

Politechnika Łódzka

ZESZYTY NAUKOWE Nr 1119

PIOTR M. SZCZYPIŃSKI

KOMPUTEROWA ANALIZA OBRAZÓW Z ENDOSKOPU BEZPRZEWODOWEGO DLA DIAGNOSTYKI MEDYCZNEJ

ŁÓDŹ 2012

POLITECHNIKA ŁÓDZKA

ZESZYTY NAUKOWE Nr 1119

ROZPRAWY NAUKOWE, Z. 426

PIOTR M. SZCZYPIŃSKI

KOMPUTEROWA ANALIZA OBRAZÓW
Z ENDOSKOPU BEZPRZEWODOWEGO
DLA DIAGNOSTYKI MEDYCZNEJ

ŁÓDŹ 2012

ZESZYTY NAUKOWE POLITECHNIKI ŁÓDZKIEJ
SCIENTIFIC BULLETIN OF THE LODZ UNIVERSITY
OF TECHNOLOGY
BULLETIN SCIENTIFIQUE
DE L'UNIVERSITÉ POLYTECHNIQUE DE LODZ
НАУЧНЫЕ ЗАПИСКИ
ЛОДЗИНСКОГО ПОЛИТЕХНИЧЕСКОГО УНИВЕРСИТЕТА
WISSENSCHAFTLICHE HEFTE
DER TECHNISCHEN UNIVERSITÄT IN LODZ

Redaktor Działu: **prof. dr hab. inż. Piotr Wodziński**

Recenzenci: **prof. dr hab. inż. Michał Strzelecki**
prof. dr hab. inż. Tomasz Zieliński

© Copyright by Politechnika Łódzka 2012

Adres Redakcji – Адрес Редакции – Editor's Office
Adresse de Redaction – Schriftleitungsadresse:

WYDAWNICTWO POLITECHNIKI ŁÓDZKIEJ
90-924 Łódź, ul. Wólczańska 223
tel./fax 42-684-07-93
e-mail: zamowienia@info.p.lodz.pl
www.wydawnictwa.p.lodz.pl

ISSN 0137-4834

Nakład 140 egz. Ark. druk. 9. Papier offset. 80 g, 70 x 100
Druk ukończono w czerwcu 2012 r.
Wykonano w Drukarni Offsetowej Quick-Druk, 90-562 Łódź, ul. Łąkowa 11

Spis treści

Streszczenie.....	5
Wykaz skrótów i oznaczeń.....	6
1. Wstęp.....	9
2. Problemy, cel i motywacja.....	13
2.1. Endoskopia kapsułkowa.....	13
2.2. Niedostatki narzędzi do analizy filmów.....	19
2.3. Przegląd kierunków badań.....	20
2.4. Motywacja i cel badań.....	22
3. Analiza ruchu endoskopu.....	24
3.1. Model ruchu endoskopu.....	24
3.2. Modele deformowalne.....	27
3.3. Koncepcja deformowalnych pierścieni.....	32
3.4. Wpływ obrazu na MDR.....	35
3.5. Obliczanie naprężeń w MDR.....	38
3.5.1. Metoda transformacji sąsiedztwa.....	39
3.5.2. Podział dopasowania na etapy.....	42
3.5.3. Transformacja z ograniczaniem składowych.....	44
3.5.4. Właściwości metody transformacji sąsiedztwa.....	45
3.6. Dopasowanie MDR.....	46
3.7. Dane wynikowe MDR.....	47
3.7.1. Deskryptory ruchu.....	47
3.7.2. Mapa przewodu pokarmowego.....	48
4. Walidacja MDR.....	51
4.1. Dobór metody obliczania wpływu obrazu.....	52
4.2. Model ruchów jelita.....	61
4.3. Sterowanie odtwarzaniem filmu.....	64
4.4. Zastosowanie mapy przewodu pokarmowego.....	65
5. Detekcja zmian chorobowych.....	68
5.1. Tekstura	69
5.2. Cechy barwy.....	72
5.3. Obszary zainteresowania i mapy cech.....	74
5.4. Selekcja cech i redukcja wymiarowości ich przestrzeni.....	76
5.5. Metoda powierzchni wypukłej.....	78
6. Walidacja metod analizy tekstury i barwy.....	82
6.1. Detekcja zmian chorobowych – program MaZda.....	82
6.2. Detekcja zmian chorobowych – deskryptory MPEG-7.....	92
6.3. Wykorzystanie map cech do segmentacji obrazów.....	94

6.3.1. Segmentacja za pomocą cech wybranych do detekcji.....	95
6.3.2. Segmentacja za pomocą cech wybranych do segmentacji.....	100
Podsumowanie.....	105
Bibliografia.....	108
Załącznik A.....	126
Symbole.....	126
Transformacja skalowania, rozciągania, obrotu i przesunięcia.....	126
Transformacja skalowania, obrotu i przesunięcia.....	127
Transformacja obrotu i przesunięcia.....	128
Dekompozycja macierzy transformacji.....	130
Załącznik B.....	132
Nazwy cech barwy i tekstury w programie MaZda 5.0.....	132
Konwersja do obrazu monochromatycznego.....	132
Modyfikacja histogramu obrazu monochromatycznego.....	134
Algorytmy obliczania cech tekstury.....	135
Cechy statystyczne histogramu.....	135
Cechy mapy wartości bezwzględnych gradientu jasności.....	136
Cechy macierzy zdarzeń.....	136
Cechy macierzy długości ciągów pikseli.....	138
Parametry modelu autoregresji.....	138
Energie w pasmach częstotliwościowych transformaty falkowej Haara. .	139
O autorze.....	140
Abstract.....	142

Streszczenie

Jednym z badań medycznych stosowanych w diagnostyce chorób przewodu pokarmowego jest bezprzewodowa endoscopia kapsułkowa. Wynikiem badania jest film, którego interpretacja przeprowadzana przez lekarza wymaga dużego skupienia uwagi, jest długotrwała i męcząca. Wyniki interpretacji nie są powtarzalne – zależą od wiedzy i doświadczenia konkretnego lekarza.

Przedmiotem niniejszej monografii są opracowane przez autora metody numeryczne, których celem jest analiza obrazów cyfrowych z endoskopu bezprzewodowego zwiększające powtarzalność, wiarygodność oraz obiektywizm diagnozy medycznej. Mają one ograniczyć nakład pracy lekarza podczas oglądania filmu, zautomatyzować procedurę interpretacji oraz umożliwić analizę ilościową wybranych zmian chorobowych. Zaproponowane rozwiązania pozwalają między innymi na scharakteryzowanie ruchu własnego endoskopu względem przewodu pokarmowego, automatyczną regulację szybkości odtwarzania filmu, rekonstrukcję obrazu powierzchni przewodu pokarmowego, detekcję wybranych zmian chorobowych na podstawie cech barwy i tekstury obrazu oraz segmentację obrazów w celu wyodrębnienia obszarów objętych zmianami patologicznymi.

Do najważniejszych osiągnięć opisanych w niniejszej monografii należą: model deformowalny do analizy ruchu własnego kamery, uniwersalna metoda obliczania naprężeń w modelach deformowalnych oraz nowy, szybki algorytm selekcji cech i klasyfikacji trudno rozdzielnych skupień wykorzystujący wielotop wypukły. Wszystkie opracowane metody poddano ocenie jakościowej i ilościowej, w szczególności zbadano ich przydatność do wspomagania interpretacji filmów endoskopowych. Efektem prowadzonych prac są udostępnione w internecie programy komputerowe, w których zastosowano opracowane algorytmy: WCE Player do analizy ruchu własnego endoskopu i rekonstrukcji powierzchni przewodu pokarmowego, a także MaZda do analizy cech barwy i tekstury obrazów oraz do klasyfikacji danych.

Wykaz skrótów i oznaczeń

CLD	deskryptor rozkładu przestrzennego koloru (ang. <i>color layout descriptor</i>)
COM	macierz zdarzeń (ang. <i>co-occurrence matrix</i>)
DCD	deskryptor koloru dominującego (ang. <i>dominant color descriptor</i>)
EHD	deskryptor histogramów linii brzegowych (ang. <i>edge histogram descriptor</i>)
HMMD	przestrzeń koloru (ang. <i>hue min max difference</i>)
HSV	przestrzeń kolorów odcień, nasycenie i wartość (ang. <i>hue, saturation, value</i>)
HTD	deskryptor tekstury homogenicznej (ang. <i>homogeneous texture descriptor</i>)
L*a*b*	przestrzeń kolorów objęta normą Międzynarodowej Komisji Oświetleniowej
LDA	liniowa analiza dyskryminacyjna (ang. <i>linear discriminant analysis</i>)
MAD	średnia różnic bezwzględnych (ang. <i>mean absolute difference</i>)
MDF	atrybuty najlepiej różnicujące (ang. <i>most discriminative features</i>)
MDR	model deformowalny o kształcie koncentrycznych pierścieni (ang. <i>model of deformable rings</i>)
MI	informacja wzajemna (ang. <i>mutual information</i>)
MSE	błąd średniokwadratowy (ang. <i>mean square error</i>)
NCC	znormalizowana korelacja wzajemna (ang. <i>normalized cross-correlation</i>)
NNCH	najbliższy wielotop wypukły (ang. <i>nearest neighbour convex hull</i>)
RBFN	radialna sieć neuronowa (ang. <i>radial basis function network</i>)
RGB	przestrzeń kolorów składowych czerwonej, zielonej i niebieskiej (ang. <i>red, green, blue</i>)
RLM	macierz długości ciągów pikseli (ang. <i>run-length matrix</i>)
ROI	obszar zainteresowania (ang. <i>region of interest</i>)
SCD	skalowalny deskryptor koloru (ang. <i>scalable color descriptor</i>)
SFFS	sekwencyjne przeszukiwanie w przód (ang. <i>sequential floating forward selection</i>)
SMO	minimalna optymalizacja sekwencyjna (ang. <i>sequential minimal optimization</i>)
SOI	filtr o skończonej odpowiedzi impulsowej
SVM	maszyna wektorów nośnych (ang. <i>support vector machine</i>)
TBD	deskryptor przeszukiwania tekstury (ang. <i>texture browsing descriptor</i>)
VSCH	wielotop wypukły oparty na wektorach (ang. <i>vector supported convex hull</i>)

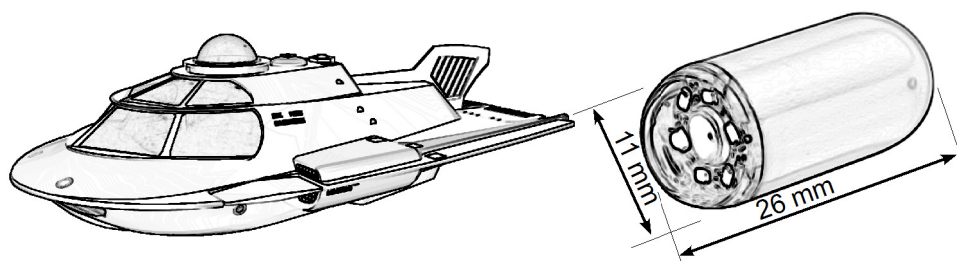
WCE	beprzewodowa endoskopia kapsułkowa (ang. <i>wireless capsule endoscopy</i>)
XYZ	przestrzeń kolorów objęta normą Międzynarodowej Komisji Oświetleniowej
YIQ	przestrzeń kolorów standardu telewizji NTSC (ang. <i>National Television System Committee</i>)
YUV	przestrzeń kolorów wykorzystywana w standardzie telewizji PAL (ang. <i>Phase Alternating Line</i>)
∇	operator nabra
a	współczynnik skalowania
A	zdarzenie losowe
B	zdarzenie losowe
$\mathbf{B}_j$	wersor j -tej hiperpłaszczyzny o zwrocie do wnętrza wielotopu
b_{j0}	przesunięcie hiperpłaszczyzny względem początku układu współrzędnych zgodnie z wektorem $\mathbf{B}_j$
$\mathbf{C}_0$	środek ciężkości wielotopu
d	odległość obrazu od środka przetwornika optoelektronicznego; symbol różniczkowania
D	średnica pola widzenia w obrazie endoskopowym
E	funkcjonał, tzw. energia wewnętrzna aktywnego konturu
E_e	składnik energii wewnętrznej modelujący oddziaływanie obrazu
E_i	składnik energii wewnętrznej modelujący naprężenia
$\mathbf{e}_k$	odległość między położeniem węzła a jego położeniem otrzymanym w wyniku transformacji
E_{MDR}	energia wewnętrzna deformowalnych pierścieni
$E_{p\ q}$	energia cząstkowa oddziaływań między węzłami
f	ogniskowa obiektywu kamery
G	deskryptor ruchu modelu deformowalnym
(i)	numer kolejnego przybliżenia procedury optymalizacji
$I(\cdot)$	funkcja jasności obrazu
I_m	maksymalna jasność obrazu
j_{kl}	elementy macierzy transformacji afinicznej
$\mathbf{J}_s$	macierz transformacji afinicznej dla sąsiedztwa węzła s
k	indeks, zmienna pomocnicza
l	indeks, zmienna pomocnicza
m	indeks, zmienna pomocnicza
n	indeks, zmienna pomocnicza
$\mathbf{M}$	macierz stała wykorzystywana do obliczania naprężeń
$M_{p\ q}$	energia cząstkowa modelowania oddziaływań obrazu
N	wielkość bloku pikseli
N_s	liczebność zbioru wektorów cech

$O(.)$	wskaźnik złożoności obliczeniowej algorytmu
P	liczba pierścieni
p, q	parametry powierzchni parametrycznej; indeksy węzłów modelu dyskretnego
$P(.)$	prawdopodobieństwo zdarzenia
Q	liczba węzłów w pierścieniu; wskaźniki niedoskonałości odseparowania klas za pomocą powierzchni wypukłej
r	promień najmniejszego pierścienia po zainicjowaniu modelu
r_D	połowa średnicy kapsułki
S	droga pokonana przez kapsułkę względem ścian przewodu pokarmowego
s	parametr krzywej parametrycznej; indeks węzła modelu dyskretnego
t	czas
Δt	czas prezentacji obrazu sekwencji wideo
t_k	element wektora przesunięcia w transformacji afinicznej
$\mathbf{T}_s$	wektor przesunięcia w transformacji afinicznej dla sąsiedztwa węzła s
$\mathbf{v}$	wektor współrzędnych węzła dyskretnego modelu deformowalnego
$\hat{\mathbf{v}}$	wektor współrzędnych węzła dyskretnego w modelu odniesienia
V_D, V_P	estymowana prędkość wzdłużna endoskopu
w	stosunek promieni sąsiednich pierścieni
$\mathbf{X}$	wektor współrzędnych wszystkich węzłów modelu deformowalnego
x, y	współrzędne węzła dyskretnego modelu deformowalnego na płaszczyźnie
x^o, y^o	wersory w kierunkach wyznaczonych przez osie współrzędnych x, y
$\hat{x}, \hat{y}$	współrzędne węzła dyskretnego, niezdeformowanego modelu deformowalnego
α	kąt obrotu
γ	kąt widzenia kamery
β	kąt określający kierunek rozciągnięcia względem osi poziomej
Δg	czas prezentacji obrazu w automatycznym sterowaniu szybkością odtwarzania filmu
ε_s	błąd średniokwadratowy położenia węzłów
η	parametr określający wielkość sąsiedztwa
H_s	sąsiedztwo węzła s
λ	wartości własne macierzy
μ	wartość średnia jasności obrazu
ξ	parametr wpływu obrazu na model deformowalny
ρ	parametr odporności modelu deformowalnego na deformowanie
ρ_A	parametr podatności modelu na skalowanie
ρ_R	parametr podatności modelu na obrót
ρ_S	parametr podatności modelu na rozciąganie
ρ_T	parametr podatności modelu na przemieszczanie
σ	odchylenie standardowe jasności w obrazie
τ	parametr odporności aktywnego konturu na rozciąganie
φ	kąt pod jakim widziany jest obrazowany obiekt
ψ	parametr sterowania szybkością odtwarzania filmu
ω	estymowana prędkość obrotowa endoskopu wokół jego osi podłużnej

1. Wstęp

W 1966 roku na ekranach kin pojawił się film *Fantastyczna podróż* (*Fantastic Voyage*) [1], [2] w reżyserii Richarda Fleishera. Fabuła oparta jest na oryginalnym pomysłu – w wyniku wyścigu zbrojeń i rywalizacji dwóch supermocarstw powstaje technologia umożliwiająca zmniejszanie atomów. Dzięki tej technologii łódź podwodna pomniejszona do wielkości mikrometra zostaje wprowadzona do ludzkiego ciała, aby wykonać skomplikowany zabieg chirurgiczny.

Trzydzieści cztery lata później, w „Nature” [3], pojawia się wzmianka o nowej technice endoskopowej, w której wykorzystuje się miniaturową kapsułkę do obserwacji wnętrza układu pokarmowego człowieka. Kapsułka, co ciekawe, powstała w wyniku badań i projektów inżyniera konstruującego wcześniej pociski rakietowe. Jego doświadczenie, zdobyte w zupełnie innej dziedzinie, zostało wykorzystane do opracowania narzędzia stosowanego w diagnostyce medycznej [4]. Bezprzewodowy endoskop kapsułkowy (ang. *Wireless Capsule Endoscope*), zwany też po prostu endoskopem bezprzewodowym lub kapsułkowym, to pierwszy krok ku realizacji fantazji z 1966 roku.



Rys. 1.1. Rysunki łodzi podwodnej Proteus z filmu *Fantastic Voyage* (po lewej) i kapsułki endoskopu bezprzewodowego (po prawej)

Przed pojawieniem się endoskopu bezprzewodowego obrazowanie wizyjne wnętrza układu pokarmowego było możliwe tylko dzięki technikom wzornikowym, gastrokopii i kolonoskopii. Gastrokopia [5–8] to metoda diagnostyki górnego odcinka układu pokarmowego, która polega na wprowadzeniu do niego giętkiego przewodu ze źródłem światła i układem przesyłającym obraz z wnętrza ciała. Za pomocą gastrokopu obejrzeć można wnętrze przełyku, żołądka oraz dwunastnicy. Kolonoskopia [5] to analogiczna technika umożliwiająca obrazowanie jelita grubego. Gastrokop i kolonoskop nie pozwalają jednak zobrazować wnętrza jelita cienkiego. Problem ten częściowo rozwiązuje technika enteroskopii dwubalonowej [9], [10], która daje możliwość wizualizacji całości układu pokarmowego, jednakże wiąże się ona z dużą uciążliwością dla pacjenta – wymaga

często znieczulenia ogólnego, jest długotrwała, a jej zastosowanie może prowadzić do powikłań.

Natomiast przeprowadzenie badania za pomocą endoskopu bezprzewodowego, po pierwsze jest metodą mało inwazyjną, a po drugie umożliwia kompleksową wizualizację całego przewodu pokarmowego. Dodatkowo należy podkreślić fakt, że jest ono mniej uciążliwe dla pacjenta niż badanie wzornikowe, gdyż kapsułka endoskopu bezprzewodowego z zamontowaną w jej wnętrzu kamerą jest połykana przez pacjenta i w sposób naturalny wędruje wewnątrz przewodu pokarmowego. Wspomniany wcześniej film *Fantastic Voyage* trwa sto minut – wynikiem działania kapsułki jest również film, ale trwa on blisko osiem godzin, a następne kilka godzin poświęca lekarz gastroenterolog na jego dokładne obejrzenie i ocenę.

Podczas analizowania danych lekarz używa komputera ze specjalistycznym oprogramowaniem, które umożliwia odtwarzanie zarejestrowanego filmu z różnymi prędkościami, zatrzymywanie go oraz cofanie, oznaczanie charakterystycznych miejsc przewodu pokarmowego, a także wskazywanie obrazów z odnalezionymi zmianami chorobowymi. Dopiero po obejrzeniu całości nagrania gastroenterolog może w pełni ocenić jego jakość i stopień przydatności do celów diagnostycznych. Narzędzia do automatycznego wykrywania obrazów ze zmianami chorobowymi dostępne w komercyjnych programach komputerowych są niedoskonałe. Przykładowo procedury wskazujące obrazy o dominującym kolorze czerwonym, mające sugerować miejsca potencjalnych krwawień, w praktyce przydatne są tylko w niewielkim stopniu. Ilościowa analiza rozległych zmian chorobowych, której wyniki stanowią kluczowy czynnik decydujący o sposobie leczenia także nie jest wspomagana automatycznie. Niewystarczające są też informacje dotyczące lokalizacji kapsułki w ciele pacjenta, a sama zawartość pojedynczego obrazu ze zmianą chorobową nie pozwala na wystarczająco dokładne wskazanie jej położenia w ciele pacjenta. Wyposażenie omawianych programów komputerowych w brakujące elementy automatycznego wspomaganie diagnozy staje się jednym z bardziej pożądanym działań poprawiających w rezultacie skuteczność endoskopii bezprzewodowej.

Szybka lokalizacja zmian chorobowych, możliwość oceny przydatności całego filmu przed przystąpieniem do jego analizowania, szybkie przywoływanie istotnych fragmentów filmu, odrzucenie nieistotnych odcinków materiału wideo oraz automatyczne sterowanie odtwarzaniem filmu mogą poprawić efektywność badania, jak i trafność samej diagnozy.

Niniejsza praca jest poświęcona zagadnieniom komputerowej analizy obrazów cyfrowych z endoskopu bezprzewodowego. Jej celem jest opracowanie i walidacja algorytmów komputerowych zwiększających powtarzalność, wiarygodność oraz obiektywizm diagnozy medycznej. Mają one również ograniczyć nakład pracy lekarza podczas oglądania filmu endoskopowego oraz zwiększyć komfort jego pracy, a ponadto zautomatyzować procedurę interpretacji oraz umożliwić analizę ilościową wybranych zmian chorobowych. Zakłada się wykorzystanie endoskopu bezprzewodowego o typowej budowie bez wprowadzania w nim zmian konstrukcyjnych i bez stosowania uzupełniających technik obrazowania.

Rozwiązanie przedstawionych problemów jest możliwe za pomocą metod komputerowej analizy obrazów, takich jak automatyczna detekcja i klasyfikacja zmian chorobowych, segmentacja obrazów w celu wyodrębnienia obszarów tych zmian oraz śledzenie ruchu własnego endoskopu względem przewodu pokarmowego.

Zarówno zmiany konstrukcyjne kapsułki, w szczególności opracowanie napędu i zdalnego sterowania, jak i stosowanie dodatkowych metod obrazowania, takich jak ultrasonografia lub tomografia komputerowa, są albo technicznie niewykonalne, albo wiążą się ze zwiększeniem kosztów badania lub zwiększają dyskomfort pacjenta.

Osiągnięcie przedstawionych celów wymaga opracowania metod numerycznych przetwarzania i analizy obrazów. Pierwszym krokiem jest analiza kształtu geometrycznego kapsułki i charakteru jej poruszania się wewnątrz przewodu pokarmowego. Umożliwia ona sformułowanie założeń upraszczających, które pozwalają na zdefiniowanie modelu ruchu kapsułki wewnątrz układu pokarmowego i wyprowadzenie przybliżonych zależności wiążących informację niesioną przez film z szybkością przemieszczania się endoskopu. W celu rozwiązania problemu analizy ruchu w obrazie i szacowania szybkości opracowano metodę deformowalnych pierścieni [11–18] należących do rodziny tzw. modeli deformowalnych [19–24]. Deformowalne pierścienie charakteryzuje oryginalna budowa, nie występująca w innych znanych modelach. Ponadto, zastosowano w nich nowatorską metodę korekcji kształtu i położenia, która rozwiązuje problem korekcji kształtu w sposób ogólny i może być stosowana w innych modelach deformowalnych o arbitralnie zdefiniowanej budowie.

Jednym z efektów działania deformowalnych pierścieni jest opis ruchu własnego kapsułki za pomocą deskryptorów. Deskryptory są wielkościami liczbowymi charakteryzującymi wybrane cechy i składowe ruchu kapsułki. Wykazano, że deskryptory umożliwiają rozpoznawanie rodzajów skurczów jelita cienkiego, które determinują ruch kapsułki oraz intensywność zmian w rejestrowanym przez endoskop obrazie. Wykorzystując zależność pomiędzy wartościami deskryptorów a intensywnością zmian w obrazie, opracowano metodę automatycznego doboru szybkości odtwarzania filmu, dzięki której obserwowana przez lekarza zmiana treści obrazu jest bardziej płynna.

Kolejnym efektem działania modelu jest pojedynczy obraz (mapa) będący reprezentacją całego przewodu pokarmowego. Obraz ten umożliwia szybką ocenę kompletności i jakości uzyskanego materiału wideo oraz wykrycie niektórych rozległych zmian chorobowych. Zaproponowana metoda generowania obrazu-mapy na podstawie filmu z endoskopu kapsułkowego stanowi także oryginalne opracowanie.

Dla potrzeb detekcji wybranych zmian patologicznych sformułowano algorytm łączący analizę tekstury i barwy obrazu oraz selekcję cech i klasyfikację. Zaprojektowana metoda selekcji i klasyfikacji, wykorzystująca wielotop [25] wypukły wektorów, będąca również własnym opracowaniem autora, pozwala na szybkie i pełne przeszukanie podzbiorów cech, co nie było możliwe w praktyce przy użyciu innych znanych metod selekcji.

Przedstawione powyżej zagadnienia zaprezentowano w pracy według określonej kolejności. Informacje dotyczące konstrukcji endoskopu bezprzewodowego, charakterystyki uzyskiwanych danych obrazowych oraz narzędzia wykorzystywane przez lekarza do analizowania danych przedstawiono w rozdziale 2. Przyjęty w rozważaniach model ruchu endoskopu kapsułkowego oraz metody analizy ruchu własnego endoskopu bezprzewodowego i automatycznego doboru szybkości odtwarzania filmu, a także algorytm generowania obrazu reprezentującego całość powierzchni wewnętrznej układu pokarmowego omówiono w rozdziale 3.

W rozdziale 5. przedstawiono metody służące do detekcji obrazów ze zmianami chorobowymi. Opisano w nim także przyjętą metodę obliczania lokalnych cech (deskryptorów) barwy i tekstury obrazu. Ponadto scharakteryzowano algorytmy selekcji cech, umożliwiające odróżnienie obrazów przedstawiających zmiany chorobowe od pokazujących prawidłowy stan powierzchni przewodu pokarmowego.

Walidacja metod, opinie lekarzy, ocena ilościowa wyników zastosowania metody w diagnostyce medycznej zawarte są w rozdziałach 4. i 6. W rozdziale 7. sformułowano wnioski oraz podsumowano całość pracy.

2. Problemy, cel i motywacja

2.1. Endoskopia kapsułkowa

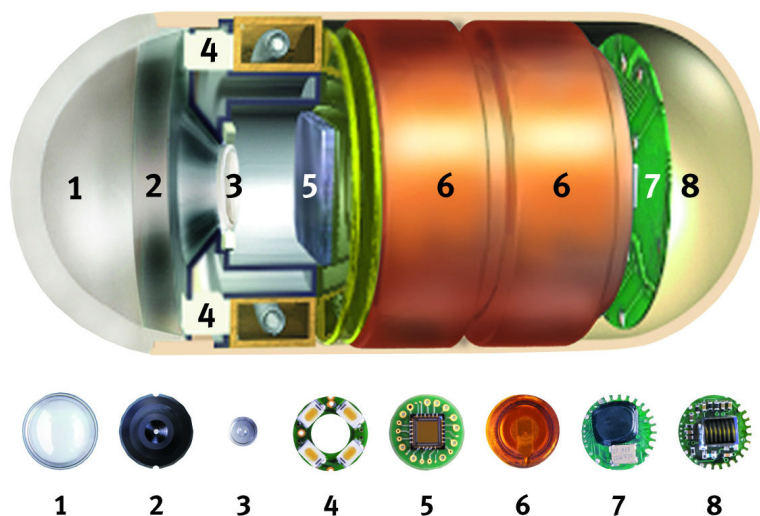
W diagnostyce układu pokarmowego wykorzystuje się metody obrazowania wizyjnego, nazywane ogólnie metodami endoskopowymi. Pierwszym urządzeniem endoskopowym, umożliwiającym takie obrazowanie w organizmach żywych był fiberskop [8]. Elementami fiberskopu są wiązka światłowodów z układami optycznymi, obiektywem na jednym końcu i okulem na drugim oraz układ oświetlający. Jeden kraniec wiązki, z obiektywem, wprowadza się do organizmu pacjenta, wykorzystując do tego naturalne otwory, podczas gdy drugi, z okulem, pozostaje na zewnątrz. Źródło światła umożliwia oświetlenie badanego narządu, a do przesyłania obrazu wnętrza układu pokarmowego na zewnątrz wykorzystuje się wiązkę światłowodów, przy czym poszczególne światłowody umożliwiają przekazanie pojedynczych elementów obrazu. Fiberskop należy do grupy endoskopów wziernikowych (enteroskopów) – urządzeń, których konstrukcja obejmuje półsztywny kabel wpychany do organizmu pacjenta – stąd, używana potocznie w języku angielskim, nazwa *push-endoscopy*.

Miniaturyzacja kamer wizyjnych spowodowała, że w aparatach endoskopowych stosowanie wiązek światłowodów przestało być konieczne – obecnie na końcu przewodu enteroskopu instaluje się miniaturową kamerę. Nowoczesne endoskopy zaopatrzone są również w manipulatory i urządzenia umożliwiające dokonywanie drobnych zabiegów wewnątrz ciała pacjenta, a także w wewnętrzny kanał służący do wprowadzania oraz wyprowadzania powietrza lub płynów.

W przypadku badań układu pokarmowego stosowane są dwie techniki enteroskopowe [8], gastroscopia [26] oraz kolonoskopia [27]. W pierwszej z nich, enteroskop wprowadzany jest przez usta i umożliwia obrazowanie przełyku, żołądka oraz dwunastnicy. W drugiej technice enteroskop wprowadzany jest przez odbyt i umożliwia obrazowanie jelita grubego. Ze względu na ograniczenia techniczne dotyczące długości i giętkości przewodu enteroskopu, klasyczne techniki enteroskopowe nie umożliwiają obrazowania jelita cienkiego. Jak wiadomo, przebieg jelita cienkiego charakteryzuje się wieloma zakrętami, a długość jelita u człowieka wynosi przeciętnie sześć metrów – utrudnia to znacznie obrazowanie jelita cienkiego za pomocą enteroskopu.

Konstrukcja enteroskopu do badania jelita cienkiego została opracowana w pierwszych latach XXI wieku [9], [10], [28]. Enteroskop ten składa się z elastycznego przewodu umieszczonego wewnątrz giętkiej rurki – zarówno na końcu przewodu, jak i na końcu rurki znajdują się woreczki (balony), które mogą być

napełniane sprężonym powietrzem. Napełnienie balonu powietrzem powoduje unieruchomienie końca rurki lub przewodu w wybranym miejscu przewodu pokarmowego. Enteroskop, zwany dwubalonowym, można wprowadzić do jelita cienkiego przez naprzemienne napełnianie i opróżnianie balonów, przy jednoczesnym naprzemiennym wpychaniu rurki lub przewodu do wnętrza ciała. Technika ta umożliwia obserwację wybranych miejsc całego przewodu pokarmowego, ale procedura wizualizacji całości jelita cienkiego jest długotrwała – wielogodzinna. Istotne jest również to, że badanie tą metodą jest uciążliwe dla pacjenta i zazwyczaj wymaga znieczulenia ogólnego. Nie można także bagatelizować faktu zagrożenia przypadkowym uszkodzeniem jelita, co w konsekwencji prowadzi do powikłań [29], [30].



Rys. 2.1. Przekrój kapsułki endoskopu bezprzewodowego typu M2A: 1. przezroczysta kopułka, 2. mocowanie soczewki, 3. soczewka, 4. diody oświetlające, 5. przetwornik optoelektryczny, 6. baterie, 7. układ mikroprocesorowy, 8. antena nadajnika (ilustracja udostępniona przez firmę *GivenImaging*)

Stosunkowo małoinwazyjną techniką obrazowania wizyjnego jelita cienkiego jest bezprzewodowa endoscopia kapsułkowa (ang. *Wireless Capsule Endoscopy* – WCE) [3], [31], [32], znana również jako endoscopia kapsułkowa lub endoscopia bezprzewodowa. Umożliwia ona uzyskanie obrazów wizyjnych całego układu pokarmowego, w tym również jelita cienkiego. Jej niewątpliwą zaletą jest niewielka uciążliwość dla pacjenta, gdyż kapsułka ma małe rozmiary, a jej połknięcie nie jest tak nieprzyjemne jak poddanie się badaniu enteroskopowemu.

Endoskop bezprzewodowy został wprowadzony na rynek przez izraelską firmę *GivenImaging*. Jak już wspomniano, kapsułka została zaprojektowana przez jednego z jej założycieli, którego doświadczenie przy projektowaniu pocisków ra-

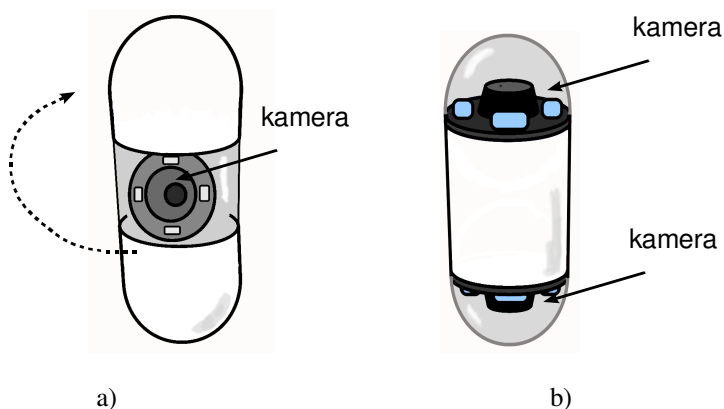
kietowych okazało się bardzo przydatne. W roku 2000 produkt otrzymał aprobatę amerykańskiej agencji FDA [3], [33]. Technologia ta (rysunki 2.1 i 2.2) otrzymała początkowo nazwę handlową *M2A*, potem gamę produktów firmy nazwano *PillCam*, a kapsułkę do badań jelita cienkiego oznaczono skrótem *SB* (ang. *small bowel*). Obecnie podobne produkty dostarczane są przez japońską firmę *Olympus* pod nazwą *EndoCapsule* [34–36], przez koreańską firmę *Intromedic* pod nazwą *MicroCam* [36], [37], a także przez chińską firmę *Chongqingjinshan Science & Technology* pod nazwą *OMOM* [38]. Oryginalną konstrukcję endoskopu kapsułkowego o nazwie *Sayaka* [36], [39–41] (rysunek 2.3a) zaproponowała też japońska firma *RF System*. Prace nad sterowaną kapsułką endoskopową były również prowadzone w Europie w ramach projektu *Vector* [39], [42], [43] (rysunek 2.4) oraz *Nemo* [43], [44].



Rys. 2.2. Elementy systemu endoskopii bezprzewodowej: pas biodrowy, zestaw anten, odbiornik-rejestrator i kapsułka w opakowaniu (ilustracja udostępniona przez firmę *GivenImaging*)

System endoskopii bezprzewodowej obejmuje kapsułkę z wbudowaną kamerą, diodowym źródłem światła białego, nadajnikiem sygnału wizyjnego oraz źródłem zasilania. Ponadto w skład systemu wchodzi odbiornik-rejestrator, system anten odbiorczych montowanych na ciele pacjenta oraz komputer z oprogramowaniem do analizy otrzymanego filmu.

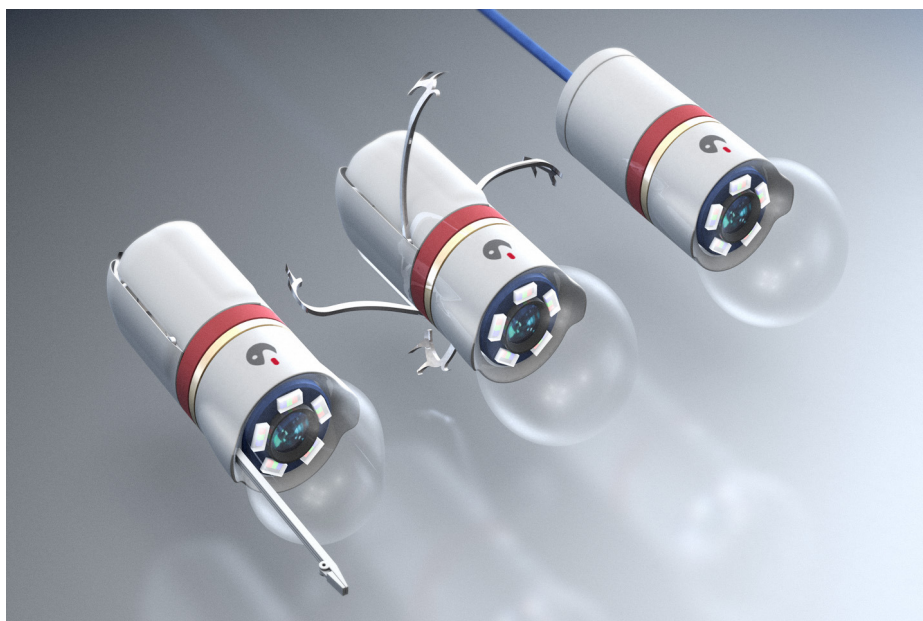
Kapsułka uruchamia się automatycznie w chwili wyjęcia jej z opakowania. Jest ona połykana przez pacjenta, po czym przemieszcza się w przewodzie pokarmowym w sposób naturalny, podobnie jak pokarm, wykorzystując naturalne ruchy układu trawiennego oraz w pewnym stopniu siły grawitacji i bezwładności. W dostępnych na rynku urządzeniach nie stosuje się mechanizmów napędowych, czy też jakiegokolwiek sterowania przemieszczaniem lub orientacją endoskopu, choć jak wspomniano prowadzone są prace nad bezprzewodowymi endoskopami z własnym napędem [45–50]. Należy zauważyć, że mankamentem wykorzystywanych obecnie systemów endoskopii bezprzewodowej jest brak możliwości precyzyjnej lokalizacji kapsułki wewnątrz organizmu pacjenta czy też określenia jej orientacji. Wady te powodują, że diagnosta nie może w sposób czynny uczestniczyć w badaniu i kierować endoskopem bezprzewodowym, tak jak ma to miejsce w przypadku technik wzornikowych. Niemożliwe jest też precyzyjne określenie położenia zmian patologicznych wykrytych za pomocą kapsułki.



Rys. 2.3. Rysunki kapsułek o specyficznych rozwiązaniach konstrukcyjnych:

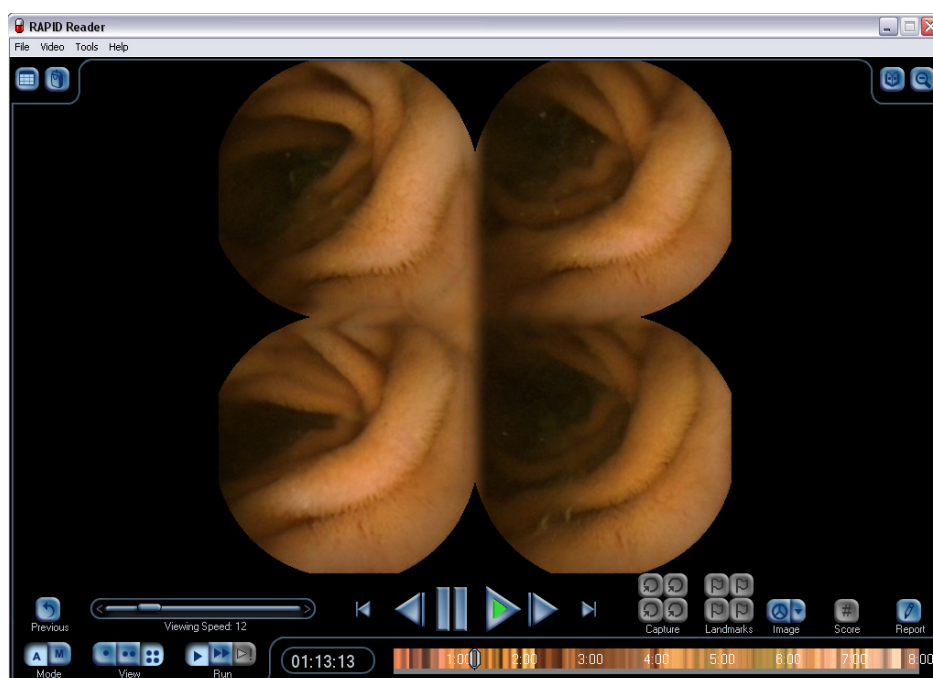
- a) *Sayaka* – z kamerą obrotową skierowaną prostopadłe do osi oraz
- b) *Eso 2* – z dwiema kamerami do obrazowania przetyku

Obrazy wykorzystywane w niniejszej pracy pochodzą z kapsułki typu *Pill-Cam SB*. Kapsułka ma 26 mm długości, 11 mm średnicy i przesyła do noszonego przez pacjenta odbiornika obrazy kolorowe o wielkości 256×256 pikseli z częstotliwością dwóch obrazów na sekundę, zaś pole widzenia kapsułki jest okrągłe o kącie widzenia 140°. Badanie trwa około ośmiu godzin i w tym czasie kapsułka jest w stanie przemieścić się do jelita grubego oraz przesyłać do rejestratora ponad 50 000 obrazów.



Rys. 2.4. Wizja artystyczna kapsułek endoskopowych umożliwiających przeprowadzanie zabiegów wewnątrz przewodu pokarmowego (ilustracja ze strony projektu *Vector* [42])

Po zakończeniu badania obrazy z rejestratora kopiuje się na dysk komputera, gdzie tworzą one ośmiogodzinny film wideo. Interpretacja tego materiału jest zadaniem wykwalifikowanego lekarza gastroenterologa. Wymaga ona obejrzenia całego filmu w celu odnalezienia zmian patologicznych, takich jak krwawienia, erozje wewnętrznej powierzchni jelit, wrzody, polipy, nieprawidłowe przewężenia światła przewodu pokarmowego i tym podobne. Opisany proces diagnostyczny wymaga skupienia uwagi, ponieważ niektóre zmiany mogą być niewielkie, widoczne jedynie w pojedynczych obrazach i łatwo jest je przeoczyć. Bardzo duża liczba zarejestrowanych obrazów sprawia, że interpretacja filmu zajmuje zazwyczaj około dwóch godzin – nawet jeśli odtwarzanie materiału odbywa się częściowo w przyspieszonym tempie. Właściwa ocena materiału wizualnego wymaga czasami kilkukrotnego obejrzenia wybranego fragmentu filmu, zatrzymywania oraz przewijania wstecz. Jest to wyczerpująca i nużąca praca, co w konsekwencji wpływa negatywnie na zdolność koncentracji i może prowadzić do nieścisłej lub nawet błędnej diagnozy. Należy zaznaczyć, że odnalezienie wszystkich zmian chorobowych, nawet tego samego rodzaju, jest istotne i ma wpływ na sposób leczenia, gdyż pojedyncze zmiany mogą być leczone z wykorzystaniem metod enteroskopowych, a liczne lub rozległe wymagają interwencji chirurgicznej.



Rys. 2.5. Okno programu *RapidReader* z włączonym jednoczesnym podglądem czterech kolejnych obrazów filmu

RapidReader [33] (rysunek 2.5) firmy *GivenImaging* jest przykładem programu komercyjnego, który służy do interpretacji danych wizyjnych z endoskopu kapsułkowego. Za jego pomocą możliwe jest odtwarzanie materiału diagnostycznego z różnymi prędkościami w przód i w tył. Program ten jest także wyposażony w pasek zaawansowania, na podstawie którego można oszacować, jaka część filmu została już odtworzona. Wskazując miejsce na pasku zaawansowania można, choć nieprecyzyjnie, wskazać wybraną część filmu, którą chce się obejrzeć. Dodatkową funkcją programu *RapidReader* jest możliwość jednoczesnego wyświetlania na ekranie dwóch lub czterech kolejnych obrazów filmu – ta wielokrotna prezentacja pozwala na dwu- lub czterokrotne wydłużenie czasu percepcji pojedynczego obrazu przy tej samej szybkości odtwarzania. W dokumentacji programu twierdzi się, że ma to wpływ na szybkość uzyskania i jakość wyników interpretacji. Innym narzędziem udostępnianym w programie jest detekcja obrazów o dużym udziale barwy czerwonej, ma to być w założeniu użyteczne w wykrywaniu krwawień wewnątrz przewodu pokarmowego [31], [51–53]. Narzędzie to jednak jest używane przez lekarzy stosunkowo niechętnie i z dużą rezerwą ze względu na znaczną liczbę błędnych pozytywnych detekcji (mała swoistość).

Przykładem niekomercyjnego programu komputerowego do analizy obrazów z endoskopu bezprzewodowego jest CapView [54], [55]. Jego głównym zadaniem jest wskazywanie obrazów przedstawiających wybrane rodzaje zmian chorobowych. Program jest wynikiem prac prowadzonych w ramach projektu badaw-

czego na Uniwersytecie Aveiro w Portugalii, którego celem jest wykorzystanie metod przetwarzania i analizy obrazów do wspomagania interpretacji filmów z endoskopu bezprzewodowego. W ramach projektu jest również tworzona kolekcja obrazów przedstawiających różne zmiany chorobowe wraz z odpowiednimi opisami.

Programy wspomagające interpretację filmów z endoskopu bezprzewodowego umożliwiają tworzenie listy tzw. znalezisk (ang. *findings*). Zadaniem diagnosty jest wprowadzenie na listę identyfikatorów obrazów charakterystycznych miejsc przewodu pokarmowego (np. początku i końca jelita cienkiego), obrazów, w których wykryto zmiany chorobowe, a także syntetycznego opisu tego, co te obrazy przedstawiają. Tak sporządzona lista staje się zazwyczaj elementem raportu z przeprowadzonego badania.

2.2. Niedostatki narzędzi do analizy filmów

Jednym z utrudnień na jakie napotyka lekarz gastroenterolog podczas interpretacji danych obrazowych jest nierównomierność ruchu endoskopu. Ruch ten wywołany jest jedynie przez skurcze przewodu pokarmowego oraz siłę grawitacji, gdyż kapsułka nie ma własnego napędu ani innego układu umożliwiającego sterowanie nią. Sposób jej przemieszczania się jest więc ściśle uzależniony od tego, jak poruszają się jelita. W jelicie cienkim wyróżnić można dwa rodzaje skurczy [56]: perystaltyczne oraz odcinkowe, zwane też segmentowymi. Pierwszy rodzaj odpowiedzialny jest za przesuwanie pokarmu – jest to skurcz obwodowy, przesuwający się w sposób jednostajny w stronę odbytu. Drugi rodzaj skurczu ma bardziej chaotyczny charakter i jest odpowiedzialny za mieszanie pokarmu. Skurcze rozdzielone są w czasie fazami bezruchu – odpoczynku mięśni jelita. W wyniku oddziaływania tego typu bodźców na endoskop bezprzewodowy, uzyskiwane obrazy charakteryzują się zróżnicowaną zmiennością. W fazie bezruchu kolejne obrazy pokazują ten sam fragment powierzchni jelita i praktycznie się nie zmieniają. Dopiero gdy endoskop jest przesuwany przez skurcze perystaltyczne, zmiany obrazu stają się płynne i w sposób jednoznaczny można zaobserwować charakter ruchu kapsułki względem otoczenia. Skurcze mieszające natomiast powodują gwałtowne zmiany w obrazie, kolejne sekwencje obrazów mogą się znacznie różnić między sobą i pokazywać zupełnie inne fragmenty ścian przewodu pokarmowego. Trudno jest wówczas powiązać ze sobą przestrzennie kolejne obrazy i określić rodzaj ruchu, w jakim znajduje się endoskop.

Lekarz dokonujący oceny diagnostycznej materiału stara się przeglądać fragmenty uzyskane w fazie odpoczynku jelita w przyspieszonym tempie, a te fragmenty, które uzyskano podczas skurczy perystaltycznych zazwyczaj ogląda z mniejszą szybkością odtwarzania, natomiast uzyskane podczas skurczy mieszających ogląda uważnie „klatka po klatce”. W przypadku, w którym po fazie bezruchu następują skurcze mieszające, lekarz nie jest w stanie wystarczająco szybko zareagować i musi zatrzymać odtwarzanie, aby przewinąć film wstecz, do chwili, w której następuje gwałtowna zmiana widoku z kamery. Należy przyznać, że ten sposób pracy z materiałem diagnostycznym wymaga ciągłego skupienia uwagi, co w konsekwencji jest bardzo męczące.

Zdarza się także często, że obraz z kapsułki zostaje przesłonięty przez pokarm, przez co obserwacja powierzchni przewodu pokarmowego staje się niemożliwa. W zapisie wideo niektórych badań, długie fragmenty filmu składają się właśnie z takich obrazów i są nieprzydatne z punktu widzenia diagnostyki medycznej. Efektywność interpretowania danych mogłaby wzrosnąć, gdyby możliwe było wykrycie takich fragmentów jeszcze przed przystąpieniem do oglądania filmu. Brakuje jednak narzędzi oceny całości nagrania wideo pod kątem jego jakości lub kompletności. Ocena taka byłaby możliwa w przypadku innej prezentacji danych obrazowych – nie w postaci filmu, ale w postaci pojedynczego obrazu przedstawiającego całość wewnętrznej powierzchni przewodu pokarmowego. Opracowania metody generowania takiego obrazu podjęła się firma *RF System*, projektując kapsułkę endoskopową o nazwie *Sayaka* [41]. W zaproponowanym rozwiązaniu kamera jest skierowana prostopadle i obraca się dookoła głównej osi kapsułki, a obrazy są nakładane na siebie tak, aby uzyskać pojedynczy obraz całości przewodu pokarmowego. Wadą tego rozwiązania jest jednak duża złożoność konstrukcyjna, ruchome elementy oraz konieczność zasilania zewnętrznym polem elektromagnetycznym. Koncepcja tworzenia pojedynczego obrazu przedstawiającego całość powierzchni przewodu pokarmowego jest natomiast warta bliższej uwagi – w dalszej części pracy zostanie przedstawiona metoda generowania podobnego obrazu na podstawie filmu z kapsułki o nieruchomej kamerze.

W wykorzystywanych obecnie programach komputerowych do analizy filmów endoskopowych brakuje także narzędzi analizy ilościowej, nie ma między innymi procedur do automatycznej detekcji zmian chorobowych i zliczania liczby ich wystąpień. W przypadku występowania rozległych zmian chorobowych, obejmujących znaczne powierzchnie, niemożliwe jest obliczenie udziału tych zmian w całkowitej powierzchni przewodu pokarmowego.

2.3. Przegląd kierunków badań

Funkcje programów służących do wspomagania interpretacji filmów z endoskopii bezprzewodowej należy uznać za niewystarczające. Poprawa ich funkcjonalności stanowi wyzwanie, ale jest zadaniem możliwym do zrealizowania. W tym celu konieczne jest opracowanie nowych metod numerycznych do wspomagania analizy filmów uzyskiwanych za pomocą endoskopu bezprzewodowego. Ich wynikiem powinno być skrócenie czasu poświęcanego przez lekarza na interpretację filmu, zmniejszenie prawdopodobieństwa przeoczenia zmian patologicznych, zwiększenie komfortu pracy diagnosty oraz umożliwienie przeprowadzenia analiz ilościowych.

Przegląd publikacji naukowych poświęconych temu zagadnieniu pozwala zauważyć kilka kierunków, w których są prowadzone badania. Jednym z nich jest opracowanie metod klasyfikacji obrazów endoskopowych [52], [57–72] na przedstawiające stan normalny, zmiany chorobowe oraz obrazy nieprzydatne z punktu widzenia diagnostyki medycznej (np. obraz treści pokarmowej przesłaniającej światło i ścianki przewodu pokarmowego). Dzięki temu uwaga gastroenterologa może skupić się głównie na interpretacji obrazów zmian chorobowych. Inne po-

dejsie polega na kategoryzacji obrazów ze względu na ich przynależność do konkretnego odcinka przewodu pokarmowego [55], [73–80] (przełyku, żołądka, dwunastnicy, jelita cienkiego, jelita grubego) i jeśli tylko wiadomo, w którym odcinku przewodu pokarmowego występuje poszukiwana zmiana patologiczna, wówczas oglądanie filmu można ograniczyć do obrazów przedstawiających właśnie ten odcinek. Trzeci kierunek polega na wstępnej analizie filmu pod względem intensywności zmian kolejnych obrazów [11–18], [81–84], co ma związek z nierównomierną prędkością przemieszczania się endoskopu. W tym przypadku w czasie odtwarzania filmu fragmenty, w których obraz się nie zmienia, są odtwarzane z większą szybkością, natomiast fragmenty, w których następują szybkie zmiany treści obrazu, są odtwarzane w zwolnionym tempie. Czwarte podejście zaś polega na przetworzeniu filmu do postaci pojedynczego obrazu stanowiącego odwzorowanie powierzchni całości układu pokarmowego [11–13], [15], [16], [18], [41]. Ocena takiego obrazu zajmuje mniej czasu niż obejrzenie filmu.

Prace nad sposobami wspomagania interpretacji wideo z endoskopii, prowadzone przez niezależne grupy badawcze, dotyczą metod detekcji obrazów przedstawiających wybrane zmiany patologiczne w układzie pokarmowym, algorytmów do segmentacji topograficznej oraz analizy ruchu własnego kapsułki. W pracach [55], [73–75], [85] deskryptory koloru i tekstury opisane standardem MPEG 7 [86–88] zostały wykorzystane do tzw. segmentacji topograficznej, czyli klasyfikacji obrazów ze względu na ich przynależność do określonej sekcji układu pokarmowego. Podobne podejście zaprezentowano w publikacjach [64], [77–80], w których deskryptory koloru i tekstury wykorzystano między innymi do detekcji przejścia kapsuły przez odźwiernik żołądka.

W artykułach [81–83] przedstawiono metodę, w której jako znaczących użyto cech obrazu, takich jak jasność, kontrast oraz wybrane parametry geometryczne obszarów obrazu poddanego segmentacji. Porównanie wektorów tych cech opisujących kolejne obrazy filmu umożliwiło podział jego fragmentów ze względu na fazy ruchu jelit (skurcz, odpoczynek). Wykazano, że fazy skurczu powodują szybsze przemieszczanie się kapsuły wewnątrz układu pokarmowego, a odpowiadające im obrazy wymagają większej uwagi podczas interpretacji.

Metody służące do opisu parametrów ruchu własnego kamery endoskopowej zostały przedstawione w pracach [89–91]. Wykorzystują one równania geometrii epipolarnej w połączeniu z dopasowaniem (tzw. rejestracją) obrazów pochodzących z endoskopu oraz grafiki komputerowej uzyskanej na podstawie obrazów objętościowych tomografii rentgenowskiej lub rezonansu magnetycznego. Wykazano przydatność tej metody w bronchoskopii (endoskopii dróg oddechowych), jednak dla potrzeb opisu ruchu własnego endoskopu kapsułkowego opisywany sposób jest niewystarczający. Wynika to z faktu, że w przeciwieństwie do stosunkowo sztywnych i nieruchomych narządów jakimi są tchawica i oskrzela, dla których utworzenie trójwymiarowego modelu wydaje się łatwe, model elastycznego i giętkiego jelita cienkiego, będącego dodatkowo w ciągłym ruchu, jest w praktyce niewykonalny – choć prace nad takim modelem są prowadzone [92], [93]. Ponadto nawet gdyby było to możliwe, to pacjentów poddawanych badaniu endoskopem bezprzewodowym rzadko bada się dodatkowo z wykorzystaniem

technik obrazowania trójwymiarowego, gdyż takie uzupełniające badanie jest zbyt kosztowne.

Na podstawie przeglądu publikacji można stwierdzić, że mamy do czynienia ze wstępnym stadium badań nad metodami analizy filmów z endoskopu kapsułkowego. Publikacje te wykazują potencjalne możliwości prezentowanych metod, pod warunkiem ich dalszego rozwijania i udoskonalania. Wykazuje się związki pomiędzy badanymi cechami obrazów a występowaniem zmian chorobowych, charakterystycznych miejsc przewodu pokarmowego oraz fazami ruchu jelit. Prezentowane w literaturze wyniki nie dowodzą jednak, że proponowane rozwiązania mogą zostać bezpośrednio wykorzystane w praktyce, a są jednocześnie obciążone dużymi błędami klasyfikacji lub ograniczają się do niewielkiej liczby wybranych przypadków, przez co są statystycznie niewiarygodne. Pomimo, że publikowane materiały nie przedstawiają jeszcze gotowych rozwiązań, to wskazują na potencjalne możliwości metod analizy obrazu w endoskopii bezprzewodowej.

2.4. Motywacja i cel badań

Choroby układu pokarmowego występują stosunkowo często i są poważnym problemem na całym świecie, w tym również w naszym kraju. Powodzenie leczenia chorób nowotworowych jelita, choroby Leśniowskiego-Crohna [94], [95] czy innych chorób wrzodowych układu pokarmowego, jest uzależnione od ich wczesnego rozpoznania – wymaga zatem stosowania nowoczesnych metod obrazowania umożliwiających odnalezienie i identyfikację zmian patologicznych wewnątrz układu pokarmowego. Potrzebne są zwłaszcza metody umożliwiające badania przesiewowe, które byłyby zautomatyzowane i jednocześnie tanie w zastosowaniu.

