

POLITECHNIKA WARSZAWSKA

WYDZIAŁ MECHATRONIKI

Rozprawa doktorska

mgr inż. Jan Klimaszewski

Metoda samokalibracji układu kamer: wizyjnej
i termowizyjnej na podstawie dopasowania obrazów

Promotor:
dr hab. inż. Barbara Putz, prof. PW

WARSZAWA, 2019

Praca częściowo finansowana z umowy między PIT-RADWAR S.A. a Politechniką Warszawską, Wydział Mechatroniki pt. "Projekt i realizacja adaptacji układu fuzji obrazów IR-TV na nową platformę sprzętową", 2015-2018, z grantu Dziekana Wydziału Mechatroniki pt. "Algorytmy wspomagające dopasowanie i fuzję obrazów wizyjnych i termowizyjnych", realizowanego na Wydziale Mechatroniki w roku 2017 oraz z grantu rozwojowego NCBR pt. "System fuzji obrazów wizyjnych i termowizyjnych dla potrzeb monitorowania otoczenia w czasie rzeczywistym", realizowanego na Wydziale Mechatroniki w latach 2009-2012.

Streszczenie

W rozprawie przedstawiono nowy sposób rozwiązania problemu kalibracji układu kamer dalekiego zasięgu: wizyjnej i termowizyjnej. Problem ten powstał w trakcie prac badawczo-rozwojowych nad układem fuzji obrazów dla polskiego przemysłu obronnego i dotyczy głowicy optoelektronicznej zestawu przeciwlotniczego Poprad. Rozwiązanie problemu kalibracji z zastosowaniem klasycznych metod z wykorzystaniem wzorca kalibracyjnego okazało się niepraktyczne w warunkach poligonowych. W rozprawie zaproponowano algorytm samokalibracji układu kamer na podstawie automatycznie wyznaczanego dopasowania obrazów z tych kamer.

Wykazano, że kalibrację można podzielić na dwa etapy. Pierwszy etap ma na celu ustalenie parametrów skali obrazu termowizyjnego względem wizyjnego. W etapie drugim ustalana jest zależność wiążąca parametry dopasowania odpowiadających sobie obiektów obrazu wizyjnego i termowizyjnego. Samokalibracja realizowana jest na podstawie automatycznego dopasowania obrazów w sekwencji wideo. Zaproponowane rozwiązanie stanowi nowość w zakresie kalibracji układu kamer wizyjnej i termowizyjnej.

W zakresie automatycznego dopasowania obrazów multimodalnych, w rozprawie zaproponowano nowe podejście do rozwiązania problemu kalibracji obrazów z wykorzystaniem metody korelacji fazowej do wyznaczenia wartości dopasowań dla wielu obiektów jednocześnie widocznych na obrazach z kamery wizyjnej i termowizyjnej. Istotną zaletą zaproponowanej metody jest fakt, że nie wymaga ona przyporządkowania odpowiadających sobie obiektów na obu obrazach.

W celu weryfikacji opracowanej metody dokonano jej implementacji i przetestowano jej działanie na sekwencjach obrazów z głowicy Poprad oraz na ogólnodostępnej bazie obrazów KAIST. W wyniku testów potwierdzono skuteczność metody i przedstawiono warunki jej stosowania w zakresie obserwowanych scen.

Zastosowanie samokalibracji do układu kamer wizyjnej i termowizyjnej dalekiego zasięgu oraz nowe ujęcie metody korelacji fazowej i wykorzystanie wyznaczonego za jej pomocą dopasowania w procesie kalibracji są oryginalnym wkładem autora w dziedzinie przetwarzania obrazów. Testy przeprowadzone dla bazy danych KAIST, gdzie ani układ kamer, ani ich wykorzystanie nie spełniają przyjętych w rozprawie założeń, świadczą o tym, że opracowana metoda samokalibracji może znaleźć zastosowanie nie tylko w przemyśle obronnym, ale też na przykład w procesie kalibracji układu kamer multimodalnych pojazdów autonomicznych.

Słowa kluczowe: *kalibracja układu kamer, samokalibracja, obrazy multimodalne, obrazy wizyjne i termowizyjne, dopasowanie obrazów, korelacja fazowa*

Abstract

Self-calibration of thermal and visible camera rig based on image registration

In this dissertation, the problem of calibration of the long-range thermal and visible camera rig is addressed. This problem arose during the research and development works on the image fusion system for the Polish defence industry. It concerns the optoelectronic head of Poprad anti-aircraft system. The solution to the calibration problem using classical methods using a calibration pattern has proved impractical in field conditions. The dissertation proposed an algorithm for self-calibration of the camera set based on automatically determined image registration.

It has been shown that the calibration can be divided into two stages. The first step is to determine the parameters of the scale of the thermal image relative to the visible image. In the second stage, the dependence constraining the image registration parameters of the corresponding thermal and visible image objects is determined. Self-calibration is carried out based on the automatic image registration in the video sequence. The proposed solution is a novelty in the field of calibration of video and infrared camera systems.

In the field of multimodal image registration, the dissertation proposes new approach to solving the problem of image calibration using a phase correlation method. It was used to determine the registration parameters for many objects simultaneously visible on images from a video and thermal imaging camera. An important advantage of the proposed method is the fact that it does not require assigning corresponding objects on both images.

In order to verify the developed method, it was implemented and tested on the sequences of images from the Poprad head and on the public database of KAIST images. As a result of the tests, the effectiveness of the method was confirmed and the conditions of its application in the scope of observed scenes were presented.

The application of self-calibration to the system of long-range vision and thermovision cameras, as well as the new approach to the phase correlation method are the original contribution of the author in the field of image processing. Tests carried out for the KAIST database, where neither the camera system nor its use meet the assumptions adopted in the dissertation, prove that the self-calibration method can be applied not only in the defence industry, but also in the process of calibration of a multi-modal camera system for autonomous vehicles.

Key words: *camera rig calibration, self-calibration, multimodal images, thermal and visible images, image registration, phase correlation*

Spis treści

| | |
|---|-----------|
| Wykaz oznaczeń | 9 |
| 1 Wprowadzenie | 11 |
| 1.1 Wstęp | 11 |
| 1.2 Przyjęte założenia | 13 |
| 1.3 Opis problemu, cel i teza pracy | 14 |
| 1.4 Układ pracy | 17 |
| 2 Przegląd metod kalibracji układu dwóch kamer | 19 |
| 2.1 Ogólny model układu dwóch kamer | 20 |
| 2.1.1 Parametry wewnętrzne kamer | 20 |
| 2.1.2 Parametry zewnętrzne układu kamer - macierz zasadnicza | 22 |
| 2.1.3 Macierz fundamentalna | 23 |
| 2.2 Ogólny przegląd metod samokalibracji | 24 |
| 2.3 Przegląd metod kalibracji układu kamer: wizyjnej i termowizyjnej | 30 |
| 2.4 Podsumowanie | 31 |
| 3 Przegląd metod automatycznego dopasowania obrazów | 33 |
| 3.1 Model dopasowania obrazów | 34 |
| 3.2 Metody wyznaczania dopasowania obrazów multimodalnych | 38 |
| 3.3 Metody wyznaczania dopasowania obrazów multimodalnych z zastosowaniem korelacji fazowej | 42 |
| 3.4 Podsumowanie | 44 |
| 4 Opracowana metoda samokalibracji układu kamer: wizyjnej i termowizyjnej | 47 |
| 4.1 Podstawy teoretyczne metody samokalibracji | 48 |
| 4.1.1 Przyjęty model układu kamer | 48 |
| 4.1.2 Podział modelu na dwie części | 51 |
| 4.1.3 Etap pierwszy kalibracji | 54 |
| 4.1.4 Etap drugi kalibracji | 57 |
| 4.2 Dopasowanie automatyczne obrazów metodą korelacji fazowej | 59 |

| | | |
|----------|--|------------|
| 4.2.1 | Dopasowanie obrazów monomodalnych | 60 |
| 4.2.2 | Dopasowanie dwóch obiektów na obrazach monomodalnych | 63 |
| 4.2.3 | Dopasowanie wielu obiektów na obrazach monomodalnych | 67 |
| 4.2.4 | Dopasowanie wielu obiektów na obrazach multimodalnych | 71 |
| 4.2.5 | Ogólny algorytm dopasowania | 74 |
| 4.2.6 | Filtracja krawędziowa obrazów | 75 |
| 4.2.7 | Wnioski i podsumowanie | 80 |
| 4.3 | Procedura samokalibracji | 80 |
| 4.3.1 | Etap pierwszy | 81 |
| 4.3.2 | Etap drugi | 82 |
| 4.3.3 | Procedura kalibracji | 83 |
| 4.3.4 | Opracowana metoda na tle innych publikacji | 86 |
| 5 | Przeprowadzone badania | 91 |
| 5.1 | Stanowisko badawcze i bazy danych obrazów | 91 |
| 5.1.1 | Głowica kamer | 91 |
| 5.1.2 | Sekwencje testowe | 92 |
| 5.2 | Wyniki testów samokalibracji układu kamer | 99 |
| 5.2.1 | Sposób przeprowadzenia testów samokalibracji | 99 |
| 5.2.2 | Zasady opisu wyników działania samokalibracji | 102 |
| 5.2.3 | Wyniki I etapu - wprowadzenie | 103 |
| 5.2.4 | Wyniki I etapu dla sekwencji wygenerowanych komputerowo | 103 |
| 5.2.5 | Wyniki I etapu dla sekwencji testowych Poprad | 103 |
| 5.2.6 | Wyniki I etapu dla sekwencji testowych KAIST | 110 |
| 5.2.7 | Wyniki II etapu - wprowadzenie | 112 |
| 5.2.8 | Wyniki II etapu dla sekwencji wygenerowanych komputerowo | 115 |
| 5.2.9 | Wyniki II etapu dla sekwencji testowych Poprad | 115 |
| 5.2.10 | Wyniki II etapu dla sekwencji testowych KAIST | 121 |
| 5.2.11 | Podsumowanie | 124 |
| 6 | Podsumowanie i wnioski | 133 |
| | Bibliografia | 137 |

Wykaz oznaczeń

c_x, c_y przesunięcie osi optycznej względem środka matrycy CCD kamery

$t(\mathbf{x})$ funkcja przesunięcia pikseli \mathbf{x} obrazu

X, Y, Z współrzędne rzeczywistego punktu otoczenia

\exists kwantyfikator szczegółowy

D zbiór parametrów określający dopasowanie

E macierz zasadnicza układu kamer

F macierz fundamentalna układu kamer

\mathbf{g}_i obraz o numerze i

R macierz opisująca obrót w euklidesowej przestrzeni trójwymiarowej jednej kamery względem drugiej

K macierz parametrów wewnętrznych kamery

\mathbf{X}_i i -ty rzeczywisty punkt otoczenia w przestrzeni 3D

\mathbf{x}_i i -ty punkt obrazu w przestrzeni 2D

$\mathcal{F}(\cdot)$ transformacja Fouriera

$\mathcal{F}^{-1}(\cdot)$ odwrotna transformacja Fouriera

\hat{w} wartość zmierzona, estymacja wielkości w

$d(\mathbf{x})$ funkcja dopasowania pikseli \mathbf{x} obrazu

$\det(\cdot)$ wyznacznik macierzy

f_x, f_y ogniskowa kamery w kierunku X i Y

$s(\mathbf{g}_1, \mathbf{g}_2)$ funkcja podobieństwa obrazów $\mathbf{g}_1, \mathbf{g}_2$

t_x, t_y, t_z współrzędne przesunięcia w euklidesowej przestrzeni trójwymiarowej jednej kamery
względem drugiej

$trace(\cdot)$ ślad macierzy

u, v współrzędne współczynników transformaty Fouriera obrazu

x, y współrzędne punktu obrazu

Rozdział 1

Wprowadzenie

1.1 Wstęp

Wraz z postępującym rozwojem techniki coraz powszechniejsze stają się systemy monitorowania i nadzoru wizyjnego, w szczególności stosowanie wielu kamer przy obserwacji tego samego obszaru. Jeśli obrazy rejestrowane są przez sensory operujące w różnych fragmentach widma promieniowania, przykładowo w świetle widzialnym i w promieniowaniu podczerwonym, otrzymujemy obrazy multimodalne. W celu usprawnienia wielokamerowego monitorowania często stosuje się operację fuzji obrazów. Fuzja obrazów, w szczególności obrazów multimodalnych, umożliwia uzyskanie jednego obrazu wynikowego zawierającego najistotniejsze cechy (informacje) obrazów składowych. Przykładowo fuzja obrazów z rys. 1.1a i 1.1b ukazuje na obrazie wynikowym szczegóły drzew widoczne za pośrednictwem kamery telewizyjnej i rozgrzane silniki samolotu odrzutowego widoczne tylko na obrazie z kamery termowizyjnej (rys. 1.2).

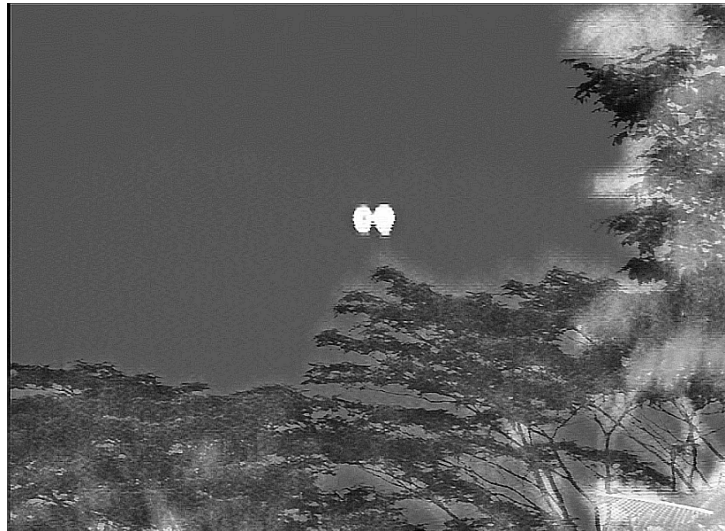


(a) Obraz z kamery wizyjnej.



(b) Obraz z kamery termowizyjnej.

Rysunek 1.1: Obrazy z kamery wizyjnej i termowizyjnej.



Rysunek 1.2: Wynik fuzji obrazów z rysunków 1.1a oraz 1.1b.

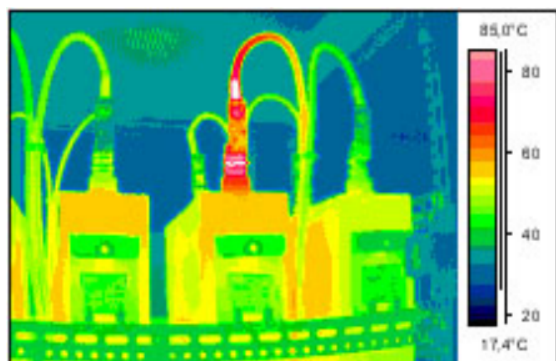
Fuzja obrazów agreguje dane z wielu źródeł informacji i przedstawia je w postaci połączonej. Prowadzi to do poprawienia jakości informacji przedstawionej na obrazie wynikowym. W związku z ograniczeniami percepcyjnymi człowieka, któremu trudno jest obserwować wiele obrazów naraz, fuzja stanowi dla niego bardzo dobrą reprezentację monitorowanego otoczenia. W celu przeprowadzenia fuzji online, czyli w trakcie działania systemu wizyjnego, konieczne jest zazwyczaj jednorazowe przeprowadzenie kalibracji układu kamer oraz ciągłe ustalenie dopasowania obrazów online. Zarówno jednorazowa kalibracja głowicy, jak i dopasowywanie obrazów z głowicy w trybie ciągłym, są skomplikowanymi operacjami, które wpływają znacząco na jakość fuzji.

Badania nad kalibracją, dopasowaniem i fuzją obrazów multimodalnych zostały rozpoczęte przez autora rozprawy w ramach projektu rozwojowego w odpowiedzi na zapotrzebowanie przemysłu obronnego. Rys. 1.1a, 1.1b i 1.2 przedstawiają obrazy z głowicy raketowego zestawu przeciwlotniczego Poprad¹² wyposażonego w system monitorowania nieba w celu wykrycia samolotów. Opisane w rozprawie metody przetwarzania obrazów opracowano w trakcie kontynuacji tych badań. Wykorzystanie wyników badań może być jednak znacznie bardziej uniwersalne. Przykładowo fuzja danych z kamery światła dziennego i termowizyjnej znajduje zastosowanie w diagnostyce maszyn przemysłowych (rys. 1.3a) i inspekcjach cieplnych budynków (rys. 1.3b), umożliwiając wykrycie nieszczelności w izolacji cieplnej lub wykrycie elementów o niebezpiecznie wysokiej temperaturze. Może zostać również wykorzystana do inspekcji linii wysokiego napięcia [1], czy nadzoru bezpieczeństwa przez lokalizację ukrytych przed kontrolą niebezpiecznych przedmiotów, na przykład na lotniskach [2]. Fuzja może służyć dla potrzeb ochrony mienia czy ratownictwa pozwalając na odnalezienie ludzi w niesprzyjających warunkach pogodowych lub przy niedostatecznym oświetleniu. Z uwagi na

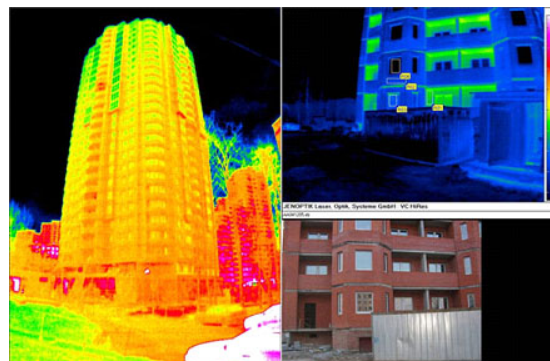
¹[https://pl.wikipedia.org/wiki/Poprad_\(przeciwlotniczy_zestaw_raketowy\)](https://pl.wikipedia.org/wiki/Poprad_(przeciwlotniczy_zestaw_raketowy))

²<https://www.defence24.pl/zestaw-przeciwlotniczy-poprad-i-jego-miejsce-w-tarczy-polski>

zastosowanie w celu monitorowania systemu tego typu powinny na ogół działać w czasie rzeczywistym.



(a) Diagnostyka urządzeń przemysłowych, źródło <http://www.tediko.cz/index.php?sub=04cz&lang=cz&p=0407cz>



(b) Inspekcja ocieplenia budynków, źródło http://ecs39.ru/index/teplovizionnoe_obsledovanie_zdaniy_i_sooruzhenij/0-20

Rysunek 1.3: Zastosowanie termowizji w diagnostyce i inspekcjach.

1.2 Przyjęte założenia

Zakłada się, że opracowane w ramach rozprawy algorytmy powinny stanowić rozwinięcie istniejącego systemu monitorowania otoczenia za pomocą głowicy optoelektronicznej raketowego zestawu przeciwlotniczego Poprad stosowanego w polskim przemyśle obronnym. W związku z tym metody przetwarzania obrazów dostosowane do tego systemu muszą być możliwe do przeprowadzenia w warunkach poligonowych. System wizyjny powinien realizować fuzję obrazów w celu prezentacji wynikowego obrazu operatorowi.

W zakresie sensoryki istotnym elementem tego systemu jest obrotowa głowica optoelektroniczna z układem dwóch kamer: wizyjnej i termowizyjnej. Na rys. 1.4 przedstawiono wygląd tej głowicy. Głowica może obracać się wokół swojej osi pionowej i wokół osi poziomej umieszczonej prostopadle do osi optycznych kamer. Głowica jest mechanicznie sztywna - żadna z kamer nie zmienia swojego położenia względem głowicy w trakcie pracy systemu. Kamery zwrócone są w tym samym kierunku z zachowaniem równoległości osi optycznych.

Obydwie zastosowane kamery są dalekiego zasięgu. Oznacza to, że obrazy pochodzące z tych kamer są wyraźne dla obiektów umieszczonych dopiero powyżej kilkudziesięciu/kilkuset metrów od głowicy. Układy optyczne kamer zestawu Poprad są wysokiej jakości, dlatego dystorsje tych kamer można pominąć w procesie kalibracji. Kamery dostarczają obraz bez zniekształceń.

Każda kamera z głowicy posiada możliwość akwizycji obrazów przy dwóch różnych kątach widzenia (elektronicznie sterowana zmiana optyki kamery). Praktyczne wykorzystanie głowicy

obejmuje ustawienie obydwu kamer na szeroki kąt, lub obydwu na wąski kąt postrzegania. Inne konfiguracje ustawienia optyki występują jedynie jako stany przejściowe.

W głowicy zastosowano kamery analogowe, które udostępniają obrazy w formacie CCIR (625 linii) o rozdzielczościach 756×581 pikseli obrazu TV oraz 384×256 pikseli obrazu IR. Przykładowe obrazy przedstawiono na rys. 1.1a i 1.1b.



Rysunek 1.4: Wyrzutnia rakiet wyposażona w głowicę z kamerami IR i TV.

Obrazy z głowicy kamer zestawu Poprad są wykorzystywane w przemyśle obronnym do śledzenia obiektów będących elementami monitorowanego otoczenia. Obraz z kamery TV wykorzystywany jest w celu wyznaczenia odchyłki w układzie sterującym ruchem głowicy. W rozprawie zakłada się, że w procesie kalibracji obraz wizyjny powinien być traktowany jako referencyjny - nie można poddać go żadnym transformacjom.

W odniesieniu do komputerowej jednostki obliczeniowej, która ma realizować zadania systemu wizyjnego, zakłada się wykorzystanie mobilnej platformy obliczeniowej Jetson TK1. Badania w zakresie wyboru odpowiedniej jednostki obliczeniowej rozpoczęto w ramach grantu NCBR pod tytułem „System fuzji obrazów wizyjnych i termowizyjnych dla potrzeb monitorowania otoczenia w czasie rzeczywistym”, który został zrealizowany w na Wydziale Mechatroniki w latach 2009-2012. Minikomputer Jetson TK1 został wybrany w wyniku tych i późniejszych prac oraz w związku z wymogami technicznymi przemysłu obronnego (obejmującymi m.in. moc układu i wymuszającymi jego niewielkie rozmiary). W rezultacie przyjęcia takiego założenia prezentowane w rozprawie metody muszą być dostosowane do implementacji na układzie Jetson TK1.

1.3 Opis problemu, cel i teza pracy

Proces monitorowania otoczenia za pośrednictwem systemu wizyjnego można podzielić na dwa etapy. Na początku wykonywana jest kalibracja układu kamer obserwujących otoczenie.

Kalibracja przeprowadzana jest jednorazowo w trybie offline działania systemu. W wyniku tej procedury następuje wyznaczenie zestawu parametrów opisujących model układu kamer. Znajomość tych parametrów umożliwia realizację etapu drugiego, w którym system wizyjny zostaje przełączony w tryb online i wykonuje w trybie ciągłym dopasowanie i fuzję obrazów pochodzących z jego kamer. Warto podkreślić, że kalibracja układu kamer jest często konieczna nawet w przypadku znajomości danych katalogowych kamer, czy parametrów konstrukcji układu. Podane informacje są nieraz mało dokładne i wymagają dostrojenia.

Celem rozprawy jest opracowanie metody samokalibracji układu dwóch kamer: wizyjnej i termowizyjnej, wykonywanej offline i umożliwiającej monitorowanie otoczenia poprzez fuzję obrazów pobieranych w czasie rzeczywistym z obu kamer. Samokalibracja jest tu rozumiana jako kalibracja automatyczna bez wykorzystania wzorca kalibracyjnego. Sposób wyznaczania wybranych parametrów układu kamer w trakcie samokalibracji musi bazować na automatycznie wyznaczonych cechach obrazów pochodzących z obserwacji otoczenia.

Cel rozprawy został ustalony wobec założeń dotyczących aplikacji systemu wizyjnego - monitorowanie otoczenia przez głowicę kamer dalekiego zasięgu zestawu przeciwlotniczego Poprad. Proces kalibracji układu kamer jest skomplikowany i ma znaczny wpływ na dalsze działanie systemu monitorowania. Uzasadnione jest opracowanie metody samokalibracji, dzięki której kalibracja zostanie częściowo zautomatyzowana i uproszczona. Metody klasyczne rozwiązują zadanie kalibracji z wykorzystaniem tzw. wzorca kalibracyjnego - ułatwia to wykonanie kalibracji. W świetle przyjętych założeń zastosowanie wzorca kalibracyjnego jest bardzo utrudnione, co opisano szczegółowo w rozdziale 2. W takim przypadku uzasadnione jest wykorzystanie metod samokalibracji, które rozwiązują problem kalibracji bez ułatwienia w postaci zastosowania wzorca kalibracyjnego.

Realizację celu rozprawy, czyli opracowanie metody samokalibracji sprowadzono do dwóch zadań. Pierwsze z nich to ustalenie, które parametry opisujące układ kamer są konieczne do identyfikacji w celu wykonania fuzji obrazów. Drugie zadanie to określenie metody estymacji tych parametrów.

Wybór parametrów układu kamer wykonano w odniesieniu do docelowej aplikacji systemu wizyjnego - fuzji obrazów. Dobre wyniki fuzji można osiągnąć jedynie w przypadku, gdy obrazy składowe są do siebie dopasowane, tzn. odpowiadają sobie położeniem, rotacją, kształtem i skalą. W przypadku fuzji obrazów w trybie online, dopasowanie obrazów również musi być wykonywane w trybie online, czyli w czasie rzeczywistym rozumianym jako czas pomiędzy akwizycją kolejnych par obrazów z układu kamer. Wyznaczenie parametrów układu kamer nie jest jednak wystarczające do uzyskania dobrych wyników fuzji, ponieważ dla kamer, które obserwują otoczenie z różnych punktów widzenia, dopasowanie odpowiadających sobie na obrazach obiektów zależy od fizycznej odległości obserwowanych obiektów od układu kamer. Odpowiadające sobie na obrazach obiekty rozumiane są jako fragmenty obrazów reprezentujące ten sam fizyczny obiekt w monitorowanym otoczeniu. Obiektom znajdującym się w różnej odległości od układu kamer należy przypisać różne wartości dopasowania

w trybie online. Znajomość funkcji opisującej dopasowanie w trybie online umożliwia uproszczenie zadania wyznaczenia dopasowania online i pośrednio pozwala na uproszczenie procesu przeprowadzenia fuzji obrazów. Przykładowo jeśli funkcja dopasowania zostanie sparametryzowana, możliwe jest wyznaczenie dopasowania online w wyniku przeszukiwania najlepszej wartości jednego tylko parametru, a nie dwóch współrzędnych w osi X i Y obrazu. W rozprawie wykazano, że zależność opisująca dopasowanie obiektów w różnych odległościach od układu kamer może zostać opisana w postaci funkcji liniowej, której wykres jest określony w rozprawie jako *linia dopasowania*. Ponadto pokazano, że kalibrację można podzielić na dwa etapy. Na pierwszym etapie wyznacza się parametry układu kamer reprezentujące różnice w rozmiarze odpowiadających sobie obiektów, natomiast na drugim etapie wyznacza się parametry opisu linii dopasowania.

Sposób wyznaczenia tych parametrów stanowi drugie zadanie konieczne do rozwiązania w celu opracowania metody samokalibracji układu kamer wizyjnej i termowizyjnej. W opracowanej metodzie zaproponowano, że parametry wyznaczone na pierwszym etapie kalibracji będą estymowane na podstawie dopasowania w trybie kalibracji obrazów z chwili bieżącej do obrazów z chwili poprzedniej odpowiednio dla kamery wizyjnej i termowizyjnej. Dopasowanie obrazów pochodzących z kamer tego samego typu określa się jako dopasowanie obrazów monomodalnych. Na drugim etapie kalibracji odpowiednie parametry wyznaczone są na podstawie dopasowania w trybie kalibracji odpowiadających sobie obiektów na obrazach wizyjnym i termowizyjnym. Takie dopasowanie określa się jako dopasowanie obrazów multimodalnych.

Warto podkreślić, że algorytmy wykonujące się w trybie kalibracji nie muszą działać w czasie rzeczywistym. Jednakże, aby proces kalibracji mógł być nadzorowany przez człowieka sterującego głowicą obrotową Poprad oraz w celu zapewnienia płynności obrazów w sekwencji, wyznaczanie dopasowania w trybie kalibracji powinno cechować się możliwie krótkim czasem działania, bez znacznego opóźniania akwizycji nowych obrazów wejściowych. Aby spełnić to założenie oraz założenia związane z przyjętą jednostką obliczeniową, metody dopasowania powinny cechować się niską złożonością obliczeniową.

Podsumowując, w rozprawie opracowano nową metodę samokalibracji bazującą na automatycznie wyznaczanym dopasowaniu obrazów w trybie kalibracji.

W przypadku automatycznego zadania dopasowania obrazów największą trudność przedstawia automatyczne wyznaczenie dopasowania obrazów multimodalnych: wizyjnego i termowizyjnego. Trudność problemu można tłumaczyć znacznymi różnicami w charakterystykach obrazowania kamer wizyjnej i termowizyjnej. W rozprawie przedstawiono sposób wykorzystania klasycznej metody korelacji fazowej do wyznaczenia w trybie kalibracji dopasowania wielu obiektów obserwowanych jednocześnie przez układ kamer. Zastosowanie takiego rozwiązania pozwoliło poprawić, w porównaniu do stosowania jednej wartości dopasowania, jakość działania opracowanej metody samokalibracji w zakresie wyznaczenia parametrów linii dopasowania.

W związku z tym proponuje się następującą tezę rozprawy:

Teza. *Możliwa jest samokalibracja układu kamer: wizyjnej i termowizyjnej na podstawie automatycznego dopasowania obrazów z tych kamer, wyznaczonego z wykorzystaniem metody korelacji fazowej dla wielu obiektów jednocześnie widocznych na tych obrazach.*

Na koniec warto podkreślić, że metodę samokalibracji zaprojektowano w ten sposób, aby wykorzystać najlepsze cechy korelacji fazowej bez negatywnych skutków związanych z jej wadami. Największymi zaletami korelacji fazowej jest szybkość działania metody oraz odporność na jednorodne zmiany intensywności pomiędzy dopasowywanymi obrazami. Największą wadą jej jest fakt, że wyznaczone w dziedzinie częstotliwości dopasowania trudno przyporządkować do konkretnych obszarów (obiektów) w dziedzinie obrazu. W związku z tym trudno uzyskać informację o odpowiadających sobie na obrazach obiektach, którą wykorzystuje większość metod samokalibracji opisywanych w literaturze. Metoda samokalibracji opracowana w ramach rozprawy została zaprojektowana tak, że nie potrzebuje takiego przyporządkowania, co stanowi jej istotną zaletę wobec wykorzystania korelacji fazowej.

W rozprawie opisano opracowaną metodę samokalibracji, którą przetestowano na przygotowanych sekwencjach obrazów testowych. Otrzymane wyniki potwierdzają skuteczność metody i prawdziwość przyjętej tezy.

1.4 Układ pracy

Pierwsza część rozprawy zawiera wprowadzenie teoretyczne do poruszanych zagadnień badawczych. W rozdziale 2 opisano stan badań dotyczący metod kalibracji i samokalibracji układu dwóch kamer. Na początku, w podrozdziale 2.1, przedstawiono ogólny matematyczny sposób opisu układu dwóch kamer. Następnie, w podrozdziałach 2.2 i 2.3 zamieszczono przegląd literatury ze szczególnym uwzględnieniem metod samokalibracji mogących mieć zastosowanie do realizacji przedstawionego powyżej celu rozprawy.

W rozdziale 3 opisano przegląd metod automatycznego dopasowania obrazów. Na początku przedstawiono przyjęty sposób opisu dopasowania obrazów. Dalsza treść rozdziału 3 prezentuje opisywane w literaturze metody dopasowania w zakresie dopasowania obrazów wizyjnego i termowizyjnego. Przedstawione metody stanowią istotne uzupełnienie zagadnień opisanych w odniesieniu do metod samokalibracji układu dwóch kamer.

Druga część rozprawy zawiera opis opracowanych przez autora rozwiązań. Wykazano tam prawdziwość przyjętej tezy. W rozdziale 4 przedstawiono sposób realizacji celu rozprawy, czyli opracowaną metodę samokalibracji układu kamer wizyjnej i termowizyjnej. W podrozdziale 4.1 przedstawiono, w jaki sposób model ogólny układu kamer dostosowano do przyjętych w rozprawie założeń oraz ustalono, które parametry opisujące układ kamer są konieczne do identyfikacji. W podrozdziale 4.2 opisano metody estymacji tych parametrów. Podrozdział

ten zawiera również uzasadnienie wykorzystania metody bazującej na korelacji fazowej do wyznaczenia dopasowania w trybie kalibracji dla wielu obiektów jednocześnie widocznych na obrazach z kamery wizyjnej i termowizyjnej. Ostatecznie w podrozdziale 4.3 przedstawiono nową metodę samokalibracji pozwalającą na wyznaczenie wybranych parametrów układu kamer: wizyjnej i termowizyjnej w celu wykonania fuzji obrazów.

Badania przeprowadzone w celu weryfikacji i oceny skuteczności opracowanej metody przedstawiono w rozdziale 5. W podrozdziale 5.1 opisano stanowisko badawcze i wykorzystane w trakcie testów bazy danych. W podrozdziale 5.2 przedstawiono i omówiono wyniki przeprowadzonych testów potwierdzające skuteczność metody w większości przypadków.

Rozdział 6 zawiera podsumowanie rozprawy oraz przedstawia możliwe kierunki dalszych prac w zakresie poruszanych w rozprawie zagadnień.

Rozdział 2

Przegląd metod kalibracji układu dwóch kamer

W bieżącym rozdziale zaprezentowany został przegląd najważniejszych metod, które mogą znaleźć zastosowanie do kalibracji układu dwóch kamer: wizyjnej i termowizyjnej. Na początku opisano model układu dwóch kamer. Są to podstawowe informacje, na tle których możliwe jest przedstawienie w dalszej części rozdziału zasad działania różnych metod kalibracji.

Standardowe podejście do kalibracji zakłada zastosowanie wzorca kalibracyjnego [3, 4, 5]. Wzorzec jest skonstruowany tak, że jego fizyczne wymiary są znane, a ich reprezentacja na obrazie jest łatwa do wykrycia. Zastosowanie wzorca ułatwia znacznie proces kalibracji. Przyjęte w rozprawie założenia (podane wstępnie w rozdziale 1.2) wiążą się w szczególności z dwoma utrudnieniami wykorzystania wzorca kalibracyjnego.

Po pierwsze zakłada się, że kamery będące elementem układu kamer są kamerami dalekiego zasięgu. Oznacza to, że obrazy dla obiektów umieszczonych blisko takiej kamery (poniżej kilkuset metrów) są rozmyte. Aby obraz wzorca miał wyraźne krawędzie, wzorzec kalibracyjny należałoby umieścić w znacznej odległości od układu kamer. Jednakże w takiej sytuacji konieczne byłoby skonstruowanie wzorca o znacznych rozmiarach - niewielki obiekt umieszczony w znacznej odległości od układu kamer o niskiej rozdzielczości byłby niewyraźny. Konstrukcja dużego wzorca jest niewskazana z uwagi na problemy związane ze stosowaniem takiego wzorca w warunkach polowych. Wzorzec musiałby być elementem wyposażenia zestawu Poprad.

Po drugie zakłada się, że układ kamer składa się z kamery wizyjnej oraz termowizyjnej. Konstrukcja wzorca kalibracyjnego dobrze widocznego w tych dwóch pasmach promieniowania elektromagnetycznego powoduje dodatkowe poważne problemy, co opisano bardziej szczegółowo w podrozdziale 2.3.

W takim przypadku uzasadnione jest wykorzystanie metod samokalibracji. Przegląd metod samokalibracji zamieszczono w rozdziale 2.2. Następnie, w rozdziale 2.3, przegląd

uzupełniono o proponowane przez różnych autorów metody kalibracji układu kamer: wizyjnej i termowizyjnej. Ostatecznie przegląd metod podsumowano w podrozdziale 2.4.

2.1 Ogólny model układu dwóch kamer

Ponieważ kalibracja prowadzi do wyznaczenia wybranych parametrów opisu układu kamer ustalenie, jakimi parametrami opisuje się ten układ, pozwoli na lepsze przedstawienie metod kalibracji w kolejnych podrozdziałach.

W ogólnym przypadku dla układu dwóch kamer konieczne jest ustalenie dwóch zestawów parametrów wewnętrznych opisujących model obrazowania każdej kamery z osobna oraz jednego zestawu parametrów zewnętrznych opisujących położenie wzajemne tych kamer. Parametry wewnętrzne opisują przekształcenie perspektywiczne obserwowanego fizycznego otoczenia 3D na cyfrowy obraz 2D. Parametry zewnętrzne opisują położenie i orientację kamer względem siebie. W związku z założeniami opisanymi w rozdziale 1.2 przyjmuje się, że dla rozważanego w rozprawie układu kamer nie jest konieczna korekcja dystorsji układu optycznego kamer. W związku z tym dystorsje zostały wykluczone z opisu modelu układu kamer.

Opis modelu układu kamer dostosowano do przyjętych w rozprawie założeń i przedstawiono w nawiązaniu do [6]. Opis podzielono na trzy etapy. Na początku, w podrozdziale 2.1.1 wprowadzono macierze opisujące parametry wewnętrzne definiujące sposób obrazowania każdej kamery. Następnie w podrozdziale 2.1.2 przedstawiono macierz zasadniczą zawierającą parametry zewnętrzne układu dwóch kamer. Macierz zasadnicza opisuje położenie i orientację jednej kamery względem drugiej. Na koniec w podrozdziale 2.1.3 zdefiniowano macierz fundamentalną integrującą zarówno parametry wewnętrzne jak i zewnętrzne układu kamer. Proces kalibracji w odniesieniu do jednej kamery prowadzi do wyznaczenia wybranych parametrów wewnętrznych tej kamery. Natomiast w odniesieniu do układu kamer proces kalibracji prowadzi do wyznaczenia wybranych parametrów wewnętrznych i zewnętrznych tego układu.

2.1.1 Parametry wewnętrzne kamer

Podstawowa zasada opisująca działanie kamery zakłada, że jest to urządzenie realizujące przyporządkowanie rzeczywistych punktów X otoczenia w przestrzeni 3D punktom x należącym do obrazu w przestrzeni 2D. Przykładowo na rys. 2.1 punkt 3D o współrzędnych (X, Y, Z) został przyporządkowany punktowi obrazu o współrzędnych (x, y) . Wygodnie jest przyjąć, że punkty są zapisane we współrzędnych jednorodnych:

$$\mathbf{X} = \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

$$\mathbf{x} = \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (2.1)$$

gdzie:

X, Y, Z - współrzędne rzeczywistego punktu otoczenia,

x, y - współrzędne punktu obrazu.

W celu zdefiniowania przyporządkowania punktom \mathbf{X} z przestrzeni 3D punktów \mathbf{x} z przestrzeni 2D przyjmijmy, że środek projekcji kamery znajduje się w środku układu współrzędnych odniesienia (jak na rys. 2.1). Dalej przyjmijmy, że istnieje specjalna płaszczyzna obrazowania o równaniu $Z = f$, gdzie f jest odległością ogniskową kamery. Jeśli punkt obrazu oraz rzeczywisty punkt otoczenia są wyrażone we współrzędnych jednorodnych, przyporządkowanie opisujące obrazowanie można zdefiniować zgodnie z rozszerzonym modelem kamery otworkowej jako model kamery typu CCD o skali nierównomiernej w kierunkach X, Y [6, rozdz. 6]. Parametry opisujące model określone są przez macierz wewnętrzną \mathbf{K} , a model obrazowania opisuje się we współrzędnych jednorodnych wzorem (2.2).

$$\mathbf{x} = \mathbf{P}\mathbf{X} = \begin{bmatrix} \mathbf{K} & \mathbf{0} \end{bmatrix} \mathbf{X} \quad (2.2)$$

gdzie:

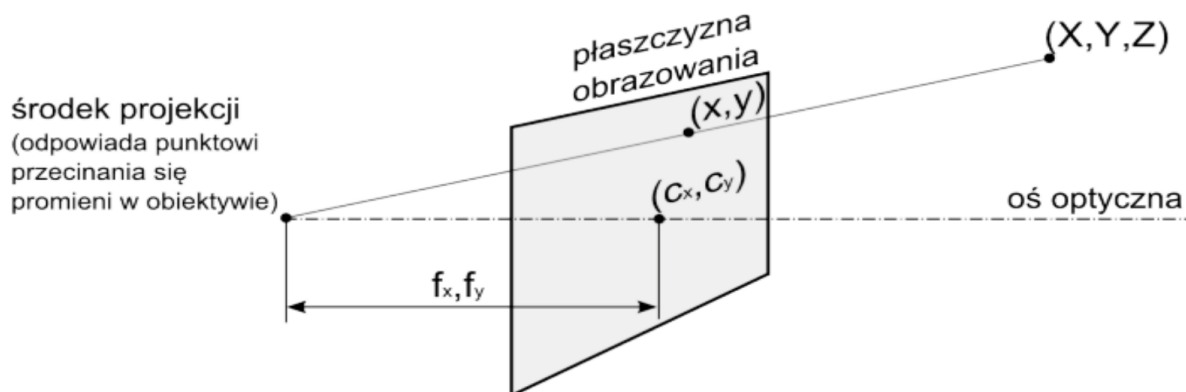
\mathbf{P} - macierz rzutowania kamery,

\mathbf{K} - macierz parametrów wewnętrznych kamery.

$\begin{bmatrix} \mathbf{K} & \mathbf{0} \end{bmatrix}$ - macierz utworzona przez dodanie jednej kolumny z wartościami 0 z prawej strony macierzy \mathbf{K} .

Dokładną postać macierzy wewnętrznej w modelu kamery CCD (rys. 2.1) definiuje wzór (2.3).

$$\mathbf{K} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \quad (2.3)$$



Rysunek 2.1: Rozszerzony model kamery.

gdzie:

\mathbf{K} – macierz wewnętrzna kamery,

$f_x = fm_x$ – ogniskowa kamery w kierunku X,

$f_y = fm_y$ – ogniskowa kamery w kierunku Y,

c_x, c_y – przesunięcie osi optycznej względem środka matrycy CCD,

f - odległość ogniskowa kamery,

m_x – ilość pikseli na jednostkę odległości w osi X,

m_y – ilość pikseli na jednostkę odległości w osi Y.

2.1.2 Parametry zewnętrzne układu kamer - macierz zasadnicza

Zgodnie z [6, rozdz. 9] układ dwóch kamer, jak na rys. 2.2, może zostać opisany przez macierz zasadniczą. Charakteryzuje ona relacje geometryczne pomiędzy kamerami - wzajemne przemieszczenie i orientację jednej z kamer względem drugiej. Macierz zasadnicza \mathbf{E} zdefiniowana jest wzorem (2.4).

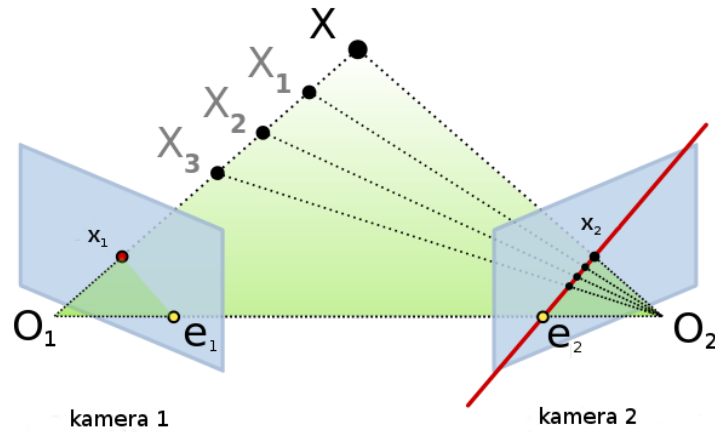
$$\mathbf{E} = [\mathbf{t}]_{\times} \mathbf{R} \quad (2.4)$$

gdzie:

\mathbf{R} – macierz opisująca obrót w euklidesowej przestrzeni trójwymiarowej jednej kamery względem drugiej,

oraz:

$$[\mathbf{t}]_{\times} = \begin{bmatrix} 0 & -t_z & t_y \\ t_z & 0 & -t_x \\ -t_y & t_x & 0 \end{bmatrix} \quad (2.5)$$



Rysunek 2.2: Wizualizacja geometrii epipolarniej układu dwóch kamer. Punkty ogniskowe kamer oznaczone są punktami O_1 oraz O_2 . Obserwowany punkt rzeczywistej sceny 3D oznaczony jest jako X . Punkt jest obrazowany przez każdą z kamer odpowiednio jako x_1 oraz x_2 . Punkty epipolarnie oznaczone są jako e_1 (obraz z kamery 1 punktu ogniskowego O_2) oraz e_2 (obraz z kamery 2 punktu ogniskowego O_1). Linia epipolarna odpowiadająca punktowi x_1 zdefiniowana jest przez punkty e_2 i x_2 . Punkty X_1 , X_2 oraz X_3 reprezentują przykładowe położenia punktu X widoczne na obrazie z lewej kamery jako punkt x_1 . Rys. na podstawie Arne Nordmann (norro) - CC BY-SA 3.0, <https://commons.wikimedia.org/w/index.php?curid=1702052>

gdzie:

$\mathbf{t} = \begin{bmatrix} t_x & t_y & t_z \end{bmatrix}^T$ - wektor przesunięcia w euklidesowej przestrzeni trójwymiarowej jednej kamery względem drugiej,

$[\mathbf{t}]_{\times}$ - reprezentacja macierzowa wektora \mathbf{t} w celu zastąpienia mnożenia wektorowego przez mnożenie macierzy, czyli $[\mathbf{t}]_{\times}$ jest to macierz antysymetryczna o własności $[\mathbf{t}]_{\times} \mathbf{R} = \mathbf{t} \times \mathbf{R}$.

2.1.3 Macierz fundamentalna

Macierz fundamentalną można przedstawić jako uogólnienie macierzy zasadniczej [6, str. 257]. Powołując się na [6, str. 246] można ją opisać wzorem (2.6).

$$\mathbf{F} = \mathbf{K}_2^{-T} [\mathbf{t}]_{\times} \mathbf{R} \mathbf{K}_1^{-1} \quad (2.6)$$

gdzie:

\mathbf{F} - macierz fundamentalna,

\mathbf{K}_1 - macierz wewnętrzna kamery pierwszej,

\mathbf{K}_2 - macierz wewnętrzna kamery drugiej,

$[\mathbf{t}]_{\times}$ - macierz przesunięcia jednej kamery względem drugiej o wektor \mathbf{t} ,

\mathbf{R} - macierz obrotu.

Biorąc pod uwagę wzór (2.4), wzór (2.6) można przedstawić w postaci (2.7).

$$\mathbf{F} = \mathbf{K}_2^{-T} \mathbf{E} \mathbf{K}_1^{-1} \quad (2.7)$$

gdzie:

\mathbf{E} – macierz zasadnicza.

W zakresie kalibracji układu kamer istotna jest szczególnie jedna z własności macierzy fundamentalnej [6, str. 245] opisująca odwzorowanie punktu na obrazie jednej kamery w odpowiadającą punktowi prostą na obrazie drugiej kamery. Taką prostą określa się mianem “linia epipolarna”. Linię epipolarną można zdefiniować wzorem (2.8). Graficznie prezentuje to rysunek 2.2. Nawiązując do wzoru (2.8), możemy określić kolejną własność (wzór (2.9)) macierzy fundamentalnej \mathbf{F} wiążącą ze sobą odpowiadające sobie punkty obydwu obrazów.

$$l_2 = \mathbf{F} \mathbf{x}_1 \quad (2.8)$$

gdzie:

\mathbf{F} - macierz fundamentalna,

\mathbf{x}_1 – punkt na obrazie z pierwszej kamery,

l_2 – linia epipolarna na obrazie z drugiej kamery odpowiadająca punktowi \mathbf{x}_1 na obrazie z pierwszej kamery.

$$\mathbf{x}_2^T \mathbf{F} \mathbf{x}_1 = 0 \quad (2.9)$$

gdzie:

\mathbf{x}_1 – punkt na obrazie z pierwszej kamery,

\mathbf{x}_2 – odpowiadający punkt na obrazie z drugiej kamery.

Punkty na obrazach odpowiadają sobie, jeśli są zobrazowaniem tego samego punktu rzeczywistego sceny obserwowanej przez głowicę kamer.

Macierz fundamentalną opisaną wzorem (2.7) oraz jej własności wykorzystano do przedstawienia różnych metod samokalibracji w kolejnych podrozdziałach rozdziału 2.

2.2 Ogólny przegląd metod samokalibracji

Często spotykane w literaturze podejście do samokalibracji zakłada, że kalibrowany układ kamer rejestruje pewne otoczenie i na podstawie sekwencji nagranych obrazów wykonywana jest automatyczna estymacja wybranych parametrów modelu układu. Z punktu widzenia przetwarzania obrazów oznacza to, że samokalibracja wykonywana jest na podstawie

automatycznego przetwarzania sekwencji obrazów cyfrowych. W celu estymacji parametrów układu konieczne jest ustalenie na różnych obrazach obiektów odpowiadających sobie. Są to projekcje tego samego fizycznego obiektu otoczenia 3D na różne płaszczyzny obrazowania kamer. W zależności od aplikacji dotyczy to obrazów z różnych kamer lub obrazów z tej samej kamery, ale obserwującej otoczenie z różnych punktów widzenia albo w różnym czasie.

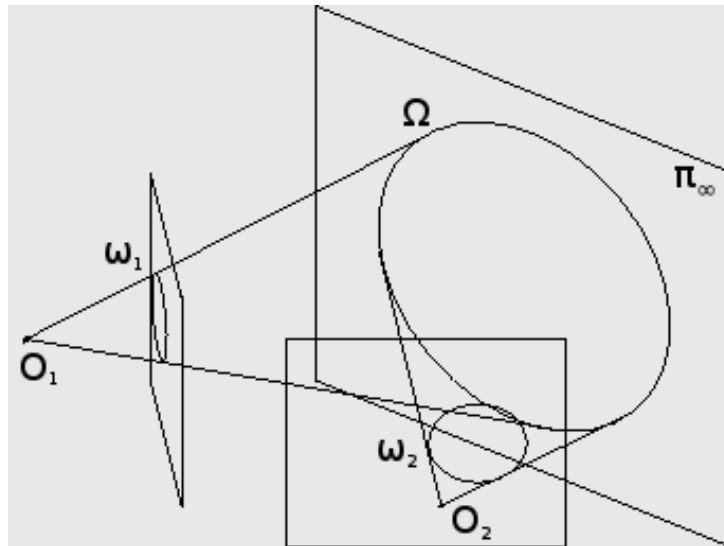
Znacząca ilość opisywanych w literaturze metod samokalibracji dotyczy jednej kamery o stałych parametrach w sytuacji ruchu - zmiany punktu obserwacji otoczenia. Istnieje jednak grupa metod samokalibracji przeznaczonych dla układu kamer stereowizyjnych. Założenia towarzyszące tym metodom są zbliżone do założeń przyjętych w rozprawie (rozdział 1.2).

Przy omawianiu metod samokalibracji warto dodatkowo wspomnieć o metodach kalibracji układu sensorów na pokładzie robota [7]. W tym zakresie robot rozumiany jest jako pewien mechanizm wieloczłonowy, na którym zamocowane są sensory, w szczególności kamery. Rozważany w rozprawie problem można sformułować w ramach kalibracji układu sensorów na pokładzie robota, podobnie jak ma to miejsce w pracy [8]. Układ kamer rozważany w rozprawie można w takim przypadku potraktować jako robot - mechanizm wieloczłonowy o dwóch obrotowych stopniach swobody. W ramach rozprawy przyjęto jednak definicje problemu kalibracji stosowane powszechnie w dziedzinie przetwarzania obrazów.

Pierwsze doniesienia opisujące procedurę automatycznej kalibracji (samokalibracji) dotyczącą jednej kamery można znaleźć w pracach [9, 3]. Autorzy przedstawiają metodę samokalibracji dla jednej kamery bazującą na równaniach Kruppa. Są to równania wiążące ze sobą macierz fundamentalną, punkty epipolarne oraz parametry wewnętrzne kamer [10]. Wykorzystanie tych równań jest standardowym podejściem do samokalibracji kamer (przykładowo [9, 3, 11, 12, 13, 14]). Celem tych metod jest często kalibracja wszystkich parametrów kamery do poziomu zależności metrycznych. Autorzy pracy [15] podkreślają dodatkowo, że w przypadku samokalibracji zakłada się zazwyczaj stałość parametrów wewnętrznych kalibrowanej kamery. Istnieją jednak metody, gdzie takie założenie nie jest konieczne.

Teoretyczne podstawy opisywanych w literaturze metod najczęściej przedstawiane są z wykorzystaniem własności geometrii rzutowej i epipolarnej związanych ze specjalnym obiektem w tej przestrzeni - *absolute conic*. Jest to krzywa stożkowa Ω leżąca na płaszczyźnie w nieskończoności Π_∞ , co przedstawiono na rys. 2.3. Rzut tej krzywej na płaszczyznę obrazowania kamery jest funkcją wyłącznie parametrów wewnętrznych kamery i nie zależy od punktu widzenia kamery. W przypadku dwóch kamer o punktach ogniskowych O_1, O_2 rzut tej krzywej na płaszczyznę obrazowania kamer można oznaczyć odpowiednio ω_1, ω_2 .

W ramach przeglądu metod samokalibracji warto wyróżnić kilka grup metod: metody bazujące wyłącznie na stałych związanych z modelem kamery, metody wykonujące dodatkowe założenia związane z ruchem kamer oraz metody zakładające zależności wiążące obserwowaną scenę.



Rysunek 2.3: Przedstawienie krzywej stożkowej *absolute conic* Ω leżącej na płaszczyźnie w nieskończoności Π_∞ . Rzut tej krzywej na płaszczyznę obrazowania kamer 1,2 oznaczono ω_1, ω_2 . O_1, O_2 to punkty ogniskowe kamer. Rys. na podstawie <https://www.cs.unc.edu/~marc/tutorial/node87.html>

W pierwszej grupie znajdują się metody posiadające najmniejszą ilość założeń. Są przez to najbardziej ogólne, ale również najbardziej skomplikowane z uwagi na znaczną ilość parametrów koniecznych do estymacji. W [9, 3] po raz pierwszy opisano metodę kalibracji jednej kamery bazującą na równaniach Kruppa na podstawie trzech obrazów sceny. W artykułach opisano estymację parametrów wewnętrznych kamery na podstawie odpowiadających sobie na obrazach punktów. Procedura zakłada wyznaczenie na tej podstawie punktów epipolarnych i dalszą kalibrację z wykorzystaniem równań Kruppa. Macierz fundamentalna przekształcenia może być znaleziona na podstawie równania (2.9). W [11] rozwinięto tę metodę samokalibracji, również z zastosowaniem trzech obrazów. Jej podstawową zaletą jest wyznaczanie dopasowań niezależnie między parami obrazów bez ustalania jednego układu odniesienia. Autorzy proponują znalezienie wszystkich rozwiązań sformułowanego problemu z wykorzystaniem metod *numerical continuation*. Różne inne modyfikacje metody z [9] przedstawiono w [12].

W [13] podkreślono trudność w kalibracji kamer z uwagi na nieliniowość ogólnych równań koniecznych do rozwiązania w celu jednoczesnej estymacji wielu szukanych parametrów. Zarówno w [13] jak i w [14] oraz podobnie jak w [16] i [17] zaproponowano wieloetapowe podejście do samokalibracji. Początkowo wyznaczana jest kalibracja perspektywiczna. W kolejnym etapie można wyznaczyć kalibrację afiniczną za pomocą przedstawionego w artykule równania więzów. W [18] pokazano natomiast jak kalibrację afiniczną doprowadzić do kalibracji metrycznej. Podobna metoda została zastosowana do kalibracji układu kamer stereowizyjnych w [19] oraz [20]

Bazując na równaniach Kruppa w [21] przedstawiono algorytm samokalibracji dla więcej niż trzech kamer z wykorzystaniem technik optymalizacji. Optymalizację zastosowano również w [22] bazując na wyznaczeniu i własnościach specjalnego obiektu geometrii rzutowej - *absolute quadric*. Wprowadzenie go upraszcza w niektórych przypadkach stosowanie równań Kruppa.

