

Politechnika Poznańska

Wydział Elektroniki i Telekomunikacji



Katedra Telekomunikacji Multimedialnej i Mikroelektroniki

Streszczenie rozprawy

**Techniki kodowania i wyznaczania map  
głębni stereoskopowej w systemach  
telewizji wielowidokowej**

Stereoscopic depth map estimation and coding techniques  
for multiview video systems

Olgierd Stankiewicz

*Promotor: prof. dr hab. inż. Marek Domański*

Poznań, 2013

POLITECHNIKA POZNAŃSKA  
Wydział Elektroniki i Telekomunikacji  
Katedra Telekomunikacji Multimedialnej i Mikroelektroniki

Pl. M. Skłodowskiej-Curie 5  
60-965 Poznań

[www.et.put.poznan.pl](http://www.et.put.poznan.pl)  
[www.multimedia.edu.pl](http://www.multimedia.edu.pl)

Copyright © Olgierd Stankiewicz, 2013  
Wszelkie prawa zastrzeżone  
Wydrukowano w Polsce

Praca doktorska wspierana przez publiczne fundusze przeznaczone na badania naukowe.

Realizacja rozdziałów pracy doktorskiej, dotyczących estymacji głębi, była częściowo sfinansowana ze środków Narodowego Centrum Nauki przyznanych na podstawie decyzji numer DEC-2012/07/N/ST6/02267.

Realizacja rozdziałów pracy doktorskiej, dotyczących kodowania głębi, była częściowo sfinansowana ze środków Narodowego Centrum Nauki przyznanych na podstawie decyzji numer DEC-2012/05/B/ST7/01279.

Praca doktorska współfinansowana ze środków Unii Europejskiej w ramach Europejskiego Funduszu Społecznego.

## Spis treści streszczenia

<b>Rozdział 1. Wstęp.....</b>	<b>1</b>
1.1. Zakres rozprawy .....	1
1.2. Cele i tezy rozprawy .....	4
1.3. Przegląd rozprawy.....	5
1.4. Metodologia pracy .....	6
1.5. Wielowidokowe sekwencje testowe .....	6
<b>Rozdział 2. Stan wiedzy w zakresie estymacji map głębi.....</b>	<b>8</b>
<b>Rozdział 3. Proponowane nowe metody estymacji głębi .....</b>	<b>10</b>
3.1. Wyprowadzenie składnika <i>Data Cost</i> funkcji celu wykorzystujące regułę MAP .....	11
3.2. Funkcja celu <i>Data Cost</i> jako uogólnienie miar SSD i SAD .....	11
3.3. Weryfikacja empiryczna założeń.....	12
3.4. Proponowane nowe sformułowanie dla składnika <i>Data Cost</i> funkcji celu .....	13
3.5. Proponowane nowe sformułowanie dla składnika <i>Transition Cost</i> funkcji celu .....	14
3.6. Estymacja głębi z wykorzystaniem zaproponowanej funkcji celu .....	14
3.7. Algorytm Mid-Level Hypothesis (MLH).....	15
3.8. Poprawa spójności czasowej głębi za pomocą redukcji szumu .....	16
<b>Rozdział 4. Stan wiedzy w zakresie kodowania map głębi .....</b>	<b>18</b>
<b>Rozdział 5. Zaproponowana nieliniowa reprezentacja głębi dla kodowania .....</b>	<b>19</b>
5.1. Idea nieliniowej reprezentacji głębi.....	19
5.2. Propozycja prostego przekształcenia nieliniowego .....	20
5.3. Wyprowadzenie teoretyczne przekształcenia nieliniowego .....	20
5.4. Aproksymacja funkcji przekształcenia nieliniowego .....	21
5.5. Wyniki eksperymentalne kodowania głębi.....	22
5.6. Nieliniowa reprezentacja głębi w międzynarodowych normach .....	22
<b>Rozdział 6. Nowa technika kodowania sekwencji trójwymiarowych.....</b>	<b>23</b>
6.1. Porównanie z innymi kodekami.....	23
6.2. Struktura zaproponowanego kodeka .....	24
6.3. Udział autora w tworzeniu kodeka .....	25
6.4. Wyniki eksperymentalne wydajności kodeka.....	25
<b>Rozdział 7. Podsumowanie rozprawy.....</b>	<b>26</b>
7.1. Osiągnięcia związane z weryfikacją tez rozprawy .....	26
7.2. Przegląd przeprowadzonych prac.....	28
7.3. Głównie osiągnięcia .....	30
7.4. Drugoplanowe osiągnięcia.....	31
<b>Dokumentacja dorobku naukowego autora.....</b>	<b>32</b>

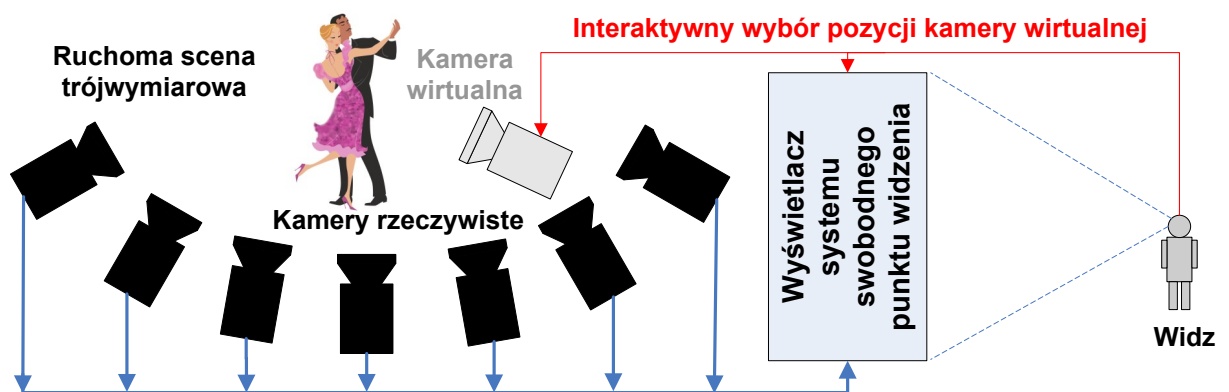


## Rozdział 1. Wstęp

### 1.1. Zakres rozprawy

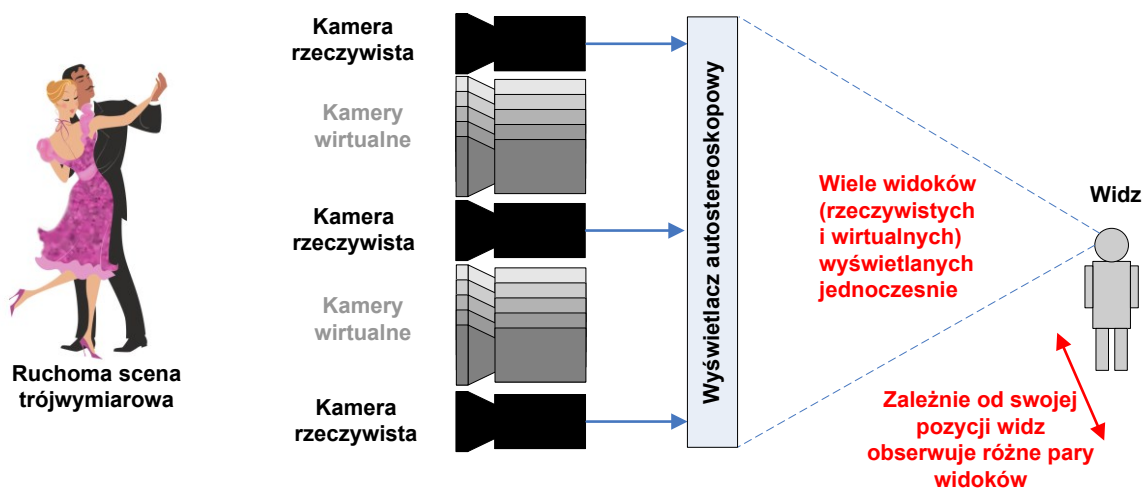
Rozprawa dotyczy nowej generacji systemów obrazu trójwymiarowego. W stosunku do dotychczas znanych i stosowanych rozwiązań wykorzystujących stereoskopię, nowe systemy mają zapewnić lepsze wrażenia obrazu przestrzennego, lepszą reprodukcję struktury sceny oraz umożliwić widzowi oddziaływanie na widzianą scenę. Zastosowania nowej generacji trójwymiarowych systemów wizyjnych obejmują zwłaszcza systemy swobodnego punktu widzenia i autostereoskopię.

**Systemy swobodnego punktu widzenia** (free viewpoint navigation) pozwalają widzowi wirtualnie poruszać się po scenie i interaktywnie wybierać punkt obserwacji (rys. 1). Oczekuje się, że wybór wirtualnego punktu obserwacji (zwanego kamerą wirtualną) będzie odbywać się w sposób płynny, co wyklucza wykorzystanie niewielkiej, przez co praktycznej, liczby widoków. Przy transmisji na odległość, zastosowanie to określa się mianem telewizji swobodnego punktu widzenia FTV (Free viewpoint TeleVision).



Rys. 1. Idea obrazu swobodnego punktu widzenia

**Autostereoskopia** jest techniką wyświetlania obrazu trójwymiarowego nie wymagającą używania specjalnych okularów. Efekt ten uzyskuje się przez jednoczesne wyświetlenie wielu widoków na monitorze autostereoskopowym (rys 2). W zależności od swojego położenia widz widzi różne pary widoków, dzięki czemu odnosi wrażenie przemieszczenia się w scenie trójwymiarowej. Dla uzyskania zadowalających efektów, liczba widoków wyświetlanych przez monitor autostereoskopowy powinna być duża. Obecnie najbardziej zaawansowane konstrukcyjnie monitory autostereoskopowe wyświetlają jednocześnie ponad 100 widoków.



Rys 2. Idea obrazu autostereoskopowego

Zarówno nawigacja swobodnego punktu widzenia, jak i autostereoskopia wymagają dostarczenia znacznej liczby widoków, których akwizycja, przetwarzanie i transmisja byłyby niepraktyczne. Sposobem, który umożliwia efektywny opis sceny trójwymiarowej jest zastosowanie reprezentacji **MVD (Multiview Video plus Depth)**. Reprezentacja sceny trójwymiarowej w formacie MVD składa się z ograniczonej liczby widoków oraz odpowiadających im map głębi, które przenoszą dodatkową informację na temat trójwymiarowej struktury sceny.

**Mapy głębi** są obrazami, których wartości odzwierciedlają odległości pomiędzy daną kamerą a punktami w scenie. Mapy głębi najczęściej ilustruje się w stopniach szarości (rys. 3), gdzie bliższe obiekty mają większą jasność, zaś dalsze – mniejszą. **Estymacja map głębi i ich kodowanie są fundamentalnymi zagadnieniami prezentowanej rozprawy.**



Rys 3. Wybrany widok z sekwencji trójwymiarowej „Poznan Carpark” (po lewej) oraz odpowiadająca mu mapa głębi (po prawej)

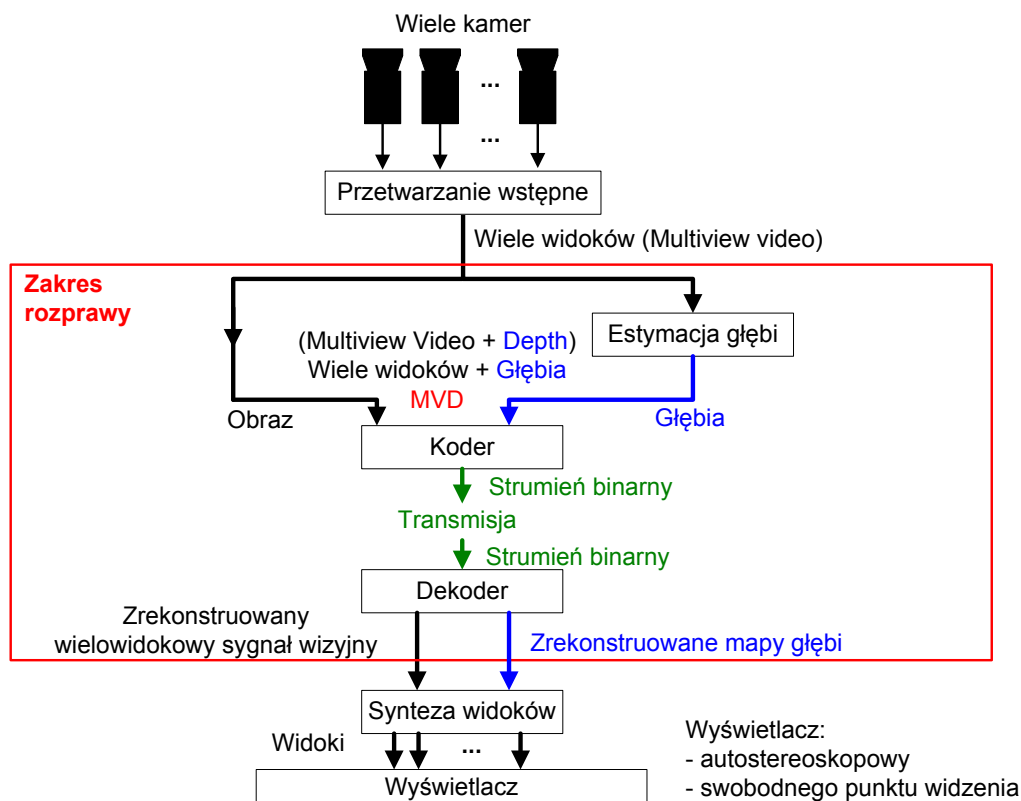
Wykorzystanie map głębi pozwala na generowanie tzw. **widoków wirtualnych**. Widok wirtualny, czyli widok pochodzący z kamery umiejscowionej wirtualnie w scenie, nie jest rejestrowany, a powstaje syntetycznie na podstawie treści zarejestrowanej przez kamery

rzeczywiste. Taka wirtualna kamera może być umieszczona w dowolnym położeniu – także w położeniu nieosiągalnym w rzeczywistym świecie. Zastosowanie reprezentacji MVD umożliwia znaczną redukcję liczby kodowanych widoków np. do 3 widoków, które są przesyłane wraz z odpowiadającymi im mapami głębi. Na podstawie tych widoków i towarzyszących im map głębi, pozostałe widoki mogą zostać zrekonstruowane po stronie dekodera/odbiornika.

