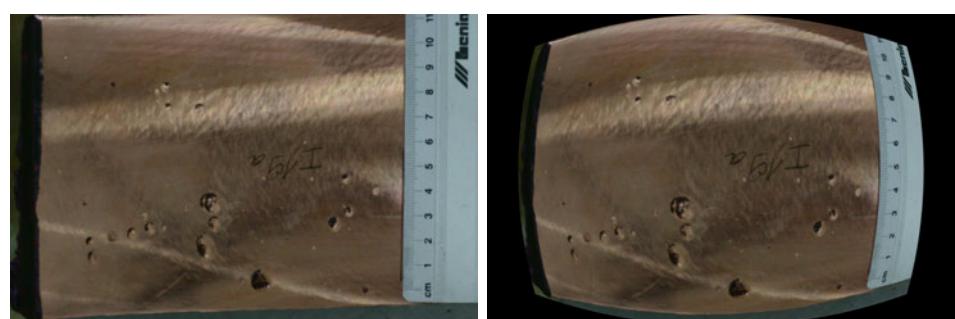
Ważnym parametrem obiektywu jest jego ogniskowa. To czy obiektyw jest szerokokątny, standardowy, czy nazywamy go teleobiektywem, zależy od jego



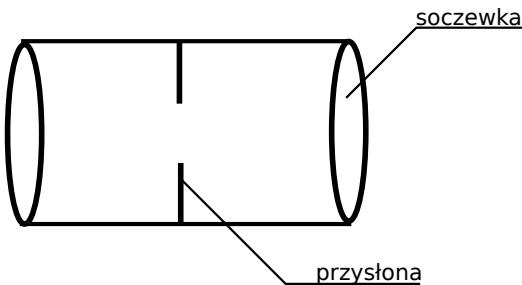
Rysunek 2.3: Budowa kamery



Rysunek 2.4: Dystorsja beczkowata – uzyskana sztucznie w celu pokazania zjawiska.



Rysunek 2.5: Dystorsja typu rybie oko – uzyskana sztucznie w celu pokazania zjawiska.



Rysunek 2.6: Obiektyw w przekroju (uproszczony)

ogniskowej oraz wielkości matrycy. W ogólności

$$\alpha = 2 \operatorname{arc} \operatorname{tg} \left( \frac{d}{2f} \right), \quad (2.1)$$

gdzie  $d$  jest rozmiarem matrycy w odpowiednim kierunku,  $f$  ogniskową a  $\alpha$  kątem widzenia. To właśnie kąt decyduje o zaklasyfikowaniu obiektywu do jednej z tych grup. Na korpusie obiektywu znajdziemy pierścień pozwalający zmieniać wartość przysłony. Stopień jej otwarcia decyduje o ilości światła padającego na matrycę. Powszechnie stosowana skala na pierścieniu przysłony to: 1.4; 2; 2.8; 4; 5.6; 8; 11; 16. Zmiana przysłony o jedną działkę w stronę wyższych wartości powoduje dwukrotne zmniejszenie jasności obrazu. Zatem, aby po zmianie przysłony o 1 działkę uzyskać ten sam poziom oświetlenia czas powinien zmienić się dwukrotnie. Skalę przysłon rozpoczęliśmy od wartości 1.4, która osiągana jest w obiektywach o wysokiej jasności. Typowe obiektywy mogą mieć skalę zaczynającą się od 4 lub nawet 5.6. O subtelnościach związków między czasem naświetlania, ustawieniem przysłony i głębią ostrości piszemy w dalszej części tego rozdziału.

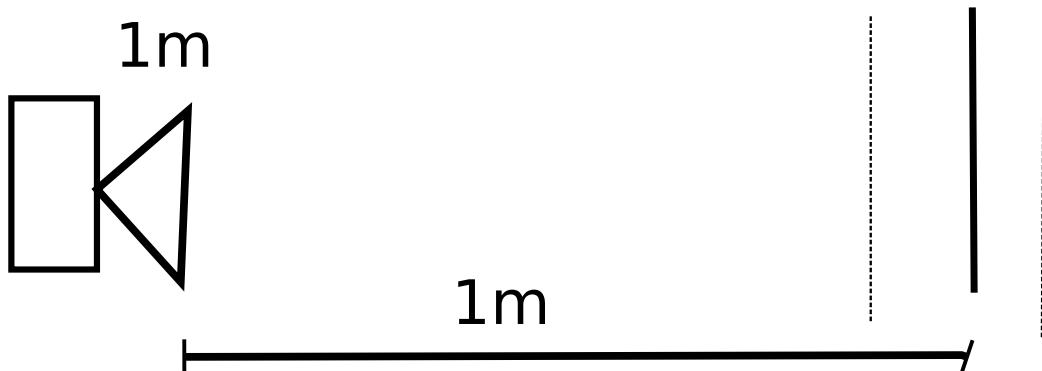
## Ostrość i głębia ostrości

Obiektyw o ogniskowej  $f$  znajdujący się w odległości  $l$  od matrycy skupia na niej światło pochodzące z nieskończoności. Jeśli obiekt znajduje się bliżej, to musimy zwiększyć odległość soczewki od matrycy, gdyż zgodnie ze wzorem, znanym z optyki, między wielkościami tymi zachodzi związek:

$$\frac{1}{l} + \frac{1}{d} = \frac{1}{f} \quad (2.2)$$

(obiektyw traktujemy jako pojedynczą soczewkę).

A jak będą odwzorowane obiekty, które znajdują się w odległości innej niż  $l$ ? Będą widziane jako nieostre, z dokładnością do głębi ostrości (por.



Rysunek 2.7: Głębia ostrości

Rys. 2.7). Natomiast, pojedynczy punkt widziany jest jako tzw. krążek rozproszenia. Głębia ostrości to zakres odległości wokół  $l$ , które mają tę właściwość, że przedmioty znajdujące się w tym przedziale odbierane są przez nasze oko jako widziane ostro.

Parametr ten ma cechy uznaniowe, gdyż zależy od przyjętej maksymalnej wielkości krążka rozproszenia. W najbardziej rygorystycznej wersji można przyjąć rozmiar pojedynczego piksela matrycy. Można też oceniać „na oko” lub zdać się na tabele dostarczone przez producentów.

Ważne: im większa wartość przysłony (mniej światła pada na matrycę), tym głębia ostrości szersza.

### Czas naświetlania i czułość matrycy

W tradycyjnej, fotochemicznej fotografii drugim podstawowym parametrem był czas naświetlania klatki będący czasem otwarcia mechanicznej migawki. Tradycyjnie czas ten podawano jako wartości pewnego szeregu:  $n = 1, 2, 4, 8, 15, 30, 60, 125$  (lub 100), 250, ..., gdzie  $1/n$  to czas naświetlania w ułamkach sekundy.

Cyfrowe sterowanie migawką pozwala na dokładniejsze określenie czasu naświetlania oraz na uzyskiwanie krótkich czasów naświetlania, rzędu  $1/1000$  sekundy.

Trzecim, niewystępującym w tradycyjnej fotografii, parametrem jest wzmacnienie sygnału pochodzącego z matrycy. W cyfrowych aparatach fotograficznych parametr ten bywa nazywany ekwiwalentną czułością i wtedy jest określany w jednostkach ISO, tak jak czułość filmu fotograficznego.

W przypadku kamer cyfrowych parametr ten zwykle jest po prostu procentowym wzmacnieniem lub osłabieniem sygnału z matrycy. Nie należy stosować maksymalnych wzmacnień, gdyż z większym wzmacnieniem wiąże się



Rysunek 2.8: Przykład szumu spowodowanego zbyt dużym wzmocnieniem.

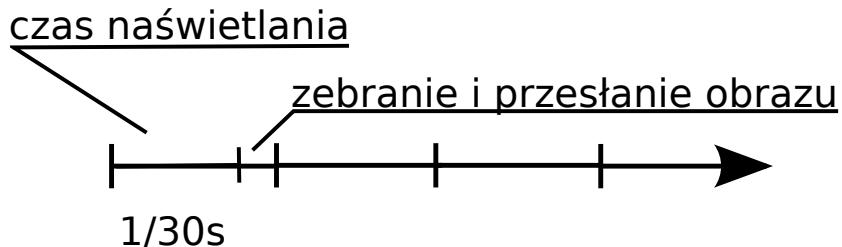
większe szумy matrycy. Zjawisko to pokazano na Rys. 2.8.

Wzmocnienie sprzętowe (*Hardware gain*) uzyskiwane jest na drodze analogowego wzmocnienia sygnału pobranego z komórek matrycy.

Pewne uwagi:

- Zwiększenie wzmocnienia rozjaśnia obraz i wpływa na jego kontrast.
- Zależnie od typu sensora, dobierać możemy:
  - całkowite wzmocnienie (*master gain*) – globalnie dla wszystkich kolorów,
  - wzmocnienie RGB – oddzielnie dla każdego z kolorów.

Na wybór trzech kluczowych parametrów (przysłona, czas naświetlania, wzmocnienie) wpływa oświetlenie. W warunkach przemysłowych oświetlenie, na ogół, można i warto kontrolować. W najprostszej wersji, stosować można krąg diod LED wokół obiektywu. W bardziej zaawansowanej – konstrukcję odcinającą światło padające z zewnątrz i oświetlenie w kolorze pomagającym w dalszym przetwarzaniu.



Rysunek 2.9: Ograniczenia czasu naświetlania, wynikające z zastosowanej prędkości odbioru obrazów z kamery.

### Szybkość pobierania klatek z kamery – *framerate*

Większość kamer ma możliwość zmiany liczby klatek pobieranych z kamery w ciągu sekundy. Parametr ten nazywany jest w dokumentacji *frame rate* i oznaczany skrótem **fps**, czyli *frames per second*. Zbyt duża liczba klatek pobieranych w ciągu sekundy może spowodować, że nie zdążymy ich przetworzyć w komputerze. W skrajnych przypadkach możemy nawet nie zdążyć ich przesyłać. Ponadto, prędkość pobierania obrazów determinuje maksymalny czas naświetlania  $t < \frac{1}{fps}$  (por. Rys. 2.9). Z drugiej strony, jeśli proces produkcyjny przebiega bardzo szybko, to zastosowanie zbyt małej wartości **fps** może spowodować, że obraz będzie zniekształcony (rozmażany) w sposób znany ze starszych fotografii sportowych.

### Wpływ innych czynników

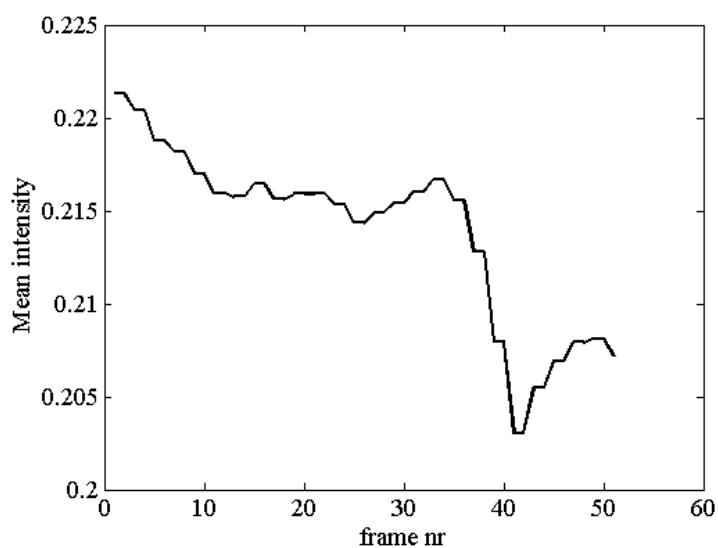
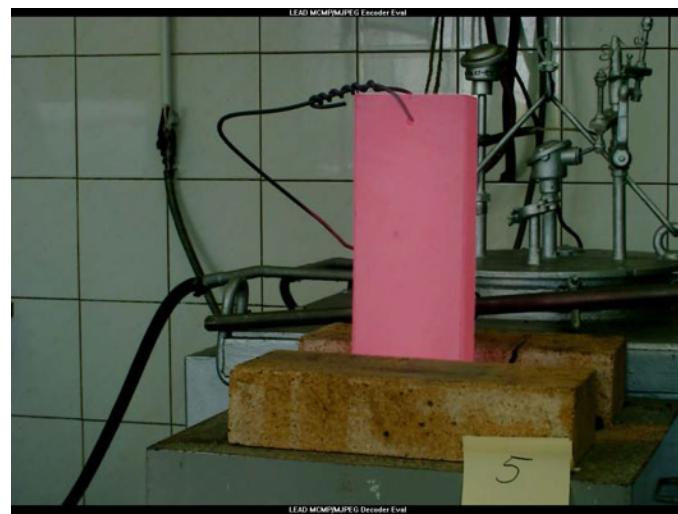
Oprócz opisanych wyżej parametrów, dodatkowymi czynnikami, które należy rozważyć są:

- temperatura,
- wilgotność,
- kurz i pył w powietrzu.

Temperatura i wilgotność wpływają na pracę elektroniki kamery. Urządzenia elektroniczne mają przecież określony zakres temperatur pracy. Nawet uszczelnione kamery mogą ulec zawilgoceniu, jeżeli wilgoć z powietrza skropli się wewnątrz obudowy kamery lub korpusu obiektywu.

Kurz ma bezpośredni wpływ na jakość otrzymywanej obrazu. W pewnych sytuacjach efekt Tyndalla ujawnia kurz w postaci smug widocznych na zdjęciu.

Ponadto, zdarzają się i inne niespodzianki, jak ta opisana ponizej. Miedzia-



Rysunek 2.10: Po lewej rozgrzana płyta, po prawej średnia intensywność na kolejnych klatkach



Rysunek 2.11: Obraz nieprawidłowy – widoczne są promienie słoneczne.

na płytę została rozgrzana do  $835^{\circ}$  a następnie rejestrowano jej stygnięcie (Rys. 2.10). Jaki jest powód nagłego wzrostu a następnie spadku średniej intensywności oświetlenia pomiędzy klatkami 30 i 40? Może mamy do czynienia z silną nieliniowością procesu stygnięcia ? Odpowiedź jest jednak dość banalna: drobna chmurka odsłoniła słońce, a chwilę potem pracownik laboratorium przeszedł wprawdzie za kamerą, ale przysłonił na chwilę światło słoneczne padające z okna.

Drugi przykład jest mniej tajemniczy, ale też się zdarza: zewnętrzne światło padało na fotografowany slab (por. Rys. 2.11). Morał: w miarę możliwości, należy zapewnić stabilne warunki oświetlenia, na przykład, przez zastosowanie obudowy lub ekranów odgradzających nas od światła naturalnego, ruchu ludzi itp.

## 2.4 Połaczenie kamery z komputerem

Jak już wspomniano, do połączenia kamery analogowej z komputerem potrzebny jest przetwornik – frame grabber. Wobec malejącej roli kamer analogowych, wątku tego nie będziemy rozwijać. Skupimy się natomiast na podstawowych informacjach dotyczących połączeń kamer cyfrowych.

**USB 2.0** Znane wszystkim łącze USB w wersji 2 zupełnie nieźle sprawdza się w warunkach przemysłowych. Warunkiem dobrej pracy tego łącza jest niezbyt duża odległość między kamerą a komputerem i względnie niski poziom zakłóceń elektromagnetycznych, pochodzących – na przykład – od silników elektrycznych dużej mocy i ich układów zasilania. Teoretyczna przepustowość tego łącza wynosi 480 megabitów na sekundę. Mimo, że w praktyce przepustowość ta rzadko jest osiągana, to łącze to pracuje poprawnie, jeśli przesyłamy obrazy o wielkości

2-3 megapiksele w tempie rzędu 10 klatek na sekundę. Praktycznie osiągane przepustowości są rzędu 240 megabitów na sekundę, a powodem tego obniżenia prędkości transmisji leżą po stronie odbiorczej, czyli po stronie sprzętu komputerowego.

Wchodzące na rynek łącze USB 3.0 powinno osiągać prędkości 5 gigabitów/sekundę.

**FireWire 400** Interfejs ten pozwala na transfer z teoretyczną szybkością 400 megabitów na sekundę. Prędkości zbliżone do maksymalnych udaje się uzyskać (źródło: Gigapedia), gdyż złącze to obsługiwane jest przez specjalizowany hardware.

FireWire w wersji 800 może osiągać dwukrotnie wyższe prędkości transmisji.

**Camera Link** Jest w pełni profesjonalnym łączem o szybkości transmisji rzędu 4 gigabitów na sekundę. Nic zatem dziwnego, że jest droższe niż USB i FireWire i warto je stosować tylko wówczas, gdy rzeczywiście mamy do czynienia z wysokimi wymaganiami szybkości transmisji obrazów.

**Ethernet** Znane wszystkim kable, pozwalające podłączyć komputer do Internetu są coraz częściej wykorzystywane do omawianych tutaj celów, gdyż pojawiły się na rynku kamery przemysłowe wyposażone w odpowiednie złącze. Osiągnięcie gigabitowych prędkości transmisji nie jest zatem takie trudne. Można spodziewać się rozwoju tego sposobu transmisji obrazów z kamery do komputera. Jedynie duży poziom zakłóceń elektromagnetycznych i bardzo duże rozdzielczości obrazów w połączeniu z dużą prędkością ich dostarczania przez kamerę może powodować konieczność sięgnięcia po kolejne narzędzie.

**Łącza światłowodowe** Zastosowanie łącza światłowodowego wiąże się z koniecznością zamiany sygnałów elektrycznych z kamery na sygnał świetlny i jego ponownej konwersji na sygnały elektryczne po stronie odbiorczej. I to właśnie ta podwójna konwersja limituje prędkość transmisji i podnosi koszty.



# Rozdział 3

## Odwzorowanie obrazu w pamięci

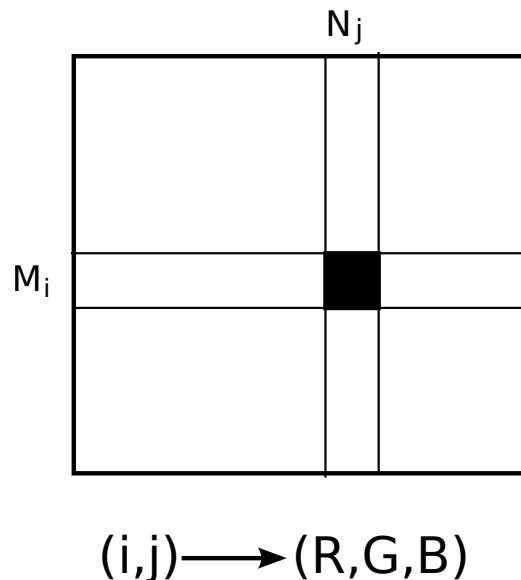
### 3.1 Obraz jako macierz

Obraz z matrycy trafia do bufora, a stąd przesyłany jest do pamięci komputera. W tym momencie tracimy informację o tym, że elementy światłoczuły matrycy CCD lub CMOS mają konkretne rozmiary. Oraz o związkach tych rozmiarów z geometrycznymi rozmiarami odwzorowanych na matrycy obiektów. Związki te zależą od odległości poszczególnych obiektów od kamery oraz od ogniskowej obiektywu. Mimo, że informacje te nie są przesyłane do komputera, powinniśmy o nich pamiętać, gdyż przydać się mogą do oceny rozmiarów obiektów lub defektów.

Przesyłana do komputera informacja dotyczy jedynie współrzędnych elementów światłoczułych oraz wartości liczbowych kolorów: czerwonego, zielonego i niebieskiego, które dalej oznaczać będziemy jako RGB. W tym momencie wygodnie jest ją traktować jako tablicę o rozmiarach  $M \times N$ . Na przecięciu  $i$ -tego wiersza i  $j$ -tej kolumny tej tablicy znajduje się trójka liczb, które reprezentują poziomy RGB w tym punkcie (por. Rys. 3.1).

W większości przypadków będziemy rozważać przetwarzanie w odcieniach szarości, gdyż zwykle to wystarcza, a przetwarzanie trwa krócej. Większość algorytmów można łatwo rozszerzyć na RGB przetwarzając każdy kolor oddzielnie.

W przypadku obrazów w odcieniach szarości, omawiana wyżej tablica staje się macierzą  $M \times N$  i z takimi właśnie reprezentacjami obrazów będziemy się najczęściej stykać w tej książce. Poziom szarości elementu  $(i, j)$  oznaczać będziemy przez  $I(i, j)$ . Posługując się będziemy także pojęciem piksela (od ang. *pixel*), rozumianego jako trójka  $(i, j)$  oraz  $I(i, j)$ . Parę  $(i, j)$  nazywać



Rysunek 3.1: Bufor, określenie współrzędnych

będziemy adresem piksela, a  $I(i, j)$  jego wartością. Warto zwrócić uwagę na fakt, że naturalna numeracja kolumn i wierszy macierzy zgodna jest z ogólną konwencją, stosowaną w przetwarzaniu obrazów, iż początek układu współrzędnych obrazu znajduje się w jego lewym górnym rogu.

Konwersja kolorów RGB na odcienie szarości realizowana jest według schematu:

$$I(i, j) = w_R \cdot P_R + w_B \cdot P_B + w_G \cdot P_G, \quad (3.1)$$

gdzie  $P_R, P_G, P_B$ , oznaczają intensywność barw RGB, natomiast  $w_R, w_G, w_B$  są współczynnikami liczbowymi – wagami. Proponuje się różne wagi, najprostsze to  $w_R = w_G = w_B = \frac{1}{3}$ , co zwykle daje sensowny rezultat. Wersja bardziej dopasowana do wrażliwości ludzkiego oka to:  $w_R = 0.29$ ,  $w_G = 0.587$ ,  $w_B = 0.114$ .

Do tej pory nie wskazaliśmy typu i zakresu liczb, które są używane jako wartości dla oznaczenia intensywaności barw i poziomów szarości. Najczęściej stosowana jest 8 bitowa arytmetyka z zakresem liczb naturalnych od zera do 255. W całej książce przyjęto następującą konwencję:

0 = kolor czarny

255 = kolor biały

Z powodu ograniczenia się do 8 bitów trzeba bardzo uważać z wykonywaniem operacji arytmetycznych (nawet dodawania i odejmowania).

W zależności od procesora, kompilatora itd. operacja (3.1) z wagami 1/3, wykonana dla wartości RGB = (100,100, 100) może dać w wyniku 100 lub

18 lub 82, jeśli nie zadbamy o prawidłowe deklaracje zmiennych.

Operacje mnożenia i dzielenia najprościej wykonać stosując wynikową zmienną 16-bitową. Następnie, należy dokonać normalizacji do zakresu [0, 255].

Pakiet typu MATLAB pozwalają stosować zarówno konwencję [0, 255] jak i liczby rzeczywiste z zakresu [0, 1], co wymaga dodatkowej uwagi od programisty.

Wśród wielu możliwych reprezentacji obrazu w odcieniach szarości warto odnotować także traktowanie obrazu jako funkcji  $f(x, y)$ , zależnej od dwóch współrzędnych przestrzennych  $(x, y)$ . Wówczas  $f(x, y)$  oznacza poziom szarości w punkcie  $(x, y)$ . Reprezentacja taka jest użyteczna głównie na etapie konstruowania lub objaśniania metody przetwarzania obrazów.

## 3.2 ROI

ROI (*region of interest*) – jest wirtualnie wyodrębnionym fragmentem obrazu, który zawiera tylko tę jego część, która jest interesująca i na której dokonywane jest przetwarzanie.

ROI zwykle definiowane przez czwórkę współrzędnych:

$i_x, i_y$ , szerokość, wysokość,

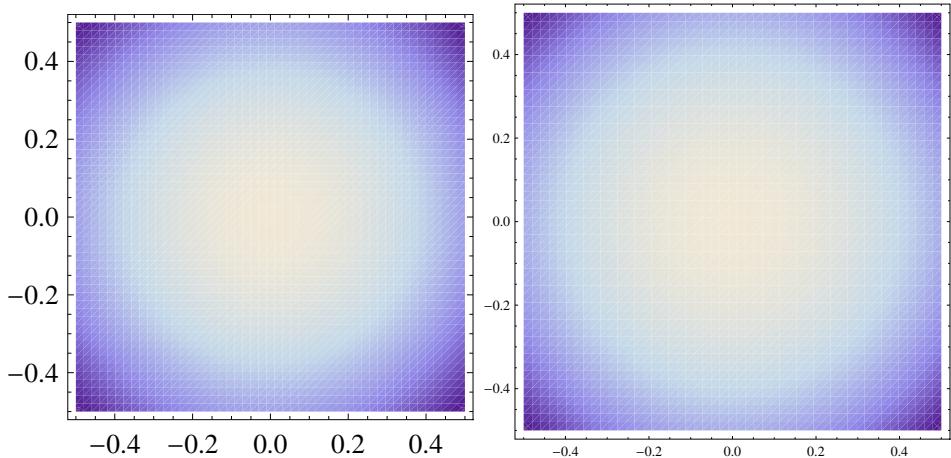
gdzie  $i_x$  oraz  $i_y$ , to współrzędne lewego górnego rogu ROI (we współrzędnych z  $(0, 0)$  w lewym górnym rogu oryginalnego obrazu).

Przy przetwarzaniu obrazów przemysłowych, gdy kamera i obiekty znajdują się stale w tym samym położeniu, warto skopiować ROI do osobnej tablicy i dalej traktować ją jako obraz i zwolnić pamięć zajmowaną przez obraz oryginalny. Podejście takie może znacznie zmniejszyć zajętość pamięci i przyspieszyć obliczenia. Dalej zakładając będziemy, że operacja taka została już wykonana i przyjęte wyżej rozmiary macierzy odnoszą się do tak przyjętego obrazu.

## 3.3 Uwagi o próbkowaniu i kwantyzacji

Znane twierdzenie Shannona o próbkowaniu, w wersji 2D, jest często przytaczane w podręcznikach jako narzędzie do wybierania rozdzielczości. Trzeba jednak zauważyć, że obrazy (traktowane jako funkcje  $f(x, y)$  na  $R^2$ ) nie mają ograniczonego przestrzennie pasma, ponieważ zwykle są nieciągłe, tam gdzie następuje zmiana intensywności pomiędzy obiektem a tłem. Skokowa zmiana intensywności implikuje teoretycznie nieograniczone widmo, które nie pozwala bezpośrednio stosować twierdzenia Shannona.

Co więcej, nawet tak gładkie obrazy jak ten po lewej na Rys. 3.2 nie mają



Rysunek 3.2: Obraz funkcji gaussowskiej, która – mimo gładkości – nie ma ograniczonego widma (rysunek po lewej) i funkcji  $\sin(r)/r$ , która ma ograniczone widmo ( $r = \sqrt{x^2 + y^2}$ ).

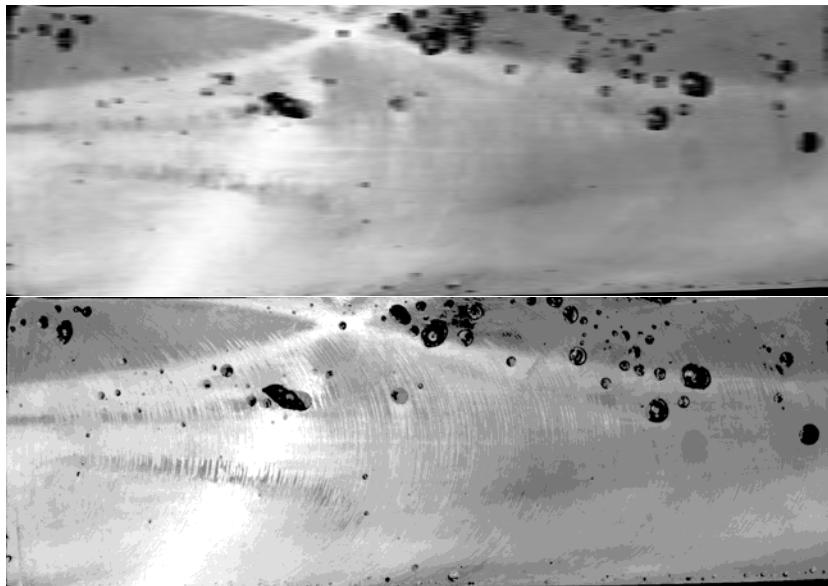
ograniczonego pasma. Dopiero obraz po prawej ma ograniczone pasmo. Jak zatem należy wybrać rozdzielczość? Podpowiedź pierwsza: ocenić jak małe obiekty (defekty) chcemy móc wykryć, zmierzyć itp. Wówczas, przyjmując odległość obrazu od kamery i ogniskową jej obiektywu, możemy określić niezbędną rozdzielczość obrazu.

Podpowiedź druga: obejrzeć widmo typowych obrazów, które będziemy przetwarzać i wybrać taką wartość częstotliwości przestrzennych powyżej, których energia zawarta w widmie jest już bardzo mała. Dla tak wybranej wartości można zastosować (w przybliżeniu) twierdzenie Shannona.

Jeśli na podstawie jednej lub obu wyżej zarysowanych metod uzyskamy rozdzielczość, która jest niższa niż ta, której dostarcza nam kamer, to można rozważyć redukcję przestrzennej rozdzielczości obrazu. Jednakże, pamiętać należy, że operacja taka zawsze pozbawia nas pewnej informacji i może wnosić dodatkowe zniekształcenia obrazu.

Przyjęty powszechnie, 8-bitowy zakres [0, 255] liczb naturalnych dość grubo odwzorowuje kolory RGB i odcienie szarości. Większość dostępnych obecnie kamer wykonuje więc za nas kwantyzację, czyli grupowanie wartości sąsiednich. Ten poziom kwantyzacji wystarcza do wielu zastosowań. Jednakże, zdarzają się przypadki, w których stosować trzeba drobniejsze ziarno kwantyzacji. Na rynku spotkać można kamery z przetwornikami 16-bitowymi.

Jeżeli defekty różnią się zasadniczo od tła, to zmniejszenie liczby poziomów kwantyzacji może być użyteczne. W innych przypadkach redukcja taka może zmniejszyć szanse lub uniemożliwić ich wykrycie. Na Rys. 3.3 pokazano



Rysunek 3.3: Przykładowe defekty w wersji 8-bitowej i 4-bitowej.

ten sam fragment slabu odwzorowany w skali 8-bitowej i 4-bitowej (zdjęcie dolne). Jak widać, redukcja do 4-bitowego przedziału poziomów szarości spowodowała, że widoczne z zdjęciu górnym w centrum slabu ślady po obróbce mechanicznej uległy zatarciu na zdjęciu dolnym.

## 3.4 Przechowywanie obrazów

### Obraz w pamięci RAM

Sposób przechowywania obrazów w pamięci operacyjnej ma bardzo duży wpływ na szybkość jego przetwarzania. Wybór sposobu zależy od organizacji dostępu programów do pamięci RAM.

Pamięć komputera ma charakter liniowy. Fakt ten w pełni wykorzystuje język C. Wybierając ten język, należy więc przyjąć konwencję zapisu macierzy w postaci ciągu liczb. Zwykle przyjmuje się zapis poszczególnych wierszy jeden za drugim, zaczynając od górnego wiersza. Rozpatrzmy przykład, macierz

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

zapisujemy jako

$$a_{11}, a_{12}, a_{13}, a_{21}, a_{22}, a_{23}, a_{31}, a_{32}, a_{33}$$

Sytuacja dla obrazów kolorowych jest analogiczna, jednak elementami macierzy są trójki liczb.

Dostęp do elementów macierzy wymaga prostych operacji arytmetycznych<sup>1</sup>, które symbolicznie zapisać można następująco:

$$a_{xy} = A[y * \text{width} + x].$$

### **Wzmianka o przechowywaniu obrazów na dysku**

W trakcie opracowywania systemu wizyjnego do monitorowania jakości produkcji często zachodzi potrzeba przechowywania obrazów przykładowych i testowych. Ponadto, już w trakcie pracy systemu, celowe jest archiwizowanie obrazów zawierających defekty.

Powstaje zatem konieczność pamiętania dużej liczby obrazów. Oprócz tablicy kolorów lub poziomów szarości, obrazy przechowywane na dysku muszą zawierać dodatkowe informacje. Stosuje się szereg formatów plików graficznych. Oto niektóre z nich

1. bitmapy Windows – BMP, DIB
2. JPEG – JPEG, JPG, JPE
3. Portable Network Graphics – PNG
4. Portable image format – PBM, PGM, PPM
5. Sun rasters – SR, RAS
6. TIFF – TIFF, TIF

Niewątpliwie najprostszym formatem jest format pgm – *portable gray map*, pochodzący z grupy formatów netpbm. W wersji tekstowej wygląda on następująco:

```
P2
# Komentarz
100 100
255
0 0 0 ...
.
.
.
```

---

<sup>1</sup>Wygodnie jest przyjąć konwencję, zgodnie z którą *width* oznacza szerokość, a *height* – wysokość.

Rozpoczyna się od oznaczenia formatu P2, następnie pamietane są wymiary obrazu i maksymalna wartość odcienia szarości. Odcienie szarości poszczególnych pikseli zapisywane są w formacie tekstowym w kolejnych liniach.

Opisany format pgm jest przykładem formatu, który nie stosuje kompresji. Przy przetwarzaniu istotne jest, by przetwarzać i przechowywać obrazy bez kompresji stratnej. W przeciwnym razie, utracie bądź rozmyciu ulec mogą istotne informacje o defektach. Ponadto, algorytmy kompresji stratnej często wnoszą artefakty, czyli widoczne ślady, które widoczne są na obrazie po dekompresji mimo, że nie było ich na oryginalnym obrazie (przykładem jest znany efekt blokowy, powstający w wyniku zastosowania kompresji z użyciem falek Haar'a). Można i warto natomiast stosować kompresję bezstratną. Jednym z popularnych formatów, oferujących kompresję bezstratną jest HuffYUV.



## Rozdział 4

# Transformacje punktowe

Trasformacje punktowe przekształcają bezpośrednio każdy punkt obrazu źródłowego w odpowiadający mu punkt obrazu docelowego. Oznacza to, że transformację taką można zapisać za pomocą funkcji. Ogólny schemat takiego algorytmu ma postać:

dla każdego punktu  $(i, j)$  wykonaj  
 $I'(i, j) = f(I(i, j)),$

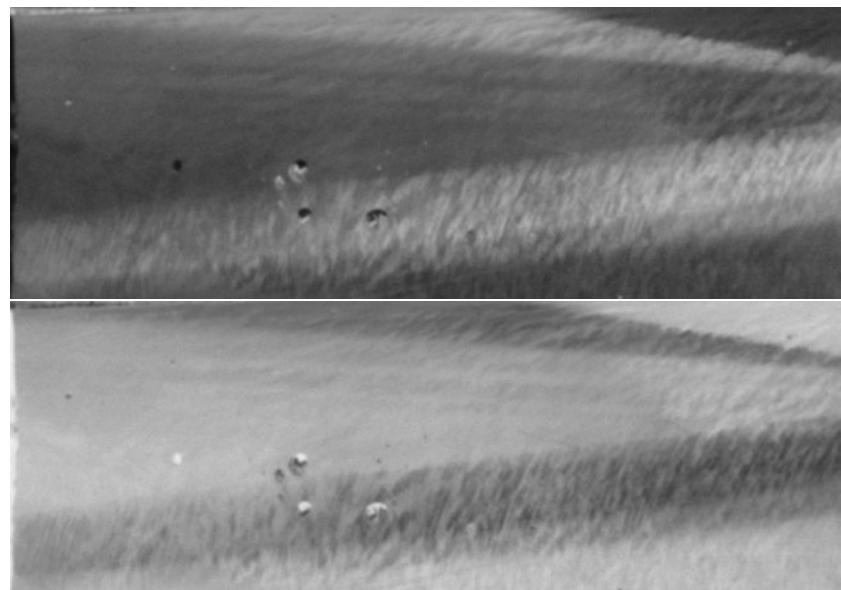
gdzie  $I(i, j)$  – obraz przed transformacją,  $I'(i, j)$  – obraz po transformacji. Dlatego też w tym rozdziale ograniczymy się do opisywania funkcji  $f$ . Należy tylko pamiętać, że  $(0, 0)$  jest w lewym, górnym rogu obrazu.

### 4.1 Obroty, odbicie lustrzane

Następujące operacje na całych obrazach, chociaż bardzo proste, są zawsze użyteczne:

- obrót,
- odbicie względem zadanej osi,
- negatyw obrazu  $I'(i, j) = 255 - I(i, j)$  (por. Rys. 4.1).

Obraz negatywowy z Rys. 4.1 nadal nie jest zadowalający, ale można na nim zobaczyć więcej defektów.



Rysunek 4.1: Obraz orginalny (u góry) i obraz z odwróconymi poziomami szarości (u dołu).



Rysunek 4.2: Obraz orginalny i o powiększonej jasności.

## 4.2 Jasność

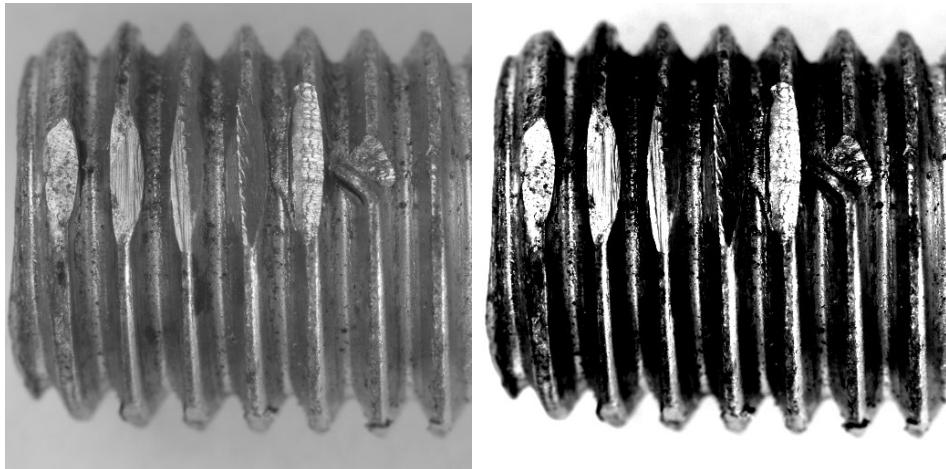
Jasność obrazu jest subiektywną oceną luminancji. Jasność obrazu można zmienić, zwiększając wartość poziomu szarości każdego piksela.

$$f(u) = \begin{cases} 0 & gdy u + i < 0 \\ 255 & gdy u + i > 255 \\ u + i & w pozostałych przypadkach \end{cases} \quad (4.1)$$

Skutek zmiany jasności pokazano na Rys. 4.2.



Rysunek 4.3: Obraz oryginalny i o zwiększym i zmniejszonym kontraście



Rysunek 4.4: Obraz oryginalny i o zwiększym kontraście, w tym przypadku zmiana jest bardzo potrzebna

## 4.3 Kontrast

Kontrast jest różnicą w jasności pomiędzy poszczególnymi obiektemi a tłem na obrazie.

Zasadniczo, zmiana kontrastu jest przeskalowaniem (rozsunięciem lub ścinięciem odcieni szarości).

$$f(u) = 127 + (u - 127) * 255 / (255 - 2 \cdot contr); \quad (4.2)$$

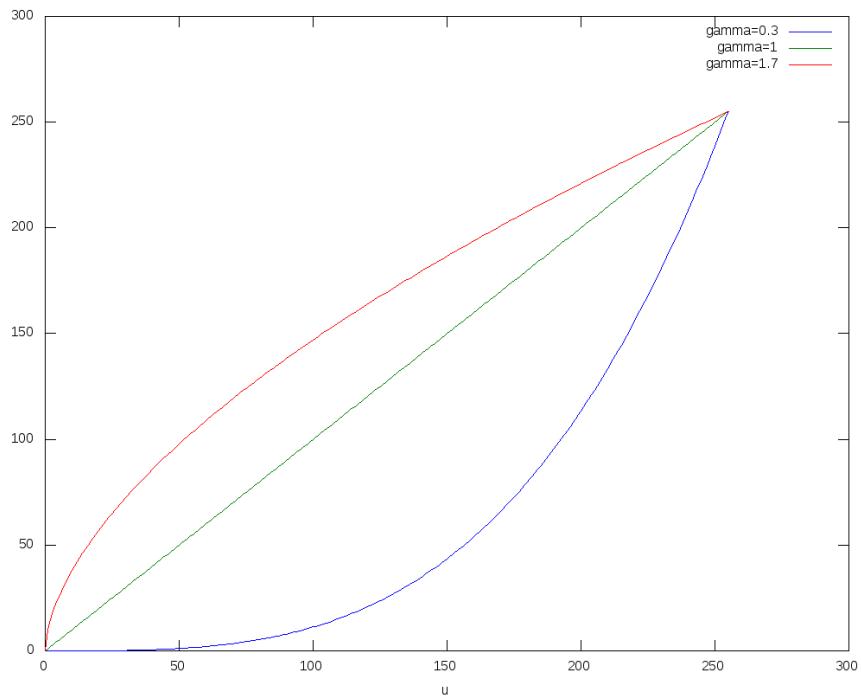
Przykłady zmiany kontrastu pokazano na Rys. 4.3 i Rys. 4.4.

## 4.4 Gamma

Gamma jest nieliniową (wykładniczą) zmianą jasności obrazu, realizowaną według wzoru:

$$f(u) = u^{1/\gamma} \cdot \frac{255}{255^{1/\gamma}}. \quad (4.3)$$

W celu zachowania zakresu [0,255] wynik musi być przeskalowany, co tłumaczy obecność drugiego czynnika we wzorze (4.3). Przykładowe przebiegi



Rysunek 4.5: Wykresy transformacji gamma dla  $\gamma = 0.3$ ,  $\gamma = 1.7$  oraz, dla porównania, dla  $\gamma = 1$

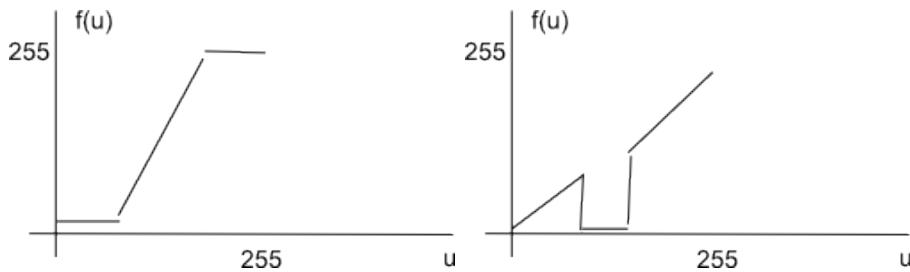
tej funkcji pokazano na Rys. 4.5, a wyniki zastosowania tych przekształceń pokazano na Rys 4.6.

## 4.5 LUT

LUT to powszechnie używany skrót od *look-up table*. Z jednej strony LUT jest w pewnym sensie najbardziej elastyczną z transformacji punktowych, gdyż funkcję  $f(u)$  podaje się jako tablicę, która podaje w jaki sposób poziomom



Rysunek 4.6: Obraz orginalny i po transformacji gamma dla  $\gamma = 0.3$  i  $1.7$



Rysunek 4.7: Przykładowe transformacje punktowe

szarości każdego piksela obrazu wejściowego przypisać wartości w obrazie po transformacji. Ponieważ zakres wejścia jest ograniczony (w przypadku odcieni szarości do  $[0, 255]$ ), to nie pojawiają się problemy obliczeniowe.

Na tym jednak nie kończą się zalety LUT. Dodatkową zaletą stosowania LUT jest szybkość realizacji transformacji, nawet wówczas, gdy pierwotnie funkcja  $f(u)$  dana jest skomplikowanym wzorem. Zysk czasu obliczeń może być jeszcze zwiększy, wtedy, gdy tablica trafi do pamięci cache. Tablice takie są stosowane także przy sprzętowym przetwarzaniu w układach FPGA.

**Zadanie.** Jak będzie wyglądał obraz w odcieniach szarości po zastosowaniu do każdego piksela transformacji za pomocą funkcji pokazanych na Rys. 4.7 ?



## Część II

**Wyszukiwanie obiektów i defektów:**

**Histogram poziomów szarości**

**Operacje arytmetyczne**

**Segmentacja – progowanie**

**Etykietowanie obiektów**

**Kontury obiektów**

**Transformacja Hough'a**



Część II niniejszej książki zajmuje w niej centralne miejsce, gdyż dotyczy wykrywania, etykietowania (numerowania) obiektów lub defektów.

Podstawowe klasy metod, które służą do wydzielania obiektów lub defektów to:

- segmentacja:
  - przez progowanie
  - morfologiczna – metodą działa wodnego,
- znajdowanie konturów,
- wykrywanie kształtów poprzez zastosowanie:
  - masek obiektów geometrycznych
  - transformacji Hough'a.

W części II omówimy większość z wymienionych wyżej metod wraz z algorytmami i przykładami ich działania na rzeczywistych przykładach.

Ograniczamy się do metod względnie prostych do zaprogramowania i takich, które mogą być stosowane on-line, gdyż wymagania czasowe w części zastosowań przemysłowych mogą być bardzo wysokie.



# Rozdział 5

## Proste narzędzia

W rozdziale tym omawiamy proste narzędzia statystyczne, które okazują się bardzo przydatne zarówno na etapie opracowywania algorytmów wykrywania defektów jak i w trakcie pracy on-line – jako elementy składowe całej procedury monitorowania jakości.

Zaczynamy od histogramu, który jest bezcennym narzędziem opisu obrazu. Następnie, przypominamy proste statystyki globalne, a w końcowym podrozdziale pokazujemy praktyczną użyteczność statystyk stosowanych lokalnie.

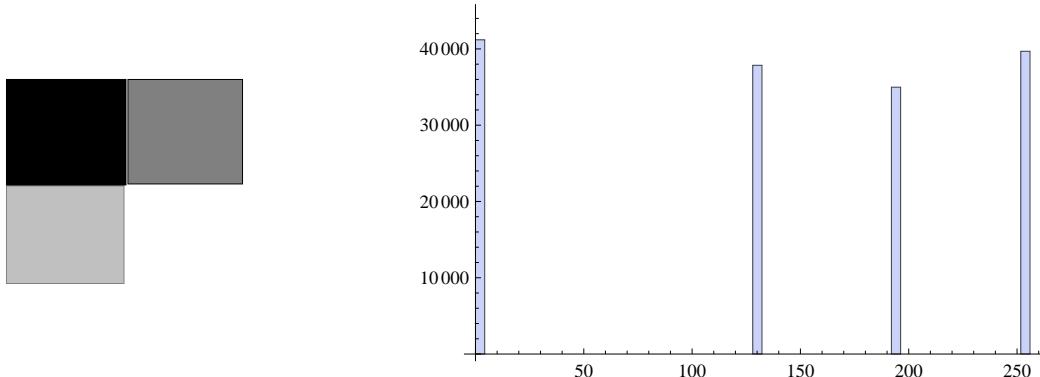
### 5.1 Histogram

Ze statystycznego punktu widzenia, empiryczny histogram jest nieparametrycznym estymatorem gęstości rozkładu prawdopodobieństwa. Tutaj histogram wprowadzony zostanie w sposób intuicyjny.

Histogram jest można traktować jako częstotliwości występowania danych odcięci (poziomów szarości) na obrazie. Można też (i tak robi się częściej) zliczać elementy z pewnego zakresu.

Histogram to narzędzie bardzo ważne i użyteczne w przetwarzaniu obrazów.

- Weź poziomy szarości  $I(i, j)$ , pomijając ich pozycje.
- Podziel przedział  $[0, 255]$  na podprzedziały  $[k \Delta, (k + 1) \Delta]$ , gdzie  $\Delta \geq 1$  wybrana szerokość przedziałów.
- Policz ile spośród  $I(i, j)$  wpada do przedziału  $[k \Delta, (k + 1) \Delta]$  i oznacz tę liczbę przez  $n_k$ .



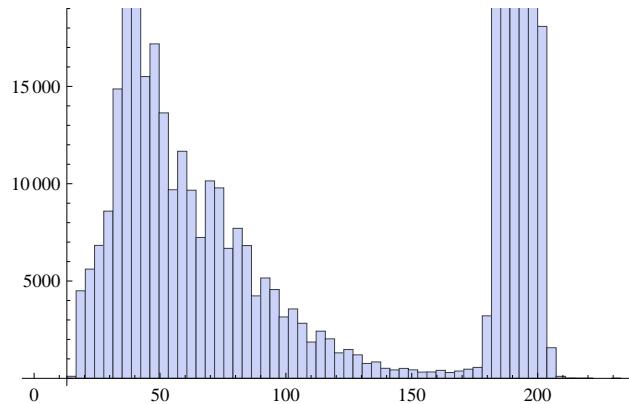
Rysunek 5.1: Nie patrząc na lewy obraz, możemy powiedzieć, że zawiera on tylko cztery odcienie szarości w podobnych proporcjach (ale nie możemy powiedzieć gdzie się znajdują)

- Wyświetl  $n_k$  jako funkcję  $k$  w postaci wykresu słupkowego. Alternatywnie, wyświetl  $n_k/(NM)$ , gdzie  $M$  i  $N$  to liczba wierszy i kolumn obrazu.

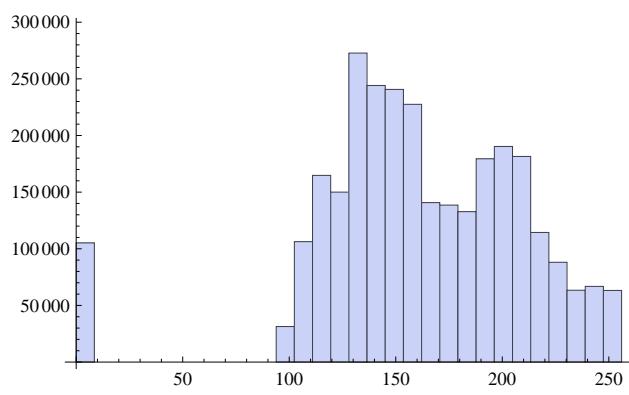
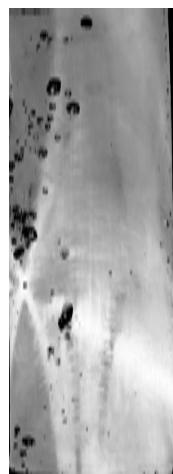
Z histogramem pozimów szarości (lub intensywności poszczególnych kolorów) będziemy się w tej książce jeszcze niejednokrotnie spotykać. Histogramów będziemy używać, między innymi, do:

- oceny poprawności naświetlenia obrazu i, w razie potrzeby, korekcji,
- odróżniania obiektów i defektów od tła,
- doboru progu (lub progów).

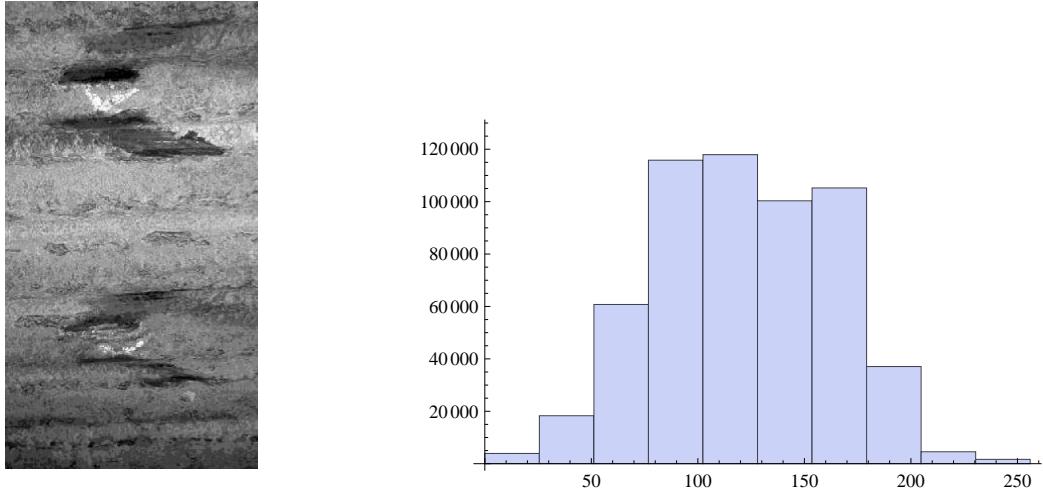
Prosty przykład histogramu pokazano na Rys. 5.1. Mimo swej prostoty pozwala on wykryć, że pozornie jednakowe kwadraty w istocie nie mają równych pól – kwadrat ciemno-szary jest wyraźnie mniejszy. Histogram pokazany na Rys. 5.2 dość łatwo pozwala odróżnić obiekt od tła. Natomiast histogram pokazany na Rys. 5.3 wskazuje, że duże czarne defekty różnią się znacząco od gładkiej powierzchni metalu. Następne dwa wykresy, pokazane na Rys. 5.4 i Rys. 5.5, ilustrują znany fakt, że dobór szerokości przedziałów histogramu jest bardzo ważny z punktu widzenia jego użyteczności. Zbyt szerokie przedziały (Rys. 5.4) niosą zbyt mało informacji o tym, co znajduje się na obrazie. Przedziały zbyt wąskie dają również nieinformatywny histogram, gdyż duża część z nich jest pusta. Na Rys. 5.5 pokazano histogram z dość dobrze dobraną szerokością przedziałów. Można z niego odczytać, że na obrazie oryginalnym występują cztery punkty skupienia poziomów szarości – lokalne



Rysunek 5.2: Prosty przypadek – wzgórze po lewej odpowiada obiektowi, wieża po prawej to poziomy szarości tła



Rysunek 5.3: Zdarza się też sytuacja idealna – poziomy szarości defektów idealnie odseparowane od tła



Rysunek 5.4: Teraz odgadnięcie poziomu defektów nie jest tak łatwe –  $\Delta = 25$

maksima histogramu. Patrząc na zdjęcie po lewej stronie Rys. 5.5, stwierdzamy, że klasyfikacja sugerowana przez histogram w istocie odpowiada różnym stadium rdzewienia metalu.

## 5.2 Statystyki globalne dla obrazów

Na etapie opracowywania algorytmów wykrywania defektów na obrazach zawsze warto obliczyć podstawowe statystyki dla całego obrazu. Do grupy tej należy

$$I_{\max} = \max_{i,j} I(i,j), \quad I_{\min} = \min_{i,j} I(i,j). \quad (5.1)$$

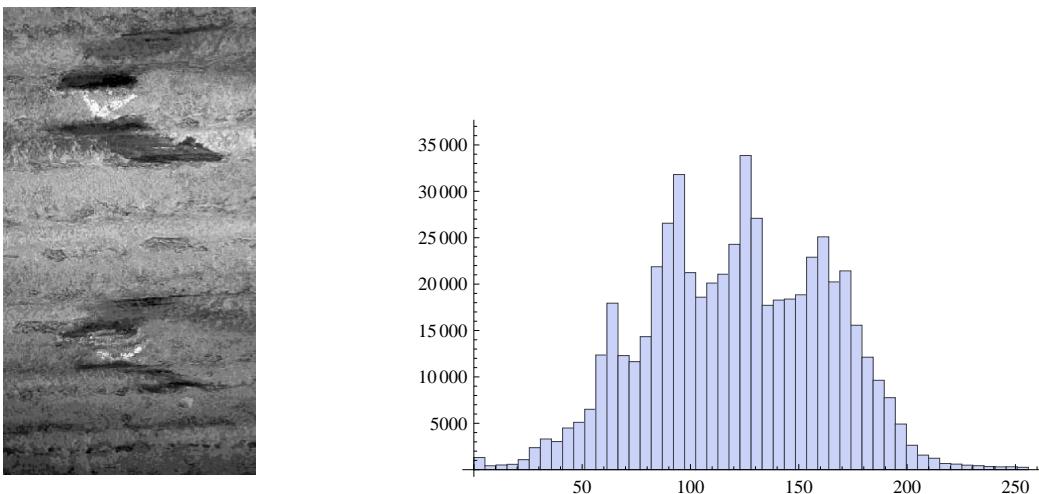
Maksymalna i minimalna wartość poziomów szarości pozwala się zorientować czy wykorzystujemy całą gamę poziomów szarości.