Ciekawe podejście do problemu przedstawiono w [23]. Uproszczone równania Kruppa i przedstawiono ich algebraiczne wyprowadzenie nie opierając się na interpretacji geometrycznej i specjalnych obiektach geometrii rzutowej. W celu uniknięcia bezpośredniego wyznaczenia punktów epipolarnych posłużono się metodą SVD. Dzięki zastosowanym zabiegom zwiększona została odporność samokalibracji na błędne wyznaczenie odpowiadających sobie punktów.

Metoda opisana w [24] zakłada z kolei wyznaczenie współczynników skali z równań Kruppa w pierwszej kolejności. W tym celu zastosowano optymalizację Levenburga-Marquardta. Ułatwiło to wyznaczenie parametrów wewnętrznych kamery, ponieważ doprowadziło do prostszych, liniowych więzów ograniczających je.

Drugą istotną grupą metod samokalibracji są algorytmy wykorzystujące dodatkowo zależności nałożone na ruch kalibrowanej kamery [15]. Dzięki tym dodatkowym założeniom można uzyskać prostsze i bardziej odporne metody. Uproszczenia mogą jednak prowadzić do problemów z ustaleniem parametrów na poziomie metrycznym. W [16] oraz [25] wyznaczono homografię płaszczyzny w nieskończoności poprzez początkową kalibrację afiniczną bazując tylko na translacji [16] lub tylko na rotacji [25]. W [26] analizowana jest metoda samokalibracji przy założeniu ruchu złożonego z przesunięcia w płaszczyźnie i rotacji wokół osi prostopadłej do tej płaszczyzny. Dodatkowo pokazane jest wykorzystanie *trifocal tensors* do kalibracji - zależności wiążącej trzy odpowiadające sobie linie na obrazach z trzech różnych punktów widzenia. W innej publikacji wykazano [27], że w przypadku ruchów płaszczyznowych kamerę 2D można sprowadzić do przypadku 1D i uprościć w ten sposób rozwiązanie problemu kalibracji. W [28] przeanalizowano problemy samokalibracji w sytuacji, gdy założenia związane z ruchem kamery tylko w postaci obrotu wokół jej osi optycznej nie są spełnione.

Ostatnią wyróżnioną grupą metod samokalibracji są algorytmy zakładające spełnienie dodatkowych założeń związanych z obserwowaną sceną. Są to założenia słabsze niż w przypadku standardowej kalibracji kamer z wykorzystaniem wzorca. W [29] przedstawiono algorytm zakładający obserwowanie płaskiej sceny i kalibrację z wykorzystaniem prostopadłych do siebie przemieszczeń kamery. Obserwowanie płaskiej sceny zakładają również autorzy [30]. Opracowano tam metodę z zastosowaniem specjalnego obiektu geometrii rzutowej - *absolute quadric*. Metoda zakłada przyjęcie jednego obrazu jako referencyjnego względem wszystkich innych. Niektórzy autorzy podkreślają, że takie podejście może prowadzić do niskiej odporności metody. Jeśli obraz referencyjny będzie niskiej jakości i wystąpią na nim problemy z lokalizacją punktów, może to doprowadzić do braku odporności w działaniu metody. Przy podobnych założeniach dotyczących sceny, opisany problem nie występuje natomiast w [31, 32]. Przyjęta została tam metoda uśredniania błędów dla

wszystkich analizowanych obrazów. Zarówno w [33] jak i w [34] możemy znaleźć propozycję wykorzystania znanych, płaskich wzorców w celu kalibracji kamery. W [33] analizowane są przypadki szczególne i osobliwości metody. Przedstawiono tam ciekawe podejście do równania Kruppa - określono je w odniesieniu do homografii pomiędzy obserwowanymi płaskimi powierzchniami. W [34] przedstawiony jest sposób na korekcję dystorsji kalibrowanej kamery. W [35] przedstawiono technikę pozwalającą na metryczną rektyfikację obrazów perspektywicznych płaszczyzn. W tym celu konieczne jest założenie: znajomości jednej wartości kąta pomiędzy obserwowanymi prostymi, równości dwóch nieznanymi kątów lub znajomości stosunku długości odcinków na obserwowanej płaszczyźnie. W [36] przedstawiono metodę odtwarzania ruchu kamery i jej parametrów wewnętrznych bazującą na znajomości dwóch homografii pomiędzy płaszczyznami z obserwowanej sceny. W [37] wykorzystano znajomość dopasowania punktów, linii oraz płaszczyzn. W procesie kalibracji autorzy wykorzystują równania wynikające z przynależności punktów do prostych lub płaszczyzn oraz prostych do płaszczyzn. Ustalenie równoległości oraz prostopadłości linii oraz płaszczyzn pozwala ustalić dodatkowe więzy. Metoda ma niewielkie znaczenie praktyczne ponieważ wszystkie te zależności autorzy wprowadzają manualnie. W [38] autorzy przedstawiają sposób łączenia różnych równań więzów. Założenia obejmują więzy wynikające ze struktury sceny: punkty w nieskończoności leżące na prostopadłych liniach, płaszczyzny o homografii jednostkowej w stosunku do płaszczyzny obrazowania oraz wynikające z modelu kamery i przemieszczeń pomiędzy poszczególnymi widokami sceny. Wszystkie równania sformułowane zostały w nawiązaniu do obrazu *absolute conic* (rys. 2.3). W [39] autorzy wykorzystują z kolei w celu samokalibracji własności równoległoboku, jeśli taki można zaobserwować na występujących w otoczeniu kamer płaszczyznach. Podobne założenie można znaleźć w [40].

Omawiając metody samokalibracji należy podkreślić, że istnieją sekwencje przemieszczeń kamery, które nie pozwalają na jednoznaczne ustalenie jej parametrów wewnętrznych. W literaturze określa się je jako *critical motion sequences*. Dla przemieszczeń przedstawionych w postaci takich sekwencji nie istnieje jednoznaczne rozwiązanie problemu samokalibracji. Analizę problemu dla wybranych metod można znaleźć w [41, 33]. Niektóre wybrane przypadki omówiono również w [42, 43, 44]. Warto zwrócić uwagę na przypadek przedstawiany w [45]. Jeśli rozważymy ogólny problem samokalibracji z wykorzystaniem równań Kruppa, napotkamy na problem z estymacją parametrów kamer, jeśli ich środek optyczny będzie leżał na pewnej sferze, a ich oś optyczna będzie przechodzić przez środek tej sfery.

Wobec przyjętych w rozprawie założeń nie można pominąć metod samokalibracji opracowanych dla potrzeb kamer stereowizyjnych. Większość autorów w celu kalibracji układu kamer stereowizyjnych wykorzystuje równanie macierzy fundamentalnej (2.9). Określane jest ono w literaturze jako więzy epipolarne i opisuje relację pomiędzy odpowiadającymi sobie punktami z różnych kamer układu stereowizyjnego.

W [46, 47] przedstawiono metodę bardzo zbliżoną do [39] czy [40] - zastosowano więzy epipolarne dla kamery stereowizyjnej do dopasowania odpowiadających sobie punktów.

Obydwie kamery są tej samej modalności, ale o różnych parametrach wewnętrznych. Estymacja parametrów układu kamer następuje w wyniku obserwacji nieznannej sceny płaskiej na podstawie dwóch obrazów. Pierwszy etap kalibracji to ustalenie macierzy homografii wiążącej dopasowane do siebie punkty za pomocą algorytmu RANSAC. W kolejnym kroku stosowana jest optymalizacja Levenberga-Marquardta w celu rozwiązania równań nieliniowych wiążących wyznaczone macierze homografii z parametrami układu kamer.

W [48] wykorzystano więzy epipolarne, a następnie równania Kruppa do kalibracji układu kamer stereo na podstawie obserwacji jednego przemieszczenia układu. Przyjęto tam uproszczony model kamery - środki obrazu kamer są znane, oraz nie ma przekoszenia kamer. Poszukiwanymi parametrami wewnętrznymi jest jedynie skala obrazów w osi X i Y. Estymacja parametrów wewnętrznych i zewnętrznych następuje poprzez minimalizację metodą Gaussa-Newtona z wykorzystaniem ciekawie skonstruowanego błędu uwzględniającego odległości punktów dopasowywanych od linii epipolarnej.

W [49] przedstawiono metodę wykorzystującą również więzy epipolarne w odniesieniu do odpowiadających sobie punktów z kamer tej samej modalności. Dodatkowo w metodzie wykorzystano ciekawe relacje trójliniowe wiążące układ trzech obrazów i odpowiadające sobie na nich punkty. Autorzy przekonują, że metoda kalibracji powinna działać on-line w trakcie pracy systemu wizyjnego w celu ciągłego monitorowania i poprawy jakości jego działania. Filtr Kalmana wykorzystano do estymacji parametrów. Kryteria optymalizacji Gaussa-Helmerta zastosowano w celu poprawienia jakości estymowanych parametrów. Autorzy zakładają, że znane są zgrubne parametry kalibracji układu kamer. W wyniku ciągłej pracy algorytmu są one dostrajane. W algorytmie wykorzystują więzy epipolarne do chwilowej estymacji parametrów i ciągłą rekursywną pracę metody, aby agregować dane z poprzednich chwil czasowych. Podobne podejście - wstępne określenie parametrów zgrubnych, a następnie dostrajanie ich - przestawiono np. w [48], czy w [11]. Wykorzystano tam więzy epipolarne i metody optymalizacji do ustalenia macierzy fundamentalnych układu kamer oraz przesunięcia tego układu w trakcie ruchu.

Metoda przedstawiona w [49] została rozwinięta w [50]. Autorzy opisują samokalibrację on-line w celu ciągłego monitorowania parametrów zewnętrznych głowicy kamer. Skupiają się jedynie na kalibracji parametrów wpływających na błędy dopasowania punktów w poziomie. Są to ich zdaniem błędy najbardziej pogarszające jakość algorytmów stereowizyjnej rekonstrukcji sceny 3D. Kalibracja przebiega dwuetapowo. W pierwszym etapie następuje dopasowanie punktów z wykorzystaniem metody RANSAC w celu estymacji macierzy fundamentalnej. W kolejnym etapie przeprowadzona jest optymalizacja z wykorzystaniem wspomnianego błędu dopasowania punktów w poziomie.

2.3 Przegląd metod kalibracji układu kamer: wizyjnej i termowizyjnej

Opisywane w literaturze metody samokalibracji układu kamer wizyjnej i termowizyjnej są nieliczne [51]. W zdecydowanej większości autorzy decydują się na standardową metodę kalibracji z wykorzystaniem płaskiego wzorca kalibracyjnego [4, 52, 53, 54, 55, 56, 57, 58].

W [51] autorzy opisują proces samokalibracji układu dwóch sensorów: kamery typu RGBD (kamera wizyjna z informacją o głębi) oraz kamery termowizyjnej. Układ można traktować jako złożony z trzech sensorów: kamery wizyjnej (RGB), kamery głębi (D) oraz kamery termowizyjnej. Zakładana jest znajomość parametrów wewnętrznych każdej z kamer oraz znajomość parametrów zewnętrznych układu kamera wizyjna i kamera głębi. Celem przedstawionej metody kalibracji jest wyznaczenie parametrów zewnętrznych układu sensorów kamera typu RGBD i kamera termowizyjna. Samokalibracja przeprowadzona jest jako optymalizacja funkcji energii związanej ze stopniem nakładania się krawędzi wykrytych dla obrazu wizyjnego, termowizyjnego i mapy głębi.

Większość opisywanych w ostatnich latach algorytmów samokalibracji układu sensorów multimodalnych dotyczy kamery w układzie ze skanerem powierzchni 3D. Takie metody samokalibracji nie odpowiadają założeniom przyjętym w ramach rozprawy. Kalibracja układu kamer wizyjnej i termowizyjnej opisywana w literaturze najczęściej zakłada wykorzystanie wzorca kalibracyjnego. Głównym problemem w takim przypadku są różnice w obrazie wzorca pomiędzy kamerą wizyjną a termowizyjną. Konieczne jest zapewnienie takiej jego konstrukcji, żeby cechy charakterystyczne wzorca były dobrze widoczne na obrazach z obydwu kamer. Pozwala to na znalezienie odpowiadających sobie punktów na obrazach z tych kamer i dalszą kalibrację na tej podstawie.

Z uwagi na te problemy autorzy [56] zastosowali żarówki halogenowe o dużej mocy w celu podniesienia temperatury wzorca przez naświetlenie go. Z uwagi na inne właściwości refleksyjne pól czarnych i białych szachownicy możliwe było poprawienie widoczności punktów charakterystycznych wzorca i przeprowadzenie kalibracji. Podobne rozwiązanie zastosowano w [59] oraz [60]. Układ sprzętowy konieczny do ustawienia w celu przeprowadzenia takiej kalibracji wydaje się jednak kłopotliwy i utrudnia wygodne stosowanie zaproponowanej metody w warunkach poligonowych.

W [55] zaproponowano konstrukcję dedykowanego wzorca wykorzystującego płaską tablicę z wyciętymi okrągłymi otworami. Za tablicą umieszczono matę grzewczą. Po włączeniu maty tablica z otworami pozostanie w temperaturze pokojowej, a przez okrągłe otwory widoczna jest mata grzewcza o podniesionej temperaturze. Otwory w tablicy są łatwo dostrzegalne przez kamerę termowizyjną dzięki znaczącej różnicy temperatury względem maty. W [58] przedstawiono podobne podejście. Zastosowano tam kartonową tablicę z wyciętymi kwadratowymi otworami. Przyjęto, że tablica będzie umieszczona na tle materiału o innej

emisyjności podczerwieni lub o innej temperaturze. W tym celu wykorzystano np. włączony monitor komputerowy. Podobne rozwiązanie zastosowano również w [1], gdzie tło dla tablicy z wyciętymi otworami przyjęto jako chłodniejsze od tablicy. Wzorzec kalibracyjny w postaci płaskiej tablicy z okręgami zastosowano również w [61], jednak szczegóły konstrukcji wzorca nie są tam opisane.

W celu zapewnienia dobrej widoczności punktów charakterystycznych wzorca niektórzy autorzy [4] zdecydowali się na konstrukcję tablicy kalibracyjnej z miedzi pomalowanej czarną farbą. Wysoki kontrast pól szachownicy jest zapewniany przez różnice w emisyjności w pasmie podczerwonym miedzi i czarnej farby. Podobną ideę wykorzystano w [62], gdzie tablicę kalibracyjną skonstruowano z białego kartonu kontrastującego z folią aluminiową. Podobny efekt w [63] osiągnięto budując szachownicę kalibracyjną z czarnych i białych kwadratowych fragmentów płyty z melaminowej.

Wzorzec można również wykonać jako układ dokładnie pozycjonowanych miniaturowych żarówek, co przedstawiono w [64]. Opisano tam dodatkowo sposób zwiększenia dokładności wykrywania lokalizacji żarówek przez kamery: wizyjną i termowizyjną.

Zamiast szachownicy autorzy [57] zastosowali podgrzewany znacznik w kształcie niewielkiego okręgu koloru czerwonego. Umieszczając znacznik w różnych miejscach sceny otrzymali dopasowania odpowiadających sobie punktów obrazu termowizyjnego i wizyjnego. Następnie możliwe było wykorzystanie RANSAC w celu estymacji parametrów macierzy fundamentalnej. Podobne rozwiązanie opisano w artykule [54]. Z uwagi na wykorzystanie tych samych rysunków oraz autorstwo, artykuł przedstawia najprawdopodobniej treści podobne do publikacji [65], którą trudno analizować z uwagi na język obcy materiału. W artykule pokazano, jak zbudować wzorzec kalibracyjny dla wspomnianego układu kamer. Wzorzec skonstruowano w postaci przestrzennego obiektu z elementami grzewczymi, aby poprawić jego widoczność dla kamery termowizyjnej. W celu kalibracji autorzy przedstawili metodę bazującą na minimalizacji sumy wartości dopasowań odpowiadających sobie punktów wzorca dla wielu par obrazów. W implementacji zastosowano OpenCV i standardową metodę DLT.

W [66] przedstawiono m. in. automatyczną metodę kalibracji układu kamer wizyjnej i termowizyjnej podstawie wzorca kalibracyjnego. Wzorzec zaproponowano w postaci szachownicy uzupełnionej o rezystory grzewcze dodane w celu poprawienia lokalizacji wzorca dla kamery termowizyjnej. W artykule autorzy skupili się głównie na aplikacji współpracy kamery termowizyjnej i skanera laserowego 3D.

2.4 Podsumowanie

Zastosowanie wzorca kalibracyjnego jest najbardziej precyzyjną metodą kalibracji układu kamer. W świetle założeń przyjętych w rozprawie wykorzystanie wzorca jest trudne - wzorzec kalibracyjny musiałby być znacznych rozmiarów, oraz pozwalać na dobrą lokalizację

jego wymiarów na podstawie obrazu zarówno z kamery wizyjnej jak i termowizyjnej. W celu kalibracji układu kamer zgodnego z przyjętymi założeniami zastosowanie wzorca kalibracyjnego nie jest konieczne. Przedstawione w podrozdziale 2.2 metody samokalibracji potrafią estymować parametry opisujące model kamery i układu kamer bez stosowania wzorca kalibracyjnego.

Metody bardziej ogólne cechują się większą złożonością obliczeniową i mniejszą odpornością na błędy w wyznaczeniu cech obrazów, na podstawie których estymowane są parametry układu kamer. W celu poniesienia odporności należy stosować metody estymacji tylko wybranych parametrów układu kamer, koniecznych do poprawnego działania przyjętej aplikacji systemu wizyjnego. Ponadto oryginalne równania Kruppa są częściowo od siebie zależne. Ograniczenie ilości równań pozwala podjąć lepszą decyzję, które z nich stosować w celu estymacji szukanych parametrów. Wyprowadzenie nowych równań może mieć na celu również uniknięcie wyznaczania punktów epipolarnych - ich dokładne wyznaczenie jest trudne w obecności szumu pomiarowego lub w przypadku ruchów zdegenerowanych [67].

Dodatkowo, jeśli to tylko możliwe, warto podzielić metodę na podproblemy [13, 14, 18]. Pozwala to uprościć rozwiązanie zadania estymacji szukanych parametrów. Dodatkowo podejście wieloetapowe jest bardziej elastyczne - każdy z etapów w zależności od potrzeb może zostać rozwiązany różnymi metodami.

Na koniec warto podkreślić, że nie należy stosować jednego obrazu jako układu odniesienia na potrzeby całej metody. Najlepiej wyznaczać kluczowe zależności pomiędzy różnymi obrazami bez wyróżniania dowolnego z nich. Zwiększa to odporność i skuteczność metody.

W nawiązaniu do założeń rozprawy i opisanych w bieżącym rozdziale metod kalibracji, proponowanym sposobem realizacji celu rozprawy jest opracowanie metody samokalibracji układu dwóch kamer: wizyjnej i termowizyjnej. Estymacja parametrów układu następuje w takim przypadku na podstawie odnalezienia na różnych obrazach odpowiadających sobie obiektów. W przypadku kamer rejestrujących obrazy w różnym paśmie promieniowania elektromagnetycznego automatyczne znalezienie takich obiektów na potrzeby samokalibracji jest istotnym problemem. W celu rozwiązania go, w rozdziale 3 przedstawiono analizę problemu dopasowania obrazów.

Rozdział 3

Przegląd metod automatycznego dopasowania obrazów

Jednym z kluczowych elementów większości metod samokalibracji jest automatyczne odnalezienie na różnych obrazach odpowiadających sobie obiektów. Automatyczne dopasowanie obrazów jest jednym ze sposobów realizacji tego zadania. Dopasowanie określa, w jaki sposób przekształcić obiekty jednego obrazu tak, aby odpowiadały położeniu i kształtom obiektów obrazu drugiego. W rozdziale przeprowadzona została klasyfikacja metod oraz wykonany został przegląd najważniejszych algorytmów dopasowania obrazów.

Przegląd algorytmów dopasowania poprzedzono opisem modelu dopasowania dwóch obrazów w podrozdziale 3.1. Są to podstawowe informacje konieczne do przedstawienia zasady działania różnych metod dopasowania obrazów w dalszych podrozdziałach bieżącego rozdziału.

W podrozdziale 3.2 znajduje się ogólny przegląd algorytmów dopasowania. Opisano metody mające znaczny wpływ na rozwój dziedziny w ostatnich latach. Ze szczególną uwagą potraktowano publikacje z zakresu dopasowania sekwencji obrazów wizyjnych i termowizyjnych dla potrzeb monitorowania.

W podrozdziale 3.3 przedstawiono natomiast przegląd ze szczególnym uwzględnieniem metod bazujących na korelacji fazowej. Zgodnie z badaniami [68, 69] autora rozprawy, metody te cechują się możliwością wydajnej implementacji. Dodatkowo w ramach tych prac udało się potwierdzić ich skuteczność działania w zakresie dopasowania obrazów wizyjnych i termowizyjnych zgodnie z przyjętymi w rozprawie założeniami.

Na koniec w podrozdziale 3.4 znajduje się podsumowanie dotyczące metod automatycznego dopasowania obrazów.

3.1 Model dopasowania obrazów

Opierając się na definicjach zawartych w [6, 70, 71], jeśli przyjmiemy, że punkt obrazu będzie zapisany we współrzędnych jednorodnych zgodnie ze wzorem (2.1), operację przesunięcia punktu obrazu można przedstawić w postaci równania:

$$\mathbf{x}_2 = \mathbf{T}\mathbf{x}_1, \quad (3.1)$$
$$\mathbf{T} = \begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix}$$

gdzie:

t_x, t_y – parametry opisujące przesunięcie punktu \mathbf{x}_1 w kierunkach osi X i Y.

Funkcja $t(\mathbf{x})$ opisuje operację przesunięcia wszystkich punktów obrazu według przyjętych założeń:

$$t(\mathbf{x}) = \begin{bmatrix} 1 & 0 & t_x(\mathbf{x}) \\ 0 & 1 & t_y(\mathbf{x}) \\ 0 & 0 & 1 \end{bmatrix} \quad (3.2)$$

Oznacza to, że funkcja t zostanie określona, jeśli ustalimy parametry $t_x(\mathbf{x}), t_y(\mathbf{x})$, dla każdego punktu \mathbf{x} obrazu. W związku z tym przesunięcie opisane za pomocą funkcji t przykładowo dla obrazu dwuwymiarowego o rozmiarze 256×256 pikseli posiada maksymalnie $256 \times 256 = 131072$ stopni swobody.

Niech dany będzie obraz monochromatyczny g_i opisany funkcją intensywności jego pikseli:

$$\mathbf{g}_i = g_i(\mathbf{x}_i) = g_i(\mathbf{x}_{i1}), g_i(\mathbf{x}_{i2}), \dots, g_i(\mathbf{x}_{im}) \quad (3.3)$$

gdzie:

$i = 0, 1, \dots$ - oznaczenie obrazu, m - ilość pikseli w obrazie, $g_i(\mathbf{x}_{i1}), g_i(\mathbf{x}_{i2}), \dots, g_i(\mathbf{x}_{im})$ - intensywność (np. w zakresie wartości $0 \dots 255$) obrazu g_i dla wszystkich punktów (pikseli) obrazu $\mathbf{x}_{i1}, \mathbf{x}_{i2}, \dots, \mathbf{x}_{im}$.

Dla dwóch obrazów $g_1(\mathbf{x}_1)$ oraz $g_2(\mathbf{x}_2)$ pochodzących z kamer obserwujących otoczenie z różnych punktów widzenia możemy określić, które punkty obrazów odpowiadają sobie. Nawiązując do równania (2.2), punkty $\mathbf{x}_{1j}, \mathbf{x}_{2k}, j \in (1, \dots, m), k \in (1, \dots, m)$ odpowiadają sobie, jeżeli są obrazem tego samego rzeczywistego punktu \mathbf{X} przestrzeni 3D:

$$\begin{aligned} \mathbf{x}_{1j} &= \mathbf{P}_1 \mathbf{X} \\ \mathbf{x}_{2k} &= \mathbf{P}_2 \mathbf{X} \end{aligned} \quad (3.4)$$

gdzie:

$\mathbf{P}_1, \mathbf{P}_2$ - macierze opisujące obrazowanie za pomocą kamery 1 oraz 2,

\mathbf{X} - rzeczywisty punkt 3D obserwowanej przez kamery 1 oraz 2 sceny,

$\mathbf{x}_{1j}, \mathbf{x}_{2k}$ - współrzędne pewnego punktu obrazu odpowiednio \mathbf{g}_1 oraz \mathbf{g}_2 .

Punkty odpowiadające sobie nie muszą mieć tych samych wartości współrzędnych. Z wykorzystaniem funkcji przesunięcia punktów opisanej wzorem (3.2) możliwe jest sprowadzenie różnic w położeniu odpowiadających sobie punktów do zera. W rozprawie przyjmuje się, że wektory przesunięcia odpowiadających sobie punktów będą określane jako dopasowanie tych punktów. Dopasowanie można więc opisać wzorem (3.5).

$$d(\mathbf{x}) = \begin{bmatrix} d_x(\mathbf{x}) \\ d_y(\mathbf{x}) \end{bmatrix} = \begin{bmatrix} t_x(\mathbf{x}) \\ t_y(\mathbf{x}) \end{bmatrix} \quad (3.5)$$

Rozwiązaniem zadania dopasowania obrazów \mathbf{g}_1 oraz \mathbf{g}_2 jest znalezienie takiej funkcji $d(\mathbf{x})$ określającej jednoznacznie funkcję $t(\mathbf{x})$, która pozwoli dopasować odpowiadające sobie punkty obydwu obrazów. Można to zapisać w postaci wzoru (3.6).

$$\exists j \in (1, 2, \dots, m) \exists k \in (1, 2, \dots, m) : \mathbf{x}_{2k} = t(\mathbf{x}_{1j}) \mathbf{x}_{1j} \quad (3.6)$$

W związku z tym, że funkcja $t(\mathbf{x})$ jest wzajemnie jednoznaczna z funkcją $d(\mathbf{x})$, w dalszej części rozprawy ustalenie funkcji $t(\mathbf{x}_1)$ będzie traktowane równoważnie z ustaleniem funkcji $d(\mathbf{x}_1)$. Warto zaznaczyć, że znając współrzędne odpowiadających sobie punktów można łatwo wyznaczyć dopasowanie tych punktów. Podstawiając (3.2) do wybranego punktu ze wzoru (3.6), a następnie biorąc pod uwagę wzór (3.5) otrzymujemy:

$$\begin{aligned}
\mathbf{x}_{2k} = t(\mathbf{x}_{1j})\mathbf{x}_{1j} &= \begin{bmatrix} 1 & 0 & t_x(\mathbf{x}_{1j}) \\ 0 & 1 & t_y(\mathbf{x}_{1j}) \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_{1j} \\ y_{1j} \\ 1 \end{bmatrix} = \begin{bmatrix} x_{1j} + t_x(\mathbf{x}_{1j}) \\ y_{1j} + t_y(\mathbf{x}_{1j}) \\ 1 \end{bmatrix} \Leftrightarrow \\
& \begin{bmatrix} x_{2k} \\ y_{2k} \\ 1 \end{bmatrix} = \begin{bmatrix} x_{1j} + t_x(\mathbf{x}_{1j}) \\ y_{1j} + t_y(\mathbf{x}_{1j}) \\ 1 \end{bmatrix} \Rightarrow \\
d(\mathbf{x}) &= \begin{bmatrix} t_x(\mathbf{x}_{1j}) \\ t_y(\mathbf{x}_{1j}) \end{bmatrix} = \begin{bmatrix} x_{2k} - x_{1j} \\ y_{2k} - y_{1j} \end{bmatrix} = \begin{bmatrix} x_{2k} \\ y_{2k} \end{bmatrix} - \begin{bmatrix} x_{1j} \\ y_{1j} \end{bmatrix} \quad (3.7)
\end{aligned}$$

Zazwyczaj nie jest znane położenie punktów \mathbf{X} , ani ich przyporządkowanie do punktów $\mathbf{x}_{1j}, \mathbf{x}_{2k}$. Dopasowanie $d(\mathbf{x})$ można jednak wyznaczyć przy założeniu, że obrazy rzeczywistego punktu otoczenia \mathbf{X} są do siebie podobne. W rozprawie podobieństwo punktów obrazu definiuje się wprowadzając funkcję podobieństwa.

Niech dana będzie funkcja podobieństwa obrazów:

$$S = s(g_1(\mathbf{x}_1), g_2(\mathbf{x}_2)) \quad (3.8)$$

gdzie:

$g_1(\mathbf{x}_1)$ - intensywności obrazu g_1 dla wszystkich punktów \mathbf{x}_1 tego obrazu,

$g_2(\mathbf{x}_2)$ - intensywności obrazu g_2 dla wszystkich punktów \mathbf{x}_2 tego obrazu,

która osiąga swoje maksimum, gdy dla wszystkich punktów \mathbf{x}_1 obrazu g_1 względem odpowiednich punktów \mathbf{x}_2 obrazu g_2 spełnione są warunki (3.4).

Rozwiązanie zadania dopasowania obrazów można wtedy zdefiniować stosując (3.6) jako znalezienie takiej funkcji $d(\mathbf{x}_1)$, że:

$$S = s(g_1(\mathbf{x}_1), g_2(t(\mathbf{x}_1)\mathbf{x}_1)) \quad (3.9)$$

osiąga maksimum względem $d(\mathbf{x}_1)$.

Ilość stopni swobody funkcji $d(\mathbf{x})$ jest często redukowana w celu uproszczenia problemu. Dla ustalonej aplikacji można założyć z góry postać parametrów $t_x(\mathbf{x}), t_y(\mathbf{x})$. Wybór postaci funkcji $t(\mathbf{x})$ wpływa wprost na wymiarowość przestrzeni parametrów dopasowania i na złożoność obliczeniową algorytmu dopasowania obrazów.

Dwa przykłady różnych postaci funkcji $t_x(\mathbf{x}), t_y(\mathbf{x})$, przedstawiają równania (3.10) oraz (3.11), gdzie funkcja $t(\mathbf{x})$ posiada dwa stopnie swobody: c_x i c_y oraz równania (3.10) oraz (3.12), gdzie funkcja $t(\mathbf{x})$ posiada trzy stopnie swobody: a_x, a_y i α . Nazwy parametrów c_x i c_y zostały określone nieprzypadkowo w nawiązaniu do (2.3). W ten sposób

można zademonstrować, że za pomocą odpowiednio określonej funkcji dopasowania można sprowadzić obraz o znanym środku optycznym c_x i c_y do układu współrzędnych, gdzie środek optyczny znajdzie się w początku tego układu. Taka operacja znajduje zastosowanie w obliczu przyjętych założeń związanych z układem kamer. Przeprowadzenie jej dla dwóch obrazów o znanych środkach optycznych odpowiadających im kamer może być elementem składowym dopasowania w trybie online w celu fuzji obrazów.

Równania (3.10) oraz (3.11) przedstawiają wzory na sztywną translację obrazu. Równania (3.10) oraz (3.12) przedstawiają wzory na obrót obrazu wokół punktu.

Niech przykładowo:

$$t(x, y) = \begin{bmatrix} 1 & 0 & t_x(x, y) \\ 0 & 1 & t_y(x, y) \\ 0 & 0 & 1 \end{bmatrix}, d(x, y) = \begin{bmatrix} t_x(x, y) \\ t_y(x, y) \end{bmatrix} \quad (3.10)$$

Jeśli do (3.10) podstawimy:

$$\begin{aligned} t_x(x, y) &= c_x \\ t_y(x, y) &= c_y \end{aligned} \quad (3.11)$$

otrzymamy wzór na przekształcenie opisujące tzw. sztywne przesunięcie obrazu o wektor $[c_x, c_y]$.

Jeśli do (3.10) podstawimy natomiast:

$$\begin{aligned} t_x(x, y) &= a_x * \cos(\alpha) - a_y * \sin(\alpha) \\ t_y(x, y) &= a_x * \sin(\alpha) + a_y * \cos(\alpha) \end{aligned} \quad (3.12)$$

otrzymamy wzór na obrót obrazu wokół punktu $A = [a_x, a_y]$ o kąt α .

Niech \mathbf{D} będzie zbiorem wszystkich parametrów definiujących postać funkcji $d(\mathbf{x})$ pewnego przekształcenia obrazu $t(\mathbf{x})$. Korzystając z równania (3.9) możemy zdefiniować funkcję wyznaczającą parametry dopasowania \mathbf{D} jako:

$$\mathbf{D} = \arg \max_{\mathbf{D}} s(g_1(\mathbf{x}_1), g_2(\mathbf{t}(\mathbf{x}_1)\mathbf{x}_1)) \quad (3.13)$$

3.2 Metody wyznaczania dopasowania obrazów multimodalnych

Istnieje wiele metod dopasowania obrazów. Bez względu na algorytm, problem dopasowania obrazów multimodalnych można sprowadzić do znalezienia wiarygodnej funkcji $s(\mathbf{g}_1, \mathbf{g}_2)$ podobieństwa ((3.8)) obrazów różnego rodzaju $\mathbf{g}_1, \mathbf{g}_2$ oraz do ustalenia odpowiedniej postaci funkcji dopasowania $d(\mathbf{x})$ opisanej wzorem (3.5). Funkcja może posłużyć do przyporządkowania odpowiadających sobie punktów obrazów dopasowywanych.

Do standardowych, podręcznikowych [72, 73] funkcji wyznaczenia podobieństwa obrazów można zaliczyć wiele wskaźników mających proste podłoże statystyczne. Każdy z nich może implementować równanie (3.8) i posłużyć do znalezienia dopasowania obrazów zgodnie z równaniem (3.13). Do najważniejszych grup wskaźników podobieństwa można zaliczyć metody bazujące na sumie bezwzględnych różnic, sumie kwadratów różnic, konwolucji (czyli splocie), kowariancji (czyli konwolucji o średniej zerowej), znormalizowanej korelacji wzajemnej, czy sumie iloczynów wektorowych. W zakresie sumy bezwzględnych różnic można wyróżnić: sumę bezwzględnych różnic (SAD), znormalizowaną sumę bezwzględnych różnic (NSAD), sumę bezwzględnych różnic o średniej zerowej (ZSAD), znormalizowaną sumę bezwzględnych różnic o średniej zerowej (ZNSAD), ważoną sumę bezwzględnych różnic (WSAD), znormalizowaną ważoną sumę bezwzględnych różnic (WNSAD). W zakresie sumy kwadratów różnic można wyróżnić: sumę kwadratów różnic (SSD), znormalizowaną sumę kwadratów różnic (NSSD), sumę kwadratów różnic o średniej zerowej (ZSSD), znormalizowaną sumę kwadratów różnic o średniej zerowej (ZNSSD). Można też wyróżnić znormalizowaną korelację wzajemną (NCC) oraz znormalizowaną korelację wzajemną o średniej zerowej. Ostatecznie można wspomnieć o sumie iloczynów wektorowych (SCP) oraz znormalizowanej sumie iloczynów wektorowych (NSCP).

Warto zauważyć, że większość algorytmów dopasowania obrazów multimodalnych przeznaczona jest do zastosowania dla obrazów medycznych [74, 75]. Jednakże z uwagi na rosnące w ostatnich latach zainteresowanie systemami monitorowania wizyjnego oraz większą dostępność kamer termowizyjnych, liczba artykułów opisujących rozwiązania techniczne związane z dopasowaniem i fuzją obrazów wizyjnych i termowizyjnych znacznie wzrosła. Wśród tych publikacji znajduje się wiele algorytmów zaprojektowanych do wykrywania i analizy sylwetki ludzkiej [76, 52, 57, 77], czasem przy założeniu, że układ kamer multimodalnych nie zmienia swojego położenia [78, 52]. Niektóre artykuły przedstawiają rozwiązania w zakresie teledetekcji z wykorzystaniem układu kamer wizyjnej i termowizyjnej [79, 80]. W literaturze można również znaleźć artykuły przedstawiające problemy związane z dopasowaniem i fuzją obrazów multimodalnych dla potrzeb monitorowania [81, 82, 83, 84]. Wśród nich znajdują się również prace autora opisane w [69]. Warto podkreślić, że problem dopasowania obrazów: wizyjnego i termowizyjnego jest nadal aktualny i stanowi wyzwanie

w obliczu bieżącego stanu wiedzy [80]. Na szczególną uwagę zasługują również prace prowadzące do ujednoczenia podejścia badawczego w zakresie oceny algorytmów. Zmierzają one do opracowania spójnej metodyki porównywania algorytmów na podstawie wspólnej bazy danych testowych uzupełnionej o idealne wartości dopasowania. Jeszcze kilka lat temu tego typu bazy danych dostępne były jedynie dla multimodalnych obrazów medycznych i nie mogły zostać zastosowane do porównywania wyników algorytmów rozwijanych w celu monitorowania.

Jedna z najczęściej spotykanych w literaturze klasyfikacji stawia naprzeciwko siebie metody bazujące bezpośrednio na intensywnościach pikseli (wokseli) oraz metody bazujące na ekstrakcji cech charakterystycznych obrazów, takich jak np. narożniki czy krawędzie. Każde z tych podejść ma swoje wady i zalety opisywane szeroko w literaturze.

Z jednej strony metody bazujące na wyznaczeniu cech charakterystycznych wymagają zastosowania niewrażliwego na szumy algorytmu ekstrakcji cech charakterystycznych [85]. Operacja ta jest wykonywana w początkowej fazie dopasowania w celu wyznaczenia wspólnej reprezentacji obrazów w ustalonej przestrzeni wartości (cech) [74]. W celu realizacji postulatu niewrażliwości na szumy, jedne z najbardziej rozpowszechnionych deskryptorów cech, takie jak np. SIFT [86] czy GLOH [87], opierają swoje działanie na wyznaczaniu gradientów intensywności i ich histogramów. Oczywistym problemem, który musi zostać rozwiązany przy próbie zastosowania tej klasy metod, jest wybór cech, które będą reprezentowały obraz. Dlatego w celu poprawienia jakości działania opracowanego algorytmu wymagana jest czasem interwencja człowieka i ręczne wskazanie poprawnie dopasowanych cech [88, 89, 90]. Dodatkowe trudności pojawiają się przy próbach rozszerzenia działania algorytmu na dokładność subpikselową lub w przypadku bardziej skomplikowanych transformacji obrazów. Okazuje się to często niemożliwe z uwagi na złożoność obliczeniową tak zmodyfikowanego problemu [90]. Z uwagi na te i inne problemy wielu autorów podziela opinię, że standardowe deskryptory cech nie dają się łatwo zastosować do dopasowania obrazów wizyjnych i termowizyjnych [78, 91, 92, 93, 94]. Propozycję rozwiązania tego problemu przedstawiono w postaci deskryptora cech MM-SURF w [82] bazując na klasycznym deskrypcie SURF. Inny deskryptor cech o akronimie SI-PIIFD przedstawiono w [83]. Autorzy przedstawiają jego działanie jako niezmiennicze w odniesieniu do różnic w skali pomiędzy dopasowywanymi obrazami wizyjnymi i termowizyjnymi.

Pomimo to w zakresie dopasowania tego typu obrazów często stosowane są metody filtracji krawędziowej. W [95] opisano algorytm dopasowania obrazów pochodzących z kamery wizyjnej i termowizyjnej na podstawie reprezentacji wielorozdzielczej z wykorzystaniem ekstrakcji krawędzi. W celu wyznaczenia dopasowania autorzy minimalizują błąd skonstruowany na podstawie przyporządkowania wartości odpowiadających sobie krawędzi na obrazach wizyjnym i termowizyjnym w ujęciu wielorozdzielczym. Później w [96] opisano zagadnienia podobne do [95], jednakże poza algorytmem filtracji krawędziowej opracowanym w celu dopasowania obrazów wizyjnych do termowizyjnych przedstawiono

nowe podejście do wyznaczenia dopasowania na podstawie tej filtracji. Podobne podejście przedstawiono również w [97], gdzie autorzy wykorzystują algorytm bazujący na algorytmie Canny'ego w celu dopasowania pojedynczych statków pomiędzy obrazami z kamer wizyjnej i termowizyjnej. W [81] przedstawiono nowy deskryptor bazujący na ekstrakcji fragmentów krawędzi przy wykorzystaniu metody *phase congruency*. Zaproponowano tam również metodę dopasowania odpowiadających sobie fragmentów krawędzi. W [79, 80] przedstawiono podobne podejście z wykorzystaniem metody *phase congruency* do detekcji krawędzi, a następnie pokazano, jak dopasować odpowiadające sobie punkty charakterystyczne obrazów. W [84] przedstawiono natomiast sposób wykorzystania biblioteki OpenCV oraz ekstrakcję krawędzi z zastosowaniem metod morfologicznych w celu dopasowania odpowiadających sobie punktów obrazów wizyjnych i termowizyjnych za pomocą SURF.

Inną grupę stanowią metody bazujące na intensywności. Opierają one swoje działanie na ustalonej zależności pomiędzy intensywnościami pikseli obrazów [98]. Większość algorytmów zamiast stosować bezpośrednio intensywność pikseli (wokseli) zakłada istnienie pewnej statystycznej bądź określonej funkcją (np. wielomianową [99, 100] lub przedziałami liniową [101]) zależności między intensywnościami obrazów dopasowywanych [85]. Zależność tę można wyrazić w postaci funkcji podobieństwa obrazów. Trudność w takim podejściu stanowi to, że założona zależność musi bardzo dobrze odzwierciedlać rzeczywistość. Dodatkowo trudno zazwyczaj zmodyfikować algorytm tak, żeby radził sobie z obrazami o innej niż założona zależności pomiędzy intensywnościami obrazów [102, 90]. Pomimo problemów metody z tej grupy okazują się najskuteczniejsze w zastosowaniu do obrazów wizyjnych i termowizyjnych. Wśród najchętniej stosowanych metod w ostatnich latach można znaleźć algorytmy bazujące na wyznaczaniu entropii. Wśród nich znajduje się najbardziej popularna - metoda informacji wzajemnej (ang. *Mutual Information*, w skrócie MI). Do grupy najczęściej stosowanych funkcji podobieństwa obrazów można zaliczyć ponadto miary SSD, *Correlation Coefficient* (CC), *Correlation Ratio* (CR) [99] oraz metody bazujące na wyznaczaniu regresji [100, 101].

Podsumowując różne algorytmy dopasowania obrazów warto zauważyć, że większość metod zakłada niezmienną względem przesunięcia zależność pomiędzy intensywnościami obrazów porównywanych [103]. Jednym z algorytmów, które rozluźniają to założenie, jest *Statistical Parametric Mapping* (SPM), w którym dopasowanie przebiega równocześnie z procesem normalizacji intensywności obrazów [103, 104, 105]. Na szczególną uwagę zasługują algorytmy opierające swoje działanie na MI. Bazują one na statystycznej zależności pomiędzy wartościami intensywności obrazów. Warto zauważyć, że nie zakładają istnienia funkcjonalnej zależności pomiędzy tymi wartościami [106]. Po raz pierwszy zostały zastosowane do dopasowania obrazów niezależnie przez dwie grupy autorów: Viola i Wells [107, 108] oraz Collignon i Maes [109, 110]. Rozwiązują wiele problemów związanych z dopasowaniem obrazów multimodalnych. Zarówno MI jak i metody pochodne (np. maksymalizacja informacji

wzajemnej, MMI) początkowo znalazły zastosowanie w dopasowaniu obrazów dla celów medycznych [111, 112]. Dopiero później ich zakres aplikacji został rozszerzony np. do teledetekcji [113, 114]. Istotną zaletą MI w odniesieniu do obrazów multimodalnych jest wyznaczanie dopasowania bezpośrednio na podstawie intensywności pikseli [115]. Pomimo swojej popularności, istnieje wiele zastrzeżeń dotyczących stosowania tej grupy metod. Miara MI opiera się tylko na zależnościach statystycznych pomiędzy obrazami, bez brania pod uwagę informacji niesionej przez widoczne struktury [116, 90, 117]. Dodatkowo przestrzeń dopasowań wyznaczona przy wykorzystaniu MI może być znacząco niemonotoniczna [103, 74]. Utrudnia to znalezienie maksimum funkcji podobieństwa – algorytmy przeszukujące mogą łatwo utknąć w lokalnych ekstremach funkcji [115, 75, 117]. Niektórzy autorzy wspominają również o tym, że metody bazujące na MI są złożone obliczeniowo z uwagi na konieczność obliczenia entropii wzajemnych i brzegowych [90, 118, 119, 120]. Szczególnie podkreślane są wady wykorzystania MI w celu dopasowania obrazów termowizyjnych – takie dopasowanie często okazuje się niemożliwe do wyznaczenia z uwagi na płaską przestrzeń poszukiwań [90]. W związku z tymi wadami wielu autorów podziela opinię, że należy nadal poszukiwać lepszych algorytmów [117, 115]. Wiele obecnie pojawiających się prac na ten temat skupionych jest na rozszerzeniu metod MI o możliwość wykorzystania informacji o widocznych strukturach [90].

Odbiegając od powyższej klasyfikacji warto odnieść się do proponowanych w ostatnich latach rozwiązań problemu dopasowania obrazów z wykorzystaniem głębokich sieci konwolucyjnych (ang. *Convolutional Neural Networks*, w skrócie CNN). Sieci te stanowią pewną klasę głębokich sieci uczących (ang. *Deep Neural Networks*, w skrócie DNN) i w ostatnich latach ich stosowanie do rozwiązania problemu dopasowania obrazów cechuje się dużą popularnością ([121, 122, 123, 124, 125, 126]). Z uwagi na rozproszony sposób implementacji sieci CNN podkreśla się ich zalety w zakresie wyznaczania gęstej mapy dopasowań [127]. Dopasowanie obrazów za pomocą CNN można wyznaczać zarówno pomiędzy obrazami monomodalnymi (przykładowo [128], czy [129]), jak i pomiędzy obrazami multimodalnymi, gdzie są głównie wykorzystywane do dopasowania obrazów medycznych (przykładowo [130, 131, 132]). Sieci CNN znajdują również zastosowanie do wyznaczenia samej funkcji podobieństwa [127] pomiędzy fragmentami porównywanych obrazów. Niektórzy autorzy podkreślają ich zalety z uwagi na to, że potrafią uchwycić skomplikowane zależności pomiędzy intensywnościami obrazów multimodalnych [130].

Ciekawą pracę w zakresie CNN przedstawiono w [127], gdzie autorzy pokazali, w jaki sposób wyznaczyć dopasowanie obrazów z wykorzystaniem ich cech w dziedzinie częstotliwości. Podobne podejście w zakresie wykorzystania zapisu obrazu w dziedzinie częstotliwości przedstawiono w [121] w zakresie współczynników transformaty Fouriera obrazu.

W ramach prac nad tematem rozprawy przetestowano wybrane podejścia do dopasowania obrazów z wykorzystaniem CNN, ale nie przyniosły one zadowalających rezultatów. Napotkane

problemy dotyczyły głównie jakości działania testowanych sieci, ale były związane również z ich znaczną złożonością obliczeniową.

3.3 Metody wyznaczania dopasowania obrazów multimodalnych z zastosowaniem korelacji fazowej

Szczególną metodą obliczeniową stosowaną do dopasowania obrazów jest metoda korelacji fazowa (w rozprawie przyjęto zgodnie z pracą [133] stosowanie skrótu POC od ang. *Phase-Only Correlation*). Metoda ta ma wyjątkowe znaczenie, ponieważ skuteczność jej działania dla obrazów zgodnych z założeniami przyjętymi w rozprawie została potwierdzona i zaprezentowana przez autora w publikacji [68]. Klasyczna metoda POC służy do wyznaczenia jednego wektora przesunięcia $[d_x, d_y]$ w procesie dopasowania pewnego obrazu do innego obrazu referencyjnego. Pierwsze wzmianki o jej zastosowaniu dla potrzeb dopasowania obrazów pochodzą z roku 1975, w którym autorzy Kuglin i Hines [134, 135] opublikowali opis metody wyznaczenia współczynników translacji dopasowania obrazów. W roku kolejnym opublikowane zostało [136] stosowane do dziś rozwinięcie tego podejścia z wykorzystaniem transformaty Fouriera-Mellina. Autorzy zaproponowali metodę wyznaczenia dopasowania niewrażliwą na różnice we współczynniku skali i obrotu porównywanych obrazów (ang. *Rotation Invariant Phase-Only Correlation*, w skrócie RIPOC).

W latach kolejnych metoda korelacji fazowej znalazła zastosowanie do estymacji ruchu w sekwencji obrazów z jednej kamery [137]. W zakresie estymacji ruchu wielu obiektów od początku stosowania metody korelacji fazowej autorzy borykali się z problemem przypisania wartości konkretnych przesunięć do odpowiadających im obiektów (fragmentów obrazu).

Metodę korelacji fazowej rozwijano w kolejnych latach po jej pierwszym opublikowaniu, a opisywane w literaturze sytuacje dotyczą głównie dopasowania obrazów monomodalnych. Podobne jak w pracy [136] wykorzystanie korelacji fazowej przedstawiono w publikacji [138]. Sposób wyznaczenia rotacji i przesunięcia przedstawiono również w [139]. Opisana procedura dopasowania polega na wykorzystaniu własności transformaty 2D Fouriera dotyczącej obrotu w dziedzinie obrazu i w dziedzinie częstotliwości.

Metoda korelacji fazowej jest nadal rozwijana w celu wykorzystania do rozwiązania aktualnych problemów z dziedziny przetwarzania obrazów, przykładowo w pracy [140] przedstawiono podstawy związane z korelacją fazową i jej zastosowanie do wyznaczenia dopasowania na monomodalnych obrazach termowizyjnych w zastosowaniach militarnych. W pracy [141] opisano dodatkowo metodę wyznaczenia skali za pomocą reprezentacji obrazu w układzie współrzędnych log-polar. W pracy [142] zaproponowano poprawę jakości wyznaczenia największego maksimum korelacji fazowej z wykorzystaniem metod segmentacji tła. W [143] opisano kompleksowo różne sposoby wykorzystania korelacji fazowej wliczając w to wyznaczenie wielu wartości dopasowania, lokalizację przestrzenną tych dopasowań, metody

interpolacji oraz wyznaczenie rotacji i skali pomiędzy obrazami. Później w [144, 145] zaproponowano sposób na zmniejszenie kosztu obliczeniowego metody poprzez zastosowanie binaryzacji obrazów. W [146] zamieszczono sposób opisu matematycznego wyznaczenia dopasowania dla dwóch przesunięć pomiędzy obrazami z tej samej kamery. W [147] możemy znaleźć zaawansowane podejście do ustalenia filtracji w dziedzinie częstotliwości dla potrzeb poprawy jakości rozpoznawania twarzy dla obrazów monomodalnych. Ciekawą pracę można również znaleźć w [148]. Autorzy próbują zastosować technikę RIPOC do dopasowania rotacji i skali dla obrazów monomodalnych. Zauważyli, że jest to technika zwracająca dużo błędnych rezultatów. Stosują ją do zawężenia przestrzeni poszukiwań właściwego dopasowania.

Dopasowanie obrazów multimodalnych metodą korelacji fazowej jest o wiele trudniejsze. Dlatego dużo rzadziej można znaleźć jej zastosowanie dla takich aplikacji. Jeden z przykładów przedstawiony jest w [149], gdzie pokazane zostało dopasowanie różnych składowych obrazów multispektralnych (MSI). Z kolei w [150] autorzy przedstawiają system fuzji obrazów wizyjnych i termowizyjnych z wykorzystaniem wyznaczenia dopasowania metodą RIPOC dla potrzeb rozpoznawania twarzy. Obrazy wejściowe poddane są początkowo ekstrakcji krawędzi metodą Canny'ego. Korelację fazową do dopasowania obrazów wizyjnych i termowizyjnych zastosowano również w [151], gdzie została wykorzystana do wyznaczenia lokalnego podobieństwa obrazów w okolicy punktów charakterystycznych wyznaczonych za pomocą SURF.

Dodatkową zaletą POC na tle wymagań przedstawionych w rozdziale 1.2 stanowi jej stosunkowo niska złożoność obliczeniowa oraz łatwość w implementacji na układach rozproszonego przetwarzania danych typu GPGPU. Wśród artykułów z ostatnich lat znaleźć można wiele dotyczących implementacji metody korelacji fazowej na platformie tego typu. Należą do nich między innymi prace [152, 153, 154]. W [152, 154] autorzy wykorzystują metodę do wyznaczenia dopasowania stereowizyjnego pomiędzy obrazami monomodalnymi w zastosowaniu do pomiarów kształtu obiektów 3D. Do oceny wyników stosują pomiar czasu wykonywania obliczeń dla implementacji na różnych kartach graficznych - GeForce i Radeon. Z kolei w [153] zademonstrowana jest implementacja algorytmu dopasowania obrazów monomodalnych dla potrzeb badania rozdzielania faz ciekłych i gazowych gazów w środowisku o niskiej grawitacji. Ewaluacja przedstawionej metody przeprowadzona jest na podstawie porównania czasu wykonania obliczeń w zależności od ilości pikseli obrazów dla jednostek typu CPU i GPGPU. Ewaluacja poprawności funkcjonowania algorytmu została przeprowadzona na podstawie zgodności wyników ze znanym modelem teoretycznym. W [155] autorzy wykorzystują metodę POC w zadaniu estymacji ruchu dla obrazów monomodalnych o relatywnie dużych rozdzielczościach. W ramach publikacji opisana jest implementacja na GPU. Oceny działania opracowanej przez siebie metody autorzy dokonują standardowo na podstawie czasu obliczeń implementacji na GPU. W [156] można znaleźć przykład implementacji algorytmu na bazie RIPOC przeznaczonej na jednostkę typu GPGPU z wykorzystaniem biblioteki programistycznej

CUFFT. Autorzy przedstawiają wyniki dopasowania dla wyraźnych obrazów monomodalnych. W [157] znajdziemy natomiast porównanie sprzętowych akceleracji obliczeń napisanych dla architektur typu PLD i GPGPU względem implementacji na CPU. Autorzy porównują dwa standardowe algorytmy dopasowania. Pierwszy algorytm wykorzystuje metodę POC do wyznaczenia dwóch parametrów translacji. Drugi natomiast wykorzystuje metodę RIPOC do dodatkowego wyznaczenia parametru obrotu obrazu. Proponowaną w artykule aplikacją opisanych algorytmów jest śledzenie ruchu na monomodalnych obrazach światła dziennego.

Pomimo opisanych zalet stosowania metody korelacji fazowej, ma ona również istotne wady. Metody nie można wykorzystać bezpośrednio do dopasowania obrazów wizyjnych i termowizyjnych z uwagi na brak jednorodnej w dziedzinie obrazu zależności pomiędzy intensywnościami odpowiadających sobie obiektów. Dlatego konieczne jest zastosowanie takich operacji wstępnego przetwarzania obrazów, które poprawią jednorodność tej zależności. Dodatkowo stosując korelację fazową w sposób klasyczny otrzymujemy zestaw dopasowań bez przyporządkowania do niego konkretnych obszarów na obrazach dopasowywanych. Stanowi to istotne utrudnienie w zakresie kalibracji układu kamer. W rozdziale 4 pokazano, w jaki sposób zaprojektować metodę samokalibracji w celu rozwiązania tego problemu.

W odniesieniu do korelacji fazowej wielu autorów podkreśla, że w celu zwiększenia dokładności metody konieczne jest stosowanie techniki oceny jakości wyznaczonego dopasowania na podstawie wysokości znalezionej maksimum. Dzięki temu można odrzucić potencjalnie błędne wartości wyznaczonego dopasowania. Przykład takiego podejścia można znaleźć w [155] w odniesieniu do obrazów monomodalnych. W [147] przedstawiona jest aplikacja metody POC do rozpoznawania twarzy dla obrazów monomodalnych. Ocena jakości wyznaczana jest jako stosunek wartości energii maksimum do całkowitej energii w przestrzeni odwrotnej transformaty Fouriera. Z kolei w [148] autorzy stosują prostą miarę podobieństwa bazującą na SSD w odniesieniu do oceny działania metody RIPOC do dopasowania rotacji i skali dla obrazów monomodalnych.