Wymagane przez reprezentację MVD **mapy głębi trzeba wyznaczyć oraz efektywnie przesłać. Problemy te są rozważane w rozprawie. W szczególności:**

1. Mapy głębi wysokiej jakości są potrzebne przy tworzeniu treści zawierających obraz trójwymiarowy. W rozprawie rozpatrywana jest **algorytmiczna estymacja głębi** (rozdziały 6 i 7), zaś opracowanie nowych algorytmów i technik estymacji map głębi jest jednym z celów pracy. Jest to zagadnienie skomplikowane, stanowiące spore **wyzwanie badawcze**, gdyż w przeciwieństwie do np. grafiki komputerowej, rozpatrywane są tutaj złożone sceny naturalne, w których występują liczne problemy związane m.in. z odbiciami i przesłonięciami. Wprawdzie znane są sposoby pozyskiwania głębi z wykorzystaniem specjalnych urządzeń (np. z pomocą kamer time-of-flight, Microsoft Kinect, etc.) to ze względu na ograniczenia w ich stosowaniu (np. trudności w ich stosowaniu w otwartej przestrzeni) nie są one rozpatrywane w rozprawie.
2. Temat **spójności map głębi w dziedzinie czasu** jest związany zarówno z estymacją głębi, jak i jej kompresją. Brak spójności czasowej wyznaczanych map głębi objawia się migotaniem w widokach wirtualnych syntezowanych z ich wykorzystaniem. Ponadto, ograniczona spójność map głębi w dziedzinie czasu wpływa negatywnie na sprawność ich kompresji. Poprawa spójności czasowej estymowanych map głębi jest kolejnym z celów rozprawy (temat poruszany jest w podrozdziale 2.6).
3. Stosownie formatu MVD stwarza potrzebę efektywnego przesyłania map głębi. Ponieważ mapy głębi różnią się pod względem charakteru swojej treści od obrazów naturalnych, konieczne jest poszukiwanie **nowych, wydajnych metod ich kompresji**. Temat ten jest poruszany w rozdziałach 4 i 5 rozprawy.

Rozwiązania powyżej przedstawionych problemów są w rozprawie rozważane w kontekście trójwymiarowych systemów wizyjnych, rozumianych jako struktura przedstawiona na rys. 4. W szczególności, **rozprawa dotyczy więc estymacji głębi oraz kodowania i dekodowania głębi.**



Rys. 4. Zakres pracy (zaznaczony czerwoną linią) w kontekście systemu trójwymiarowego

## 1.2. Cele i tezy rozprawy

Rozprawa ma służyć rozwojowi stanu wiedzy dotyczącej nowej generacji systemów trójwymiarowych, wykorzystującej reprezentację MVD (Multiview Video plus Depth). **W rozprawie proponowane są nowe algorytmy estymacji głębi** mające zapewnić lepszą jakość głębi pod względem precyzji, dokładności, spójności czasowej i jakości syntezy widoków. **W zakresie kodowania głębi zaprezentowana jest nowa technika reprezentacji i kompresji głębi.** Tezy rozprawy są następujące:

- T1) Estymacja głębi może być usprawniona dzięki wykorzystaniu modelowania funkcji kosztu bazującego na regule maksymalizacji prawdopodobieństwa a posteriori.*
- T2) Precyzja i dokładność estymowanych map głębi może być zwiększona w kroku przetwarzania końcowego poprzez iteracyjne wstawianie wartości pośrednich głębi, kontrolowane za pomocą syntezy widoków.*
- T3) Spójność w dziedzinie czasu estymowanych map głębi może być poprawiona poprzez redukcję szumów w wejściowych sekwencjach wielowidokowych.*
- T4) Nieliniowa reprezentacja głębi umożliwia zwiększenie sprawności kompresji map głębi w systemach obrazu trójwymiarowego.*



### 1.3. Przegląd rozprawy

**Rozdział 1** wprowadza temat systemów wielowidokowych i trójwymiarowych. Zaprezentowano również metodologię, której użyto podczas badań. Szczególną uwagę poświęcono weryfikacji eksperymentalnej. Zaprezentowano wykorzystany zbiór sekwencji testowych, szeroko stosowanych także w literaturze. Opisano wkład autora w przygotowanie tych sekwencji, obejmujący adaptację algorytmów przetwarzania wstępnego.

**Rozdział 2** przedstawia przegląd współczesnego stanu wiedzy w zakresie estymacji głębi, zaś w szczególności zagadnienia istotne dla tematyki rozprawy: metody optymalizacji globalnej stosowane w estymacji map głębi i funkcje celu w nich wykorzystywane.

W **rozdziale 3** przedstawione są dokonania autora w dziedzinie badań dotyczących estymacji głębi. Pokazano nowe, teoretyczne sformułowanie funkcji celu algorytmów optymalizacyjnych estymujących mapy głębi, bazujące na zasadzie maksymalizacji prawdopodobieństwa a posteriori MAP (Maximum A posteriori Probability) wraz z nową techniką estymacji głębi wykorzystującą to sformułowanie. Następnie zaprezentowane są dwie nowe metody estymacji głębi. Pierwsza z nich wykorzystuje autorski algorytm MLH (Mid-Level Hypothesis) w celu zwiększenia precyzji i dokładności estymowanych map głębi. Druga z metod w celu poprawy spójności czasowej map głębi wykorzystuje redukcję szumów w wejściowych sekwencjach wizyjnych.

**Rozdział 4** omawia znane techniki kodowania głębi według współczesnego stanu wiedzy.

W **rozdziale 5** przedstawione są badania przeprowadzone przez autora w zakresie kodowania głębi. Zaprezentowana jest nowa idea zastosowania nieliniowej reprezentacji głębi dla kodowania głębi. Najpierw pokazana jest propozycja prostej funkcji przekształcenia nieliniowego, a następnie wyprowadzenie będące wynikiem rozważań systematycznych. Dla obu propozycji przedstawione są wyniki badań eksperymentalnych. Pod koniec, omówiono włączenie zaproponowanej techniki do międzynarodowych norm określających nową generację techniki kodowania ruchomego obrazu trójwymiarowego (ISO/IEC oraz ITU).

**Rozdział 6** przedstawia osiągnięcia autora związane z nową techniką kodowania obrazu trójwymiarowego, powstałą w Katedrze Telekomunikacji Multimedialnej i Mikroelektroniki. Wskazano wybitnie dobre właściwości kodeka KTMiM w porównaniu do innych propozycji zgłoszonych do konkursu „*Call for Proposals for 3D Video Coding Technology*” zorganizowanego przez grupę ISO/IEC MPEG.

**Rozdział 7** podsumowuje rozprawę. Wymienione są tutaj oryginalne osiągnięcia autora wraz z odniesieniem do postawionych tez.

## 1.4. Metodologia pracy

Celem rozprawy są badania nad poprawą efektywności kodowania map głębi i poprawą jakości map głębi estymowanych algorytmicznie.

W przypadku obu tych problemów **jedyną metodą rzetelnej oceny propozycji rozwiązań jest przeprowadzenie serii eksperymentów wykorzystujących sekwencje testowe**. Aby przeprowadzić te eksperymenty, propozycje przedstawione w rozprawie zaimplementowano i zintegrowano z istniejącymi pakietami oprogramowania, wymienionymi poniżej.

Ze względów wymienionych w rozdziale 2 rozprawy, za punkt odniesienia dla propozycji algorytmów estymacji głębi (zarówno dla ich implementacji jak i porównania) wybrano **algorytm DERS** zawarty w oprogramowaniu wzorcowym Depth Estimation Reference Software, stworzonym w ramach międzynarodowych prac grupy ISO/IEC MPEG. Algorytm ten jest powszechnie stosowany jako odniesienie w pracach naukowych na całym świecie.

W części rozprawy dotyczącej kodowania głębi, wyniki przedstawione są w odniesieniu do trzech technik kodowania trójwymiarowych sekwencji wizyjnych: **MVC+D** i **AVC-3D** (które powstawały równoległe z badaniami autora) oraz w odniesieniu do bazującej na **HEVC** techniki kodowania obrazu trójwymiarowego, współtworzonej przez autora i zgłoszonej przez KTMiM do wspomnianego konkursu MPEG. Na podstawie pakietów oprogramowania implementujących te trzy techniki kodowania przeprowadzono również implementację propozycji technik zawartych w rozprawie.

W obu przypadkach (estymacji głębi i kodowania głębi) propozycje zaimplementowane we wspomnianych pakietach oprogramowania wykorzystano do przeprowadzenia licznych i żmudnych eksperymentów. Wydajność zaproponowanych metod i algorytmów została przetestowana przede wszystkim za pośrednictwem pomiaru jakości uzyskiwanych widoków wirtualnych, syntezowanych z użyciem przetwarzanych map głębi (estymowanych lub kodowanych). Posłużono się miarami obiektywnymi (współczynnik PSNR) i subiektywnymi (testy subiektywne i miara MOS - Mean Opinion Score). Jest to sposób przyjęty w literaturze dla eksperymentów w zakresie nowych systemów obrazu trójwymiarowego. Wyniki tych badań posłużyły do sformułowania konkluzji i określenia kierunków dalszych prac.

## 1.5. Wielowidokowe sekwencje testowe

Rozprawa przedstawia wyniki badań naukowych autora, przeprowadzonych z wykorzystaniem szeroko stosowanych w literaturze sekwencji testowych. Zbiór ten, wraz

z wytycznymi dotyczącymi przeprowadzania eksperymentów został opracowany w ramach prac grupy MPEG przy aktywnym udziale autora.

Jednym z elementów wspomnianych prac było zbudowanie w Katedrze Telekomunikacji Multimedialnej i Mikroelektroniki eksperymentalnego systemu wielokamerowego (rys. 5). System ten posłużył do rejestracji sekwencji testowych o cechach unikalnych w skali światowej, tj. obraz wysokiej jakości HD, zarejestrowane sceny naturalne, wykorzystanie powszechnie używanych kamer telewizyjnych. Cechy te zadecydowały, że dwie z zarejestrowanych sekwencji („*Poznan Street*” i „*Poznan Hall*” – rys. 6) zostały włączone do zbioru sekwencji testowych rekomendowanego przez grupę MPEG.



Rys. 5. Części składowe eksperymentalnego systemu wielokamerowego powstałego w Katedrze Telekomunikacji Multimedialnej i Mikroelektroniki z aktywnym udziałem autora rozprawy: statyw wielokamerowy (po lewej) oraz system rejestracji (po prawej)



Rys. 6. Przykładowy widok pochodzący z sekwencji „*Poznan Hall*” (po lewej) oraz przygotowania do jej produkcji (po prawej) przez Katedrę Telekomunikacji Multimedialnej i Mikroelektroniki z aktywnym udziałem autora rozprawy

## Rozdział 2. Stan wiedzy w zakresie estymacji map głębi

Algorytmiczna estymacja głębi jest zagadnieniem badanym już od wielu lat. Mimo tego, jakość głębi estymowanej przez współczesne techniki nie jest satysfakcjonująca dla wielu zastosowań, np. dla systemów trójwymiarowych nowej generacji poruszanych w pracy.

W rozprawie przedstawione jest krótkie wprowadzenie do zagadnienia estymacji głębi wraz z przeglądem istotnej literatury. Nacisk położony jest na przypadek rozważany w pracy, czyli estymację gęstych map głębi (dense depth map estimation) przy liniowym ułożeniu kamer o równoległe ułożonych osiach optycznych. W takim przypadku estymację głębi realizuje się przez wyznaczanie jednowymiarowych rozbieżności stereoskopowych. Dlatego, mimo iż rozprawa odnosi się do terminu „głębia” (estymacji i jej kodowania), wiele z przedstawianych rozważań posługuje się terminem rozbieżności lub znormalizowanej rozbieżności.

Rozprawa przedstawia współcześnie znane metody lokalne (local methods) i globalne (global optimization methods) estymacji głębi. Metody lokalne wykorzystują miary podobieństwa fragmentów obrazów w celu znalezienia lokalnie optymalnego rozwiązania problemu odpowiedniości. W pracy wymienione są szeroko stosowane w literaturze miary takie jak SAD czy SSD, wraz z bardziej zaawansowanymi rozwiązaniami, takimi jak wykorzystanie mieszanek wielu miar [20]. Metody znane z lokalnej estymacji głębi są najczęściej wspomagane przez działający globalnie algorytm optymalizacyjny, tj. Graph Cuts lub Belief Propagation. Algorytm optymalizacyjny ma na celu tzw. regularyzację a więc zwiększenie spójności przestrzennej generowanych map głębi. Algorytmy te wykorzystują specjalnie zdefiniowaną funkcję celu, która może być zdekomponowana na dwa składniki: *DataCost* i *TransitionCost*. Funkcje te są szczegółowo rozważane w rozprawie i dlatego rozdział 2 przedstawia znane z literatury sposoby ich formułowania. Podkreślony jest niedomiar prac, które zajmują się wnikliwą analizą i wyprowadzeniem postaci tych funkcji, w szczególności z zastosowaniem reguły maksymalizacji prawdopodobieństwa a posteriori MAP (Maximum A posteriori Probability). Wspomniany **niedomiar prac naukowych stał się motywacją tego, aby w rozprawie podjąć badania dotyczące nowych sposobów sformułowania składników funkcji celu *DataCost* i *TransitionCost* z wykorzystaniem reguły MAP.**

Kolejnym zagadnieniem poruszonym w rozprawie jest dokładność i precyzja estymacji rozbieżności. W większości prac stosuje się tzw. punktową precyzję, która odpowiada poszukiwaniu odpowiedniości między obrazami z dokładnością do pojedynczego okresu

próbki obrazów. Znane są metody uzyskiwania lepszej dokładności, zwanej podpunktową, jednakże są one bardzo złożone. **Dlatego, w rozprawie proponuje się oryginalną metodę estymacji głębi z dokładnością podpunktową, opartą na autorskim algorytmie MLH (Mid-level Hypothesis), która cechuje się mniejszymi wymaganiami na nakłady obliczeń, jednocześnie bez uszczerbku dla jakości uzyskiwanych map głębi.**

Następny temat poruszany w rozprawie to spójność map głębi w dziedzinie czasu. Termin ten jest powszechnie stosowany w literaturze dla wyrażenia poziomu losowych fluktuacji wartości głębi w wygenerowanych mapach głębi. Mapa głębi spójna czasowo to więc taka, w której zmiany wartości głębi są związane jedynie z prawdziwym ruchem w scenie. Nie jest to definicja ścisła, jednakże rozprawa nie podejmuje próby definiowania pojęć ogólnych, gdy nie jest to bezpośrednio potrzebne. Zamiast tego, w pracy wykorzystywane są praktyczne wskaźniki mierzące spójność czasową map głębi, tj. korelacja w czasie, czy podatność na kompresję stratną z kompensacją ruchu.