Metoda badania z wykorzystaniem endoskopu bezprzewodowego jest praktycznie nieinwazyjna, a jeśli porównać ją z klasyczną enteroskopią, to jest ona również znikomo uciążliwa dla pacjenta. Dla wielu objawów wskazujących na potencjalne schorzenia jelita cienkiego endoskopia kapsułkowa jest obecnie zalecana jako podstawowe badanie diagnostyczne w przypadku niewyjaśnionych krwawień wewnątrz przewodu pokarmowego [32]. Pomimo swoich zalet i potencjału diagnostycznego endoskopia kapsułkowa jest mało znana i stosunkowo rzadko stosowana. W Polsce takie badanie nie jest refundowane przez Narodowy Fundusz Zdrowia, a pacjent zmuszony jest zapłacić za nie około trzy tysiące złotych, przy czym znaczny udział w cenie badania ma praca lekarza.

Autor niniejszej rozprawy opracował wcześniej metody wykorzystujące modele deformowalne [23], [24], [96–102] oraz analizę tekstury i barwy obrazów [103–112]. Uzyskiwane wyniki pozwoliły postawić hipotezę, że właśnie te metody umożliwią rozwiązanie problemów występujących podczas oceny diagnostycznej filmów endoskopowych. We wcześniejszych pracach wykazano przydatność opracowanych deformowalnych modeli siatkowych do rozpoznawania obrazów i śledzenia ruchu [23], [24], [96], [97], podobne modele mogą być przydatne w śledzeniu ruchu własnego endoskopu bezprzewodowego. Prace dotyczące analizy tekstury i barwy doprowadziły do powstania oryginalnego pro-

gramu komputerowego [103–106], w którym zaimplementowane procedury służące do obliczania cech obrazów wykraczają poza specyfikację określoną standardem MPEG 7. Program ten z powodzeniem wykorzystano w wielu ośrodkach badawczych do analizy obrazów biomedycznych [113–139] i wykazano, że zaimplementowane w nim metody są użyteczne do wykrywania lub oceny zmian chorobowych. Można więc przyjąć, że będą one równie użyteczne w przypadku analizy obrazów endoskopowych.

Dzięki współpracy z Narodowym Instytutem Standardów i Technologii (USA), Azjatyckim Instytutem Gastroenterologii (Indie) oraz ze Szpitalem Klinicznym im. N. Barlickiego Uniwersytetu Medycznego w Łodzi, autor uzyskał dostęp do ponad 70 filmów będących wynikiem badań przeprowadzonych za pomocą endoskopu bezprzewodowego. Filmy te poddano interpretacji medycznej, a odnalezione obrazy zmian chorobowych zostały oznaczone i opisane przez wyspecjalizowanych diagnostów. Dostęp do takiego materiału umożliwił przeprowadzenie wiarygodnej analizy statystycznej, optymalizację parametrów proponowanych metod, wykonanie testów dla zróżnicowanych, rzeczywistych przypadków oraz walidację zaproponowanych rozwiązań. Współpraca z gastroenterologami z ośrodków medycznych umożliwiła również wykorzystanie opinii lekarzy, uzyskanie oceny eksperta-praktyka oraz testowanie metod w warunkach zbliżonych do klinicznych.

Celem opisanych tu badań było opracowanie metod komputerowej analizy obrazów, użytecznych do wspomagania diagnostyki medycznej przewodu pokarmowego z wykorzystaniem endoskopu bezprzewodowego z pojedynczą, nieruchomą kamerą zwróconą wzdłuż głównej osi kapsułki. Przyjęto, że cel powinien zostać osiągnięty dla danych obrazowych pochodzących z obecnie używanych endoskopów bezprzewodowych, bez wprowadzania w nich zmian konstrukcyjnych i bez konieczności używania innych, wspomagających metod wizualizacji lub lokalizacji.

3. Analiza ruchu endoskopu

3.1. Model ruchu endoskopu

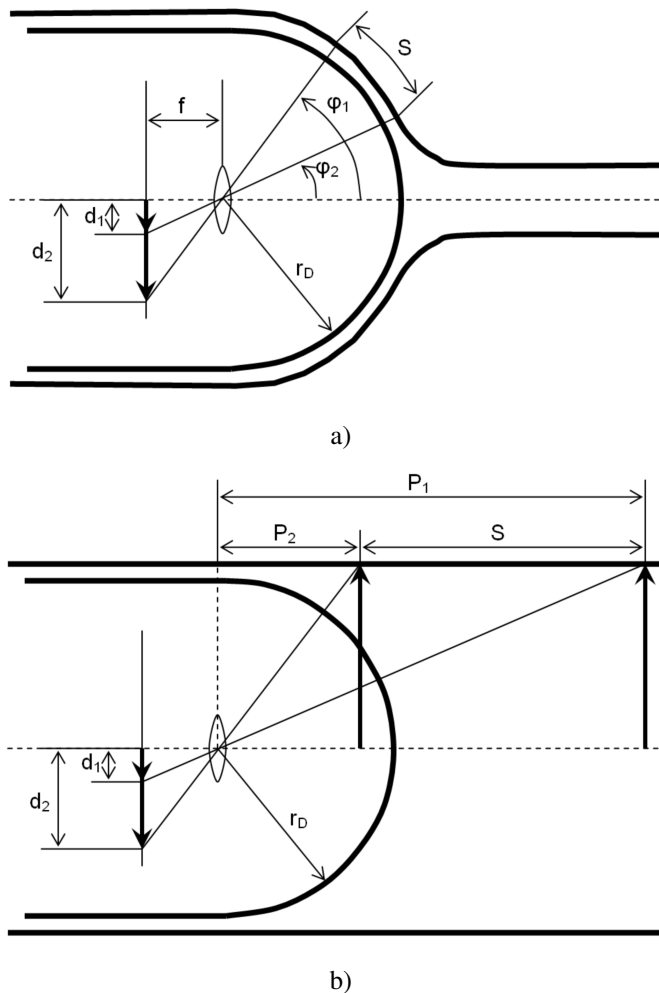
Bezpośrednim wynikiem działania kapsuły endoskopowej jest sekwencja wideo złożona z około 50 000 obrazów. Biorąc pod uwagę, że nie wykorzystuje się innych metod obrazowania, film ten stanowi jedyne źródło informacji o ruchu kapsuły względem ścian przewodu pokarmowego. Określenie sposobu poruszania się kamery na podstawie obrazów rejestrowanych przez tę kamerę należy do grupy zagadnień analizy ruchu własnego (ang. *egomotion*). W klasycznym podejściu do rozwiązania tego zagadnienia [140–142] ruch kamery określany jest przy założeniu, że porusza się ona w przestrzeni trójwymiarowej w nieruchomym środowisku. Kamera ma w takim przypadku sześć stopni swobody ruchu względem tego środowiska i może się przesuwać w trzech kierunkach: przód-tył, lewo-prawo i góra-dół oraz obracać w osi podłużnej i w osiach poprzecznych poziomej i pionowej.

Rozwiązanie problemu ruchu własnego polega na odnalezieniu w dwóch kolejnych obrazach tych samych charakterystycznych punktów sceny. Kiedy znane są właściwości układu optycznego, przesunięcia charakterystycznych punktów sceny w czasie oraz ich odległości od kamery, można sformułować równania, w których niewiadomymi są wektor przesunięcia i kąty obrotu kamery. Rozwiązanie tych równań prowadzi do opisu ruchu kamery względem sceny w sposób ilościowy.

W przypadku endoskopu bezprzewodowego problem estymacji ruchu własnego jest jednak bardziej złożony, ponieważ środowisko, w którym porusza się kamera, nie jest nieruchome. W wyniku skurczów jelit powierzchnia przewodu pokarmowego porusza się i zmienia swój kształt. Nie są też znane odległości pomiędzy punktami sceny i nie jest również możliwe zmierzenie odległości pomiędzy punktem sceny a kamerą. Ponadto należy pamiętać, że obrazy w endoskopie *SB* (ang. *small bowel*) są generowane z częstotliwością dwóch na sekundę, a kolejne obrazy w serii mogą przedstawiać zupełnie inne fragmenty przewodu pokarmowego. Taka sytuacja ma miejsce przede wszystkim podczas ruchów miesających, które sprawiają, że w kolejnych obrazach nie występują te same punkty charakterystyczne.

Rozwiązanie zagadnienia analizy ruchu kapsuły endoskopowej wymaga więc sformułowania dodatkowych założeń. W zaproponowanym modelu ruchu zakłada się, że kapsułka utrzymuje określoną orientację względem otaczającego ją środowiska, z czego wynikają ograniczenia dotyczące stopni swobody jej ruchu.

Ponadto zakłada się, że przewód pokarmowy, w którym porusza się endoskop bezprzewodowy charakteryzuje się określonym kształtem, rozmiarami oraz właściwościami mechanicznymi.



Rys. 3.1. Rzutowanie obrazu w endoskopie bezprzewodowym: a) wewnątrz elastycznego ścisłego przewodu i b) wewnątrz sztywnej rury

Prędkość i orientacja endoskopu zależą od kształtu i ruchów perystaltycznych układu trawiennego. Jednak ze względu na wydłużony kształt kapsułki i stosunkowo małe światło jelita cienkiego, które zaciska się wokół niej, przez większość czasu jej orientacja jest zgodna z kierunkiem przebiegu jelita. Skurcze perystaltyczne powodują przepychanie kapsułki w kierunku zgodnym z przebiegiem przewodu pokarmowego; biorąc pod uwagę konstrukcję endoskopu również oś kamery jest z nim zgodna. W rzeczywistości, w wyniku pofałdowania powierzchni przewodu pokarmowego i ruchów mieszających, endoskop może się ustawić

skośnie lub poprzecznie do kierunku ruchu. Można jednak przyjąć, że taka orientacja jest krótkotrwała i występuje stosunkowo rzadko.

Ograniczenia co do orientacji kapsułki względem przewodu pokarmowego powodują, że ruch obserwowany w obrazach z endoskopu bezprzewodowego ma specyficzny charakter. Gdy kamera endoskopu zwrócona jest zgodnie ze zwrotem wektora jej ruchu, to ściany przewodu pokarmowego w obrazie przesuwają się od jego środka na zewnątrz. W odwrotnym przypadku ściany w obrazie przesuwają się od zewnątrz do jego środka. Oznacza to, że szybkość przemieszczania się kapsułki endoskopowej można określić na podstawie przemieszczeń treści obrazu wzdłuż półprostych o początkach w środku obrazu. Dla porównania, podobne rozwiązania związane z ograniczeniem stopni swobody kamery można znaleźć w publikacjach [143], [144].

Powiązanie prędkości przemieszczania się endoskopu z ruchem w obrazie wymaga przyjęcia jednego z dwóch proponowanych modeli geometrii przewodu pokarmowego [145]. W pierwszym z nich przyjmuje się, że ściany przewodu pokarmowego ściśle przylegają do obudowy kapsułki, w tym również do przezroczystej kopułki osłaniającej kamerę. W takim przypadku (rys. 3.1a), ponieważ kopułka jest sferyczna, to odległość S , o jaką przesuwa się kapsułka, jest proporcjonalna do różnicy kąta φ , pod jakim widoczny jest wybrany punkt otaczającego środowiska w kolejnych obrazach. Obraz tego punktu na powierzchni rzutowania (przetwornika optoelektrycznego) tworzy się w odległości d od osi układu optycznego kamery (środku przetwornika). Parametr d jest powiązany z kątem φ zależnością $d = f \tan(\varphi)$, gdzie f jest ogniskową obiektywu kamery. Stąd prędkość kapsułki można przybliżyć równaniem (3.1.1).

$$V_D = \frac{r_D}{\Delta t} \left(\tan^{-1} \left(\frac{d_{t-\Delta t}}{f} \right) - \tan^{-1} \left(\frac{d_t}{f} \right) \right) \quad (3.1.1)$$

Wielkość d jest odległością obrazu punktu przewodu pokarmowego mierzoną od środka powierzchni przetwornika w chwili t . Wielkość $d_{t-\Delta t}$ jest odległością obrazu tego samego punktu mierzoną w czasie $t - \Delta t$. Różnica Δt jest czasem pomiędzy uzyskaniem dwóch kolejnych obrazów i w przypadku kapsułki wykorzystywanej w badaniach wynosi 0,5 s.

W drugim modelu zastosowano równania rzutowania perspektywicznego przy założeniu, że ściany przewodu pokarmowego są sztywną rurą o okrągłym przekroju i znanej średnicy. Przyjęto, że średnica przewodu jest równa średnicy ($2r_D$) kapsułki endoskopu. Odległość S kapsułki od punktu na powierzchni przewodu pokarmowego jest w tym przypadku odwrotnie proporcjonalna do odległości obrazu tego punktu od środka powierzchni przetwornika. Prędkość kapsułki można zatem estymować za pomocą równania (3.1.2).

$$V_P = \frac{r_D f}{\Delta t} \left(\frac{1}{d_t} - \frac{1}{d_{t-\Delta t}} \right) \quad (3.1.2)$$

Z technicznego punktu widzenia problemem jest nieznajomość rozdzielczości przetwornika optoelektrycznego oraz ogniskowej obiektywu. Producent nie udostępnia niestety szczegółowych danych dotyczących parametrów układu

optycznego kapsułki. Jednakże na podstawie ograniczonych informacji dotyczących kąta widzenia γ oraz średnicy obrazu D można oszacować, że:

$$f = \frac{D}{\tan(\gamma/2)} \quad (3.1.3)$$

Należy podkreślić, że obydwa uproszczenia pozwalają jedynie na zgrubne oszacowanie prędkości endoskopu bezprzewodowego wewnątrz układu pokarmowego. Zastosowane uproszczenia są kompromisem pomiędzy dokładnością obliczeń a prostotą modelu ruchu. Wynikają one w dużym stopniu z braku informacji potrzebnych do pełnej analizy ruchu własnego, np. informacji o rzeczywistej odległości punktu sceny od kamery.

3.2. Modele deformowalne

Tradycyjne techniki przetwarzania obrazów wymagają na ogół, aby wyniki jednej procedury stanowiły dane wejściowe następnej. Stosunkowo proste techniki i algorytmy łączone są kolejno, tworząc bardziej skomplikowane systemy przetwarzania i analizy obrazów. Przykładowo problem śledzenia ruchu lub dopasowywania do siebie treści dwu podobnych obrazów (ang. *image registration*) tradycyjnie rozwiązuje się, dzieląc go na dwa mniejsze zagadnienia rozwiązywane kolejno.

Pierwsze z nich polega na obliczeniu pola ruchu (ang. *motion field*) lub przepływu optycznego (ang. *optical flow*) [146]. Pole ruchu oznacza rzut wektorów ruchu w przestrzeni trójwymiarowej na płaszczyznę obrazu. Przepływ optyczny z kolei oznacza pole wektorowe określające przesunięcia treści dwóch obrazów między sobą i umożliwiające przekształcenie jednego obrazu w drugi. Z punktu widzenia przetwarzania obrazów oba określenia są zbliżone i często przyjmuje się, że pole ruchu można estymować przepływem optycznym. Obliczenie przepływu optycznego polega na wyznaczeniu w jednym z obrazów charakterystycznych miejsc, a następnie odnalezieniu punktów podobnych do nich w drugim obrazie. Poszukiwane wektory przepływu określają przesunięcie między odpowiadającymi sobie miejscami w jednym i drugim obrazie. Istnieje szereg metod [146–154] określających sposób wyboru takich miejsc oraz mierzenia podobieństwa między nimi. Stosuje się je zależnie od konkretnego zagadnienia, wymaganej dokładności czy rodzaju obrazów. Istotne jest, że w otrzymanym polu wektorowym występować mogą błędy, które wynikają z niewłaściwego przypisania do siebie punktów z jednego i drugiego obrazu. Może to być spowodowane zakłóceniami tworzącymi przypadkowo podobne wzory w obu obrazach lub tym, że w obrazie występują powtarzające się fragmenty. Przykładowo, różne choć identyczne okna jadącego autobusu lub pociągu mogą zostać uznane w kolejnych obrazach za to samo okno.

Drugim etapem analizy obrazu, po obliczeniu przepływu optycznego, jest uwzględnienie relacji przestrzennych w obrazie. Wykorzystuje się tu pewną wiedzę wstępną o tym jakiego rodzaju przesunięcia w obrazie są możliwe lub oczekiwane. W przypadku sztywnych obiektów przyjmuje się np. że ich ruch można opisać równaniem transformacji afinicznej, gdzie na podstawie pola wek-

torowego estymowane są macierz transformacji i wektor przesunięcia. Należy zwrócić uwagę, że przy sekwencyjnym stosowaniu obu metod błędy obliczenia pola wektorowego mają istotny udział w obliczonych parametrach transformacji, a te nie są wykorzystywane do korygowania błędów podczas obliczania pola wektorowego.

Analiza obrazów obejmuje często zagadnienia i problemy, których rozwiązanie wymaga zastosowania równolegle kilku tradycyjnych sposobów przetwarzania. Takim przypadkiem jest omawiana diagnostyka obrazowa, która wymaga zastosowania kilku algorytmów przetwarzania i analizy obrazów. Modele deformowalne stanowią właśnie jedno z możliwych podejść do łączenia odrębnych metod, które funkcjonując łącznie uzupełniają się i korygują efekty swoich działań. Cechą wspólną i wyróżnikiem modeli deformowalnych jest między innymi umożliwienie połączenia wybranego algorytmu przetwarzania obrazu oraz metody narzucającej ograniczenia wynikające z wiedzy o zależnościach przestrzennych występujących w tym obrazie. Oba elementy działają w modelu współbieżnie i wpływają na siebie wzajemnie w celu korygowania błędów. Podejście takie prowadzi do powstania jakościowo nowych rozwiązań o zwiększonej odporności na zakłócenia oraz mogące się pojawiać błędy (ang. *robustness*).

Ideę i sposób działania modeli deformowalnych najprościej można przedstawić na przykładzie aktywnego konturu [21], [100–102], [155–164]. Model ten jest przykładem połączenia algorytmu obliczania gradientu jasności obrazu z wiedzą o tym, że kontur jest ciągłą i gładką krzywą. W matematycznym opisie aktywny kontur jest krzywą daną funkcją parametryczną $\mathbf{v}(s) = [x(s), y(s)]^T$, gdzie $\mathbf{v}$ jest wektorem współrzędnych punktów krzywej, a $s \in \langle 0, 1 \rangle$ jest parametrem. Aktywnemu konturowi przypisuje się pewien funkcjonał nazywany energią wewnętrzną (3.2.1). Ta energia jest całką lokalnych gęstości energii liczoną wzdłuż konturu. Gęstość energii w każdym punkcie konturu jest sumą dwóch składników. Pierwszy, E_e (tzw. energia zewnętrzna), reprezentuje udział tradycyjnego algorytmu przetwarzania obrazu. Drugi natomiast, E_i (tzw. energia wewnętrzna), reprezentuje wpływ narzuconych ograniczeń wynikających z zależności przestrzennych.

$$E = \int_0^1 (E_e(s) + E_i(s)) ds \quad (3.2.1)$$

W oryginalnym modelu aktywnego konturu [21] energia E_e (3.2.2) jest proporcjonalna do wartości bezwzględnej gradientu jasności $I(\cdot)$ obrazu w punkcie $\mathbf{v}(s)$ ze znakiem minus. Energia E_i (3.2.3) jest obliczana z wykorzystaniem równania energii potencjalnej struny. Współczynniki ρ , τ i ζ są parametrami, parametr ρ określa odporność modelu na zginanie (sprężystość), parametr τ – odporność na rozciąganie (rozciągliwość), natomiast ζ – siłę, z jaką oddziałuje obraz.

$$E_e(s) = -\zeta \nabla I(\mathbf{v}(s)) \quad (3.2.2)$$

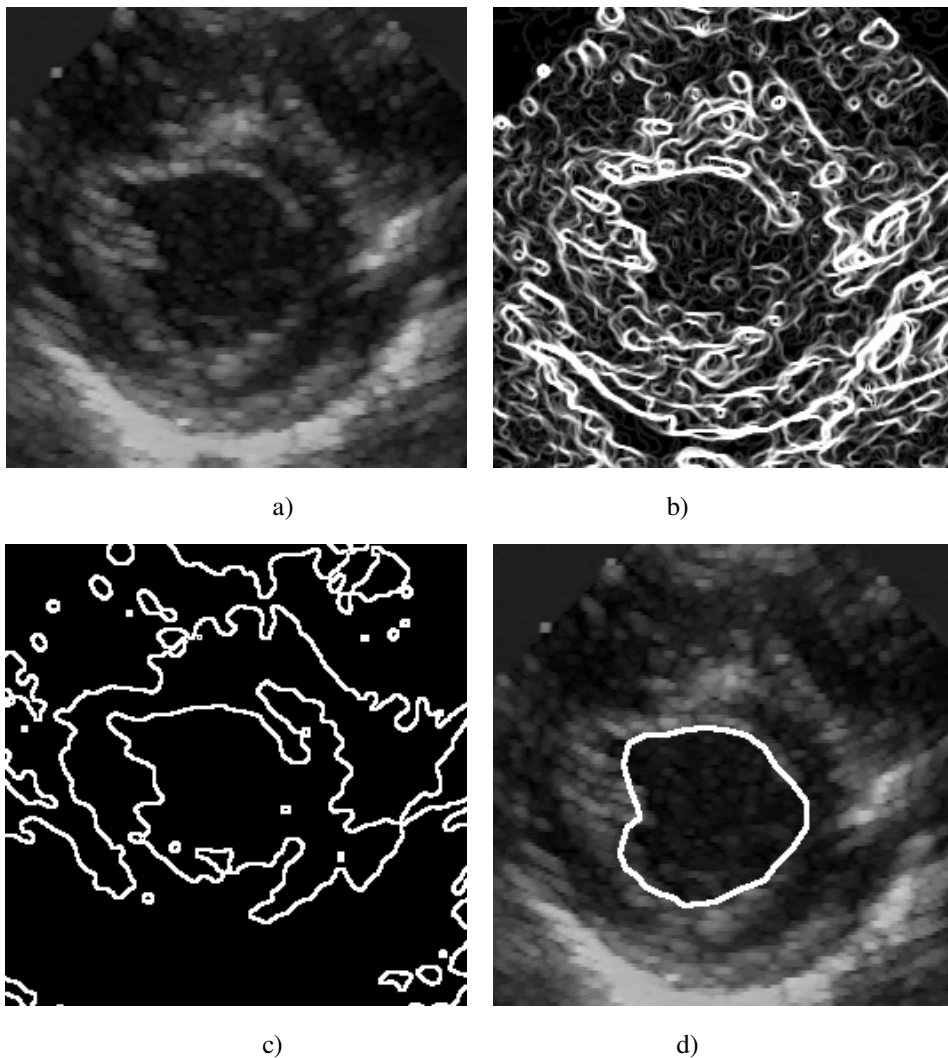
$$E_i(s) = \tau \left| \frac{d\mathbf{v}}{ds} \right|^2 + \rho \left| \frac{d^2\mathbf{v}}{ds^2} \right|^2 \quad (3.2.3)$$

Poszukiwanie rozwiązania, w tym przypadku konturu w obrazie, polega na minimalizacji wartości funkcjonału (3.2.1) przez modyfikowanie funkcji $\mathbf{v}(s)$ współrzędnych x i y punktów krzywej.

W praktyce problem jest rozwiązywany numerycznie, co wymaga dyskretyzacji parametru s . Wówczas całka równania (3.2.1) zostaje zmieniona w sumę energii cząstkowych w dyskretnych węzłach modelu. Funkcjonał staje się funkcją wielu zmiennych, gdzie zmiennymi są współrzędne poszczególnych punktów zalamania krzywej. Rozwiązanie problemu polega na zmianie tych współrzędnych, w taki sposób, aby zminimalizować wartość funkcji. Jednym ze sposobów minimalizacji funkcji jest zastosowanie **optymalizacji gradientowej**.

Analizując równania aktywnego konturu można zauważyć, że mają one następujące właściwości. Minimalizacja wielkości E_i , a więc wartości bezwzględnej pierwszej i drugiej pochodnej współrzędnych po parametrze s , powoduje wygładzanie przebiegu konturu. Składnik ten przeciwdziała zginaniu i rozciąganiu konturu. Minimalizacja składnika E_e powoduje lokalne przemieszczanie się punktów konturu (węzłów) do miejsc o dużej wartości bezwzględnej gradientu jasności obrazu.

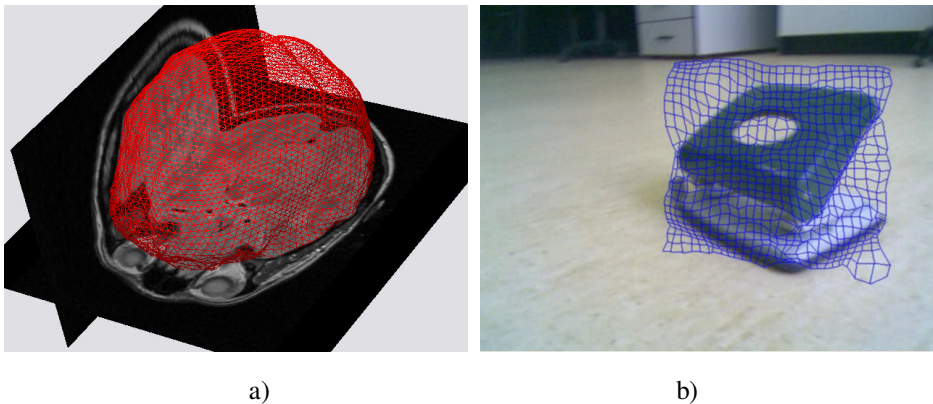
Optymalizacja gradientowa jest metodą iteracyjną – wymaga ona zainicjowania konturu, czyli arbitralnego wyboru jego wstępnego kształtu i położenia. W kolejnych krokach uzyskuje się rozwiązania coraz bliższe lokalnego minimum funkcji. Metoda ta wymaga obliczenia gradientu minimalizowanej funkcji w każdej iteracji – jest to wielowymiarowy wektor w przestrzeni współrzędnych wszystkich węzłów modelu. Rzut tego wektora na podprzestrzeń współrzędnych pojedynczego węzła można interpretować jako siłę oddziałującą na ten węzeł i wywołującą jego ruch. Wizualizacja kolejnych etapów działania modelu deformowalnego pokazuje, że przemieszcza się on i zmienia kształt pod wpływem sił oddziałujących na poszczególne węzły. Model taki, zmieniający swój kształt w czasie, nazywany jest **dynamicznym**. Zmiana kształtu i położenia modelu w wyniku minimalizacji funkcji energii będzie dalej nazywana **dopasowaniem**.



Rys. 3.2. Przykłady prób oznaczania konturu lewej komory serca w obrazie ultrasonograficznym: a) obraz oryginalny, b) wartość bezwzględna gradientu jasności obliczona za pomocą operatora Sobela, c) kontury uzyskane po zastosowaniu filtracji medianowej, progowania jasności i obliczania wartości bezwzględnej gradientu jasności oraz d) model aktywnego konturu

Właściwością modelu dynamicznego jest to, że w wyniku optymalizacji znajdowane jest lokalne minimum energii. Minimum to występuje w przypadku, w którym siły będące wynikiem obliczania gradientu wielkości E_i oraz E_e równoważą się w każdym węźle modelu. W modelu dynamicznym kształt konturu zmienia się więc pod wpływem sił oddziaływania obrazu (zewnętrznego pola) oraz sił oddziaływań wewnętrznych nazywanych **naprężeniami**. Wynikowy kształt jest więc osiągany po zrównoważeniu się tych sił i zależy między innymi od początkowego kształtu modelu. Przykładem zastosowania aktywnego konturu

jest segmentacja obrazów ultrasonograficznych (rysunek 3.2) w celu wyznaczenia obszaru przekroju komór serca. Wykorzystanie klasycznych algorytmów, takich jak transformacja za pomocą operatora gradientowego, filtracji i progowania jasności (rysunek 3.2b i 3.2c) nie prowadzi do uzyskania jednospójnego obszaru, tak jak ma to miejsce w przypadku zastosowania aktywnego konturu (rysunek 3.2d).



Rys. 3.3. Interpretacja graficzna modeli deformowalnych: a) powierzchnia do segmentacji obrazów trójwymiarowych i b) płaska sieć do śledzenia ruchu w obrazach wideo

Modele deformowalne różnią się między sobą pod wieloma względami. Pierwszym z wyróżników jest ich zastosowanie i wynikające stąd wymagania dotyczące wymiaru przestrzeni, w której są zdefiniowane oraz wymiaru przestrzeni parametrów. Modele konturów są narzędziami wykorzystywanymi głównie do segmentacji obrazów, operują w przestrzeni dwuwymiarowego obrazu i są opisywane jednym parametrem. Deformowalna powierzchnia zamknięta [102], [165–171] jest uogólnieniem modelu konturu na przypadek przestrzeni trójwymiarowej. Powierzchnia taka jest opisana funkcją parametryczną dwóch parametrów. Jej zastosowaniem jest np. segmentacja obrazów trójwymiarowych rezonansu magnetycznego, tomografii komputerowej lub ultrasonografii trójwymiarowej. Modelem w pewnym sensie pośrednim między konturem a powierzchnią jest model wycinka płaszczyzny [96–99], [172–180]. Opisany jest on parą parametrów, podobnie jak w przypadku deformowalnej powierzchni, a operuje w przestrzeni dwuwymiarowego obrazu płaskiego, tak jak aktywny kontur. Po dyskretyzacji parametrów model wycinka płaszczyzny można przedstawić w postaci płaskiej sieci połączonych węzłów, umieszczonej na powierzchni obrazu. Modele tego rodzaju są wykorzystywane do śledzenia ruchu lub dopasowywania do siebie treści dwóch podobnych obrazów.

W przypadku modeli konturu oraz deformowalnej powierzchni opracowano modele nieparametryczne wykorzystujące tzw. metodę poziomicy [22], [93], [181–192] (ang. *level set method*). W metodzie tej jest wykorzystywana funkcja, nazywana poziomiciową, której argumentami są współrzędne analizowanego obrazu. Funkcja ta ewoluuje – zmienia się, pod wpływem obrazu oraz oddziaływań wygładzających. Wynikowy kontur (w przestrzeni dwuwymiarowej) lub powierzchnię (w przestrzeni trójwymiarowej) wyznaczają miejsca zerowe tej

funkcji. W odróżnieniu od modelu parametrycznego, zachowującego topologię, modele wykorzystujące metodę poziomic mogą topologię modyfikować. W czasie ewolucji może się zmienić liczba konturów i mogą pojawić się kontury zamknięte wewnątrz innych.

W przypadku modeli parametrycznych dyskretyzowane są parametry modelu. W przypadku modelu niesparametryzowanego, dyskretyzacja dokonywana jest w dziedzinie funkcji poziomicowej, zazwyczaj zgodnie z rastrem analizowanego obrazu.

Innym stosowanym podziałem modeli konturów i powierzchni, zarówno parametrycznych, jak i nieparametrycznych, jest podział na modele brzegowo zależne (ang. *edge-based*) i obszarowo zależne (ang. *region-based*). Pierwsza grupa wykorzystuje do sterowania kształtem modelu cechy charakteryzujące granice między obszarami (np. gradient jasności), w drugiej grupie zmiana kształtu jest wywoływana przez cechy wyróżniające obszary obrazu (np. poziom jasności obrazu).

Cechą wspólną wszystkich wymienionych kategorii modeli deformowalnych jest połączenie oddziaływania lokalnych własności obrazu oraz narzuconych własności modelu. Właściwość ta może być przydatna w analizie ruchu endoskopu bezprzewodowego, gdyż pozwala wprowadzić do procesu analizy ograniczenia wynikające z przyjętego modelu ruchu.

3.3. Koncepcja deformowalnych pierścieni

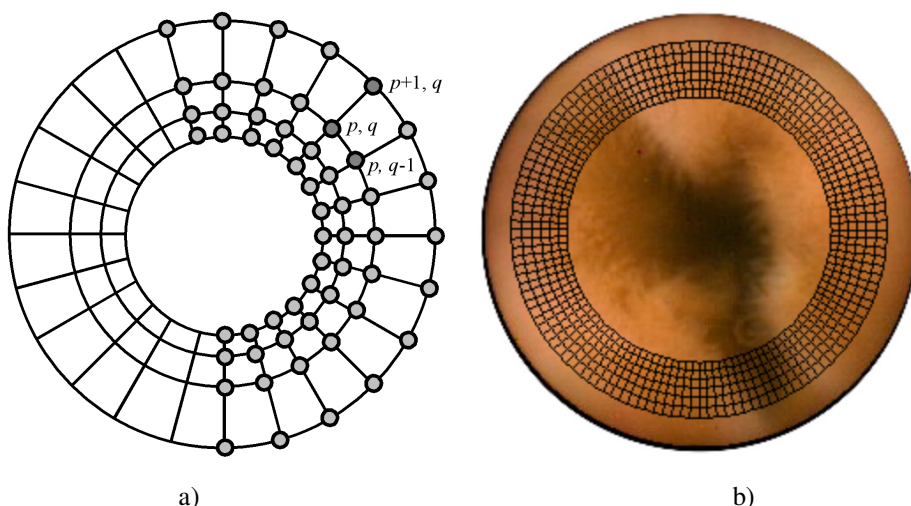
Rozwiązaniem pozwalającym połączyć metody obliczania przepływu optycznego, ograniczenia przestrzenne i zaproponowany model ruchu endoskopu bezprzewodowego, jest opracowany model deformowalny o kształcie koncentrycznych pierścieni (ang. *model of deformable rings* – MDR). Model ten ma za zadanie wstępne przetworzenie danych wideo z endoskopu i wygenerowanie dodatkowej informacji, która wspomogłaby ich interpretację. Jest to model dyskretny wycinka płaszczyzny z dwoma parametrami, działający w przestrzeni płaskiego obrazu. Model ten można przedstawić w postaci płaskiej sieci łączonych węzłów, której oryginalną cechą jest wzajemne rozmieszczenie węzłów oraz sposób ich łączenia.

W zaproponowanym rozwiązaniu węzły określone parą indeksów (dyskretnych parametrów) $p = 1, 2, \dots, P$ i $q = 1, 2, \dots, Q$ wraz z połączeniami tworzą sieć (graf planarny) o kształcie koncentrycznych pierścieni (rys. 3.4a), zdefiniowaną w dwuwymiarowej przestrzeni analizowanego obrazu. Współrzędne węzłów po zainicjowaniu modelu dane są równaniem (3.3.1), parametr r występujący w tym równaniu jest promieniem wewnętrznego pierścienia, zaś parametr w jest stosunkiem promieni dwóch sąsiednich pierścieni (większego do mniejszego).

$$\hat{\mathbf{v}}_{pq} = \begin{bmatrix} \hat{x}_{pq} & \hat{y}_{pq} \end{bmatrix}^T = r w^{p-1} \begin{bmatrix} \cos\left(\frac{2\pi q}{Q}\right) & \sin\left(\frac{2\pi q}{Q}\right) \end{bmatrix}^T \quad (3.3.1)$$

W pracy przyjęto, że znak $\hat{}$ (daszek) będzie oznaczać wektory i wartości współrzędnych węzłów po zainicjowaniu modelu, przed rozpoczęciem procesu dopasowania. Są to współrzędne węzłów w modelu nieznieskształconym jeszcze

oddziaływaniem obrazu. Przyjęto także, że początek układu współrzędnych znajduje się w środku analizowanego obrazu. Parametry r i w są tak dobrane, aby zmieścić model wewnątrz pola widzenia kamery endoskopu z zachowaniem pewnego marginesu dookoła (rys. 3.4b). Potrzeba zapewnienia tego marginesu wynika stąd, że w czasie procesu dopasowania, punkty węzłowe mogą zmieniać swoje położenie, siatka może ulegać deformacjom i może się też powiększać.



Rys. 3.4. Budowa modelu złożonego z deformowalnych pierścieni: a) schemat połączeń węzłów oraz ich indeksowanie, b) model po zainicjowaniu na powierzchni analizowanego obrazu

Podstawową zasadę działania modelu można opisać w następujący sposób: dla każdego węzła są zapamiętywane lokalne charakterystyki obrazu występujące w jego aktualnym położeniu. Po zmianie obrazu na kolejny, węzły poszukują takich miejsc, których charakterystyki są najbardziej zbliżone do zapamiętanych. W wyniku tych poszukiwań węzły przesuwają się w stronę tych miejsc, przy czym ich ruch jest do pewnego stopnia ograniczony przez naprężenia – narzucone ograniczenia dotyczące zmiany kształtu siatki. Wraz z zakończeniem ruchu węzłów są zapamiętywane lokalne charakterystyki aktualnego obrazu i procedura zostaje powtórzona dla kolejnych obrazów sekwencji wideo.

Ruch punktów węzłowych jest wynikiem minimalizacji energii danej równaniem (3.3.2). Energia ta jest zdefiniowana dla modelu z dyskretnymi parametrami (indeksami) p i q – nie występuje w nim więc całka powierzchniowa, lecz sumy składowych energii cząstkowych poszczególnych węzłów oznaczonych parą tych indeksów. Składnikami energii cząstkowych są energie wpływu obrazu na położenie węzła M_{pq} oraz energie naprężeń (korekcji kształtu) E_{pq} .

$$E_{MDR} = \sum_{p=1}^P \sum_{q=1}^Q (M_{pq} + E_{pq}) \quad (3.3.2)$$

Energia MDR jest funkcją współrzędnych wszystkich punktów węzłowych modelu. Model znajduje się w przestrzeni dwuwymiarowej, więc liczba zmien-

nych tej funkcji jest dwukrotnie większa od liczby wszystkich punktów węzłowych i jest równa $2PQ$.

Dopasowanie modelu polega na takiej zmianie współrzędnych x i y wszystkich węzłów $p q$, aby zminimalizować wartość funkcji celu E_{MDR} . Do optymalizacji tej funkcji można zastosować minimalizację gradientową. W metodzie tej, dla określonego wektora zmiennych $\mathbf{X}^{(i)}$ (wektor zawiera współrzędne wszystkich węzłów) jest obliczany gradient funkcji. Indeks (i) oznacza numer kolejnego przybliżenia (kroku, iteracji) procedury optymalizacji. Następnie od wektora $\mathbf{X}^{(i)}$ jest odejmowana wartość proporcjonalna do obliczonego wektora gradientu. Dla wynikowego wektora $\mathbf{X}^{(i+1)}$ ponownie jest obliczany gradient minimalizowanej funkcji. Procedura ta jest powtarzana wielokrotnie, aż do uzyskania odpowiednio małej wartości bezwzględnej wektora gradientu lub do wykonania założonej liczby powtórzeń. Dla funkcji E_{MDR} można zapisać równanie (3.3.3) modyfikujące wektor jej zmiennych w celu minimalizacji jej wartości.

$$\mathbf{X}^{(i+1)} = \mathbf{X}^{(i)} - \nabla E_{MDR}(\mathbf{X}^{(i)}) \quad (3.3.3)$$

$$\nabla E_{MDR} = \nabla \left(\sum_{p=1}^P \sum_{q=1}^Q (M_{pq} + E_{pq}) \right) = \left(\sum_{p=1}^P \sum_{q=1}^Q (\nabla M_{pq} + \nabla E_{pq}) \right) \quad (3.3.4)$$

W celu skrócenia czasu koniecznego do przeprowadzenia powyższej minimalizacji przyjęto uproszczenie dotyczące przybliżonego wyznaczania gradientu funkcji. Można zauważyć, że gradient wielkości M_{pq} zależy wyłącznie od położenia węzła $p q$ i nie zależy od położenia pozostałych węzłów. Wpływ obrazu na węzły może być więc obliczony indywidualnie dla każdego węzła. Wartość gradientu E_{pq} zależy natomiast nie tylko od współrzędnych punktu węzłowego p, q , ale również od współrzędnych innych węzłów, z którymi węzeł $p q$ pozostaje w interakcji. Można jednak przyjąć, że istotny wpływ na wektor gradientu energii E_{pq} mają wyłącznie współrzędne węzła $p q$, a współrzędne węzłów pozostających z nim w interakcji mają wpływ możliwy do pominięcia. Przy takim uproszczeniu równanie gradientu funkcji energii można zapisać w postaci (3.3.5). Symbolami x_{pq}^o oraz y_{pq}^o oznaczono wersory w kierunkach wyznaczonych przez osie odpowiednich współrzędnych.

$$\nabla E_{MDR} \approx \sum_{p=1}^P \sum_{q=1}^Q \left(x_{pq}^o \frac{\partial M_{pq}}{\partial x_{pq}} + y_{pq}^o \frac{\partial M_{pq}}{\partial y_{pq}} + x_{pq}^o \frac{\partial E_{pq}}{\partial x_{pq}} + y_{pq}^o \frac{\partial E_{pq}}{\partial y_{pq}} \right) \quad (3.3.5)$$

Ta uproszczona formuła, w której zmienne będące współrzędnymi węzła p, q są modyfikowane wyłącznie na podstawie cząstkowych pochodnych wielkości M_{pq} i E_{pq} , zależnych – jak się uznaje – wyłącznie od współrzędnych węzła $p q$, pozwala wielokrotnie skrócić czas potrzebny do obliczenia gradientu funkcji E_{MDR} . Podczas prowadzonych eksperymentów nie stwierdzono, aby uproszczenie to miało istotny wpływ na zbieżność procesu optymalizacji.

3.4. Wpływ obrazu na MDR

Podstawowym zadaniem deformowalnych pierścieni jest podążanie za ruchem ścian przewodu pokarmowego, a to z kolei jest powodowane przez składnik M_{pq} funkcji celu E_{MDR} . Wybór odpowiedniej formuły energii cząstkowej M_{pq} oraz dobór jej udziału w E_{MDR} ma więc istotny wpływ na sposób działania całego modelu.

Jak już wspomniano, celem modelu jest umieszczenie węzłów w miejscach aktualnie analizowanego obrazu, które są podobne do miejsc odnalezionych w obrazie poprzednim. Oznacza to konieczność wykorzystania miary podobieństwa lub różnicy między porównywanymi fragmentami obrazów. W przypadku MDR rozważano wykorzystanie kilku takich miar, wcześniej wykorzystywanych w zastosowaniach medycznych do dopasowywania obrazów [151], [152], [193–201] oraz w kompresji wideo do kompensacji ruchu [202–209], m.in.: znormalizowaną korelację wzajemną (ang. *normalized cross-correlation* – NCC) [210–212], informację wzajemną (ang. *mutual information* – MI) [152], [197], [213], błąd średniokwadratowy (ang. *mean square error* – MSE) [214] oraz średnią różnic bezwzględnych (ang. *mean absolute difference* – MAD) [207–209], [215].

Miary MAD i MSE stosowane do kompensacji ruchu określają różnicę pomiędzy dwoma blokami $N \times N$ pikseli. Pierwszy z bloków jest konstruowany na podstawie obrazu uzyskanego w czasie $t - \Delta t$ we współrzędnych $[x_{t-\Delta t}, y_{t-\Delta t}]^T$, drugi zaś powstaje z użyciem obrazu uzyskanego w czasie t we współrzędnych $[x_t, y_t]^T$. Miary MAD i MSE są zdefiniowane odpowiednio równaniami (3.4.1) i (3.4.2), gdzie N jest liczbą naturalną nieparzystą, a funkcja $I(\cdot)$ reprezentuje jasność obrazu we współrzędnych x, y .

$$\text{MAD}(x, y) = \frac{1}{N^2} \sum_{m=-(N-1)/2}^{(N-1)/2} \sum_{n=-(N-1)/2}^{(N-1)/2} \left| I_t(x+m, y+n) - I_{t-\Delta t}(x_{t-\Delta t}+m, y_{t-\Delta t}+n) \right| \quad (3.4.1)$$

$$\text{MSE}(x, y) = \frac{1}{N^2} \sum_{m=-(N-1)/2}^{(N-1)/2} \sum_{n=-(N-1)/2}^{(N-1)/2} \left(I_t(x+m, y+n) - I_{t-\Delta t}(x_{t-\Delta t}+m, y_{t-\Delta t}+n) \right)^2 \quad (3.4.2)$$

Miara podobieństwa NCC dwóch bloków pikseli z kolejnych obrazów wideo zdefiniowana jest natomiast równaniem (3.4.3), w którym μ oznacza wartość średnią jasności, a σ jest jej odchyleniem standardowym. Wielkości te są obliczane wewnątrz odpowiednich bloków obrazu.

$$\text{NCC}(x, y) = \frac{\sum_{m=-(N-1)/2}^{(N-1)/2} \sum_{n=-(N-1)/2}^{(N-1)/2} \left((I_t(x+m, y+n) - \mu_t)(I_{t-\Delta t}(x_{t-\Delta t}+m, y_{t-\Delta t}+n) - \mu_{t-\Delta t}) \right)}{\sigma_t \sigma_{t-\Delta t}} \quad (3.4.3)$$

$$\mu_t = \frac{1}{N^2} \sum_{m=-(N-1)/2}^{(N-1)/2} \sum_{n=-(N-1)/2}^{(N-1)/2} I_t(x+m, y+n) \quad (3.4.4)$$

$$\sigma_t = \sqrt{\frac{1}{N} \left(\sum_{m=-(N-1)/2}^{(N-1)/2} \sum_{n=-(N-1)/2}^{(N-1)/2} I_t^2(x+m, y+n) \right) - \mu_t^2} \quad (3.4.5)$$

$$\mu_{t-\Delta t} = \frac{1}{N^2} \sum_{m=-(N-1)/2}^{(N-1)/2} \sum_{n=-(N-1)/2}^{(N-1)/2} I_{t-\Delta t}(x_{t-\Delta t}+m, y_{t-\Delta t}+n) \quad (3.4.6)$$

$$\sigma_{t-\Delta t} = \sqrt{\frac{1}{N} \left(\sum_{m=-(N-1)/2}^{(N-1)/2} \sum_{n=-(N-1)/2}^{(N-1)/2} I_{t-\Delta t}^2(x_{t-\Delta t}+m, y_{t-\Delta t}+n) \right) - \mu_{t-\Delta t}^2} \quad (3.4.7)$$

Ostatnia z analizowanych tu miar, informacja wzajemna MI, jest zdefiniowana równaniem (3.4.8) i traktuje porównywane bloki pikseli jako realizacje pewnych zmiennych losowych A oraz B . Funkcja $P(A, B)$ oznacza prawdopodobieństwo równoczesnego wystąpienia zdarzeń A i B , $P(A)$ oraz $P(B)$ są odpowiednio prawdopodobieństwami zdarzenia A i zdarzenia B . W zastosowaniach związanych z analizą obrazów zdarzenia A oraz B oznaczają wystąpienia określonej jasności wewnątrz bloku obrazu. Przyjęto, że jasność może zmieniać się od zera do jasności maksymalnej I_m , a w takim przypadku funkcje prawdopodobieństw są odpowiednimi histogramami jasności.

$$MI(A, B) = \sum_{A=0}^{I_m} \sum_{B=0}^{I_m} P(A, B) \log \frac{P(A, B)}{P(A)P(B)} \quad (3.4.8)$$

$$P(A) = \frac{1}{N^2} \sum_{m=-(N-1)/2}^{(N-1)/2} \sum_{n=-(N-1)/2}^{(N-1)/2} \begin{cases} 1, & I_t(x+m, y+n) = A \\ 0, & I_t(x+m, y+n) \neq A \end{cases} \quad (3.4.9)$$

$$P(B) = \frac{1}{N^2} \sum_{m=-(N-1)/2}^{(N-1)/2} \sum_{n=-(N-1)/2}^{(N-1)/2} \begin{cases} 1, & I_{t-\Delta t}(x_{t-\Delta t}+m, y_{t-\Delta t}+n) = A \\ 0, & I_{t-\Delta t}(x_{t-\Delta t}+m, y_{t-\Delta t}+n) \neq A \end{cases} \quad (3.4.10)$$

$$P(A, B) = \frac{1}{N^2} \sum_{m=-(N-1)/2}^{(N-1)/2} \sum_{n=-(N-1)/2}^{(N-1)/2} \begin{cases} 1, & I_t(x+m, y+n) = A \wedge I_{t-\Delta t}(x_{t-\Delta t}+m, y_{t-\Delta t}+n) = B \\ 0, & I_t(x+m, y+n) \neq A \vee I_{t-\Delta t}(x_{t-\Delta t}+m, y_{t-\Delta t}+n) \neq B \end{cases} \quad (3.4.11)$$

Należy zauważyć, że wartości funkcji MAD i MSE są liczbami wymiernymi, większymi lub równymi zero. Im mniejsze są ich wartości, tym większe jest podobieństwo pomiędzy porównywanymi blokami obrazów. Wynika stąd, że dopasowanie obrazów można uzyskać przez minimalizację wartości tych funkcji. W przypadku miar MI i NCC mamy do czynienia z sytuacją odwrotną – wartości tych funkcji są tym większe, im większe jest podobieństwo pomiędzy porównywanymi blokami obrazów. Procedura dopasowania obrazów wymaga tu więc maksymalizacji wartości tych funkcji. Jednakże w modelu deformowalnym funkcja energii podlega minimalizowaniu – przyjęto więc, że postać energii cząstkowej $M_{p,q}$ będzie równa wartościom funkcji MAD lub MSE albo zanegowanym wartościom funkcji MI lub NCC. Funkcje te są dodatkowo przemnożone przez pewien parametr ξ , którego znaczenie wyjaśniono w dalszej części rozdziału.

tu. Stąd energię cząstkową M_{pq} można zdefiniować jednym z czterech równań, stosowanych opcjonalnie (3.4.12).

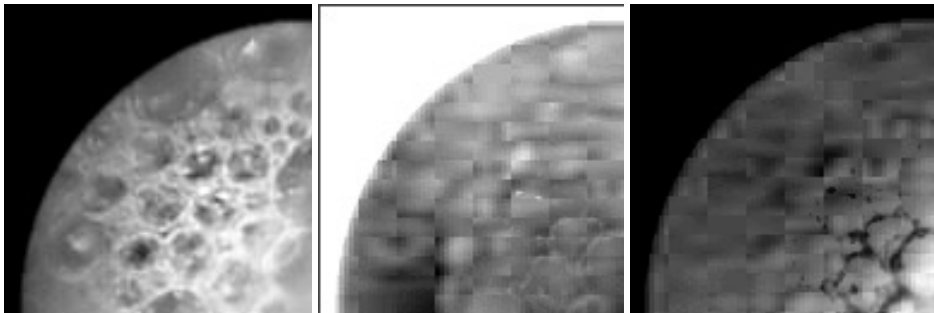
$$M_{pq} = \zeta \text{MAD}(x_{pq}, y_{pq}) \quad (3.4.12a)$$

$$M_{pq} = \zeta \text{MSE}(x_{pq}, y_{pq}) \quad (3.4.12b)$$

$$M_{pq} = -\zeta \text{MI}(x_{pq}, y_{pq}) \quad (3.4.12c)$$

$$M_{pq} = -\zeta \text{NCC}(x_{pq}, y_{pq}) \quad (3.4.12d)$$

Zagadnieniem o istotnym znaczeniu dla komputerowej implementacji algorytmu dopasowania MDR jest dyskretny charakter współrzędnych x, y rastrowych obrazów cyfrowych. Z drugiej strony współrzędne węzła modelu deformowalnego są w ogólnym przypadku wielkościami rzeczywistymi, przybliżanymi w implementacji numerycznej za pomocą liczb zmiennoprzecinkowych. Oznacza to konieczność interpolacji funkcji określonych w przestrzeni dyskretniej obrazu cyfrowego oraz, w przeciwną stronę, dyskretyzacji współrzędnych węzłów. W MDR zastosowano interpolację biliniową do przybliżania wartości funkcji MAD, MSE, MI oraz NCC. W miarach blokowych, fragment obrazu wybierany jest tak, żeby piksel środkowy bloku miał współrzędne określone współrzędnymi węzła przybliżonymi liczbami całkowitymi.



a)

b)

c)

Rys. 3.5. Fragment obrazu z endoskopu bezprzewodowego w modelu barw YUV:

a) składowa luminancji Y, b) składowa U chrominancji oraz c) składowa V chrominancji.

W składowych chrominancji widoczne są efekty kompresji blokowej i kwantyzacji

Kolejnym ważnym aspektem analizowanych obrazów endoskopowych jest fakt, że są to obrazy barwne, gdzie każdy piksel obrazu określony jest wektorem trzech wartości składowych koloru: R (czerwony), G (zielony) i B (niebieski). Przyjęto, że funkcje miar blokowych są obliczane dla poziomu jasności obrazu wyznaczonego zgodnie z rekomendacją *ITU-R BT.601* – równanie (3.4.13). Określanie podobieństwa obszarów oddzielnie dla poszczególnych składowych barw jest niecelowe, ponieważ jakość składowych chrominancji jest niewystarczająca. W systemie firmy *GivenImaging* składowe chrominancji wykorzystywanych obrazów endoskopowych są kompresowane metodami stratnymi (*Motion Jpeg*),

czego wynikiem jest występowanie zniekształceń – wyraźny jest podział obrazu na bloki (rysunek 3.5b i 3.5c).

$$I = 0.299 R + 0.587 G + 0.114 B \quad (3.4.13)$$

Podsumowując, w niniejszym podrozdziale zaproponowano cztery różne miary podobieństwa lub różnicy między porównywanymi fragmentami obrazów, które mogą być wykorzystywane do obliczania cząstkowej energii oddziaływania obrazu na węzły. Udział tej energii w modelu jest regulowany parametrem ξ . Strategia doboru tego parametru, jak również wnioski dotyczące zastosowania jednej z miar do analizy filmów endoskopowych, stanowią przedmiot rozważań rozdziału 4.

3.5. Obliczanie naprężeń w MDR

Jednym ze sposobów modelowania naprężeń w modelu parametrycznym określonym parą ciągłych parametrów (p, q) jest zastosowanie równania energii potencjalnej membrany (3.5.1) [24], [96], [165], [167], [216], będącego odpowiednikiem równania energii struny (3.2.3), wykorzystywanego w aktywnym konturze.

$$E = \tau \left[\left| \frac{\partial \mathbf{v}}{\partial p} \right|^2 + \left| \frac{\partial \mathbf{v}}{\partial q} \right|^2 \right] + \rho \left[\left| \frac{\partial^2 \mathbf{v}}{\partial p^2} \right|^2 + 2 \left| \frac{\partial^2 \mathbf{v}}{\partial p \partial q} \right|^2 + \left| \frac{\partial^2 \mathbf{v}}{\partial q^2} \right|^2 \right] \quad (3.5.1)$$

W przypadku modeli siatkowych, do których należy omawiany tu MDR, zastosowanie tego równania napotyka kilka poważnych trudności. Po pierwsze, obliczanie pochodnych, lub w modelu dyskretnym odpowiednich ilorazów różnicowych, nie jest możliwe dla punktów leżących na brzegu modelu. Po drugie, w modelu takim jak MDR, przy stałym parametrze p , zmieniający się parametr q wykreśla krzywą zamkniętą (pierścień). W przypadku modelu niezdeformowanego pierścień ten ma kształt okręgu – w takim przypadku naprężenia nie powinny występować, a funkcja E powinna być równa zero. W rzeczywistości wartości bezwzględne pochodnych są wówczas większe od zera. Jeśli zastosuje się równanie (3.5.1), wówczas nawet w niezdeformowanym MDR wartość funkcji energii E będzie większa od zera. Pochodne ulegają wyzerowaniu dopiero wówczas, gdy wszystkie węzły modelu znajdują się dokładnie w tym samym położeniu. Okazuje się, że próba minimalizacji wartości funkcji E prowadzi do niepożądanego „kurczenia się” modelu – węzły zbiegają się do jednego punktu. W przypadku zastosowania równania energii struny podobny problem „kurczenia” występuje również w modelu aktywnego konturu, o krzywej zamkniętej [101]. Problem ten został częściowo rozwiązany poprzez zastosowanie dodatkowych składników kompensujących niepożądane oddziaływanie naprężeń [162], [217]. Po trzecie, uważa się [187], [218], że parametry ρ i τ w równaniach energii struny i membrany nie dają pełnej kontroli nad sposobem, w jaki korygowany jest kształt modelu. Sposób modelowania naprężeń z zastosowaniem pochodnych powoduje, że oddziaływania korygujące kształt mają charakter lokalny. Zwiększanie wartości parametrów zmienia udział energii naprężeń w całkowitej energii modelu, ale nie zmienia zasięgu tych oddziaływań. Problem można porównać do filtracji za po-

mocą filtrów cyfrowych o skończonej odpowiedzi impulsowej (SOI) [24] – równanie (3.2.1) po dyskretyzacji parametrów jest odpowiednikiem filtru SOI o rzędzie równym 5, co wynika z aproksymacji pochodnych drugiego rzędu. Możliwa jest zmiana wartości współczynników filtru, ale nie jest możliwe zwiększenie jego rzędu.

Dotychczasowe próby opracowania innych metod obliczania naprężeń w modelach deformowalnych [23], [24] nie rozwiązywały wszystkich z wymienionych problemów. Ponadto większość z opracowanych metod jest dedykowana dla konkretnego typu modelu albo dla specyficznego rodzaju połączeń pomiędzy węzłami. Na potrzeby MDR opracowano oryginalny algorytm obliczania naprężeń, nazwany transformacją sąsiedztwa. Algorytm ten, choć zaproponowany do zastosowania w modelu o specyficznym rodzaju połączeń, może być wykorzystywany w dowolnym modelu parametrycznym, operującym w przestrzeni dwu- lub trójwymiarowej. Co więcej, metoda ta może być stosowana w modelach, w których węzły są rozmieszczone nierównomiernie i w arbitralny sposób.