Średnia intensywność:

$$\tilde{I} = \frac{1}{M N} \sum_{i=1}^M \sum_{j=1}^N I(i, j),$$

może być użyta jako miara intensywności całego obrazu. Zwykle nie jest wystarczająco czuła, aby wykryć małe zmiany.

Empiryczną medianę intensywności obliczamy następująco: sortujemy wszystkie  $I(I, j)$ , pomijając ich położenie i uwzględniając powtarzające się elementy. Jako medianę wybieramy ten poziom szarości, który znajduje się w środku



Rysunek 5.5: Wybór  $\Delta$  jest bardzo ważny – teraz  $\Delta = 5$

listy, a jeśli liczba elementów jest parzysta, to liczymy średnią z dwóch poziomów szarości elementów najbliższych środka listy z lewej i prawej strony. Mediana, jako miara położenia rozkładu poziomów szarości, jest odporna na małe i średnie zmiany.

Rozrzut można mierzyć jako  $I_{max} - I_{min}$ . Inną możliwością jest empiryczna wariancja jasności:

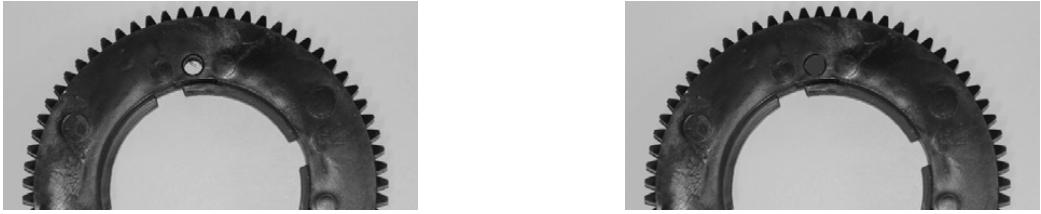
$$var(I) = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N (I(i, j) - \tilde{I})^2.$$

Wariancja jest miarą zmienności na obrazie. Zwykle nie jest wystarczająco czuła, aby wykryć małe zmiany.

Na Rys. 5.6 pokazano detal poprawny i detal uszkodzony (po lewej). Uszkodzenia polega na tym, że otwór nie został w pełni wykonany. W Tabeli 5.1 podano wartości podstawowych statystyk globalnych dla tych obrazów. Jak widać, statystyki te nie pozwalają wykryć defektu. Jednakże, jak pokazujemy w następnym podrozdziale, lokalne ich stosowanie okazuje się użyteczne, gdyż statystyki te stają się całkiem wrażliwymi wskaźnikami.

### 5.3 Lokalne zastosowanie prostych statystyk

Opisane w poprzednim podrozdziale proste statystki można i warto stosować lokalnie, obliczając je dla fragmentów obrazu. Fragmenty te można wybierać na wiele sposobów. Najprostsze z nich to:



Rysunek 5.6: Część dobra (po lewej) oraz uszkodzona (po prawej)

	Lewy	Prawy
Średnia	0.516	0.515
Medianą	0.721	0.721
Odczytanie standardowe	0.269	0.270

Tabela 5.1: Podstawowe statystyki obrazów pokazanych na Rys. 5.6.

1. kwadraty  $4 \times 4$  (lub nieco większe – dla obrazów o dużej rozdzielczości), pokrywające cały obraz, bez przecinania się,
2. całe wiersze obrazu,
3. całe kolumny obrazu.

W dalszej części tego podrozdziału zilustrujemy zastosowanie wersji 3 na przykładzie wykrywania defektów w slabie pokazanym w górnej części Rys. 5.7. Poniżej, na tymże rysunku, pokazano przebiegi poziomów szarości w przekrojach pionowych, czyli wartości wzduż kolumn macierzy  $I$ . idoczna jest duża zmienność tych przebiegów, ale daje się zaobserwować fragmnty o duzo niższych wartosciach poziomów szarości, które świadczyć mogą o występowaniu defektów powierzchni.

W celu ich wykrycia obliczono empiryczne średnie wzduż kolumn:

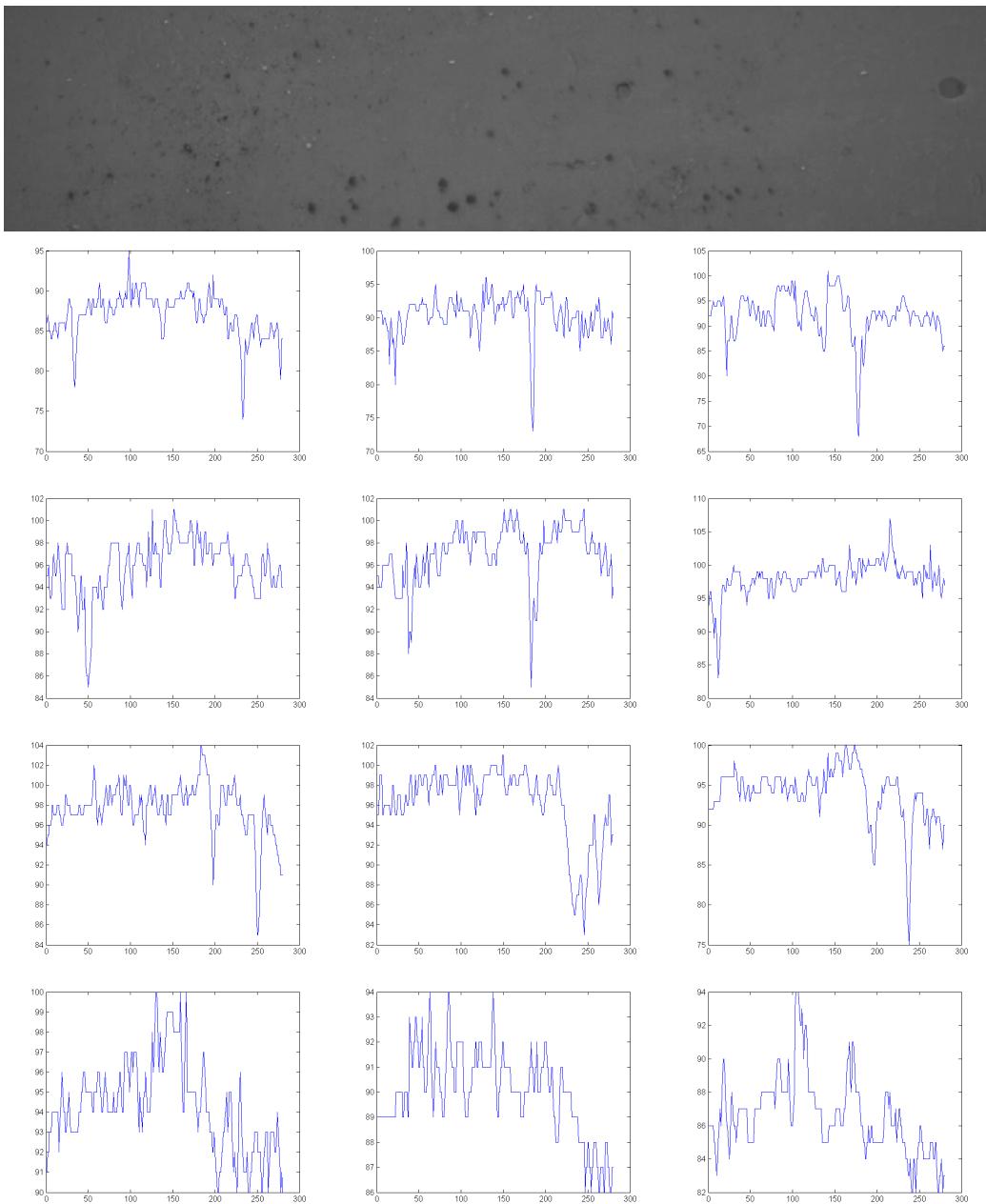
$$m_i = \frac{1}{N} \sum_{j=1}^N I(i, j), \quad i = 1, 2, \dots, M \quad (5.2)$$

oraz empiryczne mediany:

$$\mu_i = MED\{I(i, j), j = 1, 2, \dots, N\}, \quad i = 1, 2, \dots, M, \quad (5.3)$$

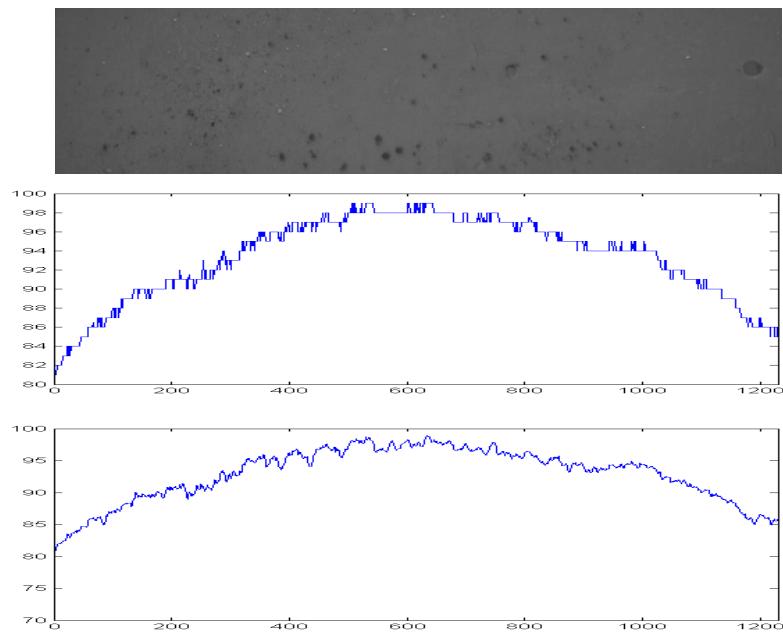
gdzie MED oznacza medianę liczoną tak, jak to opisano w poprzednim podrozdziale.

Wykresy  $m_i$  oraz  $\mu_i$  pokazano na Rys. 5.8 razem z powtórzonym zdjęciem slabu. Jak widać, lokalne średnie i mediany mają nieco niższe wartości tam,

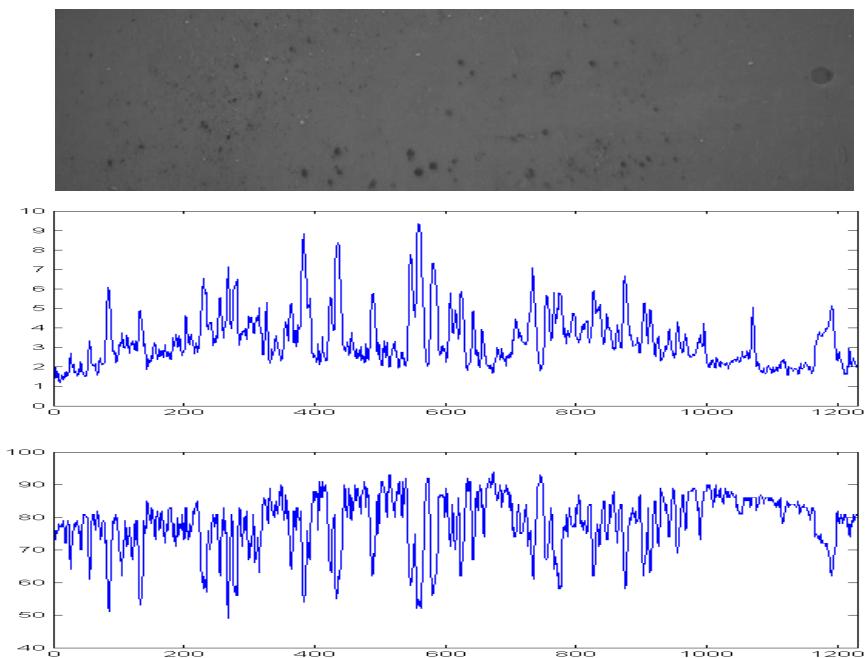


Rysunek 5.7: Obraz testowy i przekroje poziomów szarości – pionowe, co 100 pikseli wzdłuż slabu

gdzie występują defekty, lecz obniżenie to nie jest wyraziste. Mediana i średnia z poziomów szarości w przekrojach pionowych (por. Rys. 5.8) okazały się zbyt mało czułe, by wykryć defekty o mniejszych rozmiarach (wykrywają



Rysunek 5.8: Obraz testowy i przebiegi mediany (u góry) oraz średniej.



Rysunek 5.9: Obraz testowy i dyspersje oraz minima w przekrojach.

otwór w prawej części zdjęcia oraz – na przykład – grupę trzech charakterystycznych defektów przy dolnym brzegu, w środkowej części obrazu).

Obliczono także dyspersje wzduż kolumn:

$$s_i = \sqrt{\frac{1}{N} \sum_{j=1}^N (I(i, j) - m_i)^2}, \quad i = 1, 2, \dots, M \quad (5.4)$$

oraz minima wzduż kolumn

$$l_i = \text{MIN}\{I(i, j), j = 1, 2, \dots, N\}, \quad i = 1, 2, \dots, M. \quad (5.5)$$

obraz testowy i pokazano wykresy dyspersji i minimalnej wartości poziomów szarości, liczonej dla każdego przekroju pionowego. Maksima dyspersji i minima  $l_i$  wskazując, że w danej kolumnie spodziewać się można defektów. Statystyki te są bardziej wrażliwe na występowanie defektów niż średnia i mediana. Po odpowiedniej filtracji mogłyby służyć jako indykatory defektów.

W pracy [71] pokazano, że również lokalnie liczona entropia jest dobrym wskaźnikiem w omawianej sytuacji. Fraktalny charakter przebiegów z Rys. 5.7 był inspiracją zastosowania wykładnika Hursta, liczzonego dla każdego przekroju, jako wskaźnika istnienia defektu w danej okolicy.



# Rozdział 6

## Wykrywanie obiektów przez operacje arytmetyczne na obrazach

Celem tego rozdziału jest omówienie najprostszych operacji arytmetycznych na parach obrazów i pokazanie w jaki sposób mogą być one wykorzystane do wykrywania obiektów. Wyróżnić można dwa podejścia. Pierwsze z nich zakłada, że kolejno napływające obrazy porównujemy z obrazem wzorcowym lub z obrazem zawierającym uprzednio przygotowane lub na bieżąco aktualizowane tło. Z tego względu, rozdział rozpoczynamy od prostych metod przygotowania tła. Podejście drugie bazuje na porównywaniu aktualnego obrazu z poprzednim.

W obu podejściach podstawową wersją są obrazy w odcieniach szarości lub poszczególne składowe RGB. Większość omawianych tu operacji ma sens również dla obrazów binarnych – czarno-białych. Omówienie operacji logicznych na obrazach binarnych przedstawimy w dalszych rozdziałach.

### 6.1 Obraz wzorcowy i szacowanie tła

Obrazem wzorcowym nazywać będziemy taki obraz, który służy do bezpośredniego porównania z kolejnymi napływającymi obrazami. Obraz taki zawiera zwykle wzorzec montażu danego produktu lub jego fragmentu. Z tego względu powinien być on szczególnie starannie przygotowany, a porównywane z nim obrazy powinny być pobierane w tym samym położeniu i przy takim samym oświetleniu. Jeśli spełnienie tych warunków nie jest możliwe, niezbędne są zabiegi polegające na dopasowaniu obrazów (ang. *registration*).

Napływające obrazy porównywać można ze specjalnie przygotowanym

obrazem tła. Tłem nazywać będziemy obraz, który zawiera jedynie elementy stałe, czyli takie, które zawsze pojawiać się będą także na napływających obrazach. Oprócz nich spodziewamy się, że pojawiać się będą nowe elementy, które mamy wykryć.

Najprostszym podejściem do uzyskania obrazu tła jest uśrednienie kilku lub kilkunastu kolejnych obrazów. Uśredniając trzeba pamiętać o ograniczeniach wynikających z faktu, że korzystamy z 8-bitowych zmiennych. Zatem, macierz służąca do uśredniania musi dopuszczać dłuższą reprezentację liczb.

W przypadku zmieniających się warunków (na przykład oświetlenia), celowe może być stałe aktualizowanie obrazu tła. Aby uniknąć nasumowywania kolejnych obrazów, celowe jest skorzystanie z rekurencyjnej wersji średniej arytmetycznej.

Wersję taka łatwo można wyprowadzić. Założymy, że mamy liczby  $z_i$ ,  $i = 1, 2, \dots, n$ . Ich średnia ma postać:

$$\bar{Z}_n = \frac{1}{n} \sum_{i=1}^n z_i.$$

Po pojawiению się liczby  $(n+1)$ -szej

$$\bar{Z}_{n+1} = \frac{1}{n+1} \sum_{i=1}^{n+1} z_i = \frac{1}{n+1} z_{n+1} + \frac{1}{n+1} \sum_{i=1}^n z_i = \frac{z_{n+1}}{n+1} + \frac{1}{n+1} \cdot \frac{n}{n} \sum_{i=1}^n z_i,$$

co w wyniku daje:

$$\bar{Z}_{n+1} = \frac{n}{n+1} \bar{Z}_n + \frac{1}{n+1} z_{n+1}, \quad (6.1)$$

Dla  $n = 0$  mamy  $\bar{Z}_1 = z_1$ .

Stosując (6.1) do tworzenia tła, wystarczy przechowywać obraz aktualnego tła i dokonywać aktualizacji. Wadą tej metody jest powstawanie błędów zaokrągleń.

Zauważmy, że w (6.1) współczynniki wagowe sumują się do 1. Nasuwa to myśl zastosowania tzw. wykładniczo ważonej ruchomej średniej, która ma postać:

$$\tilde{Z}_{n+1} = (1 - \alpha) \tilde{Z}_n + \alpha z_{n+1}, \quad (6.2)$$

gdzie  $0 < \alpha < 1$  jest wybieranym przez nas współczynnikiem wygładzania, a iteracje zaczynają się od  $\tilde{Z}_0 = z_1$ , co – dla  $n = 0$  – daje także  $\tilde{Z}_1 = z_1$ . Zwracamy uwagę, że ten sposób uśredniania daje inne wartości niż średnia arytmetyczna i – w zależności od wyboru  $\alpha$  – pozwala wolniej lub szybciej zapominać o przeszłych danych. Stosując (6.2) do tworzenia tła, pamiętać trzeba o dostosowaniu wyniku do arytmetyki 8-bitowej.

## 6.2 Operacje na parach obrazów

Operacje na parach obrazów zwykle wymagają dwóch obrazów tego samego typu (oba w odcieniach szarości, binarne itd.). Oznaczmy obrazy wejściowe przez  $I, J$ , a obraz wynikowy przez  $Y$ . W praktyce, wynik można zapisać w miejsce jednego z obrazów wejściowych. Proste operacje obejmują:

- dodawanie,
- odejmowanie,
- kombinacje liniowe,
- dzielenie.

Przy ich wykonywaniu, należy brać pod uwagę możliwe problemy z zakresem zmiennych.

### 6.2.1 Dodawanie i odejmowanie par obrazów

**Dodawanie.** Dla każdego piksela

$$Y(i, j) = \text{Normalizacja}[I(i, j) + J(i, j)],$$

*Normalizacja* to operacja obcięcia wyniku do [0, 255].

**Odejmowanie** Dla każdego piksela:

$$Y(i, j) = \text{Normalizacja}[I(i, j) - J(i, j)],$$

gdzie *Normalizacja* do [0, 255] wykonywana jest przez zastąpieniem zerem wartości ujemnych.

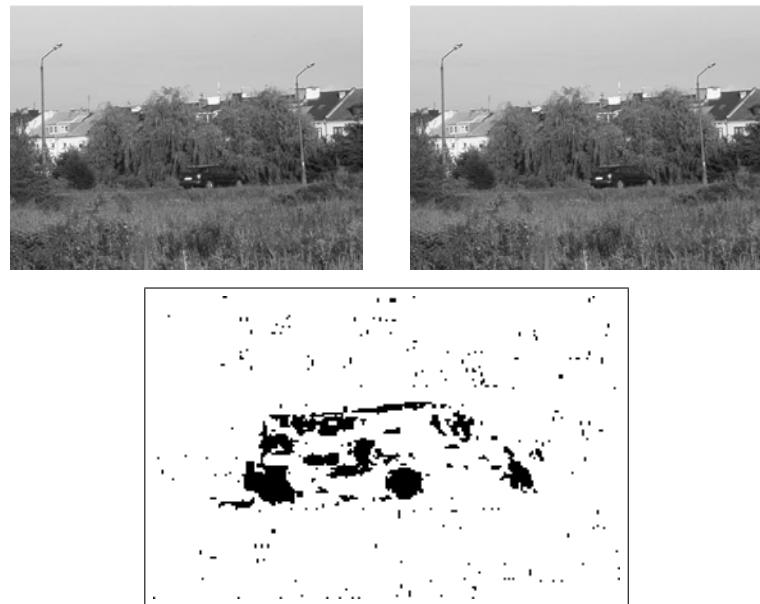
Odejmowanie jest bardzo użyteczne do wykrywania ruchu. Na Rys. 6.1 pokazano przykład wykrywania ruchu samochodu. Pod odjęciu obrazów obraz został przycięty, odwrócono poziomy szarości i wykonano binaryzację, w celu poprawienia widoczności wyniku.).

### 6.2.2 Dzielenie dwóch obrazów

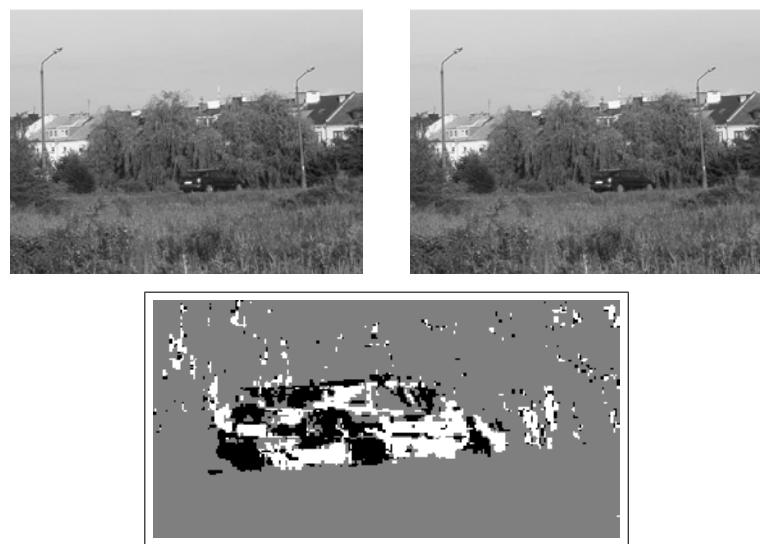
Dla każdego piksela

$$Y(i, j) = I(i, j)/J(i, j), \quad \text{gdy } J(i, j) \neq 0.$$

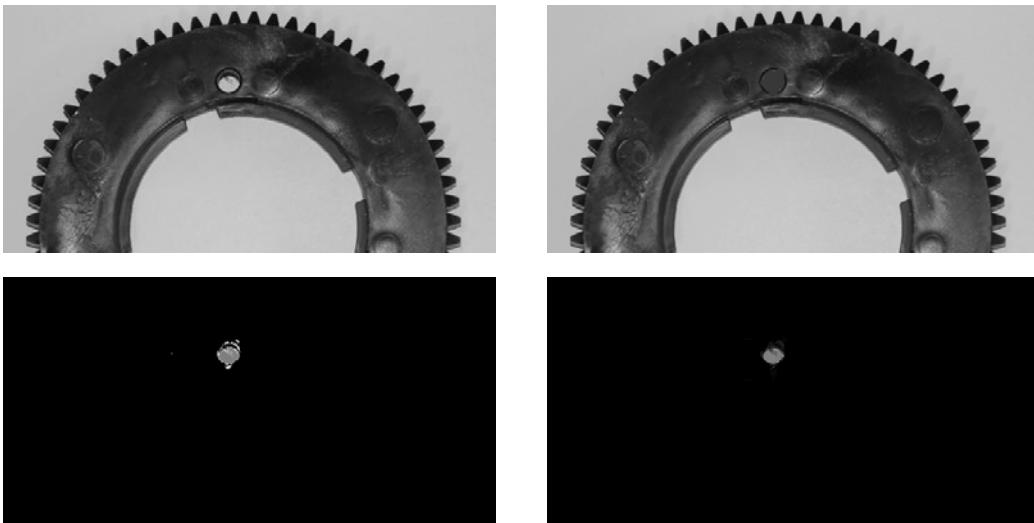
Również bardzo użyteczna operacja do wykrywania ruchu (por. Rys. 6.2). Poniżej zamieszczamy kod w systemie Matlab do wykrywania ruchu przez dzielenie.



Rysunek 6.1: Wynik odejmowania obrazów



Rysunek 6.2: Wynik dzielenia (przycięte i poprawione przed podzieleniem – dla lepszej widoczności)



Rysunek 6.3: Wynik odejmowania obrazów (po lewej) i wartości bezwzględnej ich różnicy.

```

nrfr=527;
mov = aviread('c:film.avi', nrfr);
[RGB,colormap]=frame2im(mov);
Icur=imadjust(rgb2gray(RGB));
mov = aviread(film, nrfr-2);

[RGB,colormap]=frame2im(mov);
Iback=imadjust(rgb2gray(RGB));
idiv=imdivide(Icur,Iback);
figure, imshow(Iback);
figure, imshow(Icur);
d=imcomplement(idiv);
figure, imshow(d)

```

**Poprawa tła przez dzielenie.** Przyjmijmy, że niejednorodność tła zakłóżca obraz  $I(i, j)$  w następujący sposób:

$$Y(i, j) = L(i, j) \cdot I(i, j), \quad (*)$$

gdzie  $L_{ij}$  jest (zwykle nieznaną) funkcją zakłócającą  $Y(i, j)$  – piksele obserwowanego obrazu. Aby w przybliżeniu odtworzyć  $I(i, j)$  postępujemy w następujący sposób:

- w tych samych warunkach jak dla (\*) pobieramy zdjęcie jednolitego obrazu o intensywności  $C > 0$  i zapamiętujemy odpowiednie  $Y'(i, j)$ .

- Potem  $L(i, j) = Y'(i, j)/C$ .
- Następnie, dzielimy  $Y(i, j)/L(i, j)$ .

### 6.2.3 Inne operacje na parach obrazów:

- mnożenie:

$$Y(i, j) = \text{Normalizacja}[I(i, j) \cdot J(i, j)],$$

- wartość bezwzględna różnicy

$$Y(i, j) = \text{Abs}[I(i, j) - J(i, j)],$$

- maksimum (wybierz jaśniejsze):

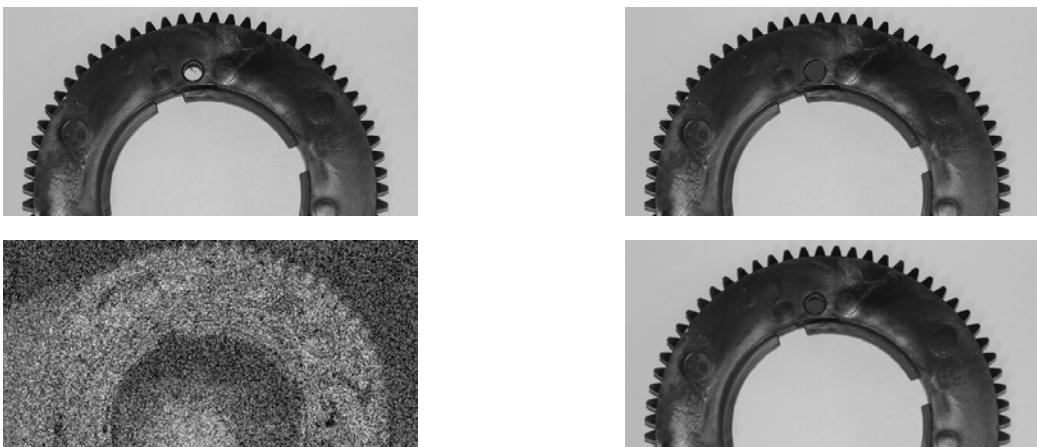
$$Y(i, j) = \text{Max}[I(i, j), J(i, j)],$$

- minimum (wybierz ciemniejsze):

$$Y(i, j) = \text{Min}[I(i, j), J(i, j)],$$

### 6.2.4 Przykłady operacji na dwóch obrazach

Na Rys. 6.3 i 6.4 pokazano efekty zastosowania prostych operacji arytmetycznych do wykrywania defektu. Część prawidłowa wyświetlna jest po stronie lewej, nieprawidłowa – po prawej. W części nieprawidłowej otwór w górnej części nie został wykonany.



Rysunek 6.4: Wynik mnożenia obrazów (po lewej) i operacji minimum (po prawej)

Operacje pokazane na Rys. 6.3 doprowadziły do wykrycia defektu. Uwaga: sukces był możliwy ze względu na bardzo dokładne pozycjonowanie.

Jak można było się spodziewać, mnożenie i minimum, pokazane na Rys. 6.4, nie pozwoliły na wykrycie defektu.

**Pytanie:** która z opisanych wyżej operacji da w tym przypadku najlepszy wynik ?



# Rozdział 7

## Obiekty i defekty jako grupy pikseli

Celem tego rozdziału jest przedstawienie podstawowych metod służących do wydzielenia i zaznaczenia tych pikseli, które należą do obiektów (lub defektów). Aby było to możliwe, obiekty muszą różnić się od otoczenia poziomem szarości lub kolorami, które zwykle i tak przekształcamy do poziomów szarości. Tak więc, obiekty (defekty) traktujemy jako jako przestrzennie spójne grupy pikseli, różniące się od otoczenia poziomem szarości.

### 7.1 Wstęp

#### 7.1.1 Segmentacja

Proces wyróżnianie w obrazie obszarów o zbliżonym kolorze lub poziomie szarości nazywać będziemy segmentacją.

Pojecie segmentacji w takim rozumieniu prowadzi do wielu nieporozumień. Źródłem ich jest ludzka percepcja. Nasze mózgi mają zdolność natychmiastowej interpretacji tego co widzimy. Stąd także po procedurach segmentacji spodziewamy się często zbyt wiele. Z drugiej strony, ograniczenie się tylko do kolorów lub poziomów szarości może nie pozwalić na odróżnienie zielonych krzewów od trawy – potrzebna jest analiza struktury liści, którą człowiek wykonuje niemal bezwiednie.

Na szczęście, w odniesieniu do obrazów przemysłowych rzadziej mamy do czynienia z takimi sytuacjami i postrzelenie segmentacji zwykle wystarcza.

Segmentacja służy do znajdowanie obiektów takich jak:

- części do skontrolowania,

- defektów

i jest jednym z ważniejszych zadań w kontroli jakości za pomocą kamer.

Segmentacja przez progowanie i znajdowanie konturu to metody, które zaznaczają obiekty niezależnie od ich kształtów. Korzystają wyłącznie z poziomów szarości lub kolorów.

### 7.1.2 Przetwarzanie obrazu przed progowaniem

Poprawa tła Tło można poprawić stosując jedno (bądź obie) z poniższych metod

- przez transformację prowadzącą do wyrównania (equalizacji) histogramu,
- przez techniki poprawy tła wspomniane poniżej.

Jeśli możemy pobrać sam obraz tła, nazwijmy go  $B(i, j)$ , to możemy:

- bezpośrednio odjąć  $B(i, j)$  od następnego obrazu, lub
- dopasować (metodą najmniejszych kwadratów) prosty model  $a \cdot i + b \cdot j + c$  do  $B(i, j)$  i użyć go do poprawek.

Stosuje się również korekcję przez dzielenie.

Filtrowanie obrazów daje następujące efekty:

- redukcję szumu,
- mody histogramu są wyższe i lepiej rozdzielone, jednak tylko wtedy, gdy używane są filtry w małym stopniu rozmywające krawędzie – jak, na przykład, filtr medianowy.

## 7.2 Progowanie

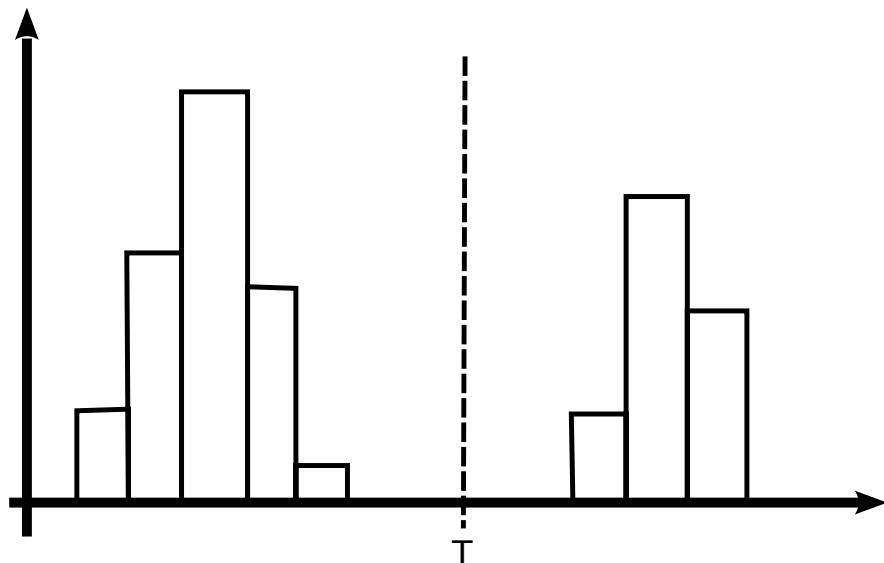
Progowanie jest operacją zamiany obrazu w odcieniach szarości na czarno-biały, kodowany zwykle jako obraz binarny: 0 (czarny) i 1 (biały). Koncepcyjnie, progowanie jest prostą operacją: dla każdego piksela sprawdzamy:

$$\begin{aligned} I(I, j) < T &\mapsto Y(i, j) = 0 \\ I(i, j) \geq T &\mapsto Y(i, j) = 1 \end{aligned} \tag{7.1}$$

gdzie  $Y$  jest obrazem wynikowym. Trudność polega na doborze progu  $T > 0$ .

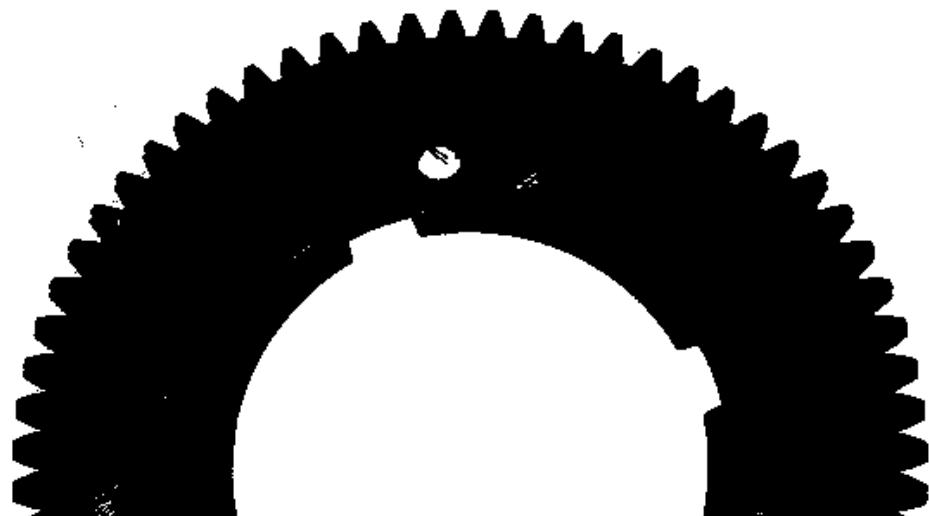
### 7.2.1 Dobór progu na postawie histogramu

W podręcznikach proponuje się zwykle, by w celu doboru progu sporządzić histogram (por. Rozdział 5.1) poziomów szarości. Jeśli histogram ten jest dwu-modalny (posiada dwa lokalne maksima), to  $T$  wybieramy pomiędzy nimi. Taką modelową sytuację pokazano na Rys. 7.1.

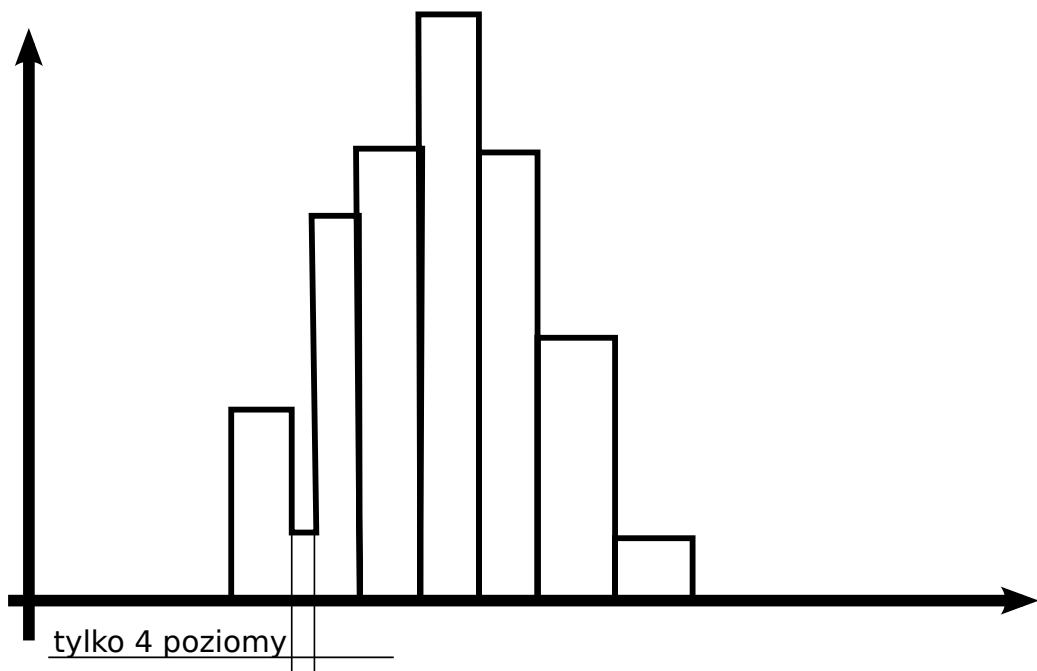


Rysunek 7.1: Histogram, pozwalający na łatwe ustalenie progu

Podobna sytuacja występuje na Rys. 5.2. Jest to histogram znanego nam detalu. Wyznaczamy próg o wartości  $T = 177$ . Binarny obraz, będący rezultatem progowania, przedstawiono na Rys. 7.2



Rysunek 7.2: Obraz po progowaniu –  $T = 177$



Rysunek 7.3: Histogram obrazu, dla którego znalezienie progu jest trudne.

Często jednak wybór progu nie jest tak jednoznaczny. Sytuacja jest bardziej skomplikowana, gdy płynne przejścia tonalne utrudniają znalezienie progu. Przykład takiego histogramu pokazano na Rys. 7.3. Ponadto, dobrze zbyt dużej szerokości pojedynczego słupka histogramu może taki punkt ukryć.

### 7.2.2 Inne warianty progowania

W przypadku, gdy obiekt lub defekt otoczony jest tłem o różnych poziomach szarości, warto rozważyć progowanie z dwoma progami  $T_1 > T_2 > 0$  o postaci:

$$\begin{array}{ll} T_1 \geq I(i, j) \geq T_2 & \mapsto Y(i, j) = 0 \\ w.p.p. & \mapsto Y(i, j) = 1. \end{array} \quad (7.2)$$

Stosowane bywa także tzw. progowanie „z histerezą”. Polega ono na tym, że badamy nie tylko czy  $I(i, j)$  spełnia wymóg, ale też czy poziomy szarości sąsiednich pikseli mieszczą się w odpowiednim przedziale.

Pewnym wariantem jest wcześniejsze zmniejszenie wpływu tła. Możemy utworzyć wzorzec tła poprzez uśrednienie kilku obrazów bez defektów. Należy jednak zwrócić szczególną uwagę na sposób uśredniania. Możliwe są dwa warianty zmniejszenia wpływu tła:

- poprzez odejmowanie: (obraz bieżący minus tło)
- poprzez dzielenie: (obraz bieżący / tło).

Opisane wyżej techniki doboru progu mogą być stosowane wówczas, gdy mamy do czynienia ze stabilnymi warunkami oświetlenia, a poszukiwane defekty (obiekty) na kolejnych obrazach mają podobne poziomy szarości. Jeśli warunki te nie są spełnione, konieczny jest automatyczny dobór progu.

## 7.3 Automatyczny dobór progu

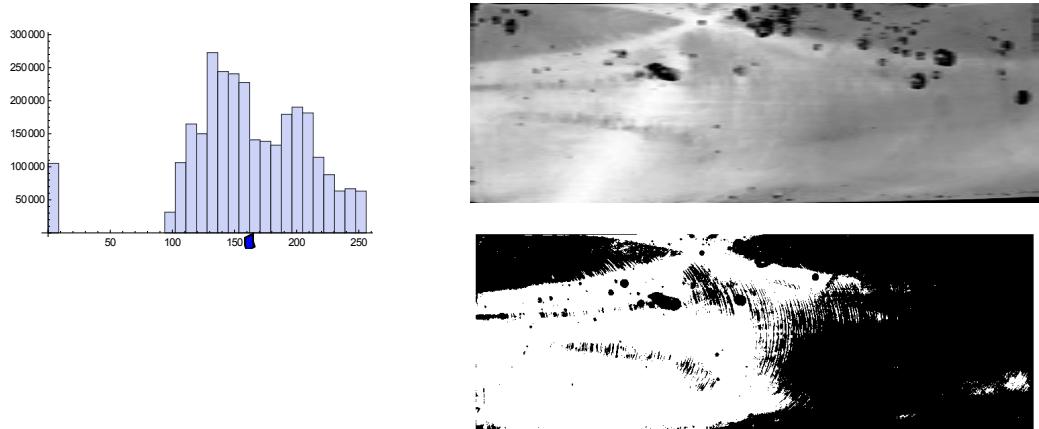
Termin automatyczny dobór progu oznacza iż dobór progu wykonywany jest programowow, zgodnie z wcześniej wybranym algorytmem, a wynik może się zmieniać, zależnie od zawartości obrazu.

Metody automatycznego wyboru progu bazują zwykle na następujących założeniach:

- histogram jest dwumodalny (dwa maksima),
- obiekty (defekty) są czarne lub ciemne,
- tło jest białe lub jasne.

Role jasnej i ciemnej barwy w dwóch ostatnich założeniach mogą być zamienione.

Jeśli te założenia nie są spełnione, należy spodziewać się złego wyboru progu. Dalej rozważymy pewne możliwości poprawy sytuacji.



Rysunek 7.4: Slab z defektami i jego histogram z zaznaczonym progiem o wartości  $T = 164$ , który został wybrany metodą Otsu. Pokazano także wynik progowania.

### 7.3.1 Metoda Otsu

Wybór  $T$  metodą Otsu bazuje na maksymalizacji wariancji między klasami: poziomów szarości pikseli zaliczonych do klas poniżej i powyżej progu (por. [13] str. 64). Dokładniej omówimy tę metodę na stronie 90. Wspominamy o niej teraz, gdyż – wobec jej popularności – stanowi dobry punkt odniesienia do porównań.

Metoda ta może dawać nieakceptowalne wartości progów wówczas, gdy liczba pikseli tła jest wielokrotnie wyższa niż liczba pikseli odpowiadających defektom. Sytuację taką zilustrowano na Rys.rysotsu, gdzie zaznaczono próg o wartości  $T = 163$  i niezadowalający wynik progowania.

### 7.3.2 Podstawowa metoda wyboru progu

Podstawowa metodą wyboru progu jest podobna do metody 2-średnich, ale nie jest z nią identyczną. W metodzie tej próg dobierany jest iteracyjnie w następujący sposób.

**Krok 1** Przyjmij  $n = 1$  i wybierz próg startowy  $T_0$ . Wybierz dokładność  $\epsilon \geq 1$ .

**Krok 2** Wylicz  $m_1$  jako średnią tych poziomów szarości dla których:  $I(i, j) \leq T_{n-1}$  oraz  $m_2$  jako średnią tych, dla których  $I(i, j) > T_{n-1}$ .

**Krok 3** Ustal kolejne przybliżenie jako  $T_n = \text{Round}[(m_1 + m_2)/2]$ .

**Krok 4** Jeśli  $|T_n - T_{n-1}| < \epsilon$ , to STOP, w przeciwnym przypadku,  $n := n + 1$  i wróć do Kroku 2.

Uwagi do metody podstawowej

- Słowo „podstawowa“ nie ma znaczenia pejoratywnego.
- Rezultat zależy od punktu startowego  $T_0$ .
- W naszym poprzednim przykładzie: dla  $T_0 = 50$  metoda ta daje dość dobry rezultat  $T = 81$ , natomiast dla  $T_0 = 150$  uzyskujemy  $T = 165$ , które nie jest bardzo użyteczne w naszym przykładzie.

### 7.3.3 Metoda 2-średnich

Dobrze znana i stosowana do klasteryzacji metoda k-średnich może być zadaptowana jako metoda 2-średnich do doboru progu. Jej przebieg jest następujący.

**Krok 1** Ustaw  $n = 1$  i wybierz dwie średnie startowe  $\mu_1 \neq \mu_2$ . Przyjmij  $T_0 = \text{Round}[(\mu_1 + \mu_2)/2]$ . Wybierz dokładność  $\epsilon \geq 1$ .

**Krok 2** Wylicz  $m_1$  jako średnią tych  $I(i, j)$ , które są bliżej  $\mu_1$  niż  $\mu_2$ . Wylicz  $m_2$  jako średnią tych  $I(i, j)$ , które są bliżej  $\mu_2$  niż  $\mu_1$ .

**Krok 3** Ustaw  $T_n = \text{Round}[(m_1 + m_2)/2]$ . Jeśli  $|T_n - T_{n-1}| < \epsilon$  to STOP.

**Krok 4** W przeciwnym razie,  $n := n + 1$ ,  $\mu_1 := m_1$ ,  $\mu_2 := m_2$  i wróć do Kroku 2.

Uwagi o metodzie 2-średnich:

- Rezultat zależy od punktów startowych  $\mu_1$  i  $\mu_2$ .
- Metoda ta jest bardziej czasochłonne niż metoda podstawowa.
- Może być wolnozbieżna.
- Zastosowana do naszego przykładu:
  - dla punktów startowych  $\mu_1 = 25$  i  $\mu_2 = 153$  daje całkiem dobry rezultat  $T = 83$  (od razu w pierwszej iteracji),
  - jednak dla punktów startowych  $\mu_1 = 77$  i  $\mu_2 = 204$  prowadzi do  $T = 163$  (po 10 iteracjach), co nie jest wartością użyteczną w naszym przykładzie.

### 7.3.4 Wybór progu jako bayesowski problem decyzyjny

Wybór progu to problem decyzyjny. Można zatem zastosować klasyfikator bayesowski w wersji dostosowanej do przetwarzania obraów przez Davies'a [13].

Przymijmy, że gęstości rozkładów prawdopodobieństwa poziomów szarości obiektu  $k = 1$  i tła  $k = 2$  są gaussowskie:

$$g_k(x) = \frac{1}{\sqrt{2\pi}\sigma_k} \exp\left[-\frac{(x-\mu_k)^2}{2\sigma_k^2}\right]$$

Niech  $p_1$  i  $p_2$  będą prawdopodobieństwami a priori przynależności danego piksela do obiektu i tła, odpowiednio. Oczywiście:  $p_1 + p_2 = 1$ . Klasyfikator, który minimalizuje prawdopodobieństwo błędu klasyfikacji piksela jako tło lub obiekt ma postać: zaklasyfikuj piksel o poziomie szarości  $x$  do klasy *obiekt*, jeśli  $p_1 g_1(x) \geq p_2 g_2(x)$  i do klasy *tło* w przeciwnym przypadku. Punkt  $T$ , dla którego

$$p_1 g_1(x) = p_2 g_2(x) \quad (7.3)$$

jest progiem oddzielającym obie klasy. Dla rozkładu gaussowskiego równanie 7.3 ma dwa rozwiązania, ale tylko jedno z nich daje poprawne wartości progu:

$$T = \frac{\pm\sqrt{\sigma_1^2\sigma_2^2\left((\sigma_1^2 - \sigma_2^2)\log\left(\frac{p_1^2\sigma_1^2}{p_2^2\sigma_2^2}\right) + (\mu_1 - \mu_2)^2\right)} + \mu_2\sigma_1^2 - \mu_1\sigma_2^2}{\sigma_1^2 - \sigma_2^2}.$$

Jeśli  $\sigma_1^2 = \sigma_2^2$ , co rzadko zdarza się w praktyce, powyższy klasyfikator upraszcza się do:

$$T = \frac{1}{2} \left( \frac{\sigma^2 \log\left(\frac{p_1^2}{p_2^2}\right)}{\mu_1 - \mu_2} + \mu_1 + \mu_2 \right)$$

Jeśli dodatkowo  $p_1 = p_2$  to otrzymujemy całkiem intuicyjny wynik.

$$T = \frac{1}{2} (\mu_1 + \mu_2).$$

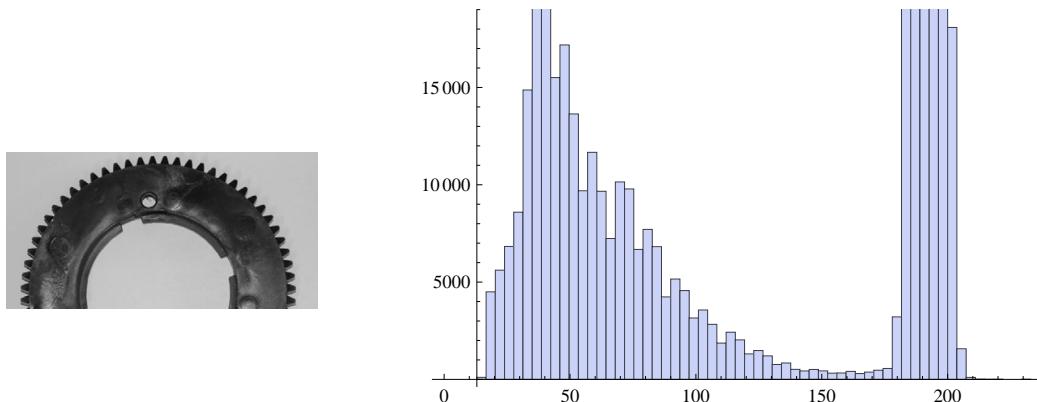
W zastosowaniach warto jednak brać pod uwagę również wariancje  $\sigma_1^2$ ,  $\sigma_2^2$ .

Uwagi na temat bayesowskiej metody wyboru progu.

- W praktyce przemysłowej obrazy są bardzo podobne. Nie jest trudno oszacować  $p_k$ ,  $\mu_k$  i  $\sigma_k$ ,  $k = 1, 2$  dla obrazów testowych.
- Rola prawdopodobieństw a priori jest kluczowa – to one kompensują zbyt dużą liczbę pikseli tła. Inne metody nie stosują takich prawdopodobieństw i próg jest zbyt blisko cięzszeego z rozkładów.

### 7.3.5 Przykład – bayesowski wybór progu

Na Rys. 7.5 powtórzono zdjęcie detalu i odpowiadający mu histogramu. Na oko wybraliśmy próg  $T \in [160, 170]$ .



Rysunek 7.5: Detal i histogram

Wybór obiektu z progiem  $T = 123$ , obliczonym metodą Otsu i 2-średnich pokazano na Rys. 7.6. Rezultat nie jest w pełni zadowalający – widoczne są dodatkowe białe miejsca u dołu detalu.



Rysunek 7.6: Progowanie metodą Otsu i 2-średnich;  $T = 123$

Spróbujmy zatem zastosować bayesowski dobór progu. Wprawdzie drugi z rozkładów („wieża”) nie jest gaussowski, lecz mimo tego potraktujemy go tak jakby był gaussowski. Parametry szacujemy na podstawie obrazu:

- $\mu_1 = 58.4, \sigma_1^2 = 687.8$ .
- $\mu_2 = 191.9, \sigma_2^2 = 33.3$ .
- $p_1 = 0.45, p_2 = 0.55$ .

Rozwiążając 7.3 otrzymujemy  $T = 231$ , które odrzucamy i  $T = 166$ , które jest poprawnym rozwiązaniem. Wynik zastosowania tego progu, pokazany na Rys. 7.7, jest znacznie lepszy niż uzyskany metoda Otsu i 2-srednich.



Rysunek 7.7: Wynik stosowania progu  $T = 166$ , otrzymanego za pomocą podejścia bayesowskiego

### 7.3.6 Jeszcze raz metoda Otsu

Niech  $n_l$  będzie liczbą pikseli mających poziom szarości  $l \in \{0, 1, \dots, 255\}$ . Jest jasne, że  $\sum_{l=1}^2 55n_l = N \cdot M$  – całkowita ilość pikseli. Zdefinujmy:

- $\hat{p}_l = n_l / (N \cdot M)$ ,
- $\hat{\pi}_0(t) = \sum_{l=1}^t \hat{p}_l$ ,  $\hat{\pi}_1(t) = 1 - \hat{\pi}_0(t)$ ,
- $\hat{\mu}_0(t) = \hat{\pi}_0^{-1}(t) \sum_{l=0}^t l \hat{p}_l$
- $\hat{\mu}_1(t) = \hat{\pi}_1^{-1}(t) \sum_{l=t+1}^{255} l \hat{p}_l$
- $\hat{\mu} = \sum_{l=0}^{255} l \hat{p}_l$  – grand mean.

Metoda Otsu: wybierz  $\hat{T}$ , dla którego wariancja międzyklasowa:

$$\hat{\sigma}_{IC}(t) = \hat{\pi}_0(t) (\hat{\mu}_0(t) - \hat{\mu})^2 + \hat{\pi}_1(t) (\hat{\mu}_1(t) - \hat{\mu})^2$$

osiąga maksimum względem  $t \in \{0, 1, \dots, 255\}$ . Widzieliśmy już jak ta metoda działa.

### 7.3.7 Zasada maksymalnej entropii

Zachowujemy poprzednie definicje i dodatkowo definiejemy.

- $H_0(t) = - \sum_{l=0}^t \left( \frac{\hat{p}_l}{\hat{\pi}_0(t)} \right) \cdot \ln \left( \frac{\hat{p}_l}{\hat{\pi}_0(t)} \right)$

- $H_1(t) = -\sum_{l=t+1}^{255} \left( \frac{\hat{p}_l}{\hat{\pi}_1(t)} \right) \cdot \ln \left( \frac{\hat{p}_l}{\hat{\pi}_1(t)} \right)$

- $H_{sum}(t) = H_0(t) + H_1(t).$

Zgodnie z zasadą maksymalnej entropii wybieramy takie  $\tilde{T}$ , dla którego  $H_{sum}(t)$  osiąga maksimum względem  $t \in \{0, 1, \dots, 255\}$ .

W celu uzasadnienia takiego sposobu wyboru progu, przypomnijmy pojęcie entropii. Dla dyskretnego rozkładu prawdopodobieństwa  $q_l \geq 0$ ,  $l = 1, 2, \dots, Q$ ,  $\sum_{l=1}^Q q_l = 1$  entropia  $H$  jest definiowana jako

$$H = -\sum_{l=1}^Q q_l \cdot \ln(q_l). \quad (**).$$

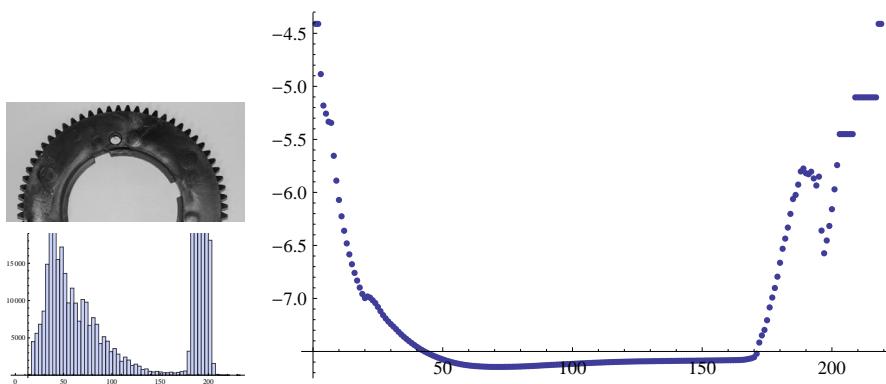
Potrzebne nam będzie jeszcze jedno pojęcie. Dla wektora losowego  $Y = (Y_1, Y_2)$ , wzajemna informacja  $I(Y)$  pomiędzy komponentami (nazywane odlegością Kullback-Leibler'a) pomiędzy rozkładem  $Y$  i rozkładami jego komponentów ma postać  $I(Y) = H(Y_1) + H(Y_2) - H(Y)$  i służy jako miara niezależności pomiędzy  $Y_1$  i  $Y_2$ .