Z literatury znane są różne techniki zwiększania spójności czasowej i wykorzystujące różne modyfikacje algorytmów estymujących głębię tak, aby uwzględniały one poprzednio wyznaczone wartości map głębi z poprzednich chwil czasowych. **W znanych pracach brak jednak technik wykorzystujących ideę zaproponowaną przez autora rozprawy, po raz pierwszy opublikowaną w artykułach [3] i [23], gdzie problem niespójności map głębi jest rozwiązywany z wykorzystaniem redukcji szumów.** Niniejsza rozprawa prezentuje badania (rozdział 3.8) będące kontynuacją tych wcześniejszych prac autora.

## Rozdział 3. Proponowane nowe metody estymacji głębi

Rozdział omawia nowe, autorskie metody estymacji map głębi.

W pierwszej kolejności zaproponowano, aby funkcję celu, stosowaną w algorytmach optymalizacyjnych stosowanych do estymacji map głębi, przeformułować w stosunku do określeń znanych z literatury i dzięki temu usprawnić estymację map głębi. Najpierw, przedstawione jest teoretyczne sformułowanie estymacji map głębi wykorzystujące regułę maksymalizacji prawdopodobieństwa a posteriori MAP (Maximum A posteriori Probability). Pokazane są założenia, jakie muszą zostać spełnione, aby możliwe było otrzymanie wyprowadzenia znanych z literatury miar podobieństwa punktów SAD i SSD. Założenia te są następnie weryfikowane. Na podstawie uzyskanych wyników, pokazano bardziej ogólne wyprowadzenie wzoru określającego składnik *DataCost* funkcji celu. Przedstawiona jest empiryczna metoda pomiaru parametrów zaproponowanego sformułowania. Zademonstrowano ją z wykorzystaniem sekwencji testowych.

Następnie, w dalszej części rozdziału, w podobny sposób proponowane jest probabilistyczne sformułowanie składników funkcji celu *TransitionCost*. Także tutaj niezbędne parametry wykorzystywane w sformułowaniu zmierzono w sposób empiryczny z użyciem sekwencji testowych.

Ostatecznie, na podstawie przedstawionych składników funkcji celu - *DataCost* i *TransitionCost*, w rozprawie zaproponowano **nową autorską metodę estymacji głębi**, którą także poddano badaniom eksperymentalnym. Pokazano, iż dzięki zaproponowanemu nowemu sformułowaniu **estymacja map głębi została usprawniona**.

W podrozdziale 3.7 zaproponowano nową metodę estymacji głębi, **wykorzystującą autorski algorytm MLH (Mid-Level Hypothesis)**. Umożliwia on **zwiększenie precyzji i dokładności uzyskiwanych map głębi** a także **redukcję złożoności obliczeniowej** samej estymacji (map głębi) o dokładności podpunktowej.

Ostatnim punktem rozdziału jest prezentacja metod poprawy spójności map głębi w dziedzinie czasu z wykorzystaniem technik redukcji szumu. Pokazane są wyniki oceny eksperymentalnej zaproponowanej metody. Wskazują one, że **redukcja szumu prowadzi do poprawy spójności czasowej estymowanych map głębi**.

### 3.1. Wyprowadzenie składnika *Data Cost* funkcji celu wykorzystujące regułę MAP

W pracy przedstawione jest nowe sformułowanie składnika *DataCost* funkcji celu algorytmów optymalizacyjnych wyznaczających mapy głębi. Sformułowanie to wykorzystuje przedstawione wyprowadzenie teoretyczne bazujące na regule maksymalizacji prawdopodobieństwa a posteriori MAP (Maximum A posteriori Probability).

Wspomniane wyprowadzenie rozważa przypadek dwóch kamer z równoległymi osiami optycznymi. Obrazy z lewego  $L_{x,y}$  i z prawego widoku  $R_{x,y}$  posiadają tę samą szerokość  $W$  i tę samą wysokość  $H$ .

Zgodnie z regułą MAP, poszukiwana jest rozbieżność  $d_{x,y}$  dla każdej pary współrzędnych  $x, y$  (w prawym widoku) która maksymalizuje prawdopodobieństwo  $p(d_{x,y})$  pod warunkiem zaobserwowania wartości luminancji w obu widokach (a posteriori). Prawdopodobieństwo to oznaczone jest w rozprawie przez  $p_{x,y,d}$ :

$$p_{x,y,d} \equiv p\left(d_{x,y} \mid (L_{1,y}, L_{2,y}, \dots, L_{W,y}, R_{1,y}, R_{2,y}, \dots, R_{W,y})\right), \quad (1)$$

zaś poszukiwana wartość rozbieżności  $d_{x,y}^*$ , wedle reguły MAP:

$$d_{x,y}^* = \max_{\arg d} (p_{x,y,d}). \quad (2)$$

Rozprawa, korzystając z założeń dotyczących m.in. niezależności statystycznej próbek luminancji, przedstawia wyprowadzenie wzoru na  $p_{x,y,d}$ , który może być wykorzystany w praktycznym algorytmie estymacji głębi:

$$p_{x,y,d} = \frac{1}{p(L_{x,y}) \cdot p(R_{x,y})} \cdot p\left(\left(L_{x+d_{x,y},y}, R_{x,y}\right) \mid d_{x,y}\right) \cdot p(d_{x,y}), \quad (3)$$

$$\begin{aligned} DataCost_{x,y}(d_{x,y}) = & 10 \cdot \log\left(p\left(\left(L_{x+d_{x,y},y}, R_{x,y}\right) \mid d_{x,y}\right)\right) + \\ & -10 \cdot \log\left(p(d_{x,y})\right) + 10 \cdot \log\left(p(L_{x,y})\right) + 10 \cdot \log\left(p(R_{x,y})\right). \end{aligned} \quad (4)$$

Powyższe wyprowadzenie i jego wynik należą do **istotnych osiągnięć rozprawy**.

### 3.2. Funkcja celu *Data Cost* jako uogólnienie miar SSD i SAD

Rozdział ten pokazuje, iż wyprowadzony wzór składnika funkcji celu *DataCost* stanowi uogólnienie znanych z literatury miar: kwadratów różnic (a także SSD – Sum of Squared Differences) oraz wartości bezwzględnych różnic (a także SAD – Sum of Absolute

Differences). W tym celu, w rozprawie rozpatruje się założenie o występowaniu addytywnego szumu o rozkładzie gaussowskim (dla pierwszej z nich) oraz założenie o występowaniu addytywnego szumu o rozkładzie zgodnym z dystrybucją Laplace'a (dla drugiej). Otrzymywana po wyprowadzeniu postać wzorów:

$$\log(p_{x,y,d}) = -\frac{1}{2\sigma^2} (L_{x+d,y} - R_{x,y})^2, \quad (5)$$

$$\log(p_{x,y,d}) = -\frac{1}{b} |L_{x+d_{x,y},y} - R_{x,y}|, \quad (6)$$

wskazuje, iż składnik **DataCost** funkcji celu sformułowany w podrozdziale 3.1 jest uogólnieniem miar SAD i SSD. Co więcej, jego wyprowadzenie wskazuje, iż wykorzystanie wspomnianych miar jest optymalne pod warunkiem spełnienia wyżej wymienionych założeń dotyczących szumu oraz założeń przyjętych w rozprawie w trakcie wyprowadzenia tych zależności tj. założeń o występowaniu lambertowskiego odbicia w scenie oraz o kompatybilności profili kolorów między kamerami, a także jednorodności rozkładów prawdopodobieństwa wartości próbek luminancji i rozbieżności. Założenia te poddano weryfikacji w następnym rozdziale.

### 3.3. Weryfikacja empiryczna założeń

W tym rozdziale przebadano prawdziwość założeń poczynionych w rozdziale 3.2.

Pierwsze z tych założeń dotyczy **niezależności próbek szumu pomiędzy ramkami** (w dziedzinie czasu i między widokowami). W celu jego przebadania, posłużoną się prostą metodą ekstrakcji szumu, opisaną szczegółowo w rozprawie. Użyto ją na zbiorze sekwencji testowych a wynik w postaci sekwencji szumowych poddano badaniom, kolejno wykorzystując współczynnik korelacji liniowej dla par próbek szumu pobranych z kolejnych ramek, wykorzystując 2-wymiarowe histogramy par próbek szumu pobranych z kolejnych ramek, oraz przeprowadzając test  $\chi^2$  na niezależność próbek szumu pobranych z kolejnych ramek. Wyniki przeprowadzonych eksperymentów były spójne i wykazały, że wartości próbek szumu w kolejnych ramkach **są niezależne**. Warto nadmienić, że zasadniczo ostatni z przeprowadzonych testów ( $\chi^2$ ) byłby tu wystarczającym dowodem, zaś dwa pierwsze testy przytoczono w rozprawie dla ilustracji uzyskanych wyników.

Następnie zajęto się badaniami rozkładów szumu. Stwierdzono, że rozkłady szumu przebadanych sekwencji zdecydowanie nie mają postaci dystrybucji Laplace'a, natomiast hipoteza o tym, iż są to rozkłady gaussowskie wymagała weryfikacji. Przeprowadzono ją z pomocą testu  $\chi^2$  zgodności rozkładów. Wykazał on, że mimo, iż kształt rozkładów szumu



dla niektórych sekwencji testowych jest zbliżony do krzywej gaussa to jednak hipotezę tą należy odrzucić. Jest to **istotny wynik rozprawy, gdyż podważono w ten sposób zasadność wykorzystywania miar SAD i SSD**, powszechnie stosowanych przy estymacji głębi.

W dalszej części rozdziału **przebadano rozkłady prawdopodobieństwa** występowania różnych wartości próbek luminancji i rozbieżności w rozważanych sekwencjach testowych. Stwierdzono, że bez wątplenia nie mają one rozkładu równomiernego, co również **podważyło zasadność wykorzystania miar SAD i SSD** przy estymacji głębi.

Zagadnienie **lambertowskiego modelu odbić w scenie i kompatybilności kolorów** przebadano zbiorczo. W tym celu wyznaczono dwuwymiarowe histogramy wartości luminancji punktów odpowiadających sobie w parze widoków. W procesie tym wykorzystano mechanizm syntezy widoków korzystający z map głębi odniesienia sekwencji testowych używanych w rozprawie. Zauważono, że założenia o lambertowskim modelu odbić w scenie i kompatybilności kolorów między widokami można uznać za prawdziwe jedynie w przypadku sekwencji syntetycznych, wygenerowanych komputerowo. W przypadku pozostałych sekwencji testowych, założeń tych nie można uznać za prawdziwe, co po raz kolejny **podważyło zasadność wykorzystania miar SAD i SSD** przy estymacji głębi.

### **3.4. Proponowane nowe sformułowanie dla składnika *Data Cost* funkcji celu**

W poprzednich rozdziałach wskazano, jakie założenia powinny być spełnione, aby można było uznać zastosowanie miar podobieństwa SAD i SSD za optymalne. Wykazano następnie, że wiele z tych założeń nie jest spełnionych dla sekwencji naturalnych. Fakt ten stał się inspiracją dla rozważań rozdziału 3.4.

W niniejszym rozdziale, odrzucono założenia, które nie przeszły weryfikacji eksperymentalnej i zaproponowano bezpośrednio wykorzystanie zaproponowanego na początku ogólnego sformułowania składnika *DataCost* funkcji celu. W szczególności, rozważono bezpośrednio wykorzystanie wzoru (21) wyprowadzonego w rozdziale 3.1 rozprawy (wzór (4) w niniejszym streszczeniu).

Następnym krokiem rozprawy było pokazanie metody wyznaczania parametrów zaproponowanego sformułowania, którą zademonstrowano z wykorzystaniem sekwencji testowych. Wskazano, iż **zaproponowane podejście nadaje się do praktycznej implementacji w algorytmach estymacji map głębi**.

### 3.5. Proponowane nowe sformułowanie dla składnika *Transition Cost* funkcji celu

Jak wspomniano w rozdziale 2.2 rozprawy, funkcja *TransitionCost* jest jednym z addytywnych składników funkcji celu wykorzystywanej przez algorytmy optymalizacji globalnej stosowane przy wyznaczaniu map głębi. Funkcja ta wprowadza dodatkowy koszt w funkcji celu dla map głębi, które nie są spójne przestrzennie. Najczęściej koszt ten oblicza się na podstawie wartości głębi sąsiadujących punktów. Im bardziej różni się głębokości takich sąsiadujących punktów, tym większy koszt wprowadzany jest do funkcji celu. Niestety, w metodach znanych z literatury koszt ten jest często wprowadzany heurystycznie, bez przeprowadzenia głębszych badań i analiz. Stało się to motywacją dla rozważań przeprowadzonych w rozdziale 3.5 rozprawy.

Rozdział przedstawia wyprowadzenie sformułowania probabilistycznego funkcji *TransitionCost*. Wykorzystano założenie, że funkcja ta może być zamodelowana na podstawie rozkładu prawdopodobieństwa dwóch sąsiadujących punktów  $p$  i  $q$  o rozbieżnościach odpowiednio  $d_p$  i  $d_q$ . Rozkład taki, oznaczony w rozprawie  $p_{2D}(d_p, d_q)$ :

$$TransitionCost_{p \rightarrow q}(d_p, d_q) = -10 \cdot \log(p_{2D}(d_p, d_q)) \quad , \quad (7)$$

został wyznaczony z wykorzystaniem rozpatrywanych sekwencji testowych i ich map głębi odniesienia. Uzyskane rozkłady empiryczne zilustrowano graficznie i porównano je ze stosowanymi powszechnie sformułowaniami składnika *TransitionCost* funkcji celu. Zauważono podobieństwo proponowanego sformułowania do modelu przyciętego-liniowego (truncated-linear). Podkreślono fakt, iż przewagą proponowanego sformułowania jest to, że nie wymaga ona ręcznej kalibracji parametrów, często mającej miejsce we współczesnych technikach estymacji głębi. Przytoczony jest przykład algorytmu DERS (Depth Estimation Reference Software), wykorzystującego heurystyczne sformułowanie funkcji celu i wymagającego ręcznego podania wartości parametrów sterujących, przykładowo tzw. „*Smoothing Coefficient*”.

