

Politechnika Poznańska  
Wydział Elektroniki i Telekomunikacji  
Katedra Telekomunikacji Multimedialnej i Mikroelektroniki

Rozprawa doktorska

# **Algorytmy kompresji sekwencji wielowidokowych**

Krzysztof Klimaszewski

Promotor: prof. dr hab. inż. Marek Domański

Poznań 2012



# Streszczenie

Rozprawa dotyczy efektywnej kompresji sekwencji wielowidokowych wraz z mapami głębi odpowiadającymi kodowanym widokom. Prace prowadzone są z wykorzystaniem technologii MVC, powstałej jako technologia kompresji sekwencji wielowidokowych bez map głębi. Badane są możliwości wykorzystania MVC, bezpośrednio lub z niewielkimi modyfikacjami, do kompresji sekwencji wielowidokowych z mapami głębi.

W pracy zbadano metody doboru odpowiednich parametrów kodowania kodera MVC pod kątem uzyskiwania jak najlepszej jakości obrazów syntezowanych z użyciem zdekompresowanych danych. Opracowano techniki alokacji bitów w strumieniach wynikowych, pozwalające na automatyczne wyznaczenie parametrów kodowania map głębi w przypadku, gdy dane o mapach głębi i dane o obrazach kodowane są niezależnie, z użyciem kodera MVC.

Opracowano i przebadano metody zwiększenia efektywności kodowania sekwencji wielowidokowych z mapami głębi poprzez wykorzystanie korelacji pomiędzy obrazami i mapami głębi. Wykorzystany został nowy sposób reprezentacji obrazów z mapami głębi – reprezentacja czterokomponentowa. Uzyskane wyniki świadczą o możliwości zwiększenia efektywności kodowania dzięki wykorzystaniu korelacji pomiędzy obrazami i mapami głębi, głównie dla kodowania w zakresie małych prędkości bitowych.

# **Abstract**

The dissertation relates to efficient compression of multiview sequences with corresponding depth maps. The works concentrate on MVC – a technology intended primarily for compression of multiview video without depth. The possible applications, direct or including small changes, of MVC technology for compression of multiview sequences with depth maps are evaluated.

The research was focused on obtaining the appropriate MVC encoder settings for obtaining the best possible quality of a view synthesized with the use of decompressed data. Bit allocation techniques were developed, allowing automatic assignment of depth map encoding parameters for cases where views and depth maps are encoded separately with the use of MVC encoder.

Methods for increasing the efficiency of encoding of multiview sequences with depth maps were developed and tested. The methods are taking advantage of the correlation between views and depth maps. A new representation of views and depth maps was used, namely, the four component representation. The results obtained during tests prove the possibility of improving the efficiency of coding process by the use of correlation between views and depth maps, especially for low bitrate encoding cases.

## Spis akronimów wykorzystywanych w tekście

- AVC - *Advanced Video Coding* - zaawansowany kodek wizyjny
- CABAC - *Context-based Adaptive Binary Arithmetic Coding* - metoda kodowania danych wykorzystująca adaptacyjne kodowanie arytmetyczne
- CAVLC - *Context Adaptive Variable Length Coding* - metoda kodowania danych wykorzystująca adaptacyjne kodowanie kodami o zmiennej długości
- CBP - *Coded Block Pattern* - składnik strumienia wynikowego kodera MVC zawierający informacje o obecności niezerowych próbek transformaty zapisywanych do strumienia przy kodowaniu obrazów z kamer
- DCBP - *Depth Coded Block Pattern* - składnik strumienia wynikowego zmodyfikowanego kodera MVC zawierający informacje o obecności niezerowych próbek transformaty zapisywanych do strumienia przy kodowaniu map głębi
- FTV - *Free viewpoint TV* - telewizja swobodnego punktu widzenia
- HEVC - *High Efficiency Video Coding* - kodek nowego typu, charakteryzujący się bardzo wysoką efektywnością
- ISO - *International Organization for Standardization* - Międzynarodowa Organizacja Normalizacyjna
- MPEG - *Moving Pictures Experts Group* - grupa ekspertów pracujących nad metodami przetwarzania sekwencji wizyjnych w ramach organizacji ISO
- MVC - *Multiview Video Coding* - kodek przeznaczony do przetwarzania wielowidokowych sekwencji wizyjnych
- PSNR - *Peak Signal to Noise Ratio* - miara jakości obrazu bazująca na stosunku sygnału do szumu
- QD - parametr kwantyzacji wykorzystywany w kodowaniu map głębi przy zastosowaniu kodera MVC
- QP - parametr kwantyzacji wykorzystywany w kodowaniu obrazów z kamer przy zastosowaniu kodera MVC
- VLC - *Variable Length Coding* - metoda kodowania danych wykorzystująca kodowanie kodami o zmiennej długości

## Spis treści

1. Wprowadzenie.....	9
1.1. Zakres pracy.....	9
1.2. Założenia pracy.....	11
1.3. Cel i teza pracy.....	12
1.4. Przegląd pracy.....	12
2. Sekwencje wielowidokowe i mapy głębi.....	15
2.1. Otrzymywanie sekwencji wielowidokowych.....	15
2.2. Mapy głębi.....	16
2.3. Wyznaczanie map głębi.....	17
3. Kodowanie ruchomych obrazów wielowidokowych.....	21
3.1. Kompresja – wprowadzenie.....	21
3.2. Technologia zaawansowanego kodera wizyjnego AVC.....	24
3.2.1. Kodowanie wewnątrzobrazowe.....	25
3.2.2. Kodowanie międzyobrazowe.....	28
3.3. Metody kompresji sekwencji wielowidokowych.....	31
3.4. Koder MVC.....	33
3.5. Efektywność mechanizmów wprowadzonych w MVC.....	35
3.6. Przetwarzanie wstępne sekwencji wielowidokowych.....	35
3.7. Podsumowanie.....	40
4. Kodowanie map głębi.....	43
4.1. Przegląd literatury.....	43
4.2. Szczególne właściwości mapy głębi.....	45
4.3. Ocena jakości map głębi.....	47
5. Bezpośrednia metoda kodowania obrazów i map głębi.....	51
5.1. Wpływ zniekształceń kodowania obrazów i głębi na jakość sekwencji syntezowanych.....	51
5.1.1. Przegląd literatury.....	51
5.1.2. Cel badań.....	52
5.1.3. Metodologia badań.....	52
5.1.4. Wyniki eksperymentu.....	53
5.1.4.1. Efektywność kompresji obrazów i map głębi.....	53
5.1.4.2. Badanie jakości syntezy ze zdekodowanych danych.....	56

5.1.4.3. Badanie jakości syntezy ze zdecydowanych map głębi.....	65
5.1.5. Podsumowanie i wnioski.....	74
5.2. Wybór optymalnych par QP-QD.....	77
5.2.1. Metodologia badań.....	77
5.2.2. Wyniki eksperymentu.....	78
5.2.2.1. Efektywność kompresji w zależności od wyboru par QP-QD.....	79
5.2.2.2. Optymalne pary QP-QD.....	89
5.2.2.3. Wyznaczenie zależności $QD = f(QP)$ .....	112
Pomiar jakości przez porównanie z widokiem syntetycznym.....	113
Pomiar jakości przez porównanie z widokiem z rzeczywistej kamery.....	119
5.2.3. Weryfikacja opracowanych wzorów.....	123
5.2.4. Podsumowanie i wnioski.....	124
5.3. Podsumowanie.....	127
6. Łączne kodowanie obrazu i głębi.....	129
6.1. Wprowadzenie.....	129
6.2. Przegląd literatury.....	129
6.3. Kompresja głębi jako czwartej składowej obrazu.....	131
6.4. Metodologia badań.....	131
6.5. Mapy głębi o zmniejszonej rozdzielczości.....	133
6.6. Mapy głębi w pełnej rozdzielczości.....	135
6.6.1. Korelacja dla wewnątrzobrazowego kodowania luminancji i głębi.....	136
6.6.1.1. Podziały makrobloku i kierunki predykcji dla luminancji i głębi.....	137
6.6.1.2. Podział na bloki dla luminancji i głębi dla kodowania wewnątrzobrazowego.....	142
6.6.1.3. Kierunki predykcji dla luminancji i głębi w kodowaniu wewnątrzobrazowym.....	155
6.6.2. Wewnątrzobrazowe kodowanie głębi z wykorzystaniem trybów dla luminancji.....	161
6.6.2.1. Wyniki dla stałego QP.....	162
6.6.2.2. Wyniki dla stałego QD.....	169
6.6.2.3. Podsumowanie.....	176
6.6.3. Wewnątrzobrazowe kodowanie głębi z możliwością wyboru odrębnych trybów.....	177
6.6.3.1. Wyniki dla stałego QP.....	184
6.6.3.2. Wyniki dla stałego QD.....	195
6.6.3.3. Podsumowanie.....	206
6.6.4. Wspólne kodowanie wewnątrzobrazowe - podsumowanie.....	206
6.7. Predykcja międzyobrazowa w kodowaniu czterokomponentowym.....	207

