

Algorytmy przetwarzania i analizy obrazów trójwymiarowych w zastosowaniach biomedycznych

Aleksander Grzyb

2015

Politechnika Poznańska



POLITECHNIKA POZNANSKA

Wydział Informatyki

**Algorytmy przetwarzania i analizy obrazów
trójwymiarowych w zastosowaniach
biomedycznych**

Aleksander Grzyb

Promotor dr hab. inż. Krzysztof Krawiec

2015

Aleksander Grzyb

Algorytmy przetwarzania i analizy obrazów trójwymiarowych w zastosowaniach biomedycznych

2015

Promotor: dr hab. inż. Krzysztof Krawiec

Politechnika Poznańska

Wydział Informatyki

Poznań

Spis treści

1 Wstęp	1
1.1 Opis problemu	1
1.2 Cel i zakres pracy	1
1.3 Projekt RIMO	1
2 Obrazowanie OCT	3
2.1 Zasada działania OCT	4
2.1.1 Metoda uzyskania trójwymiarowego obrazu tkanki	5
2.1.2 OCT w domenie częstotliwości	6
2.2 Angiografia OCT	7
2.2.1 Sposób powstania obrazu angiograficznego	7
2.3 Zastosowania OCT	8
2.3.1 OCT w okulistyczne	9
2.3.2 OCT w gastroenterologii i dermatologii	9
2.3.3 OCT w przemyśle	9
3 Algorytmy korejestracji przestrzennej obrazów OCT	11
3.1 Model deformacji kafelków	12
3.1.1 Matematyczny zapis modelu transformacji ciała sztywnego . .	13
3.2 Korejestracja kafelków	13
3.2.1 Korejestracja na podstawie wartości pikseli	15
3.2.2 Korejestracja na podstawie cech	16
3.3 Łączenie kafelków	19
4 Proponowane algorytmy	21
4.1 Oprogramowanie	21
4.1.1 Plik konfiguracyjny	21
4.1.2 OpenCV	22
4.2 Wiedza dziedzinowa	22
4.3 Rejestracja kafelków z wykorzystaniem cech	24
4.3.1 Algorytm SIFT	24
4.3.2 Dopasowanie wyekstrahowanych cech	29
4.3.3 Filtrowanie dopasowań	30

4.4 Rejestracja kafelków poprzez wykrycie położen naczyń krwionośnych w kafelkach	30
4.5 Estymacja macierzy transformacji pomiędzy kafelkami	30
4.6 Globalna rejestracja kafelków	30
4.7 Łączenie kafelków	30
5 Wyniki eksperymentów obliczeniowych	31
6 Podsumowanie i wnioski końcowe	33
Bibliografia	35

Wstęp

1.1 Opis problemu

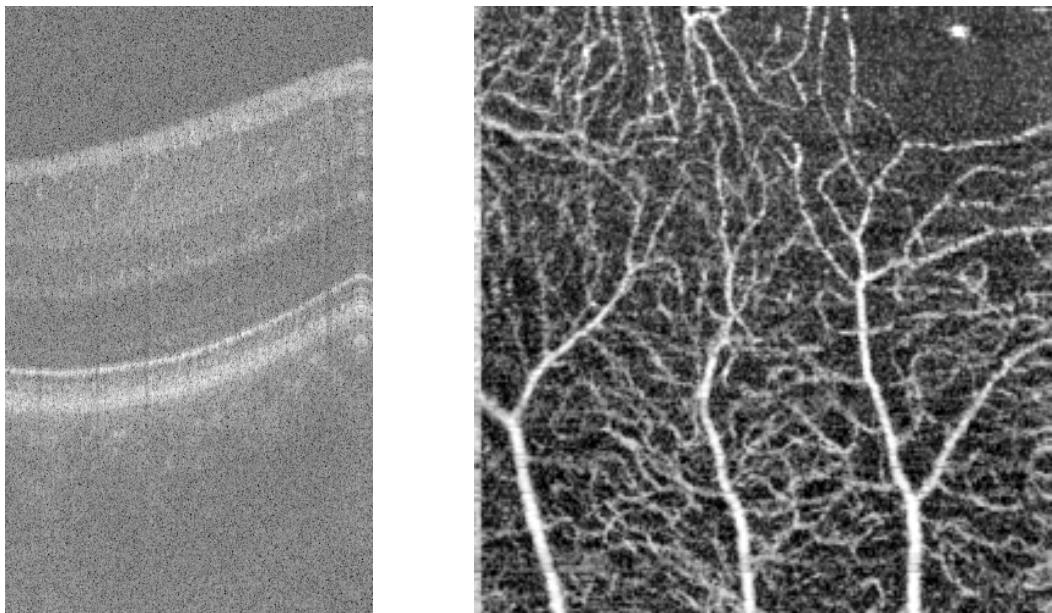
1.2 Cel i zakres pracy

1.3 Projekt RIMO

Obrazowanie OCT

Optyczna tomografia koherencyjna (ang. *optical coherence tomography, OCT*) jest metodą umożliwiającą nieinwazyjne oraz *in vivo* przechwycenie obrazu wnętrza tkanki biologicznej. Zasada działania OCT opiera się na wykorzystaniu fal świetlnych. Dzięki temu rozdzielcość obrazów jest o wiele wyższa niż w ultrasonografii (wykorzystanie fal dźwiękowych), czy rezonansie magnetycznym (wykorzystanie pola magnetycznego). Następnym powodem dużej popularności OCT w medycynie jest bezkontaktowe badanie pacjenta oraz brak wymogu przygotowania pacjenta przed badaniem.

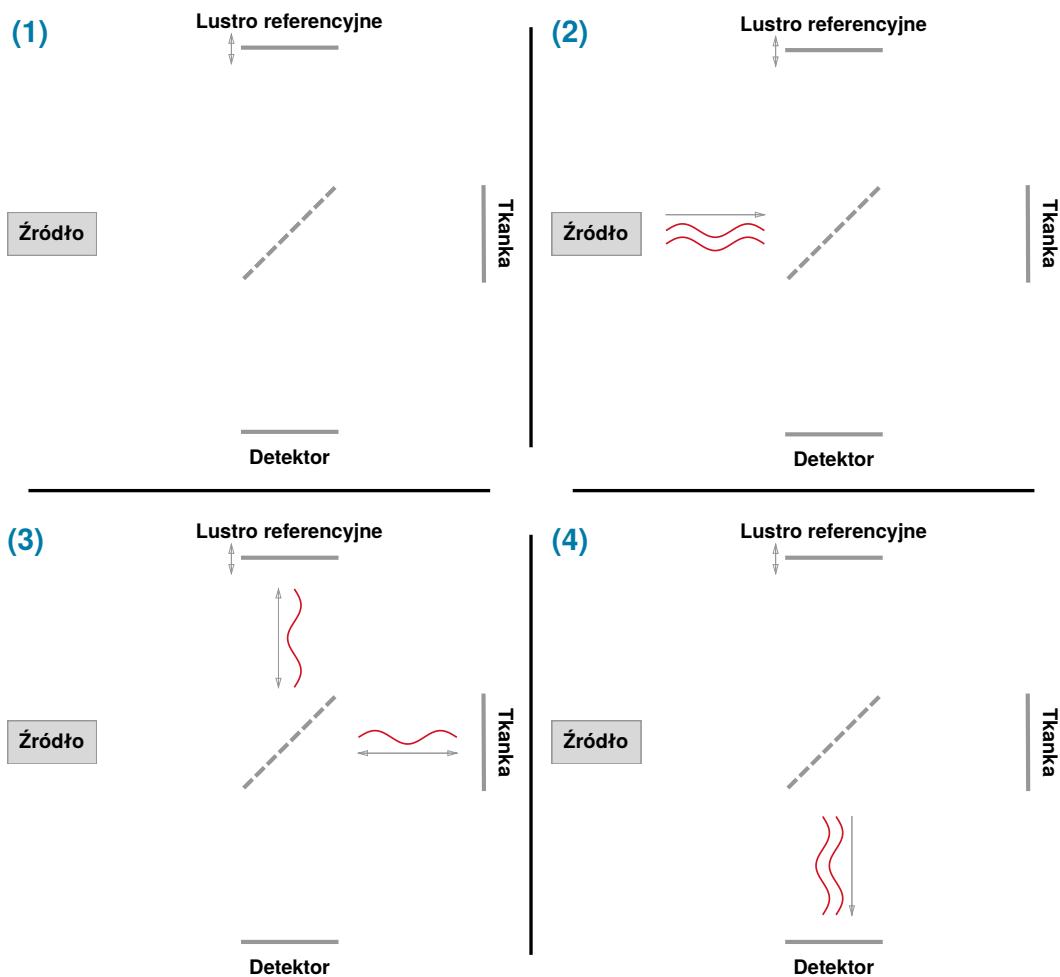
W projekcie RIMO OCT zostało wykorzystane do uzyskania szczegółowych obrazów naczyń krwionośnych siatkówki oka. Rysunek 2.1 przedstawia przykładowe obrazy siatkówki oka uzyskane dzięki wykorzystaniu OCT. Jednym z tych przykładowych obrazów jest angiografia siatkówki oka, która jest obrazem wejściowym do algorytmów omawianych w niniejszej pracy. Sposób powstania angiografii z danych OCT jest wyjaśniony w podrozdziale 2.2, natomiast ogólna zasada działania OCT jest wyjaśniona w podrozdziale 2.1. Na koniec rozdziału w podrozdziale 2.3 zostaną opisane zastosowania OCT.



Rys. 2.1: **Lewy obraz:** Dwuwymiarowy przekrój siatkówki oka (B-skan). Obraz został uzyskany poprzez połączenie jednowymiarowych A-skanów, które zawierają informację o strukturze tkanki w głąb siatkówki oka. **Prawy obraz:** Angiografia siatkówki oka uzyskana dzięki przetworzeniu danych z OCT.

2.1 Zasada działania OCT

Jednym z najważniejszych parametrów metod tomografii w medycynie jest ich rozdzielcość. Optyczna tomografia koherencyjna przechwytuje obrazy wnętrza tkanki poprzez wykorzystanie fal świetlnych. OCT za pomocą generatora tworzy falę światłową, która jest skierowana na tkankę pacjenta. Następnie po odbiciu fali poprzez tkankę wiązka jest przechwycona przez detektor. Jedną z dostępnych metod, która umożliwiłaby zlokalizowanie miejsca odbicia fali byłoby zmierzenie czasu pomiędzy wygenerowaniem fali, a zarejestrowaniem jej przez detektor (mechanizm stosowany np. w ultrasonografii z wykorzystaniem fal dźwiękowych), natomiast prędkość światła ($3 \times 10^8 \frac{m}{s}$) wyklucza zastosowanie tego mechanizmu. Zjawisko, które umożliwia dokładne zlokalizowanie miejsca odbicia to **interferencja światła o niskiej spójności**.