### 3.6. Estymacja głębi z wykorzystaniem zaproponowanej funkcji celu

W rozdziałach 3.4 i 3.5 zaproponowano nowe sformułowanie składników *DataCost* i *TransitionCost* funkcji celu stosowanej przez algorytmy optymalizacyjne wyznaczające mapy głębi. Rozdział 3.6 przedstawia badania eksperymentalne algorytmów estymujących głębię, wykorzystujących to nowe sformułowanie. Badania te przeprowadzono na podstawie

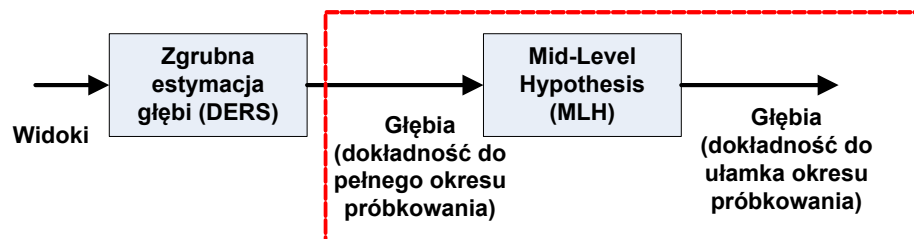
algorytmu wzorcowego DERS, będącego stosowanym w literaturze odniesieniem dla tego typu eksperymentów. Przeprowadzone eksperymenty wykazały, że **zaproponowane w rozprawie sformułowanie funkcji celu prowadzi do poprawy jakości estymowanej głębi**, co zmierzono przez pomiar jakości widoków wirtualnych, **wyrażony poprawą współczynnika PSNR do 2,8dB**.

Należy zaznaczyć, że poprawę jakości estymowanych map głębi uzyskano **bez ręcznej konfiguracji** algorytmu estymującego mapy głębi – wszystkie parametry potrzebne dla nowego sformułowania są wyznaczone automatycznie, podczas gdy oryginalny, niezmodyfikowany algorytm DERS wymaga ręcznej konfiguracji parametru „*Smoothing Coefficient*” wykorzystywanej w nim funkcji celu. **Jest to wielka zaleta zaproponowanego rozwiązania.**

### 3.7. Algorytm Mid-Level Hypothesis (MLH)

Z literatury znane są metody estymacji z dokładnością do ułamka okresu próbkowania, jednakże wymagają one dużych nakładów obliczeń. Jak wskazano w rozprawie, niestety, nakłady te nie są uzasadnione pod względem uzyskiwanej poprawy jakości estymowanych map głębi i syntezowanych z ich użyciem widoków wirtualnych.

Z tego względu, w rozprawie proponowana jest **nowa, nieznana z literatury, wydajniejsza metoda uzyskiwania map głębi o wysokiej dokładności**. Polega ona na dwukrokowej estymacji głębi. Najpierw mapa głębi wyznaczana jest zgrubnie a potem jest ona poprawiana za pomocą oryginalnego algorytmu MLH (Rys. 7).



Rys. 7. Idea zastosowania zaproponowanego algorytmu MLH

Idea algorytmu MLH polega na wyszukiwaniu, na przetwarzanej mapie głębi, sąsiadujących ze sobą punktów, różniących się wartościami głębi o jeden poziom kwantyzacji. Stawiana jest hipoteza, że punkty te mają w rzeczywistości wartość głębi pośrednią, a więc niereprezentowalną przy początkowej, ograniczonej dokładności mapy głębi. Hipoteza ta jest następnie weryfikowana poprzez porównanie jakości widoków

wirtualnych syntezowanych z uwzględnieniem początkowej i hipotetycznej wartości głębi. Następnie, jeśli hipoteza zostaje zweryfikowana pozytywnie, hipotetyczna wartość głębi staje się wartością faktyczną, zaś rozpatrywana hipoteza rozszerza się na kolejne sąsiadujące punkty, aż do iteracyjnego sprawdzenia wszystkich hipotez. Zaproponowaną, dwukrokową technikę estymacji głębi, wykorzystującą autorski algorytm MLH porównano eksperymentalnie z oryginalnym, niezmodyfikowanym, jednokrokovym algorytmem DERS. Oba algorytmy wykorzystano do uzyskania map głębi o dokładności do  $\frac{1}{4}$  okresu próbkowania obrazu. Wyniki przedstawione w rozprawie, wskazują dwojaki rodzaj zysków możliwych do otrzymania dzięki zastosowaniu algorytmu MLH. Zostały one wykazane poniżej.

Po pierwsze, zastosowanie algorytmu MLH umożliwia uzyskanie map głębi o dokładności do  $\frac{1}{4}$  okresu próbkowania o tej samej jakości co w podejściu klasycznym jednokrokovym przy dużo mniejszych kosztach obliczeniowych. **Czas wymagany dla estymacji głębi z wykorzystaniem algorytmu MLH jest około 3,4 do 3,8 razy krótszy (średnio 3,7 razy) niż w przypadku wykorzystania jednokrokovego algorytmu odniesienia DERS.**

Po drugie, zastosowanie algorytmu MLH pozwala na poprawę dokładności estymowanych map głębi, mierzoną **poprawą jakości PSNR widoków wirtualnych o średnio około 2 dB**, przy zachowaniu porównywalnego czasu obliczeń jak w przypadku zastosowania algorytmu odniesienia DERS, przy rozpatrywaniu głębi o dokładności wyrażonej pełnym okresem, próbkowania obrazu.

### **3.8. Poprawa spójności czasowej głębi za pomocą redukcji szumu**

W rozdziale tym przedstawiono wyniki badań dotyczących poprawy spójności czasowej estymowanych map głębi. Problem ten jest niedoceniony w badaniach, zaś w literaturze brak jest technik odniesienia. W rozprawie wykazano,  **iż poprawa spójności czasowej estymowanych map głębi możliwa jest poprzez redukcję szumu w obrazach wejściowych**. Wyniki prezentowane w rozprawie są kontynuacją wcześniejszych prac autora [3][23].

Rozprawa prezentuje dwie **autorskie metody redukcji szumu zaproponowane przez autora**. Pierwsza z nich to technika redukcji szumu w nieruchomym tle SBNR (Still Background Noise Reduction). Druga z wykorzystanych technik to redukcja szumu z kompensacją ruchu i przetwarzaniem kończącym MCNRR (Motion-Compensated Noise Reduction with Refinement).

Z wykorzystaniem opracowanych, wymienionych technik dokonano redukcji szumów w rozpatrywanych sekwencjach testowych i użyto ich do estymacji głębi. Wynikowe mapy głębi porównano z mapami głębi uzyskiwanymi bez zastosowania redukcji szumu.

Stwierdzono, iż zastosowanie redukcji szumu nie zmienia współczynnika PSNR synteżowanych widoków wirtualnych, co oznacza, że obie wersje map głębi (otrzymane z wykorzystaniem redukcji szumi i bez) **opisują tą samą scenę trójwymiarową**.

W dalszej części rozdziału przedstawiono wyniki testów subiektywnych, które wykazały wzrost jakości widoków wirtualnych o około 0,7 do 1,2 punktów w 11-punktowej skali MOS, co wskazuje na **zwiększenie spójności czasowej uzyskiwanych map głębi** przy zastosowaniu zaproponowanej redukcji szumu w sekwencjach wejściowych.

Pod koniec rozdziału zmierzono stopień kompresji danych map głębi przy wykorzystaniu hybrydowej techniki kodowania wielowidokowego obrazu ruchomego (MVC), wykorzystującej kompensację ruchu. Badania te wykazały, że **zastosowanie proponowanej metody estymacji głębi z redukcją szumu w sekwencjach wejściowych, skutkuje wyznaczeniem map głębi, które przy średnio 30% silniejszej kompresji pozwalają zachować tą samą jakość reprezentacji sceny**. Ponieważ porównywano kompresję map głębi reprezentujących tę samą scenę, zaś obie wersje map głębi miały pozwalały na uzyskanie tej samej jakości obrazów wirtualnych przy syntezie widoków, to różnice w sprawności kompresji tych map głębi świadczą więc przede wszystkim o losowych, nie dających się przewidzieć, fluktuacjach w kodowanych mapach głębi. Jest to zgodne z rozumieniem terminu „spójności czasowej” w rozprawie i wskazuje, iż **osiągana 30% redukcja prędkości bitowej odpowiada zwiększeniu spójności czasowej estymowanych map głębi**.

## Rozdział 4. Stan wiedzy w zakresie kodowania map głębi

Współcześnie znane i powszechnie wykorzystywane są techniki kodowania sekwencji wielowidokowych z pomocą techniki MVC. Sprawność tej techniki jest niewystarczająca przy bardzo dużej liczbie kodowanych widoków i dlatego dla nowej generacji systemów obrazu trójwymiarowego rozważa się reprezentację MVD. Format MVD (Multiview Video plus Depth), rozważany w rozprawie, wykorzystuje mapy głębi do reprezentacji informacji o strukturze przestrzennej sceny trójwymiarowej. Stwarza to pole dla badań naukowych nad zwiększeniem efektywności kompresji map głębi.

**W trakcie międzynarodowych prac nad wyżej wymienionymi normami, kompresja głębi została uznana za ważny problem badawczy.** Specyfika kompresji map głębi różni się od kompresji obrazów naturalnych, przy której oczekuje się, że zdekodowany obraz będzie możliwie subiektywnie podobny do obrazu oryginalnego. W przypadku MVD mapy głębi nie są bezpośrednio oglądane przez widza a służą do algorytmicznej syntezy widoków wirtualnych np. dla monitorów autostereoskopowych lub nawigacji swobodnego punktu widzenia. Dlatego powszechnie przyjętą w literaturze metodą jest wyrażenie jakości zdekodowanych map głębi przez jakość uzyskiwanych widoków wirtualnych.

Istnieje wiele technik nadających się do kompresji map. Skrótowy przegląd tych technik można znaleźć w rozprawie, jednakże należy zaznaczyć, że same techniki kodowania głębi znajdują się poza zakresem pracy, gdyż ta skupia się na reprezentacji map głębi. W szczególności, proponowane jest zastosowanie nieliniowej reprezentacji wartości map głębi, co jest równoważne wprowadzeniu niejednorodnej kwantyzacji w procesie kodowania.

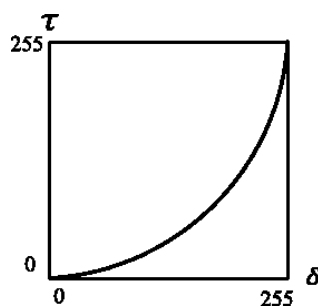
Idea nieliniowego przekształcenia map głębi występuje w literaturze, **jednakże opisywane metody różnią się od tej opisywanej w rozprawie.** W literaturze rozważa się nieliniowe przekształcenie głębi bez jakiegokolwiek odniesienia do kompresji (badając jedynie wpływ na jakość wizualną reprezentowanych map głębi) lub w odniesieniu kompresji, jednakże w celu uzyskania dokładniejszej kwantyzacji głębi na tle sceny lub innej reprezentacji sygnału predykcji (sygnał różnicowy) głębi a nie samych wartości próbek głębi, jak w przypadku rozprawy.

## Rozdział 5. Zaproponowana nieliniowa reprezentacja głębi dla kodowania

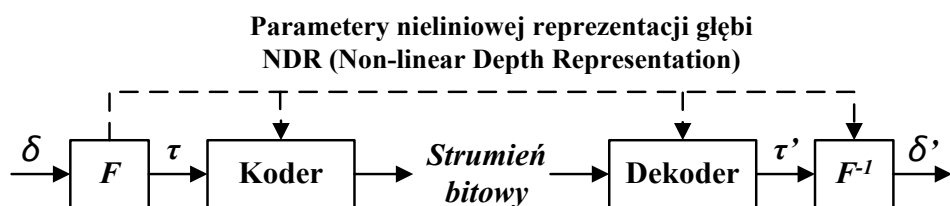
### 5.1. Idea nieliniowej reprezentacji głębi

Przy przesyłaniu i kompresji map głębi szeroko stosuje się kodowanie wartości głębi w postaci równomiernie skwantowanych wartości rozbieżności, znormalizowanych do zakresu 0...255. Jest to rozwiązanie uzasadnione, w przypadku, gdy pożądana jest jakość samych map głębi, jednakże w nowej generacji systemów obrazu trójwymiarowego, wykorzystujących reprezentację MVD (Multiview Video plus Depth), oczekuje się przede wszystkim wysokiej jakości widoków wirtualnych. Widoki te syntezuje się za pomocą map głębi i proces ten wykorzystuje się do oceny jakości map głębi. Niestety, wspomniana jednorodna kwantyzacja nie koresponduje z ludzkim sposobem postrzegania, który jest bardziej tolerancyjny dla błędów pojawiających się w tle syntezowanych widoków niż na ich pierwszym planie.

Z wyżej wymienionych względów, w rozprawie proponuje się zastosowanie nieliniowej reprezentacji głębi, która pozwala odwzorowywać bliskie obiekty dokładniej niż te dalsze (rys. 8). Rozprawa, z pomocą badań naukowych, wykazuje, że **proponowane podejście prowadzi do poprawy sprawności kompresji map głębi**. Dodatkową zaletą proponowanego rozwiązania jest to, iż proponowana technika może być łączona z innymi technikami kodowania (rys. 9).



Rys. 8. Nieliniowe przekształcenie głębi dokonywane przed kodowaniem



Rys. 9. Schemat stosowania zaproponowanej nieliniowej reprezentacji głębi

Proponowana niejednorodna kwantyzacja głębi dokonywana jest poprzez zastosowanie funkcji  $F$  przekształcającej wartości głębi na wejściu kodera oraz przez zastosowanie funkcji odwrotnej  $F^{-1}$  na wyjściu dekodera:

$$\tau = F(\delta) \quad ; \quad \delta' = F^{-1}(\tau') \quad , \quad (8)$$

gdzie  $\delta$  - nieprzekształcona wartość głębi na wejściu kodera,  $\tau$  – przekształcona wartość głębi w koderze,  $\delta'$  i  $\tau'$  - odpowiednie wartości w dekodrze.

## 5.2. Propozycja prostego przekształcenia nieliniowego

Dla sprawdzenia autorskiej koncepcji zastosowania przekształcenia nieliniowego przy kodowaniu głębi, w rozprawie najpierw zaproponowano definicję funkcji nieliniowej, w postaci funkcji potęgowej wzorowanej na korekcji gamma znanej z kodowania luminancji w klasycznych systemach wizyjnych:

$$\tau = \left( \frac{\delta}{\delta_{max}} \right)^\gamma \cdot \tau_{max} \quad , \quad (9)$$

gdzie  $\delta_{max}$  oraz  $\tau_{max}$  są maksymalnymi wartościami  $\delta$  oraz  $\tau$  (przykładowo 255 dla 8-bitowej precyzji reprezentacji próbek).