6.8. Podsumowanie i wnioski.....	209
7. Podsumowanie.....	211
7.1. Podsumowanie pracy.....	211
7.2. Główne osiągnięcia.....	212
7.3. Możliwe kierunki dalszych badań.....	214
Aneks I.....	217
Aneks II.....	219
Aneks III.....	223
Aneks IV.....	225
Bibliografia.....	227



# 1. Wprowadzenie

## 1.1. Zakres pracy

Prezentowana rozprawa dotyczy zagadnień związanych z ruchomymi obrazami wielowidokowymi (ang. *multiview video*), które rejestrowane są przy użyciu wielu dokładnie ze sobą zsynchronizowanych kamer. Sekwencje obrazów wielowidokowych, nazywane w dalszej części pracy sekwencjami wielowidokowymi, są wykorzystywane w systemach wizji wielowidokowej drugiej generacji [Kubo\_07][Doma\_10a][Doma\_11]. W takich systemach, będących wciąż w początkowej fazie rozwoju, wykorzystuje się złożone metody prezentacji obrazu za pomocą wyświetlaczy autostereoskopowych [Park\_95][Sext\_99][Dose\_12] czy holoskopowych [Aggo\_10]. Wyświetlacze takie pozwalają na uzyskiwanie realistycznego wrażenia trójwymiarowości obserwowanej sceny, jak również pozwalają na oddanie efektu paralaksy, a dla poprawnego postrzegania efektu trójwymiarowego nie jest konieczne zakładanie specjalnych okularów. Mając na uwadze wymagania tego typu wyświetlaczy opracowuje się scenariusze wykorzystania technologii wizji wielowidokowej [Red\_02][Pfis\_04][Smol\_06]. Szczególnym zastosowaniem sekwencji wielowidokowych jest telewizja swobodnego punktu widzenia (ang. *Free viewpoint TV*) [Tani\_06][Tani\_09][Tani\_11][Tani\_12]. Daje ona widzowi możliwość swobodnego wyboru punktu, z którego obserwowana będzie scena.

Dużym wyzwaniem we wspomnianych zastosowaniach jest dostarczenie widoków w liczbie wystarczającej do poprawnego działania wyświetlaczy. Jakość wyświetlanego na nich obrazu zależy bowiem od liczby dostarczonych do wyświetlacza widoków (im więcej, tym lepsza jakość wyświetlanej sceny). Bardzo aktualnym zagadnieniem jest zatem przetwarzanie, przesyłanie i przechowywanie dużej ilości danych, pochodzących z wielu kamer. Liczba kamer w działających obecnie systemach wielowidokowych sięga 100 sztuk, jak na przykład w systemie opisanym w [Fuji\_06].

Do kompresji sekwencji wielowidokowych opracowywane są specjalizowane metody. Najbardziej rozpowszechnioną jest, będąca rozszerzeniem technologii zaawansowanego kodowania wizyjnego AVC (ang. *Advanced Video Coding*), technologia znana pod nazwą MVC (ang. *Multiview Video Coding*). Technologia AVC opisana jest w normie ISO/IEC 14496-10 MPEG-4 AVC oraz ITU-T H.264 [AVC]. Technologia MVC opisana jest w aneksie H wspomnianej normy.

Ze względu na niewielkie różnice między kamerami pracującymi w systemie, jak również ze

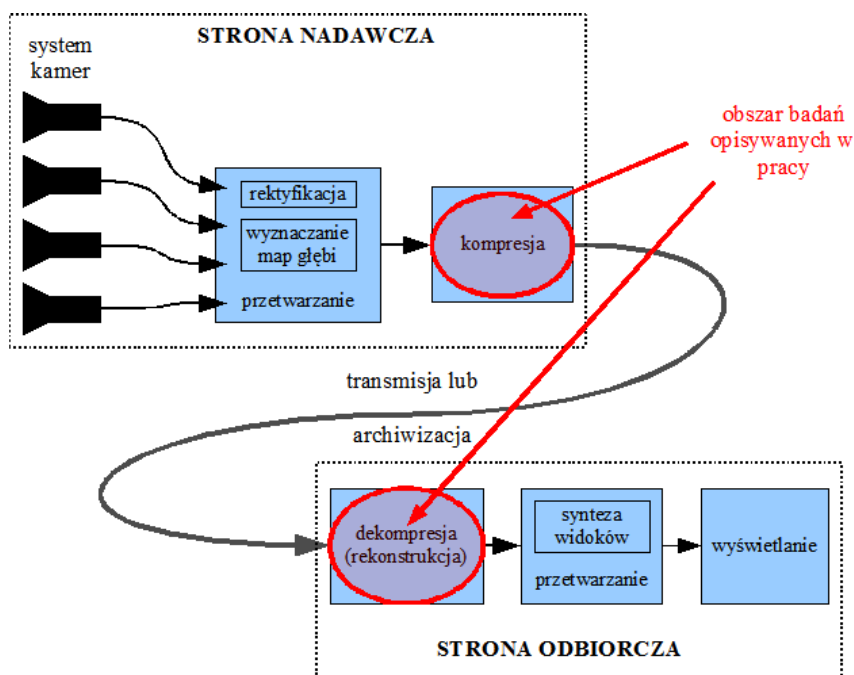
względu na różnice wynikające z niedokładnego ustawienia kamer, we współczesnych systemach przeprowadza się przetwarzanie wstępne sekwencji, polegające na rektyfikacji obrazów pochodzących z kamer [Stanko\_10].

Przesyłanie obrazów z kamer, w liczbie wystarczającej do poprawnego działania współcześnie wykorzystywanych wyświetlaczy autostereoskopowych, wymaga jednak zbyt dużych prędkości bitowych – nawet z wykorzystaniem dedykowanej technologii MVC [Mull\_11]. Zamiast przesyłania kilkunastu lub kilkudziesięciu widoków, możliwe jest jednak przesłanie jedynie kilku z nich wraz z dodatkowymi informacjami na temat przestrzennej struktury sceny. Badania nad takim sposobem transmisji prowadzone są obecnie przez grupę ekspertów MPEG (ang. *Moving Pictures Experts Group*) [MPEG\_CfP]. Przestrzenną strukturę sceny reprezentuje się za pomocą mapy głębi [Mull\_11]. Mapa głębi zawiera informacje o położeniu obiektów w scenie, a ściślej – o ich odległości od kamery. Mapa głębi przedstawiana jest z reguły jako monochromatyczny obraz, dla którego wartość próbki zależy od odległości obiektów sceny od kamery [Mull\_11]. Na podstawie obrazów z kamer wraz z przyporządkowanymi im mapami głębi możliwe jest wygenerowanie dodatkowych widoków syntetycznych [Mull\_08][Mori\_09], które mogą być wykorzystane do zapewnienia poprawnej pracy wyświetlacza.

W związku z przeprowadzaniem procesu syntezy w wyświetlaczu trójwymiarowym, konieczne jest przesłanie do odbiornika, poza obrazami z kamer, również map głębi. Istotne jest, aby w jak najmniejszym stopniu powiększać prędkość bitową ponad tę wymaganą do przesłania obrazu, przy wprowadzaniu jak najmniejszych zniekształceń do dekompresowanych map głębi.

Można przewidywać, że systemy wykorzystujące proces syntezy widoków przeprowadzany na podstawie ograniczonej liczby obrazów z kamer wraz z odpowiadającymi im mapami głębi będą w dalszym ciągu rozwijane i upowszechniane. Celowe jest zatem rozważanie problemów badawczych związanych z rozwojem tego typu systemów. Wśród najważniejszych zagadnień wymienić można: rozwój metod otrzymywania map głębi, rozwój metod syntezy widoków wirtualnych i kompresję obrazów wraz z mapami głębi.

Badania, których wyniki prezentowane są w pracy, dotyczą w szczególności ostatniego z wymienionych zagadnień – wydajnych sposobów kompresji sekwencji wielowidokowych wraz z mapami głębi (patrz rysunek 1.1).