3.4 Podsumowanie

Algorytmy automatycznego dopasowania obrazów multimodalnych mogą znaleźć zastosowanie do odnalezienia na obrazach wizyjnym i termowizyjnym odpowiadających sobie obiektów. Dzięki temu możliwa jest estymacja wybranych parametrów układu kamer. W związku z tym możliwe jest wykorzystanie metod automatycznego dopasowania obrazów, aby zrealizować cel rozprawy, czyli opracować metodę samokalibracji układu kamer wizyjnej i termowizyjnej. Mimo potwierdzonej skuteczności różnych metod automatycznego dopasowania obrazów, wielu autorów ciągle poszukuje nowych, lepszych algorytmów. Szczególnie problem dopasowania obrazów: wizyjnego i termowizyjnego jest nadal aktualny i stanowi wyzwanie w obliczu bieżącego stanu wiedzy, co bardzo mocno podkreślono w

publikacji [80] z roku 2019. Autorzy tej publikacji tłumaczą trudności w dopasowaniu obrazów tego typu znaczącymi różnicami w kontraście obrazów oraz dużymi zakresami wartości dopasowań (w związku z tym metody muszą przeszukiwać duże przestrzenie parametrów dopasowania pod kątem maksymalizacji funkcji podobieństwa obrazów). Wśród algorytmów automatycznego dopasowania obrazów na szczególną uwagę zasługują metody bazujące na korelacji fazowej. Stosowanie tych metod w odniesieniu do obrazów multimodalnych jest utrudnione, ale możliwe, co zostało potwierdzone w ramach badań autora rozprawy opisanych w [68].

Ocena jakości dopasowania w zakresie walidacji i porównania różnych metod w trybie offline nie jest problemem przy sztywnej transformacji [158]. Inaczej sprawa wygląda z algorytmami mającymi działać i wykorzystywać ocenę jakości online. Dla obrazów wizyjnych i termowizyjnych do najpopularniejszych metod tego typu zalicza się wyznaczenie części wspólnej nakładanych obiektów [76, 91, 159, 160, 77, 161], czy porównywanie wyników wyznaczonej i odwrotnej do niej transformacji dopasowania [162, 56]. W odniesieniu do korelacji fazowej znaleźć można różne podejścia, w których standardowo wykorzystuje się wysokość znalezionej maksimum w macierzy wyznaczonej przez odwrotną transformatę Fouriera [155, 147, 148]. Innym wartym uwagi rozwiązaniem problemu jest wykorzystanie informacji o ruchu obrazów dopasowywanych w sekwencji wideo [94, 163, 160]. W szczególności należy wyróżnić tu prace integrujące śledzenie obiektów z dopasowaniem [77, 57]. Popularną metodą odrzucania parametrów błędnie dopasowanych jest RANSAC. Przykłady jego użycia można znaleźć w pracach [76, 162, 91, 159, 77, 161, 160].

Rozdział 4

Opracowana metoda samokalibracji układu kamer: wizyjnej i termowizyjnej

W poprzednich rozdziałach rozprawy opisano wprowadzenie teoretyczne do poruszanych zagadnień badawczych oraz dokonano przeglądu metod samokalibracji i automatycznego dopasowania obrazów. W bieżącym rozdziale opisano sposób realizacji celu rozprawy, czyli opracowanie metody samokalibracji układu kamer wizyjnej i termowizyjnej przy ograniczeniach wynikających z przedstawionych w rozdziale 1.2 założeń. Dalsza część bieżącego rozdziału stanowi opis opracowanych oryginalnych rozwiązań autora i nawiązuje do przyjętej w rozprawie tezy.

W podrozdziale 4.1 przedstawiono od strony teoretycznej opracowaną metodę samokalibracji układu dwóch kamer wizyjnej i termowizyjnej. Wykazano, że w takim przypadku kalibrację można podzielić na dwa etapy. Pierwszy etap ma na celu ustalenie parametrów skali obrazu termowizyjnego względem wizyjnego. Estymacja tych parametrów następuje na podstawie automatycznego dopasowania w trybie kalibracji obrazów z chwili bieżącej do obrazów z chwili poprzedniej odpowiednio dla kamery wizyjnej i termowizyjnej (dopasowanie monomodalne). W etapie drugim ustalane są parametry opisujące linię dopasowania online odpowiadających sobie obiektów obrazu wizyjnego i termowizyjnego. Estymacja tych parametrów następuje na podstawie automatycznego dopasowania w trybie kalibracji odpowiadających sobie obiektów obrazów wizyjnego i termowizyjnego (dopasowanie multimodalne).

Wobec konieczności wyznaczenia automatycznego dopasowania obrazów w trybie kalibracji w podrozdziale 4.2 przedstawiono opracowaną metodę dopasowania dla potrzeb automatycznej kalibracji układu kamer wizyjnej i termowizyjnej. Zaprezentowano tam również uzasadnienie wykorzystania korelacji fazowej do wyznaczenia dopasowania w trybie kalibracji dla wielu obiektów jednocześnie widocznych na obrazach z kamery wizyjnej i termowizyjnej.

W podrozdziale 4.3 przedstawiono opracowaną metodę samokalibracji od strony algorytmicznej.

4.1 Podstawy teoretyczne metody samokalibracji

W bieżącym podrozdziale zaprezentowano podstawy teoretyczne opracowanej metody samokalibracji układu kamer multimodalnych. Przedstawiono, w jaki sposób ogólny model układu dwóch kamer dostosowano do założeń przyjętych w rozprawie i podzielono na dwie części, co ujęto w twierdzeniu 1 (p. 4.1.2). Pokazano, że dzięki temu samokalibrację można podzielić na dwa etapy: (I) ustalenie skali obrazu termowizyjnego względem wizyjnego oraz (II) ustalenie parametrów linii dopasowania obrazu wizyjnego do termowizyjnego po dostosowaniu skali.

Udowodniono następnie w twierdzeniu 2 (p. 4.1.4), że dopasowanie odpowiadających sobie punktów obrazów multimodalnych dla założeń przyjętych w rozprawie zmienia się liniowo w zależności od odległości układu kamer od obserwowanego obiektu generującego te punkty.

Stanowi to istotny rezultat rozprawy, ponieważ dzięki temu problem dopasowania online obrazów w celu dalszej fuzji online można zredukować do poszukiwania największego ich podobieństwa z jednym stopniem swobody. Upraszcza to znacząco fuzję obrazów wizyjnego i termowizyjnego. Dzięki temu upraszcza również zadanie monitorowania otoczenia przy użyciu głowicy kamer wizyjnej i termowizyjnej.

4.1.1 Przyjęty model układu kamer

W bieżącym podrozdziale przedstawiono w jaki sposób ogólny model układu kamer ujęty w podrozdziale 2.1 można uprościć z uwzględnieniem założeń przyjętych w rozprawie.

Parametry wewnętrzne

W podrozdziale 2.1.1 przedstawiono ogólny opis parametrów wewnętrznych układu kamer. W odniesieniu do założeń przyjętych w rozprawie, zgodnie ze wzorem (2.3) parametry wewnętrzne kamery wizyjnej i termowizyjnej można opisać za pomocą macierzy K_{tv} zdefiniowanej wzorem (4.1) oraz macierzy K_{ir} zdefiniowanej wzorem (4.2).

$$\mathbf{K}_{tv} = \begin{bmatrix} f_{tvx} & 0 & c_{tvx} \\ 0 & f_{tvy} & c_{tvy} \\ 0 & 0 & 1 \end{bmatrix} \quad (4.1)$$

gdzie:

\mathbf{K}_{tv} – macierz wewnętrzna kamery wizyjnej,

f_{tvx}, f_{tvy} – ogniskowe kamery wizyjnej w kierunkach x i y ,

c_{tvx}, c_{tvy} – przesunięcie osi optycznej względem środka matrycy CCD kamery wizyjnej.

Oraz:

$$\mathbf{K}_{ir} = \begin{bmatrix} f_{irx} & 0 & c_{irx} \\ 0 & f_{iry} & c_{iry} \\ 0 & 0 & 1 \end{bmatrix} \quad (4.2)$$

gdzie:

\mathbf{K}_{ir} – macierz wewnętrzna kamery termowizyjnej,

f_{irx}, f_{iry} – ogniskowe kamery termowizyjnej w kierunkach x i y ,

c_{irx}, c_{iry} – przesunięcie osi optycznej względem środka matrycy CCD kamery termowizyjnej.

Macierz zasadnicza

W podrozdziale 2.1.2 przedstawiono macierz zasadniczą \mathbf{E} opisującą wzajemne położenie układu dwóch kamer. W ogólnym przypadku macierz \mathbf{R} opisuje dowolny obrót w euklidesowej przestrzeni trójwymiarowej. Jednakże dla założonego układu kamer przyjmuje się przypadek uproszczony. Zakłada się, że macierz obrotu jest jednostkowa, tzn. nie występuje obrót między kamerami. Takie założenie uzasadnione jest parametrami konstrukcyjnymi głowicy kamer multimodalnych przedstawionymi w postaci ogólnych założeń w rozdziale 1.2. W związku z tym macierz \mathbf{R} jest opisana wzorem (4.3).

$$\mathbf{R} = \mathbf{I} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (4.3)$$

Biorąc pod uwagę (2.5) oraz (4.3), wzór (2.4) można rozwinąć do postaci (4.4).

$$\mathbf{E} = [\mathbf{t}]_{\times} \mathbf{R} = \begin{bmatrix} 0 & -t_z & t_y \\ t_z & 0 & -t_x \\ -t_y & t_x & 0 \end{bmatrix} \quad (4.4)$$

Macierz fundamentalna

W podrozdziale 2.1.3 przedstawiono ogólną postać macierzy fundamentalnej. Nawiązując do założeń przyjętych w rozprawie z zastosowaniem oznaczeń ir i tv równanie (2.9) można przedstawić w postaci (4.5).

$$\mathbf{x}_{ir}^T \mathbf{F} \mathbf{x}_{tv} = 0 \quad (4.5)$$

gdzie:

\mathbf{x}_{tv} – punkt na obrazie z kamery wizyjnej,

\mathbf{x}_{ir} – odpowiadający punkt na obrazie z kamery termowizyjnej.

oraz:

$$\mathbf{x}_{tv} = \begin{bmatrix} x_{tv} \\ y_{tv} \\ 1 \end{bmatrix} \quad (4.6)$$

$$\mathbf{x}_{ir} = \begin{bmatrix} x_{ir} \\ y_{ir} \\ 1 \end{bmatrix} \quad (4.7)$$

gdzie:

x_{ir}, y_{ir} – współrzędne w dziedzinie obrazu punktu na obrazie z kamery termowizyjnej,

x_{tv}, y_{tv} – współrzędne w dziedzinie obrazu punktu na obrazie z kamery wizyjnej odpowiadającego punktowi (x_{ir}, y_{ir}) .

Analogicznie wzór (2.7) można przekształcić do postaci (4.8).

$$\mathbf{F} = \mathbf{K}_{ir}^{-T} \mathbf{E} \mathbf{K}_{tv}^{-1} \quad (4.8)$$

gdzie:

\mathbf{K}_{ir} – macierz wewnętrzna kamery termowizyjnej,

\mathbf{K}_{tv} – macierz wewnętrzna kamery wizyjnej.

Wzór (4.8) można rozwinąć stosując wzory (4.4), (4.1), (4.2), co można przedstawić w postaci wyprowadzenia (4.9).

$$\begin{aligned} \mathbf{F} &= \mathbf{K}_{ir}^{-T} \mathbf{E} \mathbf{K}_{tv}^{-1} = \\ &= \begin{bmatrix} f_{irx} & 0 & c_{irx} \\ 0 & f_{iry} & c_{iry} \\ 0 & 0 & 1 \end{bmatrix}^{-T} \begin{bmatrix} 0 & -t_z & t_y \\ t_z & 0 & -t_x \\ -t_y & t_x & 0 \end{bmatrix} \begin{bmatrix} f_{tvx} & 0 & c_{tvx} \\ 0 & f_{tvy} & c_{tvy} \\ 0 & 0 & 1 \end{bmatrix}^{-1} = \\ &= \begin{bmatrix} 1/f_{irx} & 0 & -c_{irx}/f_{irx} \\ 0 & 1/f_{iry} & -c_{iry}/f_{iry} \\ 0 & 0 & 1 \end{bmatrix}^T \begin{bmatrix} 0 & -t_z & t_y \\ t_z & 0 & -t_x \\ -t_y & t_x & 0 \end{bmatrix} \begin{bmatrix} 1/f_{tvx} & 0 & -c_{tvx}/f_{tvx} \\ 0 & 1/f_{tvy} & -c_{tvy}/f_{tvy} \\ 0 & 0 & 1 \end{bmatrix} \quad (4.9) \end{aligned}$$

W przypadku gdy przesunięcie wzajemne kamer wzdłuż osi Z jest zerowe, możemy w (4.9) przyjąć, że parametr t_z jest pomijalny. Macierz \mathbf{F} przyjmuje wtedy nieco prostszą postać (4.10). Takie założenie uzasadnione jest parametrami konstrukcyjnymi głowicy kamer multimodalnych przedstawionymi w postaci ogólnych założeń w rozdziale 1.2.

$$\mathbf{F} = \begin{bmatrix} 1/f_{irx} & 0 & -c_{irx}/f_{irx} \\ 0 & 1/f_{iry} & -c_{iry}/f_{iry} \\ 0 & 0 & 1 \end{bmatrix}^T \begin{bmatrix} 0 & 0 & t_y \\ 0 & 0 & -t_x \\ -t_y & t_x & 0 \end{bmatrix} \begin{bmatrix} 1/f_{tvx} & 0 & -c_{tvx}/f_{tvx} \\ 0 & 1/f_{tvy} & -c_{tvy}/f_{tvy} \\ 0 & 0 & 1 \end{bmatrix} \quad (4.10)$$

W modelu (4.10) występuje 10 parametrów koniecznych do estymacji podczas kalibracji. Ilość tych parametrów można jednak jeszcze zredukować. W rozprawie przyjęto założenie, że na podstawie obrazów z kalibrowanego układu kamer wyznaczona zostanie ich fuzja. W związku z tym parametry ustalane w trakcie kalibracji muszą pozwolić jedynie na takie przekształcenie obrazów, aby możliwe było nałożenie odpowiadających sobie obiektów w ramach fuzji. W celu uproszczenia procesu kalibracji przyjęto obraz wizyjny jako referencyjny - jego położenie i skala nie ulegają zmianie. W związku z tym, aby nałożyć na siebie obrazy wizyjny i termowizyjny, konieczne jest dobranie skali i przesunięcia obrazu termowizyjnego względem wizyjnego. Takie założenia można wyrazić w postaci równań (4.11).

$$\begin{aligned} f_{tvx} &= 1 \\ f_{tvy} &= 1 \end{aligned} \quad (4.11)$$

i ostatecznie otrzymujemy model układu kamer przedstawiony w postaci wzoru (4.12), który odpowiada założeniom przyjętym w rozprawie.

$$\mathbf{F} = \begin{bmatrix} 1/f_{irx} & 0 & -c_{irx}/f_{irx} \\ 0 & 1/f_{iry} & -c_{iry}/f_{iry} \\ 0 & 0 & 1 \end{bmatrix}^T \begin{bmatrix} 0 & 0 & t_y \\ 0 & 0 & -t_x \\ -t_y & t_x & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & -c_{tvx} \\ 0 & 1 & -c_{tvy} \\ 0 & 0 & 1 \end{bmatrix} \quad (4.12)$$

W dalszej części rozdziału 4 przedstawiono, w jaki sposób macierz fundamentalną opisaną wzorem (4.12) wykorzystano w metodzie samokalibracji układu kamer wizyjnej i termowizyjnej.

4.1.2 Podział modelu na dwie części

Ustalenie parametrów macierzy \mathbf{F} równania (4.12) można podzielić na dwa etapy i wykonać sekwencyjnie. W pierwszym etapie dobrane zostaną parametry związane ze względną skalą pomiędzy obrazami, a w drugim etapie ustalone zostaną parametry związane ze względnym przesunięciem. Twierdzenie 1 przedstawia matematyczne uzasadnienie takiego podejścia.

Twierdzenie 1. *Macierz fundamentalną układu kamer można przedstawić jako iloczyn dwóch macierzy \mathbf{F}_S oraz \mathbf{F}_T :*

$$\mathbf{F} = \mathbf{F}_S \mathbf{F}_T \quad (4.13)$$

gdzie:

\mathbf{F}_S – macierz skali opisana wzorem (4.14),

\mathbf{F}_T – macierz pozostałych parametrów opisana wzorem (4.15),

oraz:

$$\mathbf{F}_S = \begin{bmatrix} 1/f_{irx} & 0 & 0 \\ 0 & 1/f_{iry} & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (4.14)$$

$$\mathbf{F}_T = \begin{bmatrix} 0 & 0 & t_y \\ 0 & 0 & -t_x \\ -t_y & t_x & c \end{bmatrix} \quad (4.15)$$

gdzie:

$$c = t_y * c_{tvx} - t_x * c_{tvy} + t_x * c_{iry}/f_{iry} - t_y * c_{irx}/f_{irx}.$$

Dowód. Ze wzoru (4.12):

$$\begin{aligned} \mathbf{F} &= \begin{bmatrix} 1/f_{irx} & 0 & -c_{irx}/f_{irx} \\ 0 & 1/f_{iry} & -c_{iry}/f_{iry} \\ 0 & 0 & 1 \end{bmatrix}^T \begin{bmatrix} 0 & 0 & t_y \\ 0 & 0 & -t_x \\ -t_y & t_x & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & -c_{tvx} \\ 0 & 1 & -c_{tvy} \\ 0 & 0 & 1 \end{bmatrix} = \\ &= \begin{bmatrix} 1/f_{irx} & 0 & 0 \\ 0 & 1/f_{iry} & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & -c_{irx}/f_{irx} \\ 0 & 1 & -c_{iry}/f_{iry} \\ 0 & 0 & 1 \end{bmatrix}^T \begin{bmatrix} 0 & 0 & t_y \\ 0 & 0 & -t_x \\ -t_y & t_x & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & -c_{tvx} \\ 0 & 1 & -c_{tvy} \\ 0 & 0 & 1 \end{bmatrix} = \\ &= \mathbf{F}_S \mathbf{F}_T \quad (4.16) \end{aligned}$$

Przedstawiając macierz \mathbf{F} zgodnie ze wzorem (4.16), równanie (4.5) możemy przedstawić w postaci (4.17). Wykonując przekształcenia prowadzące do równania (4.17) skorzystano z własności symetryczności macierzy \mathbf{F}_S , tzn. $\mathbf{F}_S = \mathbf{F}_S^T$:

$$\begin{aligned}
& \mathbf{x}_{ir}^T \mathbf{F} \mathbf{x}_{tv} = 0 \Leftrightarrow \\
& \Leftrightarrow \mathbf{x}_{ir}^T \mathbf{F}_S \mathbf{F}_T \mathbf{x}_{tv} = 0 \Leftrightarrow \\
& \Leftrightarrow \mathbf{x}_{ir}^T \mathbf{F}_S^T \mathbf{F}_T \mathbf{x}_{tv} = 0 \Leftrightarrow \\
& \Leftrightarrow (\mathbf{F}_S \mathbf{x}_{ir})^T \mathbf{F}_T \mathbf{x}_{tv} = 0 \Leftrightarrow \\
& \Leftrightarrow \mathbf{x}_{sir}^T \mathbf{F}_T \mathbf{x}_{tv} = 0
\end{aligned}$$

Ostatecznie:

$$\mathbf{x}_{sir}^T \mathbf{F}_T \mathbf{x}_{tv} = 0 \quad (4.17)$$

gdzie:

\mathbf{x}_{sir} - przeskalowany punkt obrazu termowizyjnego, opisany wzorem (4.18).

$$\mathbf{x}_{sir} = \begin{bmatrix} x_{sir} \\ y_{sir} \\ 1 \end{bmatrix} = \mathbf{F}_S \mathbf{x}_{ir} \quad (4.18)$$

gdzie:

x_{sir}, y_{sir} – współrzędne w dziedzinie obrazu punktu na obrazie z kamery termowizyjnej ze zmienioną skalą.

□

W równaniu (4.17) pokazano, że wobec przyjętych założeń więzy epipolarne odpowiadających sobie punktów obrazu wizyjnego (\mathbf{x}_{tv}) i termowizyjnego (\mathbf{x}_{ir}) można przedstawić jako zależność wiążącą odpowiadające sobie punkty przeskalowanego obrazu termowizyjnego (\mathbf{x}_{sir}) oraz obrazu wizyjnego (\mathbf{x}_{tv}). Skalowanie obrazu termowizyjnego opisane macierzą \mathbf{F}_S jest opisane współczynnikami $1/f_{irx}$ i $1/f_{iry}$.

Twierdzenie 1 pozwala na przeprowadzenie kalibracji w dwóch etapach:

- ustalenie skali obrazu termowizyjnego, związanej wyłącznie z parametrami f_{irx} i f_{iry} i opisanej równaniem $\mathbf{x}_{sir} = \mathbf{F}_S \mathbf{x}_{ir}$
- ustalenie parametrów zależności wiążącej odpowiadające sobie punkty przeskalowanego obrazu termowizyjnego (\mathbf{x}_{sir}) oraz obrazu wizyjnego (\mathbf{x}_{tv}) opisanej równaniem $\mathbf{x}_{sir}^T \mathbf{F}_T \mathbf{x}_{tv} = 0$.

W kolejnych podrozdziałach pokazano, w jaki sposób ten fakt został wykorzystany w ramach opracowanej metody samokalibracji układu kamer. Etap pierwszy samokalibracji opisany został w podrozdziale 4.1.3, natomiast etap drugi w podrozdziale 4.1.4.

4.1.3 Etap pierwszy kalibracji

Pierwszy etap kalibracji prowadzi do ustalenia parametrów skali obrazu termowizyjnego względem wizyjnego, czyli do wyznaczenia macierzy F_S opisanej wzorem (4.14). W opracowanej metodzie samokalibracji wykorzystano w tym celu dopasowanie obrazu z chwili bieżącej do obrazu z chwili poprzedniej dla sekwencji obrazów odpowiednio wizyjnych i termowizyjnych. Aby uzasadnić takie podejście, należy przyjąć dwa założenia. Po pierwsze ruch obiektów względem głowicy kamer przebiega głównie w płaszczyźnie prostopadłej do osi optycznej każdej z kamer. W rzeczywistej aplikacji takie założenie jest proste do spełnienia. Wystarczy zarejestrować scenę przedstawiającą obiekty nieruchome podczas ruchu obrotowego głowicy kamer wokół osi X lub Y. Takie założenie jest uzasadnione, ponieważ przy unieruchomionym zestawie Poprad, umieszczona na nim głowica ma możliwość tylko i wyłącznie ruchu obrotowego wokół osi X i Y. Po drugie ruchy obiektów względem głowicy są powolne. To oznacza, że zmiana kąтового położenia obiektu względem głowicy wynikająca z jej ruchu obrotowego jest pomijalnie mała.

Przyjęcie powyższych założeń pozwala określić, że dopasowanie obrazów wizyjnego i termowizyjnego będzie stałe w czasie, ponieważ w trakcie ruchu odległość obiektów od głowicy nie będzie się zmieniać. W nawiązaniu do wzoru (3.5) można to zapisać w postaci równania (4.19).

$$\mathbf{d} = \begin{bmatrix} d_x \\ d_y \end{bmatrix} = const. \quad (4.19)$$

gdzie:

\mathbf{d} - dopasowanie zgodne ze wzorem (3.5).

Dodatkowo nawiązując do wyprowadzenia (3.7) w odniesieniu do obrazów wizyjnego i termowizyjnego po korekcji skali otrzymujemy:

$$\begin{bmatrix} d_x \\ d_y \end{bmatrix} = \begin{bmatrix} x_{sir} - x_{tv} \\ y_{sir} - y_{tv} \end{bmatrix} = \begin{bmatrix} x_{sir} \\ y_{sir} \end{bmatrix} - \begin{bmatrix} x_{tv} \\ y_{tv} \end{bmatrix} \quad (4.20)$$

gdzie:

x_{sir}, y_{sir} – współrzędne punktu na obrazie termowizyjnym po skalowaniu,

x_{tv}, y_{tv} – współrzędne punktu na obrazie wizyjnym.

Wstawiając dopasowanie opisane wzorem (4.20) do równania (4.19) otrzymujemy:

$$\begin{aligned} \begin{bmatrix} d_x \\ d_y \end{bmatrix} &= const. \Leftrightarrow \\ \begin{bmatrix} x_{sir} - x_{tv} \\ y_{sir} - y_{tv} \end{bmatrix} &= const. \end{aligned} \quad (4.21)$$

Różniczkując równanie (4.21) stronami po czasie otrzymujemy:

$$\begin{aligned} \begin{bmatrix} \frac{d}{dt}d_x \\ \frac{d}{dt}d_y \end{bmatrix} &= 0 \Leftrightarrow \\ \Leftrightarrow \begin{bmatrix} \frac{d}{dt}x_{sir} \\ \frac{d}{dt}y_{sir} \end{bmatrix} &= \begin{bmatrix} \frac{d}{dt}x_{tv} \\ \frac{d}{dt}y_{tv} \end{bmatrix} \end{aligned} \quad (4.22)$$

W celu dalszego wyprowadzenia zależności konieczne jest najpierw rozwinięcie równania (4.18). W wyniku tego otrzymujemy:

$$\begin{bmatrix} x_{sir} \\ y_{sir} \\ 1 \end{bmatrix} = \mathbf{F}_S \begin{bmatrix} x_{ir} \\ y_{ir} \\ 1 \end{bmatrix} = \begin{bmatrix} 1/f_{irx} & 0 & 0 \\ 0 & 1/f_{iry} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_{ir} \\ y_{ir} \\ 1 \end{bmatrix} = \begin{bmatrix} x_{ir}/f_{irx} \\ y_{ir}/f_{iry} \\ 1 \end{bmatrix} \quad (4.23)$$

Wstawiając x_{ir} oraz y_{ir} z równania (4.23) do równania (4.22) można je rozwinąć w postaci równań (4.24) oraz (4.25) przy założeniach: $f_{irx}(t) = const$ oraz $f_{iry}(t) = const$.

$$\begin{aligned} \frac{d}{dt}x_{sir} &= \frac{d}{dt}x_{tv} \Leftrightarrow \\ \frac{d}{dt}(x_{ir}/f_{irx}) &= \frac{d}{dt}x_{tv} \Leftrightarrow \\ 1/f_{irx} \frac{d}{dt}(x_{ir}) &= \frac{d}{dt}(x_{tv}) \Leftrightarrow \\ \frac{d}{dt}(x_{ir}) &= f_{irx} \frac{d}{dt}(x_{tv}) \Leftrightarrow \\ x_{ir} &= f_{irx}x_{tv} \end{aligned} \quad (4.24)$$

gdzie:

f_{irx} - skala obrazu termowizyjnego względem wizyjnego w osi X.

Analogicznie dla ruchu w osi Y:

$$\begin{aligned}\frac{d}{dt}y_{sir} &= \frac{d}{dt}y_{tv} \Leftrightarrow \\ y_{ir} &= f_{iry}y_{tv}\end{aligned}\quad (4.25)$$

gdzie:

f_{iry} - skala obrazu termowizyjnego względem wizyjnego w osi Y.

Rozważając sekwencję par obrazów ułożoną chronologicznie (np. dwa zsynchronizowane filmy wideo) składniki y_{ir} , y_{tv} , x_{ir} oraz x_{tv} obecne w równaniach (4.24) oraz (4.25) można przedstawić w następujący sposób:

$$\begin{aligned}\begin{bmatrix} \dot{x}_{ir} \\ \dot{y}_{ir} \end{bmatrix} &= \begin{bmatrix} x_{ir} - x_{irprev} \\ y_{ir} - y_{irprev} \end{bmatrix} \\ \begin{bmatrix} \dot{x}_{tv} \\ \dot{y}_{tv} \end{bmatrix} &= \begin{bmatrix} x_{tv} - x_{tvprev} \\ y_{tv} - y_{tvprev} \end{bmatrix}\end{aligned}\quad (4.26)$$

gdzie:

x_{tvprev}, y_{tvprev} - współrzędne punktu obrazu wizyjnego z poprzedniej klatki z sekwencji,

x_{irprev}, y_{irprev} - współrzędne punktu obrazu termowizyjnego z poprzedniej klatki z sekwencji.

Nawiązując do wzoru (3.7), estymację składników y_{ir} , y_{tv} , x_{ir} oraz x_{tv} równań (4.26) dla każdej kamery niezależnie można wyznaczyć obliczając dopasowanie bieżącego obrazu g_{tv} , g_{ir} względem odpowiedniego obrazu poprzedniego g_{tvprev} , g_{irprev} z sekwencji wideo. Zostało to przedstawione w postaci równań (4.27).

$$\begin{aligned}\dot{x}_{ir} &= x_{ir} - x_{irprev} = d_{irx} \\ \dot{y}_{ir} &= y_{ir} - y_{irprev} = d_{iry} \\ \dot{x}_{tv} &= x_{tv} - x_{tvprev} = d_{tvx} \\ \dot{y}_{tv} &= y_{tv} - y_{tvprev} = d_{tvy}\end{aligned}\quad (4.27)$$

gdzie:

$(d_{tvx}, d_{tvy}) = \mathbf{d}_{tv}$ - dopasowanie obrazu wizyjnego z chwili bieżącej do poprzedniego z sekwencji wideo,

$(d_{irx}, d_{iry}) = \mathbf{d}_{ir}$ - dopasowanie obrazu termowizyjnego z chwili bieżącej do poprzedniego z

sekwencji wideo.

Etap pierwszy kalibracji ma na celu ustalenie skali obrazu termowizyjnego względem wizyjnego zgodnie z równaniami (4.24) oraz (4.25). Skalę definiują parametry f_{irx} oraz f_{iry} . Ich estymacja może nastąpić na podstawie zbioru pomiarów dopasowania obrazu bieżącego do obrazu z chwili poprzedniej dla obydwu obrazów z sekwencji obrazów wizyjnych i termowizyjnych. Aby było to możliwe, konieczne jest przyjęcie dodatkowo warunków $\dot{d}_{ir} \neq 0$ oraz $\dot{d}_{tv} \neq 0$. Pomiarzy mogą być wykonywane z wykorzystaniem metody dopasowania obrazów. Szczegółowy algorytm przedstawiający etap pierwszy kalibracji opisano w rozdziale 4.3 po omówieniu metod dopasowania obrazów w rozdziale 4.2.

4.1.4 Etap drugi kalibracji

Drugi etap kalibracji prowadzi do ustalenia zależności wiążącej dopasowanie odpowiadających sobie punktów obrazu wizyjnego i przeskalowanego obrazu termowizyjnego w osi X i w osi Y. Można ją sformułować w postaci następującego twierdzenia.

Twierdzenie 2. *Zależność wiążąca dopasowanie (d_x, d_y) odpowiadających sobie punktów obrazu wizyjnego i przeskalowanego obrazu termowizyjnego w osi X i w osi Y jest liniowa i można ją opisać równaniem (4.28).*

$$ad_x + bd_y + c = 0 \quad (4.28)$$

gdzie:

d_x, d_y – dopasowanie punktów odpowiednio w osi X oraz Y opisane wzorem (4.20),

a, b, c – stałe dla układu kamer parametry:

$$a = t_y,$$

$$b = -t_x,$$

$$c = t_y c_{tvx} - t_x c_{tvy} + t_x c_{iry} / f_{iry} - t_y c_{irx} / f_{irx}.$$

Oraz:

t_x, t_y – przesunięcie kamer względem siebie, wzór (2.5).

Dowód. Równanie (4.17) przedstawia relację wiążącą współrzędne odpowiadających sobie punktów x_{sir} oraz x_{tv} . Przekształcając to równanie można doprowadzić do postaci zgodnej z (4.28). Podstawiając do (4.17) wzory (4.18), (4.6) oraz (4.15) możemy wyprowadzić równanie (4.29).

$$\begin{aligned}
& \mathbf{x}_{sir}^T \mathbf{F}_T \mathbf{x}_{tv} = 0 \Leftrightarrow \\
& \begin{bmatrix} x_{sir} \\ y_{sir} \\ 1 \end{bmatrix}^T \begin{bmatrix} 0 & 0 & t_y \\ 0 & 0 & -t_x \\ -t_y & t_x & c \end{bmatrix}^T \begin{bmatrix} x_{tv} \\ y_{tv} \\ 1 \end{bmatrix} = 0 \Leftrightarrow \\
& c + t_y x_{sir} - t_y x_{tv} - t_x y_{sir} + t_x y_{tv} = 0 \Leftrightarrow \\
& t_y (x_{sir} - x_{tv}) - t_x (y_{sir} - y_{tv}) + c = 0
\end{aligned} \tag{4.29}$$

a stąd:

$$\begin{aligned}
& t_y d_x - t_x d_y + c = 0 \Leftrightarrow \\
& \begin{bmatrix} d_x & d_y & 1 \end{bmatrix} \begin{bmatrix} t_y \\ -t_x \\ c \end{bmatrix} = 0
\end{aligned} \tag{4.30}$$

gdzie:

$$c = t_y c_{tvx} - t_x c_{tvy} + t_x c_{iry} / f_{iry} - t_y c_{irx} / f_{irx}.$$

Równanie (4.30) dowodzi tezy twierdzenia, ponieważ jest równoważne z równaniem (4.28). \square

Linie prostą opisującą tę zależność będziemy w rozprawie nazywać *linią dopasowania*. Etap drugi kalibracji ma na celu ustalenie parametrów linii dopasowania obrazów wizyjnych i termowizyjnych zgodne z równaniem (4.28). Zostało wykazane, że relację definiują parametry t_x, t_y, c . Ich estymacja może nastąpić na podstawie zbioru pomiarów dopasowania dla sekwencji obrazów wizyjnych i termowizyjnych po dostosowaniu skali obrazu termowizyjnego do obrazu wizyjnego. Jest to możliwe do wykonania po zakończeniu etapu pierwszego kalibracji - po estymacji parametrów f_x oraz f_y . Pomiarzy mogą być wykonywane z wykorzystaniem dowolnej metody dopasowania obrazów. Szczegółowy algorytm przedstawiający etap drugi kalibracji opisano w rozdziale 4.3 po omówieniu metod dopasowania obrazów w rozdziale 4.2.

4.2 Dopasowanie automatyczne obrazów metodą korelacji fazowej

W bieżącym podrozdziale przedstawiono opracowaną metodę automatycznego dopasowania obrazów z wykorzystaniem metody korelacji fazowej dla wielu obiektów widocznych na obrazach wizyjnym i termowizyjnym. Opracowaną metodę wykorzystano przy samokalibracji układu kamer zgodnie z przyjętymi w rozprawie założeniami. Zaproponowana w rozprawie metoda samokalibracji służy wyznaczeniu odpowiednich parametrów opisu układu kamer na podstawie analizy sekwencji obrazów wizyjnych i termowizyjnych. Samokalibracja wyznaczana jest na podstawie automatycznie obliczanych estymacji ruchu obrazów odpowiednio wizyjnego i termowizyjnego (etap I) oraz na podstawie automatycznie wyznaczanego dopasowania w trybie kalibracji obrazów termowizyjnego i wizyjnego (etap II). Na etapie I estymacja ruchu obrazów obliczana jest jako dopasowanie monomodalne w trybie kalibracji obrazu z chwili bieżącej do obrazu z chwili poprzedniej odpowiednio dla kamery wizyjnej i termowizyjnej. Opisywana w bieżącym podrozdziale metoda dopasowania została więc wykorzystana zarówno na etapie I jak i II samokalibracji.

Opracowana metoda bazuje na obliczeniach w dziedzinie częstotliwości na podstawie transformacji Fouriera obrazów. W podrozdziałach 4.2.1-4.2.4 znajduje się analiza metody korelacji fazowej pod kątem wykorzystania w zaprojektowanej metodzie samokalibracji. Pokazano tam, że metodę tę można zastosować do znalezienia dopasowania wielu odpowiadających sobie obiektów jednocześnie występujących na obrazach dopasowywanych. Dodatkowo zaprezentowano zastosowanie metody w przypadku wyznaczania dopasowania multimodalnego obrazów wizyjnego i termowizyjnego. W porównaniu do zastosowań metody korelacji fazowej opisywanych w literaturze, gdzie w zakresie dopasowania obrazów multimodalnych wyznaczana jest tylko jedna wartość dopasowania, wyznaczenie wielu wartości dopasowania dla jednej pary obrazów pozwoliło na poprawę jakości działania metody samokalibracji. To uzasadnia tezę rozprawy.

W przypadku omawianym w rozprawie obrazy pochodzące z układu kamer są monochromatyczne, więc barwę ich składowych pikseli można rozumieć jako intensywność. Problemem przy stosowaniu metody korelacji fazowej bezpośrednio jest fakt, że intensywność odpowiadających sobie obiektów obrazu wizyjnego oraz termowizyjnego nie jest związana bezpośrednią zależnością. W podrozdziałach 4.2.1-4.2.4 przeanalizowano tę sytuację i opisano podstawy rozwiązania tego problemu - przed przystąpieniem do korelacji fazowej konieczne jest zastosowanie odpowiedniej metody przetwarzania wstępnego w celu zniwelowania różnic w poziomach intensywności odpowiadających sobie obiektów.

Na podstawie tej analizy w podrozdziale 4.2.5 zaproponowano ogólny algorytm dopasowania w trybie kalibracji. W rozdziale 4.2.6 pokazano, że przetwarzanie wstępne

obrazów można zrealizować poprzez zastosowanie odpowiedniej filtracji krawędziowej, na podstawie progowania wartości modułu zmian intensywności w dziedzinie obrazu.

4.2.1 Dopasowanie obrazów monomodalnych

Opracowana w rozprawie metoda samokalibracji estymuje odpowiednie parametry modelu układu kamer na podstawie: (etap I) automatycznego dopasowania monomodalnego w trybie kalibracji obrazu z chwili bieżącej do obrazu z chwili poprzedniej odpowiednio dla kamery wizyjnej i termowizyjnej oraz (etap II) automatycznego dopasowania multimodalnego w trybie kalibracji obrazów termowizyjnego i wizyjnego. Biorąc pod uwagę założenia przedstawione w rozdziale 1.2 oraz na podstawie przeglądu metod dopasowania wykonanego w rozdziale 3 proponuje się wykorzystanie korelacji fazowej w celu wyznaczenia dopasowania w trybie kalibracji na obu tych etapach. W bieżącym podrozdziale pokazano problemy i możliwe sposoby ich rozwiązania związane ze stosowaniem tej metody do dopasowania obrazów multimodalnych w trybie kalibracji na etapie II. W przypadku dopasowania monomodalnego na etapie I nie występują przeszkody teoretyczne związane ze stosowaniem korelacji fazowej.

W rozdziale 4.1.4 wykazano, że zależność opisującą dopasowanie online odpowiadających sobie punktów omawianego w rozprawie układu dwóch kamer: wizyjnej i termowizyjnej można przedstawić w postaci linii prostej. Można ją zapisać w postaci równania (4.30). Ustalenie parametrów tej linii pozwoli uprościć dopasowanie online obrazów wizyjnego i termowizyjnego. Skoro dopasowanie online odpowiadających sobie obiektów w osi X i Y obrazu będzie zmieniało się wzdłuż linii prostej, można je jednoznacznie określić przy pomocy jednego parametru określającego położenie punktu na prostej. Punkt odpowiadający temu parametrowi będzie posiadał współrzędne $\begin{bmatrix} d_x & d_y \end{bmatrix}^T$, czyli będzie reprezentował dopasowanie online odpowiadających sobie obiektów. W celu ustalenia parametrów omawianej linii konieczne jest wyznaczenie dopasowania w trybie kalibracji obiektów w różnych odległościach od głowicy kamer multimodalnych. Można tego dokonać stosując w odpowiedni sposób metodę korelacji fazowej. W bieżącym podrozdziale zamieszczono analizę i teoretyczne uzasadnienie stosowania tej metody. To uzasadnia tezę rozprawy w zakresie wykorzystania metody korelacji fazowej dla wielu obiektów jednocześnie widocznych na obrazach wizyjnym i termowizyjnym.

Ogólna zasada działania metody korelacji fazowej przedstawiona zostanie na przykładzie dopasowania dwóch obrazów monomodalnych przedstawiających jeden obiekt. Niech funkcja $g(x, y)$ reprezentuje intensywności obrazu monochromatycznego o rozmiarach $M \times N$ pikseli. Dyskretną transformatę Fouriera $\mathcal{F}(g(x, y))$ tego obrazu można zapisać w następujący sposób [164]:

$$G(u, v) = \mathcal{F}(g(x, y)) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} g(x, y) e^{-j2\pi(\frac{ux}{M} + \frac{vy}{N})} = |G(u, v)| e^{-j\phi(u, v)} \quad (4.31)$$

gdzie:

x, y - współrzędne pikseli obrazu,

u, v - współrzędne współczynników transformaty Fouriera,

$|G(u, v)|$ - amplituda sygnału w dziedzinie częstotliwości - charakteryzuje częstotliwość każdego z poziomów intensywności obrazu,

$\phi(u, v)$ - faza sygnału w dziedzinie częstotliwości - charakteryzuje przestrzenną strukturę obrazu.

Założmy, że dysponujemy dwoma obrazami g_1 i g_2 , przesuniętymi względem siebie o (x_0, y_0) . Można to zapisać w postaci równania:

$$g_2(x, y) = g_1(x - x_0, y - y_0) \quad (4.32)$$

gdzie:

(x_0, y_0) - nieznan wektor przesunięcia obiektu pomiędzy obrazami.

W celu wyznaczenia przesunięcia (x_0, y_0) możliwe jest skorzystanie z metody bazującej na wyznaczeniu transformaty Fouriera. W przypadku opisanym wzorem (4.32), zgodnie z własnością transformacji Fouriera opisującą przesunięcie w dziedzinie obrazu [139] możemy napisać:

$$G_2(u, v) = G_1(u, v)e^{-j2\pi(\frac{ux_0}{M} + \frac{vy_0}{N})} = G_1(u, v)e^{j\Delta(u, v)} \quad (4.33)$$

gdzie:

$$G_1(u, v) = \mathcal{F}(g_1(x, y))$$

$$G_2(u, v) = \mathcal{F}(g_2(x, y))$$

$$\Delta(u, v) = -2\pi(\frac{ux_0}{M} + \frac{vy_0}{N})$$

Odwołując się do wzorów (4.31) oraz (4.33) możemy dodatkowo określić następujące zależności:

$$\begin{aligned} G_1(u, v) &= |G_1(u, v)|e^{-j\phi_1(u, v)} \\ G_2(u, v) &= G_1(u, v)e^{j\Delta(u, v)} = |G_1(u, v)|e^{-j\phi_1(u, v)}e^{j\Delta(u, v)} = |G_1(u, v)|e^{(-j\phi_1(u, v) + \Delta(u, v))} \end{aligned} \quad (4.34)$$

W celu uwypuklenia składnika Δ zawierającego informację o przesunięciu (x_0, y_0) można wyznaczyć znormalizowaną widmową gęstość mocy obrazów g_1 oraz g_2 . Jest ona zdefiniowana w następujący sposób:

$$P_w(u, v) = \frac{G_1(u, v)G_2^*(u, v)}{|G_1(u, v)G_2^*(u, v)|} \quad (4.35)$$

gdzie:

$G_2^*(u, v)$ - transformata sprzężona względem $G_2(u, v)$.

Korzystając z (4.34) możemy wzór (4.35) sprowadzić do postaci:

$$\begin{aligned}
 P_w(u, v) &= \frac{|G_1(u, v)|e^{-j\phi_1(u, v)}|G_1(u, v)|e^{j(\phi_1(u, v) - \Delta(u, v))}}{||G_1(u, v)|e^{-j\phi_1(u, v)}|G_1(u, v)|e^{j(\phi_1(u, v) - \Delta(u, v))}|} = \\
 &= \frac{|G_1(u, v)|^2 e^{-j\phi_1(u, v) + j(\phi_1(u, v) - \Delta(u, v))}}{||G_1(u, v)|^2 e^{-j\phi_1(u, v) + j(\phi_1(u, v) - \Delta(u, v))}|} = \\
 &= \frac{|G_1(u, v)|^2 e^{-j\Delta(u, v)}}{||G_1(u, v)|^2 e^{-j\Delta(u, v)}|} = \\
 &= \frac{|G_1(u, v)|^2 e^{-j\Delta(u, v)}}{|G_1(u, v)|^2} = \\
 &= e^{-j\Delta(u, v)}
 \end{aligned} \tag{4.36}$$

Następnie możemy obliczyć odwrotną transformatę Fouriera względem P_w :

$$p_w(x, y) = \mathcal{F}^{-1}(P_w(u, v)) = \mathcal{F}^{-1}(e^{-j\Delta(u, v)}) = \delta(x - x_0, y - y_0) \tag{4.37}$$

gdzie:

$\delta(x - x_0, y - y_0)$ - impuls Diraca o współrzędnych (x_0, y_0) .

Zatem funkcja $p_w(x, y)$ może być traktowana jako funkcja podobieństwa obrazów $s(\mathbf{g}_1, \mathbf{g}_2)$, którą zdefiniowano w rozdziale 3.1 w postaci wzoru (3.9). Wyznaczenie wartości dopasowania pomiędzy obrazami $\mathbf{g}_1, \mathbf{g}_2$ można określić w nawiązaniu do wzoru (3.13) w sposób następujący:

$$\begin{bmatrix} \mathbf{D} & 1 \end{bmatrix} = \begin{bmatrix} d_x & d_y & 1 \end{bmatrix} = \arg \max_{x, y} p_w \tag{4.38}$$

gdzie:

$\mathbf{D} = \begin{bmatrix} d_x & d_y \end{bmatrix}$ - lokalizacja maksimum macierzy p_w .

Oznacza to, że przesunięcie między obrazami \mathbf{g}_1 oraz \mathbf{g}_2 można wyznaczyć znajdując lokalizację maksimum macierzy p_w przedstawionej we wzorze (4.37):

$$(\widehat{x}_0, \widehat{y}_0) = (d_x, d_y) \tag{4.39}$$

W praktyce realizacja metody korelacji fazowej sprowadza się do wyznaczenia macierzy wartości $p_w(x, y)$ oraz analizy tych wartości pod kątem lokalizacji maksimum. Wyznaczenie $p_w(x, y)$ proponuje się w postaci Algorytmu 4.1, który zaimplementowano dla potrzeb badań metody. W algorytmie zastosowano okno Hanna [165], w celu zapobiegania wpływom krawędzi obrazu na działanie korelacji fazowej.

Algorytm 4.1 Metoda POC - wyznacz macierz korelacji

```
1: procedure WYZNACZMACIERZPOC( $\mathbf{g}_1, \mathbf{g}_2$ )
2:   niech  $(w, h)$  oznacza rozmiar obrazu  $\mathbf{g}_1$  i  $\mathbf{g}_2$ 
3:   do obrazów  $\mathbf{g}_1$  i  $\mathbf{g}_2$  zastosuj okno Hanna
4:   inicjuj  $\mathbf{P}_w$  jako zespoloną macierz zerową o rozmiarze  $(w, h)$ 
5:    $\mathbf{G}_1 \leftarrow \mathcal{F}(\mathbf{g}_1)$ 
6:    $\mathbf{G}_2 \leftarrow \mathcal{F}(\mathbf{g}_2)$ 
7:   for all piksel obrazu  $g_1$  do
8:      $a \leftarrow \text{Re}(G_1(\text{piksel}))$ 
9:      $b \leftarrow \text{Im}(G_1(\text{piksel}))$ 
10:     $c \leftarrow \text{Re}(G_2(\text{piksel}))$ 
11:     $d \leftarrow \text{Im}(G_2(\text{piksel}))$ 
12:     $a_1 \leftarrow \text{atan2}(b, a)$ 
13:     $a_2 \leftarrow \text{atan2}(d, c)$ 
14:     $\text{Re}(P_w(\text{piksel})) \leftarrow \cos(a_1 - a_2)$ 
15:     $\text{Im}(P_w(\text{piksel})) \leftarrow \sin(a_1 - a_2)$ 
16:    $\mathbf{p}_w \leftarrow \mathcal{F}^{-1}(\mathbf{P}_w)$ 
17:    $\mathbf{p}_w \leftarrow \text{fftshift}(\mathbf{p}_w)$   $\triangleright$  przenieś zerowe częstotliwości do środka macierzy R
18:   return  $\mathbf{p}_w$ 
```

4.2.2 Dopasowanie dwóch obiektów na obrazach monomodalnych

Praktyczne stosowanie metody korelacji fazowej można uogólnić rozszerzając założenie (4.32). Załóżmy, że dysponujemy obrazami \mathbf{g}_1 i \mathbf{g}_2 , z których każdy składa się z dwóch części (widocznych obiektów): a i b . Załóżmy dalej, że obraz \mathbf{g}_2 przedstawia część a obrazu \mathbf{g}_1 przesuniętą o nieznaną wektor (x_a, y_a) oraz część b obrazu \mathbf{g}_1 przesuniętą o nieznaną wektor (x_b, y_b) . Sytuację taką można zapisać w postaci równań:

$$\begin{aligned}g_1(x, y) &= g_{1a}(x, y) + g_{1b}(x, y) \\g_2(x, y) &= g_{2a}(x, y) + g_{2b}(x, y) = g_{1a}(x - x_a, y - y_a) + g_{1b}(x - x_b, y - y_b) \\g_{2a}(x, y) &= g_{1a}(x - x_a, y - y_a) \\g_{2b}(x, y) &= g_{1b}(x - x_b, y - y_b)\end{aligned}\tag{4.40}$$

gdzie:

(x_a, y_a) - nieznaną wektor przesunięcia obiektu a ,

(x_b, y_b) - nieznaną wektor przesunięcia obiektu b .

Odwołując się do wzoru (4.31) oraz powołując się na liniowość transformacji Fouriera możemy dodatkowo określić następujące zależności:

$$\begin{aligned}
G_{1a}(u, v) &= |G_{1a}(u, v)|e^{-j\phi_{1a}(u, v)} \\
G_{1b}(u, v) &= |G_{1b}(u, v)|e^{-j\phi_{1b}(u, v)} \\
G_1(u, v) &= G_{1a}(u, v) + G_{1b}(u, v) = |G_{1a}(u, v)|e^{-j\phi_{1a}(u, v)} + |G_{1b}(u, v)|e^{-j\phi_{1b}(u, v)} \\
G_2(u, v) &= G_{2a}(u, v) + G_{2b}(u, v) = G_{1a}(u, v)e^{j\Delta_a(u, v)} + G_{1a}(u, v)e^{j\Delta_a(u, v)} = \\
&= |G_{1a}(u, v)|e^{-j(\phi_{1a}(u, v) - \Delta_a(u, v))} + |G_{1b}(u, v)|e^{-j(\phi_{1b}(u, v) - \Delta_b(u, v))} \quad (4.41)
\end{aligned}$$

gdzie:

$$\begin{aligned}
G_{1a}(u, v) &= \mathcal{F}(g_{1a}(x, y)) \\
G_{1b}(u, v) &= \mathcal{F}(g_{1b}(x, y)) \\
G_{2a}(u, v) &= \mathcal{F}(g_{2a}(x, y)) \\
G_{2b}(u, v) &= \mathcal{F}(g_{2b}(x, y)) \\
\Delta_a(u, v) &= -2\pi\left(\frac{ux_a}{M} + \frac{vy_a}{N}\right) \\
\Delta_b(u, v) &= -2\pi\left(\frac{ux_b}{M} + \frac{vy_b}{N}\right)
\end{aligned}$$

W celu uwypuklenia składników Δ_a oraz Δ_b zawierających informację o przesunięciu (x_a, y_a) oraz (x_b, y_b) można wyznaczyć znormalizowaną widmową gęstość mocy obrazów g_1 oraz g_2 w sposób analogiczny do (4.35).