3.5.1. Metoda transformacji sąsiedztwa

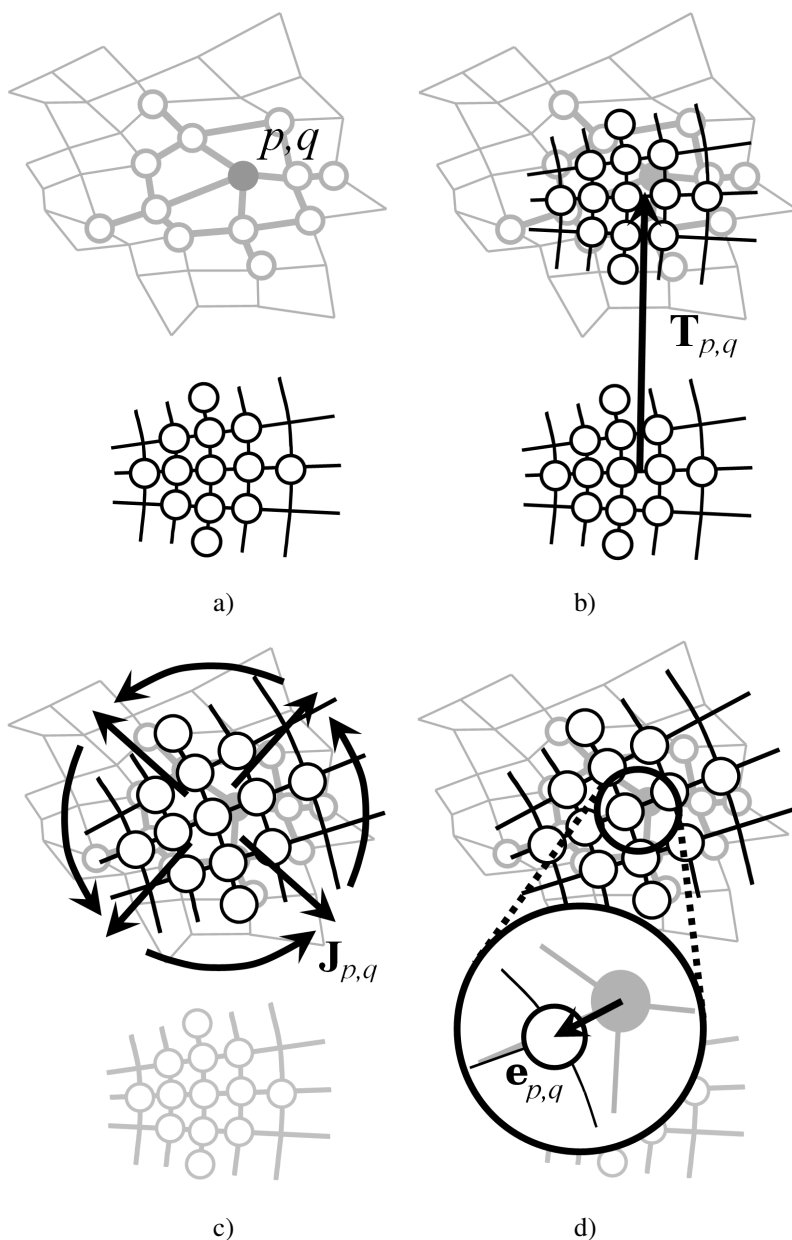
Metoda transformacji zakłada, że model deformowalny jest zbiorem węzłów rozmieszczonych w przestrzeni analizowanego obrazu. Przyjmuje się tu również, że istnieje model w postaci niezdeformowanej, nazywany dalej modelem odniesienia. Każdy węzeł w modelu odniesienia ma przypisane stałe współrzędne określające jego położenie w przestrzeni obrazu. Ponadto każdy węzeł w modelu odniesienia ma inne współrzędne, co oznacza, że nie występuje więcej niż jeden węzeł o takich samych współrzędnych. Model odniesienia jest podstawą do obliczania stopnia deformacji w odpowiadającym mu modelu deformowanym – naprężenia modelu są obliczane na zasadzie porównania modelu zdeformowanego z jego modelem odniesienia (niezdeformowanym).

Model deformowalny, zdefiniowany jako zbiór węzłów, posiada skończoną liczbę elementów. Dla uproszczenia zapisu przyjęto, że węzły identyfikowane tu będą za pomocą pojedynczego indeksu s , niezależnie od topologii modelu czy wymiarowości przestrzeni, w której ten model operuje. W przypadku deformowalnych pierścieni parametr s przyjmuje wartości od 1 do PQ i może być powiązany z parametrami p, q zależnością $s = p + P(q-1)$.

Współrzędne węzłów w modelu odniesienia, które pozostają niezmiennie, są oznaczane symbolem $\hat{\mathbf{v}}_s$, w odróżnieniu od współrzędnych w modelu deformowalnym, opisanych symbolem $\mathbf{v}_s$. W czasie analizowania obrazu poszczególne węzły przesuwać się – w ogólnym przypadku ich współrzędne $\mathbf{v}_s^{(i)}$ zmieniają się w kolejnych krokach (i) dopasowania.

Dla każdego węzła definiuje się jego sąsiedztwo, którym jest pewien podzbiór innych węzłów modelu, leżących w jego pobliżu, obejmujący również ten węzeł. Sąsiedztwo każdego z węzłów może być określane w sposób arbitralny. Istotne jest natomiast, żeby w skład każdego sąsiedztwa wchodziły przynajmniej trzy punkty węzłowe o współrzędnych niezależnych liniowo w przestrzeni dwuwymiarowej lub cztery takie punkty w przestrzeni trójwymiarowej. W MDR proponuje się wprowadzenie definicji η -sąsiedztwa (H_s) węzła s . Takie sąsiedztwo

jest zbiorem wszystkich węzłów połączonych z węzłem s za pomocą η lub mniejszej liczby połączeń, przy czym w skład tego sąsiedztwa wchodzi również węzeł s . Przyjmuje się, że połączenie występuje między parą węzłów tylko wtedy, gdy jeden z indeksów p lub q ma tę samą wartość, natomiast drugi różni się o jeden (rys. 3.4a).



Rys. 3.6. Ilustracja metody transformacji sąsiedztwa: a) sąsiedztwo węzła p, q w modelu deformowalnym i jego modelu odniesienia, b) przesunięcie, c) obrót, skalowanie i rozciągnięcie oraz d) wynikowy wektor naprężenia

Położenie węzłów wchodzących w skład sąsiedztwa H_s węzła s można, z pewnym błędem, przedstawić jako transformatę afiniczną współrzędnych tych węzłów w modelu odniesienia. Innymi słowy, dla każdego sąsiedztwa można określić stopień jego przesunięcia, obrotu, rozciągnięcia i przeskalowania w stosunku do odpowiadającego mu sąsiedztwa w modelu odniesienia. Transformacja afiniczna jest określona za pomocą macierzy transformacji, która niesie informację o skalowaniu, obrocie i rozciągnięciu, oraz za pomocą wektora przesunięcia. Jeśli macierz transformacji sąsiedztwa węzła s oznaczmy jako $\mathbf{J}_s$, a wektor przesunięcia jako $\mathbf{T}_s$, to można zapisać równanie (3.5.2) wiążące współrzędne punktu węzłowego k należącego do sąsiedztwa H_s modelu deformowalnego z odpowiadającym mu punktem modelu odniesienia. Zaznaczyć jednak należy, że występuje tu pewien błąd oznaczony symbolem $\mathbf{e}_k$.

$$\mathbf{v}_k = \mathbf{J}_s \hat{\mathbf{v}}_k + \mathbf{T}_s + \mathbf{e}_k; \quad i \in H_s \quad (3.5.2)$$

Macierz transformacji i wektor przesunięcia są obliczane w taki sposób, aby uzyskać minimalny błąd średniokwadratowy długości wektora $|\mathbf{e}_k|$. Błąd ten dany jest równaniem (3.5.3).

$$\varepsilon_s = \sum_{k \in N_s} (|\mathbf{e}_k|^2) = \sum_{k \in H_s} \left(\left| \mathbf{v}_k - (\mathbf{J}_s \hat{\mathbf{v}}_k + \mathbf{T}_s) \right|^2 \right) \quad (3.5.3)$$

Błąd ε_s jest funkcją elementów macierzy $\mathbf{J}_s$ oraz wektora $\mathbf{T}_s$. Funkcja ta ma jedno minimum występujące w przypadku, w którym jej pochodne po każdym z argumentów (elementów macierzy $\mathbf{J}_s$ oraz wektora $\mathbf{T}_s$) są równe zero. Tak więc, aby obliczyć elementy macierzy $\mathbf{J}_s$ oraz wektora $\mathbf{T}_s$ minimalizujące funkcję błędu, należy rozwiązać układ równań składających się z pochodnych funkcji błędu przyrównanych do zera.

W ogólnym przypadku transformacji afinicznej przestrzeni dwuwymiarowej, macierz $\mathbf{J}_s$ jest macierzą 2×2 , a wektor $\mathbf{T}_s$ jest wektorem dwuelementowym (3.5.4). W tym przypadku błąd średniokwadratowy ε_s można zapisać równaniem (3.5.5). Obliczając pochodne błędu ε_s po zmiennych $t_1, t_2, j_{11}, j_{12}, j_{21}$ oraz j_{22} , a następnie przyrównując te pochodne do zera, otrzymujemy układ sześciu równań. Rozwiązaniem tego układu są wartości zmiennych $t_1, t_2, j_{11}, j_{12}, j_{21}$ oraz j_{22} , dane równaniem (3.5.6).

$$\mathbf{J}_s = \begin{bmatrix} j_{11} & j_{12} \\ j_{21} & j_{22} \end{bmatrix}; \quad \mathbf{T}_s = \begin{bmatrix} t_1 \\ t_2 \end{bmatrix} \quad (3.5.4)$$

$$\varepsilon_s = \sum \left((x - t_1 + j_{11} \hat{x} + j_{12} \hat{y})^2 + (y - t_2 + j_{21} \hat{x} + j_{22} \hat{y})^2 \right) \quad (3.5.5)$$

$$\begin{bmatrix} j_{11} \\ j_{12} \\ t_1 \end{bmatrix} = \mathbf{M} \begin{bmatrix} \sum \hat{x} x \\ \sum \hat{y} x \\ \sum x \end{bmatrix}; \quad \begin{bmatrix} j_{21} \\ j_{22} \\ t_2 \end{bmatrix} = \mathbf{M} \begin{bmatrix} \sum \hat{x} y \\ \sum \hat{y} y \\ \sum y \end{bmatrix} \quad (3.5.6)$$

$$\mathbf{M} = \frac{\begin{bmatrix} \sum \hat{y}^2 \sum 1 - (\sum \hat{y})^2 & \sum \hat{y} \sum \hat{x} - \sum \hat{x} \hat{y} \sum 1 & \sum \hat{x} \hat{y} \sum \hat{y} - \sum \hat{y}^2 \sum \hat{x} \\ \sum \hat{y} \sum \hat{x} - \sum \hat{x} \hat{y} \sum 1 & \sum \hat{x}^2 \sum 1 - (\sum \hat{x})^2 & \sum \hat{x} \hat{y} \sum \hat{x} - \sum \hat{x}^2 \sum \hat{y} \\ \sum \hat{x} \hat{y} \sum \hat{y} - \sum \hat{y}^2 \sum \hat{x} & \sum \hat{x} \hat{y} \sum \hat{x} - \sum \hat{x}^2 \sum \hat{y} & \sum \hat{x}^2 \sum \hat{y}^2 - (\sum \hat{x} \hat{y})^2 \end{bmatrix}}{\sum \hat{x}^2 \sum \hat{y}^2 \sum 1 - (\sum \hat{y})^2 \sum \hat{x}^2 - (\sum \hat{x} \hat{y})^2 \sum 1 - (\sum \hat{x})^2 \sum \hat{y}^2 + 2 \sum \hat{y} \sum \hat{x} \sum \hat{x} \hat{y}} \quad (3.5.7)$$

W równaniach wykorzystano symbole x oraz y do oznaczenia elementów wektora $\mathbf{v}$. Dla uproszczenia zapisu pod symbolami sumowania pominięto zakres sumowania, przy czym sumowanie jest wykonywane dla wszystkich punktów węzłowych k należących do sąsiedztwa H_s . Pominięto również indeks k przy wektorach współrzędnych.

Można zauważyć, że w rozwiązaniu (3.5.6) występuje macierz $\mathbf{M}$ o postaci danej równaniem (3.5.7). Macierz ta jest stała, ponieważ zależy wyłącznie od współrzędnych węzłów w modelu odniesienia. Oznacza to, że macierzy $\mathbf{M}$ nie trzeba obliczać w czasie dopasowywania modelu. Ma to istotny wpływ na szybkość obliczeniową proponowanej metody.

Jeśli zależności przestrzenne między węzłami w sąsiedztwie H_s są zachowane w stosunku do modelu odniesienia, wówczas błędy $\mathbf{e}_k$ są równe zeru. W szczególności dotyczy to też błędu $\mathbf{e}_k$ obliczanego dla węzła s ($k = s$). Jeśli natomiast występuje lokalna deformacja w punkcie s , to wówczas $\mathbf{e}_k$ przy $k = s$ jest różne od zera. Błąd $\mathbf{e}_{k=s}$ można więc wykorzystać jako lokalną miarę deformacji modelu w stosunku do jego modelu odniesienia.

W proponowanej metodzie energię cząstkową naprężeń w węźle s definiuje się równaniem (3.5.8). Jest ona proporcjonalna do kwadratu odległości pomiędzy rzeczywistym położeniem $\mathbf{v}_s$ węzła a transformatą afiniczną jego położenia w modelu odniesienia. Transformacja afiniczna jest obliczona dla sąsiedztwa tego węzła, zaś parametrem ρ określa się udział energii naprężeń w energii modelu deformowalnego.

$$E_s = \rho \left\| \mathbf{v}_s - (\mathbf{J}_s \hat{\mathbf{v}}_s + \mathbf{T}_s) \right\|^2 \quad (3.5.8)$$

Szczegółowe wyprowadzenia macierzy i wektora transformacji, również dla przypadku przestrzeni trójwymiarowej, zamieszczono w załączniku A. W załączniku tym rozważono też przypadki transformacji ograniczonych do obrotu i przesunięcia oraz obrotu, skalowania i przesunięcia. Przypadki rozważane w opisywanym załączniku umożliwiają znaczne ograniczenie kierunkowego rozciągania modelu (zachowanie proporcji) lub ograniczenie skalowania (zachowanie pola powierzchni lub objętości).

3.5.2. Podział dopasowania na etapy

W metodzie transformacji sąsiedztwa możliwe jest regulowanie zasięgu oddziaływań naprężeń poprzez dobór parametru η określającego wielkość η -sąsiedztwa. Parametr η ma więc wpływ na sposób, w jaki model przeciwstawia się niepożądanym zniekształceniom. Niewielkie wartości parametru η powodują, że korygowane są lokalne zniekształcenia, natomiast zmiany kształtu o globalnym

charakterze są dopuszczane. Duże wartości parametru η powodują, że zasięg modelowania naprężeń jest większy – jeśli parametr η jest odpowiednio duży, taki że sąsiedztwo każdego węzła obejmuje wszystkie węzły modelu, to korygowane są deformacje o zasięgu zarówno lokalnym, jak i globalnym.

Jednym z kluczowych zagadnień w przetwarzaniu obrazów biomedycznych jest dopasowanie do siebie treści dwóch obrazów. Obrazy takie zazwyczaj pokazują to samo miejsce, które zostało zobrazowane w różnym czasie albo za pomocą różnych technik. W zależności od potrzeb stosowane są dwa opcjonalne podejścia – dopasowanie sztywne (ang. *rigid image registration*) lub elastyczne (ang. *non-rigid image registration*). Okazuje się, że metoda transformacji sąsiedztwa przy małych wartościach parametru η , ma cechy elastycznego dopasowania obrazów, zaś przy dużych wartościach – sztywnego.

W metodzie modelowania naprężeń z wykorzystaniem transformacji sąsiedztwa są obliczane macierze $\mathbf{J}_s$ i wektory $\mathbf{T}_s$ dla sąsiedztwa każdego punktu węzłowego s należącego do modelu. Do ich obliczenia konieczne jest wykonanie pewnej liczby sumowań i mnożeń proporcjonalnej do liczby wszystkich punktów węzłowych η -sąsiedztwa. Takie obliczenia są wykonywane indywidualnie dla każdego z węzłów. Liczba sumowań i mnożeń w pojedynczym kroku dopasowania jest zatem proporcjonalna do iloczynu liczby węzłów oraz średniej liczby węzłów w η -sąsiedztwach. Koszt obliczeniowy algorytmu zwiększa się więc wraz ze wzrostem wartości parametru η . Jeśli jednak parametr η osiągnie wartość, przy której dla każdego punktu węzłowego jego sąsiedztwo obejmuje wszystkie węzły modelu, to η -sąsiedztwa poszczególnych węzłów s stają się identyczne. Identyczne są również macierze $\mathbf{J}_s$ i wektory $\mathbf{T}_s$ tych sąsiedztw. Oznacza to, że nie jest już konieczne obliczanie parametrów transformacji dla każdego z węzłów indywidualnie. Wystarczy obliczyć parametry transformacji jednokrotnie w pojedynczym kroku procesu dopasowania. W takim przypadku nakład obliczeniowy znacznie się zmniejsza.

Z powyższych spostrzeżeń wynika, że istnieje kilka sposobów skrócenia czasu obliczeń, z których każdy jest obciążony pewną wadą. Pierwszy to zmniejszenie liczby wszystkich węzłów modelu, a drugi to zwiększenie liczby węzłów sąsiedztwa tak, aby obejmowało ono wszystkie punkty węzłowe modelu. Trzeci to redukcja liczby węzłów sąsiedztwa przez zmniejszenie parametru η .

Pierwszy sposób oszczędza czas obliczeń kosztem szczegółowości – wynika to stąd, że im większa jest liczba wszystkich punktów węzłowych, tym gęściej zbierane są informacje z powierzchni obrazu. Zastosowanie drugiej metody umożliwia szybkie dopasowanie modelu – model taki dopasowuje się w sposób sztywny, co nie jest odpowiednie do śledzenia ruchów będących np. wynikiem skurczów elastycznego przewodu pokarmowego. Zastosowanie trzeciej metody powoduje, że kształt modelu jest korygowany lokalnie, przy czym proces dopasowania bywa wówczas nieefektywny w przypadku występowania w obrazie zakłóceń. Niektóre węzły mogą wówczas utkwić w niepoprawnych położeniach – lokalnych minimach energetycznych – modelowanie naprężeń o lokalnym zakresie nie jest wówczas wystarczające do ich „wyciągnięcia” z takiego położenia.

Rozwiązaniem problemu umożliwiającym jednoczesne skrócenie czasu obliczeń, przy zachowaniu zalet rejestracji sztywnej i elastycznej, jest podział procesu

dopasowania na etapy [24], [98]. W pierwszym etapie model jest dopasowywany z sąsiedztwem obejmującym wszystkie węzły, jako całość przesuwa się on, zmienia orientację, rośnie lub maleje. Po tym etapie uzyskuje się przybliżone dopasowanie modelu do analizowanego obrazu. W drugim etapie przyjmuje się odpowiednio małe η -sąsiedztwo, czego efektem jest lokalna deformacja i dopasowanie modelu do szczegółów obrazu.

3.5.3. Transformacja z ograniczaniem składowych

Model deformowalny, wykorzystujący metodę transformacji sąsiedztwa z sąsiedztwem obejmującym wszystkie węzły, ma swobodę co do zmiany swej wielkości, obrotu, przesunięcia i zmiany proporcji (rozciągnięcia w jednym kierunku i skurczenia w kierunku do niego prostopadłym). W przypadku siatki deformowalnych pierścieni natomiast nie wszystkie ze stopni swobody są pożądane, gdyż model ten powinien pozostawać w polu widzenia kamery – nie powinien się więc zbyt swobodnie przesuwać poza to pole. Pożądane jest też utrzymanie jego proporcji. Zmiana jego wielkości oraz obrót są natomiast potrzebne w śledzeniu prędkości postępowej i prędkości obrotowej wokół osi kamery. Praktycznym rozwiązaniem jest więc tu ograniczenie pewnych wybranych elementów swobody modelu – jest to możliwe po przeprowadzeniu dekompozycji macierzy transformacji afinicznej na współczynniki osobno określające stopień obrotu, rozciągnięcia kierunkowego i skalowania. Składnik translacji jest już wydzielony i określony wektorem przesunięcia, stosując rozkład biegunowy macierzy można wydzielić czynnik odpowiedzialny za obrót, stosując rozkład własny macierzy można wydzielić czynniki odpowiedzialne za rozciągnięcie kierunkowe, a obliczając wyznacznik macierzy określić czynnik skalowania. Szczegóły dekompozycji macierzy transformacji przedstawiono w załączniku A. Równanie (3.5.9) jest wynikiem takiej dekompozycji.

$$\mathbf{J} = a \begin{bmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{bmatrix} \begin{bmatrix} \cos \beta & -\sin \beta \\ \sin \beta & \cos \beta \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} \cos \beta & \sin \beta \\ -\sin \beta & \cos \beta \end{bmatrix} \quad (3.5.9)$$

Parametry λ są wartościami własnymi macierzy, ich iloczyn jest równy jedności, a proporcje między nimi określają stopień rozciągnięcia kierunkowego. Kierunek, w którym model jest rozciągnięty, dany jest kątem β . Parametr α to kąt obrotu modelu deformowalnego w stosunku do modelu odniesienia. Parametr a jest pierwiastkiem kwadratowym wartości bezwzględnej wyznacznika macierzy transformacji i określa wielkość modelu deformowalnego względem jego modelu odniesienia.

W równaniu (3.5.8), służącym do obliczenia energii naprężeń, zamiast macierzy $\mathbf{J}$ można zastosować zmodyfikowaną macierz $\mathbf{J}'$ (3.5.10), a zamiast wektora $\mathbf{T}$ zmodyfikowany wektor $\mathbf{T}'$ (3.5.12).

$$\mathbf{J}' = a' \begin{bmatrix} \cos \alpha' & \sin \alpha' \\ -\sin \alpha' & \cos \alpha' \end{bmatrix} \begin{bmatrix} \cos \beta' & -\sin \beta' \\ \sin \beta' & \cos \beta' \end{bmatrix} \begin{bmatrix} \lambda_1' & 0 \\ 0 & \lambda_2' \end{bmatrix} \begin{bmatrix} \cos \beta' & \sin \beta' \\ -\sin \beta' & \cos \beta' \end{bmatrix} \quad (3.5.10)$$

$$a' = (a)^{\rho_A}; \quad \alpha' = \rho_R \alpha; \quad \lambda_1' = \lambda_1^{\rho_S}; \quad \lambda_2' = \lambda_2^{\rho_S} \quad (3.5.11)$$

$$\mathbf{T}' = \rho_T \mathbf{T} \quad (3.5.12)$$

Komponenty i współczynniki transformacji zostają zmodyfikowane za pomocą parametrów ρ_A , ρ_R , ρ_S i ρ_T (3.5.11). Parametry te dobiera się w zakresie od zera do jedności, a ich wartości określają udział danego komponentu lub współczynnika w wynikowej transformacji. Parametr ρ_A jest odpowiedzialny za skalowanie, ρ_R za obrót, ρ_S za rozciągnięcie, natomiast ρ_T za przemieszczenie. W MDR parametry ρ_A i ρ_R są równe jedności, natomiast pozostałe mają nieco mniejsze wartości.

3.5.4. Właściwości metody transformacji sąsiedztwa

Metoda transformacji sąsiedztwa ma właściwości, które predestynują ją do zastosowania w różnych parametrycznych modelach deformowalnych, a nie tylko w MDR – umożliwia ona modelowanie naprężeń w modelach o różnych topologiach. Mogą to być kontury otwarte i zamknięte oraz modele siatkowe. Co więcej, rozmieszczenie węzłów nie musi już mieć charakteru regularnego, ale może być dobrane arbitralnie. W przypadku aktywnego konturu, na podstawie wstępnej wiedzy lub oczekiwań, można nadać mu preferowany kształt zbliżony do kształtu linii brzegowej poszukiwanego obszaru. Zaznaczyć należy także, że metoda transformacji sąsiedztwa otwiera możliwości projektowania nowych, dotychczas nieznanych modeli. Możliwe staje się wtedy np. projektowanie modeli hybrydowych, będących połączeniem modeli siatkowych oraz aktywnych konturów.

W analogiczny sposób, jak dla przypadku dwuwymiarowego, można wyprowadzić równania dla przestrzeni trójwymiarowej (załącznik A) – umożliwia to zastosowanie tej metody do analizy obiektów przestrzennych, np. za pomocą deformowalnej powierzchni w obrazach tomografii komputerowej.

Ze względu na liczbę parametrów, metoda transformacji sąsiedztwa wydaje się być bardziej skomplikowana niż metoda wykorzystująca równanie struny lub membrany, w których występują tylko dwa parametry modelujące rozciągliwość oraz sprężystość. Interpretacja tych parametrów jest jednak niejednoznaczna, a w niektórych publikacjach pojawia się stwierdzenie, że nie ma zauważalnej różnicy w modelowaniu rozciągliwości i sprężystości lub że modelowanie sprężystości jest zbędne [187], [218]. W przypadku metody transformacji sąsiedztwa mamy natomiast do czynienia ze zbiorem od dwóch do dziesięciu parametrów, które można jednoznacznie interpretować jako zasięg oddziaływań (η), oporność modelu na przesunięcie (ρ_T), na zmianę rozmiaru (ρ_A), na obracanie (ρ_R) itd. Część z tych parametrów umożliwia kontrolę zachowań modelu niedostępną w innych metodach modelowania naprężeń.

Złożoność numeryczna algorytmu nie jest znacząco większa od innych sposobów modelowania naprężeń. Należy zauważyć, że elementy macierzy $\mathbf{M}$ (3.5.7) zawierają czynniki związane ze współrzędnymi modelu odniesienia – oznacza to, że macierz taką dla określonego modelu oblicza się jednokrotnie, zaś podczas całego procesu dopasowania wartości elementów tej macierzy nie ulegają zmianie.

W przypadku zastosowania metod z ograniczaniem składowych złożoność numeryczna zwiększa się, ponieważ jest przeprowadzana dekompozycja macierzy osobno dla każdego sąsiedztwa. Jeśli jednak ograniczyć stosowanie tej metody do przypadku, w którym sąsiedztwo obejmuje wszystkie punkty węzłowe modelu, wówczas dekompozycję trzeba przeprowadzić raz w ciągu pojedynczego kroku procesu dopasowania. W takim przypadku złożoność obliczeniowa zwiększa się jedynie w nieznacznym stopniu.

3.6. Dopasowanie MDR

Jak wspomniano, dopasowanie modelu polega na takiej zmianie współrzędnych poszczególnych punktów węzłowych, aby zminimalizować wartość funkcji E_{MDR} . Do jej minimalizacji wykorzystano optymalizację gradientową, dla której punktem startowym są współrzędne węzłów uzyskane w wyniku dopasowania modelu do poprzedniego obrazu sekwencji. Dzięki temu wynik dopasowania do analizowanego obrazu zależy od wyniku uzyskanego w obrazie poprzednim, co w efekcie prowadzi do płynnego podążania MDR za przemieszczającymi się elementami obserwowanej sceny. W praktyce okazało się, że zastosowanie metod minimalizacji o globalnym charakterze powoduje zwiększenie wartości błędu oszacowania prędkości ruchu postępowego endoskopu [14].

Równanie (3.3.5) pokazuje, że współrzędne wektora gradientu można obliczać niezależnie dla energii cząstkowych zdefiniowanych dla poszczególnych węzłów – w takim przypadku równanie iteracyjne (3.3.3) można rozbić na równania niezależnie opisujące zmianę współrzędnych dla poszczególnych węzłów p, q (3.6.1).

$$\begin{bmatrix} x_{p,q}^{(i+1)} \\ y_{p,q}^{(i+1)} \end{bmatrix} = \begin{bmatrix} x_{p,q}^{(i)} \\ y_{p,q}^{(i)} \end{bmatrix} - \left[\frac{\partial M_{p,q}^{(i)}}{\partial x} \quad \frac{\partial M_{p,q}^{(i)}}{\partial y} \right]^T - \left[\frac{\partial E_{p,q}^{(i)}}{\partial x} \quad \frac{\partial E_{p,q}^{(i)}}{\partial y} \right]^T \quad (3.6.1)$$

Podczas stosowania transformacji sąsiedztwa nie ma potrzeby numerycznego obliczania gradientu cząstkowej energii naprężeń w węzle modelu. Jeśli energia ta dana jest równaniem (3.5.8), to odpowiedni wektor gradientu można opisać równaniem (3.6.2).

$$\left[\frac{\partial E_{p,q}^{(i)}}{\partial x} \quad \frac{\partial E_{p,q}^{(i)}}{\partial y} \right]^T = 2\rho \left(\begin{bmatrix} x_{p,q}^{(i)} \\ y_{p,q}^{(i)} \end{bmatrix} - \left(\mathbf{J}_{p,q}^{(i)} \begin{bmatrix} x_{p,q}^{(i)} \\ y_{p,q}^{(i)} \end{bmatrix} + \mathbf{T}_{p,q}^{(i)} \right) \right) \quad (3.6.2)$$

Do obliczania energii oddziaływania obrazu zaproponowano miary blokowe różnicy lub podobieństwa fragmentów obrazu, które są funkcjami dyskretnych współrzędnych obrazu cyfrowego. Do obliczania gradientów takich miar stosuje się dyskretne operatory różniczkujące, np. operator Sobela [211], [219]. W wyniku zastosowania tego operatora otrzymuje się dyskretne pole wektorowe. Jak już wspomniano, na potrzeby modelu z węzłami o ciągłych współrzędnych, pole to jest interpolowane za pomocą interpolacji biliniowej.

W MDR dopasowanie jest podzielone na dwa etapy – w każdym z nich jest minimalizowana inna postać funkcji energii modelu. Pierwszy etap, obejmujący

pewną liczbę iteracji, jest przeprowadzany dla postaci funkcji energii, w której naprężenia są modelowane z η -sąsiedztwem obejmującym wszystkie węzły modelu. W drugim etapie jest minimalizowana funkcja energii w zmienionej postaci, w której naprężenia są modelowane w lokalnych sąsiedztwach.

Proces dopasowania jest powtarzany w kolejnych obrazach sekwencji wideo. Położenie i kształt modelu, uzyskane w wyniku dopasowania dla jednego obrazu, są wykorzystywane jako początkowe w procesie dopasowania dla kolejnego. Wyjątkiem od tej zasady są przypadki, w których model powiększy się nadmiernie lub skurczy. W pierwszym z nich, po zakończeniu dopasowania, zewnętrzny pierścień modelu jest usuwany i dodawany jest dodatkowy pierścień wewnętrzny. W drugim przypadku jest usuwany pierścień wewnętrzny, a dodawany jest dodatkowy pierścień zewnętrzny. W obu przypadkach wszystkim węzłom są przypisywane nowe wartości indeksu p , tak aby wartość tego indeksu była równa jeden dla wewnętrznego pierścienia i rosła dla kolejnych. Szczegóły dotyczące procedury usuwania i dodawania pierścieni są podane w podrozdziale 3.7.2 dotyczącym tworzenia mapy przewodu pokarmowego.

3.7. Dane wynikowe MDR

Zastosowanie modelu deformowalnych pierścieni prowadzi do uzyskania danych dwojakiego rodzaju. Po pierwsze, model generuje deskryptory ruchu – wartości liczbowe charakteryzujące ruch kapsułki endoskopu względem przewodu pokarmowego oraz w pewnym stopniu ruch samego przewodu. Deskryptory te są obliczane na podstawie położenia i kształtu MDR po zakończeniu procesu dopasowania, dla każdego obrazu sekwencji wideo – są więc one funkcjami zmiennymi w czasie. Po drugie, wynikiem działania MDR jest pojedynczy obraz przewodu pokarmowego, będący złożeniem informacji wyodrębnionej z kolejnych obrazów sekwencji.

3.7.1. Deskryptory ruchu

W wyniku działania MDR, dla każdego obrazu sekwencji wideo oblicza się sześć różnych deskryptorów. Deskryptor G_1 jest odległością środka modelu od środka obrazu, G_2 jest miarą globalnej deformacji modelu, G_3 jest miarą podobieństwa dwóch kolejnych obrazów sekwencji wideo. Deskryptory V_D (3.1.1) i V_P (3.1.2) estymują ruch kapsuły wzdłuż przewodu pokarmowego. Obliczana jest również prędkość kątowna ω obrotu kapsułki wokół jej osi podłużnej.

Deskryptory G_2 i G_3 są obliczane na podstawie składowych energii modelu wyznaczonych po zakończeniu procesu dopasowania. Deskryptor G_2 jest obliczany (3.7.1) jako średnia cząstkowych energii naprężeń w węzłach modelu podzielona przez parametr ρ . Deskryptor G_3 jest obliczany analogicznie (3.7.2) jako średnia cząstkowych energii oddziaływania obrazu podzielona przez parametr ζ .

$$G_2 = \frac{1}{\rho P Q} \sum_{p=1}^P \sum_{q=1}^Q E_{p,q} = \frac{1}{P Q} \sum_{p=1}^P \sum_{q=1}^Q \left| \mathbf{v}_{p,q} - (\mathbf{J}_{p,q} \hat{\mathbf{v}}_{p,q} + \mathbf{T}_{p,q}) \right|^2 \quad (3.7.1)$$

$$G_3 = \frac{1}{\xi P Q} \sum_{p=1}^P \sum_{q=1}^Q M_{p,q} \quad (3.7.2)$$

W celu obliczenia wartości pozostałych deskryptorów, po zakończeniu dopasowania, są obliczane macierz $\mathbf{J}$ i wektor $\mathbf{T}$ transformacji afinicznej modelu w stosunku do modelu odniesienia. Obliczane są one za pomocą równania (3.5.6) dla η -sąsiedztwa obejmującego wszystkie węzły modelu. Następnie macierz transformacji jest rozkładana na czynniki (3.5.9), w wyniku czego otrzymywane są parametry transformacji, określające, między innymi wielkość modelu $\sqrt{|\det \mathbf{J}|}$ i kąt α obrotu w stosunku do modelu odniesienia.

Deskryptor G_I jest odległością środka modelu od środka obrazu, a ponieważ przyjęto, że początek układu współrzędnych i środek modelu odniesienia leżą w środku obrazu, to G_I jest długością wektora $\mathbf{T}$ (3.7.3). Prędkość kątową można obliczyć (3.7.4) jako różnicę kątów α obliczonych dla modelu w aktualnym i poprzednim obrazie sekwencji, podzieloną przez czas pomiędzy uzyskaniem dwóch kolejnych obrazów. Przy czym obliczona różnica kątów jest sprowadzana do zakresu $(-\pi, +\pi)$.

$$G_I = |\mathbf{T}| \quad (3.7.3)$$

$$\omega = \frac{\alpha_t - \alpha_{t-\Delta t}}{\Delta t}; \quad -\pi < \alpha_t - \alpha_{t-\Delta t} \leq +\pi \quad (3.7.4)$$

Deskryptory V_D i V_P są obliczane na podstawie porównania względnej wielkości modelu w aktualnym i poprzednim obrazie. Odległości d z równania (3.1.1) i (3.1.2) są wyznaczone jako przeciętny promień modelu deformowalnych pierścieni (3.7.5). Po podstawieniu zależności (3.6.5) do równań (3.1.1) i (3.1.2) otrzymuje się równania (3.7.6) i (3.7.7) służące do obliczania wartości deskryptorów V_D i V_P .

$$d = r \sqrt{w^{P-1} |\det \mathbf{J}|} \quad (3.7.5)$$

$$V_D = \frac{r_D}{\Delta t} \left(\tan^{-1} \left(r \frac{\sqrt{w^{P-1} |\det \mathbf{J}_{t-\Delta t}|}}{f} \right) - \tan^{-1} \left(r \frac{\sqrt{w^{P-1} |\det \mathbf{J}_t|}}{f} \right) \right) \quad (3.7.6)$$

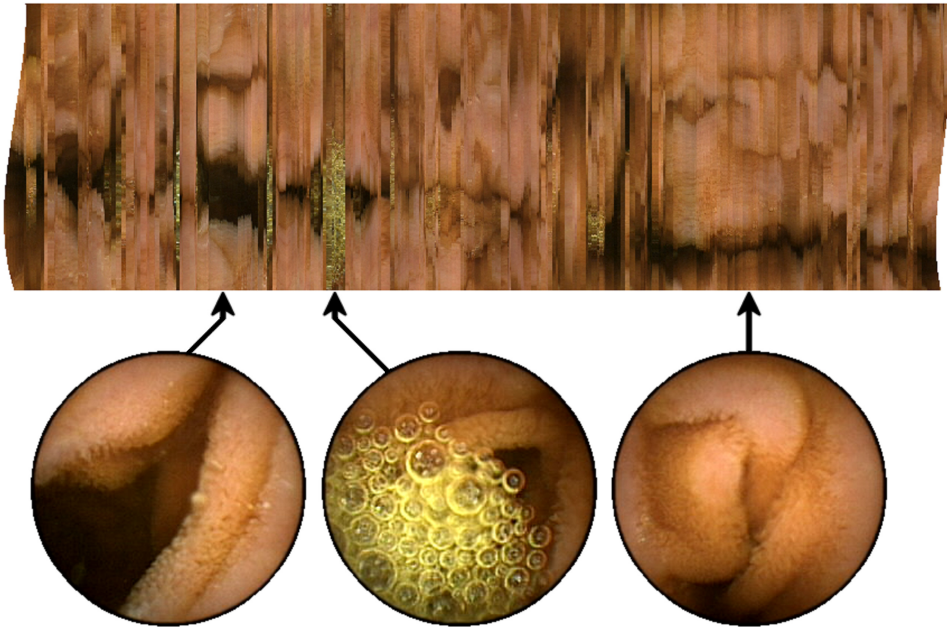
$$V_P = \frac{r_D f}{\Delta t} \left(\frac{1}{r} \sqrt{w^{P-1} |\det \mathbf{J}_t|} - \frac{1}{r} \sqrt{w^{P-1} |\det \mathbf{J}_{t-\Delta t}|} \right) \quad (3.7.7)$$

Jak już wspomniano, parametr r jest promieniem wewnętrznego pierścienia w modelu odniesienia, parametr w jest stosunkiem promieni dwóch sąsiednich pierścieni (większego do mniejszego), f jest ogniskową obiektywu kamery, a r_D jest połową średnicy kapsułki.

3.7.2. Mapa przewodu pokarmowego

W czasie analizowania filmu z endoskopu kapsułkowego model deformowalny może się kurczyć lub rosnąć. Jeśli zwrot wektora prędkości kapsułki jest

zgodny z wektorem skierowania kamery, to wówczas fragmenty obrazu przesuwają się od wewnątrz na zewnątrz, a model podążający za tymi fragmentami rośnie. W przypadku, w którym wektor prędkości ma zwrot przeciwny do zwrotu kamery, sytuacja jest odwrotna i model się kurczy.



Rys. 3.7. Przykładowe fragmenty mapy przewodu pokarmowego oraz powiązane z nimi obrazy

Wielkość modelu jest obliczana na podstawie macierzy $\mathbf{J}$ (3.5.6) transformacji afinicznej modelu w stosunku do modelu odniesienia. Macierz ta jest obliczana po zakończeniu dopasowania modelu dla η -sąsiedztwa obejmującego wszystkie jego węzły. Pierwiastek kwadratowy modułu wyznacznika tej macierzy oznacza proporcję wielkości modelu deformowalnego do wielkości modelu odniesienia.

MDR jest chroniony przed nadmiernym rozrostem (poza ramy obrazu) oraz przed nadmiernym skurczeniem. Jeśli model powiększy się o odległość między sąsiednimi pierścieniami, czyli gdy $\sqrt{|\det \mathbf{J}|} > w$, wówczas zewnętrzny pierścień modelu jest usuwany i zamiast niego dodawany jest odpowiedniej wielkości pierścień wewnętrzny. Jeśli model zmniejszy się tak, że $\sqrt{|\det \mathbf{J}|} < w^{-1}$, wówczas wewnętrzny pierścień jest usuwany i zamiast niego dodawany jest odpowiedniej wielkości pierścień zewnętrzny. Współrzędne węzłów dodawanych pierścieni są ekstrapolacją współrzędnych węzłów w dwóch najbliższych pierścieniach.

W obydwu przypadkach, po usunięciu lub dodaniu pierścienia, zapamiętywane są składowe koloru obrazu odczytane we współrzędnych kolejnych węzłów q zewnętrznego pierścienia. Wektory te zostają ułożone w szereg, tworząc pojedynczą linię obrazu cyfrowego. Poszczególne linie, uzyskane podczas usuwania i dodawania pierścieni, składane są kolejno razem, tworząc mapę – obraz będący reprezentacją wewnętrznej powierzchni przewodu pokarmowego. Rozdzielczość

poprzeczna mapy może zostać zwiększona przez odczytanie składowych koloru dodatkowo pomiędzy węzłami pierścienia. Rozdzielczość podłużną można zwiększyć, odczytując składowe koloru pomiędzy dwoma pierścieniami modelu, w obrazach pomiędzy zdarzeniami usunięcia i dodania pierścienia.

Procedura tworzenia mapy wiąże się z koniecznością przeprowadzenia interpolacji dla trzech różnych przypadków. Po pierwsze, odczyt składowych koloru pomiędzy węzłami wymaga interpolacji współrzędnych węzłów, funkcji $x(\cdot)$ i $y(\cdot)$, których wartości znane są w dziedzinie dyskretnych parametrów p i q . Po drugie, składowe koloru RGB są określone w dyskretnych współrzędnych obrazu, tak więc ich odczytanie w quasi ciągłych współrzędnych (x, y) wymaga interpolacji tych składowych z dyskretnych współrzędnych na quasi ciągłe współrzędne węzłów modelu. Po trzecie, odczyt składowych RGB pomiędzy pierścieniami powinien być dokonywany w czasie pomiędzy kolejnymi zdarzeniami usunięcia i dodania pierścienia, a to wymaga dzielenia okresów pomiędzy zdarzeniami na krótsze, możliwie równe odcinki czasu.

Wraz z każdą linią mapy jest zapisywany kolejny numer obrazu sekwencji wideo, dla którego linia ta została utworzona – informacja ta umożliwia powiązanie współrzędnych mapy z obrazami wideo. W procesie interpretacji filmu z endoskopu umożliwia to szybkie przywołanie jego fragmentu poprzez wskazanie wybranego miejsca na mapie.

4. Walidacja MDR

W celu walidacji MDR i optymalizacji jego działania przeprowadzono trzy eksperymenty. Pierwszy z nich umożliwia ocenę działania modelu z wykorzystaniem jednej z miar podobieństwa lub różnicy między blokami obrazów. W wyniku tego eksperymentu określono zasady optymalizowania parametrów związanych z obliczeniami wpływu obrazu na MDR. Drugi z eksperymentów wykazuje, że deskryptory ruchu obliczane przez model umożliwiają rozróżnienie rodzajów skurczów przewodu pokarmowego. Na podstawie uzyskanych wyników zaproponowano postać równania umożliwiającego automatyczne sterowanie odtwarzaniem filmu na podstawie deskryptorów ruchu wygenerowanych przez model. Trzeci z eksperymentów weryfikuje przydatność mapy przewodu pokarmowego generowanej przez MDR w praktyce klinicznej.

Model deformowalnych pierścieni został zaimplementowany w języku C++ w postaci filtru w technologii *DirectShow* (DS) [220–222]. Technologia ta jest dostępna w systemach operacyjnych Microsoft Windows®. Filtr DS jest modulem programowym mającym wykonywać wybraną operację przetwarzania sygnałów audiowizualnych. Poszczególne moduły mogą być zestawiane jeden za drugim, umożliwiając potokowe przetwarzanie danych. Rozwiązanie to zapewnia dużą elastyczność w wykorzystaniu gotowych filtrów. Umożliwia to między innymi wykorzystanie gotowych modułów służących do odczytu danych z plików, rozdzielenia danych na strumienie wideo i audio, dekodowania i prezentacji obrazu (renderowania). Korzystanie z gotowych modułów, zazwyczaj zoptymalizowanych pod kątem szybkości działania, pozwala w krótkim czasie opracować sprawnie działający system przetwarzania strumieni danych. Filtr z zaimplementowanym MDR może być wykorzystany bez konieczności ponownej kompilacji jego kodu źródłowego.

W opracowanym filtrze można dobierać wartości parametrów modelu – możliwy jest wybór jednej z metod obliczania wpływu obrazu na model oraz podział procesu dopasowania na poszczególne etapy. W każdym z nich można niezależnie definiować parametry modelu i określić liczbę iteracji procesu dopasowania.

W dalszych badaniach modelu przyjęto arbitralnie liczbę pierścieni $P = 7$, liczbę węzłów w pierścieniu $Q = 128$, początkowy promień wewnętrznego pierścienia $r = 70$ oraz proporcję między promieniami sąsiednich pierścieni $w = 1,04$. Dobór parametrów jest uzasadniony wielkością obrazu – model z parametrami o takich wartościach mieści się bowiem w polu widzenia kamery endoskopu z pewnym marginesem pola widzenia dookoła, a odległości między pierścieniami oraz węzłami w pierścieniu są porównywalne.

Działanie modelu testowano na komputerze z procesorem Intel Pentium 4 tak-towanym zegarem o częstotliwości 3 GHz i systemem operacyjnym Microsoft Windows XP Professional. Badania przeprowadzono w latach 2006-2008, a uzyskane wyniki opublikowano w pracach [15], [145]. Wielkości dotyczące uzyskiwanych czasów obliczeń podano dla takiej konfiguracji komputera. Algorytm obliczeniowy został zaimplementowany w pojedynczym wątku. Należy jednak zwrócić uwagę, że energie cząstkowe obliczane są niezależnie dla poszczególnych węzłów. Węzły modelu można podzielić na podzbiory i dokonywać w nich obliczeń niezależnie, w osobnych wątkach i w osobnych rdzeniach procesorów. Pozwoli to wielokrotnie przyspieszyć analizę filmów w stosunku do wyników podanych w niniejszym rozdziale.

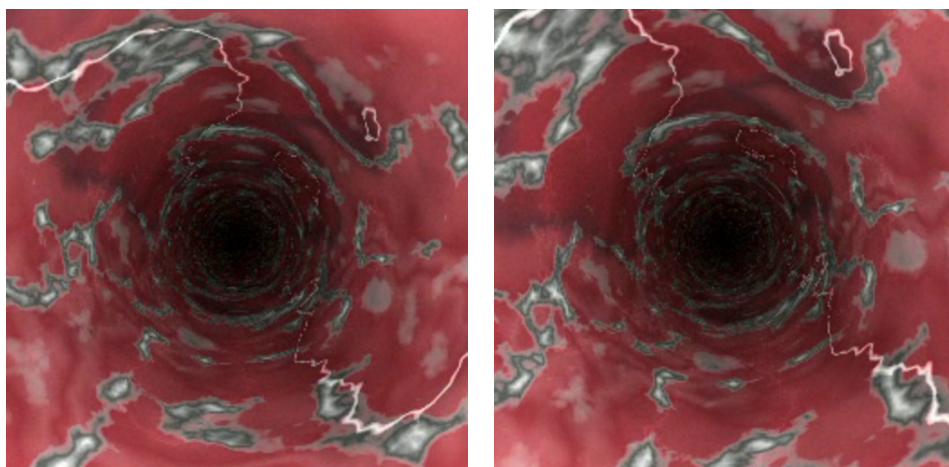
4.1. Dobór metody obliczania wpływu obrazu

Do obliczania wpływu obrazu w MDR zaproponowano opcjonalne wykorzystanie jednej z czterech miar określających różnicę pomiędzy fragmentami obrazu – są to funkcja informacji wzajemnej ze znakiem minus ($-MI$), funkcja znormalizowanej korelacji wzajemnej ze znakiem minus ($-NCC$), średnia wartość różnicy bezwzględnej (MAD) oraz błąd średniokwadratowy (MSE). W związku z tym przeprowadzono eksperyment [11], [14], którego celem była walidacja działania modelu wykorzystującego zaproponowane metody oraz wybór jednej z nich, najbardziej przydatnej w przetwarzaniu wideo z endoskopu bezprzewodowego. Przyjęto, że wybrana metoda przetwarzania musi spełnić następujące warunki: 1) wynikowa mapa przewodu pokarmowego musi przedstawiać obraz odpowiadający właściwościom powierzchni przewodu pokarmowego, 2) model powinien prawidłowo śledzić ruch posuwisty kapsułki wzdłuż przewodu pokarmowego w kierunku zgodnym z osią kapsułki, a także ruch obrotowy wokół tej osi, 3) model musi być odporny na zdarzenia nieprzewidziane w zaproponowanym modelu ruchu, podczas ich występowania algorytm nie może przerywać przetwarzania danych, oraz 4) czas wymagany do przeanalizowania filmu powinien być akceptowalny i nie przekraczać czasu poświęcanego typowo na przeprowadzenie badania. Przyjęto także, że do analizy będą wykorzystane popularne komputery klasy PC.

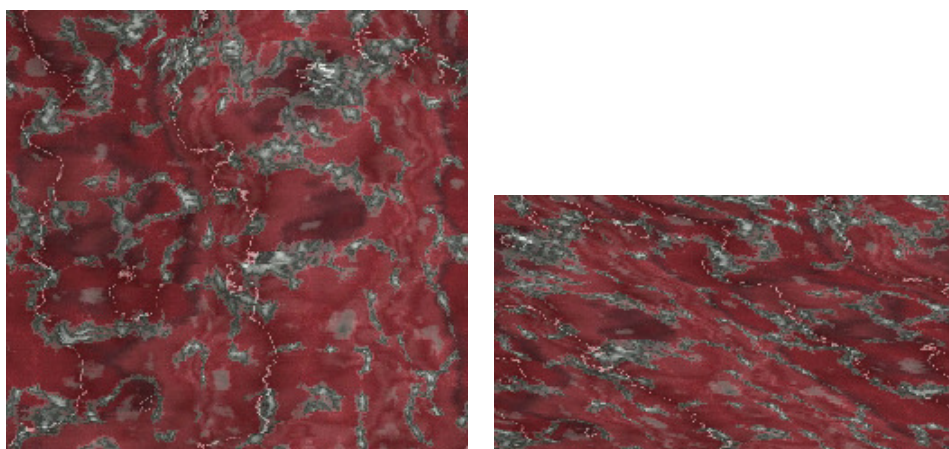
Na potrzeby eksperymentu wygenerowano film przedstawiający syntetyzowany obraz z kamery znajdującej się wewnątrz sztywnej rury o średnicy odpowiadającej średnicy kapsułki endoskopu bezprzewodowego. Na wewnętrzzną stronę wirtualnej rury nałożono teksturę, która barwą i wzorem przypomina powierzchnię ścian układu pokarmowego. Film ten został wygenerowany dla kapsułki poruszającej się ze stałą prędkością 0,19 mm/s, z osią kamery ustawioną równolegle do osi rury i z prędkością obrotową wokół tej osi równą 7,2 °/s. Film składa się z sekwencji 101 obrazów, co oznacza, że dla tej sekwencji proces dopasowania powtarzany jest sto razy. Do wygenerowania filmu wykorzystano program *Persistence of Vision Raytracer (POV-Ray)* [223].

Wygenerowany film jest bardzo dużym uproszczeniem w stosunku do rzeczywistych filmów uzyskiwanych z endoskopu bezprzewodowego. Nie odzwierciedla on w sposób realistyczny zróżnicowania ruchu endoskopu ani zmienności

kształtu przewodu pokarmowego. Celem eksperymentu nie jest jednak wykazanie skuteczności działania proponowanego algorytmu w warunkach rzeczywistych, ale ilościowe porównanie jego czterech wariantów do czego nawet tak wyidealizowany przypadek jest wystarczający.



Rys. 4.1. Obrazy 15. i 20. wygenerowanej sztucznie sekwencji wideo przedstawiającej wnętrze teksturowanej rury



Rys. 4.2. Wynikowe mapy uzyskane z wykorzystaniem funkcji MAD, z blokiem obrazu 3×3 z parametrem $\xi = 0,1$ (po lewej) i $\xi = 0,02$ (po prawej). Mapa po prawej stronie przedstawia przekrzywioną teksturę, co jest wynikiem zaniżenia wartości prędkości kątowej, a także jest krótsza w wyniku zaniżenia wartości szybkości postępowej

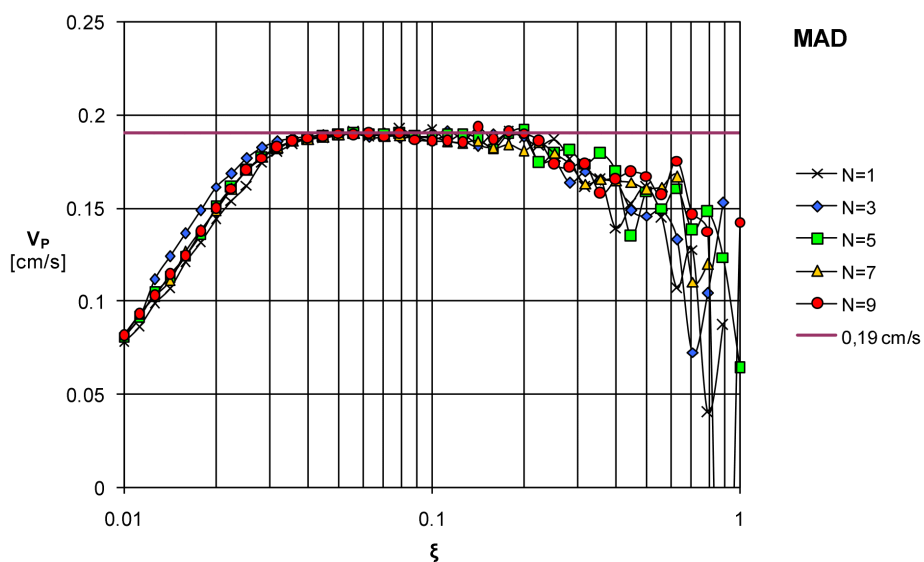
Wygenerowany film poddano analizie za pomocą deformowalnych pierścieni. Przyjęto, że dopasowanie modelu dla każdego obrazu przebiega w dwóch etapach. W pierwszym z nich liczba iteracji została ustalona na $I_{ml} = 10$, parametr określający wielkość η -sąsiedztwa dobrano tak, aby dla każdego węzła jego η -są-

siedztwo obejmowało wszystkie węzły modelu. Parametry określające sposób obliczania naprężeń ustalono na $\rho = 1$, $\rho_S = 0,9$ i $\rho_T = 0,95$. W drugim etapie obejmującym $I_{m2} = 4$ iteracje, wartości odnośnych parametrów wynosiły $\eta = 2$, $\rho = 0,5$, $\rho_S = 0,9$ i $\rho_T = 1$.

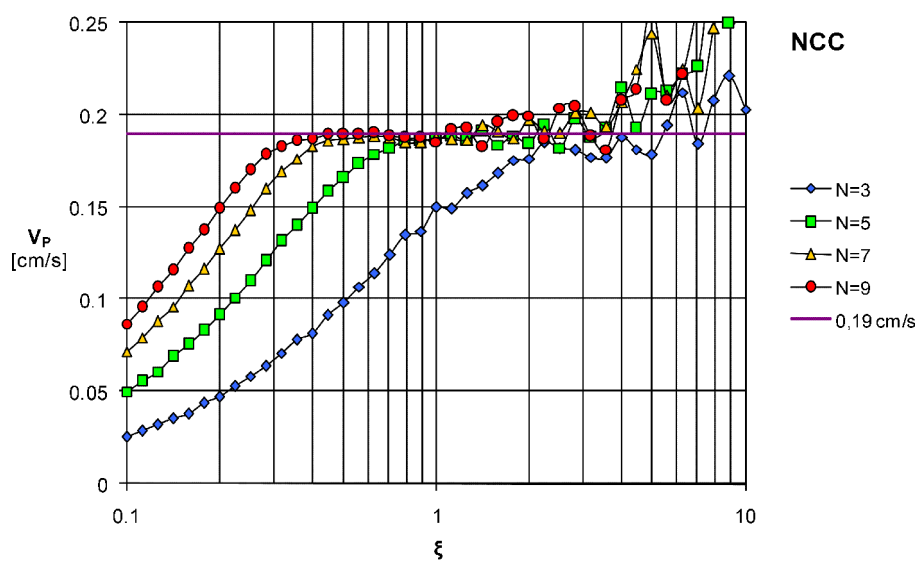
Dla tak ustalonych parametrów, związanych z obliczaniem naprężeń, przetestowano model dla czterech zaproponowanych metod obliczania wpływu obrazu, kilku wartości parametru określającego wielkość bloku obrazu ($N = 1, 3, 5, 7$ i 9) oraz różnych wartości parametru udziału energii obrazu ζ . Zakres zmienności parametru ζ był dobierany odrębnie dla każdej badanej metody obliczania wpływu obrazu, w taki sposób, aby uzyskać prawidłowe działanie modelu wewnątrz zakresu zmienności tego parametru. Wartość parametru ζ w drugim etapie dopasowania była zmniejszana do połowy. Zróżnicowanie wartości parametrów w obu etapach umożliwia zgrubne i szybkie dopasowanie modelu w pierwszym etapie oraz bardziej subtelne i szczegółowe w drugim.

W przypadku zastosowania funkcji NCC, MAD i MSE analizowane są obrazy w skali szarości o 256. poziomach, kodowane ośmioma bitami na piksel. W przypadku miary MI zastosowanie takiej liczby poziomów jasności wymaga długich czasów obliczeń, sięgających 2 minut na obraz (dwa miesiące dla filmu złożonego z 50 000 obrazów). Dlatego w przypadku funkcji MI liczbę poziomów jasności ograniczono do 8. przy kodowaniu jasności trzema bitami na piksel, co skraca czas obliczeń blisko trzydziestokrotnie.