Rysunek 1.1. Schemat przykładowego systemu wielowidokowego z zaznaczonym obszarem badań przedstawionych w pracy.

## 1.2. Założenia pracy

Przeprowadzone prace badawcze dotyczą kompresji sekwencji wielowidokowych z mapami głębi. Opisywane w pracy metody dotyczą wykorzystania kodera przeznaczonego do kompresji sekwencji wielowidokowych jako kodera sekwencji wielowidokowych z mapami głębi. Podstawą prowadzonych prac jest technologia zgodna z normą ISO/IEC 14496-10 MPEG-4 AVC oraz ITU-T H.264 [AVC][Rich\_10]. Technologia AVC jest powszechnie stosowana i zaadaptowana przez przemysł. Jest to wypróbowana technologia, bazująca na hybrydowej kompresji sekwencji wizyjnych. Technologia AVC podlega ciągłemu procesowi ulepszania i dostosowywania do nowych wymagań. W roku 2009 opublikowany został aneks H do wspomnianej normy, przystosowujący koder AVC do kodowania sekwencji wielowidokowych. W roku 2011 rozpoczęte zostały prace nad modyfikacją umożliwiającą kodowanie sekwencji wielowidokowych z mapami głębi (nazywanymi również sekwencjami różnowymiarowymi) [3DV\_WD][Rusa\_12]. W pracach tych, poza metodami opisywanymi w prezentowanej rozprawie, rozpatrywane są również nowe metody, wykorzystujące złożone algorytmy predykcji trójwymiarowej. Metody predykcji trójwymiarowej nie są przedmiotem prezentowanej rozprawy.

### 1.3. Cel i teza pracy

Celem pracy jest *analiza możliwości wykorzystania techniki wielowidokowego kodowania wizyjnego (MVC) do kodowania sekwencji wielowidokowych wraz z mapami głębi.*

Celami szczegółowymi są: *określenie sposobu doboru parametrów kompresji pod kątem uzyskiwania jak najlepszej efektywności kompresji w założonych warunkach oraz zaproponowanie i przebadanie metody kompresji wykorzystującej zależność pomiędzy sekwencjami wielowidokowymi i mapami głębi.* Opracowane sposoby mają zostać poddane ocenie efektywności kompresji z wykorzystaniem badań eksperymentalnych.

Teza pracy jest następująca:

*Istnieją proste reguły, umożliwiające ustalenie proporcji pomiędzy prędkościami bitowymi obrazów i głębi, pozwalające na maksymalizację efektywności kompresji sekwencji wielowidokowych z mapami głębi przeprowadzanej przy użyciu kodera MVC. Mapy głębi mogą być w procesie kompresji traktowane jako dodatkowa składowa barwna, co pozwala na proste wykorzystanie zależności pomiędzy obrazami i odpowiadającymi im mapami głębi.*

### 1.4. Przegląd pracy

Dalsza część pracy podzielona jest na rozdziały opisujące kolejne zagadnienia będące tematyką niniejszej rozprawy.

W **rozdziale 2** opisane są sekwencje wielowidokowe, sposoby ich otrzymywania i możliwe zastosowania. Omówione są również mapy głębi wraz ze sposobami ich otrzymywania.

W **rozdziale 3** omówione są standardowe, dotychczas stosowane metody kompresji sekwencji wielowidokowych. Wyjaśnione są zasady działania kodeka MVC, zgodnego z normą ISO/IEC 14496-10 MPEG-4 AVC oraz ITU-T H.264 wraz z aneksem H. Opisane są również pokrótce problemy, jakie można napotkać przy kompresji sekwencji wielowidokowych. W rozdziale tym przedstawione jest porównanie efektywności kompresji sekwencji wielowidokowych poddanych różnym procesom obróbki wstępnej.

W **rozdziale 4** opisane są metody kompresji sekwencji map głębi. Opisane są również używane w pracy metody pomiaru jakości dla sekwencji wielowidokowych z głębią.

W **rozdziale 5** przedstawione są oryginalne wyniki badań i metody pozwalające na uzyskanie

jak największej efektywności kodowania sekwencji wielowidokowych z wykorzystaniem istniejącego kodera sekwencji wielowidokowych, przy zachowaniu jak najlepszej jakości materiałów rekonstruowanych. Wyznaczone są wzory pozwalające na automatyczny dobór parametrów kodowania map głębi.

W **rozdziale 6** przedstawiona jest autorska idea kodowania czterokomponentowych sekwencji wielowidokowych. Przeanalizowane są w nim także dane z badań statystycznych nad kompresją obrazów i map głębi. Opisane są oryginalne metody wykorzystania zależności istniejących pomiędzy danymi w mapach głębi i w obrazach w celu poprawy efektywności kodowania w porównaniu z obecnie stosowanymi rozwiązaniami.

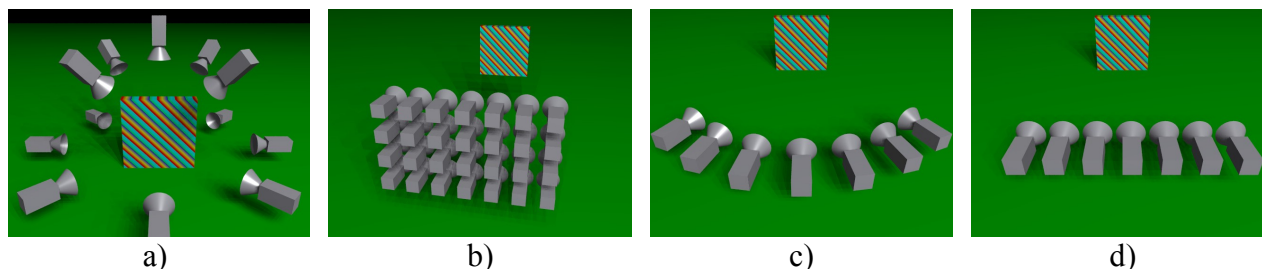
W **rozdziale 7**, podsumowującym pracę, przedstawione są wnioski z przeprowadzonych badań oraz wskazane są główne osiągnięcia badawcze pracy. Wytyczone są kierunki dalszych badań w dziedzinie kodowania sekwencji wielowidokowych.



## 2. Sekwencje wielowidokowe i mapy głębi

### 2.1. Otrzymywanie sekwencji wielowidokowych

Do otrzymywania sekwencji wielowidokowych wykorzystuje się systemy złożone z wielu zsynchronizowanych ze sobą kamer. Im więcej kamer użytych jest do rejestracji obrazu sceny, tym dokładniejsza jest jej trójwymiarowa reprezentacja. Duże znaczenie ma też sposób rozłożenia kamer w przestrzeni. W zależności od tego, w jakim celu ma być wykorzystywany system wielowidokowy (a zatem – jakim ograniczeniom podlega), rozpatruje się różne sposoby ustawienia kamer. Największą swobodę wyboru punktu obserwacji daje się uzyskać w przypadku rozłożenia kamer na powierzchni sfery (lub półsfery) otaczającej scenę, dzięki czemu uzyskuje się możliwość rejestrowania informacji o wyglądzie sceny z każdej strony. W większości przypadków takie ułożenie kamer jest jednak bądź fizycznie niemożliwe, bądź zupełnie niepraktyczne ze względu na rozmiary takiego systemu kamer. Innymi, łatwiejszymi w realizacji aranżacjami systemu kamer są ustawienia liniowe, łukowe i macierzowe. Przykładowe sposoby ustawienia kamer w systemie akwizycji sekwencji wielowidokowej przedstawione są na rysunku 2.1.



Rysunek 2.1. Różne sposoby ustawienia kamer w systemie akwizycji sekwencji wielowidokowych: a – na półsfery, b – w macierzy, c – na łuku, d – w linii.

W praktyce, ze względu na oczywiste ograniczenia wynikające głównie z kosztu urządzeń, rozważa się jak najprostsze systemy o jak najmniejszej liczbie kamer. W budowanych obecnie systemach najczęściej wykorzystuje się liniowe ustawienie kamer [Vetr\_05][Feld\_08][Doma\_09a], nieco rzadziej ustawienie na łuku [Zitn\_04]. Do rzadkości należą systemy z ustawieniem kamer w formie macierzy (z użyciem takiego systemu powstała na przykład sekwencja testowa „Akko&Kayo” opisana w [Tani\_05]) lub na planie krzyża [Kawa\_04].

W badaniach, których wyniki prezentowane są w rozprawie wykorzystane były ogólnie

dostępne sekwencje testowe lub też sekwencje udostępniane przez grupę ekspertów MPEG w ramach prac związanych z ustanowieniem standardu kompresji sekwencji wielowidokowych.