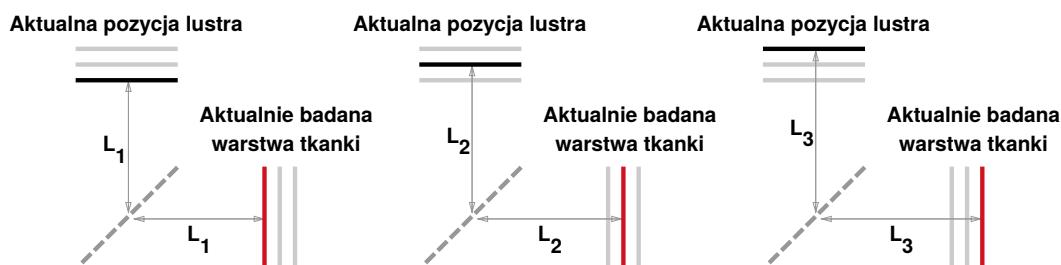


Rys. 2.2: Kolejne etapy działania metody OCT. (1) - Etap początkowy. (2) - Źródło wyemitowało wiązkę światła. (3) - Fala rozdzieliła się za pomocą interferometru na wiązkę referencyjną (skierowaną na lustro referencyjne) oraz na wiązkę próbki (skierowaną na tkankę). (4) - Wiązki po odbiciu od lustra referencyjnego i tkanki ponownie łączą się za pomocą interferometru. W tej części występuje zjawisko interferencji, które jest zarejestrowane przez detektor.

Rysunek 2.2 składa się z bardzo uproszczonych schematów OCT, które obrazują kolejne etapy działania metody. Schematy na rysunku 2.2 składają się z pięciu elementów:

- **Źródła** - Źródło (np. dioda superluminescencyjna) światła podczerwonego, które jest falą o niskiej spójności.
- **Rozdzielacza wiązek** (na rysunku 2.2 przedstawiony za pomocą przerywanych linii na środku każdego schematu) - Interferometr (np. Michelsona) umożliwiający rozdzielenie fali na dwie wiązki oraz następne ich połączenie.
- **Lustro referencyjne** - Lustro, które odbija wiązkę referencyjną. Posiada możliwość oddalania oraz przybliżania się względem interferometru.
- **Tkanka** - Badana tkanka, która odbija wiązkę próbki.
- **Detektor** - Rejestruje zjawisko interferencji związek.

Najbardziej istotnym etapem wymaganym do zrozumienia mechanizmu OCT jest etap (4) pokazany na rysunku 2.2. W tym kroku wiązka referencyjna i wiązka próbki łączą się i zachodzi zjawisko interferencji. Dzięki temu, że wiązki są falami o niskiej spójności interferencja zachodzi tylko na małej długości (ang. *coherence length*). Odczytując za pomocą detektora charakterystyczny wzorzec interferencji występującej na *coherence length* jesteśmy w stanie wydobyć informacje na temat próbki wnętrza tkanki oraz wiemy dzięki położeniu lustra referencyjnego położenie próbki. Poszczególne badanie głębszych warstw tkanki przedstawia rysunek 2.3.

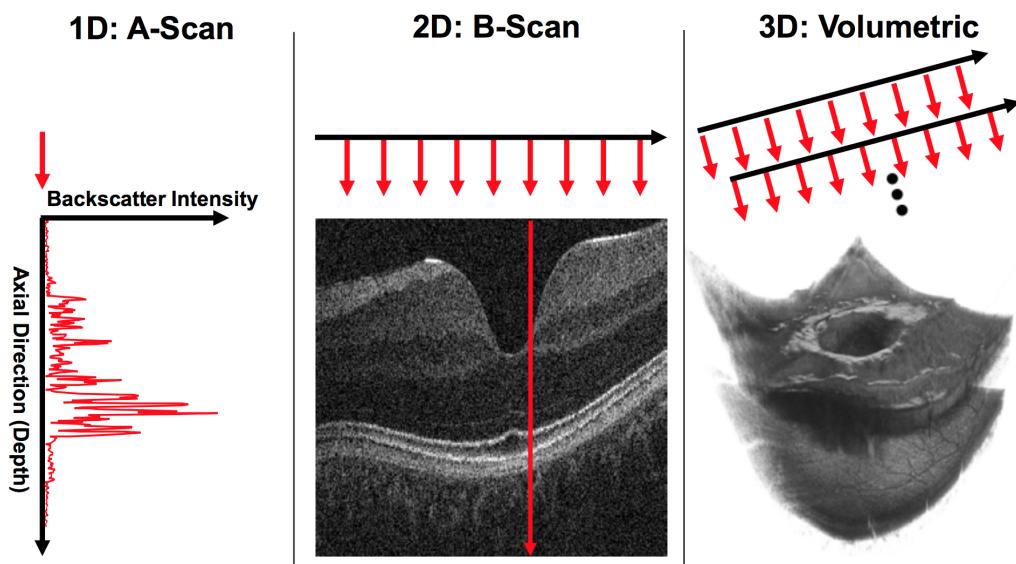


Rys. 2.3: Kolejne etapy badania głębszych warstw tkanki dzięki przesuwaniu lustra referencyjnego. Aktualna pozycja lustra referencyjnego jest zaznaczona kolorem czarnym, natomiast aktualnie badana warstwa tkanki jest zaznaczona kolorem czerwonym.

2.1.1 Metoda uzyskania trójwymiarowego obrazu tkanki

Poprzez poruszenie lustra referencyjnego pomiary przeprowadzane są w głąb tkanki (wzdłuż osi Z). Zbiór pomiarów w głąb tkanki nazywa się A-skanem (obraz jednowymiarowy). Powtarzając ten proces w osi X lub Y i następnie poprzez połączenie sąsiadujących A-skanów otrzymuje się przekrój tkanki zwany B-skanem (obraz

dwuwymiarowy). Cały proces uzyskania B-skanów można powtórzyć dla sąsiadujących przekrojów. Poprzez połączenie otrzymanych B-skanów otrzymuje się obraz trójwymiarowy tkanki. Na rysunku 2.4 [8] przedstawione są poszczególne skany.



Rys. 2.4: [8] Lewy obraz: Pojedynczy A-skan wykonany w głąb tkanki. Środkowy obraz: Otrzymany B-skan poprzez połączenie A-skanów. Prawy obraz: Trójwymiarowy obraz tkanki stworzony poprzez połączenie B-skanów.

2.1.2 OCT w domenie częstotliwości

Metoda OCT, która wykonuje ruch lustrem referencyjnym jest zwana metodą OCT w domenie czasu (ang. *time-domain OCT, TdOCT*). Alternatywną oraz nowszą metodą od TdOCT jest OCT w domenie częstotliwości (ang. *Fourier-domain OCT, FdOCT*). FdOCT umożliwia 100 razy szybsze [@14] skanowanie w porównaniu do TdOCT. TdOCT jest w stanie wykonać ok. 400 A-skanów w przeciągu sekundy, natomiast FdOCT jest ich w stanie wykonać dziesiątki tysięcy. Szybsze skanowanie poprawia również jakość skanów ze względu na to, że pacjent ma mniejszą szansę poruszenia okiem podczas skanowania (ruch oka w trakcie skanowania przyczynia się do powstania artefaktów ruchu na obrazach OCT). Oprócz poprawy szybkości skanowania FdOCT ma wyższą rozdzielcość w przedziale od 3-7 mikrometrów [@14]. Jest to poprawa względem TdOCT o 8-10 mikrometrów.

Większa szybkość oraz rozdzielcość FdOCT w porównaniu do TdOCT jest możliwa dzięki dwóm modyfikacjom technicznym:

1. FdOCT jako źródło światła wykorzystuje laser o wysokiej szerokości pasma, co znacząco zwiększa rozdzielcość.

2. FdOCT jako detektor wykorzystuje spektrometr, który przeprowadza analizę widma fali (połączona wiązka referencyjna i próbki), która dotarła do spektrometru z interferometrem. Wykonanie transformaty Fouriera na widmie fali tworzy A-skan tkanki. Dzięki tej technice FdOCT wyeliminowało potrzebę ruszania lustrem referencyjnym co znacząco zwiększa szybkość skanowania.

FdOCT posiada również wady. Jest droższy od TdOCT ze względu na wykorzystanie drogiego lasera jako źródła światła co powoduje, że jest wykorzystywany tylko w celach badawczych. Następna wada wynika z szybkości powstawania A-skanów. Szybsze skanowanie może prowadzić do utraty jakości obrazów OCT, natomiast utrata tej jakości może być zmniejszona poprzez użycie techniki przetwarzania sygnałów zwanej *oversampling*.

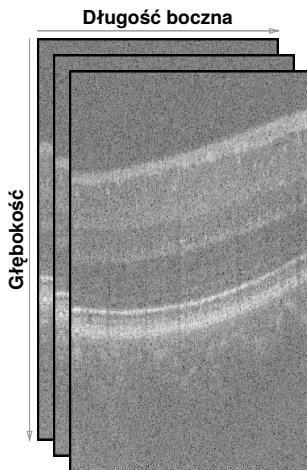
2.2 Angiografia OCT

Angiografia to technika służąca do wizualizacji naczyń krwionośnych i organów ciała. W większości przypadków w medycynie angiografia jest wykonywana poprzez wstrzygnięcie pacjentowi środka kontrastowego, który nie przepuszcza promieni rentgenowskich. Następnie wykorzystuje się jedną z technik obrazowania (np. fluoroskopię) opartą o promienie rentgenowskie. Ta metoda ma jedną znaczącą wadę, jako technika inwazyjna może prowadzić do reakcji alergicznych na środek kontrastowy oraz jest przeciwwskazana kobietom w ciąży i dzieciom. Z tego powodu nieustannie poszukiwane są metody, które jednocześnie są nieinwazyjne i tworzą obrazy o jakości porównywalnej do metod inwazyjnych.

Angiografia OCT jest metodą, która opiera się na A-skanach i B-skanach OCT i jest w stanie stworzyć obraz *en face* naczyń krwionośnych siatkówki oka. Przykład takiego obrazu znajduje się na rysunku 2.2.

2.2.1 Sposób powstania obrazu angiograficznego

Obrazy angiograficzne, które zostały użyte w niniejszej pracy powstały za pomocą metody zwanej *Speckle Variance Detection*, która polega na liczeniu wariancji dla każdego piksela pomiędzy sąsiadującymi B-skanami. Ta technika została wykorzystana, ponieważ lokalna wariancja dla obszarów gdzie występuje przypływ krwi (naczynia krwionośne) ma wyższą wartość w porównaniu do obszarów gdzie występuje struktura statyczna. Rezultatem tej metody jest obraz przepływowaty, którego piksele mają wartość wariancji sąsiadujących B-skanów. Na rysunku 2.5 został przedstawiony przykładowy zbiór sąsiadujących B-skanów.



Rys. 2.5: Trzy sąsiadujące B-skany na postawie których powstaje jeden obraz przepływowowy.