Ciąg częstości występowania w danym obrazie poszczególnych poziomów szarości  $\hat{p}_l$ ,  $l = 0, 1, \dots, 255$  dzielony jest na dwa rozkłady przez wstępny próg  $t$ :

$$\hat{p}_l / \hat{\pi}_0(t), \quad l = 0, 1, \dots, t \quad (\text{lewy})$$

$$\hat{p}_l / \hat{\pi}_1(t), \quad l = (t+1), 1, \dots, 255 \quad (\text{prawy})$$

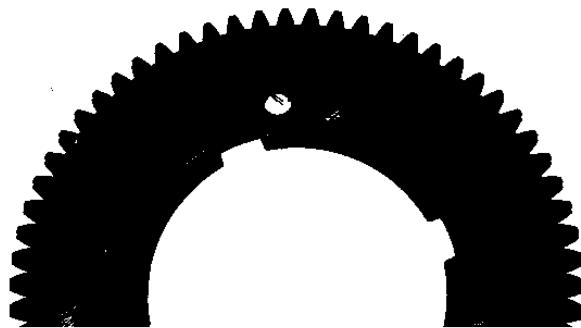
Trzeba zwrócić uwagę, że są one normalizowane przez ich sumy  $\hat{\pi}_0(t)$ ,  $\hat{\pi}_1(t)$ , odpowiednio. Tak więc, w naszym przypadku maksymalizacja  $H_{sum}(t)$  jest odpowiednikiem maksymalizacji stopnia statystycznej niezależności pomiędzy lewym i prawym rozkładem.



Rysunek 7.8: Histogram i odpowiadające mu  $H_{sum}(t)$

Uwaga: w praktyce maksimum  $H_{sum}(t)$  powinno być ograniczone do węższego zakresu odcieni (np. 40–210), gdyż dla  $t < 40$  i  $t > 210$  obliczana jest entropia prawie całego obrazu, która jest większa od entropii jego części.

Na Rys. 7.8 pokazano przebieg  $H_{sum}(t)$  – lokalne maksimum  $H_{sum}(t)$  w  $\tilde{T} = 174$  jest wyraźnie widoczne.



Rysunek 7.9: Próg  $\tilde{T} = 174$  – obliczony metodą maksimum entropii

Wynik progowania z  $\tilde{T} = 174$  – prawie idealnie wybiera obiekt.

### 7.3.8 Uwagi o wyborze metody progowania

- Metoda podstawowa, 2-średnich, Bayesowski próg i maksimum entropii są czasochłonne.
- Najlepiej stosować ją off-line aby wybrać próg, którego można używać on-line tylko wtedy, gdy:
  - analizujemy bardzo podobne obiekty (z możliwie małymi defektami)
  - zachowane są takie same warunki oświetlenia.
- Histogram typowego obrazu jest dwumodalny i pojedynczy próg wystarczy do odpowiedniego wybrania obiektu.

Jeśli te warunki nie są spełnione możemy wybrać jedno z poniższych podejść:

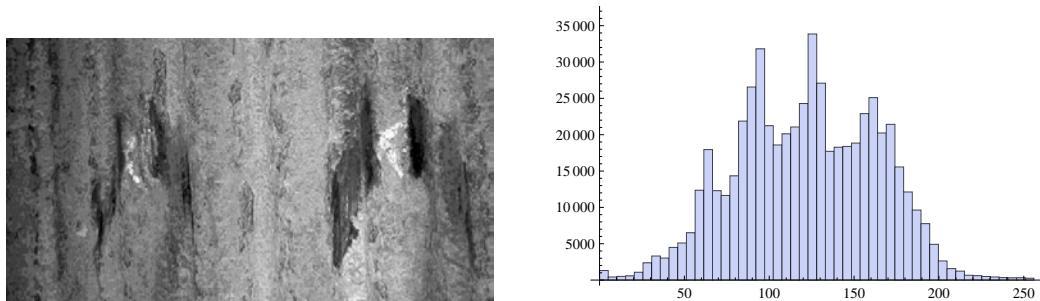
- zastosować prostszą metodę selekcji progu,
- podzielić obraz w podobrazy i wybrać progi lokalnie,

- wykorzystać więcej niż jeden próg do każdego piksela,
- użyć więcej niż jednego progu do piksela, biorąc pod uwagę poziomy szarości jego sąsiadów (progowanie z histerezą).

W dalszej części tego rozdziału omówimy:

- progowanie adaptacyjne,
- wpływ szumu na progowanie,
- konsekwencje istnienia więcej niż dwóch maksimów lokalnych.

Aby zilustrować ostatni przypadek, przypomnijmy histogram pokazany na Rys. 7.10.



Rysunek 7.10: Histogram rdzy

## 7.4 Progowanie adaptacyjne

Globalne progowanie zawodzi jeśli:

- kontrast między obiektem a tłem jest mały,
- obraz jest zaszumiony,
- jasność tła zmienia się znacząco w różnych częściach obrazu.

### 7.4.1 Progowanie adaptacyjne

Wróćmy do rozważań na temat doboru  $T$ . Metody „adaptacyjne“ lub „dynamiczne“ progowania zwykle nie są ani dynamiczne ani adaptacyjne – nazwy te są zwykle używane do różnych wariantów progowania lokalnego, które pozwala na różne  $T$  dla podobrazów lub nawet pojedynczych pikseli. W drugiej

z tych wersji informacja do wyboru lokalnego progu pobierana jest z zachodzących na siebie otoczeń. Pomijamy ten wariant, jako stosowany głównie do obrazów medycznych.

Ogólny schemat tak zwanego adaptacyjnego doboru progu jest następujący:

- podzielić obraz na podobrazy.
- dla każdego podobrazu zastosować jedną z opisanych wcześniej, prostych metod doboru progu,
- zapamiętaj obliczone progi próg i stosuj je do nadchodzących obrazów.

Nie należy stosować zbyt małych podobrazów (mniejszych niż 16x16 lub 32x32). W przeciwnym razie dane statystyczne nie są reprezentatywne.

#### 7.4.2 Progowanie wielopoziomowe

Podobnie jak poprzednio, zajmujemy się progowaniem pewnego fragmentu obrazu. W tym przypadku regiony te nie koniecznie są prostokątne – zależą od danych w pewien określony sposób.

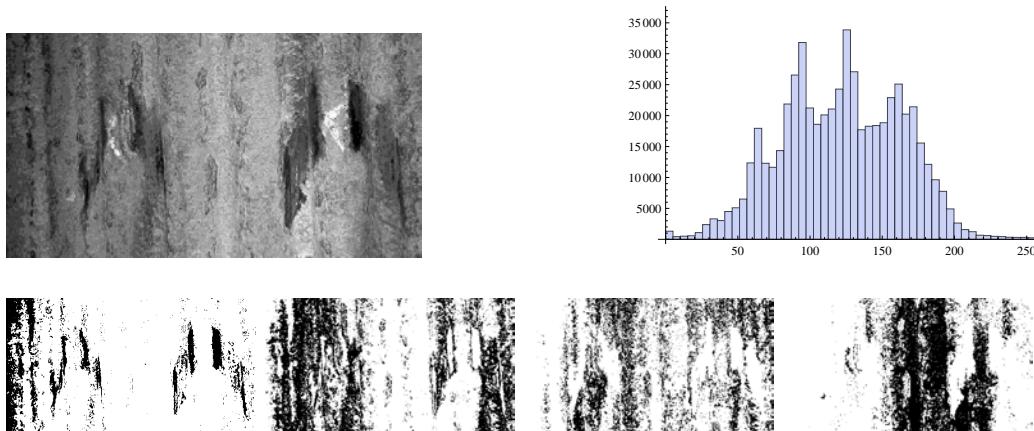
- Wybierz próg na jeden ze sposobów opisanych poprzednio i zastosuj go dla całego obrazu. Zapamiętaj położenia białych i czarnych obszarów.
- Wyznacz histogramy poziomów szarości tych fragmentów.
- Jeśli histogram jednego z tych obszarów jest dwumodalny, oblicz nowy próg i podziel jak powyżej ten obszar na mniejsze podobszary.
- Dziel podobszary dopuły, dopóki wszystkie staną się unimodelne.

#### 7.4.3 Progowanie dwustronne – przykład

Wspominaliśmy już, że czasem warto móc wybrać poziomy szarości obiektu zawarte pomiędzy  $T_1$  a  $T_2 > T_1$ . Przypomnijmy, że w tym przypadku, dla każdego piksela

$$Y(i, j) = \begin{cases} 0 \text{ (czarny)} & \text{gdy } T_1 \leq I(i, j) \leq T_2, \\ 1 \text{ (biały)} & \text{w.p.p.} \end{cases}$$

Jeśli histogram ma trzy mody i obiekty mają poziomy szarości blisko środkowej z nich, to można wybrać  $T_1$  w połowie między lewą a środkową modą. Analogicznie postępujemy dla  $T_2$ .



Rysunek 7.11: Wybór obszarów o różnym stopniu korozji (od lewej): [0-75], [75-120], [120-150], [150-255]

Na Rys. 7.11 pokazano użyteczność podwójnego progowania w sytuacjach, gdy defekty mogą być stopniowane. W tym przypadku można wybrać obszary w różnym stopniu skorodowania.

## 7.5 Jeszcze o progowaniu

### 7.5.1 Proste progowanie – przykład pozytywny

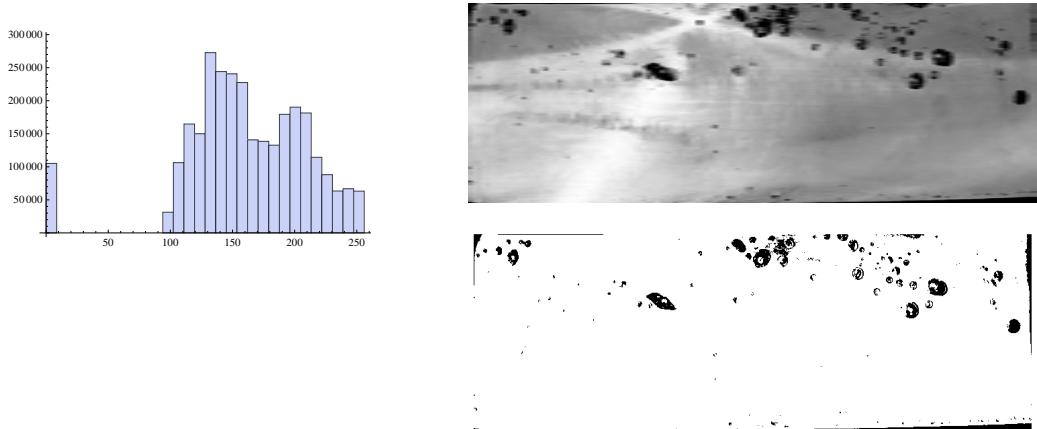
Na Rys. 7.12 pokazano przykład wykrywania defektów w slabie. Po analizie histogramu wybrano  $T = 50$ . Wszystkie defekty są zaznaczone, ale brzegi slabu także zostały zaznaczone jako defekty, gdyż obraz był źle skadrowany.

W podrozdziale tym koncentrujemy się głównie na metodach, które są używane w rozwiązaniach przemysłowej kontroli jakości. W takich rozwiązaniach przykładą się większą wagę do powtarzalności niż do całkowitej dokładności.

### 7.5.2 Dwie proste reguły

Najprostsza reguła, stosowalna jeśli mamy doczynienia z powtarzalnym procesem i sztucznym, dobrze dobranym oświetleniem, jest następująca.

- Gdy istnieją tylko dwa maksima lokalne w histogramie, to możemy wybrać  $T$  w połowie pomiędzy nimi..
- Reguła ta nie powinna być stosowana, gdy dwa wzgórza histogramu są zawierają bardzo różniące się liczby pikseli.



Rysunek 7.12: Przykład skuteczności prostego progowania

Następujące dalej metody działają dobrze w takich zastosowaniach gdzie położenie obiektu (np. czarnego) może się różnić, ale jego całkowity obszar pozostaje stały.

Druga z bardzo prostych reguł opiera się na powtarzalności procesów przemysłowych.

- Oblicz jaka część obrazu jest zwykle zajmowana przez obiekty (defekty). Wartość w procentach oznacz przez  $D$ .
- Wyznacz histogram  $\hat{p}_l$ ,  $l = 0, \dots, 255$ .
- Znajdź najmniejsze  $k$ , dla którego

$$\sum_{l=1}^k \hat{p}_l > D$$

i weź je jako  $T$ .

### 7.5.3 Metoda Trussell'a

Niech  $n_l$  będzie liczbą pikseli mających poziom szarości  $l \in \{0, 1, \dots, 255\}$ . Oczywiście jest, że  $\sum_{l=1}^{255} n_l = N \cdot M$  – całkowita liczba pikseli. Zdefiniujmy:

- $\hat{p}_l = n_l / (N \cdot M)$ ,
- $\hat{\pi}_0(t) = \sum_{l=1}^t \hat{p}_l$ ,  $\hat{\pi}_1(t) = 1 - \hat{\pi}_0(t)$ ,
- $\hat{\mu}_0(t) = \hat{\pi}_0^{-1}(t) \sum_{l=0}^t l \hat{p}_l$

- $\hat{\mu}_1(t) = \hat{\pi}_1^{-1}(t) \sum_{l=t+1}^{255} l \hat{p}_l$
- $\sigma_0^2(t) = \hat{\pi}_0^{-1}(t) \sum_{l=0}^t (l - \hat{\mu}_0(t))^2 \hat{p}_l$
- $\hat{\sigma}_1^2(t) = \hat{\pi}_1^{-1}(t) \sum_{l=t+1}^{255} (l - \hat{\mu}_1(t))^2 \hat{p}_l$

Zdefiniujmy także powszechnie znaną statystykę do testowania hipotezy o różnicę wartości oczekiwanych (średnich) dwóch populacji o rozkładzie normalnym:

$$Student(t) = \frac{|\hat{\mu}_0(t) - \hat{\mu}_1(t)|}{\sqrt{\frac{\sigma_0^2(t)}{N_0(t)} + \frac{\sigma_1^2(t)}{N_1(t)}}},$$

where  $N_0(t) = \sum_{l=0}^t n_l$ ,  $N_1(t) = \sum_{l=t+1}^{255} n_l$ . Statystyka ta mierzy odstęp między wartościami średnimi dwóch populacji, które tutaj wydzielamy z poziomów szarości obrazu poprzez wskazanie bieżącego progu  $t$ .

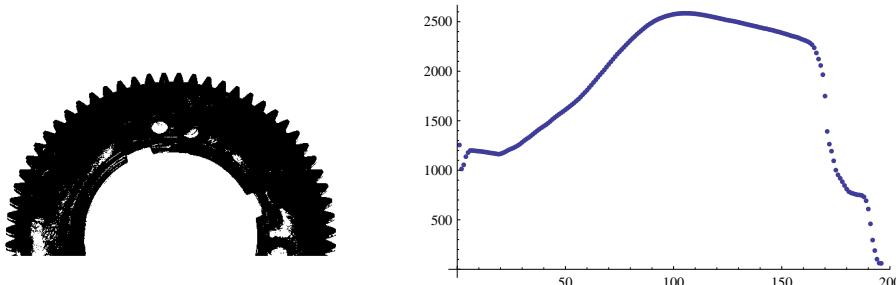
Zgodnie z propozycją Trussell'a: wybieramy takie  $t \in \{0, 1, \dots, 255\}$ , dla którego te dwie populacje różnią się jak tak bardzo jak to możliwe, tzn.

$$T = \arg \max_t Student(t),$$

gdzie  $\arg \max_t$  oznacza tę wartość  $t$ , dla której maksimum jest osiągane.

Algorytm Trussell'a jest prawdopodobnie najszerzej wykorzystywany z metod automatycznych, gdyż zwykle daje całkiem dobre wyniki i jest łatwy w implementacji (opinia przytoczona za *The Image Processing Handbook* [68]).

Przykład zastosowania metody Trussell'a pokazano na Rys. 7.13.



Rysunek 7.13: Wykres statystyki t-Studenta ma maksimum dla  $T = 110$  (po prawej). Wynik progowania dla  $T = 110$  (po lewej)

Rezultat progowania nie jest zadowalający, gdyż refleksy światła zostały potraktowane jako tło. Jak pamiętamy, progowanie bayesowskie dało w tym przypadku znacznie lepszy efekt.

Metoda Trussella jest prosta w implementacji, ale dość czasochłonna. W wielu przypadkach jej uproszczona wersja:

$$T = \arg \max_t |\mu_0(t) - \mu_1(t)|$$

daje podobne wyniki przy znacznie mniejszym zapotrzebowaniu na moc obliczeniową.

Uproszczenie można uzasadnić tym, że funkcja

$$\sqrt{\frac{\sigma_0^2(t)}{N_0(t)} + \frac{\sigma_1^2(t)}{N_1(t)}}$$

często jest płaska, za wyjątkiem miejsc blisko brzegu obszaru.

#### 7.5.4 Kolejne cztery proste metody

##### Metody bazujące na statystykach położenia i skali

Trzy bardzo proste metody to:

- $T = \hat{\mu}$  – emipryczna średnia poziomów szarości
- $T =$  emipryczna mediana poziomów szarości
- $T = k_1 \hat{\mu} + k_2 \hat{\sigma}$ , gdzie  $\hat{\sigma}$  – emipryczna dyspersja.

W ostatnim przypadku, stałe  $k_1 > 0$ ,  $k_2 > 0$  – trzeba dobrać. Zalecane wartości  $k_1 = 1$ ,  $k_2 = 0.5 \dots 2$ . nie zawsze dają dobry wynik.

Na Rys. 7.14 pokazano wyniki zastosowania powyższych metod doboru progu do znanych już zdjęć.



$T = \text{średnia}$ ,

$T = \text{mediana}$ ,

$T = \hat{\mu} + 0.5 \cdot \hat{\sigma}$

Rysunek 7.14: Wyniki zastosowania trzech prostych metod doboru progu

Wyniki zastosowania średniej i mediany nie zawsze są akceptowalne. Rezultat zastosowania kombinacji średniej i dyspersji jest dobry, ale osiągnięto go dopiero wtedy, gdy  $k_1, k_2$  wybrano metodą prób i błędów.

### Metoda Kittler'a-Illingworth'a

Metoda Kittler'a-Illingworth'a (nazywana również *SIS – simple image statistic*) pozwala uniknąć stosowania histogramu przy wyborze  $T$ . Dla każdego piksela (poza brzegiem) obliczamy wagi:

$$D_{ij} = \max[|I((i-1), j) - I((i+1), j)|, |I(i, (j-1)) - I(i, (j+1))|].$$

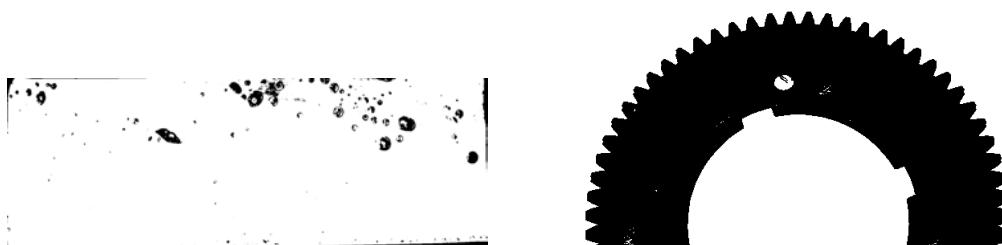
Próg obliczamy następująco:

$$\hat{T} = \frac{\sum_{i=2}^{M-1} \sum_{j=2}^{N-1} I(i, j) \cdot D_{ij}}{\sum_{i=2}^{M-1} \sum_{j=2}^{N-1} D_{ij}}$$

Uzasadnienie metody:

Należy zauważyć, że  $D_{ij}$  jest duże jedynie wtedy, gdy gradient poziomów szarości jest duży w okolicy  $(i, j)$ . W obszarach płaskich  $D_{ij} \approx 0$ . Rozważmy idealny obraz, zawierający tylko tło (poziom szarości = "B") i obiekt (poziom szarości = "O"), wówczas  $\hat{T} = \frac{B+O}{2}$ , co jest zgodne z intuicją. Zauważmy, że wynik nie zależy od wartości "B" i "O", a więc nie musimy ich znać.

Na Rys. 7.15 pokazano wynik zastosowania tej metody dla obu naszych obrazów testowych. Uzyskaliśmy całkiem dobre wyniki w dwóch różnych sytuacjach.

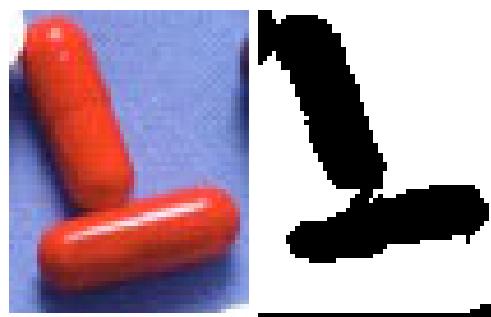


Rysunek 7.15: Wynik progowania metodą Metoda Kittler'a-Illingworth'a

Należy jedynie zaznaczyć, że chociaż unika się obliczania histogramu, to koszt obliczeniowy jest wciąż znaczny.

## 7.6 Uwagi dotyczące segmentacji obrazów kolorowych

- Dlaczego zakrętki butelek z mlekiem, kefirem itd. są zwykle czerwone, zielone lub niebieskie.
- Segmentacja obrazów kolorowych może być bardzo szybki, jeśli kolor obiektu jest zdefiniowany precyzyjnie w trójwymiarowej przestrzeni kolorów RGB.
- Pozwala to sprawdzić 200-300 butelek na minutę.
- Dla innych kolorów lepiej jest stosować przestrzeń barw HSV (HSI) – (Hue, Saturation, Value (Intensity))



Rysunek 7.16: Możemy szybko wykryć czerwone tabletki, cienie są automatycznie pomijane.

Trzy oddzielne histogramy (dla każdego z kolorów odzielnie) mogą być użyte do wyboru tła.

## Rozdział 8

# Lokalizacja, etykietowanie i pomiar

W wyniku progowania otrzymujemy obraz binarny, na którym jedynką oznaczoną zostały grupy pikseli. grupy te najprawdopodobniej odpowiadają obiektem lub defektom, bo tak staraliśmy się wybrać próg. Warto sobie jednak uświadomić, że to my widzimy w nich ważne dla nas kształty. Z punktu widzenia programów, które je przetwarzają są one nadal tylko grupami pikseli. W rozdziale tym omówimy metody, których celem jest przekształcenie takich grup w obiekty lub przynajmniej ich zlokalizowanie i dokonanie prostych pomiarów.

Zaczniemy jednak od metody, która nie wymaga binaryzacji i opierają się na znajomości kształtu poszukiwanego obiektu.

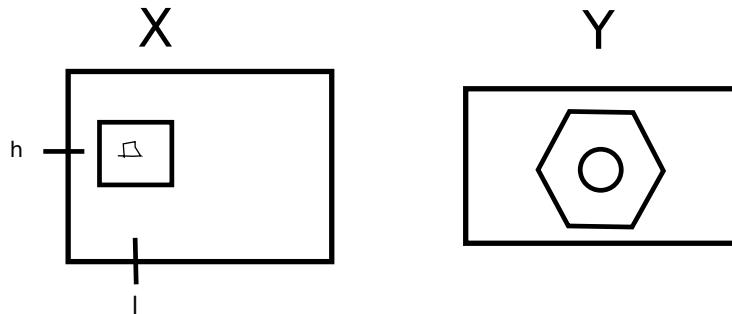
### 8.1 Poszukiwanie obiektu o znanym kształcie

Jeżeli dysponujemy obrazem poszukiwanego obiektu to można postępować według następującego schematu:

- przeglądamy obraz, przesuwając po nim wzorzec obiektu,
- liczymy miarę dopasowania fragmentu obrazu do danego kształtu,
- wybieramy fragment najbardziej podobny (o ile istnieje).

Ideę przeszukiwania naszkicowano na Rys. 8.1. Jedną z częściej stosowanych miar dopasowania jest empiryczny współczynnik korelacji:

$$\rho_{kl} = \frac{W^{-2} \sum_{i=-W}^W \sum_{j=-W}^W (x_{k+i,l+j} - \bar{x})(y_{k+i,l+j} - \bar{y})}{\sigma_x \cdot \sigma_y}, \quad (8.1)$$



Rysunek 8.1: Obiekt o znanym kształcie

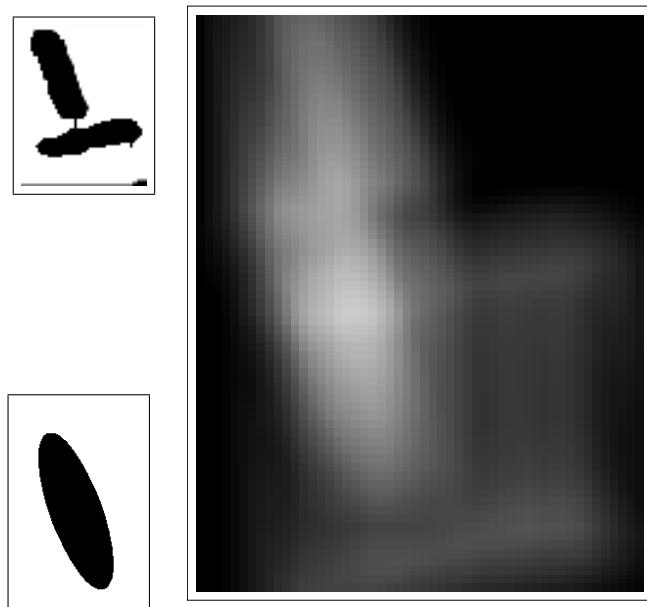
gdzie  $x_{kl}$  oznacza poziom szarości obrazu, a  $y_{kl}$  – wzorca. We wzorze (8.1) odstąpiliśmy od stosowanej w tej książce notacji, na rzecz notacji stosowanej powszechnie w statystyce. Zauważmy, że wzorzec ma rozmiary  $W \times$ , gdzie  $W$  jest zwykle dużo mniejsze od rozmiarów obrazu. Zgodnie z powyższym opisem, poszukujemy maksimum  $|\rho_{kl}|$  względem  $k$  i  $l$ , przebiegających cały obraz, za wyjątkiem pasów szerokości  $W$  wzdłuż jego brzegów. Zatem, jeśli część szukanego obiektu znajdzie się w obszarze tych pasów, to odpowiadające mu  $|\rho_{kl}|$  będzie mniejsze niż wynikłoby to ze faktycznego stopnia dopasowania. W takiej sytuacji może się zdarzyć, że obiekt nie zostanie wykryty. Sposobem zapobieżenia takiej sytuacji jest poszerzenie obrazu na tyle, by obiekty nie znajdowały się w pasie brzegowym. Jeśli nie jest to możliwe, można próbować sztucznie poszerzyć obraz, przez dodanie odpowiednich liczb wierszy oraz kolumn i próbować wypełnić ten obszar wartościami liczbowymi, które będą słabo skorelowane ze wzorcem (na przykład, liczbami pseudo-losowymi).

Trzeba jeszcze odpowiedzieć na pytanie, czy znalezione maksimum  $|\rho_{kl}|$  jest na tyle duże, że rzeczywiście świadczy o znalezieniu obiektu. Jeśli maksymalna wartość  $|\rho_{kl}|$  jest mała (bliska zero), to możemy mieć wątpliwości, czy dany obiekt znajduje się na obrazie. Wartość progową dla  $|\rho_{kl}|$  wybrać można na podstawie jeden ze statystyk testowych, służących do testowania hipotezy  $|\rho_{kl}| = 0$ .

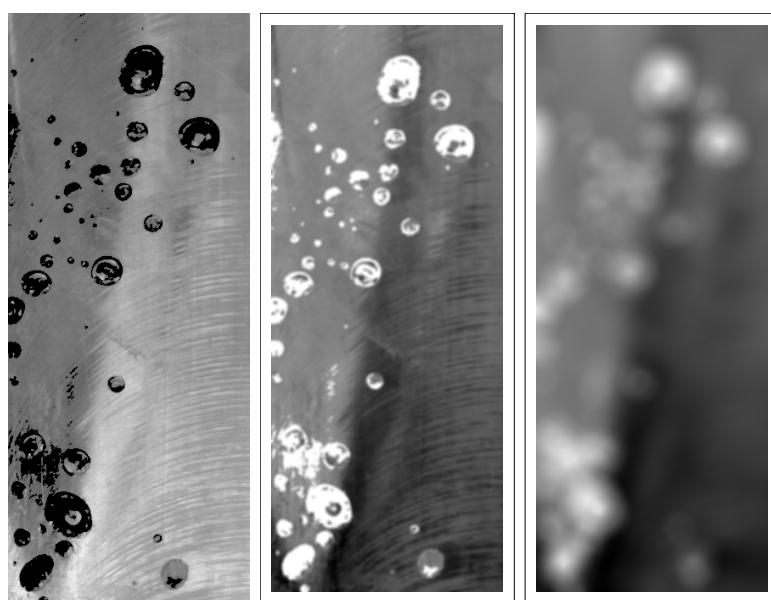
Przeglądanie całego obrazu oknem  $W \times W$  jest dość czasochłonne. Naukły obliczeniowe można zmniejszyć, jeśli zamiast współczynnika korelacji zastosujemy

$$c_{k,l} = \sum_{i=-W}^W \sum_{j=-W}^W x_{k+i,l+j} \cdot y_{k+i,l+j}. \quad (8.2)$$

Racjonalnym argumentem za stosowaniem tej uproszczonej miary dopasowania jest fakt, że (8.2) jest głównym źródłem zmienności (8.1), przy założeniu, że poziomy szarości przeszukiwanego obrazu nie są bardzo różne w różnych



Rysunek 8.2: Przykład zastosowania metody wyszukiwania wzorca – opis w tekście



Rysunek 8.3: Przykład zastosowania metody wyszukiwania wzorca do obrazu w odcieniach szarości – opis w tekście

podobszarach.

Opisane powyżej podejście, zarówno w wersji (8.1) jak i (8.2), stosować można także do obrazów binarnych, zakładając, że wzorzec także jest binarny. Jeśli przyjmiemy konwencję, że jedynki na obrazie i we wzorcu odpowiadają obiekowi, to miara (8.2) zlicza ile razy – w danym położeniu okna – wystąpiła koincydencja pikseli czarnych. Jest to dodatkowy argument za stosowaniem wersji (8.2), gdyż dla obrazów binarnych często zdarza się będzie, że w danym położeniu  $\sigma_x$  będzie równać się zero.

Metoda dopasowywania wzorca jest:

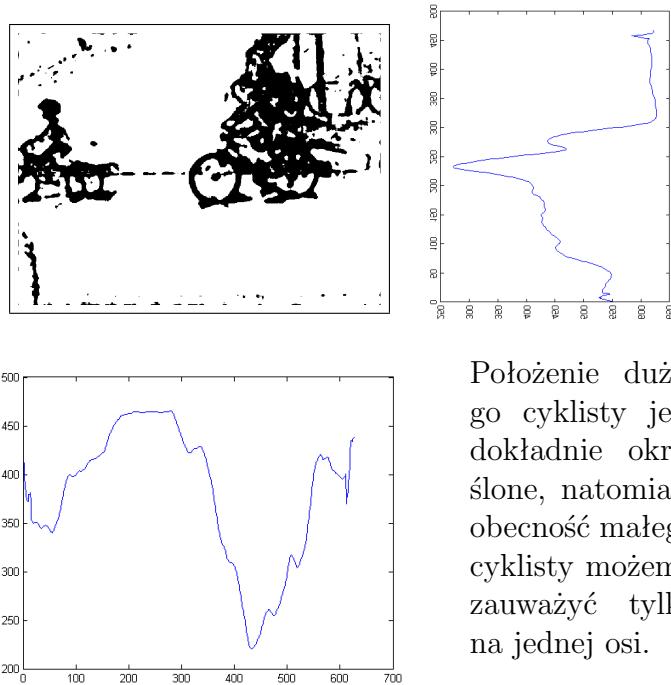
- czasochłonna,
- wrażliwa na rozmiar poszukiwanego obiektu i jego obrót,
- wciąż jednak użyteczna i możliwa do zastosowania dla obrazów binarnych i w odcieniach szarości
- nie jest nadmiernie wrażliwa na zmiany oświetlenia.

W celu przyspieszenia obliczeń można zastosować następujące przeszukiwanie zgrudne-do-dokładnego  $\rho_{kl}$  jest obliczane najpierw dla co 8-mego piksela, biorąc pod uwagę, że wzór nie pasuje dokładnie do obiektu. Następnie obliczane jest maksimum  $|\rho_{kl}|$ , a dalsze poszukiwania ograniczają się do jego sąsiadów.

Jako pierwszy przykład zastosowania tej metody rozpatrzmy zadanie znalezienia kapsułek. Obraz binarny pokazano w lewym górnym rogu Rys. 8.2. Wzorzec kapsułki znajduje się w lewym dolnym rogu. Po prawej stronie pokazano macierz  $|\rho_{kl}|$ , przedstawioną jako obraz, na którym białe obszary odpowiadają  $|\rho_{kl}| \approx 1$ . Zauważmy, że maksimum jasnych punktów odpowiada położeniu górnej kapsułki. Kapsułka dolna nie została wykryta, gdyż – jak wspomniano wyżej – metoda dopasowywania wzorca nie jest odporna na obruty.

Przykład drugi to zastosowanie metody korelacyjnej do obrazu w odcieniach szarości Rys. 8.3 (po lewej). Metodę stosowano dwukrotnie: raz wzorcem było koło o promieniu 2 pikseli, za drugim razem koło o promieniu 20 pikseli. Macierze korelacji wyświetlono jako obrazy: środkowy i po prawej, odpowiednio. Jasne obszary na tych obrazach odpowiadają korelacji bliskiej 1. Jak widać, zastosowanie mniejszego wzorca pozwala wykryć zarówno drobne jak i grubsze defekty. Zastosowanie wzorca od dużym promieniu wyławia tylko defekty o większych rozmiarach.

## 8.2 Histogramy brzegowe



Rysunek 8.4: Ilustracja działania metody histogramów brzegowych

Bardzo szczegółowy opis i dyskusję techniki histogramów brzegowych znaleźć można w [13]. Technika ta służy do szybkiego wykrycia, ustalenia pozycji i w pewnych przypadkach do oceny rozmiarów obiektu. A oto główne cechy tego podejścia.

- Zawsze można je stosować do obrazu binarnego (po segmentacji).
- W wielu przypadkach można stosować wprost do obrazów w odcieniach szarości.
- Główna zaleta – szybkość i prostota.
- Wady – wymaga dodatkowego sprawdzania, gdy występuje wiele obiektów.

Nazwa tej metody nawiązuje do znanego z rachunku prawdopodobieństwa faktu iż mając łączne prawdopodobieństwa dla pary dyskretnych zmiennych

losowych, potrafimy obliczyć rozkłady brzegowe (czyli rozkłady każdej z nich) przez sumowanie po jednym z indeksów.

W technice histogramów brzegowych zwykle nie posługujemy się częstotliwościami lecz wprost sumami nieunormowanymi w sposób opisany poniżej.

- Oblicz

$$h_j = \sum_{i=1}^M I(j, i), \quad j = 1, 2, \dots, N$$

$$v_i = \sum_{j=1}^N I(j, i), \quad i = 1, 2, \dots, M$$

- to znaczy, sumy wartości kolumn przechowujemy w  $h_j$ , a wierszy w  $v_i$ .

Cała istota metody histogramów brzegowych sprowadza się do stwierdzenia:

- gdy obiekt jest czarny tzn.  $I(i, j) = 0$  lub szary (około 0) to w odpowiednich miejscach  $h_j$  i  $v_i$  będziemy mieć niższe wartości.

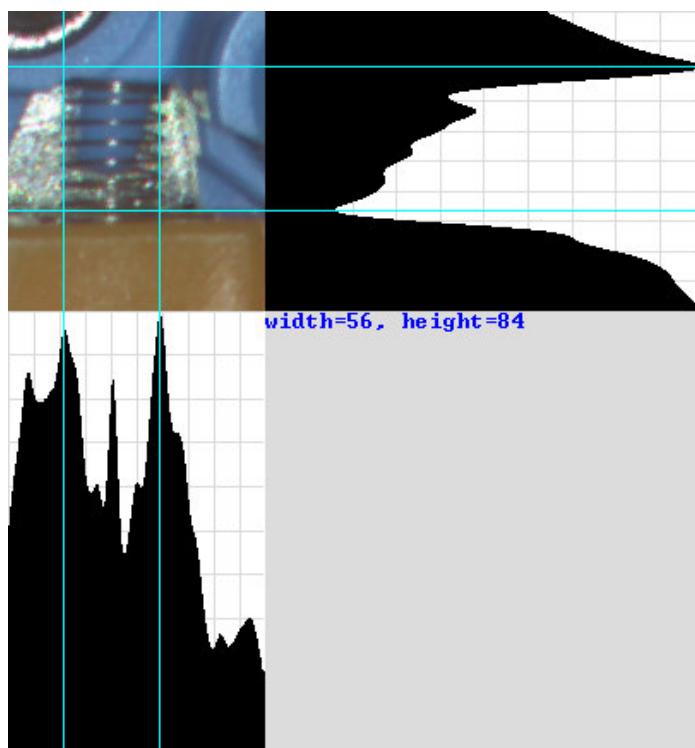
Działanie tej metody pokazano na Rys. 8.4. Lokalne minima obu wykresów jednoznacznie lokalizują dorosłego cyklistę. Natomiast mały cyklista został wykryty tylko na osi x-ów. Jest to typowa sytuacja w histogramie brzegowym, gdy występuje więcej niż jeden obiekt.

Co możemy zrobić?

1. Wziąć próbki ze wszystkich możliwych pozycji (tutaj 4) i zdecydować, które są czarne.
2. Użyć metody podobrazów, w nadziei, że nie przetniemy obiektu (specjalne sklejanie jest potrzebne gdy wykryto fragmenty tego samego obiektu w podobrazach).

**Zagadka:** na Rys. 8.4 jest błąd. Na czym on polega ?

### 8.3 Histogramy brzegowe – przykład



Rysunek 8.5: Histogramy brzegowe – sprężynka 3

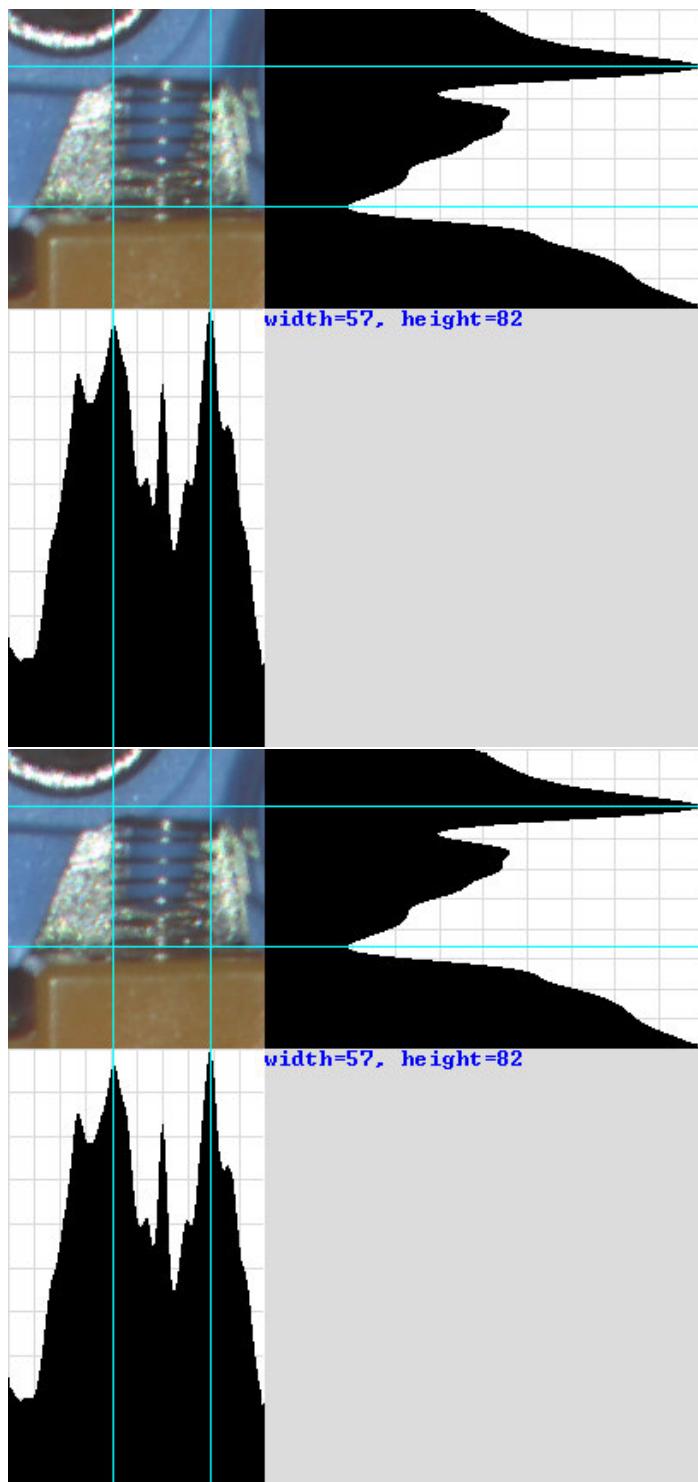
Zastosowanie histogramów brzegowych do wykrywania i oceny rozmiarów sprężynki pokazano na Rys. 8.5, 8.6 i 8.7.

**Autorzy dziękują firmie OPTOSOFT za udostępnienie tych zdjęć.**

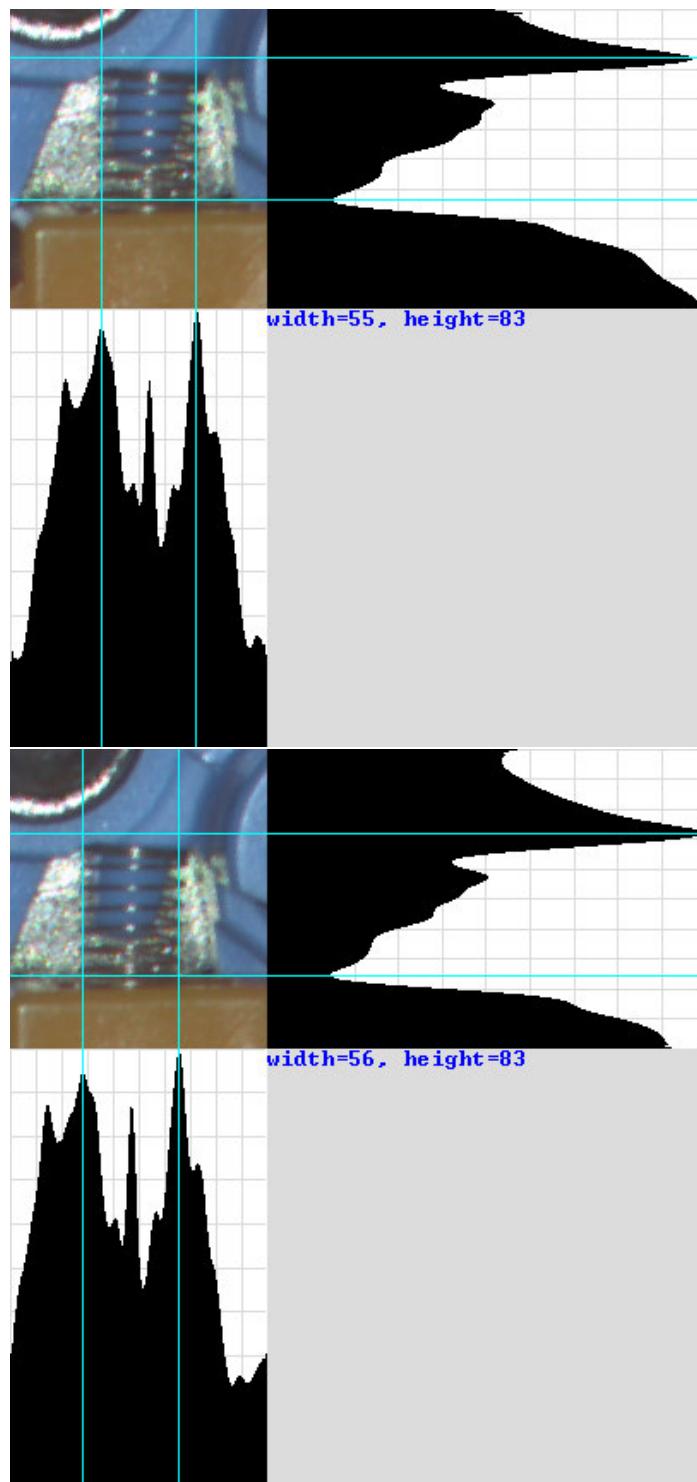
Zdjęcia te pokazują, że mimo swej prostoty, histogramy brzegowe mogą być skutecznym i precyzyjnym narzędziem lokalizacji i pomiaru obiektów. Dlaczego zamieszczamy tak dużą liczbę zdjęć? Odpowiedź jest prosta – zawsze należy zbadać, czy dana metoda daje stabilne, mało różniące się wyniki. Przytoczone zdjęcia dokumentują fakt, że w tym zastosowaniu tak właśnie jest.

### 8.4 Uwagi o analizie grup pikseli

Obraz binarny uzyskany jako wynik progowania niesie informacje istotne dla nas, ale nie dla komputera. Dlaczego? Ponieważ zawiera czarne piksele i nic więcej (por. Rys 8.8).



Rysunek 8.6: Histogramy brzegowe – sprężynek 1



Rysunek 8.7: Histogramy brzegowe – sprężynek 2



Rysunek 8.8: Binarny obraz kapsułek

Bez dalszej analizy możemy tylko policzyć czarne piksele (w pewnych zastosowaniach to wystarczy – mówi o całkowitej powierzchni defektu). Grupa zaznaczonych (czarnych) pikseli, które są odzielone od innych grup białymi pikselami nazywana będzie dalej skupiskiem<sup>1</sup>. Grupy te są kandydatami do bycia obiektem. Niektóre z nich nimi nie są i nie interesują nas. Inne są artefaktami, które pojawiły się w wyniku przetwarzania obrazu. Co możemy zrobić z obrazem po wydzieleniu i nadaniu skupiskom etykiet?

- Możemy je policzyć.
- Wypełnić puste miejsca, spowodowane tym, że części obiektu nie zostały wykryte.
- Możemy je zmierzyć (aby sprawdzić, czy produkt odpowiada wymaganiom).
- Rozpoznać skupisko jako obiekt danej klasy.

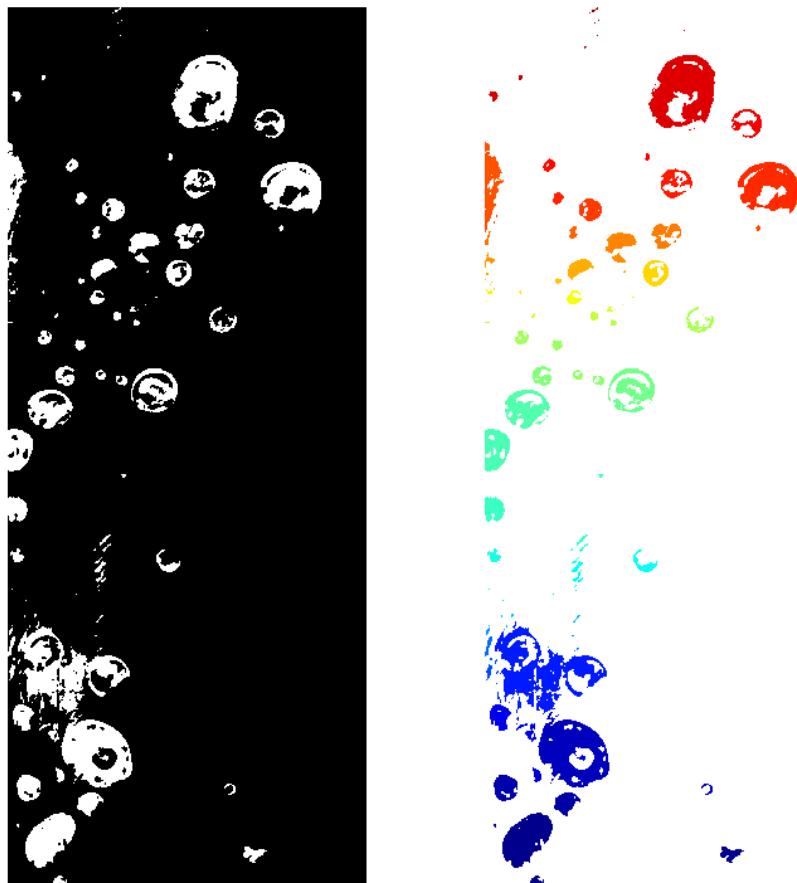
Kolorowanie skupisk pikseli nie służy wyłącznie do czytelnego ich wyświetlenia, ale głównie do zaznaczenia do którego ze skupisk należy dany piksel i do oznaczenia ich taką samą etykietą. Po etykietowaniu powstaje struktura danych dogodna do dalszej obróbki.

Wykrycie skupiska jako obiektu (defektu) jest wielopoziomowym procesem. Najpierw należy wybrać cechy do klasyfikatora, następnie nauczyć go rozpoznawać obiekty. Dopiero po tym można używać go on-line.

---

<sup>1</sup>Stosujemy ten termin jako odpowiednik angielskiego wyrazu *blob*.

## 8.5 Etykietowanie obiektów

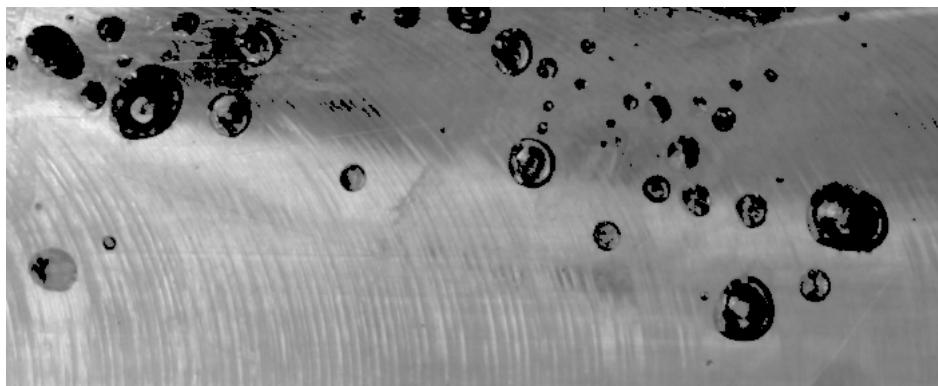


Rysunek 8.9: Obraz defektów po przypisaniu kolorowych etykiet.

W odrozdzie tym zmienimy konwencję konwencje oznaczeń – zaznaczone skupiska oznać jako białe, a więc jako mające wartość 255 (lub 1 na skali  $[0,1]$ ), gdyż taka konwencja stosowana jest w większości algorytmów i dostępnego oprogramowania. zaznaczane są jako białe = 1 (lub 255).

Naszym celem jest przypisanie unikatowej etykiety do każdego skupiska<sup>2</sup>

<sup>2</sup>Operacje taka nazywa się w literaturze *labelling*.



Rysunek 8.10: Kratery widoczne w części defektów po progowaniu nie powinny być uzupełniane.

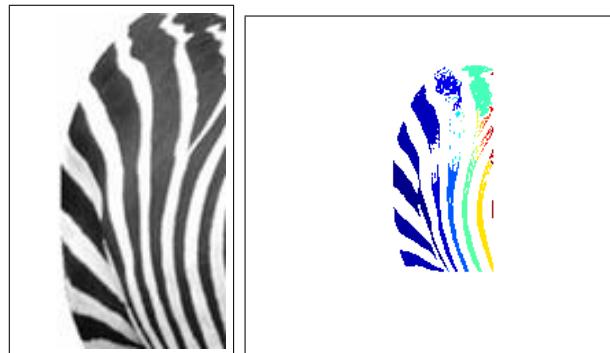
(które jest obiektem, defektem). Przyjęło się stosować dodatnie liczby całkowite jako etykiety i wyświetlać je jako obrazy sztucznie kolorowane (por. Rys. 8.9).

Zanim przejdziemy do opisu algorytmu etykietowania, poczynimy kilka uwag.

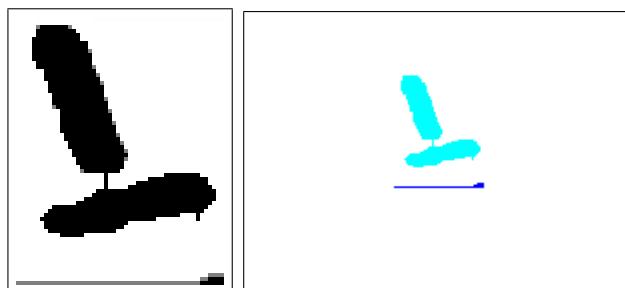
- Zwykle wypełnia się luki w skupiskach przed ich zaetykietowaniem. Nie zawsze jest to właściwe, gdyż może zniekształcić rzeczywistość, na przykład, wywiercone otwory mogą zostać potraktowane jak luki. Inny przykład pokazano na Rys. 8.10. Po progowaniu wewnątrz defektów pozostają kratery. Wypełnienie ich jako luki byłoby niewłaściwe, ponieważ kratery są widoczne i informatywne.
- Skupiska zaznaczonych pikseli nie zawsze są obiektami )por. Rys. 8.11)
- Dlaczego dwie kapsułki na Rys. 8.12 są kolorowane jako jedna? Uwagę zwraca cienka linia pomiędzy tabletkami – progowanie nie zawsze jest wystarczające, czasami trzeba stosować dalsze przetwarzanie (*postprocessing*) (na przykład, usuwające małe grupy operacje morfologiczne).

### 8.5.1 Spójne obszary

Trudności z etykietowaniem kapsułek na Rys.8.12 wynikają z faktu, że nie odpowiedzieliśmy sobie na pytanie: co to znaczy, że grupa pikseli stanowi spójny obszar ? W Tabeli 8.1 pokazano obszar, który niewątpliwie jest spójny (po lewej) i dwa obszary stykające się narożnikiem (po prawej). Decyzja o tym czy uznać je za obszary spójne (ang. *connected components*) należy do nas.



Rysunek 8.11: Zebra nie byłaby szczęśliwa, gdyby jej paski traktować jako oddzielne obiekty.



Rysunek 8.12: Dwie kapsułki nie zostały rozróżnione w procesie etykietowania

M	M	M	M		
M	M	M	M		
M	M	M	M		
M	M	M	M		

M	M				
M	M				
		M	M		
		M	M		

Tabela 8.1: Obszar spójny (po lewej) i obszary, których traktowanie jako spójne zależy od naszej decyzji.