## 2.2. Mapy głębi

W literaturze opisującej mapy głębi spotyka się dwa sposoby reprezentacji odległości obiektów od kamery. W pierwszym wartość próbki jest proporcjonalna do odległości obiektu od kamery, w drugim – wartość próbki jest proporcjonalna do odwrotności tej odległości. Wybór sposobu reprezentacji ma znaczenie w przypadku, gdy wartości próbek w mapie głębi podlegają kwantowaniu. W przypadku zapisywania odległości od kamery uzyskujemy równomierne rozłożenie przedziałów o tej samej szerokości w całym zakresie pomiaru. Jeśli zapisywana jest odwrotność odległości od kamery, uzyskiwane są mniejsze przedziały odległości (większa rozdzielczość) w obszarze bliższym kamerze, a większe (mniejsza rozdzielczość) w zakresie największych odległości. Taki sposób zapisu odległości jest charakterystyczny dla tak zwanej mapy rozbieżności. Zgodnie z zależnościami geometrycznymi w modelu kamery otworkowej [Cyg\_02], odwrotność odległości obiektu od linii kamer jest proporcjonalna do różnicy jego położenia w obrazach z sąsiednich kamer (czyli rozbieżności). Odpowiednie zależności przedstawione są na rysunku 2.2 i we wzorach od (2.1) do (2.4).

Ze wzoru (2.4) wynika, że rozbieżność dla obiektu ( $D_{I-II}$ ) pomiędzy obrazami z kamery I i II jest odwrotnie proporcjonalna do jego odległości od linii kamer, a wprost proporcjonalna do odległości między kamerami  $D_k$ .

$$\frac{\Delta_1}{f} = \frac{A}{d} \quad (2.1)$$

$$\frac{\Delta_2}{f} = \frac{B}{d} \quad (2.2)$$

$$D_{I-II} = \Delta_1 + \frac{w}{2} - \left( \frac{w}{2} - \Delta_2 \right) = \Delta_1 - \Delta_2 \quad (2.3)$$

$$D_{I-II} = \frac{A+B}{d} \cdot f = \frac{D_k}{d} \cdot f \quad (2.4)$$

gdzie:

I, II, III – numery kolejnych kamer,

$f$  – ogniskowa kamer,

$d$  – odległość obiektu od linii kamer,

$D_k$  – odległość między kamerami,

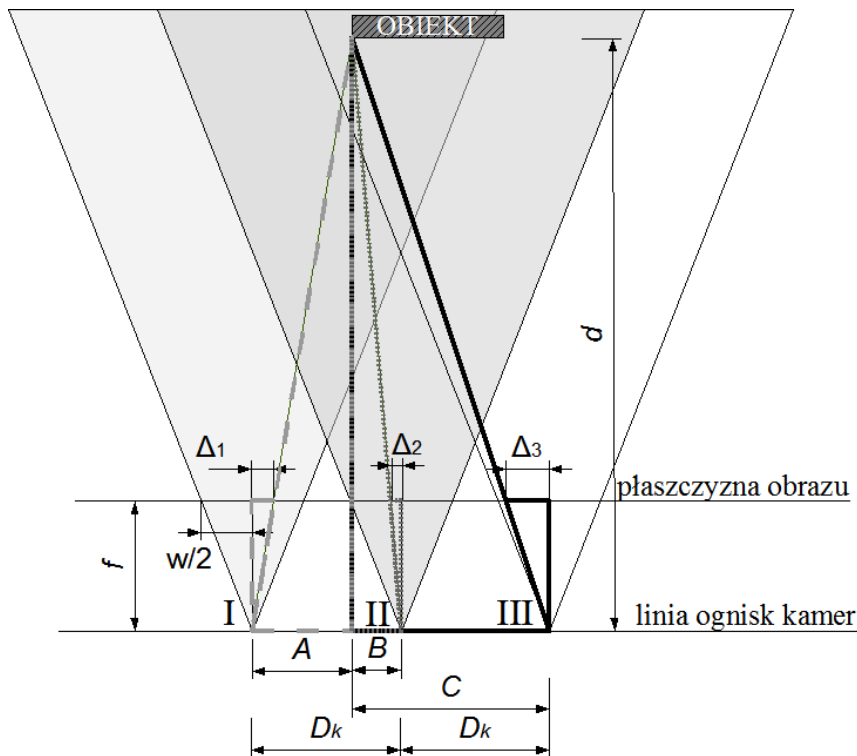
$A, B, C$  – odległość rzutu obiektu na linię ognisk kamer od ognisk kamer, odpowiednio I, II oraz III

$w$  – szerokość obrazów kamer,



$D_{i-j}$  – rozbieżność dla obiektu między obrazami z kamer i oraz j,

$\Delta_i$  – odległość obrazu obiektu w kamerze i od osi optycznej kamery i.



Rysunek 2.2. Zależność między rozbieżnością a odległością obiektu od linii kamer. Kolejne kamery oznaczone cyframi rzymskimi.

Reprezentacja głębi z wykorzystaniem rozbieżności jest metodą stosowaną w telewizji trójwymiarowej, ze względu na możliwość dokładniejszej reprezentacji głębi dla obiektów położonych bliżej kamery, które powinny być dokładniej odwzorowane ze względu na właściwości ludzkiego systemu widzenia [Li\_11]. W dostępnych materiałach testowych do reprezentacji tej odległości używa się zwykle tak zwanej znormalizowanej rozbieżności [Senoh\_07], dla której wartości rozbieżności ograniczone są do ustalonego przedziału wartości.

W dalszej części pracy pod pojęciem „mapa głębi” rozumiana będzie zawsze mapa rozbieżności.

### 2.3. Wyznaczanie map głębi

Rozwinięto wiele różnych technik wyznaczania map głębi, a opisy najczęściej wykorzystywanych obecnie metod odnaleźć można w literaturze [Stre\_02][Sun\_03][Schar\_03][Ince\_07][Lee\_08]. Wśród nich wymienić warto następujące.

- Zastosowanie skanera wykorzystującego technikę światła strukturalnego (*structured light*) [Skar\_09][Lavo\_04][Chen\_87][Free\_10]. Metoda wykorzystuje oświetlenie całej sceny za pomocą oświetlacza, rzucającego charakterystyczny wzorec świetlny na obiekty sceny. Obraz tak oświetlonej sceny rejestrowany jest przy pomocy kamery, której położenie i orientacja względem oświetlacza są znane. Zarejestrowany obraz podlega analizie, w wyniku której uzyskiwany jest trójwymiarowy model sceny. Niektóre systemy tego typu umożliwiają uzyskiwane dokładności pomiaru głębi rzędu ułamka milimetra [Forst\_06], jednak takie wyniki okupione są długim czasem akwizycji. Metoda ta zapewnia wierne odwzorowanie głębi, jednak nie nadaje się do wykorzystania przy akwizycji sekwencji wizyjnych z uwagi na powolne działanie oraz konieczność oświetlania sceny w nienaturalny sposób.
- Zastosowanie modyfikacji systemu wykorzystującego wspomnianą technikę światła strukturalnego. Charakterystyczne oświetlenie sceny uzyskiwane jest przy pomocy niewidzialnego dla ludzkiego oka światła podczerwonego. Technika ta nadaje się do akwizycji map głębi w czasie rzeczywistym i zyskuje ostatnio dużą popularność w przemyśle rozrywkowym, głównie za sprawą wykorzystania jej do budowy przystawki Kinect [Kinect][Smis\_11].
- Użycie specjalizowanych kamer głębi pracujących na zasadzie wykorzystania czasu powrotu impulsu świetlnego (*time of flight cameras*) [Lange\_01][Oggi\_03][Anders\_05]. Jest to metoda alternatywna wobec poprzedniej, o podobnych osiągnięciach i ograniczeniach, jednak o innej zasadzie działania. Również wykorzystuje do działania oświetlenie sceny przy użyciu światła podczerwonego. Nadaje się do akwizycji map głębi w czasie rzeczywistym.
- Przetwarzanie obrazów ze zwykłych kamer [Souri\_10]. Spośród wykorzystywanych metod wymienić można metody polegające na estymacji rozbieżności pomiędzy odpowiednimi punktami obrazu z sąsiednich kamer systemu wielokamerowego, opisywane na przykład w [Kolmo\_02][Kauff\_07][Stank\_08][Tani\_08b][LeeSB\_08][LiT\_09][Cigla\_09] oraz takie, w których estymacja map głębi dokonywana jest na podstawie ruchu obiektów lub kamery [Li\_06][Gord\_06][Zabu\_08] bądź też na podstawie pojedynczego obrazu sceny [Han\_03][Hoi\_05][Sax\_07]. Istnieją również rozwiązania, w których mapa głębi wyznaczana jest z wykorzystaniem kombinacji tych metod, na przykład na podstawie estymacji rozbieżności pomiędzy obrazami z kamer i podobieństwa między polami ruchu [Suarez\_12]. Co ważne, zastosowanie tych metod nie wiąże się z koniecznością zakupu dodatkowego sprzętu. Nie

zakłóca również w żaden sposób procesu akwizycji sekwencji, a sam proces estymacji map głębi może być odłożony w czasie i przeprowadzony już po akwizycji materiału filmowego. Dokładność metod można w znaczący sposób poprawić poprzez dostarczenie dodatkowych informacji o scenie, wprowadzonych przez operatora systemu. Tymi dodatkowymi informacjami mogą być na przykład mapy położenia krawędzi obiektów w obrazie [Li\_09] czy też wartości startowe do estymacji dla poszczególnych obiektów sceny. Główną wadą wymienionych metod jest duża złożoność obliczeniowa.