Wartości intensywności pikseli V_{jk} obrazu przepływowego liczone są na podstawie wartości intensywności pikseli I_{ijk} sąsiadujących N B-skanów za pomocą wzoru (na rysunku 2.5 $N = 3$):

$$V_{jk} = \frac{1}{N} \sum_{i=1}^N (I_{ijk} - I_{mean})^2 \quad (2.1)$$

Gdzie j i k to indeksy boczne i głębokościowe B-skanu (współrzędne pikseli), a I_{mean} to średnia intensywność zbioru tych samych pikseli, dla których liczona jest wariancja.

Rezultatem połączenia sąsiadujących obrazów przepłybowych obliczonych ze wszystkich B-skanów jest trójwymiarowy obraz przepływowego. Następnie poprzez użycie projekcji maksymalnego natężenia (ang. *maximum intensity projection, MIP*), która polega na projekcji woksela o najwyższej intensywności z obrazu 3D na obraz 2D. W przypadku trójwymiarowego obrazu przepływowego będzie to woksel, którego wartość była największą wartością wariancji pikseli sąsiadujących B-skanów. Po zastosowaniu projekcji maksymalnego natężenia na trójwymiarowym obrazie przepływowym otrzymujemy angiograficzny obraz *en face* naczyń krwionośnych.

2.3 Zastosowania OCT

Optyczna tomografia koherencyjna ze względu na swoje właściwości (badanie nie-inwazyjne oraz *in vivo*) jest metodą, która ma szerokie zastosowanie w medycynie oraz w innych specjalizacjach.

2.3.1 OCT w okulistyczce

Najbardziej popularnym zastosowaniem OCT w medycynie jest badanie oka [6]. Technika OCT umożliwia przechwycenie trójwymiarowych obrazów części oka takich jak dno, czy warstwy przednie. Dzięki temu jest wykorzystywane do diagnozowania takich chorób jak stwardnienie rozsiane, zwyrodnienie plamki żółtej, czy jaskra.

2.3.2 OCT w gastroenterologii i dermatologii

OCT w porównaniu do innych metod diagnostycznych w medycynie jest techniką nową. Lekarze i naukowcy nieustannie starają się znaleźć nowe zastosowania dla OCT, która jest metodą obiecującą i szybko rozwijającą się. OCT jest potencjalnym kandydatem by w niektórych diagnozach zastąpić konwencjonalną biopsję, która wymaga usunięcia kawałka tkanki z organizmu. Przykładem takiego zastosowania jest badanie struktury błon śluzowych i podśluzowych w układzie pokarmowym. W tym przypadku OCT dostarczyło czyste obrazy, które dostarczają lekarzom dużo diagnostycznej informacji [12]. Innym przykładem są próby wykorzystania OCT do wczesnej diagnozy raka skóry, który obecnie również diagnozowany jest poprzez biopsję. W tym przypadku OCT w obecnym stopniu zaawansowania nie jest w stanie dostarczyć na tyle dokładnych danych by stać się jedną metodą diagnozy. Następnym przykładem zastosowania OCT w dermatologii jest diagnoza zapalnych chorób skóry [7].

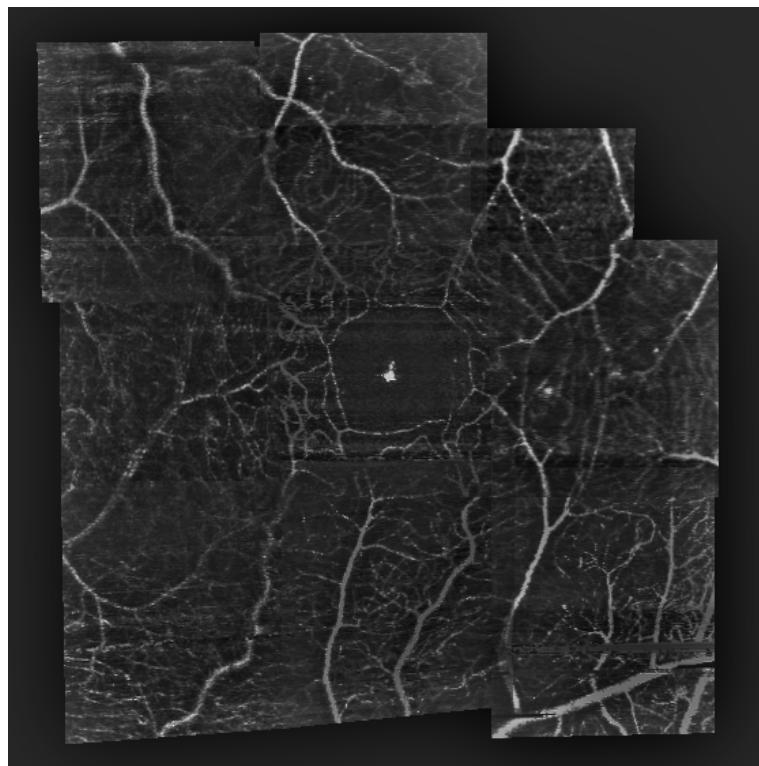
2.3.3 OCT w przemyśle

OCT wykorzystywane jest również do zastosowań przemysłowych. Umożliwia badanie np. grubości materiałów [17], czy badanie grubości warstwy pancerza tabletek podczas ich produkcji w przemyśle farmaceutycznym [11].

Algorytmy korejestracji przestrzennej obrazów OCT

Mozaiką nazywa się obraz, który powstaje poprzez połączenie grupy obrazów zwanych kafelkami na podstawie ich wzajemnych relacji. Znanym oraz popularnym przykładem łączenia obrazów w jeden większy jest funkcja panoramy w telefonach komórkowych, czy aparatach fotograficznych. Od strony użytkownika proces tworzenia panoramy polega na powolnym przesuwaniu telefonem po linii poziomej do momentu aż żądany krajobraz zostanie uchwycony. Od strony urządzenia proces polega na wykonywaniu serii zdjęć oraz następnie łączenie nachodzących klatek w jeden obraz. Rezultatem jest jednolita panorama, która składa się z grupy mniejszych węższych zdjęć.

Celem niniejszej pracy jest stworzenie mozaiki OCT (przykład na rysunku 3.1), która powstaje z połączenia mniejszych nachodzących na siebie nawzajem angiograficznych obrazów OCT (przykład obrazu angiograficznego OCT znajduje się z prawej strony na rysunku 2.1).

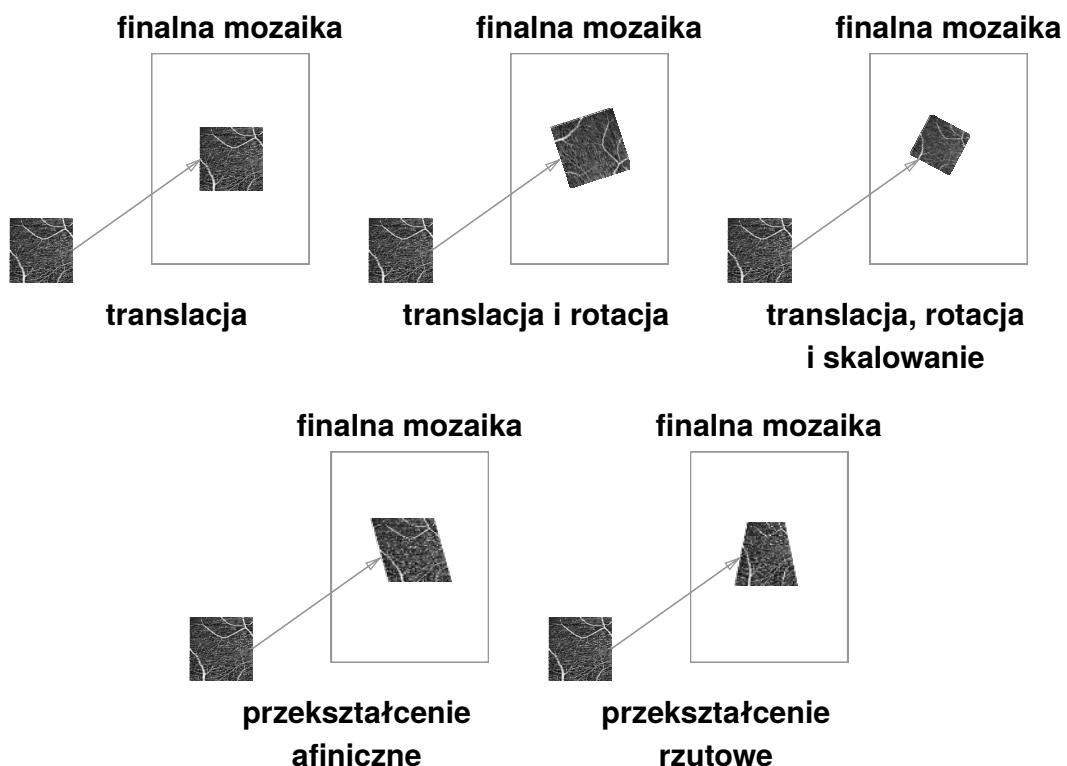


Rys. 3.1: Mozaika OCT stworzona z połączenia angiograficznych obrazów OCT.

Proces automatycznego stworzenia mozaiki takiej jak na rysunku 3.1 jest zadaniem nietrywialnym i wymaga dokładnej analizy dziedzinowej oraz precyzyjnego wyboru metod. Pierwszym krokiem jest wybór modelu deformacji kafelków (sekcja 3.1), następnym etapem, który jest jednocześnie najbardziej wymagającym jest wybór metody wzajemnej korejestracji kafelków (sekcja 3.2). Posiadając zdefiniowane wzajemne relacje kafelków oraz ich docelowe położenie w finalnej mozaice należy wykonać proces łączenia kafelków (sekcja 3.3). W każdej z tych sekcji została wyjaśniona idea metody w kontekście stworzenia mozaiki OCT, natomiast szczegółowy opis zaimplementowanych metod znajduje się w rozdziale 4.

3.1 Model deformacji kafelków

Model deformacji kafelków określa dozwolone przekształcenia geometryczne, które odwzorują piksele kafelka do pikseli kafelka w finalnej mozaice. Ze względu na to, że angiograficzne obrazy OCT znajdują się na jednej płaszczyźnie możliwy zbiór modeli deformacji ogranicza nam się do transformacji dwuwymiarowych, które zostały zobrazowane na rysunku 3.2.



Rys. 3.2: Zbiór transformacji dwuwymiarowych dla przykładowego angiograficznego obrazu OCT.

Idealnie OCT powinno tworzyć angiograficzne obrazy, które są względem siebie tylko przesunięte, natomiast w rzeczywistości pojawiają się zniekształcenia wynikające

z niedokładności urządzenia oraz ruchu oka pacjenta, przez co niektóre kafelki są nieznacznie obrócone względem siebie. Z tego względu wybranym modelem deformacji kafelków został **model transformacji ciała sztywnego**, czyli połączenie translacji i rotacji.

