

Politechnika Poznańska
Wydział Elektroniki i Telekomunikacji
Katedra Telekomunikacji Multimedialnej i Mikroelektroniki

Streszczenie rozprawy doktorskiej

Hybrydowe techniki wyznaczania map głębi i ich wykorzystanie w systemach obrazu trójwymiarowego

Hybrid techniques of depth map estimation and their application in
three-dimensional video systems

Maciej Kurc

Promotor: Prof. dr hab. inż. Marek Domański

Poznań, 2019

Spis treści

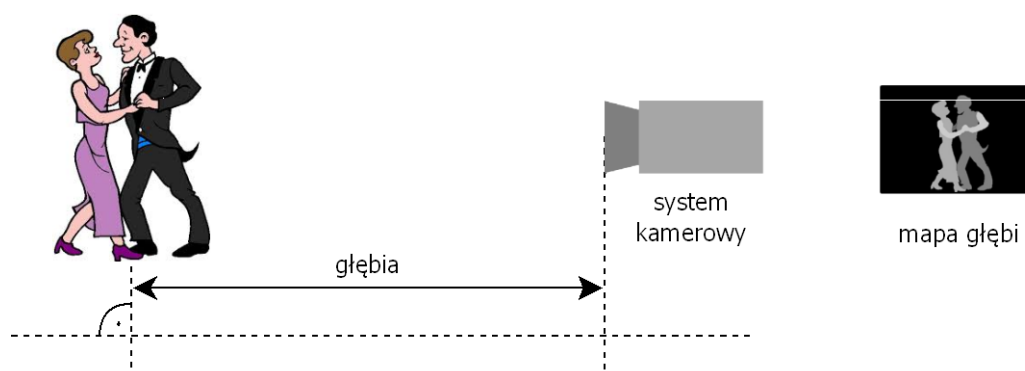
1. Wprowadzenie.....	3
1.1 Zakres rozprawy	3
1.2 Cele i tezy rozprawy	5
1.3 Przegląd pracy	6
2. Najważniejsze osiągnięcia rozprawy.....	10
2.1 Fuzja danych z kamer głębi ToF oraz danych z kamer wizyjnych.....	10
2.2 Poprawa spójności międzywidokowej map głębi.....	11
2.3 Synchronizacja kamer w hybrydowym systemie wielokamerowym.....	13
3. Wyniki eksperymentalne.....	15
3.1 Ocena jakości map głębi poprzez jakość widoku wirtualnego.....	15
3.2 Bezpośrednie porównanie jakości map głębi z danymi odniesienia (ang. ground-truth).....	19
3.3 Ocena poprawy spójności międzywidokowej map głębi	22
3.4 Ocena wpływu spójności międzywidokowej map głębi na efektywność ich kompresji.....	24
4. Podsumowanie.....	27
4.1 Oryginalne osiągnięcia przedstawione w rozprawie	27
4.2 Przeprowadzone prace badawcze	29
4.3 Wnioski	29
Literatura	30
Dorobek naukowy Autora	33

1. Wprowadzenie

1.1 Zakres rozprawy

Rozprawa doktorska porusza zagadnienia związane z **wyznaczaniem głębi przy pomocy kamer głębi, fuzją danych wizyjnych z danymi z kamer głębi oraz kompresją wielowidokowych sekwencji wizyjnych wraz z towarzyszącymi im mapami głębi.**

Nowoczesne, zaawansowane systemy multimedialne bardzo często wykorzystują mapy głębi. Mapa głębi jest to dwuwymiarowa macierz próbek, w której każda z nich reprezentuje odległości między kamerą a danym punktem w przestrzeni trójwymiarowej. Zilustrowane to jest na rysunku 1.1.1.



Rysunek 1.1.1–Ilustracja zagadnienia głębi oraz mapy głębi

Mapa głębi może być wyznaczona na podstawie analizy danych wizyjnych ze stereoskopowej pary kamer bądź z systemu wielokamerowego składającego się z większej liczby kamer. Proces ten nosi nazwę wyznaczania głębi (ang. depth estimation). Bazuje on na poszukiwaniu podobieństw pomiędzy obrazami z dwóch (lub więcej) sąsiadujących ze sobą kamer. Podobieństwa te, w połączeniu z parametrami kamer, pozwalają na wnioskowanie o głębi (odległości). Proces wymaga znajomości parametrów kamer, w szczególności parametrów określających ich położenie.

Alternatywną metodą pozyskiwania map głębi jest bezpośredni pomiar odległości za pomocą techniki aktywnego pomiaru (ang. active sensing). Technikami takimi są metoda oświetlenia strukturalnego (ang. Structured Lightning) [1] [3] lub pomiar czasu propagacji fali świetlnej (ang. Time-of-Flight - ToF) [5]. W rozprawie **autor skupia się na pomiarze czasu propagacji fali świetlnej przy pomocy specjalistycznych kamer głębi ToF.**

Zarówno wyznaczanie jak i akwizycja głębi posiadają swoje wady. Mapy głębi wyznaczane na podstawie analizy danych wizyjnych są często niedokładne. Jest to skutkiem braku charakterystycznych punktów w obrazach lub niemożliwości ich jednoznacznego dopasowania [6] [7]. Wyznaczanie głębi wymaga również znacznego nakładu mocy obliczeniowej. Z tego względu trudno ją stosować w systemach czasu rzeczywistego. Z kolei mapy głębi uzyskane przy pomocy techniki

oświetlenia strukturalnego są bardzo dokładne, lecz sama technika jest powolna. Wymaga ona rejestracji wielu obrazów oraz ich przetworzenia w celu obliczenia mapy głębi. Natomiast technika pomiaru czasu propagacji fali świetlnej pozwala na uzyskanie map głębi o małej rozdzielczości przestrzennej oraz cechujących się wysokim poziomem szumu.

Autor rozprawy proponuje **metodę fuzji zarejestrowanych oraz wyznaczanych map głębi** w celu **uzyskania map głębi o lepszej jakości** niż te jakie można uzyskać stosując każdą z tych technik z osobna. Przez lepszą jakość należy rozumieć większą zgodność reprezentacji odległości z rzeczywistą odległością pomiędzy kamerą a punktami w scenie. Oprócz tej metody autor prezentuje **szereg usprawnień dla algorytmów wyznaczania parametrów kamer** jak i **metody redukcji szumu dla danych z kamer głębi typu ToF**.

Dokładność wyznaczania parametrów kamer, ma kluczowe znaczenie dla wyżej wspomnianej metody fuzji danych jak i dla samego procesu wyznaczania głębi. W rozprawie autor skupia się wyłącznie na wyznaczaniu parametrów dla liniowych systemach wielokamerowych oraz dla systemów hybrydowych, posiadających liniowy układ kamer wizyjnych oraz co najmniej jedną kamerę głębi zorientowaną w tym samym kierunku, co kamery wizyjne. Autor proponuje szereg usprawnień dla algorytmów wyznaczania parametrów kamer między innymi **metodę wyznaczania parametrów zewnętrznych kamer głębi wykorzystującą trójwymiarowe współrzędne punktów obrazu**.

W systemie wielokamerowym, poszczególne kamery muszą być ze sobą zsynchronizowane. W rozprawie analizowane jest **znaczenie problemu synchronizacji dla poprawności wyznaczania głębi**. Ponadto autor przedstawia oryginalną **metodę konwersji sygnału synchronizacyjnego wykorzystywanego przez kamery wizyjne na sygnał wymagany przez kamery głębi**. Prezentowaną metodę można zastosować nie tylko do kamer głębi, ale także dla innych typów kamer przemysłowych. Autor prezentuje też **autorski projekt oraz realizację urządzenia, pozwalającego na dokonanie wspomnianej konwersji sygnałów synchronizacyjnych**.

Niezwykle istotnym zagadnieniem, jakie zostało poruszone w rozprawie, jest **kompresja wielowidokowych sekwencji wizyjnych wraz z towarzyszącymi im mapami głębi**. Algorytmy kompresji sekwencji wielowidokowych, zgodne ze stanem wiedzy, takie jak MVC [29], MV-HEVC [30] i 3D-HEVC [24] silnie polegają na podobieństwach międzywidokowych zarówno w obrazach jak i mapach głębi. Podobieństwa te naturalnie występują w sekwencjach wielowidokowych, ponieważ każdy widok reprezentuje tę samą scenę. Niestety mapy głębi wyznaczone są niezależnie dla poszczególnych widoków przez co ich spójność międzywidokowa jest niska co z kolei przekłada się na niską efektywność kompresji. Autor rozprawy porusza ten problem poprzez **proponując algorytmu poprawy spójności międzywidokowej map głębi**. Celem zaproponowanej metody jest poprawa efektywności kompresji map głębi bez degradacji ich jakości oraz jakości widoków wirtualnych do syntezy których są one wykorzystywane.

1.2 Cele i tezy rozprawy

Głównym rozprawy jest **opracowanie metody fuzji danych z kamer wizyjnych z danymi z kamer głębi**. W założeniu fuzja ta ma na celu umożliwienie uzyskania map głębi o lepszej jakości niż przy użyciu wyłącznie wyznaczania głębi bądź bezpośredniego pomiaru odległości.

W rozprawie przedstawiony jest również szereg celów pobocznych związanych z przetwarzaniem wstępnym danych dla celów fuzji. Cele te dotyczą: **synchronizacji różnych typów kamer, wyznaczania parametrów kamer wizyjnych, wyznaczania parametrów kamer głębi oraz redukcji szumu dla danych z kamer głębi**.

Drugim celem rozprawy jest **poprawa spójności międzywidokowej map głębi**. Algorytmy kompresji sekwencji wielowidokowych polegają na podobieństwach sąsiadujących ze sobą widoków, zarówno co do obrazu jak i głębi. Niestety, mapy głębi wyznaczane są niezależnie dla każdego widoku z osobna, co często czyni je niespójnymi. Autor rozprawy porusza ten problem poprzez propozycję algorytmu przetwarzającego, który pozwala na zwiększenie tej spójności bez konieczności ponownego wyznaczania głębi.

W rozprawie przedstawione są dwie tezy:

- **“Stosunkowo prosty system hybrydowej akwizycji głębi przy użyciu kamer głębi ToF oraz kamer wizyjnych umożliwia pozyskanie map głębi o wyższej jakości porównując do systemów bazujących jedynie na kamerach ToF lub analizie obrazu”**
- **“Poprawa międzywidokowej spójności map głębi, która zwiększa jakość map głębi oraz ich efektywność kompresji przy użyciu algorytmów wykorzystujących międzywidokowe zależności w mapach głębi”**.

1.3 Przegląd pracy

Rozdział 1

Rozdział pierwszy przybliża czytelnikowi tematykę i zakres rozprawy. Wprowadza też podstawowe pojęcia.

Rozdział 2

W rozdziale drugim autor przedstawia swoje oryginalne osiągnięcia dotyczące usprawnienia algorytmów wyznaczania parametrów kamer oraz rektyfikacji obrazów. Modyfikacje te mają swoje zastosowanie dla hybrydowych systemów wielokamerowych złożonych z liniowo rozmieszczonych kamer wizyjnych oraz kamer głębi typu ToF. Są to:

- **Metoda wyznaczania kierunku linii, wzdłuż której rozmieszczone są kamery wizyjne.** Prezentowana metoda nie bazuje na założeniach odnośnie orientacji kamer względem tej linii w przeciwieństwie do rozwiązań znanych z literatury [28], co czyni ją uniwersalną. W połączeniu ze znajomością kierunków osi optycznych kamer pozwala ona na kompensację rotacji, co umożliwia późniejsze wykonanie rektyfikacji obrazów bez ryzyka utraty danych z części pola widzenia kamery.
- **Metoda wyznaczania parametrów zewnętrznych kamer głębi bazująca na trójwymiarowych współrzędnych punktów obrazu.** Proponowana metoda wykorzystuje transformację sztywną (ang. rigid transform), która umożliwia wyznaczenie względnego położenia chmur punktów pochodzących z dwóch różnych kamer głębi. Znajomość takiej transformacji implikuje znajomość zewnętrznych parametrów tych kamer.