M	M	M			
M	$\mathcal{M}$	M			
M	M	M			

		M			
	M	$\mathcal{M}$	M		
		M			

Tabela 8.2: Rola sąsiedztwa w definiowaniu spójności obszarów

		M				
M	$\mathcal{M}$	M				
	M	A				
			A			
				A		

Tabela 8.3: Piksele oznaczone przez "A" nie są traktowane jako ta sama grupa co "M" jeśli stosujemy 4-sąsiedztwo. Należą to tej samej grupy przy stosowaniu 8-sąsiedztwa.

W Tabeli 8.2 pokazano rolę definicji sąsiedztwa w określeniu, które obszary uważa się za spójne. Dla centralnego (obecnie rozpatrywanego) piksela (zaznaczonego jako  $\mathcal{M}$ ) możemy wyróżnić dwa przypadki:  
 po lewej – 8 połączonych elementów,  
 po prawej – 4 połączone elementy. Piksele zaznaczone przez M w otoczeniu  $\mathcal{M}$  są traktowane jako należące do tej samej grupy.

Sposób traktowania sąsiadów zależy zatem od wybranego sposobu interpretowania połączeń między pikselami. Stwierdzenie to dodatkowo ilustruje Tabela 8.3.

- Powyższy sposób definiowania 4 lub 8 sąsiadów jest jedynym używanym w literaturze, gdyż nie ma innego wyboru na poziomie pikseli, jeśli z góry założymy symetrię.
- Rosnąca rozdzielcość kamer przemysłowych (10 MPix) rodzi pytanie: czy nie należy problemu spójności obszarów przemyśleć ponownie ? **Zadanie:** zaproponować kilka wersji zdefiniowania sąsiedztw.

### 8.5.2 Algorytm etykietowania

Zanim omówimy jeden z algorytmów etykietowania pikseli należących do spójnych obszarów, warto poczynić kilka uwag.

- Znanych jest wiele metod etykietowania, gdyż:
  - zadanie jest trudne,
  - żadna z metod nie jest najlepsza we wszystkich kategoriach.
- Znane metody różnią się:
  - stopniem skomplikowania programu,
  - prędkością obliczeń,
  - wymaganą pamięcią.
- Dwa ostatnie wymagania najczęściej konkurują ze sobą i można zwiększyć szybkość kosztem zajętości pamięci lub odwrotnie. .

Przyjmujemy następujące założenia.

- Dla uproszczenia rozważamy wyłącznie 4-sąsiedztwo.
- Etykietowaniu podlega obraz binarny o rozmiarze  $M \times N$  i wartościach  $b_{ij} \in \{0, 1\}$ .
- Alokujemy pamięć na etykiety (albo  $M \times N$  macierz  $L_{ij}$  lub listę etykiet i adresów).
- Alokujemy również dodatkowo  $2 \times \mathcal{N}$  tablicę  $C$  do rozstrzygania konfliktu etykiet, gdzie  $\mathcal{N}$  jest maksymalną liczbą spodziewanych skupisk do zaetykietowania.

W rzeczywistości  $\mathcal{N}$  powinno być trochę większe, żeby zachować fragmenty skupisk, które są później sklejane.

#### Algorytm – etap 1

- Startując z lewego, górnego rogu przeglądamy obraz  $b_{ij}$  do napotkania pierwszego zaznaczonego piksela i ustawiamy odpowiadającą etykietę w  $L$  na "1". Niezaznaczone piksele  $b_{ij} = 0$  są pomijane (np. ustawiamy ich  $L_{ij}$  na "0").
- Gdy algorytm znajdzie się gdzieś wewnątrz tablicy  $b$ , powiedzmy, w punkcie  $(i, j)$ ,  $i \geq 1, j > 1$ , to piksele powyżej i po lewej już są zaetykietowane. Niech " $n$ " oznacza największy z nich. Teraz mogą zdarzyć się trzy przypadki

- niezaznaczony piksel  $b_{ij} = 0$  – ustaw  $L_{ij} = "0"$ ,  $j := j + 1$  (lub zmień wiersz).
- zaznaczono piksels  $b_{ij} = 1$  oraz  $L_{i(j-1)} = "0"$  oraz  $L_{(i-1)j} = "0"$ , wówczas, ustaw  $L_{ij} = "n + 1"$  (następną etykietę) i  $j := j + 1$  (lub zmień wiersz).
- zaznaczony piksel  $b_{ij} = 1$  oraz  $[L_{i(j-1)} > "0" \text{ LUB } L_{(i-1)j} > "0"]$  rozdziela się na dwa przypadki:

**Przypadek 1** poprzednie etykiety są zgodne t.j.  $L_{i(j-1)} = "m"$  oraz  $L_{(i-1)j} = "m"$  gdzie " $m$ " jest jedną z wcześniej przypisanych etykiet " $1$ ", " $2$ ", ..., " $n$ ".

**Przypadek 2** nie wszystkie poprzednie etykiety zgadzają się

		j	
	"m"	"m"	
i	"m"	nowe "m"	

$L$  - macierz – Przypadek 1

		j	
	"0"	"k"	
i	"m"	nowe "???"	

$L$  - macierz – Przypadek 2

**Przypadek 1** Ustaw  $L_{ij} = "m"$ ,

**Przypadek 2** Ustaw  $L_{ij} = "???"$  – przypisz najmniejszą etykietę pomiędzy konfliktami np. " $k$ " or " $m$ " (nie wykonuj, jeśli " $0$ " jest obecne)

- Ustaw  $j := j + 1$  (lub zmień wiersz).
- Stop jeśli  $(i, j)$  w prawym, dolnym rogu.

Czy to już koniec ? Jeszcze nie. Wyobraźmy sobie trójkąt:  $\triangle$ . Jak zostanie on zaznaczony przez powyższy algorytm ?

### 8.5.3 Rozwiązywanie konfliktów

W poprzednim podrozdziale zetknęliśmy się z sytuacją konfliktu etykiet. Aby móc je rozwiązać,  $2 \times \mathcal{N}$  macierz  $C$  wypełniana jest równocześnie z macierzą  $L$  w powyższym algorytmie. Macierz  $C$  wypełniamy zgodnie z następującymi regułami.

- Pierwszy wiersz  $C$  zawsze zawiera  $1, 2, \dots, \mathcal{N}$  (używany wyłącznie dla wygody). Na początku drugi wiersz  $C$  zawiera " $0$ " na każdej pozycji.

- Kiedy pojawia się nowe skupisko pikseli:

		j	
	”0”	”0”	
i	”0”	nowe ”n+1”	

to w drugim wierszu ( $n + 1$ ) kolumny w macierzy  $C$  ustawiamy ” $n + 1$ ” (zgodnie z  $L$ ).

- Jeśli występuje konflikt pomiędzy etykietami w drugim wierszu  $C$ , ustawiamy największą z etykiet (” $m$ ” lub ” $k$ ”), będących w konflikcie.

Warto zauważyć, że  $C$  zawiera informacje o będących w konflikcie etykietach, która jest komplementarna do tej przechowywanej w  $L$ . Informacja ta wystarcza do rozwiązania konfliktu przez ponowne przejrzenie  $L$  i  $C$  i zmianę etykiet, kiedy zachodzi taka potrzeba.

**Przykład.** Teraz możemy etykierować i mierzyć kapsułki, gdyż dokładne wiemy, który pixel<sup>3</sup> należy do którego skupiska. W Tabeli 8.4 pokazano etykietowanie kapsułek, a po prawej stronie podano tabelkę z wynikami ich pomiarów.



Powierzchnia	493	396
Centroid	[16.2 21.0]	[29.5 47.5]
Długość osi wielkiej	44.1	47
Ekscentryczność	0.94	0.97
Orientacja	-67.9	12.6
Liczba Eulera	1	1
Obwód	102.7	100.8

Tabela 8.4: Przykład etykietowania i pomiaru obiektów

Zamieszczne w tej tabeli parametry zdefiniujemy w następnym podrózdziale.

## 8.6 Opisywanie obiektów

Po zaetykietowaniu obiektów na obrazie binarnym, dysponujemy obrazem wraz z dodatkową informacją o adresach pikseli należących do każdego z obiektów (lub tylko przypadkowych skupisk pikseli). Może ona być przechowywana jako:

---

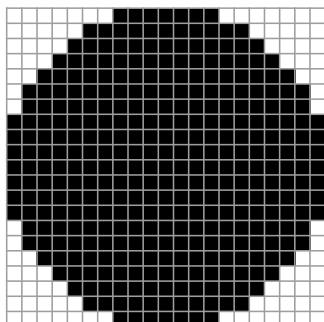
<sup>3</sup>Aby zapewnić efekt dydaktyczny, usunięto cienki rząd pikseli, które poprzednio łączyły obie kapsułki.

- macierz  $L_{ij}$ , tego samego rozmiaru co obraz binarny, zawierająca liczby całkowite, które są etykietami przypisanymi każdemu pikselowi ( $L_{ij} = "k"$  oznacza, że piksel  $(i, j)$  należy do obiektu " $k$ ", jeśli " $k$ " = "0" to uznajemy, że należy do tła),
- jako listę pikseli, które tworzą obiekt (ta wersja stosowana jest w Matlabie).

Naszym celem jest opisanie tych obiektów w sposób odpowiedni do danego zastosowania. Stosujemy konwencję  $L_{ij} = 1$  dla bieżącego obiektu i  $L_{ij} = 0$  dla tła (w praktyce nie stosuje się takiego sposobu przechowywania obiektów). Aby opisać obiekty, nawet proste pojęcia musimy zdefiniować w języku pikseli.

- Powierzchnia  $A$  to liczba pikseli  $L_{ij} = 1$ .
- Tak rozumiana powierzchnia może być użyta do przybliżenia powierzchni rzeczywistego obiektu  $c_A \cdot A$ , gdzie  $c_A$ , [ $mm^2/pixel$ ] jest powierzchnią odpowiadającą jednemu pikselowi w rzeczywistości.
- Stała  $c_A$  może być uzyskana w wyniku kalibracji kamery.
- Najprostszym sposobem jest pobranie obrazu kwadratu o znanej powierzchni i obliczenie ilości zawartych w nim pikseli. Trzeba zauważyć, że  $c_A$  zależy od odległości oraz zastosowanej kamery i obiektywu (2.3).
- Dla małych obiektów, oszacowanie powierzchni nie może być zbyt dokładne.

Powód jest widoczny na rysunku:



Koło o promieniu  $r = 10$  pikseli. Można poprawić dokładność zmniejszając odległość pomiędzy obiektem a kamerą lub stosując wyższą rozdzielczość

Około 50 pikseli może być błędnie dołączonych do obszaru koła. Jeśli  $c_A = 0.01 mm^2/pixel$ , to możemy popełnić błąd o ok.  $5 mm^2$ . Dlaczego dokładność określenia powierzchni jest ważna w pewnych zastosowaniach?

Jeśli pobieramy próbki do badania i mnożymy ich powierzchnię przez tysiące elementów produkowanych w ciągu dnia to nawet błąd o  $5\text{ mm}^2$  może mieć wpływ na końcowy wynik.

- Z poprzedniego rysunku jasno wynika, że obwód koła również będzie obliczony z błędem.
- Jako obwód  $P$  możemy przyjąć liczbę pikseli dotykających krawędzi obiektu. Obliczając  $P$  możemy zacząć w dowolnym punkcie początkowym i ostatecznie wrócić do niego.

### 8.6.1 Cechy geometryczne

Wymienione niżej cechy geometryczne obiektów mogą służyć jako wskaźniki, które służą do ich szybkiego rozpoznawania, w sposób niezależny od skali.

- Współczynnik cyrkularności (podobieństwa do koła) obiektu:

$$C = \frac{4\pi \text{ Powierzchnia}}{(Obwod)^2}$$

- $C = 1$  dla koła
- dla wydłużonych (cienkich) obiektów  $C < 1$  np. dla trójkąta równobocznego  $C \approx 0.6$ .
- Dla kwadratu  $C = \pi/4$ , ale dla wydłużonego prostokąta  $c = 10 \cdot b$  uzyskujemy  $C = 0.26$ .
- $C$  jest niezależne od skali.
- Znanych jest wiele innych wskaźników (Haralicka, Ferreta, solidność = (Powierzchnia/Powierzchnię wypukłą obszaru) itd. (por. [78]) Pozwalają łatwo i szybko rozróżnić obiekty o różnych kształtach, np. ziarna mające kształt elipsy od niechcianych składników).

Lista innych własności geometrycznych jest bardzo długa. Przykładowo, Matlab oblicza następujące: Area, Centroid, BoundingBox, SubarrayIdx, MajorAxisLength, MinorAxisLength, Eccentricity, Orientation, ConvexHull, ConvexImage, ConvexArea, FilledArea, EulerNumber, Extrema, EquivDiameter, Solidity, Extent, PixelIdxList, PixelList, Perimeter. Są one bardzo użyteczne, ale warto zapoznać się z nimi dopiero, gdy stanie się to potrzebne.

### 8.6.2 Momenty geometryczne

Ze statystyki pamiętamy, że dla dwóch zmiennych losowych  $X, Y$ , ich momenty rzędu  $m \geq 0, n \geq 0$  zdefiniowane są następująco:

- $q(m, k) = E(X^k Y^m)$  – niecentralny moment
- $Q(m, k) = E((X - E(X))^m (Y - E(Y))^n)$  – centralny moment

o ile istnieją. Wprawdzie obrazy binarne mogą nie spełniać wymagań teorii prawdopodobieństwa, jednak empiryczne wersje  $q, Q$  (różnie znormalizowane) są stosowane jako deskryptory kształtu.

Rozważmy obiekty zaetykietowane jako "1".

- Niech  $l_{ij} = 1$  jeżeli  $L_{ij} = "1"$  i  $l_{ij} = 0$ , w przeciwnym przypadku.
- Niech  $\mathcal{L}$  będzie całkowitą liczbą pikseli oznaczonych jako "1". Wtedy  $\mathcal{L} = \sum_{i=1}^N \sum_{j=1}^M l_{ij}$ .
- Niech  $\beta(m, n) = 1 + (m + n)/2$  i

$$v(m, n) = \frac{1}{\mathcal{L}^{\beta(m, n)}} \cdot \sum_{i=1}^N \sum_{j=1}^M l_{ij} (i - \hat{x})^m (j - \hat{y})^n,$$

gdzie  $\hat{x}_i$ 's and  $\hat{y}_j$  są emprycznymi średnimi:

- $\hat{x} = \mathcal{L}^{-1} \sum_{i=1}^N \sum_{j=1}^M i l_{ij}$
- $\hat{y} = \mathcal{L}^{-1} \sum_{i=1}^N \sum_{j=1}^M j l_{ij}$
- Używając kombinacji  $v(m, n)$  Hu opracował 7 momentów, które (w wersji ciągiej) są niezmienne względem translacji, rotacji i skalowania. Są one dobrymi kandydatami jako cechy podawane do klasyfikatora. W 1996 Pawlak i Liao [37] wykazali że momenty te są wrażliwe na błędy zaokrągleń i zaproponowali momenty oparte o wielomiany Zernika, które są ortogonalne na kole jednostowym i niezmienne względem rotacji.

Pierwsze cztery momenty Hu  $\mu$  są zdefiniowane następująco:

- $\mu_1 = v(2, 0) + v(0, 2)$ ,
- $\mu_2 = (v(2, 0) - v(0, 2))^2 + 4 v^2(1, 1)$ ,
- $\mu_3 = (v(3, 0) - 3 v(1, 2))^2 + (v(0, 3) - 3 v(2, 1))^2$ ,

- $\mu_4 = (v(3,0) + v(1,2))^2 + (v(0,3) - v(2,1))^2$ .

Pozostałe są zbyt skomplikowane, aby je tu przytaczać (patrz [54] str. 609). Problem deskryptorów będziemy omawiać także w dalszej części tej książki jako deksryptory Fouriera dla konturów.

## 8.7 Znajdowanie obiektów

Oprócz omówionego szeroko progowania, segmentacji można również dokonywać przez:

- wzrost regionów (*region growing*),
- metody działu wodnego (*watershed methods*):
  - podejście opadowe (*rainfall approach*),
  - podejście zalewowe (*flooding approach*).

Metody te mają dobrą interpretację, ale są rzadko stosowane w przetwarzaniu on-line – pominiemy ich dokładny opis.

### 8.7.1 Przykład

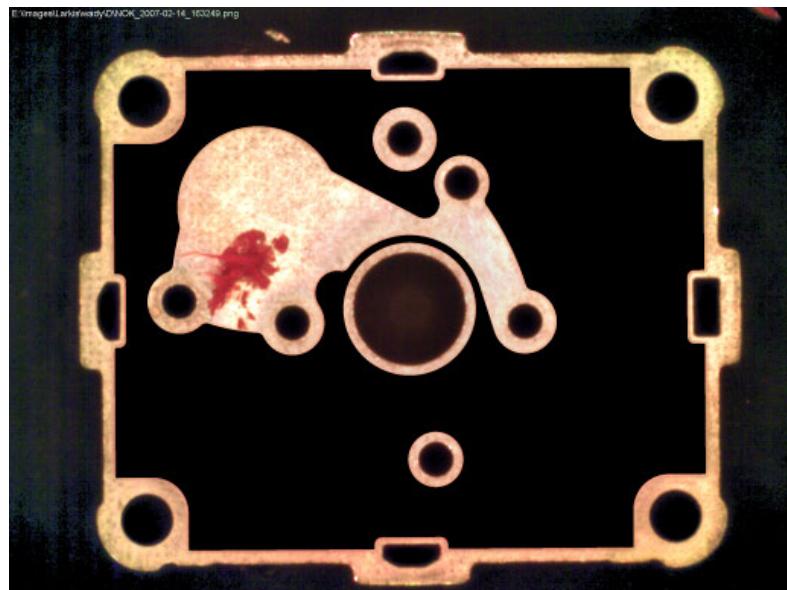
Autorzy dziękują firmie OPTOSOFT Sp. z o.o. za udostępnienie obrazów. Autorem oprogramowania i sprzętu jest W. Rohleder, prezes spółki.

Wykrywanie wycieku.

- Zadanie: wykryć potencjalny wyciek płynu.
- Pokazano wszystkie etapy procesu przetwarzania.
- Wyciek jest wykrywalny ludzkiem okiem. Jak go wykryć, dobierając właściwą sekwencję zadań częściowych ?
- Znamy teorię do każdego kroku, ale połączenie ich w całość wciąż jest sztuką.

Zastosowana do obrazu z Rys. 8.13 sekwencja przetwarzania składa się z następujących etapów.

1. Przekształcenie od odcieni szarości – Rys. 8.14.
2. Poprawy kontrastu – Rys. 8.15.



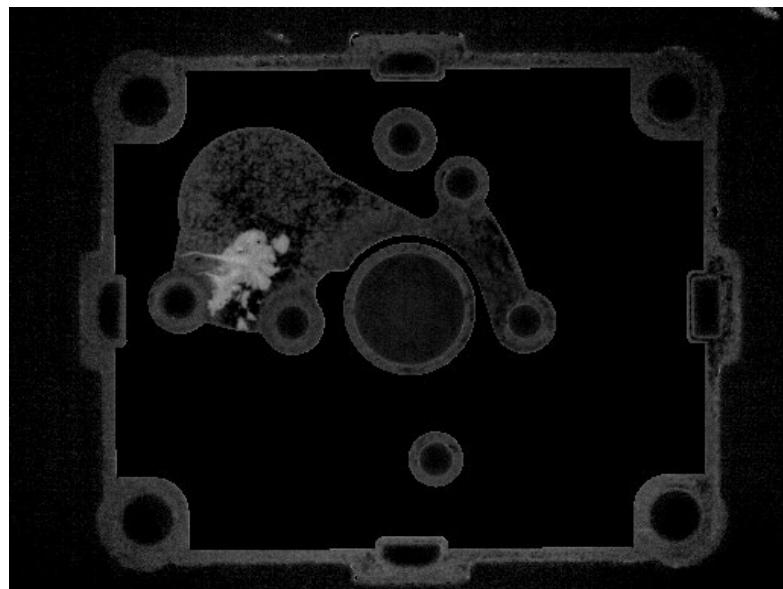
Rysunek 8.13: Wykrywanie wycieku – orginalny obraz

3. Progowania – Rys. 8.16.
4. Usunięcia drobnych grup pikseli, które zostały oznaczone w wyniku progowania (można to zrobić stosując podejście morfologiczne lub filtr medianowy) – Rys. 8.17.
5. Lokalizacja wycieku (na przykład, metodą histogramów brzegowych) – Rys. 8.18.

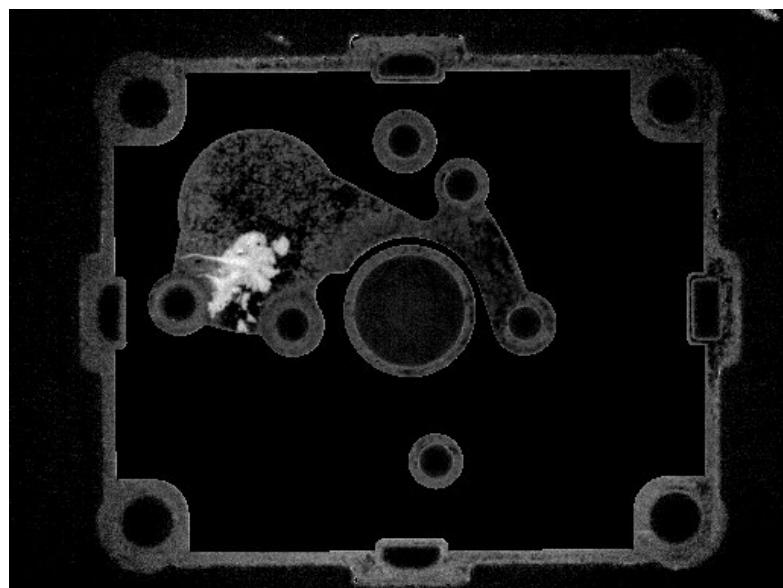
Ten sam szereg operacji powtórzony dla innego obrazu z wyciekiem (Rys. 8.19). Powody są następujące.

- Powinniśmy sprawdzić proponowaną metodę w zmiennych warunkach.
- Zawsze istnieje niebezpieczeństwo, że metoda będzie zbyt dostrojona do przykładu, a nie sprawdza się, gdy warunki się zmieniają.

Wyniki kolejnych etapów przetwarzania pokazano na Rys. 8.20- 8.24. Przykład ten był trudniejszy, gdyż wyciek wystąpił w dwóch miejscach i był mniejszy. Mimo tego, został prawidłowo wykryty i oznaczony w sposób, który pozwala na ocenę jego wielkości.



Rysunek 8.14: Wykrywanie wycieku – przekształcenie od odcieni szarości



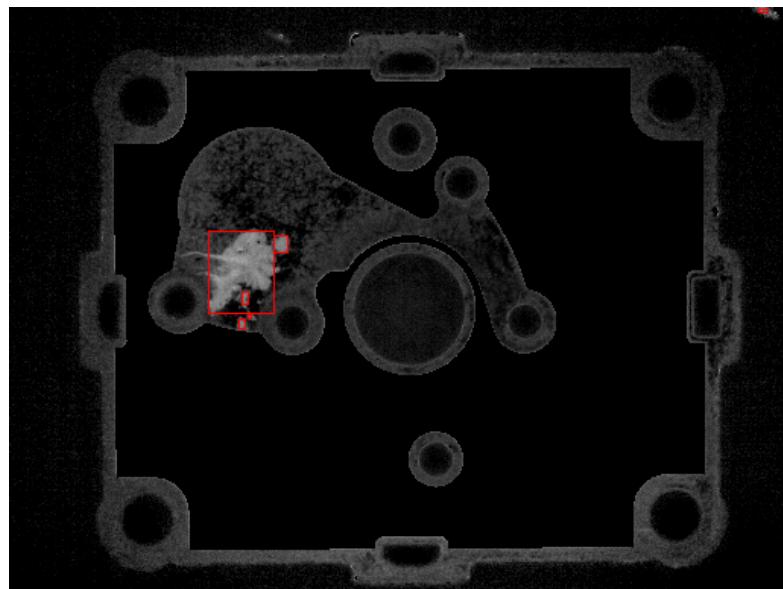
Rysunek 8.15: Wykrywanie wycieku – zwiększenie kontrastu



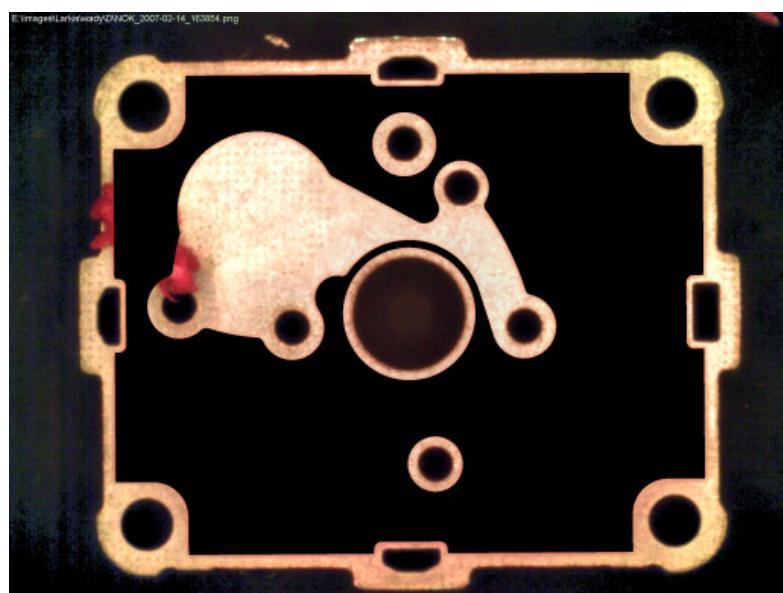
Rysunek 8.16: Wykrywanie wycieku – segmentacja



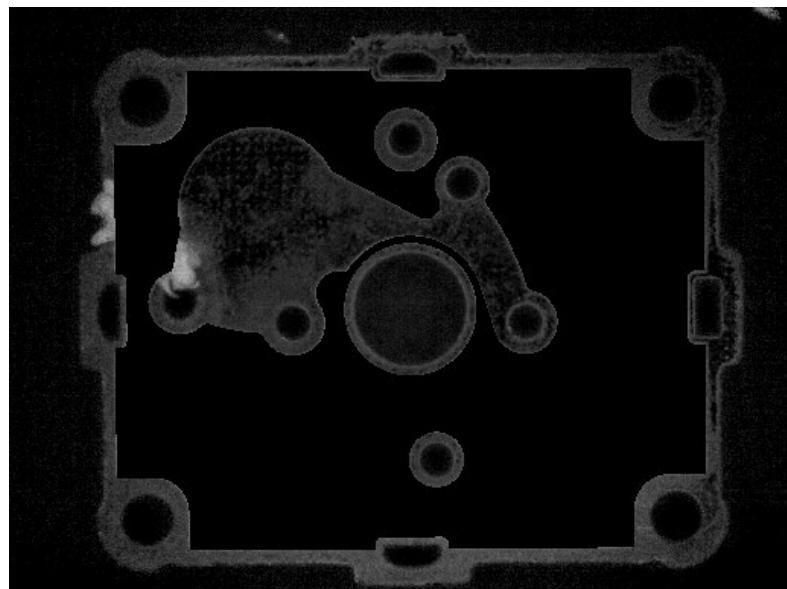
Rysunek 8.17: Wykrywanie wycieku – poprawa rezultatów segmentacji



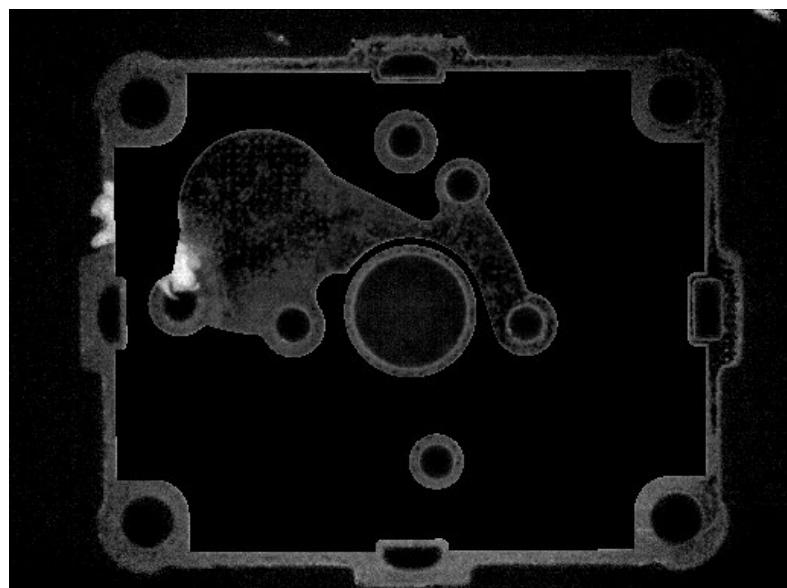
Rysunek 8.18: Wykrywanie wycieku – lokalizacja



Rysunek 8.19: Wykrywanie wycieku 2 – orginalny obraz



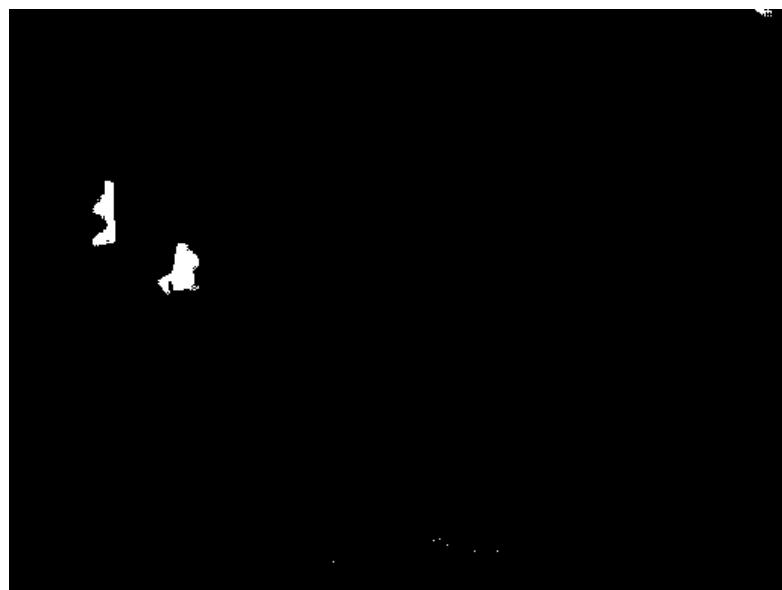
Rysunek 8.20: Wykrywanie wycieku 2 – przekształcenie od odcieni szarości



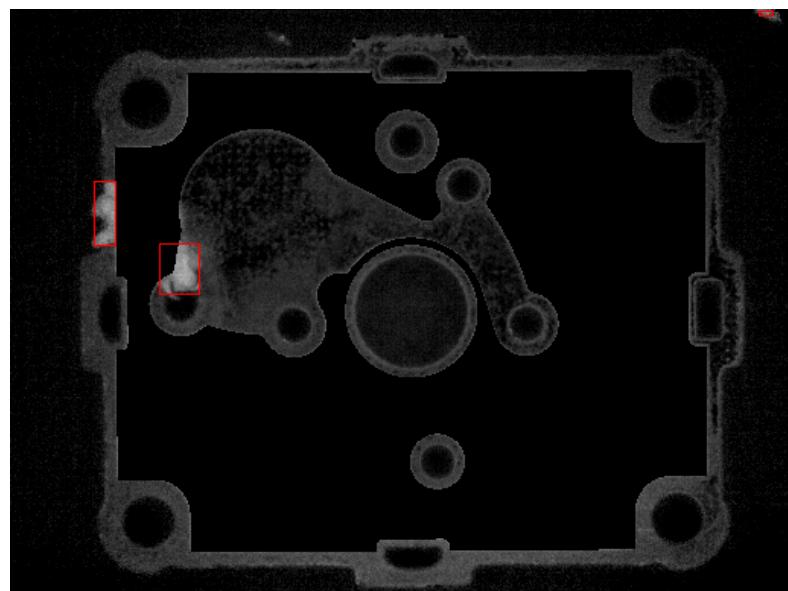
Rysunek 8.21: Wykrywanie wycieku 2 – zwiększenie kontrastu



Rysunek 8.22: Wykrywanie wycieku 2 – segmentacja



Rysunek 8.23: Wykrywanie wycieku 2 – poprawa rezultatów segmentacji



Rysunek 8.24: Wykrywanie wycieku 2 – lokalizacja

# Rozdział 9

## Kontury obiektów i defektów

W poprzednich rozdziałach staraliśmy się wykrywać całe obiekty lub defekty, traktowane jako duże grupy pikseli. W niniejszym rozdziale omawiamy podejście alternatywne, polegające na wykryciu jedynie konturów. Przedstawiamy także podstawowe metody opisu konturów, które pozwalają na ich pomiar.

### 9.1 Okonturowywanie

Najprostszą definicją konturu to krzywa, która oddziela obszary znaczaco różniące się poziomami szarości.

Znajdowanie obiektów przez wykrywanie krawędzi to bardzo szeroka klasa metod, które próbuje opisywać (a później rozpoznawać lub mierzyć) obiekty na podstawie ich brzegów. Obejmuje wykrywanie krawędzi przez:

- aproksymację pierwszego rzędu gradientu intesywności obrazu,
- aproksymacja Laplasjanu (2-go rzędu)
- operacje morfologiczne (patrz 13).

Oto lista najczęściej używanych metod wykrywania krawędzi.

1. Gradientowe – klasyczne Prewitt, Sobel, Kirsch,
2. metody drugiego rzędu: Laplasjan i LoG,
3. Canny, SUSAN, mod-SUSAN.

Zagadnieniem zbliżonym do wykrywania krawędzi jest wyszukiwanie narożników.

$h(-1, -1)$	$h(-1, 0)$	$h(-1, 1)$
$h(0, -1)$	$h(0, 0)$	$h(0, 1)$
$h(1, -1)$	$h(1, 0)$	$h(1, 1)$

Tabela 9.1: Ogólna postać maski  $3 \times 3$ .

$I(1, 1)$	$I(1, 2)$	$I(1, 3)$				
$I(2, 1)$	•	$I(2, 3)$				
$I(3, 1)$	$I(3, 2)$	$I(3, 3)$				

Tabela 9.2: Macierz obrazu z zaznaczonym punktem przyłożenia maski w  $(2, 2)$ .

### 9.1.1 Maski

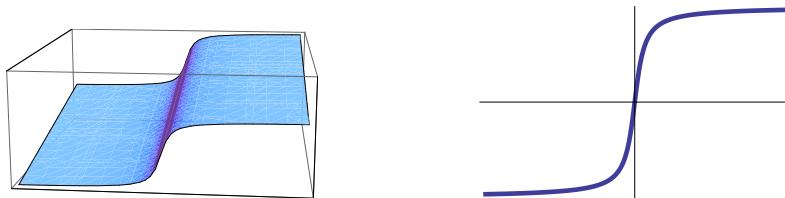
Przed rozważeniem szczegółów wykrywania krawędzi, musimy zapoznać się z ogólną i bardzo ważną koncepcją splotu obrazu z maskami.

- Maską nazywamy  $(2K + 1) \times (2K + 1)$  macierz z elementami  $h(k, l)$ , które określają wagę stosowane lokalnie do obrazu  $I(i, j)$ .
- Obraz wyjściowy  $Y$  jest tworzony jako splot  $I$  z  $h$  w następujący sposób

$$Y(i, j) = \sum_{k=-K}^{K} \sum_{l=-K}^{K} I((i+k), (j+l)) \cdot h(k, l) \quad (9.1)$$

Maskę  $3 \times 3$ , ( $K = 1$ ) pokazano w Tabeli 9.1, natomiast w Tabeli 9.2 pokazano sposób przyłożenia jej do obrazu.

- Przyjęto (i jest to wygodne) stosować maski  $(2K + 1) \times (2K + 1)$ ,  $K = 1, 2, \dots$ , ale można stosować też maski o parzystym rozmiarze.
- Dwuwymiarowa szybka transformata Fouriera (FFT) może być zastosowana do efektywnego obliczenia splotu 9.1 (w czasie  $O(\log(M \cdot N))$ ). Podejście to jest zalecane, gdy maska jest duża.
- Jak zwykle, napotykamy problem na brzegu obrazu.
- Będziemy stosować maski do wykrywania krawędzi, ale ich zastosowania są dużo szersze.

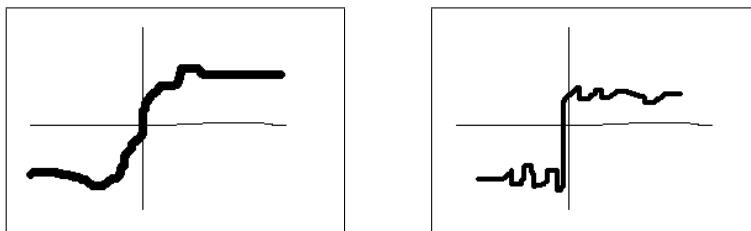


Rysunek 9.1: Powierzchnia obrazu z krawędzią i jej przekrój

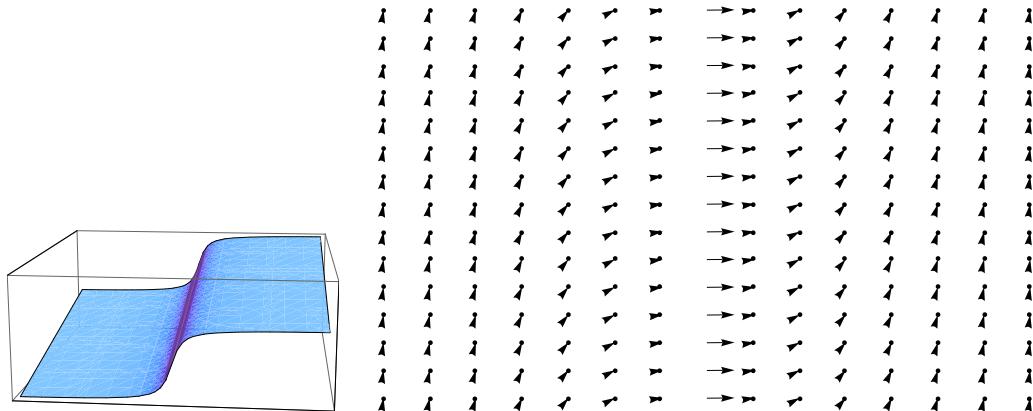
Rysunek 9.2: Krawędzie o coraz mniejszym nachyleniu

### 9.1.2 Definicja krawędzi

Przyjmijmy, że obecna jest tylko jedna krawędź. Na Rys. 9.1 pokazano ją jako powierzchnię oraz w postaci przekroju. Dalej, dla uproszczenia, pokazywać będziemy tylko przekroje krawędzi. Krawędzi jest nagłą zmianą poziomu szarości obrazu, ale kiedy mamy uznać, że zmiana jest nagła? na Rys. 9.2 pokazano na czym polega problem. Trudności z jednoznaczną definicją kraędzi rosną, gdy pojawiają się niejednorodności lub zakłócenia (por. Rys 9.3). Po prostu znamy zazwyczaj przy powyższym ogólnym sformułowaniu tego pojęcia. Można korzystać z robocza definicji: krawędzią jest to, co wykrywa dany detektor krawędzi, przy konkretnych wartościach parametrów. Jak zobaczymy, detektory krawędzi na ogół wykrywają takie zmiany poziomów szarości, które skłonni jesteśmy uznać za krawędzie. Jedynie przy ustawieniu bardzo czołowej detektora pojawia się problem, wynikający z większej czułości kamery i detektora, niż ma nasze oko.



Rysunek 9.3: Krawędzie z zakłóceniami



Rysunek 9.4: Powierzchnia obrazu z krawędzią i odpowiadające jej pole gradientów

### 9.1.3 Pole gradientu

Przypomnijmy z analizy matematycznej, że w kartezjańskim układzie współrzędnych operator gradientu ma postać:

$$\text{grad } f(x, y) = [f_x(x, y), f_y(x, y)]^T,$$

gdzie  $T$  oznacza transpozycję, a pochodne cząstkowe definiujemy następująco:

$$f_x(x, y) = \frac{\partial f(x, y)}{\partial x}, \quad f_y(x, y) = \frac{\partial f(x, y)}{\partial y}.$$

Na Rys. 9.4 pokazano powierzchnię poziomów szarości obrazu z krawędzią i odpowiadające jej pole gradientów. Analiza tego pola pozwala wysnuć dwa wnioski.

1. Wniosek zasadniczy – długość wektora gradientu jest największa na krawędzi.
2. Przekrój tej samej powierzchni, pokazany na Rys. 9.1 będzie nieco różny, gdy przecięcia dokonamy w różnych punktach.

Istotny jest wniosek pierwszy, gdyż nasuwa on pomysł, by jako krawędzie traktować te miejsca, gdzie długość wektora gradientu jest wystarczająco duża (powyżej zadanego progu). Zwykle, jako wskaźnik krawędzi, stosowany jest kwadrat długości  $\text{grad } f(x, y)$

$$G(x, y) = f_x^2(x, y) + f_y^2(x, y)$$

lub  $G_1 = |f_x| + |f_y|$  albo  $G_2 = \max[|f_x|, |f_y|]$ .

Zwracamy uwagę Czytelnika na fakt, że w części książek z przetwarzania obrazów gradientem bywa nazywany kwadrat jego długości:

$$G = ||\text{grad } f|| = \left( \frac{\partial f(x, y)}{\partial x} \right)^2 + \left( \frac{\partial f(x, y)}{\partial y} \right)^2 \quad (9.2)$$

Pochodną w kierunku osi x-ów aproksymuje się zwykle różnicami skończonymi:

$$\frac{\partial f(x, y)}{\partial x} = f_x(x, y) \approx \Delta^{-1} [f(x + \Delta, y) - f(x, y)]$$

lub w wersji symetrycznej:

$$\frac{\partial f(x, y)}{\partial x} = f_x(x, y) \approx (2\Delta)^{-1} [f(x + \Delta, y) - f(x - \Delta, y)],$$

przy czym w praktyce czynnik  $\Delta^{-1}$  jest pomijany – uwzględnia się go przy ustalaniu wartości progu.

Analogicznie przybliża się pochodne cząstkowe po  $y$ .

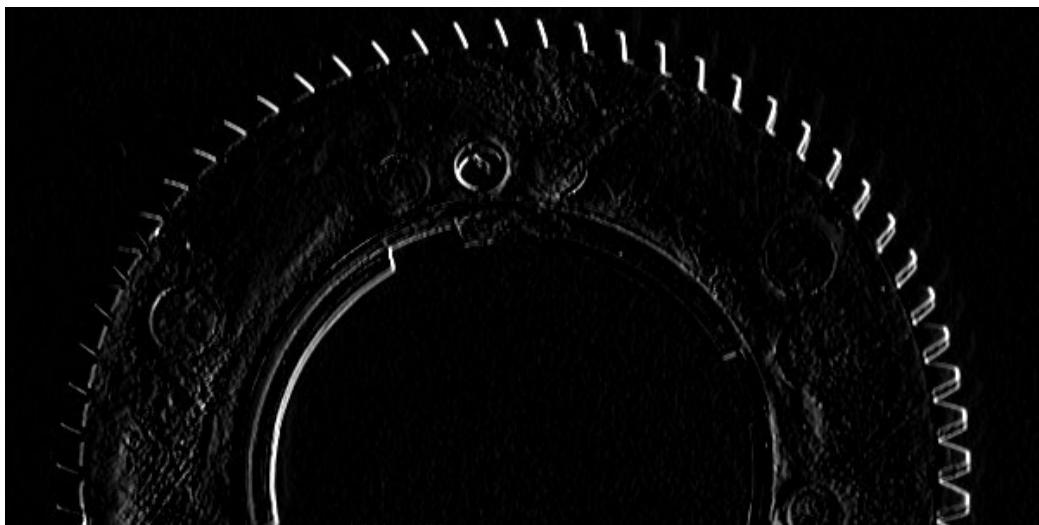
#### 9.1.4 Metody gradientowe

Oznaczając przez  $f_{ij} = f(i\Delta, j\Delta)$ , powyższe przybliżenia oblicza się jako:  $f_{i+1,j} - f_{i,j}$  lub  $f_{i+1,j} - f_{i-1,j}$ , które wygodnie jest zapisywać w postaci masek:  $[0 1]$ ,  $[1 0 -1]$ , odpowiednio.

Jedną ze znanych masek jest maska Prewitta, która ma postać:

$$\begin{bmatrix} 1 & 0 & -1 \\ 1 & 0 & -1 \\ 1 & 0 & -1 \end{bmatrix}$$

Zauważmy, że maska ta jest potrojona w sotsunku do pierwowzoru. Wynika to z wrażliwości operacji różniczkowania na szumy. Poszerzając maskę, uzyskujemy uśrednioną wartość pochodnej. Jeśli skok jest ostatecznie duży to zaznaczamy kontur. Tradycyjnie, kontur zaznacza się kolorem białym na czarnym tle. Efekt zastosowania maski Prewitta pokazano na Rys. 9.5.

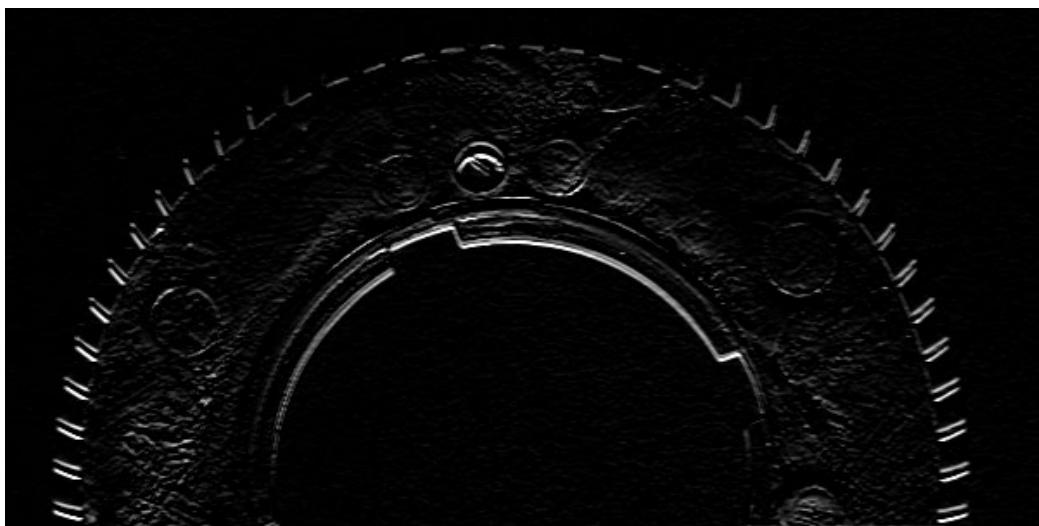


Rysunek 9.5: Obraz po zastosowaniu maski Prewitta

Analogicznie wygląda maska względem  $y$ :

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix}$$

Ten klasyczny detektor konturów jest wyraźnie kierunkowy – wykrywa linie pionowe lub poziome (por. Rys. 9.6).



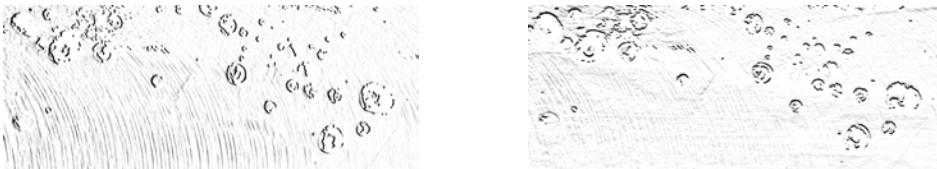
Rysunek 9.6: Obraz po drugiej masce Prewitta

-1	0	1
-2	0	2
-1	0	1

1	2	1
0	0	0
-1	-2	-1

Tabela 9.3: Maski operatora Sobela

Rysunek 9.7: Kierunek  $x$  (po lewej) i  $y$  (po prawej). Bez progowania, ale z odwróceniem poziomów szarości

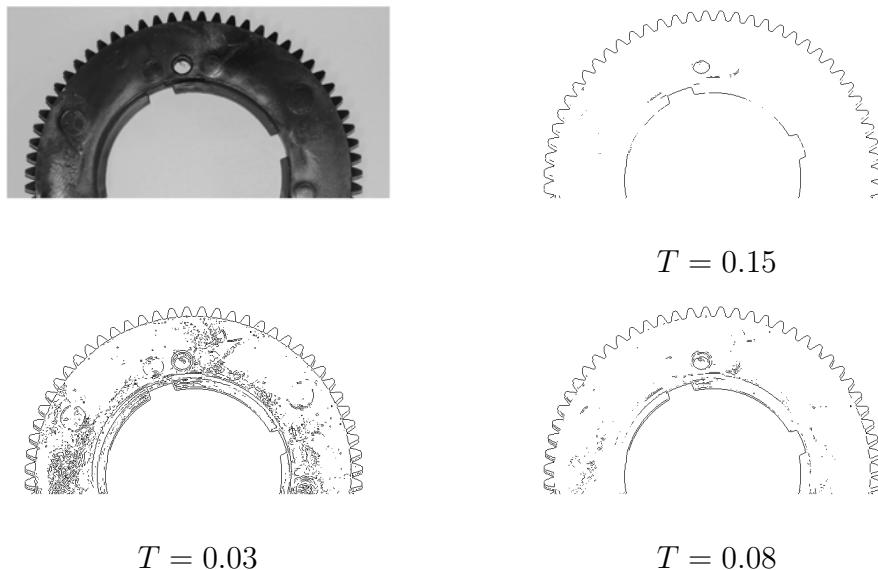
Maski Sobela  $3 \times 3$  kierunkach  $x$  i  $y$  pokazano na Rys 9.3. Wersja Sobela detektora krawędzi kładzie nacisk na piksel środkowy. Efekty zastosowania masek Solea pokazano na Rys. 9.7.

Dobry detektor krawędzi musi dawać na wyjściu zero w obszarze o takiej samej jasności. Implikuje to wymaganie, by suma wszystkich współczynników musi wynosiła zero. **Alogrytm wykrywania krawędzi:**

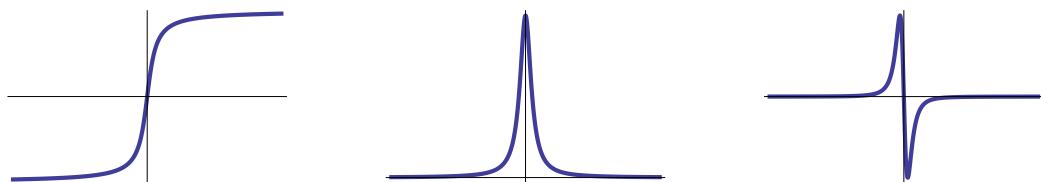
1. Wybierz próg  $T$ .
2. Zastosuj do każdego piksela  $(i, j)$  (poza brzegami obrazu) maskę Sobela (lub Prewitta) w obu kierunkach, otrzymując przybliżenia  $g_x(i, j)$  i  $g_y(i, j)$  funkcji  $f_x$  and  $f_y$ .
3. Jeśli  $G_{ij} = |g_x(i, j)| + |g_y(i, j)| > T$  to uznaj, że piksel należy do krawędzi.

Jak wybrać  $T$ ? Najprostsza wersja:  $T = \delta f \cdot C_{norm}$ , gdzie  $\delta f > 0$  – wartość bezwględna minimalnej zmiany poziomów szarości, którą chcemy traktować jako krawędź,  $C_{norm}$  – stała normalizująca,  $C_{norm}$  zależy od typu maski, sposobu oceny długości wektora gradientu, i od reprezentacji obrazu: skala  $[0, 255]$  lub  $[0, 1]$ .

Rolę doboru progu w wykrywaniu (uznawaniu) konturów ilustruje Rys. 9.8. Obrazy były reprezentowane jako  $[0,1]$  w odcieniach szarości.



Rysunek 9.8: Wykrywanie krawędzi detektorem Sobela z różnym poziomem progu  $T$



Rysunek 9.9: Na wykresach (od lewej) pokazano przekrój krawędzi, pochodną i drugą pochodną

### 9.1.5 Metody drugiego rzędu

W podrozdziale tym omówimy detektory oparte o dyskretyzację operatora Laplace'a.

$$\Delta f(x, y) = \frac{\partial^2 f(x, y)}{\partial x^2} + \frac{\partial^2 f(x, y)}{\partial y^2}, \quad (9.3)$$

który jest 2-wymiarowym odpowiednikiem drugiej pochodnej.

Intuicyjne wyjaśnienie: dlaczego operator podwójnego różniczkowania pozwala wykrywać krawędzie, pokazano na Rys. 9.9

- Z analizy tych wykresów wynika następujący wniosek: miejsce zerowe drugiej pochodnej dokładnie wskazuje położenie krawędzi.

0	1	0
1	-4	1
0	1	0

Tabela 9.4: Maska detektora Laplace'a

$$\frac{1}{4} \cdot \begin{array}{|c|c|c|} \hline 0 & -1 & 0 \\ \hline -1 & 4 & -1 \\ \hline 0 & -1 & 0 \\ \hline \end{array} \quad \frac{1}{8} \cdot \begin{array}{|c|c|c|} \hline -1 & -1 & -1 \\ \hline -1 & 8 & -1 \\ \hline -1 & -1 & -1 \\ \hline \end{array}$$

Tabela 9.5: Znormalizowane maski operatora Laplace'a dla 4 i 8 sąsiadów

- Maska opiera się na następującym przybliżeniu drugiej pochodnej:

z rozwinięcia w szereg Taylora  $\phi'(x - \Delta) \approx \phi'(x) - \Delta \phi''(x)$ ,  $\phi'(x + \Delta) \approx \phi'(x) + \Delta \phi''(x)$ . Odejmując, uzyskujemy:

$$\frac{\partial^2 f(x, y)}{\partial x^2} \approx 0.25 \Delta^{-2} [f(x + \Delta, y) - 2 \cdot f(x, y) + f(x - \Delta, y)]$$

Powtarzając to również dla  $y$ , otrzymujemy maskę pokazaną w Tabeli 9.4. Trzeba zauważyć, że  $-\Delta f$  przecina zero w tym samym punkcie co  $\Delta f$ .

Znormalizowane maski Laplace'a dla 4-sąsiadztwa i 8-sąsiadztwa pokazano w Tabeli 9.5

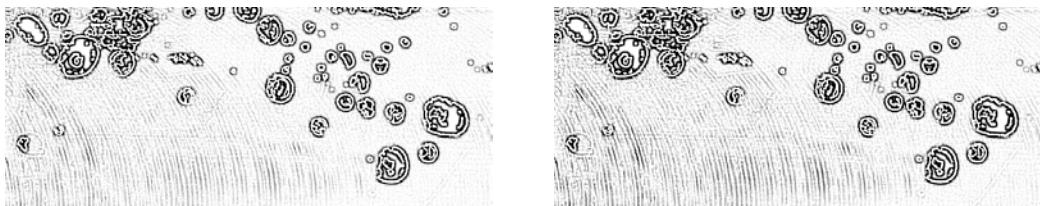
Pozostaje tylko wyjaśnić: jak znaleźć punkty przecięć zera? Oznaczmy przez  $R(i, j)$  wynik zastosowania jednej z masek pokazanych w Tabeli 9.4 do obrazu  $f_{ij}$ . Zdefiniujmy:  $R^+(i, j) = \max[0, R(i, j)]$ ,  $R^-(i, j) = \min[0, R(i, j)]$ . Ze względu na obecność szumu i jego wzmacnianie w wyniku przybliżonego różniczkowania, definiujemy także:

$$r_{ij}^+ = \max [R^+(i, j) \text{ w oknie } 3 \times 3]$$

$$r_{ij}^- = \min [R^-(i, j) \text{ w oknie } 3 \times 3].$$

Zauważmy, że  $r_{ij}^+$  jest największą spośród dodatnich wartości  $R(i, j)$  w danej okolicy, a  $r_{ij}^-$  to najmniejsza spośród ujemnych wartości  $R(i, j)$  w tejże okolicy. Jeśli zatem wartości te różnią się znacznie, to możemy przyjąć, że nastąpiło przejście przez zero w okolicy punktu  $(i, j)$  i oznaczyć go jako potencjalnie przynależny do krawędzi.