3.1.1 Matematyczny zapis modelu transformacji ciała sztywnego

Współrzędne piksela w kafelku możemy określić jako trójelementowy wektor $\tilde{x} = (x, y, 1)$, gdzie x i y to współrzędne piksela w układzie współrzędnych kafelka. Tak zdefiniowany piksel poddaje się transformacji by uzyskać współrzędne tego piksela $\hat{x} = (x', y')$ w układzie współrzędnych finalnej mozaiki. W sekcji 3.1 została wybrana transformacja ciała sztywnego, który zakłada tylko translację oraz rotację i może być zapisana jako:

$$\hat{x} = \begin{bmatrix} R & t \end{bmatrix} \tilde{x} \quad (3.1)$$

gdzie:

$$R = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \quad \text{i} \quad t = \begin{bmatrix} t_x \\ t_y \end{bmatrix} \quad (3.2)$$

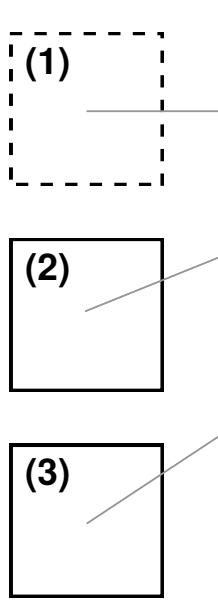
W równaniu 3.2 θ to kąt obrotu (rotacja) względem początku układu współrzędnych, a t_x i t_y to odpowiednio przesunięcia względem osi x i osi y. Parametry θ , t_x i t_y są niewiadomymi równania 3.1, których obliczenie jest tematem sekcji 3.2.

3.2 Korejestracja kafelków

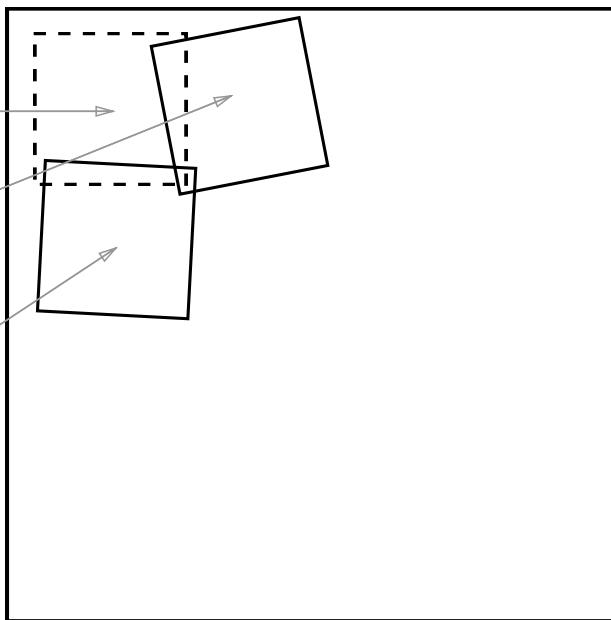
Po wyborze modelu deformacji kafelków można przejść do wyboru metody, która będzie określać jego parametry (w przypadku niniejszej pracy są to parametry przesunięcia i obrotu). Metoda ta powinna zwrócić takie wartości by kafelek znalazł się w odpowiednim miejscu w finalnej mozaice z możliwie najmniejszym błędem. By rozwiązać ten problem najpierw trzeba poznać położenie kafelków względem siebie oraz ustalić kafelek referencyjny. Rysunek 3.3 przedstawia przykładowe rozmieszczenie kafelków. Kafelek referencyjny oznaczony poprzez przerywane linie jest przesuwany do finalnej mozaiki, natomiast żeby odpowiednio umieścić kafelki (2) i (3) trzeba najpierw znać dopasowanie kafelków (2) i (3) do kafelka referencyjnego (1). W przykładzie na rysunku 3.3 zostało założone, że kafelki (2) i (3) mogą być do-

pasowane do kafelka (1). Informacja na temat tego, które kafelki należy dopasować do których kafelek wynika z wiedzy dziedzinowej i opisane jest to w sekcji 4.2.

Zbiór kafelków



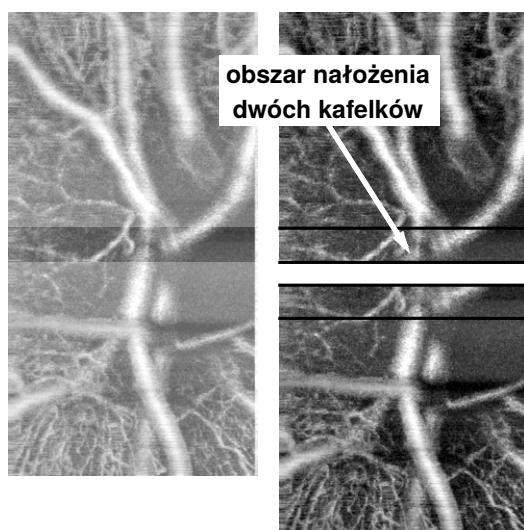
Finalna mozaika



Rys. 3.3: Przykładowe rozmieszczenie kafelków na finalnej mozaice. Kafelek referencyjny został wyróżniony przerywaną linią.

Dopasowanie dwóch kafelków do siebie nazywa się również ich korejestracją, czyli przeniesieniem dwóch kafelków do wspólnego układu współrzędnych w taki sposób by były względem siebie dopasowane. Prosty przykład korejestracji przedstawiony jest na rysunku 3.4, gdzie dwie kafelki mają wspólny obszar nałożenia.

kafelki po procesie korejestracji kafelki z nakładającym się obszarem



Rys. 3.4: Przykładowa korejestracja dwóch kafelków ze wspólnym obszarem nałożenia.

Dopasowanie kafelków z rysunku 3.4 jest przykładem bardzo prostym, ponieważ wymaga tylko translacji jednego kafelka względem drugiego. W rzeczywistości pomiary OCT nie są tak dokładne przez co prosta translacja wzdłuż jednej z osi nie wystarcza i wymagane jest automatyczne wyznaczenie parametrów translacji i rotacji, które może być wykonane na wiele sposobów. Dwie najpopularniejsze metody to:

1. **Korejestracja na podstawie wartości pikseli** (sekcja 3.2.1) - Metoda przesuwająca jeden obraz względem drugiego i następnie oceniająca jak nałożone piksele obrazów pasują do siebie na podstawie ich wartości.
2. **Korejestracja na podstawie cech** (sekcja 3.2.2) - Metoda wyodrębniająca charakterystyczne cechy z każdego obrazu i następnie dopasowującą te cechy by osiągnąć globalną spójność. Dopasowania są później użyte przez algorytm do wyznaczenia macierzy transformacji pomiędzy obrazami.

W niniejszej pracy do osiągnięcia żądanego rezultatu zostało wykorzystane połączenie podejścia tych dwóch metod. Wykorzystane techniki dokładnie opisane są w rozdziale 4.

3.2.1 Korejestracja na podstawie wartości pikseli

Korejestracja na podstawie wartości pikseli jest często nazywana metodą bezpośrednią, ponieważ wymaga bezpośredniego porównania wartości pikseli nakładających się obrazów. Pierwszym krokiem w tej metodzie jest wybranie odpowiedniej miary błędu. Taką miarą błędu może być funkcja suma błędów kwadratowych E_{RSS} (ang. *residual sum of squares, RSS*). Mając daną funkcję $I_S(w_i)$, która zwraca wartość piksela w obrazie źródłowym (przesuwany) w lokacji $w_i = (x_i, y_i)$ oraz funkcję $I_T(w_i)$, która zwraca wartość piksela w obrazie docelowym (referencyjnym, pozostającym na miejscu) E_{RSS} można zdefiniować jako:

$$E_{RSS} = \sum_i [I_S(w_i + u) - I_T(w_i)]^2 = \sum_i e_i^2 \quad (3.3)$$

gdzie $u = (t_x, t_y)$ to wartość przesunięcia obrazu źródłowego, a $e_i = I_S(w_i + u) - I_T(w_i)$ to błąd resztowy pomiędzy wartościami pikseli przesuniętego obrazu źródłowego i docelowego.

Mając tak zdefiniowaną miarę błędu najprostszą metodą, która wyznaczyłaby najlepsze dopasowanie obrazów byłoby obliczenie błędu dla każdego możliwego przesunięcia obrazu źródłowego. Posiadając wartości wszystkich błędów z odpowiednimi dla

nich wartościami przesunięcia można by wybrać wartość najmniejszego błędu. Takie podejście natomiast wiąże się z ogromnym kosztem obliczeniowym w przypadku, gdy obrazy są dużej rozdzielczości i jest ich wiele. Proces można przyspieszyć używając algorytm optymalizacji Levenberga-Marquardta [16] albo poprzez wykorzystanie piramid obrazu w metodzie *Hierarchical Motion Estimation* [15].

Ze względu na specyficzny charakter problemu niniejszej pracy została zaimplementowana autorska metoda, która również bazuje na bezpośrednim odczytaniu wartości pikseli i jest opisana w sekcji 4.4, natomiast różni się znaczaco od metod wymienionych w niniejszej sekcji.

3.2.2 Korejestracja na podstawie cech

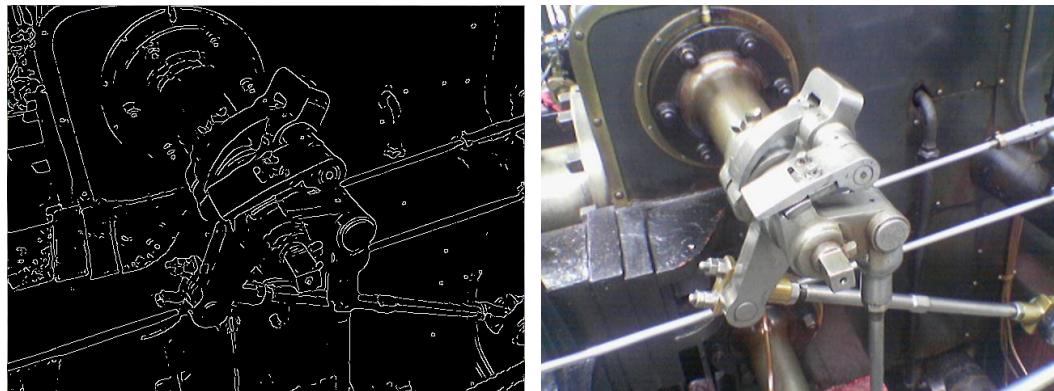
Korejestracja na podstawie cech jest techniką, której pierwszym krokiem jest wyekstrahowanie charakterystycznych cech z każdego obrazu i następnie ich dopasowanie. Na podstawie tej informacji można wyestymować macierz transformacji jednego obrazu względem drugiego. W niniejszej sekcji zostaną opisane metody, które realizują wymienione kroki.