Osiągnięcia, przedstawione w rozdziale 2 nie stanowią głównych osiągnięć rozprawy, ale są niezbędne dla przygotowania danych do fuzji obrazu oraz głębi. Z tego względu nie zostały one tak gruntownie przebadane jak sam proces fuzji. Wyniki potwierdzające poprawność usprawnień zaproponowanych przez autora mogą nie być wystarczające, aby stwierdzić, że te usprawnienia są generyczne.

Rozdział 3

W rozdziale trzecim autor prezentuje swoje osiągnięcia dotyczące wstępnego przetwarzania danych z kamer głębi, które jest niezbędnym krokiem w procesie fuzji tych danych z danymi z kamer wizyjnych. W szczególności należy wymienić:

- **Kompensacja pomiaru odległości.** Kamera głębi ToF dokonuje pomiaru odległości wzdłuż drogi propagacji fali świetlnej. Odległość, jaką reprezentuje mapa głębi, jest mierzona wzdłuż osi optycznej kamery. Zmiana reprezentacji z jednej na drugą wymaga znajomości parametrów wewnętrznych kamery głębi. Dostępne na rynku kamery głębi posiadają wbudowaną możliwość wykonania takiej transformacji, niestety przy użyciu nieznanego zestawu parametrów. Ponieważ

transformacja ta powinna być przeprowadzona przy użyciu tych samych parametrów, jakie zostaną użyte w procesie fuzji danych, autor zaproponował alternatywną metodę. Proponowana metoda korzysta z parametrów wewnętrznych wyznaczonych przez autora, co czyni ją zgodną z pozostałą częścią procesu fuzji danych [35].

- **Kompensacja systematycznego błędu pomiaru odległości.** Kamera głębi ToF, jak każde urządzenie pomiarowe, wprowadza pewien błąd systematyczny do wyników pomiarów. Autor zaproponował metodę kompensacji tego błędu przy użyciu identycznego, dwuwymiarowego wzorca kalibracyjnego jak stosowany podczas wyznaczania parametrów kamer. Trójwymiarowe współrzędne cech wzorca, wyznaczone przy użyciu parametrów kamery oraz matematycznego modelu kamery otworkowej [22] [23] [21] są porównywane z wartościami zmierzonymi przez kamerę ToF. Pozwala to na wyznaczenie transformacji niwelującej wspomniany błąd systematyczny.
- **Autorski algorytm redukcji szumu dla danych z kamer głębi ToF.** Szum ten można podzielić na dwie kategorie: szum przestrzenny (ang. spatial noise) oraz szum czasowy (ang. temporal noise). W rozprawie autor proponuje metodę redukcji szumu czasowego poprzez filtrację czasową wybranych punktów za pomocą filtru IIR pierwszego rzędu. Wybór punktów dokonywany jest na podstawie zmiany ich luminancji (nie odległości) w czasie. Punkty wykazujące wysoką zmienność luminancji nie są filtrowane w przeciwieństwie do pozostałych. Ma to na celu uniknięcie efektu rozmycia poruszających się obiektów. Proponowana metoda redukcji szumu przestrzennego bazuje na dwuwymiarowym, bilateralnym filtrze o skończonej odpowiedzi impulsowej (FIR), w którym wagi maski zależą od luminancji powiązanej z filtrowanymi punktami. Wagi maski filtru dobierane są w taki sposób, aby filtr pozwalał na wygładzenie ciągłych obszarów mapy odległości przy jednoczesnym zachowaniu ostrych krawędzi.

Rozdział 4

Rozdział czwarty poświęcony jest w całości zagadnieniu synchronizacji czasowej w systemach wielokamerowych. Autor przedstawia istotność synchronizacji dla procesu wyznaczania map głębi. Szczególny nacisk położony został na zagadnienie synchronizacji kamer różnego typu - profesjonalnych kamer telewizyjnych oraz przemysłowych kamer głębi ToF. Oba rodzaje urządzeń wymagają sygnałów synchronizacyjnych różnego rodzaju.

Rozdział 5

Rozdział piąty został w całości poświęcony głównemu zagadnieniu rozprawy, jakim jest fuzja danych z kamer głębi oraz danych z kamer wizyjnych.

Autor opisuje swoją propozycję modyfikacji algorytmu wyznaczania głębi zgodnego ze stanem wiedzy. Algorytm ten bazuje na minimalizacji energii za pomocą metody globalnej optymalizacji

takiej jak propagacja wierzeń (ang. Belief Propagation) [18] lub cięcie grafów (ang. Graph Cuts) [19]. Propozycja autora polega na modyfikacji funkcji kosztu podlegającej minimalizacji w taki sposób, aby odzwierciedlała zarówno zależności między danymi z kamer wizyjnych jak i z kamer głębi.

W rozdziale piątym autor prezentuje także swoją propozycję algorytmu fuzji danych z wielu kamer głębi. Zaproponowane podejście pozwala uzyskać szerszy zakres przestrzenny reprezentacji sceny przy użyciu danych z kilku kamer głębi oraz ich parametrów.

W końcowej części rozdziału zawarty jest opis eksperymentów oraz ich wyników mających kluczowe znaczenie dla udowodnienia pierwszej tezy rozprawy. Autor prezentuje wyniki porównania jakości map głębi wyznaczonych za pomocą algorytmów zgodnych ze stanem wiedzy oraz tych otrzymanych przy użyciu wprowadzonych przez niego modyfikacji. Przedstawione zostało porównanie przez obiektywną miarę jakości widoku wirtualnego, zgodnie z zaleceniami grupy MPEG [16], oraz przez porównanie z danymi referencyjnymi (ang. ground-truth) [4].

Rozdział 6

Szósty rozdział rozprawy jest poświęcony zaproponowanej przez autora metodzie przetwarzania map głębi, która pozwala zwiększyć ich spójność międzywidokową. Spójność międzywidokowa odgrywa kluczową rolę dla algorytmów kompresji sekwencji wielowidokowych takich jak 3D-HEVC [24] oraz 3D-AVC [25], które to wykorzystują ją do usunięcia informacji nadmiarowej z reprezentacji trójwymiarowej sceny. Autor prezentuje:

- **Autorską miarę niespójności międzywidokowej map głębi.** Proponowana miara bazuje na wariancji wartości głębi obliczanej na zestawach odpowiadających sobie punktów w całym zestawie widoków. Jest ona wykorzystywana w prezentowanym algorytmie poprawy spójności jak kryterium oceny podczas jego działania [34].
- **Algorytm zwiększający spójność międzywidokową map głębi.** Algorytm, proponowany przez autora, bazuje na iteracyjnej wymianie informacji pomiędzy widokami. Wymiana następuje pomiędzy odpowiadającym sobie punktom z różnych widoków, które to dopasowywane są do siebie poprzez projekcje między przestrzeniami poszczególnych widoków [9]. Za kryterium zakończenia działania algorytmu uznawany jest brak wzrostu spójności międzywidokowej określany zgodnie z jego autorską miarą [34].
- Opis oraz wyniki eksperymentów potwierdzających redukcję prędkości bitowej skompresowanego strumienia danych spowodowaną wzrostem spójności międzywidokowej kodowanych map głębi. W eksperymencie zastosowane zostały algorytmy kompresji zgodne ze stanem wiedzy takie jak 3D-HEVC [24] i 3D-AVC [25] oraz starsze techniki takie jak MVC+D [26]. Eksperymenty zostały przeprowadzone na reprezentatywnym zestawie wielowidokowych sekwencji wizyjnych wraz z mapami głębi [8]. Autor tej rozprawy jest jednocześnie

współautorem kilku tych sekwencji [52]. Redukcję prędkości bitowej przy jednoczesnym braku pogorszenia jakości zrekonstruowanego obrazu wyznaczono przy użyciu miary Bjontegaarda [20]. Wyniki eksperymentalne posłużyły autorowi do potwierdzenia drugiej tezy rozprawy doktorskiej.

Rozdział 7

Rozdział siódmy podsumowuje rozprawę poprzez wskazanie wszystkich oryginalnych osiągnięć autora wraz z możliwymi ścieżkami dalszych badań w przyszłości.

Aneks A

W pierwszym aneksie do tekstu rozprawy zaprezentowane zostało autorskie urządzenie sprzętowe, które pozwala na tłumaczenie sygnału synchronizacji pomiędzy dwoma odmiennymi standardami.

Kamery wizyjne, użyte przez autora w badaniach, synchronizowane są sygnałem Genlock. Sygnał taki jest ciągły i składa się z ciągu pustych ramek zakodowanych analogowo przy użyciu odpowiedniego standardu (np. SMPTE 274M [27] dla obrazu w rozdzielczości Full HD). Kamery głębi ToF są urządzeniami przemysłowymi i wymagają ciągu pojedynczych impulsów wyzwalających.

Prezentowane urządzenie dokonuje analizy wejściowego sygnału Genlock w czasie rzeczywistym i na jego podstawie generuje ciąg impulsów dla kamer głębi. Pozwala ono również na zadanie przesunięcia fazowego pomiędzy momentami akwizycji obrazu przez kamery wizyjne oraz kamery głębi.

Aneks B

Aneks B zawiera wykresy krzywych prędkość bitowa – zniekształcenie (ang. rate-distortion) dla wszystkich eksperymentów kompresji map głębi dotyczących poprawy ich spójności międzywidokowej.

Aneks C

W aneksie C zebrane są przykładowe kadry z sekwencji wizyjnych nagranych przez autora przy użyciu systemu wielokamerowego jego konstrukcji. Użyty system wielokamerowy zawiera szereg rozmieszczonych liniowo kamer wizyjnych oraz jedną lub więcej kamer głębi.

Aneks D

Aneks D zawiera szczegółowe wyniki liczbowe dla eksperymentów związanych z fuzją danych wizyjnych oraz map głębi z kamer ToF. Wyniki przedstawiają wartości miar jakości dla map głębi wyznaczonych przy użyciu algorytmów zgodnych ze stanem wiedzy oraz przy użyciu algorytmów zmodyfikowanych przez autora.

2. Najważniejsze osiągnięcia rozprawy

2.1 Fuzja danych z kamer głębi ToF oraz danych z kamer wizyjnych

Celem fuzji danych z kamer wizyjnych oraz danych z kamer głębi jest **uzyskanie lepszej jakości map głębi, niż jest to możliwe przy użyciu wyłącznie jednego z tych dwóch rodzajów danych.**

Autor rozprawy proponuje oryginalną metodę fuzji danych wizyjnych z danymi z kamer głębi ToF. Proponowana metoda bazuje na modyfikacji funkcji hipotetycznej energii, podlegającej minimalizacji w algorytmie cięcia grafów (ang. Graph Cuts) [19]. Umożliwia to wprowadzenie do algorytmu zarówno danych będących wynikiem analizy obrazu jak i danych z kamer głębi ToF.

Funkcja hipotetycznej energii dana jest wzorem:

$$E(f) = \sum_{\forall p_i \in P} D(p_i, f(p_i)) + \sum_{\forall p_i, p_j \in Q} V(p_i, p_j, f(p_i), f(p_j)), \quad (2.1.1)$$

gdzie $E(f)$ jest funkcją energii, P jest zbiorem wszystkich punktów a Q jest zbiorem wszystkich par punktów. Funkcja f definiuje zależność pomiędzy fizyczną wartością odległości a wartością głębi. Funkcja D definiuje koszt posiadania wartości głębi $f(p_i)$ dla punktu p_i a funkcja V definiuje koszt posiadania wartości głębi $f(p_i)$ i $f(p_j)$ przez parę punktów p_i oraz p_j .

Proponowane przez autora modele kosztu dopasowania punktu dla danych z kamer głębi dane są następującymi wzorami. Są to odpowiednio modele: liniowy, kwadratowy, gaussowski oraz skokowy.

$$G(d) = a \cdot |d - d_0|, \quad (2.1.2)$$

$$G(d) = a \cdot (d - d_0)^2, \quad (2.1.3)$$

$$G(d) = 1 - e^{\frac{-(d-d_0)^2}{2\sigma^2}}, \quad (2.1.4)$$

$$G(d) = \begin{cases} 0 & \text{dla} \\ 1 & \text{w przeciwnym razie} \end{cases} \quad |d - d_0| < b, \quad (2.1.5)$$

gdzie $G(d)$ jest funkcją modelującą koszt posiadania przez dany punkt wartości głębi równej d , d_0 jest wartością głębi zmierzoną przez kamerę ToF natomiast a , b oraz σ są parametrami tych modeli. Proponowane, metody integracji kosztu z kosztem wynikającym z analizy obrazu dane są wzorami:

$$D(d) = c \cdot G(d) + (1 - c) \cdot D_{SAD}(d), \quad (2.1.6)$$

$$D(d) = c \cdot D_{SAD}(d) \cdot G(d) + (1 - c) \cdot D_{SAD}(d), \quad (2.1.7)$$

gdzie $D(d)$ jest końcową wartością kosztu przypisania wartości głębi d dla danego punktu, $D_{sad}(d)$ jest wartością kosztu wynikającego z analizy obrazów, $G(d)$ jest funkcją modelującą koszt dla danych z kamery ToF a c jest znormalizowaną miarą pewności pomiaru odległości przez kamerę ToF.