W ten sposób obiekty bliższe są reprezentowane dokładniej aniżeli odległe. Wspomniane nieliniowe przekształcenie z powodzeniem zastosowano w technice kodowania obrazu przestrzennego, zgłoszonej przez Katedrę Telekomunikacji Multimedialnej i Mikroelektroniki na konkurs „*Call for Proposals for 3D Video Coding Technology*” zorganizowany przez MPEG. Technika ta została później niezależnie przetestowana zgodnie z metodologią zawartą w punkcie 1.5.3 rozprawy.

Wyniki przedstawione w rozprawie, pokazują, że zastosowanie zaproponowanego przekształcenia potęgowego daje **poprawę jakości synteżowanych widoków wynoszącą około 1 punktu** w 11-punktowej skali MOS (Mean Objective Score).

## 5.3. Wyprowadzenie teoretyczne przekształcenia nieliniowego

Obiecujące wyniki, uzyskane z wykorzystaniem wstępnej propozycji nieliniowego przekształcenia przy kodowaniu głębi, stały się motywacją do podjęcia w rozprawie rozważań systematycznych. W punkcie 5.3 zaproponowano wyprowadzenie teoretyczne przekształcenia liniowego, przy którym założono, że krok kwantyzacji  $s(\tilde{\delta})$  powinien się zmniejszać wraz z rosnącą wartością rozbieżności  $\delta$  (znormalizowanej do przedziału  $[0;1]$ ). Przyjęto również,



że krok kwantyzacji  $s(\tilde{\delta})$  powinien się zmieniać jednorodnie w całym zakresie rozbieżności. Przekłada się to na sytuację, gdy przesunięcie się wzdłuż osi rozbieżności  $\tilde{\delta}$  o stałą wartość  $\Delta\tilde{\delta}$  skutkuje taką samą kompresją/ekspansją dynamiki, niezależnie od samej wartości  $\tilde{\delta}$ . Funkcją, która spełnia te wymagania jest funkcja eksponentialna:

$$s(\tilde{\delta}) = A \cdot e^{-\alpha \cdot \tilde{\delta}} \quad , \quad (10)$$

gdym:

$$s(\tilde{\delta} + \Delta\tilde{\delta}) = A \cdot e^{-\alpha \cdot (\tilde{\delta} + \Delta\tilde{\delta})} = s(\tilde{\delta}) \cdot e^{-\alpha \cdot \Delta\tilde{\delta}} \quad . \quad (11)$$

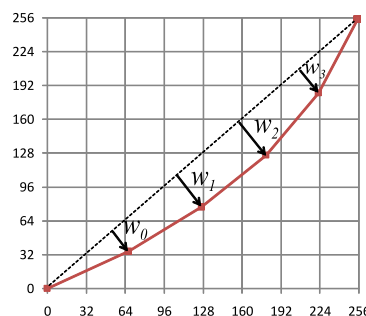
Na podstawie tych założeń, wyprowadzono wzór na poszukiwane przekształcenie:

$$\tau = F(\delta) = -\frac{\tau_{max}}{\alpha} \cdot \ln \left( 1 - \frac{\delta}{\delta_{max}} \cdot (1 - e^{-\alpha}) \right) \quad . \quad (12)$$

Jak stwierdzono, zastosowanie wymienionego przekształcenia w kompresji map głębi nie tylko **zapewnia poprawę jakości subiektywnej** (jak było to w przypadku przekształcenia potęgowego, prezentowanego w punkcie 5.3), ale także **poprawę jakości obiektywnej**, wyrażoną wzrostem współczynnika PSNR.

#### 5.4. Aproksymacja funkcji przekształcenia nieliniowego

Z pomocą badań w rozprawie wykazano, iż obie z zaproponowanych funkcji przekształcenia nieliniowego zwiększają sprawność kompresji map głębi. Ich zastosowanie w praktyce wymaga aby stronie dekodera zasygnalizowano jakiego kształtu funkcji nieliniowej użyto w koderze. Definiowanie, w sposób explicite, możliwego kształtu funkcji przekształcenia nieliniowego, któremu poddawana miałyby być głębia, ograniczyłoby elastyczność zaproponowanego rozwiązania. Dlatego, autor rozprawy zaproponował sposób aproksymacji kształtu funkcji przekształcenia z pomocą funkcji odcinkowo-liniowej (rys. 10). W rozwiązaniu tym, jedynie odchylenia węzłów są niezbędne dla odtworzenia kształtu krzywej przekształcającej, zarówno w koderze jak i dekodrze.



Rys. 10. Funkcja przekształcenia nieliniowego zdefiniowana przez wartości odchylenia od diagonalnej mierzonych względem równomiernie rozłożonych węzłów

## 5.5. Wyniki eksperymentalne kodowania głębi

Zaproponowana technika kodowania głębi z wykorzystaniem nieliniowego przekształcenia została przebadana subiektywnie i obiektywnie. Uzyskane wyniki, pokazane w rozprawie, przedstawiono grupie ISO MPEG wraz z propozycją nowej techniki kompresji map głębi, pod nazwą NDR (Non-linear Depth Representation). W serii eksperymentów eksploracyjnych [51][53], międzynarodowe gremium ekspertów potwierdziło uzyskane wyniki i oceniło pozytywnie przydatność zgłoszonej propozycji pod kątem jej włączenia do powstających norm kodowania treści trójwymiarowych, dotyczących technik MVC+D i AVC-3D.

Przeprowadzone eksperymenty wykazały, że zastosowanie proponowanej techniki NDR pozwala w znacznym stopniu usprawnić kodowanie głębi. Przy uwzględnieniu sekwencji, gdzie technika NDR była użyta (czym sterowała heurystyczna reguła zaproponowana w podrozdziale 5.4 rozprawy), osiągnięta **redukcja prędkości bitowej wynosi średnio około 24%** dla MVC+D i **22%** dla AVC-3D, przy zachowaniu tej samej jakości kodowanej głębi. Przy uwzględnieniu wszystkich sekwencji w zbiorze (także tych, gdzie technika NDR nie była użyta) **średnia osiągnięta redukcja prędkości bitowej wynosi około 10%** (zarówno dla MVC+D i AVC-3D).

Dla weryfikacji zaproponowanej metody NDR przeprowadzono również testy subiektywne [50], zgodne z metodologią DSIS (Double Stimulus Impairment Scale). Testy te, zapewniły dodatkowe potwierdzenie, że **proponowana metoda zwiększa efektywność kompresji**.

## 5.6. Nieliniowa reprezentacja głębi w międzynarodowych normach

Zaproponowana przez autora technika nieliniowej reprezentacji głębi NDR (Non-linear Depth Representation) **została przyjęta do rozszerzeń międzynarodowej normy opisującej technikę kodowania AVC (ISO/IEC IS 14496-10 i ITU Rec. H.264 ), znanych pod nazwami MVC+D i AVC-3D**. Pierwsze z tych rozszerzeń zostało już sfinalizowane jako Aneks „I” specyfikacji AVC, zaś drugie, zapewne pod koniec roku 2014, oficjalnie stanie się Aneksem „J” tejże specyfikacji.

## Rozdział 6. Nowa technika kodowania sekwencji trójwymiarowych

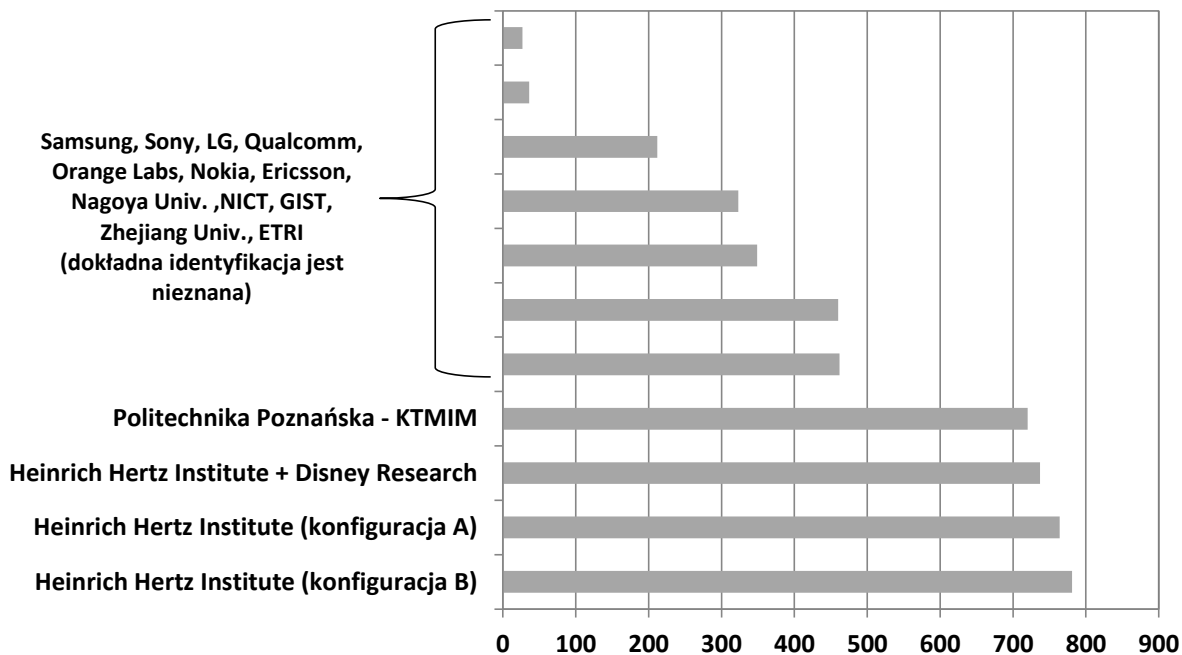
W roku 2011 grupa MPEG, działająca z ramienia ISO/IEC, ogłosiła konkurs „*Call for Proposals on 3D Video Coding Technology*” (CfP). Jego celem było wyłonienie najlepszej z dostępnych technik kompresji scen trójwymiarowych.

W konkursie tym wzięły udział instytuty badawcze i firmy takie jak: Microsoft, Samsung, Sony, LG, Qualcomm, Orange Labs, Nokia, Ericsson, Disney Research, Fraunhofer Institute for Telecommunications - Heinrich Hertz Institute, jak i uniwersytety: Aachen University (Niemcy), Nagoya (Japonia), National Institute of Information and Communications Technology (NICT, Japonia), Gwangju Institute of Science and Technology (GIST, Chiny), Zhejiang University of Science and Technology (Chiny), Electronics and Telecommunications Research Institute (ETRI, Korea) i Katedra Telekomunikacji Multimedialnej i Mikroelektroniki (KTMiM) Politechniki Poznańskiej.

**Autor rozprawy był jednym z twórców propozycji kodeka z KTMiM.** Propozycję tą opisano już w dokumencie MPEG [55] i publikacjach konferencyjnych [14][15][12][17][13], zaś w pracy [1] zawarto jej szczegółowy opis. Dlatego w rozprawie zawarto jedynie krótki opis propozycji tego kodeka, uwypuklający osiągnięcia samego autora rozprawy.

### 6.1. Porównanie z innymi kodekami

We wspomnianym konkursie wyszczególniono dwie kategorie propozycji: opierających się na technice kodowania AVC i opierających się na technice HEVC. W pierwszej z nich było 12 zgłoszeń, zaś w drugiej 11. Wszystkie zgłoszone propozycje były poddane ocenie obiektywnej i subiektywnej przez niezależne laboratoria badawcze. Wyniki, podsumowane przez grupę MPEG pod koniec roku 2011, wyłoniły dwa najlepsze rozwiązania w kategorii HEVC, przygotowane przez Heinrich Hertz Institute (HHI) we współpracy z Disney Research Labs (w kilku wariantach) oraz przez Katedrę Telekomunikacji Multimedialnej i Mikroelektroniki (KTMiM) Politechniki Poznańskiej. Wysoki poziom kodeka KTMiM został w rozprawie przedstawiony na tle innych zgłoszeń na konkurs za pomocą tzw. relacji przewyższania (outranking). Relacja ta (Rys. 11) przedstawia ilość przypadków, w testach subiektywnych, w których dana propozycja była najlepsza wśród badanych. Porównano również narzędzia kompresji wykorzystywane przez czołowe propozycje konkursowe.



Rys. 11. Relacja przewyższania (outranking) propozycji różnych uczestników konkursu

## 6.2. Struktura zaproponowanego kodeka

Kodek trójwymiarowy opracowany w KTMiM został zgłoszony do konkursu w kategorii kompatybilności z HEVC. W szczególności, jeden z widoków (zwany widokiem bazowym) jest kodowany w sposób kompatybilny ze składnią HEVC. Pozwala to na ekstrakcję widoku bazowego przez dekodery monoskopowej techniki HEVC. Pozostałe widoki (zwane widokami pobocznymi) są kompresowane z wykorzystaniem nowych, zaproponowanych technik kompresji.

W kodowaniu i dekodowaniu widoków pobocznych wykorzystuje się informację zawartą w widokach już zakodowanych, które używane są jako widoki odniesienia. Zasadniczo, tylko widok bazowy jest kodowany w całości. W widokach pobocznych jedynie niewielka część powierzchni obrazu jest kodowana, zaś pozostała jej część jest rekonstruowana za pomocą syntezy widoków wirtualnych. Taki schemat kompresji nazwano kodowaniem regionów odsłoniętych (Disoccluded Region coding [1][15]).

Szczegółowy opis innych technik kompresji można znaleźć [1][2][14][15] oraz w zasadniczym tekście rozprawy. Poniżej przedstawiono udział autora w powstawaniu tych technik.

### 6.3. Udział autora w tworzeniu kodeka

Autor w znaczny sposób przyczynił się do opracowania kodera trójwymiarowego powstałego w Katedrze Telekomunikacji Multimedialnej i Mikroelektroniki w następujący sposób:

- Autor współtworzył ogólną **strukturę kodeka**.
- Autor wniósł istotny wkład w **dobór algorytmów**, które zostały zawarte w kodeku.
- Autor współtworzył schemat kompresji wykorzystujący podział kodowanej treści na **warstwę główną i warstwę szumu (narzędzie Layer separation)**, m.in. przez propozycję oryginalnej koncepcji i dostarczenie autorskiego algorytmu redukcji szumów (algorytm MCNRR opisany m.in. w rozdziale 3.8.2 rozprawy).
- Autor współuczestniczył w powstawaniu metody kodowania wykorzystującej **zunifikowaną reprezentację głębi (Unified Depth Representation)**, dostarczając m.in. autorski algorytm MLH (Mid-Level Hypothesis) opisany w rozdziale 3.7 rozprawy.
- Autor zaproponował i zaimplementował algorytm kodowania wykorzystujący **nieliniową reprezentację głębi NDR** opisany w rozdziale 5 rozprawy.