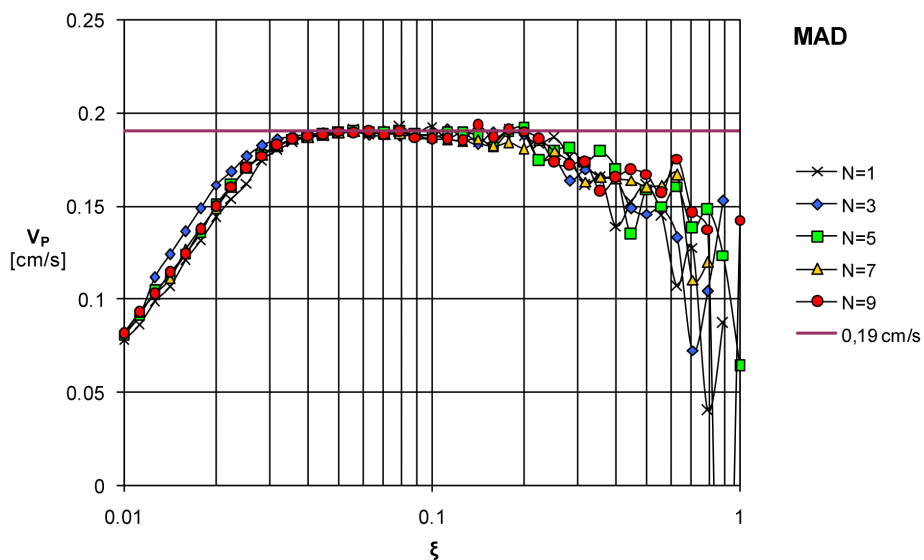
Wynikiem działania opisywanego modelu jest mapa powierzchni wewnętrznej oraz deskryptory ruchu. W przypadku mapy dokonano subiektywnej oceny proporcji uzyskanej tekstury oraz jej ciągłości. W przypadku deskryptorów porównano uzyskane wartości prędkości postępowej V_P oraz prędkości obrotowej ω ze znanymi wartościami tych prędkości wykorzystanymi podczas generowania filmu. Zmierzono również czas obliczeń wymagany do przeprowadzenia procesu dopasowania, czyli dla pojedynczego obrazu sekwencji. Rysunek 4.1 przedstawia przykładowe obrazy wygenerowanej sekwencji wideo, a na rysunku 4.2 przedstawiono wybrane przykłady map uzyskanych za pomocą MDR. Na rysunku 4.3 zamieszczono wykresy prędkości postępowej obliczone dla testowanych metod przy różnych wartościach parametrów N i ζ , a na rysunku 4.4 zamieszczono analogiczne wykresy prędkości obrotowej. Tabela 4.1 przedstawia wartości czasów wymaganych do przeprowadzenia obliczeń dla poszczególnych metod przy różnych wartościach parametru N .



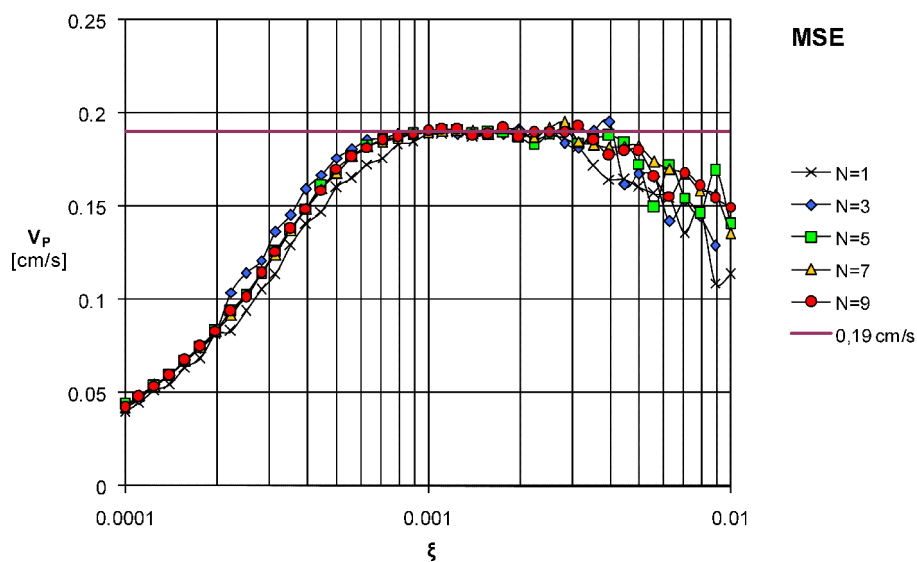
a)



b)

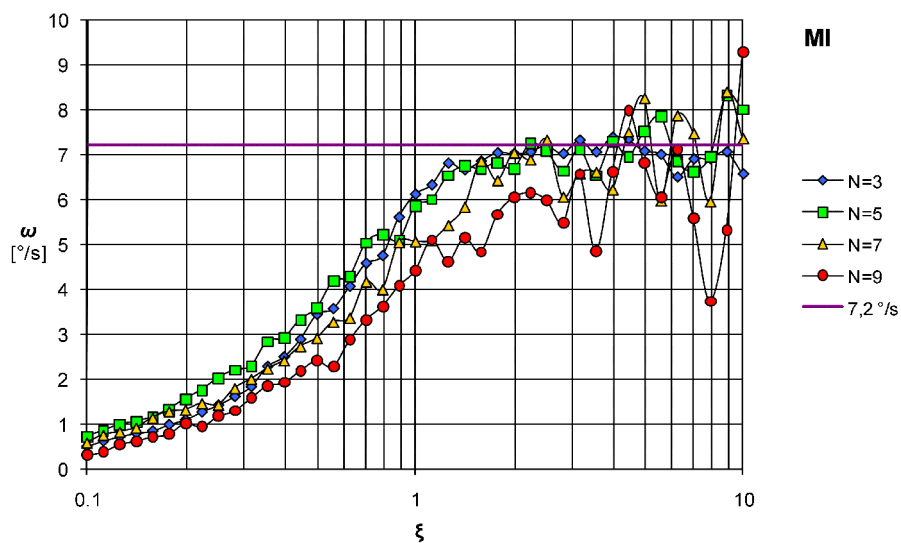


c)

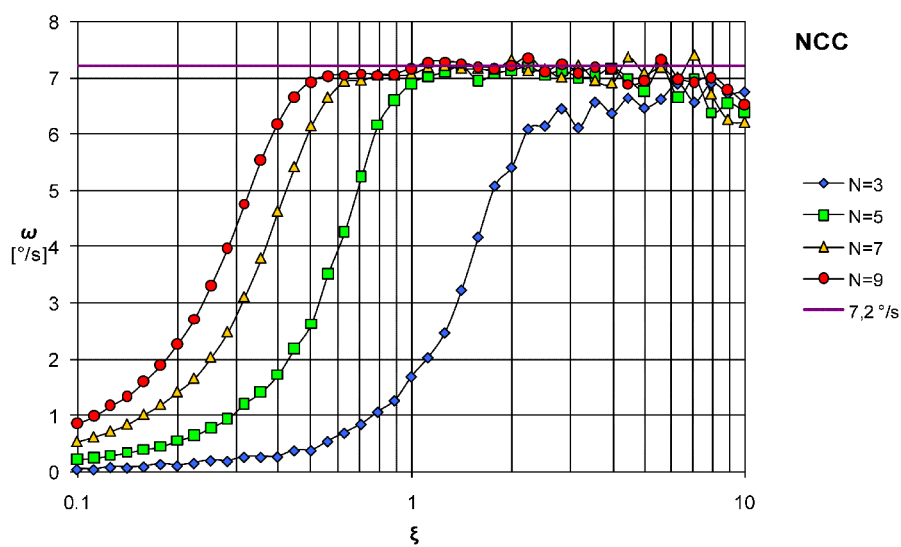


d)

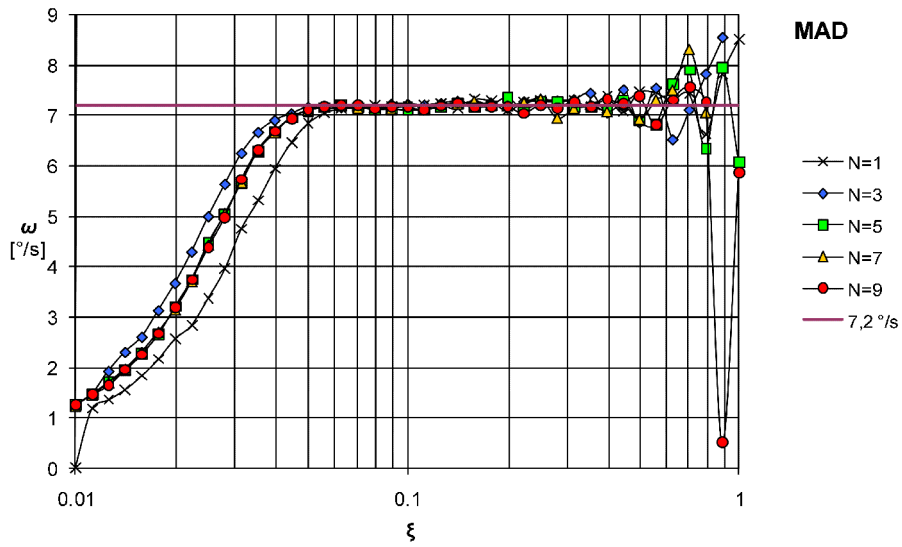
Rys. 4.3. Wykresy estymowanej prędkości postępowej V_P w funkcji ξ uzyskane dla miar: a) MI, b) NCC, c) MAD i d) MSE



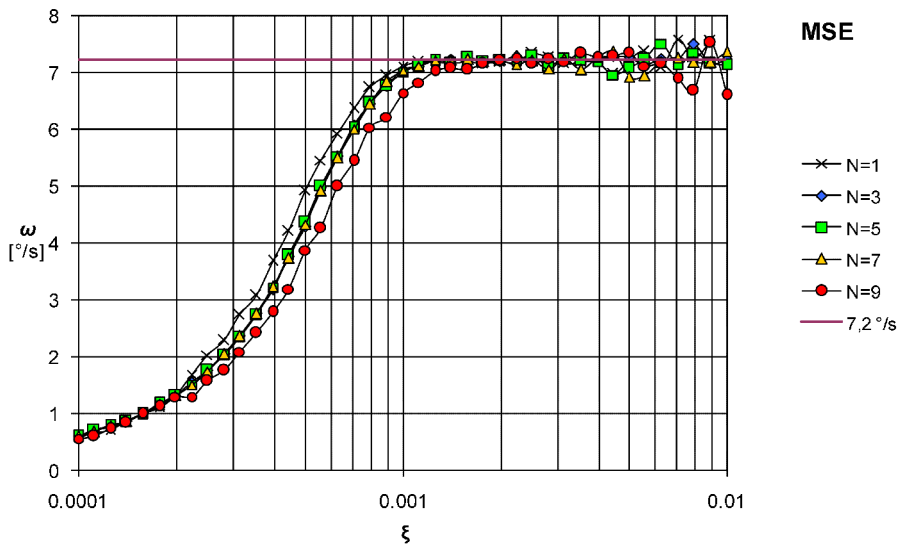
a)



b)



c)



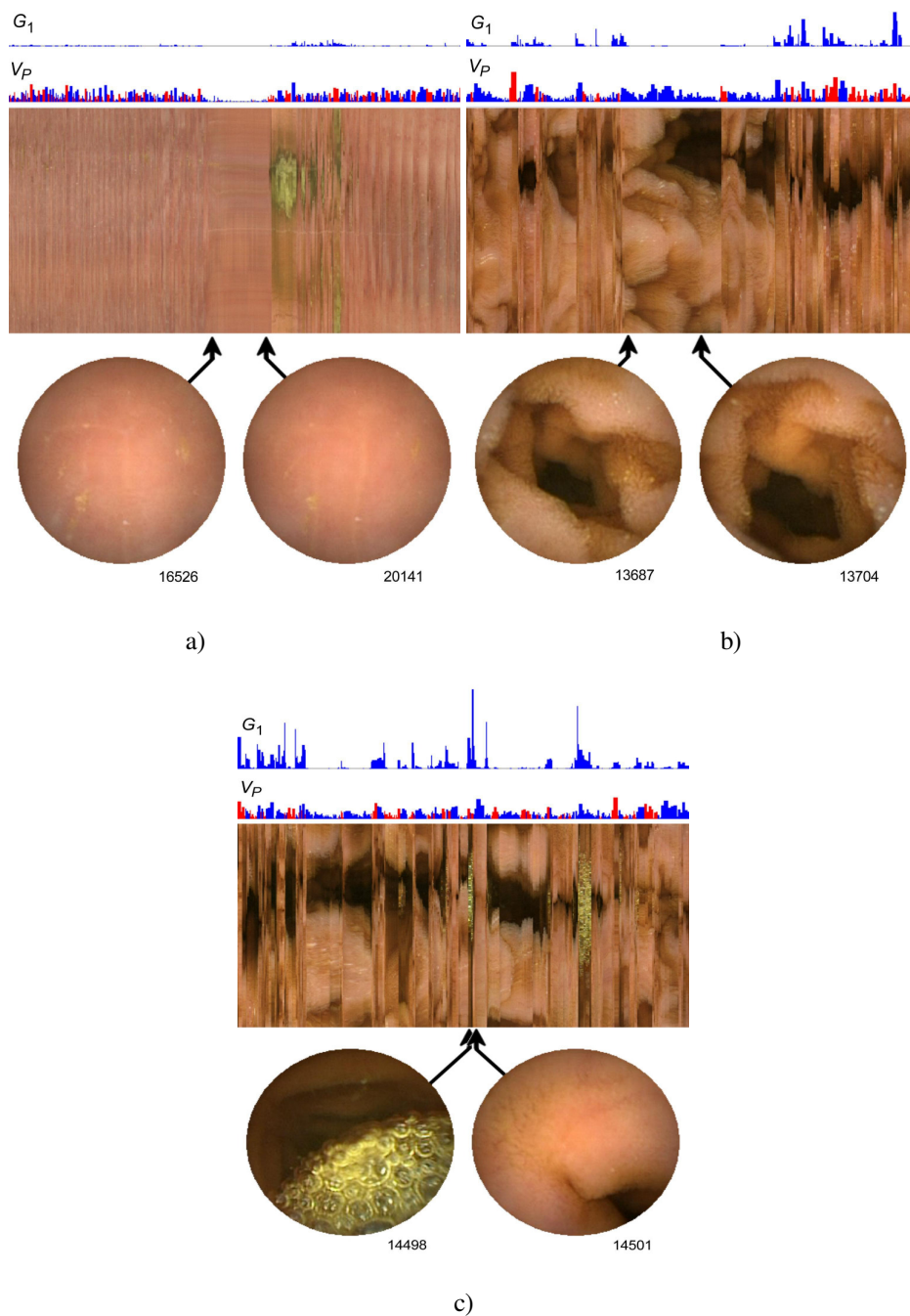
d)

Rys. 4.4. Wykresy estymowanej prędkości obrotowej ω w funkcji ξ uzyskane dla miar: a) MI, b) NCC, c) MAD i d) MSE

Tabela. 4.1. Czas analizy jednego obrazu w milisekundach

N	MAD	MSE	MI	NCC
1	18	18	N.D.	N.D.
3	37	38	132	84
5	75	75	184	214
7	127	132	238	397
9	200	209	327	654

Analiza wyników zamieszczonych na rysunkach 4.3 i 4.4 pozwala zaobserwować, że dla wszystkich testowanych metod obliczania wpływu obrazu istnieje taka para parametrów N i ξ , dla których wartości prędkości obliczone przez MDR są zbliżone do oczekiwanych. W przypadku funkcji MAD, MSE i NCC zaobserwować można pewien zakres parametru ξ , dla którego na wykresach szybkości występuje plateau o wartościach bliskich oczekiwany. Oznacza to, że dla tych trzech metod istnieje pewien zakres parametru ξ , w którym działanie modelu można uznać za prawidłowe. W przypadku funkcji NCC takie płaskie przebiegi wykresu uzyskano dla bloków o wielkości przynajmniej 5×5 . Dla funkcji MAD i MSE płaskie przebiegi bliskie wartości oczekiwanej uzyskano dla mniejszych bloków, a nawet dla „bloków” o wielkości 1×1 (pojedyncze piksele). Jeśli porównać czasy wymagane do przeprowadzenia dopasowania, to są one najkrótsze dla funkcji MAD i MSE. Porównując płaskie fragmenty wykresów funkcji MAD i MSE można zaobserwować, że są one najszersze dla funkcji MAD. W tym przypadku wartości prędkości bliskie prędkościom oczekiwany, z błędem mniejszym niż 5%, uzyskuje się dla parametru ξ w zakresie od 0,03 do 0,2, co stanowi ponad sześciokrotną zmienność. W przypadku funkcji MSE zakres zmienności parametru ξ , przy którym uzyskuje się wartości bliskie oczekiwany, jest mniejszy – czterokrotny. Dla obu metod, MAD i MSE, wartości funkcji V_p w płaskim obszarze wykresu są największe, a poza nim maleją. Jest to istotne spostrzeżenie, które umożliwia optymalizację parametru ξ dla rzeczywistych sekwencji wideo przez maksymalizację wartości V_p . Biorąc pod uwagę te spostrzeżenia, do obliczania wpływu obrazu na MDR, wybrano funkcję MAD wykorzystującą blok obrazu o wielkościach 3×3 piksele. Dla potrzeb analizy filmów endoskopowych parametr ξ dobrano tak, aby maksymalizował wartości funkcji szybkości postępowej V_p w wybranych kilku rzeczywistych filmach uzyskanych z kapsułki endoskopowej. W dalszych eksperymentach przyjęto wartość tego parametru równą 0,1 w pierwszym etapie i 0,05 w drugim etapie dopasowania.



Rys. 4.5. Przykładowe mapy przewodu pokarmowego, wybrane obrazy, dla których zostały wygenerowane oraz poglądowe wykresy deskryptorów ruchu: a) bezruch, b) wynik skurczów perystaltycznych i c) wynik skurczów segmentowych. Oznaczenia liczbowe przy obrazach oznaczają numer kolejny obrazu w filmie

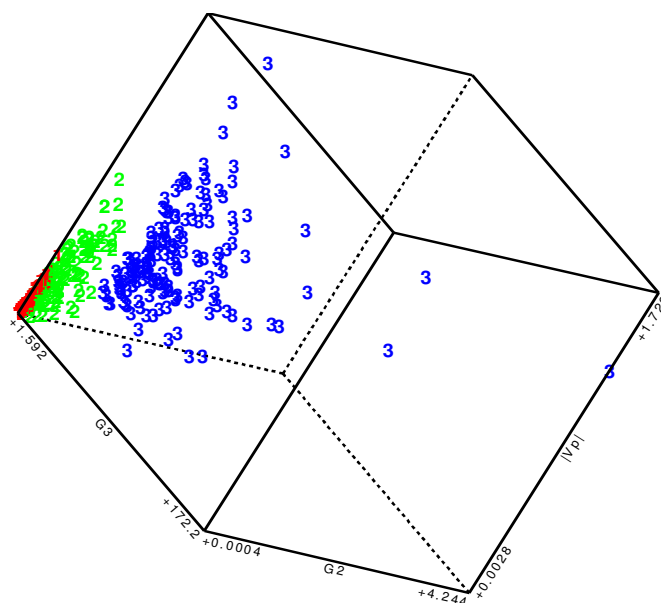
Działanie modelu zostało przetestowane na kilkudziesięciu filmach uzyskanych z endoskopu bezprzewodowego. We wszystkich przypadkach ekspert gastroenterolog uznał, że wynikowa mapa przedstawia przydatne informacje, model śledzi ruch posuwisty i nie przerywa analizy filmu w przypadku występowania innych rodzajów ruchu.

4.2. Model ruchów jelita

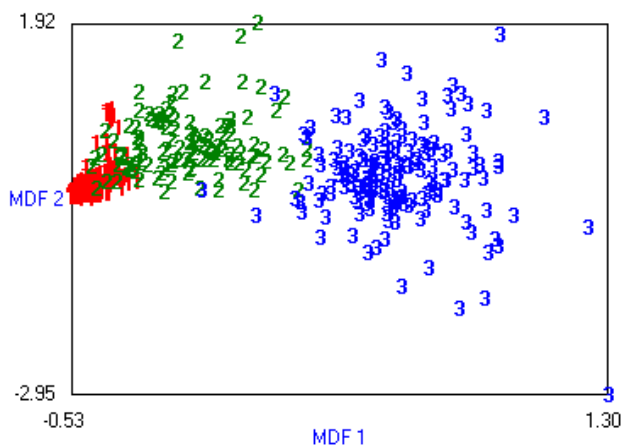
Ruch kapsułki endoskopu bezprzewodowego w jelicie cienkim jest wynikiem ruchów jelita wywoływanych skurczami mięśni oraz w pewnym stopniu sił grawitacji. Ruch jelita cienkiego nie jest jednorodny, a ze względu na jego intensywność można wyróżnić w nim trzy przypadki. Najbardziej intensywny i chaotyczny jest ruch wywołany skurczami segmentowymi, mieszającymi pokarm. Mniej intensywne ruchy perystaltyczne, których celem jest przesuwanie pokarmu, wywołane są skurczem obwodowym, którego miejsce przemieszcza się w sposób jednostajny w stronę odbytu. Trzecim przypadkiem jest bezruch, w którym jelito pozostaje w stanie spoczynku.

Rodzaj ruchu jelita ma wpływ na intensywność zmian treści obrazu. W fazie bezruchu kolejne obrazy praktycznie się nie zmieniają, podczas ruchów perystaltycznych endoskop jest przesuwany stosunkowo łagodnie, a obraz zmienia się płynnie i jednostajnie. Skurcze mieszające powodują gwałtowne zmiany w obrazie. Jak wspomniano, praktyczne zastosowanie analizy ruchu we wspomaganie interpretacji medycznej filmu z endoskopu bezprzewodowego polega na dostosowaniu jego odtwarzania do intensywności zmian w obrazie. Film może być odtwarzany szybko w przypadku braku zmian w obrazie (faza spoczynku), wolniej w przypadku ruchów jednostajnych (skurcze perystaltyczne) oraz bardzo wolno w przypadku nagłych i chaotycznych zmian obrazu (ruchy segmentowe).

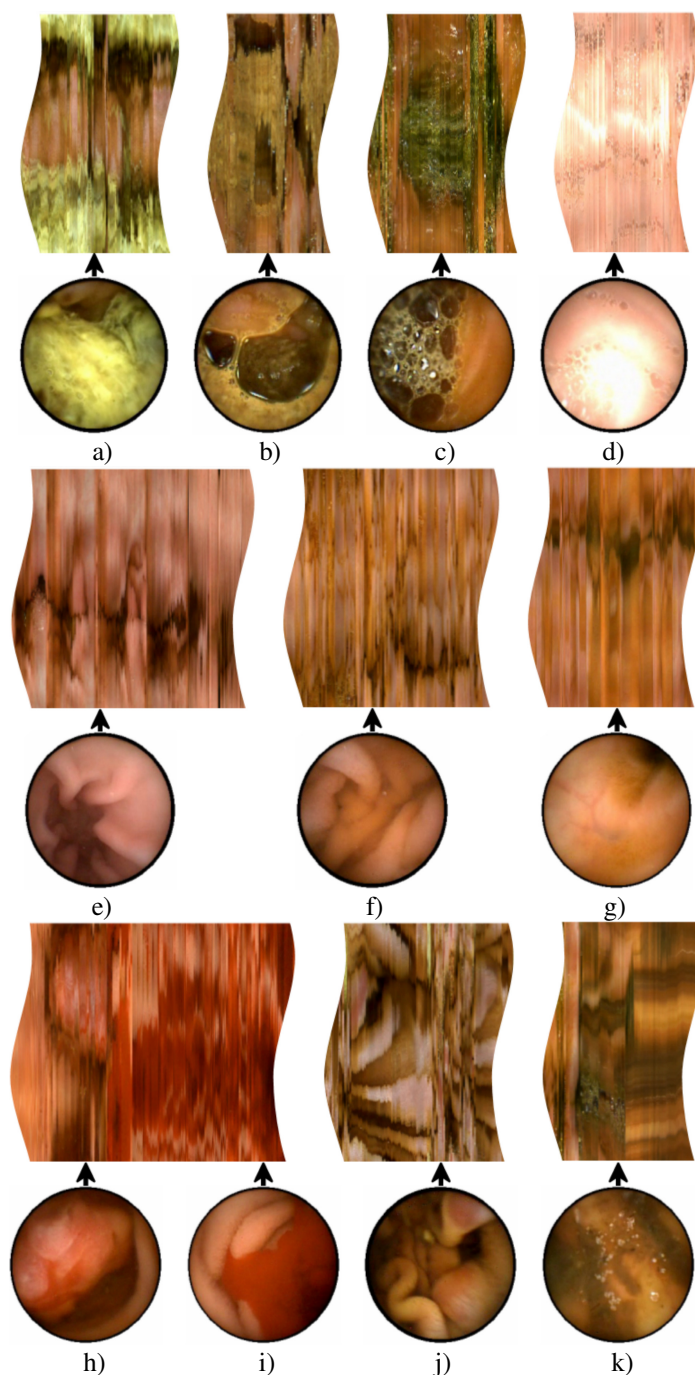
Celem kolejnego eksperymentu [11], [17] było potwierdzenie, że deskryptory ruchu obliczane przez MDR umożliwiają klasyfikację ruchów jelita cienkiego. W eksperymencie wykorzystano 50 filmów pochodzących z badań medycznych wykonanych za pomocą endoskopu bezprzewodowego. Filmy te zostały poddane analizie za pomocą MDR. Następnie odszukano w nich fragmenty przedstawiające różne fazy ruchu jelita cienkiego. Po konsultacji z lekarzem, wybrano po sto obrazów należących do fragmentów reprezentujących bezruch, ruch wywołany skurczami przepychającymi oraz ruch wywołany skurczami mieszającymi. Wektory deskryptorów ruchu V_p , ω , G_1 , G_2 i G_3 obliczone dla tych obrazów, z oznaczeniem ich przynależności do klas, poddano liniowej analizie dyskryminacyjnej (ang. *Linear discriminant analysis* – *LDA*) [224], [225].



Rys. 4.6. Rozkład wektorów deskryptorów ruchu. Cyframi oznaczono przynależność do poszczególnych klas: 1 – bezruch, 2 – skurcze perystaltyczne i 3 – skurcze segmentowe



Rys. 4.7. Rozkład wektorów w przestrzeni MDF



Rys. 4.8. Przykłady fragmentów map oraz wybrane obrazy, na podstawie których mapy zostały wygenerowane. Przykłady pokazują: a) treść pokarmową, b) i c) pianę, d) prześwietlony obraz, e) wejściowy odcinek żołądka, f) jelito cienkie, g) jelito grube, h) wrzód w chorobie Crohna-Leśniewskiego, i) krwawienie, j) erozję oraz k) miejsce zatrzymania się kapsułki z bliżej nieokreślonego powodu

Liniowa analiza dyskryminacyjna umożliwia znalezienie liniowej transformacji przestrzeni cech (tutaj deskryptorów ruchu) w nową przestrzeń atrybutów najlepiej różnicujących klasy (ang. *Most discriminative features – MDF*). Cechy MDF będące kombinacjami liniowymi cech oryginalnych są uporządkowane w kolejności określonej ich zdolnością dyskryminacyjną. Przy czym pierwsza z nich najlepiej różnicuje (dyskryminuje) klasy zdarzeń (tutaj fazy ruchu), zaś każda kolejna tworzy kierunek ortogonalny w stosunku do wszystkich poprzednich. Określenie „najlepiej różnicujące” odnosi się do kryterium Fishera i oznacza maksymalizację wartości, tzw. współczynnik Fishera [226], który wyraża zdolność dyskryminacyjną danej cechy jako iloraz średnich wariancji cechy pomiędzy klasami do jej średniej wariancji wewnątrz poszczególnych klas.

Rysunek 4.6 przedstawia rozkład wektorów trzech wybranych deskryptorów ruchu $|V_p|$, G_1 i G_3 . Współrzędne wektorów klasy bezruchu oznaczono cyfrą „1”, klasy skurczy perystaltycznych cyfrą „2” i klasy skurczy segmentowych cyfrą „3”. Można zaobserwować, że wartości deskryptorów są najmniejsze w klasie bezruchu i największe w klasie skurczy segmentowych, a także, że wektory poszczególnych klas znajdują się w osobnych obszarach przestrzeni parametrów. Odseparowanie wektorów należących do osobnych klas jest możliwe – z założeniem pewnego akceptowalnego błędu – za pomocą hiperpłaszczyzny, czyli klasyfikacji liniowej.

Rysunek 4.7 przedstawia natomiast przestrzeń dwóch pierwszych cech MDF uzyskanych za pomocą transformacji liniowej przestrzeni wszystkich pięciu wymienionych deskryptorów ruchu. Klasyfikacja liniowa z wykorzystaniem wyłącznie pierwszej cechy MDF umożliwia prawidłowe przyporządkowanie ponad 90% wektorów. Wyniki te potwierdzono za pomocą zbioru testowego wektorów cech, stosując otrzymany klasyfikator liniowy do klasyfikacji faz ruchu w filmach innych niż wykorzystane na etapie jego uczenia.

4.3. Sterowanie odtwarzaniem filmu

Jednym z zastosowań MDR jest wykorzystanie wygenerowanych deskryptorów ruchu do sterowania szybkością odtwarzania filmu. Po konsultacjach z gastroenterologiem ustalono, że obrazy przedstawiające intensywny ruch będący wynikiem skurczy segmentowych powinny być prezentowane przez około pół sekundy każdy. Jest to czas wystarczająco długi, aby spojrzeć na obraz, ocenić go i ewentualnie podjąć decyzję o zatrzymaniu odtwarzania w celu przeprowadzenia dokładniejszej analizy. Ustalono również, że obrazy powstałe podczas jednostajnego i płynnego ruchu, będącego wynikiem skurczy perystaltycznych, mogą być odtwarzane szybciej, przez czas około 1/10 sekundy. Z kolei w przypadku obrazów uzyskanych podczas bezruchu uznano, że mogą być prezentowane przez możliwie krótki czas – ze względu na częstotliwość odświeżania typowych współczesnych monitorów komputerowych, czas prezentacji pojedynczego obrazu wyniesie 1/75 sekundy.

Do estymacji zależności między deskryptorami ruchu a czasem prezentacji obrazu wykorzystano regresję liniową [227–229], w której zakłada się, że wartość oczekiwana funkcji może być przybliżona kombinacją liniową zmiennych nieza-

leżnych. Poszukuje się tu takiej kombinacji, która zminimalizuje błąd pomiędzy wartością oczekiwaną a jej przybliżeniem. W przypadku estymacji funkcji czasu Δg prezentacji obrazu minimalizowano błąd średniokwadratowy. Uzyskaną kombinację liniową deskryptorów przedstawia równanie (4.3.1).

$$\Delta g = \psi (0.1332 |V_p| + 0.0037 |\omega| - 0.3871 G_1 + 0.0007 G_2 + 0.0135 G_3 - 0.0315) \quad (4.3.1)$$

Parametr ψ domyślnie jest równy jeden, jego wartość może być modyfikowana przez użytkownika w celu względnego przyspieszenia lub spowolnienia szybkości odtwarzania filmu. Ponadto w opracowanym programie wartość Δg jest ograniczona z dołu przez czas odświeżania obrazu wykorzystywanego monitora oraz z góry przez wartość pół sekundy.

4.4. Zastosowanie mapy przewodu pokarmowego

MDR przetestowano na blisko 60 filmach z endoskopu bezprzewodowego. Dla każdego z filmów wygenerowano mapę przewodu pokarmowego o szerokości 256 pikseli. Każdy z wierszy mapy powstał w wyniku zebrania informacji z obrazu we współrzędnych 128 węzłów i w 128 położeniach pomiędzy sąsiednimi węzłami w pierścieniu. Długość map waha się pomiędzy 15000 a 25000 pikseli.

Przydatność map została poddana subiektywnej ocenie diagnostów gastroenterologów [11], [18] – stwierdza się w niej, że mapy umożliwiają szybkie rozpoznanie obszarów krwawień, rozległych owrzodzeń oraz fragmentów, w których powierzchnia jelita jest przesłonięta treścią pokarmową. W czasie interpretacji filmu, dzięki zaproponowanej mapie, można więc pominąć fragmenty nieużyteczne z punktu widzenia diagnostyki medycznej i skupić uwagę na istotnych obszarach. Ponadto wykres prędkości V_p obliczanej przez MDR przydatny jest podczas lokalizacji miejsc, w których kapsułka się zatrzymała – może to wskazywać na przewężenia lub dysfunkcje przewodu pokarmowego. Spostrzeżenia te pozwalają sądzić, że korzystanie z mapy przyspiesza wyszukiwanie zmian chorobowych układu pokarmowego lub zmniejsza liczbę zmian przeoczonych przez lekarza podczas interpretacji filmu.

W celu ilościowej weryfikacji powyższych spostrzeżeń przeprowadzono badania doświadczalne [11], w których udział wzięło dwóch gastroenterologów zajmujących się zawodowo badaniami z zastosowaniem endoskopu bezprzewodowego. Na potrzeby eksperymentu przygotowano 10 nowych filmów wideo, których lekarze biorący udział w eksperymencie wcześniej nie znali. Wszystkie filmy poddano analizie z zastosowaniem MDR, dzieląc je następnie na dwie grupy po pięć filmów. Pierwszy z lekarzy otrzymał filmy z pierwszej grupy wraz z odpowiadającymi im mapami przewodu pokarmowego oraz filmy z drugiej grupy bez map i dodatkowych informacji wygenerowanych przez MDR. Drugi lekarz otrzymał filmy z drugiej grupy wraz z odpowiednimi mapami, natomiast filmy z pierwszej grupy bez map i dodatkowych informacji.

Tabela 4.2. Liczby znalezisk i czasy interpretacji

Numer filmu	Brak mapy		Wykorzystanie mapy	
	Liczba znalezisk	Czas interpretacji [min]	Liczba znalezisk	Czas interpretacji [min]
	Lekarz 1		Lekarz 2	
1	9	54	11	38
2	11	32	12	11
3	18	38	6	18
4	22	55	12	24
5	10	35	11	11
	Lekarz 2		Lekarz 1	
6	14	25	24	44
7	25	45	55	49
8	4	19	14	40
9	16	34	8	65
10	18	30	28	60
Sumarycznie	147	367	181	360

Zadaniem lekarzy była interpretacja otrzymanych filmów, odnalezienie charakterystycznych miejsc przewodu pokarmowego – punktów orientacyjnych (ang. *landmarks*) – takich jak początek i koniec jelita cienkiego oraz obrazów przedstawiających zmiany chorobowe. Odnalezione punkty orientacyjne i obrazy ze zmianami chorobowymi zostały przez lekarzy opisane, a opisy wraz z numerami obrazów umieszczone na tzw. liście znalezisk (ang. *findings*). W przypadku każdego z filmów zmierzono czas, jaki lekarz poświęcił na jego interpretację. Po zakończeniu eksperymentu, wytworzone listy znalezisk zostały wzajemnie zweryfikowane przez lekarzy biorących udział w eksperymencie. Potwierdzili oni, że odnalezione obrazy rzeczywiście przedstawiają punkty orientacyjne, zmiany chorobowe oraz że opisy umieszczone na liście są adekwatne do odwzorowanych w tych obrazach informacji medycznych.

Eksperyment pokazał, że przeciętny czas poświęcony przez lekarzy na przeprowadzenie interpretacji był nieznacznie krótszy w przypadku zastosowania mapy aniżeli w przypadku, w którym nie była ona wykorzystana. Sumaryczny czas analizy filmów bez zastosowania mapy wyniósł 367 minut, natomiast z jej wykorzystaniem 360 minut. Sumaryczna liczba znalezisk dla filmów wyposażonych w mapy przewodu pokarmowego wyniosła 181, natomiast dla filmów bez dodatkowych informacji wygenerowanych przez MDR wyniosła 141. Dane szczegółowe dotyczące liczby znalezisk i czasu interpretacji poszczególnych filmów przedstawia tabela 4.2.

Zamieszczone tam wyniki wykazały przydatność mapy przewodu pokarmowego – co prawda nie uzyskano znaczącej redukcji czasu poświęconego na interpretację, zaledwie 2%, ale zwiększenie liczby znalezisk o 20% jest istotnym osiągnięciem. Należy zauważyć, że odnalezienie jak największej liczby obrazów ze zmianami chorobowymi umożliwia prawidłową ocenę ilościową schorzenia. Ocena ta ma natomiast wpływ na zalecenia dotyczące sposobu leczenia. Tak więc, metoda zmniejszająca szanse przeoczenia zmiany chorobowej ma istotne znaczenie w uzyskaniu prawidłowej diagnozy.

5. Detekcja zmian chorobowych

Automatyczna detekcja zmian chorobowych uwidocznionych w obrazach biomedycznych jest jednym z istotnych zagadnień diagnostyki w medycynie. Przyjmuje się, że obrazy zmian różnią się w znaczący sposób od obrazów przedstawiających stan normalny. Metody omawiane w niniejszym rozdziale mają za zadanie sprecyzowanie co jest istotą tych różnic i określenie cech będących wyróżnikami między obrazami jednej i drugiej grupy. Cechy takie, nazywane też deskryptorami, muszą zostać zdefiniowane w sposób umożliwiający ich wyodrębnienie metodami numerycznymi – są one wartościami liczbowymi charakteryzującymi obraz w sposób ilościowy.

Można wyróżnić trzy metody poszukiwania cech znaczących, czyli takich, które odróżniają obrazy należące do różnych grup – klas obrazów. Są to metody wykorzystujące model obrazu, transformację obrazu oraz analizę statystyczną. W pierwszym przypadku opracowywane są modele obrazów obu klas uwzględniające fizyczne różnice między obrazowanymi tkankami oraz właściwości wykorzystanej techniki obrazowania. Przykładowo, przy detekcji krwawień można rozpatrywać skład chemiczny krwi, jej cechy optyczne, takie jak barwa i mętność, oraz jej cechy fizyczne, jak płynność i zdolność dyfuzji. Następnie, uwzględniając charakter układu optycznego i oświetlenia, określa się cechy jakie powinien mieć obraz, w którym widoczna jest krew. Cechy takie, np. udział barwy czerwonej odpowiedniej dla hemoglobiny, można wykorzystywać jako wyróżnik obrazów przedstawiających krwawienie. Drugim podejściem do problemu jest zastosowanie metod transformacyjnych. Wykorzystują one transformacje obrazów, które w łatwiejszy sposób niż oryginały pozwalają odróżnić obrazy należące do różnych klas – np. transformata falkowa umożliwia obliczenia składowych częstotliwościowych dla różnych skal obrazu. Trzecim podejściem jest zastosowanie metod analizy statystycznej, w których cechami są np. wielkości statystyczne charakteryzujące histogramy obrazu.

W dalszej części rozdziału uwagę poświęcono głównie metodom transformacyjnym i analizie statystycznej oraz zagadnieniom selekcji cech znaczących za pomocą wybranych metod klasyfikacji i uczenia nadzorowanego. Opracowanie modelu obrazu uwzględniającego dużą różnorodność rozważanych zmian chorobowych oraz zmienność warunków pozyskiwania obrazów tych zmian w przypadku obrazów z endoskopu wydaje się niecelowe – model taki musiałby być bardzo skomplikowany i byłby trudny w interpretacji.

5.1. Tekstura

Tekstura jest pewnym wzorem w obrazie, często złożonym z powtarzających się regularnie drobniejszych, podobnych do siebie elementów. Obszar obrazu obejmujący taki wzór jest postrzegany przez człowieka jako jednorodny w sensie subiektywnie przypisywanych mu cech. Tekstura w przypadku obrazów wizyjnych może być obrazem fakturowanych powierzchni (np. tynk, powierzchnia betonu), powierzchni będącej złożeniem drobnych elementów o przypadkowym charakterze (trawnik, piasek, żwir, futro) oraz wzorów tworzonych przez połączenie elementów regularnych (ściana z cegły, pleciona tkanina). Człowiek postrzega tego typu obszary w sposób całościowy, nie zwracając uwagi na poszczególne elementy, które go tworzą. Tekstura ma cechy wynikające z charakteru wizualizowanych obiektów oraz warunków panujących podczas ich wizualizacji. Zazwyczaj jest więc charakteryzowana poprzez cechy wizualizowanego obiektu, który może mieć określoną granulację (piasek i żwir), cechy kierunkowe (trawa, futro), chropowatość (tynk), regularność (ściana z cegły) itp. Cechami wynikającymi z warunków obrazowania jest np. kontrast tworzonego wzoru, który w obrazowaniu powierzchni fakturowanych jest po części efektem sposobu ich oświetlenia.

Patrząc przez pryzmat pojęcia tekstury, obraz można potraktować jako mozaikę obszarów o jednorodnych wzorach. Jeśli celem jest **segmentacja** tak rozumianego obrazu, to analiza polega na lokalnym określaniu cech tekstury, a następnie oznaczaniu obszarów jednorodnych ze względu na wartości tych cech. Innym celem jest rozpoznawanie – **detekcja** obszarów charakteryzujących się określonym wzorem – również wymagające wyznaczenia cech, a następnie stwierdzenia ich podobieństwa do cech poszukiwanego wzoru. W obydwu przypadkach konieczne jest obliczenie cech charakteryzujących teksturę w sposób ilościowy – cechy charakteryzujące ją w sposób opisowy lub jakościowy są niewystarczające w przypadku stosowania metod numerycznych do ich analizy.

Publikacje poświęcone ilościowemu opisowi tekstury [230–237] są źródłem wielu algorytmów umożliwiających scharakteryzowanie rozkładu jasności w obrazie. Wśród nich są metody analizy statystycznej histogramu, metody transformacyjne, umożliwiające analizę składowych częstotliwościowych obrazu traktowanego jak sygnał dwuwymiarowy, oraz oparte na modelach akwizycji i transmisji. Do analizy obrazów z endoskopu bezprzewodowego wykorzystano wybrane metody analizy statystycznej, metody transformacyjne oraz opierające się na modelu transmisji obrazu. Zostały one zaimplementowane w programie MaZda [103–106] i wykazały swoją przydatność w analizie różnorodnych obrazów biomedycznych [113], [116–119], [121], [128], [129], [131–133], [135–139], [238–250]. Wyniki uzyskane tymi metodami porównano z metodami objętymi standardem MPEG-7 [86], [87], [251–253], które były wykorzystywane do badania obrazów endoskopowych przez inne zespoły badawcze [75], [79], [85] na świecie. W dalszej części tego podrozdziału podano krótką charakterystykę metod ilościowego opisu tekstury wykorzystywanych w programie MaZda oraz określonych standardem MPEG-7.

Jedną z metod wykorzystywanych w programie MaZda jest analiza histogramu pierwszego rzędu obrazu monochromatycznego [112], [230]. Histogram jest

funkcją określoną w dziedzinie jasności obrazu, która podaje liczbę pikseli o danej jasności występujących w wybranym obszarze. Cechami są wielkości statystyczne histogramu, takie jak: średnia, wariancja, współczynnik skośności, kurtoza oraz kwantyle. Niedoskonałością opisu obrazu za pomocą cech histogramu pierwszego rzędu jest nieuwzględnienie zależności przestrzennych rozkładu jasności – cechy obliczone dla obrazu, w którym piksele są pozamieniane miejscami, np. uporządkowane od najciemniejszego do najjaśniejszego, będą miały wartości identyczne z obliczonymi dla obrazu oryginalnego.

Aby ocenić zależności przestrzenne występujące w obrazie można poddać go transformacji i obliczyć cechy statystyczne współczynników uzyskanej transformaty. Jedną z możliwości jest obliczenie wartości bezwzględnej gradientu jasności w każdym punkcie obrazu [112]. Powstaje w wyniku tego nowy obraz, w którym są wyeksponowane kontury rozumiane jako znaczące, lokalne zmiany jasności pomiędzy jednorodnymi obszarami w obrazie oryginalnym. Gradient obliczany jest jako pierwiastek drugiego stopnia sumy kwadratów pochodnych jasności w prostokątnych do siebie kierunkach, np. poziomym i pionowym. Pochodne w obrazie cyfrowym są zwykle obliczane jako połowa różnicy jasności sąsiednich pikseli po lewej i po prawej (dla kierunku poziomego) oraz od góry i od dołu (dla kierunku pionowego). Wielkości statystyczne obliczone dla histogramu wartości bezwzględnej gradientu, w odróżnieniu od cech histogramu obrazu oryginalnego, niosą już informację o pewnych zależnościach przestrzennych rozkładu jasności.

Kolejna metoda wykorzystana w analizie tekstury obrazów endoskopowych wykorzystuje macierz zdarzeń (ang. *co-occurrence matrix* – COM, *gray-level co-occurrence matrix* – GLCM) – histogram drugiego rzędu rozkładu jasności obrazu [235], [254]. Elementy tej macierzy są liczbami par pikseli obrazu o określonych poziomach jasności. Piksele w parach znajdują się w określonej wzajemnej relacji przestrzennej – jeden z nich znajduje się we współrzędnych przesuniętych o określony wektor względem współrzędnych drugiego. COM jest macierzą kwadratową o wymiarach odpowiadających liczbie poziomów jasności obrazu i określa liczbę par pikseli obrazu, w których jeden z pikseli ma jasność daną numerem wiersza, a drugi numerem kolumny. Cechami są wielkości statystyczne macierzy, takie jak energia, kontrast, korelacja, średnia sumacyjna, wariancja sumacyjna i różnicowa oraz entropia i entropia różnicowa. W programie MaZda cechy mogą być obliczone dla dwudziestu różnych macierzy zdarzeń, uzyskanych dla różnych wektorów przesunięcia między pikselami w parach: $[0\ k]^T$, $[k\ 0]^T$, $[k\ k]^T$ i $[k\ -k]^T$, gdzie k przyjmuje wartości naturalne w zakresie od 1 do 5.

W programie MaZda zaimplementowano również algorytm obliczania cech na podstawie macierzy długości ciągów pikseli (ang. *run-length matrix* – RLM, *grey-level run-length matrix* – GRLM) [234]. Macierz ta niesie informację o liczbie ciągów pikseli o identycznej jasności i określonej długości, gdzie długość dana jest numerem kolumny, a jasność numerem wiersza. RLM jest obliczana dla czterech różnych kierunków ciągów pikseli: poziomego, pionowego, 45° i 135° . Na jej podstawie są obliczane: moment odwrotny emfazy krótkich ciągów, mo-

ment emfazy długich ciągów, niejednorodność poziomów jasności, niejednorodność długości ciągów oraz udział powierzchni obrazu w ciągach.

Kolejną grupę cech stanowią parametry modelu autoregresji [231]. Model ten zakłada, że obraz jest transmitowany liniami od górnej do dolnej, a każda z linii jest przesyłana piksel po pikselu od strony lewej do prawej oraz że jasność piksela można przewidywać na podstawie jasności pikseli przesłanych wcześniej. W zaimplementowanym algorytmie są poszukiwane cztery parametry wiążące jasność piksela z jego sąsiadami z lewej, z lewej u góry, u góry i z prawej u góry. Parametry te są obliczane tak, aby dla uzyskanego równania liniowego, błąd średniokwadratowy pomiędzy jasnością przewidywaną i rzeczywistą był minimalny – wartość tego błędu jest piątą cechą.

Program MaZda umożliwia również rozkład obrazu za pomocą transformacji falkowej Haara [255], [256] w pięciu różnych skalach. Na podstawie wynikowej transformaty jest obliczana energia obrazu w pasmach częstotliwości. Obliczone wartości są następnie wykorzystywane jako cechy charakteryzujące teksturę.

Przed obliczeniem cech tekstury, histogram jasności analizowanego obrazu można opcjonalnie poddać normalizacji na dwa sposoby [112]. W pierwszym przypadku zakres jasności obrazu jest zmieniany liniowo tak, aby pełny zakres jasności, nowego obrazu obejmował zakres od średniej jasności obrazu oryginalnego, pomniejszonej o potrójną wartość odchylenia standardowego, do średniej jasności, powiększonej o potrójną wartość odchylenia standardowego. W drugim przypadku normalizacja polega na takiej zmianie zakresu jasności, aby pełny zakres jasności nowego obrazu obejmował zakres od pierwszego do 99 percentyla obliczonych dla histogramu jasności oryginalnego obrazu. W obydwu przypadkach wartości jasności wybiegające poza zakres zostają ograniczone wartościami minimum i maksimum jasności nowego obrazu. Ponadto w nowym obrazie można zmodyfikować parametr określający liczbę bitów kodujących poziom jasności obrazu, co ma istotny wpływ na wartości cech COM i RLM, ponieważ decyduje o rozmiarach tych macierzy. Dodatkowo w przypadku metody RLM zmniejszenie liczby bitów zwiększa długości ciągów pikseli o tym samym poziomie jasności. Zastosowanie normalizacji i standaryzacji ma w założeniu uniezależnić wartości cech od intensywności oświetlenia, co ma znaczenie w analizie obrazów endoskopowych, w których odległe ściany przewodu pokarmowego są słabiej oświetlone od tych znajdujących się blisko. Szczegóły algorytmów obliczania cech tekstury obrazu wykorzystanych w programie MaZda zamieszczono w załączniku B.

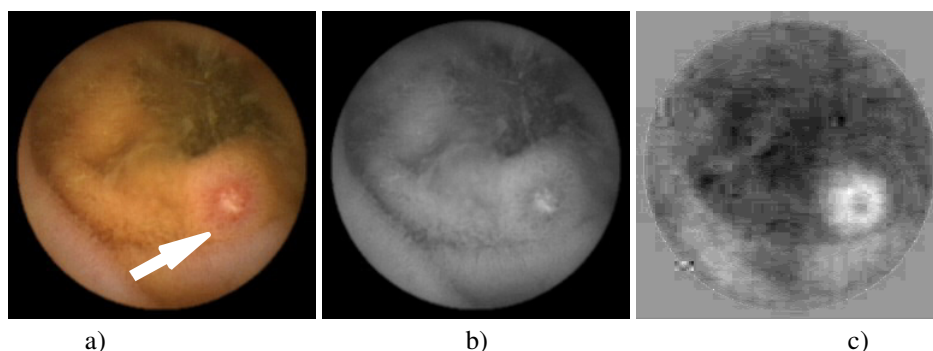
MPEG-7 jest standardem opisu zawartości multimediów (ang. *Multimedia Content Description Interface*), który w założeniu ma charakteryzować treść obrazów, filmów i dźwięków [86], [87], [251], [252]. Do opisu obrazów są wykorzystywane między innymi cechy tekstury, należące do trzech różnych grup: deskryptor tekstury homogenicznej (ang. *homogeneous texture descriptor* – HTD), deskryptor przeszukiwania tekstury (ang. *texture browsing descriptor* – TBD) oraz deskryptor histogramów linii brzegowych (ang. *edge histogram descriptor* – EHD). W dokumentacji standardu przyjęto nazewnictwo, w którym deskryptorami są określane grupy cech, a nie pojedyncze cechy. W celu obliczenia deskryptora tekstury homogenicznej obraz jest filtrowany za pomocą zestawu

filtrów Gabora o sześciu orientacjach i pięciu skalach, a następnie dla wynikowych pasm częstotliwości są obliczane momenty energii pierwszego i drugiego rzędu. Metoda obliczania deskryptora przeszukiwania tekstury wykorzystuje filtry Gabora do wyznaczenia dwóch (tzw. dominujących) kierunków filtracji, w których energia obrazu jest największa. Metoda cech histogramów linii brzegowych analizuje bloki pikseli w poszukiwaniu konturów pomiędzy obszarami obrazu o różnych jasnościach i przypisuje te kontury do jednej z pięciu grup: skierowanych poziomo, pionowo, pod kątem 45° i 135° oraz nieokreślonych (niekierunkowych). Udział ilościowy bloków określonego rodzaju jest cechą charakteryzującą teksturę obrazu. Implementacje metod do obliczania cech tekstury ujętych w standardzie MPEG-7 są łatwo dostępne i mogą być wykorzystywane również w badaniach naukowych – w tym do analizy obrazów z endoskopu bezprzewodowego, czego dowodem są liczne publikacje [55], [58], [64], [75], [77], [85], [88], [257].

5.2. Cechy barwy

Kolor w wizyjnym obrazowaniu przewodu pokarmowego jest informacją istotną z punktu widzenia diagnozowania stanów chorobowych oraz lokalizacji endoskopu bezprzewodowego. Przykładami zmian chorobowych o specyficznej barwie są krwawienia (odcienie czerwieni), wrzody z włóknione (jasnoszare i białe) czy też nadmierne wydzielanie żółci (zabarwienie żółtobrazowe). Ponadto barwa obrazu zależna jest od tego, w którym odcinku przewodu pokarmowego znajduje się endoskop [55], [73], [77], [80] – oznacza to, że informacja o kolorze umożliwia w pewnym stopniu lokalizację kapsułki endoskopowej w poszczególnych odcinkach przewodu pokarmowego. Posiłkowanie się wyłącznie jasnością obrazu endoskopowego może więc być niewystarczające – niezbędna wydaje się analiza barwy.

Jedną z możliwości charakteryzowania obrazu ze względu na barwę jest analiza jego składowych RGB i może polegać na obliczeniu cech niezależnie dla poszczególnych składowych. Przykładowo, składowa zielona niesie informację związaną z krwawieniem, w odróżnieniu od żółto-różowych ścianek przewodu pokarmowego, w przypadku krwi udział zieleni jest niewielki. Zauważyć można jednak, że każda ze składowych RGB zależna jest proporcjonalnie od intensywności oświetlenia, a tym samym od odległości między obrazowanym obiektem a źródłem oświetlenia zamontowanym w endoskopie. Wartości cech obliczonych na podstawie składowych RGB są więc wzajemnie zależne, a to może być nieadekwatne do celu analizy. Słuszne wydaje się więc (rys. 5.1) przeprowadzenie analizy również w innych przestrzeniach kolorów, w których rozdzielane są składowe chrominancji – niosące informację o barwie (niezależne od jasności) oraz luminancję (określającą poziom jasności). Przestrzenie o rozdzielonych składowych luminancji i chrominancji są uzyskiwane w wyniku transformacji przestrzeni RGB.



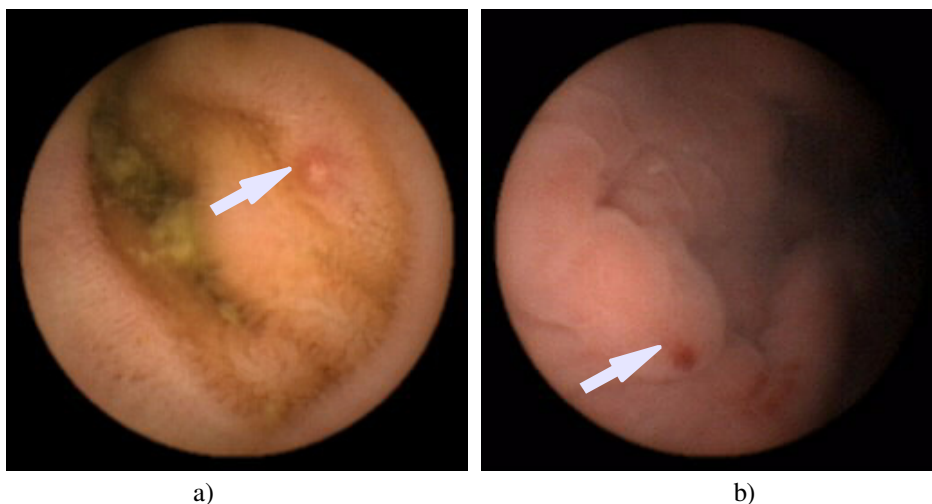
Rys. 5.1. Obraz niewielkiej zmiany wrzodowej w jelicie cienkim: a) oryginalny obraz kolorowy, b) obraz monochromatyczny jasności oraz c) obraz składowej Q w przestrzeni YIQ eksponującej zaróżowiony obszar otoczenia wrzodu

W programie MaZda zaproponowano metodę łączącą analizę barwy z analizą tekstury. Polega ona na zamianie oryginalnego obrazu kolorowego na obraz monochromatyczny reprezentujący jasność lub jedną ze składowych wybranej przestrzeni koloru. Dla uzyskanego w ten sposób obrazu są obliczane cechy tekstury, w tym cechy statystyczne histogramu. W programie zaimplementowano procedury transformacji przestrzeni RGB [258–261] na przestrzenie YUV, YIQ, HSV, CIE $L^*a^*b^*$ oraz CIE XYZ. Ponadto, są też wykorzystywane przestrzenie, w których składowe chrominancje UV oraz IQ są normalizowane względem poziomu jasności Y – uniezależnia to zakres zmienności tych składowych od intensywności oświetlenia. Przykładowo można obliczyć cechy macierzy RLM dla obrazu monochromatycznego będącego reprezentacją składowej Q w przestrzeni YIQ. Równania opisujące zależności między zastosowanymi przestrzeniami barw a przestrzenią RGB zamieszczono w załączniku B.

Standard MPEG-7 definiuje pięć deskryptorów koloru [86], [87], [251–253] oraz dwa zestawy parametrów, dla których deskryptory te są obliczane. Pierwszym deskryptorem jest kolor dominujący (ang. *Dominant Color Descriptor* – DCD), który wskazuje położenie maksimum funkcji histogramu obrazu określonego w trójwymiarowej przestrzeni koloru oraz podaje jej wartość. Kolejny deskryptor to skalowalny deskryptor koloru (ang. *Scalable Color Descriptor* – SCD), który jest zbiorem współczynników transformaty Haara histogramu koloru w przestrzeni HSV. Jego skalowalność polega na możliwości modyfikowania liczby współczynników przez zmianę sposobu dyskretyzacji przestrzeni kolorów lub zmianę liczby bitów wykorzystywanych do kodowania ich wartości. Deskryptor rozkładu przestrzennego koloru (ang. *Color Layout Descriptor* – CLD) jest z kolei zbiorem współczynników dyskretnej transformaty kosinusowej obliczonej dla obrazu pomniejszonego do rozmiaru 8×8 pikseli, uszeregowanych zgodnie z zasadami porządkowania zygzakowego. Ostatni z deskryptorów – deskryptor struktury kolorów (*Color-Structure Descriptor* – CSD) opisuje histogram strukturalny kolorów w przestrzeni HMMD (ang. *Hue Min Max Difference*). Deskryptor koloru dominującego zależy od parametrów określających przestrzeń koloru (ang. *Color Space*) oraz sposób kwantyzacji tej przestrzeni (ang. *Color Quantization*).