Wynik zastosowania masek z Tabeli 9.5 pokazano na Rys. 9.10. W celu zwiększenia czytelności tych obrazów wyświetlono je w negatywie.



Rysunek 9.10: Obrazy (macierze  $R$ ), otrzymane w wyniku zastosowania masek operatora Laplace'a do obrazu testowego

### Detektory krawędzi oparte na operatorze Laplace'a

Podsumowując, algorytm detekcji krawędzi działa według poniższego schematu.

- Wybierz próg  $T > 0$ .
- Dla każdego  $(i, j)$  oblicz:

$$r_{ij}^+ = \max [R^+(i, j) \text{ w oknie } 3 \times 3]$$

$$r_{ij}^- = \min [R^-(i, j) \text{ w oknie } 3 \times 3]$$

- Jeśli  $r_{ij}^+ - r_{ij}^- > T$ , to uznajemy, że  $(i, j)$  należy do krawędzi.

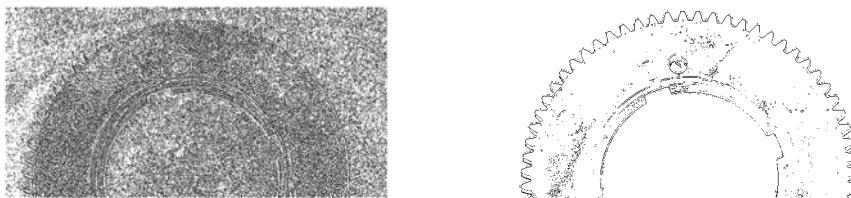
**Pytanie:** Dlaczego w ostatnim kroku tego algorytmu użycie  $|.|$  nie jest konieczne?

### Dobór progu

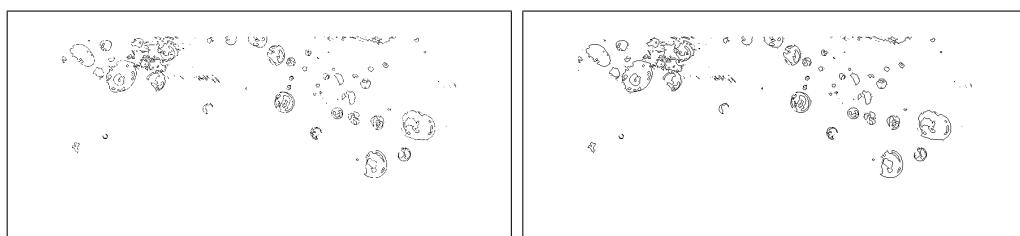
Jak już wspomniano, fluktuacje wartości  $R_{ij}$  są zwykle duże, gdyż dwukrotnie różniczkujemy nie tylko obraz, ale także szum. Utrudnia to właściwy dobór progu, co pokazano na Rys. 9.11 i 9.12. Na Rys. 9.11 widać, że nawet w tak prostym przykładzie trudno wybrać właściwy próg dla przecięcia zera. Nieuchciane ślady są wciąż widoczne. Rozpowszechnione w literaturze i w implementacjach programowych ulepszenie detektora Laplace'a przedstawiamy w następującym podrozdziale.

### Detektor LoG

Idea ulepszenia nadmiernie czułego detektora Laplace'a jest bardzo prosta: zastosować filtr wygładzający w celu pozostawienia jedynie dużych (wysokich) krawędzi, a następnie zastosować detektor Laplace'a. Do wygładzania zastosować można każdy wygładzający (dolnoprzepustowy) filtr. Jednakże,

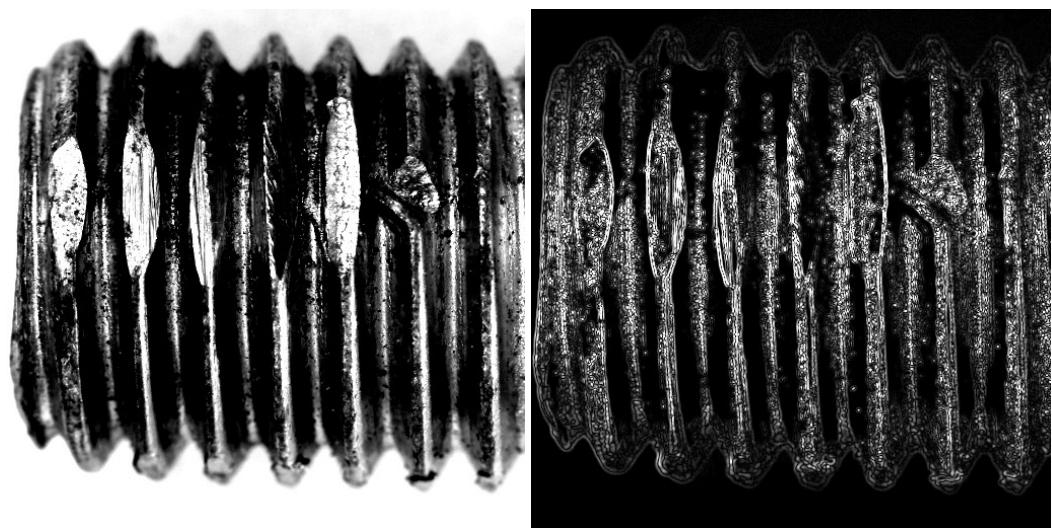


Rysunek 9.11: Po lewej – krawędzie wykryte dla  $T = 0.005$  (wszystko jest krawędzią). Po prawej –  $T = 0.1$  (wynik wciąż nie jest zadowalający)



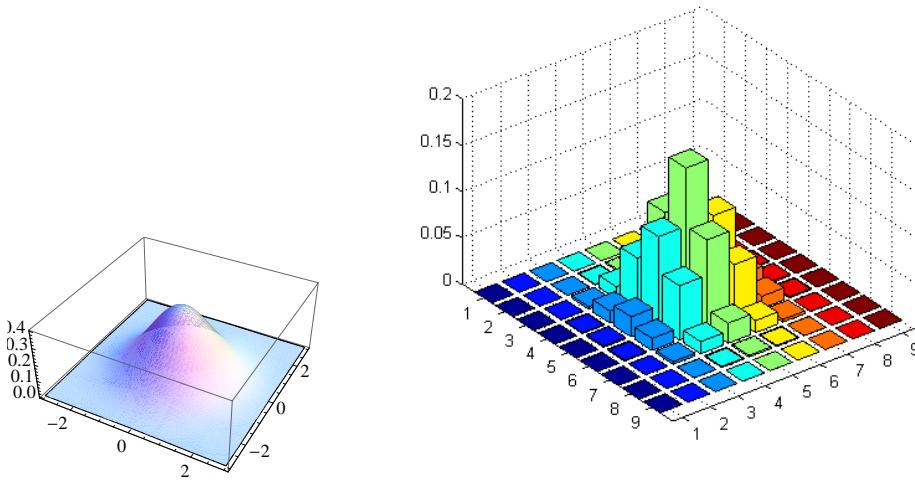
Rysunek 9.12: Po lewej – krawędzie dla  $T = 0.25$ . Po prawej –  $T = 0.12$ .

w tym zastosowani najbardziej popularny jest filtr z jadrem gaussowskim, stąd metoda ta znana jest pod nazwą **LoG** (od ang. *Laplacian of Gaussian*).



Rysunek 9.13: Obraz oryginalny (po lewej) i po LoG (po prawej)

W wersji ciągłej (względem zmiennych przestrzennych) filtr gaussowski



Rysunek 9.14: Jądro gaussowskie i jego zdyskretyzowana wersja w oknie  $9 \times 9$  dla  $\sigma = 1$

ma postać:

$$g(x, y) = \int \int H(x - x', y - y') f(x', y') dx' dy' , \quad (9.4)$$

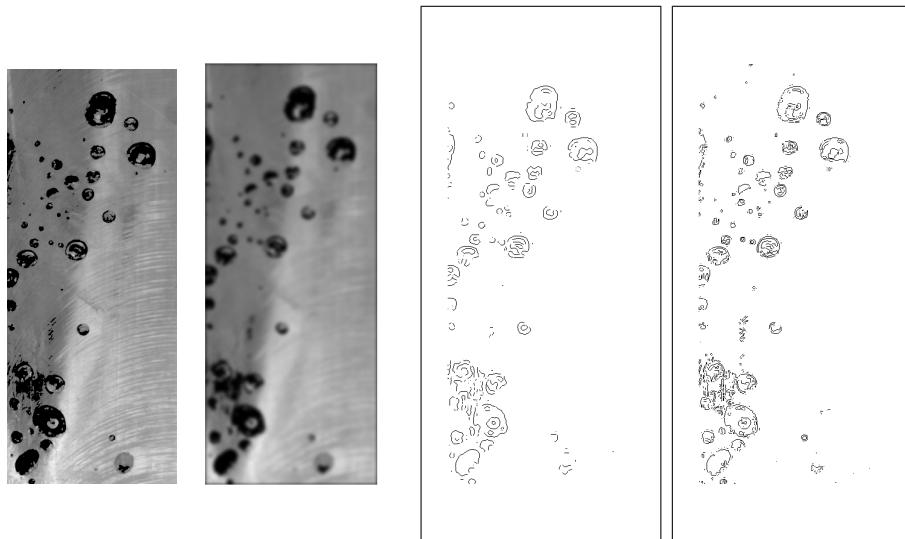
gdzie:

$$H(x, y) = \frac{1}{\sqrt{2 \pi} \sigma} \exp \left[ -\frac{x^2 + y^2}{2 \sigma^2} \right] . \quad (9.5)$$

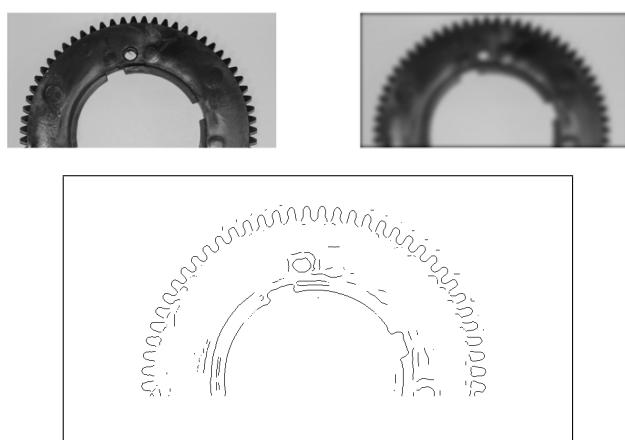
W praktyce, stosowana jest dyskretna wersja splotu (9.4) wraz z próbko-waną na siatce równomiernej wersją (9.5), oznaczaną dalej przez  $h(k, l)$ . Po dyskretyzacji dostajemy:

$$g_{ij} = \sum_{k=-K}^K \sum_{l=-K}^K f_{(i+k)(j+l)} \cdot h(k, l) \quad (*)$$

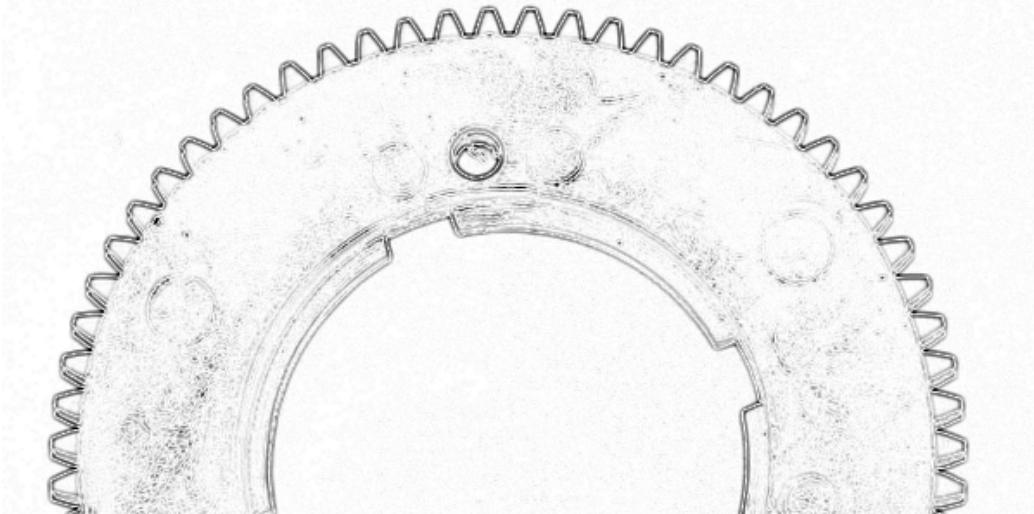
Po filtracji stosujemy maske operatora Laplace'a, a następnie metodę wyszukiwania przecięć zera. Ceną za zmniejszenie wrażliwości detektora Laplace'a jest wprowadzenie dodatkowego parametru  $\sigma > 0$ , od którego zależy stopień rozmywania szczegółów obrazu. Zatem, jego wybór wpływa także na decyzje, co uznajemy za krawędzie. Jako wskazówkę do wyboru  $\sigma$ , przypomnijmy, że  $H(x, y)$  jest prawie zerowe dla  $|x|$  i  $|y|$  większych od  $3\sigma$ . Detektor LoG jest bardzo czuły na parametrów. Czasem konieczna jest subtelna zmiana wielkości okna,  $\sigma$  lub  $T$ . Ilustruje to następny przekład. Na Rys. 9.16 pokazano wynik detekcji LoG z zastosowaniem filtracji w oknie  $21 \times 21$  z parametrem



Rysunek 9.15: Od lewej: ogrinalny slab, slab wygładzony filtrem gaussowskim w oknie 13 przy  $\sigma = 3$ , wynik detekcji krawędzi dla  $T = 0.015$ . Po prawej – koncowy efekt detekcji z filtrem w oknie 5,  $\sigma = 5$  i progiem  $T = 0.001$



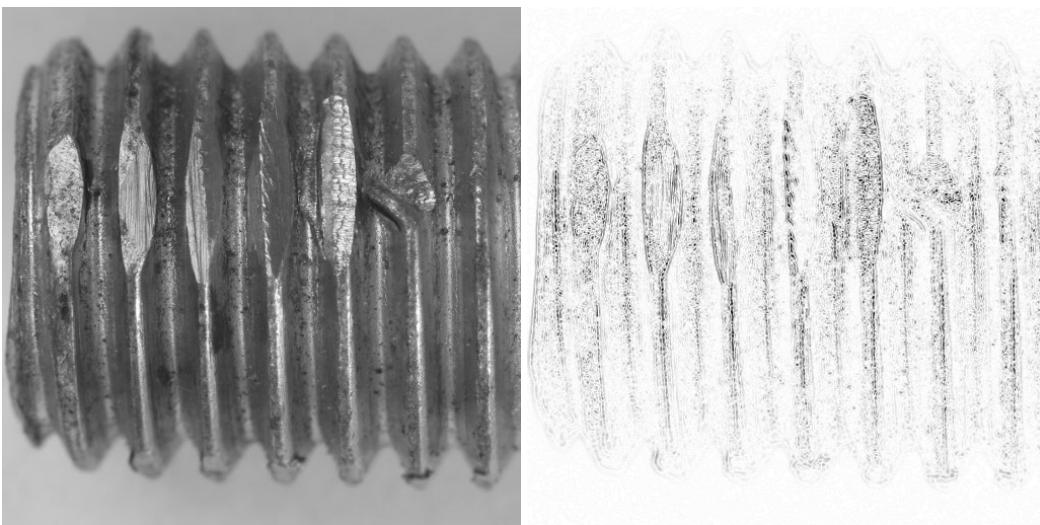
Rysunek 9.16: Od lewej: obraz oryginalny, obraz po filtracji gaussowskiej, wynik wykrywania krawędzi detektorem LoG



Rysunek 9.17: Obraz po detekcji LoG – okno o rozmiarze  $3 \times 3$

wygładzania  $\sigma = 6$  i progiem  $T = 0.001$ . Użycie do filtracji okna  $3 \times 3$  daje obraz pokazany na Rys. 9.17.

Jeśli rezultat nie jest zadowalający (jak na Rys. 9.18)



Rysunek 9.18: Obraz przed i po LoG – rezultat jest niezadowalający

to jedną z przyczyn, oprócz złego doboru parametrów, może być zbyt mały kontrast.

### 9.1.6 Inne detektory krawędzi

Znanych jest wiele innych detektorów krawędzi:

- Roberts'a,
- Canny (w pewnym sensie optymalny),
- Kirsch'a,
- Marr-Hildreth,
- Robinsona,
- Mallat-Zhong (oparty o falki).

Powyższa lista nie jest pełna. Można do niej dodać detektor o wdzięcznym akronimie SUSAN (por. [72]) oraz jego modyfikację [59] i wiele innych, co świadczy o tym, że zadanie detekcji konturów jest trudne.

Poniższa rada może być użyteczne w połączeniu z dowolnym z wcześniejszej wymienionych detektorów.

- Zwiększyć rozmiary maski w celu silniejszego wygładzenia i ograniczenia się do wykrywania tylko konturów wzdłuż, których przechodzi granica oddzielająca obrzy o bardzo dużych zmianach jasności. (dla detektorów Sobela, Prewitta, ... operatory masek  $5 \times 5$ ,  $7 \times 7$ , ... są znane).

### 9.1.7 Ocena detektorów krawędzi

Poniżej zestawiono cechy, którymi powinien charakteryzować się dobry i dobrze zestrojony detektor krawędzi.

- Niezawodność wykrycia – powinien zapewniać kompromis między niskim prawdopodobieństwem nie wykrycia rzeczywistej krawędzi a niskim prawdopodobieństwem wykrycia nie istniejącej krawędzi. Oba te prawdopodobieństwa nie mogą być jednocześnie minimalizowane.
- Właściwa lokalizacja krawędzi – wykryte punkty powinny być tak blisko rzeczywistych punktów na krawędzi jak to tylko możliwe. Wymaganie to wynika z faktu, że detektory krawędzi korzystają z porównań poziomów szarości w pewnej okolicy badanego piksela i mogą mu przypisać krawędź, która w istocie znajduje się o kilka pikseli dalej.

- Szerokość krawędzi – jest wskazane, aby detektor dostarczał tylko pojedynczy punkt w środku krawędzi. Powód tego wymagania jest taki sam jak powyżej – skoro punkty zaliczne do krawędzi mogą leżeć o kilka pikseli dalej, to zarówno one jak i właściwa krawędź zostaną zaznaczone na obrazie wynikowym.

Po dokładnych badaniach statystycznych Pratt [54], na stronie 489, tak charakteryzuje detektory z punktu widzenia dużego prawdopodobieństwa wykrycia.

- Operatory  $3 \times 3$  Sobela i Prewitta są lepsze od detektora Robertsa  $2 \times 2$ .
- Operator Prewitta jest lepszy niż Sobela dla pionowych krawędzi. Detektor Sobela jest lepszy dla krawędzi ukośnych.
- Operatory różniczkowe Sobela i Prewitta są odrobinę lepsze niż operatory Robinsona (trój- i pięcio-poziomowy).
- Detektor, który dostarcza cienkich krawędzi może być bardzo zawodny. Stosujemy więc operacje pocieniające krawędzie, które wykryto za pomocą bardziej czułego detektora.
- Czułość na położenie krawędzi (patrz [54] str 486): operator Sobela zapewnia najbardziej liniową odpowiedź. Laplasjan jest symetryczny więc niewrażliwy na orientację krawędzi.

## 9.2 Opisy krawędzi

Podobnie jak w przypadku progowania, wykryte na krawędziach punkty zostają zaznaczone. Człowiek łatwo może zobaczyć ich kształt, ale – bez dodatkowego oprogramowania – komputer tego nie potrafi. Co zatem możemy zrobić z punktami oznaczonymi jako należące do krawędzi ?

- Zakodować je do bardziej zwartej, mniej pamięcio-chłonnej, postaci.
- Połączyć metodą dopasowywania krzywych.
- Obliczyć deskryptory fourierowskie, które pozwolą rozpoznać obiekty.
- Szukać obiektów o zadanym kształcie.

W dalszej części tego rozdziału omawiamy techniki, służące do realizacji wyżej wymienionych zadań (por. [54], [13], [25]).

NW	N(orth)	NE
W(est)	(C)enter	E(ast)
SW	S(outh)	SE

Tabela 9.6: Symbole używane do kodowania krawędzi – pozostawiamy je jako skróty angielskich wyrazów

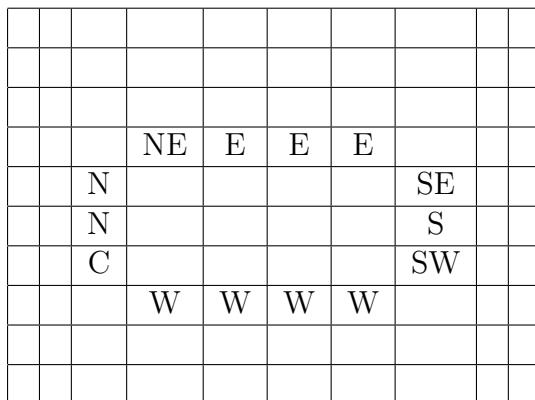


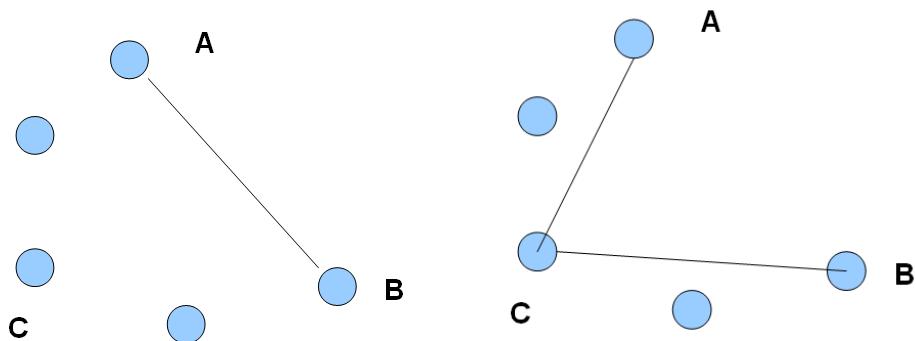
Tabela 9.7: Łańcuch kodujący kontur w kształcie elipsy

### 9.2.1 Kodowanie łańcuchowe

Kodowanie łańcuchowe jest prostą techniką przechowywania krawędzi w zwarnej formie. Dla 8-sąsiedzwa stosujemy 8 symboli, zwykle są to cyfry 1,2,...,8,, ale my – ze względów dydaktycznych – będziemy stosować symbole pokazane w Tabeli 9.6. **Schemat kodowania** punktów należących do krawędzi jest następujący.

1. Wśród punktów zaznaczonych jako krawędzie znajdź punkt, który nie został jeszcze przypisany do żadnego łańcucha. Jeśli takich punktów nie ma, to STOP.
2. W punkcie tym ustaw centrum maski z Tabeli 9.6 i rozpoczęź nowy łańcuch od litery C.
3. Jeśli powyższa maska pokrywa kolejny punkt (punkty) zaznaczony jako krawędź, to dopisz do łańcucha odpowiedni symbol (symbole) z Tabeli 9.6 i przesuń centrum maski do tego punktu.
4. Gdy w kolejnym położeniu maska nie pokrywa żadnego z punktów označonych jako krawędź – zakończ łańcuch i przejdź do Kroku 1.

Działanie powyższego schematu pokazano w Tabeli 9.7. Uwaga: w opisanej



Rysunek 9.19: Idea łączenia punktów krawędzi

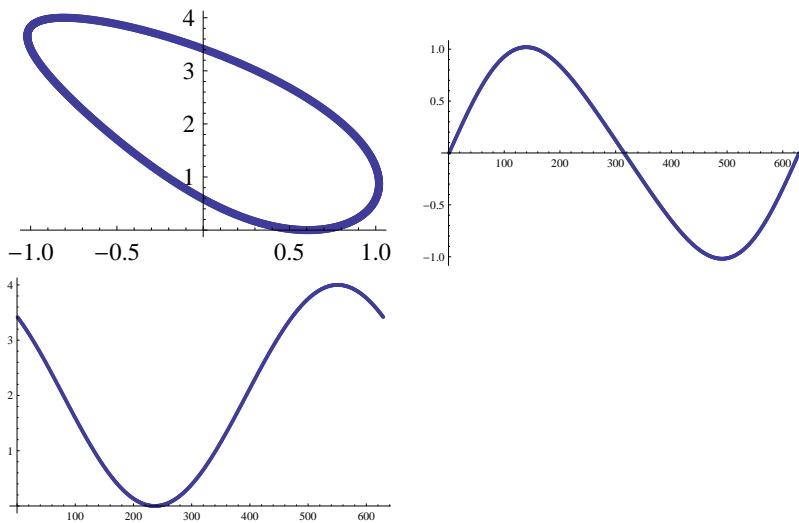
wersji algorytm nie gwarantuje niezależności kodu łańcuchów od obrotów obiektów. Aby taką niezależność uzyskać potrzebne są dodatkowe operacje (por. [54]).

### 9.2.2 Odcinkowo-liniowe łączenie krawędzi

Pochodząca od Forsena, prosta i skuteczna metoda odcinkami liniowego łączenia krawędzi przebiega następująco (por. [54]).

- Zaczynamy od dwóch wybranych punktów A, B, które były zaznaczone jako fragment krawędzi.
- Połącz A i B linią prostą.
- Znajdź punkt, C – który jest najbardziej odległy od linii A–B.
- Jeśli odległość ta jest większa niż zadany próg  $\epsilon > 0$ , to podziel linię A,B na dwa segmenty A–C i C–B i usuń linię A–B.
- Powtarzaj powyższe kroki dopóki wszystkie punkty zaznaczone jako krawędź nie zostaną połączone lub ich odległość od połączeń jest mniejsza niż  $\epsilon > 0$ . Te ostatnie usuń jako zbędne dla osiągnięcia zadanej dokładności przybliżenia.

Ta prosta idea została pokazana na Rys. 9.19.



Rysunek 9.20: Parametryczny opis krzywej zbliżonej kształtem do elipsy

### 9.2.3 Fourierowski opis konturu

Pomysł zastosowania rozwinięć w szereg Fouriera do opisu konturów przedstawił Cosgriff (1960) (por. [54] str. 610). Ogólna idea jest następująca.

- Wyraź współrzędne kartezjańskie pikseli zaznaczonych jako krawędzie we współrzędnych biegunkowych.
- Oblicz przybliżone wartości współczynników Fouriera dla tego konturu.
- Odrzuć współczynniki o małych wartościach bezwględnych, jako reprezentujące szum.
- Zachowaj te, które opisują główny kształt.

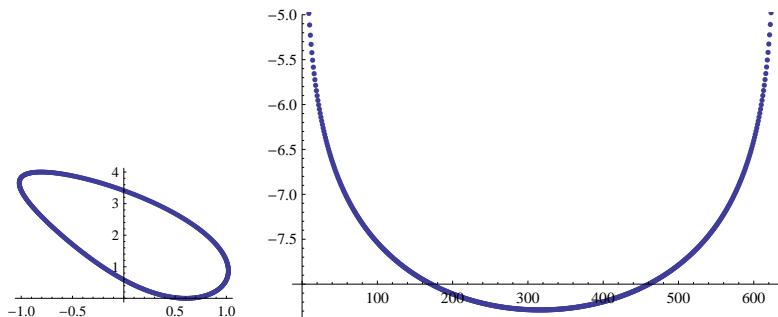
Przejdźmy do bardziej szczegółowego opisu. Wyobraźmy sobie idealny kontur obiektu opisywany przez

$$c(t) = x(t) + j \cdot y(t), \quad j^2 = -1,$$

$t$  = długość łuku, zwykle  $t \in [0, 2\pi]$ . Rozwińmy  $c(t)$  w szereg Fouriera:

$$c(t) = \sum_{k=-\infty}^{\infty} C_k \exp(j k t), \quad (9.6)$$

$$C_k = \frac{1}{2\pi} \int_0^1 c(t) \exp(-j k t) dt.$$



Rysunek 9.21: Krzywa  $c(t)$  o kształcie zbliżonym do elipsy (po lewej) i wykres modułów współczynników  $|C_k|$

- Pozostawmy w szeregu 9.6 tylko  $(2K + 1)$  elementów o numerach od  $-K$  do  $K$ .
- Obliczmy oszacowania dla  $C_k$ , stosując dyskretną transformatę Fouriera (w wersji FFT) do punktów konturu.

Współczynniki  $C_k$ ,  $k = -K, \dots, 0, 1, \dots, K$  nazywane są deskryptorami. Pozwalają przechowywać najważniejsze cechy  $c(t)$  i mogą być stosowane do rozpoznawania kształtu oraz do rekonstrukcji – przez zsumowanie szeregu:

$$rec(t) \sum_{k=-K}^K C_k \exp(j k t). \quad (9.7)$$

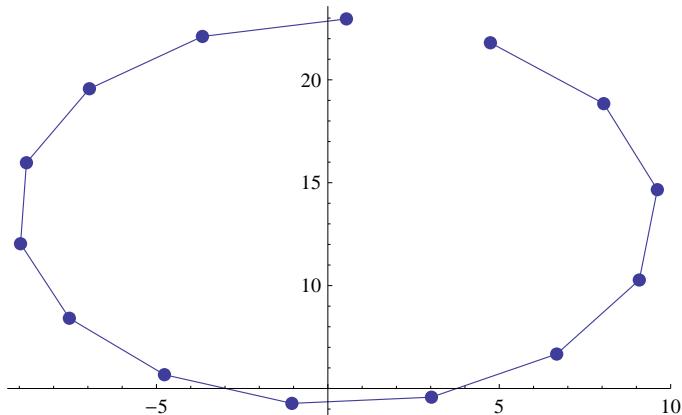
W praktyce  $rec(t)$  obliczane jest w dyskretnych, równoodległych punktach i do sumowania można użyć szybkiej transformaty Fouriera.

**Przykład 1** Na Rys. 9.21 pokazano kontur i wykres współczynników  $\log(|C_k|)$ . Na Rys. 9.22 pokazano wynik rekonstrukcji kształtu na podstawie 14 deskryptorów. Zauważmy, że mała deformacja orginalnej elipsy nie jest widoczna. Kąt obrotu nie jest zachowany.

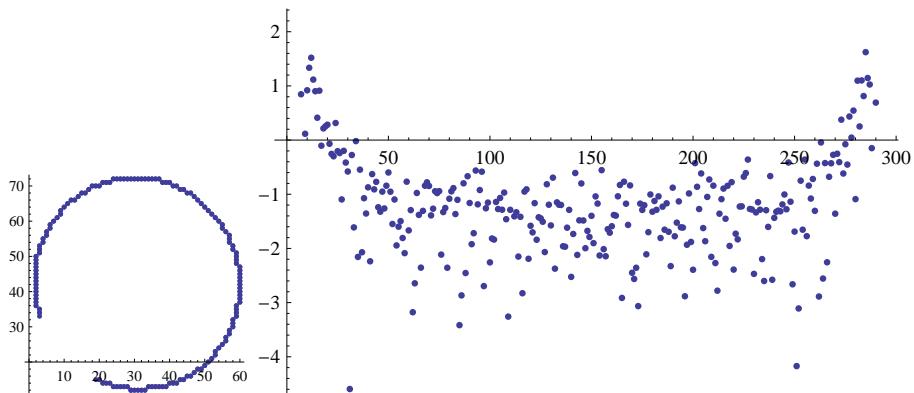
**Pytanie:** czy niezachowanie kata obrotu pierwotnej, zniekształcone, elipsy jest cechą pozytywną, czy negatywną?

**Przykład 2** Poprzedni przykład był wyidealizowany w tym sensie, że znaliśmy dokładnie współrzędne punktów leżących na sztucznie wygenerowanej krzywej.

Zobaczmy w jakim stopniu deskryptory fourierowskie pozwolą rekonstruować kształt, który powstał w wyniku faktycznego wykrywania krawędzi monety – por. Rys. 9.23. W prawej części tego rysunku pokazano wykres modułów wartości współczynników Fouriera.



Rysunek 9.22: Rekonstrukcja kształtu na podstawie 14 deskryptorów.

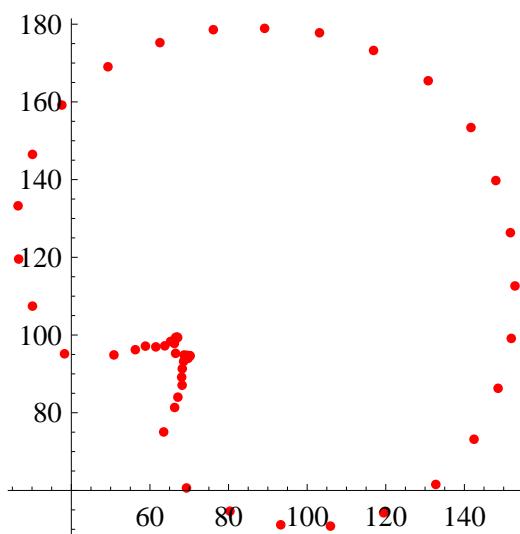


Rysunek 9.23: Krawędzie monety i wykres 295 modułów wartości współczynników Fouriera

Do rekonstrukcji kształtu wykorzystano z 52 z 285 współczynników. Efekt pokazano na Rys. 9.24. Jak widać, kontur monety został prawidłowo zrekonstruowany w tych obszarach, w których były obserwacje. Natomiast tam, gdzie ich zabrakło, rekonstrukcja kształtu jest nieprawidłowa.

Na zakończenie rozważań o deskryptorach warto poczynić kilka uwag.

- Rozwinięcie krawędzi w szereg Fouriera jest przykładem bardzo szerokiej klasy deskryptorów.
- Istnieją deskryptory oparte o inne szeregi ortogonalne jak wielomiany Legendra, falki itd.
- Deskryptory są wrażliwe na duże przerwy w wykrytych krawędziach.



Rysunek 9.24: Wynik rekonstrukcji brzegu monety na podstawie 52 deskrypcji Fouriera

Należy je stosować, gdy wykrycie krawędzi jest pewne (zapewnić właściwe oświetlenie).

# Rozdział 10

## Wykrywanie obiektów na podstawie konturów i transformacja Hougha

### 10.1 Proste metody wykrywania

#### 10.1.1 Operacje logiczne na obrazach binarnych

Jeśli  $f_{ij}$  i  $g_{ij}$  (wejścia) są obrazami binarnymi o wartościach  $\{0, 1\}$ , to wszystkie operacje boolowskie, zastosowane do każdego piksela, są dobrze zdefiniowane i użyteczne:

- $f_{ij} \text{ AND } g_{ij}$  – do porównywaniapróbki odniesienia z przychodzącym produktem,
- $f_{ij} \text{ OR } g_{ij}$  – do zbieraniaobiektów (defektów) z serii obrazów,
- $f_{ij} \text{ XOR } g_{ij}$  – do wyboru nietypowych obiektów, które są obecne na  $f_{ij}$  lub na  $g_{ij}$  ale nie na obu z nich.

Uwagi.

- Powyżej przyjęto, że tło jest czarne = 0, podczas gdy obiekty są białe = 1
- Czasami dla lepszej widoczności powinno się obrócić obraz ( $\text{NOT } f_{ij}$ ).
- Użycie AND do porównywania obiektów wymaga bardzo precyzyjnego pozycjonowania, inaczej uzyskane wyniki są mylące.

$\frac{1}{6}$	-1	2	-1	$\frac{1}{6}$	-1	-1	-1	$\frac{1}{6}$	2	-1	-1
	-1	2	-1		2	2	2		-1	2	-1
	-1	2	-1		-1	-1	-1		-1	-1	2

Tabela 10.1: Maski do wykrywania linii

### 10.1.2 Wykrywanie fragmentów linii – maski

Proste, lokalne detektory linii korzystają z masek pokazanych w Tabeli 10.1. Ich użycie jest podobne do wykrywania krawędzi – przesuwaj je po powierzchni obrazu i porównuj wynik z zadanym progiem.

- Zalety metody – prostota.
- Wada – w wyniku nie uzyskujemy matematycznego opisu linii.

#### Pytania

1. Jak będzie wyglądać czwarta, brakująca w Tabeli 10.1, maska ?
2. Jaki będzie efekt, jeśli zastosujemy maski o większych rozmiarach ?

## 10.2 Współrzędne biegunkowe – sygnatury obiektów

### 10.2.1 Sygnatury

Wykryte krawędzie są jedynie punktami. Podobnie jak w przypadku segmentacji, komputer nie jest w stanie ich zinterpretować. Co zatem możemy zrobić z wykrytymi krawędziami?

- Poprawić je – gdyż jest wiele niezaznaczonych punktów na krawędzi pomiędzy tymi zaznaczonymi.
- Opisać je jako krzywe.
- Próbować stosować je w takiej surowej postaci do rozpoznawania obiektów.
- Można próbować dopasować chmurę punktów reprezentujących krawędź do zadanego kształtu (jak przy segmentacji obiektów). Nie jest wskazane stosowanie bezpośrednio, gdyż krawędzie są bardziej nierównie niż obiekty.



Rysunek 10.1: Koło i jego wykres we współrzędnych biegunowych (sygnatura).

- Dla typowego, niezbyt skomplikowanego kształtu można stosować dopasowanie 1-D poprzez transformację krawędzi od współrzędnych biegunowych i porównanie ich z zadanym kształtem (także we współrzędnych biegunowych) Taka wzorcowa krzywa nazywana bywa sygnaturą obiektu.

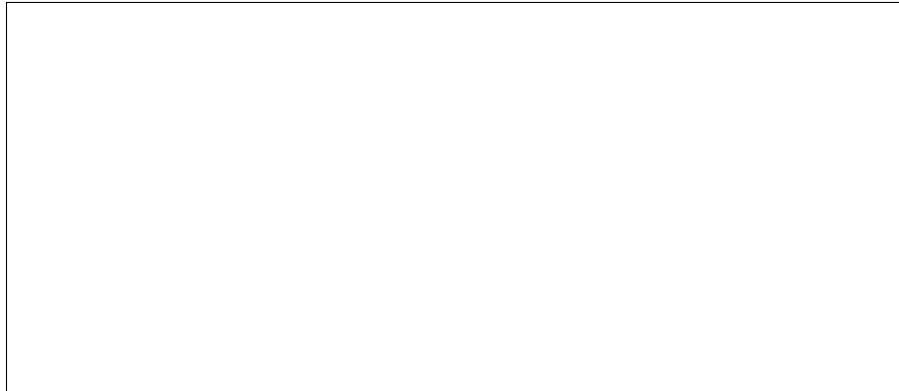
Klasa krzywych może być reprezentowana przez promień  $r(\theta)$  jako funkcję kąta  $\theta$ .

Punkt  $(x, y)$  we współrzędnych kartezjańskich może być reprezentowany jako:  $r = \sqrt{x^2 + y^2}$ ,  $\theta = \text{arc tg}(y/x)$  we współrzędnych biegunowych. Przekształcenie odwrotne ma postać:  $x = r \cos(\theta)$ ,  $y = r \sin(\theta)$ . Zasadniczą trudność stanowi wykres  $(r, \theta)$  – może mieć wiele wartości (na przykład, czołg z lufą).

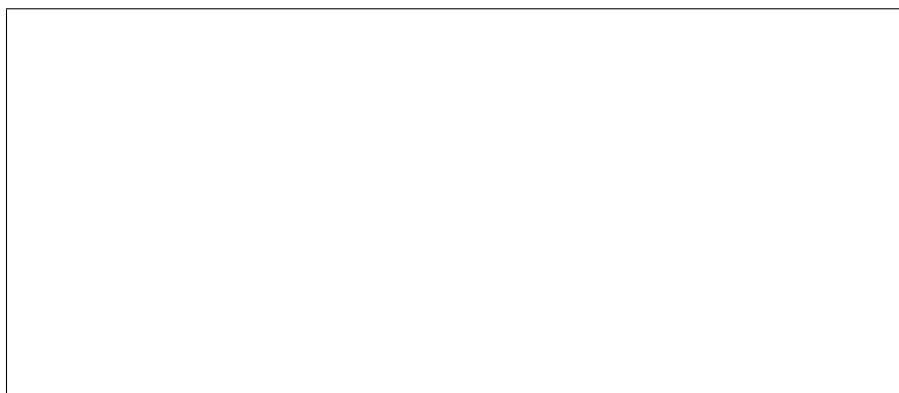
Freeman (1978) proponuje przechowywanie dodatkowej funkcji  $\Phi(s) \in \{-1, 1\}$ . Funkcja  $\Phi(s)$  jest nazywana funkcją fazową, gdzie  $\Phi(s) = -1$  odpowiada krzywej „zamkniętej”. Podejście Freemana jest znane jako wykres  $(r, s)$ .

Dalej wyjaśnimy wykresy  $(r, \theta)$  o podobnych własnościach. Najłatwiej zgromadzić intuicje, oglądając sygnatury prostych kształtów geometrycznych. Na Rys. 10.1 pokazano sygnaturę okręgu, na Rys. 10.2 – sygnatury dwóch elips a na Rys. 10.3 dwie figury zbliżone do trójkąta o zaokrąglonych wierzchołkach. W zasadzie nie powinno się wyciągać ogólnych wniosków na podstawie przykładów, lecz w tym przypadku wystarczy chwila refleksji, aby wyciągnąć następujące wnioski:

1. Obrotom figury dopowiadają sygnatury o tym samym kształcie, lecz przesunięte wzdłuż osi  $\theta$ .
2. Podobne kształty mają gładko zmieniające się sygnatury. Trójkąt ma



Rysunek 10.2: Elipsy i ich sygnatury. Obroty wyłącznie je przesuwają.



Rysunek 10.3: Trójkąt o zaokrąglonych kątach.

sygnaturę piłozębna, natomiast lekko zniekształconym trójkątom na Rys. 10.3 odpowiadają funkcje o lekko wygładzonych ekstremach.

### 10.2.2 Centroid

Ważnym pojęciem dla zastosowań sygnatur jest centroid, czyli centralny punkt w którym umieszcza się współrzędne (0,0) biegunowego układu współrzędnych. To właśnie poprawne wyznaczenie położenia centroidu warunkuje dobre wyniki zastosowania sygnatur do klasyfikacji kształtów.

- Wyznaczenie centroidu jest proste, ale tylko wtedy gdy kontur ma wyłącznie regularne kształty.
- Jak znaleźć znaleźć, gdy znamy tylko wykryte punkty na brzegu ?
- Najprościej – jako empiryczną średnią punktów tworzących krawędzie.



Rysunek 10.4: Część krawędzi koła i jej wykres we współrzędnych biegunowych z centroidem na (0,0)

- Jednakże, jak wiadomo, średnia arytmetyczna nie jest odporna na duże odstępstwa w danych, a tutaj często mamy do czynienia z taką sytuacją. Powstaje ona wtedy, gdy znaczny fragment konturu nie zostanie wykryty.

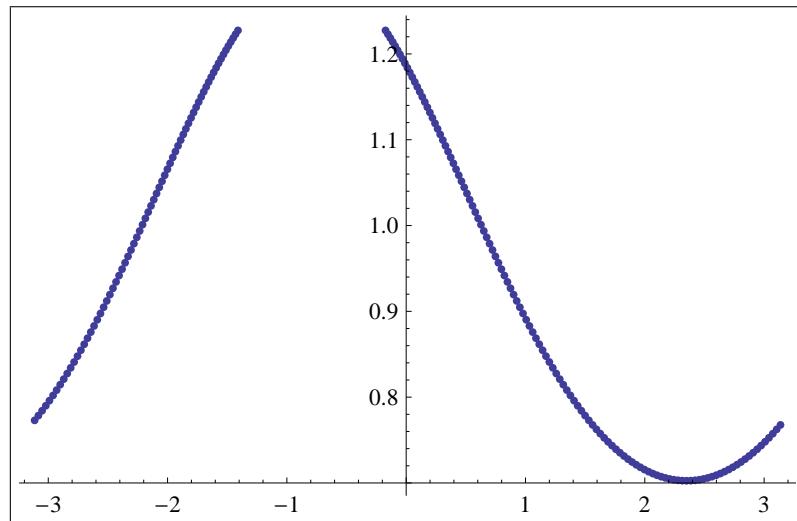
Powyzsze kłopoty z doborem centroidu pokazano na Rys. 10.4 i Rys. 10.5. Na pierwszym z nich pokazano wykryty fragment okręgu, a centroid umieszczono w punkcie (0, 0). W rezultacie, wykres we współrzędnych biegunowych jest całkiem poprawny (prawa część Rys. 10.4). W lewej części tego rysunku zaznaczono (grupą kropą) centroid uzyskany jako średnia arytmetyczna wykrytych fragmentów brzegu. Jeśli w punkcie tym umieścimy centroid, to wykres we współrzędnych biegunowych będzie wyglądał tak, jak Rys. 10.5 i dość trudno będzie zakwalifikować ten obiekt jako okrąg.

### 10.2.3 Dopasowywanie wykresów

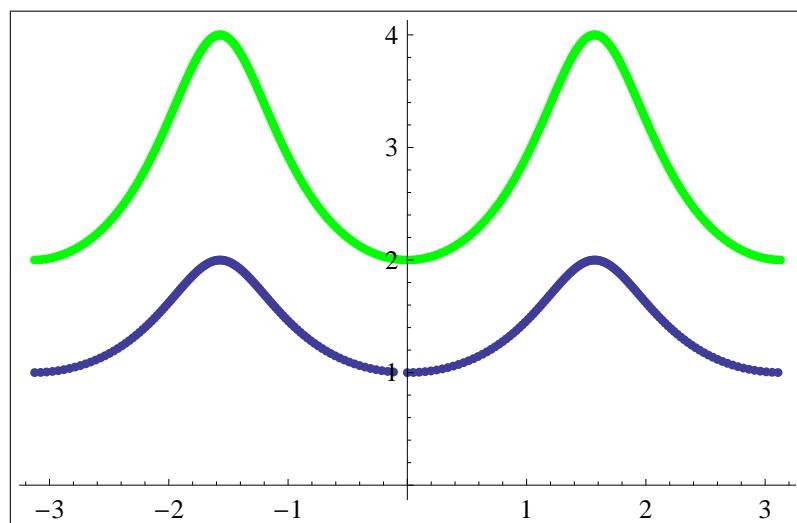
Jak zatem znaleźć dobry centroid? Nie ma ogólnej metody. Można stosować metodę dopasowywania, którą opisano poniżej.

- Oznaczmy teoretyczną sygnaturę przez  $\xi(\theta)$ ,  $\theta \in [0, 2\pi]$ . Przed dopasowaniem jej do wykrytych punktów brzegowych należy przekształcić ją oraz dane do tej samej skali. Potrzeba taka widoczna jest na Rys. 10.6.
- W tym celu najprościej jest przyjąć, że idealny profil  $\xi(\theta)$ , ma skalę  $\max_{\theta \in [0, 2\pi]} \xi(\theta) = 1$ .
- Skalowanie wykrytych punktów brzegowych.
  1. Dane są  $(x'_i, y'_i)$ ,  $i = 1, 2, \dots, n$  – punkty na brzegu.
  2. Określić centroid  $c = [c^{(x)}, c^{(y)}]$ .

156 ROZDZIAŁ 10. WYKRYWANIE OBIEKTÓW NA PODSTAWIE KONTURÓW



Rysunek 10.5: Sygnatura okręgu ze źle dobranym środkiem.

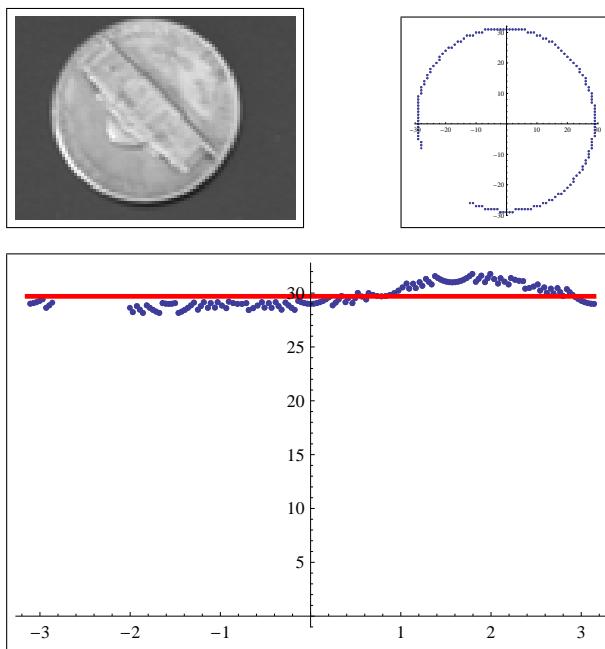


Rysunek 10.6: Ilustracja potrzeby skalowania – dwie elipsy, których jedna jest dwa razy większa od drugiej

3. Wycentrować dane:  $x_i(c) = x'_i - c^{(x)}$ ,  $y_i(c) = y'_i - c^{(y)}$ .
  4. Przekształcić je do współrzędnych biegunowych:  $r_i^2(c) = x_i^2(c) + y_i^2(c)$ ,  $\theta_i(c) = \arctg\left(\frac{y_i(c)}{x_i(c)}\right)$ .
  5. Obliczyć  $r^*(c) = \max_i[r_i(c)]$ .
  6. Przeskalować punkty  $R_i(c) = r_i(c)/r^*(c)$  i stworzyć  $(\theta_i(c), R_i(c))$ ,  $i = 1, 2, \dots, n$ .
- Obliczyć kryterium dopasowania sygnatury do przeskalowanych danych
- $$Q(\alpha, c) = n^{-1} \sum_{i=1}^n |\xi(\theta_i(c) + \alpha) - R_i(c)|$$
- Dla danego  $c$ , znaleźć  $\alpha$ , które minimalizuje  $Q(\alpha, c)$ . (Pamiętamy, że przesunięciu wzdłuż osi  $\theta$  odpowiada obrót obiektu. Zatem, minimalizując  $Q(\alpha, c)$  próbujemy dopasować kąt obrotu obiektu tak, by jak najbardziej pasował do wykrytych punktów brzegowych.)
- Jeśli nie mamy pod ręką zaprogramowanego algorytmu minimalizacji funkcji jednej zmiennej (na przykład, metody złotego podziału odcinka), to minimalizację tę można przeprowadzić następująco:  $Q^*(c) = \min_j [Q(j \cdot \Delta\alpha, c)]$ , gdzie  $\Delta\alpha$  jest wybranym krokiem.
- Jeśli  $Q^*(c)$  jest mniejsze niż wybrana wcześniej dokładność, to STOP – obiekt został dopasowany.
  - W przeciwnym przypadku, mogą zajść dwie sytuacje:
    1. centroid został źle dobrany,
    2. wybrana sygnatura  $\xi(\theta)$  nie pasuje do wykrytych krawędzi.

Co więcej, nie ma ogólnej metody rozróżnienia między nimi. Pozostaje zatem:

1. Zoptymalizować  $Q(\alpha, c)$  względem  $c$ , przyjmując jako punkt startowy, na przykład, średnią z  $(x'_i, y'_j)$ . W każdym kroku minimalizacji względem  $c$  musimy powtórzyć powyższą procedurę minimalizacji względem  $\alpha$ , lub użyć procedury optymalizacji dla dwóch zmiennych.
2. Jeśli opisana powyżej minimalizacja względem  $c$  i  $\alpha$  nie da zadowalającej wartości kryterium dopasowania, to pozostaje wybór kolejnego obiektu i powtórzenie całości obliczeń dla odpowiadającej mu sygnatury.



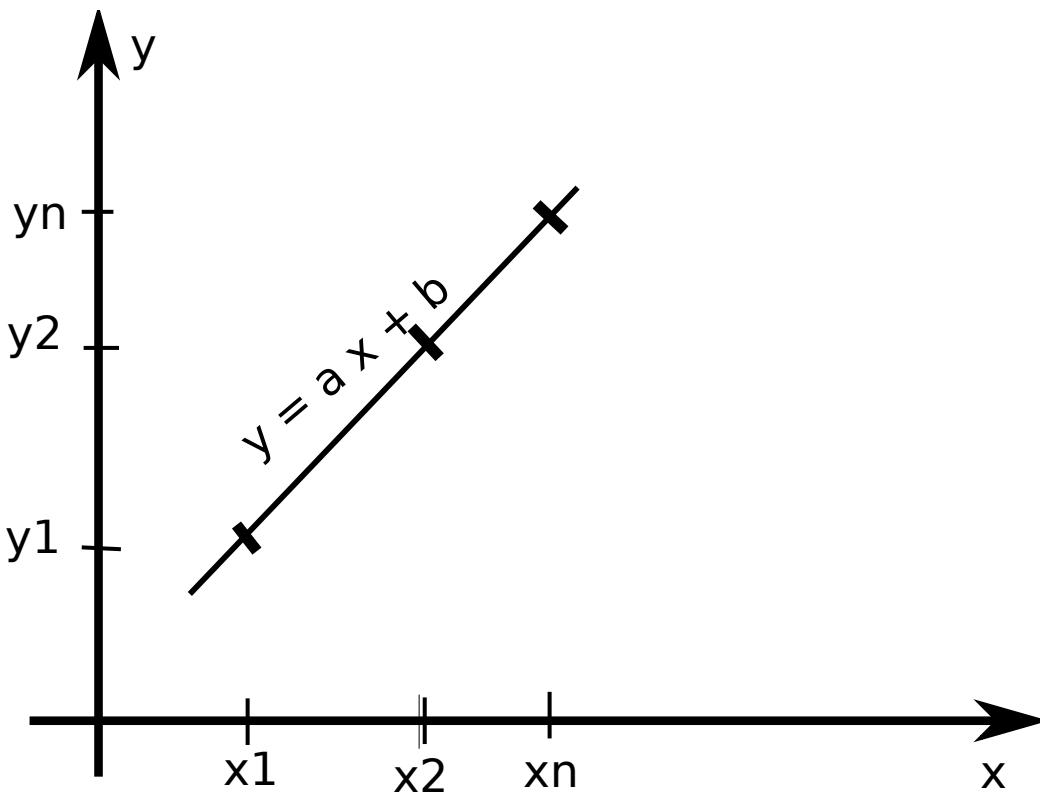
Rysunek 10.7: Moneta, wykryty brzeg i dopasowanie ich profili.  $Q = 0.88$ .

Przykład działania tej procedury pokazano na Rys 10.7. Uzyskaną wartość wskaznika dopasowania, wynoszącą 0.88 uznano za wystarczającą, aby stwierdzić, że mamy do czynienia z okręgiem lub jego znacznym fragmentem.

### 10.3 Transformacja Hougha

W podrozdziale tym przedstawimy podstawowe idee zastosowania transformacji Hough'a do wykrywania i znajdowania opisu linii, okręgów, elpis i innych tworów geometrycznych. Transformacja Hougha została opublikowana w 1962 roku. Od tamtej pory powstała ogromna liczba publikacji, które zawierają uogólnienia tej metody i jej usprawnienia, głównie pod kątem redukcji nakładów obliczeniowych. W swej najogólniejszej postaci transformacja ta pozwala znajdować "dowolne" kształty na obrazie. My ograniczymy się do takich wersji, które do wykrycia kształtu wymagają, by znany był jego opis matematyczny, z dokładnością do nieznanych parametrów.

- Transformata ta nie jest kolejną transformatą całkową.
- Aby zrozumieć transformację Hough'a, wystarczy przypomnieć sobie wiele ze szkoły i aktywnie ją zastosować.



Rysunek 10.8: Obraz z linią

### 10.3.1 Wykrywanie prostych – wersja dydaktyczna

Jako wstęp, rozpatrzmy wykrywanie prostych na obrazie binarnym, na którym – w wyniku detekcji krawędzi – zaznaczono punkty. Na początku, przyjmować będziemy, że na obrazie znajduje się tylko jedna prosta, lecz nie znamy jej położenia.

Drogę do odkrycia transformacji Hough'a na własna rękę przedstawimy w postaci kilkunastu kroków.