Ekstrakcja cech

Ekstrakcja cech (ang. *keypoints detection*) jest techniką, w której algorytm poprzez skanowanie obrazu piksel po pikselu podejmuje decyzję, czy dany punkt jest określona cechą, czy nie. Rezultatem tej techniki może być zbiór punktów, krawędzi, rogów, czy obszarów. Każda cecha ma swój deskryptor (ang. *keypoint descriptor*), który ją określa. W dziedzinie wizji komputerowej zostało zaimplementowanych wiele algorytmów, które ekstrahują cechy obrazów. Poniżej zostały wymienione niektóre algorytmy, które są zaimplementowane przez bibliotekę przetwarzania obrazów OpenCV:

- **SIFT** [10] - algorytm, który potrafi znaleźć charakterystyczne cechy niezależnie od lokalizacji, rotacji, czy skali obrazu. Jest jednym z najpopularniejszych algorytmów i jest wykorzystywany w niniejszej pracy. Został opisany dokładniej w sekcji 4.3.1.
- **SURF** [1] - SURF jest oparty na tych samych zasadach co algorytm SIFT, natomiast detale poszczególnych kroków są różne. Autorzy algorytmu w swojej pracy pokazali, że SURF potrafi być szybszy of SIFT nie tracąc przy tym jakości. Tak jak SIFT jest niezależny od lokalizacji, rotacji, czy skali obrazu.

- **Canny [5]** - Algorytm, który potrafi wykryć krawędzie obiektów w obrazie, przez co dostarcza dużo informacji o ich strukturze. Na rysunku 3.5 zostało zaprezentowane działanie algorytmu Canny.

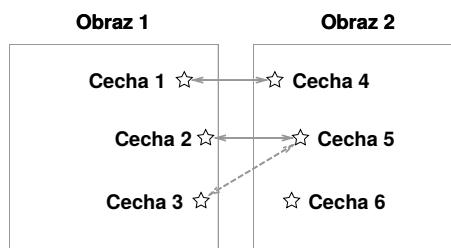


Rys. 3.5: Lewy obraz: Wynik zaaplikowanego algorytmu Canny na kolorowym obrazie z prawej strony. Prawy obraz: Oryginalny obraz.

Dopasowanie wyekstrahowanych cech

Następnym krokiem po wyekstrahowaniu cech w dwóch obrazach jest ich dopasowanie na podstawie deskryptorów. Najprostszym rozwiązaniem byłoby porównanie każdej cechy z pierwszego obrazu do każdej cechy z drugiego obrazu, natomiast dla niektórych zastosowań jest to rozwiązanie zbyt wymagające obliczeniowo. Przez co zostały opracowane bardziej wydajne techniki. Przykładem takiej metody jest wykorzystanie drzewa k-d przez Beis i Lowe [2].

Kolejnym wyzwaniem w dopasowaniu cech jest wykrycie obserwacji odstających (ang. *outliers*). Problem obrazuje rysunek 3.6.



Rys. 3.6: Dwa obrazy, w których zostały wykryte po trzy cechy (na rysunku oznaczone jako gwiazdy). Algorytm szukający dopasowań wykrył dopasowania (na rysunku oznaczone jako strzałki) pomiędzy cechami: 1 i 4, 2 i 5, 3 i 5. Z wiedzy dziedzicznej wiemy, że obrazy nachodzą na siebie (tak jak angiograficzne obrazy OCT). Patrząc na dopasowania pomiędzy obrazami można zauważyć, że jedno z nich nie pasuje do reszty (strzałka narysowana przerywaną linią). Zakładając, że możemy tylko obracać i przesuwać obrazy niemożliwa jest estymacja macierzy transformacji biorąc pod uwagę wszystkie dopasowania. Niepoprawne dopasowanie trzeba odrzucić by otrzymać poprawną macierz transformacji.

Problem z rysunku 3.6 można rozwiązać poprzez odpowiednie filtrowanie dopasowań. Jedną z takich metod jest RANSAC (ang. *random sample consensus*, *RANSAC*), który jest algorytmem iteracyjnym i jego zadaniem jest eliminacja obserwacji odstających (w naszym przypadku są to dopasowania, które nie pasują do większości). RANSAC jest wykorzystywany w niniejszej pracy i jest dokładniej opisy w sekcji 4.3.3. Do rozwiązania tego problemu można również wykorzystać relacje geometryczne (wiedza dziedzinowa) pomiędzy obrazami, których cechy są dopasowywane. Taką relację geometryczną może być np. informacja, że obraz nie może się przesunąć dalej niż 10 pikseli. Dzięki temu wiemy, że wszystkie dopasowania cech, które znajdują się dalej niż 10 pikseli są niepoprawne. Filtracja na podstawie wiedzy dziedzinowej również jest wykorzystywana w niniejszej pracy i jest opisana dokładniej w sekcji 4.3.3.

Estymacja macierzy transformacji na podstawie dopasowań

Po wyznaczeniu dopasowań pomiędzy dwoma obrazami można przejść do estymacji macierzy transformacji jednego obrazu względem drugiego. Estymacja macierzy transformacji odbywa się na podstawie ustalonego modelu deformacji (sekcja 3.1). Biorąc pod uwagę jako model transformacji ciało sztywne (dozwolona rotacja i translacja) równanie pozwalające obliczyć parametry rotacji i translacji można zapisać jako:

$$T = SR + t \quad (3.4)$$

gdzie T to zbiór cech w obrazie docelowym, S to zbiór cech w obrazie źródłowym, R to macierz rotacji, a t to wektor translacji. Wyznaczenie R i t z tego równania byłoby proste, jeżeli wszystkie cechy z obrazu źródłowego dałoby się idealnie przekształcić na cechy obrazu docelowego. Założmy, że do obliczenia parametrów rotacji i translacji zostanie wzięte pod uwagę tylko jedno dopasowanie. Następnie wyznaczoną macierz transformacji stosujemy do cechy z obrazu źródłowego innego dopasowania. Istnieje bardzo mało szansa, że wynikiem będzie cecha obrazu docelowego z tego dopasowania (będzie blisko niej, ale nie idealnie w tym samym miejscu). Ta różnica wynika z tego, że obrazy posiadają szum oraz algorytmy zwracające cechy nie są tak dokładne. Jednym z prostszych rozwiązań tego problemu jest wykorzystanie metody minimalizacji błędu średniokwadratowego, który polega na znalezieniu takiej macierzy transformacji dla której błąd średniokwadratowy pomiędzy cechą w obrazie docelowym, a cechą z obrazu źródłowego po transformacji jest najmniejszy. Mając dane współrzędne cechy w obrazie docelowym T_i oraz współrzędne

odpowiadającej cechy w obrazie źródłowym S_i oraz macierz transformacji M błąd średniokwadratowy można zapisać jako:

$$E_{LS} = \sum_i d(T_i, MS_i)^2 \quad (3.5)$$

gdzie E_{LS} to minimalny błąd średniokwadratowy, a funkcja $d(\dots)$ to odległość euklidesowa pomiędzy dwoma punktami.

Równanie 3.5 można rozwiązać algorytmami optymalizacji. W niniejszej pracy do wyznaczenia parametrów rotacji i translacji została wykorzystana funkcja OpenCV, której użycie jest opisane w sekcji 4.5.

3.3 Łączenie kafelków

Po wyznaczeniu macierzy translacji pomiędzy dwoma kafelkami następnym krokiem jest ich złączenie w obszarze, w którym na siebie nachodzą. Problem by nie istniał jeżeli dopasowanie byłoby idealne oraz ekspozycja dwóch kafelków byłaby identyczna. Niestety w rzeczywistym świecie po złączeniu widoczne są krawędzie kafelków oraz inne artefakty wynikające z błędu rejestracji, czy niedokładności aparatu. Stworzenie ładnej oraz jednorodnej mozaiki wymaga określenia, które piksele z dwóch kafelków należy wykorzystać w procesie łączenia oraz jakie przypisać im wagę. Najprostszą metodą jest wzięcie średniej ważonej każdego piksela z dwóchłączonych kafelków:

$$f(x) = \frac{\sum_i \omega_i(x) f_i(x)}{\sum_i \omega_i(x)} \quad (3.6)$$

gdzie i to liczba nakładających się kafelków, $f(x)$ to wartość piksela x w finalnej mozaice, $f_i(x)$ to wartość piksela x w kolejnych nakładających się kafelkach, a $\omega(x)$ to waga piksela x . Następnym krokiem w tej metodzie jest ustalenie wartości wag. Jednym z pomysłów jest wykorzystanie zwykłej średniej ($\omega(x) = 1$ dla każdego piksela x). Nie jest to jednak najlepszy sposób, ponieważ krawędzie kafelków są dalej widoczne. Lepszym podejściem jest uzależnienie wartości wag od odległości piksela od środka kafelka do którego należy tzn. piksel znajdujący się przy krawędzi swojego kafelka będzie miał mniejszą wagę niż piksel znajdujący się blisko środka swojego kafelka (metoda wykorzystana jest w niniejszej pracy i dokładniej jest opisana w sekcji 4.7). Istnieją również bardziej zaawansowane metody takie jak algorytm *multi-band blending* zaproponowany przez Burt i Adelson [4], który daje jeszcze lepsze rezultaty [3] niż średnia ważona.

Proponowane algorytmy

Niniejszy rozdział opisuje szczegółowo kolejne kroki rozwiązania problemu, który został dokładnie przedstawiony w sekcji 1.1. Sekcja 4.1 opisuje sposób implementacji oprogramowania oraz wykorzystane technologie. Sekcja 4.2 opisuje wiedzę dziedzinową na temat danych wejściowych (kafelków). Następnie sekcje od 4.3.1 do 4.7 opisują zasadę działania poszczególnych metod w kolejności zgodnej z ich wykonywaniem w programie.

4.1 Oprogramowanie

Program o nazwie `mostitch`, który jest celem niniejszej pracy jest całkowicie napisany w języku C++. Ten język został wybrany ze względu na to, że wykorzystywana biblioteka przetwarzania obrazów OpenCV (opisana w skrócie w sekcji 4.1.2) posiada interfejs w języku C++. Program można uruchomić z konsoli systemu za pomocą komendy:

```
./mostitch path_to_config_file
```

gdzie `path_to_config_file` to obowiązkowy argument do programu, który wskazuje ścieżkę do pliku konfiguracyjnego (opisanego dokładniej w sekcji 4.1.1), który określa wszystkie parametry niezbędne do prawidłowego działania programu.

Program umożliwia stworzenie mozaik dla dowolnej ilości zbiorów kafelków. Dodatkowo dla każdego zbioru kafelków powstają różne wersje mozaik, ze względu na wykorzystanie różnych metod ich tworzenia. Dzięki tej funkcjonalności użytkownik może wybrać mozaikę, która jest najbardziej odpowiednia. Wyjściem programu więc jest zbiór mozaik (obrazy .png), w których każdy z nich odpowiada wybranemu zbioru kafelków oraz metodzie tworzenia.