Korzystając z (4.41) możemy wzór (4.35) sprowadzić do postaci:

$$\begin{aligned}
P_w(u, v) &= \frac{G_1(u, v)G_2^*(u, v)}{|G_1(u, v)G_2^*(u, v)|} = \\
&= \frac{|G_{1a}(u, v)|e^{-j\phi_{1a}(u, v)} + |G_{1b}(u, v)|e^{-j\phi_{1b}(u, v)}}{|G_1(u, v)|} \\
&= \frac{|G_{1a}(u, v)|e^{-j(\phi_{1a}(u, v) - \Delta_a(u, v))} + |G_{1b}(u, v)|e^{-j(\phi_{1b}(u, v) - \Delta_b(u, v))}}{|G_2(u, v)|} = \\
&= \frac{|G_{1a}(u, v)|^2 e^{-j\Delta_a(u, v)} + |G_{1b}(u, v)|^2 e^{-j\Delta_b(u, v)}}{|G_1(u, v)||G_2(u, v)|} + \\
&= \frac{|G_{1a}(u, v)||G_{1b}(u, v)|(e^{-j(\phi_{1a} - \phi_{1b} + \Delta_b(u, v))} + e^{-j(\phi_{1b} - \phi_{1a} + \Delta_a(u, v))})}{|G_1(u, v)||G_2(u, v)|} \quad (4.42)
\end{aligned}$$

Korzystając z własności liniowości transformacji Fouriera można wyznaczyć p_w , czyli odwrotną transformatę Fouriera z P_w zgodnego ze wzorem (4.42):

$$\begin{aligned}
p_w(x, y) = \mathcal{F}^{-1}(P_w(u, v)) &= \mathcal{F}^{-1}\left(\frac{|G_{1a}(u, v)|^2 e^{-j\Delta\phi_a(u, v)}}{|G_1(u, v)||G_2(u, v)|}\right) + \\
&\quad \mathcal{F}^{-1}\left(\frac{|G_{1b}(u, v)|^2 e^{-j\Delta\phi_b(u, v)}}{|G_1(u, v)||G_2(u, v)|}\right) + \\
\mathcal{F}^{-1}\left(\frac{|G_{1a}(u, v)||G_{1b}(u, v)|(e^{-j(\phi_{1a}-\phi_{1b}+\Delta_b(u, v))} + e^{-j(\phi_{1b}-\phi_{1a}+\Delta_a(u, v))})}{|G_1(u, v)||G_2(u, v)|}\right) &= \\
&\quad \mathcal{F}^{-1}\left(e^{-j\Delta\phi_a(u, v)} \frac{|G_{1a}(u, v)|^2}{|G_1(u, v)||G_2(u, v)|}\right) + \\
&\quad \mathcal{F}^{-1}\left(e^{-j\Delta\phi_b(u, v)} \frac{|G_{1b}(u, v)|^2}{|G_1(u, v)||G_2(u, v)|}\right) + \\
\mathcal{F}^{-1}\left(\frac{|G_{1a}(u, v)||G_{1b}(u, v)|(e^{-j(\phi_{1a}-\phi_{1b}+\Delta_b(u, v))} + e^{-j(\phi_{1b}-\phi_{1a}+\Delta_a(u, v))})}{|G_1(u, v)||G_2(u, v)|}\right) &= \quad (4.43)
\end{aligned}$$

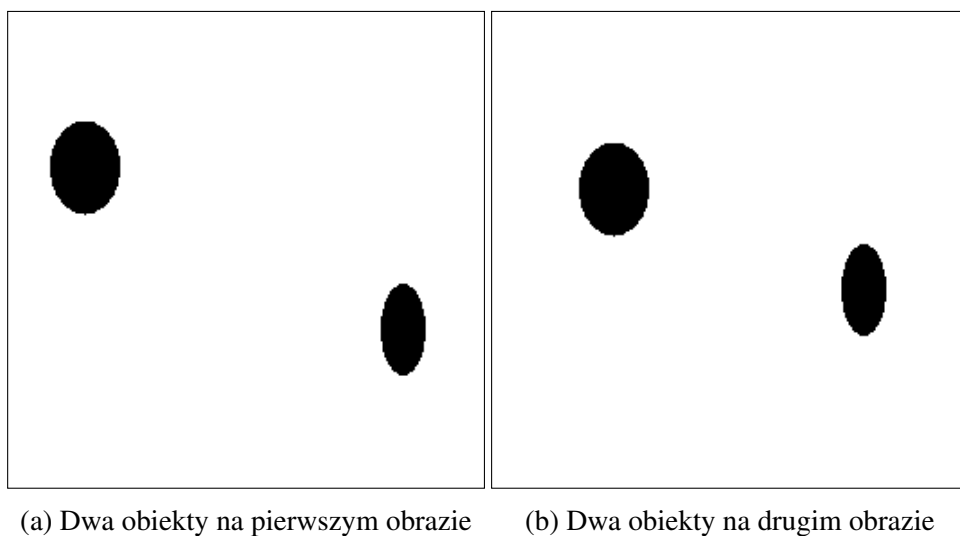
Jeśli weźmiemy pod uwagę twierdzenie o transformacji Fouriera splotu funkcji, otrzymamy:

$$\begin{aligned}
p_w(x, y) = \mathcal{F}^{-1}(P_w(u, v)) &= \mathcal{F}^{-1}(e^{-j\Delta\phi_a(u, v)}) * \mathcal{F}^{-1}\left(\frac{|G_{1a}(u, v)|^2}{|G_1(u, v)||G_2(u, v)|}\right) + \\
&\quad \mathcal{F}^{-1}(e^{-j\Delta\phi_b(u, v)}) * \mathcal{F}^{-1}\left(\frac{|G_{1b}(u, v)|^2}{|G_1(u, v)||G_2(u, v)|}\right) + \\
\mathcal{F}^{-1}\left(\frac{|G_{1a}(u, v)||G_{1b}(u, v)|(e^{-j(\phi_{1a}-\phi_{1b}+\Delta_b(u, v))} + e^{-j(\phi_{1b}-\phi_{1a}+\Delta_a(u, v))})}{|G_1(u, v)||G_2(u, v)|}\right) &= \\
&\quad \delta(x - x_a, y - y_a) * \mathcal{F}^{-1}\left(\frac{|G_{1a}(u, v)|^2}{|G_1(u, v)||G_2(u, v)|}\right) + \\
&\quad \delta(x - x_b, y - y_b) * \mathcal{F}^{-1}\left(\frac{|G_{1b}(u, v)|^2}{|G_1(u, v)||G_2(u, v)|}\right) + \\
\mathcal{F}^{-1}\left(\frac{|G_{1a}(u, v)||G_{1b}(u, v)|(e^{-j(\phi_{1a}-\phi_{1b}+\Delta_b(u, v))} + e^{-j(\phi_{1b}-\phi_{1a}+\Delta_a(u, v))})}{|G_1(u, v)||G_2(u, v)|}\right) &= \quad (4.44)
\end{aligned}$$

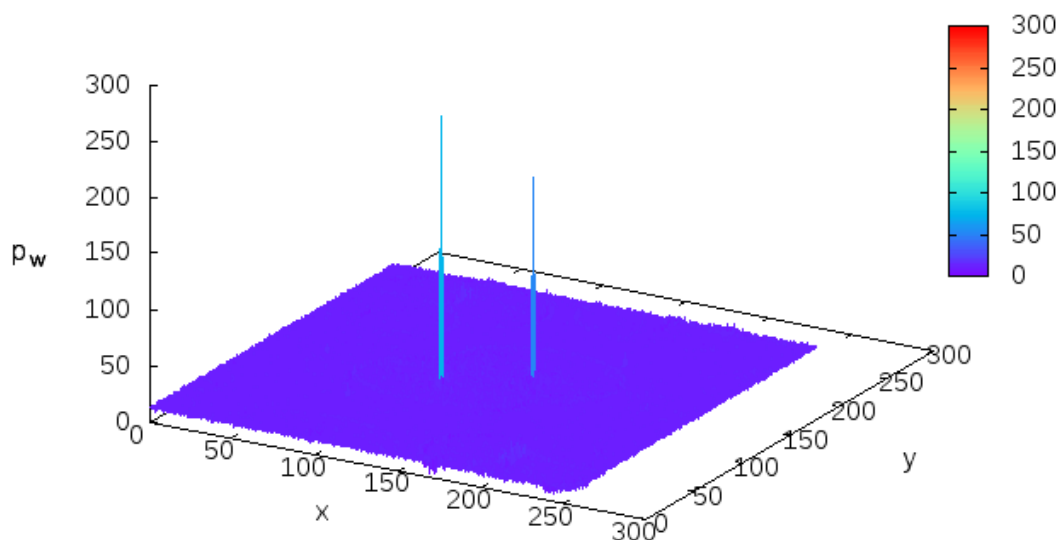
Podsumowując równanie (4.44) można stwierdzić, że macierz \mathbf{p}_w posiada lokalne maksima o lokalizacji (x_a, y_a) oraz (x_b, y_b) . Macierz zawiera również trzeci składnik opisujący korelację i przesunięcie obiektu b na obrazie \mathbf{g}_2 względem obiektu a na obrazie \mathbf{g}_1 . Wyniki działania tej metody można przedstawić na prostym przykładzie.

Przykład 1. Przykładowe obrazy wejściowe \mathbf{g}_1 oraz \mathbf{g}_2 przedstawione zostały na rysunkach 4.1a oraz 4.1b. Odwrotna transformata Fouriera z \mathbf{P}_w znajduje się natomiast na rysunku 4.2.

Obrazy 4.1b, oraz 4.1a zostały wygenerowane komputerowo, a wektory przesunięcia kształtów widocznych na rysunku 4.1b względem rysunku 4.1a wynoszą $(-13, -22)$ oraz $(25, 12)$. Rysunek 4.2 przedstawia macierz korelacji fazowej \mathbf{p}_w wyznaczoną z zastosowaniem algorytmu 4.1. Przesunięcia pomiędzy obiektami widocznymi na obrazach można wyznaczyć



Rysunek 4.1: Przykład dwóch obiektów - obrazy wygenerowane komputerowo



Rysunek 4.2: Przykład działania korelacji fazowej dla dwóch obiektów z rys 4.1b - macierz korelacji fazowej p_w

jako lokalizację dwóch największych maksimów macierzy p_w względem jej środka otrzymując $(-13.03, -21.93)$ oraz $(24.98, 12.00)$.

Podsumowując, metodę korelacji fazowej można stosunkowo łatwo zastosować do dopasowania dwóch monomodalnych obiektów. W kolejnym fragmencie podrzdziału metoda zostanie rozszerzona o uzasadnienie stosowania tej metody dla więcej niż dwóch obiektów monomodalnych.

4.2.3 Dopasowanie wielu obiektów na obrazach monomodalnych

Założmy teraz, że dysponujemy znanymi obrazami g_1 i g_2 , z których każdy składa się z n policzalnych części (widocznych obiektów) o numerach: $i = 1, 2, \dots, n$. Założmy, że obraz g_2 przedstawia część i obrazu g_1 przesuniętą o nieznaną wektor (x_i, y_i) . Sytuację taką można zapisać w postaci:

$$\begin{aligned} g_1(x, y) &= \sum_{i=1}^n g_{1i}(x, y) \\ g_2(x, y) &= \sum_{i=1}^n g_{2i}(x, y) = \sum_{i=1}^n g_{1i}(x - x_i, y - y_i) \\ g_{2i}(x, y) &= g_{1i}(x - x_i, y - y_i) \end{aligned} \quad (4.45)$$

Odwołując się do wzoru (4.31) oraz powołując się na liniowość transformacji Fouriera możemy dodatkowo określić następujące zależności:

$$\begin{aligned} G_{1i}(u, v) &= |G_{1i}(u, v)| e^{-j\phi_{1i}(u, v)} \\ G_{2i}(u, v) &= |G_{2i}(u, v)| e^{-j\phi_{2i}(u, v)} \\ G_1(u, v) &= \sum_{i=1}^n G_{1i}(u, v) = \sum_{i=1}^n |G_{1i}(u, v)| e^{-j\phi_{1i}(u, v)} \\ G_2(u, v) &= \sum_{i=1}^n G_{2i}(u, v) = \sum_{i=1}^n |G_{2i}(u, v)| e^{-j\phi_{2i}(u, v)} = \\ &= \sum_{i=1}^n |G_{1i}(u, v)| e^{-j(\phi_{1i}(u, v) - \Delta_i(u, v))} \end{aligned} \quad (4.46)$$

gdzie:

$$G_{1i}(u, v) = \mathcal{F}(g_{1i}(x, y))$$

$$G_{2i}(u, v) = \mathcal{F}(g_{2i}(x, y))$$

$$\Delta_i(u, v) = -2\pi\left(\frac{ux_i}{M} + \frac{vy_i}{N}\right)$$

W celu uwypuklenia składników Δ_i zawierających informację o przesunięciu (x_i, y_i) można wyznaczyć znormalizowaną widmową gęstość mocy obrazów g_1 oraz g_2 w sposób analogiczny do (4.35):

$$P_w(u, v) = \frac{G_1(u, v)G_2^*(u, v)}{|G_1(u, v)G_2^*(u, v)|} = \frac{(\sum_{i=1}^n G_{1i}(u, v))(\sum_{i=1}^n G_{2i}^*(u, v))}{|G_1(u, v)G_2^*(u, v)|} \quad (4.47)$$

Korzystając z (4.46) możemy rozwinąć składnik $(\sum_{i=1}^n G_{1i}(u, v))(\sum_{i=1}^n G_{2i}^*(u, v))$ ze wzoru (4.47) w następujący sposób:

$$\begin{aligned}
& \left(\sum_{i=1}^n G_{1i}(u, v) \right) \left(\sum_{i=1}^n G_{2i}^*(u, v) \right) = \\
& = \sum_{i=1}^n G_{1i}(u, v) G_{2i}^*(u, v) + \sum_{i=1, k=1, i \neq k}^n G_{1i}(u, v) G_{2k}^*(u, v) = \\
& = \sum_{i=1}^n |G_{1i}(u, v)| e^{-j\phi_{1i}(u, v)} |G_{1i}(u, v)| e^{j(\phi_{1i}(u, v) - \Delta_i(u, v))} + \sum_{i=1, k=1, i \neq k}^n G_{1i}(u, v) G_{2k}^*(u, v) = \\
& = \sum_{i=1}^n |G_{1i}(u, v)|^2 e^{-j\Delta_i(u, v)} + \sum_{i=1, k=1, i \neq k}^n G_{1i}(u, v) G_{2k}^*(u, v)
\end{aligned} \tag{4.48}$$

Podstawiając wyniki wyprowadzenia (4.48) do (4.47) otrzymujemy:

$$P_w(u, v) = \frac{\sum_{i=1}^n |G_{1i}(u, v)|^2 e^{-j\Delta_i(u, v)} + \sum_{i=1, k=1, i \neq k}^n G_{1i}(u, v) G_{2k}^*(u, v)}{|G_1(u, v) G_2^*(u, v)|} \tag{4.49}$$

Przeprowadzając rozumowanie analogiczne jak w równaniach (4.43) i (4.44) możemy wyznaczyć transformatę odwrotną z \mathbf{P}_w :

$$\begin{aligned}
p_w(x, y) &= \mathcal{F}^{-1}(P_w(u, v)) = \\
&= \sum_{i=1}^n \delta(x - x_i, y - y_i) * \mathcal{F}^{-1}\left(\frac{|G_{1i}(u, v)|^2}{|G_1(u, v)| |G_2(u, v)|}\right) + \sum_{i=1, k=1, i \neq k}^n \mathcal{F}^{-1}\left(\frac{G_{1i}(u, v) G_{2k}^*(u, v)}{|G_i(u, v)| |G_2(u, v)|}\right)
\end{aligned} \tag{4.50}$$

W równaniu (4.50) składnik $\sum_{i=1}^n \delta(x - x_i, y - y_i) * \mathcal{F}^{-1}\left(\frac{|G_{1i}(u, v)|^2}{|G_1(u, v)| |G_2(u, v)|}\right)$ odpowiada lokalnym maksimum o lokalizacji zgodnej z szukanymi wektorami (x_i, y_i) przemieszczenia obiektów $i = 1, 2, \dots, n$.

Składnik $\sum_{i=1, k=1, i \neq k}^n \mathcal{F}^{-1}\left(\frac{G_{1i}(u, v) G_{2k}^*(u, v)}{|G_i(u, v)| |G_2(u, v)|}\right)$ opisuje zależności wiążące przesunięcie obiektów i względem k dla sytuacji $i \neq k$. Praktyczne znaczenie tej części równania można interpretować posługując się definicją korelacji wzajemnej (4.51), oraz związanym z nią twierdzeniem dotyczącym transformacji Fouriera (4.52).

$$(g_2 * g_1) = (g_2(x, y) * g_1(x, y))(n, m) = \sum_n \sum_m (g_2^*(x, y) g_1(x + n, y + m)) \tag{4.51}$$

$$\mathcal{F}^{-1}(G_1 G_2^*) = (g_2 * g_1) \tag{4.52}$$

W związku z tym, składnik $\sum_{i=1, k=1, i \neq k}^n \mathcal{F}^{-1}\left(\frac{G_{1i}(u,v)G_{2k}^*(u,v)}{|G_1(u,v)||G_2(u,v)|}\right)$ odpowiada znormalizowanej korelacji wzajemnej obiektów i, k dla sytuacji $i \neq k$.

Dopasowanie wielu obiektów może być wyznaczone jako znalezienie lokalizacji n największych maksimumów. Niech ta operacja będzie zdefiniowana w postaci wzoru (4.53).

$$\mathbf{D} = \begin{bmatrix} d_{x1} & d_{y1} \\ d_{x2} & d_{y2} \\ \dots & \\ d_{xn} & d_{yn} \end{bmatrix} = n \arg \max_{x,y} p_w(x, y) \quad (4.53)$$

gdzie:

$n \arg \max$ - ekstrakcja lokalizacji n największych maksimumów,

\mathbf{D} - zbiór lokalizacji n największych maksimumów,

d_{xj}, d_{yj} - współrzędne lokalizacji maksimum o numerze $j = 1, 2, \dots, n$.

Dodatkowo można założyć, że wyznaczone w ten sposób lokalizacje maksimumów można traktować jako estymację wzajemnego przemieszczenia obiektów (x_i, y_i) :

$$(\hat{x}_i, \hat{y}_i) = (d_{xj}, d_{yj}) \quad (4.54)$$

gdzie:

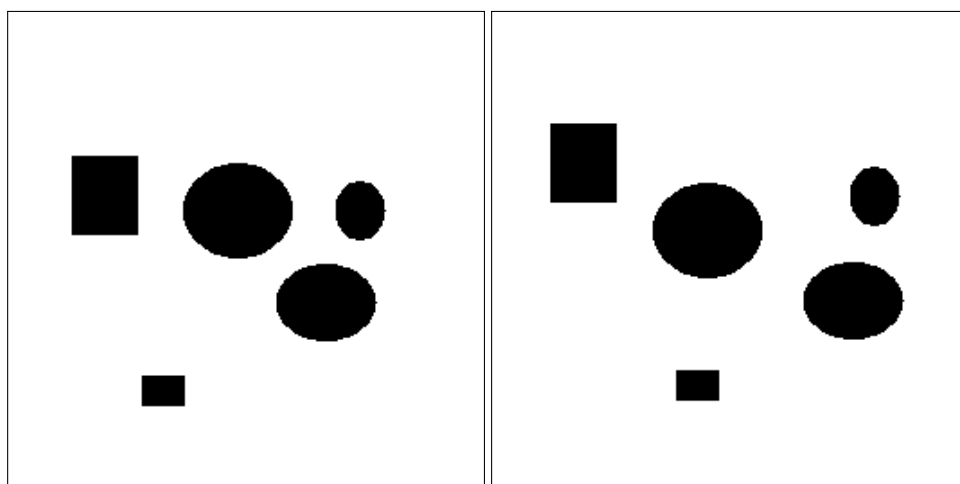
(\hat{x}_i, \hat{y}_i) - estymacja przesunięcia (x_i, y_i) pomiędzy obiektami i, i - numer obiektu widocznego na obrazach,

j - numer maksimum w kolejności od największego do najmniejszego, numery i oraz j nie muszą sobie odpowiadać wartościami.

Korzystając ze wzoru (4.53) metodę można zademonstrować na prostym przykładzie.

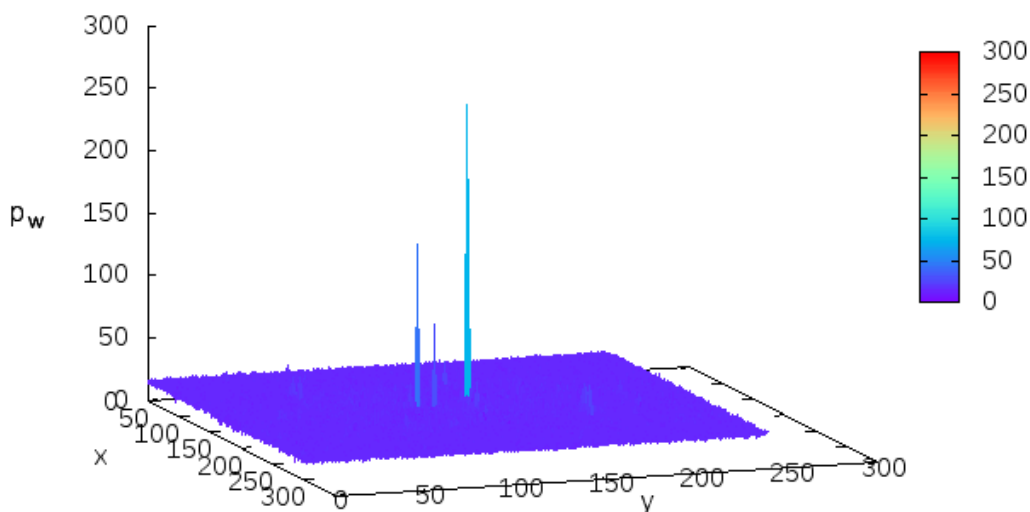
Przykład 2. Przykładowe obrazy wejściowe g_1 oraz g_2 przedstawione zostały na rysunkach 4.3a oraz 4.3b. Odwrotna transformata Fouriera z \mathbf{P}_w znajduje się natomiast na rysunku 4.4.

Obrazy 4.3b, oraz 4.3a zostały wygenerowane komputerowo, a wektory przesunięcia kształtów widocznych na rysunku 4.3b względem rysunku 4.3a wynoszą $(-3, -18), (-8, 11), (28, -3), (17, -8), (24, -1)$. Rysunek 4.4 przedstawia macierz korelacji fazowej \mathbf{p}_w wyznaczoną z wykorzystaniem algorytmu 4.1. Przesunięcia pomiędzy obiektami widocznymi na obrazach można wyznaczyć jako lokalizację największych maksimumów macierzy \mathbf{p}_w względem jej środka. Z pięciu ustalonych w ten sposób lokalizacji jedynie trzy reprezentują przesunięcia pomiędzy odpowiadającymi sobie obiektami, są to: $(-7.99, 11.12), (24.03, -1.30), (16.07, -7.92)$, które pasują do przesunięć $(-8, 11), (24, -1), (17, -8)$. Resztę przesunięć trudno zidentyfikować dysponując tylko macierzą \mathbf{p}_w .



(a) Pięć obiektów na pierwszym obrazie (b) Pięć obiektów na drugim obrazie

Rysunek 4.3: Przykład dla pięciu obiektów - obrazy wygenerowane komputerowo



Rysunek 4.4: Przykład działania korelacji fazowej dla pięciu obiektów z rys. 4.3b - macierz korelacji fazowej p_w

Podsumowując, metodę korelacji fazowej można zastosować do dopasowania więcej niż dwóch monomodalnych obiektów, jednakże może to być utrudnione w sytuacji występowania wielu obiektów. W niektórych sytuacjach może nie być możliwe wyznaczenie przesunięcia pomiędzy wszystkimi obiektami widocznymi na dopasowywanych obrazach. Ponadto nie istnieje oczywisty sposób przyporządkowania wyznaczonych dopasowań do obiektów, którym one odpowiadają. W kolejnym podrozdziale metoda zostanie rozszerzona o uzasadnienie jej stosowania dla wielu obiektów na obrazach multimodalnych. Obrazami multimodalnymi mogą być w szczególności obrazy wizyjny i termowizyjny.

4.2.4 Dopasowanie wielu obiektów na obrazach multimodalnych

Metoda korelacji fazowej jest odporna na niektóre rodzaje zakłóceń obrazu [166] np.: jednolite zmiany intensywności, zmiany średniego poziomu jasności itp. Można to uzasadnić porównując własność transformacji Fouriera (4.55) ze wzorem (4.35). Na tej podstawie można wywnioskować, że korelacja fazowa jest metodą opierającą swoje działanie na znormalizowanym w dziedzinie częstotliwości współczynniku korelacji wzajemnej.

$$KorelacjaWzajemna(g_1, g_2) = \mathcal{F}^{-1}[\mathcal{F}^*(g_1)\mathcal{F}(g_2)] \quad (4.55)$$

Odporność metody na jednorodną zmianę poziomu intensywności można wykazać na podstawie następującego wyprowadzenia. Załóżmy, że dysponujemy znanymi obrazami termowizyjnym g_{ir} i wizyjnym g_{tv} , z których każdy składa się z n policzalnych części (widocznych obiektów) o numerach: $i = 1, 2, \dots, n$. Załóżmy też, że obraz g_{ir} przedstawia część i obrazu g_{tv} przesuniętą o nieznaną wektor (x_i, y_i) . Dodatkowo załóżmy, że intensywność każdego obiektu z obrazu termowizyjnego jest związana prostą zależnością liniową (współczynniki $a_i > 0$) z intensywnością odpowiadającego obiektu na obrazie wizyjnym. Sytuację taką można zapisać w postaci równań:

$$\begin{aligned} g_{tv}(x, y) &= \sum_{i=1}^n g_{tvi}(x, y) \\ g_{ir}(x, y) &= \sum_{i=1}^n g_{iri}(x, y) = \sum_{i=1}^n a_i g_{tvi}(x - x_i, y - y_i) \\ g_{iri}(x, y) &= a_i g_{tvi}(x - x_i, y - y_i) \end{aligned} \quad (4.56)$$

Odwołując się dalej do wzorów (4.31), (4.33) oraz powołując się na liniowość transformacji Fouriera możemy dodatkowo określić następujące zależności:

$$\begin{aligned} G_{tvi}(u, v) &= |G_{tvi}(u, v)|e^{-j\phi_{tvi}(u, v)} \\ G_{iri}(u, v) &= |G_{iri}(u, v)|e^{-j\phi_{iri}(u, v)} \\ G_{tv}(u, v) &= \sum_{i=1}^n G_{tvi}(u, v) = \sum_{i=1}^n |G_{tvi}(u, v)|e^{-j\phi_{tvi}(u, v)} \\ G_{ir}(u, v) &= \sum_{i=1}^n G_{iri}(u, v) = \sum_{i=1}^n |G_{iri}(u, v)|e^{-j\phi_{iri}(u, v)} = \\ &= \sum_{i=1}^n a_i |G_{tvi}(u, v)|e^{-j(\phi_{tvi}(u, v) - \Delta_i(u, v))} \end{aligned} \quad (4.57)$$

gdzie:

$$G_{tvi}(u, v) = \mathcal{F}(g_{tvi}(x, y))$$

$$G_{iri}(u, v) = \mathcal{F}(g_{iri}(x, y))$$

$$\Delta_i(u, v) = -2\pi\left(\frac{ux_i}{M} + \frac{vy_i}{N}\right)$$

W celu uwypuklenia składników Δ_i zawierających informację o przesunięciu (x_i, y_i) można wyznaczyć znormalizowaną widmową gęstość mocy obrazów g_{tv} oraz g_{ir} w sposób analogiczny do (4.35):

$$P_w(u, v) = \frac{G_{tv}(u, v)G_{ir}^*(u, v)}{|G_{tv}(u, v)G_{ir}^*(u, v)|} = \frac{(\sum_{i=1}^n G_{tvi}(u, v))(\sum_{i=1}^n G_{iri}^*(u, v))}{|(\sum_{i=1}^n G_{tvi}(u, v))(\sum_{i=1}^n G_{iri}^*(u, v))|} \quad (4.58)$$

Korzystając dalej z (4.57) możemy rozwinąć składnik $(\sum_{i=1}^n G_{tvi}(u, v))(\sum_{i=1}^n G_{iri}^*(u, v))$ ze wzoru (4.58) w następujący sposób:

$$\begin{aligned} & \left(\sum_{i=1}^n G_{tvi}(u, v)\right)\left(\sum_{i=1}^n G_{iri}^*(u, v)\right) = \\ & = \left(\sum_{i=1}^n |G_{tvi}(u, v)|e^{-j\phi_{tvi}(u, v)}\right)\left(\sum_{i=1}^n a_i |G_{tvi}(u, v)|e^{j(\phi_{tvi}(u, v) - \Delta_i(u, v))}\right) \end{aligned} \quad (4.59)$$

Współczynniki a_i reprezentują zmianę intensywności obiektu na obrazie termowizyjnym względem obrazu wizyjnego. Jeśli założymy, że zmiana intensywności jest jednorodna na całym obrazie termowizyjnym, tzn. $a = a_i, i = 1, 2, \dots, n$, możemy współczynnik a wyrugować z równania (4.58):

$$\begin{aligned} P_w(u, v) & = \frac{(\sum_{i=1}^n G_{tvi}(u, v))(\sum_{i=1}^n G_{iri}^*(u, v))}{|(\sum_{i=1}^n G_{tvi}(u, v))(\sum_{i=1}^n G_{iri}^*(u, v))|} = \\ & = \frac{a(\sum_{i=1}^n |G_{tvi}(u, v)|e^{-j\phi_{tvi}(u, v)})(\sum_{i=1}^n |G_{tvi}(u, v)|e^{j(\phi_{tvi}(u, v) - \Delta_i(u, v))})}{|a(\sum_{i=1}^n |G_{tvi}(u, v)|e^{-j\phi_{tvi}(u, v)})(\sum_{i=1}^n |G_{tvi}(u, v)|e^{j(\phi_{tvi}(u, v) - \Delta_i(u, v))})|} = \\ & = \frac{(\sum_{i=1}^n |G_{tvi}(u, v)|e^{-j\phi_{tvi}(u, v)})(\sum_{i=1}^n |G_{tvi}(u, v)|e^{j(\phi_{tvi}(u, v) - \Delta_i(u, v))})}{|(\sum_{i=1}^n |G_{tvi}(u, v)|e^{-j\phi_{tvi}(u, v)})(\sum_{i=1}^n |G_{tvi}(u, v)|e^{j(\phi_{tvi}(u, v) - \Delta_i(u, v))})|} \end{aligned} \quad (4.60)$$

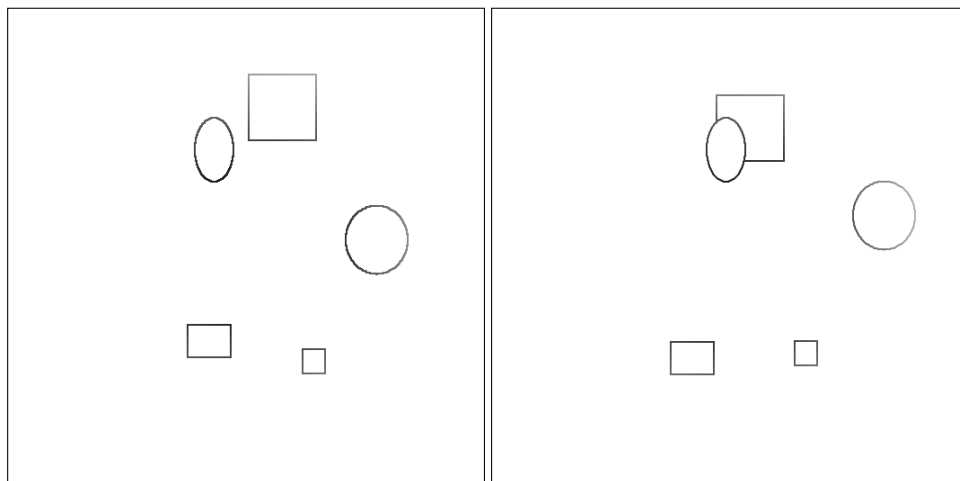
W ten sposób można sprowadzić zależność (4.60) do przypadku dopasowania wielu obiektów obserwowanych na obrazach monomodalnych (4.47). Metodę korelacji fazowej można więc zastosować do dopasowania wielu obiektów na obrazach multimodalnych pod warunkiem rozwiązania problemu z niejednorodnymi w dziedzinie obrazu zmianami intensywności pomiędzy porównywanymi obrazami.

W tym celu można przykładowo wykorzystać wybrane cechy obrazów: gradienty lub ekstrakcję krawędzi [68, 98]. Takie podejście wiąże się z założeniem, że odpowiadające sobie obiekty na obrazach wizyjnych i termowizyjnych są do siebie podobne strukturalnie, ale niekoniecznie odpowiadają sobie poziomami intensywności. Oznacza to, że zastosowanie

ekstrakcji krawędzi powinno uzasadnić założenie (4.45) w przypadku obrazów wizyjnych i termowizyjnych. Krawędzie powinny jednak spełniać określone założenia, aby pozbyć się problemu z różnymi, niejednorodnymi poziomami intensywności nowych obiektów (krawędzi obiektów sprzed filtracji). W tym celu można wykorzystać wartość bezwzględną gradientu intensywności i zastosować progowanie wartości bezwzględnej w celu dalszego ujednoczenia poziomów poziomów intensywności nowych obiektów krawędziowych. Po zastosowaniu tych zabiegów korelację fazową można wykorzystać do dopasowania wielu obiektów multimodalnych.

Dokładny opis opracowanej metody filtracji krawędziowej przedstawiono w rozdziale 4.2.6 jednakże już tutaj można przedstawić jej przykładowe działanie.

Przykład 4. Przykładowe obrazy wejściowe g_1 oraz g_2 przedstawione zostały na rysunkach 4.5a oraz 4.5b. Odwrotna transformata Fouriera z P_w znajduje się natomiast na rysunku 4.6.

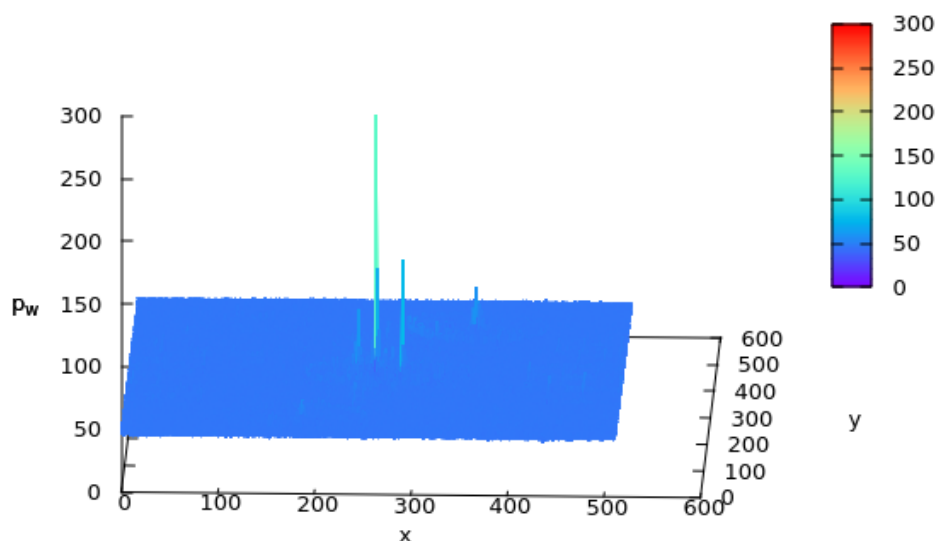


(a) Pięć obiektów w postaci krawędzi na pierwszym obrazie (b) Pięć obiektów w postaci krawędzi na drugim obrazie

Rysunek 4.5: Przykład działania filtracji krawędziowej dla pięciu obiektów - obrazy wygenerowane komputerowo

Wektory przesunięcia kształtów widocznych na rysunku 4.5b względem rysunku 4.5a wynoszą $(-18, 23)$, $(26, -27)$, $(-1, 19)$, $(9, -9)$, $(31, 0)$. Rysunek 4.6 przedstawia macierz korelacji fazowej p_w wyznaczoną z wykorzystaniem algorytm 4.1. Przesunięcia pomiędzy obiektami widocznymi na obrazach można wyznaczyć jako lokalizację pięciu największych maksimum macierzy p_w względem jej środka. Ustalone w ten sposób lokalizacje reprezentują przesunięcia pomiędzy odpowiadającymi sobie obiektami, są to: $(30.94, 0.06)$, $(25.88, -26.89)$, $(-0.98, 18.72)$, $(8.98, -9.01)$, $(-18.24, 22.89)$, które pasują do przesunięć $(31, 0)$, $(26, -27)$, $(-1, 19)$, $(9, -9)$, $(-18, 23)$.

Podsumowując, zarówno analiza teoretyczna, jak i przedstawiony przykład świadczą o tym, że możliwe jest ustalenie dopasowania wielu obiektów z wykorzystaniem metody korelacji fazowej dla obrazów o niejednorodnych w dziedzinie obrazu różnicach w poziomach intensywności. Takie obrazy można uzyskać z układu kamer: wizyjnej i termowizyjnej, gdzie



Rysunek 4.6: Przykład działania korelacji fazowej dla pięciu obiektów po filtracji krawędziowej z rys. 4.5b - macierz korelacji fazowej p_w

intensywność odpowiadających sobie obiektów nie jest związana bezpośrednią zależnością. Przed przystąpieniem do korelacji fazowej konieczne jest jednak zastosowanie odpowiedniej metody przetwarzania wstępnego w celu zniwelowania różnic w poziomach intensywności odpowiadających sobie obiektów.

4.2.5 Ogólny algorytm dopasowania

Zgodnie z analizą przedstawioną w poprzednim podrozdziale ustalono, że dopasowanie wielu obiektów znajdujących się w różnej odległości od układu kamer wizyjnej i termowizyjnej może być wyznaczone na podstawie metody korelacji fazowej. W tym celu konieczne jest przeprowadzenie wstępnego przetwarzania obrazu w celu uwypuklenia krawędzi, a przez to ujednolicenia poziomów intensywności odpowiadających sobie obiektów. W związku z tym proponuje się ogólny schemat automatycznego wyznaczenia dopasowania obrazów w postaci Algorytmu 4.2.

Algorytm 4.2 Wyznaczenie dopasowania sekwencji obrazów - algorytm ogólny

- 1: **procedure** DOPASUJSEKWENCJE
- 2: **while true do**
- 3: pobierz obrazy g_{ir}, g_{tv} z głowicy kamer
- 4: $(g_{iredge}, g_{tvedge}) \leftarrow \text{przetworzWstepnieObrazy}(g_{ir}, g_{tv})$ ▷ Patrz rozdział 4.2.6.
- 5: $p_w \leftarrow \text{wyznaczMacierzPoc}(g_{iredge}, g_{tvedge})$ ▷ Algorytm 4.1
- 6: znajdź lokalizacje D największych maksimum w p_w

S

Zaproponowany algorytm bazuje na metodzie POC poprzedzonej filtracją krawędziową w celu upodobnienia do siebie dwóch obrazów o różnej modalności. Podejście takie pozwala na dopasowanie obrazów na podstawie zawartej w nich informacji strukturalnej.

Praktyczna realizacja znalezienia lokalizacji największych maksimum \mathbf{D} w p_w w przypadku jednego obiektu może zostać zrealizowana za pomocą funkcji $\arg \max$, a w przypadku wielu obiektów za pomocą funkcji $n \arg \max$ (zależność (4.53)). W praktyce nie wiadomo ile przesunięć występuje na obrazach dopasowywanych dlatego w rozprawie proponuje się modyfikację funkcji $n \arg \max$ o dodatkowe założenia:

- ilość znalezionych maksimum będzie określona automatycznie, za pomocą algorytmu,
- do znalezionych lokalizacji maksimum \mathbf{D} zaliczamy tylko maksima, których stosunek wartości do wartości największego znalezionej maksimum jest większy od pewnej wartości granicznej¹.

Otrzymujemy wtedy funkcję $k \arg \max$:

$$\left[\mathbf{D} \quad \mathbf{1} \right] = \begin{bmatrix} d_{x1} & d_{y1} & 1 \\ d_{x2} & d_{y2} & 1 \\ \dots & & \\ d_{xk} & d_{yk} & 1 \end{bmatrix} = k \arg \max_{x,y} p_w \quad (4.61)$$

gdzie:

(d_{xj}, d_{yj}) - lokalizacja wybranego maksimum $j = 1, 2, \dots, k$,

oraz:

$$p_w(x, y) = \mathcal{F}^{-1}(P_w(u, v)) = \mathcal{F}^{-1}\left(\frac{G_{tv}(u, v)G_{ir}^*(u, v)}{|G_{tv}(u, v)G_{ir}^*(u, v)|}\right)$$

W następnym podrozdziale opisano dokładniej operację *przetworzWstepnieObrazy* z Algorytmu 4.2.

4.2.6 Filtracja krawędziowa obrazów

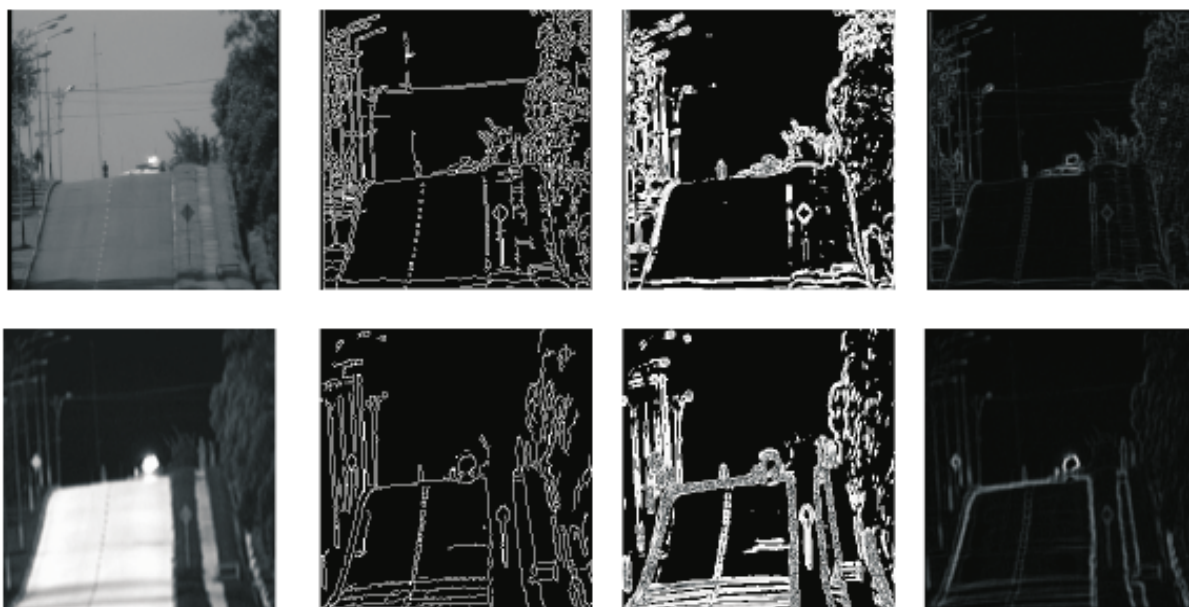
Wstępne przetwarzanie obrazów w celu dopasowania obrazów wizyjnych i termowizyjnych za pomocą korelacji fazowej w praktyce oznacza wyróżnienie informacji strukturalnych zawartych na tych obrazach. W celu wybrania metody najlepiej odpowiadającej założeniom przyjętym w rozprawie wykonane zostało porównanie wybranych metod ekstrakcji krawędzi. Szczegółowe informacje na ten temat autor rozprawy przedstawił w publikacji [68]. Poniżej znajduje się podsumowanie wykonanych prac i osiągniętych wyników.

¹w rozprawie przyjęto doświadczalnie wartość 0.4

W ramach tych prac przeprowadzone zostały testy implementacji działania różnych metod ekstrakcji krawędzi. Sprawdzano jakość wyznaczenia pojedynczego dopasowania obrazów wizyjnych i termowizyjnych, wstępnie przetworzonych z wykorzystaniem różnych metod filtracji krawędziowej. Testy wykonano dla zestawu osiemnastu różnych sekwencji obrazów i dla pięciu wybranych metod detekcji krawędzi:

- detektor Canny w implementacji z OpenCV (oznaczony jako CannyCV1),
- uproszczony detektor Canny (oznaczony jako CannyCV2),
- detektor Harris'a w implementacji z OpenCV (oznaczony jako HarrisCV),
- detektor morfologiczny (Morf1),
- detektor morfologiczny zmodyfikowany (Morf2).

Przykładowe obrazy przedstawiające krawędzie dla par z wybranej sekwencji i dla różnych metod ekstrakcji przedstawione zostały na rysunku 4.7.



Rysunek 4.7: Przykładowe wyniki filtracji krawędziowej. Górny rząd - obrazy TV, dolny rząd - obrazy IR. Od lewej odpowiednio: obrazy źródłowe, obrazy po filtracji CannyCV2, obrazy po filtracji HarrisCV oraz obrazy po filtracji morfologicznej oznaczone jako Morf1.

Detektor krawędzi typu Canny

W trakcie badań przetestowano dwie implementacje detektora Canny'ego opisanego w [167]. Pierwsza obejmuje implementację zawartą w bibliotece programistycznej OpenCV, a druga upraszcza wykorzystany tam kod źródłowy. Implementacja z biblioteki OpenCV oznaczona jako CannyCV1 ma standardowy przebieg:

- filtracja odsumiająca obrazu źródłowego $g(x, y)$ z wykorzystaniem filtru Gaussa,
- wyznaczenie obrazów gradientów poziomych $g_{ex} = \frac{\partial g_x}{\partial g}$ oraz pionowych $g_{ey} = \frac{\partial g_y}{\partial g}$ z wykorzystaniem filtracji typu Sobel,

- detekcja jednego z czterech kierunków gradientu: 0, 45, 90 lub 135 stopni wg wzoru $\alpha = \arctan\left(\frac{g_{ex}}{g_{ey}}\right)$,
- wyznaczenie maksimum gradientu $g_e = \sqrt{g_{ex}^2 + g_{ey}^2}$ dla czterech możliwych kierunków kąta α ,
- oznaczenie pikseli, dla których maksimum gradientu jest większe od pewnej wartości progowej, jako pikseli należących do krawędzi - wartość progowa posiada histerezę stosowaną w celu uzyskania krawędzi o grubości jednego piksela,
- iteracyjne przedłużanie i łączenie krawędzi w celu uzyskania ich ciągłości, z dodatkowym wykluczeniem pojedynczych pikseli jako krawędzi.

Uproszczony przebieg algorytmu (CannyCV2) pomija ostatni z wymienionych wyżej etapów.

Detektor krawędzi Harrisa

Kolejna metoda uwypuklenia informacji strukturalnej dla potrzeb dopasowania obrazów multimodalnych obejmuje detektor Harrisa [168, 169] w implementacji z biblioteki OpenCV. Zasada działania algorytmu polega na wyznaczeniu macierzy autokorelacji \mathbf{M} . Wyjściowy obraz krawędzi wyznaczany jest na podstawie wielkości (4.62).

$$\mathbf{Hr} = \det(\mathbf{M}) - k * \text{trace}(\mathbf{M})^2 = \lambda_1 \lambda_2 - k(\lambda_1 + \lambda_2)^2 \quad (4.62)$$

gdzie:

k - tzw. współczynnik Harrisa, λ_1, λ_2 - wartości własne macierzy \mathbf{M} .

Krawędź występuje dla $\mathbf{Hr} < 0$, czyli gdy $\lambda_1 \gg \lambda_2$ lub odwrotnie.

Morfologiczny detektor krawędzi

Morfologiczny detektor krawędzi zaprojektowany i opisany w pracach [170, 68] polega na wykorzystaniu operacji erozji i dylatacji. W tym celu dla obrazu $g(x, y)$ zdefiniowano operacje min i max oraz element strukturalny ES . Element ES stanowi zazwyczaj kwadrat 3×3 (zastosowany w rozprawie) lub 5×5 itp. pikseli o stałej intensywności oraz początku układu współrzędnych zdefiniowanym w pikselu centralnym elementu. Dylatacja g_D dla obrazu monochromatycznego zdefiniowana jest jako transformacja przypisująca wartość maksimum z intensywności $g(m, n)$ pikseli objętych elementem ES dla każdego piksela (x, y) obrazu w sytuacji, gdy początek układu elementu ES umieszczony jest w środku piksela o współrzędnych (x, y) (wzory (4.63) i (4.64)). Erozja g_E zdefiniowana jest analogicznie.

$$g_D(x, y) = \max_{m, n \in ES} g(m, n) \circ ES \quad (4.63)$$

$$g_E(x, y) = \min_{m, n \in ES} g(m, n) \circ ES \quad (4.64)$$

Operator “ \circ ” jest zdefiniowany dla dwóch wartości: intensywności pikseli obrazu oraz wartości elementu strukturalnego ES . W celu wykorzystania tak zdefiniowanych operacji morfologicznych dla potrzeb ekstrakcji krawędzi skorzystano z definicji morfologicznego gradientu opisanego w pracy [171], przedstawionej wzorem (4.65).

$$g_{e1}(x, y) = g_D(x, y) - g_E(x, y) \quad (4.65)$$

Ekstrakcja krawędzi sprowadza się w tym przypadku praktycznie do Algorytmu 4.3. Można zauważyć, że w algorytmie uwzględniono wyznaczenie wartości bezwzględnej z gradientu krawędzi w wierszu 7 oraz progowanie tej wartości na podstawie stałych parametrów t_1, t_2 w wierszach od 9 do 14.

Algorytm 4.3 Przetwarzanie wstępne obrazów z wykorzystaniem filtracji morfologicznej

```

1: procedure PRZETWORZWSSTEPNIEOBRAZY( $\mathbf{g}_{ir}, \mathbf{g}_{tv}$ )
2:   utwórz obrazy krawędziowe  $\mathbf{g}_{iredge}, \mathbf{g}_{tvedge}$  o zerowych wartościach krawędzi
3:   for all obraz  $\mathbf{g}_{ir}, \mathbf{g}_{tv}$  do
4:     for all piksel  $p$  obrazu  $\mathbf{g}$  do
5:        $sum \leftarrow 0$ 
6:       for all piksel  $p_o$  otoczenia piksela  $p$  do
7:         wyznacz bezwzględny gradient intensywności  $p_g \leftarrow abs(p_o - p)$ 
8:          $sum \leftarrow sum + p_g$ 
9:       if  $sum < t_1$  then
10:         $g_{edge}(p) \leftarrow 0$ 
11:       else if  $sum > t_2$  then
12:         $g_{edge} \leftarrow max$ 
13:       else
14:         $g_{edge}(p) \leftarrow (sum * const)^2$ 
15:   return ( $\mathbf{g}_{iredge}, \mathbf{g}_{tvedge}$ )

```

Dodatkowo, w celu zmniejszenia wpływu szumów na końcowy efekt algorytmu zaproponowano modyfikację metody z wykorzystaniem operacji wyrównania histogramu przed ekstrakcją krawędzi. Algorytm podstawowy oznaczono jako Morf1, a zmodyfikowany jako Morf2.

Omówienie wyników

Uzyskane w wyniku badań wyniki działania różnych metod ekstrakcji krawędzi podsumowano w tabeli 4.1.

W przypadku algorytmów ekstrakcji krawędzi metoda wskazywana najczęściej jako najlepsza to detektor Canny w implementacji z OpenCV. Zastosowanie takiej filtracji w świetle przyjętych założeń stwarza jednak problemy z uwagi na jej wysoką złożoność obliczeniową. W rozprawie przyjmuje się stosowanie metody morfologicznej Morf1. Cechuje się ona niską złożonością obliczeniową, a w kolejnych rozdziałach pokazano, że wyniki samokalibracji

Tabela 4.1: Porównanie wyników otrzymanych przy dopasowaniu obrazów algorytmem opisanym w rozdziale 4.2.5 z najlepszymi metodami ekstrakcji krawędzi oraz bez ekstrakcji krawędzi. err_{max} oznacza największy błąd dla testowanej sekwencji. Oznaczenia metod: 0 - POC, 1 - CannyCV1+POC, 2- HarrisCV+POC, 3 - Morf1+POC, 4 - CannyCV2+POC, 5 - Morf2+POC.

| sekwencja | najlepsza metoda | druga metoda | err_{max} najlepszej metody | err_{max} drugiej metody | err_{max} POC bez krawędzi |
|------------------------------------|------------------|--------------|-------------------------------|----------------------------|------------------------------|
| 1 | 2 | 1 | 2.02 | 2.11 | 39.11 |
| 2 | 4 | 2 | 0.72 | 1.74 | 1.40 |
| 3 | 1 | 5 | 1.15 | 1.15 | 16.13 |
| 4 | 3 | 1 | 0.30 | 0.98 | 1.26 |
| 5 | 0 | 1 | 0.60 | 1.55 | 0.60 |
| 6 | 1 | 2 | 0.74 | 0.74 | 52.49 |
| 7 | 2 | 4 | 1.06 | 3.40 | 3.64 |
| 8 | 5 | 1 | 1.26 | 1.38 | 1.53 |
| 9 | 5 | 4 | 2.32 | 32.70 | 8.55 |
| 10 | 1 | 4 | 0.57 | 0.57 | 1.47 |
| 11 | 1 | 5 | 1.87 | 1.87 | 2.66 |
| 12 | 5 | 1 | 2.53 | 3.06 | 3.06 |
| 13 | 3 | 0 | 1.79 | 1.83 | 1.79 |
| 14 | 1 | 4 | 1.74 | 1.74 | 3.23 |
| 15 | 3 | 0 | 0.30 | 0.32 | 0.30 |
| 16 | 1 | 0 | 4.34 | 4.34 | 5.00 |
| 17 | 1 | 3 | 4.94 | 4.94 | 12.60 |
| 18 | 0 | 1 | 3.15 | 3.36 | 3.15 |
| najczęstszy zwycięzca | 1 | 1 | | | |
| najczęstszy drugi zwycięzca | 3 i 5 | 4 | | | |

uzyskane z jej udziałem są zadowalające. Niezależnie od tego, czy jest to obraz wizyjny, czy termowizyjny, wyznaczenie krawędzi można przeprowadzić na podstawie zmian intensywności w dziedzinie obrazu. W bieżącym podrozdziale pokazano, że zniwelowanie różnic w poziomach intensywności odpowiadających sobie obiektów (reprezentowanych w postaci krawędzi) można zrealizować przez wyznaczanie krawędzi jako progowanie wartości modułu zmian intensywności w dziedzinie obrazu. Zarówno metoda Canny'ego jak i morfologiczna zawierają te operacje. Zastosowanie wartości bezwzględnej powoduje zaniedbanie znaku krawędzi (nieważne czy krawędź jest granicą między obszarem jasnym i ciemnym, czy ciemnym i jasnym), a operacja progowania niweluje dalsze różnice w poziomach intensywności krawędzi odpowiadających sobie obiektów (nieważne, czy intensywność krawędzi ma wartość wysoką, czy niską).

4.2.7 Wnioski i podsumowanie

Opracowana metoda dopasowania stanowi kontynuację prac autora rozpoczętych w ramach [68]. W literaturze występują nieliczne artykuły opisujące stosowanie metody korelacji fazowej do dopasowania obrazów multimodalnych, a w szczególności obrazów wizyjnego i termowizyjnego. Ze względu na założenia związane z multimodalnością proponowany w rozprawie sposób wyznaczenia dopasowania jest najbardziej podobny do [150] - bazuje na metodzie RIPOC poprzedzonej filtracją krawędziową dla obrazów wizyjnego i termowizyjnego. Jednakże dopasowywane we wspomnianej pracy obrazy są mało skomplikowane w porównaniu z tymi testowanymi w rozprawie. Dodatkowo zademonstrowane w artykule [150] przykłady przedstawiają pojedynczy obiekt (twarz) na tle o jednolitej barwie, zaś metoda zaproponowana w niniejszym rozdziale umożliwia wyznaczenie wartości dopasowania dla wielu obiektów. Inne przykłady wykorzystania filtracji krawędziowej z metodą POC można znaleźć w [149] w zastosowaniu do obrazów multispektralnych (MSI). Korelację fazową do dopasowania obrazów wizyjnych i termowizyjnych zastosowano również w [151], gdzie została wykorzystana w inny sposób do wyznaczenia lokalnego podobieństwa obrazów w okolicy punktów charakterystycznych wyznaczonych za pomocą algorytmu SURF.

4.3 Procedura samokalibracji

W bieżącym podrozdziale przedstawiono procedurę opracowanej metody samokalibracji z wykorzystaniem automatycznego dopasowania obrazów wizyjnego i termowizyjnego. Kolejno opisano skrótowo sposób realizacji etapu pierwszego, a następnie sposób realizacji etapu drugiego samokalibracji, po czym przedstawiono procedurę samokalibracji z uwzględnieniem dwóch opisanych etapów. Na koniec przedstawiono opracowaną metodę na tle innych publikacji dotyczących kalibracji.

Założono dodatkowo, że przed przystąpieniem do procedury samokalibracji zostanie przeprowadzony techniczny etap wstępny przygotowania obrazów. Na tym etapie wstępnym nastąpi ręczne wybranie obszarów zainteresowania dla obydwu kamer: wizyjnej i termowizyjnej. Kamery często zwracają obraz uzupełniony o “ramkę” zawierającą informacje na przykład o trybie akwizycji. Jest to czarny lub zapisany różnymi informacjami pas pikseli nie zawierający pikseli reprezentujących obiekty sceny. Obszar zainteresowania wyklucza tę ramkę z dalszej analizy.

4.3.1 Etap pierwszy

Na pierwszym etapie samokalibracji należy ustalić parametry macierzy F_S opisanej wzorem (4.14). Można to zrealizować na podstawie wzorów (4.24) oraz (4.25), które można przedstawić w postaci jednego równania (4.66).

$$\begin{bmatrix} \dot{x}_{ir} \\ \dot{y}_{ir} \end{bmatrix} = \begin{bmatrix} f_{irx} \dot{x}_{tv} \\ f_{iry} \dot{y}_{tv} \end{bmatrix} \quad (4.66)$$

gdzie:

$$\begin{bmatrix} \dot{x}_{tv} \\ \dot{y}_{tv} \end{bmatrix} - \text{wektor ruchu kamery TV względem obserwowanej sceny,}$$

$$\begin{bmatrix} \dot{x}_{ir} \\ \dot{y}_{ir} \end{bmatrix} - \text{wektor ruchu kamery IR względem obserwowanej sceny.}$$

Estymację ruchu obrazów IR i TV można łatwo wyznaczyć jako dopasowanie monomodalne obrazu z chwili bieżącej g_{ir}, g_{tv} do obrazu z chwili poprzedniej g_{irpop}, g_{tvpop} odpowiednio dla kamery IR i TV. Takie dopasowanie można wyznaczyć stosując w odpowiedni sposób metodę korelacji fazowej opisaną równaniem (4.37), co pokazują zależności (4.67).

$$\begin{bmatrix} \widehat{\dot{x}}_{ir} \\ \widehat{\dot{y}}_{ir} \end{bmatrix} = \arg \max_{x,y} \mathcal{F}^{-1} \left(\frac{G_{ir}(u,v)G_{irpop}^*(u,v)}{|G_{ir}(u,v)G_{irpop}^*(u,v)|} \right)$$

$$\begin{bmatrix} \widehat{\dot{x}}_{tv} \\ \widehat{\dot{y}}_{tv} \end{bmatrix} = \arg \max_{x,y} \mathcal{F}^{-1} \left(\frac{G_{tv}(u,v)G_{tvpop}^*(u,v)}{|G_{tv}(u,v)G_{tvpop}^*(u,v)|} \right) \quad (4.67)$$

gdzie:

$$\begin{bmatrix} \widehat{\dot{x}}_{ir} \\ \widehat{\dot{y}}_{ir} \end{bmatrix} - \text{estymacja ruchu } (\dot{x}_{ir}, \dot{y}_{ir}) \text{ obrazu IR,}$$

$\begin{bmatrix} \widehat{x_{tv}} \\ \widehat{y_{tv}} \end{bmatrix}$ - estymacja ruchu (x_{tv}, y_{tv}) obrazu TV,
 $\widehat{G}(u, v) = \mathcal{F}(g(x, y))$.