5.3. Obszary zainteresowania i mapy cech

Obszar zainteresowania (ang. *Region of Interest* – ROI) jest zbiorem pikseli wyznaczających fragment obrazu, który ma być poddany analizie. Wstępne oznaczenie takiego fragmentu umożliwia skoncentrowanie działania algorytmów obliczeniowych w miejscu istotnym dla osiągnięcia zakładanych celów analizy, a to z kolei pozwala uniknąć zbędnych obliczeń dla pozostałej, nieistotnej części obrazu. Oznaczanie obszarów zainteresowania jest szczególnie ważne w obrazach biomedycznych, gdzie obiektem zainteresowania jest wybrany organ, tkanka lub struktura, a pojedynczy obraz przedstawia również inne, nieadekwatne elementy. Aby wyodrębnić cechy związane wyłącznie z określonym obiektem zainteresowania, należy ograniczyć działanie algorytmu analizy obrazu do obszaru zainteresowania odpowiadającego wyłącznie temu obiektowi.

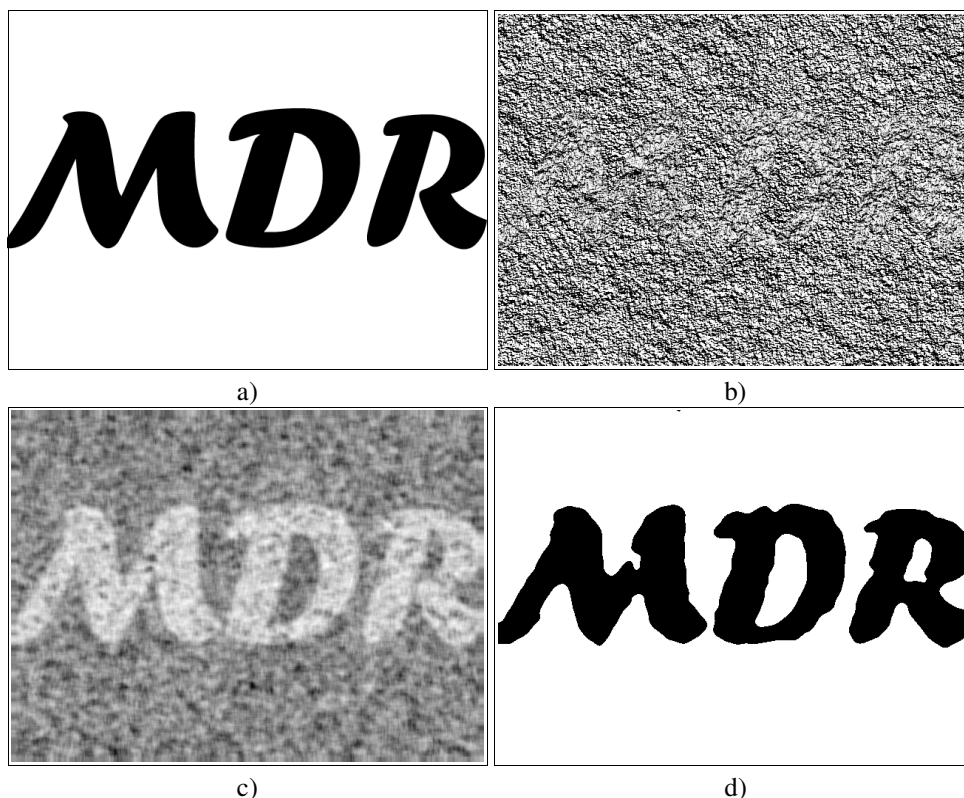


Rys. 5.2. Obrazy drobnych zmian chorobowych: a) wrzód i b) angiektazja

Niektóre zmiany chorobowe przewodu pokarmowego mają charakter miejscowy, są drobne i zajmują niewielki fragment obrazu endoskopowego. Przykładem takich zmian (rys 5.2) są drobne owrzodzenia, wybroczyny i angiektazje (widoczne prześwitujące naczynia krwionośne, podskórne zaczerwienienia), które w przedstawiającym je obrazie mogą obejmować kilkanaście do kilkudziesięciu pikseli – stanowi to mniej niż tysięczną część całej jego powierzchni. Obliczenie cechy tekstury lub koloru dla całego obrazu spowodowałoby, że charakteryzowałaby ona w znacznie większym stopniu obszar otaczający schorzenie niż samą zmianę chorobową. Z tego powodu niecelowe jest analizowanie obrazu w całości – bardziej odpowiednie wydaje się podzielenie go na mniejsze fragmenty w taki sposób, aby udział powierzchniowy uwidocznionej zmiany chorobowej we fragmencie stał się znaczący, a następnie przeprowadzenie analizy indywidualnie w każdym z nich.

Obliczenie cechy obrazu w pewnym niewielkim obszarze zainteresowania, sąsiedztwie pojedynczego piksela powoduje, że jej wartość ma charakter lokalny,

powiązany z pewnym konkretnym miejscem. Jeśli wartość takiej cechy zostanie obliczona w sąsiedztwie każdego piksela analizowanego obrazu, to wynikiem będzie funkcja wartości tej cechy określona w dziedzinie współrzędnych obrazu – mapa cechy. Jeżeli wartość cechy zostanie przedstawiona w postaci poziomu jasności, to odwzorowanie funkcji staje się obrazem monochromatycznym. Obraz ten można z kolei poddać dalszej analizie – np. zastosować metodę segmentacji przez progowanie jasności, co w efekcie umożliwi wydzielenie obszarów tekstur jednorodnych (rys. 5.3). Wyzwaniem jest znalezienie odpowiedniej cechy, której mapa skontrastuje różne tekstury jednorodne i umożliwi ich rozróżnienie.



Rys. 5.3. Segmentacja obrazu zawierającego tekstury homogeniczne: a) maska określająca obszary teksturowania, b) sztucznie wygenerowany obraz zawierający dwie jednorodne tekstury, c) mapa entropii macierzy zdarzeń eksponująca różnice między obszarami, d) obraz po segmentacji – wynik progowania mapy poprzedzony filtracją medianową

Możliwość obliczania wartości cech teksturowych i barwowych w obszarach zainteresowania jest zatem przydatną techniką analizowania obrazów biomedycznych. Po pierwsze, umożliwia obliczanie cech w obszarach obrazu powiązanych z konkretnym przedmiotem analizy. Porównanie wartości cech obliczanych w takich obszarach oraz poza nimi warunkuje odnalezienie wśród nich takich, które są charakterystyczne dla wybranego obiektu i umożliwiają jego rozpoznawanie. Po drugie, zawężenie obszaru analizy do małych fragmentów obrazu pozwala obliczać cechy w taki sposób, aby charakteryzowały one drobne zmiany chorobowe, których udział powierzchniowy w całym obrazie jest niewielki. Po trzecie, obli-

czenia wykonywane w małym obszarze zainteresowania, przemieszczającym się we współrzędnych obrazu, pozwalają wyznaczyć mapę cech, a za jej pomocą segmentację obrazu na obszary jednorodnej tekstury.

Porównując opisane wcześniej narzędzia generowania cech tekstury i koloru, należy więc wziąć pod uwagę ich możliwości związane z obliczaniem cech w obszarach zainteresowania. Można zauważyć, że jest to możliwe zarówno w przypadku programu MaZda, jak i w przypadku implementacji standardu MPEG-7. W przypadku metod opisanych standardem MPEG-7 stosowanie obszarów zainteresowania napotyka jednak wiele ograniczeń. Nie przewidziano stosowania obszarów innych niż jednospójne prostokątne, a ponadto do obliczenia HTD wymagany jest obszar nie mniejszy niż 128×128 pikseli. Obszar taki stanowi 33% powierzchni pokrywającej pole widzenia kamery endoskopu bezprzewodowego. Obliczenie cech w tak dużym obszarze umożliwiającej identyfikację drobnych zmian chorobowych wydaje się mało prawdopodobne. W programie MaZda natomiast obszar zainteresowania może być zdefiniowany w sposób arbitralny, może on mieć dowolnie nieregularny kształt (jednospójny, wielospójny lub niespójny) i wielkość. Program udostępnia również procedurę umożliwiającą automatyczne obliczenie map wybranych cech tekstury i barwy.

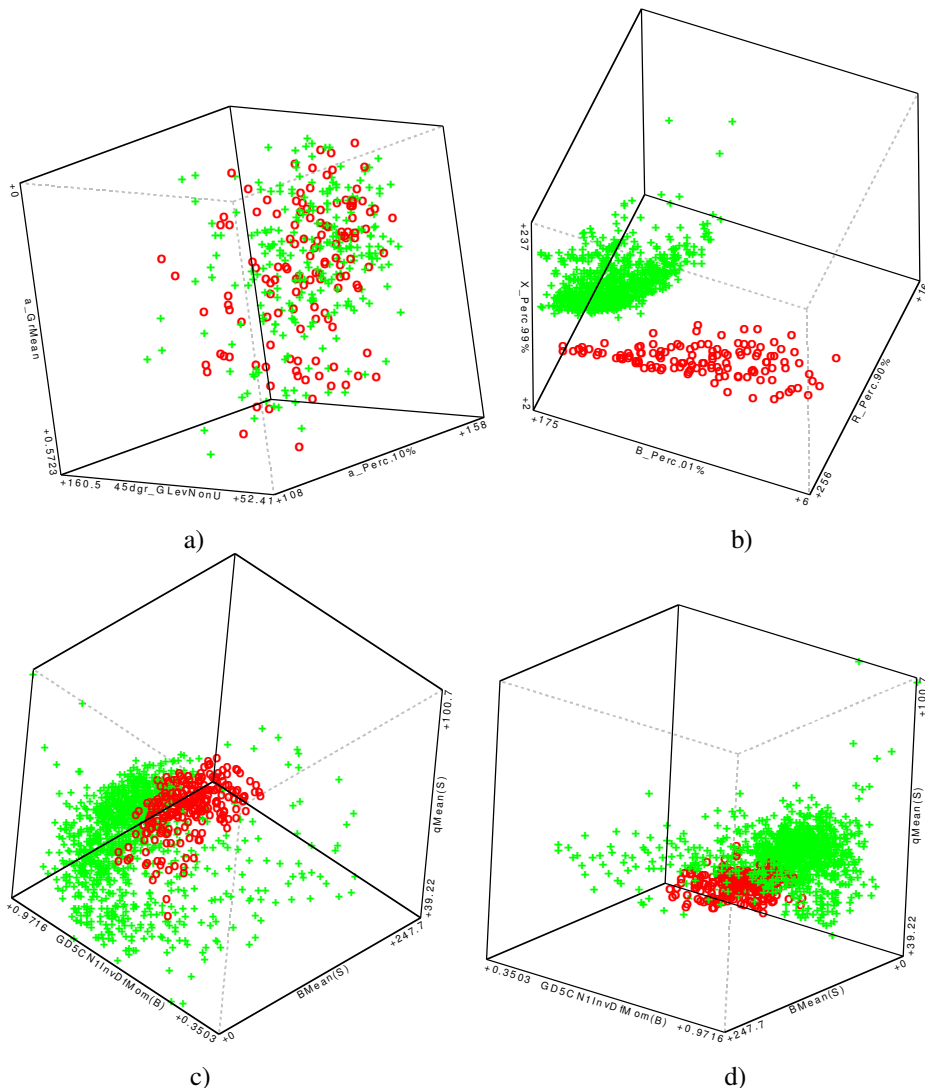
5.4. Selekcja cech i redukcja wymiarowości ich przestrzeni

Metody obliczania cech tekstury i barwy obrazu są na tyle złożone, że interpretacja wartości cech polegająca na odniesieniu ich wartości do tego jak człowiek postrzega obraz jest bardzo trudna. Duża liczba i różnorodność cech oraz zależności występujące między ich wartościami powodują, że określenie a-priori przydatności danej cechy do rozróżniania różnych rodzajów (klas) obrazów jest praktycznie niemożliwe. Przyjęto zatem metodę polegającą na obliczeniu wielu różnych cech, automatycznym wyborze spośród nich cech znaczących (umożliwiających rozróżnienie obrazów ze względu na prezentowaną przez nie treść), a następnie określeniu reguł klasyfikowania danych obrazowych na podstawie wartości wybranych cech.

Detekcja zmian chorobowych wymaga określenia związku pomiędzy wartościami obliczanych cech a tym, czy obraz przedstawia rozważaną zmianę czy też nie. Obrazy z endoskopu bezprzewodowego wykorzystywane w badaniach zostały wstępnie ocenione przez diagnostów pod kątem rodzaju występujących w nich zmian chorobowych – oznaczyli oni również obszary obejmujące te zmiany. Dzięki temu możliwe stało się powiązanie zbiorów cech wygenerowanych we fragmentach obrazów z informacją o tym co te fragmenty przedstawiają. Tak przygotowane dane (zbiór uczący) umożliwiają zastosowanie metod nadzorowanego uczenia maszynowego [228], [262], [263], wynikiem którego są reguły klasyfikacji danych – zasad przypisywania znaczeń „stan normalny” lub „stan chorobowy”.

Liczba cech obliczanych w obszarze zainteresowania jest znaczna i w przypadku wykorzystania algorytmów zaimplementowanych w programie MaZda może sięgać kilku tysięcy. Uporządkowany zbiór cech tworzy wektor, który w sposób ilościowy i wielowymiarowy opisuje taki obszar. Rozróżnienie (dyskry-

minacja) wymaga przeprowadzenia analizy rozkładów takich wektorów w ich wielowymiarowej przestrzeni. Biorąc pod uwagę, że przestrzeń ma kilka tysięcy wymiarów, jej analiza wymaga dużych nakładów obliczeniowych. Ponadto, z punktu widzenia możliwości ludzkiej percepcji, analiza taka jest złożonym problemem – uzyskane wyniki trudno jest świadomie ocenić, poddać krytyce lub wnioskować na ich podstawie. Celowa jest zatem redukcja wymiarowości przestrzeni cech, czyli zmniejszenie liczby cech wykorzystywanych do dyskryminacji.



Rys. 5.4. Rozkłady wektorów wybranych cech dwóch klas (wektory należące do klas oznaczono różnymi symbolami): a) wektory obu klas przemieszane – cechy nieprzydatne w dyskryminacji, b) wektory tworzą osobne skupiska, które można rozdzielić płaszczyzną – możliwość zastosowania klasyfikacji liniowej oraz c-d) wektory jednej klasy tworzą skupisko otoczone przez wektory drugiej klasy – wymagana jest klasyfikacja nieliniowa

Istnieją dwie metody umożliwiające redukcję wymiarowości przestrzeni cech: pierwsza polega na selekcji cech, a druga na rzutowaniu ich przestrzeni [264]. Okazuje się zazwyczaj, że ograniczony podzbiór cech (tzw. cech znaczących) niesie informacje istotne z punktu widzenia dyskryminacji. Pozostałe cechy są nieistotne, można je odrzucić i nie brać pod uwagę. Zabieg polegający na wyborze cech znaczących i odrzuceniu pozostałych nazywany jest **selekcją**. Drugą metodą jest **ekstrakcja** cech polegająca na rzutowaniu przestrzeni cech na nową przestrzeń o mniejszej liczbie wymiarów. Wymaga to obliczenia pewnej liczby nowych cech, których wartości są funkcjami wartości cech oryginalnych. Liczba nowych cech powinna być znacząco mniejsza od liczby cech oryginalnych, przy czym nowy zbiór cech powinien się charakteryzować niezmniejszoną lub jedynie nieznacznie zmniejszoną zdolnością dyskryminacyjną.

Zazwyczaj cechy wybierane są przy założeniu, że zostaną one wykorzystane w konkretnym rodzaju klasyfikatora – algorytmie klasyfikacji. Cechy wybierane są tak, aby maksymalizować zdolność dyskryminacyjną, przy czym zdolność ta musi być wielkością liczbową, zdefiniowaną dla wybranego typu klasyfikatora, mierzącą poprawność jego działania. Jedną z metod redukcji wymiarowości przestrzeni cech jest liniowa analiza dyskryminacyjna omówiona w rozdziale 4.2, w której jako kryterium zdolności dyskryminacyjnej zastosowano liniowy dyskryminator Fishera. Dyskryminator ten można również wykorzystać w procesie selekcji cech, wybierając taki podzbiór cech, który zapewni jego dużą wartość. Należy być jednak świadomym, że wybrane w ten sposób cechy będą, przy pewnych założeniach, optymalne dla klasyfikatora liniowego [225], [262], natomiast mogą nie być właściwe do zastosowania w innej metodzie klasyfikacji.

W ogólnym przypadku selekcja cech dla wybranego typu klasyfikatora polega na: 1) wybraniu pewnego podzbioru cech, 2) optymalizacji parametrów (uczeniu) klasyfikatora dla tego podzbioru oraz 3) ilościowej ocenie działania klasyfikatora (obliczeniu zdolności dyskryminacyjnej). Proces ten jest powtarzany wielokrotnie dla różnych podzbiorów cech. Wybierany jest ten podzbiór, dla którego uzyskana zdolność dyskryminacyjna jest największa. Należy zwrócić uwagę, że jeśli celem selekcji jest wybór optymalnego podzbioru k cech ze zbioru n cech ($k < n$), to liczba wszystkich możliwych podprzestrzeni równa jest liczbie k -elementowych kombinacji bez powtórzeń ze zbioru n -elementowego.

$$C_n^k = \binom{n}{k} = \frac{n!}{k!(n-k)!} \quad (5.4.1)$$

Przykładowo, w przypadku liczby cech $n = 3000$ i $k = 3$ oznacza to konieczność powtórzenia procesu uczenia klasyfikatora blisko 27 miliardów razy.

5.5. Metoda powierzchni wypukłej

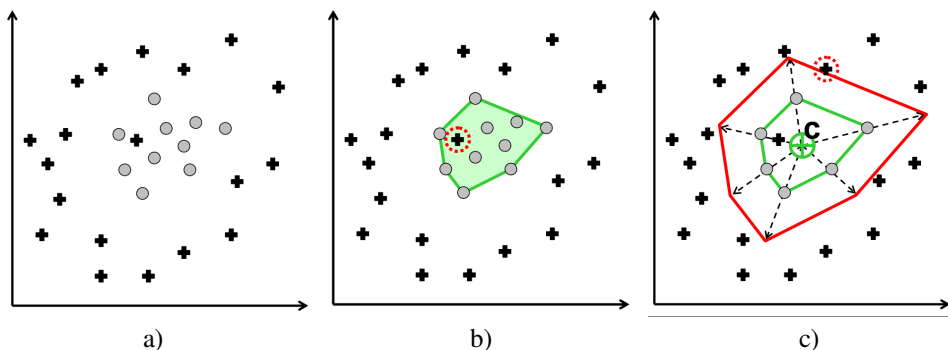
Jeśli przyjrzeć się obrazom wnętrza układu pokarmowego nie przedstawiających widocznych zmian chorobowych, to widać że są one bardzo zróżnicowane. Mogą one przedstawiać powierzchnię wewnętrzną przełyku, żołądka, jelita cienkiego lub grubego, mogą przedstawiać zróżnicowaną zawartość, jak niestrawiony pokarm, pianę, soki trawienne itp. Natomiast w przypadku obrazów wybranej

zmiany chorobowej ich różnorodność nie jest tak duża – poszczególne przypadki są do siebie podobne. Wynika stąd, że wartości cech w przypadku obrazów bez zmian chorobowych mają duży zakres zmienności, a w przypadku obrazów wybranej zmiany chorobowej mniejszy i często zawierający się wewnątrz tego pierwszego. W przestrzeni o ograniczonej liczbie wymiarów rozkład wektorów cech dla wybranej zmiany chorobowej może mieć charakter skupiska o stosunkowo niewielkich rozmiarach, otoczonego przez rozproszone wektory obliczone dla pozostałych obrazów. W takim przypadku dyskryminacja danych za pomocą klasyfikatora liniowego nie jest możliwa, ponieważ nie można rozdzielić obydwu grup obserwacji za pomocą hiperpłaszczyzny (rys. 5.4).

Przestrzenie cech, w których wektory jednej klasy stanowią skupisko otoczone przez wektory drugiej klasy, można analizować za pomocą metod, takich jak maszyna wektorów nośnych (ang. *support vector machine* – SVM) z radialną funkcją bazową [263], [265] lub radialna sieć neuronowa (ang. *radial basis function network* – RBFN) [266–268]. W metodach tych, w odróżnieniu od standardowej SVM oraz perceptronowej sieci neuronowej, liniowe równania, na podstawie których są rozdzielane wektory obu klas zamieniano na równości z funkcjami radialnymi. Funkcja radialna to taka, której wartość zależy wyłącznie od odległości od pewnego wybranego punktu. Ograniczeniem obu metod jest konieczność normalizacji zakresów zmienności cech. Ponadto w procesie uczenia wymagają one równolicznych zbiorów obu klas do uzyskania optymalnego rozwiązania. W przypadku wykorzystywanych danych, liczność zbioru wektorów cech obliczonych dla obrazów z wybranymi zmianami chorobowymi jest znacznie mniejsza niż zbioru uzyskanego dla obrazu bez widocznych zmian. Należy też zauważyć, że obie metody w procesie uczenia – optymalizacji klasyfikatora – wykorzystują algorytmy iteracyjne wymagające znacznych nakładów obliczeniowych. Dodatkowo, zastosowanie funkcji radialnych wiąże się z dodatkową koniecznością ich parametryzacji.

Celowe wydaje się więc zastosowanie metody selekcji cech i klasyfikacji danych umożliwiającej analizowanie rozkładów, w których wektory jednej klasy tworzą skupisko otoczone przez wektory innej klasy, niewymagającej równoważenia liczności zbiorów oraz nie wymagającej długotrwałego, iteracyjnego procesu uczenia klasyfikatora. Rozwiązaniem spełniającym te wymagania jest opracowana przez autora metoda analizy przestrzeni cech za pomocą powierzchni wypukłej opartej na wektorach (ang. *vector supported convex hull* – VSCH) [65], [66], [70]. Dane wejściowe metody to zbiory n -wymiarowych wektorów dwóch klas. Poszukiwana jest k -wymiarowa podprzestrzeń przestrzeni n -wymiarowej, w której wektory klasy pierwszej formują skupisko otoczone przez wektory klasy drugiej. Dla każdej testowanej podprzestrzeni tworzony jest wielotop (dla $k = 2$ – wielokąt, dla $k = 3$ – wielobok) wypukły wektorów klasy pierwszej (rys. 5.5). Następnie zliczane są wektory klasy drugiej zawarte wewnątrz uzyskanej powierzchni. Wybierana jest ta podprzestrzeń, w której liczba Q_1 zliczonych w ten sposób wektorów jest najmniejsza. Jeśli istnieje więcej niż jedna taka podprzestrzeń, wówczas wykorzystuje się dodatkowe kryterium oceny. Określa ono jak bardzo można powiększyć wielotop wypukły wokół jego środka ciężkości, aby nie obejmował dodatkowych wektorów drugiej klasy. Powiększanie wielotopu

polega na jego skalowaniu, przy czym środek ciężkości jest punktem stałym przekształcenia. Skala a powiększenia wielotopu jest podstawą drugiego kryterium decydującego o wyborze podprzestrzeni cech – im jest ona większa, tym większe jest odseparowanie wektorów jednej klasy od drugiej.



Rys. 5.5. Ilustracja graficzna algorytmu obliczania zdolności dyskryminacyjnej za pomocą powierzchni wypukłej opartej na wektorach: a) wektory jednej klasy tworzą skupisko otoczone przez wektory drugiej klasy, b) wielobok wypukły utworzony na wektorach jednej z klas, wielobok obejmuje jeden z wektorów klasy drugiej ($Q_1 = 1$), c) wielobok maksymalnie powiększony wokół jego środka ciężkości C , tak aby nie obejmował innego wektora drugiej klasy

Na podstawie obu kryteriów można zdefiniować pojedynczą wartość liczbową zdolności dyskryminacyjnej VSCH. Można zauważyć, że liczba Q_1 jest nieujemną liczbą całkowitą. Skala a jest wielkością rzeczywistą większą niż jeden. Odwrotność skali $Q_2 = a^{-1}$ jest zatem wielkością ułamkową z zakresu $(0, 1)$. Suma obu wielkości $Q = Q_1 + Q_2$ jest wskaźnikiem niedoskonałości odseparowania klas. Zatem odwrotność wskaźnika Q^{-1} może być wykorzystywana jako zdolność dyskryminacyjna VSCH.

Wielotop k -wymiarowej przestrzeni [269] oparty na skończonej liczbie wektorów jest określony za pomocą układu m liniowych nierówności (5.5.1), w których wektor $\mathbf{B}_j$ jest wersorem j -tej ($j = 1, 2, \dots, m$) hiperpłaszczyzny (ścianki) zwróconym do wnętrza wielotopu, a b_{j0} określa przesunięcie hiperpłaszczyzny względem początku układu współrzędnych w kierunku wyznaczanym przez wektor $\mathbf{B}_j$, ale o przeciwnym zwrocie. Wektory $\mathbf{B}_j$ i parametry b_{j0} są uzyskiwane w wyniku rozwiązania równania (5.5.2) dla k liniowo niezależnych wektorów cech leżących na brzegu wielotopu, przy czym rozwiązanie musi spełniać warunek (5.5.3) oraz nierówność (5.5.1) dla wszystkich wektorów cech klasy pierwszej.

$$\mathbf{B}_j^T \mathbf{x} + b_{j0} \leq 0 \quad (5.5.1)$$

$$\mathbf{B}_j^T \mathbf{x} + b_{j0} = 0 \quad (5.5.2)$$

$$|\mathbf{B}_j^T| = 1 \quad (5.5.3)$$

Metoda VSCH umożliwia konstrukcję prostych reguł klasyfikacji danych. W tym celu uzyskany wielotop wypukły wektorów klasy pierwszej jest powiększany wokół swojego środka ciężkości w skali $\sqrt{a}$. Powiększony w ten sposób wielotop obejmuje wektory klasy pierwszej, tworząc dookoła nich pewien margines i jednocześnie zachowując odległość od niezawartych w nim wektorów klasy drugiej. Powiększony wielotop opisany jest układem nierówności (5.5.4), w których $\mathbf{C}_0$ jest wektorem współrzędnych jego środka ciężkości. Spełnienie przez wektor wszystkich nierówności tego układu oznacza jego przynależność do klasy pierwszej, niespełnienie przynajmniej jednej z nich oznacza jego przynależność do klasy drugiej.

$$\mathbf{B}_j^T \mathbf{x} + (\mathbf{B}^T \mathbf{C}_0)(\sqrt{A}-1) + b_{j0} \sqrt{a} \leq 0 \quad (5.5.4)$$

Do wyznaczania wielotopu wypukłego wykorzystano szybki algorytm, którego czasowa złożoność obliczeniowa wynosi $O(N_{sl} \log N_{sl})$ [270–272], gdzie N_{sl} oznacza licznosc wektorów klasy pierwszej. Dla porównania złożoność obliczeniowa SVM zawiera się w przedziale od $O(N_s^2)$ do $O(N_s^3)$ [265], gdzie N_s jest licznoscą wektorów obu klas. Ponadto, SVM w przeciwieństwie do VSCH jest optymalizowany w sposób iteracyjny, co dodatkowo wydłuża czas obliczeń.

Podsumowując, zaproponowana metoda, należąca do grupy klasyfikatorów nieliniowych, umożliwia analizę specyficznych rozkładów wektorów, w których wektory jednej klasy tworzą skupisko otoczone przez wektory klasy drugiej. Uzyskiwane wyniki nie zależą od wstępnej normalizacji lub standaryzacji zakresów danych ponieważ podstawą obliczania zdolności dyskryminacyjnej jest skalowanie wielotopu, którego długość wzdłuż wybranej osi przestrzeni jest proporcjonalna do zakresu zmienności cechy w klasie pierwszej. Metoda charakteryzuje się stosunkowo niewielką złożonością obliczeniową i, co ważne, złożoność ta zależy od liczby wektorów klasy pierwszej, a nie od liczby wektorów obydwu klas – VSCH jest więc tym bardziej efektywna im większy jest stosunek licznosci wektorów klasy drugiej do ich licznosci w klasie pierwszej.

Idea wykorzystania powierzchni wypukłych w klasyfikacji nie jest nowa i została wykorzystana w klasyfikatorze najbliższej powierzchni wypukłej (ang. *nearest convex hull classifier* – NCHC) [273], [274] oraz w podobnej metodzie najbliższego sąsiada powierzchni wypukłej (ang. *nearest neighbour convex hull* – NNCH) [275]. Jednakże NNCH buduje wielotopy wypukłe dla wektorów wszystkich rozważanych klas i nie zapewnia poprawnej klasyfikacji w przypadku gdy jeden z wielotopów zawiera się wewnątrz drugiego. W takim przypadku NNCH wymagałby zastosowania radialnych funkcji bazowych, podczas gdy VSCH wykorzystuje naturalną właściwość pojedynczego wielotopu do rozgraniczania wektorów w rozważanym tutaj przypadku. Poza tym w obu metodach nie proponowano miary zdolności dyskryminacyjnej klasyfikatora, która umożliwiłaby jego zastosowanie do selekcji cech.

6. Walidacja metod analizy tekstury i barwy

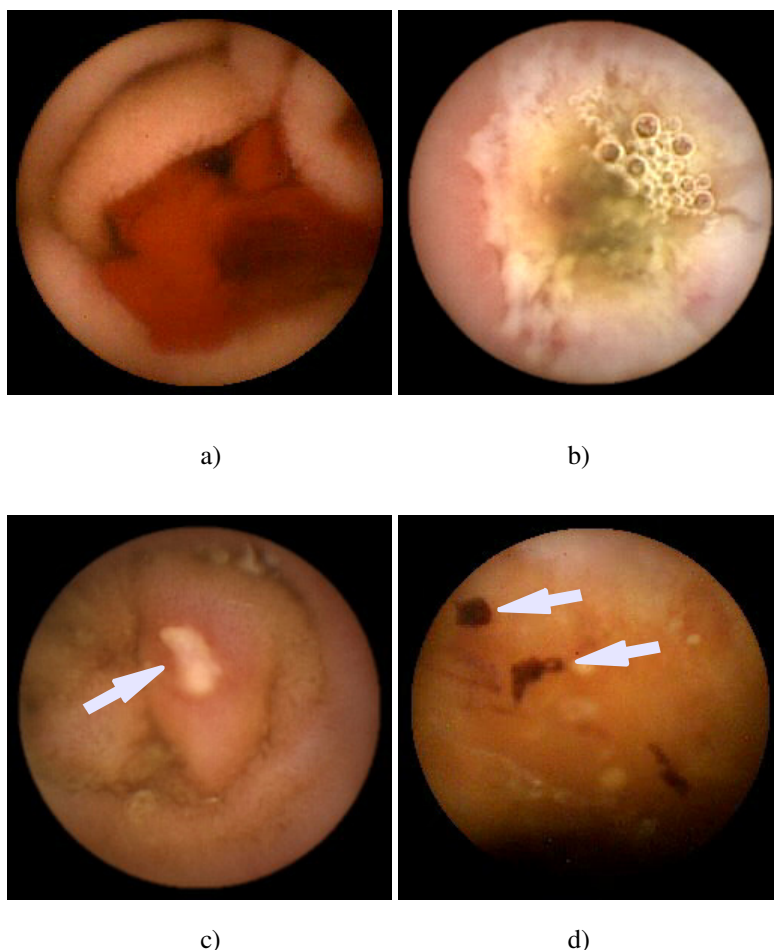
W poprzednim rozdziale opisano metody analizy obrazów wykorzystujące obliczane numerycznie cechy tekstury i barwy, selekcji cech znaczących oraz uczenia nadzorowanego. Wydaje się, że właśnie one mogą być przydatne do detekcji zmian chorobowych przewodu pokarmowego w obrazach endoskopowych. W celu walidacji tych metod oraz dla potwierdzenia hipotezy o ich przydatności do analizy obrazów z endoskopu bezprzewodowego przeprowadzono trzy eksperymenty. Eksperymenty te opisano w kolejnych podrozdziałach. W pierwszym z nich wykorzystano program MaZda do wygenerowania cech tekstury i barwy opisujących wybrane obszary obrazów endoskopowych. Dokonano selekcji tych cech oraz uczenia nadzorowanego, w wyniku czego sformułowano reguły klasyfikacji obrazów na przedstawiające stan zdrowy i patologiczny przewodu pokarmowego. Następnie za pomocą tych reguł sklasyfikowano obrazy, a wyniki poddano ocenie jakościowej i ilościowej. W drugim eksperymencie wykorzystano cechy barwy i tekstury zdefiniowane w standardzie MPEG-7. Podobnie jak w pierwszym eksperymencie określono reguły klasyfikacji obrazów oraz dokonano oceny uzyskanych wyników. Przeprowadzenie obu eksperymentów umożliwiło również porównanie zdolności dyskryminacyjnej cech generowanych przez program MaZda oraz opisanych standardem MPEG-7. Celem trzeciego eksperymentu była weryfikacja hipotezy, że mapy cech i reguły wykorzystywane do detekcji zmian chorobowych mogą również być wykorzystane do segmentacji obrazów – pozwolą wydzielić obszary zmian chorobowych oraz ocenić ich wielkość obiektywnie i w sposób ilościowy.

6.1. Detekcja zmian chorobowych – program MaZda

Celem pierwszego eksperymentu była ocena zdolności dyskryminacyjnej generowanych przez program MaZda cech tekstury oraz barwy w zastosowaniu do obrazów endoskopii bezprzewodowej. W ramach eksperymentu oceniono również działanie metody VSCH służącej do selekcji cech. Porównano wyniki klasyfikacji wykorzystującej cechy wyselekcjonowane tą metodą oraz cechy wyselekcjonowane za pomocą innych uznanych algorytmów, takich jak RBFN czy SVM.

W wyniku konsultacji z lekarzami, ekspertami w zakresie diagnostyki medycznej układu pokarmowego, skupiono się na czterech rodzajach anomalii występujących w przewodzie pokarmowym: krwawieniach, wybroczynach, drobnych wrzodach o charakterze skupionym oraz rozległych owrzodzeniach zwłókniałych. Anomalie te są efektem chorób wrzodowych, między innymi

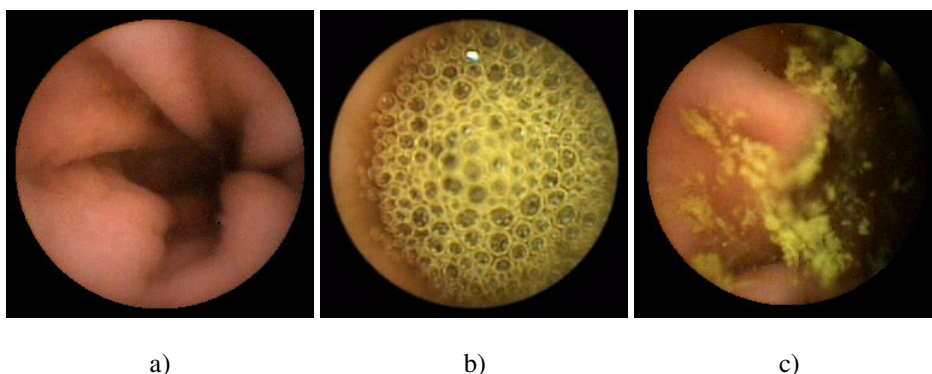
choroby Leśniewskiego-Crohna [8], [26], [94] – przewlekłego, nieswoistego zapalenia jelita cienkiego. Etiologia choroby nie jest znana, a efekty leczenia są zależne od jej wczesnego zdiagnozowania. Dlatego opracowanie techniki wykrywania wymienionych zmian chorobowych wydaje się ważne i potrzebne.



Rys. 6.1. Przykłady obrazów z widocznymi zmianami chorobowymi: a) krwawienie (obszar o czerwonym zabarwieniu), b) owrzodzenie rozległe i zwłókniałe (biały obszar o postrzępionym brzegu), c) drobny wrzód oraz d) wybroczyny

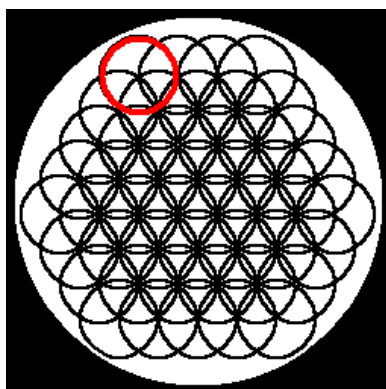
Spśród ponad 50 filmów, którymi dysponowano w eksperymencie, zidentyfikowano trzy, w których zaobserwowano owrzodzenia o rozległym i zwłókniałym charakterze – wyodrębniono z nich 52 obrazy tego typu zmian, z czterech innych filmów wyselekcjonowano 32 obrazy wrzodów obejmujących niewielki obszar, w trzech innych filmach zidentyfikowano 23 obrazy przedstawiające miejsca krwawień, a w pięciu filmach 50 obrazów drobnych, punktowych wybroczyn i angiektazji. Dodatkowo, dla celów porównawczych, wylosowano ponad 500 obrazów bez widocznych zmian chorobowych – przedstawiających prawi-

dłowy stan ścianek przewodu pokarmowego, pianę lub treść pokarmową. Gastroenterolodzy biorący udział w badaniu potwierdzili występowanie określonej zmiany patologicznej lub brak takich zmian w wyselekcjonowanych obrazach. Ponadto w każdym z obrazów przedstawiających zmianę chorobową oznaczyli objęty nią fragment – dokonując tym samym ręcznej segmentacji obrazu. Przykłady obrazów wykorzystanych w eksperymencie przedstawiono na rysunkach 6.1 i 6.2.



Rys. 6.2. Przykłady obrazów klasy referencyjnej: a) obraz prawidłowej powierzchni przewodu pokarmowego, b) piana i c) treść pokarmowa przypominająca owrzodzenia

Dla każdego obrazu utworzono po 48 kolistych, równomiernie rozmieszczonych w polu widzenia, nakładających się obszarów zainteresowania (rys. 6.3) obejmujących po 2009 pikseli każdy. Za pomocą programu MaZda, w poszczególnych obszarach zainteresowania każdego z wyselekcjonowanych obrazów, obliczono cechy tekstury i barwy. Cechy obliczono dla monochromatycznych reprezentacji obrazów, takich jak luminancja, składowe przestrzeni kolorów RGB, CIE $L^*a^*b^*$ i XYZ, składowe chrominancji przestrzeni YUV, YIQ, HSI oraz składowe UV i IQ znormalizowane względem wartości luminancji. Dla każdego z otrzymanych obrazów monochromatycznych obliczono cechy histogramu, parametry statystyczne gradientu jasności, cechy macierzy zdarzeń, macierzy ciągów pikseli oraz parametry modelu autoregresji (załącznik B). W przypadku luminancji histogram obrazu poddano wstępnej normalizacji oraz dodatkowo obliczono wartości energii w pasmach częstotliwościowych transformaty falkowej Haara. W wyniku obliczeń, dla każdego z obszarów zainteresowania, uzyskano wektor cech składający się z 3049. elementów opisujących ten obszar pod względem barwy i tekstury.



Rys. 6.3. Kształty i rozkład obszarów zainteresowania, w których obliczono cechy koloru i tekstury za pomocą programu MaZda. Kolorem czerwonym oznaczono wybrany, pojedynczy obszar zainteresowania

W przypadku obrazów przedstawiających rozległe owrzodzenia oraz krwawienia, do dalszej analizy wybrano wektory cech, które obliczono w obszarach zainteresowania całkowicie zawartych w obszarach obejmujących zmianę chorobową. W przypadku drobnych wrzodów oraz wybroczyn i angiektazji do dalszych analiz wybrano wektory cech obliczonych w obszarach zawierających w sobie zmianę chorobową. W efekcie uzyskano 283 wektory cech obszarów aktywnych krwawień (oznaczone jako klasa B – *ang. Bleeding*), 533 wektory cech obszarów angiektazji i wybroczyn (klasa P – *ang. Petechia*), 312 wektorów dla wrzodów drobnych (klasa F – *ang. Focal ulcer*), 392 dla rozległych zwłókniałych owrzodzeń przewodu pokarmowego (klasa E – *ang. Extensive ulcer*) oraz ponad 20 000 wektorów dla obszarów bez widocznych zmian zwyrodnieniowych (klasa R – *ang. Reference*). W przypadku wektorów cech klas B, F, P i E wylosowano po 128 wektorów wykorzystywanych następnie w procesie uczenia oraz po 100 wektorów na potrzeby testowania uzyskanych reguł klasyfikacji. W przypadku klasy R wylosowano odpowiednio 1024 wektory zbioru uczącego i 100 wektorów zbioru testującego.

W następnym etapie eksperymentu dokonano selekcji cech za pomocą opracowanej metody VSCH. Obliczono zdolność dyskryminacyjną tej metody dla wszystkich dwu- i trójwymiarowych podprzestrzeni cech (wszystkich par i trójek cech), przy czym wielotop wypukły był budowany na wektorach jednej z klas B, F, P albo E, podczas gdy wektory klasy referencyjnej R miały go otaczać. Selekcji cech dokonano więc osobno dla wektorów poszczególnych klas zmian chorobowych w odniesieniu do klasy referencyjnej. Wyniki selekcji uzyskane za pomocą VSCH porównano z wynikami uzyskanymi za pomocą metod odniesienia – RBFN oraz SVM z radialnymi funkcjami bazowymi. Ze względu na większą złożoność obliczeniową i związany z tym dłuższy czas analizy, w przypadku obu tych metod zastosowano przeszukiwanie podprzestrzeni dwuwymiarowych oraz przeszukiwanie przestrzeni trójwymiarowych za pomocą suboptymalnego algorytmu sekwencyjnego przeszukiwania w przód (*ang. sequential floating forward selection – SFFS*) [276], [277].

Aby porównać szybkość działania algorytmów selekcji cech wykonano niżej opisany eksperyment. Wybrano dwie klasy po 128 wektorów składających się z 3049. cech każdy. Następnie dokonano pełnego przeszukania dwuwymiarowych podprzestrzeni cech za pomocą każdej z trzech testowanych metod. Eksperyment przeprowadzono wykorzystując komputer klasy PC z procesorem Intel Core 2 Duo taktowanym zegarem o częstotliwości 2.8 GHz i zainstalowanym systemem operacyjnym Windows XP Professional. Algorytm VSCH zaimplementowano w pojedynczym wątku w języku programowania C++ i wykorzystano w nim kod źródłowy szybkiego algorytmu konstruowania wielotopu wypukłego [270], [278]. W przypadku algorytmów RBFN i SVM wykorzystano ich implementacje dostępne w programie Weka [279]. W przypadku metod SVM oraz RBFN uzyskano czas analizy równy odpowiednio 2240 i 2630 sekund. W przypadku metody VSCH analiza zajęła 420 sekund – obliczenia trwały 5 razy krócej.

W przypadku podprzestrzeni trójwymiarowych, spośród których starano się docelowo wybrać podzbiór cech znaczących, pełne przeszukanie za pomocą metody VSCH zajęło 2 tygodnie. Mierzac czas przeszukania przestrzeni cech o mniejszej wymiarowości oszacowano, że pełne przeszukanie przestrzeni 3049. wymiarowej za pomocą RBFN lub SVM zajęłoby ponad dwa miesiące. Uznano, że czas ten jest nieakceptowalnie długi i podjęto decyzję o zastosowaniu wspomnianego wcześniej algorytmu przeszukiwania SFFS.

Tabela. 6.1. Wyniki selekcji cech

Przedmiot detekcji	Metoda selekcji	Wyselekcjonowane cechy
Krwawienia	VSCH	qCZ5D5SumAverg, iD5GrNonZeros, uMean
	SVM	vMean, Skewness, RMean
	RBFN	vPerc99, Mean, qCH1D5SumAverg
Wrzody rozległe	VSCH	XCZ1D5DifEntrp, Perc90, GPerc90
	SVM	GPerc90, QPerc01, SZ2S5DifEntrp
	RBFN	ZCH1D5Contrast, vPerc50, UMean
Wrzody drobne	VSCH	hCV3D5DifEntrp, UCH1D5DifEntrp, hCV3D5InvDfMom
	SVM	RCZ1D5DifEntrp, GCH5D5DifEntrp, SCHD5InvDfMom
	RBFN	GCH3D5InvDfMom, WavEnLH1, SCH1D5InvDfMom
wybroczyny	VSCH	qPerc99, GCZD5DifEntrp, qCH3D5DifEntrp
	SVM	qGrKurtosis, hPerc90, qMean
	RBFN	qMean, XCZ5D5InvDfMom, HPerc90

W tabeli 6.1 przedstawiono trójki cech będące wynikiem selekcji przeprowadzonej metodami VSCH, SVM oraz RBFN. Można zauważyć, że wynikowe trójki są różne dla różnych metod selekcji. Trójki te najlepiej spełniają kryteria dyskryminacji definiowane przez te metody. Poszczególne trójki cech przeznaczone do detekcji różnych klas anomalii również różnią się między sobą. Nie występuje zatem taka trójka cech, która z punktu widzenia więcej niż jednego kryterium selekcji byłaby optymalna do detekcji więcej niż jednego rodzaju zmiany chorobowej.

Tabela. 6.2. Wyniki klasyfikacji zbiorów uczących i testowych

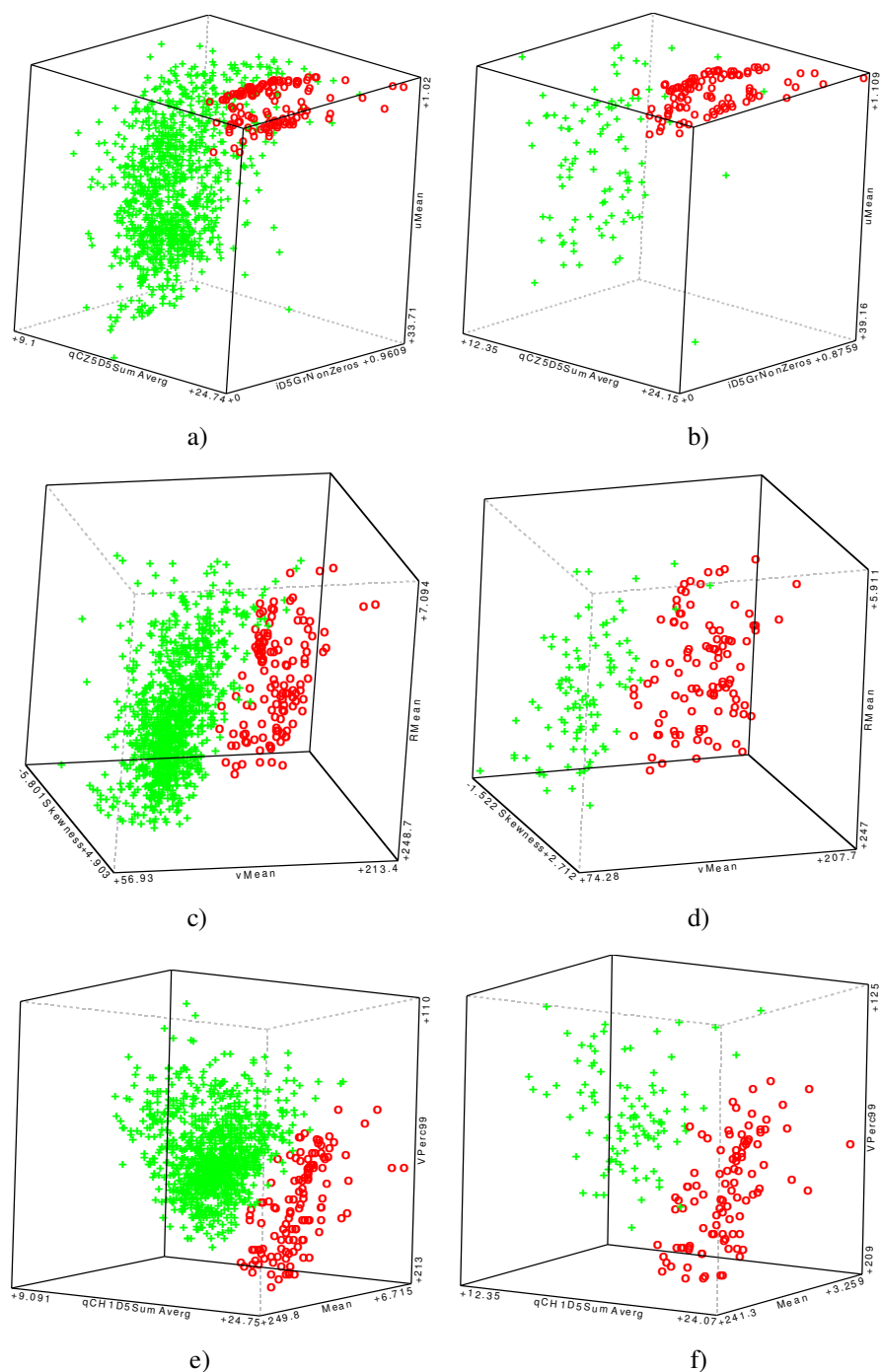
Przedmiot detekcji	Wskaźnik	VSCH		SVM		RBFN	
		Zbiór uczący	Zbiór testowy	Zbiór uczący	Zbiór testowy	Zbiór uczący	Zbiór testowy
Krwawienia	swoistość	1,000	0,990	0,980	0,970	1,000	0,990
	czułość	1,000	0,980	1,000	1,000	1,000	1,000
Wrzody rozległe	swoistość	0,990	1,000	0,950	0,880	0,960	0,760
	czułość	1,000	0,840	0,910	0,980	0,970	0,930
Wrzody drobne	swoistość	0,960	0,960	0,920	0,890	0,980	0,870
	czułość	1,000	0,880	0,970	0,970	0,970	0,940
Wybroczyny	swoistość	0,970	0,770	0,940	0,940	0,890	0,870
	czułość	1,000	0,820	0,990	0,900	0,980	0,920

Zastosowane algorytmy selekcji cech są metodami uczenia nadzorowanego, umożliwiającymi nie tylko selekcję cech, ale również sformułowanie reguł klasyfikacji danych. Reguły uzyskane za ich pomocą wykorzystano w związku z tym do klasyfikacji wektorów zbioru uczącego oraz zbioru testowego. W tabeli 6.2 przedstawiono wartości swoistości (stosunek liczby prawidłowych zaliczeń do klasy referencyjnej do rzeczywistej liczby wektorów klasy referencyjnej) i czułości (stosunek liczby prawidłowych zaliczeń do klasy zmiany chorobowej do rzeczywistej liczby wektorów klasy chorobowej) obliczonych dla zbiorów uczących oraz testowych. Wyniki zamieszczone w tabeli potwierdzają właściwość klasyfikatora VSCH, który dla zbioru uczącego gwarantuje uzyskanie maksymalnej wartości czułości. Co ciekawe, dla trzech spośród czterech analizowanych przypadków anomalii klasyfikator ten umożliwił uzyskanie wskaźników swoistości o wartościach większych niż w przypadku klasyfikatorów SVM i RBFN. Ponadto klasyfikacja krwawień za pomocą VSCH charakteryzuje się wartościami swoistości i czułości równymi jedności, co oznacza bezbłędny wynik. Okazuje się jednak, że w przypadku klasyfikacji wektorów zbiorów testowych wskaźnik czułości jest dla VSCH mniejszy niż dla klasyfikatorów SVM i RBFN. Jednocześnie wskaźniki swoistości uzyskane za pomocą klasyfikatora

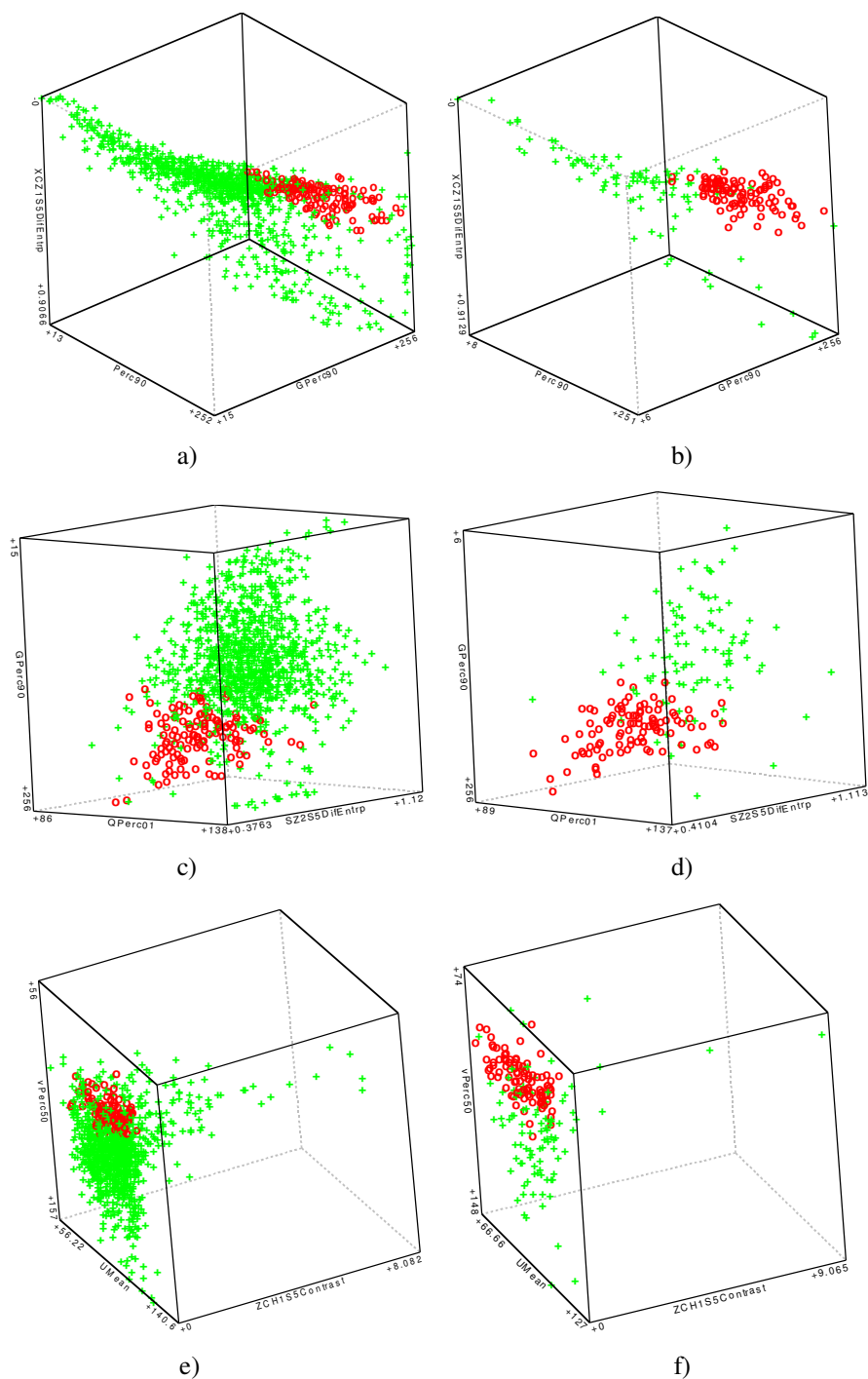
VSCH w trzech przypadkach przewyższają wskaźniki uzyskane za pomocą metod SVM i RBFN. Natomiast w czwartym przypadku, detekcji wybroczyn, wskaźnik swoistości VSCH jest znacząco mniejszy niż wskaźniki uzyskane za pomocą pozostałych klasyfikatorów – VSCH „wykrywa” zmianę chorobową w obrazach, w których ona nie występuje. Metoda klasyfikacji, która maksymalizuje wskaźnik czułości dla danych uczących, w przypadku danych testowych wykazuje się stosunkowo małą czułością przy jednocześnie dużej wartości wskaźnika swoistości klasyfikacji. Wyniki te potwierdzają przydatność zaproponowanej metody selekcji i klasyfikacji danych pokazując, że w niektórych przypadkach umożliwia ona znalezienie reguł klasyfikacji, które pod względem swoistości i czułości są porównywalne lub przewyższają te uzyskiwane za pomocą SVM i RBFN.

Na rysunkach 6.4 i 6.5 przedstawiono wybrane rozkłady wektorów zbiorów uczących i testowych w przestrzeniach wyselekcjonowanych trójek cech. Kąty rzutowania przestrzeni trójwymiarowej starano się dobrać tak, aby w najlepszy sposób wyeksponować rozdzielenie wektorów należących do poszczególnych klas. Wektory obliczone dla obrazów zmian chorobowych oznaczono czerwonymi krzyżykami, natomiast wektory klasy referencyjnej – zielonymi kółeczkami. Subiektywna ocena rozkładów wektorów potwierdza, że uzyskane wyniki selekcji są porównywalne dla wszystkich zastosowanych metod.