Ponadto, autor proponuje także sposób modyfikacji kosztu przypisania głębi dla pary punktów, która to ma na celu zwiększenie istotności krawędzi obrazu w procesie wyznaczania głębi. Proponowana, zmodyfikowana funkcja kosztu V dana jest wzorem:

$$V(p_i, p_j, f(p_i), f(p_j)) = \lambda \cdot |f(p_i) - f(p_j)|, \quad (2.1.8)$$

gdzie λ jest współczynnikiem gładkości. Pozostałe oznaczenia są identyczne jak we wzorze 2.1.1.

Zaproponowana przez autora metoda została zaimplementowana przez niego w oprogramowaniu DERS [15]. Oprogramowanie, w oryginalnej wersji, zawiera w sobie implementację algorytmu wyznaczania głębi zgodną ze stanem wiedzy.

2.2 Poprawa spójności międzywidokowej map głębi

Algorytmy wyznaczające mapy głębi zgodne ze stanem wiedzy, wykorzystują w tym celu zestaw dwóch lub większej liczby obrazów. Jednakże dla każdej kamery wykorzystywany jest inny ich zestaw. Skutkuje to tym, że mapy głębi wyznaczane są na podstawie potencjalnie różnych danych, pomimo tego, że te dane przedstawiają tę samą scenę. Powoduje to powstanie niespójności międzywidokowych pomiędzy tak wyznaczonymi mapami głębi.

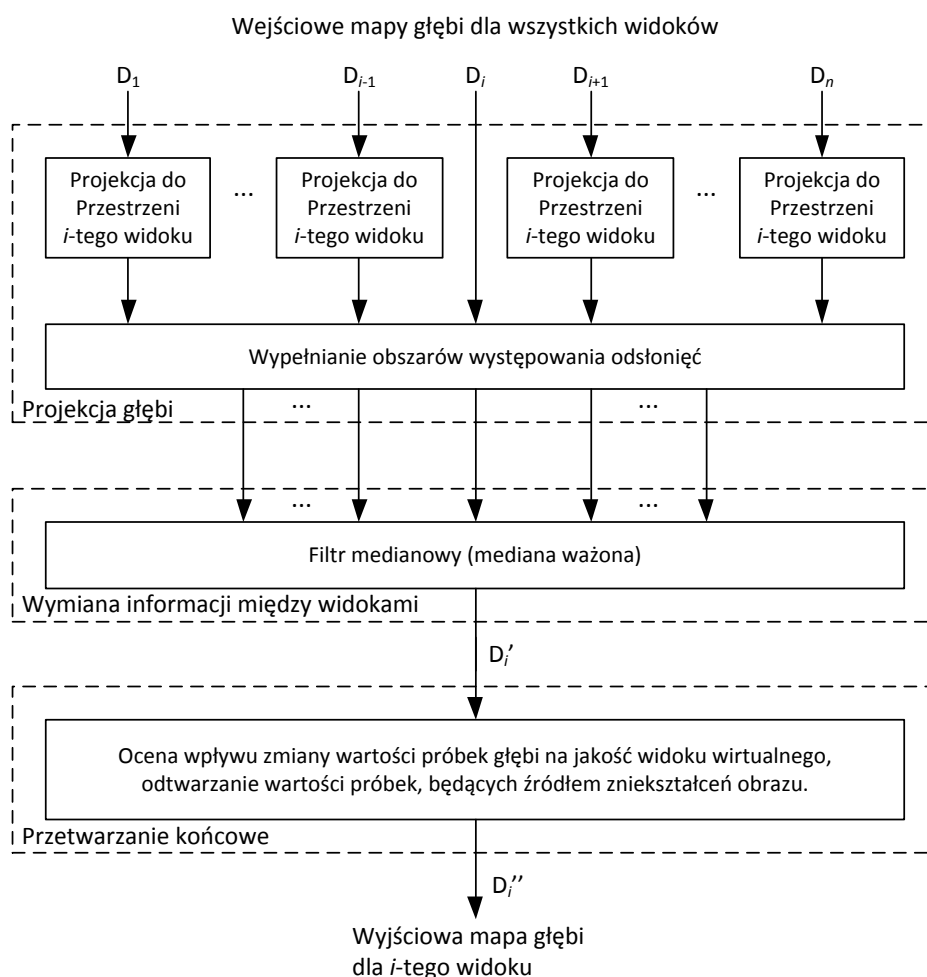
Przykład niespójności międzywidokowej map głębi można zaobserwować na rysunku 2.2.1. Jest to zestaw trzech map głębi, dla różnych widoków, wyznaczonych dla sekwencji „Newspaper” ze zbioru sekwencji reprezentatywnych [8].



Rysunek 2.2.1 – Przykład niespójności międzywidokowej map głębi. Obszary występowania niespójności zaznaczone zostały kolorem czerwonym.

Autor rozprawy proponuje oryginalny algorytm poprawy spójności międzywidokowej map głębi [34]. Jest to algorytm iteracyjny, podczas działania którego informacje ze wszystkich dostępnych map głębi dla danej sekwencji są ze sobą wymieniane.

Szczegółowy schemat blokowy przepływu danych dla pojedynczej iteracji proponowanego algorytmu jest przedstawiony na rysunku 2.2.2.



Rysunek 2.2.2 – Schemat blokowy przepływu danych podczas pojedynczej iteracji w proponowanym algorytmie poprawy spójności międzywidokowej map głębi.

Podczas każdej iteracji algorytmu wejściowe mapy głębi $D_1 - D_n$ są transformowane do przestrzeni i -tego widoku. Obszary występowania odsłonieć są wypełniane. Następnie każda próbka filtrowana jest za pomocą ważonego filtru medianowego, działającego na przestrzeni widoków. W rezultacie otrzymujemy nową mapę głębi D_i' . Proces ten jest powtarzany dla każdego widoku dla i przyjmującego wartości od 1 do N , gdzie N to liczba widoków.

Nowa mapa głębi D_i' poddawana jest obiektywnej ocenie jakości poprzez wygenerowanie przy jej użyciu widoku wirtualnego a następnie porównanie go z odpowiadającym mu widokiem rzeczywistym. Porównanie odbywa się dla każdego punktu, wraz z jego sąsiedztwem, niezależnie przy użyciu miary SSIM [11]. Dla obszarów obrazu, w których porównanie wykazuje zbyt dużą różnicę w stosunku do obrazu odniesienia, przywracana jest wartość głębi z poprzedniej iteracji.

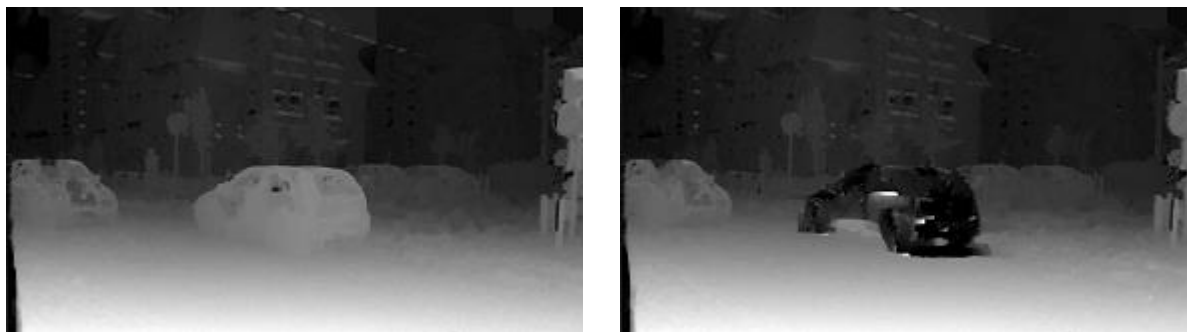
Zaproponowany przez autora algorytm stał się częścią technologii kompresji wielowidokowych sekwencji wizyjnych oraz map głębi zgłoszonej przez Politechnikę Poznańską

do konkursu zorganizowanego przez grupę MPEG w 2011 roku [36] [42]. Prezentowana metoda pozwala zwiększyć efektywność wielowidokowej kompresji map głębi o ponad 5%.

2.3 Synchronizacja kamer w hybrydowym systemie wielokamerowym

W systemie wielokamerowym, wszystkie kamery wizyjne muszą pracować synchronicznie. Wymusza to na nich akwizycję kolejnych ramek z tą samą częstotliwością oraz w tych samych chwilach czasowych. Zagadnienie to jest bardzo istotne z punktu widzenia procesu estymacji głębi.

Mapa głębi, wyznaczona przy użyciu niesynchronicznych sekwencji wizyjnych, będzie zawierała zafałszowane wartości próbek w obszarach, w których występuje poprzeczny ruch. Przykład mapy głębi, wyznaczonej dla sceny z poruszającym się pojazdem, dla sekwencji synchronicznej i niesynchronicznej jest przedstawiony na rysunku 2.3.1.



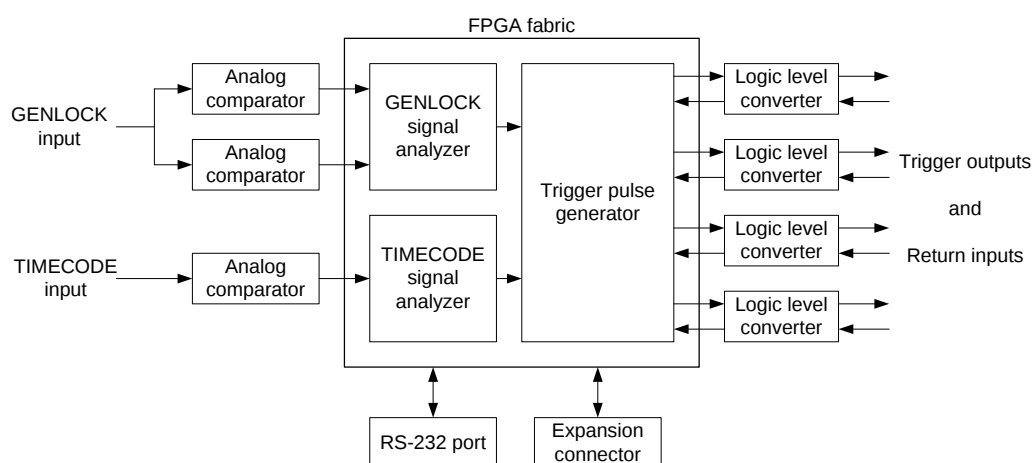
Rysunek 2.3.1 – Mapa głębi wyznaczona dla synchronicznego systemu wielokamerowego (lewy obraz) oraz niesynchronicznego (prawy obraz).

Zastosowanie w systemie wielokamerowym kamer głębi wymusza ich synchronizację z kamerami wizyjnymi. Niestety, bardzo często różne rodzaje kamer wymagają różnego rodzaju sygnałów synchronizacyjnych. Dla przykładu: profesjonalne kamery telewizyjne są synchronizowane poprzez sygnał Genlock, który zawiera impulsy synchronizacji pionowej dla każdej linii obrazu oraz poziomej dla całych ramek. Natomiast kamery głębi ToF wymagają pojedynczych impulsów synchronizacyjnych, które inicjują proces akwizycji mapy głębi. **Synchronizacja tych dwóch rodzajów kamer wymaga konwersji pomiędzy różnymi rodzajami sygnałów synchronizacji.** Na rynku brak jest urządzeń realizujących to zadanie. W literaturze również występuje niedostatek publikacji poświęconych temu zagadnieniu [17].