4.1.1 Plik konfiguracyjny

Do zarządzania plikiem konfiguracyjnym została wykorzystana biblioteka **libconfig**¹, która umożliwia bezproblemowy odczyt pliku konfiguracyjnego z rozszerzeniem

¹<http://www.hyperrealm.com/libconfig/>

.cfg. Format pliku konfiguracyjnego jest bardziej czytelny w porównaniu do powszechnie wykorzystywanych plików XML. Biblioteka również jest świadoma typu zmiennej przez co unika się konwersji typu string na typy takie jak int, czy float.

W pliku konfiguracyjnym zawarte są informacje:

- Ścieżka do folderu z folderami zawierającymi zbiory kafelków dołączenia.
- Ścieżka do miejsca, w którym będą zapisane wynikowe mozaiki.
- Parametry pozwalające na automatyczne wczytanie obrazów kafelków.
- Parametry modyfikujące działanie metod przetwarzania kafelków.

Wszystkie parametry zawarte w pliku konfiguracyjnym są szczegółowo opisane w przykładzie pliku konfiguracyjnego o nazwie config.cfg, który jest dołączony do niniejszej pracy.

4.1.2 OpenCV

Wszystkie rozwiązania zaimplementowane w niniejszej pracy bazują na bibliotece przetwarzania obrazów o nazwie **OpenCV**. Biblioteka jest dozwolona do bezpłatnego wykorzystania w projektach prywatnych i komercyjnych. OpenCV jest używane w ogromnej ilości projektów z różnych dziedzin i została ściągnięta już ponad 9 milionów razy². OpenCV cieszy się taką popularnością ze względu na szybkość działania, implementację większości metod przetwarzania obrazu oraz umiejętnością pracy na wielu rdzeniach. Kod źródłowy biblioteki jest dostępny publicznie w serwisie Github³ przez co każdy może rozwijać OpenCV i ma wgląd do zaimplementowanych metod.

4.2 Wiedza dziedzinowa

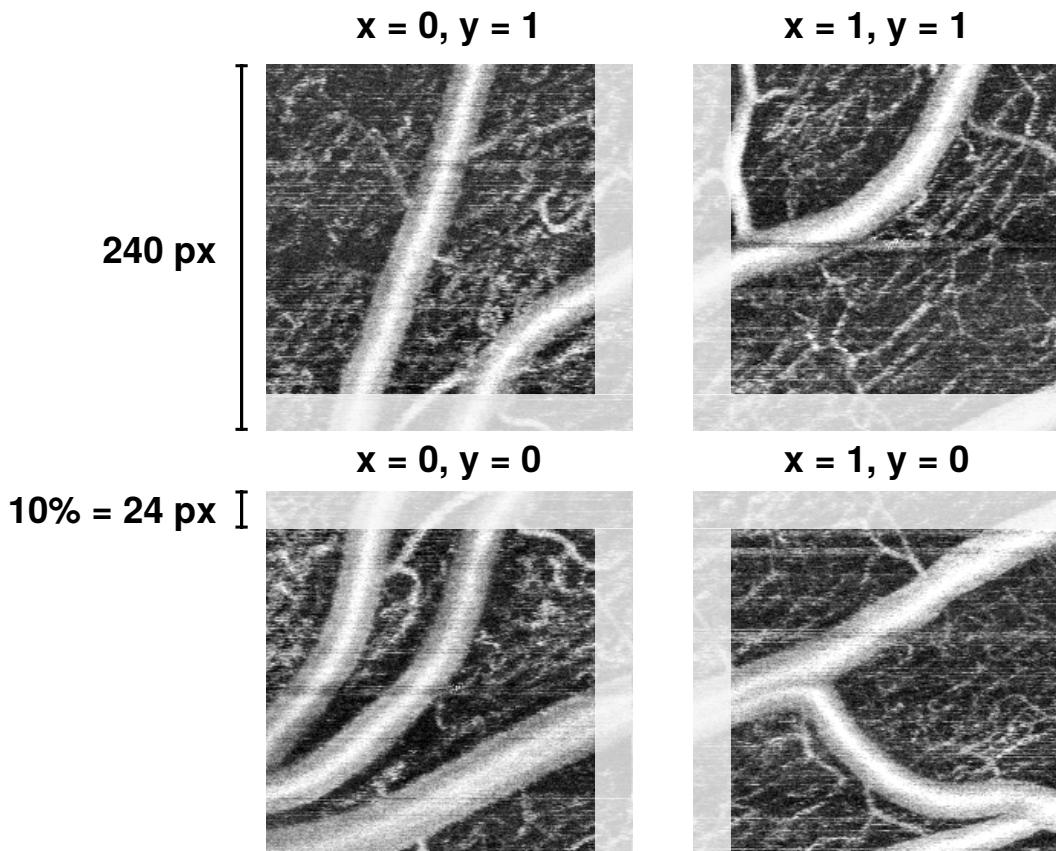
Wiedza dziedzinowa określa zbiór informacji związanych z problemem i danymi wejściowymi. Dokładne zdefiniowanie oraz analiza wiedzy dziedzinowej prowadzi do ułatwienia problemu oraz lepszych rezultatów. W przypadku problemu niniejszej pracy danymi wejściowymi do stworzenia mozaiki są angiograficzne obrazy OCT (kafelki). Są to monochromatyczne obrazy o rozmiarze 240 na 240 pikseli. Wraz z obrazami do programu dostarczane są jeszcze dwie informacje:

²<http://opencv.org>

³<https://github.com/Itseez/opencv>

- Względna pozycja kafelków względem siebie, która jest zapisana w nazwie obrazu. Lokalizacja jest podana za pomocą współrzędnych x i y układu kartezjańskiego.
- Szacowany obszar nałożenia kafelków na siebie wyrażony w procentach szerokości obrazu.

Na rysunku 4.1 jest przedstawiony przykładowy zbiór kafelków przeznaczony do stworzenia jednej mozaiki wraz z dostarczoną wiedzą dziedzinową.



Rys. 4.1: Cztery kafelki rozmieszczone zgodnie z ich pozycją w docelowej mozaice. Na każdym obrazie białym przezroczystym prostokątem zaznaczony jest szacowany obszar nałożenia sąsiadujących kafelków.

Informacje dostarczone wraz z kafelkami mają istotny wpływ na działanie algorytmu. Jasno podane położenie kafelek informuje algorytm, który kafelek należy dopasowywać z którym (sekcja 4.3.3), a szacowany obszar nałożenia określa obszar, w którym należy ekstrahować cechy (sekcja 4.3.1). Bez tych informacji stworzenie mozaiki byłoby o wiele bardziej skomplikowane.

4.3 Rejestracja kafelków z wykorzystaniem cech

Idea ekstrakcji cech w celu rejestracji obrazów została opisana w sekcji 3.2.2 i jest to pierwszy krok w tworzeniu mozaiki. W niniejszej sekcji najpierw zostanie wyjaśniony wybrany algorytm do ekstrakcji cech oraz sposób jego użycia w programie (sekcja 4.3.1). Następnie przedstawiony zostanie proces dopasowania wyekstrahowanych cech do siebie (sekcja 4.3.2), a na koniec tej sekcji wyjaśnione zostanie filtrowanie znalezionych dopasowań (sekcja 4.3.3).

4.3.1 Algorytm SIFT

SIFT (ang. *scale-invariant feature transform*) to algorytm służący do wykrycia i opisania cech w obrazie. Został opublikowany przez David Lowe w 1999 roku [10]. Dzięki umiejętności wyekstrahowania cech, które są niewrażliwe na rotacje, skalowanie, lokalizacje i przekształcenie afoniczne obrazu jest stosowany w wielu aplikacjach (np. lokalizacja robota [13], czy rozpoznawanie zachowań człowieka [9]).

Działanie algorytmu można podzielić na 6 kroków, które są opisane w kolejności wykonania poniżej.

1. Reprezentacja obszaru za pomocą przestrzeni *scale-space*

Algorytm rozpoczyna swoją pracę tworząc przestrzeń *scale-space* obszaru obrazu, w którym ma wyekstrahować cechy. *Scale-space* generuje się poprzez kilkakrotne użycie filtra Gaussa z różną siłą na docelowym obszarze. Następnie oryginalny obraz próbujemy zmniejszając jego wielkość dwukrotnie i powtarzamy proces użycia filtru Gaussa. Obrazy o tej samej wielkości tworzą oktawę. Na rysunku 4.2 pokazana jest przestrzeń *scale-space* dla przykładowego obrazu.

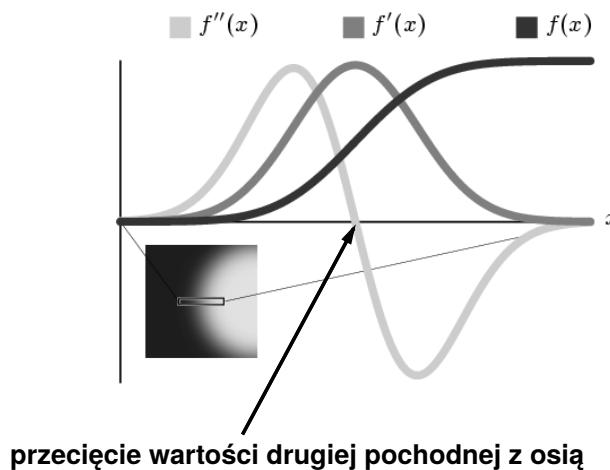


Rys. 4.2: Przestrzeń *scale-space* dla przykładowego obrazu. Skala rośnie wraz z siłą filtru Gaussa. Obrazy tej samej wielkości tworzą oktawę.

Autor algorytmu SIFT zalecana użycie czterech oktaw i pięciu poziomów siły filtru Gaussa.

2. Obliczenie *difference of Gaussians*

Charakterystycznymi punktami obrazu są krawędzie i rogi. Są to miejsca, które najlepiej nadają się na znalezienie cech. Operacja *laplacian of Gaussian*, *LoG* świetnie sobie radzi w wykrywaniu krawędzi i rogów w obrazie. Polega ona na wyeliminowaniu szumu poprzez zaaplikowanie filtru Gaussa, a następnie obliczenie drugiej pochodnej. Na rysunku 4.3 przedstawione jest działanie LoG na jednowymiarowym fragmencie obrazu.