Ogólnym wnioskiem z przeprowadzonego eksperymentu jest to, że metody analizy tekstury i barwy obrazu oraz metody selekcji i klasyfikacji cech umożliwiają detekcję zmian chorobowych przewodu pokarmowego w obrazach z endoskopu bezprzewodowego. Szczególnie w przypadku krwawień i rozległych owrzodzeń zaproponowane podejście umożliwia ich wiarygodną detekcję z czułością oraz swoistością większą niż 90%. Należy zwrócić uwagę, że krwawienia i rozległe owrzodzenia widoczne są zazwyczaj w kilku kolejnych obrazach sekwencji wideo, co dodatkowo minimalizuje ryzyko ich przeoczenia, szczególnie w przypadku czułości klasyfikacji na poziomie 0,98-0,99. Wartości swoistości na poziomie 0,94-0,99 również wydają się być satysfakcjonujące. Wartość swoistości 0,94 oznacza, że 6% obrazów przedstawiających normalny stan przewodu pokarmowego może zostać błędnie wskazanych jako przedstawiające anomalie. W przypadku filmu składającego się z 50000 obrazów oznacza to zbiór 3000 zdjęć. Ich przejrzanie nadal jest uciążliwe, lecz wydaje się mniej pracochłonne niż przejrzanie całego filmu.



Rys. 6.4. Rozkłady wektorów cech klas B (oznaczone $\circ$) i R (oznaczone $+$) wyselekcjonowanych metodą: a) VSCH – zbiorcy uczące, b) VSCH – zbiorcy testowe, c) SVM – zbiorcy uczące, d) SVM – zbiorcy testowe, e) RBFN – zbiorcy uczące oraz f) RBFN – zbiorcy testowe



Rys. 6.5. Rozkłady wektorów cech klas E (oznaczone $\circ$) i R (oznaczone $+$) wyselekcjonowanych metodą: a) VSCH – zbiory uczące, b) VSCH – zbiory testowe, c) SVM – zbiory uczące, d) SVM – zbiory testowe, e) RBFN – zbiory uczące oraz f) RBFN – zbiory testowe

W poprzednim teście przeprowadzono selekcję trójek cech (trójwymiarowych podprzestrzeni przestrzeni cech) umożliwiających najlepszą dyskryminację wektorów należących do różnych klas. Ciekawe jest, czy w przypadku algorytmu VSCH możliwe jest uzyskanie lepszych wskaźników klasyfikacji dla podprzestrzeni cech o innej wymiarowości. Kolejny test przeprowadzono więc w celu określenia zależności wskaźników swoistości oraz czułości klasyfikatora VSCH od wymiaru podprzestrzeni cech. Przeprowadzono sprawdzian krzyżowy w taki sposób, że każdy ze zbiorów wektorów analizowanych klas podzielono na pięć możliwie równolicznych podzbiorów. Cztery połączone podzbiory wykorzystywano w procesie uczenia (selekcja cech oraz formułowanie reguł klasyfikacji), a pozostały zbiór do testowania uzyskanego klasyfikatora. Eksperyment powtórzono pięciokrotnie tak, że za każdym razem inny z podzbiorów wykorzystano do testowania.

W eksperymencie wykorzystano cztery klasy obrazów, w tym trzy klasy obrazów ze zmianami chorobowymi, krwawieniami, wrzodami rozległymi i drobnymi oraz klasę referencyjną. Celem było, tak jak w poprzednim eksperymencie, uzyskanie reguł klasyfikacji umożliwiających oddzielenie wektorów cech obrazów zmian chorobowych od wektorów cech obrazów przedstawiających stan normalny.

Selekcji pojedynczych cech oraz par cech (podprzestrzenie jedno i dwuwymiarowe) dokonano za pomocą pełnego przeszukania wszystkich możliwych podprzestrzeni przestrzeni cech. Aby ograniczyć czas analizy, selekcję trójek, czwórek i piątek cech przeprowadzono w przestrzeni pięćdziesięciu wstępnie wyselekcjonowanych cech, które wykazały się największą zdolnością dyskryminacyjną obliczoną podczas selekcji podprzestrzeni dwuwymiarowych.

W każdym z kolejnych pięciu eksperymentów sprawdzianu krzyżowego uzyskano różne kombinacje wyselekcjonowanych cech oraz nieznacznie różniące się wartości wskaźników czułości i swoistości. Uśrednione wartości tych wskaźników zamieszczono w tabeli 6.3. Wynika z niej, że w przypadku zbiorów uczących wartość swoistości rośnie wraz ze wzrostem wymiarowości analizowanej podprzestrzeni cech. W przypadku zbiorów testowych uzyskano podobny efekt – swoistość wzrasta wraz z wymiarowością przestrzeni. Zauważyć można jednak, że czułość, która przy klasyfikacji zbiorów uczących jest zawsze równa jedności, w przypadku klasyfikacji zbiorów testowych maleje wraz ze wzrostem wymiarowości wyselekcjonowanych cech.

Porównując wyniki uzyskane w metodzie VSCH dla trójek cech w tabelach 6.2 i 6.3 można zauważyć, że uzyskane wartości uzyskanych wskaźników są mniejsze w przypadku danych w tabeli 6.3. Może to wynikać z przyjętej metody niepełnego przeszukiwania zbioru cech.

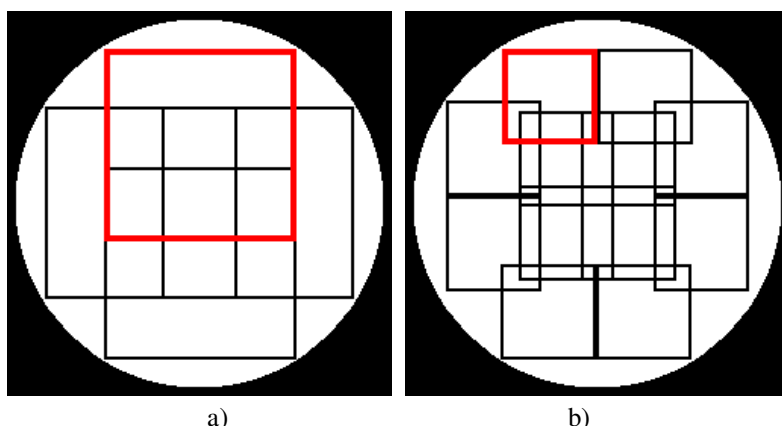
Tabela. 6.3. Uśrednione wyniki klasyfikacji za pomocą VSCH walidacji krzyżowej dla podprzestrzeni cech o wymiarowości od 1 do 5

Zbiory uczące						
Przedmiot detekcji	Wskaźnik	VSCH-1D	VSCH-2D	VSCH-3D	VSCH-4D	VSCH-5D
krwawienia	swoistość	0,962	0,999	1,000	1,000	1,000
	czułość	1,000	1,000	1,000	1,000	1,000
wrzody rozległe	swoistość	0,666	0,916	0,986	0,999	1,000
	czułość	1,000	1,000	1,000	1,000	1,000
wrzody drobne	swoistość	0,601	0,859	1,000	1,000	1,000
	czułość	1,000	1,000	1,000	1,000	1,000

Zbiory testowe						
Przedmiot detekcji	Wskaźnik	VSCH-1D	VSCH-2D	VSCH-3D	VSCH-4D	VSCH-5D
krwawienia	swoistość	0,958	0,997	0,997	0,999	0,999
	czułość	0,991	0,969	0,956	0,943	0,921
wrzody rozległe	swoistość	0,665	0,909	0,978	0,996	0,993
	czułość	0,996	0,956	0,684	0,522	0,675
wrzody drobne	swoistość	0,602	0,838	0,989	0,998	0,998
	czułość	0,982	0,794	0,263	0,246	0,241

6.2. Detekcja zmian chorobowych – deskryptory MPEG-7

Celem kolejnego eksperymentu była ocena zdolności dyskryminacyjnej cech generowanych przez program *eXperimentation Model (XM)* [252] implementujący algorytmy zdefiniowane standardem MPEG-7 oraz porównanie uzyskanych wyników z wynikami pierwszego eksperymentu. Cechy obliczono dla obrazów zmian chorobowych wykorzystanych w pierwszym eksperymencie. Konieczne jednak okazało się zdefiniowanie obszarów zainteresowania o innym kształcie i wielkości. Inaczej niż w pierwszym eksperymencie nie przeprowadzono tutaj selekcji cech, a podczas procedury uczenia klasyfikatora wykorzystano wszystkie obliczone deskryptory. Przeprowadzenie eksperymentu było przedmiotem pracy [280].



Rys. 6.6. Rozkład kwadratowych obszarów zainteresowania w obrazie, w których obliczono deskryptory koloru i tekstury standardu MPEG-7: a) zestaw pierwszy i b) zestaw drugi. Kolorem czerwonym oznaczono wybrane, pojedyncze obszary zainteresowania

Algorytmy zaimplementowane w programie XM umożliwiają obliczanie cech (deskryptorów) tekstury i barwy w prostokątnych obszarach zainteresowania. Obszary zainteresowania nie mogą być zbyt małe – większość metod wymaga, aby rozmiar obszarów był co najmniej kwadratem o rozmiarach 64×64 pikseli, a do obliczenia deskryptorów tekstury homogenicznej (HTD) jest wymagany obszar 128×128 pikseli. W związku z tym przygotowano dwa zestawy kwadratowych, nakładających się na siebie obszarów zainteresowania. W pierwszym zestawie określono cztery obszary o rozmiarach 128×128 pikseli – ich rozkład w obrazie przedstawia rysunek 6.6a. W obszarach tego zestawu obliczono wszystkie deskryptory tekstury i barwy zdefiniowane standardem MPEG-7 – włącznie z HTD. W drugim zestawie określono dwanaście kwadratowych, nakładających się obszarów o rozmiarach 64×64 pikseli, ich rozkład przedstawia rysunek 6.6b. W obszarach drugiego zestawu możliwe było obliczenie deskryptorów tekstury i barwy zdefiniowanych standardem z wyjątkiem HTD.

Wygenerowane wektory cech podzielono losowo na zbiory uczące i testowe, przy czym w skład zbiorów uczących weszło około $2/3$, a w skład testowych około $1/3$ wszystkich wektorów. Do klasyfikacji danych wykorzystano metody zaimplementowane w programie Weka [279]. Najlepsze wyniki pod względem wskaźników swoistości i czułości uzyskano stosując nieliniowy klasyfikator SVM wykorzystujący do rozdzielania wektorów różnych klas funkcje wielomianowe drugiego stopnia. W procesie uczenia nadzorowanego wykorzystano algorytm minimalnej optymalizacji sekwencyjnej (ang. *Sequential Minimal Optimization – SMO*) [276], [277]. Należy nadmienić, że w drugim eksperymencie nie przeprowadzono wstępnej selekcji cech, a w procesie uczenia i klasyfikacji wykorzystywano komplet wygenerowanych deskryptorów. Było to możliwe ponieważ liczba deskryptorów barwy i tekstury generowanych przez program XM jest o rząd wielkości mniejsza niż liczba cech generowanych przez program MaZda. Wyniki klasyfikacji zbiorów testowych przedstawiono w tabeli 6.4.

Tabela. 6.4. Wyniki klasyfikacji zbiorów testowych

Zbiory testowe			
Przedmiot detekcji	Wskaźnik	Obszary 128×128	Obszary 64×64
krwawienia	swoistość	0,570	0,870
	czułość	0,570	0,870
wrzody rozległe	swoistość	0,670	0,670
	czułość	1,000	1,000
wrzody drobne	swoistość	0,750	0,910
	czułość	0,800	1,000
wybroczyny	swoistość	0,630	0,940
	czułość	0,870	0,880

Analizując uzyskane wskaźniki można zaobserwować, że są one większe w przypadku zastosowania mniejszego obszaru zainteresowania – pomimo niewykorzystania w tym przypadku deskryptorów tekstury homogenicznej (HTD). Potwierdza to hipotezę, że istotne jest obliczanie cech w stosunkowo małych obszarach zainteresowania, dzięki czemu drobne zmiany chorobowe reprezentowane są w większym stopniu. Porównując wyniki uzyskane w pierwszym (tabela 6.2 – klasyfikator SVM) i drugim eksperymencie (tabela 6.4 – obszary 64×64) można zauważyć, że w przypadku krwawień i wybroczyn lepsze wyniki detekcji uzyskano wykorzystując program MaZda. Przykładowo dla krwawień uzyskano czułość = 1 i swoistość = 0,97, podczas gdy stosując deskryptory standardu MPEG-7, uzyskano czułość i swoistość równe 0,87. W przypadku detekcji drobnych wrzodów wartości czułości i swoistości są natomiast większe w przypadku zastosowania deskryptorów MPEG-7. Należy jednak zwrócić uwagę, że w pierwszym eksperymencie (podrozdział 6.1) do klasyfikacji wykorzystano jedynie trójkę wybranych cech, podczas gdy w eksperymencie opisanym w niniejszym podrozdziale wykorzystano wszystkie deskryptory tekstury i koloru standardu MPEG-7 z wyjątkiem HTD.

6.3. Wykorzystanie map cech do segmentacji obrazów

Celem trzeciego eksperymentu była odpowiedź na pytanie, czy cechy tekstury i koloru są przydatne do segmentacji obrazów endoskopowych. Jeśli takie cechy przedstawić w postaci map i dla każdego piksela zastosować reguły klasyfikacji, to powinno to umożliwić podział obrazu na rozłączne obszary określające powierzchnię zdrową oraz zmienioną chorobowo. Jeśli tak jest w istocie, to dostępne staje się narzędzie pozwalające opisać zmianę chorobową w sposób ilościowy pod względem zajmowanego przez nią pola powierzchni.

W kolejnych podrozdziałach przedstawiono wyniki uzyskane dla zestawów cech wyselekcjonowanych na dwa różne sposoby. W podrozdziale 6.3.1 wykorzy-

stano cechy tekstury i koloru wyselekcjonowane do detekcji schorzeń w pierwszym eksperymencie (tabela 6.1). W podrozdziale 6.3.2 zaproponowano odmienne podejście zarówno co do sposobu generowania cech, jak i ich selekcji.

6.3.1. Segmentacja za pomocą cech wybranych do detekcji

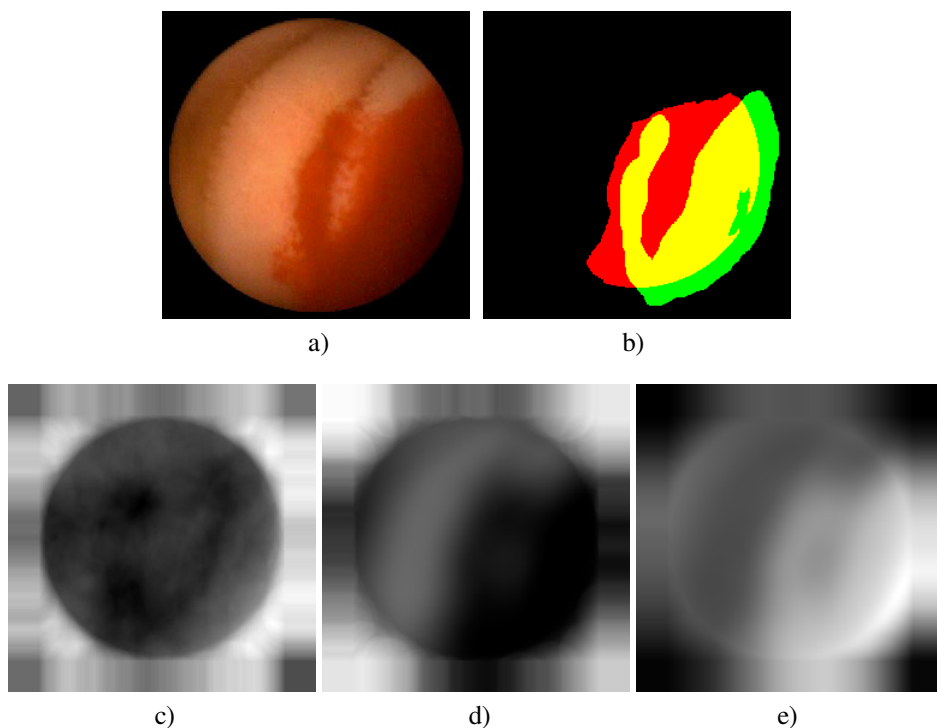
Na potrzeby eksperymentu wybrano po szesnaście obrazów przedstawiających krwawienia, wrzody drobne oraz wrzody rozległe, a także 64 obrazy, w których nie stwierdzono zmian chorobowych. Następnie za pomocą programu MaZda dla każdego z obrazów wygenerowano mapy tych cech, które zostały wyselekcjonowane metodą VSCH w pierwszym eksperymencie (tabela 6.1). Cechy obliczano w obszarach zainteresowania w kształcie zbliżonym do koła składającym się z 2009 pikseli, to jest identycznych jak obszary zainteresowania z pierwszego eksperymentu (podrozdział 6.1). Następnie, wykorzystując odpowiednie trójki cech oraz reguły klasyfikacji otrzymane za pomocą VSCH dokonano segmentacji obrazów – klasyfikacji poszczególnych pikseli ze względu na ich przynależność do obszarów zmiany chorobowej lub fragmentu obrazu nie mającego charakteru takiej zmiany. Uzyskane wyniki segmentacji porównano z wynikami ręcznej segmentacji wykonanej przez ekspertów gastroenterologów.

Tabela. 6.5. Porównanie obszarów wyznaczonych automatycznie z obszarami oznaczonymi przez eksperta

Obszary	Wskaźnik	Średnia ($\pm$ odchylenie standardowe)
krwawień	dokładność	0,597 ($\pm$ 0,221)
	kompletność	0,946 ($\pm$ 0,126)
	wsp. Jaccarda	0,353 ($\pm$ 0,238)
wrzodów rozległych	dokładność	0,929 ($\pm$ 0,061)
	kompletność	0,306 ($\pm$ 0,248)
	wsp. Jaccarda	0,268 ($\pm$ 0,208)
wrzodów drobnych	dokładność	0,576 ($\pm$ 0,134)
	kompletność	0,608 ($\pm$ 0,286)
	wsp. Jaccarda	0,043 ($\pm$ 0,036)

Na rysunku 6.7 przedstawiono przykładowy obraz oraz obliczone dla niego mapy cech służące do klasyfikacji krwawień. Rysunek 6.8 przedstawia wybrane obrazy zmian chorobowych wraz z wynikami segmentacji ręcznej i automatycznej, natomiast rysunek 6.9 przedstawia obrazy, w których lekarz nie stwierdził występowania zmian chorobowych, natomiast zaproponowany algorytm wskazał obszary domniemanego stanu chorobowego. W tabeli 6.5 zamieszczono uśrednione wskaźniki ilościowe porównania obszarów zaznaczonych ręcznie i automatycznie, takie jak kompletność, dokładność oraz współczynnik Jaccarda. Współczynniki te obliczono analizując wyniki klasyfikacji 31417 pikseli zawartych w kole znajdującym się w środku obrazu endoskopowego. Pole powierzchni tego

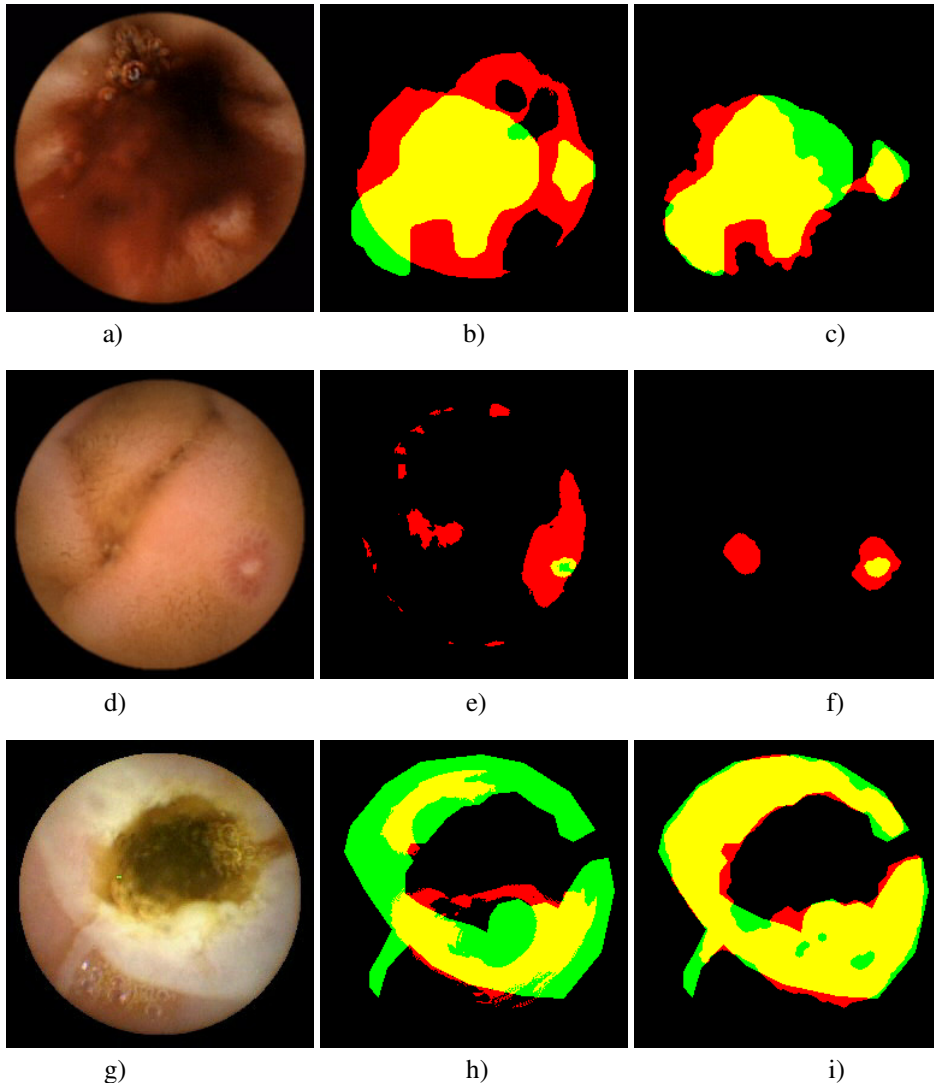
koła stanowi około połowy powierzchni całego obrazu. Obejmuje ono piksele, dla których możliwe było obliczenie cech tekstury wewnątrz obszarów zainteresowania całkowicie zawartych w polu widzenia kamery endoskopu bezprzewodowego. Wskaźnik dokładności obliczono jako stosunek liczby pikseli uznanych jednocześnie za należące do obszaru zmiany chorobowej przez zastosowany algorytm klasyfikacji oraz eksperta do liczby pikseli obszaru uznanego przez algorytm za zmieniony chorobowo. Kompletność jest stosunkiem liczby pikseli części wspólnej porównywanych obszarów zmiany chorobowej do liczby pikseli należących do obszaru tej zmiany wskazanego przez eksperta. Współczynnik Jaccarda jest natomiast stosunkiem liczby pikseli części wspólnej porównywanych obszarów zmiany chorobowej do liczby pikseli należących do przynajmniej jednego z tych obszarów.



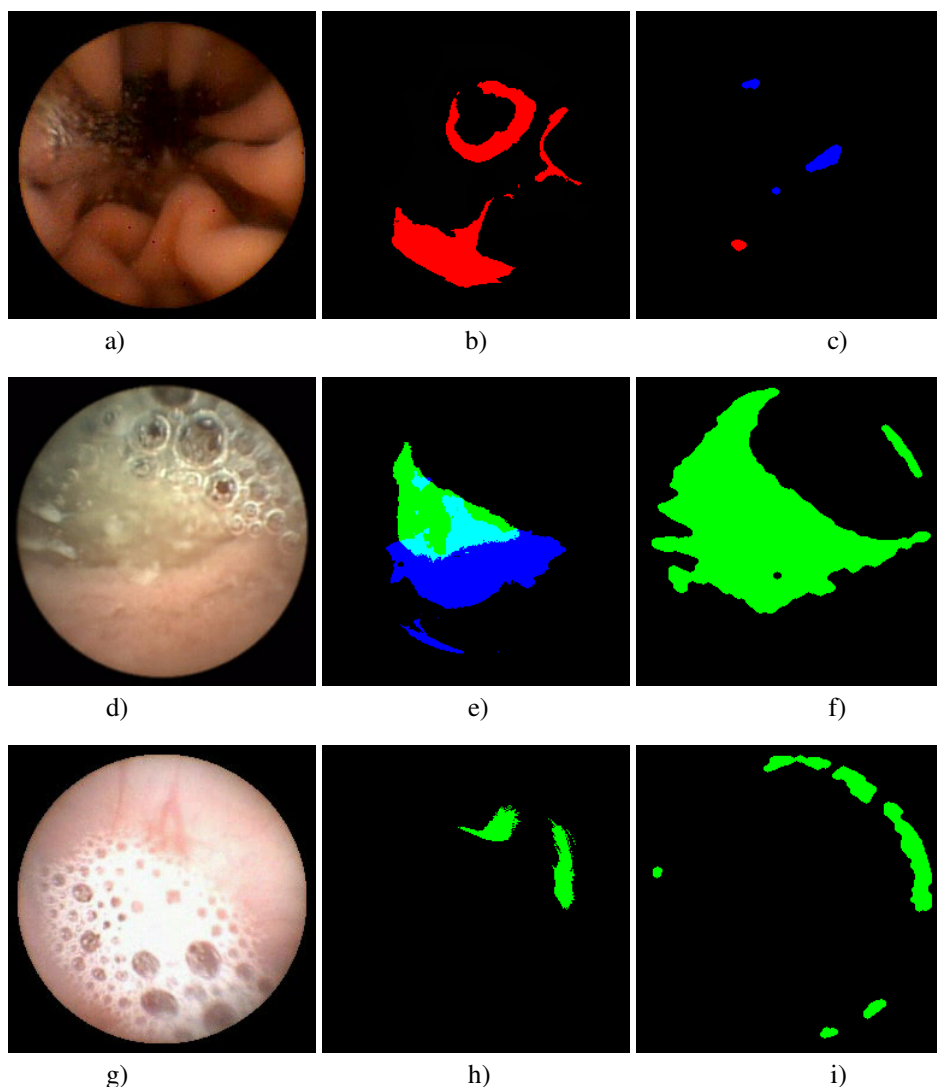
Rys. 6.7. Segmentacja obrazu krwawienia: a) obraz endoskopowy, b) obszar krwawienia zaznaczony przez lekarza (zielony i żółty) oraz wynik segmentacji automatycznej (czerwony i żółty), c) mapa cechy qCH1D5DifEntrp, d) mapa cechy uMean oraz e) mapa cechy vCN3D5SumAverg

Współczynniki Jaccarda dla krwawień i wrzodów rozległych wynoszą odpowiednio 0,353 i 0,268. Wartości te oznaczają, że przeciętnie jedynie około jedna trzecia powierzchni obu obszarów jest dla nich wspólna. W przypadku segmentacji wrzodów drobnych wartość współczynnika Jaccarda jest wyjątkowo mała i wynosi średnio 0,043. Liczby te świadczą o tym, że wyselekcjonowane cechy i reguły klasyfikacji opracowane do detekcji zmian chorobowych nie mogą być bezpośrednio zastosowane do wiarygodnej segmentacji takich obrazów.

Z czego mogą wynikać tak znaczne różnice między wynikami uzyskanymi dla detekcji i segmentacji? Prawdopodobnych przyczyn może być kilka. Po pierwsze, w procesie uczenia wykorzystano wektory cech klasy odniesienia (R) wygenerowane dla obrazów innych niż obrazy przedstawiające zmianę chorobową. Wydaje się, że aby dokonać procesu uczenia na potrzeby segmentacji, wektory cech klasy odniesienia powinny zostać wygenerowane również dla obszarów niezmiennych chorobą występujących w obrazach przedstawiających takie zmiany.



Rys. 6.8. Obrazy endoskopowe oraz obszary zmiany chorobowej, obszar zielony i żółty wskazany przez lekarza oraz obszar czerwony i żółty wyznaczony automatycznie (żółty oznacza część wspólną): a-c) krwawienie, d-f) wrzód drobny, g-i) rozległy wrzód z włókniałą. W środkowej kolumnie (b, e, h) zamieszczono wyniki segmentacji za pomocą cech wyselekcjonowanych do detekcji, w prawej kolumnie (c, e, i) za pomocą cech wyselekcjonowanych do segmentacji.



Rys. 6.9. Obrazy endoskopowe z domniemanymi obszarami zmian chorobowych wyznaczonymi automatycznie – ekspert takich zmian nie stwierdził: a-c) domniemane krwawienie (oznaczony na czerwono) i drobne owrzodzenie (obszary niebieskie), d-f) domniemany wrzód drobny (obszary niebieskie) i zwłókniały (obszary zielone) oraz g-i) domniemany wrzód zwłókniały (oznaczony na zielono). W środkowej kolumnie (b, e, h) zamieszczono wyniki segmentacji za pomocą cech wyselekcjonowanych do detekcji, w prawej kolumnie (c, e, i) za pomocą cech wyselekcjonowanych do segmentacji.

Po drugie, wektory cech drobnych wrzodów były obliczane w obszarach zainteresowania obejmujących powierzchnię takich wrzodów, przy czym sam obszar zainteresowania był zazwyczaj większy od obszaru wrzodu. Spowodowało to sytuację, w której podczas segmentacji jako zmienione chorobowo były oznaczane również te fragmenty, które znajdują się odpowiednio blisko rzeczywistego owrzodzenia. Potwierdzają to wyniki segmentacji, w których pola powierzchni

wyznaczone automatycznie były znacznie większe od obszarów zmiany wskazanej przez lekarza i zazwyczaj je obejmowały. Potwierdzają to również wyniki ilościowe segmentacji obrazów krwawień (duża kompletność równa 0,946) oraz w pewnym stopniu wyniki segmentacji drobnych wrzodów (kompletność 0,608). Ciekawym przykładem jest obraz z rysunku 6.7a – obszar oznaczony przez lekarza nie obejmuje fragmentu obrazu, w którym krew jest obecna, choć rozrzedzona. Procedura automatycznej segmentacji oznaczyła również ten fragment.

Po trzecie, wyselekcjonowane cechy mogą charakteryzować nie tyle sam obszar zmiany chorobowej, co pewne otoczenie tej zmiany. Potwierdzającym to przypadkiem jest wynik segmentacji z rysunku 6.7b. Część wspólna obszarów obejmuje pierścień otaczający owrzodzenie, algorytm automatycznej segmentacji nie oznacza wnętrza wrzodu, rozciąga się natomiast na otaczającej go powierzchni. Może to oznaczać, że wyselekcjonowane cechy oraz opracowane reguły detekcji drobnych wrzodów wykrywają raczej strefę przejściową pomiędzy powierzchnią zdrową a zmienioną chorobowo. Nie stanowi to problemu podczas detekcji, natomiast prowadzi do błędów w przypadku segmentacji.

Po czwarte, w przypadku rozległych owrzodzeń wektory cech charakteryzujących tę zmianę były obliczane w obszarach zainteresowania całkowicie zawartych wewnątrz takich obszarów zmienionych chorobowo. Odwrotnie zatem niż w przypadku drobnych owrzodzeń, wyznaczane tu obszary mogą być mniejsze niż rzeczywiste obszary zwyrodnień, przy czym powinny się one w nich zawierać. Potwierdzają to duża dokładność (0,929) przy jednoczesnej małej kompletności (0,306).

W eksperymencie dokonano również segmentacji 64. obrazów, w których nie stwierdzono występowania anomalii. Pomimo braku zmian chorobowych zastosowane algorytmy segmentacji oznaczyły obszary domniemanych krwawień w czterech, obszary owrzodzeń zwłókniałych w trzech oraz obszary wrzodów drobnych w dwudziestu jeden obrazach. Analizując oznaczone obszary domniemanych krwawień można zauważyć, że rzeczywiście występują w nich fragmenty zaczerwienione. W jednym z nich oznaczony obszar pokrywał się z zaczerwienionym konturem widocznym pomiędzy silnie oświetloną fałdą przewodu pokarmowego a ciemnym tłem – prawdopodobnie wynik aberracji chromatycznej układu optycznego soczewki. W pozostałych obrazach występowały czerwone punkty będące prawdopodobnie wynikiem zakłóceń lub defektów przetwornika optoelektrycznego – przykładem czego jest obraz z rysunku 6.8b. W przypadku błędów segmentacji polegających na oznaczeniu fragmentów domniemanych owrzodzeń przyczyną były jasne obiekty treści pokarmowej przypominające owrzodzenie (rysunek 6.8e) lub prześwietlone fragmenty obrazu (rysunek 6.8h). Wydaje się, że tego typu błędów nie da się wyeliminować analizując obrazy statyczne. Lekarz weryfikuje, czy dany element jest wrzodem obserwując jego ruch względem ścianek przewodu pokarmowego. Jeśli obiekt przesuwają się względem nich, to nie stanowi części przewodu pokarmowego i nie jest uznawany za owrzodzenie. Oznacza to, że aby prawidłowo dokonać detekcji lub segmentacji, niezbędna jest analiza ruchu obszaru domniemanej anomalii względem ścianek układu trawiennego – co stanowi postulat dla przyszłych opracowań i eksperymentów.

6.3.2. Segmentacja za pomocą cech wybranych do segmentacji

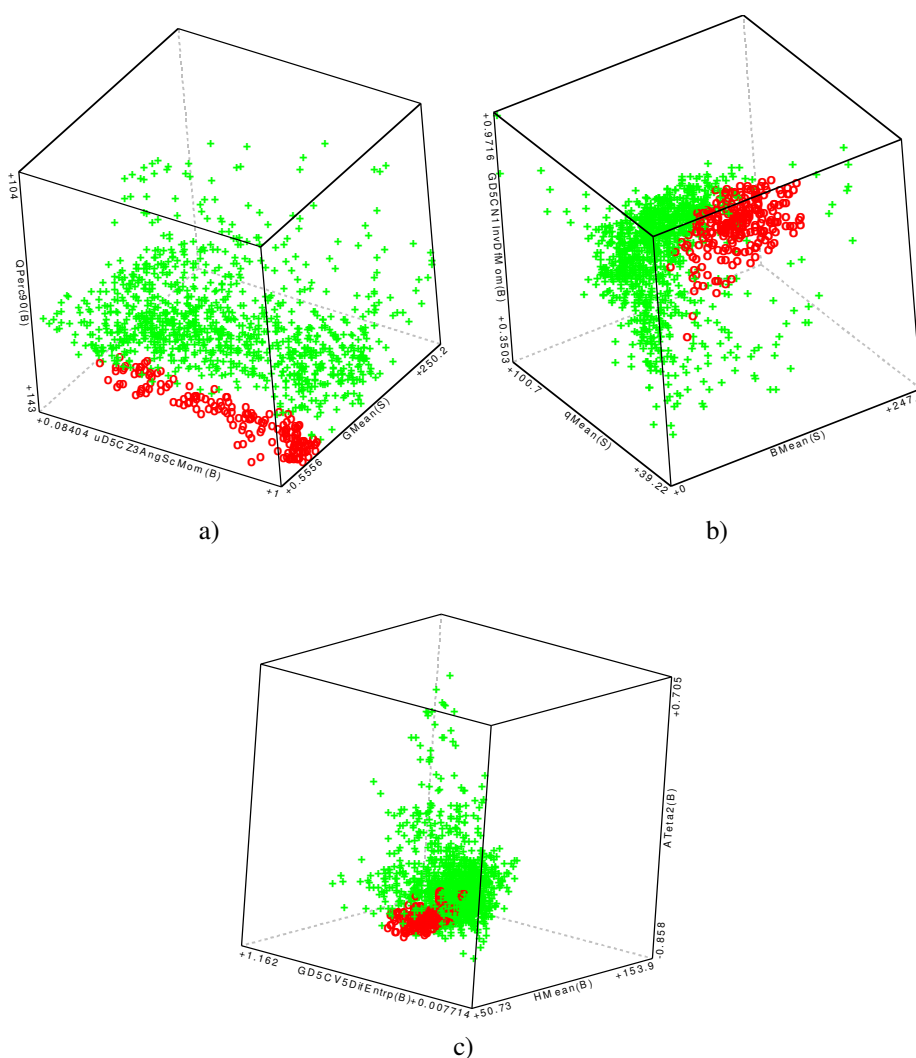
Celem kolejnego badania było uzyskanie odpowiedzi na pytanie, czy i w jakim stopniu możliwe jest poprawienie wyników segmentacji po zastosowaniu cech specjalnie do tego celu wyselekcjonowanych. Różnice w procedurze generowania i selekcji cech w stosunku do eksperymentu opisanego w podrozdziale 6.1 są następujące: (1) Wygenerowano nowy zbiór uczący wektorów cech. W tym celu w sposób przypadkowy wybrano punkty należące oraz nienależące do obszarów zmian chorobowych wskazanych przez lekarzy. Cechy zostały obliczone w otoczeniach tych punktów bez względu na to jak duża część tego otoczenia była zawarta w obszarze zmiany chorobowej. (2) Zbiór cech obliczanych w otoczeniach o kształcie koła składającego się z 2009 pikseli uzupełniono o cechy średnich oraz wariancji składowych kolorów obliczane w otoczeniach o kształcie kwadratu 3×3 piksele. (3) Zbiór uczący został uzupełniony o wektory cech wygenerowane dla punktów z obszarów niezmiennych chorobowo, w obrazach przedstawiających takie zmiany. Wektory te oznaczono jako należące do klasy referencyjnej – bez zmiany chorobowej.

Tabela. 6.6. Wyniki selekcji cech do zastosowania w segmentacji

Obszary	Wyselekcjonowane cechy
Krwawień	QPerc90(B), uD5CZ3AngScMom(B), GMean(S)
Wrzodów rozległych	GD5CN1InvDfMom(B), qHMean(S), BMean(S)
Wrzodów drobnych	GD5CV5DifEntrp(B), HMean(B), ATeta2(B)

Do wyselekcjonowania cech znaczących wykorzystano metodę VSCH – nazwy wybranych trójek cech zamieszczono w tabeli 6.7, natomiast rozkłady wektorów zbioru uczącego w wynikowych podprzestrzeniach cech przedstawiono na rysunku 6.10. W nazwach cech wymienionych w tabeli i na rysunku literą (B) oznaczono, że cecha została obliczona w otoczeniu o kształcie koła, natomiast litera (S) oznacza, że zastosowano otoczenie 3×3 piksele.

Segmentacji poddano te same obrazy, które wykorzystano w eksperymencie opisanym w podrozdziale 6.3.1. Wykorzystano mapy wyselekcjonowanych cech z tabeli 6.6 oraz reguły klasyfikacji otrzymane za pomocą metody VSCH. Podobnie jak w poprzednim przypadku uzyskane wyniki porównano z wynikami ręcznej segmentacji wykonanej przez ekspertów gastroenterologów. W tabeli 6.7 zamieszczono wskaźniki ilościowe porównujące obszary zaznaczone ręcznie i automatycznie, takie jak kompletność, dokładność oraz współczynnik Jaccarda. Na rysunku 6.11 przedstawiono przykładowy obraz z widocznymi krwawieniami, obliczone dla niego mapy cech oraz wynik segmentacji. Zarówno mapy, jak i wynik segmentacji można porównać z wynikiem segmentacji tego samego przypadku z zastosowaniem cech wyselekcjonowanych do detekcji zamieszczonego na rysunku 6.7. Ponadto wybrane wyniki segmentacji przedstawia prawa kolumna obrazów na rysunkach 6.8 i 6.9.



Rys. 6.10. Podprzestrzenie cech . Rozkłady wektorów cech wyselekcjonowane do celów segmentacji obrazów. Podprzestrzenie cech wyselekcjonowano do segmentacji obrazów a) krwawień, b) owrzodzeń rozległych oraz c) wrzodów drobnych. Wektory zmian chorobowych oznaczone o, wektory klasy referencyjnej oznaczone +

Tabela. 6.7. Porównanie obszarów wyznaczonych: automatycznie oraz przez eksperta (A), automatycznie oraz przez innego eksperta (B), przez obu ekspertów

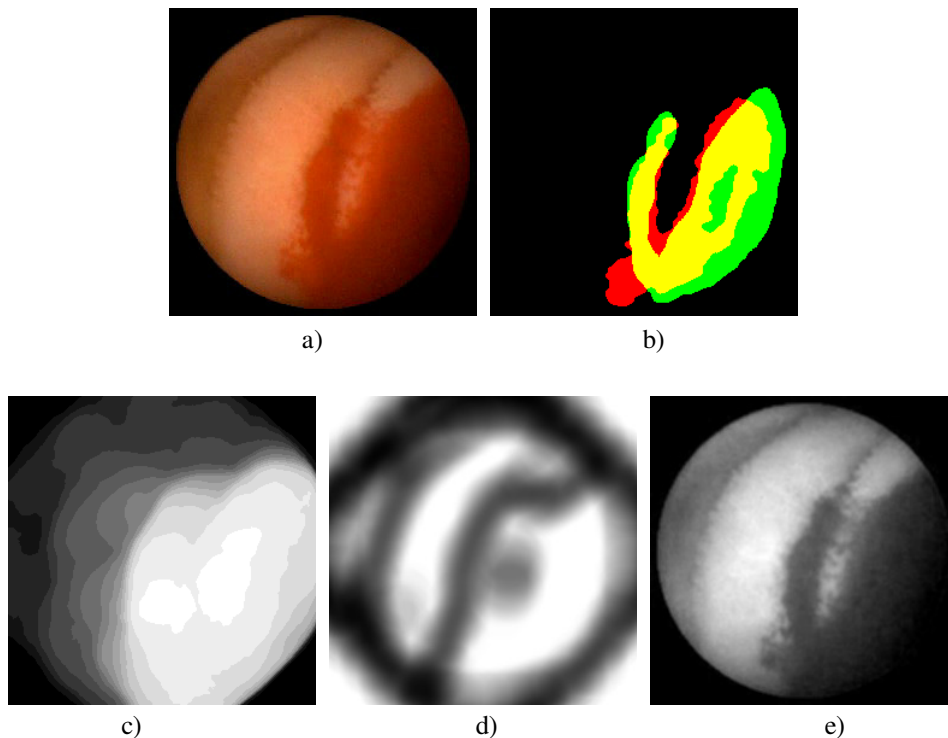
Obszary	Wskaźnik	Średnia (odchylenie standardowe)		
		A	B	C
krwawień	dokładność	0,605 ($\pm 0,226$)	0,682 ($\pm 0,382$)	0,461 ($\pm 0,296$)
	kompletność	0,732 ($\pm 0,177$)	0,485 ($\pm 0,291$)	0,797 ($\pm 0,385$)
	wsp. Jaccarda	0,487 ($\pm 0,179$)	0,447 ($\pm 0,277$)	0,459 ($\pm 0,277$)
wrzodów rozległych	dokładność	0,905 ($\pm 0,101$)	0,656 ($\pm 0,219$)	0,859 ($\pm 0,244$)
	kompletność	0,669 ($\pm 0,184$)	0,653 ($\pm 0,231$)	0,638 ($\pm 0,263$)
	wsp. Jaccarda	0,621 ($\pm 0,172$)	0,513 ($\pm 0,209$)	0,604 ($\pm 0,250$)
wrzodów drobnych	dokładność	0,267 ($\pm 0,114$)	0,688 ($\pm 0,241$)	0,289 ($\pm 0,103$)
	kompletność	0,696 ($\pm 0,257$)	0,514 ($\pm 0,167$)	0,981 ($\pm 0,050$)
	wsp. Jaccarda	0,235 ($\pm 0,142$)	0,395 ($\pm 0,133$)	0,287 ($\pm 0,101$)

Kolumna A tabeli 6.7 zawiera wielkości porównania obszarów wydzielonych automatycznie oraz obszarów referencyjnych oznaczonych przez jednego z ekspertów. Po przeprowadzeniu segmentacji automatycznej poproszono innego gastroenterologa o ponowne, niezależne oznaczenie obszarów zmian chorobowych w obrazach wykorzystywanych w eksperymencie. Wyniki porównania obszarów wydzielonych automatycznie oraz oznaczonych przez drugiego z ekspertów zawiera kolumna B tabeli. W kolumnie C zamieszczono wskaźniki będące wynikiem porównania obszarów wskazanych przez obie osoby. Uzyskane wyniki pokazują, że rozbieżności pomiędzy obszarami wskazanymi przez ekspertów są porównywalne z rozbieżnościami pomiędzy obszarami wyznaczonymi automatycznie a wskazanymi przez lekarzy.

Ponadto można zauważyć, że wyniki segmentacji obrazów przedstawiających zmiany chorobowe są dokładniejsze w przypadku zastosowania cech wyselekcjonowanych specjalnie dla celów segmentacji. Potwierdza to zarówno subiektywna ocena kształtu i położenia otrzymanych obszarów (przykłady z rysunku 6.8), jak i wskaźniki ilościowe (porównanie danych w tabelach 6.5 oraz 6.7). W przypadku współczynnika Jaccarda uzyskano wartości ponad dwukrotnie większe w przypadku segmentacji obrazów krwawień i wrzodów rozległych oraz ponad pięciokrotnie większe w przypadku wrzodów drobnych.

Tak jak i w poprzednim eksperymencie, dokonano segmentacji 64 obrazów, w których nie stwierdzono występowania anomalii. Pomimo braku zmian chorobowych zastosowane algorytmy segmentacji oznaczyły obszary domniemanych krwawień w czterech, obszary owrzodzeń zwłókniałych w 28 oraz obszary wrzodów drobnych w 23 obrazach. Dla przypomnienia w poprzednim eksperymencie uzyskano liczby odpowiednio czterech, trzech i 21 takich przypadków. Uzyskany wynik jest więc gorszy. Oznacza to, że gdyby cechy wyselekcjonowane do seg-

mentacji wykorzystać w detekcji zmian chorobowych, swoistość takiej metody byłaby stosunkowo niewielka.



Rys. 6.11. Segmentacja obrazu krwawienia: a) obraz endoskopowy, b) obszar krwawienia zaznaczony przez lekarza (zielony i żółty) oraz wynik segmentacji automatycznej (czerwony i żółty), c) mapa cechy QPerc90(B), d) mapa cechy uD5CZ3AngScMom(B) oraz e) mapa cechy GMean(S)

Czy możliwa jest poprawa jakości segmentacji statycznych obrazów endoskopowych? W przypadku zastosowania proponowanych metod selekcji cech i uczenia nadzorowanego wymagałoby to przeprowadzenia nowych obliczeń wektorów cech – prawdopodobnie w znacznie większej liczbie niż miało to miejsce w pierwszym eksperymencie, optymalizacji wielkości obszaru zainteresowania oraz optymalizacji parametrów algorytmów obliczania map cech, takich jak normalizacja i standaryzacja histogramu jasności czy też dobór liczby poziomów jasności. Proces taki byłby pracochłonny, długotrwały i wymagający pod względem obliczeniowym. Dlatego w [67–69] zaproponowano inne podejście do problemu segmentacji obrazów endoskopowych. Wykorzystano w nim mapy cech statystycznych (średniej, wariancji, skośności i kurtozy) histogramów składowych przestrzeni kolorów RGB, YUV i HSV. Następnie przeprowadzono segmentację obrazów na podstawie wygenerowanych map, wykorzystując algorytm *k*-średnich (ang. *k-means*) [210], [281] należący do grupy metod uczenia nienadzorowanego – analizy skupień, tzw. klasteryzacji danych (ang. *cluster analysis*) [210], [282]. Jeśli potraktować każdy piksel analizowanego obrazu jako osobne zdarzenie opisane wektorem cech, to algorytm ten umożliwia segmentację obrazu na pewną

liczbę różniących się istotnie między sobą obszarów. Okazuje się, że kształt i położenie uzyskiwanych obszarów w istotny sposób są powiązane z kształtem i położeniem obszarów zmian chorobowych. Ocena przynależności danego obszaru do klasy anomalii lub klasy prawidłowego stanu przewodu pokarmowego może być następnie przeprowadzona metodami uczenia nadzorowanego i klasyfikacji, np. za pomocą wyselekcjonowanych cech i klasyfikatorów VSCH, RBFN lub SVM. Wstępne, obiecujące wyniki segmentacji obrazów endoskopowych za pomocą metod uczenia nienadzorowanego przedstawiono w [68].

Podsumowanie

W pracy przedstawiono wyniki działania oryginalnych algorytmów analizy i przetwarzania obrazów cyfrowych, których celem jest wspomaganie diagnostyki medycznej przewodu pokarmowego wykorzystującej technikę endoskopii bezprzewodowej. W rozdziale trzecim przedstawiono nowatorską metodę analizy ruchu w filmie endoskopowym, która umożliwia również utworzenie pojedynczego obrazu będącego reprezentacją całego przewodu pokarmowego. W rozdziale czwartym przedstawiono wyniki oceny metody. Wyniki testów o charakterze zarówno jakościowym, jak i ilościowym wykazały, że umożliwia ona szybką lokalizację wielu rodzajów zmian chorobowych oraz ocenę przydatności całego filmu jeszcze przed przystąpieniem do jego analizowania. W szczególności metoda ta umożliwia szybką identyfikację obszarów krwawień, owrzodzeń oraz fragmentów filmu nieprzydatnych do celów diagnostycznych ze względu na widoczną treść pokarmową przesłaniającą powierzchnię przewodu pokarmowego.

Wyniki analizy ruchu zastosowano do sterowania szybkością odtwarzania wideo. Dzięki zastosowanemu rozwiązaniu fragmenty filmu, w których ruch endoskopu był powolny są przewijane z większą szybkością, natomiast te, które charakteryzuje szybkie przemieszczanie się kapsułki endoskopu lub wzmożona aktywność ruchową jelit są wyświetlane w spowolnionym tempie. Zwalnia to osobę oglądającą film od konieczności ręcznego sterowania szybkością odtwarzania i pozwala uniknąć konieczności przewijania filmu wstecz w przypadku nagłego wystąpienia skurczów mieszających – w przypadku takich skurczów film należy oglądać bardzo wnikliwie, obraz po obrazie. Wyniki ilościowe porównujące efekty procesów diagnostycznych z zastosowaniem i bez zastosowania proponowanej metody wykazały zwiększenie efektywności procesu w przypadku jej zastosowania – nastąpiło nieznaczne skrócenie czasu poświęcanego na diagnozę oraz wyraźny, bo dwudziestoprocentowy, wzrost liczby zidentyfikowanych zmian chorobowych oraz miejsc charakterystycznych układu trawiennego.

W celu detekcji wybranych rodzajów anomalii zaproponowano algorytm łączący analizę tekstury i barwy obrazu oraz selekcję cech obrazu i klasyfikację z wykorzystaniem metod uczenia nadzorowanego. Zaproponowano też zastosowanie uzyskanych reguł klasyfikacji do segmentacji obrazów endoskopowych, której celem miało być wydzielenie obszarów zmienionych chorobowo. Aspekty techniczne zaproponowanych rozwiązań przedstawiono w rozdziale piątym, natomiast ich walidacja stała się przedmiotem rozdziału szóstego. Wykazano, że zaproponowane metody są przydatne do detekcji krwawień oraz owrzodzeń zwłókniałych, zarówno rozległych, jak i drobnych. Uzyskano także satysfakcjonujące wyniki detekcji obrazów przedstawiających wybroczyny i angiektazje.

Zastosowanie metod klasyfikacji z wykorzystaniem cech barwy i tekstury umożliwiło również segmentację obrazów i wyodrębnienie obszarów wybranych zmian chorobowych. Uzyskane wyniki segmentacji, choć nie w pełni zadowalające, dały podstawy do wnioskowania o właściwościach obrazów z endoskopu bezprzewodowego i wskazały kierunki dalszych badań.

Z przeprowadzonych testów wynika, że opracowane metody komputerowej analizy obrazów cyfrowych z endoskopu bezprzewodowego pozwalają zmniejszyć nakład pracy lekarza podczas oglądania filmu endoskopowego oraz zwiększają komfort jego pracy, ponadto automatyzują one procedurę interpretacji filmu, a także umożliwiają analizę ilościową wybranych zmian chorobowych. W badaniach wykorzystano filmy uzyskane z endoskopu bezprzewodowego wykorzystywanego w praktyce klinicznej, nie wprowadzono w nim żadnych zmian konstrukcyjnych ani nie wykorzystano żadnych uzupełniających metod obrazowania. Wykazano zatem, że zwiększenie efektywności diagnostyki jest możliwe wyłącznie poprzez opracowanie odpowiednich algorytmów komputerowych.

Opracowane metody numeryczne są oryginalnymi rozwiązaniami. Ich zastosowanie nie ogranicza się wyłącznie do analizy obrazów z endoskopu bezprzewodowego, chociaż ich powstanie było wynikiem takiej właśnie potrzeby.

Do analizy ruchu własnego kamery znajdującej się wewnątrz przewodu pokarmowego opracowano model deformowalny o kształcie połączonych ze sobą koncentrycznych pierścieni. Do obliczenia szybkości postępowej wykorzystywano uproszczone równania uzyskane przy założeniu ograniczeń stopni swobody ruchu kapsułki endoskopu. W modelu deformowalnym oryginalna jest nie tylko koncepcja jego kształtu połączeń, ale również metoda obliczania stopnia deformacji i tzw. naprężeń. Metoda ta jest na tyle ogólna i uniwersalna, że może być zastosowana w modelach o innym kształcie i topologii, takich jak aktywny kontur, deformowalny wzorec czy powierzchnia trójwymiarowa. Co istotne, jest ona pozbawiona wad dotychczas wykorzystywanych metod bazujących na równaniach struny lub membrany oraz umożliwia nadanie modelowi preferowanego kształtu, który może wynikać ze wstępnych oczekiwań lub wiedzy a-priori.

Opracowane algorytmy analizy filmu endoskopowego zostały zaimplementowane w programie komputerowym WCE Player. Wykorzystano w tym celu środowisko programistyczne Visual Studio, język programowania C++ oraz technologię multimedialną DirectShow.

Kolejnym oryginalnym osiągnięciem jest propozycja połączenia w jednym algorytmie transformacji przestrzeni kolorów, obliczania cech tekstury oraz uczenia nadzorowanego i klasyfikacji. Porównanie wygenerowanych tą metodą cech obrazu z deskryptorami standardu MPEG-7 wykazało ich przewagę w przypadku analizy obrazów endoskopowych. Zaproponowany algorytm analizy obrazów kolorowych zawierających teksturę zaimplementowano w programie komputerowym MaZda. Wykorzystano w tym celu środowisko programistyczne Embarcadero RAD Studio® oraz biblioteki OpenGL, OpenCV i qhull. Należy nadmienić, że możliwości programu MaZda nie ograniczają się wyłącznie do analizy obrazów płaskich – umożliwia on także analizę obrazów sekwencji wideo oraz tekstur przestrzennych w obrazach trójwymiarowych, np. w tomografii komputerowej.

W programie MaZda zaimplementowano również metodę selekcji cech wykorzystującą wielotop wypukły. Jej opracowanie jest wynikiem prac związanych z analizą obrazów endoskopowych, jednak jej zastosowanie nie ogranicza się wyłącznie do tego zagadnienia. Metoda ta umożliwia wyselekcjonowanie cech o dużej zdolności dyskryminacyjnej w celu ich późniejszego zastosowania do klasyfikacji danych. Wyniki selekcji uzyskane za jej pomocą okazały się równie dobre jak te, które uzyskano za pomocą maszyny wektorów nośnych i radialnej sieci neuronowej. Jednocześnie okazała się ona od nich bardziej efektywna, co umożliwiło wyselekcjonowanie cech znaczących w znacznie krótszym czasie. Ponadto w przeciwieństwie do maszyny wektorów nośnych i radialnej sieci neuronowej nie wymaga ona standaryzacji zakresów danych i może być stosowana z powodzeniem w sytuacji, w której zbiory wektorów różnych klas są nierównoliczne.

Obydwa programy, zarówno MaZda, jak i WCE Player, zostały udostępnione w Internecie [283–285] i mogą być wykorzystywane bezpłatnie. Opracowane algorytmy i programy komputerowe wzbudziły zainteresowanie i okazały się przydatne w badaniach naukowych prowadzonych w różnych ośrodkach badawczych na świecie. Metoda analizy obrazów endoskopowych wykorzystująca deformowalne pierścienie stała się przedmiotem przeglądu w pracach [76], [77], [92], [286]. Program MaZda natomiast jest wykorzystywany w wielu ośrodkach badawczych na świecie nie tylko do analizy obrazów biomedycznych (wymienionych w podrozdziale 5.1), ale również obrazów żywności [110], [111], [127], [240], [287–290].

Podsumowując, do najważniejszych osiągnięć tej pracy należy zaliczyć opracowanie oryginalnego modelu deformowalnego do analizy ruchu własnego kamery, uniwersalną metodę obliczania naprężeń w modelach deformowalnych, nowy, szybki algorytm selekcji cech i klasyfikacji trudno rozdzielnych skupień, wykorzystujący wielotop wypukły, a także programy komputerowe WCE Player oraz MaZda udostępnione w internecie i cytowane w literaturze. Zastosowanie opracowanych algorytmów umożliwia automatyczną regulację szybkości odtwarzania filmu, ułatwia dostęp do wybranych fragmentów filmu, zwiększa liczbę wykrytych zmian chorobowych oraz wiarygodność i powtarzalność wyników diagnozowania. Należy nadmienić, że ich stosowanie nie wymaga wprowadzania zmian konstrukcyjnych w endoskopie oraz nie wymaga zastosowania dodatkowych sposobów obrazowania.