Z tego względu autor rozprawy zaproponował oryginalną metodę konwersji sygnału synchronizacyjnego Genlock na ciąg impulsów elektrycznych wyzwalających jedną lub więcej kamer głębi. Proponowane przez autora rozwiązanie polega na analizie sygnału Genlock w czasie rzeczywistym i generacji impulsów wyzwalających. Impulsy te generowane są z zadany przesunięciem fazowym w stosunku do impulsów wyznaczających granice ramek w sygnale Genlock. Pozwala to na dokładne zsynchronizowanie momentów akwizycji obrazu oraz głębi.

Autor nie ograniczył się wyłącznie do propozycji metody, ale również zaprojektował i wykonał prototyp urządzenia sprzętowego, które ją implementuje. Urządzenie działa w oparciu o układ FPGA. Posłużyło ono autorowi do konstrukcji hybrydowego systemu wielokamerowego, przy pomocy którego zarejestrowane zostały sekwencje użyte w ocenie zaproponowanej przez niego metody fuzji danych.

Na rysunku 2.3.2 przedstawiony jest schemat blokowy proponowanego konwertera sygnałów synchronizacyjnych, natomiast na rysunku 2.3.3 przedstawione jest zdjęcie wykonanego prototypu podczas pracy.



Rysunek 2.3.2 – Schemat blokowy konwertera sygnałów synchronizacji kamer

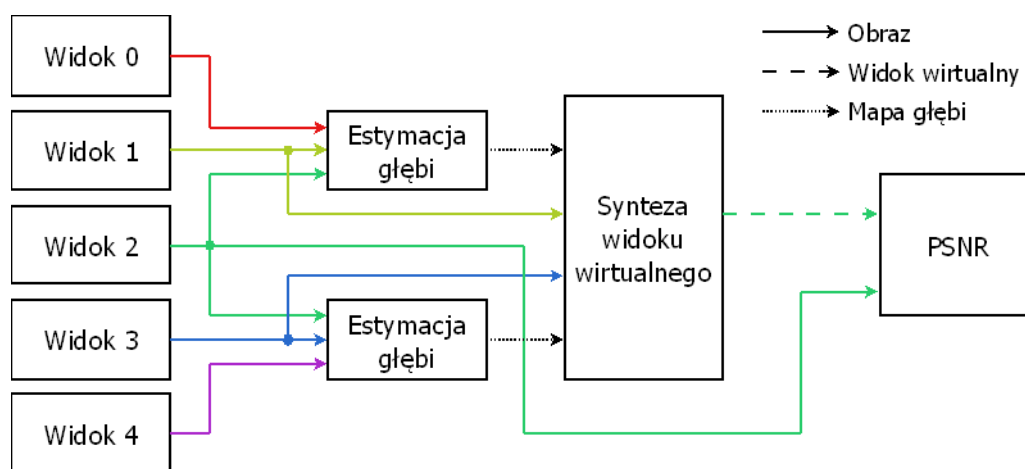


Rysunek 2.3.3 – Zdjęcie prototypu konwertera sygnałów synchronizacyjnych.

3. Wyniki eksperymentalne

3.1 Ocena jakości map głębi poprzez jakość widoku wirtualnego

W pracy, podobnie jak w literaturze, jakość mapy głębi oceniona została poprzez ocenę jakości widoku wirtualnego. Metoda ta polega na wygenerowaniu widoku wirtualnego - obrazu odpowiadającego wirtualnej kamerze znajdującej się w identycznej lokalizacji przestrzennej, co wybrana kamera rzeczywista. Służy do tego technika DIBR (ang. Depth Image Based Rendering) [9]. Tak wygenerowany widok wirtualny jest porównywany z obrazem odniesienia, pochodzącym z rzeczywistej kamery, za pomocą obiektywnej miary podobieństwa obrazów. Najczęściej stosowanymi miarami są PSNR [10] oraz SSIM [11]. Ocena przy pomocy miary PSNR jest częściej spotykana w opracowaniach w literaturze. Na rysunku 3.1.1 przedstawiony jest schemat działania tej metody.



Rysunek 3.1.1 – Schemat blokowy przepływu informacji w metodzie oceny jakości map głębi przy użyciu syntezy widoku wirtualnego. Kolory linii oznaczają poszczególne widoki.

Na potrzeby udowodnienia tez rozprawy autor dokonał rejestracji wielu sekwencji wielowidokowych oraz map głębi. W tym celu wykorzystany został system wielokamerowy, którego konstrukcja jest w istotnej części dziełem autora. System ten został opracowany w laboratoriach Katedry Telekomunikacji Multimedialnej i Mikroelektroniki wydziału Elektroniki i Telekomunikacji Politechniki Poznańskiej. W skład systemu wchodzi kamery wizyjne Canon XH-G1 [31] oraz kamery głębi ToF SR4000 [32].

Autor przygotował pięć sekwencji wielowidokowych wraz z mapami głębi. Na rysunku 3.1.2 przedstawione są wybrane kadry z tych sekwencji.



Boards_01



Boards_02



Boards_03



Office



Flower

Rysunek 3.1.2– Wybrane kadry z sekwencji zarejestrowanych przez autora rozprawy.

Dla każdej sekwencji wielowidokowej, wyznaczony został szereg zestawów map głębi. W tym celu wykorzystane zostało oprogramowanie DERS (ang. Depth Estimation Reference Software) [15]. Oprogramowanie zawierało modyfikacje realizujące implementację algorytm fuzji danych wizyjnych oraz map głębi zaproponowanego przez autora rozprawy. Synteza widoków wirtualnych przeprowadzona została przy pomocy algorytmów zaimplementowanych w oprogramowaniu VSRS (ang. View Synthesis Reference Software) [VSRS].

Każdy zestaw map głębi wyznaczony został przy użyciu innego modelu kosztu dla danych z kamery głębi ToF oraz innego sposobu integracji tego kosztu z kosztem obliczonym na podstawie analizy obrazu. Dodatkowo wzięte pod uwagę zostało uwzględnienie bądź nieuwzględnienie proponowanej modyfikacji funkcji kosztu V dla par punktów. Warianty ważenia kosztu dopasowania dla indywidualnych punktów oraz modyfikacji kosztu dopasowania dla par punktów podsumowane są w tabeli 3.1.1.

Tabela 3.1.1. – Podsumowanie wariantów modyfikacji funkcji kosztu.

Wariant	Opis wariantu
Wariant 1	• Brak modyfikacji funkcji dopasowania par punktów • Ważenie kosztu dopasowania indywidualnych punktów za pomocą równania 2.1.6
Wariant 2	• Brak modyfikacji funkcji dopasowania par punktów • Ważenie kosztu dopasowania indywidualnych punktów za pomocą równania 2.1.7
Wariant 3	• Zmodyfikowana funkcja dopasowania par punktów, uwzględniająca krawędzie obrazów • Ważenie kosztu dopasowania indywidualnych punktów za pomocą równania 2.1.6
Wariant 4	• Zmodyfikowana funkcja dopasowania par punktów, uwzględniająca krawędzie obrazów • Ważenie kosztu dopasowania indywidualnych punktów za pomocą równania 2.1.7

W tabeli 3.1.2 przedstawione zostały uśrednione wartości miary PSNR dla widoków wirtualnych wygenerowanych na podstawie map głębi wyznaczonych proponowaną metodą przy użyciu sekwencji wielowidokowych zarejestrowanych przez autora.

Tabela 3.1.2. – Średnie wartości miary PSNR dla widoków wirtualnych dla sekwencji zarejestrowanych przez autora.

	Wyłącznie analiza obrazu	Proponowany algorytm											
		Model skokowy			Model liniowy			Model kwadratowy			Model Gaussowski		
		b=±1	b=±2	b=±4	a= $\frac{1}{5}$	a= $\frac{1}{12.5}$	a= $\frac{1}{20}$	a= $\frac{1}{25}$	a= $\frac{1}{150}$	a= $\frac{1}{400}$	σ=3	σ=6	σ=10
Wariant 1	27.75	25.90	25.87	26.00	25.75	25.77	25.80	25.80	25.89	25.93	25.78	25.86	25.90
Wariant 2	27.75	27.63	27.61	27.59	27.61	27.61	27.60	27.62	27.60	27.64	27.61	27.59	27.64
Wariant 3	27.70	25.90	25.86	25.99	25.73	25.75	25.80	25.76	25.90	25.99	25.76	25.85	25.93
Wariant 4	27.70	27.61	27.63	27.59	27.63	27.66	27.60	27.64	27.59	27.57	27.65	27.60	27.57

Niestety żaden z eksperymentów nie pozwolił na uzyskanie poprawy jakości widoku wirtualnego w sensie jakim oddaje ją miara PSNR. Jest ku temu wiele powodów. Najważniejszym z nich jest niska korelacja wartości PSNR z subiektywną oceną jakości widoków wirtualnych o czym traktuje literatura [12] [13]. Miara PSNR jest również bardzo czuła na przesunięcie punktów w obrazie a przesuwanie punktów jest podstawą funkcjonowania techniki DIBR.

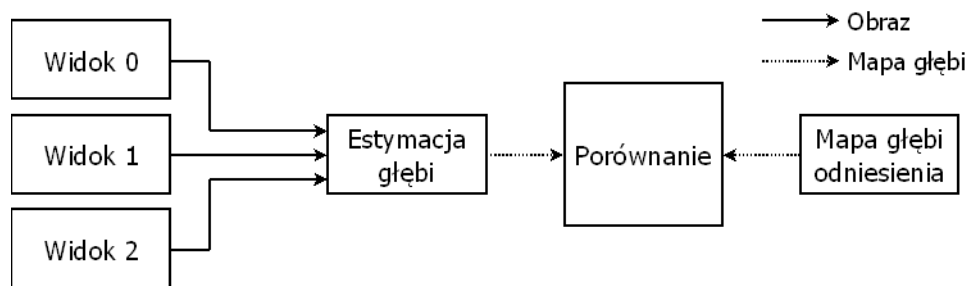
Metoda oceny jakości map głębi poprzez ocenę jakości widoku wirtualnego posiada istotne ograniczenia. Głównym z nich jest niska korelacja miar takich jak PSNR czy SSIM z subiektywną jakością takiego obrazu. Problem ten występuje szczególnie, gdy porównywane są obrazy wirtualne wygenerowane przy użyciu map głębi wyznaczonych w różny sposób. Zniekształcenia występujące w nich mają inny charakter niż te dla porównywania których stworzone zostały te miary.

Problem ten poruszany jest w literaturze: Tikanmaki i inni [12] opisuje eksperymenty dotyczące podziału prędkości bitowej pomiędzy poszczególne elementy składowe strumienia danych podczas kompresji wielowidokowych sekwencji wizyjnych wraz z mapami głębi. Obiektywna ocena jakości zrekonstruowanych sekwencji wizyjnych została zestawiona z oceną subiektywną. Wnioski z eksperymentów potwierdzają słabą korelację miar obiektywnych oraz subiektywnych w tym zastosowaniu. Z kolei Banitalebi-Dehkordi i inni [13] przedstawia w publikacji przegląd miar jakości stereoskopowych sekwencji wizyjnych. Autorzy wprowadzają także nową obiektywną miarę jakości. Wnioski z przedstawionych porównań sugerują, że korelacja pomiędzy miarami obiektywnymi a subiektywnymi jest niewystarczająca, aby te pierwsze można zastosować dla oceny jakości widoku wirtualnego.

Dodatkowo, należy tutaj nadmienić, że **mapa głębi, wyznaczana w procesie analizy obrazu, niekoniecznie odzwierciedla fizyczną odległość**. Dzieje się tak gdyż proces dopasowywania do siebie fragmentów dwóch lub więcej obrazów jest wrażliwy na występowanie podobnych do siebie struktur, które nie odpowiadają tym samym elementom filmowanej sceny. Połączenie potencjalnie zafałszowanych danych o odległości z pomiarami rzeczywistej odległości, jakie dostarcza kamera ToF, może skutkować pogorszeniem jakości finalnej mapy głębi.