### 6.4. Wyniki eksperymentalne wydajności kodeka

Wszystkie propozycje kodeków, zgłoszone na konkurs zorganizowany przez grupę MPEG, zostały przebadane w testach subiektywnych i obiektywnych przez niezależne laboratoria badawcze. Wyniki tych badań, wskazujących wysoki poziom poznańskiego kodeka z KTMiM, zaprezentowano na początku rozdziału 6 rozprawy.

W części 6.4 tego rozdziału, omówione są także wyniki badań przeprowadzonych przy współudziale autora już po rozstrzygnięciu konkursu. Miały one na celu ocenę udziału poszczególnych technik kompresji w wynikach osiągniętych przez kodek. Badania te przeprowadzono tak, aby możliwe było ich porównanie z wynikami konkursu MPEG. W szczególności, na szeroką skalę przeprowadzono badania subiektywne z użyciem monitorów trójwymiarowych.

Wyniki przeprowadzonych badań **wykazały, że zastosowanie techniki kodowania nieliniową reprezentacją głębi NDR (Non-linear Depth Representation) przyniosło redukcję prędkości bitowej o około 20% względem HEVC (simulcast), przy zachowaniu tej samej jakości subiektywnej**. Wszystkie techniki kompresji (włączając te nie objęte niniejszą rozprawą) **umożliwiają ok. 50%÷60% redukcję prędkości bitowej**. Można więc ocenić, że znaczna część zysków w kodeku trójwymiarowym pochodzi od opisanej w rozprawie autorskiej technice nieliniowej reprezentacji głębi (NDR).

## Rozdział 7. Podsumowanie rozprawy

### 7.1. Osiągnięcia związane z weryfikacją tez rozprawy

*Ad. T1) Estymacja głębi może być usprawniona dzięki wykorzystaniu modelowania funkcji kosztu bazującego na regule maksymalizacji prawdopodobieństwa a posteriori.*

W rozprawie zaproponowano nowe sformułowanie funkcji celu dla algorytmów optymalizacyjnych wyznaczających mapy głębi. Dwa składniki tego sformułowania - *DataCost* i *TransitionCost* - przedstawiono odpowiednio w rozdziałach 3.4 i 3.5. Posługując się wyprowadzeniem teoretycznym, wykorzystującym regułę maksymalizacji prawdopodobieństwa a posteriori MAP (Maximum A posteriori Probability), oraz badaniami empirycznymi, wykazano, że **wykorzystanie zaproponowanego, nowego sformułowania prowadzi do poprawy jakości estymowanych map głębi, wyrażonej wzrostem współczynnika PSNR do 2,8 dB**, liczonego dla widoków wirtualnych syntezowanych z użyciem map głębi wyznaczonych za pomocą zaproponowanej techniki. Odniesieniem był algorytm DERS reprezentujący aktualny stan wiedzy w zakresie estymacji głębi.

Wspomnieć należy, że jedną z głównych **zalet zaproponowanej metody** jest to, iż **nie wymaga ona manualnego doboru parametrów sterujących**, tak jak ma to miejsce np. w algorytmie odniesienia DERS.

Wyniki te **wspierają postawioną tezę o usprawnieniu estymacji głębi.**

*Ad. T2) Precyzja i dokładność estymowanych map głębi może być zwiększona w kroku przetwarzania końcowego poprzez iteracyjne wstawianie wartości pośrednich głębi, kontrolowane za pomocą syntezy widoków.*

W rozdziale 3.7 zaproponowano nową metodę estymacji głębi wykorzystującą autorski algorytm MLH (Mid-level Hypothesis). W rozprawie wykazano, iż jego użycie **poprawia precyzję i dokładność estymowanych map głębi.**

Zaproponowana metoda jako dane wejściowe wykorzystuje wyznaczoną wcześniej mapę głębi o niskiej precyzji np. wyznaczoną z dokładnością do pełnego okresu próbkowania. Następnie, taka mapa głębi jest przetwarzana za pomocą zaproponowanego algorytmu MLH. Wyszukiwane są sąsiadujące ze sobą punkty, różniące się wartościami głębi o jeden poziom kwantyzacji. Stawiana jest hipoteza, że punkty te mają w rzeczywistości wartość głębi pośrednią, a więc niereprezentowalną przy początkowej, ograniczonej dokładności mapy

głębi. Hipoteza ta jest następnie weryfikowana poprzez porównanie jakości widoków wirtualnych syntezowanych z uwzględnieniem początkowej i hipotetycznej wartości głębi. Następnie, jeśli hipoteza zostaje zweryfikowana pozytywnie, hipotetyczna wartość głębi staje się wartością faktyczną, zaś rozpatrywana hipoteza rozszerza się na kolejne sąsiadujące punkty, aż do iteracyjnego sprawdzenia wszystkich hipotez.

Uzyskane wyniki wskazują, że zastosowanie algorytmu MLH prowadzi do dwojakiego rodzaju korzyści, zależnie od sposobu interpretacji wyników. Z jednej strony, zaproponowany algorytm MLH pozwala uzyskać wyższą precyzję i dokładność estymowanych map głębi, co skutkuje **poprawą jakości widoku wirtualnego, zmierzoną wzrostem współczynnika PSNR o około 2dB**. Z drugiej strony, algorytm MLH pozwala uzyskać porównywalną taką samą jakość jak algorytm estymujący głębię z wyższą precyzją jednokrokowo, **przy znacznie - około 3,7 razy - mniejszych nakładach obliczeniowych**.

Wyniki te wspierają postawioną w rozprawie tezę.

***Ad. T3) Spójność w dziedzinie czasu estymowanych map głębi może być poprawiona poprzez redukcję szumów w wejściowych sekwencjach wielowidokowych.***

W rozdziale 3.8 rozprawy przedstawiono wyniki badań dotyczących zwiększania spójności czasowej estymowanych map głębi. Zaprezentowano autorską metodę estymacji głębi, polegającą na wykorzystaniu redukcji szumu w sekwencjach wejściowych a priori samej estymacji. Zaproponowaną metodę przebadano eksperymentalnie z wykorzystaniem dwóch, opracowanych przez autora, technik redukcji szumu: SBNR (Still Background Noise Reduction) i MCNRR (Motion-Compensated Noise Reduction with Refinement).

Posługując się zbiorem sekwencji testowych, wykazano iż zaproponowana metoda umożliwia uzyskanie **poprawy spójności czasowej estymowanych map głębi**, która zaowocowała **30% redukcją prędkości bitowej**, potrzebnej na przesłanie wyestymowanych map głębi, opisujących tę samą scenę trójwymiarową z tą samą jakością.

***Ad. T4) Nieliniowa reprezentacja głębi umożliwia zwiększenie sprawności kompresji map głębi w systemach obrazu trójwymiarowego.***

W rozdziale 5 zaprezentowano oryginalną ideę zastosowania nieliniowego przekształcenia dla reprezentacji i kodowania map głębi. Wykazano, że jej wykorzystanie przynosi znaczną **redukcję prędkości bitowej, polepszając tym samym sprawność kompresji**.

W pierwszej kolejności, pokazano proste przekształcenie, które znalazło zastosowanie w propozycji kodeka trójwymiarowego Katedry Telekomunikacji Multimedialnej

i Mikroelektroniki (KTMiM) zgłoszonej na konkurs „*Call for Proposals on 3D Video Coding Technology*” zorganizowany przez grupę MPEG w 2011r.

Sukces kodeka odniesiony w konkursie, jak i późniejsze badania badające udział poszczególnych technik kompresji, stały się motywacją dla systematycznego opracowania propozycji nieliniowego przekształcenia. Propozycja ta, korzystająca z wyprowadzenia teoretycznego, przedstawiona jest w rozprawie wraz szczegółami implementacji i wynikami weryfikacji eksperymentalnej. Jak pokazano, **wykorzystanie zaproponowanej techniki kodowania wykorzystującej nieliniową reprezentację głębi NDR (Non-linear Depth Representation) pozwala na uzyskanie około 25% redukcji prędkości bitowej strumienia głębi**. Przekłada się to na **poprawę jakości subiektywnej obrazu wirtualnego o około 1 punkt w skali MOS (11-punktowej)**.

Pod koniec rozdziału 5 podkreślono przyjęcie zaproponowanej techniki kodowania NDR do nowych, trójwymiarowych rozszerzeń **międzynarodowej normy ISO/IEC 14496-10 i rekomendacji ITU H.264 (występujących pod nazwami MVC+D i AVC-3D)**. Warto nadmienić, że idea nieliniowej reprezentacji głębi jest tematem zgłoszenia patentowego Politechniki Poznańskiej w Stanach Zjednoczonych i w Polsce [71].

## 7.2. Przegląd przeprowadzonych prac

W celu przeprowadzenia rzetelnej oceny osiągnięć rozprawy, autor przeprowadził **wyczerpującą serię żmudnych eksperymentów z wykorzystaniem sekwencji testowych**. Aby to umożliwić, zaimplementował i zintegrował zaproponowane metody estymacji głębi i kodowania głębi do następujących pakietów oprogramowania:

- oprogramowanie wzorcowe dla estymacji głębi DERS (Depth Estimation Reference Software) grupy ISO/IEC MPEG,
- oprogramowanie wzorcowe 3D-ATM dla technik kodowania obrazu trójwymiarowego bazujących na AVC (3D-ATM), stworzone przez grupę ISO/IEC MPEG,
- kodek trójwymiarowy stworzony przez Katedrę Telekomunikacji Multimedialnej i Mikroelektroniki (KTMiM), opierających się na technice kodowania monoskopowego HEVC.

Wykorzystane pakiety oprogramowania dla estymacji głębi (DERS) i syntezy widoków wirtualnych (VSRS) zawierają około 15 000 linii kodu źródłowego w języku C++. Fragmenty dodane lub dodane przez autora (zawierające zaproponowane techniki) obejmują około 1000 linii.



Oprogramowanie wzorcowe 3D-ATM zawiera około 230 000 linii kodu źródłowego w języku C++. Fragmenty zmodyfikowane lub dodane przez autora rozprawy (zawierające zaproponowane techniki) obejmują około 5 000 linii.

Kodek trójwymiarowy KTMiM [55], bazujący na oprogramowaniu wzorcowym ISO/IEC MPEG dla techniki kodowania HEVC, zawiera około 115 000 linii kodu źródłowego w języku C++, z których około 60 000 zostało dodanych podczas powstawania technik kompresji opracowanych przez zespół z Politechniki Poznańskiej. Fragmenty zmodyfikowane lub dodane przez autora rozprawy (zawierające proponowaną w rozprawie technikę nieliniowej reprezentacji głębi NDR) obejmują około 10 000 linii.

Niezależnie od powyższych prac, dla przeprowadzenia eksperymentów, autor stworzył zbiór narzędzi, bibliotek i skryptów programistycznych, zawierających około 20 000 linii kodu źródłowego w C++ i około 5 000 linii kodu źródłowego w języku Python.

**Sumarycznie, podczas przeprowadzania badań podsumowanych rozprawą, autor pracował z oprogramowaniem liczącym około 385 000 linii kodu źródłowego, wśród których około 41 000 zostało stworzonych przez autora.**

Eksperymenty przeprowadzone przez autora wymagały **sporych nakładów mocy obliczeniowej**.

Badania w zakresie analizy szumu w sekwencjach wizyjnych zajęły około 10 dni nieustannej pracy jednowątkowego komputera dużej mocy. Eksperymenty w zakresie estymacji głębi zajęły około 5 000 dni nieustannej pracy jednowątkowego komputera dużej mocy, zaś w zakresie kodowania głębi około 1 600 dni.

Sumarycznie, przeprowadzenie eksperymentów niezbędnych dla ukończenia rozprawy zajęło około **6 610 dni nieustannej pracy jednowątkowego komputera dużej mocy**.

Wykonanie tak dużej liczby badań było możliwe jedynie dzięki zastosowaniu wieloprocessorowych klastrów obliczeniowych, które umożliwiły zrównoleglenie obliczeń około 72-krotnie (sześć stacji roboczych, 12 wątków każda). Wspomnieć należy, iż większy stopień zrównoleglenia nie był możliwy z powodu ograniczeń pamięciowych – estymacja głębi dla sekwencji wysokiej jakości HD wymagała około 2 GB pamięci operacyjnej zaś każdy przebieg kodowania takiej sekwencji wymagał około 1 GB pamięci operacyjnej.

Nawet z wykorzystaniem klastra obliczeniowego, nakład pracy niezbędny dla przygotowania tak dużej liczby eksperymentów (np. dzielenie pracy na wiele maszyn i nadzorowanie obliczeń) stanowił znaczącą część prac nad rozprawą.

### 7.3. Głównie osiągnięcia

1. **Nieliniowa reprezentacja głębi NDR (Non-linear Depth Representation)**, która może być wykorzystana do zwiększenia efektywności kompresji map głębi w systemach trójwymiarowych. Propozycja została włączona do nowych rozszerzeń trójwymiarowych **międzynarodowej normy kodowania ISO/IEC 14496-10 i ITU H.264 - MVC+D i AVC-3D**. W rozprawie pokazano, że NDR przynosi znaczącą poprawę jakości subiektywnej syntezowanych widoków wirtualnych, wynoszącą nawet do **1 punktu w skali MOS (11-punktowej)**, oraz znaczącą poprawę sprawności kompresji, zapewniając **redukcję prędkości bitowej strumienia głębi o około 25%**.
2. **Metoda pozyskiwania map głębi spójnych w dziedzinie czasu z wykorzystaniem redukcji szumów w sekwencjach wejściowych**. W rozprawie wykazano, że użycie autorskich technik redukcji szumu: SBNR (Still Background Noise Reduction) i MCNRR (Motion-Compensated Noise Reduction with Refinement), prowadzi do zwiększenia spójności czasowej uzyskiwanych map głębi. Poprawę zmierzono badając jakość subiektywną widoków wirtualnych syntezowanych z użyciem uzyskiwanych map głębi i wyniosła ona około **1 punktu w skali MOS (11-punktowej)**. Odnotowano również poprawę spójności czasowej map głębi wyrażoną wydajnością kompresji – w tym przypadku stwierdzono **uzyskanie około 30% redukcji prędkości bitowej**.
3. **Algorytm poprawiania dokładności estymowanej głębi MLH (depth refinement by Mid-Level Hypothesis)**, który umożliwia uzyskanie głębi o wysokiej precyzji w kroku przetwarzania końcowego. Algorytm MLH, po pierwsze, umożliwia zwiększenie precyzji i dokładności głębi, wyrażone wzrostem **jakości do około 2dB PSNR** (jakość syntezowanego widoku wirtualnego). Po drugie, **MLH pozwala na wyznaczenie map głębi, o precyzji podpunktowej, w około 3,7 razy krótszym czasie** niż przy zastosowaniu klasycznej, jednokrokowej estymacji map głębi.
4. **Znaczny udział w przygotowaniu kodeka trójwymiarowego, stworzonego w Katedrze Telekomunikacji Multimedialnej i Mikroelektroniki**. Autor rozprawy wniósł istotny wkład w opracowanie i dobór technik które zostały ostatecznie zastosowane w kodeku. W szczególności, wymienić tu można zastosowanie algorytmów MLH i NDR wspomnianych powyżej. Poznański kodek trójwymiarowy został doskonale oceniony podczas rozwiązania konkursu „*Call for Proposals for 3D Video Coding Technology*” zorganizowanego przez grupę MPEG, **umożliwiając około 50%÷60% redukcję prędkości bitowej**.