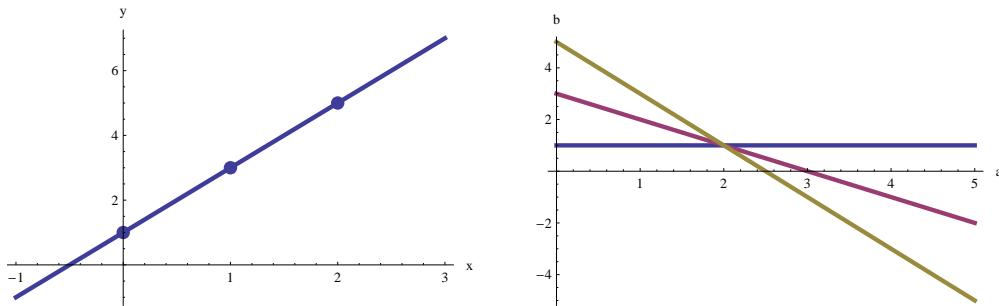
- Co to znaczy, że prosta (por. znany Rys. 10.8)  $y = a \cdot x + b$  przechodzi przez punkty  $(x_n, y_n)$ ,  $n = 1, 2, \dots, N$  ?
- Spodziewana odpowiedź: dla każdego z tych punktów spełnione jest równanie

$$y_n = a \cdot x_n + b, \quad n = 1, 2, \dots, N \quad (10.1)$$

- Spójrzmy na 10.1 w inny sposób

$$b = (-x_n) \cdot a + y_n, \quad n = 1, 2, \dots, N \quad (10.2)$$

## 160 ROZDZIAŁ 10. WYKRYWANIE OBIEKTÓW NA PODSTAWIE KONTURÓW



Rysunek 10.9: Po lewej – prosta  $y = 2 \cdot x + 1$  i punkty  $\{0,1\}, \{1,3\}, \{2,5\}$ . Po prawej proste  $b = 1$ ,  $b = -a + 3$ ,  $b = -2 \cdot a + 5$ , z punktem wspólnym  $(2, 1)$ .

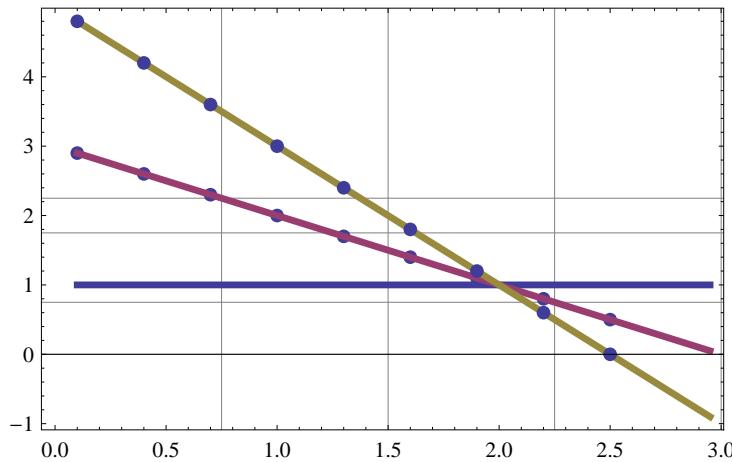
- Jak interpretować 10.2 ? We współrzędnych  $a$ ,  $b$  każde z równań 10.2 może być interpretowane jako linia prosta o nachyleniu  $(-x_n)$ , przesunięta o  $y_n$ . Powstała w ten sposób dualność punkt–linia.
- Co wspólnego mają wszystkie proste 10.2 ? Wszystkie proste 10.2 mają wspólny punkt  $(a_0, b_0)$  – dokładnie jeden – ten, który odpowiada prostej, dla której spełnione jest 10.1.
- Musielimy nieco nagiąć notację, aby zmusić nasz mózg do traktowania pary  $a$ ,  $b$  raz jako konkretny punkt  $(a_0, b_0)$ , a drugi raz – jako współrzędne w przestrzeni parametrów.

Powyzsze rozważania i dualność punkt–prosta zilustrowano na Rys. 10.9. Proste i ładne – wystarczy znaleźć punkty przecięcia  $(a_0, b_0)$  w przestrzeni parametrów i mamy prostą  $y = a_0 x + b_0$ . Pozostaje odpowiedzieć na pytanie: jak znaleźć ten punkt ?

Trzeba zauważyć, że wykres po lewej stronie Rys. 10.9 jest trochę mylący – wykryte krawędzie składają się przecież z wielu punktów. Część z nich może należeć do pewnych prostych, ale nie wiemy które z nich.

Przepis na przybliżone znajdowanie przecięć linii w przestrzeni parametrów  $(a, b)$  przedstawiamy poniżej.

- Podziel przestrzeń  $(a, b)$  prostokątną siatką złożoną z  $m_1 \times m_2$  linii. Oczywiście, w praktyce musimy się odgraniczyć do pewnego prostokąta w tej przestrzeni. Oczka tej siatki nazywane będą komórkami lub celami, kojarzonymi z akumulatorem, opisanym w następnym kroku.
- Przygotuj macierz  $A$  o rozmiarze  $m_1 \times m_2$  wypełnioną zerami. Macierz  $A$  nazywać będziemy akumulatorem, z powodów, które staną się za chwile jasne.

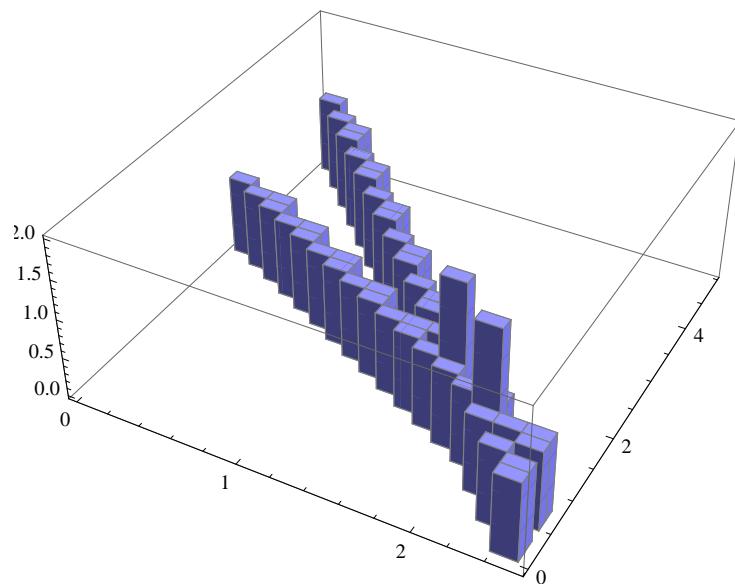


Rysunek 10.10: Proste w przestrzeni parametrów i komórki akumulatora

- Wybierz rozmiar kroku  $\Delta a > 0$  i dla każdego  $n = 1, 2, \dots, N$  i  $i = 1, 2, \dots, m_1$  wykonaj następujące kroki:
  1. Oblicz  $b_n = -(x_n)(i \Delta a) + y_n$
  2. Znajdź komórkę  $(i, j)$ , do której wpada  $(i \Delta a, b_j)$ .
  3. Dodaj +1 do  $A(i, j)$  – tu właśnie akumulujemy wartości, które można nazwać wynikami głosowania na proste, mające swoje współczynniki w komórce o numerach  $(i, j)$ .
- Po zakończeniu zapełniania akumulatora, znajdź komórki  $(i^*, j^*)$  w  $A$  z największą (lub wystarczająco dużą) liczbą elementów.  $(i^* \Delta a, j^* \Delta b)$  dostarcza aproksymacji parametrów prostej.

W celu zilustrowania powyższej procedury na Rys 10.10 pokazano proste w przestrzeni parametrów (te same, które otrzymaliśmy na Rys. 10.9) wraz z komórkami akumulatora. Na Rys 10.10 pokazano stan akumulatora po zakończeniu wypełniania. Wysokość słupków odpowiada stanowi poszczególnych komórek. Widoczne jest maksimum w okolicy punktu  $(2, 1)$ . Drugie maksimum w tej okolicy może wprowadzać w błąd. Zapobiec temu można zmniejszając rozmiar komórek akumulatora. Ładna koncepcja, jednakże

- powyższy algorytm nie jest stosowany w praktyce. Dlaczego ?
- Jak można go ulepszyć ?



Rysunek 10.11: Akumulator w naszym przykładzie. Maksimum w okolicy punktu  $(2, 1)$  jest wyraźnie widoczne

### 10.3.2 Dogodniejsza parametryzacja prostej

Powyższy algorytm nie jest w praktyce stosowany, gdyż równania o postaci  $y = a x + b$  nie pozwalają opisać prostych pionowych, a w praktyce często z takimi właśnie prostymi mamy do czynienia (krawędzie budynków, prostokątne obudowy wyrobów przemysłowych, ścieżki na płytach drukowanych etc.)

Alternatywne rozwiązanie jest dość proste – wystarczy zastosować inną parametryzację prostej, mianowicie:

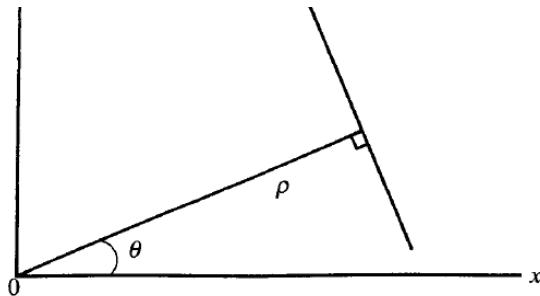
$$\rho = x \cos(\theta) + y \sin(\theta), \quad (10.3)$$

gdzie  $\rho > 0$  jest odległością między prostą a początkiem układu współrzędnych. Odległość ta jest długością odcinka, który jest prostopadły do prostej i łączy ją z punktem  $(0, 0)$ . Natomiast  $\theta \in [0, \pi)$  jest kątem (w radianach) pomiędzy tym odcinkiem a dodatnią półosią osi  $x$ -ów (por. Rys. 10.12). Dodatkową zaletą reprezentacji (10.3), jest to, że w naturalny sposób ograniczony jest obszar zmienności  $\theta$  i  $\rho$ .

Po szczegółowym przedstawieniu poprzedniej wersji algorytmu, możemy ograniczyć się do przedstawienie poniższego szkicu metody dla parametryzacji (10.3).

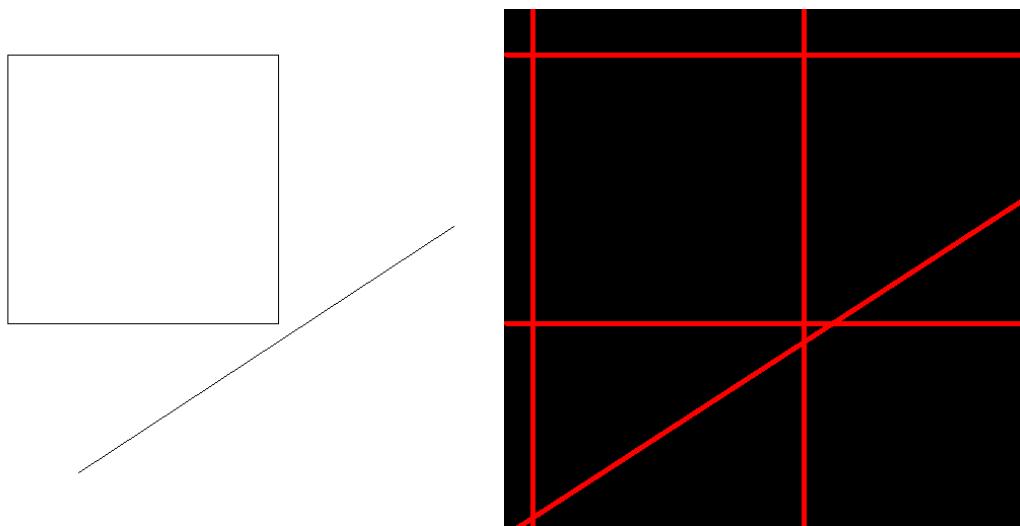
**Szkic algorytmu** wykrywania prostych:

- Wyzeruj akumulator  $A$

Rysunek 10.12: Przestrzeń parametrów:  $(\rho, \theta)$ .

- Dla każdego punktu  $(x_n, y_n)$ ,  $n = 1, 2, \dots, N$  wykonaj:
  1. dla  $\theta$  zmieniającego się od 0 do  $\pi$  wykonaj kroki 2 i 3
  2.  $\rho = x \cos \theta + y \sin \theta$
  3.  $A(\theta, \rho) = A(\theta, \rho) + 1$
- Znajdź  $(\theta, \rho)$ , będące lokalnymi maksimami  $A(\theta, \rho)$

**Przykład** zastosowania transformacji Hougha. Rys. 10.13 (po lewej) został poddany powyższej procedurze.



Rysunek 10.13: Przykładowy rysunek (po lewej) i wykryte linie (po prawej)

W takich wyidealizowanych warunkach wszystkie linie zostają wykryte (Rys. 10.13 (po prawej)), a ich współrzędne pokazano w Tabeli 10.2.

**Pytanie:** Jaki funkcje w przestrzeni parametrów odpowiadają punktom w przestrzeni obrazu, w przypadku gdy stosujemy parametryzację (10.3) ?

$\rho$	$\theta$
310	1.57
45	1.57
295	0.
28	0.
436	0.994

Tabela 10.2: Współrzędne linii wykrytych za pomocą transformacji Hougha

### 10.3.3 Uwagi na temat transformacji Hougha

Oto kilka uwag praktycznych na temat stosowania transformacji Hougha.

- Ze względu na duży nakład obliczeniowy, do wykrywania linii najpierw warto oszacować ich gradienty i rozpatrywać tylko krawędzie o odpowiednio dużej długości wektora gradientu (por. rozdział 10.3.4).
- Odpowiednio do wymaganej dokładności lokalizacji linii, dobierać krok dyskretny w przestrzeni parametrów.
- Jeśli stosować musimy bardzo drobną siatkę w przestrzeni parametrów, to warto dodawać "1" także do otoczenia bieżącej komórki.
- Jak, po zakończeniu wykrywania linii w przestrzeni parametrów, rozstrzygnąć, który zestaw  $(\rho, \theta)$  odpowiada danej prostej w przestrzenionobrazu? W tym celu warto opisywać punkty  $(x_n, y_n)$  numerami linii w przestrzeni parametrów.

Zalety transformacji Hougha można podsumować następująco.

- Ma cechy analizy globalnej obrazu, w odróżnieniu od omawianych dotąd metod, które bazują głównie na informacji z pewnego otoczenia rozpatrywanego piksela.
- Pozwala wykrywać wiele różnych rodzajów obiektów w jednym przebiegu.
- Metoda jest niezmiennicza ze względu na translacje.
- Do pewnego stopnia jest odporna na zakłócenia. Zależnie od intensywności szumu, lokalne masima ulegają rozmyciu.
- Pozwala automatycznie wykryć linie z pewnymi przerwami. Daje to pewien zakres odporności i możliwość poprawienia krawędzi.
- Można rozpoznać zdeformowane, w pewnym stopniu, figury geometryczne.

### 10.3.4 Modyfikacje i uogólnienia

Główna wadą transformacji Hougha jest jej złożoność obliczeniowa, która rośnie wykładniczo wraz z liczbą parametrów modelu, opisującego poszukiwane obiekty geometryczne.

W odniesieniu do poszukiwania prostych, wadę tę można nieco zredukować, ograniczając liczbę przeglądanych wartości  $\theta$ . Pamiętajac o tym, że wektor gradientu prostej jest do nie prostopadły, możemy ograniczyć się do okolic

$$\theta = \arctg \left( \frac{\partial f}{\partial y} / \frac{\partial f}{\partial x} \right). \quad (10.4)$$

W praktyce, przybliżamy pochodne cząstkowe występujące we wzorze (10.4) za pomocą różnic skończonych. W celu zaoszczędzenia nakładów na przegladanie obrazu, warto wyliczać odpowiednie wartości  $\theta$  już na etapie wykrywania krawędzi. Dla wartości  $\theta$  przybliżających (10.4) wyszukujemy odpowiednie  $\rho$ .

Transformację Hougha ogólniono na wiele sposobów. Oto niektóre z nich.

- Wykrywanie okręgów, opisanych jako  $(x - a)^2 + (y - b)^2 = r^2$  – można powtórzyć powtórzyć wcześniejszy proces. Jest tylko jedna różnica – tym razem mamy trzy parametry, a więc akumulator musi być tablicą o trzech indeksach.
- Transformację Hougha można stosować do wykrycia każdego kształtu, który może być opisany równaniem:  $F(x, y, a, b, c, \dots) = 0$ , gdzie  $F$  jest znaną funkcją, natomiast  $a, b, \dots$  parametrami, określającymi położenie i wielkość danego obiektu. Dobrym przykładem jest wykrywanie elips, które stosowane jest do znajdowania twarzy na obrazie.
- Transformację Hougha ogólniono na dowolne kształty ( $F$  nie jest określone).

Jak można sobie wyobrazić wykrywanie kształtu, którego nie potrafimy opisać wzorem? Wystarczy mieć wzorzec brzegu i jeden punkt wewnętrzny. Wówczas, podobnie jak w metodzie centroidu, potrafimy stworzyć tabelę  $(\rho, \theta)$  i dalej postępujemy tak jak poprzednio

*166 ROZDZIAŁ 10. WYKRYWANIE OBIEKTÓW NA PODSTAWIE KONTURÓW*

## Część III

**Poprawianie obrazów:**

**Histogram – wyrównywanie**

**Filtracja**

**Operacje morfologiczne**



Przetwarzając obrazy w warunkach przemysłowych powinniśmy dążyć do tego, aby poprawianie obrazów nie było potrzebne. Warto zadbać o:

- stabilne warunki oświetlenia,
- dobrą jakość obiektywu,
- kamerę z matrycą o niskim poziomie szumów,

po to, by nie tracić cennych ułamków sekundy na poprawianie obrazów.

W pewnych sytuacjach jednak korekcji obrazów uniknąć się nie da. Dla tego w tej części książki omówimy:

1. wyrównywanie histogramu,
2. proste filtry cyfrowe,
3. operacje morfologiczne.

Wyrównywanie histogramu służy, między innymi, do korekcji naświetlenia obrazu. Przypomnijmy, że prostą technikę korekcji gamma omawialiśmy w Rozdziale 3.

Ograniczymy się do bardzo prostych, a przez to relatywnie szybkich, filtrów cyfrowych. Zakładamy, że Czytelnikowi znane są klasyczne filtry liniowe, wykładane w ramach teorii sygnałów. Sporo uwagi poświęcimy filtrom zachowującym krawędzie.

W ostatnim rozdziale tej części przedstawimy wstęp do morfologicznych technik przetwarzania obrazów. Podejście morfologiczne zdobywa coraz więcej zwolenników, a teoria nadal się rozwija. Tutaj, skupimy się na tych operacjach morfologicznych, które służą do poprawiania obrazów binarnych, lecz warto zaznaczyć, że podejście to ma wiele innych zastosowań.



## Rozdział 11

# Wyrównywanie histogramu

Wyrównywanie histogramu służy, jak już wspomnieliśmy, do korekcji naświetlenia obrazu. Rozpatrzmy niepoprawnie naświetlony obraz (prześwietlony lub niedoświetlony) w sposób zmniejszający jego czytelność, ale nie zupełnie czarny bądź biały. Wyrównywanie poziomów (nazywane equalizacją histogramu (od *histogram equalization*)) polega na rozsunięciu odcieni na obrazie tak, aby zajmowały cały odcinek [0,255].

### 11.0.5 Przykład wstępny

Na Rys. 11.1 widzimy fragment gorącej, metalowej powierzchni. Jego histogram jest zbyt skoncentrowany – poziomy szarości nie zajmują całego odcinka [0,255]. Po zastosowaniu wyrównywania obraz staje się bardziej wyraźny. Na Rys. 11.2 widać teraz histogram zajmujący cały odcinek [0,255].

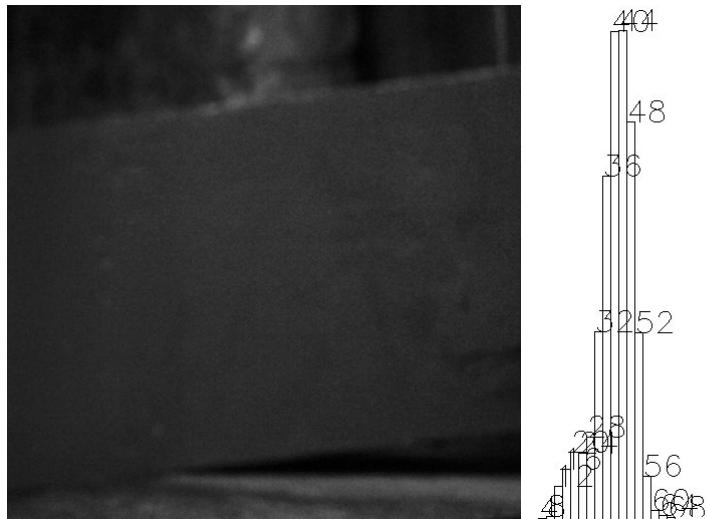
Przymijmy, że poziomy szarości niepoprawnie naświetlonego obrazu znajdują się w zakresie  $[a,b]$ , a my chcemy je rozciągnąć do  $[0,255]$ . Oznaczmy przez  $S$  poziom szarości obrazu źródłowego, a przez  $I$  poziom szarości obrazu docelowego. Wówczas przekształcenie

$$I = \frac{255}{b-a} \cdot (S - a) \quad (11.1)$$

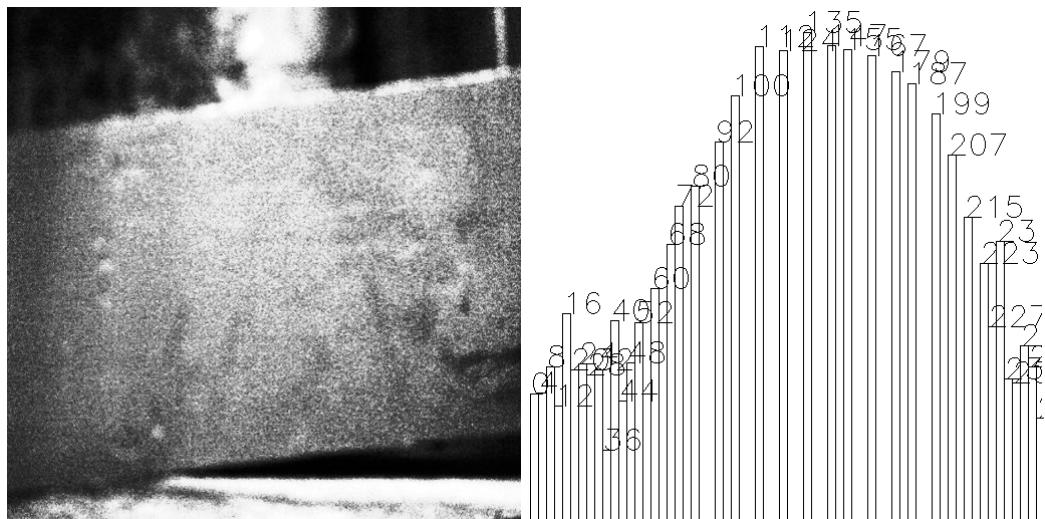
równoważne jest równoczesnej zmianie jasności i kontrastu obrazu.

### 11.0.6 Podstawy teoretyczne

Teoretyczne uzasadnienie techniki wyrównywania histogramu opiera się na prostej, lecz ważnej, własności dystrybuant rozkładów prawdopodobieństw. Własność tę wyprowadzamy poniżej.



Rysunek 11.1: Fragment niepoprawnie naświetlonej metalowej płyty i jego histogram



Rysunek 11.2: Fragment metalowej płyty i jego histogram

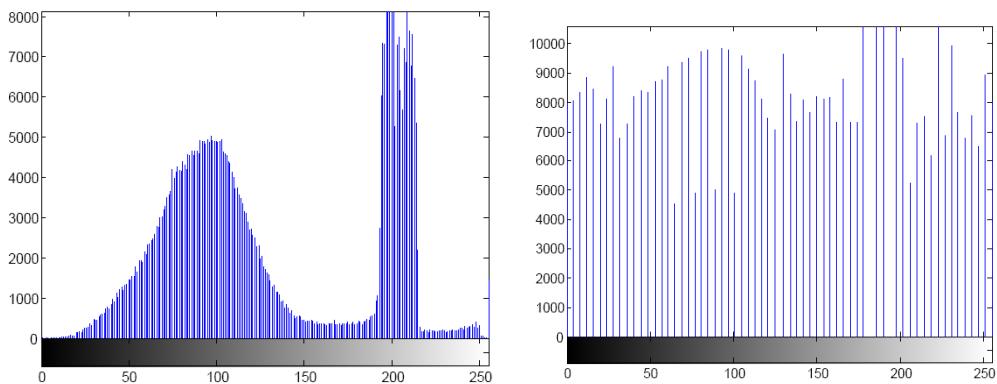


Rysunek 11.3: Ciemny samochód na ciemnym tle (po lewej) i wynik wyrównywania histogramu (po prawej)

- Niech  $X$  będzie zmienną losową o rozkładzie określonym dystrybuantą  $F_X(x)$ . Przypomnijmy, że  $F_X(x) = P\{\omega : X(\omega) \leq x\}$  (wersja prawostronne ciągła).
- Twierdzenie: Zmienna losowa  $Y$  zdefiniowana wzorem  $Y = F_X(X)$  jest zmienną losową o rozkładzie równomiernym w  $[0, 1]$ .
- Inaczej mówiąc: przekształcając zmienną losową  $X$  przez jej własną dystrybuantę  $F_X$  uzyskujemy zmienną losową o rozkładzie równomiernym.
- Nie nakładamy ograniczeń na  $X$  i jej dystrybuantę. Tak więc, twierdzenie to możemy zastosować do poziomów szarości dowolnego obrazu, jeśli jego poziomy szarości mogą być traktowane jako realizacje pewnej (zwykle nieznanej) zmiennej losowej.
- Aby uprościć dowód, zakładamy dodatkowo, że  $F_X$  jest ścisłe rosnącą. Istnieje wówczas funkcja odwrotna do dystrybuanty  $F_X$ , którą oznaczymy przez  $F_X^{-1}(\cdot)$ .
- Zauważmy, że zmienna losowa  $Y = F_X(X)$  ma następującą dystrybuantę:

$$P\{Y \leq y\} = P\{F_X(X) \leq y\} = P\{X \leq F_X^{-1}(y)\}.$$

Ostatnia równość wynika z faktu, że  $F_X$  jest ścisłe rosnącą.



Rysunek 11.4: Histogramy zdjęć z Rys. 11.3

- Teraz wystarczy ponownie odczytać wyrażenie  $P\{X \leq F_X^{-1}(y)\}$ . Jest to przecież wartość dystrybuanty  $F_X$  wzięta w punkcie  $F_X^{-1}(y)$ . Zatem

$$P\{Y \leq y\} = P\{X \leq F_X^{-1}(y)\} = F_X(F_X^{-1}(y)) = y$$

- To znaczy, że  $Y = F_X(X)$  ma dystrybuantę  $F_Y(y) = y$  dla  $0 \leq y \leq 1$  i zero poza tym odcinkiem. Jest to właśnie dystrybuanta zmiennej losowej o rozkładzie równomiernym na  $[0, 1]$ .

### 11.0.7 Wersja empiryczna i przykłady

Mamy ładne twierdzenie, ale nie znamy  $F_X$  dla konkretnego obrazu. W praktyce obliczany jest histogram poziomów szarości. Oznaczmy empiryczne prawdopodobieństwa każdego poziomu szarości z  $[0, 255]$  przez  $p_k$ ,  $k = 0, 1, 2, \dots, 255$ .

- Utwórzmy emipryczny histogram:

$$F_k = \sum_{j=0}^k p_j, \quad k = 0, 1, 2, \dots, 255 \quad (11.2)$$

i użyjmy go zamiast  $F_X$ .

- Jak to zrobić w praktyce? W tablicy LUT piksele o poziomie szarosci  $k$  przekształcane są do  $F_k$ , zdefiniowanego przez 11.2. Następnie obraz zostaje przeskalowany do 0 – 255.

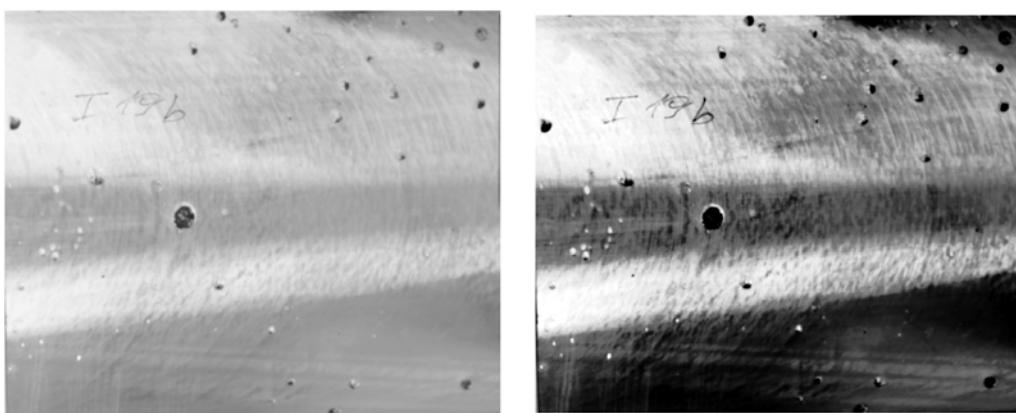
#### Przykład 1.

Po lewej stronie Rys. 11.3 pokazano zdjęcie ciemnego samochodu na ciemnym

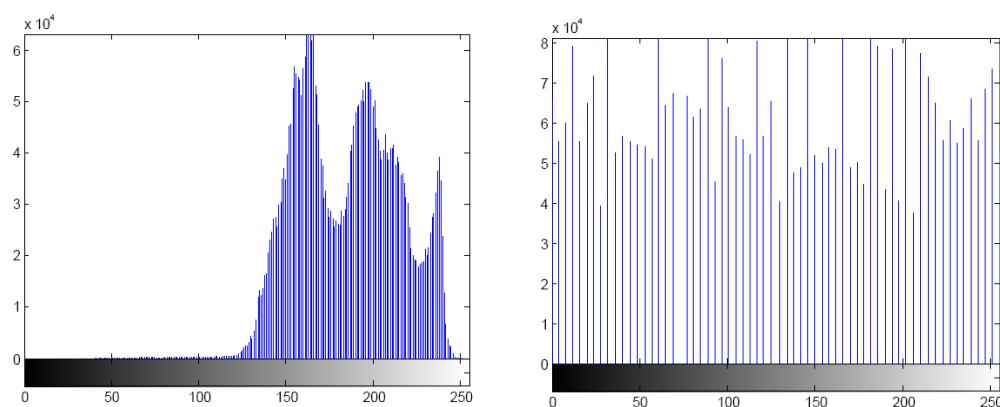
tle, a na Rys. 11.4 histogram tego obrazu. Na tychże rysunkach, po prawych stronach, pokazano zdjęcie tegoż samochodu po wyrównaniu histogramu oraz odpowiadający mu histogram. Efekt wyrównywania jest widoczny, chociaż nie w pełni zadowalający.

### Przykład 2.

Na Rys. 11.5 pokazano znany nam już fragment slabu przed i po wyrównaniu histogramu. Odpowiadające tym zdjęciom histogramy przedstawiono na Rys. 11.6. Prawie niewidoczny w pierwotnej wersji napis I-196 stał się wyraźniejszy. Pokazywalismy wcześniej to zdjęcie jako przykład niepoprawnego naświetlenia, gdyż widoczne są smugi promieni słonecznych. Nic dziwnego, że po wyrównywaniu histogramu smugi te stały się jeszcze bardziej widoczne.



Rysunek 11.5: Slab przed i po wyrównaniu histogramu



Rysunek 11.6: Histogramy obrazów z Rys. 11.5



## Rozdział 12

# Proste metody filtracji obrazów

Podstawowym zadaniem filtracji jest usuwanie szumów z obrazu. Może ona dodatkowo uwypuklić to, co nas na nim interesuje. Ogólne metody projektowania filtrów cyfrowych mogą być stosowane także do obrazów, ale fakt, że są one dwu wymiarowe powoduje, iż filtry mają swoją specyfikę. Odsyłamy Czytelnika do [54], gdzie przedstawiono podstawy projektowania liniowych filtrów cyfrowych z uwzględnieniem metod częstotliwościowych. W rozdziale tym przedstawimy bardzo proste, łatwo implementowalne filtry cyfrowe, z naciskiem na nieliniowe filtry, do pewnego stopnia, zachowujące krawędzie.

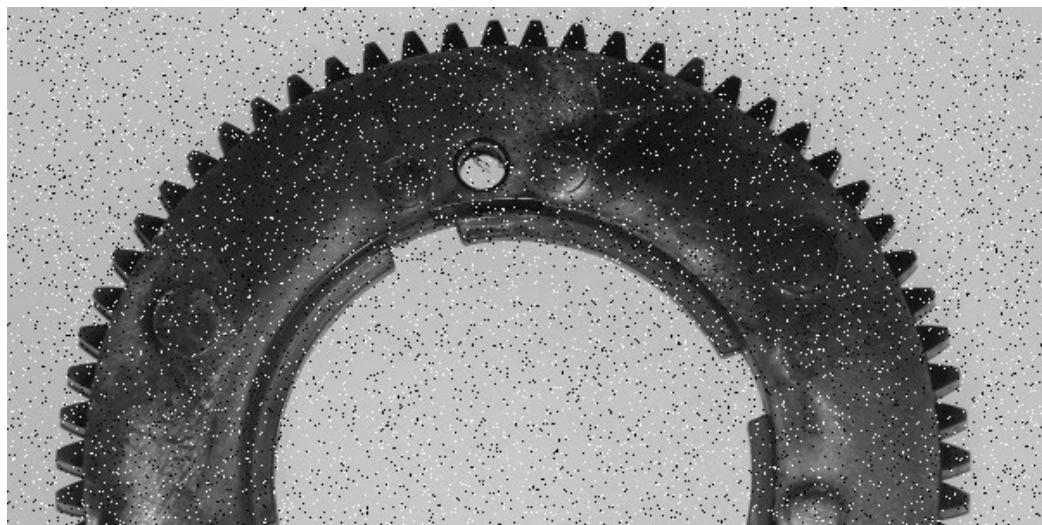
Postawowym zagadnieniem przy wyborze filtra jest informacja o zakłóceńach. Statystyczne związki między szumami powstającymi w poszczególnych sensorach na ogół istnieją, ale czasem – pod presją konieczności stosowania bardzo prostych filtrów – sa one pomijane.

Zwykle przyjmuje się addytywne modele szumów. Próbuje się też modelować je mnożnikowo, jednak jest to trudniejsze, także z teoretycznego punktu widzenia.

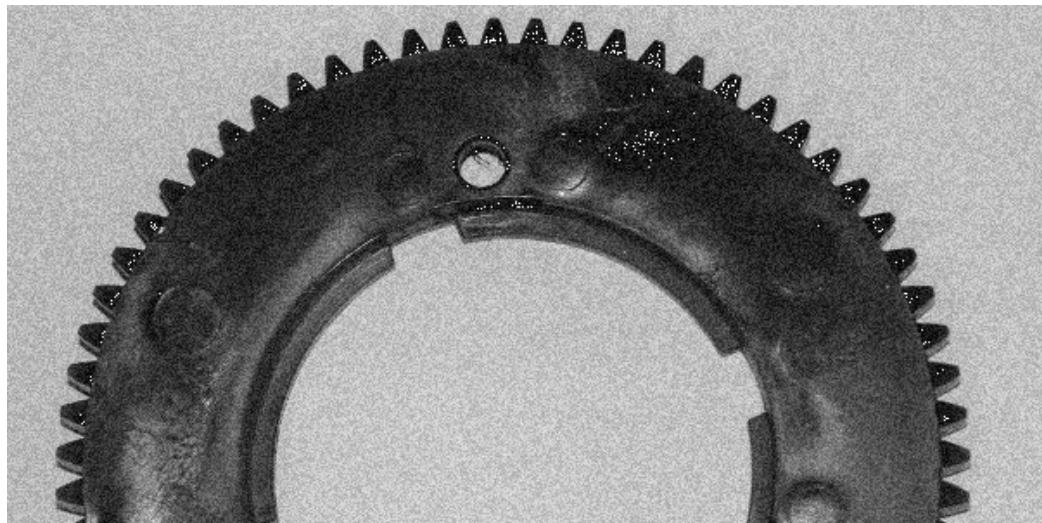
Wyróżnia się różne rodzaje szumów, przy czym klasyfikacja ta nie jest metodologicznie jednorodna:

- gaussowskie (patrz Rys. 12.3)
- o rozkładzie równomiernym (patrz Rys. 12.2)
- sól z pieprzem – dany piksel jest zamieniany na czarny lub biały, występuje np. w złych warunkach oświetleniowych, przy dużym wzmacnieniu (patrz Rys. 12.1).

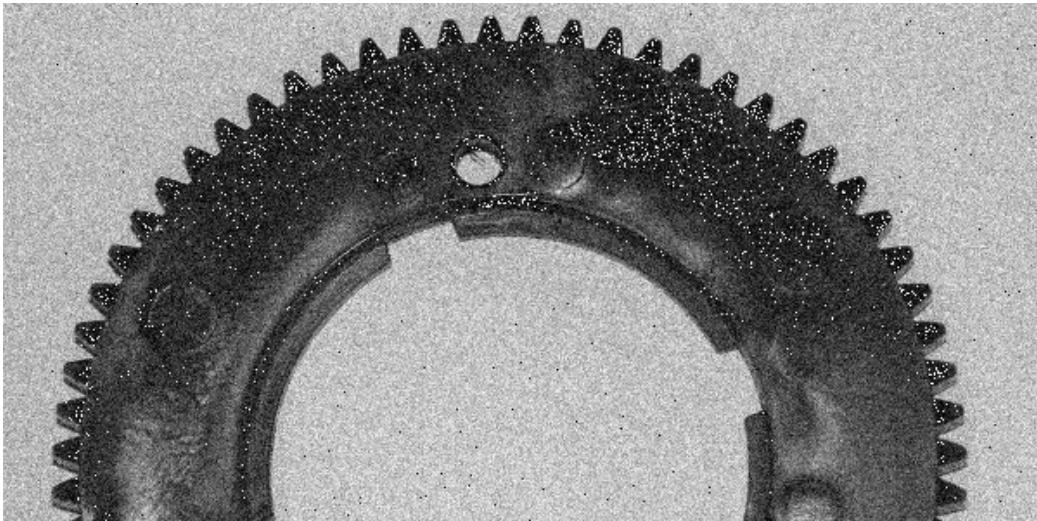
Na Rys. 12.1, 12.2, 12.3 pokazano ten sam obraz, który został poddany wyżej wymienionym rodzajom zakłóceń. Zakłócenia te były generowane algorytmicznie i dodane do obrazu. Informacje o szumie można szacować na podstawie



Rysunek 12.1: Szum typu sól z pieprzem



Rysunek 12.2: Szum o rozkładzie równomiernym



Rysunek 12.3: Szum gaussowski

wybranego fragmentu, na przykład, równomiernego tła.

Szумy występują także w obrazach kolorowych. Jeśli obraz reprezentowany jest w przestrzeni barw RGB, to opisane poniżej filtry stosować można osobno do każdego koloru podstawowego.

Dodatkowym źródłem szumu jest rozmywanie związanego z nieostrością. Kolejnym źródłem zniekształceń jest obiektyw. Najprostszym modelem takich zniekształceń jest splot obrazu z odpowiedzia impulsową układu optycznego:

$$J(u, v) = I(u, v) * H(u, v). \quad (12.1)$$

Można próbować estymować  $H$ , badając empirycznie jak zniekształcone są punktowe źródła. Jeśli mamy możliwość przeprowadzenia prostego eksperymentu, to możemy umieścić proste figury geometryczne w różnych częściach obrazu i zarejestrować je kamerą, a następnie ocenić  $H(u, v)$ . Jeśli nie ma takiej możliwości, to szukamy znanego obiektu na obrazie i na tej podstawie oceniamy funkcję  $H(u, v)$ .

Warto wspomnieć także o zakłóceniach okresowych. Można je wykryć w postaci punktów o dużej wartości amplitudy, po poddaniu obrazu transformacji Fouriera.

## 12.1 Proste filtry

Proste filtry operują na otoczeniu punktu  $(i, j)$  i wynik umieszcza w odpowiadającym mu punkcie obrazu wyjściowego. Zgodnie z tradycją wielu

dziedzin, w których stosuje się filtrację cyfrową, poziomy szarości obrazu wejściowego w punkcie  $(i, j)$  oznaczać będziemy przez  $x_{ij}$ . Odpowiednie poziomy szarości obrazu wyjściowego (wyniku filtracji) oznaczać będziemy przez  $y_{ij}$ .

### 12.1.1 Filtr ruchomej średniej

#### Filtr średniej arytmetycznej

Najprostszy filtr dolnoprzepustowy to filtr tzw. ruchomej średniej:

$$y_{ij} = \frac{1}{(2m+1)^2} \sum_{k=-m}^m \sum_{l=-m}^m x_{(i-k),(j-l)}. \quad (12.2)$$

Okno  $(2m+1) \times (2m+1)$  przesuwane jest po tablicy obrazu i w każdym jego położeniu obliczna jest średnia arytmetyczna. Średnia ta wstawiana jest do obrazu wynikowego w punkcie, który odpowiada centrum okna.

#### Efekt brzegowy

Jak zwykle, także i tutaj występuje tzw. efekt brzegowy. Polega on na tym, że w punktach  $(i, j)$  bliskich granicy tablicy nie mamy danych do uśredniania. Proste metody radzenia sobie z tą trudnością to:

1. pozostawienie obszarów brzegowych bez filtracji i przepisanie wartości z obrazu źródłowego,
2. zmniejszenie  $m$  do wartości 1 w okolicy brzegu – wówczas nieodfiltrowany zostanie pas szerokości jednego piksela,
3. jeśli na brzegach występuje jednolite tło – można sięgnąć do wartości pikseli na przeciwnym brzegu.

Odradzamy zastępowanie zerami pasa brzegowego, gdyż daje to czarne fragmenty, które potem mogą być potraktowane jako krawędzie. Powyższe rady obowiązują w odniesieniu do wszystkich omawianych w tym rozdziale filtrów i nie będziemy ich powtarzać.

#### Filtr ważonej ruchomej średniej

Uogólnieniem filtru (12.2) jest filtr z wagami:

$$y_{ij} = \sum_{k=-m}^m \sum_{l=-m}^m x_{(i-k),(j-l)} w_{kl}, \quad (12.3)$$



Rysunek 12.4: Wynik działania filtru gaussowskiego o oknie 11x11

gdzie  $w_{kl} \geq 0$  są wagami takimi, że

$$\sum \sum w_{kl} = 1.$$

Dobierając wagi możemy nadawać większe znaczenie poziomom szarości tych pikseli, które leżą bliżej piksela aktualnie centralnego. Ważnym przykładem takiego filtru jest filtr gaussowski – przypadek ruchomej średniej, który omawialiśmy już w podrozdziale 9.1.5.

Jak pamiętamy,  $w_{kl}$  pochodzą z dyskretyzacji gęstości rozkładu normalnego. Zatem, filtr ten przywiązuje dużo większą wagę do centrum, a pozostałe współczynniki szybko maleją.

Na Rys. 12.4 pokazano efekt zastosowania filtru gaussowskiego z oknem  $11 \times 11$  do obrazu z Rys. 12.3. Jak widać, filtr znacznie stłumił zakłócenia. Ubocznym efektem jego zastosowania jest nadmierne wygładzenie krawędzi. Jak pamiętamy, efekt ten nie zawsze jest szkodliwy, gdyż stosuje się go do wykrywania krawędzi metodą Log. Jednakże, w sytuacjach, gdy zmiana poziomu szarości na krawędzi jest bardzo mała, wówczas zastosowanie takiego filtru może utrudnić jej wykrycie.

### 12.1.2 Filtry bazujące na innych średnich

Oprócz średniej arytmetycznej, która – ze względu na liniowość - ma charakter podstawowy, znanych jest wiele innych technik uśredniania. Każda z nich może podstawą do stworzenia filtru cyfrowego do przetwarzania obrazów. Podobnie jak w poprzednim podrozdziale, przesuwamy okno, pobieramy z niego

poziomy szarości i uśredniamy. Nie jest to zabieg czysto formalny, gdyż różne rodzaje średnich mają nieco różne własności, które dziedziczą po nich filtry cyfrowe.

W podrozdziale tym omówimy kilka takich filtrów, przypominając wcześniej definicje odpowiednich średnich. Definiując odpowiednie średnie używać będziemy oznaczeń  $a_1, a_2, \dots$  dla ciągu uśrednianych liczb.

### Filtr średniej geometrycznej

Załóżmy, że  $a_n > 0$ ,  $n = 1, \dots, N$ . Wówczas średnia geometryczna dana jest wzorem:

$$A_g = \left( \prod_{n=1}^N a_n \right)^{\frac{1}{N}}. \quad (12.4)$$

Łatwo zauważyc, że logarytm średniej geometrycznej jest średnią arytmetyczną logarytmów uśrednianych liczb:

$$\ln A_g = \frac{1}{N} \sum_{n=1}^N \ln(a_n).$$

Można pokazać, że  $A_g \leq A_{sr}$ , gdzie  $A_{sr} = \frac{1}{N} \sum_{n=1}^N a_n$ , przy czym równość zachodzi wtedy i tylko wtedy, gdy  $a_1 = a_2 = \dots = a_n$ .

Filtr średniej geometrycznej ma postać:

$$y_{i,j} = \left( \prod_{k,l=-M}^M x_{i+k,j+l} \right)^{\frac{1}{(2M+1)^2}}, \quad (12.5)$$

lecz pamiętać trzeba, że można go stosować tylko wtedy, gdy  $x_{ij} > 0$ .

**Pytanie.** Porównać filtr średniej arytmetycznej, z filtrem geometrycznej ruchomej średniej. Jakie konsekwencje dla wyniku filtracji ma nierówność  $A_g \leq A_{sr}$ ?

### Filtr średniej harmonicznej

Załóżmy, że  $a_n > 0$ ,  $n = 1, \dots, N$ . Średnia harmoniczna tego ciągu zdefiniowana jest następująco:

$$A_h = \frac{1}{\frac{1}{N} \sum_{n=1}^N \frac{1}{a_n}}. \quad (12.6)$$

Można wykazać, że zachodzą następujące nierówności:

$$A_h \leq A_g \leq A_{sr}. \quad (12.7)$$

Odpowiedni filtr cyfrowy ma postać:

$$y_{i,j} = \frac{1}{\frac{1}{(2M+1)^2} \sum_{k,l=-M}^M \frac{1}{x_{i+k,j+l}}} \quad (12.8)$$

W zastosowaniu do obrazów, na których występują zakłócenia typu *sól i pieprz* filtr ten skutecznie odfiltrowuje *sól*, czyli białe wtrącenia. Nie jest skuteczny dla wtrąceń czarnych, czyli *pieprzu*. Jest to konsekwencja nierówności (12.7).

### Filtr średniej kontr-harmonicznej

Podobnie jak poprzednio, zakładamy, że  $a_n > 0$ ,  $n = 1, \dots, N$ . Średnia kontr-harmoniczna zdefiniowana jest wzorem filtru średniej kontr-harmonicznej:

$$A_c = \frac{\sum_{n=1}^N (a_n)^{Q+1}}{\sum_{n=1}^N (a_n)^Q}, \quad (12.9)$$

przy czym dopuszcza się zarówno dodatnie jak i ujemne wykładniki  $Q$ . Obecność parametru  $Q$  powoduje, że mamy do czynienia z rodziną filtrów, które w odniesieniu do obrazów mają postać:

$$y_{i,j} = \frac{\sum_{k,l=-M}^M (x_{i+k,j+l})^{Q+1}}{\sum_{k,l=-M}^M (x_{i+k,j+l})^Q}. \quad (12.10)$$

Wybierając  $Q > 0$  otrzymujemy filtr, który redukuje przebarwienia ciemne – *pieprz*, natomiast dla  $Q < 0$  redukuje *sól*.

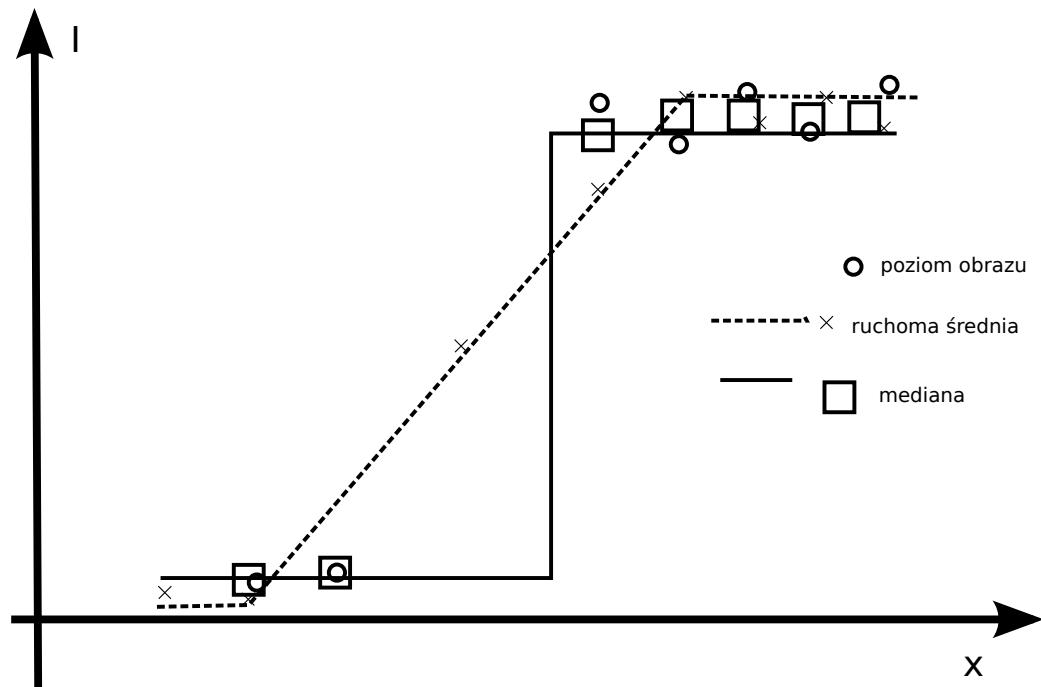
### 12.1.3 Filtr ruchomej mediany

Wszystkie dotąd omawiane filtry bazowały na różnie definiowanych średnich. Innym kierunkiem poszukiwań skutecznych filtrów są metody oparte na statystykach porządkowych. W wersji empirycznej, statystyki porządkowe korzystają z uporządkowanego ciągu danych  $a_n$ ,  $n = 1, \dots, N$ . Taki uporządkowany ciąg danych oznaczać będziemy następująco:

$$a_{(1)} \leq a_{(2)} \dots \leq a_{(N)}, \quad (12.11)$$

gdzie  $a_{(i)}$  oznacza, że – po uporządkowaniu – liczba ta znajduje się na  $i$ -tej pozycji, niezależnie od tego, na której pozycji była w pierwotnym ciągu danych. W uporządkowanym ciągu danych powtarzające się liczby zostają zachowane.

Gdy  $N$  jest nieparzyste, to medianą (empiryczną) ciągu (12.11) nazywamy liczbę o numerze  $(N+1)/2$ , pochodzącą z tego ciągu. Dla  $N = 2$  mediana równa jest średniej arytmetycznej. Dla pozostałych parzystych  $N$  w podręcznikach ze statystyki matematycznej definiuje się jako  $(a_{(N/2)} + a_{(1+N/2)})/2$ . Jeśli poziomy szarości obrazu reprezentowane są jako liczby 8-bitowe, to albo powyższy wynik trzeba zaokrąglić, albo jako medianę przyjąć jedną z



Rysunek 12.5: Porównanie zachowania mediany i średniej na krawędzi

liczb  $a_{(N/2)}$  lub  $a_{(1+N/2)}$ . Medianę empiryczną oznaczać będziemy symbolem  $Med[]$ , a w nawiasach kwadratowych wskazywać będziemy ciąg, którego operacja obliczanie mediany dotyczy.

W zastosowaniu do przetwarzania obrazów filtr medianowy!medianowy ma postać:

$$y_{i,j} = Med [x_{i+k,j+l} \mid k \leq M, |l| \leq M] \quad (12.12)$$

Filtr medianowy jest skuteczny i jednocześnie wprowadza znacznie mniejsze zniekształcenia krawędzi niż filtry ruchomej średniej. Powodem niezłągo zachowywania krawędzi jest fakt, że w centrum każdego okna tego filtru wstawiamy jedną z wartości poziomów szarości pikseli należących do tego okna. W rezultacie, skok prawie nie zostaje rozmazany. Wyjaśnienie tej własności filtrów medianowych pokazano na Rys. 12.5, natomiast efekt zastosowania filtru do znanego nam detalu przedstawiono na Rys. 12.6.

Wadą filtrów medianowych jest nakład obliczeniowy. Wartości poziomów szarości trzeba posortować i to wielokrotnie. Czasы filtracji dla obrazów o dużej rozdzielczości i dużych rozmiarach okien są bardzo długie.

Wiele prac poświęconych było zmniejszaniu nakładów obliczeniowych potrzebnych do medianowej filtracji. Jeden z nurtów bazuje na obserwacji, że po przesunięciu centrum okna o jeden piksel nie musimy wszystkich wartości



Rysunek 12.6: Wynik działania filtru medianowego o oknie 9x9

sortować od początku. Drugi nurt – to próby przybliżonego liczenia mediany. Przykładowo, zamiast liczyć dokładnie medianę w oknie  $9 \times 9$ , liczymy mediany w oknach o mniejszych rozmiarach (na przykład,  $3 \times 3$  lub nawet  $3 \times 1$ ) i wyniki tych obliczeń traktujemy jako dane do następnego obliczenia mediany.

Medianę liczy się bardzo łatwo wówczas, gdy obraz jest binarny (czarno-biały). Wystarczy bowiem zliczyć liczbę czarnych pikseli. Jeśli jest ich więcej niż połowa, to wynik filtracji zaznaczamy na czarno. Innymi słowy, jako wynik filtracji w danym oknie bierzemy ten kolor, który jest częściej reprezentowany. Filtr medianowy w zastosowaniu do obrazów binarnych pozwala skutecznie usunąć drobne pozostałości progowania lub wykrywania krawędzi, które nie są obiektami.

## 12.2 Filtry z wertykalnymi wagami zachowujące krawędzie

- Filtr ruchomej mediany zachowuje krawędzie lepiej niż filtr ruchomej średniej. Wciąż nie jest jednak zadowalający.
- Opracowano szeroką klasę filtrów bazujących na dyskretyzacji równań różniczkowych o pochodnych cząstkowych. Opierają się na traktowaniu poziomów szarości obrazu jako funkcji, która pełni rolę warunków początkowych dla równań dyfuzji. Równania takie zachowują się jak

filtr podobny do filtru gaussowskiego. Następnym – bardzo ważnym – krokiem badań była modyfikacja współczynników tak, aby dyfuzja zatrzymywała się automatycznie w miejscach, gdzie wystepują krawędzie, a po ich przekroczeniu dyfuzyjne rozmycie jest kontynuowane.