Powyższy wzór pozwala estymować przemieszczenie, ponieważ zakładamy, że w przypadku dopasowania monomodalnego prawdziwe są następujące równości:

$$\mathcal{F}^{-1}\left(\frac{G_{ir}(u, v)G_{irpop}^*(u, v)}{|G_{ir}(u, v)G_{irpop}^*(u, v)|}\right) = \mathcal{F}^{-1}(e^{-j\Delta_{ir}(u, v)}) = \delta(x - x_{ir}, y - y_{ir})$$

$$\mathcal{F}^{-1}\left(\frac{G_{tv}(u, v)G_{tvpop}^*(u, v)}{|G_{tv}(u, v)G_{tvpop}^*(u, v)|}\right) = \mathcal{F}^{-1}(e^{-j\Delta_{tv}(u, v)}) = \delta(x - x_{tv}, y - y_{tv})$$

Podsumowując równania (4.66) oraz (4.67) etap pierwszy przebiega zgodnie z równaniem (4.68).

$$\arg \max_{x, y} (\mathcal{F}^{-1}\left(\frac{G_{ir}(u, v)G_{irpop}^*(u, v)}{|G_{ir}(u, v)G_{irpop}^*(u, v)|}\right)) = \begin{bmatrix} f_{irx} & 0 \\ 0 & f_{iry} \end{bmatrix} \arg \max_{x, y} (\mathcal{F}^{-1}\left(\frac{G_{tv}(u, v)G_{tvpop}^*(u, v)}{|G_{tv}(u, v)G_{tvpop}^*(u, v)|}\right)) \quad (4.68)$$

4.3.2 Etap drugi

Na drugim etapie kalibracji należy ustalić parametry macierzy F_T opisanej wzorem (4.15) posługując się równaniem (4.30) na podstawie dopasowania przeskalowanego obrazu IR do obrazu TV. Dopasowanie obrazów rozumiane jest jako dopasowanie obiektów widocznych na przeskalowanym obrazie IR do odpowiadających im obiektów widocznych na obrazie TV.

Estymacja dopasowania wspomnianych obrazów multimodalnych może być wyznaczona z wykorzystaniem zależności (4.61) w celu znalezienia lokalizacji kilku największych maksimów macierzy $p_w(x, y)$ opisanej wzorem (4.58). Można to podsumować wzorem (4.69).

$$\begin{bmatrix} \mathbf{D} & \mathbf{1} \end{bmatrix} = \begin{bmatrix} \widehat{d_{x1}} & \widehat{d_{y1}} & 1 \\ \widehat{d_{x2}} & \widehat{d_{y2}} & 1 \\ \dots & & \\ \widehat{d_{xk}} & \widehat{d_{yk}} & 1 \end{bmatrix} = k \arg \max_{x, y} \mathcal{F}^{-1}\left(\frac{G_{tv}(u, v)G_{ir}^*(u, v)}{|G_{tv}(u, v)G_{ir}^*(u, v)|}\right) \quad (4.69)$$

Znalezienie lokalizacji kilku największych maksimów $p_w(x, y)$ pozwala estymować dopasowanie obiektów na obrazach multimodalnych, ponieważ zakładamy, że w takim przypadku prawdziwa jest równość:

$$= \sum_{i=0}^n \delta(x - d_{xi}, y - d_{yi}) * \mathcal{F}^{-1}\left(\frac{|G_{tvi}(u, v)|^2}{|G_{tv}(u, v)||G_{ir}(u, v)|}\right) + \sum_{i=0, k=0, i \neq k}^n \mathcal{F}^{-1}\left(\frac{G_{tvi}(u, v)G_{irk}^*(u, v)}{|G_{ir}(u, v)||G_{tv}(u, v)|}\right) =$$

Podsumowując zależności (4.30) oraz (4.69) etap drugi przebiega zgodnie z równaniem (4.70).

$$k \arg \max_{x, y} \left(\mathcal{F}^{-1}\left(\frac{G_{tv}(u, v)G_{ir}^*(u, v)}{|G_{tv}(u, v)||G_{ir}(u, v)|}\right) \right) \begin{bmatrix} t_y \\ -t_x \\ c \end{bmatrix} = 0 \quad (4.70)$$

4.3.3 Procedura kalibracji

Procedura samokalibracji na etapie pierwszym według opracowanej metody przebiega w następujący sposób:

- **Etap wstępny**

- Na etapie wstępnym przygotowania obrazów następuje ręczny wybór obszarów zainteresowania dla obydwu kamer: wizyjnej i termowizyjnej. Obszar zainteresowania obejmuje tylko piksele obrazu z danej kamery, które reprezentują elementy obserwowanego otoczenia. Zostało to opisane w rozdziale 4.1.

- **Etap I**

- Uruchomienie algorytmu realizacji pierwszego etapu samokalibracji, czyli automatycznego wyznaczania rozmiaru obrazu termowizyjnego względem wizyjnego. Jest to realizowane przez estymację parametrów f_{irx}, f_{iry} opisujących skalę obrazu IR względem TV. Parametry te spełniają zależności (4.17) oraz (4.66).
- W celu ich estymacji opracowana metoda wykorzystuje w trybie kalibracji automatyczne pomiary dopasowania monomodalnego obrazów z chwili bieżącej do obrazów z chwili poprzedniej zgodnie z zależnością (4.68). Wyznaczenie parametrów modelu opisanego równaniem (4.68) tak, aby odpowiadały pomiarom dopasowania wykonano na podstawie algorytmu RANSAC.
- Proces estymacji parametrów f_{irx}, f_{iry} został opisany dokładniej w postaci algorytmu 4.4.

Procedura samokalibracji na etapie drugim według opracowanej metody przebiega w następujący sposób:

- **Etap wstępny**

- Na etapie wstępnym przygotowania obrazów następuje ręczne wybranie obszarów zainteresowania dla obydwu kamer: wizyjnej i termowizyjnej. Obszar

zainteresowania obejmuje tylko piksele obrazu z danej kamery, które reprezentują elementy obserwowanego otoczenia. Zostało to opisane w rozdziale 4.1.

- Dodatkowo względem wybrania obszaru zainteresowania konieczne jest ręczne ustalenie parametrów f_{irx}, f_{iry} opisujących rozmiar obrazu termowizyjnego względem wizyjnego.

• **Etap II**

- Uruchomienie algorytmu realizacji etapu drugiego samokalibracji, czyli automatycznego wyznaczania parametrów opisu linii dopasowania. Jest to realizowane przez estymację parametrów t_x, t_y, c . Parametry te spełniają zależności (4.17) oraz (4.30).
- W celu ich estymacji opracowana metoda wykorzystuje w trybie kalibracji automatyczne pomiary dopasowania multimodalnego obrazów zgodnie z zależnością (4.70). Wyznaczenie parametrów modelu opisanego równaniem (4.70) tak, aby odpowiadały pomiarom dopasowania, wykonano na podstawie algorytmu RANSAC.

W nawiązaniu do powyższego opisu zaproponowano dwa algorytmy realizacji samokalibracji odpowiednio na etapie pierwszym i drugim. Algorytm 4.4 realizacji etapu pierwszego można ogólnie opisać w następujący sposób:

1. Pobierz kolejną parę obrazów IR i TV.
2. Przeprowadź filtrację krawędziową dla tych obrazów (wiersz 10).
3. Wyznacz macierz korelacji fazowej \mathbf{p}_{wir} i \mathbf{p}_{wtv} przetworzonych wstępnie obrazów wizyjnego i termowizyjnego względem odpowiednich obrazów z chwili poprzedniej (wiersze 12 oraz 13).
4. Znajdź największe maksima \mathbf{p}_{wir} i \mathbf{p}_{wtv} (wiersze 14 oraz 15) i zapisz ich lokalizację jako \mathbf{d}_{ir} oraz \mathbf{d}_{tv} . Lokalizacja tych maksimów określa dopasowanie obrazów monomodalnych, czyli estymuje ruch obrazów w sekwencji wideo.
5. Jeśli pomiary dopasowania odpowiednio w osi X oraz Y są niezerowe, dodaj je do zbioru odpowiednio \mathbf{DO}_x oraz \mathbf{DO}_y (wiersze od 16 do 21).
6. Jeśli:
 - liczba zebranych pomiarów w zbiorach \mathbf{DO}_x oraz \mathbf{DO}_y jest jednocześnie większa od dobranej eksperymentalnie wartości minimalnej (odpowiednio $\min_{LDO_x}, \min_{LDO_y}$)
 - lub
 - ogólna liczba pomiarów dodanych do zbiorów \mathbf{DO}_x oraz \mathbf{DO}_y wyrażona przez $licznik_{DP}$ jest mała w stosunku do liczby wszystkich wykonanych pomiarów $licznik_{WP}$, ale pomiarów dodanych do jednego ze zbiorów \mathbf{DO}_x oraz \mathbf{DO}_y jest wystarczająco dużo,

przeprowadź estymację parametrów f_{irx} oraz f_{iry} stosując algorytm RANSAC.

Algorytm 4.4 Etap pierwszy samokalibracji

```
1: procedure SAMOKALIBRACJAETAPPIERWSZY
2:    $\mathbf{g}_{irpop} \leftarrow \emptyset$ 
3:    $\mathbf{g}_{tvpop} \leftarrow \emptyset$ 
4:    $\mathbf{DO}_x \leftarrow \emptyset$ 
5:    $\mathbf{DO}_y \leftarrow \emptyset$ 
6:    $licznik_{WP} \leftarrow 0$ 
7:    $licznik_{DP} \leftarrow 0$ 
8:   while true do
9:     pobierz obrazy  $\mathbf{g}_{ir}, \mathbf{g}_{tv}$  z głowicy kamer
10:    ( $\mathbf{g}_{iredge}, \mathbf{g}_{tvedge}$ )  $\leftarrow$  przetworzWstepnieObrazy( $\mathbf{g}_{ir}, \mathbf{g}_{tv}$ )
11:    if  $\mathbf{g}_{irprev} \neq \emptyset$  and  $\mathbf{g}_{tvpref} \neq \emptyset$  then
12:       $\mathbf{p}_{wir} \leftarrow$  wyznaczMacierzPOC( $\mathbf{g}_{ir}, \mathbf{g}_{irpop}$ )
13:       $\mathbf{p}_{wtv} \leftarrow$  wyznaczMacierzPOC( $\mathbf{g}_{tv}, \mathbf{g}_{tvpop}$ )
14:       $\mathbf{d}_{ir} = (d_{irx}, d_{iry}) \leftarrow$  arg max  $\mathbf{p}_{wir}$ 
15:       $\mathbf{d}_{tv} = (d_{tvx}, d_{tvy}) \leftarrow$  arg max  $\mathbf{p}_{wtv}$ 
16:      if ( $d_{tvx} \neq 0$  and  $d_{irx} \neq 0$ ) or ( $d_{tvy} \neq 0$  and  $d_{iry} \neq 0$ ) then
17:        if  $d_{tvx} \neq 0$  and  $d_{irx} \neq 0$  then
18:          dodaj parę ( $d_{irx}, d_{tvx}$ ) do zbioru  $\mathbf{DO}_x$ 
19:        if  $d_{tvy} \neq 0$  and  $d_{iry} \neq 0$  then
20:          dodaj parę ( $d_{iry}, d_{iry}$ ) do zbioru  $\mathbf{DO}_y$ 
21:           $licznik_{DP} \leftarrow licznik_{DP} + 1$ 
22:           $licznik_{WP} \leftarrow licznik_{WP} + 1$ 
23:           $war_{MLP} \leftarrow$  (liczność  $\mathbf{DO}_x > min_{LDO_x}$ ) and (liczność  $\mathbf{DO}_y > min_{LDO_y}$ )
24:           $war_{ZLP} \leftarrow$  (liczność  $\mathbf{DO}_x \geq max_{LDO_x}$ ) or (liczność  $\mathbf{DO}_y \geq max_{LDO_y}$ )
25:           $war_{NLWP} \leftarrow$  ( $licznik_{DP}/licznik_{WP} < min_{LW}$ )
26:          if  $war_{MLP} = true$  or ( $war_{NLWP} = true$  and  $war_{ZLP} = true$ ) then
27:            wyznacz skalę obrazu  $\mathbf{g}_{ir}$  względem  $\mathbf{g}_{tv}$  w osi X  $f_{irx} \leftarrow RANSAC(\mathbf{DO}_x)$ 
28:            wyznacz skalę obrazu  $\mathbf{g}_{ir}$  względem  $\mathbf{g}_{tv}$  w osi Y  $f_{iry} \leftarrow RANSAC(\mathbf{DO}_y)$ 
29:            zakończ samokalibrację return
30:       $\mathbf{g}_{irpop} \leftarrow \mathbf{g}_{iredge}$ 
31:       $\mathbf{g}_{tvpop} \leftarrow \mathbf{g}_{tvedge}$ 
```

Z kolei Algorytm 4.5 realizacji etapu drugiego samokalibracji obrazów można ogólnie opisać w następujący sposób:

1. Pobierz kolejną parę obrazów IR i TV.
2. Przeprowadź filtrację krawędziową dla tych obrazów (wiersz 6).
3. Wyznacz macierz korelacji fazowej \mathbf{p}_w przetworzonych wstępnie obrazów (wiersz 8).
4. Znajdź kilka największych maksimum macierzy \mathbf{p}_w (wiersz 8) i zapisz ich lokalizację w tablicy \mathbf{D} . Lokalizacja tych maksimum określa dopasowanie obrazów multimodalnych.
5. Jeśli \mathbf{D} zawiera jedno dopasowanie i odpowiadająca mu wysokość maksimum jest większa od średniej wartości wszystkich dopasowań zebranych w \mathbf{DO} (wiersz 11), dodaj to dopasowanie do zbioru \mathbf{DO} (wiersz 12).
6. Jeśli \mathbf{D} zawiera wiele dopasowań, dodaj wszystkie do zbioru \mathbf{DO} .

7. Jeśli:

- ilość zebranych pomiarów dopasowania w DO jest większa od eksperymentalnie dobranej wartości minimalnej
oraz
- geometryczne odchylenie standardowe zebranych w DO pomiarów dopasowania jest większe od dobranej eksperymentalnie wartości minimalnej,

przeprowadź estymację parametrów linii dopasowania stosując algorytm RANSAC. Algorytm RANSAC wyznacza bezpośrednio parametry t_x, t_y, c z równania (4.70). Na tej podstawie wyznaczane jest nachylenie $slope$ oraz odległość do środka układu współrzędnych $shift$ linii dopasowania.

Algorytm 4.5 Etap drugi samokalibracji

```
1: procedure SAMOKALIBRACJAETAPDRUGI
2:   wyczyść zbiór dopasowań DO  $\leftarrow \emptyset$ 
3:   while true do
4:     pobierz obrazy  $\mathbf{g}_{ir}, \mathbf{g}_{tv}$  z głowicy kamer
5:     dopasuj skalę obrazów  $\mathbf{g}_{ir}, \mathbf{g}_{tv}$  zgodnie z parametrami  $s_x$  oraz  $s_y$ 
6:      $(\mathbf{g}_{tvedge}, \mathbf{g}_{vedge}) \leftarrow \text{przetworzWstepnieObrazy}(\mathbf{g}_{ir}, \mathbf{g}_{tv})$ 
7:      $\mathbf{p}_w \leftarrow \text{wyznaczMacierzPOC}(\mathbf{g}_{iredge}, \mathbf{g}_{tvedge})$ 
8:      $\mathbf{D} \leftarrow k \arg \max \mathbf{p}_w$ 
9:      $min_{QD} \leftarrow$  średnia wysokość maksimum z DO
10:    if liczność  $\mathbf{D} = 1$  then
11:      if wysokość maksimum jedyne dopasowania w  $\mathbf{D} > min_{QD}$  then
12:        dodaj jedyne dopasowanie z  $\mathbf{D}$  do zbioru DO
13:    else if liczność  $\mathbf{D} > 1$  then
14:      dodaj wszystkie dopasowania z  $\mathbf{D}$  do zbioru DO
15:     $w_x \leftarrow$  wariancja wartości dopasowania ze zbioru DO w osi X
16:     $w_y \leftarrow$  wariancja wartości dopasowania ze zbioru DO w osi Y
17:     $w_g \leftarrow \sqrt{w_x^2 + w_y^2}$ 
18:    if liczność DO  $> min_{LDO}$  and  $w_g > min_{WGDO}$  then
19:      wyznacz parametry linii dopasowania  $(slope, shift) \leftarrow RANSAC(\text{DO})$ 
20:      zakończ samokalibrację return
```

4.3.4 Opracowana metoda na tle innych publikacji

W literaturze opisano nieliczne metody samokalibracji układu kamer wizyjnej i termowizyjnej. W szczególności należy podkreślić pracę [51], gdzie autorzy opisują proces automatycznej kalibracji układu dwóch sensorów: kamery typu RGBD oraz kamery termowizyjnej. Aplikacja opisana w tym artykule oraz przyjęte założenia są jednak różne od założeń przyjętych w rozprawie - autorzy wyznaczają tylko macierz parametrów zewnętrznych układu kamer. W rozprawie pokazano, że w świetle przyjętych założeń konieczna jest kalibracja innych parametrów definiujących układ kamer - współczynniki skali obrazu termowizyjnego względem

wizyjnego oraz wybrane parametry macierzy fundamentalnej opisujące linię dopasowania. W [51] w podobnie jak w rozprawie wykorzystana została informacja o przemieszczeniu układu sensorów względem obserwowanej sceny oraz filtracja krawędziowa w celu upodobnienia do siebie obrazów o różnych modalnościach. Istnieją jednak znaczące różnice. Obserwowany ruch wykorzystany został w [51] do synchronizacji sekwencji danych z sensorów. Zastosowana filtracja krawędziowa dotyczy nie tylko obrazów wizyjnych i termowizyjnych, ale również mapy głębi. Można przypuszczać, że ułatwia to proces kalibracji, ponieważ kamera głębi i kamera wizyjna są skalibrowane. Dodatkowo filtracja krawędziowa obrazu termowizyjnego wyznaczona jest jako lokalne maksimum różnic intensywności jego negatywu w otoczeniu piksela. Takie podejście rozważano w ramach badań nad tematem rozprawy, jednakże ostatecznie wykorzystanie bezwzględnych różnic w intensywnościach okazało się dawać lepsze rezultaty.

Z uwagi na nieliczne publikacje dotyczące samokalibracji układu kamer spełniających założenia przyjęte w rozprawie, opracowaną metodę można odnieść do metod samokalibracji układu kamer monomodalnych. W tym zakresie istnieje wiele metod kalibracji podobnych do tej opracowanej w rozprawie. W znaczącej ilości przypadków w celu samokalibracji stosowane jest podejście wieloetapowe [13, 14, 18]. Czasami najpierw znajdowane są parametry opisujące przekształcenie perspektywiczne w postaci macierzy fundamentalnej [172] lub tensorów trójogniskowych (*trifocal tensor*) [173, 26]. W szczególności podobne podejście do wyznaczenia współczynnika skali jak w pierwszym etapie opracowanej metody przedstawiono w [8] w odniesieniu do stereowizyjnego układu kamer monomodalnych z wykorzystaniem cech charakterystycznych wybranych punktów obrazów. W kolejnym etapie autorzy szukają możliwości określenia parametrów pozwalających na metryczną rekonstrukcję obserwowanej sceny [9, 3, 174]. Wieloetapową samokalibrację można znaleźć również w [35]. W pracy tej macierz perspektywiczna rozkładana jest na trzy przekształcenia. Przedstawiona metoda jest jednak dość niepraktyczna - wymaga znajomości jednego kąta, równości dwóch nieznanymi kątów lub znajomości stosunku długości odcinków na pewnej obserwowanej płaszczyźnie. W przedstawionej w rozprawie metodzie macierz fundamentalna rozkładana jest na dwa przekształcenia. W pierwszym etapie ustalane są parametry macierzy wewnętrznej i na bazie tych efektów w etapie drugim estymowane są parametry macierzy zewnętrznej.

Przyjęte w rozprawie założenia związane z ruchem kamery względem obserwowanej sceny można znaleźć np. w pracy [16], gdzie autorzy prezentują samokalibrację na podstawie ruchów translacyjnych kamery. W swojej metodzie nie estymują jednak parametrów wewnętrznych. Niektórzy autorzy podobnie jak w opracowanej metodzie przyjęli, że ruchy kamery będą płaszczyznowe [26]. Ruchy płaszczyznowe zakładają przesunięcia w ramach pewnej płaszczyzny 2D i obrót wokół osi prostopadłej do tej płaszczyzny. W opracowanej w rozprawie metodzie jest podobne założenie. W założeniu kalibrowane kamery są umieszczone na głowicy obrotowej, a obserwowana scena jest w znacznej odległości. W wyniku tego

przyjęto, że ruchy kamery można opisać jako translację w ramach płaszczyzny równoległej do obserwowanej sceny. Rotacje są pomijane.

Podobne założenie dotyczące obserwowanej sceny można znaleźć w [30]. Jednakże tam autorzy określają parametry układu kamer na podstawie układu płaszczyzn wzajemnie ortogonalnych. Takie założenie jest nie do przyjęcia w warunkach poligonowych. Kalibracja na podstawie obserwacji płaskich wzorców przeprowadzona jest również w [33]. W celu kalibracji wyznaczane są homografie pomiędzy płaszczyznami i na podstawie tych macierzy określono równanie Kruppa, które posłużyło następnie w celu kalibracji. Obserwację płaskiego wzorca kalibracyjnego zakłada również praca [34]. Autor przedstawił relację homografii płaszczyzn z macierzami wewnętrznymi kamer i przeprowadza wykrywanie odpowiadających sobie punktów na podstawie wydrukowanego wzorca. We wszystkich tych przypadkach samokalibracji na podstawie obserwacji płaskich obiektów metody zostały przetestowane tylko na podstawie obrazu wzorców kalibracyjnych w warunkach laboratoryjnych. W pracach [46, 47, 39, 40] przedstawiono również metody samokalibracji układu kamer bazujące na obserwacji scen płaskich.

W przypadku samokalibracji układu kamer, podobnie jak w opracowanej metodzie, większość autorów wykorzystuje więzy epipolarne [46, 47, 39, 40, 48, 49]. Metody kalibracji wykorzystujące tę zależność są zazwyczaj szybkie w działaniu i dobrze radzą sobie z obiektami ruchomymi sceny 3D [49]. W pracy [48] przedstawiono samokalibrację na podstawie jednego przemieszczenia układu kamer przyjmując uproszczony model kamery. Autorzy zakładają, że środki obrazu kamer są znane oraz, że nie ma przekoszenia kamer. Poszukiwana jest jedynie skala obrazów w osi X i Y. Podobne założenia są przyjęte w rozprawie. Jednakże dla potrzeb fuzji pokazano, że nie jest istotna skala każdego z obrazów układu kamer, a jedynie wzajemny stosunek tych skal. W pracy [47] pokazano dopasowanie do siebie wybranych punktów obrazów układu kamer, a nie jak w przypadku opracowanej w rozprawie metody, całych obrazów. Dopasowanie obrazów, a nie punktów zwiększa odporność opracowanej metody. Jest to szczególnie istotne w przypadku dopasowania obrazów multimodalnych. Ponadto w publikacji [47] nie wykorzystano zalet podziału samokalibracji na kilka etapów, zaś relacja będąca podstawą kalibracji przedstawiona jest w funkcji współrzędnych odpowiadających punktów, a nie w funkcji dopasowania. W pracy [49] autorzy przekonują, że metoda kalibracji powinna działać on-line w trakcie pracy systemu wizyjnego w celu ciągłego monitorowania i poprawy jakości jego działania. W opracowanej w rozprawie metodzie zastosowano szybkie algorytmy pozwalające na krótki czas działania. Pod tym kątem opracowana metoda spełnia wymagania czasowe przedstawione w pracy [49] i mogłaby zostać wykorzystana w trybie pracy on-line. We wspomnianym artykule autorzy wykorzystują więzy epipolarne do chwilowej estymacji parametrów i ciągłą rekursywną pracę metody, aby agregować dane z wielu różnych chwil czasowych. Jest to strategia zbliżona do przyjętej w opracowanej w rozprawie metodzie. Metoda z publikacji [49] została rozwinięta w pracy [50]. Autorzy nie przeprowadzają tam jednak estymacji parametrów wewnętrznych kamer. W rozprawie wykazano, że na potrzeby

fuzji układu kamer multimodalnych konieczne jest dodatkowo ustalenie stosunku skali obrazów z różnych kamer. Metoda opisana w pracy [50] nie znajduje więc zastosowania w takim przypadku.

W zakresie metod automatycznego dopasowania obrazów należy podkreślić, że w literaturze nie znaleziono żadnych przypadków aplikacji metody korelacji fazowej do dopasowania kilku obiektów na obrazach z kamer: wizyjnej i termowizyjnej. Autor rozprawy przedstawił również sposób wykorzystania tych dopasowań w procesie samokalibracji w sytuacji, gdy nie są znane obiekty, do których te dopasowania można przyporządkować. Analiza działania metody korelacji fazowej w przypadku wielu obiektów stanowi wykonane przez autora rozprawy rozwinięcie podejścia przedstawionego w pracy [146] dla dwóch obiektów. W literaturze nie znaleziono analizy działania metody korelacji fazowej dla wielu obiektów o różnych intensywnościach. W rozprawie zamieszczono opracowane przez autora uzasadnienie problemów z wyznaczeniem dopasowania przy niejednorodnych zmianach intensywności w dziedzinie obrazu pomiędzy obrazami odpowiadających sobie obiektów. Obiekty mogą być rozumiane jako krawędzie obserwowanych na obrazach kształtów. Samo stosowanie metody korelacji fazowej do znalezienia wartości kilku dopasowań na obrazach z jednej kamery nie jest nowością. W literaturze [137, 143, 144, 145, 142] opisywane są sposoby estymacji ruchu wielu obiektów w sekwencji obrazów z wykorzystaniem korelacji fazowej. Wiele maksimum korelacji wyznaczanych jest w podobny sposób [143] jak w rozprawie. Warto podkreślić, że standardowe wykorzystanie korelacji fazowej do estymacji ruchu wielu obiektów obarczone jest istotnym problemem związanym z brakiem informacji o tym, którym fragmentom obrazu odpowiadają wyznaczone estymacje ruchu. W rozprawie przedstawiono sposób wykorzystania dopasowań bez konieczności posiadania tej informacji. W ten sposób zaprojektowano metodę samokalibracji, która korzysta z zalet korelacji fazowej bez negatywnych skutków wad tej metody wyznaczenia dopasowań w zakresie określenia ich lokalizacji przestrzennej.

W zakresie stosowania filtracji krawędziowej opracowana w rozprawie metoda jest podobna na przykład do rozwiązań w pracach [95, 96]. W [95] zastosowano podobnie jak w rozprawie filtrację krawędziową z wykorzystaniem progowania, ale nie zastosowano wartości bezwzględnej wyznaczonego gradientu. Układ kamer i wykorzystanie filtracji krawędziowej w celu dopasowania obrazów wizyjnego i termowizyjnego przedstawione w rozprawie są podobne do sytuacji opisanej w pracy [97]. Jednakże tam dopasowywane są pojedyncze obiekty (w postaci statków na morzu) i nie wykorzystano korelacji fazowej, ale metodę bazującą na korelacji wzajemnej i poszukiwaniu dopasowania z uwzględnieniem różnic w skali pomiędzy obrazami.

Pomimo wielu podobieństw pomiędzy opracowaną w rozprawie metodą, a metodami opisanymi w literaturze, opracowana metoda nosi istotne znamiona oryginalności. Kalibracja układu kamer wizyjnej i termowizyjnej dotychczas prawie zawsze była realizowana z wykorzystaniem wzorca. Dopiero w ostatnich latach pojawiły się publikacje dotyczące samokalibracji podobnego układu kamer, jednak w większości przypadków dotyczą one

tylko wyznaczenia parametrów zewnętrznych. Opracowana metoda wykorzystuje podobne zależności jak metody samokalibracji układu kamer stereowizyjnych. Jest ponadto dostosowana do szybkiej implementacji dla układu kamer z przeznaczeniem do fuzji obrazów w celu monitorowania odległej sceny. Odporność metody w odniesieniu do samokalibracji obrazów multimodalnych poprawia wykorzystanie dopasowania całych obrazów, a nie wybranych punktów, nie wymaga ona konieczności bezpośredniego wykrywania cech obrazów obserwowanych.

Rozdział 5

Przeprowadzone badania

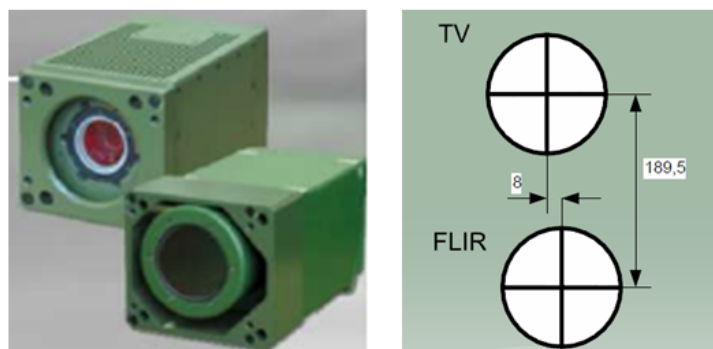
5.1 Stanowisko badawcze i bazy danych obrazów

W niniejszym rozdziale opisano wyniki działania metody samokalibracji opracowanej dla potrzeb przemysłu obronnego. Opracowane podstawy teoretyczne i analizy przetestowano wstępnie w warunkach laboratoryjnych, a następnie zaimplementowano na dedykowanym komputerze mobilnym Jetson TK1. Na podstawie przygotowanej implementacji przetestowano działanie metody w warunkach zbliżonych do poligonowych na głowicy zestawu przeciwlotniczego Poprad.

W kolejnych podrozdziałach opisano stanowisko badawcze oraz przygotowane zestawy par obrazów wizyjnych i termowizyjnych odpowiadających sobie pod względem czasu akwizycji. Przedstawiono sposób wykorzystania obrazów dla potrzeb ewaluacji opracowanej metody samokalibracji.

5.1.1 Głowica kamer

Sekwencje wideo, które składają się na podstawową bazę danych testowych nagrano za pomocą głowicy kamer przeciwlotniczego zestawu raketowego Poprad, wyposażonej w dwie kamery: wizyjną (TV) i termowizyjną (IR) (rys. 1.4 w rozdziale 1). Wzajemne rozmieszczenie kamer przedstawiono schematycznie na rysunku 5.1. Kamery udostępniały sygnały analogowe w formacie CCIR (625 linii) o rozdzielczościach odpowiednio 756×581 pikseli obrazu TV oraz 384×256 pikseli obrazu IR. Każda kamera z tej głowicy posiada możliwość akwizycji obrazów przy dwóch różnych kątach widzenia (elektronicznie sterowana zmiana optyki kamery). W pracy rozważane są najczęściej występujące sytuacje, w których obie kamery są ustawione na szeroki kąt lub obie na wąski kąt postrzegania. Inne konfiguracje ustawienia optyki występują w praktyce monitorowania otoczenia zestawu przeciwlotniczego Poprad jedynie jako stany przejściowe.



Rysunek 5.1: Kamery głowicy zestawu przeciwlotniczego Poprad i ich rozmieszczenie względem siebie.

5.1.2 Sekwencje testowe

Na potrzeby testów algorytmu samokalibracji opisanego w rozdziale 4 przygotowano zestaw sekwencji par obrazów testowych. W bieżącym podrozdziale scharakteryzowano wszystkie przygotowane sekwencje testowe. Na zestaw danych testowych składają się proste sekwencje wygenerowane komputerowo, sekwencje par obrazów pochodzące z głowicy optoelektronicznej zestawu Poprad oraz sekwencje z bazy danych KAIST [175]. Każda z sekwencji liczy po kilkaset par obrazów. Na sekwencje wygenerowane komputerowo składają się obrazy przygotowane w celu przetestowania etapu I oraz etapu II samokalibracji. Sekwencje wygenerowane komputerowo do testów poszczególnych etapów zostały przygotowane w różny sposób z uwagi na różne mechanizmy estymacji parametrów na etapie I i II samokalibracji.

Na rysunkach 5.2 przedstawiono przykładowe obrazy z sekwencji wygenerowanych komputerowo do testów wstępnych etapu II samokalibracji, zaś na rysunkach 5.3, 5.4, 5.5, 5.6 - przykładowe pary obrazów pochodzące z sekwencji testowych Poprad. Dla każdej sekwencji obrazów wygenerowanej komputerowo wartość parametrów ustalanych na etapie kalibracji była znana. Dla każdej sekwencji obrazów zarejestrowanych przez głowicę Poprad ustalono manualnie zestaw wzorcowych parametrów kalibracji. Parametry te traktowane są jako idealnie opisujące kalibrowany układ kamer i względem nich określana jest dokładność parametrów wyznaczonych przez samokalibrację.

Sekwencje testowe wygenerowane komputerowo

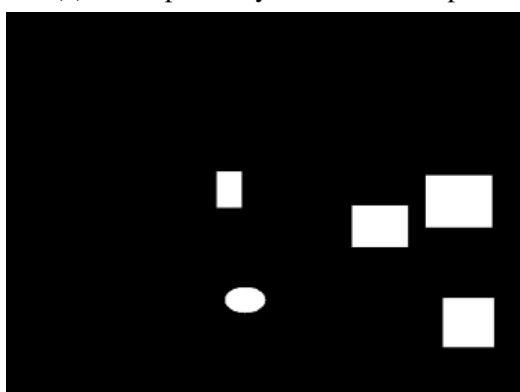
Sekwencje o numerach as036-as056 przygotowano do testów etapu I samokalibracji. Zostały one wygenerowane komputerowo jako modyfikacja wybranej sekwencji zarejestrowanej przez głowicę Poprad. Przykładowe obrazy z tych sekwencji przedstawiono na rys. 5.2a i 5.2b. Sekwencje wygenerowano tak, że każda para obrazów w sekwencji źródłowej została przesunięta o wektor zmieniający się w czasie trwania nagrania. Zmienność tego wektora symuluje ruch układu kamer względem obserwowanej sceny. Zachowano przy tym zależność



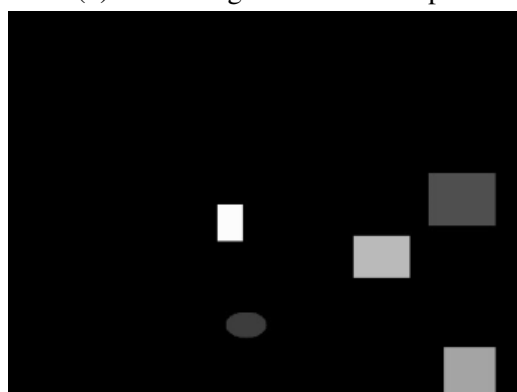
(a) obraz pierwszy do testów I etapu



(b) obraz drugi do testów I etapu



(c) obraz pierwszy do testów II etapu



(d) obraz drugi do testów II etapu

Rysunek 5.2: Przykładowe obrazy z sekwencji wygenerowanych komputerowo.

przesunięcia pomiędzy obrazem termowizyjnym a wizyjnym odpowiadającą skali obrazów termowizyjnych względem wizyjnych.

Sekwencje o numerach as001-as035 przygotowano do testów etapu II samokalibracji. Zostały one wygenerowane komputerowo i składają się z prostych kształtów geometrycznych o różnych poziomach intensywności umieszczonych na czarnym tle. Przykładowe obrazy z tych sekwencji przedstawiono na rys. 5.2c i 5.2d. Dopasowanie odpowiadających sobie obiektów zmienia się zgodnie z wygenerowanymi parametrami linii dopasowania. Parametry te są znane i wykorzystane jako wartości wzorcowe do wyznaczenia dokładności działania etapu II samokalibracji.

Sekwencje testowe Poprad

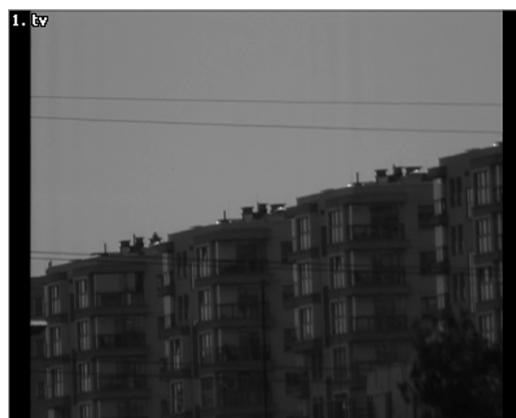
Sekwencje z zestawu Poprad widoczne na rys. 5.3, 5.4, 5.5 i 5.6 przedstawiają różne rodzaje otoczenia. Sekwencje o numerach 1, ..., 8 przedstawiają miejski teren zabudowany w różnych ujęciach. Występują na nich obiekty znajdujące się blisko układu kamer, takie jak drzewa, czy lampa uliczna, oraz obiekty znajdujące się dalej, takie jak blok mieszkalny. Na tych nagraniach wideo można zaobserwować zazwyczaj gwałtowne ruchy układu kamer względem otoczenia.

Sekwencja nr 1 przedstawia blok na planie dalekim i lampę uliczną na planie bliskim. Głowica porusza się w trakcie trwania nagrania. Przykładowe obrazy reprezentujące tę

sekwencje przedstawione są na rysunkach 5.3a, 5.3b. Sekwencje nr 2 i 3 przedstawiają drzewa, lampę uliczną oraz inne elementy zabudowy miejskiej w różnej odległości od układu kamer. W tle, na dalszym planie można przez zaobserwować bloki mieszkalne. Głowica porusza się w trakcie trwania nagrania. Przykładowe obrazy reprezentujące tę sekwencję przedstawione są na rysunkach 5.3c, 5.3d.



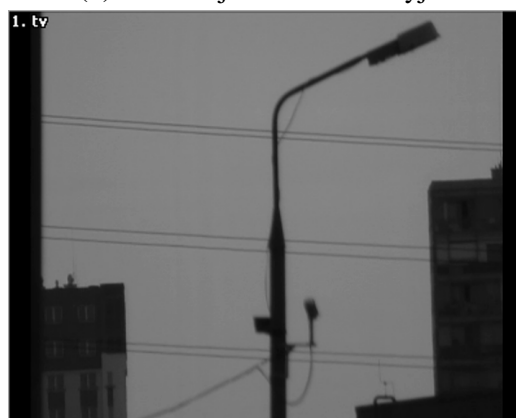
(a) sekwencja 1 kamera termowizyjna



(b) sekwencja 1 kamera wizyjna



(c) sekwencja 2 kamera termowizyjna



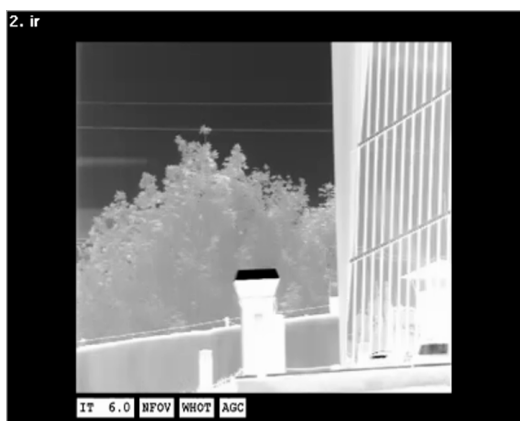
(d) sekwencja 2 kamera wizyjna

Rysunek 5.3: Wybrane obrazy z sekwencji nr 1, 2, 3.

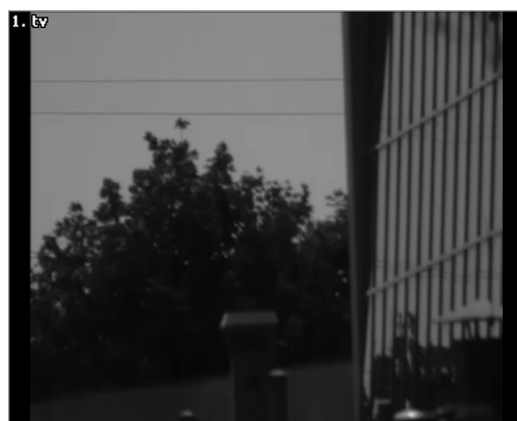
Sekwencje nr 4 i 5 przedstawiają blok na planie dalekim, drzewo w średniej odległości, elementy zabudowy miejskiej oraz lampę uliczną w bardzo bliskiej odległości od układu kamer. Głowica porusza się gwałtownie w trakcie trwania nagrania. Przykładowe obrazy reprezentujące tę sekwencję przedstawione są na rysunkach 5.4a. Podobną scenę, ale w innym ujęciu oraz przy innych ruchach układu kamer można zobaczyć na sekwencji nr 6. Przykładowe obrazy z tej sekwencji przedstawione są na rysunkach 5.4c, 5.4d.

Sekwencje nr 7 i 8 przedstawiają blok na planie dalekim, drzewo w średniej odległości, elementy zabudowy miejskiej w bardzo bliskiej odległości od układu kamer. Głowica porusza się w trakcie trwania nagrania, ale nie tak gwałtownie jak w przypadku sekwencji 4, 5, czy 6. Przykładowe obrazy reprezentujące tę sekwencję przedstawione są na rysunkach 5.5a, 5.5b.

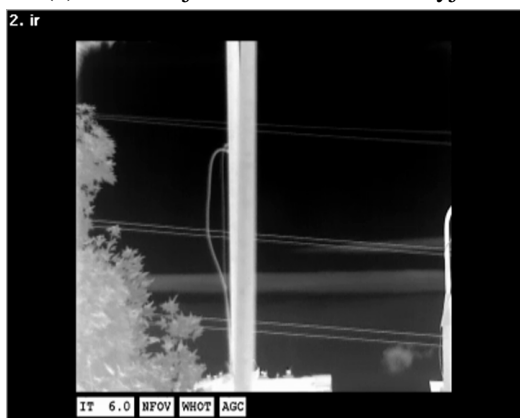
Sekwencje o numerach 9, 10, 11 przedstawiają inny rodzaj sceny. Głównym jej elementem jest samochód przemieszczający się połą drogą. W trakcie trwania tych nagrań samochód



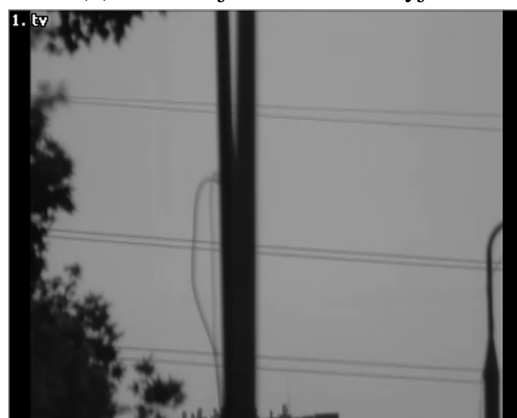
(a) sekwencja 4 kamera termowizyjna



(b) sekwencja 4 kamera wizyjna



(c) sekwencja 6 kamera termowizyjna



(d) sekwencja 6 kamera wizyjna

Rysunek 5.4: Wybrane obrazy z sekwencji nr 4, 5, 6.

oddala się od układu kamer. Sekwencja nr 9 przedstawia samochód w bardzo bliskiej odległości, sekwencja nr 10 przedstawia samochód w średniej odległości, a sekwencja nr 11 samochód w najdalszej odległości od układu kamer. Samochód i głowica poruszają się w trakcie trwania nagrania. Obraz z kamery wizyjnej zawiera bardzo mało szczegółów. Przykładowe obrazy reprezentujące te sekwencje przedstawione są na rysunkach 5.5c oraz 5.5d.

Sekwencje o numerach 12,13 obrazują scenę wiejską. Głównym elementem obserwowanej sceny jest wóz ciągnięty przez konia. Ruchy kamery na tych nagraniach są stosunkowo powolne. Obserwowany wóz przejeżdża powolnie polną drogą w stałej odległości od układu kamer. Na polu widoczna jest flaga powiewająca na wietrze. Na pierwszym planie widzimy pole i słup. W trakcie trwania sekwencji układ kamer porusza się nieznacznie. Przykładowe obrazy reprezentujące tę sekwencję przedstawione są na rysunkach 5.6a, 5.6b.

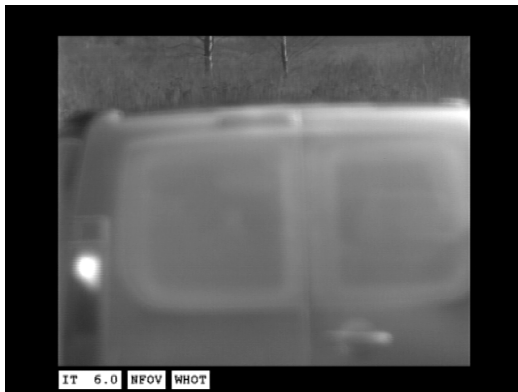
Sekwencje o numerach 14, ..., 17 przedstawiają zabudowę miejską podobną do sekwencji 1, ..., 8, ale obserwowaną w innej skali. Na sekwencji 14 obserwowany jest głównie blok na planie dalekim. Układ kamer nie porusza się w trakcie nagrania. Na sekwencji 15 można zauważyć ruch układu kamer. W wyniku tego obserwowane są obiekty w bliższej odległości od układu kamer - drzewa i komin. Podobna sytuacja występuje na sekwencji 16, ale elementami



(a) sekwencja 7 kamera termowizyjna



(b) sekwencja 7 kamera wizyjna



(c) sekwencja 9 kamera termowizyjna



(d) sekwencja 9 kamera wizyjna

Rysunek 5.5: Wybrane obrazy z sekwencji nr 7, 8, 9.



(a) sekwencja 12 kamera termowizyjna



(b) sekwencja 12 kamera wizyjna

Rysunek 5.6: Wybrane obrazy z sekwencji nr 10, 11, 12.

w bliskiej odległości są tylko drzewa. Sekwencja 17 przedstawia ujęcie innej korony drzewa oraz bliską ścianę budynku na tle nieba.

Sekwencje testowe KAIST

Dostępna publicznie baza danych KAIST [175] zawiera sekwencje obrazów nagrane przy użyciu układu kamer: wizyjnej i termowizyjnej umieszczonego na samochodzie osobowym. Baza powstała w celu ewaluacji algorytmów detekcji przechodniów w ruchu miejskim.

Założenia związane z układem kamer nie są w całości zgodne z założeniami przyjętymi w rozprawie. Kamery pozwalają na rejestrację obiektów w bliskiej odległości od układu. Dodatkowo samochód porusza się w trakcie rejestracji nagrań wideo. Pomimo to uznano, że przetestowanie opracowanej metody z wykorzystaniem tej bazy danych będzie stanowiło istotne odniesienie do obecnego stanu wiedzy z zakresu monitorowania otoczenia przy użyciu kamer wizyjnych i termowizyjnych.

Baza danych KAIST składa się z 12 zestawów zawierających od jednej do dziewięciu sekwencji par obrazów IR-TV. Na tej podstawie przygotowano sekwencje o numerach s18-s58. Przykładowe obrazy z tych sekwencji przedstawiono na rysunkach 5.7.



(a) Zestaw set00, sekwencja v000, kamera termowizyjna.



(b) Zestaw set00, sekwencja v000, kamera wizyjna.



(c) Zestaw set00, sekwencja v007, kamera termowizyjna.



(d) Zestaw set00, sekwencja v007, kamera wizyjna.



(e) Zestaw set02, sekwencja v001, kamera termowizyjna.



(f) Zestaw set02, sekwencja v001, kamera wizyjna.

Rysunek 5.7: Wybrane obrazy z sekwencji testowej KAIST.

5.2 Wyniki testów samokalibracji układu kamer

Implementacja opracowanej metody samokalibracji układu kamer wizyjnej i termowizyjnej została przygotowana przez autora rozprawy jako moduł programistyczny większej aplikacji UFO-B realizującej zadania układu fuzji obrazów (UFO) wizyjnych (TV) i termowizyjnych (IR) w zakresie procesu kalibracji układu kamer wizyjnej i termowizyjnej. Oprogramowanie UFO-B zostało zrealizowane przez zespół badawczy, którego członkiem był autor rozprawy, jako kontynuacja prac nad grantem rozwojowym NCBR pod tytułem „System fuzji obrazów wizyjnych i termowizyjnych dla potrzeb monitorowania otoczenia w czasie rzeczywistym” wykonywanych na Wydziale Mechatroniki [69, 176].

Oprogramowanie UFO-B jest przeznaczone do wykonywania kalibracji, dopasowania i fuzji obrazów z układu kamer IR i TV celem uzyskania jak najlepszego efektu fuzji i umożliwia w szczególności przeprowadzenie samokalibracji zgodnie z opisaną w rozdziale 4 procedurą.

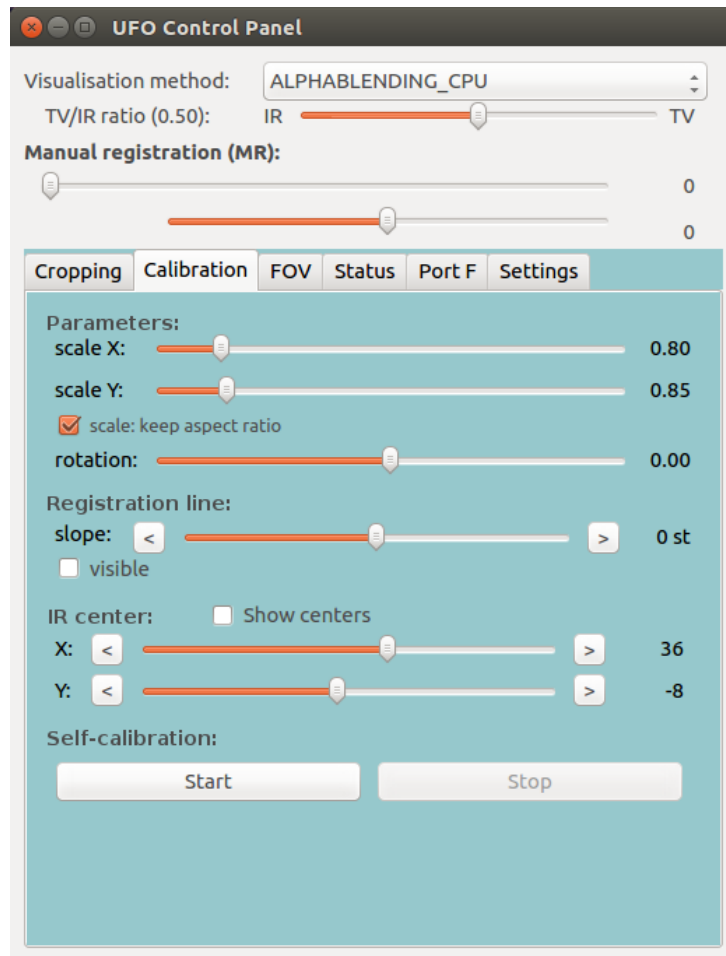
W skład oprogramowania wchodzi między innymi aplikacje *video-input*, *video-output*, *ufo-com* oraz *ufo-panel*. Aplikacje uruchomione jednocześnie komunikują się ze sobą. Program *ufo-com* realizuje główne funkcje przetwarzania obrazów i stanowi rdzeń oprogramowania UFO-B. Program *ufo-panel* (rys. 5.8) stanowi graficzny interfejs użytkownika pozwalający na sterowanie pracą *ufo-com*. Program *video-input* pozwala na wczytanie zapisanych na dysku komputera nagrań z kamer wizyjnej i termowizyjnej oraz przesłanie ich do *ufo-com*. Program *video-output* wykorzystywany jest w celu wizualizacji fuzji obrazów. Wybraną procedurę samokalibracji można przeprowadzić po uruchomieniu wszystkich opisanych programów.

Proces samokalibracji obrazów wizyjnego i termowizyjnego musi doprowadzić do ustalenia parametrów f_{ir} , f_{tv} oraz parametrów linii dopasowania opisanych dokładniej w rozdziale 4.3. Parametry f_{irx} oraz f_{iry} reprezentują rozmiar obrazu termowizyjnego względem wizyjnego. Oznaczone zostały one w uproszczeniu jako *scale X* oraz *scale Y* i umieszczone w odpowiedniej sekcji aplikacji *ufo-panel*. Ustalenie nachylenia linii dopasowania (parametr *slope* w *ufo-panel*) odpowiada ustaleniu proporcji wybranych parametrów zewnętrznych t_x , t_y układu kamer, a ustalenie odległości tej linii od środka obrazu wizyjnego odpowiada ustaleniu parametrów wewnętrznych c_{irx} , c_{iry} kamery termowizyjnej (*IR center* w *ufo-panel*).

Dodatkowo w aplikacji *ufo-panel* można ustawić obszary zainteresowania opisane jako *Crop* oraz parametr obrotu obrazu IR względem obrazu TV - opisany jako *rotation*. W ten sposób można zrealizować ręcznie etap wstępny samokalibracji opisany w rozdziale 4.

5.2.1 Sposób przeprowadzenia testów samokalibracji

Przeprowadzenie samokalibracji możliwe jest po uruchomieniu aplikacji *ufo-panel* komunikującej się z *ufo-com*. W celu przeprowadzenia testów przygotowano dwie wersje oprogramowania *ufo-com*. Jedna wersja realizuje samokalibrację na etapie pierwszym, a druga wersja realizuje samokalibrację na etapie drugim. Za pomocą programu *video-input* należy



Rysunek 5.8: Interfejs aplikacji *ufo-panel* dla zakładki *Calibration*

wybrać nagrania wideo, za pomocą których chcemy przeprowadzić kalibrację. Za pomocą programu *video-output* można obserwować wizualizację procesu samokalibracji. Zadaniu samokalibracji układu kamer przypisana jest oddzielna zakładka w interfejsie *ufo-panel* - *Calibration* (rys. 5.8).

Poniżej przedstawiono ogólną procedurę samokalibracji układu kamer.

1. Uruchom programy *ufo-com*, *ufo-panel*, *video-output* oraz *video-input*.
2. Przy użyciu programu *video-input* otwórz nagranie wideo, za pomocą którego zostanie przeprowadzona samokalibracja. Uruchom odtwarzanie nagrania.
3. W zakładce *Cropping* programu *ufo-panel* wybierz obszary zainteresowania na obrazach z kamer IR i TV (rys. 5.9). Zakładka umożliwia realizację etapu wstępnego kalibracji obrazów, czyli wykluczenie obszaru obrazów nie zawierających pikseli reprezentujących obiekty obserwowanego otoczenia.
4. W zakładce *Calibration* programu *ufo-panel* w sekcji *Parameters* ustaw manualnie parametry, których estymacja nie jest rezultatem działania samokalibracji. W przypadku samokalibracji na etapie pierwszym konieczne jest ustawienie jedynie parametru obrotu

rotation obrazu termowizyjnego względem wizyjnego. W przypadku samokalibracji na etapie drugim konieczne jest ustawienie parametrów obrotu oraz rozmiaru *scale X*, *scale Y* obrazu termowizyjnego względem wizyjnego. Opcjonalnie w sekcji *IR Center* można ustawić parametry X i Y. W przypadku sekwencji obrazów testowych wykorzystanych do weryfikacji opracowanej metody wszystkie te parametry były znane. W trakcie przygotowywania sekwencji testowych zostały one ustalone manualnie sposób ręczny jako parametry wzorcowe.