## 7.4. Drugoplanowe osiągnięcia

1. Propozycja metody estymacji głębi, wykorzystującej nowe sformułowanie funkcji celu algorytmów optymalizacyjnych, bazujące na regule maksymalizacji prawdopodobieństwa a posteriori MAP (**Maximum A posteriori Probability**). Zaproponowana metoda umożliwia nienadzorowaną estymację głębi **bez potrzeby manualnego strojenia parametrów** przy zachowaniu tej samej jakości uzyskiwanych map głębi. W porównaniu do algorytmu odniesienia DERS z suboptymalnie dobranymi parametrami, zaproponowana technika daje poprawę jakości estymowanych map głębi, skutkującą w średnim wzroście PSNR o około **2,8dB** dla widoków wirtualnych syntezy z wykorzystaniem tych map głębi.
2. Dla wspomnianej wyżej metody został zaprezentowany sposób wyznaczania parametrów zaproponowanego modelu funkcji celu algorytmów optymalizacyjnych wykorzystywanych w estymacji map głębi. Przy tej okazji, dogłębnie przeanalizowano szum występujący w wielowidokowych sekwencjach wizyjnych i zagadnienie korelacji między-widokowej pomiędzy odpowiadającymi sobie próbkami.
3. Autor miał spory udział w przygotowaniu **wielowidokowych sekwencji testowych** stworzonych w Katedrze Telekomunikacji Multimedialnej i Mikroelektroniki. Udział ten dotyczył zarówno tworzenia samej treści jak i technicznych oraz naukowych zagadnień związanych z ich powstawaniem. Sekwencje te zostały włączone do zbioru wielowidokowych sekwencji testowych grupy ISO/IEC MPEG, używanego na całym świecie w badaniach dotyczących nowej generacji technik trójwymiarowych.

## Dokumentacja dorobku naukowego autora

(Stan w lipcu 2014)

### Cytowania – ogółem ponad 100 cytowań:

- w IEEE Transactions – 7,
- w innych czasopismach o zasięgu międzynarodowym – 18,
- w materiałach konferencjach IEEE – 15,
- w materiałach konferencji MPEG – 25,
- w materiałach innych konferencji międzynarodowych – 27,
- w innych publikacjach – 17.

Według systemu Google Scholar H-indeks wynosi 4.

### Artykuły w czasopismach o zasięgu międzynarodowym

- [1] M. Domański, O. Stankiewicz, K. Wegner, M. Kurc, J. Konieczny, J. Siast, J. Stankowski, R. Ratajczak, T. Grajek, „High Efficiency 3D Video Coding Using New Tools Based on View Synthesis”, IEEE Transactions on Image Processing, tom 22, nr 9, str. 3517-3527, 2013.

### Artykuły w wydawnictwach zwartych o zasięgu międzynarodowym

- [2] M. Domański, K. Klimaszewski, O. Stankiewicz, J. Stankowski, K. Wegner, „Efficient transmission of 3D video using MPEG-4 AVC/H.264 compression technology”, in S. Zeadally, E. Cerqueira, M. Curado, M. Leszczuk (edytorzy): Computer Communication Networks and Telecommunications, Lecture Notes in Computer Science, Springer-Verlag, tom 6157, str. 145-156, 2010.
- [3] K. Wegner, O. Stankiewicz, „Generation of temporally consistent depth maps using noise removal from video”, w L. Bolc, R. Tadeusiewicz, L.J. Chmielewski (edytorzy): Computer Vision and Graphics, Lecture Notes in Computer Science, Springer-Verlag, tom 6375, str. 292-299, 2010.
- [4] M. Domański, K. Klimaszewski, J. Konieczny, M. Kurc, A. Łuczak, O. Stankiewicz, K. Wegner, “An experimental free-view television system”, Image Processing and Communications Challenges, R. Choraś, A. Zabłudowski (edytorzy), Academy Publishing House EXIT Warsaw, str. 169-176, Warszawa 2009.

### Raporty międzynarodowych konkursów

- [5] O. Stankiewicz, J. Chmiel, M. Tłuczek, W. Świtła, “ReadIT: A portable text reading system for the blind people” - final report, 2nd place in World, IEEE Computing Society International Design Competition 2005, Waszyngton D.C., lipiec 2005.

## **Materiały konferencji międzynarodowych dostępne w bazie IEEEExplore**

- [6] M. Domański, A. Dziembowski, A. Kuehn, M. Kurc, A. Łuczak, D. Mieloch, J. Siast, O. Stankiewicz, K. Wegner, "Experiments on Acquisition and Processing of Video for Free-Viewpoint Television", 3DTV-Conference 2014 The True Vision Capture, Transmission and Display of 3D Video, Budapeszt, Węgry, 2014.
- [7] O. Stankiewicz, M. Domański, K. Wegner, „Analysis of noise in multi-camera systems", 3DTV-Conference 2014 The True Vision Capture, Transmission and Display of 3D Video, Budapeszt, Węgry, 2014.
- [8] K. Wegner, O. Stankiewicz, M. Domański, „Fast View Synthesis using platelet-based depth representation” 21th International Conference on Systems, Signals and Image Processing, IWSSIP 2014, Dubrownik, Chorwacja, maj 2014.
- [9] O. Stankiewicz, K. Wegner, M. Domański, „Nonlinear Depth Representation for 3D Video Coding”, IEEE International Conference on Image Processing (ICIP), Melbourne, Australia, wrzesień, 2013.
- [10] M. Domański, T. Grajek, D. Karwowski, K. Klimaszewski, J. Konieczny, M. Kurc, A. Łuczak, R. Ratajczak, J. Siast, O. Stankiewicz, J. Stankowski, K. Wegner, „New coding technology for 3d video with depth maps as proposed for standardization within MPEG”, 19th International Conference on Systems, Signals and Image Processing (IWSSIP), Wiedeń, Austria, 11-13 kwietnia 2012.
- [11] S. Cancino-Suarez, K. Klimaszewski, O. Stankiewicz, M. Domański, „Enhancement of stereoscopic depth estimation by the use of motion information”, IEEE 44th Southeastern Symposium on System Theory (SSST), Jacksonville, USA, 2012.
- [12] M. Domański, O. Stankiewicz, K. Wegner et al. “New Coding Technology for 3D Video within Depth Maps as Proposed for Standardization within MPEG”, International Conference on Systems, Signals and Image Processing (IWSSIP 2012), str. 415-418, Wiedeń, Austria, kwiecień, 2012.
- [13] M. Kurc, O. Stankiewicz, M. Domański “Depth map inter-view consistency refinement for multiview video”, Picture Coding Symposium, Kraków, Polska, 2012.
- [14] M. Domański, T. Grajek, D. Karwowski, K. Klimaszewski, J. Konieczny, M. Kurc, A. Łuczak, R. Ratajczak, J. Siast, O. Stankiewicz, J. Stankowski, K. Wegner, „Coding of multiple video+depth using HEVC technology and reduced representations of side views and depth maps”, 29th Picture Coding Symposium PSC2012, Kraków, Polska, 2012.
- [15] M. Domański, J. Konieczny, M. Kurc, R. Ratajczak, J. Siast, O. Stankiewicz, J. Stankowski, K. Wegner, „3D video compression by coding of disoccluded regions”, 2012 IEEE International Conference on Image Processing (ICIP), Orlando, USA, 30 września – 3 października 2012.
- [16] J. Stankowski, M. Domański, O. Stankiewicz, J. Konieczny, J. Siast, K. Wegner, „Extensions of the HEVC technology for efficient multiview video coding”, 2012 IEEE International Conference on Image Processing (ICIP), Orlando, USA, 30 września – 3 października, 2012.

- [17] O. Stankiewicz, M. Domański, K. Wegner, „Stereoscopic depth refinement by mid-level hypothesis”, IEEE International Conference on Multimedia & Expo, Singapur, lipiec 2010.
- [18] K. Wegner, O. Stankiewicz M. Domański, „Stereoscopic depth estimation using fuzzy segment matching”, 28th Picture Coding Symposium (PCS2010), Nagoya, Japonia, 8-10 grudnia 2010.
- [19] O. Stankiewicz, K. Wegner, M. Domański, „Error concealment for MVC and 3D video coding”, 28th Picture Coding Symposium (PCS2010), Nagoya, Japonia, 8-10 grudnia 2010.
- [20] K. Wegner, O. Stankiewicz, „Similiarity measures for depth estimation”, 3DTV-Conference 2009 The True Vision Capture, Transmission and Display of 3D Video, Poczdam, Niemcy, 4-6 maja 2009.
- [21] O. Stankiewicz, K. Wegner, „A hybrid technique for stereoscopic depth estimation in video”, International Conference on Signals and Electronic Systems (ICSES), Kraków, wrzesień 2008.
- [22] O. Stankiewicz , A. Łuczak, „Flexible processor architecture optimized for advanced coding algorithms”, Picture Coding Symposium 2007, Lizbona, Portugalia, 7-9 listopada 2007.
- [23] O. Stankiewicz, A. Łuczak, Antoni Roszak, „Temporal noise reduction for preprocessing of video streams in monitoring systems”, International Conference on Signals and Electronic Systems (ICSES'06), str. 231-234, Łódź, Polska, 18-20 września 2006.

### **Artykuły w polskich czasopismach**

- [24] O. Stankiewicz, K. Wegner, „System telewizji stereowizyjnej z wyznaczaniem mapy głębi”, Przegląd Telekomunikacyjny 4/2008, str. 213 - 216, kwiecień 2008.
- [25] O. Stankiewicz, K. Wegner, „Analiza dokładności syntezy obrazu w systemach telewizji wielowidokowej”, Przegląd Telekomunikacyjny 6/2009, str. 376-379, czerwiec 2009.
- [26] K. Wegner, K. Klimaszewski, O. Stankiewicz, J. Stankowski, M. Domański, „Przygotowanie wielowidokowych sekwencji wizyjnych dla badań nad telewizją trójwymiarową”, Przegląd Telekomunikacyjny 6/2010, str. 304-308, czerwiec 2010.
- [27] M. Domański, T. Grajek, D. Karwowski, K. Klimaszewski, R. Ratajczak, J. Konieczny, M. Kurc, A. Łuczak, R. Ratajczak, J. Siast, O. Stankiewicz, J. Stankowski, K. Wegner, „Poznański kodek obrazów trójwymiarowych”, Przegląd Telekomunikacyjny, tom 82, nr 2-3, str. 81-83, 2013.

**Najważniejsze z artykułów opublikowanych jako dokumenty grupy MPEG (ISO/IEC Moving Pictures Experts Group) oraz JCT-3V (ITU-T/ISO/IEC Joint Collaborative Team on 3D Video Coding Extension Development)**

- [28] O. Stankiewicz, M. Domański, K. Wegner, "Noise in 3D video sequences", ISO/IEC JTC1/SC29/WG11 MPEG2014/m34305, Sapporo, Japonia, lipiec 2014.
- [29] K. Wegner, O. Stankiewicz, "DERS Software Manual", ISO/IEC JTC1/SC29/WG11 MPEG2014/m34302, Sapporo, Japonia, lipiec 2014.
- [30] K. Wegner, O. Stankiewicz, M. Domański, "Fast View Synthesis through Depth Simplification ", ISO/IEC JTC1/SC29/WG11 MPEG2014/m34298, Sapporo, Japonia, lipiec 2014.
- [31] D. Rusanovskyy, F. C. Chen, J. Y. Lee, J.-L. Lin, O. Stankiewicz, T. Suzuki, D. Tian, "JCT-3V AHG Report: 3D-AVC Software Integration (AHG3)", Joint Collaborative Team on 3D Video Coding Extensions of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, Doc. JCT3V-I0003, m34405, 9th Meeting, Sapporo, Japonia, lipiec 2014.
- [32] K. Wegner, O. Stankiewicz, Krzysztof Klimaszewski, Marek Domański, „FTV EE3: Compression of FTV video with circular camera arrangement”, ISO/IEC JTC1/SC29/WG11 MPEG2014/M33243, Walencja, Hiszpania, kwiecień 2014.
- [33] M. Domański, A. Dziembowski, A. Kuehn, M. Kurc, A. Łuczak, D. Mieloch, J. Siast, O. Stankiewicz, K. Wegner, „Poznan Blocks – a multiview video test sequence and camera parameters for Free Viewpoint Television” ISO/IEC JTC1/SC29/WG11 MPEG2014/M32243, San Jose, USA, styczeń 2014.
- [34] K. Wegner, O. Stankiewicz, Adrian Dziembowski, D. Mieloch, Marek Domański, „First version of depth maps for Poznan Blocks multiview video test sequence”, ISO/IEC JTC1/SC29/WG11 MPEG2014/M32248, San Jose, USA, styczeń 2014.
- [35] K. Wegner, O. Stankiewicz, „Z-distance-based search range in Depth Estimation Reference Software”, ISO/IEC JTC1/SC29/WG11 MPEG2014/M32249, San Jose, USA, styczeń 2014.
- [36] D. Rusanovskyy, O. Stankiewicz, D. Tian, J. Y. Lee, J. Lin, „JCT-3V AHG Report: 3D-AVC Software Integration (AHG4)", Joint Collaborative Team on 3D Video Coding Extensions of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, Doc. JCT3V-G0004, m32551, 7th Meeting, San Jose, USA, styczeń 2014.
- [37] K. Wegner, O. Stankiewicz, M. Tanimoto, M. Domański, „Enhanced Depth Estimation Reference Software (DERS) for Free-viewpoint Television”, ISO/IEC JTC1/SC29/WG11 Doc. M31518, Genewa, Szwajcaria, październik 2013.
- [38] K. Wegner, O. Stankiewicz, M. Tanimoto, M. Domański, „Enhanced View Synthesis Reference Software (VSRS) for Free-viewpoint Television”, ISO/IEC JTC1/SC29/WG11 Doc. M31520, Genewa, Szwajcaria, październik 2013.