3.2 Bezpośrednie porównanie jakości map głębi z danymi odniesienia (ang. ground-truth)

Technika ta polega na bezpośrednim porównaniu wartości odpowiadających sobie próbek mapy głębi z próbkami z mapy głębi odniesienia za pomocą jednej lub wielu obiektywnych miar. Autor rozprawy postąpił według metodologii opisanej w literaturze, jako że jest ona szeroko rozpoznawalna, jako ogólna metoda oceny jakości algorytmów wyznaczania głębi [4]. Na rysunku 3.2.1 przedstawiony jest schemat przepływu danych we wspomnianej metodzie.



Rysunek 3.2.1 – Schemat przepływu danych w metodzie bezpośredniego porównania mapy głębi z mapą głębi odniesienia.

Wadą techniki bezpośredniego porównania map głębi jest konieczność posiadania mapy głębi odniesienia, której autor nie był w stanie zapewnić dla stworzonego przez siebie materiału wizyjnego z przyczyn technicznych. Z tego względu autor postanowił użyć materiału pochodzącego z bazy danych Middlebury, który zawiera dane wizyjne jak i mapy głębi odniesienia [2][3][4][14]. Ponieważ baza danych Middlebury nie zawiera danych pochodzących z kamer głębi ToF, ich obecność została zasymulowana poprzez sztuczne ograniczenie rozdzielczości przestrzennej wybranych map głębi odniesienia. **Należy tutaj zaznaczyć, że pozycje przestrzenne symulowanych kamer głębi nie są tożsame z pozycjami kamer wizyjnych.** Odpowiada to realistycznym warunkom ich rozmieszczenia. **Przez to zagadnienie fuzji danych nie zostaje sprowadzone do zagadnienia zwiększenia rozdzielczości przestrzennej mapy głębi.**

Na potrzeby eksperymentów autor wykorzystał następujące zestawy danych z bazy danych Middlebury: Aloe, Art, Baby1, Baby2, Baby3, Books, Bowling1, Bowling2, Cloth1, Cloth2, Cloth3, Cloth4, Dolls, Flowerpots, Lampshade1, Lampshade2, Laundry, Midd1, Midd2, Moebius, Monopoly, Plastic, Reindeer, Rocks1, Rocks2, Wood1 i Wood2.

Ocena jakości map głębi przeprowadzona została zgodnie z metodologią znaną z literatury [4]. W tabeli 3.2.2 przedstawione zostały nazwy miar, ich oznaczenia oraz opisy:

Tabela 3.2.2 - Miary użyte do oceny jakości map głębi

Nazwa miary (ang)	Symbol	Opis
Bad pixels - all	B	Procent nieprawidłowych punktów
Bad pixels - nonocc	$B_{\bar{o}}$	Procent nieprawidłowych punktów na obszarze występowania odsłoneń
Bad pixels - occ	B_o	Procent nieprawidłowych punktów na obszarze występowania przesłoneń
Bad pixels - textured	B_T	Procent nieprawidłowych punktów na obszarze występowania tekstury
Bad pixels - textureless	$B_{\bar{T}}$	Procent nieprawidłowych punktów na obszarze braku występowania tekstury
Bad pixels - discontinuous	B_D	Procent nieprawidłowych punktów w pobliżu nieciągłości map głębi

Często spotykane w literaturze [4] określenie "nieprawidłowy punkt" (ang. "bad pixel") odpowiada punktom mapy głębi, w których wartość głębi różni się od głębi odniesienia o więcej niż zadaną wartość progową. Wymieniona wartość progowa oraz inne parametry niezbędne dla obliczenia wyżej wymienionych miar dobrane zostały zgodnie z metodologią przedstawioną w literaturze [4].

Zastosowanie danych ze zbioru Middlebury pozwoliło autorowi na przezwycięzenie ograniczeń technologii ToF oraz skonstruowanego przez siebie systemu wielokamerowego. Pozwoliło to na udowodnienie poprawności zaproponowanej metody fuzji danych wizyjnych z danymi z kamer głębi. Dzięki temu druga teza rozprawy została udowodniona.

Eksperymenty porównawcze przeprowadzone dla takich samych warunków jak dla porównania przez ocenę jakości widoku wirtualnego, jedynie metoda oceny jakości map głębi uległa zmianie. Ocena jakości map głębi przeprowadzona została zgodnie z metodologią opisaną w literaturze [4]. Podsumowanie wartości wybranych miar zebrane zostało w tabelach 3.2.3 – 3.2.6. Tabele przedstawiają uśrednione wyniki dla wszystkich eksperymentów.

Tabela 3.2.3. – Procent nieprawidłowych punktów w mapie głębi dla różnych klas punktów zgodnie z tabelą 3.2.2. Wyniki dla wariantu 1.

	Wyłącznie analiza obrazu	Proponowany algorytm											
		Model skokowy			Model liniowy			Model kwadratowy			Model Gaussowski		
		$b=\pm 1$	$b=\pm 2$	$b=\pm 4$	$a=\frac{1}{5}$	$a=\frac{1}{12.5}$	$a=\frac{1}{20}$	$a=\frac{1}{25}$	$a=\frac{1}{150}$	$a=\frac{1}{400}$	$\sigma=3$	$\sigma=6$	$\sigma=10$
B	9.63	4.47	6.09	8.95	4.58	4.46	4.33	4.36	4.58	6.00	4.42	4.27	4.63
$B_{\bar{T}}$	11.19	4.37	6.58	10.19	4.36	4.31	4.23	4.27	4.79	6.60	4.30	4.28	4.88
B_T	4.57	5.31	5.16	4.93	5.74	5.44	5.19	5.21	4.74	4.61	5.32	4.86	4.68
B_D	21.61	18.84	20.27	21.80	19.66	19.24	18.76	18.68	18.78	19.47	18.95	18.13	18.76

Tabela 3.2.4. – Procent nieprawidłowych punktów w mapie głębi dla różnych klas punktów zgodnie z tabelą 3.2.2. Wyniki dla wariantu 2.

	Wyłącznie analiza obrazu	Proponowany algorytm											
		Model skokowy			Model liniowy			Model kwadratowy			Model Gaussowski		
		$b=\pm 1$	$b=\pm 2$	$b=\pm 4$	$a=\frac{1}{5}$	$a=\frac{1}{12.5}$	$a=\frac{1}{20}$	$a=\frac{1}{25}$	$a=\frac{1}{150}$	$a=\frac{1}{400}$	$\sigma=3$	$\sigma=6$	$\sigma=10$
B	9.63	9.63	9.68	9.74	9.48	9.60	9.75	9.68	9.64	9.89	9.54	9.62	9.87
$B_{\bar{T}}$	11.19	11.19	11.24	11.33	10.98	11.18	11.33	11.26	11.23	11.42	11.10	11.20	11.41
B_T	4.57	4.39	4.40	4.44	4.38	4.43	4.46	4.41	4.46	4.68	4.41	4.45	4.66
B_D	21.61	20.54	20.67	20.80	20.70	20.74	20.76	20.72	20.85	20.90	20.66	20.77	20.91

Tabela 3.2.5. – Procent nieprawidłowych punktów w mapie głębi dla różnych klas punktów zgodnie z tabelą 3.2.2. Wyniki dla wariantu 3.

	Wyłącznie analiza obrazu	Proponowany algorytm											
		Model skokowy			Model liniowy			Model kwadratowy			Model Gaussowski		
		$b=\pm 1$	$b=\pm 2$	$b=\pm 4$	$a=\frac{1}{5}$	$a=\frac{1}{12.5}$	$a=\frac{1}{20}$	$a=\frac{1}{25}$	$a=\frac{1}{150}$	$a=\frac{1}{400}$	$\sigma=3$	$\sigma=6$	$\sigma=10$
B	8.91	4.35	6.02	8.63	4.43	4.31	4.19	4.25	4.49	5.90	4.28	4.17	4.55
$B_{\bar{T}}$	10.37	4.27	6.55	9.82	4.23	4.18	4.10	4.18	4.72	6.54	4.18	4.21	4.84
B_T	4.08	5.08	4.95	4.67	5.48	5.18	4.93	4.96	4.50	4.38	5.08	4.63	4.44
B_D	18.82	18.14	19.74	20.83	18.88	18.43	17.95	18.02	18.14	18.95	18.20	17.56	18.21

Tabela 3.2.6. – Procent nieprawidłowych punktów w mapie głębi dla różnych klas punktów zgodnie z tabelą 3.2.2. Wyniki dla wariantu 4.

	Wyłącznie analiza obrazu	Proponowany algorytm											
		Model skokowy			Model liniowy			Model kwadratowy			Model Gaussowski		
		b=±1	b=±2	b=±4	a= $\frac{1}{5}$	a= $\frac{1}{12.5}$	a= $\frac{1}{20}$	a= $\frac{1}{25}$	a= $\frac{1}{150}$	a= $\frac{1}{400}$	σ=3	σ=6	σ=10
B	8.91	8.89	8.95	9.01	8.92	9.02	9.05	8.99	9.07	9.28	8.98	9.07	9.10
$B_{\bar{T}}$	10.37	10.35	10.43	10.47	10.38	10.50	10.53	10.47	10.54	10.75	10.46	10.55	10.58
B_T	4.08	3.96	4.01	4.05	3.98	4.03	4.04	4.01	4.06	4.16	4.01	4.05	4.08
B_D	18.82	18.55	18.74	18.83	18.71	18.80	18.86	18.78	18.90	19.03	18.79	18.85	18.96

Wyniki wyraźnie uwidaczniają fakt, że mapy głębi uzyskane poprzez fuzję danych z kamer głębi z danymi wizyjnymi są bliższe referencyjnym mapom głębi niż te uzyskane wyłącznie na bazie analizy obrazu. Poprawa występuje dla wariantów eksperymentu 1 oraz 3. Jak należało oczekiwać największa poprawa występuje na obszarach gdzie w obrazach brakuje tekstury ($B_{\bar{T}}$). W tych obszarach analiza obrazu nie umożliwia wnioskowania o odległości punktów ze względu na brak cech, które mogłyby być do siebie dopasowane. Z drugiej strony w obszarach, gdzie tekstura występuje (B_T), poprawa jest niewielka bądź nawet jej nie ma. Są to regiony gdzie analiza obrazu pozwala dobrze dopasować cechy a przez to dokładnie wyznaczyć głębię. Minimalną poprawę można również zaobserwować w obszarach występowania odsłoneń (B_D). Niewielka jej wartość jest skutkiem konieczności zwiększenia rozdzielczości przestrzennej mapy głębi pochodzącej z kamery głębi. Podniesienie rozdzielczości wiąże się z interpolacją, która implikuje efekt rozmycia na krawędziach.

3.3 Ocena poprawy spójności międzywidokowej map głębi

Spójność międzywidokowa map głębi to kluczowy czynnik wpływający na efektywność ich kompresji przy użyciu technik kompresji sekwencji wielowidokowych. **Na potrzeby obiektywnej oceny spójności map głębi autor wprowadził miarę, która odzwierciedla różnice w reprezentacji tej samej sceny przez różne mapy głębi.** Proponowana miara bazuje na średniej wariancji obliczanej dla każdego zestawu punktów mapy głębi, które odpowiadają temu samemu punktowi w scenie. Miara ta obliczana jest dwuetapowo. W pierwszym etapie obliczane są miary cząstkowe V_i dla każdego widoku według wzoru 3.3.1:

$$V_i = \frac{1}{w \cdot h} \sum_{y=0}^w \sum_{x=0}^h var(D'_1(x, y), D'_2(x, y), \dots, D'_N(x, y)), \quad (3.3.1)$$

gdzie w i h określają przestrzenne wymiary mapy głębi natomiast $var()$ określa operator wariancji. Zmienne $D_1', D_2', \dots, D_N'$ oznaczają mapy głębi dla widoków 1,2, ...,N przetransformowane do przestrzeni i -tego widoku.

W drugim etapie obliczana jest średnia wszystkich miar cząstkowych V_i dla wszystkich widoków w sekwencji zgodnie ze wzorem 3.3.2:

$$V = \frac{1}{N} \sum_{i=1}^N V_i \quad (3.3.2)$$

Im wyższa wartość V tym mniej spójne międzywidokowo są mapy głębi.

Zastosowanie autorskiej miary do oceny poprawy spójności międzywidokowej polega na porównaniu jej wartości dla map głębi przed oraz po przetworzeniu za pomocą zaproponowanego algorytmu. Eksperymenty przeprowadzone przez autora w tym zakresie zostały wykonane na reprezentatywnym zbiorze map głębi [8]. Jako miara poprawy spójności (dokładniej redukcji niespójności) zastosowana została metryka zaproponowana przez autora.