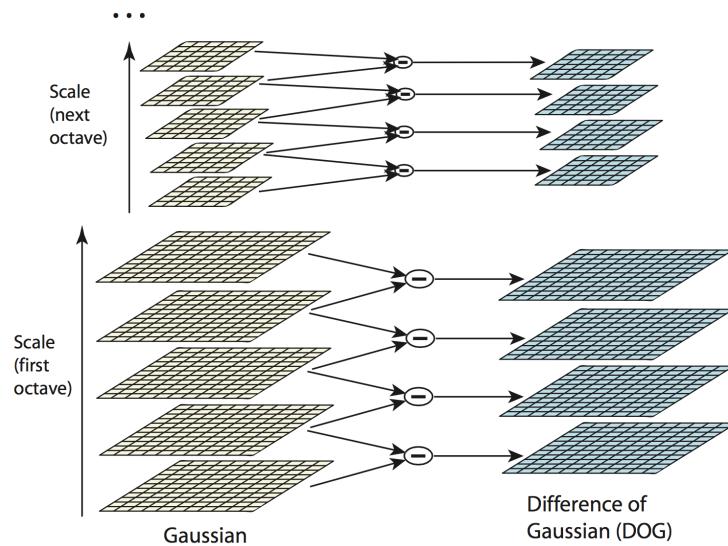


Rys. 4.3: Przykład wykrycia krawędzi przez LoG. Czarny wykres przedstawia jednowymiarowy sygnał z obrazu. Miejsce przecięcia się z osią drugiej pochodnej tzw. *zero-crossing* wskazuje miejsce krawędzi w oryginalnym obrazie.

Obliczenie drugiej pochodnej całego obrazu jest natomiast kosztownym procesem. SIFT implementuje operację *difference of Gaussians*, *DoG*, która w przybliżeniu jest równa LoG, ale jest o wiele szybsza. DoG polega na odjęciu dwóch kolejnych obrazów z tej samej oktawy. Rysunek 4.4 przedstawia operację DoG. DoG jest obliczane dla każdej pary w każdej oktawie.

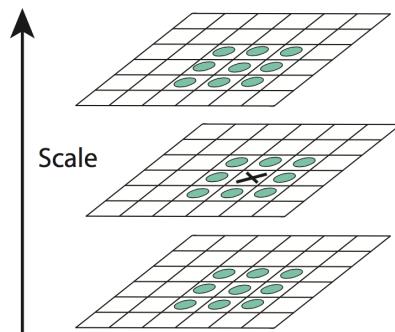
3. Ekstrakcja cech

Następnym krokiem po obliczeniu obrazów DoG jest ekstrakcja cech. Na rysunku 4.4 druga pochodna poprzez przecięcie osi (*zero-crossing*) wskazuje miejsce występowania krawędzi w oryginalnym obrazie, natomiast można też zauważać, że lokalne maksimum z lewej strony i lokalne minimum z prawej strony wykresu wskazują



Rys. 4.4: [10] Poprzez proces odjęcia kolejnych zdjęć z jednej oktawy otrzymujemy obrazy DoG, które doskonale odzwierciedlają LoG.

dwie strony krawędzi. SIFT ekstrahuje cechy poprzez wykrycie lokalnych maksimów i minimów w obrazach DoG. Proces wyjaśniony jest na rysunku 4.5.



Rys. 4.5: [10] Maksima i minima wykrywane są za pomocą porównania wartości aktualnie badanego piksela (zaznaczony czarnym krzyżykiem) z jego 26 sąsiadami powstałymi poprzez złączenie kolejnych obrazów DoG z jednej oktawy. Cecha jest wykrywana w miejscu aktualnie badanego piksela jeżeli wartość tego piksela jest największa lub najmniejsza spośród 26 sąsiadów.

4. Filtrowanie cech

W większości przypadków ekstrakcja cech może wygenerować ich ogromną ilość. SIFT poprzez zastosowanie następujących dwóch filtrów pozbywa się cech najgorszej jakości:

1. Pierwszy filtr polega na sprawdzeniu intensywności piksela obrazu DoG w miejscu gdzie wykryto cechę. Jeżeli wartość piksela jest mniejsza niż ustalony wcześniej próg cecha jest odrzucana.
2. Drugi filtr polega na obliczeniu dwóch prostopadłych do siebie gradientów w miejscu wystąpienia cechy. W zależności od otoczenia cechy możliwe są 3 opcje:
 - Wartości dwóch gradientów będą małe z czego wynika, że w otoczeniu cechy jest region płaski. Cecha jest odrzucana.
 - Wartość jednego gradientu będzie duża, a drugiego mała. Oznacza to, że wzdłuż gradientu z małą wartością występuje krawędź w oryginalnym obrazie (gradient z dużą wartością jest prostopadły do krawędzi). Cecha jest odrzucana.
 - Wartość dwóch gradientów będzie duża. Oznacza to, że w cecha znajduje się w miejscu występowania rogu w oryginalnym obrazie. Cecha jest zachowana ze względu na to, że rogi są najlepszymi cechami.

5. Przydzielenie orientacji cechom

Kolejny krok algorytmu SIFT polega na obliczeniu orientacji każdej cechy na podstawie jej otoczenia, przez co deskryptory cech mogą być przedstawione względnie do swojej orientacji. Dzięki temu algorytm SIFT jest niezależny od rotacji obrazów. Proces obliczenia orientacji zaczyna się od wyznaczenia wartości gradientu $m(x, y)$ i orientacji $\theta(x, y)$ dla każdego piksela z otoczenia wokół cechy za pomocą wzorów:

$$m(x, y) = \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2} \quad (4.1)$$

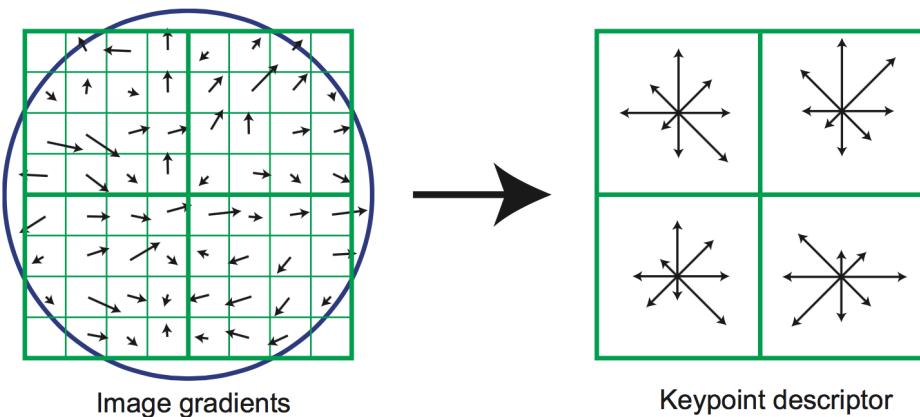
$$\theta(x, y) = \tan^{-1}((L(x, y+1) - L(x, y-1))/(L(x+1, y) - L(x-1, y))) \quad (4.2)$$

Następnie na podstawie wartości orientacji tworzony jest histogram, który posiada 36 przedziałów przez co pokrywa zakres orientacji od 0 do 360 stopni. Każdy piksel dodany do histogramu jest ważony poprzez użycie jego wartości gradientu i wartości okrągłego okna filtra Gaussa (piksele znajdujące się dalej od cechy mają mniejszą ważność niż te znajdujące się bliżej). Orientacja odpowiadająca najwyższej wartości w histogramie jest wybrana jako orientacja cechy. Dodatkowo każda orientacja, która posiada wartość większą niż 80% maksymalnej wartości w histogramie jest zapisywana jako nowa cecha obrazu.

6. Generacja deskryptorów cech

Deskryptor cechy musi ją charakterystycznie określić i być łatwy do obliczenia. SIFT przedstawia cechę jako 128 elementowy wektor, który został stworzony poprzez wykonanie następujących kroków:

1. Tworzone jest otoczenie 16 na 16 pikseli wokół cechy.
2. Dla każdego piksela w otoczeniu cechy jest liczona wartość gradientu oraz orientacja zgodnie ze wzorami 4.1 i 4.2. Każdą orientację piksela w otoczeniu odejmujemy od orientacji cechy obliczonej w poprzednim punkcie, przez co algorytm SIFT jest niezależny od rotacji obrazu.
3. Wartości gradientu oraz orientacji są ważone przez okrągłe okno filtru Gaussa i są następnie gromadzone poprzez utworzenie histogramu orientacji, który określa informacje w regionach 4 na 4 piksele. Rysunek 4.6 przedstawia proces kumulacji wartości gradientu oraz orientacji poprzez wykorzystanie histogramu.



Rys. 4.6: [10] **Lewy obraz:** Wartości orientacji oraz gradientu z uwzględnieniem okrągłego okna filtru Gaussa. **Prawy obraz:** Wartości z obszarów 4 na 4 piksele są gromadzone w histogramie, który posiada 8 przedziałów. Wartości orientacji z lewego obrazu, które wchodzą do jednego przedziału histogramu są sumowane (np. piksel z orientacją 12 stopni wchodzi do przedziału 0 - 44 stopni histogramu i jego wartość jest dodawana z każdą orientacją 0 - 44 stopni).

4. Prawy obraz z rysunku 4.6 przedstawia deskryptor z obszaru 8 na 8 pikseli, który posiada 32 elementy (4 obszary, każdy posiada 8 orientacji). Autor algorytmu SIFT podaje, że najlepsze rezultaty można uzyskać poprzez policzenie deskryptora z obszaru 16 na 16 pikseli wokół cechy, wtedy deskryptor posiada 128 elementów (16 obszarów, każdy posiada 8 orientacji).

5. Na koniec 128 elementowy wektor jest normalizowany.

Sposób wykorzystania algorytmu SIFT w programie

Biblioteka OpenCV posiada zaimplementowany algorytm SIFT i jego użycie sprowadza się do paru linii kodu:

```
SIFT featuresFinder = SIFT::SIFT(...);  
featuresFinder(...);
```

Konstruktor `SIFT::SIFT(...)` tworzy nam obiekt `SIFT`, który posiada funkcjonalność algorytmu SIFT. Do konstruktora `SIFT::SIFT(...)` podajemy parametry algorytmu tj. ilość najlepszych cech do zwrócenia, ilość oktaw do wykorzystania, progi do filtrowania wykrytych cech (sekcja 4.3.1) oraz siłę filtru Gaussa. Operator `featuresFinder(...)` przyjmuje obraz w którym mają zostać znalezione cechy i zwraca miejsca wystąpienia tych cech oraz ich deskryptory. Więcej informacji na temat obiektu `SIFT` znajduje się w oficjalnej dokumentacji OpenCV⁴.