- [39] O. Stankiewicz, K. Wegner, M. Domański, „Optimized QP/QD curve for 3D coding with half and full resolution depth maps”, Joint Collaborative Team on 3D Video Coding Extension Development of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11 Doc. JTC3V-E0269, 5th Meeting, Wiedeń, Austria, 27 lipca – 2 sierpnia 2013.
- [40] K. Wegner, O. Stankiewicz, M. Domański, „AHG14: Comparison of half resolution depth map coding versus full resolution depth map coding in 3D-ATM”, Joint Collaborative Team on 3D Video Coding Extension Development of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11 Doc. JTC3V-D0080, 4th Meeting, Incheon, Korea Południowa, 20–26 kwietnia 2013.
- [41] O. Stankiewicz, K. Wegner, F.-C. Chen, C.-C. Lin, D. Rusanovskyy, „AHG12: Recommendation on MVC+D reference software”, Joint Collaborative Team on 3D Video Coding Extension Development of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11 Doc. JTC3V-C0164, Genewa, Szwajcaria, 17–23 stycznia 2013.
- [42] D. Rusanovskyy, Y. Lee, J. Lin, D. Tian, O. Stankiewicz, „AHG4: 3D-AVC Software Integration (AHG4)", Joint Collaborative Team on 3D Video Coding Extension Development of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, Doc. JCT3V-C0004, m27931, 3rd Meeting, Genewa, Szwajcaria, 17–23 stycznia 2013.
- [43] T. Grajek, O. Stankiewicz, K. Wegner, „AHG9: Correlation analysis between MOS data collected on stereoscopic and autostereoscopic displays”, Joint Collaborative Team on 3D Video Coding Extension Development of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11 Doc. JTC3V-C0202, Genewa, Szwajcaria, 17–23 stycznia 2013.
- [44] S. Shimizu, O. Stankiewicz, S. Sugimoto, H. Kimata, K. Wegner, M. Domański, „3D-HEVC HLS on depth definition”, Joint Collaborative Team on 3D Video Coding Extension Development of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11 Doc. JTC3V-B0164, 2nd Meeting, Szanghaj, Chiny, 13–19 października 2012.
- [45] K. Wegner, O. Stankiewicz J. Siast, M. Domański, „Independent intra-period coding in HEVC”, Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11 Doc. JTCVC-K0332 11th Meeting, Szanghaj, Chiny, 10–19 października 2012.
- [46] O. Stankiewicz, K. Wegner, M. Domański, „3D-HEVC with reduced resolution of depth”, Joint Collaborative Team on 3D Video Coding Extension Development of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11 Doc. JTC3V-B0183, 2nd Meeting, Szanghaj, Chiny, 13–19 października 2012.
- [47] O. Stankiewicz, K. Wegner, M. Domański, „Impact of View Synthesis Optimization (VSO) on depth quality”, Joint Collaborative Team on 3D Video Coding Extension Development of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11 Doc. JTC2-A0090, 1st Meeting, Sztokholm, Szwecja, 16–20 lipca 2012.



- [48] J. Siast, O. Stankiewicz, K. Wegner, M. Domański, „Independent intra-period coding in 3D-HTM”, Joint Collaborative Team on 3D Video Coding Extension Development of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11 Doc. JTC2-A0091, 1st Meeting, Sztokholm, Szwecja, 16–20 lipca 2012.
- [49] K. Wegner, O. Stankiewicz, J. Siast, „3D-CE1h results on Depth Map Disocclusion Coding by Poznan University of Technology”, ISO/IEC JTC1/SC29/WG11 Doc. M25014, Genewa, Szwajcaria, maj 2012.
- [50] K. Wegner, O. Stankiewicz, „Results of subjective evaluation in 3DV-CE2”, ISO/IEC JTC1/SC29/WG11 Doc. M25244, Genewa, Szwajcaria, maj 2012.
- [51] K. Wegner, O. Stankiewicz, J. Siast, „3D-CE2a results on Nonlinear Depth Representation”, ISO/IEC JTC1/SC29/WG11 Doc. M25019, Genewa, Szwajcaria, maj 2012,
- oraz:*
- K. Wegner, O. Stankiewicz, J. Siast, „3D-CE2h results on Nonlinear Depth Representation”, ISO/IEC JTC1/SC29/WG11 Doc. M25020, Genewa, Szwajcaria, maj 2012,
- jak również:*
- O. Stankiewicz, K. Wegner, M. Kurc, „3D-AVC-CE3 results on Nonlinear Depth Representation and Coding”, ISO/IEC JTC1/SC29/WG11, Doc. M23788, San Jose, USA, 6-10 lutego 2012.
- [52] K. Wegner, J. Siast, J. Konieczny, O. Stankiewicz, M. Domański, „Poznan University of Technology tools for 3DV coding integrated into 3D-HTM”, ISO/IEC JTC1/SC29/WG11 Doc. M23783, San Jose, USA, luty 2012.
- [53] O. Stankiewicz, K. Wegner, M. Domański, „Nonlinear depth representation – extended results”, ISO/IEC JTC1/SC29/WG11 Doc. M23791, San Jose, USA, 2012.
- [54] K. Wegner, J. Siast, J. Konieczny, O. Stankiewicz, M. Domański, „Poznan University of Technology tools for 3DV coding integrated into 3D-HTM”, ISO/IEC JTC1/SC29/WG11 Doc. M23783, San Jose, USA, luty 2012.
- [55] M. Domański, T. Grajek, D. Karwowski, K. Klimaszewski, J. Konieczny, M. Kurc, A. Łuczak, R. Ratajczak, J. Siast, O. Stankiewicz, J. Stankowski, K. Wegner, „Technical Description of Poznan University of Technology proposal for Call on 3D Video Coding Technology”, ISO/IEC JTC1/SC29/WG11, Doc. M22697, Genewa, Szwajcaria, listopad 2011.
- [56] M. Domański, T. Grajek, D. Karwowski, K. Klimaszewski, J. Konieczny, M. Kurc, A. Łuczak, R. Ratajczak, J. Siast, O. Stankiewicz, J. Stankowski, K. Wegner, „Multiview HEVC – experimental results”, Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11, Doc. M22147, JCTVC-G582, Genewa, Szwajcaria, listopad 2011.

- [57] K. Wegner, O. Stankiewicz, „Improved depth estimation with advanced occlusion handling”, ISO/IEC JTC1/SC29/WG11 Doc. M21365, Torino, Włochy, lipiec 2011.
- [58] K. Wegner, O. Stankiewicz, M. Domański, „3DV/FTV EE1 report on Poznan Carpark sequence – improved depth for extended range of frames”, ISO/IEC JTC1/SC29/WG11 Doc. M19391, Daegu, Korea Południowa, styczeń 2011.
- [59] K. Wegner, O. Stankiewicz, „Frame range extension of Poznan Street and Poznan Carpark sequences (3DV/EE1)”, ISO/IEC JTC1/SC29/WG11 Doc. M18506, Guangzhou, Chiny, październik 2010.
- [60] K. Wegner, O. Stankiewicz, K. Klimaszewski, M. Domański, „Comparison of multiview compression performance using MPEG-4 MVC and prospective HVC technology”, ISO/IEC JTC1/SC29/WG11 Doc. M17913, Genewa, Szwajcaria, lipiec 2010.
- [61] O. Stankiewicz, K. Wegner, M. Domański, „Estimation of temporally consistent depth maps using noise removal from video”, ISO/IEC JTC1/SC29/WG11 Doc. M17612, Dresden, Niemcy, kwiecień 2010.
- [62] J. Stankowski, K. Klimaszewski, O. Stankiewicz, K. Wegner, M. Domański, „Preprocessing methods used for Poznan 3D/FTV test sequences”, ISO/IEC JTC1/SC29/WG11 Doc. M17174, Kyoto, Japonia, styczeń 2010.
- [63] O. Stankiewicz, K. Wegner, M. Domański, „First version of depth maps for Poznan 3D/FTV test sequences”, ISO/IEC JTC1/SC29/WG11 Doc. M17176, Kyoto, Japonia, styczeń 2010.
- [64] O. Stankiewicz, K. Wegner, M. Wildeboer, „A soft-segmentation matching in Depth Estimation Reference Software (DERS) 5.0”, ISO/IEC JTC1/SC29/WG11 Doc. M17049, Xian, Chiny, październik 2009.
- [65] M. Domański, T. Grajek, K. Klimaszewski, M. Kurc, O. Stankiewicz, J. Stankowski, K. Wegner, „Poznań Multiview Video Test Sequences and Camera Parameters”, ISO/IEC JTC1/SC29/WG11 Doc. M17050, Xian, Chiny, październik 2009.
- [66] K. Wegner, O. Stankiewicz, „An enhancement of depth estimation reference software with use of soft-segmentation”, ISO/IEC JTC1/SC29/WG11 Doc. M16757, Londyn, Wielka Brytania, lipiec 2009.
- [67] K. Wegner, O. Stankiewicz, „Analysis of sub-pixel precision in depth estimation reference software and view synthesis reference software”, ISO/IEC JTC1/SC29/WG11 Doc. M16027, Lozanna, Szwajcaria, luty 2009.
- [68] O. Stankiewicz, K. Wegner, „Application of middle level hypothesis algorithm for improvement of depth maps produced by depth estimation reference software”, ISO/IEC JTC1/SC29/WG11 Doc. M16028, Lozanna, Szwajcaria, luty 2009.
- [69] O. Stankiewicz, K. Wegner, „Depth Map estimation software version 3”, ISO/IEC JTC1/SC29/WG11 Doc. M15540, Hanower, Niemcy, lipiec 2008.

## **Patenty w Stanach Zjednoczonych**

- [70] M. Domański, K. Klimaszewski, J. Konieczny, M. Kurc, R. Ratajczak, J. Siast, O. Stankiewicz, J. Stankowski, K. Wegner, "The image encoding method", US 8,761,527 B2, zgłoszenie 2012, data udzielenia patentu - 24 czerwca 2014,

*a także:*

zgłoszenie patentowe nr P.397008 w Polsce: „Sposób kodowania obrazu”.

## **Zgłoszenia patentowe w Stanach Zjednoczonych**

- [71] M. Domański, T. Grajek, J. Konieczny, M. Kurc, A. Łuczak, J. Siast, O. Stankiewicz, J. Stankowski, K. Wegner, „Method for coding of stereoscopic depth”, zgłoszenie nr 13680822,

*a także:*

zgłoszenie patentowe nr P.397016 w Polsce: „Sposób kodowania głębi stereoskopowej”.

- [72] M. Domański, J. Konieczny, M. Kurc, R. Ratajczak, J. Siast, O. Stankiewicz, J. Stankowski, K. Wegner, „Method for predicting the shape of an encoded area based on depth map”, zgłoszenie nr 13680740,

*a także:*

zgłoszenie patentowe nr P.397010 w Polsce: „Sposób predykcji kształtu obszaru kodowanego z wykorzystaniem map głębi”.

## **Pozostałe zgłoszenia patentowe w Polsce**

- [73] O. Stankiewicz, K. Wegner, M. Domański, „Sposób wyznaczania modeli przestrzennych z wykorzystaniem redukcji szumów w wejściowych sekwencjach wizyjnych”, zgłoszenie nr P-392496.

- [74] M. Domański, J. Konieczny, M. Kurc, R. Ratajczak, J. Siast, O. Stankiewicz, J. Stankowski, K. Wegner, „Sposób przetwarzania obrazu zsyntezowanego”, zgłoszenie nr P.397012.

- [75] M. Domański, T. Grajek, J. Konieczny, M. Kurc, J. Siast, O. Stankiewicz, J. Stankowski, K. Wegner, „Sposób alokacji prędkości bitowej przy kodowaniu sekwencji wielowidokowych z informacją przestrzenną”, zgłoszenie nr P.397014.

- [76] M. Domański, T. Grajek, J. Konieczny, M. Kurc, A. Łuczak, J. Siast, O. Stankiewicz, J. Stankowski, K. Wegner, „Sposób wyznaczania parametrów kwantyzacji sterujących dekwantyzatorem sygnału resztkowego obrazu z wykorzystaniem mapy głębi odpowiadającej dekodowanemu obrazowi”, zgłoszenie nr P.397013.

- [77] M. Domański, T. Grajek, J. Konieczny, M. Kurc, J. Siast, O. Stankiewicz, J. Stankowski, K. Wegner, „Zastosowanie cech sygnału różnicowego mapy głębi do sterowania koderem sekwencji wizyjnych z informacją przestrzenną”, zgłoszenie nr P.397015.
- [78] M. Domański, T. Grajek, K. Klimaszewski, J. Konieczny, M. Kurc, J. Siast, O. Stankiewicz, J. Stankowski, K. Wegner, „Sposób międzyobrazowej predykcji mapy odległości”, zgłoszenie nr P.397011.
- [79] M. Domański, T. Grajek, D. Karwowski, J. Konieczny, M. Kurc, J. Siast, O. Stankiewicz, J. Stankowski, K. Wegner, „Sposób kodowania parametrów kamer”, zgłoszenie nr P.397009.
- [80] M. Domański, T. Grajek, K. Klimaszewski, J. Konieczny, M. Kurc, J. Siast, O. Stankiewicz, J. Stankowski, K. Wegner, „Sposób przetwarzania map rozbieżności”, zgłoszenie nr P.397017.

#### **Międzynarodowe normy uwzględniające propozycje autora rozprawy**

- Annex I “Multiview and Depth video coding” of ISO/IEC 14496-10, Int. Standard “Generic coding of audio-visual objects – Part 10: Advanced Video Coding”, 8th Edition, 2013.
- Annex I “Multiview and Depth video coding” of ITU-T Rec. H.264, Edition 8.0, 2013.
- “Study Text of ISO/IEC 14496-10:2012/PDAM3 AVC compatible video-plus-depth extension”, ISO/IEC JTC1/SC29/WG11, Doc.N13332, Genewa, Szwajcaria, styczeń 2013.