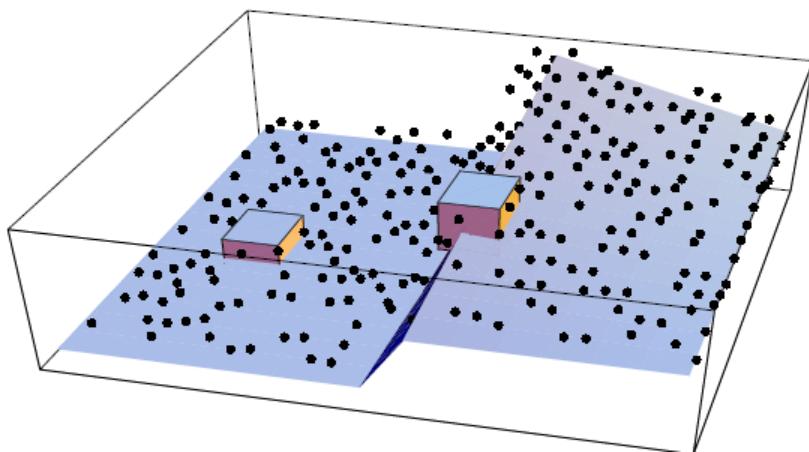
Oprócz wspomnianych wyżej filtrów opartych na hamowaniu dyfuzji pojawiło się w literaturze kilka niezależnych wątków badań nad nieliniowymi filtrami zachowującymi krawędzie. Jedne występują pod nazwą filtrów bilateralnych, inne pod akronimem SUSAN [72] oraz pod hasłem wertykalnych wag. Opiszemy ten ostatni wątek, który zapoczątkowany został w raporcie autorsta E. Rafajłowicza (1996) i przedstawiony na seminarium na Uniwersytecie Wiadrina we Fankfurcie nad Odrą w lipcu 1996 roku, a w poszerzonej wersji w pracy [63]. W pracy [59] pokazano, że podejście SUSAN można traktować jako szczególny przypadek wertykalnego ważenia.

- Inspiracją idei wertykalnego ważenia były nieparametryczne metody statystyki matematycznej.
- Ważne przypadki szczególne mogą jednak być objaśnione bez korzystania z rozbudowanego aparatu matematycznego.
- W klasycznym podejściu do nieparametrycznej estymacji regresji używa się wag działających w przestrzeni zmiennych niezależnych. Podobnie, omawiane wcześniej filtry ważonej ruchomej średniej brały pod uwagę wartości z pikseli sąsiednich i uśredniały je, stosując z góry ustalone wartości wag. W podejściu wertykalnego ważenia bierzemy pod uwagę także zaobserwowane wartości, a więc nie tylko horyzontalnych, ale także wertykalnych sąsiadów. Na tej podstawie modyfikowane są wagi (por. [63]).

### 12.2.1 Idea wertykalnie uciętego pudełka

Rozpoczniemy od najprostszej wersji wertykalnego ważenia, a mianowicie od wertykalnego odcienia obserwacji. W połączeniu z horyzontalnie poruszającym się oknem, dostajemy w efekcie pudełko, które porusza się po powierzchni poziomów szarości obrazu. Przykładowe dwa położenia takiego pudełka zilustrowano na Rys. 12.17.

- Pudełko z obciętymi obserwacjami, zwane dalej V-box, jest trójwymiarowym obiektem, mającym wysokość  $2H > 0$ .
- Pudełko to ma centrum nad aktualnie rozpatrywanym oknem obrazu, na wysokości poziomu szarosci w tym punkcie.



Rysunek 12.7: Pudełko nadążające za powierzchnią obrazu

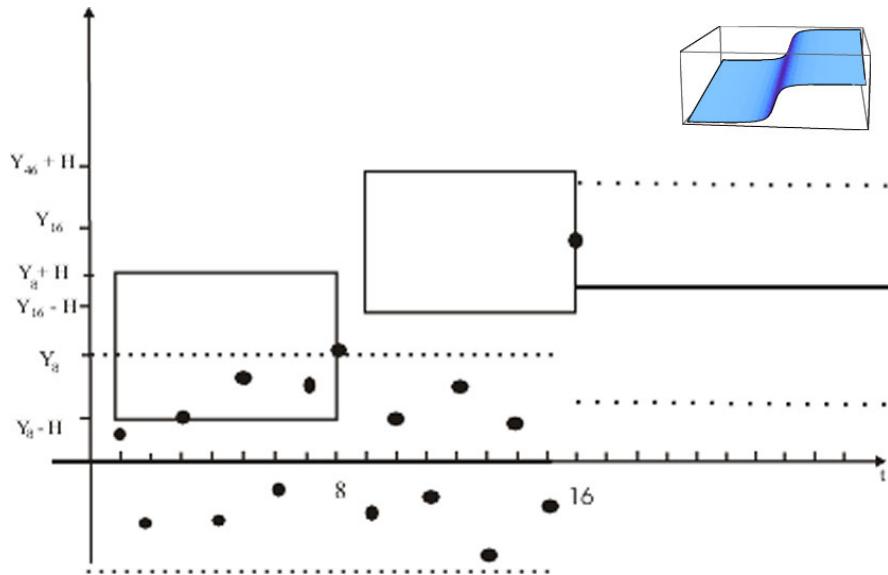
- V-box porusza się równolegle do obrazu i ”chwyta” obserwacje.
- Dopuszczamy uśrednianie poziomów szarości jedynie tych sąsiadów, których poziomy szarości wpadły do pudełka.
- Gdy wysokość pudełka  $H > 0$  jest prawidłowo wybrana i nie ma skoku jasności obrazu, to pudełko zawiera większość obserwacji, które – po uśrednieniu – dają odfiltrowany obraz.
- Jeśli wystąpi skok, to V-box zawiera tylko obserwacje przed lub po skoku. Uśrednianie bierze pod uwagę wyłącznie poziomy przed lub po skoku, co zapobiega rozmyciu krawędzi.

Powyzsze sytuacje naszkicowano na Rys. 12.8.

### 12.2.2 Ogólniejsza klasa filtrów z wertykalnymi wagami

Dość ogólną klasą nieliniowych filtrów może być opisana następująco:

$$y_{ij} = \frac{\sum_{k,l=-M}^M x_{i+k,j+l} \cdot w^2(x_{i+k,j+l})}{\sum_{k,l=-M}^M w^2(x_{i+k,j+l})}, \quad (12.13)$$



Rysunek 12.8: Przekrój przez poziomy szarości i obserwacje, które zawiera V-box w dwóch przypadkach: gdy V-box znajduje się w obszarze o małych zmianach jasności (po lewej) i w punkcie skoku (po prawej)

gdzie  $x_{ij}$  i  $y_{ij}$  to poziomy szarości obrazu wejściowego i wyjściowego, odpowiednio. Natomiast  $w(\cdot)$  to funkcja wagowa, która pozwala kształtować własności filtru. We wzorze (12.13) występuje ona w kwadracie, aby podkreślić, że powinna być ona nieujemna. W implementacji należy oszczędzać operacje arytmetyczne i wystarczy zapewnić nieujemność wag.

Poniżej pokażemy jak uzyskać różne filtry poprzez właściwy dobór  $w(\cdot)$ .

### Filtr ruchomej średniej

Omawiany wcześniej filtr ruchomej średniej otrzymamy ustalając  $w(\cdot) \equiv 1$  we wzorze (12.13)

$$y_{ij} = \frac{1}{(2M+1)^2} \sum_{k,l=-M}^M x_{i+k,j+l}. \quad (12.14)$$

### Filtr ruchomej średniej harmonicznej

Wstawiając  $w(y) = 1/\sqrt{y}$  do (12.13) otrzymujemy filtr

$$y_{ij} = (2M+1)^2 \left/ \sum_{k,l=-M}^M x_{i+k,j+l}^{-1} \right., \quad (12.15)$$

który poznaliśmy już wcześniej.

### Filtr średniej w uciętym pudełku

Opisany wcześniej algorytmicznie filtr średniej z danych w uciętym wertykalnie pudełku możemy otrzymać jako szczególny przypadek (12.13) następujaco:

$$y_{ij} = \frac{\sum_{k,l=-M}^M x_{i+k,j+l} \cdot U_H(x_{i+k,j+l} - x_{ij})}{\sum_{k,l=-M}^M U_H(x_{i+k,j+l} - x_{ij})}, \quad (12.16)$$

gdzie  $U_H(t) = 1$ , jeśli  $|t| < H$  i  $U_H(t) = 0$ , w przeciwnym wypadku. Tuż V-box jest centrowany na obserwacji  $y_{ij}$ . Filtr ten jest uogólnieniem na przypadek 2D filtru znanego zaproponowanego przez Lee w 1983. Filtr Lee znany jest w literaturze jako  $\sigma$ -filtr i – jak widać – można go traktować jako specjalny przypadek wertykalnego ważenia.

W 12.16 filtracja odbywa się przez obliczenie średniej obserwacji wpadających do pudełka:

$$[i-M, i+M] \times [j-M, j+M] \times [x_{ij} - H, x_{ij} + H], \quad (12.17)$$

które jest ustalone w środku nieprzetworzonych obserwacji  $x_{ij}$ .

### Filtr średniej w pudełku centrowanym na medianie

Pudełko w filtrze (12.16) centrowane jest na surowej obserwacji  $x_{ij}$ . Podstawiając w jej miejsce przetworzone obserwacje, pochodzące z sąsiednich pikseli, możemy otrzymać rodzinę filtrów o różnych własnościach (por. [63]). Jako ilustrację, pokażemy filtr o pudełku centrowanym na medianie:

$$\check{\theta}(x_{ij}) = \widehat{Med}\{x_{i+k,j+l}, k, l = -M, \dots, M\}. \quad (12.18)$$

Podstawiając to wyrażenie do (12.16), zamiast surowych obserwacji  $x_{ij}$ , otrzymujemy filtr mniej wrażliwy na błędy grube o postaci

$$y_{ij} = \frac{\sum_{k,l=-M}^M x_{i+k,j+l} \cdot U_H(x_{i+k,j+l} - \check{\theta}(x_{ij}))}{\sum_{k,l=-M}^M U_H(x_{i+k,j+l} - \check{\theta}(x_{ij}))}. \quad (12.19)$$

### Uwagi na temat uogólnień

W powyższych filtrach uśrednialiśmy dane w pudełku. Możemy konstruować podobne filtry zastępując uśrednianie obliczeniem mediany empirycznej z

poziomów szarości w pudełku, które może być centrowane na surowych obserwacjach, medianie, bądź średniej arytmetycznej. Otrzymamy w ten sposób filtry: mediana-mediana, mediana-dane, mediana-średnia. Filtry te lepiej zachowują krawędzie, ale wymagają większych nakładów obliczeniowych.

### Przykłady filtracji z wertykalnymi wagami

W celu zilustrowania działania omawianych filtrów zastosowano je do znanego fragmentu slabu. Na Rys. 12.9, w lewym górnym rogu, pokazano oryginalne zdjęcie. Na środkowym zdjęciu w pierwszym wierszu tego rysunku pokazano wynik filtracji za pomocą filtra z medianą w wertykalnie uciętym pudełku z  $H = 0.1$  w oknie o rozmiarze  $11 \times 11$ , czyli dla  $M = 5$ . Ten rozmiar okna stosowano we wszystkich przykładach. Obrazy reprezentowane były w skali szarości  $[0, 1]$  i do tego przedziału odnoszą się wartości  $H$ .

Wynik filtracji za pomocą średniej w pudełku o wysokości  $H = 0.1$  przedstawiono w pierwszym wierszu po prawej stronie, a w drugim wierszu po lewej mamy wynik filtracji z  $H = 0.015$ .

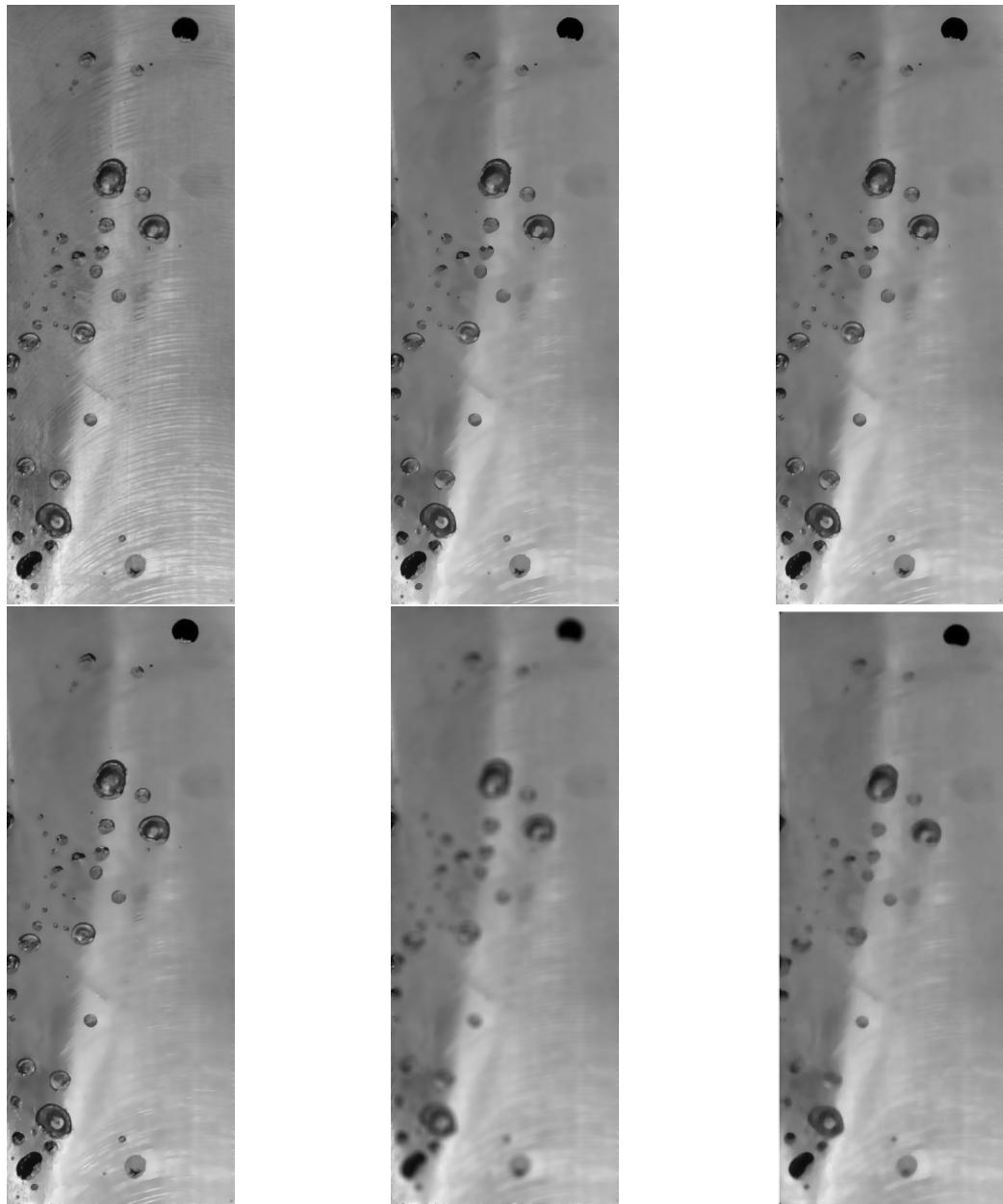
Dla porównania, obraz środkowy w drugim wierszu zawiera wynik filtracji za pomocą klasycznej ruchomej średniej, a obraz po prawej – to wynik zastosowania medianowego filtra (w obu przypadkach w oknie  $11 \times 11$ ). Na obu tych obrazach widać, że brzegi – zwłaszcza małych defektów – są znacznie bardziej rozmyte niż te na obrazach filtrowanych w wertykalnie uciętym pudełku.

Kończąc krótki przegląd filtrów z wertykalnymi wagami warto wspomnieć, że idea wertykalnego ważenia ma także inne zastosowania w przetwarzaniu obrazów i sygnałów. W pracy [65] zaproponowano algorytm wykrywania skokowych zmian w sygnałach i obrazach. Można także skonstruować algorytmy segmentacji.

### Uwagi o innych filtrach

Warto wspomnieć o kilku ważnych klasach filtrów, które nie zostały tutaj omówione.

- Filtry w dziedzinie częstotliwości 2D. Filtry te działają w oparciu o dwu-wymiarową szybką transformację Fouriera (FFT). Stosowana jest ona dwukrotnie. Po przetransformowaniu obrazu do dziedziny częstotliwości zerowane są te fragmenty widma, które odpowiadają szumom. Zastosowana następny raz, odwrotna FFT daje oryginał odfiltrowanego obrazu.



Rysunek 12.9: I wiersz Vert. mediana, V-średina  $H = 0.1, M = 5$  V-średnia  
Class-średnia Class-mediana

- Klasyczny filtr Wienera jest filtrem optymalnym, pod warunkiem, że zachowane są odpowiednie założenia o zakłóciennach, a mianowicie ich stacjonarność, a w przypadku obrazów – izotropowość. Wymagana jest też znajomość funkcji autokorelacji zakłóceń.
- Filtr Kalmana jest filtrem rekurencyjnym o walorach optymalności, jeśli znamy macierz korelacyjną zakłóceń. W wersji dogodnej do przetwarzania obrazów został on zaimplementowany w bibliotece OpenCV.
- Filtry adaptacyjne konstruowane są tak, by mogły się dostosować w miarę napływu informacji o zakłóciennach. Znane filtry tego typu oparte o zasadę *plug-in*, to znaczy, że punktem wyjścia struktura filtru optymalnego (na przykład, filtru Kalmana), a w miejsce nieznanych parametrów (na przykład, macierzy kowariancji) wstawiane są ich oszacowania, które są rekurencyjnie aktualniane.

# Rozdział 13

## Operacje morfologiczne

Teoria operacji morfologicznych bazuje na teorii zbiorów Minkowskiego. Dzięki temu podejście morfologiczne jest bardzo eleganckie, ale nie zawsze udaje się je zastosować w pracy on-line ze względu na duże wymagania dotyczące mocy obliczeniowej.

W rozdziale tym omówimy tylko podstawowe operacje morfologiczne, służące do poprawiania obrazów binarnych. Odsyłamy Czytelnika do znakomitych opracowań, dostępnych w języku polskim, [45], [42] w celu poszerzenia wiedzy na temat bardziej zaawansowanych technik morfologicznych.

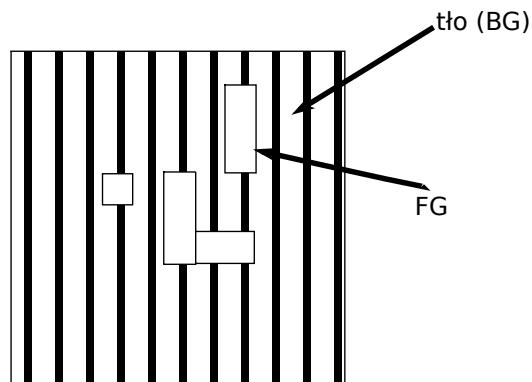
### 13.1 Podstawowe pojęcia

Skupiamy uwagę na operacjach morfologicznych dla obrazów binarnych, uzy-skanych w wyniku segmentacji lub wykrywania krawędzi. Przyjmiemy kodowanie zero-jedynkowe. Zatem, dla wartości obrazu  $I$  dla piksela o położeniu  $p$  mamy  $I(p) \in \{0, 1\}$ .

Operacje, które będziemy wykonywać są nieliniowe. Zatem, kolejność ich wykonywania jest istotna. Zwracamy uwagę, że operacje, których nazwy sugerują, że są one wzajemnym przeciwnieństwem, często nie są operacjami wzajemnie odwrotnymi w sensie matematycznym.

#### 13.1.1 Definicje i podstawowe własności

- Potrzebne nam będą trzy ważne terminy.
  1. Tło *Background* –  $BG = \{p : I(p) = 0\}$ .
  2. Nośniakiem obrazu, oznaczanym przez  $supp\{I\}$ , nazywać będziemy zbiór tych wszystkich  $p$ , dla których  $I(p) = 1$ . Zbiór ten, czasem



Rysunek 13.1: Tło i obiekt

oznaczany też jako *FG* (od ang. *foreground*), zawiera te wszystkie lokalizacje, w których znajdują się obiekty.

3. Element strukturalny (strukturalizujący) *Structuring element*, oznaczany dalej przez *Z* lub *SE*. Element strukturalny jest to obszar o względnie niewielkim rozmiarze, od kilku do kilkunastu pikseli, zwykle kwadrat lub koło, który wskazuje jak duży ma być zasięg operacji morfologicznych wykonywanych na obrazie. Dla uproszczenia zakładamy, że element *Z* jest symetryczny względem obrotu o  $\pi$  względem swego punktu centralnego.

Efekt działania elementu strukturalizującego można sobie wyobrazić jako wynik zastosowania pędzla lub gumki, w zależności od typu operacji, którą wykonujemy. Rozmiar owego pędzla lub gumki pozostaje w naszej dyspozycji. Dodajmy jeszcze, że daną operację morfologiczną wykonuje się ze stałym rozmiarem i kształtem elementu strukturalizującego. Jednakże, już następna operacja korzystać może z innego elementu strukturalizującego. Jednakże we wzorach przytaczanych poniżej, zakładamy, że cały czas korzystamy z jednego elementu strukturalizującego.

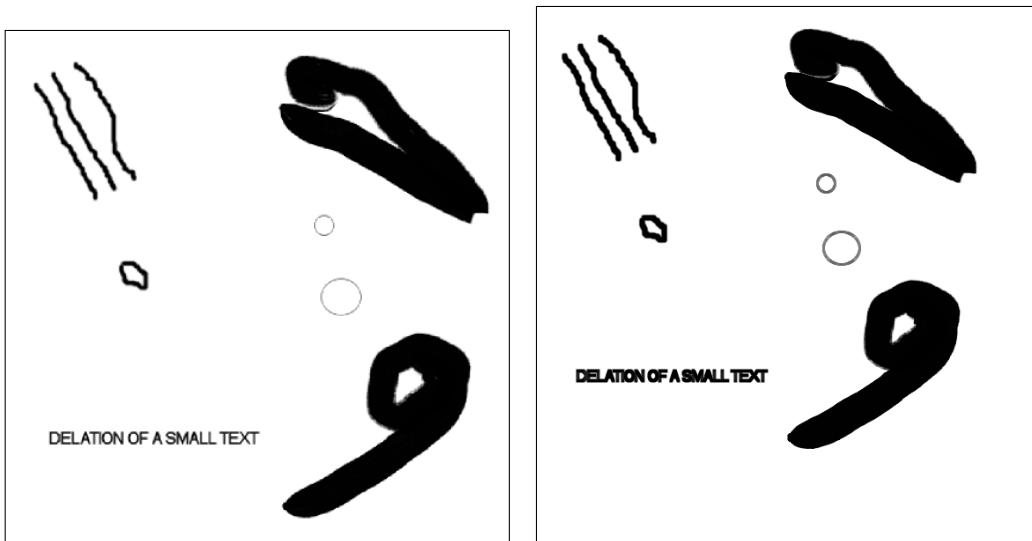
Między tłem a obiektami zachodzą oczywiste związki.

$$FG\{I\} \cup BG\{I\} = I$$

$$FG\{I\} \cap BG\{I\} = \emptyset$$

$$FG\{I\}^C = BG\{I\}$$

$$BG\{I\}^C = FG\{I\},$$



Rysunek 13.2: Dylatacja 3x3 SE

gdzie  $C$  oznacza dopełnienie zbioru.

Stosujemy notację  $Z + p$  dla oznaczenia operacji: przesuń zbiór  $Z$  tak, aby jego środek ciężkości znalazł się w  $p$ .

### 13.1.2 Dylatacja

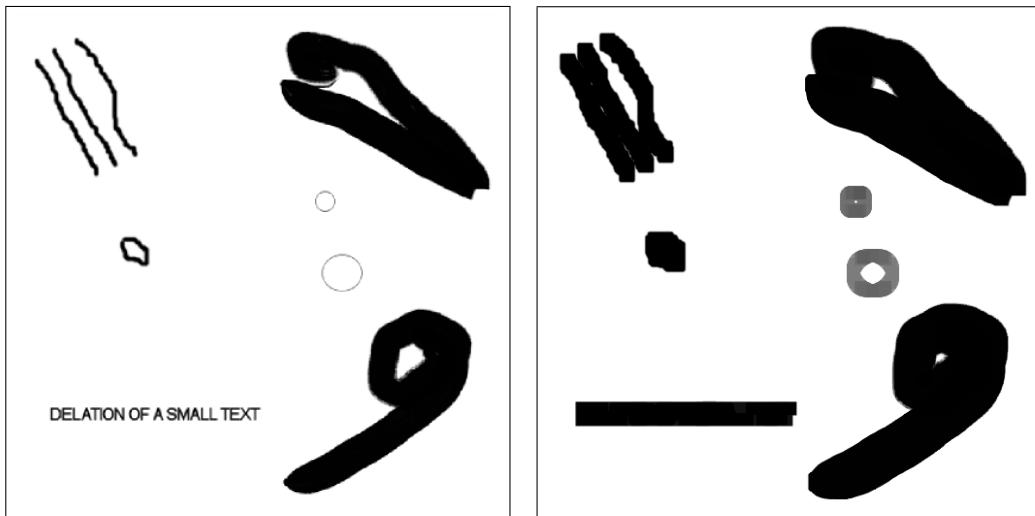
Dylatacją o symetryczny element  $Z$  nazywamy zbiór tych wszystkich położen  $p$  na obrazie takich, że  $(Z + p) \cap supp\{I\} \neq \emptyset$ . Równoważnie, dylatacja może być zdefiniowana wzorem

$$I \oplus Z = \bigcup_{p \in supp\{I\}} (Z + p). \quad (13.1)$$

Dylatacja oznacza w praktyce pewne poszerzenie obiektów na obrazie. Zgodnie ze wzorem (13.1), realizowane jest ono w ten sposób, że:

- 1) element  $z$  przesuwany jest w położenie  $p \in supp\{I\}$ , a więc tam, gdzie znajduje się fragment jednego z obiektów,
- 2) sumujemy wszystkie tak otrzymane zbiorów.

Efekt zastosowania dylatacji z coraz większym elementem strukturalizującym pokazano na Rys. 13.2 oraz 13.3.



Rysunek 13.3: Dylatacja 13x13 SE

### 13.1.3 Erozja

Erozję obrazu binarnego  $I$  o elementzie strukturalnym  $Z$  definiujemy następująco:

$$I \ominus Z = \bigcap_{p \in \text{supp}(I)} (Z + p). \quad (13.2)$$

Intuicyjna rola erozji to zmniejszanie brzegów obiektów o SE, a w przypadku bardzo małych obiektów, nawet całkowite ich usunięcie. Zgodnie z (13.2), SE przesuwany jest do z każdego punktów obiektów i obliczany jest przekrój tak powstałych zbiorów.

Erozję pokazano na Rys. 13.4 i 13.5 dla kwadratowych SE o różnych rozmiarach.

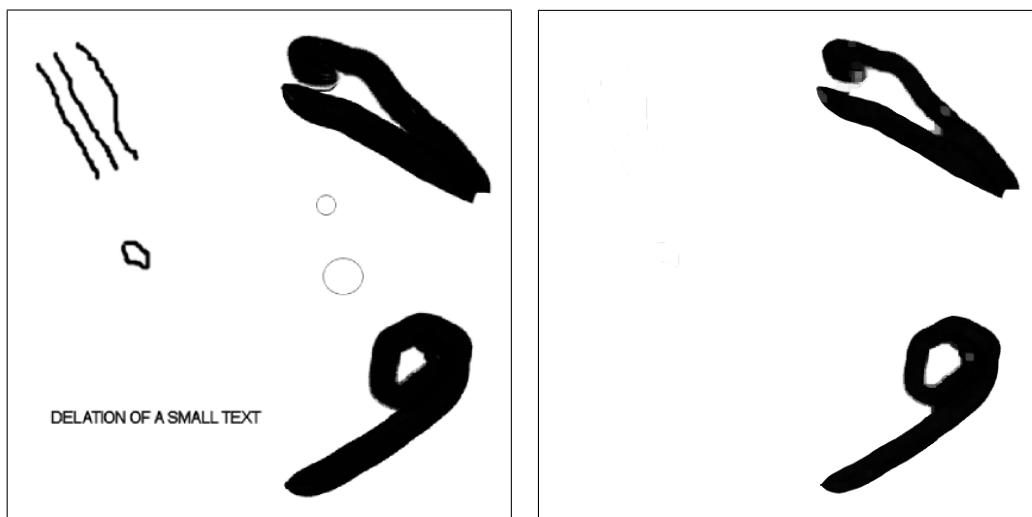
## 13.2 Otwarcie i domknięcie

Operacje otwarcia i domknięcia bazują na dylatacji i erozji. Są one bezpośrednio bardziej użyteczne niż dylatacja i erozja. Definicje tych pojęć nawiązują do intuicji związanych ze zbiorami domkniętymi i otwartymi.

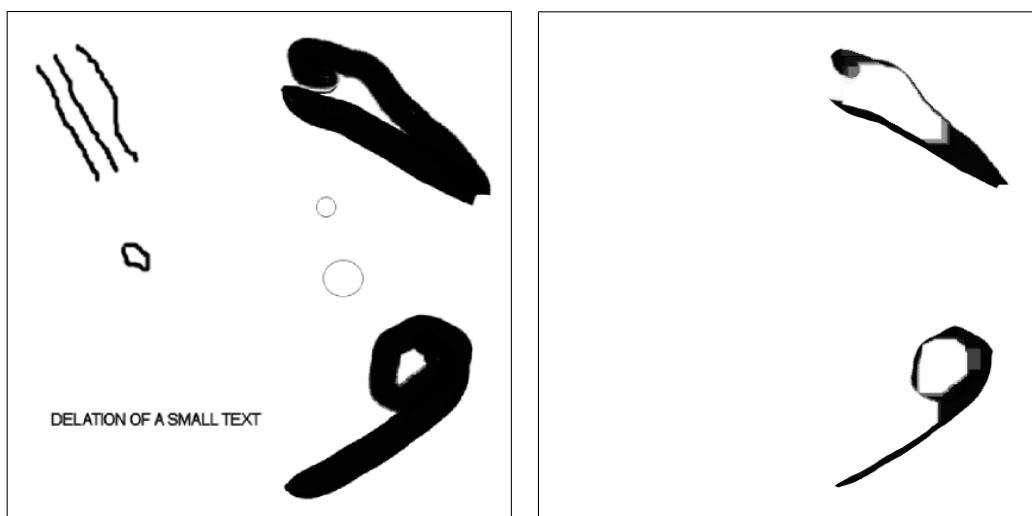
### 13.2.1 Definicje

Operacja otwarcia z elemnitem strukturalnym  $Z$  ma postać:

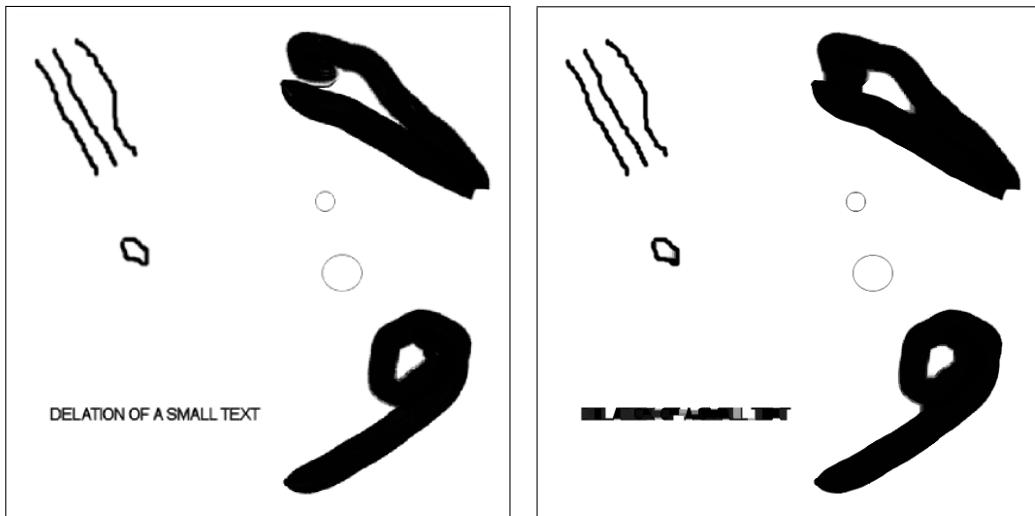
$$I \circ Z = (I \ominus Z) \oplus Z. \quad (13.3)$$



Rysunek 13.4: Erozja 7x7 SE



Rysunek 13.5: Erozja 15x15 SE



Rysunek 13.6: Domknięcie SE 7x7

Operacja ta pozwala usunąć małe obiekty.

Operacja domknięcia z elemntem strukturalnym  $Z$  zdefiniowana jest wzorem.

$$I \bullet Z = (I \oplus \check{Z}) \ominus \check{Z} \quad (13.4)$$

Operacje otwarcia i domknięcia są dualne w następującym sensie:

$$I \circ Z = [I^C \bullet Z]^C$$

$$I \bullet Z = [I^C \circ Z]^C.$$

Operacje domknięcia z niesymetrycznym elementem strukturalnym dają mało intuicyjne efekty.

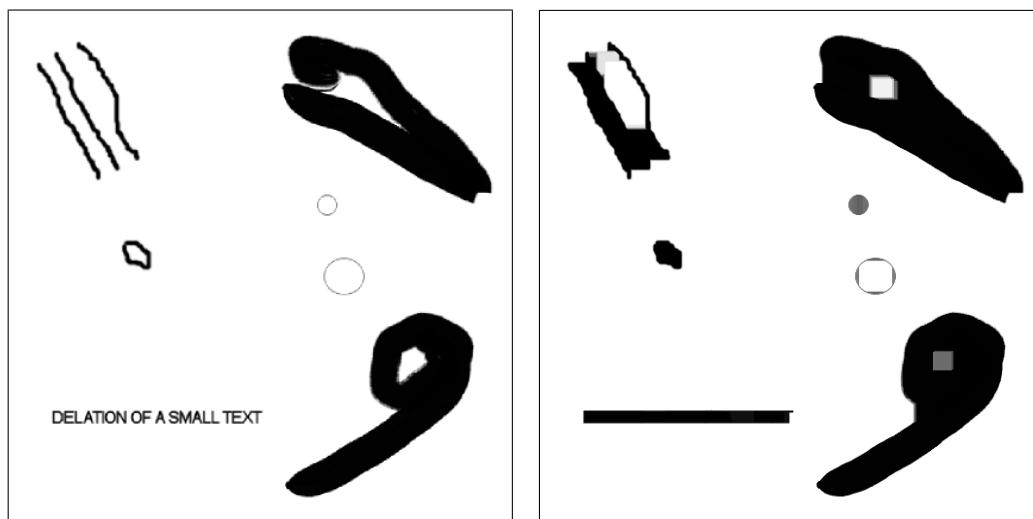
### 13.2.2 Przykłady działania operacji otwarcia i domknięcia

Przykłady działania operacji otwarcia i domknięcia pokazano na Rys. 13.8, 13.9, 13.10 oraz na Rys. 13.6, 13.7, odpowiednio.

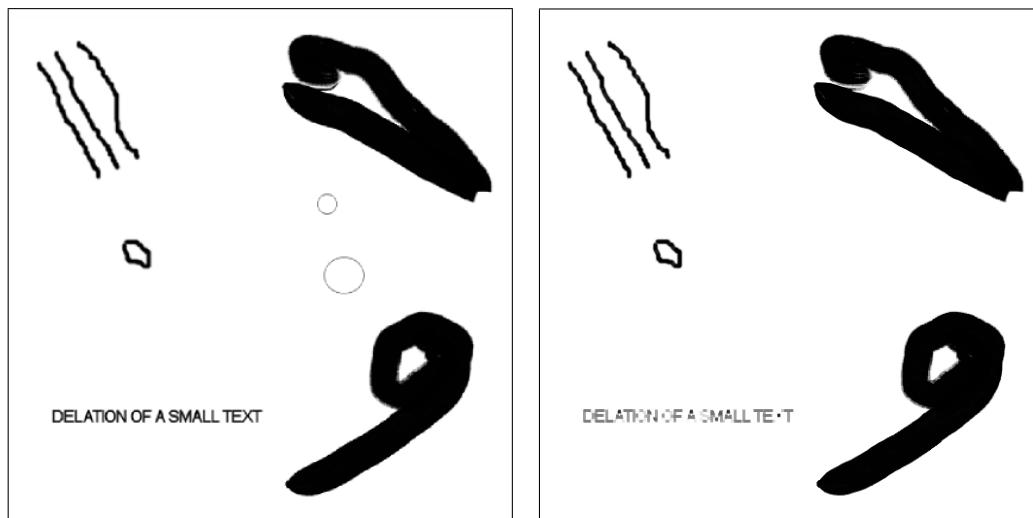
Na Rys. 13.11 porównano wyniki otwarcia i domknięcia na tym samym, znanym już, obrazie rzeczywistym.

Podobnie, na Rys. 13.13 porównano wyniki działania operacji erozji i dylatacji.

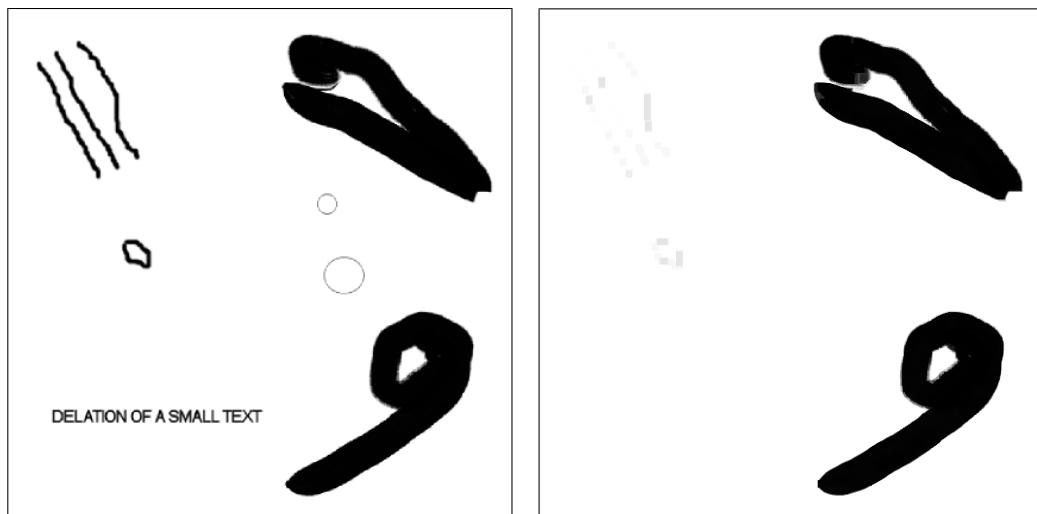
Jako przykład zastosowania otwarcia i domknięcia rozważmy zadanie ekstrakcji brzegu obiektu. Poszerzając obiekt przez domknięcie, zwężając przez



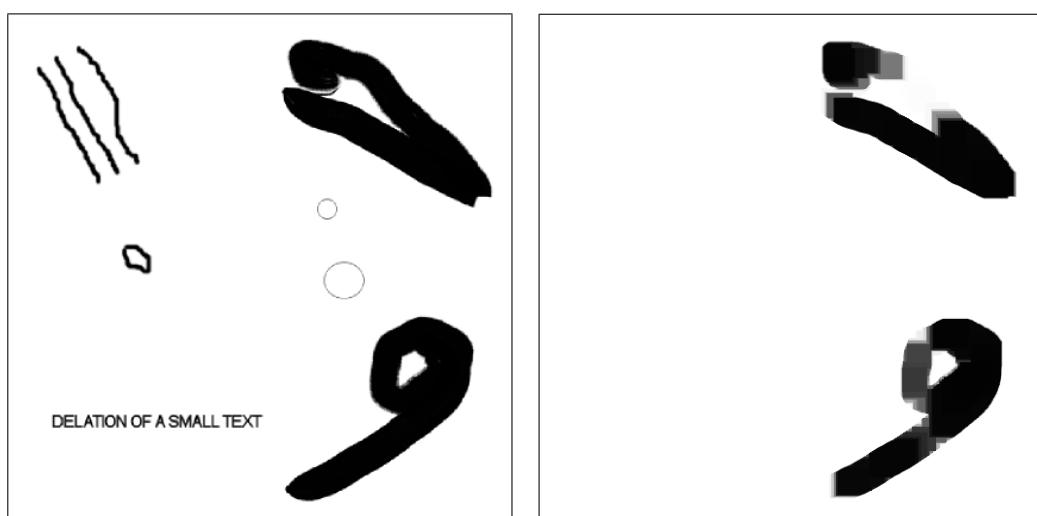
Rysunek 13.7: Domkniecie SE 19x19 – bardzo domknięty



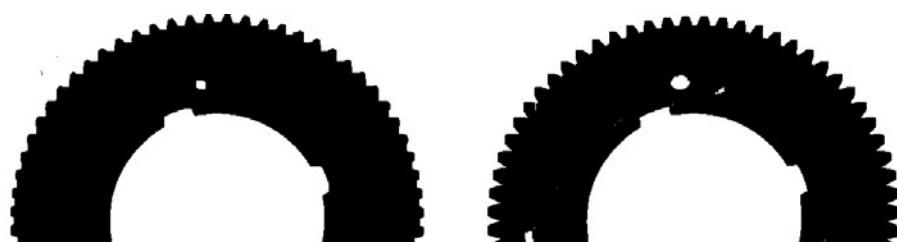
Rysunek 13.8: Otwarcie SE 3x3.



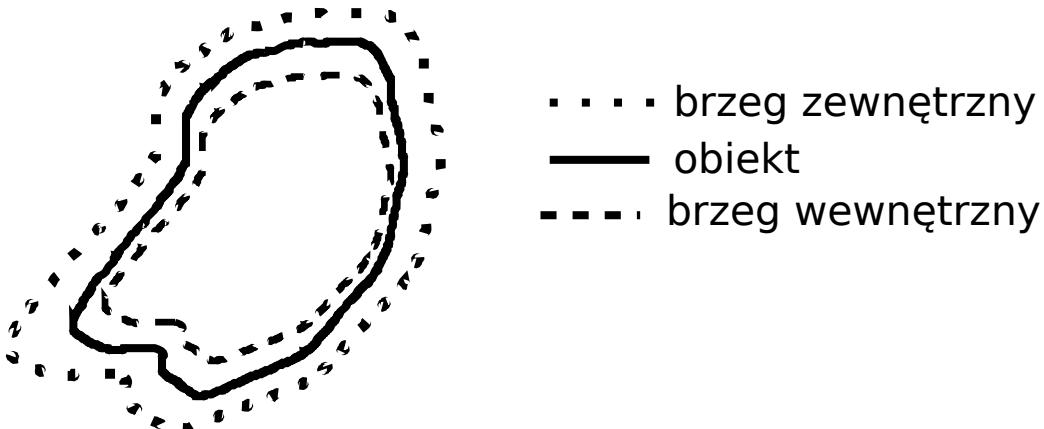
Rysunek 13.9: Otwarcie SE 7x7.



Rysunek 13.10: Otwarcie SE 25x25 – szeroko otwarty.



Rysunek 13.11: Porównanie otwarcia z domknięciem



Rysunek 13.12: Różnice na brzegu

otwarcie i dejmując, dostaniemy brzeg obiektu o grubości zależnej od wielkości elementu strukturalizującego )(por. Rys. 13.12). Subtelne efekty można uzyskać zmieniając kształt elementu strukturalnego.

### 13.3 Złożone operacje morfologiczne

Dylatacja, erozja, otwarcie i domknięcie mogą być połączone na szereg różnych sposobów w celu stworzenia użytecznych operacji takich jak:

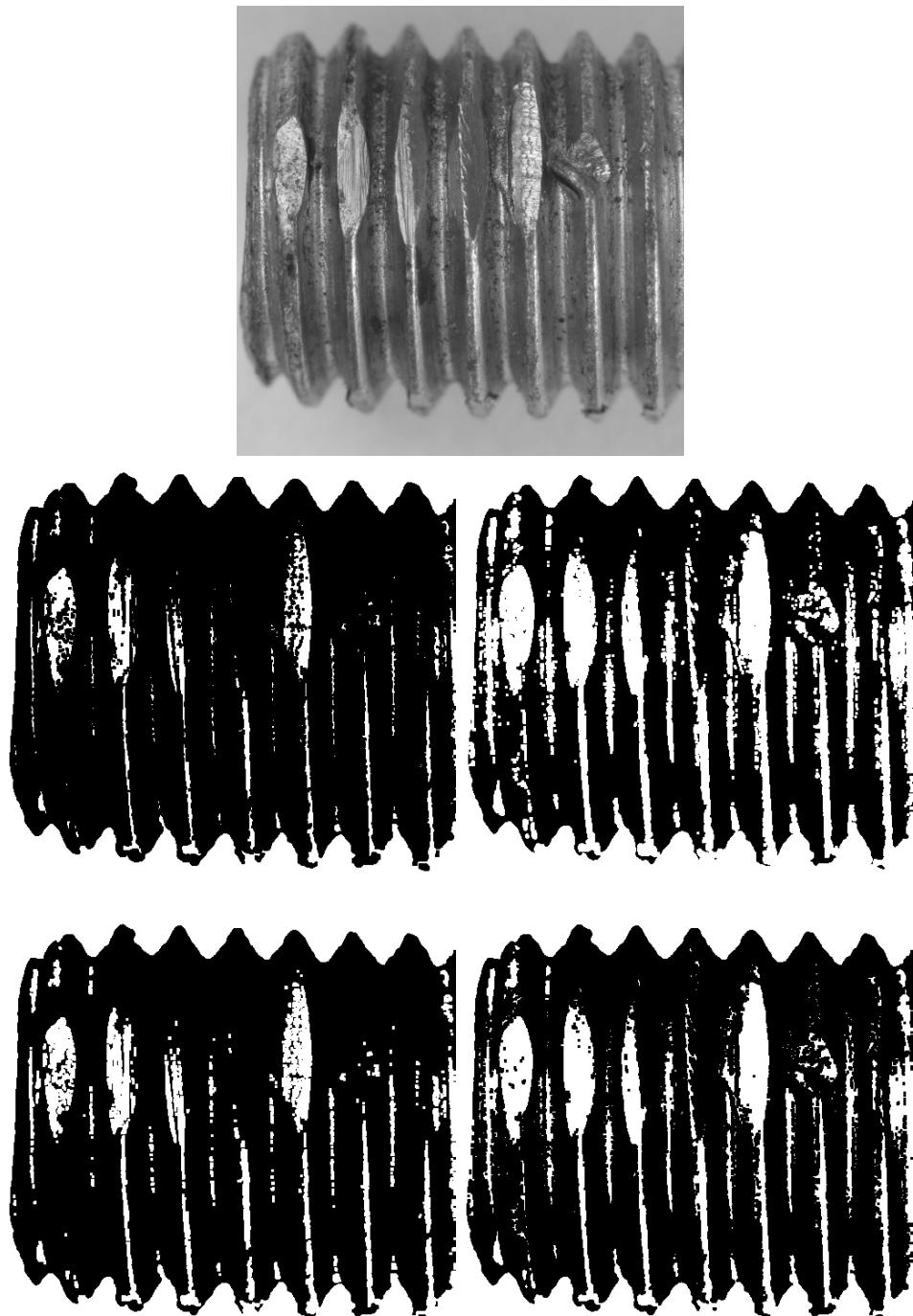
1. trafi – nie trafi (*hit-miss*),
2. filtr *top-hat*,
3. filtr *bottom-hat*,
4. rekonstrukcja,
5. wypełnianie ubytków, itd.

#### 13.3.1 Rekonstrukcja

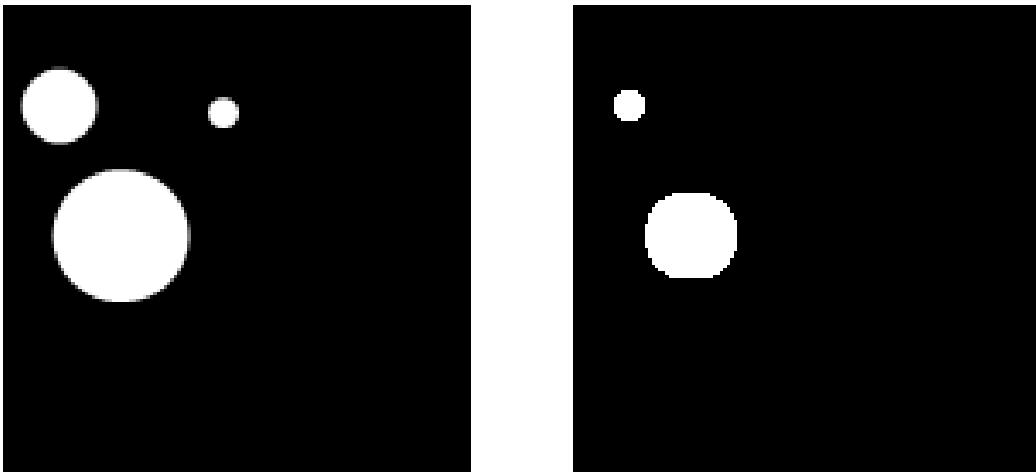
Operacja otwarcia likwiduje drobne szумy, ale też usuwa pewne fragmenty obrazu. Dokonując rekonstrukcji, możemy częściowo przywrócić ogólny kształt obiektu.

Zarys algorytmu rekonstrukcji jest następujący.

1.  $J = I \circ Z$



Rysunek 13.13: Porównanie erozji z dylatacją (wiersz środkowy). Porównanie otwarcia z domknięciem (wiersz dolny)



Rysunek 13.14: SE - koło,  $r=7$ . Znaleźć koła większe od  $r=7$ . Znaleziono prawidłowe obiekty, ale zostały one zmniejszone

2.  $T = J$
3.  $J = J \oplus Z_k, k = 4, 8$
4.  $J = I \text{ AND } J$
5. Jeśli  $J \neq T$  to wróć do kroku 2
6. stop,  $J$  zawiera zrekonstruowany obraz

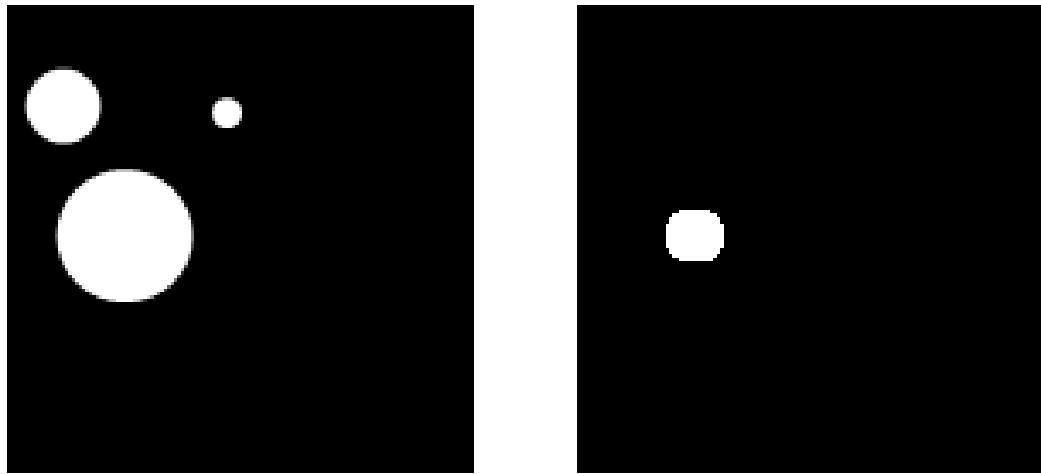
### 13.3.2 Hit or miss

Operacja *hit or miss* pozwala znaleźć obiekt o zadanym kształcie. Na Rys. 13.14 i 13.15 pokazano wyniki poszukiwania kół o promieniach większych niż zowany.

### 13.3.3 Inne zastosowania

Za pomocą operacji morfologicznych możemy także:

- wybierać brzegi obiektów,
- wypełniać obiekty,
- analizować bloki,
- szukać powłoki wypukłej.



Rysunek 13.15: SE - koło  $r = 12$ . Znaleziono koła większe niż  $r=12$ , ale również zostały one mniejszone

Jako przykład, omówimy operację wypełniania. Przypuśćmy, że znaleźliśmy kontur i punkt w środku ( $x_0$ ). Startujemy od zbioru  $X_0 = \{x_0\}$ . Poszerzamy o element strukturalizujący:

$$X_k = (X_{k-1} \oplus B) \cap A^C$$

Iterujemy tak długo aż  $X_k = X_{k-1}$ . W wyniku uzyskujemy  $A \cup X_k$ .

Operacje morfologiczne można zdefiniować także dla obrazów w odcieniach szarości. Podstawą jest interpretacja odcieni szarości w trzech wymiarach. Element strukturalizujący też może wyrażony w odcieniach szarości jako obiekt 3D. Nie rozwijamy tego wątku, gdyż operacje takie trudno zaimplementować w wersji nadającej się do pracy w czasie rzeczywistym. Zaawansowane przykłady zastosowań operacji morfologicznych znaleźć można w pracy [87].

## Część IV

### Epilog



# Rozdział 14

## Zamiast zakończenia – zastosowania termowizji

### 14.1 Bliska podczerwień – zastosowania w monitorowaniu jakości produkcji ciągłej

Kamery CMOS charakteryzują się względnie dużą czułością nie tylko w świetle widzialnym, ale także w bliskiej podczerwieni.

Wybrane klatki z zarejestrowanego filmu cyfrowego pokazano na Rys. 14.1, 14.2, 14.3.

#### 14.1.1 Obserwacje obiektu

Możliwość obróbki cyfrowych obrazów quasi-termowizyjnych on-line otwiera nowe możliwości. Należą do nich:

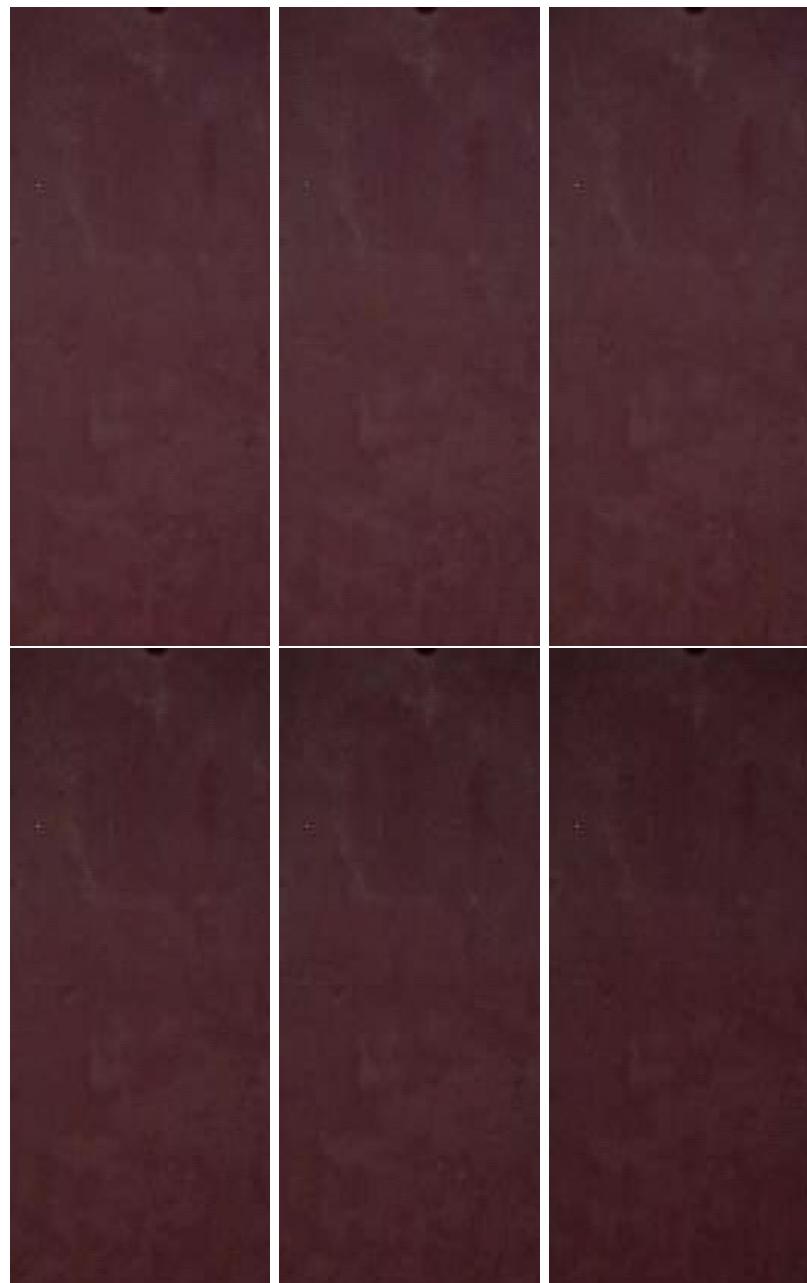
1. śledzenie obszarów o najwyższej temperaturze,
2. ocena równomierności stygnięcia.