Bibliografia

- [1] Fantastic Voyage – Wikipedia, the free encyclopedia. [Online]. Available: http://en.wikipedia.org/wiki/Fantastic_Voyage. [Accessed: 11-Mar-2012].
- [2] Fantastyczna podróż / Fantastic Voyage (1966) – Filmweb. [Online]. Available: <http://www.filmweb.pl/film/Fantastyczna+podr%C3%B3%C5%B7-1966-33127>. [Accessed: 11-Mar-2012].
- [3] G. Iddan, G. Meron, A. Glukhovsky and P. Swain, Wireless capsule endoscopy, *Nature*, vol. 405, no. 6785, p. 417, May 2000.
- [4] Invest in Israel – Given Imaging. [Online]. Available: <http://www.investinisrael.gov.il/NR/exeres/01104877-8720-4F39-8441-BD24EC58D8B9.htm>. [Accessed: 01-Oct-2011].
- [5] H. Yamamoto and H. Kita, Enteroscopy, *J Gastroenterol*, vol. 40, no. 6, pp. 555-562, Jul. 2005.
- [6] D. Koppel, Yuan-fang Wang and Hua Lee, Viewing enhancement in video-endoscopy, presented at the Applications of Computer Vision, 2002, pp. 304-308.
- [7] P. Swain, Role of video endoscopy in managing small bowel disease, *Gut*, vol. 53, no. 12, pp. 1866-1875, Dec. 2004.
- [8] M. Classen, G.N. Tytgat and C.J. Lightdale, *Gastroenterological endoscopy*. Thieme, 2002.
- [9] L.B. Gerson, Double-balloon enteroscopy: the new gold standard for small-bowel imaging?, *Gastrointestinal Endoscopy*, vol. 62, no. 1, pp. 71-75, Jul. 2005.
- [10] L.B. Gerson and J. Van Dam, Wireless Capsule Endoscopy and Double-Balloon Enteroscopy for the Diagnosis of Obscure Gastrointestinal Bleeding, *Techniques in Vascular and Interventional Radiology*, vol. 7, no. 3, pp. 130-135, Sep. 2004.
- [11] P.M. Szczypiński, R.D. Sriram, P.V.J. Sriram and D.N. Reddy, A model of deformable rings for interpretation of wireless capsule endoscopic videos, *Medical Image Analysis*, vol. 13, no. 2, pp. 312-324, Apr. 2009.
- [12] P.M. Szczypiński, P.V.J. Sriram, R.D. Sriram and D.N. Reddy, Computerized Image Analysis of Wireless Capsule Endoscopy Videos Using a Dedicated Web-like Model of Deformable Rings – A Feasibility Study, presented at the 12th United European Gastroenterology Week, Prague, 2004, vol. 36 (Suppl. I), p. A76.
- [13] P. Szczypiński, P.V.J. Sriram, R.D. Sriram and D.N. Reddy, Przetwarzanie i wspomaganie interpretacji danych z endoskopu bezprzewodowego za pomocą modelu deformowalnych pierścieni, *Zeszyty Naukowe Elektronika*, vol. 10, pp. 129-147, 2005.
- [14] P.M. Szczypinski, Selecting a motion estimation method for a Model of Deformable Rings, in *ICSES'06 - International Conference on Signals and Electronic Systems, Proceedings*, 2006, pp. 297-300.
- [15] P.M. Szczypiński, Przetwarzanie danych video z endoskopu bezprzewodowego za pomocą modelu deformowalnych pierścieni, presented at the Sympozjum TPO, Serock-Poland, 2006.

- [16] P.M. Szczypiński, P.V.J. Sriram, R.D. Sriram and D.N. Reddy, Model of Deformable Rings for Aiding the Wireless Capsule Endoscopy Video Interpretation and Reporting, in *Computer Vision and Graphics*, vol. 32, K. Wojciechowski, B. Smolka, H. Palus, R.S. Kozera, W. Skarbek and L. Noakes, Eds. Dordrecht: Kluwer Academic Publishers, 2006, pp. 167-172.
- [17] P.M. Szczypiński, Sterowanie odtwarzaniem wideo z endoskopu bezprzewodowego za pomocą współczynnika deformacji modelu deformowalnych pierścieni, presented at the XV Krajowa Konferencja Naukowa, Biocybernetyka i Inżynieria Biomedyczna, Wrocław, 2007.
- [18] P.M. Szczypiński, Computer Program for Aiding in Wireless Endoscopic Video Interpretation, in *Computers in Medical Activity*, vol. 65, E. Kącki, M. Rudnicki and J. Stempczyńska, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009, pp. 53-61.
- [19] D. Terzopoulos, J. Platt, A. Barr and K. Fleischer, Elastically deformable models, *ACM SIGGRAPH Computer Graphics*, vol. 21, pp. 205-214, Aug. 1987.
- [20] T. McInerney and D. Terzopoulos, Deformable models in medical image analysis: a survey, *Medical Image Analysis*, vol. 1, no. 2, pp. 91-108, Jun. 1996.
- [21] M. Kass, A. Witkin and D. Terzopoulos, Snakes: Active contour models, *International journal of computer vision*, vol. 1, no. 4, pp. 321-331, 1988.
- [22] S. Osher and J. Sethian, Fronts propagating with curvature-dependent speed: Algorithms based on Hamilton-Jacobi formulations, *Journal of Computational Physics*, vol. 79, pp. 12-49, Nov. 1988.
- [23] P. Szczypiński and A. Materka, Variable-Flexibility Elastic Model for Digital Image Analysis, in *XXI KKTOiUE*, Kiekrz, 1998.
- [24] P. Szczypiński, Modele deformowalne do ilościowej analizy i rozpoznawania obiektów w obrazach cyfrowych, Rozprawa doktorska, Wydział Elektrotechniki i Elektroniki Politechniki Łódzkiej, Łódź, 2000.
- [25] Wielokomórka – Wikipedia, wolna encyklopedia. [Online]. Available: <http://pl.wikipedia.org/wiki/Wielokomórka>. [Accessed: 11-Mar-2012].
- [26] B. Block, G. Schachschal and H. Schmidt, *Endoscopy of the upper GI tract: a training manual*. Thieme, 2004.
- [27] J.D. Waye, D.K. Rex and C.B. Williams, *Colonoscopy: principles and practice*. Blackwell, 2003.
- [28] A. May, L. Nachbar, A. Wardak, H. Yamamoto and C. Ell, Double-balloon enteroscopy: preliminary experience in patients with obscure gastrointestinal bleeding or chronic abdominal pain, *Endoscopy*, vol. 35, no. 12, pp. 985-991, Dec. 2003.
- [29] K. Honda, T. Mizutani, K. Nakamura, N. Higuchi, K. Kanayama, Y. Sumida, S. Yoshinaga, S. Itaba, H. Akiho, K. Kawabe, Y. Arita and T. Ito, Acute pancreatitis associated with peroral double-balloon enteroscopy: a case report, *World J. Gastroenterol.*, vol. 12, no. 11, pp. 1802-1804, Mar. 2006.
- [30] H.-H. Yen, Y.-Y. Chen, W.-W. Su, M.-S. Soon and Y.-M. Lin, Intestinal necrosis as a complication of epinephrine injection therapy during double-balloon enteroscopy, *Endoscopy*, vol. 38, no. 5, p. 542, May 2006.
- [31] D.G. Adler and C.J. Goustout, Wireless capsule endoscopy, *Hosp Physician*, vol. 39, no. 5, pp. 14-22, May 2003.
- [32] A. Bednarczuk and G. Rydzewska, Endoskopia kapsułkowa – nadal nowość czy już standard w diagnostyce gastroenterologicznej?, *Polski Merkurusz Lekarski*, vol. XXVI, no. 155, pp. 506-511, 2009.

- [33] Given Imaging Ltd., Patients: Abdominal Pain, Crohn's Disease, Gastrointestinal bleeding, Inflammatory Bowel Disease. [Online]. Available: <http://www.givenimaging.com/en-us/Patients/Pages/CapsuleEndoscopy.aspx>. [Accessed: 08-Sep-2011].
- [34] Olympus Corporation, EndoCapsule – Taking capsule endoscopy to the next level. [Online]. Available: http://www.olympus-europa.com/endoscopy/2001_5491.htm. [Accessed: 11-Mar-2012].
- [35] C. Gheorghe, R. Iacob and I. Bancila, Olympus capsule endoscopy for small bowel examination, *Journal of Gastrointestinal and Liver Diseases*, vol. 16, no. 3, p. 309, 2007.
- [36] A. Moglia, A. Menciasci, P. Dario and A. Cuschieri, Capsule endoscopy: progress update and challenges ahead, *Nature Reviews Gastroenterology & Hepatology*, vol. 6, pp. 353-361, May 2009.
- [37] Intromedic, MicroCam Info. [Online]. Available: <http://www.intromedic.com/en/product/productInfo.asp>. [Accessed: 11-Mar-2012].
- [38] C. Li, B. Zhang, C. Chen and Y. Li, OMOM capsule endoscopy in diagnosis of small bowel disease, *Journal of Zhejiang University – Science B*, vol. 9, no. 11, pp. 857-862, Nov. 2008.
- [39] C. Mc Caffrey, O. Chevalerias, C. O'Mathuna and K. Twomey, Swallowable-Capsule Technology, *IEEE Pervasive Computing*, vol. 7, pp. 23-29, Jan. 2008.
- [40] O. Ferrer-Roca, Telepathology and Optical Biopsy, *International Journal of Telemedicine and Applications*, vol. 2009, no. 740712, pp. 1-9, 2009.
- [41] RF Co. Ltd., The Next Generation Capsule Endoscope Sayaka. [Online]. Available: <http://www.rfamerica.com/sayaka/index.html> [Accessed: 11-Mar-2012].
- [42] Vector Consortium, Versatile Endoscopic Capsule for gastrointestinal TumOr Recognition and therapy. [Online]. Available: <http://www.vector-project.com/>. [Accessed: 11-Mar-2012].
- [43] P. Swain, The future of wireless capsule endoscopy, *World Journal of Gastroenterology*, vol. 14, p. 4142, 2008.
- [44] CORDIS, Nano based capsule-Endoscopy with molecular imaging and optical biopsy. [Online]. Available: http://cordis.europa.eu/search/index.cfm?fuseaction=proj.document&PJ_RCN=9776762 [Accessed: 08-Sep-2011].
- [45] Byungkyu Kim, Sukho Park, Chang Yeol Jee and Seok-Jin Yoon, An earthworm-like locomotive mechanism for capsule endoscopes, 2005, pp. 2997-3002.
- [46] Jiwoon Kwon, Sukho Park, Byungkyu Kim and Jong-Oh Park, Bio-Material Property Measurement System for Locomotive Mechanism in Gastro-Intestinal Tract, in *Robotics and Automation, 2005. ICRA 2005. Proceedings of the 2005 IEEE International Conference on*, 2005, pp. 1303-1308.
- [47] Xiaona Wang, M. Q.-H. Meng and Yawen Chan, Physiological Factors of the Small Intestine in Design of Active Capsule Endoscopy, in *Engineering in Medicine and Biology Society, 2005. IEEE-EMBS 2005. 27th Annual International Conference of the*, 2005, pp. 2942-2945.
- [48] F. Carpi, S. Galbiati and A. Carpi, Controlled Navigation of Endoscopic Capsules: Concept and Preliminary Experimental Investigations, *IEEE Transactions on Biomedical Engineering*, vol. 54, pp. 2028-2036, Nov. 2007.
- [49] A. Chiba, M. Sendoh, K. Ishiyama and K.I. Arai, Magnetic actuator for capsule endoscope navigation system, presented at the INTERMAG, 2005, pp. 626-626.

- [50] N. Bourbakis, G. Giakos and A. Karargyris, Design of new-generation robotic capsules for therapeutic and diagnostic endoscopy, in *Imaging Systems and Techniques (IST), 2010 IEEE International Conference on*, 2010, pp. 1-6.
- [51] E. Scapa, H. Jacob, S. Lewkowicz, M. Migdal, D. Gat, A. Gluckhovski, N. Gutmann and Z. Fireman, Initial experience of wireless-capsule endoscopy for evaluating occult gastrointestinal bleeding and suspected small bowel pathology, *Am. J. Gastroenterol.*, vol. 97, no. 11, pp. 2776–2779, Nov. 2002.
- [52] B. Li and M. Q.-H. Meng, Computer-Aided Detection of Bleeding Regions for Capsule Endoscopy Images, *Biomedical Engineering, IEEE Transactions on*, vol. 56, no. 4, pp. 1032-1039, 2009.
- [53] B.S. Lewis and P. Swain, Capsule endoscopy in the evaluation of patients with suspected small intestinal bleeding: Results of a pilot study, *Gastrointestinal Endoscopy*, vol. 56, pp. 349-353, Sep. 2002.
- [54] “CAPVIEW.” [Online]. Available: <http://www.capview.org/>. [Accessed: 08-May-2012].
- [55] J.P.S. Cunha, M. Coimbra, P. Campos and J.M. Soares, Automated Topographic Segmentation and Transit Time Estimation in Endoscopic Capsule Exams, *IEEE Transactions on Medical Imaging*, vol. 27, pp. 19-27, 2008.
- [56] J.J. Christensen and D.L. Wingate, *A Guide to gastrointestinal motility*. Bristol John Wright Publisher, 1983.
- [57] Phooi Yee Lau and P.L. Correia, Detection of bleeding patterns in WCE video using multiple features, in *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE*, 2007, pp. 5601-5604.
- [58] B. Giritharan, Xiaohui Yuan, Jianguo Liu, B. Buckles, JungHwan Oh and Shou Jiang Tang, Bleeding detection from capsule endoscopy videos, in *Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE*, 2008, pp. 4780-4783.
- [59] J. Bonnel, A. Khademi, S. Krishnan and C. Ioana, Small bowel image classification using cross-co-occurrence matrices on wavelet domain, *Biomedical Signal Processing and Control*, vol. 4, no. 1, pp. 7-15, 2009.
- [60] B. Li and M.Q.-H. Meng, Small bowel tumor detection for wireless capsule endoscopy images using textural features and support vector machine, in *Intelligent Robots and Systems, 2009. IROS 2009. IEEE/RSJ International Conference on*, 2009, pp. 498-503.
- [61] B. Li and M. Q.-H. Meng, Computer-based detection of bleeding and ulcer in wireless capsule endoscopy images by chromaticity moments, *Computers in Biology and Medicine*, vol. 39, no. 2, pp. 141-147, 2009.
- [62] A. Karargyris and N. Bourbakis, A methodology for detecting blood-based abnormalities in Wireless Capsule Endoscopy videos, in *BioInformatics and BioEngineering, 2008. BIBE 2008. 8th IEEE International Conference on*, 2008, pp. 1-6.
- [63] N. Bourbakis, Detecting Abnormal Patterns in WCE Images, in *Fifth IEEE Symposium on Bioinformatics and Bioengineering (BIBE'05)*, Minneapolis, MN, USA, pp. 232-238.
- [64] M.W. Mackiewicz, M. Fisher and C. Jamieson, Bleeding detection in wireless capsule endoscopy using adaptive colour histogram model and support vector classification, in *Proceedings of SPIE*, San Diego, CA, USA, 2008, p. 69140R-69140R-12.

- [65] P. Szczypinski and A. Klepaczko, Selecting texture discriminative descriptors of capsule endoscopy images, in *Image and Signal Processing and Analysis, 2009. ISPA 2009. Proceedings of 6th International Symposium on*, 2009, pp. 701-706.
- [66] P. Szczypiński and A. Klepaczko, Convex Hull-Based Feature Selection in Application to Classification of Wireless Capsule Endoscopic Images, in *Advanced Concepts for Intelligent Vision Systems*, vol. 5807, J. Blanc-Talon, W. Philips, D. Popescu and P. Scheunders, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009, pp. 664-675.
- [67] A. Klepaczko and P.M. Szczypinski, Algorytm lokalnej aproksymacji wielomianowej w zastosowaniu do segmentacji obrazów endoskopowych, in *XVI Krajowa Konferencja Biocybernetyka i Inżynieria Biomedyczna*, Warszawa, 2010.
- [68] A. Klepaczko and P. Szczypiński, Automated Segmentation of Endoscopic Images Based on Local Shape-Adaptive Filtering and Color Descriptors, in *Advanced Concepts for Intelligent Vision Systems*, vol. 6474, J. Blanc-Talon, D. Bone, W. Philips, D. Popescu and P. Scheunders, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, pp. 245-254.
- [69] A. Klepaczko, P. Szczypiński, P. Daniel and M. Pazurek, Local Polynomial Approximation for Unsupervised Segmentation of Endoscopic Images, in *Computer Vision and Graphics*, vol. 6375, L. Bolc, R. Tadeusiewicz, L.J. Chmielewski and K. Wojciechowski, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, pp. 33-40.
- [70] P. Szczypiński and A. Klepaczko, Automated Recognition of Abnormal Structures in WCE Images Based on Texture Most Discriminative Descriptors, in *Image Processing and Communications Challenges 2*, vol. 84, R.S. Choraś, Ed. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, pp. 263-270.
- [71] P.M. Szczypiński and A. Klepaczko, Wykrywanie zmian patologicznych w filmach z endoskopu bezprzewodowego, presented at the VI Sympozjum Naukowe, Techniki Przetwarzania Obrazu, Serock, 2010.
- [72] P. Szczypiński, A. Klepaczko and M. Strzelecki, An Intelligent Automated Recognition System of Abnormal Structures in WCE Images, in *Hybrid Artificial Intelligent Systems*, vol. 6678, E. Corchado, M. Kurzyński and M. Woźniak, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, pp. 140-147.
- [73] M. Coimbra, P. Campos and J.P.S. Cunha, Topographic Segmentation and Transit Time Estimation for Endoscopic Capsule Exams, presented at the IEEE ICASSP, Toulouse-France, 2006, vol. 2, pp. 1164-1167.
- [74] M. Coimbra, J. Kustra, P. Campos and J.P.S. Cunha, Combining color with spatial and temporal position of the endoscopic capsule for improved topographic classification and segmentation, *Proc. of SAMT*, pp. 17-18, 2006.
- [75] M.T. Coimbra and J.P.S. Cunha, MPEG-7 Visual Descriptors – Contributions for Automated Feature Extraction in Capsule Endoscopy, *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 16, no. 5, pp. 628-637, 2006.
- [76] A. Karargyris and N. Bourbakis, Wireless Capsule Endoscopy and Endoscopic Imaging: A Survey on Various Methodologies Presented, *IEEE Engineering in Medicine and Biology Magazine*, vol. 29, pp. 72-83, Jan. 2010.
- [77] M. Mackiewicz, J. Berens and M. Fisher, Wireless Capsule Endoscopy Color Video Segmentation, *Medical Imaging, IEEE Transactions on*, vol. 27, no. 12, pp. 1769-1781, 2008.
- [78] M.W. Mackiewicz, J. Berens and M. Fisher, Wireless capsule endoscopy video segmentation using support vector classifiers and hidden markov models, in *Proc.*

- International Conference Medical Image Understanding and Analyses*, 2006, pp. 81-85.
- [79] M. Mackiewicz, J. Berens, M. Fisher and D. Bell, Colour and Texture Based Gastrointestinal Tissue Discrimination, in *Acoustics, Speech and Signal Processing, 2006. ICASSP 2006 Proceedings. 2006 IEEE International Conference on*, 2006, vol. 2, p. II.
- [80] J. Berens, M.W. Mackiewicz and G.D. Bell, Stomach, intestine, and colon tissue discriminators for wireless capsule endoscopy images, in *Proceedings of SPIE*, San Diego, CA, USA, 2005, pp. 283-290.
- [81] P. Spyridonos, F. Vilarino, J. Vitria, F. Azpiroz and P. Radeva, Anisotropic Feature Extraction from Endoluminal Images for Detection of Intestinal Contractions, in *Medical Image Computing and Computer-Assisted Intervention – MICCAI*, vol. 4191, Springer Berlin / Heidelberg, 2006, pp. 161-168.
- [82] F. Vilarino, Automatic Detection of Intestinal Juices in Wireless Capsule Video Endoscopy, in *Pattern Recognition, International Conference on*, 2006, vol. 4, pp. 719-722.
- [83] F. Vilarino, L.I. Kuncheva and P. Radeva, ROC curves and video analysis optimization in intestinal capsule endoscopy, *Pattern Recognition Letters*, vol. 27, no. 8, pp. 875-881, Jun. 2006.
- [84] A. Karargyris and N. Bourbakis, A video-frame based registration using segmentation and graph connectivity for Wireless Capsule Endoscopy, in *Life Science Systems and Applications Workshop, 2009. LiSSA 2009. IEEE/NIH*, 2009, pp. 74-79.
- [85] M. Coimbra, P. Campos and J.P.S. Cunha, Extracting clinical information from endoscopic capsule exams using MPEG-7 visual descriptors, *IEE Seminar Digests*, vol. 2005, no. 11099, pp. 105-110, 2005.
- [86] W. Skarbek, MPEG-7, *IX Konferencja PLOUG, Kościelisko, X*, 2003.
- [87] T. Sikora, The MPEG-7 visual standard for content description-an overview, *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 11, no. 6, pp. 696-702, 2001.
- [88] H. Eidenberger, How good are the visual MPEG-7 features, in *SPIE & IEEE Visual Communications and Image Processing Conference*, Lugano, Switzerland, 2003, pp. 476-488.
- [89] K. Mori, D. Deguchi, J. Hasegawa, Y. Suenaga, J. Toriwaki, H. Takabatake and H. Natori, A Method for Tracking the Camera Motion of Real Endoscope by Epipolar Geometry Analysis and Virtual Endoscopy System, in *Medical Image Computing and Computer-Assisted Intervention*, vol. 2208, W. Niessen and M. Viergever, Eds. Springer Berlin / Heidelberg, 2001, pp. 1-8.
- [90] K. Mori, D. Deguchi, J. Sugiyama, Y. Suenaga, J. Toriwaki, C.R. Maurer, H. Takabatake and H. Natori, Tracking of a bronchoscope using epipolar geometry analysis and intensity-based image registration of real and virtual endoscopic images, *Medical Image Analysis*, vol. 6, no. 3, pp. 321-336, Sep. 2002.
- [91] J. Nagao, K. Mori, T. Enjouji, D. Deguchi, T. Kitasaka, Y. Suenaga, J.-I. Hasegawa, J.-I. Toriwaki, H. Takabatake and H. Natori, Fast and accurate bronchoscope tracking using image registration and motion prediction, in *Lecture Notes in Computer Science*, 2004, vol. 3217, pp. 551-558.
- [92] A. Karargyris, O. Karargyris and N. Bourbakis, 3D Representation of the Digestive Tract Surface in Wireless Capsule Endoscopy Videos, in *BioInformatics and BioEngineering (BIBE), 2010 IEEE International Conference on*, 2010, pp. 279-280.

- [93] A. Skalski, Segmentacja 3D danych medycznych pochodzących z tomografii komputerowej oraz endoskopowych zapisów wideo, Rozprawa Doktorska, Akademia Górniczo-Hutnicza, Wydział Elektrotechniki, Automatyki, Informatyki i Elektroniki, Kraków, 2009.
- [94] Choroba Leśniowskiego-Crohna – Wikipedia, wolna encyklopedia. [Online]. Available: http://pl.wikipedia.org/wiki/Choroba_Leśniowskiego-Crohna. [Accessed: 11-Mar-2012].
- [95] W. Latos, A. Gadowska-Cicha, K.J. Niepsuj and A. Sieroń, Choroba Leśniowskiego-Crohna w górnym odcinku przewodu pokarmowego, *Wiad. Lek.*, vol. 58, pp. 222-226, 2005.
- [96] P. Szczypiński and A. Materka, Object Tracking and Recognition Using Deformable Grid with Geometrical Templates, in *International Conference on Signals and Electronic Systems*, Ustroń-Poland, 2000, pp. 169-174.
- [97] P. Szczypiński and A. Materka, Program do analizy obrazów za pomocą deformowalnych modeli, *Zeszyty Naukowe Elektronika*, vol. 5, pp. 33-51, 2000.
- [98] P. Szczypiński and A. Materka, Variable-Flexibility Elastic Model for Digital Image Analysis, *Bulletin of the Polish Academy of Sciences*, vol. 47, no. 3, pp. 263-269, 1999.
- [99] J. Jaksa, K. Ślot and P. Szczypiński, Face Recognition Using Deformable Models, in *International Conference on Signals and Electronic Systems*, Łódź-Poland, 2001, pp. 111-116.
- [100] P. Strumiłło and P. Szczypiński, Automatic Extraction of Fuzzy and Broken Image Edges using Active Contour Model, in *XVIII KKTOiUE*, 1995.
- [101] P. Strumiłło and P. Szczypiński, Application of an Active Contour Model for Extraction of Fuzzy and Broken Image Edges, *Machine Graphics & Vision*, vol. 5, no. 4, pp. 576-594, 1996.
- [102] P. Szczypiński, Center-Point Model of Deformable Surface, in *Computer Vision and Graphics*, vol. 32, K. Wojciechowski, B. Smolka, H. Palus, R.S. Kozera, W. Skarbek and L. Noakes, Eds. Dordrecht: Kluwer Academic Publishers, 2006, pp. 343-348.
- [103] P.M. Szczypinski, M. Strzelecki, A. Materka and A. Klepaczko, MaZda – A software package for image texture analysis, *Computer Methods and Programs in Biomedicine*, vol. 94, no. 1, pp. 6676, 2009.
- [104] P.M. Szczypiński, M. Strzelecki, A. Materka and A. Klepaczko, MaZda – The Software Package for Textural Analysis of Biomedical Images, in *Computers in Medical Activity*, vol. 65, E. Kącki, M. Rudnicki and J. Stempczyńska, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009, pp. 73-84.
- [105] P.M. Szczypinski, M. Strzelecki and A. Materka, Mazda – a software for texture analysis, in *2007 International Symposium on Information Technology Convergence (ISITC 2007)*, Jeonju, Korea, 2007, pp. 245-249.
- [106] M. Strzelecki, A. Materka and P.M. Szczypinski, MaZda, in *Texture analysis for magnetic resonance imaging*, M. Hájek, Ed. Prague: Med4publishing, 2006.
- [107] M. Kociołek, P. Szczypiński, A. Materka and M. Strzelecki, Discrete Wavelet Transform-Derived Features for Digital Image Texture Analysis, in *International Conference on Signals and Electronic Systems*, Łódź-Poland, 2001, pp. 99-104.
- [108] M. Kociołek, A. Materka, M. Strzelecki and P. Szczypiński, Investigation of Wordlength Effect on Discriminative Power of Co-occurrence Matrix – Derived Features for Digital Image Texture Analysis, in *International Conference on Signals and Electronic Systems*, Ustroń-Poland, 2000, pp. 163-168.

- [109] M. Kociólek, A. Materka, M. Strzelecki and P. Szczypiński, Badanie wpływu liczby poziomów jasności obrazu na zdolność dyskryminacji tekstur przy użyciu macierzy zdarzeń, *Zeszyty Naukowe Elektronika*, vol. 5, pp. 21-33, 2000.
- [110] A.K. Thybo, P.M. Szczypinski, A.H. Karlsson, S. Donstrup, H.S. Stodkilde-Jorgensen and H.J. Andersen, Prediction of sensory texture quality attributes of cooked potatoes by NMR-imaging (MRI) of raw potatoes in combination with different image analysis methods, *Journal of Food Engineering*, vol. 61, no. 1, pp. 91-100, Jan. 2004.
- [111] A.K. Thybo, A.H. Karlsson, H.C. Bertram, H.J. Andersen, P.M. Szczypinski and Donstrup, Nuclear magnetic resonance (NMR) and magnetic resonance imaging (MRI) in texture measurement, in *Solid foods*, D. Kilcast, Ed. Woodhead Publishing, 2004, pp. 184-204.
- [112] P. Szczypiński, M. Kociólek, A. Materka and M. Strzelecki, Computer Program for Image Texture Analysis in PhD Students Laboratory, presented at the International Conference on Signals and Electronic Systems, Łódź-Poland, 2001, pp. 255-262.
- [113] A. Amin and D. Mahmoud-Ghoneim, Texture analysis of liver fibrosis microscopic images: a study on the effect of biomarkers, *Acta Biochimica et Biophysica Sinica*, vol. 43, pp. 193-203, Jan. 2011.
- [114] S.A. Waugh, R.A. Lerski, L. Bidaut and A.M. Thompson, The influence of field strength and different clinical breast MRI protocols on the outcome of texture analysis using foam phantoms, *Medical Physics*, vol. 38, p. 5058, 2011.
- [115] C. Vijayakumar and D. Gharpure, Development of image-processing software for automatic segmentation of brain tumors in MR images, *Journal of Medical Physics*, vol. 36, p. 147, 2011.
- [116] D. Mahmoud-Ghoneim, Optimizing automated characterization of liver fibrosis histological images by investigating color spaces at different resolutions, *Theor Biol Med Model*, vol. 8, p. 25, 2011.
- [117] M.S. de Oliveira, M.L.F. Balthazar, A D'Abreu, C.L. Yasuda, B.P. Damasceno, F. Cendes and G. Castellano, MR Imaging Texture Analysis of the Corpus Callosum and Thalamus in Amnesic Mild Cognitive Impairment and Mild Alzheimer Disease, *American Journal of Neuroradiology*, vol. 32, pp. 60-66, 2011.
- [118] K.K. Holli, M. Waljas, L. Harrison, S. Liimatainen, T. Luukkaala, P. Ryymin, H. Eskola, S. Soimakallio, J. Ohman and P. Dastidar, Mild Traumatic Brain Injury: Tissue Texture Analysis Correlated to Neuropsychological and DTI Findings, *Academic Radiology*, vol. 17, no. 9, pp. 1096-1102, Sep. 2010.
- [119] M.E. Mayerhoefer, W. Schima, S. Trattnig, K. Pinker, V. Berger-Kulemann and A. Ba-Ssalamah, Texture-based classification of focal liver lesions on MRI at 3.0 Tesla: A feasibility study in cysts and hemangiomas, *Journal of Magnetic Resonance Imaging*, vol. 32, pp. 352-359, Jul. 2010.
- [120] L.C.V. Harrison, M. Raunio, K.K. Holli, T. Luukkaala, S. Savio, I. Elovaara, S. Soimakallio, H.J. Eskola and P. Dastidar, MRI Texture Analysis in Multiple Sclerosis: Toward a Clinical Analysis Protocol, *Academic Radiology*, vol. 17, no. 6, pp. 696-707, 2010.
- [121] A. Kassner and R.E. Thornhill, Texture Analysis: A Review of Neurologic MR Imaging Applications, *American Journal of Neuroradiology*, vol. 31, pp. 809-816, Apr. 2010.

- [122] K. Holli, A.-L. Laaperi, L. Harrison, T. Luukkaala, T. Toivonen, P. Ryymin, P. Dastidar, S. Soimakallio and H. Eskola, Characterization of Breast Cancer Types by Texture Analysis of Magnetic Resonance Images, *Academic Radiology*, vol. 17, no. 2, pp. 135-141, 2010.
- [123] M. Kozakiewicz, A. Marciniak-Hoffman and M. Olszycki, Comparative Analysis of Three Bone Substitute Materials Based on Co-Occurrence Matrix, *Dent. Med. Probl.*, vol. 47, no. 1, pp. 23-29, 2010.
- [124] A. Amin and D. Mahmoud-Ghoneim, Zizyphus spina-christi protects against carbon tetrachloride-induced liver fibrosis in rats, *Food and Chemical Toxicology*, vol. 47, no. 8, pp. 2111-2119, Aug. 2009.
- [125] M.E. Mayerhoefer, P. Szomolanyi, D. Jirak, A. Berg, A. Materka, A. Dirisamer and S. Trattnig, Effects of Magnetic Resonance Image Interpolation on the Results of Texture-Based Pattern Classification, *Investigative Radiology*, vol. 44, pp. 405-411, Jul. 2009.
- [126] M. Abella, J.M. Zubeldia, L. Conejero, N. Malpica, J.J. Vaquero and M. Desco, Automatic quantification of histological studies in allergic asthma, *Cytometry Part A*, vol. 75A, pp. 271-277, Mar. 2009.
- [127] S. Kumar and G.S. Mittal, Textural Characteristics of Five Microorganisms for Rapid Detection Using Image Processing, *Journal of Food Process Engineering*, vol. 32, pp. 126-143, Feb. 2009.
- [128] M. Kozakiewicz, A. Marciniak-Hoffman and M. Denkowski, Long Term Comparison of Application of Two Beta-Tricalcium Phosphates in Oral Surgery, *Dent. Med. Probl.*, vol. 46, no. 4, pp. 384-388, 2009.
- [129] A. Kassner, F. Liu, R. E. Thornhill, G. Tomlinson and D. J. Mikulis, Prediction of hemorrhagic transformation in acute ischemic stroke using texture analysis of postcontrast T1-weighted MR images, *J. Magn. Reson. Imaging*, vol. 30, no. 5, pp. 933-941, 2009.
- [130] L.C.V. Harrison, T. Luukkaala, H. Pertovaara, T.O. Saarinen, T.T. Heinonen, R. Järvenpää, S. Soimakallio, P.-L.I. Kellokumpu-Lehtinen, H.J. Eskola and P. Dastidar, Non-Hodgkin lymphoma response evaluation with MRI texture classification, *J. Exp. Clin. Cancer Res.*, vol. 28, p. 87, 2009.
- [131] J. Zhang, L. Tong, L. Wang, and N.Li, Texture analysis of multiple sclerosis: a comparative study, *Magnetic Resonance Imaging*, vol. 26, no. 8, pp. 1160-1166, 2008.
- [132] P. Guggenbuhl, D. Chappard, M. Garreau, J.-Y. Bansard, G. Chales and Y. Rolland, Reproducibility of CT-based bone texture parameters of cancellous calf bone samples: Influence of slice thickness, *European Journal of Radiology*, vol. 67, no. 3, pp. 514-520, 2008.
- [133] L. Harrison, P. Dastidar, H. Eskola, R. Jarvenpaa, H. Pertovaara, T. Luukkaala, P.-L. Kellokumpu-Lehtinen and S. Soimakallio, Texture analysis on MRI images of non-Hodgkin lymphoma, *Computers in Biology and Medicine*, vol. 38, no. 4, pp. 519-524, Apr. 2008.
- [134] O. Yu, N. Parizel, L. Pain, B. Guignard, B. Eclancher, Y. Mauss and D. Grucker, Texture analysis of brain MRI evidences the amygdala activation by nociceptive stimuli under deep anesthesia in the propofol-formalin rat model, *Magnetic Resonance Imaging*, vol. 25, no. 1, pp. 144-146, Jan. 2007.
- [135] M. Kozakiewicz, M. Stefanczyk, A. Materka and P. Arkuszewski, Selected characteristic features of radiotexture of bone of the dental alveolus in the human maxilla and mandible, *Magazyn Stomatologiczny*, vol. 3, no. 181, pp. 40-43, 2007.

- [136] E. Lerouxel, H. Libouban, M.-F. Moreau, M.F. Basle, M. Audran and D. Chappard, Mandibular bone loss in an animal model of male osteoporosis (orchidectomized rat): a radiographic and densitometric study, *Osteoporosis International*, vol. 15, no. 10, pp. 814-819, Oct. 2004.
- [137] S. Herlidou, R. Grebe, F. Grados, N. Leuyer, P. Fardellone and M.-E. Meyer, Influence of age and osteoporosis on calcaneus trabecular bone structure: a preliminary in vivo MRI study by quantitative texture analysis, *Magnetic Resonance Imaging*, vol. 22, no. 2, pp. 237-243, Feb. 2004.
- [138] L. Bonilha, E. Kobayashi, G. Castellano, G. Coelho, E. Tinois, F. Cendes and L. M. Li, Texture Analysis of Hippocampal Sclerosis, *Epilepsia*, vol. 44, no. 12, pp. 1546-1550, 2003.
- [139] D. Jirák, M. Dezortová, P. Taimr and M. Hájek, Texture analysis of human liver, *J. Magn. Reson. Imaging*, vol. 15, no. 1, pp. 68-74, 2002.
- [140] J. Weng, T.S. Huang and N. Ahuja, 3-D Motion Estimation, Understanding and Prediction from Noisy Image Sequences, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-9, pp. 370-389, May 1987.
- [141] W. Burger and B. Bhanu, Estimating 3D egomotion from perspective image sequence, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, pp. 1040-1058, Nov. 1990.
- [142] T.Y. Tian, C. Tomasi and D.J. Heeger, Comparison of approaches to egomotion computation, presented at the Conference on Computer Vision and Pattern Recognition, 1996, pp. 315-320.
- [143] K. Daniilidis, Fixation Simplifies 3D Motion Estimation, *Computer Vision and Image Understanding*, vol. 68, no. 2, pp. 158-169, Nov. 1997.
- [144] T. Twardowski, T. Zielinski, K. Duda, M. Socha and M. Duplaga, Fast Estimation of Broncho-Fiberscope Egomotion for Ct-Guided Transbronchial Biopsy, in *Image Processing, 2006 IEEE International Conference on*, 2006, pp. 1189-1192.
- [145] P.M. Szczypinski, R.D. Sriram, P.V.J. Sriram and D.N. Reddy, A model of deformable rings for interpretation of wireless capsule endoscopic videos, *Medical Image Analysis*, vol. 13, no. 2, pp. 312-324, Apr. 2009.
- [146] A. Verri and T. Poggio, Motion field and optical flow: qualitative properties, *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 11, no. 5, pp. 490-498, 1989.
- [147] B.K.P. Horn and B.G. Schunck, Determining optical flow, *Artificial Intelligence*, vol. 17, no. 1-3, pp. 185-203, 1981.
- [148] B. Glocker, N. Komodakis, G. Tziritas, N. Navab and N. Paragios, Dense image registration through MRFs and efficient linear programming, *Medical Image Analysis*, vol. 12, no. 6, pp. 731-741, Dec. 2008.
- [149] S.S. Beauchemin and J.L. Barron, The computation of optical flow, *ACM Comput. Surv.*, vol. 27, no. 3, pp. 433-466, 1995.
- [150] Motion Analysis and Object Tracking – opencv v2.1 documentation. [Online]. Available: [http://opencv.willowgarage.com/documentation/cpp/video_motion_analysis_and_object_tracking.html?highlight=optical flow](http://opencv.willowgarage.com/documentation/cpp/video_motion_analysis_and_object_tracking.html?highlight=optical%20flow). [Accessed: 02-Mar-2012].
- [151] B. Zitova and J. Flusser, Image registration methods: a survey, *Image and Vision Computing*, vol. 21, no. 11, pp. 977-1000, 2003.
- [152] J.B.A. Maintz and M.A. Viergever, A survey of medical image registration, *Medical Image Analysis*, vol. 2, no. 1, pp. 1-36, Mar. 1998.

- [153] M. Holden, A Review of Geometric Transformations for Nonrigid Body Registration, *IEEE Transactions on Medical Imaging*, vol. 27, pp. 111-128, Jan. 2008.
- [154] B.D. Lucas and T. Kanade, An Iterative Image Registration Technique with an Application to Stereo Vision, in *Proc. Int'l Joint Conf. Artificial Intelligence*, 1981, pp. 674-679.
- [155] W.M. Neuenschwander, P. Fua, L. Iverson, G. Székely and O. Kübler, Ziplock Snakes, *International Journal of Computer Vision*, vol. 25, no. 3, pp. 191-201, Dec. 1997.
- [156] L.D. Cohen and I. Cohen, Finite-element methods for active contour models and balloons for 2-D and 3-D images, *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 15, no. 11, pp. 1131-1147, 1993.
- [157] C. Chesnaud, P. Refregier and V. Boulet, Statistical region snake-based segmentation adapted to different physical noise models, *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 21, no. 11, pp. 1145-1157, Nov. 1999.
- [158] S.R. Gunn and M.S. Nixon, A robust snake implementation; a dual active contour, *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 19, no. 1, pp. 63-68, Jan. 1997.
- [159] F. Leymarie and M.D. Levine, Simulating the grassfire transform using an active contour model, *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 14, no. 1, pp. 56-75, Jan. 1992.
- [160] S. Lobregt and M.A. Viergever, A discrete dynamic contour model, *IEEE Transactions on Medical Imaging*, vol. 14, pp. 12-24, Mar. 1995.
- [161] S. Menet, P. Saint-Marc and G. Medioni, Active contour models: overview, implementation and applications, in *IEEE International Conference on Systems, Man, and Cybernetics Conference Proceedings*, Los Angeles, CA, USA, 1990, pp. 194-199.
- [162] T. McInerney and D. Terzopoulos, Deformable models in medical image analysis: a survey, *Medical Image Analysis*, vol. 1, no. 2, pp. 91-108, 1996.
- [163] L.H. Staib and J.S. Duncan, Parametrically deformable contour models, presented at the Computer Vision and Pattern Recognition, 1989, pp. 98-103.
- [164] L.H. Staib and J.S. Duncan, Boundary finding with parametrically deformable models, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, pp. 1061-1075, Nov. 1992.
- [165] W. Neuenschwander, P. Fua, G. Székely, O. Kuebler, A. Gruen and O. Kuebler, From ziplock snakes to velcro surfaces, in *Automatic extraction of man-made objects from aerial and space images*, Birkhauser Verlag Basel, 1995.
- [166] Y.F. Wang and J.-F. Wang, Surface reconstruction using deformable models with interior and boundary constraints, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, pp. 572-579, May 1992.
- [167] W. Neuenschwander, P. Fua, G. Székely and O. Kübler, Deformable Velcro surfaces, presented at the Computer Vision, 1995, pp. 828-833.
- [168] P. Makowski, T.S. Sørensen, S.V. Therkildsen, A. Materka, H. Stødkilde-Jørgensen and E.M. Pedersen, Two-phase active contour method for semiautomatic segmentation of the heart and blood vessels from MRI images for 3D visualization, *Computerized Medical Imaging and Graphics*, vol. 26, no. 1, pp. 9-17, 2002.
- [169] L.D. Cohen and I. Cohen, Finite-element methods for active contour models and balloons for 2-D and 3-D images, *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 15, no. 11, pp. 1131-1147, Nov. 1993.

- [170] L.D. Cohen, E. Bardinet and N. Ayache, Surface reconstruction using active contour models, Jan. 1992.
- [171] L.D. Cohen, On active contour models and balloons, *CVGIP: Image understanding*, vol. 53, no. 2, pp. 211-218, 1991.
- [172] A.A. Amini, P. Shi, R.T. Constable, K. Johnson, J.S. Duncan and J.C. Gore, Energy-minimizing deformable grids for tracking tagged MR cardiac images, presented at the Computers in Cardiology, 1992, pp. 651-654.
- [173] Cheng-Yuan Liou and Quan-Ming Chang, Meshed snakes, presented at the Neural Networks, 1996, vol. 3, pp. 1516-1521.
- [174] H. Delingette, Adaptive and deformable models based on simplex meshes, in *Motion of Non-Rigid and Articulated Objects, 1994., Proceedings of the 1994 IEEE Workshop on*, 1994, pp. 152-157.
- [175] G.D. Hager and P.N. Belhumeur, Efficient region tracking with parametric models of geometry and illumination, *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 20, no. 10, pp. 1025-1039, Oct. 1998.
- [176] A.K. Jain, Y. Zhong and M.-P. Dubuisson-Jolly, Deformable template models: A review, *Signal Processing*, vol. 71, pp. 109-129, Dec. 1998.
- [177] P. van Beek, A. M. Tekalp, N. Zhuang, I. Celasun, and Minghui Xia, Hierarchical 2-D mesh representation, tracking, and compression for object-based video, *IEEE Trans. Circuits Syst. Video Technol.*, vol. 9, no. 2, pp. 353-369, Mar. 1999.
- [178] Yao Wang and O. Lee, Use of two-dimensional deformable mesh structures for video coding .I. The synthesis problem: mesh-based function approximation and mapping, *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 6, no. 6, pp. 636-646, 1996.
- [179] Yao Wang, O. Lee and A. Vetro, Use of two-dimensional deformable mesh structures for video coding. II. The analysis problem and a region-based coder employing an active mesh representation, *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 6, no. 6, pp. 647-659, 1996.
- [180] Yao Wang and Ouseb Lee, Active mesh-a feature seeking and tracking image sequence representation scheme, *IEEE Transactions on Image Processing*, vol. 3, pp. 610-624, Sep. 1994.
- [181] J.A. Sethian, Fast marching methods, *SIAM review*, pp. 199-235, 1999.
- [182] J. Sethian, *Level set methods and fast marching methods: evolving interfaces in computational geometry, fluid mechanics, computer vision, and materials science*, 2nd ed., reprint. Cambridge Univ. Press, 2002.
- [183] S. Osher, *Geometric level set methods in imaging, vision, and graphics*. New York: Springer, 2003.
- [184] R. Malladi, *Geometric methods in bio-medical image processing*. Berlin – New York: Springer, 2002.
- [185] Xiao Han, Chenyang Xu and J.L. Prince, Topology preserving level set method for geometric deformable models, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, pp. 755-768, Jun. 2003.
- [186] J. Chen and A.A. Amini, Quantifying 3-D Vascular Structures in MRA Images Using Hybrid PDE and Geometric Deformable Models, *IEEE Transactions on Medical Imaging*, vol. 23, pp. 1251-1262, Oct. 2004.
- [187] Chenyang Xu, A. Yezzi and J.L. Prince, On the relationship between parametric and geometric active contours, presented at the Conference on Signals, Systems and Computers, 2000, vol. 1, pp. 483-489.

- [188] Haiyan Wang and Bijoy Ghosh, Geometric active deformable models in shape modeling, *IEEE Transactions on Image Processing*, vol. 9, pp. 302-308, Feb. 2000.
- [189] C.P. Lee, *Robust Image Segmentation using Active Contours: Level Set Approaches*. North Carolina State University, Department of Electrical and Computer Engineering, 2005.
- [190] R. Malladi, J.A. Sethian and B.C. Vemuri, Shape modeling with front propagation: a level set approach, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 17, pp. 158-175, Feb. 1995.
- [191] W.J. Niessen, B.M.T.H. Romeny and M.A. Viergever, Geodesic deformable models for medical image analysis, *IEEE Transactions on Medical Imaging*, vol. 17, pp. 634-641, Aug. 1998.
- [192] T.F. Chan and L.A. Vese, Active contours without edges, *Image Processing, IEEE Transactions on*, vol. 10, no. 2, pp. 266-277, 2001.
- [193] M.E. Alexander and R.L. Somorjai, The registration of MR images using multiscale robust methods, *Magnetic Resonance Imaging*, vol. 14, no. 5, pp. 453-468, 1996.
- [194] J. Feldmar and N. Ayache, Rigid, affine and locally affine registration of free-form surfaces, *International Journal of Computer Vision*, vol. 18, no. 2, pp. 99-119, May 1996.
- [195] W.R. Crum, T. Hartkens and D. Hill, Non-rigid image registration: theory and practice, *British journal of radiology*, vol. 77, no. Special Issue 2, p. S140, 2004.
- [196] R. Baggs and D. Tamir, Non-Rigid Image Registration, in *Proceedings of the Twenty-First International FLAIRS Conference*, 2008.
- [197] J.P.W. Pluim, J.B.A. Maintz and M.A. Viergever, Mutual-information-based registration of medical images: A survey, *IEEE Transactions on Medical Imaging*, vol. 22, no. 8, pp. 986-1004, 2003.
- [198] P. A. van den Elsen, E.-J. D. Pol and M.A. Viergever, "Medical image matching-a review with classification," *IEEE Engineering in Medicine and Biology Magazine*, vol. 12, pp. 26-39, Mar. 1993.
- [199] K. Rohr, Elastic registration of multimodal medical images: A survey, *Kunstliche Intelligenz*, vol. 14, no. 3, pp. 11-17, 2000.
- [200] H. Chui and A. Rangarajan, A new point matching algorithm for non-rigid registration, *Computer Vision and Image Understanding*, vol. 89, no. 2-3, pp. 114-141, 2003.
- [201] J. Santamaria, O. Cordon and S. Damas, A comparative study of state-of-the-art evolutionary image registration methods for 3D modeling, *Computer Vision and Image Understanding*, vol. 115, no. 9, pp. 1340-1354, 2011.
- [202] J. Shi and C. Tomasi, Good features to track, presented at the Computer Vision and Pattern Recognition, Seattle, WA , USA, 1994.
- [203] Pohsiang Hsu, K.J.R. Liu and T. Chen, A low bit-rate video codec based on two-dimensional mesh motion compensation with adaptive interpolation, *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 11, no. 1, pp. 111-117, 2001.
- [204] W. Badawy and M. Bayoumi, A scalable affine core for video object motion compensation, in *Computer Architectures for Machine Perception, 2000. Proceedings. Fifth IEEE International Workshop on*, 2000, pp. 142-146.
- [205] J.Y. Tham, S. Ranganath, M. Ranganath and A.A. Kassim, A novel unrestricted center-biased diamond search algorithm for block motion estimation, *Circuits*

- and Systems for Video Technology, IEEE Transactions on*, vol. 8, no. 4, pp. 369-377, 1998.
- [206] J. Jain and A. Jain, Displacement measurement and its application in interframe image coding, *Communications, IEEE Transactions on*, vol. 29, no. 12, pp. 1799-1808, 1981.
- [207] A.M. Tourapis, Enhanced predictive zonal search for single and multiple frame motion estimation, in *Proceedings of Visual Communications and Image Processing*, 2002, vol. 1998, pp. 1069-1079.
- [208] A.M. Tourapis, O.C. Au, M.L. Liou and others, Predictive motion vector field adaptive search technique (PMVFAST)-enhancing block based motion estimation, in *Proc. SPIE Conf. Visual Communication and Image Processing*, 2001, pp. 883-892.
- [209] J. Watkinson, *The MPEG handbook: MPEG-1, MPEG-2, MPEG-4*. Elsevier/Focal Press, 2004.
- [210] R.O. Duda, P.E. Hart and D.G. Stork, *Pattern classification. 2nd ed.* Chichester: Wiley-Interscience, 2001.
- [211] R.C. Gonzalez and R.E. Woods, *Digital image processing*. Prentice-Hall, Inc., Upper Saddle River, NJ, 2006.
- [212] Cross-correlation – Wikipedia, the free encyclopedia. [Online]. Available: http://en.wikipedia.org/wiki/Cross-correlation#Normalized_cross-correlation. [Accessed: 02-Mar-2012].
- [213] Mutual information – Wikipedia, the free encyclopedia. [Online]. Available: http://en.wikipedia.org/wiki/Mutual_information. [Accessed: 02-Mar-2012].
- [214] Mean squared error – Wikipedia, the free encyclopedia. [Online]. Available: http://en.wikipedia.org/wiki/Mean_squared_error. [Accessed: 02-Mar-2012].
- [215] Mean difference – Wikipedia, the free encyclopedia. [Online]. Available: http://en.wikipedia.org/wiki/Mean_absolute_difference. [Accessed: 02-Mar-2012].
- [216] W. Neuenschwander, P. Fua, G. Szekely and O. Kuebler, Velcro Surfaces: Fast Initialization of Deformable Models, *Computer Vision and Image Understanding*, vol. 65, pp. 237-245, Feb. 1997.
- [217] I.N. Bankman, *Handbook of medical imaging: processing and analysis*. San Diego CA: Academic Press, 2000.
- [218] C. Xu, J.L. Prince and A.Y. Jr, A Summary of Geometric Level-Set Analogues for a General Class of Parametric Active Contour and Surface Models, in *Proceedings of the IEEE Workshop on Variational and Level Set Methods (VLSM'01)*, 2001, p. 104.
- [219] T. Pavlidis, *Grafika i przetwarzanie obrazów*. Warszawa: WNT, 1987.
- [220] M. Pesce, *Programming Microsoft DirectShow for digital video and television*. Redmond Wash.: Microsoft Press, 2003.
- [221] DirectShow – Microsoft Developer Network, *Microsoft Developer Network*. [Online]. Available: [http://msdn.microsoft.com/en-us/library/ms783323\(VS.85\).aspx](http://msdn.microsoft.com/en-us/library/ms783323(VS.85).aspx). [Accessed: 02-Mar-2012].
- [222] P. Szczypiński, P. Pełczyński, D. Szajerman and P. Strumiłło, Implementation of Computer Vision Algorithms in DirectShow Technology, in *Image Processing and Communications Challenges 2*, vol. 84, R.S. Choraś, Ed. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, pp. 31-38.
- [223] POV-Ray – The Persistence of Vision Raytracer. [Online]. Available: <http://www.povray.org/>. [Accessed: 03-Oct-2011].

- [224] K. Fukunaga, *Introduction to statistical pattern recognition*. Academic Press, 1990.
- [225] Linear discriminant analysis – Wikipedia, the free encyclopedia. [Online]. Available: http://en.wikipedia.org/wiki/Linear_discriminant_analysis. [Accessed: 02-Mar-2012].
- [226] R.A. Fisher, The Use of Multiple Measurements in Taxonomic Problems,”*Annals of Human Genetics*, vol. 7, pp. 179-188, Sep. 1936.
- [227] J. Cohen, *Applied multiple regression/correlation analysis for the behavioral sciences*. L. Erlbaum Associates, 2003.
- [228] D. Hand, *Principles of data mining*. Cambridge Mass.: MIT Press, 2001.
- [229] Linear regression – Wikipedia, the free encyclopedia. [Online]. Available: http://en.wikipedia.org/wiki/Linear_regression. [Accessed: 02-Mar-2012].
- [230] A. Materka and M. Strzelecki, Texture analysis methods-a review, Technical University of Lodz, Institute of Electronics, COST B11, 1998.
- [231] A.K. Jain, *Fundamentals of digital image processing*. Prentice-Hall, Inc., 1989.
- [232] M. Tuceryan, Texture Analysis, in *Handbook of pattern recognition & computer vision*, C. Chen, Ed. Singapore – River Edge NJ: World Scientific, 1993.
- [233] A.C. Bovik, M. Clark and W.S. Geisler, Multichannel texture analysis using localized spatial filters, *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 12, no. 1, pp. 55-73, 1990.
- [234] R.M. Haralick, Statistical and structural approaches to texture, *Proceedings of the IEEE*, vol. 67, no. 5, pp. 786-804, 1979.
- [235] R.M. Haralick, K. Shanmugam and I. Dinstein, Textural Features for Image Classification, *Systems, Man and Cybernetics, IEEE Transactions on*, vol. 3, no. 6, pp. 610-621, 1973.
- [236] L. Cohen, Time-frequency distributions-a review, *Proceedings of the IEEE*, vol. 77, no. 7, pp. 941-981, 1989.
- [237] J.G. Daugman, Complete discrete 2-D Gabor transforms by neural networks for image analysis and compression, *Acoustics, Speech and Signal Processing, IEEE Transactions on*, vol. 36, no. 7, pp. 1169-1179, 1988.
- [238] D. Mahmoud-Ghoneim, A. Amin and P. Corr, MRI-based texture analysis: a potential technique to assess protectors against induced-liver fibrosis in rats, *Radiology and Oncology*, vol. 43, pp. 30-40, Mar. 2009.
- [239] M. Kociński, A. Klepaczko, A. Materka, M. Chekenya and A. Lundervold, 3D image texture analysis of simulated and real-world vascular trees, *Computer Methods and Programs in Biomedicine*, vol. In Press, Corrected Proof, 2011.
- [240] P. Zapotoczny, M. Zielińska and Z. Nita, Application of image analysis for the varietal classification of barley: Morphological features, *Journal of Cereal Science*, vol. 48, no. 1, pp. 104-110, Jul. 2008.
- [241] M. Abella, J. M. Zubeldia, L. Conejero, N. Malpica, J. J. Vaquero and M. Desco, Automatic quantification of histological studies in allergic asthma, *Cytometry Part A*, vol. 75, no. 3, pp. 271-277, 2009.
- [242] K. Holli, A.-L. Lääperi, L. Harrison, S. Soimakallio, P. Dastidar and H.J. Eskola, Detection of characteristic texture parameters in breast MRI, in *4th European Conference of the International Federation for Medical and Biological Engineering*, vol. 22, J. Sloten, P. Verdonck, M. Nyssen and J. Haueisen, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009, pp. 517-521.
- [243] G. Chen, S. Jespersen, M. Pedersen, Q. Pang, M.R. Horsman and H.S. Jorgensen, Evaluation of anti-vascular therapy with texture analysis, *Anticancer research*, vol. 25, no. 5, p. 3399, 2005.

- [244] L.C.V. Harrison, P. Dastidar, K.K. Holli, S. Savio, A. Autere, A. Oinonen, V. Pylkki, S. Soimakallio and H. Eskola, Manual Segmentation of Brain Tissue and Multiple Sclerosis Lesions for Texture Analysis, in *World Congress on Medical Physics and Biomedical Engineering, September 7-12, 2009, Munich, Germany*, vol. 25/2, O. Dössel and W. C. Schlegel, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009, pp. 300-303.
- [245] V. Punys, J. Puniene, R. Jurkevicius, and J. Punys, Myocardium tissue analysis based on textures in ultrasound images, *Studies in Health Technology and Informatics*, vol. 116, pp. 435-440, 2005.
- [246] K.K. Holli, L. Harrison, P. Dastidar, M. Wäljas, J. Öhman, S. Soimakallio and H. Eskola, Texture Analysis of Corpus Callosum in Mild Traumatic Brain Injury Patients, in *World Congress on Medical Physics and Biomedical Engineering, September 7-12, 2009, Munich, Germany*, vol. 25/4, O. Dössel and W.C. Schlegel, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009, pp. 37-40.
- [247] D. Mahmoud-Ghoneim, Y. Cherel, L. Lemaire, J.D. de Certaines and A. Maniere, Texture analysis of magnetic resonance images of rat muscles during atrophy and regeneration, *Magnetic Resonance Imaging*, vol. 24, no. 2, pp. 167-171, Feb. 2006.
- [248] R. Zhou and Y. Li, Texture analysis of MR image for predicting the firmness of Huanghua pears (*Pyrus pyrifolia* Nakai, cv. Huanghua) during storage using an artificial neural network, *Magnetic Resonance Imaging*, vol. 25, no. 5, pp. 727-732, 2007.
- [249] J.J. Szymanski, J.T. Jamison and D.J. DeGracia, Texture analysis of poly-adenylated mRNA staining following global brain ischemia and reperfusion, *Computer Methods and Programs in Biomedicine*, vol. In Press, Corrected Proof, 2001.
- [250] A. Materka, P. Cichy and J. Tuliszkiwicz, Texture analysis of x-ray images for detection of changes in bone mass and structure, *Texture Analysis in Machine Vision, Series in Machine Perception & Machine Intelligence*, vol. 40, pp. 189-195, 2000.
- [251] J.M. Martínez, MPEG-7 Overview. ISO/IEC, Oct-2004.
- [252] P. Owczarek, T. Rosiński, Model eksperymentalny opisu treści wizyjnych, in *Instytut Elektroniki i Telekomunikacji, Politechnika Poznańska*, Poznań, 2003.
- [253] Y. Zhong and A.K. Jain, Object localization using color, texture and shape, *Pattern Recognition*, vol. 33, no. 4, pp. 671-684, Apr. 2000.
- [254] Co-occurrence matrix – Wikipedia, the free encyclopedia. [Online]. Available: http://en.wikipedia.org/wiki/Co-occurrence_matrix. [Accessed: 05-Oct-2011].
- [255] S.G. Mallat, Multifrequency channel decompositions of images and wavelet models, *Acoustics, Speech and Signal Processing, IEEE Transactions on DOI – 10.1109/29.45554*, vol. 37, no. 12, pp. 2091-2110, 1989.
- [256] S.G. Mallat, *Review of multifrequency channel decompositions of images and wavelet models*. New York University, Courant Institute of Mathematical Sciences, Computer Science Division, 1988.
- [257] B. Li and M.Q.-H. Meng, Ulcer recognition in capsule endoscopy images by texture features, in *Intelligent Control and Automation, 2008. WCICA 2008. 7th World Congress on*, 2008, pp. 234-239.
- [258] S.J. Sangwine and R.E. Horne, *The colour image processing handbook*. Chapman & Hall, 1998.
- [259] R.W. Hunt, *The reproduction of colour*. John Wiley & Sons, 2004.