Wartości miary niespójności międzywidokowej dla map głębi przed oraz po przetworzeniu przez zaproponowany przez autora algorytm zebrane zostały w tabeli 3.3.1.

Tabela 3.3.1. – Uśredniona wartość miary niespójności dla każdej użytej sekwencji testowej przed oraz po przetworzeniu przez proponowany algorytm.

Nazwa sekwencji	Niespójność (oryginalne mapy głębi)	Niespójność (przetworzone mapy głębi)	Liczba iteracji algorytmu
Balloons	48.15	1.78	13
Undo Dancer	0.73	0.56	3
GT Fly	3.09	1.70	3
Kendo	251.82	1.34	11
Lovebird1	21.14	2.22	8
Newspaper	155.77	2.34	15
Poznan Hall2	6.96	0.33	7
Poznan Street	5.20	0.43	9

Największą poprawę spójności można zaobserwować dla sekwencji przedstawiających sceny naturalne. Posiadają one mapy głębi wyznaczone automatycznie bądź pół-automatycznie niezależnie dla każdej kamery co jest źródłem ich niespójności. Ciekawym przypadkiem są sekwencje „Undo Dancer” oraz „GT-Fly”. Są to sekwencje syntetyczne zawierające domyślnie spójne mapy głębi. Po mimo tego w trakcie trzech iteracji proponowany algorytm spowodował dalszy wzrost ich spójności. Można ten fakt tłumaczyć błędami zaokrągleń podczas konwersji syntetycznych map głębi do 8-bitowej reprezentacji, jaka jest powszechnie stosowana.

3.4 Ocena wpływu spójności międzywidokowej map głębi na efektywność ich kompresji

Autor przeprowadził szereg eksperymentów polegających na kodowaniu nieprzetworzonych map głębi oraz przetworzonych za pomocą zaproponowanego przez siebie algorytmu zwiększającego ich spójność międzywidokową. Podczas eksperymentów wykorzystany został zbiór sekwencji wielowidokowych zalecanych przez grupę MPEG [8]. Autor zastosował techniki kompresji bazujące na technologii kompresji H.264/AVC takie jak MVC+D [26] i 3D-AVC [25] oraz techniki bazujące na technologii HEVC takie jak 3D-HEVC [24]. Eksperymenty zostały przeprowadzone zgodnie z zaleceniami grupy MPEG dla kodowania sekwencji wielowidokowych wraz z mapami głębi [16].

Na potrzeby opracowania wyników zastosowana została miara Bjontegaarda [20]. Miara ta została stworzona na potrzeby porównywania różnych techniki kompresji obrazu. Jednak z powodzeniem znajduje swoje zastosowanie również w porównaniu kompresji przetworzonych oraz nieprzetworzonych map głębi za pomocą tego samego kodeka. Istnieją dwa składniki tej miary: pierwszy z nich pozwala oszacować zmianę efektywności kompresji wyrażoną, jako zmiana przepływności bitowej. Drugi pozwala określić zmianę jakości obrazu po jego rekonstrukcji (zdekodowaniu) poprzez zastosowanie obiektywnej miary PSNR [10].

Każda sekwencja testowa została zakodowana wybranym kodekiem dla czterech różnych ustawień jakości zgodnie z zaleceniami [16]. Dla każdego ustawienia jakości obrazu dla kodowanych widoków zostały porównane z oryginalnymi za pomocą miary PSNR. Zdekodowane widoki oraz mapy głębi posłużyły do wygenerowania widoków wirtualnych. Jakość tych widoków wirtualnych została porównana z widokami rzeczywistymi również przy pomocy miary PSNR.

Wartość PSNR jak i przepływności bitowe dla czterech ustawień jakości posłużyły do sporządzenia krzywych prędkość bitowa -znieskształcenie (ang. rate-distortion). Dla każdego eksperymentu utworzone zostały dwie takie krzywe uwzględniające:

- jakość kodowanego obrazu kontra przepływność bitowa nań przypadająca,
- jakość widoku wirtualnego kontra całkowita przepływność bitowa

Punkty określone przez przepływność bitową oraz jakość dla każdego ustawienia jakości posłużyły do obliczenia miar Bjontegaarda dla każdego z eksperymentów. Tabele 3.4.2 – 3.4.4 zawierają przedstawienie wyników liczbowych tych miar.

Tabela 3.4.2 – Zysk prędkości bitowej oraz wartości PSNR dla widoków kodowanych oraz widoków wirtualnych przy zastosowaniu kodeka MVC+D.

	Widoki kodowane		Widoki wirtualne (syntezowane)	
Sekwencja testowa	Obraz ΔBR [%]	Obraz $\Delta PSNR$ [dB]	Głębia ΔBR [%]	Widoki wirtualne $\Delta PSNR$ [dB]
Poznan Hall 2	0.00%	0.00	-2.68%	0.10
Poznan Street	0.00%	0.00	-6.22%	0.22
Undo Dancer	0.00%	0.00	-3.61%	0.12
GT Fly	0.00%	0.00	-2.97%	0.12
Kendo	0.00%	0.00	-7.84%	0.39
Balloons	0.00%	0.00	-3.96%	0.20
Newspaper	0.00%	0.00	-8.05%	0.32
Średnia	0.00%	0.00	-5.05%	0.21

Tabela 3.4.3 – Zysk prędkości bitowej oraz wartości PSNR dla widoków kodowanych oraz widoków wirtualnych przy zastosowaniu kodeka 3D-AVC.

	Widoki kodowane		Widoki wirtualne (syntezowane)	
Sekwencja testowa	Obraz ΔBR [%]	Obraz $\Delta PSNR$ [dB]	Głębia ΔBR [%]	Widoki wirtualne $\Delta PSNR$ [dB]
Poznan Hall 2	0.02%	0.00	-2.90%	0.06
Poznan Street	0.02%	0.00	-7.93%	0.22
Undo Dancer	-0.02%	0.00	-1.08%	0.04
GT Fly	-0.03%	0.00	-5.53%	0.16
Kendo	-0.03%	0.00	-7.39%	0.29
Balloons	-0.01%	0.00	-3.42%	0.14
Newspaper	-0.01%	0.00	-9.84%	0.36
Średnia	-0.01%	0.00	-5.44%	0.18

Tabela 3.4.4 – Zysk prędkości bitowej oraz wartości PSNR dla widoków kodowanych oraz widoków wirtualnych przy zastosowaniu kodeka 3D-HEVC.

	Widoki kodowane		Widoki wirtualne (syntezowane)	
Sekwencja testowa	Obraz ΔBR [%]	Obraz $\Delta PSNR$ [dB]	Głębia ΔBR [%]	Widoki wirtualne $\Delta PSNR$ [dB]
Poznan Hall 2	-0.01%	0.00	-3.37%	0.11
Poznan Street	0.08%	0.00	-6.66%	0.22
Undo Dancer	-0.03%	0.00	-3.99%	0.13
GT Fly	-0.16%	0.01	-3.00%	0.11
Kendo	0.08%	0.00	-8.80%	0.39
Balloons	0.11%	-0.01	-4.39%	0.20
Newspaper	-0.21%	0.01	-8.03%	0.31
Średnia	-0.02%	0.00	-5.46%	0.21

Dla kodeka MVC+D nie można zaobserwować poprawy efektywności kompresji widoków. Dzieje się tak ze względu na fakt, że kodek ten traktuje niezależnie widoki oraz mapy głębi. W przypadku kodeka 3D-AVC można zauważyć spadek przepływności bitowej dla obrazów, w niektórych przypadkach do 0.03%. Kodek 3D-AVC posiada możliwość używania wspólnej informacji sterującej dla widoków oraz map głębi. Podobna sytuacja występuje dla kodeka 3D-HEVC.

Natomiast dla wszystkich testowanych kodeków wyraźnie zarysowuje się spadek prędkości bitowej dla części strumienia odpowiedzialnego za reprezentację map głębi. Ponadto da się zaobserwować nieznaczny wzrost jakości widoków wirtualnych. Jest to oczekiwany rezultat. Mapy głębi przetworzone za pomocą algorytmu zaproponowanego przez autora, poprzez swoją większą spójność międzywidokową, pozwalają na lepsze wykorzystanie podobieństw międzywidokowych przez algorytmy kompresji. **Pozwala to na osiągnięcie średniej redukcji prędkości bitowej o ponad 5%.**

4. Podsumowanie

4.1 Oryginalne osiągnięcia przedstawione w rozprawie

Fuzja danych z kamer wizyjnych oraz danych z kamer głębi ToF

Głównym osiągnięciem autora, przedstawionym w rozprawie, jest metoda fuzji danych z kamer wizyjnych oraz danych z kamer głębi ToF.

Proces fuzji pozwala na wprowadzenie dodatkowych cech do algorytmu estymacji głębi, zgodnego ze stanem wiedzy [15]. Dodatkowe cechy pozwalają na otrzymanie map głębi, które są dokładniejsze pod względem reprezentacji fizycznych odległości. Fuzja danych wizyjnych oraz danych z kamer głębi jest przeprowadzana przed etapem globalnej optymalizacji kosztu. Umożliwia to podjęcie decyzji na etapie jego globalnej optymalizacji, które dane są lepszym źródłem informacji o odległości: dane pochodzące z analizy obrazów czy też bezpośredni pomiar odległości z kamery głębi.

Autor proponuje również metodę fuzji danych z wielu kamer głębi. Zastosowanie wielu kamer głębi pozwala na zwiększenie zarówno zakresu przestrzennego reprezentacji sceny jak i również dokładności pomiaru odległości.

Poprawa spójności międzywidokowej map głębi

Niezależna estymacja głębi dla sąsiadujących ze sobą widoków nieuchronnie prowadzi do ich niespójności. Jest to spowodowane faktem, że dla każdego widoku głębia wyznaczana jest na podstawie różnych par bądź większej liczby obrazów [15].

Autor rozprawy zaproponował algorytm, który pozwala na poprawę spójności międzywidokowej map głębi poprzez iteracyjną wymianę informacji między nimi. Algorytm pozwala na przetworzenie map głębi bez konieczności ponownej ich estymacji.

Przetworzone mapy głębi cechują się większą spójnością przestrzenną przez co znacznie lepiej poddają się kompresji za pomocą kodeka wielowidokowego. Eksperymenty przeprowadzone przy użyciu kodeków takich jak MVC, MVC+D, 3D-AVC i 3D-HEVC pokazały, że **prędkość bitowa może zostać zredukowana o ponad 5%** w stosunku do prędkości bitowej dla oryginalnych map głębi. Co więcej jakość kodowanych danych wizyjnych, przy użyciu przetworzonych map głębi, nie ulega pogorszeniu.

Zaproponowany przez autora algorytm stał się częścią technologii kompresji wielowidokowych sekwencji wizyjnych oraz map głębi zgłoszonej przez Politechnikę Poznańską do konkursu zorganizowanego przez grupę MPEG w 2011 roku [36] [42]. **Technologia ta została uznana za drugą najlepszą na świecie pod względem efektywności kompresji.**

Synchronizacja kamer wizyjnych oraz kamer głębi

W hybrydowym systemie wielokamerowym, wszystkie kamery muszą rejestrować obrazy synchronicznie. Niestety kamery wizyjne oraz kamery głębi wykorzystują różne rodzaje sygnałów synchronizacji przez co nie jest możliwa ich bezpośrednia synchronizacja.

Autor szczegółowo zbadał problem konwersji sygnału Genlock, stosowanego w profesjonalnych kamerach telewizyjnych, na sygnał wyzwalania (ang. trigger), stosowany w kamerach głębi. W rezultacie **autor zaproponował metodę konwersji, która umożliwia synchronizację kamer wizyjnych oraz kamer głębi z jednego generatora sygnału Genlock.**

Ponadto **autor zaprojektował oraz wykonał urządzenie sprzętowe, które implementuje proponowaną metodę konwersji.** Urządzenie posiada wejście sygnału Genlock oraz wyjście sygnału wyzwalania dla kamer głębi. Posłużyło ono autorowi do konstrukcji hybrydowego systemu wielokamerowego oraz umożliwiło rejestrację szeregu sekwencji wielowidokowych.