Dalej pokazano wyniki eksperymentu studzenia miedzianego slabu, . Celem eksperymentu była ocena możliwości zastosowania kamer pracujących w bliskiej podczerwieni do:

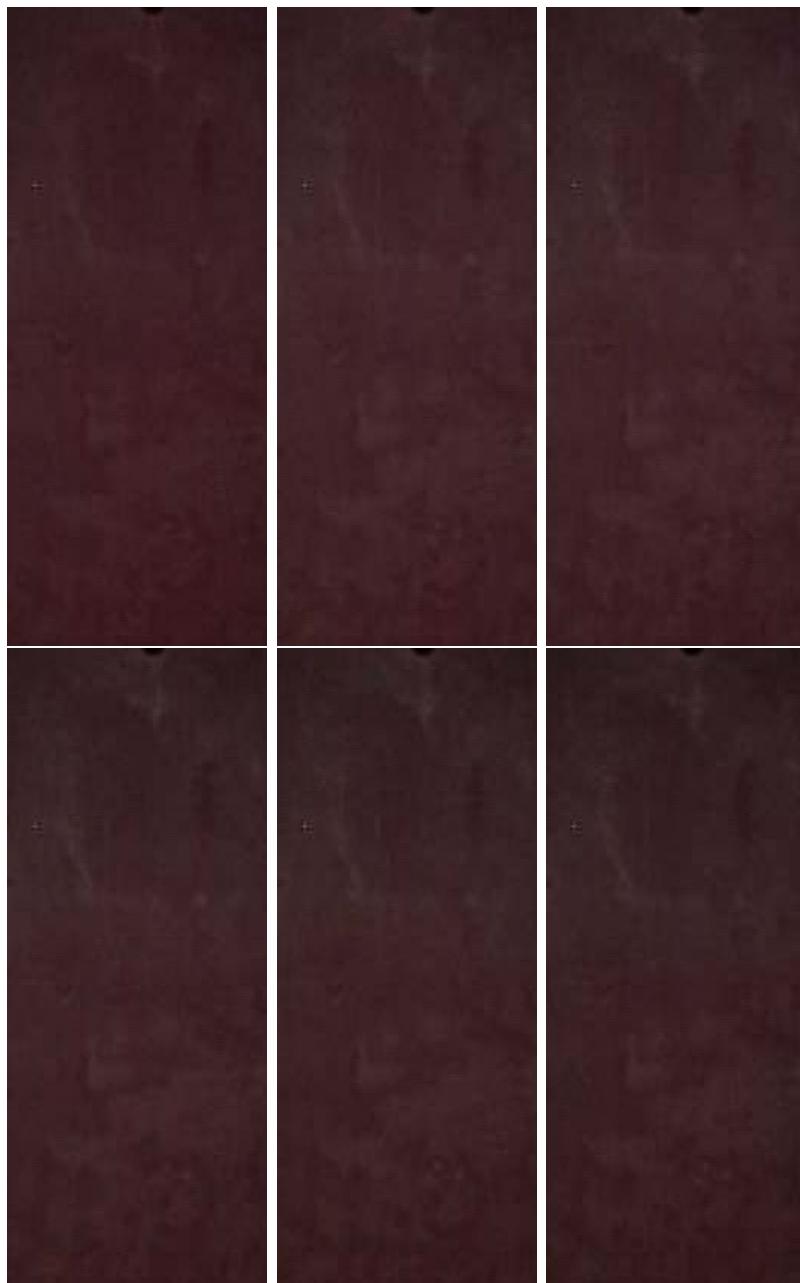
1. uzyskania danych do identyfikacji,
2. wykrywania defektów.

Metodykę prowadzenia eksperymentu opisać można w następujących krokach.

208 ROZDZIAŁ 14. ZAMIAST ZAKOŃCZENIA – ZASTOSOWANIA TERMOWIZJI



Rysunek 14.1: Wybrane zdjęcia termowizyjne w bliskiej podczerwieni – klatki od 10 do 110.



Rysunek 14.2: Wybrane zdjęcia termowizyjne w bliskiej podczerwieni – klatki od 130 do 230.

210 ROZDZIAŁ 14. ZAMIAST ZAKOŃCZENIA – ZASTOSOWANIA TERMOWIZJI



Rysunek 14.3: Wybrane zdjęcia termowizyjne w bliskiej podczerwieni – klatki od 200 do 500.

1. Wlewek miedziany (o długości ok. 30-40 cm) podgrzewano w piecu do temperatury ok.  $850^{\circ}\text{C}$ , na tyle długo, by temperaturę tę osiągnąć mogło wnętrze wlewka.
2. Wlewek wyjmowano z pieca i ustawiano w stronę przygotowanej kamery.
3. Włączano kamerę i rejestrowano cyfrowo proces stygnięcia wlewka, aż do osiągnięcia temperatury, poniżej której ustawało promieniowanie w świetle widzialnym.

### 14.1.2 Wykrywanie wad przez testowanie stałości współczynnika przewodnictwa ciepła

Rozważmy proces polegający na ogrzewaniu płaskiej płyty z materiału dobrze przewodzącego ciepło (metalu) o wymiarach  $X \times Y$ . Oznaczmy rozkład temperatury na powierzchni przez  $T(x, y, t)$ . Równanie przewodnictwa ciepła przybiera następującą postać

$$\frac{\partial T(x, y, t)}{\partial t} = \alpha \Delta T(x, y, t) - bT(x, y, t) + u(x, y, t) \quad (14.1)$$

gdzie  $\Delta T(x, y, t)$  jest laplasjanem  $T$ ,  $u(x, y, t)$  odpowiada rozproszonemu źródłu ciepła ogrzewającemu rozważaną płytę. Współczynnik  $b$  określa wymianę ciepła z otoczeniem. Rozważania następują w  $(x, y) \in (0, X) \times (0, Y)$

Dla równania 14.1 warunki brzegowe drugiego rodzaju określone są następująco

$$\frac{\partial T(0, y, t)}{\partial x} = 0, \frac{\partial T(X, y, t)}{\partial x} = 0, \frac{\partial T(x, 0, t)}{\partial x} = 0, \frac{\partial T(x, Y, t)}{\partial x} = 0 \quad (14.2)$$

dla  $t > 0$ . Dla uproszczenia obliczeń przyjmijmy  $T(x, y, 0) = 0$ . Warunki brzegowe odpowiadają sytuacji, gdy brzeg płyty jest izolowany od otoczenia.

Zastosujmy metodę momentów (opisaną np. w [61]). Przy warunkach początkowych 14.2 funkcje własne  $w_1(x, y), \dots, w_r(x, y)$  przyjmują (po normalizacji) postać

$$\frac{\pi}{2} \left[ \frac{1}{2\pi}, \cos(x), \cos(y), \cos(2x), \cos(2y), \dots \right] \quad (14.3)$$

Dla uproszczenia zapisu przyjęto  $X = Y = \pi$

Rozwiązań równania 14.1 napiszmy w postaci

$$T(x, y, t) = \sum_{k=1}^r i(k, t) w_k(x, y) \quad (14.4)$$

## 212 ROZDZIAŁ 14. ZAMIAST ZAKOŃCZENIA – ZASTOSOWANIA TERMOWIZJI

gdzie  $i(k, t)$  są funkcjami, które należy wyznaczyć. W teorii,  $r \rightarrow \infty$ , w praktyce szereg ten można stosować jako rozwiązanie przybliżone już dla  $r$  rzędu 10. Podstawiając 14.4 do 14.1, wymnażając przez  $w_1(x, y), \dots, w_r(x, y)$  całkując i biorąc pod uwagę ortogonalność funkcji własnych uzyskujemy równania różniczkowe zwyczajne na  $i(k, t)$

$$\frac{di(k, t)}{dt} = -\lambda(k)i(k, t) + U(k, t), k = 1, \dots, r \quad (14.5)$$

gdzie wartości własne  $\lambda(t)$  odpowiadające  $\cos(px) \cos(qy)$  mają postać

$$\lambda(k) = a(p^2 + q^2) + b, k = 1, \dots, r \quad (14.6)$$

warunki początkowe mają postać

$$i(k, 0) = 0 \quad (14.7)$$

Wymuszenie  $U(k, t)$  zdefinowane jest następująco

$$U(k, t) = \int_0^x \int_0^y u(x, y, t) w_k(x, y) dy dx \quad (14.8)$$

Identyfikację wektora  $\alpha$  najprościej przeprowadzić poprzez zastosowanie metody najmniejszych kwadratów dla  $\lambda(k)$ . Poźniej, ponownie metodą najmniejszych kwadratów, znajdujemy  $\alpha$  z gdyż  $\lambda(k)$  są liniowe względem  $\bar{a}$ . Dokładna metoda postępowania opisana jest w [56], a [57] zastosowano ja do testowania stałości (jednorodności) współczynnika od temperatury. Brak jednorodności świadczy o wykryciu defektu.

Zauważmy, że mimo złożonej teorii, końcowe wzory potrzebne do podjęcia decyzji są proste i mogą być stosowane on-line. Korzystamy tu z faktu, że mamy "pomiar temperatury" we wszystkich punktach powierzchni.

## 14.2 Zastosowanie do projektowania układu pomiarowego systemu automatyki

W rozdziale tym przedstawiono możliwości zastosowania kamery termowizyjnej do projektowania układu pomiarowego systemu automatyki.

Badaniu podlegał wymiennik ciepła. Medium ogrzewanym było powietrze atmosferyczne. Wymienniki takie służą do ogrzewania dużych hal fabrycznych, hal magazynowych i supermarketów. Ich moc wynosi od kilkudziesięciu do nawet 500 kW.

Konstrukcja wymiennika jest następująca. Składa się on z:

- tunelu doprowadzającego powietrze chłodne,
- wentylatora tłoczącego powietrze chłodne
- bębna wymiennika ciepła, z palnikiem gazowym,
- tunelu wylotowego ciepłego powietrza,
- rury odprowadzającej spaliny,
- przewodów doprowadzających gaz,
- układu sterowania.

Bęben zawierający palnik gazowy zasilany jest propanem technicznym. Ilość wdmuchiwanego powietrza regulowana jest poprzez zmianę prędkości obrotowej wentylatora, który zasilany jest poprzez falownik. Moc badanego wymiennika wynosiła 100kW.

W toku doświadczenia obserwowano i rejestrowano cyfrowo nagrzewanie się powierzchni bębna wymiennika. Kamerę umieszczono w kanale doprowadzającym w taki sposób, aby przepływające powietrze zapewniało jej chłodzenie.

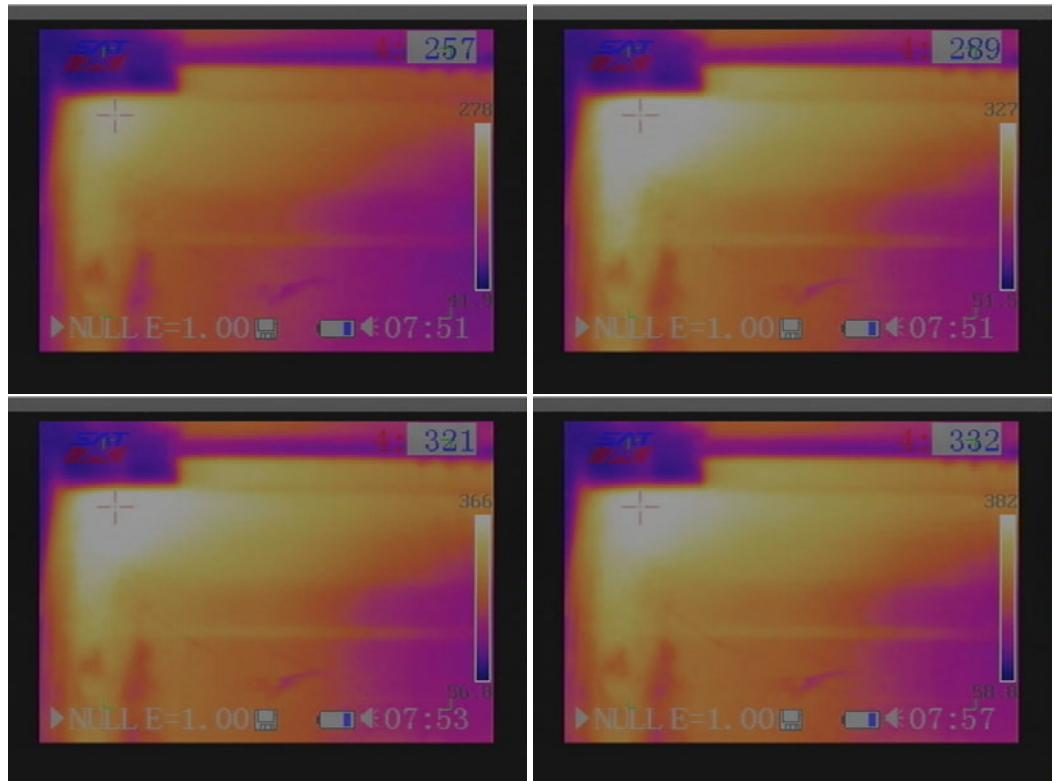
Rozkład temperatury widoczny na Rys. 14.4, 14.5, 14.6 okazał się zaskakujący dla inżynierów.

Pomiar dokonany kamerą termowizyjną pozwolił na wybranie punktu o najwyższej temperaturze (zaznaczony na obrazach "krzyżykiem") i umieszczenie w tym miejscu termopary, co znacznie zwiększyło szanse na zapobieganie stopieniu bębna, nawet wówczas gdy użytkownik nieprawidłowo eksplatawał nagrzewnicy (na przykład, wymuszał dużą moc grzania przy małym przepływie nagrzewanego powietrza).

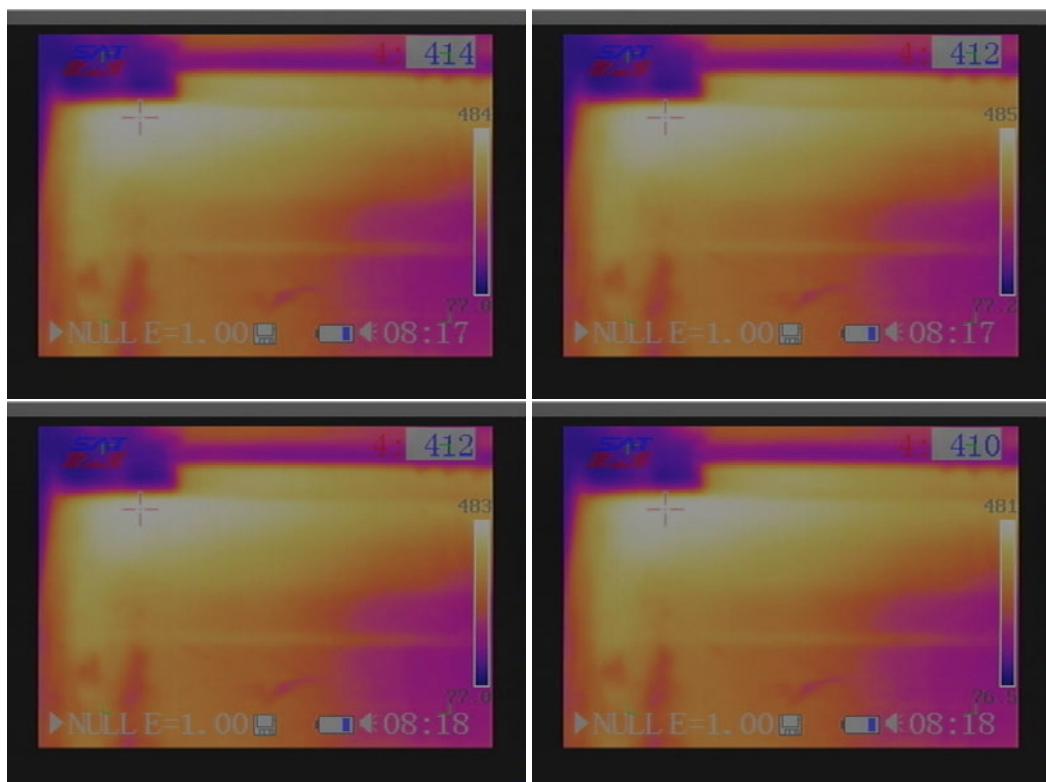
Wobec złożonej konstrukcji nagrzewnicy, która składa się z bębna i połączonych z nim elementów metalowych, przewodzących ciepło także do obudowy, przeprowadzenie teoretycznych obliczeń pola temperatury, na przykład, metodą elementów skończonych, byłoby bardzo trudne (nawet wtedy, gdyby ograniczyć się do rozprływu stacjonarnego) oraz czaso- i koszto-chłonne. Na Rys. 14.4, 14.5, 14.6 widać, że pole to zmienia się wraz z ustawioną mocą nagrzewnicy i przy dużych mocach zagrożony stopieniem może być duży obszar. Zagrożeniem dla trwałości bębna jest także duży gradient temperatury widoczny w lewym górnym rogu obrazów termowizyjnych.

Trudna do zobrazowania w książce, lecz widoczna na filmie cyfrowym, jest także dynamika zmian pola temperatury. Obserwacje zmian w czasie mogą być interesujące dla konstruktorów.

214 ROZDZIAŁ 14. ZAMIAST ZAKOŃCZENIA – ZASTOSOWANIA TERMOWIZJI

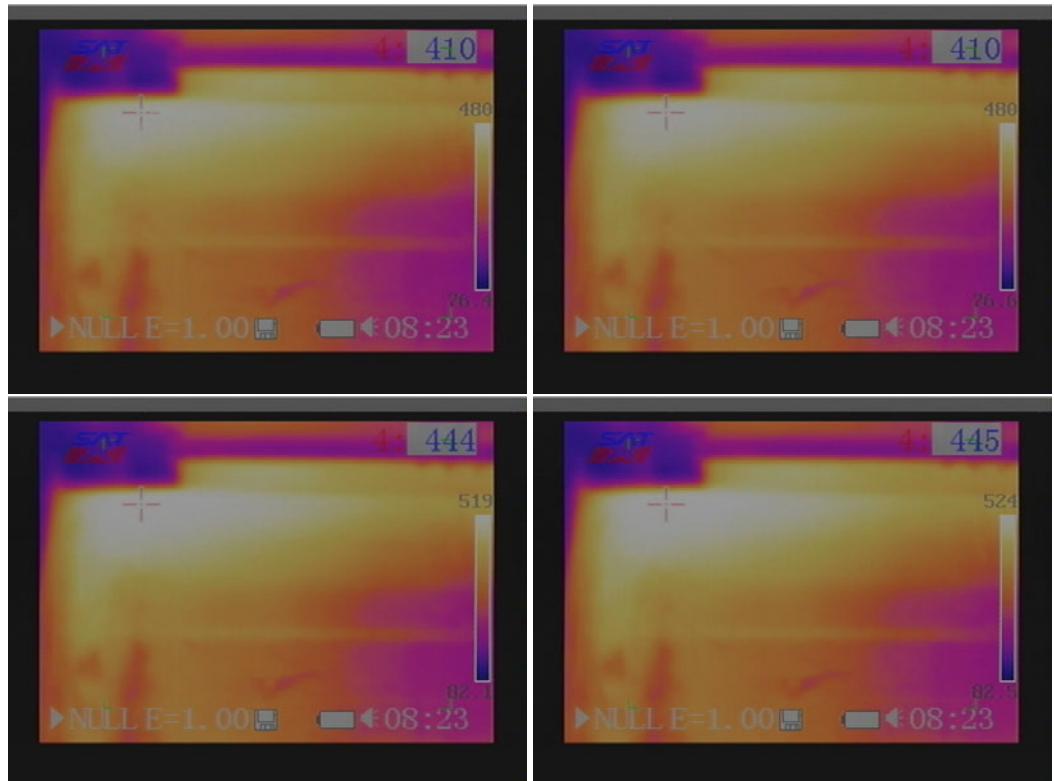


Rysunek 14.4: Rozkład temperatury bębna nagrzewnicy – ustawiona moc 10 kW.



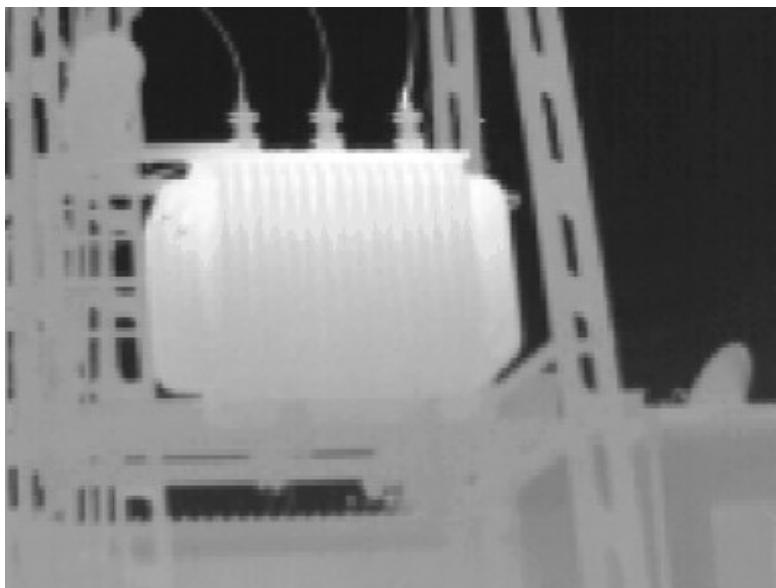
Rysunek 14.5: Rozkład temperatury bębna nagrzewnicy – ustawiona moc 30 kW.

216 ROZDZIAŁ 14. ZAMIAST ZAKOŃCZENIA – ZASTOSOWANIA TERMOWIZJI



Rysunek 14.6: Rozkład temperatury bębna nagrzewnicy – ustawiona moc 50 kW.

### 14.3. OCENA STRAT CIEPLNYCH W SYSTEMACH ENERGETYCZNYCH 217



Rysunek 14.7: Termowizyjny obraz transformatora 1.

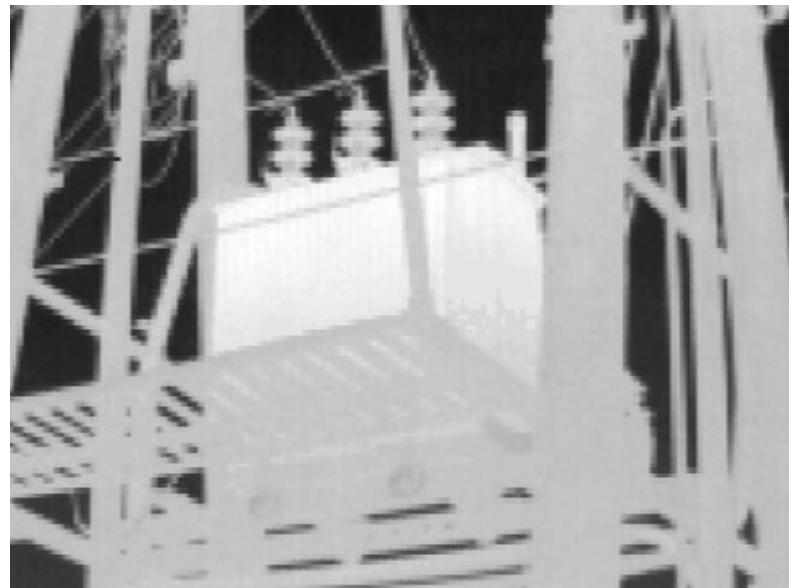
### 14.3 Ocena strat cieplnych w systemach energetycznych

Na zdjęciach 14.7 i 14.8 pokazano dwa transformatory a na zdjęciu 14.9 – metalowy słup linii energetycznej. Wszystkie te obrazy termowizyjne zostały zebrane w jednakowych warunkach termicznych otoczenia i obciążenia sieci (w odstępie kilkudziesięciu sekund). Porównanie zdjęć 14.7 i 14.8 wskazuje na znaczne różnice strat ciepła. Interesujący jest obraz termiczny słupa linii energetycznej – wskazuje on na duże straty ciepła nie tylko w przewodach, ale także w wyniku przepływu ciepła poprzez konstrukcję metalową słupa.

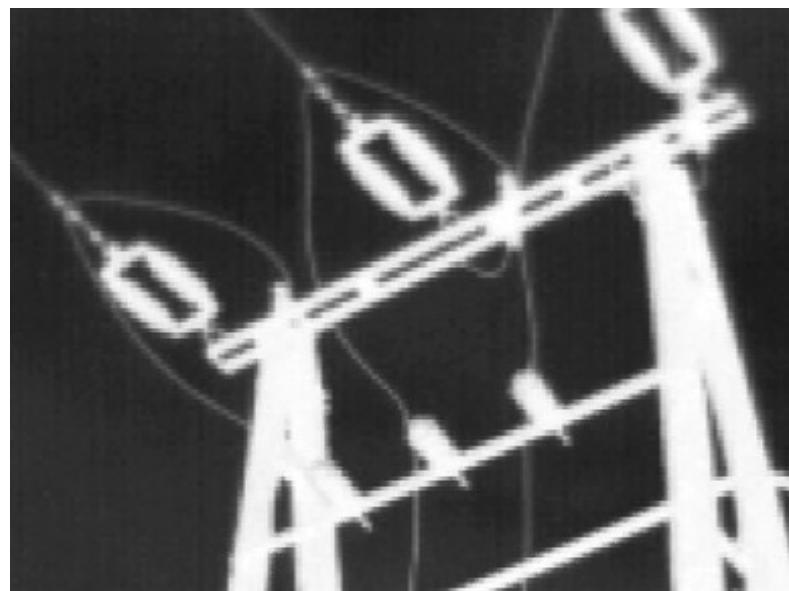
### 14.4 Pomiar strat ciepła w budynkach

Stosowanie kamer termowizyjnych do badania strat cieplnych w budynkach (na potrzeby audytu energetycznego lub uzyskiwania świadectwa charakterystyki energetycznej budynku) jest zastosowaniem dobrze znanym. Dlatego zastosowania tego rodzaju dokumentujemy w zwarty sposób, dla kompletności obrazu. Na zdjęciu 14.10 (po lewej) widzimy termowizyjny obraz wspólnego domu jednorodzinnego. Straty ciepła następują praktycznie wyłącznie poprzez okna. Widoczny w podczerwieni, na zdjęciu 14.10 (po prawej),

218 ROZDZIAŁ 14. ZAMIAST ZAKOŃCZENIA – ZASTOSOWANIA TERMOWIZJI



Rysunek 14.8: Termowizyjny obraz transformatora 2.



Rysunek 14.9: Termowizyjny obraz metalowego słupa trakcji energetycznej.



Rysunek 14.10: Termowizyjny obraz współczesnego domu mieszkalnego (po lewej) i współczesnego budynku użyteczności publicznej

współczesny budynek użyteczności publicznej traci ciepło całą powierzchnią ściany.

220 ROZDZIAŁ 14. ZAMIAST ZAKOŃCZENIA – ZASTOSOWANIA TERMOWIZJI

## Rozdział 15

# Dodatek – podstawy termowizji

### Promieniowanie cieplne

Każde ciało w temperaturze większej od zera bezwzględnego jest źródłem promieniowania o dość szerokim zakresie widma (teoretycznie w całym widmie). Dla ciała doskonale czarnego (patrz 15) prawo Planck'a określa moc promieniowania o długości fali  $\lambda$  ciała o temperaturze  $T$ , wypromienowanej w jednostkowym stożku kąta brylowego

$$I(\lambda, T) = \frac{2hc^2}{\lambda^5} \frac{1}{e^{\frac{hc}{\lambda kT}} - 1} \quad (15.1)$$

We wzorze 15.1  $h$  oznacza stałą Planck'a  $h = 6.627 \cdot 10^{-34} J \cdot s$ ,  $k$  stałą Boltzmanna  $k = 1.38 \cdot 10^{-23} \frac{J}{K}$ .

Łatwo sprawdzić, że maksimum promieniowania przypada dla  $hc = 4.97\lambda kT$ . Na rysunkach 15 oraz 15 widzimy rozkłady emisji w funkcji długości fali. Przy wzroście temperatury o  $2000^\circ$  moc zmienia się o sześć rzędów wielkości.

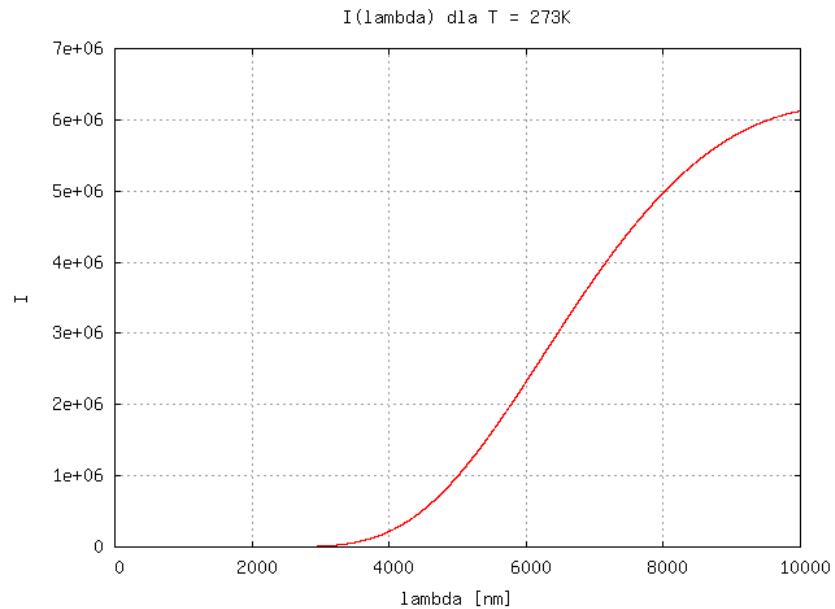
Dla uzyskania wartości całkowitej mocy w określonym zakresie wystarczy scałkować wzór 15.1 po określonym zakresie  $\lambda$ .

$$I_{\lambda_1 \lambda_2}(T) = \int_{\lambda_1}^{\lambda_2} I(\lambda, T) d\lambda \quad (15.2)$$

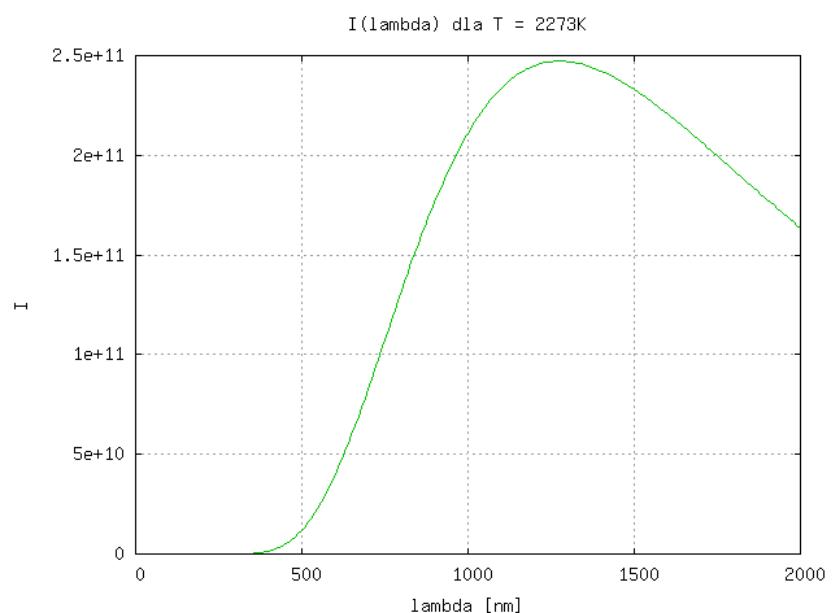
Całkowanie w zakresie  $[0, \infty]$ , a następnie scałkowanie po półsferze prowadzi do otrzymania prawa Stefana-Boltzmanna w postaci:

$$E = \sigma T^4, \quad (15.3)$$

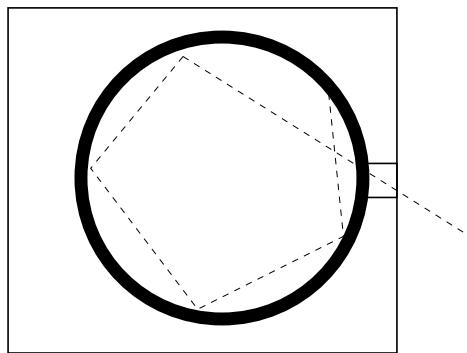
gdzie  $\sigma = \frac{2\pi^5 k^4}{15c^2 h^3}$ .



Rysunek 15.1: Wykresy  $I(\lambda)$  w temperaturze  $0^{\circ}C$



Rysunek 15.2: Wykresy  $I(\lambda)$  w temperaturze  $2000^{\circ}C$



Rysunek 15.3: Schemat ciała doskonale czarnego

### Ciało doskonale czarne

Przez ciało doskonale czarne rozumiemy obiekt pochłaniający całkowicie dostarczane do niego promieniowanie. Ciało takie emitemuje promieniowanie zgodnie z prawem Plancka omówionym w 15.

Najprostszą konstrukcją ciała doskonale czarnego jest wnęka z niewielkim otworem. Jeżeli wnęka wyłożona jest materiałem o dużym współczynniku czerni to konstrukcja ta dość dokładnie odpowiada teoretycznym założeniem dla c.d.c.

### Emisjyjność innych obiektów

Jeżeli rozważane obiekty nie są ciałami doskonale czarnymi to wprowadza się tzw. zdolność promieniowania  $\zeta$ . Dla ciała szarego stosuje się prawa 15.1 i 15.3, z uwzględnieniem omawianego współczynnika (jest on uwzględniany jako mnożnik). Współczynnik  $\zeta$  zależy od temperatury.

W przypadku zabarwienia obiektu rozkład energii znajduje się poniżej krzywej dla c.d.cz. i jest w pewnym stopniu zniekształcony. Najbardziej podobne do krzywej idealnej są krzywe dla metali trudnotopliwych (wolfram, platyna, osm). Maksimum jest przesunięte w kierunku krótszych fal.

W przypadku gazów emisja następuje wyłącznie w określonych długościach fali (np. serie widmowe w atomie wodoru).

### Podział widma promieniowania elektromagnetycznego

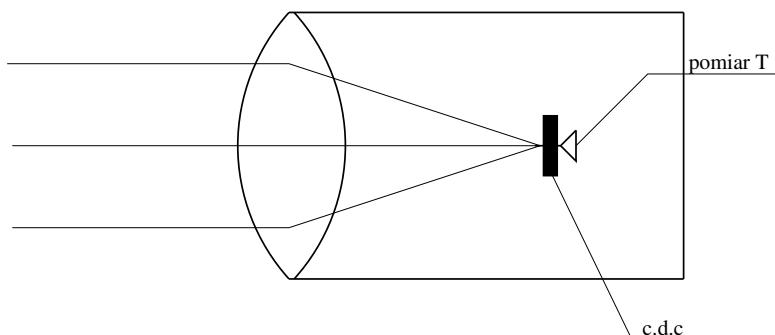
Widmo fal elektromagnetycznych dzieli się umownie na kilka zakresów przedstawionych w Tabeli 15.1 (częstotliwości określają początek zakresu).

Nazwa	$f$	$\lambda$
Promieniowanie $\gamma$	$10^{19}$	$0.01\text{nm}$
Promieniowanie Rentgenowskie	$5 \cdot 10^{17}$	$0.1\text{nm}$
Nadfiolet	$7.5 \cdot 10^{14}$	$400\text{nm}$
Światło widzialne	$4 \cdot 10^{14}$	$750\text{nm}$
Podczerwień bliska	$1.2 \cdot 10^{14}$	$2.5\mu\text{m}$
Podczerwień cieplna	$3 \cdot 10^{14}$	$10\mu\text{m}$
Podczerwień daleka	$3 \cdot 10^{11}$	$1\text{mm}$
Fale radiowe		

Tabela 15.1: Częstotliwości i długości fal elektromagnetycznych.

### Bezkontaktowy pomiar temperatury

Zagadnienie obrazowania rozkładu temperatury najprościej omówić na przykładzie przyrządu mierzącego temperaturę w jednym punkcie. Najprostszą konstrukcję przedstawiono na rysunku 15.



Rysunek 15.4: Schemat punktowego pomiaru temperatury.

Soczewka, wykonana z materiału przepuszczającego podczerwień, skupia promieniowanie termiczne z obiektu na ciele doskonale czarnym. Zaabsorbowane promieniowanie powoduje zwiększenie temperatury. Temperatura ta przeliczona jest na temperaturę obiektu.

Najczęściej do pomiaru wykorzystywany jest bolometr tj. urządzenie mierzące całe docierające do niego promieniowanie. Bolometry różnią się konstrukcją. Najstarsze, mają konstrukcję w postaci termorezistora platynowego pokrytego czernią platynową.

Matryca zawierająca znaczną liczbę bolometrów tworzy przetwornik, pozwalający na uzyskanie obrazu termicznego obserwowanego obiektu.

## Optyka

Prawa rządzące optyką w zakresie fal podczerwonych są dokładnie takie same jak w zakresie światła widzialnego. Jednakże pewne materiały wykazują odmienne właściwości. Zwykłe szkło okienne w zakresie fal podczerwonych staje się materiałem odbijającym promieniowanie. Z tego powodu optyka kamery termowizyjnych wykonywana jest zwykle ze szkła germanowego. Jest to jeden z powodów znacznych kosztów urządzeń pracujących w tym zakresie.



# Bibliografia

- [1] Anagnostopoulos C., High performance computing algorithms for textile quality control, Mathematics and Computers in Simulation 60 (2002) 389–400
- [2] Andreadis I., Automated Visual Inspection Systems Expert Systems, 2002, Pages 771-800
- [3] Asoudegi E., Quantitative automated inspection of standard parts using machine vision Computers and Industrial Engineering, Volume 23, Issues 1-4, November 1992, Pages 361-364
- [4] Bhalerao A. and R. Wilson, Unsupervised image segmentation combining region and boundary estimation, Image and Vision Computing 19 (2001) 353-368
- [5] Bradski G. and A. Kaehler, Learning OpenCV, O'REILLY, Beijing, Cambridge, Farnham, Köln, Sebastopol, Taipei, Tokyo 2008.
- [6] Brosnan T. and Da-Wen Sun, Improving quality inspection of food products by computer vision—a review, Journal of Food Engineering 61 (2004) 3–16
- [7] Campoy P. et al, InsPulp : An on-line visual inspection system for the pulp industry, Computers in Industry 56 (2005) 935–942
- [8] Choraś, R. S., Komputerowa Wizja. Metody interpretacji i identyfikacji obiektów. Akademicka Oficyna Wydawnicza EXIT 2005.
- [9] Conci A. and C. B. Proenca, A fractal image analysis system for fabric inspection based on a box-counting method, Computer Networks and ISDN Systems 30, 1998.1887–1895
- [10] Chapman, K. W. et al, A High Speed Statistical Process Control Application of Machine Vision to Electronics Manufacturing Computers and Industrial Engineering, Volume 19, Issues 1-4, 1990, Pages 234-238

- [11] Chinrungrueng C. and Suvichakorn A., Fast Edge-Preserving Noise Reduction for Ultrasound Images, *IEEE Transactions On Nuclear Science*, VOL. 48, NO. 3, JUNE 2001, 849-854
- [12] Conklin C. et al, Dichromatic color perception: a fast alternative for machine vision systems, *Engineering Applications of Artificial Intelligence* 15 (2002) 351–355
- [13] Davies, E.R., (2005) Machine Vision: Theory, Algorithms, Practicalities, Morgan Kaufmann (3rd edition).
- [14] Davies, E.R., (2000) Image Processing for the Food Industry, World Scientific, Singapore.
- [15] Davies, E.R., Low-level vision requirements, *Electronics and Communication Engineering Journal* October 2000, 197-210
- [16] Davies, E. R., A generalised approach to the use of sampling for rapid object location., *Int. J. Appl. Math. Comput. Sci.*, 2008, Vol. 18, No. 1, 7–19 DOI: 10.2478/v10006-008-0001-3.
- [17] Dahl C. K. and Kim H. Esbensen, Image analytical determination of particle size distribution characteristics of natural and industrial bulk aggregates, *Chemometrics and Intelligent Laboratory Systems* 89 (2007) 9–25
- [18] Demant C., Streicher-Abel B. and P. Waszkewitz; *Industrial Image Processing: Visual Quality Control in Manufacturing*, Springer, Berlin, 1999.
- [19] Derganc J. et al, Real-time automated visual inspection of color tablets in pharmaceutical blisters, *Real-Time Imaging* 9 (2003) 113–124
- [20] Derganc J. et al, A machine vision system for measuring the eccentricity of bearings, *Computers in Industry* 50 (2003) 103–111
- [21] Derrode S. and F. Ghorbel, Robust and Efficient Fourier–Mellin Transform Approximations for Gray-Level Image Reconstruction and Complete Invariant Description, *Computer Vision and Image Understanding* 83, 57–78 (2001)
- [22] Firmin, C., Hamad, D., Postaire, J., Zhang, R.: Gaussian neural networks for bottles inspection: a learning procedure. *International Journal of Neural System* 8 (1997) 41–46

- [23] Foster J. W. et al, Automated visual inspection: A tutorial Computers and Industrial Engineering, Volume 18, Issue 4, 1990, 493-504
- [24] Garcia H. C. et al An Automated Feature Selection Method for Visual Inspection Systems, IEEE Transactions On Automation Science And Engineering, VOL. 3, NO. 4, OCTOBER 2006 394-406
- [25] Gonzales R. C., Woods R. E., *Digital Image Processing*, 2-nd ed.,Prentice Hall 2002.
- [26] Hannah . P. and E R Davies, Automatic Detection Of Foreign Objects In Food Products, Factory 2000 - Advanced Factory Automation, 3-5 October 1994, Conference Publication No 398, 517-522
- [27] Heizman M., Aspects of Image Fusion for Automated Visual Inspection, Pattern Recognition and Image Analysis, 2008, Vol. 18, No. 2, pp. 222–230.
- [28] Hoggar S. G., MATHEMATICS OF DIGITAL IMAGES Creation, Compression, Restoration, Recognition, Cambridge Univ. Press, 2006
- [29] Houzet D., Real-Time Image Processing with a MIMD Computer Real-Time Imaging, Volume 2, Issue 6, December 1996, 383-392
- [30] Hryniwicz O., *Współczesne metody statystyczne w sterowaniu jakości*. IBS PAN, Warszawa 1996
- [31] Jähne, B. Digital Image Processing. Springer-Verlag, 2002.
- [32] Jain A.K. and Chitra D., Practicing vision: Integration, evaluation and applications Pattern Recognition, Volume 30, Issue 2, February 1997, 183-196
- [33] Jain A.K. Fundamentals of Digital Image Processing, Prentice Hall, NY, 1989.
- [34] Kepner J., A multi-threaded fast convolver for dynamically parallel image filtering Journal of Parallel and Distributed Computing, Volume 63, Issue 3, March 2003, Pages 360-372
- [35] King T., Vision-in-the-loop for control in manufacturing Mechatronics, Volume 13, Issue 10, December 2003, 1123-1147
- [36] Kummert and Rafajłowicz Editors, International Journal of Applied Mathematics and Computer Science, 2008, Vol. 18, numer specjalny – poświęcony przetwarzaniu obrazów.

- [37] S. X. Liao and M. Pawlak, "On Image Analysis by Moments," IEEE Trans. Pattern Analysis and Machine Intelligence, PAMI-18, 3, March 1996, 254–266.
- [38] Otsu, N., "A Threshold Selection Method from Gray-Level Histograms," IEEE Transactions on Systems, Man, and Cybernetics, Vol. 9, No. 1, 1979, pp. 62-66.
- [39] Lin H.-D., Automated visual inspection of ripple defects using wavelet characteristic based multivariate statistical approach, Image and Vision Computing 25 (2007) 1785–1801
- [40] Loupas T., W. N. Mcdicken, and P. L. Allen, "An adaptive weighted median filter for speckle suppression in medical ultrasound images," IEEE Trans. Circuits Syst., vol. 36, pp. 129–135, Jan. 1989.
- [41] Malamas E. N., A survey on industrial vision systems, applications and tools Image and Vision Computing, Volume 21, Issue 2, 10 February 2003, Pages 171-188
- [42] Malina, W. i Smiatacz, M., *Metody cyfrowego przetwarzania obrazów*. Akademicka Oficyna Wydawnicza EXIT, 2005.
- [43] Mery D. and Medina O., Automated Visual Inspection of Glass Bottles Using Adapted Median Filtering, A. Campilho, M. Kamel (Eds.): ICIAR 2004, LNCS 3212, pp. 818–825, 2004. Springer-Verlag Berlin Heidelberg 2004
- [44] Nielsen A. A., The Regularized Iteratively Reweighted MAD Method for Change Detection in Multi- and Hyperspectral Data, IEEE Transactions On Image Processing, VOL. 16, NO. 2, february 2007, 463-478.
- [45] Nieniewski, M., *Morfologia matematyczna w przetwarzaniu obrazów*. Akademicka Oficyna Wydawnicza PLJ, 1998.
- [46] Niessen W. J. et al, Nonlinear Multiscale Representations for Image Segmentation, Computer Vision And Image Understanding, Vol. 66, No. 2, May, pp. 233–245, 1997
- [47] Newman, T., Jain, A.: A survey of automated visual inspection. Computer Vision and Image Understanding 61 (1995) 231–262
- [48] Ma W. Y. and B. S. Manjunath, EdgeFlow: A Technique for Boundary Detection and Image Segmentation, IEEE Transactions On Image Processing, VOL. 9, NO. 8, AUGUST 2000, 1375 - 1388

- [49] Mery, D., Jaeger, T., Filbert, D.: A review of methods for automated recognition of casting defects. *Insight* 44 (2002) 428–436
- [50] O'Leary P. Machine vision for feedback control in a steel rolling mill Computers in Industry, Volume 56, Issues 8-9, December 2005, 997-1004
- [51] Pham D.T. and R.J. Alcock, Automated Visual Inspection and Artificial Intelligence Smart Inspection Systems, 2003, Pages 1-34
- [52] Pham D.T. and R.J. Alcock Industrial Inspection Systems Smart Inspection Systems, 2003, 193-208
- [53] Piironen T. et al, Automated Visual Inspection of Rolled Metal Surfaces, Machine Vision and Applications (1990) 3:247-254
- [54] Pratt, W. K., *Digital image processing*, New York, Wiley, 1991.
- [55] Prieto F. et al, Visual system for fast and automated inspection of 3D parts, *Revue Internationale de CFAO et d'informatique graphique*. 1999, 1-17
- [56] Rafajłowicz Ewaryst: Repeated least squares with inversion and its application in identifying linear distributed-parameter systems. *International Journal of Systems Science*. 2000 vol. 31, nr 8, s. 1003-1010
- [57] Rafajłowicz Ewaryst: Testing homogeneity of coefficients in distributed systems with application to quality monitoring. *IEEE Transactions on Control Systems Technology*. 2008, vol. 16, nr 2, s. 314-321
- [58] E. Rafajłowicz, Planowanie eksperymentu i jego zastosowania w monitorowaniu jakości produkcji, Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław, 2005
- [59] Rafajłowicz E., SUSAN edge detector reinterpreted, simplified and modified. International Workshop on Multidimensional (nD) Systems, Aveiro, Portugal, June 27-29, 2007.
- [60] Rafajłowicz E., Wnuk M., Dudek K., Janiczek T., Ułasiewicz J., Rafajłowicz W., Krzywański R., "Analiza możliwości zastosowania metod wizyjnych do ciągłej kontroli odlewanej pasma", Raport końcowy z realizacji I i II Etapu, Raport Instytutu Informatyki Automatyki i Robotyki, grudzień 2006.

- [61] E. Rafajłowicz, W. Rafajłowicz: Beyond Automated Visual Inspection Systems. Proc. 14th Int. Conference of WOSC, Wrocław, September 2008.
- [62] E. Rafajłowicz, W. Rafajłowicz i A. Rusiecki red. „Algorytmy przetwarzania obrazów i wstęp do pracy z biblioteką OpenCV.” Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław 2009
- [63] Rafajłowicz E., Pawlak M., Steland A., Nonlinear image processing and filtering: a unified approach based on vertically weighted regression. *Int. J. Appl. Math. Comput. Sci.*, 2008, Vol. 18, No. 1, 49–61 DOI: 10.2478/v10006-008-0005-z.
- [64] Rafajłowicz E., Skubalska-Rafajłowicz E. Equidistributed sequences along space-filling curves in sampling of images. Proc. 16th European Signal Processing Conference, EUSIPCO 2008, Lausanne, Switzerland, August 25-29, 2008.
- [65] Rafajłowicz Ewaryst, Pawlak Mirosław, Steland Ansgar: Nonparametric sequential change-point detection by a vertically trimmed box method. *IEEE Transactions on Information Theory*. 2010, vol. 56, nr 7, 3621-3634
- [66] Rivera M. and J. L. Marroquin Adaptive Rest Condition Potentials: First and Second Order Edge-Preserving Regularization, *Computer Vision and Image Understanding* 88, 76–93 (2002)
- [67] Rosin P. L., Thresholding for Change Detection, *Computer Vision and Image Understanding* 86, 79–95 (2002)
- [68] Russ J. C., *The Image Processing Handbook*, CRC Press Taylor and Francis Group, 2007
- [69] Shan Y. and G.W. Boon, Sub-pixel location of edges with non-uniform blurring: a finite closed-form approach, *Image and Vision Computing* 18 (2000) 1015–1023
- [70] Shankar N.G. and Z.W. Zhong, Defect detection on semiconductor wafer surfaces, *Microelectronic Engineering* 77 (2005) 337–346
- [71] Skubalska-Rafajłowicz E., Local correlation and entropy maps as tools for detecting defects in industrial images. *Int. J. Appl. Math. Comput. Sci.*, 2008, Vol. 18, No. 1, 41–47 DOI: 10.2478/v10006-008-0004-0.
- [72] Smith S. M. and J. M. Brady, SUSAN – A new approach to low level image processing, *Int. J. Comput. Vision*, vol.23, no.1, pp.4578, 1997.

- [73] Sobel, I., Feldman,G.: "A 3x3 Isotropic Gradient Operator for Image Processing", presented at a talk at the Stanford Artificial Project in 1968, unpublished but often cited, orig. in Pattern Classification and Scene Analysis, Duda,R. and Hart,P., John Wiley and Sons,'73, pp 271-272.
- [74] Stefani S. A. and C R Nagarajah and R Willgoss, A surface inspection technique for continuously extruded cylindrical products, Meas. Sci. Technol. 10 (1999) N21-N25.
- [75] Stojanovic et al, Real-Time Vision-Based System for Textile Fabric Inspection, Real-Time Imaging 7, 507-518 (2001)
- [76] Szczeniowski Fizyka doświadczalna t.3: Optyka, PWN Warszawa 1973.
- [77] Tai X et al editors, Image Processing Based on Partial Differential Equations, Springer-Verlag Berlin Heidelberg 2007
- [78] Tadeusiewicz R., Korohoda P., Komputerowa analiza i przetwarzanie obrazów, Wydawnictwo Fundacji Postępu Telekomunikacji Kraków 1997.
- [79] Tadeusiewicz R., Mikrut Z.: Neural Networks applied to visual pattern recognition - a comparative study. In: Applied Mathematics and Computer Science vol. 4, No. 3, 1994, pp. 397-411
- [80] Tadeusiewicz R., Wszołek W., Wszołek T.: The Selection of Distinctive Features in the Task of Pathological Speech Processing and Classification, in: Medical and Biological Engineering and Computing (Journal of the International Federation for Medical and Biological Engineering), vol. 37, suppl. 2, part I, 1999, pp. 528-529
- [81] Thompson J. R., Koronacki J., *Statystyczne sterowanie procesem. Metoda Deminga etapowej optymalizacji jakości*. Akademicka Oficyna Wydawnicza PLJ, Warszawa, 1994.
- [82] Tien, F. -C. , Yeh, C. -H. and Hsieh, K. -H.(2004) 'Automated visual inspection for microdrills in printed circuit board production', International Journal of Production Research, 42: 12, 2477 — 2495
- [83] Tolba A. S. and A.N. Abu-Rezeq, A self-organizing feature map for automated textile products, Computers in Industry 32 (1997) 319-333
- [84] Trdic R. et al, Monitoring Mineral Wool Production Using Real-Time Machine Vision, Real-Time Imaging 5, 125-140 (1999)

- [85] Tsai D.-M. and C.-Y. Hsieh, Automated surface inspection for directional textures, *Image and Vision Computing* 18 (1999) 49–62
- [86] Tsai D.-M. and T. Y. Huang, Automated surface inspection for statistical textures, *Image and Vision Computing* 21 (2003) 307–323
- [87] Wejrzanowski T., Spychalski W., Różniatowski K., Kurzydłowski K. J., *Image Based Analysis Of Complex Microstructures Of Engineering Materials*, *Int. J. Appl. Math. Comput. Sci.*, 2008, Vol. 18, No. 1, 33–39.
- [88] Wojas J., Promieniowanie termiczne i jego detekcja
- [89] Xie X., Mirmehdi M., *TEXEMS: Texture Exemplars for Defect Detection on Random Textured Surfaces*, *IEEE Transactions On Pattern Analysis And Machine Intelligence*, Vol. 29, NO. 8, August 2007, 1454-1464
- [90] Yazdi H. R. and T. G. King, Application of' Vision in the Loop' for Inspection of Lace Fabric, *Real-Time Imaging* 4, 317-332 (1998)
- [91] Yin L., R. Yang, M. Gabbouj, and Y. Neuvo, “Weighted median filters: a tutorial,” *IEEE Trans. Circuits Syst. II*, vol. 43, pp. 157–192, Mar. 1996.
- [92] Zhang L. et al, Real-time automated visual inspection system for contaminant removal from wool, *Real-Time Imaging* 11 (2005) 257–269

# Skorowidz

- centroid, 154
- domknięcie, 196
- dylatacja, 195
- erozja, 196
- filtr
  - średniej geometrycznej, 182
  - średniej harmonicznej, 182
  - ruchomej średniej, 180
  - ważonej ruchomej średniej, 180
- filtracja, 177
- głębia ostrości, 38
- histogram, 63, 171
- Hougha transformacja, 158
- kamera, 3, 19, 20, 23, 25–27, 30, 31, 33
  - termowizyjna, 3, 28
- kontur, 129
- krawędź, 131
- maska, 130
  - Laplacea, 137
  - Prewitta, 133
  - Sobela, 135
- matryca, 27, 34
- szum, 39
- obiektyw, 35, 225
  - ogniskowa, 35
- obraz
  - formaty, 50
- gamma, 55
- jasność, 54
- kontrast, 55
- operacja, 75
  - reprezentacja, 45, 49
- Otsu metoda, 86
- otwarcie, 196
- progowanie, 82
  - adaptacyjne, 93
  - dwustronne, 94
- przesłona, 37
- przetwornik, 34
- roi, 47
- segmentacja, 81
- transformacja, 53
  - obrót, 53
- Trusella metoda, 96
- wzorzec
  - lokalizacja, 101
- znieksztalcenia, 35