5. Uruchom opracowany algorytm samokalibracji za pomocą przycisku *Start* w sekcji *Self-calibration* programu *ufo-panel*. Możliwe są dwa scenariusze zakończenia procesu automatycznej kalibracji:

- (a) Samokalibracja zakończyła się samoczynnie - odpowiednie parametry zostały wyznaczone automatycznie. W przypadku etapu pierwszego samokalibracji wyznaczone zostaną parametry skali *scale X*, *scale Y* obrazu termowizyjnego względem wizyjnego. W przypadku etapu drugiego wyznaczone zostaną parametry nachylenia linii dopasowania *slope* oraz parametr *shift* odległości tej linii od środka układu współrzędnych. Z poziomu *ufo-panel* parametr można ustawić za pomocą drugiego od góry suwaka w sekcji *Manual registration (MR)*.
- (b) Samokalibracja nie zakończyła się samoczynnie - nagranie wideo nie pozwala na automatyczne wyznaczenie dopasowania w trybie kalibracji. Nagranie może zawierać zbyt mało elementów ruchomych lub przedstawiać obiekty, które trudno dopasować do siebie. Należy wtedy zakończyć ręcznie procedurę automatycznej kalibracji (sekcja *Self-calibration* przycisk *Stop* programu *ufo-panel*). W dalszym opisie i analizie wyników działania samokalibracji, takie przypadki zakończenia procesu samokalibracji zostały odpowiednio oznaczone.



(a) Biała ramka zaznacza prawidłowo obszar informatywny z kamery IR.



(b) Biała ramka zaznacza nieprawidłowo obszar informatywny z kamery IR.

Rysunek 5.9: Wizualizacja działania zakładki *Cropping*.

W kolejnym podrozdziale przedstawiono wyniki kalibracji przeprowadzonej zgodnie z przedstawioną procedurą dla sekwencji danych testowych opisanych w rozdziale 5.1.2.

5.2.2 Zasady opisu wyników działania samokalibracji

W bieżącym rozdziale przedstawiono wyniki działania opracowanych w ramach rozprawy algorytmów samokalibracji. Same metody zostały szczegółowo omówione w rozdziale 4 oraz podsumowane w podrozdziale 4.3. Opracowana metoda została podzielona na dwa etapy. Działanie każdego etapu zostało przedstawione osobno. W pierwszej kolejności przedstawiono wyniki działania etapu I kalibracji, czyli ustalenie parametrów skali obrazu termowizyjnego względem wizyjnego. Dalej znajdują się wyniki działania etapu II kalibracji, czyli ustalenie parametrów linii dopasowania opisującej zależność dopasowania online d_x, d_y odpowiadających sobie punktów obrazu wizyjnego i termowizyjnego. Wyniki w tym zakresie przedstawiono w odniesieniu do nachylenia linii dopasowania. Nachylenie jest parametrem, który najtrudniej dobrać w sposób ręczny i wyznaczenie go w sposób automatyczny przez samokalibrację stanowi istotny rezultat opracowanej metody. Testy działania każdego etapu samokalibracji rozpoczyna opis wyników dla sekwencji wygenerowanych komputerowo. Następnie znajduje się analiza wyników dla sekwencji testowych Poprad. Na koniec zamieszczono analizę wyników działania opracowanej metody samokalibracji dla sekwencji testowych KAIST i podsumowano wszystkie wyniki.

Analizę wyników działania metod samokalibracji wykonano w ujęciu dokładności i precyzji. Dokładność jest rozumiana jako zgodność wartości wyznaczonej za pomocą metody samokalibracji z wartością wzorcową. Jako miarę dokładności pomiaru (wyznaczenia wartości przez algorytm samokalibracji) przyjęto bezwzględną różnicę pomiędzy znanymi dla sekwencji testowych wartościami wzorcowymi, a wartościami wyznaczonymi przez testowane metody.

$$err_i = |w_i - \hat{w}_i| \quad (5.1)$$

gdzie:

err_i - błąd pomiaru i ,

$i = 1, \dots, n$ - numer pomiaru w serii,

n - ilość pomiarów w serii,

w_i - wartość wzorcowa i ,

\hat{w}_i - wartość zmierzona i .

Dokładność metody oceniana jest na podstawie średniego błędu działania metody dla serii pomiarów:

$$\overline{err} = \frac{\sum_{i=1}^n err_i}{n} \quad (5.2)$$

gdzie:

\overline{err} - średni błąd serii pomiarów.

Jako miarę precyzji metody przyjęto odchylenie standardowe s błędu działania metody dla serii pomiarów.

$$s = \sqrt{\frac{\sum_{i=1}^n (err_i - \overline{err})^2}{n - 1}} \quad (5.3)$$

gdzie:

s - odchylenie standardowe błędu.

Wyniki działania metod w dalszej części rozdziału przedstawiono na wykresach jako $\overline{err} \pm s$.

5.2.3 Wyniki I etapu - wprowadzenie

Działanie I etapu samokalibracji polega na wyznaczeniu skali obrazu termowizyjnego względem referencyjnego obrazu wizyjnego. Wyniki działania tej metody przedstawiono w ujęciu błędów bez określenia jednostek tych wartości. Przykładowo, wartość skali wynosząca 1.0 oznacza, że obraz termowizyjny ma taki sam rozmiar jak obraz wizyjny, a wartość 2.0 oznacza, że obraz termowizyjny jest dwukrotnie większy od obrazu wizyjnego. Błędy wyznaczenia skali należy interpretować w sposób analogiczny. Przykładowo wartość błędu na poziomie 0.1 oznacza, że metoda samokalibracji pomyliła się w estymacji o ok. 10% rozmiaru obrazu wizyjnego.

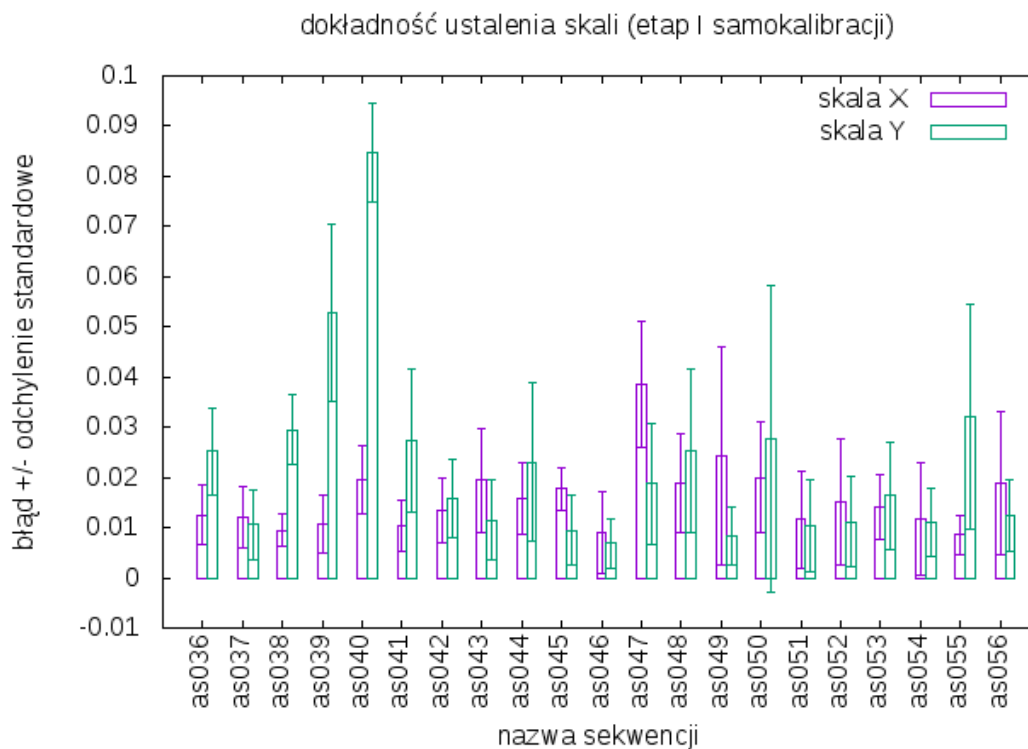
5.2.4 Wyniki I etapu dla sekwencji wygenerowanych komputerowo

Wykres 5.10 przedstawia błąd automatycznego wyznaczenia skali dla wygenerowanych komputerowo sekwencji testowych z wykorzystaniem algorytmu opisanego jako etap pierwszy samokalibracji.

Analiza wykresu pozwala zauważyć, że dla większości sekwencji wygenerowanych komputerowo błąd metody jest podobny i niższy niż 0.05. Przy estymacji skali obrazu termowizyjnego względem wizyjnego w kierunku osi Y zdarzyły się sytuacje wyróżniająco gorszej dokładności - wysokie błędy dla sekwencji as039 oraz as040. Ogólne działanie etapu I samokalibracji na obrazach wygenerowanych komputerowo można podsumować parametrami średniego błędu oraz odchylenia standardowego błędu wynoszącymi 0.016 ± 0.012 w kierunku osi X i 0.022 ± 0.022 w kierunku osi Y. Sekwencje wygenerowane komputerowo dla etapu I samokalibracji zostały przygotowane w celu potwierdzenia poprawności działania algorytmu na najprostszych przykładach i stanowią potwierdzenie ogólnie przyjętej koncepcji związanej z samokalibracją parametrów skali.

5.2.5 Wyniki I etapu dla sekwencji testowych Poprad

Wyniki testów na sekwencjach wygenerowanych komputerowo nie dostarczają jednak pełnego obrazu działania etapu I samokalibracji. Wykres 5.11 przedstawia błędy automatycznego wyznaczenia skali dla sekwencji testowych Poprad.



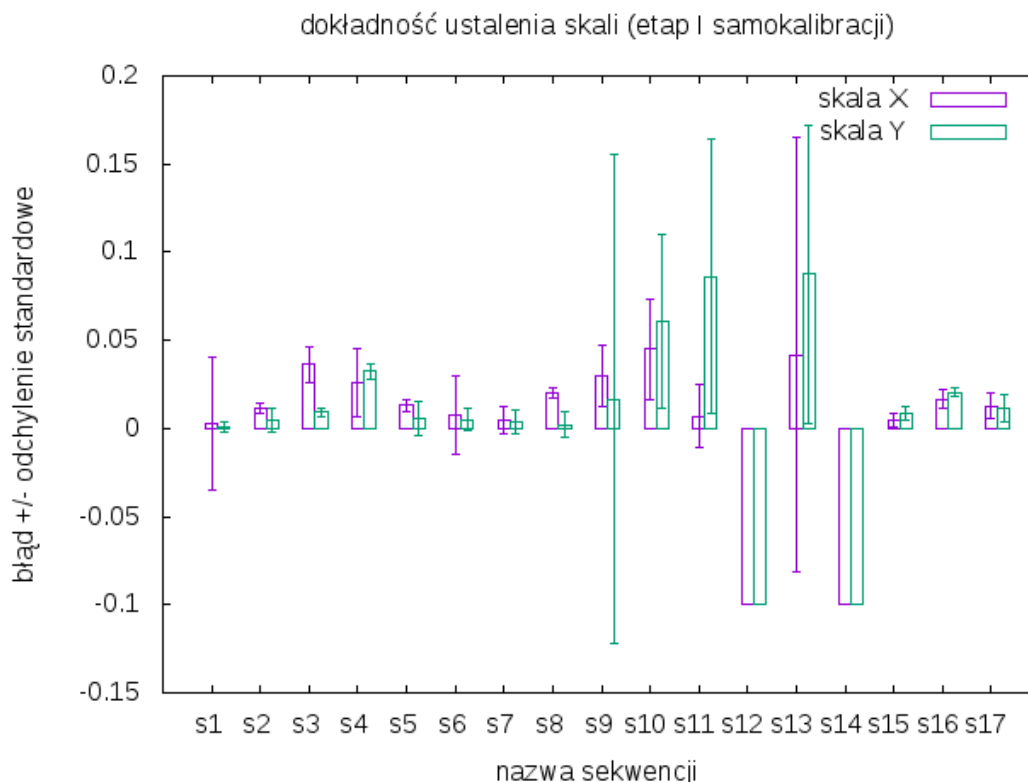
Rysunek 5.10: Błędy wyznaczenia skali dla sekwencji testowych wygenerowanych komputerowo.

Przykładowy wynik etapu I samokalibracji, czyli ustalenie rozmiaru obrazu termowizyjnego względem obrazu wizyjnego przedstawia rys. 5.12.

Na wykresie 5.11 warto zwrócić uwagę na sekwencje s12 i s14 - za pomocą ujemnych wartości błędu oznaczono, że dla tych sekwencji algorytm etapu I samokalibracji nie wyznaczył żadnych wartości skali obrazu termowizyjnego względem wizyjnego. Algorytm wyznacza parametr skali na podstawie dopasowania obrazu z chwili bieżącej do obrazu z chwili poprzedniej. Reprezentuje to ruch układu kamer względem otoczenia w trakcie nagrywania sekwencji. Na sekwencjach s12 i s14 występuje jedynie nieznaczny ruch układu kamer względem otoczenia, co spowodowało nieskuteczność działania algorytmu.

Jeśli wyliminujemy te sekwencje z dalszej analizy, można ocenić, że samokalibracja na etapie I cechuje się dobrą dokładnością. Działanie metody dla tak wybranych sekwencji testowych Poprad można podsumować parametrami średniego błędu oraz odchylenia standardowego błędu wynoszącymi 0.025 ± 0.034 w kierunku osi X i 0.031 ± 0.051 w kierunku osi Y.

W tym miejscu warto podkreślić, że aby zaproponowana metoda stanowiła użyteczną pomoc w trakcie procesu kalibracji układu kamer, nie musi ona działać dobrze w każdych możliwych warunkach. Sekwencje testowe Poprad nie zostały z kolei przygotowane specjalnie w celu uzyskania poprawnych wyników samokalibracji, ale w celu przedstawienia działania metody w różnych sytuacjach. Sekwencje reprezentują przykładowe obrazy pochodzące z



Rysunek 5.11: Błędy wyznaczenia skali dla sekwencji testowych Poprad; skali nie udało się wyznaczyć dla sekwencji s12 i s14 - oznaczono to ujemnymi wartościami dokładności.



(a) Fuzja przed kalibracją skali.



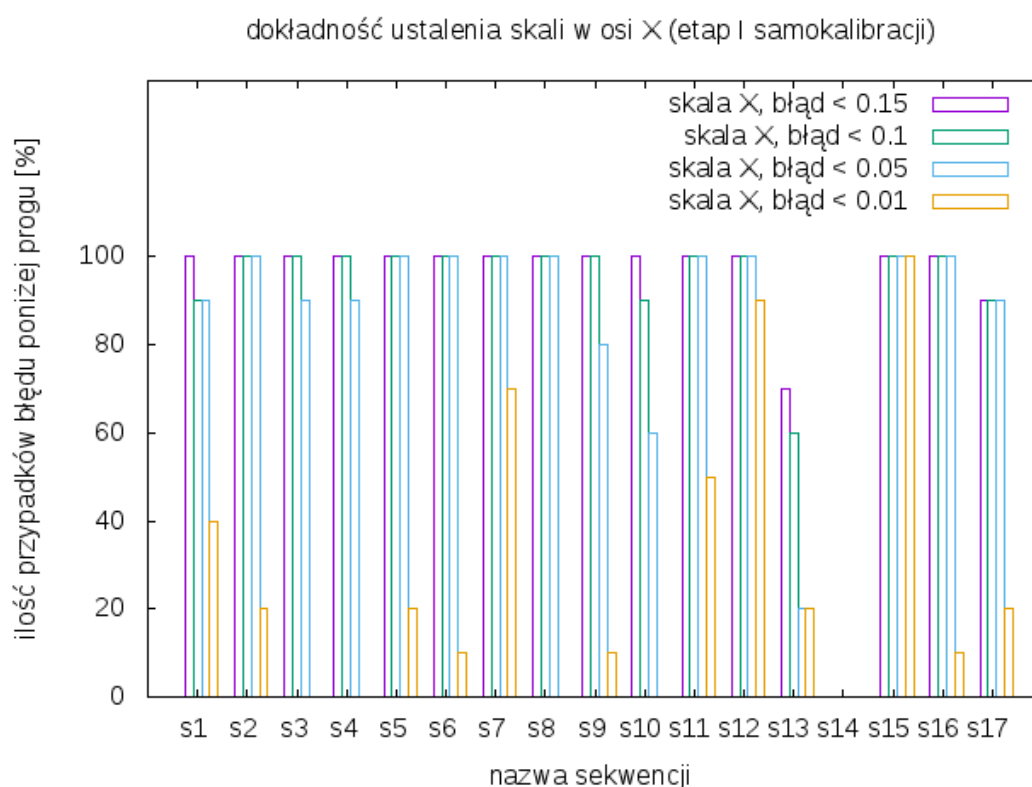
(b) Fuzja po kalibracji skali.

Rysunek 5.12: Przykładowe wyniki samokalibracji skali (etap I samokalibracji) dla sekwencji testowych Poprad.

układu kamer Poprad w trakcie monitorowania otoczenia. Na podstawie analizy wykresu 5.11 można wskazać, w jakich sytuacjach algorytmy etapu I samokalibracji działają najlepiej. Etap I samokalibracji z wykluczeniem sekwencji s12, s13 i s14 (dla których metoda w osi X nie działa, lub działa w niską dokładnością) cechuje się średnim błędem oraz odchyleniem standardowym błędu w kierunku osi X na poziomie 0.018 ± 0.015 , a w kierunku osi Y

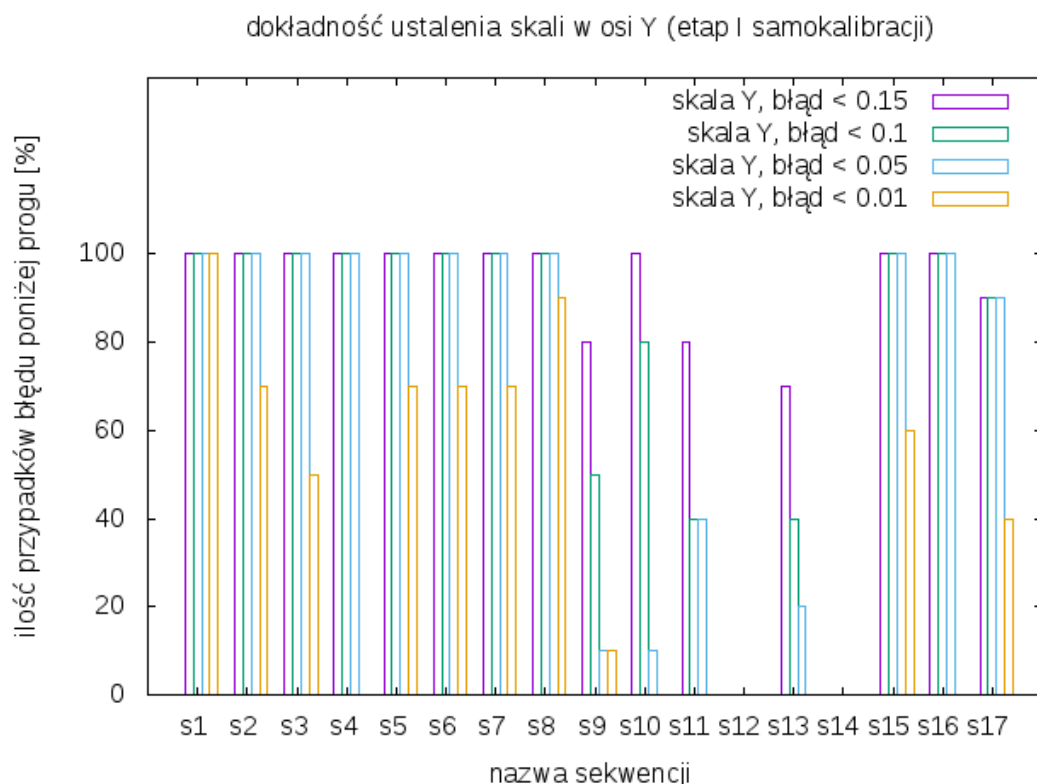
na poziomie 0.025 ± 0.046 . Jeśli z analizy wykluczmy sekwencje s9, s10, s11, s12, s13 oraz s14 (dla których metoda w osi Y nie działa, lub działa z niską dokładnością) etap I samokalibracji cechuje się średnim błędem oraz odchyleniem standardowym błędu w kierunku osi X na poziomie 0.018 ± 0.015 , a w kierunku osi Y na poziomie 0.011 ± 0.010 . Na podstawie wykonanej analizy można przedstawić rekomendację sposobu stosowania opracowanej metody. Przedstawiono ją w na końcu bieżącego rozdziału w podsumowaniu wyników.

Analiza dokładności działania algorytmu samokalibracji na etapie I pozwala jednak na sformułowanie ograniczonych wniosków. Lepszy wgląd w jakość działania samokalibracji przedstawia analiza działania algorytmu dla każdej sekwencji z wykorzystaniem klasyfikacji błędu poniżej określonego progu. W tym celu przyjęto arbitralnie cztery graniczne wartości błędu i na tej podstawie przeanalizowano działanie etapu I samokalibracji. Na wykresie 5.13 przedstawiono procentową liczbę przypadków (przy dziesięciu próbach) ustalenia skali w kierunku osi X z błędem poniżej określonego progu, dla każdej sekwencji testowej Poprad. Na wykresie 5.14 przedstawiono analogiczny wykres dla estymacji skali w kierunku osi Y.



Rysunek 5.13: Liczba przypadków ustalenia skali w osi X obrazu termowizyjnego względem wizyjnego z błędem poniżej określonego progu dla sekwencji testowych Poprad.

Analiza tych wykresów pozwala stwierdzić, że w odniesieniu do sekwencji s1-s8 oraz s15, s16 i s17 w większości przypadków ustalone w wyniku samokalibracji parametry cechują się błędami poniżej 0.05. Sytuacje, w których dokładność jest gorsza, czyli błędy mają wartość powyżej progu 0.05, zarejestrowano szczególnie w kierunku osi Y w trakcie testów sekwencji

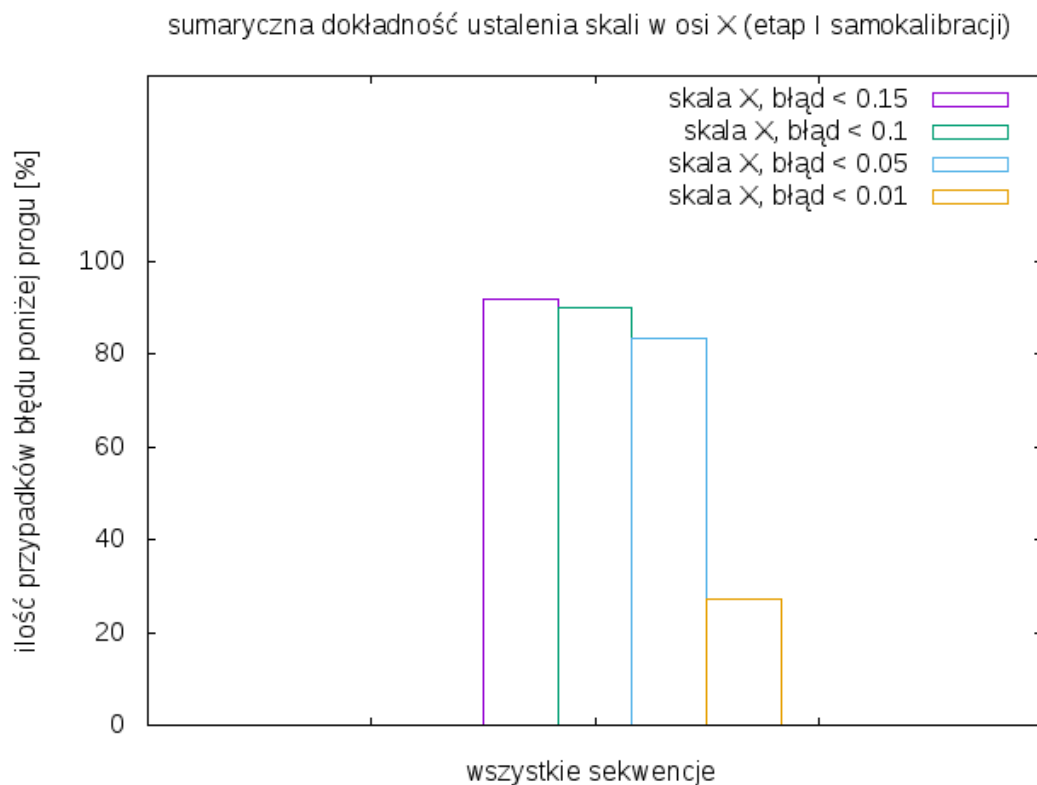


Rysunek 5.14: Liczba przypadków ustalenia skali w osi Y obrazu termowizyjnego względem wizyjnego z błędem poniżej określonego progu dla sekwencji testowych Poprad.

s9-s14. Porównując wykresy 5.13 oraz 5.14 dla tych samych progów błędów można zauważyć, że znaczące różnice w działaniu samokalibracji w kierunku osi X i Y występują dla sekwencji s9, s10, s11 i s12. Estymacja skali w kierunku osi X z błędami poniżej 0.05 następuje tam w większej liczbie przypadków niż dla osi Y. Różnicę można zaobserwować w szczególności dla sekwencji s11 i s12, gdzie liczba przypadków estymacji skali w osi X z błędami poniżej 0.05 (a nawet poniżej 0.01) jest znacząco większa niż analogiczna liczba przypadków dla osi Y. Zaproponowany algorytm estymacji skali obrazu termowizyjnego względem wizyjnego działa na podstawie pomiarów ruchu układu kamer względem obserwowanego otoczenia w trakcie nagrywania sceny. Różnice w działaniu oraz gorsze wyniki estymacji skali w osi Y dla sekwencji s9, s10, s11 i s12 można tłumaczyć tym, że dla wspomnianych sekwencji nie zarejestrowano wystarczająco zróżnicowanych ruchów układu kamer właśnie w osi Y. Dla sekwencji S13 zarówno w osi X, jak i w osi Y układ kamer porusza się nieznacznie, a dla sekwencji s14 układ kamer jest nieruchomy.

Podsumowanie danych z wykresu 5.13 przedstawiono sumarycznie dla wszystkich sekwencji testowych Poprad na wykresie 5.15. W sposób analogiczny dane z wykresu 5.14 podsumowano na wykresie 5.16.

Testy etapu I samokalibracji dla wszystkich sekwencji Poprad pokazują, że procentowa liczba przypadków uzyskania błędów poniżej określonego progu znacząco spada dla wartości

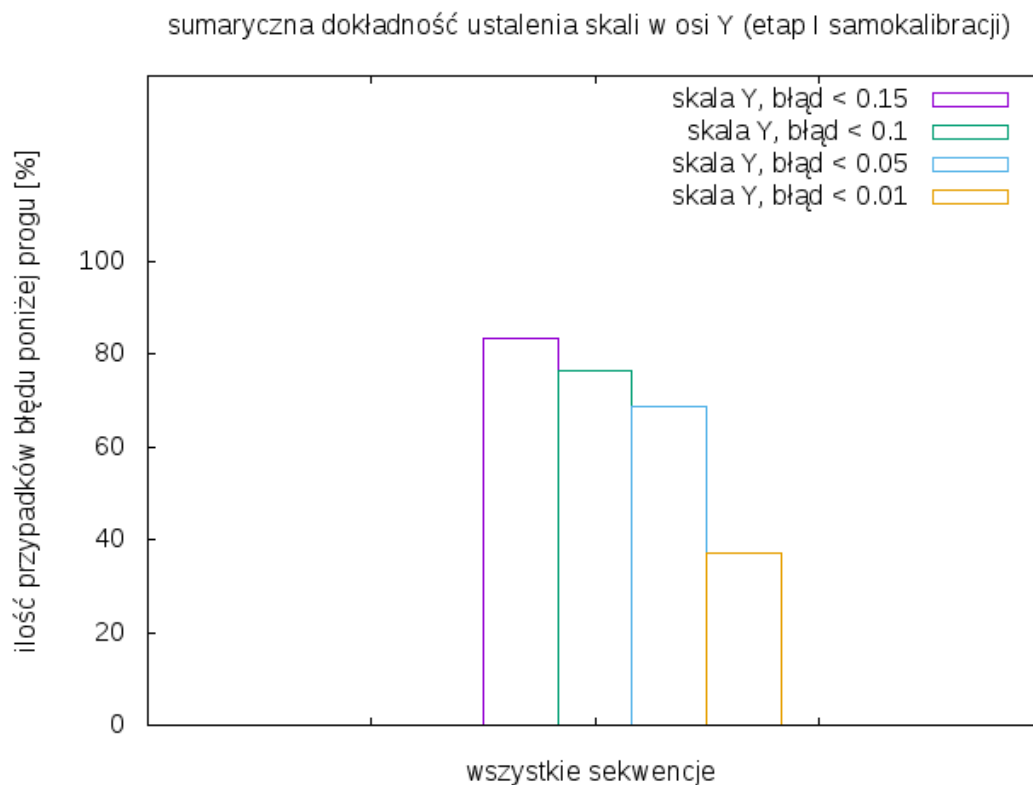


Rysunek 5.15: Sumaryczna liczba przypadków ustalenia skali w osi X z błędem poniżej określonego progu dla sekwencji testowych Poprad.

progowej błędu 0.01. Dzieje się tak zarówno przy estymacji skali w kierunku osi X, jak i w kierunku osi Y.

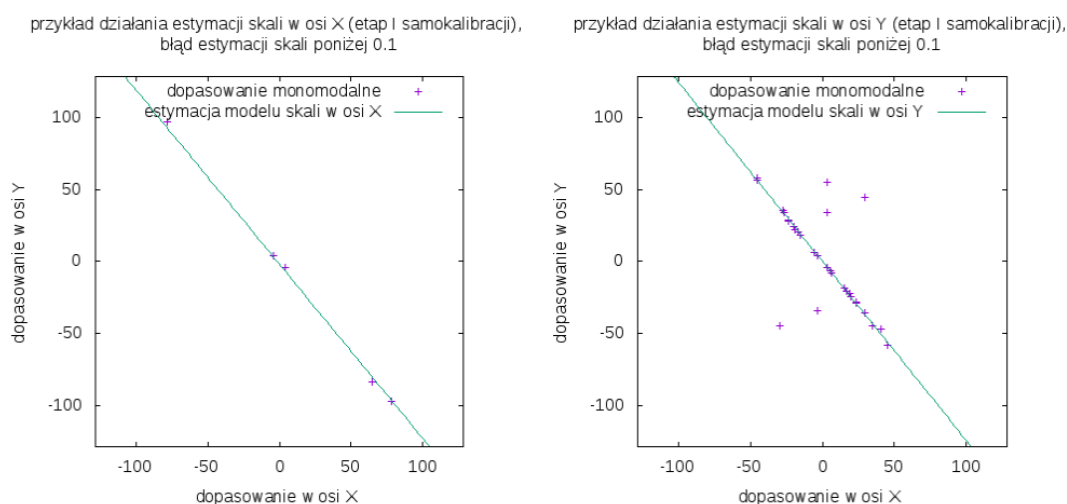
Powyższe analizy można podsumować wnioskiem, że w przypadku analizy wszystkich sekwencji estymacja w osi X cechuje się lepszą dokładnością i precyzją niż w kierunku osi Y. Można to tłumaczyć na różne sposoby. Po pierwsze dla zarejestrowanych danych testowych Poprad ruch obrazów w osi Y jest mniej zróżnicowany niż w przypadku osi X. Po drugie, wszystkie obrazy pochodzą oryginalnie z kamer analogowych. Oznacza to, że pełen obraz z dowolnej kamery składany jest z dwóch półobrazów. Może to prowadzić do błędów w cyfrowej analizie obrazu z powodu uwidoczniania się efektu przeplotu na pełnych obrazach. Taki efekt jest bardziej szkodliwy dla analizy obrazu w osi Y niż w osi X.

Przykład estymacji skali metodą RANSAC w sytuacji uzyskania błędów poniżej 0.1 przedstawiony został na rysunkach 5.17. Zarówno w przypadku estymacji skali w osi X, jak i w osi Y nastąpiło poprawne dopasowanie modelu (linii prostej przechodzącej przez środek układu współrzędnych) do pomiarów dopasowania monomodalnego w trybie kalibracji. Na wykresach warto zwrócić szczególną uwagę na zakres zmienności wartości pomiarowych (dopasowania monomodalne w trybie kalibracji). Można również zauważyć, że niektóre pomiary są ignorowane przez algorytm dopasowania modelu do danych RANSAC. Są to



Rysunek 5.16: Sumaryczna liczba przypadków ustalenia skali w osi Y z błędem poniżej określonego progu dla sekwencji testowych Poprad.

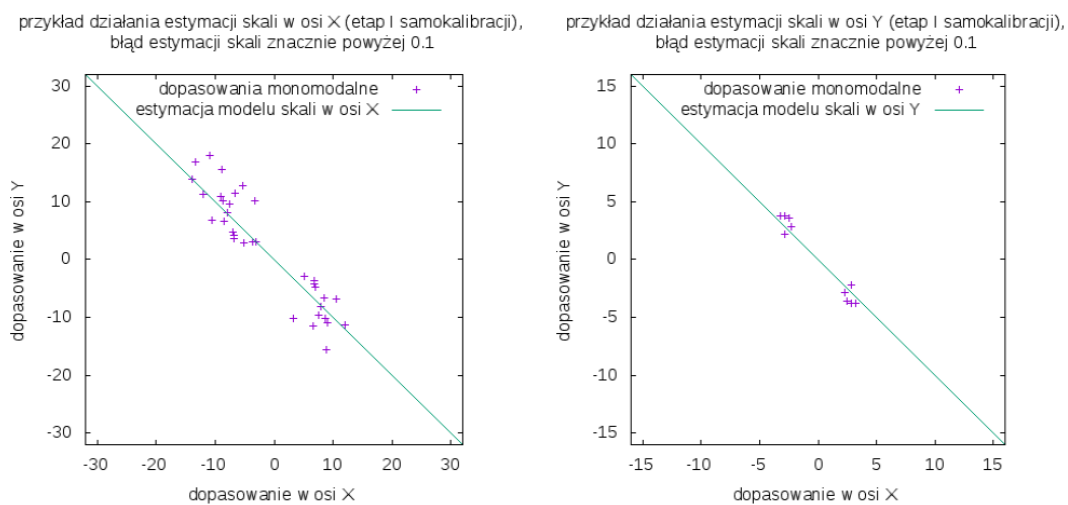
najprawdopodobniej błędy działania algorytmu dopasowania obrazów z chwili bieżącej do chwili poprzedniej, czyli pomiarów ruchu układu kamer względem otoczenia.



Rysunek 5.17: Przykłady estymacji skali z wykorzystaniem RANSAC dla sekwencji testowych Poprad.

Na rys. 5.18 widać z kolei przykład estymacji skali z błędem powyżej 0.1. Wyznaczenie skali z gorszą dokładnością nastąpiło w tym przypadku prawdopodobnie z uwagi na niską

precyzję pomiarów dopasowania monomodalnego. Na wykresie warto zwrócić uwagę na niski zakres zmienności danych pomiarowych w porównaniu do estymacji przedstawionej na wykresach 5.17. W pomiarach wydają się nie występować w tym przypadku błędy algorytmu wyznaczenia dopasowania monomodalnego. Zarówno w osi X, jak i w osi Y przyczyną gorszej dokładności estymacji skali może być zbyt niska precyzja pomiarów spowodowana niskim zakresem zmienności pomiarów. Jednostką pomiarów dopasowania monomodalnego estymującego ruch układu kamer są piksele. Obrazy cyfrowe są zdyskretyzowane w dziedzinie obrazu. Na stopień dyskretyzacji ma wpływ głównie rozdzielczość obrazów. Przy niskich wartościach przesunięcia (ruchu) można spodziewać się niskiej precyzji pomiarów. Ta niska precyzja może utrudniać estymację skali.



Rysunek 5.18: Przykłady estymacji skali z wykorzystaniem RANSAC dla sekwencji testowych Poprad.

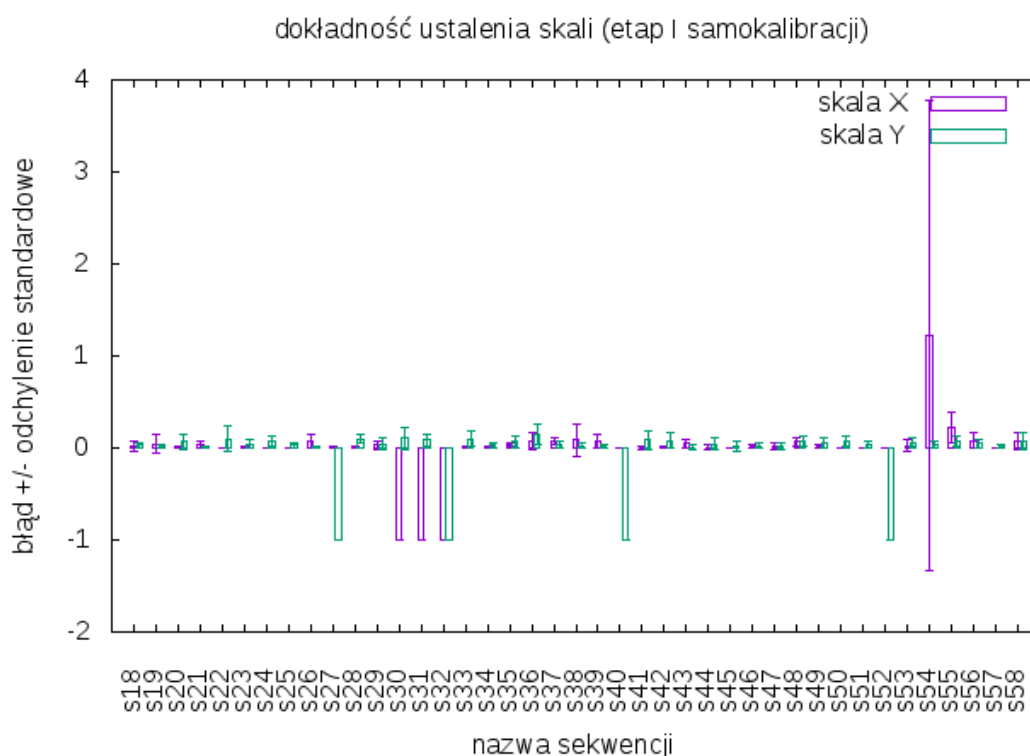
Podsumowując, przedstawione przykłady pozwalają zauważyć, że problemy z estymacją skali z wyższą dokładnością mogą występować w sytuacji niskich zakresów zmienności danych pomiarowych (nieznaczny ruch układu kamer względem otoczenia). Niewielka liczba błędnych pomiarów dopasowania monomodalnego nie pogarsza znacząco dokładności estymacji skali.

5.2.6 Wyniki I etapu dla sekwencji testowych KAIST

Opracowaną metodę samokalibracji na etapie I przetestowano dodatkowo na ogólnodostępnej bazie sekwencji testowych KAIST. Wykres 5.19 przedstawia błędy automatycznego wyznaczenia skali dla sekwencji testowych KAIST. Wartościami ujemnymi oznaczono tam sekwencje, dla których nie udało się przeprowadzić samokalibracji z powodu zbyt niskiej liczby pomiarów dopasowania. W przypadku estymacji skali w kierunku osi X są to sekwencje s30, s31 i s32, a w kierunku osi Y są to sekwencje s27, s32 i s52. Znaczne błędy w działaniu samokalibracji występują również w przypadku sekwencji s54 w kierunku osi X. W sekwencji s32 układ kamer nie porusza się względem otoczenia. W przypadku sekwencji s30 i s31

zarejestrowano przejazd samochodu głównie w trakcie jazdy na wprost - oznacza to, że ruchy układu kamer względem otoczenia są znacząco różne od przyjętych w rozprawie założeń. Problemy z działaniem samokalibracji w odniesieniu do sekwencji s54 można tłumaczyć niewielką liczbą szczegółów widocznych na nagraniu z kamery termowizyjnej.

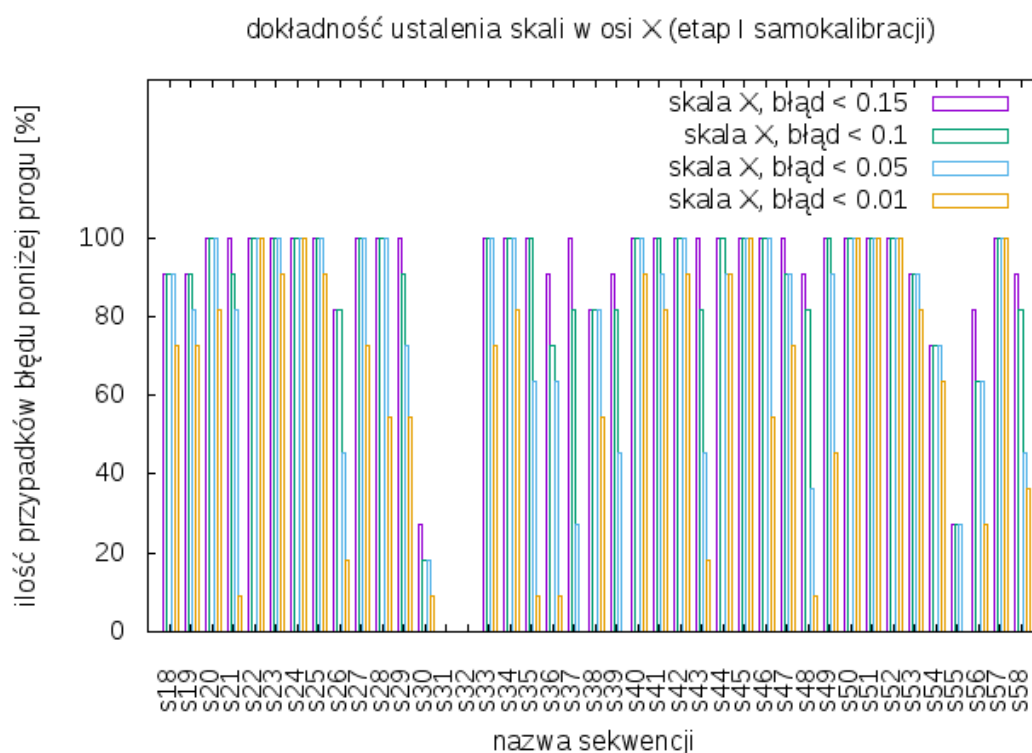
Jeśli wyeliminujemy te sekwencje z analizy, można ocenić, że samokalibracja na etapie I cechuje się dobrą dokładnością i jej działanie dla wybranych sekwencji KAIST można podsumować parametrami średniego błędu oraz standardowego odchylenia błędu wynoszącymi 0.036 ± 0.077 w kierunku osi X i 0.050 ± 0.068 w kierunku osi Y. Sekwencje KAIST nie zostały przygotowane specjalnie w celu samokalibracji, ale zawierają przykładowe nagrania z jazdy samochodem osobowym w warunkach miejskich. Wybierając określone sekwencje z bazy KAIST można przedstawić rekomendację sposobu stosowania etapu I samokalibracji. Przedstawiono ją w na końcu bieżącego rozdziału w podsumowaniu wyników.



Rysunek 5.19: Błędy wyznaczenia skali dla sekwencji testowych KAIST; skali w osi X nie udało się wyznaczyć dla sekwencji s30, s31 i s32, skali w osi Y nie udało się wyznaczyć dla sekwencji s27, s32 i s52 - oznaczono to ujemnymi wartościami błędu.

W celu wykonania bardziej szczegółowej analizy dokładności etapu I samokalibracji dla sekwencji testowych KAIST przedstawiono działanie tego algorytmu dla każdej sekwencji z klasyfikacją błędów poniżej określonego progu. Progi błędów ustalono arbitralnie w sposób analogiczny, jak przy sekwencjach testowych Poprad. Na wykresie 5.20 przedstawiono procentową liczbę przypadków (na dziesięć prób) ustalenia skali w kierunku osi X z błędem

poniżej określonego progu dla każdej sekwencji testowej z zestawu KAIST. Na wykresie 5.21 przedstawiono analogiczny wykres dla estymacji skali w kierunku osi Y.



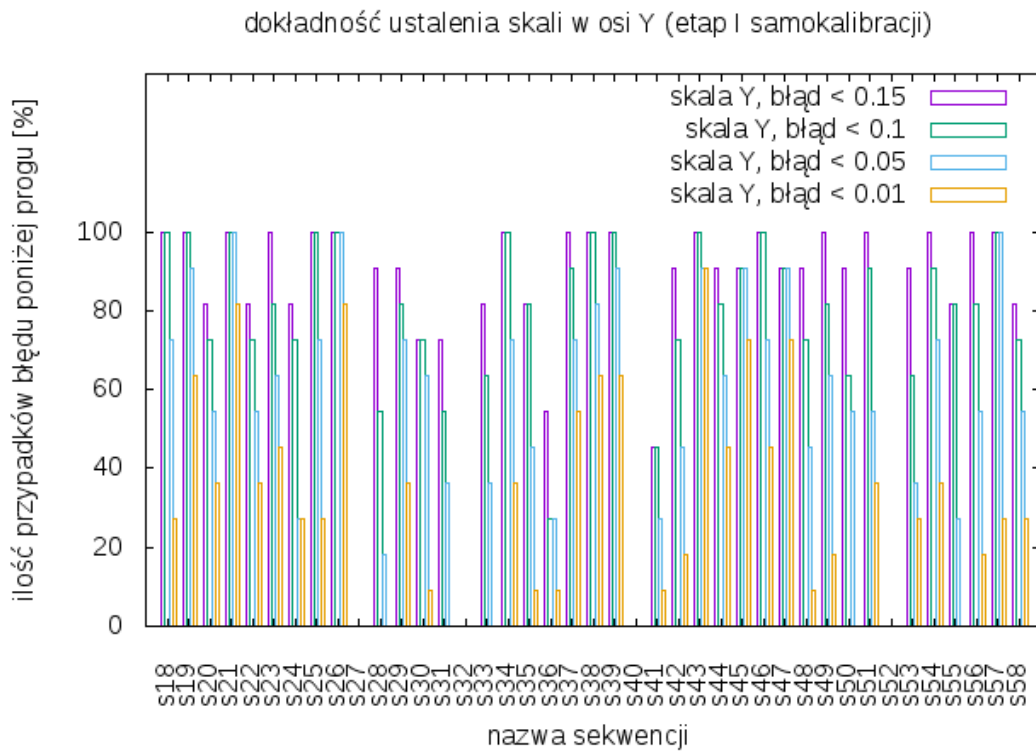
Rysunek 5.20: Liczba przypadków ustalenia skali w osi X obrazu termowizyjnego względem wizyjnego z błędem poniżej określonego progu dla sekwencji testowych KAIST.

Podsumowanie danych z wykresu 5.20 przedstawiono sumarycznie dla wszystkich sekwencji na wykresie 5.22, natomiast podsumowanie danych z wykresu 5.21 przedstawiono sumarycznie dla wszystkich sekwencji na wykresie 5.23.

Porównując dane z wykresów działania samokalibracji w kierunku osi X i Y można zauważyć, że podobnie jak w przypadku sekwencji testowych Poprad, dokładność działania w kierunku osi X jest lepsza niż w kierunku osi Y.

5.2.7 Wyniki II etapu - wprowadzenie

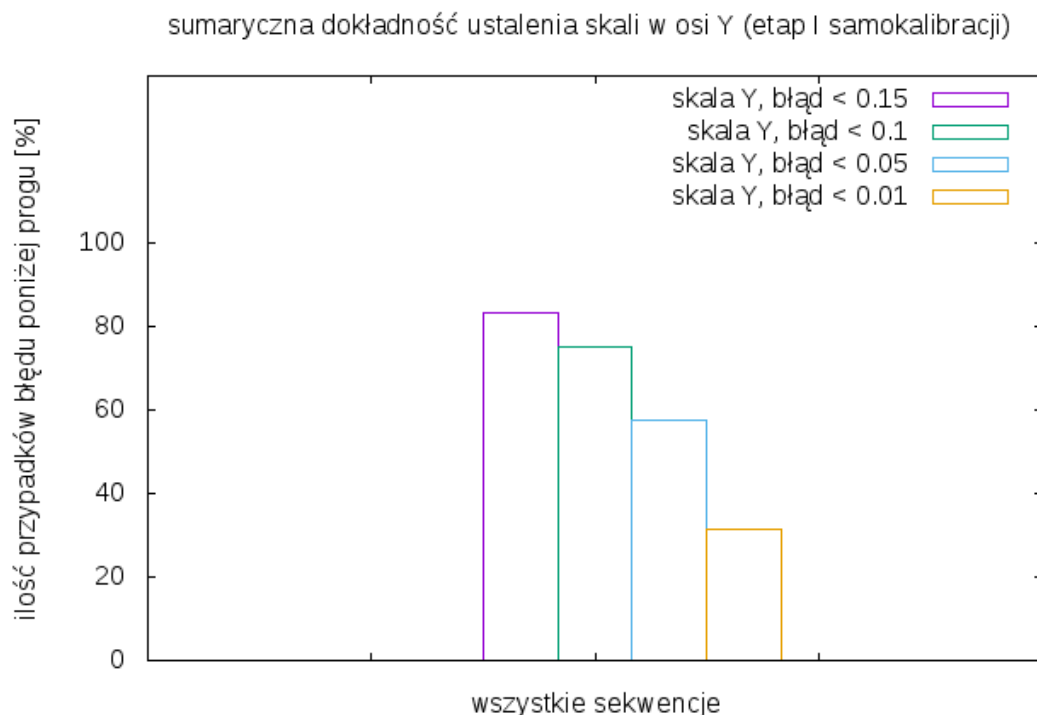
W celu lepszej demonstracji opracowanego rozwiązania problemu kalibracji przedstawiono działanie czterech wersji algorytmu realizującego etap II samokalibracji. Dzięki przedstawieniu różnych wersji algorytmu możliwe jest podkreślenie pozytywnego wpływu na samokalibrację opracowanej w ramach rozprawy metody automatycznego dopasowania wielu obiektów widocznych jednocześnie na obrazach multimodalnych. W celu zwiększenia czytelności wykonanej analizy wyników, na wykresach przedstawiono jedynie nachylenie linii dopasowania. Ręczne ustalenie nachylenia linii dopasowania jest procedurą stwarzającą



Rysunek 5.21: Liczba przypadków ustalenia skali w osi Y obrazu termowizyjnego względem wizyjnego z błędem poniżej określonego progu dla sekwencji testowych KAIST.



Rysunek 5.22: Sumaryczna liczba przypadków ustalenia skali w osi X z błędem poniżej określonego progu dla sekwencji testowych KAIST.



Rysunek 5.23: Sumaryczna liczba przypadków ustalenia skali w osi Y z błędem poniżej określonego progu dla sekwencji testowych KAIST.

najwięcej problemów w trakcie kalibracji układu kamer zestawu Poprad, w związku z czym automatyzacja tego procesu niesie ze sobą największe korzyści praktyczne.

Wersję pierwszą algorytmu etapu II samokalibracji opisano akronimem OPCT (ang. *One Peak Constant Threshold*). Przedstawia ona klasyczne wykorzystanie metody korelacji fazowej - wyznaczenie lokalizacji jednego największego maksimum. W celu odróżnienia pomiarów cechujących się lepszą jakością wykorzystano bardzo prosty mechanizm bazujący na wysokości znalezionego maksimum. Do estymacji parametrów równania (4.70) wykorzystano tylko lokalizację tych maksimum korelacji fazowej, których wysokość przewyższa pewien stały, ustalony eksperymentalnie próg¹. OPCT stanowi punkt odniesienia działania pozostałych testowanych metod.

Wersję drugą algorytmu etapu II samokalibracji opisano akronimem OPMT (ang. *One Peak Mean Threshold*). W tej wersji, podobnie jak w OPCT, następuje wyznaczenie lokalizacji jednego największego maksimum macierzy korelacji fazowej. W celu odróżnienia pomiarów cechujących się lepszą jakością wykorzystano jednak w inny sposób wysokości znalezionego maksimum. Do estymacji parametrów równania (4.70) wykorzystano tylko lokalizację tych maksimum korelacji fazowej, których wartość przewyższa liczoną na bieżąco średnią wysokość maksimum w sekwencji.

Wersję trzecią algorytmu etapu II samokalibracji opisano akronimem MPQD (ang. *Multiple Peaks Quality Drop*). W tej wersji następuje wyznaczenie lokalizacji kilku największych

¹w badaniach próg przyjęto w sposób doświadczalny

maksimów macierzy korelacji fazowej wyznaczanych zgodnie z zależnością $k \arg \max$. Wszystkie wyznaczone w ten sposób maksima wykorzystywane są do estymacji parametrów równania (4.70).

Ostatecznie, wersję czwartą algorytmu etapu II samokalibracji opisano akronimem HP (ang. *Hybrid Peaks*). Do estymacji parametrów równania (4.70) wykorzystano w niej połączone mechanizmy zastosowane w wersji OPMT oraz MPQD. Sposób integracji tych sposobów przedstawia Algorytm 4.5.

Do opisu różnych wersji algorytmu samokalibracji na etapie II zastosowano akronimy odpowiadające angielskim nazwom funkcji programistycznych implementujących te algorytmy. W kodzie źródłowym implementacji przyjęto stosowanie języka angielskiego zgodnie z ogólnym standardem.

5.2.8 Wyniki II etapu dla sekwencji wygenerowanych komputerowo

Wykres 5.24 przedstawia błędy automatycznego wyznaczenia nachylenia linii dopasowania dla wygenerowanych komputerowo sekwencji testowych z wykorzystaniem algorytmu opisanego jako etap drugi kalibracji. Analizując wykres można zauważyć, że dla większości wygenerowanych komputerowo sekwencji (pomijając sekwencje np. as005, czy as025) błędy wyznaczenia nachylenia linii dopasowania są podobne i niższe niż 10° . Ogólne działanie etapu I samokalibracji na obrazach wygenerowanych komputerowo można podsumować parametrami średniego błędu i odchylenia standardowego błędu nachylenia linii dopasowania wynoszącymi $2.8 \pm 2.7[^\circ]$ dla OPCT, $5.1 \pm 10.5[^\circ]$ dla OPMT, $2.6 \pm 1.7[^\circ]$ dla MPQD, $2.9 \pm 5.7[^\circ]$ dla HP.

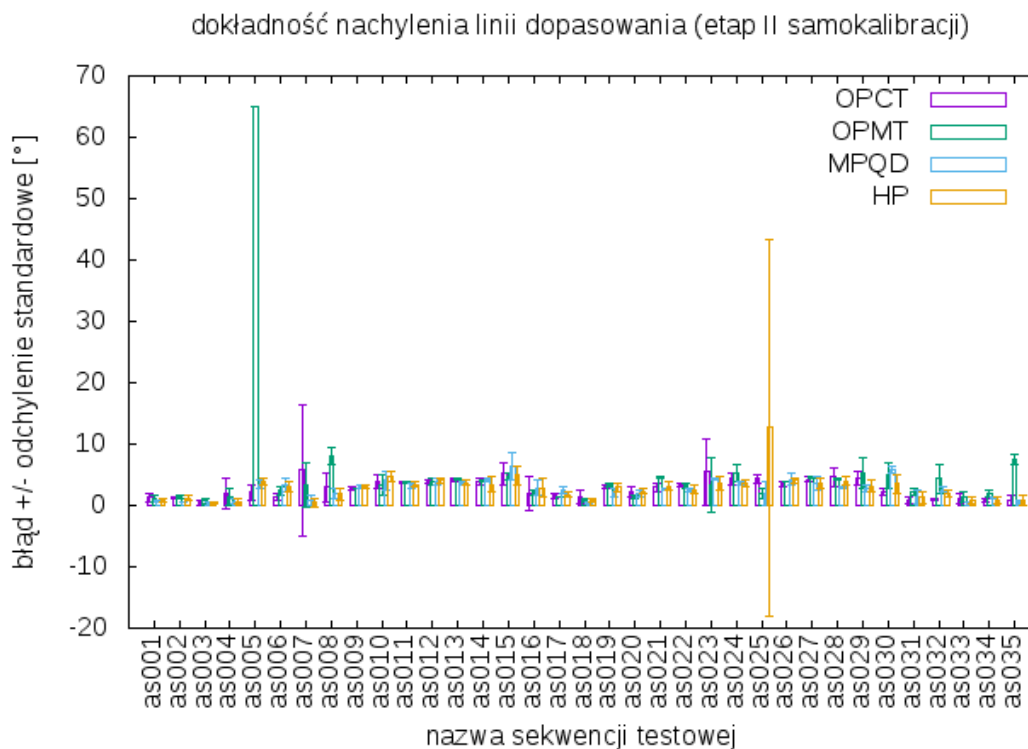
Podobnie jak w odniesieniu do etapu I samokalibracji, sekwencje wygenerowane komputerowo dla etapu II przygotowano w celu ogólnego sprawdzenia koncepcji działania metody w najprostszych sytuacjach. Dobre wyniki działania etapu II samokalibracji dla sekwencji testowych wygenerowanych komputerowo stanowią potwierdzenie koncepcji działania metody.