Obecnie prowadzone są badania mające na celu usprawnienie procesu wytwarzania map głębi do zastosowań w nowej dziedzinie, czyli w wizji wielowidokowej. Prace te zmierzają do pełnej automatyzacji procesu przy dużej szybkości działania i wystarczająco dobrej jakości. Na szczególną uwagę zasługują hybrydowe metody, korzystające z wielu źródeł informacji (na przykład kamery głębi oraz estymacja rozbieżności [Zhu\_08][Kang\_11]) w procesie estymacji mapy głębi sceny.

W pracy wykorzystywane są dwa rodzaje map głębi. Dla wcześniejszych badań wykorzystywane są mapy głębi uzyskane w pełni automatycznie, z użyciem oprogramowania [Tani\_08b] lub półautomatycznie, natomiast w badaniach rozpoczętych później wykorzystywane są wyłącznie mapy głębi stworzone w sposób półautomatyczny. Mapy głębi wyznaczone w sposób półautomatyczny (nazywane w pracy mapami głębi wysokiej jakości) dają lepsze wyniki syntezy niż mapy głębi wyznaczone w pełni automatycznie, jednak ich dostępność jest bardzo ograniczona. Dostępne są obecnie jedynie mapy głębi dla wybranych sekwencji testowych, a ich ilość jest ograniczona do trzech kamer w każdej sekwencji. Półautomatycznie wyznaczone mapy głębi korzystają z dodatkowych danych pomocniczych, jak na przykład mapa ważniejszych krawędzi czy startowe wartości głębi dla poszczególnych części obrazu, które przygotowane są ręcznie dla pierwszej klatki każdej sekwencji i każdej kamery. Dzięki dostarczeniu algorytmowi wyznaczania map głębi dodatkowych danych, możliwe staje się dokładniejsze wyznaczenie krawędzi obiektów i łatwiejsze jest poprawne umiejscowienie ich w scenie.

Mapy głębi wysokiej jakości wykorzystane w badaniach prezentowanych w rozprawie powstały na potrzeby grupy MPEG już po rozpoczęciu badań prezentowanych w rozprawie.



## 3. Kodowanie ruchomych obrazów wielowidokowych

### 3.1. Kompresja – wprowadzenie

W procesie kompresji sekwencji obrazów wykorzystuje się reprezentację próbek obrazu w tak zwanych współrzędnych transmisyjnych [Doma\_11]. Pojedynczy obraz z sekwencji reprezentowany jest wtedy za pomocą trzech składowych – luminancji  $Y$  (składowej odpowiadającej jasności punktów obrazu) i dwóch chrominancji  $C_B$  oraz  $C_R$  (składowych reprezentujących barwę punktu obrazu).

Z reguły w zastosowaniach telewizyjnych stosuje się decymację każdej chrominancji w obu osiach, według schematu 4:2:0 [Doma\_11]. Przykładowy wygląd składowych obrazu reprezentowanych w tym schemacie pokazany jest na rysunku 3.1. Można zauważyć na nim, że zdecydowana większość informacji o obrazie zachowywana jest za pomocą sygnału stanowiącego składową luminancji.

Proces kompresji sekwencji wizyjnej polega na zmianie sposobu jej reprezentacji na taki, do którego przesłania bądź przechowywania potrzebne są mniejsze zasoby (np. prędkość bitowa, przestrzeń dyskowa) niż w przypadku oryginalnej sekwencji. Proces kompresji jest więc istotny dla szybkiego przesyłania sekwencji, jak też dla archiwizacji sekwencji. Kodowanie sekwencji może być zrealizowane jako bezstratne, czyli takie, które umożliwia dokładne odtworzenie sekwencji wejściowej (sekwencja zrekonstruowana jest identyczna z sekwencją wejściową), lub jako stratne w wyniku którego część informacji zawartej w sekwencji wejściowej jest tracona (sekwencja zrekonstruowana jest podobna, ale nie identyczna z sekwencją wejściową) [Przel\_05][Doma\_11]. Wszystkie rozwiązania opisywane w prezentowanej pracy dotyczą kompresji stratnej, a zatem takiej, której zastosowanie wiąże się z wprowadzeniem zniekształceń do obrazu dekompresowanego.

W celu określenia stopnia zniekształcenia sekwencji stosuje się różne miary zniekształceń. Najbardziej wiarygodne są miary subiektywne – uwzględniające ocenę jakości zrekonstruowanych sekwencji przez grupę widzów, gdyż wprost odzwierciedlają one postrzeganą przez nich jakość sekwencji. Miary tego typu uwzględniają charakter zniekształceń i stopień ich subiektywnej dokuczliwości dla widza. Wadą tego typu rozwiązań jest kosztowność, długotrwałość i konieczność zgromadzenia odpowiedniej liczby widzów [BT500]. Alternatywnym sposobem pomiaru jakości sekwencji wizyjnej jest dokonanie oceny obiektywnej. Miary obiektywne bazują jedynie na

operacjach matematycznych wykonywanych na wartościach punktów obrazów oryginalnych i zrekonstruowanych. Co za tym idzie, mogą być zautomatyzowane i za ich pomocą można szybko i bez udziału ludzi otrzymać wynik pomiaru jakości. Wadą miar obiektywnych jest to, że nie zawsze dają one wyniki zbieżne z wynikami uzyskiwanymi z użyciem miar subiektywnych.



Rysunek 3.1. Reprezentacja obrazu za pomocą 3 składowych w schemacie 4:2:0. Oryginalny obraz kolorowy u góry, na dole – 3 składowe, od lewej: Y, C<sub>B</sub> oraz C<sub>R</sub>.

Mimo opisanych wad, metody pomiaru jakości poprzez miary obiektywne są niezastąpione w badaniach nad kompresją sekwencji wizyjnych, głównie dzięki możliwości przeprowadzania pomiarów w sposób automatyczny.

Bardzo często używaną w pracach badawczych obiektywną miarą jakości jest miara PSNR (z angielskiego *Peak Signal to Noise Ratio*), pozwalająca na szybkie i proste określenie stopnia zniekształcenia poprzez porównanie obrazu oryginalnego i zrekonstruowanego [Doma\_11].

Miara PSNR określona jest jako wyrażony w decybelach stosunek maksymalnej mocy sygnału do mocy zniekształceń, zgodnie ze wzorem (3.1). Przy jej obliczaniu bierze się zwykle pod uwagę jedynie składową luminancji porównywanych obrazów, ze względu na fakt, że jakość składowej luminancji jest dominującym czynnikiem wpływającym na postrzeganą przez widzów jakość sekwencji. W takim wypadku wartości  $f(i,j)$  i  $g(i,j)$  we wzorze (3.1) odpowiadają wartości próbki luminancji dla poszczególnych punktów obrazu.