Kalibracja kamer wizyjnych oraz kamer głębi w hybrydowych systemach wielokamerowych

Autor przeprowadził szereg badań związanych z kalibracją kamer oraz rektyfikacją obrazów. Zaproponował również zestaw modyfikacji do algorytmów wyznaczania parametrów kamer znanych z literatury, które pozwalają na ich poprawne działanie dla danych z kamer głębi ToF. Ponadto autor zaproponował oryginalny algorytm wyznaczania parametrów zewnętrznych kamer głębi, który bazuje na trójwymiarowych współrzędnych punktów obrazu, jakie dostarcza kamera ToF.

Najważniejszymi z zaproponowanych algorytmów są:

- algorytm estymacji kierunku linii, na której rozmieszczone są kamery wizyjne,
- algorytm zmiany kierunku linii rozmieszczenia kamer wizyjnych, pozwalający na lepszą kompensację ich rotacji,
- metoda wyznaczania parametrów zewnętrznych kamer głębi przy użyciu trójwymiarowych współrzędnych punktów obrazu.

Metoda redukcji szumu dla danych z kamer głębi ToF

Dane z współczesnych kamer głębi ToF cechują się wysokim poziomem szumu. Występowanie przestrzennego oraz czasowego szumu w mapie głębi ma negatywny wpływ na reprezentację sceny trójwymiarowej. Z tego względu **autor zaproponował algorytm redukcji szumu dla danych z kamer ToF.**

Proponowany algorytm uwzględnia zarówno informacje o odległości jak i amplitudzie odbitej fali świetlnej. Podstawą jego działania jest bilateralna filtracja przestrzenna oraz adaptacyjna filtracja czasowa połączona z wykrywaniem ruchu. W rezultacie proponowana metoda pozwala na redukcję czasowych fluktuacji map głębi, co przyczynia się do wzrostu ich spójności czasowej. Dzieje się to

bez wprowadzania niepożądanych zniekształceń takich jak rozmycie spowodowane ruchem oraz zniekształcenia w okolicach krawędzi.

4.2 Przeprowadzone prace badawcze

Na potrzeby udowodnienia tez rozprawy, autor dokonał serii złożonych i kompleksowych eksperymentów. Eksperymenty te obejmowały swoim zakresem zarówno mnogość konfiguracji autorskiej metody fuzji danych wizyjnych z danymi z kamer głębi jak i metody poprawy spójności międzywidokowej map głębi. Autor przeprowadził również pokaźną liczbę eksperymentów związanych z kompresją wielowidokową map głębi.

Na potrzeby eksperymentów związanych z fuzją danych, wykonanych zostało dla każdego zbioru danych wejściowych 13 jej przebiegów każdorazowo dla różnego zestawu parametrów. Danymi wejściowymi było 5 sekwencji wielowidokowych, przygotowanych przez autora, oraz 27 zbiorów danych z bazy Middlebury. W sumie przeprowadzonych zostało **1755 przebiegów fuzji danych** wizyjnych z danymi z kamer głębi połączonych z analizą jakości otrzymanych map głębi.

Eksperymenty, związane z poprawą spójności międzywidokowej map głębi, obejmowały przetwarzanie głębi dla 8 sekwencji wielowidokowych połączonych z ich kompresją za pomocą 4 kodeków. Każda sekwencja kodowana była przy użyciu 4 ustawień jakości dla oryginalnych oraz przetworzonych map głębi. W sumie przeprowadzone zostało **256 kodowań oraz dekodowań**.

4.3 Wnioski

W rozprawie udowodniono tezę jakoby możliwe było skonstruowanie hybrydowego systemu wielokamerowego, wyposażonego w kamery wizyjne oraz kamery głębi, oraz uzyskanie przy jego pomocy map głębi o lepszej jakości niż możliwe do uzyskania przy pomocy systemu wyposażonego wyłącznie w kamery wizyjne.

W celu udowodnienia tezy pierwszej autor zaproponował algorytm fuzji danych, który bazuje na modyfikacji funkcji kosztu w procesie globalnej optymalizacji mapy głębi. Modyfikacje, niezbędne do implementacji proponowanego algorytmu, zostały wprowadzone do oprogramowania DERS. Autor udowodnił możliwość wyznaczenia lepszych map głębi proponowaną metodą poprzez szereg eksperymentów. Za dane wejściowe posłużyły dane ze zbioru Middlebury oraz sekwencje wielowidokowe przygotowane przez autora.

Druga teza rozprawy, która została udowodniona, traktuje o możliwości przetworzenia wstępnego map głębi w taki sposób aby zwiększeniu uległa ich spójność międzywidokowa.

Na potrzeby udowodnienia tezy drugiej autor zaproponował algorytm poprawiający spójność międzywidokową map głębi oraz miarę tej spójności. Autor przeprowadził też szereg eksperymentów dotyczących kompresji przetworzonych map głębi i wykazał, że zastosowanie zaproponowanej przez niego metody pozwala na zwiększenie efektywności kompresji o ponad 5%.

Literatura

- [1] K. Szeląg, G. Maczkowski, R. Gierwiało, A. Gebarska, R. Sitnik, "Robust geometric, phase and color structured light projection system calibration", *Opto-Electronics Review* 25 (4), 326-336
- [2] D. Scharstein, C. Pal, "Learning conditional random fields for stereo", *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2007)*, Minneapolis, MN, USA, June 2007.
- [3] D. Scharstein, R. Szeliski, "High-accuracy stereo depth maps using structured light", *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2003)*, vol. 1, pp. 195-202, Madison, WI, USA, June 2003.
- [4] D. Scharstein, R. Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms", *International Journal of Computer Vision*, 47(1/2/3):7-42, April-June 2002.
- [5] R. Horaud, M. Hansard, G. Evangelidis, C. Ménier, "An Overview of Depth Cameras and Range Scanners Based on Time-of-Flight Technologies", *Machine Vision and Applications Journal*, 2016, 27 (7), pp.1005-1020.
- [6] D. Mieloch, "Depth Estimation in Free-Viewpoint Television" (PhD dissertation), Poznan University of Technology, Faculty of Multimedia Telecommunication and Microelectronics, Poznań, 2018.
- [7] E. Bebeslea-Sterp, R. Brad, R. Brad, "A Comparative Study of Stereovision Algorithms", (*IJACSA*) *International Journal of Advanced Computer Science and Applications*, vol. 8, no. 11, 2017.
- [8] "Call for Proposals on 3D Video Coding Technology", *ISO/IEC JTC1/SC29/WG11, MPEG2011/W12036*, Geneva, Switzerland, March 2011.
- [9] C. Fehn, "Depth-image-based Rendering (DIBR), Compression and Transmission for a New Approach on 3D-TV", in *Proc. SPIE Conf. 5291*, CA, U.S.A., pp. 93-104, Jan. 2004.
- [10] Q. Huynh-Thu, M. Ghanbari, "Scope of validity of PSNR in image/video quality assessment", *Electronics Letters*, vol. 44, no. 13, pp. 800-801, June 19 2008.
- [11] Z. Wang, L. Lu, and A.C. Bovik, "Video quality assessment based on structural distortion measurement," *Signal Processing: Image Communication*, vol. 19, no. 2, pp. 121–132, 2004
- [12] A. Tikanmaki, A. Gotchev, A. Smolic and K. Miller, "Quality assessment of 3D video in rate allocation experiments," *2008 IEEE International Symposium on Consumer Electronics*, Vilamoura, 2008, pp. 1-4. doi: 10.1109/ISCE.2008.4559441
- [13] A. Banitalebi-Dehkordi, M. T. Pourazad and P. Nasiopoulos, "3D video quality metric for 3D video compression," *IVMSP 2013*, Seoul, 2013, pp. 1-4. doi: 10.1109/IVMSPW.2013.6611930
- [14] H. Hirschmuller, D. Scharstein, "Evaluation of cost functions for stereo matching", *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2007)*, Minneapolis, MN, USA, June 2007.

- [15] O. Stankiewicz, K. Wegner, M. Wildeboer, „A soft – segmentation matching in Depth Estimation Reference Software (DERS) 5.0”, ISO/IEC JTC1/SC29/WG11 MPEG2009/M17049, Xian, China, Oct. 2009.
- [16] D. Rusanovskyy, K. Muller, A. Vetro, "Common Test Conditions of 3DV Core Experiments", ISO/IEC JTC1/SC29/WG11, MPEG2014/M27299, Shanghai, China, October 2012
- [17] Calbaza, D.E.; Cordos, I.; Seth-Smith, N.; Savaria, Y., "An ADPLL circuit using a DDPS for Genlock applications," Circuits and Systems, 2004. ISCAS '04. Proceedings of the 2004 International Symposium on , vol.4, no., pp.IV,569-72 Vol.4, 23-26 May 2004.
- [18] J.S. Yedidia, W.T. Freeman, Y. Weiss, "Understanding Belief Propagation and Its Generalisations" in Exploring Artificial Intelligence in the New Millennium, G. Lakemeyer, B. Nebel, ISBN: 1-55860-811-7, Chapter 8, pp. 239-236, Morgan Kaufmann Publishers, January 2003.
- [19] Y. Boykov, O. Veksler, R. Zabih, "Fast approximate energy minimisation via graph cuts", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.23, no.11, pp.1222,1239, November 2001.
- [20] G. Bjontegaard, "Calculation of Average PSNR Differences between RD-curves," ITU-T video Coding Experts Group document VCEG-M33, March 2001.
- [21] Z. Zhang. "A Flexible New Technique for Camera Calibration". IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(11):1330-1334, 2000.
- [22] R. Hartley, A. Zisserman, "Multiple View Geometry in Computer Vision", Cambridge University Press, March 2014, ISBN 0521540518, 9780521540513.
- [23] B. Cyganek, "Computer Processing of 3D images", Academic Publisher House EXIT, ISBN 83-87674-34-6, Warsaw 2002.
- [24] G. Tech, K. Wegner, Y. Chen, S. Yea, "3D-HEVC Test Model 1", JCT-3V of ITU-T SG 16 WP 3 and ISO/IEC JTC1/SC29/WG11, Doc. JCT3V-A1005-d0, Stockholm, Sweden, 16-20 July, 2012.
- [25] Recommendation ITU-T H.264, "Advanced video coding for generic audiovisual services", Annex J - "Multiview and depth video with enhanced non-base view coding", 04/2017.
- [26] Recommendation ITU-T H.264, "Advanced video coding for generic audiovisual services", Annex I - "Multiview and depth video coding", 04/2017.
- [27] SMPTE standard 274M-2008, "1920 × 1080 Scanning and Analog and Parallel Digital Interfaces for Multiple Picture Rates".
- [28] Y.S. Kang, C. Lee; Y.S. Ho, "An Efficient Rectification Algorithm for Multi-View Images in Parallel Camera Array," 3DTV Conference: The True Vision - Capture, Transmission and Display of 3D Video, pp. 61,64, 28-30 May 2008.
- [29] Recommendation ITU-T H.264, "Advanced video coding for generic audiovisual services", Annex H - "Multiview video coding", 04/2017.
- [30] Recommendation ITU-T H.265, "High efficiency video coding", Annex G - "Multiview high efficiency video coding", 02/2018.

- [31] Canon XH-G1 camera datasheet, retrieved on 2011 from:
<https://si.ua.es/es/lccm/documentos/av-camara-de-video-hdv-profesional-canon-xh-a1.pdf>
- [32] Mesa Imaging SR4000 ToF camera datasheet. Retrieved on 2011 from:
http://www.mesa-imaging.ch/dlm.php?fname=pdf/SR4000_Data_Sheet.pdf

Dorobek naukowy Autora

Artykuły w czasopismach o dużym znaczeniu międzynarodowym (lista A)

- [33] Marek Domański, Olgierd Stankiewicz, Krzysztof Wegner, Maciej Kurc, Jacek Konieczny, Jakub Siast, Jakub Stankowski, Robert Ratajczak, Tomasz Grajek, "High Efficiency 3D Video Coding Using New Tools Based on View Synthesis", IEEE Transactions on Image Processing, Special Issue on 3D Video Representation, Compression, & Rendering, Vol. 22, No. 9, September 2013, pp. 3517-3527.