5.2.9 Wyniki II etapu dla sekwencji testowych Poprad

Wyniki testów na sekwencjach wygenerowanych komputerowo nie dostarczają jednak pełnego obrazu działania etapu II samokalibracji. Wykres 5.25 przedstawia porównanie błędów działania różnych wersji algorytmu etapu II kalibracji automatycznej estymacji kąta nachylenia linii dopasowania. Wersję OPCT potraktowano jako referencyjną, natomiast HP jako ostatecznie proponowane, najlepsze rozwiązanie.

Przykładowy wynik etapu II samokalibracji, czyli ustalenie parametrów linii dopasowania przedstawiono na rys. 5.26.

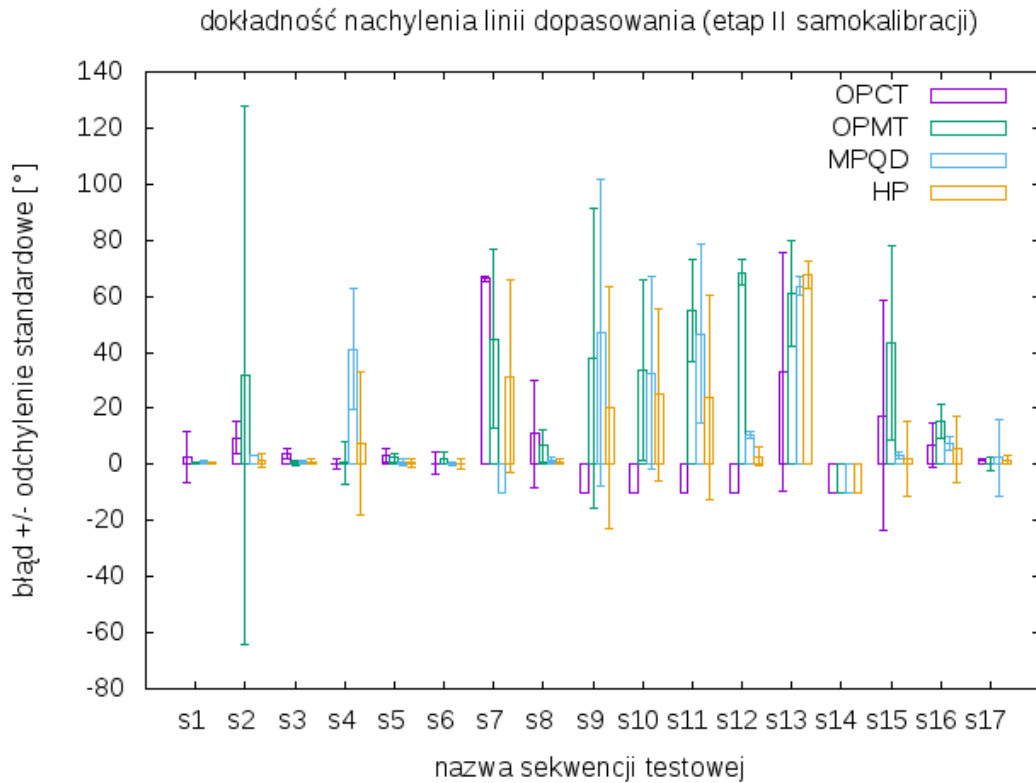
Na podstawie wykresu 5.25 można przeprowadzić ogólną analizę dokładności prezentowanych metod. Na podstawie danych przedstawionych na wykresie można zauważyć,



Rysunek 5.24: Błędy wyznaczenia nachylenia linii dopasowania dla sekwencji testowych wygenerowanych komputerowo.

że wersja referencyjna OPCT algorytmu kalibracji cechuje się średnimi błędami poniżej ok. 20° i odpowiednio niskimi odchyleniami standardowymi błędów dla ośmiu sekwencji testowych Poprad: s1-s6, s16 i s17. Wartość 20° określona została arbitralnie w celu ogólnego opisu dokładności. Druga wersja OPMT algorytmu cechuje się średnimi błędami poniżej 20° i odpowiednio niskimi odchyleniami standardowymi błędów również dla ośmiu sekwencji testowych: s1, s3-s6, s8, s16 i s17. Ta wersja działa lepiej w przypadku sekwencji s8, ale gorzej dla sekwencji s2. Trzecia wersja MPQD algorytmu cechuje się średnimi błędami poniżej 20° i odpowiednio niskimi odchyleniami standardowymi błędów dla dziesięciu sekwencji testowych: s1, s2, s3, s5, s6, s8, s12, s15, s16, s17. W porównaniu do poprzednich, ta wersja cechuje się lepszymi wynikami dla sekwencji s2, s12, s15, a gorszymi dla sekwencji s4. Ostatecznie wersja HP algorytmu cechuje się średnimi błędami poniżej 20° i odpowiednio niskimi odchyleniami standardowymi błędów również dla dziesięciu tych samych sekwencji testowych co MPQD: s1, s2, s3, s5, s6, s8, s12, s15, s16, s17. Na podstawie powyższej analizy błędów różnice w działaniu różnych wersji algorytmu można w szczególności zauważyć na przykładzie sekwencji Poprad s2, s4, s8, s12 oraz s15.

Warto podkreślić zasadność stosowania filtracji krawędziowej w celu uzyskania poprawnych wyników samokalibracji na etapie II. Działanie samokalibracji w dowolnej wersji bez wstępnego przetwarzania obrazów zawodzi w większości sytuacji. Przykładowo etap II



Rysunek 5.25: Błędy nachylenia linii dopasowania dla sekwencji testowych Poprad. Ujemne wartości błędów oznaczają, że z powodu braku pomiarów metoda nie wyznaczyła żadnej wartości dla danej sekwencji.



(a) Ustalone wzdłuż linii dopasowanie d_x, d_y pozwala nałożyć na siebie obiekty w dalekiej odległości od układu kamer.



(b) Ustalone wzdłuż linii dopasowanie d_x, d_y pozwala nałożyć na siebie obiekty w bliskiej odległości od układu kamer.

Rysunek 5.26: Przykładowe wyniki samokalibracji linii dopasowania (etap II samokalibracji) dla sekwencji testowych Poprad. Linię dopasowania zaznaczono czarną linią prostą o początku w środku obrazu.

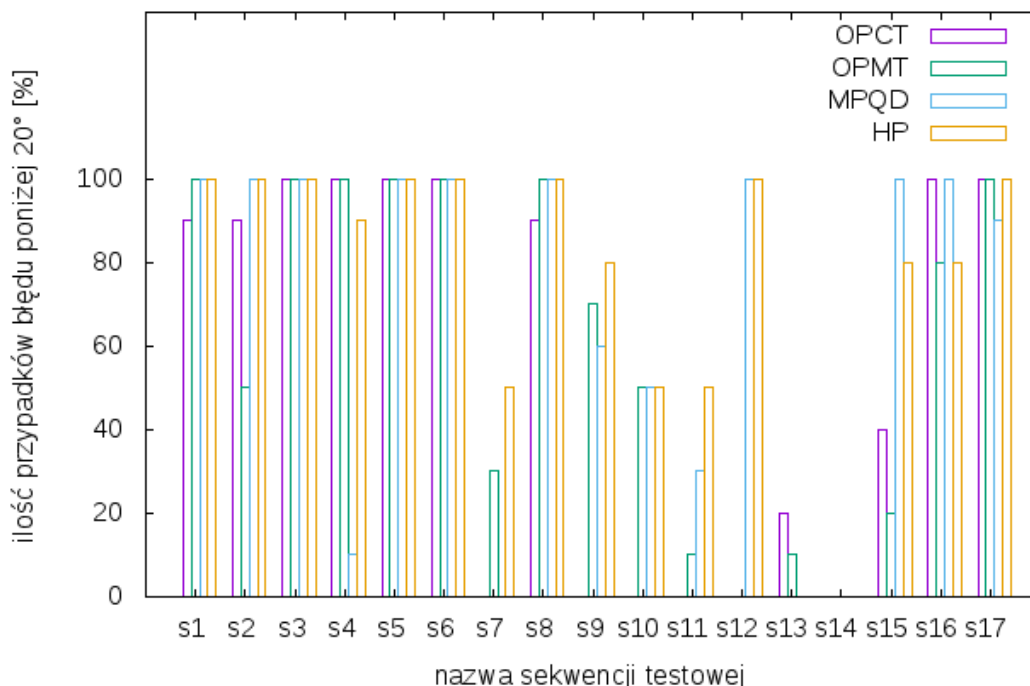
samokalibracji w wersji HP, ale bez wykorzystania filtracji krawędziowej jedynie w przypadku sekwencji s9, s10 i s15 zwraca wyniki o błędach poniżej 20° .

Podobnie jak w przypadku etapu I samokalibracji, etap II samokalibracji może stanowić użyteczną pomoc w trakcie procesu kalibracji układu kamer nawet jeśli nie zapewnia wysokiej dokładności działania w każdym możliwych warunkach. Sekwencje testowe Poprad nie zostały przygotowane specjalnie w celu poprawnych testów samokalibracji, ale stanowią przykładowe obrazy pochodzące z nagrań kamerami Poprad w trakcie monitorowania otoczenia. Na podstawie analizy wykresu 5.25 możemy wskazać, w jakich sytuacjach zaproponowana metoda samokalibracji na etapie II działa najlepiej. W ten sposób można opracować rekomendację sposobu stosowania opracowanej metody. Przedstawiono ją w na końcu bieżącego rozdziału w podsumowaniu wyników. Samokalibracja na etapie II z wykluczeniem sekwencji s4, s7, s9, s10, s11, s13 i s14 (dla których metoda nie działa, lub działa z niską dokładnością) cechuje się średnimi błędami i standardowym odchyleniem błędu nachylenia linii dopasowania na poziomie 9.46 ± 15.66 dla OPCT, 20.70 ± 37.63 dla OPMT, 3.62 ± 5.21 dla MPQD oraz 3.30 ± 5.42 dla HP. Te wyniki można podsumować ogólnym stwierdzeniem, że etap II samokalibracji dla wybranych sekwencji testowych Poprad działa z dokładnością lepszą dla metod opierających się na wyznaczaniu wielu wartości dopasowania (MPQD i HP).

Analiza błędów działania różnych wersji algorytmu automatycznej kalibracji na etapie II pozwala jednak na sformułowanie tylko ograniczonych wniosków. Lepszy wgląd w jakość działania różnych wersji samokalibracji przedstawia bardziej szczegółowa analiza wyników algorytmów dla każdej sekwencji. Podobnie jak przy analizie działania etapu I samokalibracji, działanie etapu II samokalibracji przedstawiono w dalszej części rozdziału na podstawie klasyfikacji błędów względem określonych progów. Wartości progów ustalono arbitralnie. Na wykresie 5.27 przedstawiono procentową liczbę przypadków (na dziesięć prób) ustalenia kąta nachylenia linii dopasowania z błędem poniżej 20° dla każdej sekwencji testowej Poprad.

Analizując działanie samokalibracji z wykorzystaniem filtracji krawędziowej na podstawie wykresu 5.27, należy zwrócić szczególną uwagę na sekwencje s2, s4, s8, s12 oraz s15. To na ich przykładzie można najlepiej zauważyć różnice w działaniu różnych wersji algorytmu samokalibracji na etapie II. Wprowadzenie w wersji OPMT techniki średniego progu odróżniającego pomiary dopasowania o niskiej i wysokiej jakości (te ostatnie biorą udział w estymacji przy użyciu RANSAC) pozwala poprawić dokładność kalibracji dla sekwencji s8. Poprawę działania można jednak zauważyć również w przypadku sekwencji s1, s7, s8, s9, s10 oraz s11. Technika średniego progu obniża znacząco dokładność samokalibracji dla sekwencji s2. Negatywny wpływ można zauważyć również dla sekwencji s13, s15 i s16. Zastosowanie w wersji MPQD algorytmu wyznaczania wielu maksimów korelacji fazowej za pomocą funkcji $k \arg \max$ powoduje względem wersji OPCT zwiększenie liczby poprawnie ustalonych parametrów kalibracji w szczególności dla sekwencji s12 i s15. Poprawę działania można jednak zauważyć również w przypadku sekwencji s1, s8, s9, s10 oraz s11. Pogorszenie jakości natomiast można w szczególności zaobserwować w przypadku sekwencji s4, jednakże następuje ono również nieznacznie dla sekwencji s17. Ostatecznie w wersji HP względem referencyjnej OPCT następuje poprawa działania w szczególności w przypadkach sekwencji

wyniki działania etapu II samokalibracji,
przypadki błędu nachylenia linii dopasowania poniżej 20°



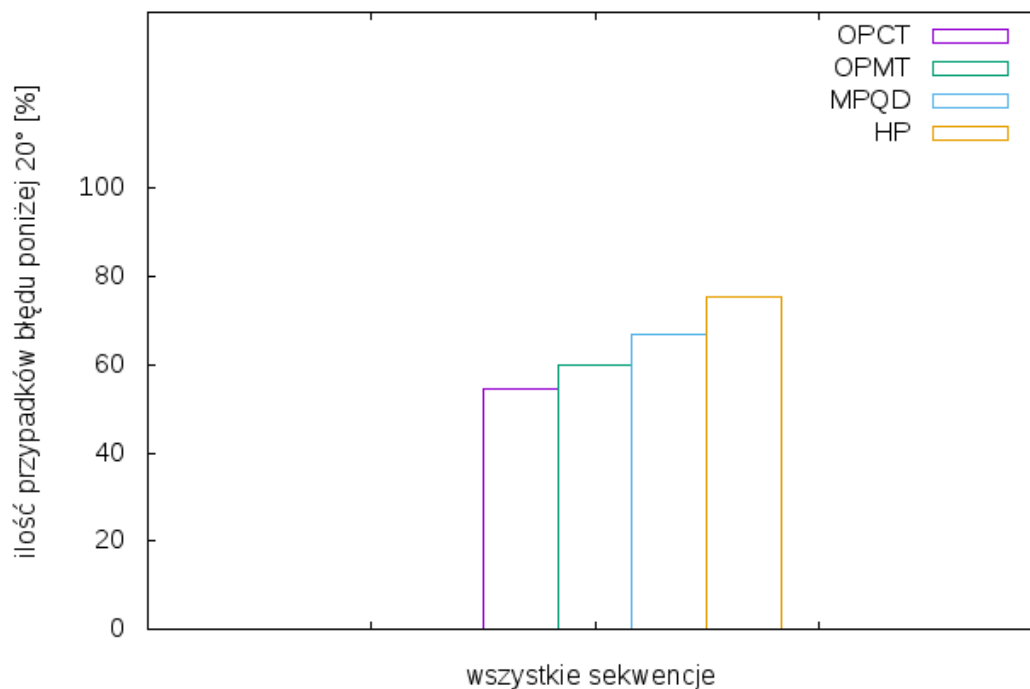
Rysunek 5.27: Wyniki działania II etapu samokalibracji dla sekwencji Poprad. Porównanie liczby przypadków ustalenia kąta nachylenia linii dopasowania różnymi metodami z błędami poniżej określonego progu.

s8, s12 i s15. Poprawę działania można jednak zauważyć również w przypadku sekwencji s1, s2, s9, s10 oraz s11. Pogorszenie działania następuje w szczególności dla sekwencji s4 jednakże można je również zauważyć w przypadku sekwencji s13 i s17.

W celu podsumowania wyników, dane z wykresu 5.27 przedstawiono sumarycznie dla wszystkich sekwencji na wykresie 5.28. Na tej podstawie można łatwo zauważyć, że dla przedstawionego zestawu danych testowych Poprad każda z zastosowanych w wersjach OPMT oraz MPQD technik poprawia jakość działania samokalibracji, a wersja HP zwraca wyniki z błędami poniżej 20° w największej liczbie przypadków. Analogicznie do wykresów działania metody samokalibracji na etapie II z błędami poniżej 20° przeprowadzono analizę dla innych zakresów błędów: poniżej 15°, 10°, 5° oraz 2°. We wszystkich tych przypadkach ogólna liczba przypadków uzyskania błędów poniżej ustalonego progu zmniejsza się, ale wnioski dotyczące pozytywnego wpływu technik zastosowanych w wersjach OPMT, MPQD i ostatecznie w HP są podobne. Z uwagi na czytelność odpowiednie wykresy nie zostały zamieszczone w rozprawie. Podsumowanie wyników metody w wersji HP dla różnych poziomów błędów przedstawiono na wykresie 5.29.

Etap drugi samokalibracji wiąże się z rozwiązaniem bardziej złożonych problemów z uwagi na konieczność wyznaczenia dopasowania multimodalnego. Na rysunku 5.30 przedstawiono przykład estymacji nachylenia linii dopasowania dla błędu poniżej i powyżej

wyniki działania etapu II samokalibracji,
wszystkie przypadki błędu nachylenia linii dopasowania poniżej 20°

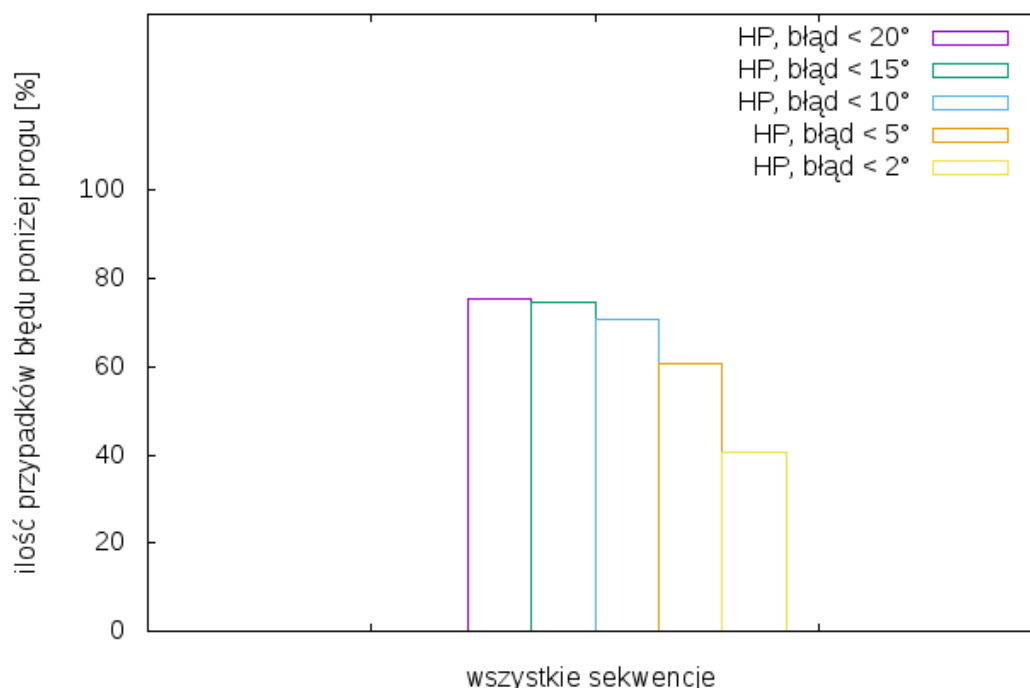


Rysunek 5.28: Sumaryczne wyniki działania II etapu samokalibracji dla sekwencji Poprad. Porównanie sumarycznej liczby przypadków ustalenia kąta nachylenia linii dopasowania różnymi metodami z błędami poniżej określonego progu.

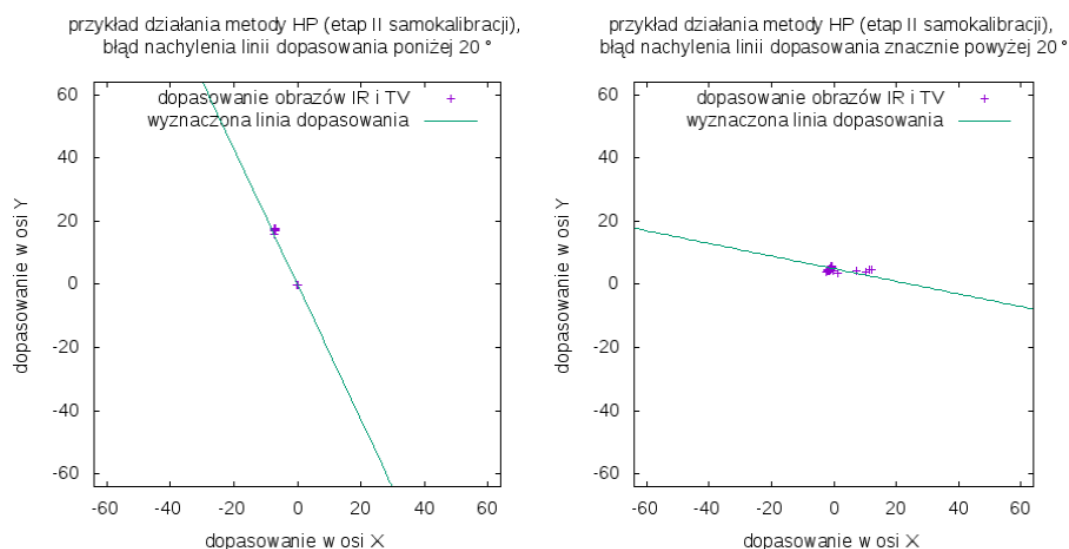
20° z wykorzystaniem dopasowania modelu do danych metodą RANSAC. W przypadku estymacji z błędem poniżej 20° można zauważyć dwie oddalone od siebie grupy wartości pomiarowych pozwalające łatwo dopasować do nich linię dopasowania. Przykład estymacji modelu z dokładnością powyżej 20° można tłumaczyć kilkoma prawdopodobnie błędnymi wartościami pomiarowymi położonymi w dodatnim kierunku osi d_x . Można przypuszczać, że grupa pomiarów umieszczona po lewej stronie opisuje dokładne wartości dopasowania multimodalnego. W przypadku estymacji z błędami poniżej 20° pomiary dopasowania w trybie kalibracji dotyczyły obiektów w bliskiej i dalekiej odległości od układu kamer. W przypadku estymacji z błędami powyżej 20° pomiary dotyczyły obiektów w stałej odległości od układu kamer, a błędy wyznaczenia dopasowania obrazów w trybie kalibracji spowodowały dalsze problemy z estymacją linii dopasowania.

Sytuacje, w których samokalibracja na etapie II osiąga niską dokładność, można tłumaczyć na dwa sposoby. Po pierwsze, problemy pojawiają się w sytuacji występowania pomiarów dotyczących obiektów w stałej odległości od układu kamer. Po drugie, w przypadku obrazów, które trudno dopasować, błędy w automatycznym wyznaczeniu dopasowania multimodalnego dodatkowo utrudniają dokładną estymację odpowiednich parametrów.

wyniki działania metody HP (etap II samokalibracji),
wszystkie przypadki błędu nachylenia linii dopasowania poniżej progu



Rysunek 5.29: Sumaryczne wyniki działania II etapu samokalibracji w wersji HP dla sekwencji Poprad. Porównanie sumarycznej liczby przypadków ustalenia kąta nachylenia linii dopasowania różnymi metodami z błędami poniżej określonego progu.

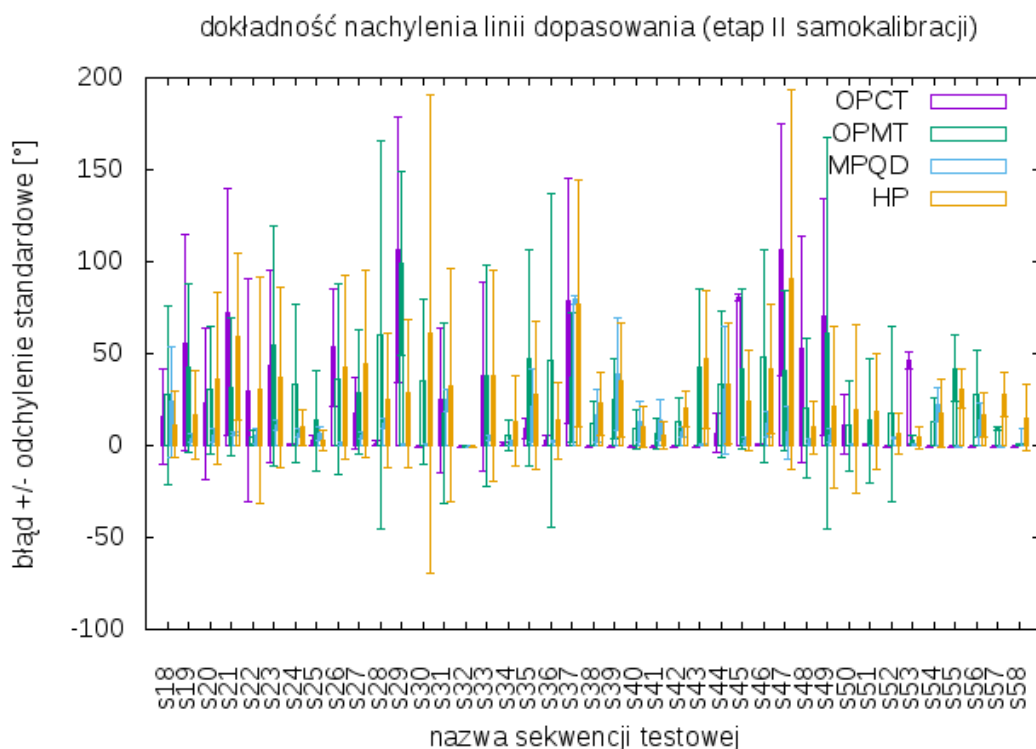


Rysunek 5.30: Przykłady estymacji parametrów linii dopasowania z wykorzystaniem RANSAC dla sekwencji testowych Poprad.

5.2.10 Wyniki II etapu dla sekwencji testowych KAIST

Wykres 5.31 przedstawia porównanie działania różnych wersji algorytmu etapu II samokalibracji kąta nachylenia linii dopasowania. W analogiczny sposób jak w podrozdziale

5.2.9 analizę wyników przeprowadzono w oparciu o porównanie działania czterech wersji algorytmu: OPCT, OPMT, MPQD i HP.



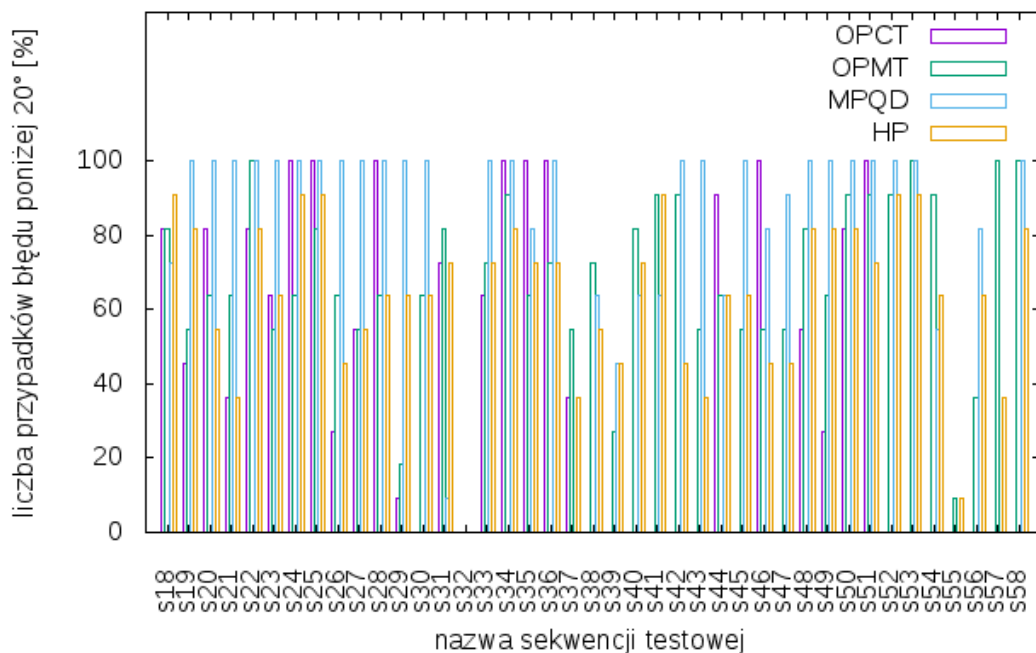
Rysunek 5.31: Błędy nachylenia linii dopasowania dla sekwencji testowych KAIST. Ujemne wartości błędów oznaczają, że z powodu braku pomiarów metoda nie wyznaczyła żadnej wartości dla danej sekwencji.

Z powodu znacznej liczby sekwencji analiza wykresu 5.31 jest utrudniona i pozwala na sformułowanie jedynie ograniczonych wniosków. Lepszy wgląd w jakość działania różnych wersji samokalibracji przedstawia bardziej szczegółowa analiza błędów algorytmów dla każdej sekwencji. Na wykresie 5.32 przedstawiono liczbę przypadków (na dziesięć prób) ustalenia kąta nachylenia linii dopasowania z błędem poniżej 20° dla każdej sekwencji testowej.

Podsumowanie danych z wykresu 5.32 przedstawiono sumarycznie dla wszystkich sekwencji na wykresie 5.33.

Analizując wykres 5.33 można zauważyć, że zastosowanie techniki MPQD pozwala osiągnąć błędy poniżej 20° dla największej liczby testowanych przypadków. Analizując wyniki przedstawione na wykresach 5.32 i 5.31 można opracować rekomendację stosowania tej metody samokalibracji i wskazać sekwencje, w których etap II samokalibracji cechuje się dobrą dokładnością. Przedstawiono ją na końcu bieżącego rozdziału w podsumowaniu wyników. Problemy z działaniem metody w przypadku sekwencji s30, s35 i s37 występują z powodów podobnych jak w przypadku sekwencji testowej Poprad s14 - sekwencje przedstawiają obiekty w stałej odległości od układu kamer znajdujące się na środku obrazów, a obiekty w innej odległości od układu widoczne są na obrzeżach obrazów. Zastosowanie okna Hanna w trakcie

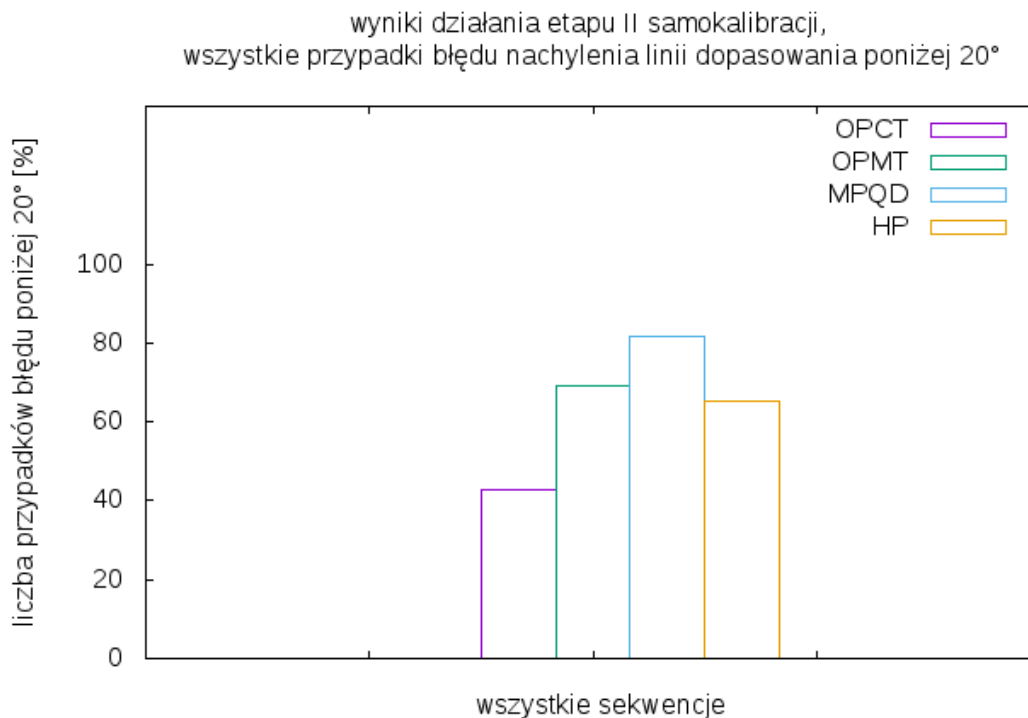
wyniki działania etapu II samokalibracji,
przypadki błędu nachylenia linii dopasowania poniżej 20°



Rysunek 5.32: Wyniki działania II etapu samokalibracji dla sekwencji KAIST. Porównanie liczby przypadków ustalenia kąta nachylenia linii dopasowania różnymi metodami z błędami poniżej określonego progu.

wyznaczania korelacji fazowej powoduje, że obiekty na obrzeżach są pomijane. Sekwencje s38-s41 zawierają obrazy o niewielkiej liczbie szczegółów, które trudno dopasować. Problemy z działaniem etapu II samokalibracji występują również w przypadku sekwencji s31 i s32, gdzie należałoby zastosować inny algorytm filtracji krawędziowej, który w lepszy sposób wyróżniłby odpowiadające sobie obiekty na obrazach wizyjnym i termowizyjnym. Podobne problemy występują w odniesieniu do sekwencji s44, s46, s47, s54-s57. Wykluczając z analizy sekwencje s31, s32, s35, s37-s41, s44, s46, s47, s54-s57 (dla których metoda nie działa lub cechuje się niską dokładnością), można stwierdzić, że etap II samokalibracji cechuje się średnim błędem i odchyleniem standardowym błędu nachylenia linii dopasowania na poziomie $34.61 \pm 50.51[^\circ]$ dla OPCT, $31.19 \pm 54.89[^\circ]$ dla OPMT, $5.06 \pm 7.97[^\circ]$ dla MPQD oraz $25.25 \pm 46.13[^\circ]$ dla HP. Dla wybranych sekwencji wyniki łatwo podsumować stwierdzeniem, że dla sekwencji testowych KAIST etap II samokalibracji działa najlepiej przy zastosowaniu techniki wyznaczenia wielu wartości dopasowania MPQD.

Podsumowanie wyników metody w wersji MPQD dla różnych progów błędów można natomiast odczytać z wykresu 5.34.

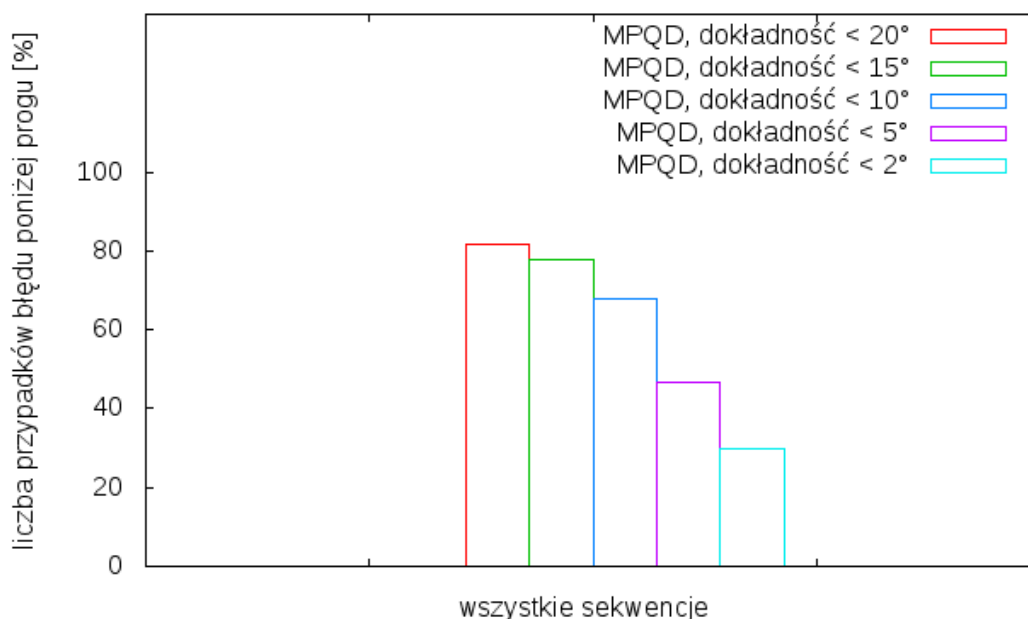


Rysunek 5.33: Sumaryczne wyniki działania II etapu samokalibracji dla sekwencji KAIST. Porównanie sumarycznej liczby przypadków ustalenia kąta nachylenia linii dopasowania różnymi metodami z błędami poniżej określonego progu.

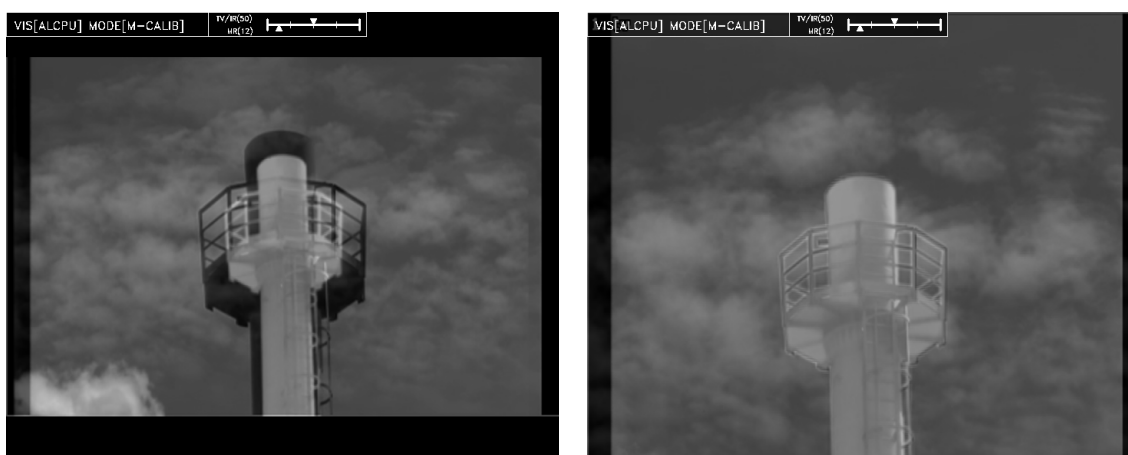
5.2.11 Podsumowanie

Opracowana metoda samokalibracji składa się z dwóch etapów i pozwala na ustalenie istotnych parametrów koniecznych do wyznaczenia fuzji obrazów online. W celu fuzji online obrazów konieczna jest w szczególności znajomość skali obrazu termowizyjnego względem wizyjnego oraz dopasowanie obrazów online. Na pierwszym etapie samokalibracji estymowana jest skala obrazu termowizyjnego względem wizyjnego. Na drugim etapie wyznaczane są parametry linii dopasowania, dzięki czemu rozwiązanie problemu dopasowania obrazów online można znacząco ułatwić. Kluczowe znaczenie w tym przypadku ma nachylenie linii dopasowania, ponieważ jest parametr najtrudniejszy do ustalenia w sposób manualny. Dla zbioru sekwencji testowych wygenerowanych komputerowo oraz dla sekwencji Poprad i KAIST pokazano, że samokalibracja może zostać wykonana dla układu dwóch kamer: wizyjnej i termowizyjnej na podstawie dopasowania tych obrazów. Dodatkowo pokazano, że możliwe jest zastosowanie metody korelacji fazowej w celu wyznaczenia tego dopasowania. Przykładowy wynik etapu I samokalibracji, czyli ustalenie rozmiaru obrazu termowizyjnego względem obrazu wizyjnego przedstawia rys. 5.35, natomiast przykładowy wynik etapu II samokalibracji, czyli ustalenie parametrów linii dopasowania przedstawia rys. 5.36.

wyniki działania metody MPQD (etap II samokalibracji),
wszystkie przypadki błędu nachylenia linii dopasowania poniżej progu



Rysunek 5.34: Sumaryczne wyniki działania II etapu samokalibracji w wersji MPQD dla sekwencji KAIST. Porównanie sumarycznej liczby przypadków ustalenia kąta nachylenia linii dopasowania różnymi metodami z błędami poniżej określonego progu.



(a) Fuzja przed kalibracją skali.

(b) Fuzja po kalibracji skali.

Rysunek 5.35: Przykładowe wyniki samokalibracji skali (etap I samokalibracji) dla sekwencji testowych Poprad.

Zarówno w odniesieniu do etapu I, jak i etapu II samokalibracji wyniki testów z wykorzystaniem sekwencji testowych wygenerowanych komputerowo potwierdziły ogólnie przyjętą koncepcję samokalibracji.

Dla etapu I samokalibracji z wykorzystaniem sekwencji testowych Poprad wykazano, że dla wybranych sekwencji (wykluczenie s9, s10, s11, s12, s13 oraz s14, przykładowe obrazy z tych sekwencji, które stwarzają najwięcej problemów, przedstawiono na rys. 5.37) metoda



(a) Ustalone wzdłuż linii dopasowanie d_x, d_y pozwala nałożyć na siebie obiekty w dalekiej odległości od układu kamer.

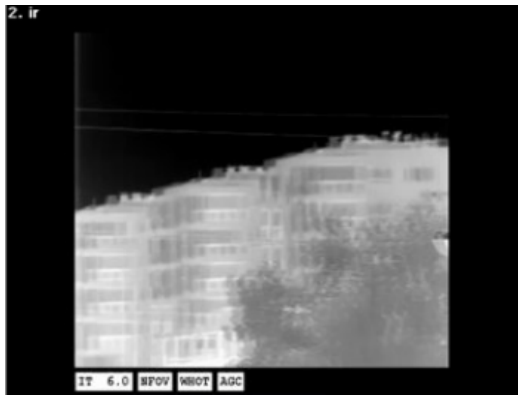


(b) Ustalone wzdłuż linii dopasowanie d_x, d_y pozwala nałożyć na siebie obiekty w bliskiej odległości od układu kamer.

Rysunek 5.36: Przykładowe wyniki samokalibracji linii dopasowania (etap II samokalibracji) dla sekwencji testowych Poprad. Linię dopasowania zaznaczono czarną linią prostą o początku w środku obrazu.

cechuje się średnim błędem i odchyleniem standardowym błędu w kierunku osi X na poziomie 0.018 ± 0.015 , a w kierunku osi Y na poziomie 0.011 ± 0.010 . Zastosowanie tego algorytmu samokalibracji zalecane jest w sytuacji, gdy na nagraniach zarejestrowany został znaczący ruch głowicy kamer względem otoczenia. Przy zbyt niskim zakresie zmienności tych ruchów (lub gdy one w ogóle nie występują) estymacja skali obrazu termowizyjnego względem wizyjnego cechuje się gorszą dokładnością i niższą precyzją.

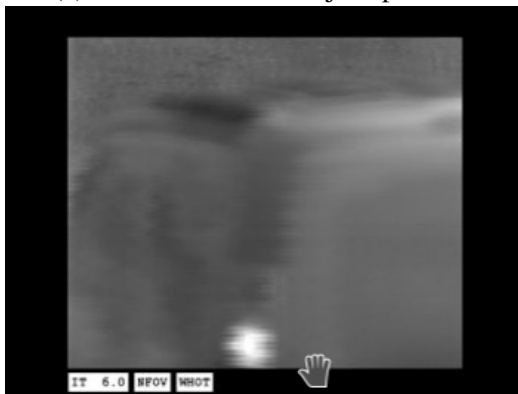
Wyniki etapu I samokalibracji dla sekwencji testowych KAIST potwierdzają powyższe wnioski. Jeśli wykluczymy z analizy sekwencje s27, s30, s31, s32, s52 i s54 (przykładowe obrazy z tych sekwencji, które stwarzają najwięcej problemów, przedstawiono na rys. 5.38 i 5.39), to możemy stwierdzić, że samokalibracja na etapie I cechuje się średnią dokładnością oraz błędem standardowym na poziomie 0.036 ± 0.077 w kierunku osi X i 0.050 ± 0.068 w kierunku osi Y. Podobnie jak dla sekwencji Poprad, zastosowanie etapu I samokalibracji jest zalecane w sytuacji zarejestrowania znacznych ruchów układu kamer względem otoczenia. W odniesieniu do bazy KAIST należy podkreślić, że nagrania KAIST nie spełniają przyjętych w rozprawie założeń dotyczących układu kamer - kamery obserwują obiekty w bliskim otoczeniu, pojazd porusza się w trakcie rejestracji nagrań, a układ kamer nie ma możliwości ruchu obrotowego. Nie spełnione są więc założenia związane z ruchem układu kamer i relacją tego ruchu do dopasowania monomodalnego pomiędzy kolejnymi obrazami w sekwencji wideo. Dla sekwencji KAIST sytuacje zbliżone do założeń przyjętych w rozprawie zachodzą głównie w przypadku jazdy samochodem osobowym po łuku (ruch układu kamer w kierunku osi X) oraz w przypadku przejeżdżania przez progi zwalniające i inne wniesienia (ruch w kierunku osi Y). Dodatkowo obrazy pochodzące z kamery termowizyjnej w wybranych przypadkach zawierają niewielką ilość informacji o obserwowanym otoczeniu. Te wszystkie spostrzeżenia tłumaczą



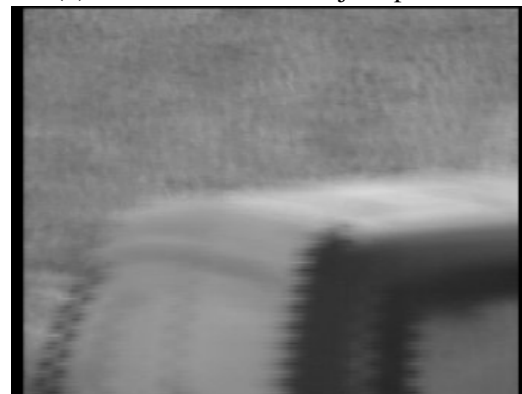
(a) Obraz IR z sekwencji Poprad s4.



(b) Obraz TV z sekwencji Poprad s4.



(c) Obraz IR z sekwencji Poprad s9.



(d) Obraz TV z sekwencji Poprad s9.



(e) Obraz IR z sekwencji Poprad s10.



(f) Obraz TV z sekwencji Poprad s10.

Rysunek 5.37: Przykładowe obrazy z sekwencji Poprad, które stwarzają najwięcej problemów.

uzyskanie gorszych wyników etapu I samokalibracji dla bazy KAIST niż w odniesieniu do sekwencji testowych Poprad.

W odniesieniu do etapu II samokalibracji testowanego dla wybranych sekwencji Poprad (z wykluczeniem sekwencji s4, s7, s9, s10, s11, s13 i s14, przykładowe obrazy z tych sekwencji przedstawiono na rys. 5.37 i 5.39) otrzymano średnią dokładność i błąd standardowy nachylenia linii dopasowania na poziomie $9.46 \pm 15.66[^\circ]$ dla OPCT, $20.70 \pm 37.63[^\circ]$ dla OPMT, $3.62 \pm 5.21[^\circ]$ dla MPQD oraz $3.30 \pm 5.42[^\circ]$ dla HP. Wyniki jasno wskazują na istotną poprawę dokładności działania metody w przypadku zastosowania wielu wartości dopasowań (techniki MPQD i HP) wyznaczonych za pomocą korelacji fazowej dla obrazów wizyjnego i



(a) Obraz IR z sekwencji KAIST s27.



(b) Obraz TV z sekwencji KAIST s27.



(c) Obraz IR z sekwencji KAIST s31.



(d) Obraz TV z sekwencji KAIST s31.

Rysunek 5.38: Przykładowe obrazy z sekwencji KAIST, które stwarzają najwięcej problemów.

termowizyjnego. Obydwie techniki dają zgodne wyniki dokładności samokalibracji, jednakże biorąc pod uwagę analizę procentowej liczby przypadków uzyskania dokładności poniżej określonego progu, można wskazać, że zastosowanie techniki HP pozwala osiągnąć lepsze rezultaty. Metoda ta pozwala uzyskać dobrą dokładność w największej liczbie przypadków w odniesieniu do sekwencji testowych Poprad. Sytuacje, w których etap II samokalibracji na bazie testowej Poprad działa z wysoką dokładnością obejmują sceny zawierające wyraźne obiekty znajdujące się w bliskiej i w dalekiej odległości od układu kamer. Podobnie jak w przypadku etapu I, w etapie II należy zapewnić odpowiedni zakres zmienności pomiarów dopasowania. Obserwowanie sceny płaskiej, gdzie wszystkie obiekty znajdują się w tej samej odległości od układu kamer, będzie prowadzić do błędnego działania metody. Problemy z dokładnymi wynikami można zaobserwować również w sytuacji, gdy obserwowana scena zawiera obiekty w różnych odległościach na obrzeżach obrazów. Z powodu zastosowania okna Hanna w trakcie wyznaczania korelacji fazowej elementy na obrzeżach obrazów dopasowywanych mogą nie być brane pod uwagę przez algorytm.

Wyniki etapu II samokalibracji dla wybranych sekwencji KAIST (z wykluczeniem s31, s32, s35, s37-s41, s44, s46, s47, s54-s57, przykładowe obrazy z tych sekwencji przedstawiono na rys. 5.38 i 5.39) cechują się średnią dokładnością i błędem standardowym nachylenia linii dopasowania na poziomie $34.61 \pm 50.51[^\circ]$ dla OPCT, $31.19 \pm 54.89[^\circ]$ dla OPMT, $5.06 \pm 7.97[^\circ]$



(a) Obraz IR z sekwencji Poprad s11.



(b) Obraz TV z sekwencji Poprad s11.



(c) Obraz IR z sekwencji Poprad s14.



(d) Obraz TV z sekwencji Poprad s14.

Rysunek 5.39: Przykładowe obrazy z sekwencji Poprad, które stwarzają najwięcej problemów.

dla MPQD oraz $25.25 \pm 46.13[^\circ]$ dla HP. Na tej podstawie można zauważyć, że stosowanie techniki wyznaczania wielu wartości dopasowań MPQD prowadzi do poprawienia jakości działania metody. W przypadku bazy KAIST technika HP nie poprawia wyników w porównaniu do MPQD. Analiza procentowej liczby przypadków uzyskania dokładności poniżej określonego progu dla bazy KAIST wskazuje również na metodę MPQD jako najlepszy wybór sposobu wyznaczenia wartości dopasowań pomiędzy obrazami wizyjnym i termowizyjnym. Podobnie jak w przypadku sekwencji testowych Poprad, etap II samokalibracji najlepiej stosować w sytuacjach, gdy na obserwowanej scenie występuje odpowiedni zakres zmienności odległości obiektów od układu kamer. Jeśli obiekty umieszczone są na obrzeżach obrazów z sekwencji, nie będą one brane pod uwagę przez algorytm korelacji fazowej z uwagi na zastosowanie okna Hanna. Dodatkowo warto podkreślić, że w odniesieniu do bazy KAIST występowały problemy z wyznaczeniem dopasowania w sytuacjach, gdy obrazy wizyjne lub termowizyjne zawierały niewielką ilość szczegółów. W niektórych sytuacjach błędne działanie można wytłumaczyć brakiem synchronizacji czasowej obrazów z kamer wizyjnej i termowizyjnej.

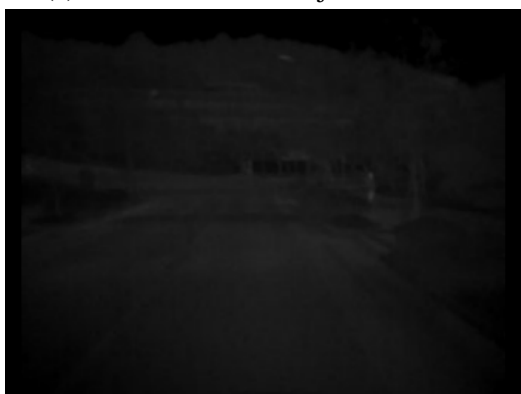
Podsumowanie dokładności działania opracowanej metody dla wybranych sekwencji przedstawiono w odniesieniu do etapu I na wykresie 5.41 i w odniesieniu do etapu II na wykresie 5.42. W szczególności warto podkreślić, że zarówno dla zbioru sekwencji testowych Poprad, jak i dla bazy danych KAIST, zastosowanie korelacji fazowej do



(a) Obraz IR z sekwencji KAIST s39.



(b) Obraz TV z sekwencji KAIST s39.



(c) Obraz IR z sekwencji KAIST s54.



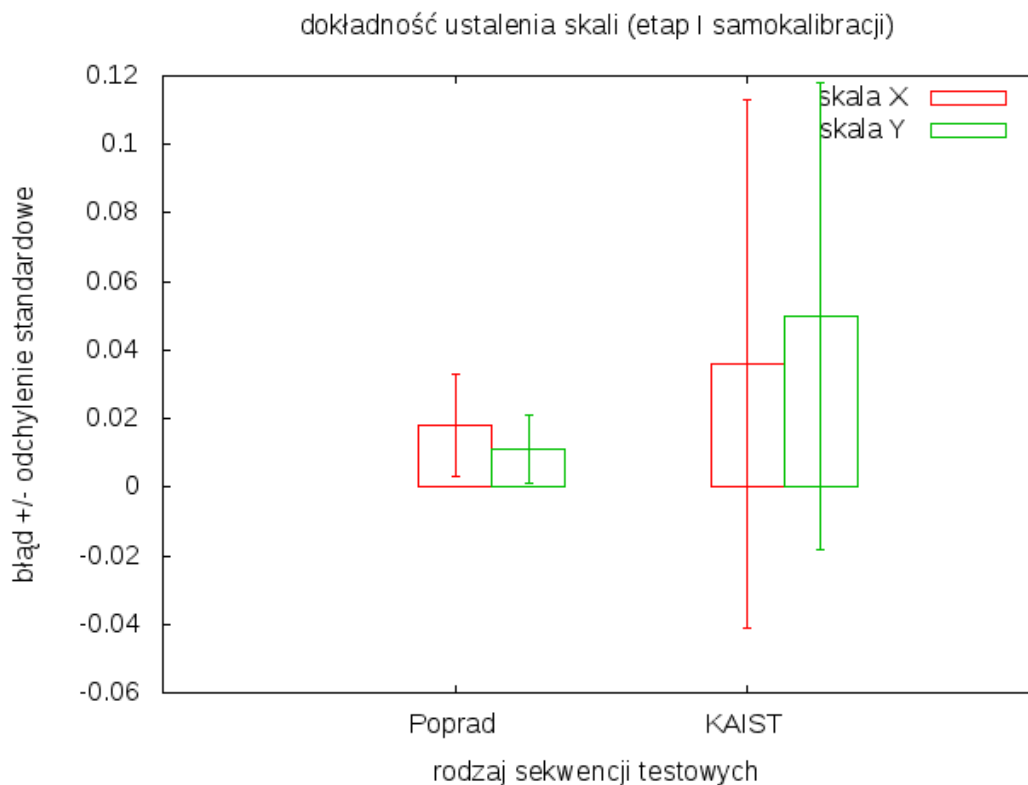
(d) Obraz TV z sekwencji KAIST s54.

Rysunek 5.40: Przykładowe obrazy z sekwencji KAIST, które stwarzają najwięcej problemów.

wyznaczania dopasowania dla wielu obiektów jednocześnie widocznych na obrazach wizyjnym i termowizyjnym pozwoliło poprawić dokładność działania zaproponowanej metody samokalibracji na etapie II. W przypadku sekwencji Poprad najlepsza okazała się technika HP wyznaczania wartości dopasowania, a w odniesieniu do bazy KAIST najskuteczniejsza okazała się technika MPQD.