W przypadku, gdy należy porównać różne dostępne metody kodowania sekwencji wizyjnych, konieczne jest opracowanie metody pozwalającej porównać ich efektywności. Warto pamiętać, że efektywność kompresji jest zależna od dwóch parametrów – od prędkości bitowej generowanego strumienia oraz od jakości rekonstruowanych (dekompresowanych) obrazów. W porównaniu metod

trzeba zatem wziąć pod uwagę oba te czynniki.

$$PSNR[ dB]=10 \log_{10} \left( \frac{(2^n - 1)^2}{\frac{1}{M \cdot N} \sum_{i=1}^N \sum_{j=1}^M (f(i, j) - g(i, j))^2} \right), \quad (3.1)$$

gdzie:

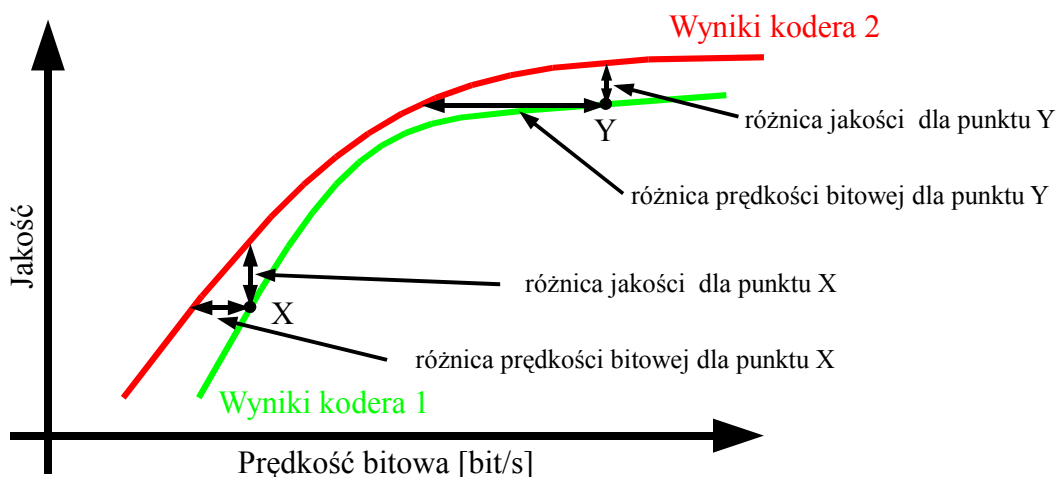
$n$  – liczba bitów na punkt obrazu – we wszystkich opisywanych w pracy przypadkach  $n$  będzie przyjmowało wartość 8,

$M, N$  – rozmiary obrazu (w punktach),

$f(i, j)$  – wartość próbki dla punktu obrazu zrekonstruowanego o współrzędnych  $i, j$ ,

$g(i, j)$  – wartość próbki dla punktu obrazu odniesienia o współrzędnych  $i, j$ .

Jednym z powszechnie stosowanych sposobów porównywania różnych metod kompresji sekwencji wizyjnych jest porównanie jakości sekwencji uzyskiwanej przy pomocy dwóch metod kompresji dla tej samej wartości prędkości bitowej strumienia (uzyskuje się wtedy porównanie w wymiarze jakości). Alternatywną metodą jest porównanie prędkości bitowych strumieni generowanych przez porównywane kodery w przypadku, gdy rekonstruowane sekwencje osiągają tę samą wartość miary jakości (uzyskuje się wtedy porównanie w wymiarze prędkości bitowej). Obydwie metody zaprezentowane zostały na rysunku 3.2. W przypadku, gdy uzyskane w czasie testów punkty pomiarowe nie mają takich samych współrzędnych w żadnej z osi, konieczna jest interpolacja mierzonych wartości.



Rysunek 3.2. Porównanie efektywności koderów – przykład. Koder z numerem 2 (linia czerwona) ma wyższą efektywność kodowania niż koder z numerem 1 (linia zielona).

Warto zauważyć, że w większości przypadków uzyskane dla koderów sekwencji wizyjnych krzywe zależności jakości od prędkości bitowej są wyraźnie nieliniowe, a wyniki porównania efektywności silnie zależą od wybranego do pomiarów punktu pomiarowego. I tak w przykładzie z rysunku 3.2 różnica w prędkości bitowej między koderami 1 i 2 mierzona w punkcie Y jest znacznie większa, niż różnica mierzona w punkcie X.

W literaturze można odnaleźć opisy metod pomiaru jakości uniezależniające wynikową wartość od wybranego punktu pomiarowego, dające w wyniku jedną, syntetyczną wartość. Przykładem takiej metody jest metoda opisana w [Bjont\_01], jednak doświadczenia z użyciem zaproponowanej tam metody wskazują, że dla niektórych przypadków, dla których porównywane krzywe prędkość bitowa/jakość leżą blisko siebie, wyniki uzyskane z użyciem tej miary mogą być mylące. Inną wadą takiej metody jest brak możliwości stwierdzenia w jakich zakresach prędkości bitowych przewagę ma jedna, a w jakich – druga spośród porównywanych metod kodowania. W związku z tym, wspomniana metoda nie jest używana do oceny wyników eksperymentów prezentowanych w pracy. W pracy wszelkie porównania dokonywane będą za pomocą porównania prędkości bitowych dla tej samej jakości sekwencji rekonstruowanej lub za pomocą porównania jakości uzyskiwanej przy tej samej prędkości bitowej. Miarą jakości stosowaną w ramach pracy będzie zawsze miara PSNR liczona zgodnie ze wzorem (3.1) dla składowej luminancji, chyba, że będzie zaznaczone inaczej.

### **3.2. Technologia zaawansowanego kodera wizyjnego AVC**

Kodowanie sekwencji wizyjnych w sposób zgodny z normą ISO/IEC 14496-10 MPEG-4 AVC oraz ITU-T H.264 (oznaczaną zwykle skrótowo jako AVC) jest jedną z wykorzystywanych obecnie metod kompresji sekwencji ruchomych. Poniżej podane zostaną podstawowe informacje na temat wspomnianego standardu kompresji, istotne w zrozumieniu dalszej części rozprawy. Szczegółowy opis kodera znaleźć można w książkach [Doma\_11][Rich\_10].

Koder według MPEG-4 AVC/H.264 jest koderem hybrydowym, to znaczy, że swe działanie opiera na dwóch podstawowych mechanizmach – predykcji z kompensacją ruchu i kodowaniu transformatowym.

W zależności od wybranego profilu pracy kodera, w procesie kodowania wykorzystuje się różne narzędzia i różne reprezentacje kodowanych obrazów. Więcej informacji na temat profili kodera AVC można znaleźć w pracach [Doma\_11][Rich\_10] jak też bezpośrednio w normie [AVC].

Sekwencja wizyjna kodowana za pomocą kodera AVC podzielona jest zawsze na ramki lub na pola. Podział na ramki oznacza, że sekwencję wizyjną rozpatruje się jako sekwencję następujących



po sobie obrazów o pełnej rozdzielczości, natomiast podział na pola oznacza, że sekwencja jest ciągiem następujących po sobie obrazów pól, czyli półobrazów, kolejno dla linii parzystych i nieparzystych wyświetlacza. Drugi przypadek wykorzystuje się do kodowania materiałów z wybieraniem międzyliniowym („z przeplotem”), wykorzystywanym powszechnie w materiałach dla telewizji analogowej. W dalszej części pracy rozpatrywane jest jedynie kodowanie ramkowe, czyli bez tak zwanego „przeplotu”.

Każda ramka przetwarzanej sekwencji kodowana jest na jeden z dwóch możliwych sposobów, bądź to całkowicie niezależnie od innych ramek – za pomocą kodowania wewnątrzobrazowego lub też z dodatkową możliwością wykorzystania informacji zawartych w innych ramkach sekwencji, czyli przy pomocy kodowania międzyobrazowego. Sposób kodowania zależy od wybranego schematu kodowania sekwencji. Poniżej przedstawione są podstawowe informacje dotyczące obu sposobów kodowania. Szczegółowe opisy z dokładnymi objaśnieniami można znaleźć w pracach [Doma\_11][Rich\_10].

Każdy obraz w trakcie przetwarzania dzielony jest na makrobloki – fragmenty o rozmiarach 16 na 16 punktów luminancji. Dla każdego makrobloku przesyła się również informacje o dwóch blokach chrominancji o wymiarach 8 na 8 punktów. Dla każdego makrobloku wybierany jest indywidualnie optymalny sposób kodowania spośród dostępnych dla danej ramki.

U podstaw efektywnego kodowania sekwencji wizyjnych w standardzie AVC leżą dobrze działające mechanizmy predykcji, czyli przewidywania zawartości kolejnych makrobloków kodowanego obrazu. Mechanizmy predykcji dla makrobloku mogą wykorzystywać informacje znajdujące się w sąsiednich obszarach obrazu, z którego pochodzi kodowany makroblok (mamy wtedy do czynienia z kodowaniem wewnątrzobrazowym), lub też informacje pochodzące z innych obrazów kodowanej sekwencji (mamy wtedy do czynienia z kodowaniem międzyobrazowym).