Artykuły opublikowane w materiałach konferencji międzynarodowych

- [34] Maciej Kurc, Olgierd Stankiewicz, Marek Domański, "Depth map inter-view consistency refinement for multiview video", 29th Picture Coding Symposium, PCS 2012, Kraków, Poland, 7-9 May 2012, pp. 137-140.
- [35] Maciej Kurc, Krzysztof Wegner, Marek Domański, "Transformation of Depth Maps Produced by ToF Cameras", International Conference on Signals and Electronic Systems ICSES 2014, Poznań, Poland, 11-13 September 2014,
- [36] Marek Domański, Adrian Dziembowski, Agnieszka Kuehn, Maciej Kurc, Adam Łuczak, Dawid Mieloch, Jakub Siast, Olgierd Stankiewicz, Krzysztof Wegner, "Experiments on acquisition and processing of video for free-viewpoint television", 3DTV Conference 2014, Budapest, Hungary, 2-4 July 2014.
- [37] Marek Domański, Adrian Dziembowski, Tomasz Grajek, Adam Grzelka, Łukasz Kowalski, Maciej Kurc, Adam Łuczak, Dawid Mieloch, Robert Ratajczak, Jarosław Samelak, Olgierd Stankiewicz, Jakub Stankowski, Krzysztof Wegner, "Methods of High Efficiency Compression for Transmission of Spatial Representation of Motion Scenes", IEEE International Conference on Multimedia and Expo ICME 2015, Torino, Italy, June 29-July 3 2015.
- [38] Marek Domański, Jacek Konieczny, Maciej Kurc, Adam Łuczak, Jakub Siast, Olgierd Stankiewicz, Krzysztof Wegner, "Fast Depth Estimation on Mobile Platforms and FPGA Devices", 3DTV Conference: The True Vision-Capture, Transmission and Display of 3D Video, 3DTV- Con 2015, Lisbon, Portugal, 8-10 July 2015.
- [39] Marek Domański, Jacek Konieczny, Maciej Kurc, Robert Ratajczak, Jakub Siast, Olgierd Stankiewicz, Jakub Stankowski, Krzysztof Wegner, "3D video compression by coding of disoccluded regions", 19th IEEE International Conference on Image Processing (ICIP 2012), Orlando, Florida, U.S.A., 30 September - 3 October 2012, pp. 1317-1320.
- [40] Marek Domański, Krzysztof Klimaszewski, Jacek Konieczny, Maciej Kurc, Adam Łuczak, Olgierd Stankiewicz, Krzysztof Wegner, "An experimental Free-view Television System", 1st International Conference on Image Processing & Communications (IPC), Bydgoszcz, Polska, September 2009, pp. 175-184.
- [41] Marek Domański, Tomasz Grajek, Damian Karwowski, Jacek Konieczny, Maciej Kurc, Adam Łuczak, Robert Ratajczak, Jakub Siast, Jakub Stankowski, Krzysztof Wegner, "Coding of multiple video+depth using HEVC technology and reduced representations of side views and depth maps", 29th Picture Coding Symposium, PCS 2012, Kraków, Poland, 7-9 May 2012, pp. 5-8.

- [42] Marek Domański, Tomasz Grajek, Damian Karwowski, Krzysztof Klimaszewski, Jacek Konieczny, Maciej Kurc, Adam Łuczak, Robert Ratajczak, Jakub Siast, Olgierd Stankiewicz, Jakub Stankowski, Krzysztof Wegner, "New coding technology for 3D video with depthmaps as proposed for standardization within MPEG", 19th International Conference on Systems, Signals and Image Processing, IWSSIP 2012, Vienna, Austria, 11-13 April 2012, pp. 401-404.
- [43] Robert Ratajczak, Tomasz Grajek, Krzysztof Wegner, Krzysztof Klimaszewski, Maciej Kurc, Marek Domański, "Vehicle Dimensions Estimation Scheme Using AAM on Stereoscopic Video", 10th IEEE International Conference on Advanced Video and Signal-Based Surveillance, AVSS 2013, Workshop on Vehicle Retrieval in Surveillance (VRS) 2013, Kraków, Poland, 27-30 August 2013, pp. 478-482.
- [44] Sławomir Maćkowiak, Jacek Konieczny, Maciej Kurc, Przemysław Maćkowiak, "Football Player Detection in Video Broadcast", Proceedings of ICCVG, Lecture Notes in Computer Science, 20-22 September 2010, pp. 118-125.
- [45] Sławomir Maćkowiak, Jacek Konieczny, Maciej Kurc, Przemysław Maćkowiak, "A complex system for football player detection in broadcasted video", 2010 International Conference on Signals and Electronic Systems (ICSES), 7-10 September 2010, pp. 119-122.

Artykuły naukowe w materiałach konferencji normalizacyjnych

- [46] Marek Domański, Adrian Dziembowski, Tomasz Grajek, Adam Grzelka, Maciej Kurc, Adam Łuczak, Dawid Mieloch, Jakub Siast, Olgierd Stankiewicz, Krzysztof Wegner, [FTV AHG] Video and depth multiview test sequences acquired with circular camera arrangement – "Poznan Service" and "Poznan People", ISO/IEC JTC1/SC29/WG11 MPEG2015, Doc. M36569, Warsaw, Poland, 20-27 June 2015.
- [47] Marek Domański, Krzysztof Klimaszewski, Maciej Kurc, Robert Ratajczak, Olgierd Stankiewicz, Krzysztof Wegner, "Super-multi-viewlight-field images from Poznan University of Technology", ISO/IEC JTC1/SC29/WG11 MPEG2015, Doc. M36566, Warsaw, Poland, 20-27 June 2015.
- [48] Marek Domański, Adrian Dziembowski, Maciej Kurc, Adam Łuczak, Dawid Mieloch, Jakub Siast, Olgierd Stankiewicz, Krzysztof Wegner, Poznan University of Technology test multiview video sequences acquired with circular camera arrangement – "Poznan Team" and "PoznanBlocks" sequences, ISO/IEC JTC1/SC29/WG11 MPEG2015, Doc. M35846, Geneva, Switzerland, 14-20 February 2015.
- [49] Marek Domański, Krzysztof Klimaszewski, Maciej Kurc, Adam Łuczak, Olgierd Stankiewicz, Krzysztof Wegner, "FTV: Poznan Laboratory - a test light-field sequence from Poznan University of Technology", ISO/IEC JTC 1/SC 29/WG 11, Doc. MPEG2014/m35071, Strasbourg, France, 20-24 October 2014.
- [50] Marek Domański, Tomasz Grajek, Damian Karwowski, Krzysztof Klimaszewski, Jacek Konieczny, Maciej Kurc, Adam Łuczak, Robert Ratajczak, Jakub Siast, Olgierd Stankiewicz, Jakub Stankowski, Krzysztof Wegner, "Technical Description of Poznan University of Technology proposal for Call on 3D Video Coding Technology", ISO/IEC JTC1/SC29/WG11, MPEG 2011 / M22697, Geneva, Switzerland, 28 November-02 December 2011.

- [51] Marek Domański, Tomasz Grajek, Damian Karwowski, Krzysztof Klimaszewski, Jacek Konieczny, Maciej Kurc, Adam Łuczak, Robert Ratajczak, Jakub Siast, Olgierd Stankiewicz, Jakub Stankowski, Krzysztof Wegner, "Multiview HEVC – experimental results", Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11, MPEG 2011 / M22147, Geneva, Switzerland, 28 November - 02 December 2011.
- [52] Marek Domański, Tomasz Grajek, Krzysztof Klimaszewski, Maciej Kurc, Olgierd Stankiewicz, Jakub Stankowski, Krzysztof Wegner, "Poznań Multiview Video Test Sequences and Camera Parameters", ISO/IEC JTC1/SC29/WG11, MPEG/M17050, Doc. m17050, Xian, China, 26-30 October 2009.

Artykuły w czasopiśmie krajowych

- [53] Marek Domański, Tomasz Grajek, Damian Karwowski, Krzysztof Klimaszewski, Jacek Konieczny, Maciej Kurc, Adam Łuczak, Robert Ratajczak, Jakub Siast, Olgierd Stankiewicz, Jakub Stankowski, Krzysztof Wegner, "Poznański kodek obrazów trójwymiarowych", Przegląd Telekomunikacyjny, No. 2-3, February/March 2013, pp. 81-83,
- [54] Adam Łuczak, Marta Stępniewska, Jakub Siast, Marek Domański, Olgierd Stankiewicz, Maciej Kurc, Jacek Konieczny, "Network-on-Multi-Chip (NoMC) with monitoring and debugging support", Journal of Telecommunications and Information Technology, No. 3/2011, 2011, pp. 81-81,
- [55] Adam Łuczak, Maciej Kurc, Marta Stępniewska, Jakub Siast, "Interfejs komunikacyjny dla układów FPGA serii Virtex", Pomiary Automatyka Kontrola, PAK, 2010, pp. 749-751,
- [56] Sławomir Maćkowiak, Jacek Konieczny, Maciej Kurc, Przemysław Mackowiak, "System detekcji i śledzenia piłkarzy w transmisjach widowisk sportowych w cyfrowym sygnale wizyjnym", Elektronika, 2010, pp. 13-15,
- [57] Adam Łuczak, Maciej Kurc, Marta Stępniewska, Krzysztof Wegner, "Platforma przetwarzania rozproszonego bazująca na sieci NoC", Pomiary Automatyka Kontrola, PAK, 2009, pp. 690-692,

Przyznane patenty

- [58] Marek Domański, Tomasz Grajek, Krzysztof Klimaszewski, Maciej Kurc, Jakub Stankowski, Robert Ratajczak, Krzysztof Wegner, "A system and method for object dimension estimation", Patent office: USPTO, Status: granted, Application number: US 14/664983, Filing date: 23.03.2015.
- [59] Marek Domański, Tomasz Grajek, Krzysztof Klimaszewski, Maciej Kurc, Jakub Stankowski, Robert Ratajczak, Krzysztof Wegner, "A system and method for object dimension estimation", Patent office: EPO, Status: granted, Application number: EP15160202.06, Filing date: 23.03.2015,
- [60] Marek Domański, Krzysztof Klimaszewski, Jacek Konieczny, Maciej Kurc, Robert Ratajczak, Jakub Siast, Olgierd Stankiewicz, Jakub Stankowski, Krzysztof Wegner, "Image coding method", Patent office: USPTO, Status: granted, Application number: US 13/680652, Filing date: 19.11.2012, Publication number: US 2013/0129235A1, Publication date: 23.05.2013, Patent number: US 8761527, Date of Patent: 24.06.2014.

Zgłoszenia patentowe

- [61] Marek Domański, Tomasz Grajek, Krzysztof Klimaszewski, Maciej Kurc, Jakub Stankowski, Robert Ratajczak, Krzysztof Wegner, "System do estymacji rozmiarów obiektu i sposób estymacji rozmiarów obiektu", Patent office: UPRP, Status: application, Application number: P.411656, Filling date: 21.03.2015,
- [62] Marek Domański, Tomasz Grajek, Jacek Konieczny, Maciej Kurc, Adam Łuczak, Jakub Siast, Olgierd Stankiewicz, Jakub Stankowski, Krzysztof Wegner, "Method for coding of stereoscopic depth", Patent office: USPTO, Status: application, Application number: US 13/680822, Filling date: 19.11.2012, Publication number: US 2013/0129244A1, Publication date: 23.05.2013,
- [63] Marek Domański, Jacek Konieczny, Maciej Kurc, Robert Ratajczak, Jakub Siast, Olgierd Stankiewicz, Jakub Stankowski, Krzysztof Wegner, "Method for predicting a shape of an encoded area using a depth map", Patent office: USPTO, Status: application, Application number: US 13/680740, Filling date: 19.11.2012, Publication number: US 2013/0128968A1, Publication date: 23.05.2013,
- [64] Tomasz Żernicki, Maciej Kurc, Marcin Chryszczanowicz, Jakub Zamojski, Piotr Makaruk, Piotr Szczechowiak, Łukasz Januszkiewicz, „Microphone probe, method, system and computer program product for audio signals processing”, Patent office: EPO, Status: request for examination was made, Application number EP3414919, Filling date: 9.02.2017.