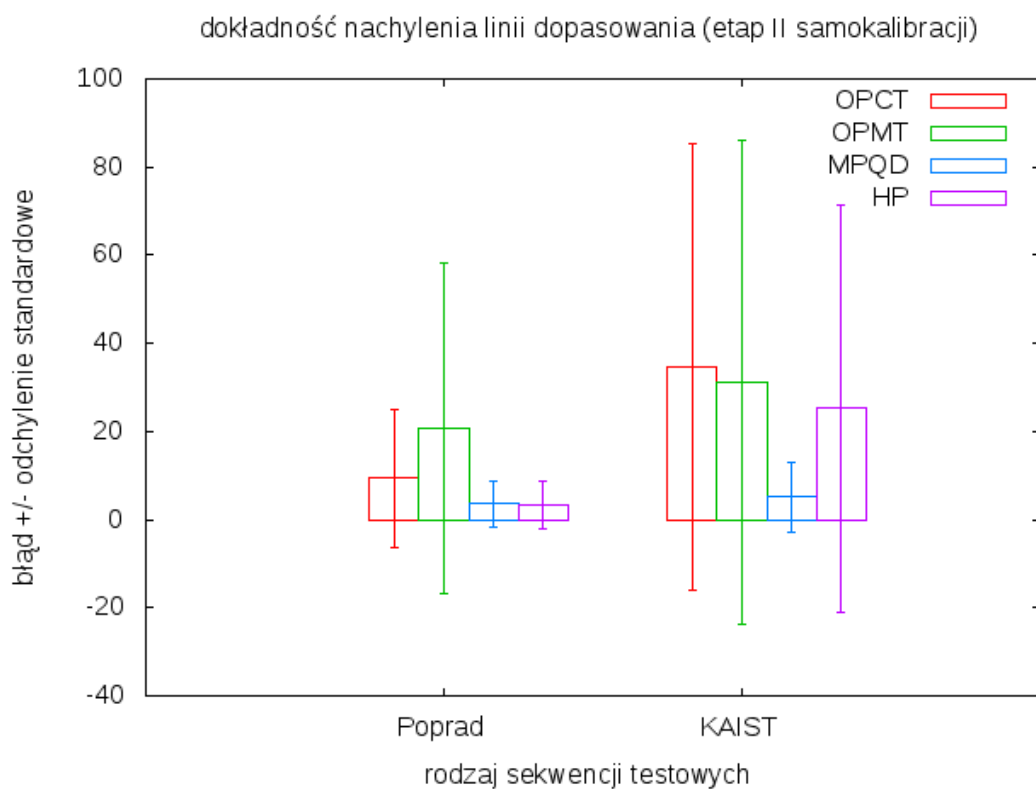
Zastosowanie etapu I samokalibracji zalecane jest w przypadku, gdy na nagraniach zarejestrowany został znaczący ruch głowicy kamer względem otoczenia w kierunku osi X i w kierunku osi Y. Zastosowanie etapu II samokalibracji zalecane jest w sytuacjach, w których obserwowane na nagraniach sceny zawierają wyraźne obiekty w różnych odległościach od układu kamer zarejestrowane w centrum obrazów wizyjnych i termowizyjnych. Sceny, gdzie zakres zmienności tej odległości jest większy pozwalają na dokładniejszą estymację parametrów układu kamer. Obiekty widoczne na obrzeżach obrazów są pomijane w trakcie dopasowania obrazów. Warto zauważyć, że problemy zarówno w działaniu etapu I, jak i etapu II mogą pojawić się również w przypadku uwidocznienia się przeplotu na obrazach (rys. 5.37). Takie sytuacje można zaobserwować głównie w trakcie ruchów głowicy w kierunku osi X.

Głównym czynnikiem powodującym wykluczenie możliwości stosowania opracowanej metody dla niektórych sekwencji są problemy z dokładnym działaniem dopasowania obrazów: wizyjnego i termowizyjnego. Wyznaczenie dopasowania pomiędzy obrazami tego typu jest



Rysunek 5.41: Podsumowanie dokładności działania etapu I samokalibracji dla wybranych sekwencji.

szczególnie trudnym zadaniem, które stanowi wyzwanie w obliczu obecnego stanu wiedzy. W najnowszej pracy z roku 2019 [80] przedstawiono algorytmy, które mogą cechować się lepszą dokładnością, ale kosztem złożoności obliczeniowej. Jednak nawet w przypadku tych algorytmów opisywane wyniki dopasowania obrazów znajdują się na poziomie 85%. Przedstawione w pracy [80] metody wyznaczają dopasowanie pomiędzy odpowiadającymi sobie punktami obrazów w czasie ok. 200 ms, natomiast w przypadku metody opisanej w rozprawie czas wyznaczenia dopasowań na podobnej jednostce obliczeniowej wynosi ok. 20 ms. W odniesieniu do implementacji na układzie Jetson TK1 czas wyznaczenia dopasowania metod opisanych w rozprawie wynosi ok. 80 ms. Czasy działania przedstawionej metody dopasowania dotyczą implementacji dla CPU i w ramach kontynuacji badań opisanych w rozprawie mogą zostać znacznie zredukowane w wyniku implementacji na układy typu GPGPU.



Rysunek 5.42: Podsumowanie dokładności działania etapu II samokalibracji dla wybranych sekwencji.

Rozdział 6

Podsumowanie i wnioski

Rozprawa dotyczy zagadnienia kalibracji układu kamer dalekiego zasięgu: wizyjnej i termowizyjnej. Problem pojawił się w trakcie prac badawczo-rozwojowych nad układem fuzji obrazów dla polskiego przemysłu obronnego i dotyczy głowicy optoelektronicznej zestawu przeciwlotniczego Poprad. Rozwiązanie problemu kalibracji z zastosowaniem klasycznych metod z wykorzystaniem wzorca kalibracyjnego okazało się niepraktyczne w warunkach poligonowych. W związku z tym za cel rozprawy przyjęto opracowanie metody samokalibracji dla przedstawionego układu dwóch kamer na podstawie automatycznie wyznaczonych cech obrazów pochodzących z obserwacji otoczenia.

Pierwsza część rozprawy zawiera wprowadzenie teoretyczne do poruszanych zagadnień badawczych. Na początku opisano stan badań dotyczący metod kalibracji układu dwóch kamer ze szczególnym uwzględnieniem metod samokalibracji mogących mieć zastosowanie do realizacji przedstawionego powyżej celu. Uzasadniono również przyjęcie samokalibracji jako sposobu rozwiązania problemu kalibracji opisanego układu kamer.

Standardowe podejście do kalibracji zakłada zastosowanie wzorca kalibracyjnego, co znacznie ułatwia proces kalibracji. Przyjęte w rozprawie założenia wiążą się w szczególności z dwoma utrudnieniami wykorzystania wzorca kalibracyjnego.

Po pierwsze zakłada się, że kamery będące elementem układu kamer są kamerami dalekiego zasięgu. Oznacza to, że obrazy dla obiektów umieszczonych blisko takiej kamery są rozmyte. Aby obraz wzorca miał wyraźne krawędzie, wzorec kalibracyjny należałoby umieścić w znacznej odległości od układu kamer. Jednakże w takiej sytuacji konieczne byłoby skonstruowanie wzorca o znacznych rozmiarach. Konstrukcja dużego wzorca jest niewskazana z uwagi na problemy związane ze stosowaniem takiego wzorca w warunkach polowych.

Po drugie zakłada się, że układ kamer składa się z kamery wizyjnej oraz termowizyjnej. Konstrukcja wzorca kalibracyjnego dobrze widocznego w tych dwóch pasmach promieniowania elektromagnetycznego powoduje dodatkowe poważne problemy.

Te dwa wnioski uzasadniły wykorzystanie metod samokalibracji. W wyniku wykonanego przeglądu ustalono, że jednym z kluczowych elementów większości metod samokalibracji

jest automatyczne odnalezienie na różnych obrazach odpowiadających sobie obiektów. Automatyczne dopasowanie obrazów jest jednym ze sposobów realizacji tego zadania. Dopasowanie określa, w jaki sposób przekształcić obiekty jednego obrazu tak, aby odpowiadały położeniu i kształtom obiektów obrazu drugiego.

W świetle powyższych ustaleń przedstawiono przegląd metod automatycznego dopasowania obrazów w odniesieniu do obrazów wizyjnego i termowizyjnego. Ze szczególną uwagą potraktowano publikacje z zakresu dopasowania sekwencji obrazów wizyjnych i termowizyjnych dla potrzeb monitorowania. Opisano również publikacje z zakresu metod bazujących na korelacji fazowej. Metody te cechują się możliwością wydajnej implementacji, a skuteczność ich działania w zakresie dopasowania obrazów wizyjnych i termowizyjnych została przez autora potwierdzona przy założeniach zbliżonych do założeń przyjętych w rozprawie i opisana w odpowiedniej publikacji.

Druga część rozprawy zawiera opis opracowanego przez autora rozwiązania problemu kalibracji. Przedstawiono opracowaną metodę samokalibracji układu kamer: wizyjnej i termowizyjnej na podstawie automatycznego dopasowania obrazów z tych kamer, wyznaczonego z wykorzystaniem metody korelacji fazowej dla wielu obiektów jednocześnie widocznych na tych obrazach. W szczególności pokazano, w jaki sposób model ogólny układu kamer dostosowano do przyjętych w rozprawie założeń i przedstawiono uzasadnienie wykorzystania metody bazującej na korelacji fazowej.

Na początku tej części przedstawiono od strony teoretycznej opracowaną metodę samokalibracji układu dwóch kamer wizyjnej i termowizyjnej. Wykazano (twierdzenie 1), że w świetle przyjętych założeń kalibrację można podzielić na dwa etapy. Pierwszy etap ma na celu ustalenie parametrów skali obrazu termowizyjnego względem wizyjnego. Estymacja tych parametrów następuje na podstawie automatycznego dopasowania w trybie kalibracji obrazów z chwili bieżącej do obrazów z chwili poprzedniej, odpowiednio dla kamery wizyjnej i termowizyjnej. W etapie drugim (twierdzenie 2) ustalane są parametry opisujące zależność parametrów dopasowania odpowiadających sobie obiektów obrazu wizyjnego i termowizyjnego. Tę zależność można przedstawić jako linię dopasowania. Estymacja tych parametrów następuje na podstawie automatycznego dopasowania w trybie kalibracji odpowiadających sobie obiektów obrazów wizyjnego i termowizyjnego.

Następnie przedstawiono opracowaną metodę automatycznego dopasowania obrazów dla potrzeb samokalibracji układu kamer wizyjnej i termowizyjnej. Zaprezentowano również uzasadnienie wykorzystania metody korelacji fazowej do wyznaczenia dopasowania w trybie kalibracji dla wielu obiektów jednocześnie widocznych na obrazach z kamery wizyjnej i termowizyjnej.

Ostatecznie, w celu weryfikacji opracowanej metody dokonano jej implementacji i przetestowano jej działanie na sekwencjach obrazów z głowicy Poprad oraz na ogólnodostępnej bazie obrazów KAIST. W wyniku testów potwierdzono skuteczność metody i przedstawiono

warunki jej stosowania w zakresie obserwowanych scen. W ten sposób potwierdzono prawdziwość tezy postawionej na wstępie rozprawy.

W szczególności warto podkreślić, że zarówno dla zbioru sekwencji testowych Poprad, jak i dla bazy danych KAIST, zastosowanie korelacji fazowej do wyznaczania dopasowania dla wielu obiektów jednocześnie widocznych na obrazach wizyjnym i termowizyjnym pozwoliło poprawić dokładność działania zaproponowanej metody samokalibracji na etapie II w stosunku do klasycznie wykorzystywanego dopasowania dla jednego obiektu.

Testy przeprowadzone dla bazy danych KAIST, gdzie ani układ kamer, ani ich wykorzystanie nie spełnia przyjętych w rozprawie założeń, pokazały, że opracowana metoda samokalibracji ma charakter uniwersalny i może znaleźć zastosowanie nie tylko w przemyśle obronnym (na przykład do celów kalibracji układu kamer multimodalnych samochodu autonomicznego).

Praca zawiera następujące istotne elementy nowości, według najlepszej wiedzy autora nie opisane wcześniej w literaturze:

- Wykazano, że kalibrację układu kamer dalekiego zasięgu: wizyjnej i termowizyjnej w świetle przyjętych założeń można podzielić na dwa etapy. Pierwszy etap ma na celu ustalenie parametrów skali obrazu termowizyjnego względem wizyjnego. W etapie drugim ustalana jest zależność wiążąca parametry dopasowania odpowiadających sobie obiektów obrazu wizyjnego i termowizyjnego. Zarówno parametry skali, jak i parametry opisujące powyższą zależność, definiują analizowany układ kamer. Takie przedstawienie problemu kalibracji pozwala na ustalenie istotnych parametrów koniecznych do wyznaczenia fuzji obrazów online.
- Estymacja parametrów układu kamer w trakcie samokalibracji została zrealizowana na podstawie automatycznego dopasowania obrazów w sekwencji wideo. Estymacja parametrów skali następuje na podstawie automatycznego dopasowania w trybie kalibracji obrazów z chwili bieżącej do obrazów z chwili poprzedniej odpowiednio dla kamery wizyjnej i termowizyjnej. W etapie drugim ustalone są parametry opisujące linię dopasowania odpowiadających sobie obiektów obrazu wizyjnego i termowizyjnego. W szczególności dla etapu drugiego samokalibracji przedstawiono, w jaki sposób metodę korelacji fazowej wykorzystać do wyznaczenia wartości dopasowań dla wielu obiektów jednocześnie widocznych na obrazach z kamery wizyjnej i termowizyjnej. Na obydwu etapach automatyczne dopasowanie obrazów zrealizowano na podstawie automatycznego dopasowania wielu par obrazów z wykorzystaniem metody korelacji fazowej i zastosowania wobec nich metody RANSAC.
- Metodę samokalibracji zaprojektowano w ten sposób, aby wykorzystać najlepsze cechy metody korelacji fazowej bez negatywnych skutków związanych z jej wadami. Największymi zaletami tej metody jest szybkość działania oraz odporność na jednorodne zmiany intensywności pomiędzy dopasowywanymi obrazami. Największą wadą jej jest

fakt, że wyznaczone w dziedzinie częstotliwości dopasowania trudno przyporządkować do konkretnych obszarów w dziedzinie obrazu. W związku z tym trudno uzyskać informację o odpowiadających sobie na obrazach obiektach, którą wykorzystuje większość metod samokalibracji opisywanych w literaturze. Metoda samokalibracji opracowana w ramach rozprawy została zaprojektowana tak, że nie potrzebuje takiego przyporządkowania, co stanowi jej istotną zaletę.

Badania nad kalibracją obrazów multimodalnych zostały rozpoczęte w ramach projektu rozwojowego w odpowiedzi na zapotrzebowanie przemysłu obronnego, ale wykorzystanie osiągniętych wyników może być znacznie bardziej uniwersalne. Przykładowo fuzja danych z kamery światła dziennego i termowizyjnej może znaleźć zastosowanie w diagnostyce maszyn przemysłowych i inspekcjach cieplnych budynków, w zakresie ochrony mienia, czy w ratownictwie. W ramach dalszych prac nad tematyką poruszaną w rozprawie możliwe jest zbadanie skuteczności opracowanej metody dla obrazów o innych modalnościach i w zakresie innych aplikacji, niż te poruszane w rozprawie.

Kolejnym kierunkiem rozwoju przedstawionych badań jest udoskonalenie algorytmu automatycznego wyznaczania dopasowania. Taki algorytm mógłby poprawić dokładność i rozszerzyć warunki stosowania zaproponowanej metody samokalibracji. Dodatkowo warto zbadać możliwość realizacji algorytmów filtracji krawędziowej w dziedzinie częstotliwości, co mogłoby pozwolić na dalsze zredukowanie czasu obliczeń. Pewne rozważania na ten temat można znaleźć w pracy [147].

W ostatnich latach, w zakresie nowoczesnych metod obliczeniowych i nowych algorytmów, dużą popularnością cieszą się metody bazujące na głębokich sieciach uczących (DNN). Rozwinięcie badań prowadzonych w ramach rozprawy można ukierunkować w stronę sprawdzenia możliwości wykorzystania sieci DNN do rozwiązania wybranych problemów związanych z kalibracją, dopasowaniem, czy przetwarzaniem wstępnym obrazów. Ciekawe informacje na ten temat można znaleźć w pracy [127], gdzie autorzy pokazali, w jaki sposób wyznaczyć dopasowanie obrazów z wykorzystaniem ich reprezentacji w dziedzinie częstotliwości. Niektórzy autorzy przedstawiają też możliwość wykorzystania niższych warstw sieci DNN, nauczonych rozwiązywania jednego problemu, w celu ułatwienia rozwiązania innego problemu. Wydaje się to również obiecującym kierunkiem dalszych prac w zakresie kalibracji i dopasowania obrazów: wizyjnego i termowizyjnego.

Bibliografia

- [1] H. Jung and J. Lyou. Matching of thermal and color images with application to power distribution line fault detection. In *2015 15th International Conference on Control, Automation and Systems (ICCAS)*, pages 1389–1392, Oct 2015.
- [2] M. Kowalski, N. Palka, M. Piszczek, and M. Szustakowski. Hidden object detection system based on fusion of thz and vis images. *Acta Physica Polonica A, ISSN 0587-4246*, 124:490–493, 09 2013.
- [3] O. D. Faugeras, Q. T. Luong, and S. J. Maybank. Camera self-calibration: Theory and experiments. In G. Sandini, editor, *Computer Vision — ECCV’92*, pages 321–334, Berlin, Heidelberg, 1992. Springer Berlin Heidelberg.
- [4] W. Ursine, F. Calado, G. Teixeira, H. Diniz, Prof. S. Silvino, and R. de Andrade. Thermal visible autonomous stereo visio system calibration methodology for non-controlled environments. In *Quantitative InfraRed Thermography, 2012 11th International Conference on*, June 2012.
- [5] Yannick Zoetgnande, Alain-Jérôme Fougères, Geoffroy Cormier, and Jean-Louis Dillenseger. Robust low resolution thermal stereo camera calibration. In *11th International Conference on Machine Vision (ICMV18)*, volume 11041 of *proc SPIE*, page 110411D, Munich, Germany, November 2018.
- [6] R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, ISBN: 0521540518, second edition, 2004.
- [7] Mili Shah, Roger D. Eastman, and Tsai Hong. An overview of robot-sensor calibration methods for evaluation of perception systems. In *Proceedings of the Workshop on Performance Metrics for Intelligent Systems, PerMIS ’12*, pages 15–20, New York, NY, USA, 2012. ACM.
- [8] Fadi Dornaika and R Chung. Self-calibration of a stereo rig without stereo correspondence. In *Vision Interface ’99, Trois-Rivières, Canada*, may 1999.
- [9] Stephen J. Maybank and Olivier D. Faugeras. A theory of self-calibration of a moving camera. *International Journal of Computer Vision*, 8(2):123–151, Aug 1992.

- [10] S. Bougnoux. From projective to euclidean space under any practical situation, a criticism of self-calibration. In *Sixth International Conference on Computer Vision (IEEE Cat. No.98CH36271)*, pages 790–796, Jan 1998.
- [11] Q.-T. Luong and O.D. Faugeras. Self-calibration of a moving camera from point correspondences and fundamental matrices. *International Journal of Computer Vision*, 22(3):261–289, Mar 1997.
- [12] A. Heyden and K. Astrom. Euclidean reconstruction from constant intrinsic parameters. In *Proceedings of 13th International Conference on Pattern Recognition*, volume 1, pages 339–343 vol.1, Aug 1996.
- [13] M. Pollefeys, L. Van Gool, and A. Oosterlinck. The modulus constraint: a new constraint for self-calibration. In *Proceedings of 13th International Conference on Pattern Recognition*, volume 1, pages 349–353 vol.1, Aug 1996.
- [14] Marc Pollefeys and Luc Van Gool. Self-calibration from the absolute conic on the plane at infinity. In Gerald Sommer, Kostas Daniilidis, and Josef Pauli, editors, *Computer Analysis of Images and Patterns*, pages 175–182, Berlin, Heidelberg, 1997. Springer Berlin Heidelberg.
- [15] E. E. Hemayed. A survey of camera self-calibration. In *Proceedings of the IEEE Conference on Advanced Video and Signal Based Surveillance, 2003.*, pages 351–357, July 2003.
- [16] T. Moons, L. Van Gool, M. Proesmans, and E. Pauwels. Affine reconstruction from perspective image pairs with a relative object-camera translation in between. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(1):77–83, Jan 1996.
- [17] Martin Armstrong, Andrew Zisserman, and Paul Beardsley. Euclidean structure from uncalibrated images. 01 1994.
- [18] M. Pollefeys and L. van Gool. Stratified self-calibration with the modulus constraint. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(8):707–724, Aug 1999.
- [19] A. Zisserman, P. A. Beardsley, and I. D. Reid. Metric calibration of a stereo rig. In *Proceedings IEEE Workshop on Representation of Visual Scenes (In Conjunction with ICCV’95)*, pages 93–100, June 1995.
- [20] R. Horaud, G. Csurka, and D. Demirdijian. Stereo calibration from rigid motions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(12):1446–1452, Dec 2000.

- [21] Richard I. Hartley. Euclidean reconstruction from uncalibrated views. In Joseph L. Mundy, Andrew Zisserman, and David Forsyth, editors, *Applications of Invariance in Computer Vision*, pages 235–256, Berlin, Heidelberg, 1994. Springer Berlin Heidelberg.
- [22] B. Triggs. Autocalibration and the absolute quadric. In *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 609–614, June 1997.
- [23] Manolis I.A. Lourakis and Rachid Deriche. Camera Self-Calibration Using the Singular Value Decomposition of the Fundamental Matrix: From Point Correspondences to 3D Measurements. Technical Report RR-3748, INRIA, August 1999.
- [24] Cheng Lei, Fuchao Wu, Zhanyi Hu, and H. T. Tsui. A new approach to solving kruppa equations for camera self-calibration. In *Object recognition supported by user interaction for service robots*, volume 2, pages 308–311 vol.2, Aug 2002.
- [25] Richard I. Hartley. Self-calibration from multiple views with a rotating camera. In Jan-Olof Eklundh, editor, *Computer Vision — ECCV '94*, pages 471–478, Berlin, Heidelberg, 1994. Springer Berlin Heidelberg.
- [26] Martin Armstrong, Andrew Zisserman, and Richard Hartley. Self-calibration from image triplets. In Bernard Buxton and Roberto Cipolla, editors, *Computer Vision — ECCV '96*, pages 1–16, Berlin, Heidelberg, 1996. Springer Berlin Heidelberg.
- [27] O. Faugeras, , and P. Strum. Self-calibration of a 1d projective camera and its application to the self-calibration of a 2d projective camera. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(10):1179–1185, Oct 2000.
- [28] and S. B. Kang, , and G. Xu. Error analysis of pure rotation-based self-calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(2):275–280, Feb 2004.
- [29] Hua Li, Huaifeng Zhang, Fuchao Wu, and Zhanyi Hu. A new linear camera self-calibration technique. In *Computer Vision, 5th Asian Conf. on*, Melbourne, Australia, Jan 2002.
- [30] Bill Triggs. Autocalibration from planar scenes. In *Proceedings of the 5th European Conference on Computer Vision-Volume I - Volume I*, ECCV '98, pages 89–105, London, UK, UK, 1998. Springer-Verlag.
- [31] Ezio Malis and Roberto Cipolla. Camera self-calibration from unknown planar structures enforcing the multiview constraints between collineations. *IEEE Trans. Pattern Anal. Mach. Intell.*, 24(9):1268–1272, September 2002.

- [32] Ezio Malis and Roberto Cipolla. Multi-view constraints between collineations: Application to self-calibration from unknown planar structures. In David Vernon, editor, *Computer Vision — ECCV 2000*, pages 610–624, Berlin, Heidelberg, 2000. Springer Berlin Heidelberg.
- [33] P. F. Sturm and S. J. Maybank. On plane-based camera calibration: A general algorithm, singularities, applications. In *Proceedings. 1999 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (Cat. No PR00149)*, volume 1, pages 432–437 Vol. 1, June 1999.
- [34] Z. Zhang. A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(11):1330–1334, Nov 2000.
- [35] D. Liebowitz and A. Zisserman. Metric rectification for perspective images of planes. In *Proceedings. 1998 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (Cat. No.98CB36231)*, pages 482–488, June 1998.
- [36] Gang Xu, J. Terai, and Heung-Yeung Shum. A linear algorithm for camera self-calibration, motion and structure recovery for multi-planar scenes from two perspective images. In *Proceedings IEEE Conference on Computer Vision and Pattern Recognition. CVPR 2000 (Cat. No.PR00662)*, volume 2, pages 474–479 vol.2, June 2000.
- [37] Didier Bondyfalat and Sylvain Bougnoux. Imposing euclidean constraints during self-calibration processes. In Reinhard Koch and Luc Van Gool, editors, *3D Structure from Multiple Images of Large-Scale Environments*, pages 224–235, Berlin, Heidelberg, 1998. Springer Berlin Heidelberg.
- [38] D. Liebowitz and A. Zisserman. Combining scene and auto-calibration constraints. In *Proceedings of the Seventh IEEE International Conference on Computer Vision*, volume 1, pages 293–300 vol.1, Sep. 1999.
- [39] Abderrahim Saaidi, Akram Halli, H Tairi, and Khalid Satori. Self-calibration using a planar scene and parallelogram. *ICGST-GVIP*, 9(1), 04 2009.
- [40] N.E. Akkad, Mostafa Merras, Abderrahim Saaidi, and Khalid Satori. Robust method for self-calibration of cameras having the varying intrinsic parameters. *Journal of Theoretical and Applied Information Technology*, 50:57–67, 01 2013.
- [41] P. Sturm. Critical motion sequences for monocular self-calibration and uncalibrated euclidean reconstruction. In *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 1100–1105, June 1997.
- [42] Marc Pollefeys. Self-calibration and metric 3d reconstruction from uncalibrated image sequences. 1999.

- [43] Y. Ma, S. Soatto, J. Kosecka, and S. Sastry. Euclidean reconstruction and reprojection up to subgroups. In *Proceedings of the Seventh IEEE International Conference on Computer Vision*, volume 2, pages 773–780 vol.2, Sep. 1999.
- [44] F. Kahl. Critical motions and ambiguous euclidean reconstructions in auto-calibration. In *Proceedings of the Seventh IEEE International Conference on Computer Vision*, volume 1, pages 469–475 vol.1, Sep. 1999.
- [45] P. Sturm. A case against kruppa’s equations for camera self-calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(10):1199–1204, Oct 2000.
- [46] Aziz Baataoui, Ismail El Batteoui, and Khalid Satori. Camera self-calibration by an equilateral triangle. *IJCA Special Issue on Software Engineering, Databases and Expert Systems*, SEDEX(1):29–34, September 2012. Full text available.
- [47] A. Baataoui, N. El Akkad, A. Saaidi, K. Satori, and M. Masrar. Robust method for camera self-calibration by an unknown planar scene. *Machine Graphics and Vision*, Vol. 23, No. 1/2:37–57, 2014.
- [48] Z. Zhang, Q. Luong, and O. Faugeras. Motion of an uncalibrated stereo rig: self-calibration and metric reconstruction. *IEEE Transactions on Robotics and Automation*, 12(1):103–113, Feb 1996.
- [49] Thao Dang, Christian Hoffmann, and Christoph Stiller. Continuous stereo self-calibration by camera parameter tracking. *IEEE transactions on image processing : a publication of the IEEE Signal Processing Society*, 18:1536–50, 08 2009.
- [50] Joonas Melin and Risto Ritala. Novel method for online stereo self-calibration. In *XXI IMEKO World Congress, Measurement in Research and Industry, Prague, Czech Republic*, August 30 - September 4 2015.
- [51] J. T. Lussier and S. Thrun. Automatic calibration of rgbd and thermal cameras. In *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 451–458, Sep. 2014.
- [52] Bilodeau G.A. Torabi A. Local self-similarity-based registration of human {ROIs} in pairs of stereo thermal-visible videos. *Pattern Recognition*, 46(2):578 – 589, 2013.
- [53] Maarten Vandersteegen. Generic thermal-visible camera calibration. In *De Nayer, KU Leuven*, November 2017.
- [54] Z. Yu, S. Lincheng, Z. Dianle, Z. Daibing, and Y. Chengping. Camera calibration of thermal-infrared stereo vision system. In *2013 Fourth International Conference on Intelligent Systems Design and Engineering Applications*, pages 197–201, Nov 2013.

- [55] L. St-Laurent, D. Prévost, and X. Maldague. Fast and accurate calibration-based thermalcolour sensors registration. In *Quantitative InfraRed Thermography, 2010 10th International Conference on, Québec(Canada), paper QIRT2010126*, 2010.
- [56] Stephen J. Krotosky and Mohan M. Trivedi. Mutual information based registration of multimodal stereo videos for person tracking. *Computer Vision and Image Understanding*, 106(2–3):270 – 287, 2007. Special issue on Advances in Vision Algorithms and Systems beyond the Visible Spectrum.
- [57] Jian Zhao and S. c. S. Cheung. Human segmentation by fusing visible-light and thermal imagery. In *Computer Vision Workshops (ICCV Workshops), 2009 IEEE 12th International Conference on*, pages 1185–1192, Sept 2009.
- [58] S. Vidas, R. Lakemond, S. Denman, C. Fookes, S. Sridharan, and T. Wark. A mask-based approach for the geometric calibration of thermal-infrared cameras. *IEEE Transactions on Instrumentation and Measurement*, 61(6):1625–1635, June 2012.
- [59] Vijay John, Shogo Tsuchizawa, Zheng Liu, and Seiichi Mita. Fusion of thermal and visible cameras for the application of pedestrian detection. *Signal, Image and Video Processing*, 11(3):517–524, Mar 2017.
- [60] P. Saponaro, S. Sorensen, S. Rhein, and C. Kambhamettu. Improving calibration of thermal stereo cameras using heated calibration board. In *2015 IEEE International Conference on Image Processing (ICIP)*, pages 4718–4722, Sep. 2015.
- [61] Ming-Der Yang, Tung-Ching Su, and Hung-Yu Lin. Fusion of infrared thermal image and visible image for 3d thermal model reconstruction using smartphone sensors. *Sensors*, 18:2003, 06 2018.
- [62] Song Han, Xiaojing Gu, and Xingsheng Gu. An accurate calibration method of a multi camera system. In Minrui Fei, Shiwei Ma, Xin Li, Xin Sun, Li Jia, and Zhou Su, editors, *Advanced Computational Methods in Life System Modeling and Simulation*, pages 491–501, Singapore, 2017. Springer Singapore.
- [63] Aravindhnan K Krishnan and Srikanth Saripalli. Cross-calibration of rgb and thermal cameras with a lidar for rgb-depth-thermal mapping. *Unmanned Systems*, 05(02):59–78, 2017.
- [64] A. Ellmauthaler, E. A. B. da Silva, C. L. Pagliari, J. N. Gois, and S. R. Neves. A novel iterative calibration approach for thermal infrared cameras. In *2013 IEEE International Conference on Image Processing*, pages 2182–2186, Sep. 2013.

- [65] Y. Zhang, L. Shen, D. Zhou, and D. Zhang. Calibration of large fov thermal/visible hybrid binocular vision system. In *Proceedings of the 32nd Chinese Control Conference*, pages 5938–5942, July 2013.
- [66] M. Gschwandtner, R. Kwitt, A. Uhl, and W. Pree. Infrared camera calibration for dense depth map construction. In *2011 IEEE Intelligent Vehicles Symposium (IV)*, pages 857–862, June 2011.
- [67] Q.-T. Luong and O.D. Faugeras. On the determination of epipoles using cross-ratios. *Computer Vision and Image Understanding*, 71(1):1 – 18, 1998.
- [68] Jan Klimaszewski, Marcin Kondej, Mariusz Kawecki, and Barbara Putz. Registration of infrared and visible images based on edge extraction and phase correlation approaches. In Ryszard S. Choraś, editor, *Image Processing and Communications Challenges 4*, pages 153–162, Berlin, Heidelberg, 2013. Springer Berlin Heidelberg.
- [69] Barbara Putz, Michał Bartyś, Adrian Antoniewicz, Jan Klimaszewski, Marcin Kondej, and Maciej Wielgus. Real-time image fusion monitoring system: Problems and solutions. In Ryszard S. Choraś, editor, *Image Processing and Communications Challenges 4*, pages 143–152, Berlin, Heidelberg, 2013. Springer Berlin Heidelberg.
- [70] D. Zosso, X. Bresson, and J. Thiran. Geodesic active fields - a geometric framework for image registration. *Image Processing, IEEE Transactions on*, 20(5):1300–1312, May 2011.
- [71] L. Ambrosio, N. Fusco, and D. Pallara. *Functions of Bounded Variation and Free Discontinuity Problems*. Oxford Science Publications. Clarendon Press, 2000.
- [72] B. Cyganek. *An Introduction to 3D Computer Vision Techniques and Algorithms*. John Wiley & Sons, 2007.
- [73] H.B. Mitchell. *Image Fusion: Theories, Techniques and Applications*. Springer Berlin Heidelberg, 2010.
- [74] A. Wong. An adaptive monte carlo approach to phase-based multimodal image registration. *Information Technology in Biomedicine, IEEE Transactions on*, 14(1):173–179, Jan 2010.
- [75] I. Reducindo, E.R. Arce-Santana, D.U. Campos-Delgado, and A. Alba. Evaluation of multimodal medical image registration based on particle filter. In *Electrical Engineering Computing Science and Automatic Control (CCE), 2010 7th International Conference on*, pages 406–411, Sept 2010.

- [76] Tanushri Chakravorty, Guillaume-Alexandre Bilodeau, and Eric Granger. Automatic image registration in infrared-visible videos using polygon vertices. *CoRR*, abs/1403.4232, 2014.
- [77] Atousa Torabi, Guillaume Massé, and Guillaume-Alexandre Bilodeau. An iterative integrated framework for thermal–visible image registration, sensor fusion, and people tracking for video surveillance applications. *Computer Vision and Image Understanding*, 116(2):210 – 221, 2012.
- [78] Camille Sutour, Jean-François Aujol, Charles-Alban Deledalle, and Baudouin Denis Senneville. Edge-based multi-modal registration and application for night vision devices. *Journal of Mathematical Imaging and Vision*, 53(2):131–150, 2015.
- [79] Xiangzeng Liu, Yunfeng Ai, Juli Zhang, and Zhuping Wang. A novel affine and contrast invariant descriptor for infrared and visible image registration. *Remote Sensing*, 10(4):658, Apr 2018.
- [80] X. Liu, Y. Ai, B. Tian, and D. Cao. Robust and fast registration of infrared and visible images for electro-optical pod. *IEEE Transactions on Industrial Electronics*, 66(2):1335–1344, Feb 2019.
- [81] Chunyang Zhao, Huaici Zhao, Jinfeng Lv, Shijie Sun, and Bo Li. Multimodal image matching based on multimodality robust line segment descriptor. *Neurocomputing*, 177:290 – 303, 2016.
- [82] Dong Zhao, Yan Yang, Zhihang Ji, and Xiaopeng Hu. Rapid multimodality registration based on mm-surf. *Neurocomputing*, 131:87 – 97, 2014.
- [83] Q. Du, A. Fan, Y. Ma, F. Fan, J. Huang, and X. Mei. Infrared and visible image registration based on scale-invariant piifd feature and locality preserving matching. *IEEE Access*, 6:64107–64121, 2018.
- [84] C. Cai, X. Ding, and H. Zhu. Registration of infrared and visible image based on opencv. In *2018 IEEE International Conference on Mechatronics and Automation (ICMA)*, pages 2086–2091, Aug 2018.
- [85] C. Wachinger and N. Navab. Structural image representation for image registration. In *Computer Vision and Pattern Recognition Workshops (CVPRW), 2010 IEEE Computer Society Conference on*, pages 23–30, June 2010.
- [86] DavidG. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2):91–110, 2004.

- [87] K. Mikolajczyk and C. Schmid. A performance evaluation of local descriptors. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 27(10):1615–1630, Oct 2005.
- [88] Lisa Gottesfeld Brown. A survey of image registration techniques. *ACM Comput. Surv.*, 24(4):325–376, December 1992.
- [89] E. De Castro and C. Morandi. Registration of translated and rotated images using finite fourier transforms. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, PAMI-9(5):700–703, Sept 1987.
- [90] Xiaofeng Fan, H. Rhody, and E. Saber. A spatial-feature-enhanced mmi algorithm for multimodal airborne image registration. *Geoscience and Remote Sensing, IEEE Transactions on*, 48(6):2580–2589, June 2010.
- [91] P. L. St-Charles, G. A. Bilodeau, and R. Bergevin. Online multimodal video registration based on shape matching. In *Computer Vision and Pattern Recognition Workshops (CVPRW), 2015 IEEE Conference on*, pages 26–34, June 2015.
- [92] Cristhian Aguilera, Fernando Barrera, Felipe Lumbreras, Angel D. Sappa, and Ricardo Toledo. Multispectral image feature points. *Sensors*, 12(9):12661, 2012.
- [93] Le Yu, Dengrong Zhang, and Eun-Jung Holden. A fast and fully automatic registration approach based on point features for multi-source remote-sensing images. *Computers and Geosciences*, 34(7):838 – 848, 2008.
- [94] Haitao Yu, Bingjian Wang, Shangqian Liu, Huixin Zhou, Hanlin Qin, Songqi Yang, and Zhiting Liu. New multimodal image registration method based on moving targets’ trajectory. volume 8907, pages 890745–890745–8, 2013.
- [95] Kristin J. Dana and P. Anandan. Registration of visible and infrared images. In *4/SPIE*, volume 1957, pages 1–13, 1993.
- [96] M. Irani and P. Anandan. Robust multi-sensor image alignment. In *Computer Vision, 1998. Sixth International Conference on*, pages 959–966, Jan 1998.
- [97] Fugen Zhou Weimin Zhang, Risheng Wang. The registration of dual-modality ship target images based on edge extraction. In *International Symposium on Optoelectronic Technology and Application 2014: Image Processing and Pattern Recognition*, volume 9301, 2014.
- [98] M. Hasan, M.R. Pickering, and Xiuping Jia. Robust automatic registration of multimodal satellite images using ccre with partial volume interpolation. *Geoscience and Remote Sensing, IEEE Transactions on*, 50(10):4050–4061, Oct 2012.

- [99] Alexis Roche, Grégoire Malandain, Xavier Pennec, and Nicholas Ayache. The correlation ratio as a new similarity measure for multimodal image registration. In William M. Wells, Alan Colchester, and Scott Delp, editors, *Medical Image Computing and Computer-Assisted Intervention* MICCAI98, volume 1496 of *Lecture Notes in Computer Science*, pages 1115–1124. Springer Berlin Heidelberg, 1998.
- [100] J. Orchard. Efficient least squares multimodal registration with a globally exhaustive alignment search. *Image Processing, IEEE Transactions on*, 16(10):2526–2534, Oct 2007.
- [101] Frank M. Candocia. Jointly registering images in domain and range by piecewise linear comparametric analysis. *Image Processing, IEEE Transactions on*, 12(4):409–419, April 2003.
- [102] Jeff Orchard. Multimodal image registration using floating regressors in the joint intensity scatter plot. *Medical Image Analysis*, 12(4):385–396, 2008.
- [103] Alexis Roche, Grégoire Malandain, and Nicholas Ayache. Unifying Maximum Likelihood Approaches in Medical Image Registration. Technical Report RR-3741, July 1999.
- [104] John Ashburner and Karl J. Friston. Unified segmentation. *NeuroImage*, 26(3):839 – 851, 2005.
- [105] Karl. J. Friston, J. Ashburner, C. D. Frith, J.-B. Poline, J. D. Heather, and R. S. J. Frackowiak. Spatial registration and normalization of images. *Human Brain Mapping*, 3(3):165–189, 1995.
- [106] Wanmei Ou and C. Chédotel. Polynomial intensity correction for multimodal image registration. In *Biomedical Imaging: From Nano to Macro, 2009. ISBI '09. IEEE International Symposium on*, pages 939–942, June 2009.
- [107] P. Viola and W.M.I.I.I. Wells. Alignment by maximization of mutual information. In *Computer Vision, 1995. Proceedings., Fifth International Conference on*, pages 16–23, Jun 1995.
- [108] William M. Wells III, Paul Viola, Hideki Atsumi, Shin Nakajima, and Ron Kikinis. Multi-modal volume registration by maximization of mutual information. *Medical Image Analysis*, 1(1):35 – 51, 1996.
- [109] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens. Multimodality image registration by maximization of mutual information. *Medical Imaging, IEEE Transactions on*, 16(2):187–198, April 1997.

- [110] A. Collignon, F. Maes, D. Delaere, D. Vandermeulen, P. Suetens, and G. Marchal. Automated multi-modality image registration based on information theory. In *Computational Imaging and Vision*, volume 3, pages 263 – 274, 1995.
- [111] Anand Rangarajan, Haili Chui, and James S. Duncan. Rigid point feature registration using mutual information. *Medical Image Analysis*, 3(4):425 – 440, 1999.
- [112] D. Rueckert, C. Hayes, C. Studholme, P. Summers, M. Leach, and D.J. Hawkes. Non-rigid registration of breast mr images using mutual information. In William M. Wells, Alan Colchester, and Scott Delp, editors, *Medical Image Computing and Computer-Assisted Intervention MICCAI 98*, volume 1496 of *Lecture Notes in Computer Science*, pages 1144–1152. Springer Berlin Heidelberg, 1998.
- [113] A.A. Cole-Rhodes, K.L. Johnson, J. LeMoigne, and I. Zavorin. Multiresolution registration of remote sensing imagery by optimization of mutual information using a stochastic gradient. *Image Processing, IEEE Transactions on*, 12(12):1495–1511, Dec 2003.
- [114] Hua-Mei Chen, P.K. Varshney, and M.K. Arora. Performance of mutual information similarity measure for registration of multitemporal remote sensing images. *Geoscience and Remote Sensing, IEEE Transactions on*, 41(11):2445–2454, Nov 2003.
- [115] Renbo Xia, Jibin Zhao, Weijun Liu, and Jianhuang Wu. Robust non-rigid registration of multimodal medical image by combining local phase coherence with residual complexity measure. In *Bioinformatics and Biomedical Engineering (iCBBE), 2010 4th International Conference on*, pages 1–4, June 2010.
- [116] Rui Gan and Albert C.S. Chung. Multi-dimensional mutual information based robust image registration using maximum distance-gradient-magnitude. In Gary E. Christensen and Milan Sonka, editors, *Information Processing in Medical Imaging*, volume 3565 of *Lecture Notes in Computer Science*, pages 210–221. Springer Berlin Heidelberg, 2005.
- [117] Peng Ye and Fang Liu. A implicit similarity method for multi-modal image registration. In *Multi-Platform/Multi-Sensor Remote Sensing and Mapping (M2RSM), 2011 International Workshop on*, pages 1–4, Jan 2011.
- [118] M. Rubeaux, J.-C. Nunes, L. Albera, and M. Garreau. Edgeworth-based approximation of mutual information for medical image registration. In *Image Processing Theory Tools and Applications (IPTA), 2010 2nd International Conference on*, pages 195–200, July 2010.
- [119] M. Zouqi, J. Samarabandu, and Yanbo Zhou. Multi-modal image registration using line features and mutual information. In *Image Processing (ICIP), 2010 17th IEEE International Conference on*, pages 129–132, Sept 2010.

- [120] R. Shams and N. Barnes. Speeding up mutual information computation using nvidia cuda hardware. In *Digital Image Computing Techniques and Applications, 9th Biennial Conference of the Australian Pattern Recognition Society on*, pages 555–560, Dec 2007.
- [121] A. B. Abche, F. Yaacoub, A. Maalouf, and E. Karam. Image registration based on neural network and fourier transform. In *2006 International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 4803–4806, Aug 2006.
- [122] I. Elhanany, M. Sheinfeld, A. Beck, Y. Kadmon, N. Tal, and D. Tirosh. Robust image registration based on feedforward neural networks. In *Systems, Man, and Cybernetics, 2000 IEEE International Conference on*, volume 2, pages 1507–1511 vol.2, 2000.
- [123] Shuxiu Wang, Sheng Lei, and Faliang Chang. Image registration based on neural network. In *2008 International Conference on Information Technology and Applications in Biomedicine*, pages 74–77, May 2008.
- [124] Jianzhen Wu and Jianying Xie. Zernike moment-based image registration scheme utilizing feedforward neural networks. In *Fifth World Congress on Intelligent Control and Automation (IEEE Cat. No.04EX788)*, volume 5, pages 4046–4048 Vol.5, June 2004.
- [125] H. Sarnel, Y. Senol, and D. Sagirlibas. Accurate and robust image registration based on radial basis neural networks. In *2008 23rd International Symposium on Computer and Information Sciences*, pages 1–5, Oct 2008.
- [126] P. Gadde and X. H. Yu. Magnetic resonance imaging registration with artificial neural networks. In *2014 Proceedings of the International Conference on Artificial Intelligence*, pages 459–464, July 2014.
- [127] P. Gadde and X. H. Yu. Image registration with artificial neural networks using spatial and frequency features. In *2016 International Joint Conference on Neural Networks (IJCNN)*, pages 4643–4649, July 2016.
- [128] Alexey Dosovitskiy, Philipp Fischer, Eddy Ilg, Philip Häusser, Caner Hazirbas, Vladimir Golkov, Patrick van der Smagt, Daniel Cremers, and Thomas Brox. FlowNet: Learning optical flow with convolutional networks. *2015 IEEE International Conference on Computer Vision (ICCV)*, pages 2758–2766, 2015.
- [129] Eddy Ilg, Nikolaus Mayer, Tonmoy Saikia, Margret Keuper, Alexey Dosovitskiy, and Thomas Brox. FlowNet 2.0: Evolution of optical flow estimation with deep networks. *CoRR*, abs/1612.01925, 2016.
- [130] Hien Van Nguyen, Kevin Zhou, and Raviteja Vemulapalli. Cross-domain synthesis of medical images using efficient location-sensitive deep network. In Nassir Navab,

- Joachim Hornegger, William M. Wells, and Alejandro Frangi, editors, *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, pages 677–684, Cham, 2015. Springer International Publishing.
- [131] Martin Simonovsky, Benjamín Gutiérrez-Becker, Diana Mateus, Nassir Navab, and Nikos Komodakis. A deep metric for multimodal registration. In Sebastien Ourselin, Leo Joskowicz, Mert R. Sabuncu, Gozde Unal, and William Wells, editors, *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016*, pages 10–18, Cham, 2016. Springer International Publishing.
- [132] Xi Cheng, Li Zhang, and Yefeng Zheng. Deep similarity learning for multimodal medical images. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, 6(3):248–252, 2016.
- [133] Fazl-e-Basit, M. Y. Javed, and U. Qayyum. "face recognition using processed histogram and phase-only correlation (poc)". In *2007 International Conference on Emerging Technologies*, pages 238–242, Nov 2007.
- [134] C. D. Kuglin and D. C. Hines. The phase correlation image alignment method. *IEEE Conference on Cybernetics and Society*, pages 163–165, 1975.
- [135] J. J. Pearson, D. C. Hines Jr., S Golosman, and C. D. Kuglin. Video-rate image correlation processor. volume 0119, pages 197–205, 1977.
- [136] David Casasent and Demetri Psaltis. Position, rotation, and scale invariant optical correlation. *Appl. Opt.*, 15(7):1795–1799, Jul 1976.
- [137] G A. Thomas. Television motion measurement for datv and other applications. *NASA STI/Recon Technical Report N*, 88:13496, 08 1987.
- [138] B.S. Reddy and B.N. Chatterji. An fft-based technique for translation, rotation, and scale-invariant image registration. *Image Processing, IEEE Transactions on*, 5(8):1266–1271, Aug 1996.
- [139] E. De Castro and C. Morandi. Registration of translated and rotated images using finite fourier transforms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-9(5):700–703, Sep. 1987.
- [140] Himanshu Singh, Millie Pant, Sudhir Khare, and Yogita Saklani. Application of unnormalized and phase correlation techniques on infrared images. In Millie Pant, Kanad Ray, Tarun K. Sharma, Sanyog Rawat, and Anirban Bandyopadhyay, editors, *Soft Computing: Theories and Applications*, pages 461–471, Singapore, 2018. Springer Singapore.

- [141] Jignesh Sarvaiya, Suprava Patnaik, and Kajal Kothari. Image registration using log polar transform and phase correlation to recover higher scale. *Journal of Pattern Recognition Research*, 7:90–105, 01 2012.
- [142] L. Hill and T. Vlachos. Motion measurement using shape adaptive phase correlation. *Electronics Letters*, 37(25):1512–1513, Dec 2001.
- [143] M.H.G. Peeters. *Implementation of the phase correlation algorithm: motion estimation in the frequency domain*. TU Eindhoven. Fac. Elektrotechniek: stageverslagen. Technische Universiteit Eindhoven, 2003.
- [144] F. Michielin, G. Calvagno, P. Sartor, and O. Erdler. A true motion estimation method based on binarized cross correlation. In *2013 IEEE Third International Conference on Consumer Electronics & Berlin (ICCE-Berlin)*, pages 1–5, Sep. 2013.
- [145] F. Michielin, G. Calvagno, P. Sartor, T. Emmerich, C. Unruh, and O. Erdler. A parallel true motion estimation method based on binarized cross correlation. In *2014 IEEE International Conference on Image Processing (ICIP)*, pages 1198–1202, Oct 2014.
- [146] Fei Wang, Zhao Zhen, Chun Liu, Zengqiang Mi, Bri-Mathias Hodge, Miadreza Shafie-khah, and João P.S. Catalão. Image phase shift invariance based cloud motion displacement vector calculation method for ultra-short-term solar pv power forecasting. *Energy Conversion and Management*, 157:123 – 135, 2018.
- [147] Y. Ouerhani, M. Jridi, A. Alfalou, and C. Brosseau. Optimized pre-processing input plane gpu implementation of an optical face recognition technique using a segmented phase only composite filter. *Optics Communications*, 289(0):33 – 44, 2013.
- [148] O. Samritjarapon and O. Chitsobhuk. An fft-based technique and best-first search for image registration. In *Communications and Information Technologies, 2008. ISCIT 2008. International Symposium on*, pages 364–367, Oct 2008.
- [149] Dheeraj Singiresu, Sanjeev Agarwal, Shivakar Vulli, and Harish Ramakrishnan. Gpu-based processing for airborne detection. volume 7664, pages 766427–766427–12, 2010.
- [150] Vaidehi. V, Ramya. R, PrasannaDevi. M, Naresh Babu N T, Balamurali P, and Girish Chandra M. Fusion of multi-scale visible and thermal images using emd for improved face recognition. In *Proceedings of the International MultiConference of Engineers and Computer Scientists 2011 Vol I, IMECS2011*, pages 543–548, 2011.
- [151] You Kyung Han and Jae Wan Choi. Matching points extraction between optical and tir images by using surf and local phase correlation. *Journal of Korean Society for Geospatial Information System*, 23(1):81–88, 03 2015.

- [152] M. Miura, K. Fudano, K. Ito, T. Aoki, H. Takizawa, and H. Kobayashi. Gpu implementation of phase-based stereo correspondence and its application. In *Image Processing (ICIP), 2012 19th IEEE International Conference on*, pages 1697–1700, Sept 2012.
- [153] Peter J. Lu, Hidekazu Oki, Catherine A. Frey, Gregory E. Chamitoff, Leroy Chiao, Edward M. Fincke, C. Michael Foale, Sandra H. Magnus, William S. McArthur, Daniel M. Tani, Peggy A. Whitson, Jeffrey N. Williams, William V. Meyer, Ronald J. Sicker, Brion J. Au, Mark Christiansen, Andrew B. Schofield, and David A. Weitz. Orders-of-magnitude performance increases in gpu-accelerated correlation of images from the international space station. *Journal of Real-Time Image Processing*, 5(3):179–193, Sep 2010.
- [154] Mamoru Miura, Kinya Fudano, Koichi Ito, Takafumi Aoki, Hiroyuki Takizawa, and Hiroaki Kobayashi. Performance evaluation of phase-based correspondence matching on gpus. volume 8856, pages 885614–885614–9, 2013.
- [155] M. Hariadi Rosa A. Asmara. Accelerating phase based motion estimation with hierarchical search technique using parallel threading in graphical processing unit(gpu). *International Journal of Computer Science and Network Security*, 9(8):140–146, 2009.
- [156] Sudhakar Sah, Jan Vanek, YoungJun Roh, and Ratul Wasnik. Gpu accelerated real time rotation, scale and translation invariant image registration method. In Aurelio Campilho and Mohamed Kamel, editors, *Image Analysis and Recognition*, volume 7324 of *Lecture Notes in Computer Science*, pages 224–233. Springer Berlin Heidelberg, 2012.
- [157] K. Matsuo, T. Hamada, M. Miyoshi, Y. Shibata, and K. Oguri. Accelerating phase correlation functions using gpu and fpga. In *Adaptive Hardware and Systems, 2009. AHS 2009. NASA/ESA Conference on*, pages 433–438, July 2009.
- [158] D. Skerl, B. Likar, and F. Pernus. A protocol for evaluation of similarity measures for rigid registration. *Medical Imaging, IEEE Transactions on*, 25(6):779–791, June 2006.
- [159] S. Sonn, G. A. Bilodeau, and P. Galinier. Fast and accurate registration of visible and infrared videos. In *Computer Vision and Pattern Recognition Workshops (CVPRW), 2013 IEEE Conference on*, pages 308–313, June 2013.
- [160] G.A. Bilodeau, A. Torabi, and F. Morin. Visible and infrared image registration using trajectories and composite foreground images. *Image and Vision Computing*, 29(1):41 – 50, 2011.
- [161] Yaron Caspi, Denis Simakov, and Michal Irani. Feature-based sequence-to-sequence matching. *Int. J. Comput. Vision*, 68(1):53–64, 2006.

- [162] Tae Eun Choe and I. Cohen. Registration of multimodal fluorescein images sequence of the retina. In *Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on*, volume 1, pages 106–113 Vol. 1, Oct 2005.
- [163] Zheng Liu and R. Laganieri. Registration of ir and eo video sequences based on frame difference. In *Computer and Robot Vision, 2007. CRV '07. Fourth Canadian Conference on*, pages 459–464, May 2007.
- [164] Rafael C. Gonzalez and Richard E. Woods. *Digital image processing*. Prentice Hall, Upper Saddle River, N.J., 2008.
- [165] F. J. Harris. On the use of windows for harmonic analysis with the discrete fourier transform. *Proceedings of the IEEE*, 66(1):51–83, Jan 1978.
- [166] H. Foroosh, J. B. Zerubia, and M. Berthod. Extension of phase correlation to subpixel registration. *IEEE Transactions on Image Processing*, 11(3):188–200, March 2002.
- [167] Canny John. A computational approach to edge detection. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, PAMI-8(6):679–698, November 1986.
- [168] Chris Harris and Mike Stephens. A combined corner and edge detector. In *In Proc. of Fourth Alvey Vision Conference*, pages 147–151, 1988.
- [169] Derpanis K.G. Harris corner detector. In *Computer and In. Science*, volume 2, pages 1–2, 2004.
- [170] Mariusz Kawecki and Barbara Putz. Real-time edge detection using dynamic structuring element. In Tomáš Březina and Ryszard Jabłoński, editors, *Mechatronics 2013*, pages 471–478. Springer International Publishing, 2014.
- [171] James SJ Lee, Robert M Haralick, and Linda G Shapiro. Morphologic edge detection. *Robotics and Automation, IEEE Journal of*, 3(2):142–156, 1987.
- [172] Zhengyou Zhang. Determining the epipolar geometry and its uncertainty: A review. *International Journal of Computer Vision*, 27(2):161–195, Mar 1998.
- [173] P. H. S. Torr and A. Zisserman. Robust parameterization and computation of the trifocal tensor. *Image and Vision Computing*, 15:591–605, 1997.
- [174] Q.-T. Luong and O. D. Faugeras. Camera calibration, scene motion and structure recovery from point correspondences and fundamental matrices. *IJCV*, 22:261–289, 1997.
- [175] S. Hwang, J. Park, N. Kim, Y. Choi, and I. S. Kweon. Multispectral pedestrian detection: Benchmark dataset and baseline. In *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1037–1045, June 2015.

- [176] Barbara Putz, Anna Timofiejczuk, Michał Bartyś, and Jacek Gwardecki. System fuzji obrazów wizyjnych i termowizyjnych dla potrzeb monitorowania otoczenia w czasie rzeczywistym. *Pomiary Automatyka Kontrola*, 57(7):784–788, 2011.