### **3.2.1. Kodowanie wewnątrzobrazowe**

Kodowanie wewnątrzobrazowe polega na odpowiednim zakodowaniu obrazu z wykorzystaniem jedynie informacji zawartych w tym samym obrazie. Dla obrazów kodowanych wewnątrzobrazowo, mechanizmy predykcji wykorzystują dane z wcześniej zakodowanych, sąsiednich makrobloków aktualnie kodowanego obrazu. Standard AVC przewiduje możliwość kodowania wewnątrzobrazowego makrobloków z wykorzystaniem kodowania całego bloku 16x16 punktów lub też z użyciem dodatkowego podziału na bloki o rozmiarach 4x4 punkty lub też (jeśli umożliwia to wybrana konfiguracja koodera) na bloki o rozmiarach 8x8 punktów. Dla każdego makrobloku

sposób podziału wybierany jest niezależnie, na podstawie algorytmu sterowania, uwzględniającego zniekształcenia wprowadzane przy kompresji i liczbę bitów niezbędnych do przesłania w danym przypadku.

W każdym przypadku dla wybranego podziału makrobloku przeprowadzana jest predykcja zawartości aktualnie kodowanego bloku z wykorzystaniem wartości próbek z kolumny punktów sąsiadujących z blokiem z lewej strony oraz z wiersza punktów sąsiadujących z blokiem od góry.

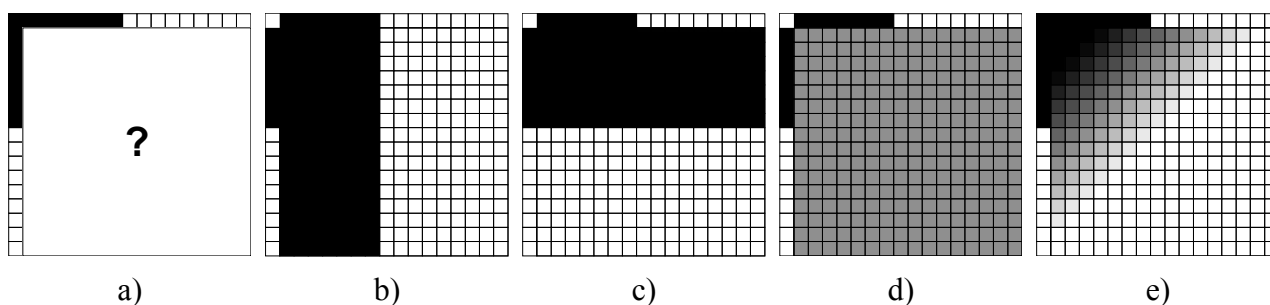
Dla bloków luminancji o rozmiarze 16x16 punktów możliwe są 4 różne sposoby predykcji, jak to pokazane zostało na rysunku 3.3:

- kopiowanie zawartości wiersza próbek powyżej makrobloku (*vertical prediction*),
- kopiowanie kolumny na lewo od makrobloku (*horizontal prediction*),
- wypnienie makrobloku średnią wartością sąsiednich próbek (*DC prediction*),
- wyznaczenie wartości próbek luminancji w bloku przy pomocy interpolacji dwuwymiarowej (*plane prediction*).

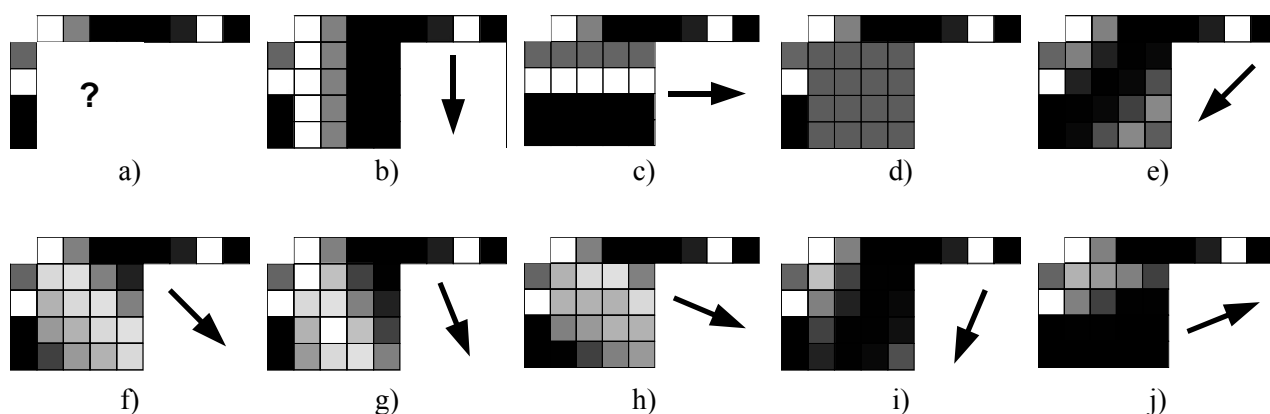
Dla bloków luminancji o rozmiarach 8x8 i 4x4 punkty kierunku predykcji wybierany jest spośród 9 możliwych, jak zostało to pokazane na rysunku 3.4:

- predykcja na podstawie średniej wartości próbek punktów sąsiadujących z kodowanym blokiem (*DC prediction*),
- kopiowanie zawartości wiersza próbek powyżej makrobloku (*vertical prediction*),
- kopiowanie kolumny na lewo od makrobloku (*horizontal prediction*),
- 6 kierunków ukośnych – predykcja odbywa się tu na drodze interpolacji odpowiednich wartości próbek sąsiadujących z kodowanym blokiem.

Dla bloków chrominancji o rozmiarze 8x8 punktów (jest to jedyny dostępny dla kodowania chrominancji rozmiar – bloki chrominancji nie podlegają podziałom), predykcja dokonywana jest na jeden z 4 sposobów, działających na identycznych zasadach jak predykcja dla próbek luminancji w bloku 16x16 punktów.



Rysunek 3.3. Predykcje w kodowaniu wewnątrzobrazowym dla bloku 16x16 punktów luminancji (na podstawie [AVC]). Blok do predykcji wraz z sąsiednimi punktami (a), predykcja pionowa (b), predykcja pozioma (c), predykcja wartością średnią (d), predykcja płaszczyzną (*plane*) (e).



Rysunek 3.4. Predykcje w kodowaniu wewnątrzobrazowym dla bloku 4x4 (na podstawie [AVC]). Blok do predykcji wraz z wykorzystywanymi sąsiednimi punktami (a), predykcja wartością średnią (d), predykcje kierunkowe z zaznaczonym kierunkiem – (b), (c), (e), (f), (g), (h), (i), (j).

Po wyznaczeniu przewidywanych wartości próbek, kolejnym krokiem jest określenie różnicy pomiędzy wartościami wyznaczonymi w drodze predykcji a rzeczywistymi wartościami próbek w aktualnie kodowanym bloku, czyli wyznaczenie tak zwanego sygnału błędu.

Próbki sygnału błędu predykcji poddawane są transformacji za pomocą przekształcenia całkowitoliczbowego, bazującego na przekształceniu kosinusowym.

Próbki wyznaczonej transformaty poddawane są kwantyzacji. Szerokość przedziałów kwantyzacji jest ustalana na etapie kodowania. Podawana jest ona w postaci indeksu (parametru) do ujętej w normie [AVC] tablicy zawierającej szerokości przedziałów kwantyzacji. Parametr kwantyzacji jest określany zwykle skrótem QP. Im szersze przedziały wykorzystane są na etapie kodowania, tym więcej skwantowanych próbek transformaty ma wartość 0 i nie jest w ogóle wysyłanych, ale również – tym niższa jest jakość zrekonstruowanego obrazu.

Kolejnym, po kwantyzacji etapem kodowania jest odpowiednia zamiana kolejności próbek transformaty a następnie kodowanie entropijne wartości próbek i umieszczenie danych w

strumieniu.

Dla każdego kodowanego wewnątrzobrazowo makrobloku przesłane są następujące informacje:

- sposób podziału makrobloku ( dla luminancji: 1 blok 16x16, 4 bloki 8x8, 16 bloków 4x4, chrominancja zawsze kodowana jest w bloku 8x8),
- kierunek predykcji, oddzielny dla każdego bloku luminancji i chrominancji,
- informacja o tym, czy dla danego bloku 8x8 punktów luminancji lub dla bloku chrominancji przesyłane są próbki transformaty sygnału błędu (tak zwane pole CBP - *coded block pattern*),
- przesłanie skwantowanych próbek transformaty błędu predykcji dla makrobloku, zarówno dla luminancji jak i dla chrominancji.