4.3.2 Dopasowanie wyekstrahowanych cech

Cecha jest uważana za dopasowaną do drugiej cechy, jeżeli deskryptory (w przypadku SIFT 128 elementowy wektor) tych cech są do siebie podobne. W świecie rzeczywistym bardzo rzadko zdarza się, że deskryptory są do siebie identyczne, dlatego algorytmy zwracają zazwyczaj deskryptor, który jest najbardziej podobny. Trudność problemu dopasowania wyekstrahowanych cech jest uzależniony od ilości wykrytych cech w każdym z obrazów. Im więcej cech do dopasowania tym problem staje się bardziej kosztowny obliczeniowo. W przypadku problemu niniejszej pracy mamy do czynienia z obrazami o rozmiarze 240 na 240 pikseli z czego w większości zbiorów danych obrazy nakładają się w 10% szerokości obrazu, czyli algorytm SIFT szuka cech w obszarze 24 na 240 pikseli. Jest to obszar na tyle mały, że ilość cech wyekstrahowanych przez SIFT jest niewielka. Dzięki temu w niniejszej pracy w procesie rejestracji kafelków każda cecha jednego obrazu jest porównywana z każdą cechą obrazu dopasowywanego. Jeżeli w jednym kafelku wykryto n cech, a w drugim m cech to algorytm wykonuje nm dopasowań. Taki mechanizm dopasowania nazywa się *Brute Force Matching* i tak jak algorytm SIFT również jest zaimplementowany w OpenCV. Użycie *Brute Force Matching* w OpenCV tak jak w przypadku SIFT sprowadza się do kilku linii kodu:

⁴http://docs.opencv.org/modules/nonfree/doc/feature_detection.html

```
BFMatcher bruteForceMatcher;  
bruteForceMatcher.match(...);
```

Obiekt BFMatcher posiada funkcjonalność, która pozwala na dopasowanie każdego deskryptora z jednego zbioru do każdego deskryptora z drugiego zbioru. Funkcja `match(...)` przyjmuje deskryptory z dwóch zbiorów i zwraca dopasowania. Więcej informacji na temat obiektu BFMatcher znajduje się w oficjalnej dokumentacji OpenCV⁵.

4.3.3 Filtrowanie dopasowań

Filtrowanie na podstawie wiedzy dziedzinowej

RANSAC

4.4 Rejestracja kafelków poprzez wykrycie położeń naczyń krwionośnych w kafelkach

4.5 Estymacja macierzy transformacji pomiędzy kafelkami

4.6 Globalna rejestracja kafelków

4.7 Łączenie kafelków

⁵http://docs.opencv.org/modules/features2d/doc/common_interfaces_of_descriptor_matchers.html

5

Wyniki eksperymentów obliczeniowych

6

Podsumowanie i wnioski końcowe

Bibliografia

- [1]Herbert Bay, Andreas Ess, Tinne Tuytelaars i Luc Van Gool. „Speeded-Up Robust Features (SURF)“. W: *Comput. Vis. Image Underst.* 110.3 (czer. 2008), s. 346–359.
- [2]Jeffrey S. Beis i David G. Lowe. „Shape Indexing Using Approximate Nearest-Neighbour Search in High-Dimensional Spaces“. W: *Proceedings of the 1997 Conference on Computer Vision and Pattern Recognition (CVPR '97)*. CVPR '97. Washington, DC, USA: IEEE Computer Society, 1997, s. 1000–.
- [3]Matthew Brown i David G. Lowe. „Automatic Panoramic Image Stitching Using Invariant Features“. W: *Int. J. Comput. Vision* 74.1 (sierp. 2007), s. 59–73.
- [4]Peter J. Burt i Edward H. Adelson. „A Multiresolution Spline with Application to Image Mosaics“. W: *ACM Trans. Graph.* 2.4 (paź. 1983), s. 217–236.
- [5]J Canny. „A Computational Approach to Edge Detection“. W: *IEEE Trans. Pattern Anal. Mach. Intell.* 8.6 (czer. 1986), s. 679–698.
- [6]A F Fercher, W Drexler, C K Hitzenberger i T Lasser. „Optical coherence tomography - principles and applications“. W: *Reports on Progress in Physics* 66.2 (2003), s. 239.
- [7]Welzel J. „Optical coherence tomography in dermatology: a review.“ W: *Skin. Res. Technol.* 7.1 (2001).
- [8]Martin F. Kraus, Benjamin Potsaid, Markus A. Mayer, Ruediger Bock, Bernhard Baumann, Jonathan J. Liu, Joachim Hornegger i James G. Fujimoto. „Motion correction in optical coherence tomography volumes on a per A-scan basis using orthogonal scan patterns“. W: *Biomed. Opt. Express* 3.6 (2012), s. 1182–1199.
- [9]Ivan Laptev, Barbara Caputo, Christian Schüldt i Tony Lindeberg. „Local Velocity-adapted Motion Events for Spatio-temporal Recognition“. W: *Comput. Vis. Image Underst.* 108.3 (grud. 2007), s. 207–229.
- [10]David G. Lowe. „Distinctive Image Features from Scale-Invariant Keypoints“. W: *Int. J. Comput. Vision* 60.2 (list. 2004), s. 91–110.
- [11]D. Markl, G. Hanneschläger, M. Leitner, S. Sacher, D. Koller i J. Khinast. *A device and a method for monitoring a property of a coating of a solid dosage form during a coating process forming the coating of the solid dosage form*. EP Patent App. EP20,140,166,452. 2014.

- [12] Andrew M. Rollins, Rujchai Ung-arunyawee, Amitabh Chak, Richard C. K. Wong, Kenji Kobayashi, Michael V. Sivak i Joseph A. Izatt. „Real-time in vivo imaging of human gastrointestinal ultrastructure by use of endoscopic optical coherence tomography with a novel efficient interferometer design“. W: *Opt. Lett.* 24.19 (1999), s. 1358–1360.
- [13] Stephen Se, David G. Lowe i James J. Little. „Vision-based Mobile Robot Localization And Mapping using Scale-Invariant Features.“ W: *ICRA*. IEEE, 2001, s. 2051–2058.
- [15] Richard Szeliski. *Image Alignment and Stitching: A Tutorial*. Spraw. tech. MSR-TR-2004-92. Microsoft Research, 2004, s. 89.
- [16] P. Thévenaz i M. Unser. „User-Friendly Semiautomated Assembly of Accurate Image Mosaics in Microscopy“. W: *Microscopy Research and Technique* 70.2 (2007), s. 135–146.
- [17] W.J. Walecki i P. Van. *Determining thickness of slabs of materials by inventors*. US Patent 7,116,429. 2006.

Strony internetowe

- [@14] James Strong. *Retinal OCT Imaging*. 2011. URL: <http://www.opsweb.org/?page=RetinalOCT>.

Spis rysunków

2.1	Lewy obraz: Dwuwymiarowy przekrój siatkówki oka (B-skan). Obraz został uzyskany poprzez połączenie jednowymiarowych A-skanów, które zawierają informację o strukturze tkanki w głąb siatkówki oka. Prawy obraz: Angiografia siatkówki oka uzyskana dzięki przetworzeniu danych z OCT.	3
2.2	Kolejne etapy działania metody OCT. (1) - Etap początkowy. (2) - Źródło wyemitowało wiązkę światła. (3) - Fala rozdzieliła się za pomocą interferometru na wiązkę referencyjną (skierowaną na lustro referencyjne) oraz na wiązkę próbki (skierowaną na tkankę). (4) - Wiązki po odbiciu od lustra referencyjnego i tkanki ponownie łączą się za pomocą interferometru. W tej części występuje zjawisko interferencji, które jest zarejestrowane przez detektor.	4
2.3	Kolejne etapy badania głębszych warstw tkanki dzięki przesuwaniu lustra referencyjnego. Aktualna pozycja lustra referencyjnego jest zaznaczona kolorem czarnym, natomiast aktualnie badana warstwa tkanki jest zaznaczona kolorem czerwonym.	5
2.4	[8] Lewy obraz: Pojedynczy A-skan wykonany w głąb tkanki. Średnioobraz: Otrzymany B-skan poprzez połączenie A-skanów. Prawy obraz: Trójwymiarowy obraz tkanki stworzony poprzez połączenie B-skanych.	6
2.5	Trzy sąsiadujące B-skany na postawie których powstaje jeden obraz przepływowego.	8
3.1	Mozaika OCT stworzona z połączenia angiograficznych obrazów OCT.	11
3.2	Zbiór transformacji dwuwymiarowych dla przykładowego angiograficznego obrazu OCT.	12
3.3	Przykładowe rozmieszczenie kafelków na finalnej mozaice. Kafelek referencyjny został wyróżniony przerywaną linią.	14
3.4	Przykładowa korejestracja dwóch kafelków ze wspólnym obszarem nałożenia.	14
3.5	Lewy obraz: Wynik zaaplikowanego algorytmu Canny na kolorowym obrazie z prawej strony. Prawy obraz: Oryginalny obraz.	17

3.6	Dwa obrazy, w których zostały wykryte po trzy cechy (na rysunku oznaczone jako gwiazdy). Algorytm szukający dopasowań wykrył dopasowania (na rysunku oznaczone jako strzałki) pomiędzy cechami: 1 i 4, 2 i 5, 3 i 5. Z wiedzy dziedzinowej wiemy, że obrazy nachodzą na siebie (tak jak angiograficzne obrazy OCT). Patrząc na dopasowania pomiędzy obrazami można zauważać, że jedno z nich nie pasuje do reszty (strzałka narysowana przerywaną linią). Zakładając, że możemy tylko obracać i przesuwać obrazy niemożliwa jest estymacja macierzy transformacji biorąc pod uwagę wszystkie dopasowania. Niepoprawne dopasowanie trzeba odrzucić by otrzymać poprawną macierz transformacji.	17
4.1	Cztery kafelki rozmieszone zgodnie z ich pozycją w docelowej mozaice. Na każdym obrazie białym przeźroczystym prostokątem zaznaczony jest oszacowany obszar nałożenia sąsiadujących kafelków.	23
4.2	Przestrzeń <i>scale-space</i> dla przykładowego obrazu. Skala rośnie wraz z siłą filtra Gaussa. Obrazy tej samej wielkości tworzą oktawę.	24
4.3	Przykład wykrycia krawędzi przez LoG. Czarny wykres przedstawia jednowymiarowy sygnał z obrazu. Miejsce przecięcia się z osią drugiej pochodnej tzw. <i>zero-crossing</i> wskazuje miejsce krawędzi w oryginalnym obrazie.	25
4.4	[10] Poprzez proces odjęcia kolejnych zdjęć z jednej oktawy otrzymujemy obrazy DoG, które doskonale odzwierciedlają LoG.	26
4.5	[10] Maksima i minima wykrywane są za pomocą porównania wartości aktualnie badanego piksela (zaznaczony czarnym krzyżykiem) z jego 26 sąsiadami powstałymi poprzez złączenie kolejnych obrazów DoG z jednej oktawy. Cecha jest wykrywana w miejscu aktualnie badanego piksela jeżeli wartość tego piksela jest największa lub najmniejsza spośród 26 sąsiadów.	26
4.6	[10] Lewy obraz: Wartości orientacji oraz gradientu z uwzględnieniem okrągłego okna filtra Gaussa. Prawy obraz: Wartości z obszarów 4 na 4 piksele są gromadzone w histogramie, który posiada 8 przedziałów. Wartości orientacji z lewego obrazu, które wchodzą do jednego przedziału histogramu są sumowane (np. piksel z orientacją 12 stopni wchodzi do przedziału 0 - 44 stopni histogramu i jego wartość jest dodawana z każdą orientacją 0 - 44 stopni).	28

Spis tablic