- [260] A. Materka and P. Strumiłło, *Wstęp do komputerowej analizy obrazów – Skrypt*. Łódź: Politechnika Łódzka, 2009.
- [261] List of color spaces and their uses – Wikipedia, the free encyclopedia. [Online]. Available: http://en.wikipedia.org/wiki/List_of_color_spaces_and_their_uses. [Accessed: 02-Mar-2012].
- [262] K. Fukunaga, *Introduction to statistical pattern recognition*. Academic Press, 1990.
- [263] V. Vapnik, *The nature of statistical learning theory*, 2nd ed. New York: Springer, 2000.
- [264] A.K. Jain, R.P.W. Duin and Jianchang Mao, Statistical pattern recognition: a review, *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 22, no. 1, pp. 4-37, 2000.
- [265] C.J.C. Burges, A Tutorial on Support Vector Machines for Pattern Recognition, *Data Mining and Knowledge Discovery*, vol. 2, no. 2, pp. 121-167, Jun. 1998.
- [266] P. Yee, *Regularized radial basis function networks: theory and applications*. New York: John Wiley, 2001.
- [267] J. Moody and C.J. Darken, Fast Learning in Networks of Locally-Tuned Processing Units, *Neural Computation*, vol. 1, pp. 281-294, Jun. 1989.
- [268] Radial basis function network – Wikipedia, the free encyclopedia. [Online]. Available: http://en.wikipedia.org/wiki/Radial_basis_function_network. [Accessed: 05-Oct-2011].
- [269] M. de Berg, *Computational geometry: algorithms and applications.*, 2. ed. Berlin: Springer Verlag, 2000.
- [270] C.B. Barber, D.P. Dobkin and H. Huhdanpaa, The Quickhull algorithm for convex hulls, *ACM Transactions on Mathematical Software*, vol. 22, no. 4, pp. 469-483, 1996.
- [271] D. Avis, How good are convex hull algorithms?, *Computational Geometry*, vol. 7, pp. 265-301, Apr. 1997.
- [272] Convex hull algorithms – Wikipedia, the free encyclopedia. [Online]. Available: http://en.wikipedia.org/wiki/Convex_hull_algorithms. [Accessed: 02-Mar-2012].
- [273] G.I. Nalbantov, P.J.F. Groenen and J.C. Bioch, Nearest Convex Hull Classification, *Econometric Institute, Rotterdam*, EI 2006-50, Jan. 2007.
- [274] G.I. Nalbantov, J.C. Bioch and P.J.F. Groenen, Classification with Support Hyperplanes, in *Machine Learning: ECML 2006*, vol. 4212, J. Fürnkranz, T. Scheffer and M. Spiliopoulou, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006, pp. 703-710.
- [275] X. Zhou and Y. Shi, Nearest Neighbor Convex Hull Classification Method for Face Recognition, in *Computational Science – ICCS 2009*, vol. 5545, G. Allen, J. Nabrzyski, E. Seidel, G. van Albada, J. Dongarra and P. Slood, Eds. Springer Berlin / Heidelberg, 2009, pp. 570-577.
- [276] P. Pudil and P. Somol, Current Feature Selection Techniques in Statistical Pattern Recognition, in *Computer Recognition Systems*, vol. 30, M. Kurzyński, E. Puchała, M. Woźniak, and A. Żołnierczyk, Eds. Springer Berlin / Heidelberg, 2005, pp. 53-68.
- [277] P. Pudil, J. Novovičová and J. Kittler, Floating search methods in feature selection, *Pattern Recognition Letters*, vol. 15, no. 11, pp. 1119-1125, Nov. 1994.
- [278] Qhull code for Convex Hull, Delaunay Triangulation, Voronoi Diagram and Halfspace Intersection about a Point. [Online]. Available: <http://www.qhull.org/>. [Accessed: 05-Oct-2011].

- [279] Weka 3 – Data Mining with Open Source Machine Learning Software in Java. [Online]. Available: <http://www.cs.waikato.ac.nz/ml/weka/> [Accessed: 05-Oct-2011].
- [280] J. Nowak, *Master of Engineering Thesis: Application of MPEG 7 Texture, Shape and Colour Descriptors for Detection of Pathological Changes of Gastrointestinal System in Wireless Endoscopy Images*. Technical University of Łódź: 2011.
- [281] k-means clustering – Wikipedia, the free encyclopedia. [Online]. Available: <http://en.wikipedia.org/wiki/K-means> [Accessed: 06-Oct-2011].
- [282] Cluster analysis – Wikipedia, the free encyclopedia. [Online]. Available: http://en.wikipedia.org/wiki/Cluster_analysis [Accessed: 06-Oct-2011].
- [283] MaZda Web Site. [Online]. Available: <http://www.eletel.p.lodz.pl/programy/mazda/> [Accessed: 02-Mar-2012].
- [284] Wspomaganie diagnostyki medycznej przewodu pokarmowego z zastosowaniem analizy komputerowej sekwencji obrazów uzyskiwanych z endoskopu kapsułkowego. [Online]. Available: <http://212.191.90.111/joomla/> [Accessed: 06-Oct-2011].
- [285] P.M. Szczypiński, Strona WWW – Downloads. [Online]. Available: http://www.eletel.p.lodz.pl/pms/prog_en.html. [Accessed: 02-Mar-2012].
- [286] A. Karargyris, A Novel Synergistic Diagnosis Methodology for identifying Abnormalities in Wireless Capsule Endoscopy videos, Wright State University, 2010.
- [287] A. Khoshroo, A. Keyhani, R.A. Zoroofi, S. Rafiee, Z. Zamani and M.R. Alsharif, Classification of Pomegranate Fruit using Texture Analysis of MR Images, *Agricultural Engineering International: CIGR Journal*, vol. XI, 2009.
- [288] S. Kumar and G.S. Mittal, Rapid Detection of Microorganisms Using Image Processing Parameters and Neural Network, *Food and Bioprocess Technology*, vol. 3, pp. 741-751, Jul. 2008.
- [289] J. Letal, D. Jirak, L. Suderlova and M. Hajek, MRI texture analysis of MR images of apples during ripening and storage, *Lebensmittel-Wissenschaft und-Technologie*, vol. 36, no. 7, pp. 719-727, 2003.
- [290] G. Collewet, M. Strzelecki and F. Mariette, Influence of MRI acquisition protocols and image intensity normalization methods on texture classification, *Magnetic Resonance Imaging*, vol. 22, no. 1, pp. 81-91, 2004.

Załącznik A

Symbole

- $\hat{\mathbf{v}}_s$ – wektor współrzędnych punktu węzłowego w modelu odniesienia, stały,
 $\mathbf{v}_s$ – wektor współrzędnych punktu węzłowego w modelu deformowalnym,
 x, y – współrzędne, elementy wektora $\mathbf{v}$,
 N_s – n -sąsiedztwo punktu węzłowego s ,
 $\mathbf{J}_s$ – macierz transformacji n -sąsiedztwa punktu węzłowego s ,
 $\mathbf{T}_s$ – wektor translacji n -sąsiedztwa punktu węzłowego s ,
 ε_s – błąd średniokwadratowy odległości między rzeczywistym położeniem węzłów a położeniem obliczonym na podstawie transformacji,
 Σ – symbol sumowania po wszystkich punktach węzłowych sąsiedztwa N_s .

Transformacja skalowania, rozciągania, obrotu i przesunięcia

W transformacji afinicznej składającej się ze skalowania, rozciągania i obrotu macierz $\mathbf{J}_s$ oraz wektor $\mathbf{T}_s$ mają następującą postać:

$$\mathbf{J}_s = \begin{bmatrix} j_{11} & j_{12} \\ j_{21} & j_{22} \end{bmatrix}; \quad \mathbf{T}_s = \begin{bmatrix} t_1 \\ t_2 \end{bmatrix}$$

Dla tego przypadku błąd ε_s można zapisać w następującej postaci:

$$\varepsilon_s = \sum \left((x - t_1 + j_{11} \hat{x} + j_{12} \hat{y})^2 + (y - t_2 + j_{21} \hat{x} + j_{22} \hat{y})^2 \right)$$

a po przekształceniu:

$$\varepsilon_s = \sum \left(x^2 - 2xt_1 + t_1^2 + y^2 - 2yt_2 + t_2^2 + (j_{11} \hat{x})^2 + (j_{21} \hat{x})^2 + (j_{12} \hat{y})^2 + (j_{22} \hat{y})^2 + \right. \\ \left. + 2 \sum (j_{11}j_{12} \hat{x} \hat{y} + j_{21}j_{22} \hat{x} \hat{y} - j_{11}(x - t_1)\hat{x} - j_{21}(y - t_2)\hat{x} - j_{12}(x - t_1)\hat{y} - j_{22}(y - t_2)\hat{y}) \right)$$

Pochodne błędu błąd ε_s po zmiennych $t_1, t_2, j_{11}, j_{12}, j_{21}$ oraz j_{22} mają postać:

$$\frac{\partial \varepsilon_s}{\partial j_{11}} = 2 \sum (j_{11} \hat{x}^2 + j_{12} \hat{x} \hat{y} - \hat{x}(x - t_1)) = 0$$

$$\frac{\partial \varepsilon_s}{\partial j_{12}} = 2 \sum (j_{12} \hat{y}^2 + j_{11} \hat{x} \hat{y} - \hat{y}(x - t_1)) = 0$$

$$\begin{aligned}\frac{\partial \varepsilon_s}{\partial j_{21}} &= 2 \sum (j_{21} \hat{x}^2 + j_{22} \hat{x} \hat{y} - \hat{x}(y - t_2)) = 0 \\ \frac{\partial \varepsilon_s}{\partial j_{22}} &= 2 \sum (j_{22} \hat{y}^2 + j_{21} \hat{x} \hat{y} - \hat{y}(y - t_2)) = 0 \\ \frac{\partial \varepsilon_s}{\partial t_1} &= 2 \sum (t_1 - x + j_{11} \hat{x} + j_{12} \hat{y}) = 0 \\ \frac{\partial \varepsilon_s}{\partial t_2} &= 2 \sum (t_2 - y + j_{21} \hat{x} + j_{22} \hat{y}) = 0\end{aligned}$$

Tym samym otrzymujemy układ równań, które można przedstawić w postaci macierzowej:

$$\begin{bmatrix} \sum \hat{x}^2 & \sum \hat{x} \hat{y} & \sum \hat{x} \\ \sum \hat{x} \hat{y} & \sum \hat{y}^2 & \sum \hat{y} \\ \sum \hat{x} & \sum \hat{y} & \sum 1 \end{bmatrix} \begin{bmatrix} j_{11} \\ j_{12} \\ t_1 \end{bmatrix} = \begin{bmatrix} \sum \hat{x} x \\ \sum \hat{y} x \\ \sum x \end{bmatrix}$$

$$\begin{bmatrix} \sum \hat{x}^2 & \sum \hat{x} \hat{y} & \sum \hat{x} \\ \sum \hat{x} \hat{y} & \sum \hat{y}^2 & \sum \hat{y} \\ \sum \hat{x} & \sum \hat{y} & \sum 1 \end{bmatrix} \begin{bmatrix} j_{12} \\ j_{21} \\ t_2 \end{bmatrix} = \begin{bmatrix} \sum \hat{x} y \\ \sum \hat{y} y \\ \sum y \end{bmatrix}$$

Rozwiązaniem tego układu równań są:

$$\begin{bmatrix} j_{11} \\ j_{12} \\ t_1 \end{bmatrix} = \mathbf{M} \begin{bmatrix} \sum \hat{x} x \\ \sum \hat{y} x \\ \sum x \end{bmatrix}; \quad \begin{bmatrix} j_{21} \\ j_{22} \\ t_2 \end{bmatrix} = \mathbf{M} \begin{bmatrix} \sum \hat{x} y \\ \sum \hat{y} y \\ \sum y \end{bmatrix}$$

gdzie:

$$\mathbf{M} = \frac{\begin{bmatrix} \sum \hat{y}^2 \sum 1 - (\sum \hat{y})^2 & \sum \hat{y} \sum \hat{x} - \sum \hat{x} \hat{y} \sum 1 & \sum \hat{x} \hat{y} \sum \hat{y} - \sum \hat{y}^2 \sum \hat{x} \\ \sum \hat{y} \sum \hat{x} - \sum \hat{x} \hat{y} \sum 1 & \sum \hat{x}^2 \sum 1 - (\sum \hat{x})^2 & \sum \hat{x} \hat{y} \sum \hat{x} - \sum \hat{x}^2 \sum \hat{y} \\ \sum \hat{x} \hat{y} \sum \hat{y} - \sum \hat{y}^2 \sum \hat{x} & \sum \hat{x} \hat{y} \sum \hat{x} - \sum \hat{x}^2 \sum \hat{y} & \sum \hat{x}^2 \sum \hat{y}^2 - (\sum \hat{x} \hat{y})^2 \end{bmatrix}}{\sum \hat{x}^2 \sum \hat{y}^2 \sum 1 - (\sum \hat{y})^2 \sum \hat{x}^2 - (\sum \hat{x} \hat{y})^2 \sum 1 - (\sum \hat{x})^2 \sum \hat{y}^2 + 2 \sum \hat{y} \sum \hat{x} \sum \hat{x} \hat{y}}$$

Transformacja skalowania, obrotu i przesunięcia

W transformacji składającej się ze skalowania, obrotu i przesunięcia, w której przestrzeń transformowana nie jest rozciągana kierunkowo, macierz $\mathbf{J}_s$ oraz wektor $\mathbf{T}_s$ mają następującą postać:

$$\mathbf{J}_s = \begin{bmatrix} j_{11} & j_{12} \\ -j_{12} & j_{11} \end{bmatrix}; \quad \mathbf{T}_s = \begin{bmatrix} t_1 \\ t_2 \end{bmatrix}$$

Dla tego przypadku błąd ε_s można zapisać w następującej postaci:

$$\begin{aligned}\varepsilon_s &= \sum (x^2 - 2x t_1 + t_1^2 + y^2 - 2y t_2 + t_2^2 + (j_{11} \hat{x})^2 + (j_{12} \hat{x})^2 + (j_{12} \hat{y})^2 + (j_{11} \hat{y})^2) + \\ &+ 2 \sum (j_{11} j_{12} \hat{x} \hat{y} - j_{12} j_{11} \hat{x} \hat{y} - j_{11} (x - t_1) \hat{x} + j_{12} (y - t_2) \hat{x} - j_{12} (x - t_1) \hat{y} - j_{11} (y - t_2) \hat{y})\end{aligned}$$

Pochodne błęd ε_s po zmiennych t_1 , t_2 , j_{11} oraz j_{12} mają następującą postać:

$$\frac{\partial \varepsilon_s}{\partial j_{11}} = 2 \sum (j_{11}(\hat{x}^2 + \hat{y}^2) - \hat{x}(x - t_1) - \hat{y}(y - t_2)) = 0$$

$$\frac{\partial \varepsilon_s}{\partial j_{12}} = 2 \sum (j_{12}(\hat{x}^2 + \hat{y}^2) - \hat{y}(x - t_1) + \hat{x}(y - t_2)) = 0$$

$$\frac{\partial \varepsilon_s}{\partial t_1} = 2 \sum (t_1 - x + j_{11}\hat{x} + j_{12}\hat{y}) = 0$$

$$\frac{\partial \varepsilon_s}{\partial t_2} = 2 \sum (t_2 - y - j_{12}\hat{x} + j_{11}\hat{y}) = 0$$

Tym samym otrzymujemy układ równań, które można przedstawić w postaci:

$$j_{11} \left(\sum \hat{x}^2 + \sum \hat{y}^2 \right) + t_1 \sum \hat{x} + t_2 \sum \hat{y} - \sum \hat{x}x - \sum \hat{y}y = 0$$

$$j_{12} \left(\sum \hat{x}^2 + \sum \hat{y}^2 \right) + t_1 \sum \hat{y} - t_2 \sum \hat{x} + \sum \hat{x}y - \sum \hat{y}x = 0$$

$$j_{11} \sum \hat{x} + j_{12} \sum \hat{y} + t_1 \sum 1 - \sum x = 0$$

$$j_{11} \sum \hat{y} - j_{12} \sum \hat{x} + t_2 \sum 1 - \sum y = 0$$

lub w postaci macierzowej:

$$\begin{bmatrix} \left(\sum \hat{x}^2 + \sum \hat{y}^2 \right) & 0 & \sum \hat{x} & \sum \hat{y} \\ 0 & \left(\sum \hat{x}^2 + \sum \hat{y}^2 \right) & \sum \hat{y} & -\sum \hat{x} \\ \sum \hat{x} & \sum \hat{y} & \sum 1 & 0 \\ \sum \hat{y} & -\sum \hat{x} & 0 & \sum 1 \end{bmatrix} \begin{bmatrix} j_{11} \\ j_{12} \\ t_1 \\ t_2 \end{bmatrix} = \begin{bmatrix} \sum \hat{x}x + \sum \hat{y}y \\ \sum \hat{y}x - \sum \hat{x}y \\ \sum x \\ \sum y \end{bmatrix}$$

Rozwiązaniem układu równań jest:

$$\begin{bmatrix} j_{11} \\ j_{12} \\ t_1 \\ t_2 \end{bmatrix} = \frac{\begin{bmatrix} \sum 1 & 0 & -\sum \hat{x} & -\sum \hat{y} \\ 0 & \sum 1 & -\sum \hat{y} & \sum \hat{x} \\ -\sum \hat{x} & -\sum \hat{y} & \left(\sum \hat{x}^2 + \sum \hat{y}^2 \right) & 0 \\ -\sum \hat{y} & \sum \hat{x} & 0 & \left(\sum \hat{x}^2 + \sum \hat{y}^2 \right) \end{bmatrix}}{\left(\sum \hat{x}^2 + \sum \hat{y}^2 \right) \sum 1 - \left(\sum \hat{x} \right)^2 - \left(\sum \hat{y} \right)^2} \begin{bmatrix} \sum \hat{x}x + \sum \hat{y}y \\ \sum \hat{y}x - \sum \hat{x}y \\ \sum x \\ \sum y \end{bmatrix}$$

Transformacja obrotu i przesunięcia

W transformacji składającej się z obrotu i przesunięcia, w której przestrzeń transformowana nie jest skalowana oraz rozciągana kierunkowo, macierz $\mathbf{J}_s$ oraz wektor $\mathbf{T}_s$ mają następującą postać:

$$\mathbf{J}_s = \begin{bmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{bmatrix}; \quad \mathbf{T}_s = \begin{bmatrix} t_1 \\ t_2 \end{bmatrix}$$

Jeśli dla przestrzeni dwuwymiarowej $\mathbf{v}_i = [x, y]^T$ oraz $\hat{\mathbf{v}}_i = [\hat{x}, \hat{y}]^T$, to błąd ε_s można zapisać w następującej postaci:

$$\varepsilon_s = \sum \left((x - t_1 + \hat{x} \cos \alpha + \hat{y} \sin \alpha)^2 + (y - t_2 - \hat{x} \sin \alpha + \hat{y} \cos \alpha)^2 \right)$$

Sumowanie wykonywane jest dla wszystkich węzłów i należących do sąsiedztwa węzła s . Po przekształceniu otrzymuje się:

$$\varepsilon_s = \sum \left(x^2 - 2xt_1 + t_1^2 + y^2 - 2yt_2 + t_2^2 + (\hat{x})^2 + (\hat{y})^2 \right) + 2 \sum \left(-(x-t_1)\hat{x} \cos \alpha + (y-t_2)\hat{x} \sin \alpha - (x-t_1)\hat{y} \sin \alpha - (y-t_2)\hat{y} \cos \alpha \right)$$

Pochodne błędu ε_s po zmiennych t_1 , t_2 oraz α mają następującą postać:

$$\frac{\partial \varepsilon_s}{\partial t_1} = 2 \sum (t_1 - x + \hat{x} \cos \alpha + \hat{y} \sin \alpha) = 0$$

$$\frac{\partial \varepsilon_s}{\partial t_2} = 2 \sum (t_2 - y - \hat{x} \sin \alpha + \hat{y} \cos \alpha) = 0$$

$$\frac{\partial \varepsilon_s}{\partial \alpha} = 2 \sum \left((\hat{y}(x-t_1) - \hat{x}(y-t_2)) \cos \alpha + (\hat{x}(x-t_1) + \hat{y}(y-t_2)) \sin \alpha \right) = 0$$

a po przekształceniu:

$$\begin{aligned} & \cos \alpha \sum (\hat{y}x - \hat{x}y) - t_1 \cos \alpha \sum \hat{y} + t_2 \cos \alpha \sum \hat{x} + \\ & + \sin \alpha \sum (\hat{x}x + \hat{y}y) - t_1 \sin \alpha \sum \hat{x} - t_2 \sin \alpha \sum \hat{y} = 0 \\ & \cos \alpha \sum \hat{x} + \sin \alpha \sum \hat{y} + t_1 \sum 1 - \sum x = 0 \\ & \cos \alpha \sum \hat{y} - \sin \alpha \sum \hat{x} + t_2 \sum 1 - \sum y = 0 \end{aligned}$$

Po wyznaczeniu niewiadomych t_1 i t_2 w drugim i trzecim równaniu układu:

$$\begin{aligned} t_1 &= \frac{\sum x - \cos \alpha \sum \hat{x} - \sin \alpha \sum \hat{y}}{\sum 1} \\ t_2 &= \frac{\sum y - \cos \alpha \sum \hat{y} + \sin \alpha \sum \hat{x}}{\sum 1} \end{aligned}$$

oraz po podstawieniu ich do równania pierwszego, otrzymuje się:

$$\begin{aligned} & \sum 1 (\cos \alpha \sum (\hat{y}x - \hat{x}y) + \sin \alpha \sum (\hat{x}x + \hat{y}y)) + \\ & - (\cos \alpha \sum \hat{y} + \sin \alpha \sum \hat{x}) (\sum x - \cos \alpha \sum \hat{x} - \sin \alpha \sum \hat{y}) + \\ & + (\cos \alpha \sum \hat{x} - \sin \alpha \sum \hat{y}) (\sum y + \cos \alpha \sum \hat{y} - \sin \alpha \sum \hat{x}) = 0 \end{aligned}$$

po dalszych przekształceniach:

$$\begin{aligned} & \cos \alpha (\sum (\hat{y}x - \hat{x}y) \sum 1 - \sum \hat{y} \sum x + \sum \hat{x} \sum y) + \\ & + \sin \alpha (\sum (\hat{x}x + \hat{y}y) \sum 1 - \sum \hat{x} \sum x + \sum \hat{y} \sum y) + \\ & + 2 \sum \hat{x} \sum \hat{y} (\sin^2 \alpha + \cos^2 \alpha) = 0 \end{aligned}$$

Korzystając z tożsamości tzw. jedynki trygonometrycznej, otrzymujemy:

$$\sin \alpha = \frac{2 \sum \hat{x} \sum \hat{y} + \cos \alpha (\sum (\hat{y}x - \hat{x}y) \sum 1 - \sum \hat{y} \sum x + \sum \hat{x} \sum y)}{(\sum \hat{x} \sum x - \sum \hat{y} \sum y - \sum (\hat{x}x + \hat{y}y) \sum 1)}$$

a następnie równanie kwadratowe z niewiadomą $\cos \alpha$:

$$1 - \cos^2 \alpha = \left(\frac{2 \sum \hat{x} \sum \hat{y} + \cos \alpha (\sum (\hat{y}x - \hat{x}y) \sum 1 - \sum \hat{y} \sum x + \sum \hat{x} \sum y)}{(\sum \hat{x} \sum x - \sum \hat{y} \sum y - \sum (\hat{x}x + \hat{y}y) \sum 1)} \right)^2$$

Rozwiązując to równanie, a następnie wyznaczając niewiadome $\sin \alpha$, t_1 i t_2 , otrzymuje się dwa rozwiązania:

$$\sin \alpha = -\frac{(RQ - P\sqrt{P^2 + Q^2 - R^2})}{P^2 + Q^2}; \quad \cos \alpha = -\frac{(RP + Q\sqrt{P^2 + Q^2 - R^2})}{P^2 + Q^2}$$

oraz

$$\sin \alpha = -\frac{(RQ + P\sqrt{P^2 + Q^2 - R^2})}{P^2 + Q^2}; \quad \cos \alpha = -\frac{(RP - Q\sqrt{P^2 + Q^2 - R^2})}{P^2 + Q^2}$$

gdzie:

$$\begin{aligned} P &= \sum (\hat{y}x - \hat{x}y) \sum 1 - \sum \hat{y} \sum x + \sum \hat{x} \sum y \\ Q &= \sum (\hat{x}x + \hat{y}y) \sum 1 - \sum \hat{x} \sum x + \sum \hat{y} \sum y \\ R &= 2 \sum \hat{x} \sum \hat{y} \end{aligned}$$

Wyboru rozwiązania dokonuje się po sprawdzeniu, dla którego z rozwiązań wartość błędu ε_s jest mniejsza.

Dekompozycja macierzy transformacji

Przyjmijmy, że obliczona została macierz $\mathbf{J}_s$ dla wariantu transformacji skalowania, rozciągania, obrotu i przesunięcia. Dla uproszczenia zapisu w dalszej części indeks s zostanie pominięty.

Macierz $\mathbf{J}$ można przedstawić jako iloczyn dwóch macierzy $\mathbf{U}$ i $\mathbf{S}$.

$$\mathbf{J} = \mathbf{U} \mathbf{S}$$

Macierz $\mathbf{U}$ jest macierzą ortonormalną (ang. *unitary matrix*), natomiast macierz $\mathbf{S}$ jest macierzą symetryczną. Rozkład macierzy $\mathbf{J}$ na macierze $\mathbf{U}$ i $\mathbf{S}$ nazywa się rozkładem biegunowym (ang. *polar decomposition*).

Własnością macierzy $\mathbf{U}$ jest, że

$$\mathbf{U}^{-1} = \mathbf{U}^T$$

Jeśli rozpatruje się przypadek przestrzeni dwuwymiarowej, wówczas macierz $\mathbf{U}$ można zapisać w postaci:

$$\mathbf{U} = \begin{bmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{bmatrix}$$

gdzie α jest kątem obrotu w transformacji.

Stąd wynika, że:

$$\mathbf{S} = \mathbf{U}^{-1} \mathbf{J} = \mathbf{U}^T \mathbf{J} = \begin{bmatrix} j_{11} \cos \alpha + j_{21} \sin \alpha & j_{12} \cos \alpha + j_{22} \sin \alpha \\ j_{21} \cos \alpha - j_{11} \sin \alpha & j_{22} \cos \alpha - j_{12} \sin \alpha \end{bmatrix}$$

Ponieważ macierz $\mathbf{S}$ jest symetryczna, to:

$$j_{12} \cos \alpha + j_{22} \sin \alpha = j_{21} \cos \alpha - j_{11} \sin \alpha$$

a po przekształceniu:

$$(j_{21} - j_{12}) \cos \alpha = (j_{11} + j_{22}) \sin \alpha$$

Umożliwia to obliczenie kąta α jako nachylenia wektora $\begin{bmatrix} j_{21} - j_{12} & j_{11} + j_{22} \end{bmatrix}^T$:

$$\alpha = \angle \begin{bmatrix} j_{21} - j_{12} \\ j_{11} + j_{22} \end{bmatrix}^T$$

przy czym kąt α podawany jest w zakresie od $-\pi$ do $+\pi$.

Macierz $\mathbf{S}$ jest obliczana za pomocą równania:

$$\mathbf{S} = \mathbf{U}^T \mathbf{J}$$

Macierz symetryczną $\mathbf{S}$ można z kolei przedstawić w postaci:

$$\mathbf{S} = \mathbf{Q}^{-1} \mathbf{D} \mathbf{Q}$$

gdzie macierz $\mathbf{D}$ jest macierzą diagonalną, natomiast macierz $\mathbf{Q}$ jest macierzą ortonormalną. Macierz ortonormalną $\mathbf{Q}$ można przedstawić w postaci:

$$\mathbf{Q} = \begin{bmatrix} \cos \beta & -\sin \beta \\ \sin \beta & \cos \beta \end{bmatrix}$$

Do rozkładu macierzy $\mathbf{S}$ na macierze $\mathbf{Q}$ oraz $\mathbf{D}$ stosuje się rozkład własny macierzy (ang. *eigen decomposition*). Wiersze macierzy $\mathbf{Q}$ złożone są ze znormalizowanych wektorów własnych macierzy $\mathbf{S}$, natomiast wartości własne macierzy $\mathbf{S}$ stanowią elementy diagonali macierzy $\mathbf{D}$. Wartości własne macierzy $\mathbf{S}$ dla przypadku przestrzeni dwuwymiarowej można obliczyć jako:

$$\lambda_{s1} = \text{trace} \mathbf{S} + \sqrt{(\text{trace} \mathbf{S})^2 - 4 \det \mathbf{S}}$$

$$\lambda_{s2} = \text{trace} \mathbf{S} - \sqrt{(\text{trace} \mathbf{S})^2 - 4 \det \mathbf{S}}$$

gdzie $\text{trace} \mathbf{S} = s_{11} + s_{22}$ oznacza ślad macierzy $\mathbf{S}$, a $\det \mathbf{S}$ oznacza jej wyznacznik.

Macierz $\mathbf{D}$ ma postać:

$$\mathbf{D} = \begin{bmatrix} \lambda_{s1} & 0 \\ 0 & \lambda_{s2} \end{bmatrix} = \sqrt{|\det \mathbf{D}|} \begin{bmatrix} \frac{\lambda_{s1}}{\sqrt{|\det \mathbf{D}|}} & 0 \\ 0 & \frac{\lambda_{s2}}{\sqrt{|\det \mathbf{D}|}} \end{bmatrix} = A \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$$

Aby wyznaczyć kąt β oraz macierz $\mathbf{Q}$ można zauważyć, że:

$$\mathbf{Q} \mathbf{S} = \mathbf{D} \mathbf{Q}$$

co można zapisać jako:

$$\begin{bmatrix} \cos \beta & -\sin \beta \\ \sin \beta & \cos \beta \end{bmatrix} \begin{bmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} \cos \beta & -\sin \beta \\ \sin \beta & \cos \beta \end{bmatrix}$$

Oznacza to, między innymi, że:

$$s_{11} \cos \beta - s_{21} \sin \beta = \lambda_1 \cos \beta$$

co umożliwia obliczenie kąta β jako:

$$\beta = \angle [j_{11} - \lambda_1 \quad j_{21}]^T$$

Finalizuje to rozkład macierzy $\mathbf{J}$ do postaci:

$$\mathbf{J} = a \begin{bmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{bmatrix} \begin{bmatrix} \cos \beta & -\sin \beta \\ \sin \beta & \cos \beta \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} \cos \beta & \sin \beta \\ -\sin \beta & \cos \beta \end{bmatrix}$$

Załącznik B

Nazwy cech barwy i tekstury w programie MaZda 5.0

Obliczanie cech w programie MaZda obejmuje trzy etapy. W pierwszym obraz kolorowy jest zamieniany na obraz monochromatyczny. W programie zaimplementowano kilkanaście algorytmów takiej konwersji do wyboru. Drugim etapem jest modyfikacja histogramu obrazu monochromatycznego obejmująca opcjonalną standaryzację lub normalizację zakresów jasności, a także zmiany liczby bitów kwantyzacji poziomów jasności. Trzecim etapem jest obliczenie cech dokonywane w oznaczonych obszarach zainteresowania o arbitralnie wybranych kształtach. W programie zaimplementowano osiem algorytmów obliczania takich cech.

Przyjęto, że nazwy wygenerowanych cech powinny być unikatowe ze względu na zastosowane algorytmy i parametry zastosowanych algorytmów. Nazwa każdej cechy składa się zatem ze znaków alfanumerycznych identyfikujących sposób konwersji na obraz monochromatyczny, sposób modyfikacji zakresów jasności i liczby bitów kwantyzacji, użyty algorytm obliczeniowy oraz nazwę cechy. Przykładowo nazwa SN5CHDifVrnc oznacza cechę uzyskaną po konwersji do obrazu monochromatycznego wskazującego nasycenie koloru (litera S), w którym histogram został poddany normalizacji (N), kwantyzacji za pomocą pięciu (cyfra 5) bitów na piksel, cecha została wygenerowana za pomocą macierzy zdarzeń (C) w kierunku poziomym (H), dla której to macierzy obliczono współczynnik wariancji różnicowej (DifVrnc).

Dalsze części załącznika przedstawiają opis metod stosowanych do obliczania cech w programie MaZda oraz definiują kody wykorzystywane do identyfikacji tych metod w nazwach cech.

Konwersja do obrazu monochromatycznego

Danymi wejściowymi programu MaZda są obrazy monochromatyczne nie wymagające konwersji lub obrazy kolorowe, w których każdy piksel dany jest wektorem trzech składowych kolorów czerwonego (ang. *red* – *R*) zielonego (*green* – *G*) i niebieskiego (*blue* – *B*). Przyjmuje się, że zakres zmienności składowych kolorów oraz zakres zmienności wynikowego obrazu monochromatycznego to 0-255. Obrazy kolorowe są zamieniane na monochromatyczne *I* zgodnie z jednym z poniższych równań. Pierwszy znak nazwy cechy (przedrostek) określa rodzaj przeprowadzonej konwersji.

Brak przedrostka – jasność (ang. *brightness*)

$$I = \frac{299R + 587G + 114B}{1000}$$

Przedrostek R – składowa czerwona

$$I = R$$

Przedrostek G – składowa zielona

$$I = G$$

Przedrostek B – składowa niebieska

$$I = B$$

Przedrostek U – składowa U przestrzeni YUV

$$I = (886B - 587G - 299R + 886 \cdot 255) / 1772$$

Przedrostek V – składowa V przestrzeni YUV

$$I = (-114B - 587G + 701R + 701 \cdot 255) / 1402$$

Przedrostek H – składowa H (ang. *hue*) przestrzeni HSB

$$I = \frac{255}{2\pi} \arg \left(\frac{886B - 587G - 299R}{886} + j \frac{-114B - 587G - 701R}{701} \right)$$

Przedrostek S – składowa nasycenia barwy S (ang. *saturation*) przestrzeni HSB

$$I = 0,937 \sqrt{\left(\frac{886B - 587G - 299R}{886} \right)^2 + \left(\frac{-114B - 587G - 701R}{701} \right)^2}$$

Przedrostek I – składowa I przestrzeni YIQ

$$I = (-3213B - 2744G - 5957R + 5958 \cdot 255) / 11916$$

Przedrostek Q – składowa Q przestrzeni YIQ

$$I = (-3111B - 5226G + 2115R + 5226 \cdot 255) / 10452$$

Przedrostek u – składowa U przestrzeni YUV znormalizowana względem jasności

$$I = 29,07 \left(\left(\frac{(886B - 587G - 299R)}{(299R + 587G + 114B + 1)} \right) + 1 \right)$$

Przedrostek v – składowa V przestrzeni YUV znormalizowana względem jasności

$$I = 76,245 \left(\left(\frac{(-114B - 587G + 701R)}{(299R + 587G + 114B + 1)} \right) + 1 \right)$$

Przedrostek i – składowa I przestrzeni YIQ znormalizowana względem jasności

$$I = 53,006 \left(\left(\frac{(-321,3B - 522,6G - 211,5R)}{(299R + 587G + 114B + 1)} \right) + 2,8185 \right)$$

Przedrostek q – składowa Q przestrzeni YIQ znormalizowana względem jasności

$$I = 70,457 \left(\left(\frac{(311,1B - 522,6G + 211,5R)}{(299R + 587G + 114B + 1)} \right) + 0,8903 \right)$$

Przedrostek h – składowa H przestrzeni HSB obrócona o kąt półpełny

$$I = \frac{255}{2\pi} \arg \left(-\frac{886B - 587G - 299R}{886} - j \frac{-114B - 587G - 701R}{701} \right)$$

Przedrostek X – składowa X przestrzeni CIE XYZ

$$I = 105,163 \bar{R} + 91,188 \bar{G} + 46,0275 \bar{B}$$

gdzie

$$\bar{R} = \begin{cases} \left(\frac{1}{1,055} \left(\frac{R}{255} + 0,055 \right) \right)^{2,4}, & \frac{R}{255} > 0,04045 \\ \left(\frac{R}{255} \frac{1}{12,92} \right), & \frac{R}{255} \leq 0,04045 \end{cases}$$

$$\bar{G} = \begin{cases} \left(\frac{1}{1,055} \left(\frac{G}{255} + 0,055 \right) \right)^{2,4}, & \frac{G}{255} > 0,04045 \\ \left(\frac{G}{255} \frac{1}{12,92} \right), & \frac{G}{255} \leq 0,04045 \end{cases}$$

$$\bar{B} = \begin{cases} \left(\frac{1}{1,055} \left(\frac{B}{255} + 0,055 \right) \right)^{2,4}, & \frac{B}{255} > 0,04045 \\ \left(\frac{B}{255} \frac{1}{12,92} \right), & \frac{B}{255} \leq 0,04045 \end{cases}$$

Przedrostek Y – składowa Y przestrzeni CIE XYZ

$$I = 54,213 \bar{R} + 182,376 \bar{G} + 18,411 \bar{B}$$

Przedrostek – składowa Y przestrzeni CIE XYZ

$$I = 4,5162 \bar{R} + 27,8928 \bar{G} + 222,417 \bar{B}$$

Przedrostek L – składowa L* przestrzeni CIE L*a*b*

$$I = 255 \left(0,2126 \bar{R} + 0,7152 \bar{G} + 0,0722 \bar{B} \right)^{\frac{1}{3}}$$

Przedrostek a – składowa a* przestrzeni CIE L*a*b*

$$I = 376,13 \left(\left(0,4339 \bar{R} + 0,3762 \bar{G} + 0,1899 \bar{B} \right)^{\frac{1}{3}} - \left(0,2126 \bar{R} + 0,7152 \bar{G} + 0,0722 \bar{B} \right)^{\frac{1}{3}} + 0,3390 \right)$$

Przedrostek b – składowa b* przestrzeni CIE L*a*b*

$$I = 159,23 \left(\left(0,2126 \bar{R} + 0,7152 \bar{G} + 0,0722 \bar{B} \right)^{\frac{1}{3}} - \left(0,0177 \bar{R} + 0,1095 \bar{G} + 0,8728 \bar{B} \right)^{\frac{1}{3}} + 0,8007 \right)$$

Modyfikacja histogramu obrazu monochromatycznego

Zakres poziomów jasności w obszarze zainteresowania (ang. *region of interest* – *ROI*) obrazu monochromatycznego może zostać poddany standaryzacji lub normalizacji. W związku z tym dla oryginalnego obrazu monochromatycznego są obliczane cechy histogramu *Mean*, *Variance*, *Perc01* i *Perc99* (definicje podano w dalszej części załącznika). Zakres jasności nowego obszaru zostaje wówczas zmieniony tak, aby odpowiadał zakresowi jasności obrazu oryginalnego w następujący sposób.

Litera N – normalizacja

$$\langle \text{Perc01}, \text{Perc99} \rangle$$

Litera S – standaryzacja

$\langle Mean - 3 Variance, Mean + 3 Variance \rangle$

Litera D – zakres jasności pozostaje niezmieniony

Liczba występująca po literze D, N lub S określa liczbę bitów wykorzystanych do kwantyzacji poziomu jasności piksela.

Algorytmy obliczania cech tekstury

Do obliczania cech tekstury obrazu wykorzystywanych jest sześć algorytmów, analizy statystycznej histogramu jasności, analizy macierzy zdarzeń, analizy macierzy długości ciągów pikseli, statystyki histogramu wartości bezwzględnych gradientu jasności, obliczanie parametrów modelu autoregresji oraz energie transformaty falkowej Haara. Wszystkie cechy obliczane są w arbitralnie zdefiniowanym obszarze zainteresowania ROI (ang. *region of interest*).

Cechy statystyczne histogramu

Podstawą obliczania cech jest znormalizowany histogram $p(k)$ jasności ROI obrazu. Zmienna $k = 1, 2, \dots, N_g$ oznacza poziom $k - 1$ jasności obrazu. Parametr N_g jest wartością maksymalną zakresu jasności obrazu zależną od liczby bitów bpp kwantyzacji $N_g = 2^{bpp} - 1$. Histogram jest obliczany dla oryginalnego obrazu monochromatycznego z niezmienionym zakresem jasności – stąd w nazwach cech statystycznych histogramu nie występuje litera oznaczająca rodzaj zmiany zakresu jasności ani liczba bitów kwantyzacji poziomu jasności.

$$Mean = \sum_{k=0}^{N_g} k p(k)$$

$$Variance = \sum_{k=0}^{N_g} (k - Mean)^2 p(k)$$

$$Skewness = Variance^{-\frac{3}{2}} \left(\sum_{k=0}^N (k - Mean)^3 p(k) \right)$$

$$Kurtosis = Variance^{-2} \left(\sum_{k=0}^N (k - Mean)^4 p(k) \right) - 3$$

$$Perc01 = \min(K): \sum_{k=1}^K p(k) \geq 0,01$$

$$Perc10 = \min(K): \sum_{k=1}^K p(k) \geq 0,10$$

$$Perc50 = \min(K): \sum_{k=1}^K p(k) \geq 0,50$$

$$Perc90 = \min(K): \sum_{k=1}^K p(k) \geq 0,90$$

$$Perc99 = \min(K): \sum_{k=1}^K p(k) \geq 0,99$$

Cechy mapy wartości bezwzględnych gradientu jasności

Cechy statystyczne wartości bezwzględnej gradientu jasności są obliczane na podstawie obrazu mapy gradientu $|G(x,y)|$. Do nazw cech tej grupy po przedrostku oznaczającym rodzaj konwersji jest dodawana litera „G”.

$$Mean = \frac{1}{Area} \sum_{x,y \in ROI} |G(x,y)|$$

$$Variance = \frac{1}{Area} \sum_{x,y \in ROI} (|G(x,y)| - Mean)^2$$

$$Skewness = Variance^{-\frac{3}{2}} \frac{1}{Area} \sum_{x,y \in ROI} (|G(x,y)| - Mean)^3$$

$$Kurtosis = Variance^{-2} \frac{1}{Area} \sum_{x,y \in ROI} (|G(x,y)| - Mean)^4 - 3$$

gdzie

$$|G(x,y)| = \sqrt{(I(x,y+1) - I(x,y-1))^2 + (I(x+1,y) - I(x-1,y))^2}$$

$$Area = \sum_{x,y \in ROI} 1$$

Cechy macierzy zdarzeń

Cechy tej grupy są obliczane na podstawie znormalizowanego histogramu jasności drugiego rzędu nazywanego macierzą zdarzeń (ang. *gray-level co-occurrence matrix* COM, GLCM). Do nazw cech po przedrostku oznaczającym rodzaj konwersji dodawana jest litera „C”. Macierz określa zależności jasności pomiędzy pikselami w parach, przy czym zależność przestrzenna pikseli w parze jest określona. Element $p(k,l)$ macierzy oznacza liczbę par, w której jeden z pikseli ma jasność $k-1$, a drugi $l-1$ podzieloną przez podwojoną liczbę wszystkich par pikseli branego pod uwagę obszaru. W programie MaZda przyjęto, że zależność przestrzenna pomiędzy pikselami w parze może być określona jednym z czterech kierunków („H”, „V”, „Z” lub „N”) oraz jedną z pięciu odległości ($d = 1, 2, 3, 4, 5$). W zależności od wybranego kierunku i odległości po literze „C” w nazwie cechy występuje oznaczenie składające się z odpowiedniej litery i cyfry.

Oznaczenie Hd – kierunek poziomy, w parze jeden piksel przesunięty względem drugiego o wektor $[d, 0]^T$

Oznaczenie Vd – kierunek pionowy, w parze jeden piksel przesunięty względem drugiego o wektor $[0, d]^T$

Oznaczenie Zd – kierunek 45°, w parze jeden piksel przesunięty względem drugiego o wektor $[d, d]^T$

Oznaczenie Nd – kierunek 135°, w parze jeden piksel przesunięty względem drugiego o wektor $[d, -d]^T$

$$AngScMom = \sum_{k=1}^{N_g} \sum_{l=1}^{N_g} p^2(k,l)$$

$$\begin{aligned}
Contrast &= \sum_{m=1}^{2N_g} \left(m^2 p_{\text{dif}}(m) \right) \\
Correlat &= \frac{1}{\rho_k \rho_l} \sum_{k=1}^{N_g} \sum_{l=1}^{N_g} (k l p(k, l) - \mu_k \mu_l) \\
SumOfSqs &= \sum_{k=1}^{N_g} \sum_{l=1}^{N_g} (k - \mu_k)^2 p(k, l) \\
InvDfMom &= \sum_{k=1}^{N_g} \sum_{l=1}^{N_g} p \frac{(k, l)}{1 + (k - l)^2} \\
SumAverg &= \sum_{m=1}^{2N_g} (m p_{\text{sum}}(m)) \\
SumVarnc &= \sum_{m=1}^{2N_g} \left((m - SumAverg)^2 p_{\text{sum}}(m) \right) \\
SumEntrp &= - \sum_{m=1}^{2N_g} p_{\text{sum}}(m) \log(p_{\text{sum}}(m)) \\
Entropy &= - \sum_{k=1}^{N_g} \sum_{l=1}^{N_g} p(k, l) \log(p(k, l)) \\
DifVarnc &= \sum_{m=1}^{N_g-1} (i - \mu_{\text{dif}})^2 p_{\text{dif}}(m) \\
DifEntrp &= - \sum_{m=1}^{N_g} p_{\text{dif}}(m) \log(p_{\text{dif}}(m))
\end{aligned}$$

gdzie

$$\begin{aligned}
\mu_k &= \sum_{k=0}^{N_g} \sum_{l=0}^{N_g} k p(k, l) \\
\mu_l &= \sum_{k=0}^{N_g} \sum_{l=0}^{N_g} l p(k, l) \\
\sigma_k &= \sum_{k=0}^{N_g} \sum_{l=0}^{N_g} (k - \mu_k)^2 p(k, l) \\
\sigma_l &= \sum_{k=0}^{N_g} \sum_{l=0}^{N_g} (l - \mu_l)^2 p(k, l) \\
p_{\text{sum}}(m) &= \sum_{k=1}^{m-1} p(k, m - k) \\
p_{\text{dif}}(m) &= \begin{cases} \sum_{k=1}^{N_g-m} (p(k, m+k) + p(m+k, k)), & m \neq 0 \\ \sum_{k=1}^{N_g-m} (p(k, k)), & m = 0 \end{cases}
\end{aligned}$$

$$\mu_{\text{dif}} = \sum_{m=0}^{2N_g} m p_{\text{dif}}(m)$$

Cechy macierzy długości ciągów pikseli

Cechy tej grupy są obliczane na podstawie macierzy długości ciągów (ang. run-length matrix – RLM). Do nazw cech po przedrostku oznaczającym rodzaj konwersji jest dodawana litera „R”. Element $p(k, l)$ macierzy oznacza liczbę grup pikseli o jasności $k-1$ tworzących w określonym kierunku nieprzerwany ciąg o długości l . Kierunki, w których są mierzone długości ciągów oznaczane są literą („H”, „V”, „Z” lub „N”) występującą po literze „R”, analogicznie jak w przypadku cech obliczanych na podstawie macierzy zdarzeń.

$$A = \sum_{k=1}^{N_g} \sum_{\forall l} p(k, l)$$

$$\text{ShrtREmph} = \frac{1}{A} \sum_{k=1}^{N_g} \sum_{\forall l} \frac{p(k, l)}{l^2}$$

$$\text{LongREmph} = \frac{1}{A} \sum_{k=1}^{N_g} \sum_{\forall l} l^2 p(k, l)$$

$$\text{GLevNonUni} = \frac{1}{A} \sum_{k=1}^{N_g} \left(\sum_{\forall l} p(k, l) \right)^2$$

$$\text{RLNonUni} = \frac{1}{A} \sum_{\forall l} \left(\sum_{k=1}^{N_g} p(k, l) \right)^2$$

$$\text{Fraction} = \frac{\sum_{k=1}^{N_g} \sum_{\forall l} p(k, l)}{\sum_{k=1}^{N_g} \sum_{\forall l} l p(k, l)}$$

Parametry modelu autoregresji

W modelu autoregresji zakłada się, że w przypadku obrazu, w którym piksele zostały zapisane sekwencyjnie, jasność piksela zależy od jasności pikseli poprzedzających. W przypadku obrazu, w którym piksele w linii zapisano w kolejności od lewej do prawej i dla kolejnych linii zależność ta dana jest liniowym równaniem w postaci:

$$I(x, y) = \theta_1 (I(x-1, y) - \mu) + \theta_2 (I(x, y-1) - \mu) + \theta_3 (I(x-1, y-1) - \mu) + \theta_4 (I(x+1, y-1) - \mu) + \mu + e(x, y)$$

gdzie

$$\mu = \frac{\sum_{x, y \in ROI} I(x, y)}{\sum_{x, y \in ROI} 1}$$

$\theta_1, \theta_2, \theta_3$ i θ_4 są parametrami, a $e(x, y)$ jest funkcją błędu. Parametry $\theta_1, \theta_2, \theta_3$ i θ_4 są obliczane w taki sposób, aby uzyskać minimalną możliwą wartość σ – średnią kwadratów błędów e w całym analizowanym obszarze zainteresowania. Parametry modelu są wykorzystywane jako cechy obszaru odpowiednio o nazwach *Teta1*, *Teta2*, *Teta3*, *Teta4* oraz *Sigma*. Do nazw cech po przedrostku oznaczającym rodzaj konwersji jest dodawana litera „A”.

Energie w pasmach częstotliwościowych transformaty falkowej Haara

Cechy tej grupy są obliczane na podstawie obrazów będących wynikiem dyskretniej transformaty falkowej Haara. Obraz jest traktowany jako sygnał dwuwymiarowy a obliczane cechy odzwierciedlają udział energii obrazu w poszczególnych pasmach częstotliwościowych. Obraz zarówno w pionie, jak i w poziomie podlega filtracji rozdzielającej jego składowe górno- (H) i dolnoczęstotliwościowe (L). W wyniku transformacji uzyskuje się cztery obrazy oznaczane jako HH, HL, LH i LL, przy czym pierwsza litera oznacza pasmo będące wynikiem filtracji w poziomie, a druga w pionie. Wynikowy obraz LL jest ponownie poddawany transformacji, która w odróżnieniu od pierwszej jest nazywana transformacją drugiej skali (lub drugiego poziomu). Proces transformowania obrazu jest powtarzany dla kolejnych skal, do piątej włącznie. Dla każdego uzyskanego obrazu jest obliczana tzw. energia obrazu.

$$WavEn = \frac{\sum_{x,y \in ROI} I^2(x, y)}{\sum_{x,y \in ROI} 1}$$

gdzie $I(x, y)$ oznacza jasność piksela wynikowego obrazu transformaty falkowej.

Do nazwy cechy *WavEn* są dodawane litery oznaczające pasmo częstotliwościowe („HH”, „HL”, „LH” lub „LL”) i cyfra oznaczająca skalę, dla której dokonano transformacji.

O autorze



Dr inż. Piotr M. Szczypiński ukończył studia magisterskie w 1995 roku na kierunku Elektronika i Telekomunikacja Wydziału Elektrotechniki, Elektroniki, Informatyki i Automatyki Politechniki Łódzkiej, a w 2001 roku uzyskał stopień naukowy doktora (wyróżniona praca doktorska) na tym samym Wydziale. Od 2001 roku jest pracownikiem Instytutu Elektroniki Politechniki Łódzkiej zatrudnionym na stanowisku adiunkta. Jego specjalnością są komputerowe metody przetwarzania i analizy obrazów biomedycznych wykorzystujące modele defor-

mowalne oraz analizę tekstury. Opracował on innowacyjne metody segmentacji obrazów biomedycznych oraz analizy obrazów endoskopii bezprzewodowej.

Dr Szczypiński jest autorem lub współautorem programów komputerowych, które są używane w wielu ośrodkach badawczych i medycznych w kraju i zagranicą, m.in.: Mammo (analiza zdjęć mammograficznych), MaZda (analiza tekstury w obrazach cyfrowych, segmentacja obrazów trójwymiarowych), More IDE (środowisko programistyczne do tworzenia systemów eksperckich) i WCE Player (wspomaganie diagnozy obrazów wideo z endoskopu bezprzewodowego). Program MaZda zajął pierwsze miejsce w konkursie Top Lista Politechniki Łódzkiej, najważniejsze osiągnięcia pierwszej dekady XXI w kategorii Nauka.

Dr Szczypiński jest autorem lub współautorem ponad 40 artykułów naukowych oraz referatów konferencyjnych, brał udział w ponad dziesięciu projektach badawczych, w tym COST i DFG, finansowanych ze środków budżetowych Unii Europejskiej Polski oraz USA, a także pełnił funkcję kierownika grantu badawczego finansowanego ze środków MNiSW.

W 2001 roku prowadził badania dotyczące metod analizy żywności w Duńskim Instytucie Agrofizyki. W latach 2002-2004 i 2007 gościł jako *Guest Researcher* w Narodowym Instytucie Standardów i Technologii w USA, gdzie zajmował się metodami automatycznego wnioskowania i projektowaniem systemów eksperckich. Tam również rozpoczął prace dotyczące analizy obrazów endoskopowych. Ponadto współpracuje z wieloma ośrodkami naukowymi na świecie, m.in. z Azjatyckim Instytutem Gastroenterologii, Norweskim Uniwersytetem w Bergen, Niemieckim Centrum do Badań nad Rakiem w Heidelbergu

i Uniwersytetem w Jenie, a także w Polsce, m.in. z Uniwersytetem Medycznym w Łodzi, Instytutem Agrofizyki PAN w Lublinie, Towarzystwem Przetwarzania Obrazów oraz Uniwersytetem Warmińsko-Mazurskim.

Dr Piotr Szczypiński pełni również liczne obowiązki dydaktyczne. Jest współautorem oraz wykładowcą przedmiotów prowadzonych w języku angielskim i polskim, m.in.: *Radio frequency circuits*, *Analog Electronics*, *Microprocessor systems*, *Computer Graphics and Image Processing*, Grafika komputerowa, Systemy mikroprocesorowe, Architektura komputerów, Projektowanie zaawansowanych aplikacji mobilnych, Systemy rzeczywistości wirtualnej oraz Systemy alarmowe – cieszących się dużym zainteresowaniem studentów. W latach 2005-2010 był pełnomocnikiem do spraw dydaktyki Zakładu Elektroniki Medycznej. Jest członkiem komisji dydaktycznej dla kierunku Elektronika i Telekomunikacja oraz Inżynieria Biomedyczna Wydziału Elektrotechniki, Elektroniki, Informatyki i Automatyki Politechniki Łódzkiej. Jest recenzentem doniesień konferencyjnych i publikacji zgłaszanych do czasopism (m.in. *IEEE Transactions on Medical Imaging* oraz *Journal of Electromagnetic Waves and Applications*).

Abstract

Wireless capsule endoscopy is one of the medical tests used in diagnosis of gastrointestinal disorders. A result is a video of internal lumen of gastrointestinal tract which interpretation carried out by an expert gastroenterologist requires a lot of attention and is time consuming. The final diagnosis is rarely reproducible – it depends on the knowledge and experience of the diagnostic experience of the expert.

The subject of this monograph is presentation and validation of novel algorithms for wireless endoscope video analysis whose purpose is to improve the reproducibility, reliability and objectivity of medical diagnosis. The algorithms are designed to reduce amount of work devoted to watching the movie, to automate the procedure of the data interpretation and to enable a quantitative description of selected lesions. The proposed methods allow to characterize the endoscope's egomotion (using a dedicated deformable model), reconstruct the intestine's internal surface, detect selected lesions (based on color and texture analysis), segment images in order to identify areas of pathological changes and dynamically adapt playback speed.

The key achievements presented in this monograph include deformable model of rings for analysis of endoscopic camera egomotion passing through the gastrointestinal tract, versatile method for calculating tensions in the deformable model grid and an efficient algorithm using convex polytopes for feature selection and classification of specifically-shaped clusters. All the developed methods were implemented in a computer programs, and thereafter evaluated qualitatively and quantitatively. The computer programs – WCE Player for egomotion estimation and MaZda for image classification based on color and texture analysis – are available from the internet.