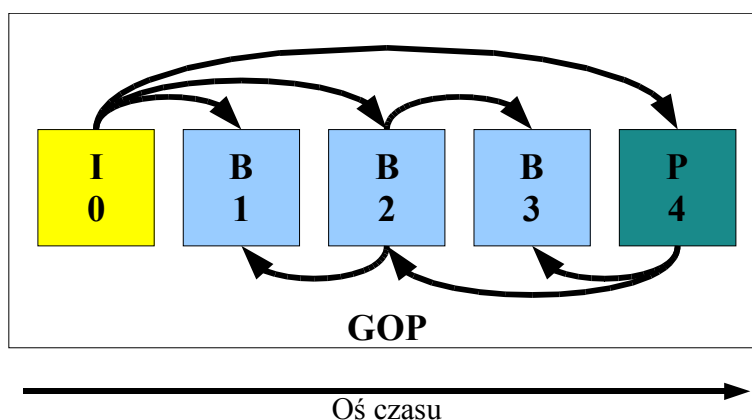
### 3.2.2. Kodowanie międzyobrazowe

Kodowanie międzyobrazowe pozwala na znaczne zwiększenie efektywności kompresji w porównaniu z kodowaniem wewnątrzobrazowym. W kodowaniu tym wykorzystuje się mechanizm predykcji z kompensacją ruchu. Jego działanie polega na wyszukaniu w poprzednio zakodowanych obrazach (tak zwanych obrazach odniesienia) obszaru najbardziej podobnego do aktualnie kodowanego obszaru. Zamiast przesyłać w strumieniu informacje o zawartości makrobloku, przesyła się jedynie współrzędne znalezionej podobnego obszaru. Współrzędne najbardziej podobnego bloku przesyła się w postaci kodowanego predykcyjnie tak zwanego wektora ruchu. W trakcie kodowania bloku wyliczany jest sygnał błędu predykcji, czyli różnica pomiędzy kodowanym blokiem a odnalezionym fragmentem obrazu odniesienia. Błąd predykcji poddawany jest takiemu samemu procesowi, jak w przypadku makrobloków kodowanych wewnątrzobrazowo. Przeprowadzone jest przekształcenie sygnału błędu z użyciem transformaty całkowitoliczbowej a wyniki poddawane są kwantyzacji. Tak obliczone próbki skwantowanej transformaty sygnału błędu predykcji poddawane są kodowaniu entropijnemu i umieszczane w strumieniu.

Ponieważ większość sekwencji wizyjnych charakteryzuje się dużą korelacją pomiędzy kolejnymi ramkami oraz, co za tym idzie – małymi wartościami sygnału błędu predykcji, kodowanie międzyobrazowe pozwala na bardzo reprezentowanie sekwencji wizyjnych. Dla bardzo wielu makrobloków sygnał błędu predykcji ma na tyle niewielką energię, że na etapie kwantyzacji wszystkie próbki jego transformaty są zerowane. W takim przypadku w strumieniu nie są przesyłane żadne dane o błędzie predykcji.

Ważnym zagadnieniem w procesie kodowania międzyobrazowego jest wybór struktury grupy obrazów (GOP – *group of pictures*). Grupa obrazów jest strukturą umożliwiającą poprawne

przeprowadzenie procesu dekodowania strumienia w wypadku, gdy dekodowanie nie rozpoczyna się od początku strumienia danych. Podział na grupy obrazów umożliwia w takim wypadku rozpoczęcie poprawnego dekodowania sekwencji. Dzięki zastosowaniu odpowiednich struktur i liczby ramek w pojedynczej grupie obrazów możliwe jest uzyskiwanie pożądanych cech strumieniowych, na przykład niskie opóźnienie dekodowania, wysoka efektywność, swobodny dostęp do dowolnego fragmentu sekwencji. Pojedyncza grupa obrazów zaczyna się zawsze od ramki kodowanej wewnątrzobrazowo, tak zwanej ramki I (kodowanie wewnątrzobrazowe opisane było w poprzednim punkcie). Dzięki temu dekodер ma możliwość rozpoczęcia dekodowania nawet, jeśli nie dysponuje żadnymi wcześniejszymi zdekodowanymi obrazami z danej sekwencji. Poza pierwszą ramką w grupie, pozostałe ramki mogą być kodowane z użyciem umożliwiającego uzyskanie znacznie mniejszych rozmiarów strumienia bitowego kodowania międzyobrazowego. Kolejność kodowania ramek wraz z odpowiednią konfiguracją kodera pozwala uzyskać wysoką efektywność kodowania, dzięki wykorzystaniu tak zwanych ramek P (ramki z jednokierunkową predykcją międzyobrazową) oraz B (ramki z dwukierunkową predykcją międzyobrazową). W czasie kodowania ramek P wykorzystywana jest predykcja ze zrekonstruowanych obrazów pochodzących z wcześniejszych chwil czasu. Dla ramek typu B możliwe jest wykorzystanie w procesie predykcji zarówno zrekonstruowanych obrazów pochodzących z wcześniejszych chwil czasu jak i obrazów pochodzących z późniejszych chwil czasu, jednak zakodowanych wcześniej. Przykładowy schemat wykorzystania predykcji międzyobrazowej przedstawiony jest na rysunku 3.5. Strzałkami oznaczono możliwe kierunki predykcji. Grot strzałki skierowany jest w kierunku obrazu kodowanego, przeciwny koniec strzałki skierowany jest w stronę ramki obrazu wykorzystywanego w procesie predykcji międzyobrazowej.



Rysunek 3.5. Przykładowa struktura grupy obrazów (GOP). Ramki ułożone są wzdłuż osi czasu w kolejności, w jakiej są one wyświetlane.

W przykładzie z rysunku 3.5 kolejność kodowania obrazów jest następująca. Jako pierwszy kodowany jest obraz z numerem 0. Jest on kodowany wewnątrzobrazowo, zatem nie korzysta z żadnych obrazów odniesienia.

Kolejnym kodowanym obrazem jest obraz z numerem 4. Jest on kodowany międzyobrazowo z wykorzystaniem obrazu z numerem 0 jako ramki odniesienia. Dla obrazu numer 4 jest to obraz sprzed 4 okresów akwizycji obrazu względem aktualnie kodowanego obrazu. Odniesienie do obrazu sprzed 4 okresów akwizycji obrazu oznaczane jest zwykle jako „ $T - 4$ ”.

Kolejnym kodowanym obrazem jest obraz z numerem 2. Jest on kodowany międzyobrazowo z wykorzystaniem obrazu z numerem 0 ( $T - 2$ ) oraz obrazu z numerem 4 (z chwili czasu późniejszej o 2 okresy, zatem oznaczane jako  $T + 2$ ) jako ramek odniesienia.

Kolejnym kodowanym obrazem jest obraz z numerem 1. Obraz 1 jest kodowany międzyobrazowo z wykorzystaniem obrazu z numerem 0 ( $T - 1$ ) oraz obrazu z numerem 2 (z późniejszej chwili czasu:  $T + 1$ ) jako ramek odniesienia.

Kolejnym kodowanym obrazem jest obraz z numerem 3. Obraz 3 jest kodowany międzyobrazowo z wykorzystaniem obrazu z numerem 2 ( $T - 1$ ) oraz obrazu z numerem 4 (z późniejszej chwili czasu:  $T + 1$ ) jako ramek odniesienia.

Warto odnotować, że standard AVC dopuszcza używanie w procesie predykcji międzyobrazowej więcej niż jednego (dla obrazu typu P) lub dwóch (dla obrazu typu B) obrazów odniesienia. W ten sposób zwiększa się prawdopodobieństwo odnalezienia w obrazach odniesienia obszaru cechującego się dużym podobieństwem do aktualnie kodowanego makrobloku.

Makroblok kodowany międzyobrazowo może być podzielony na mniejsze bloki, podobnie jak makroblok kodowany wewnątrzobrazowo. Dla kodowania międzyobrazowego wybór sposobów podziału jest szerszy i obejmuje bloki o wielkości 16x16, 16x8, 8x16, 8x8, 8x4, 4x8 i 4x4 punkty. Każdy z bloków o rozmiarze 8x8 lub większym może być kodowany z użyciem innego obrazu odniesienia. Każdy blok, nawet ten o rozmiarze 4x4 punkty może używać własnego wektora ruchu.

Sposób kodowania makrobloku charakteryzujący się najkrótszą reprezentacją jest kodowanie *skip* – w jego przypadku w strumieniu nie są przesyłane żadne poprawki do wartości otrzymywanych za pomocą mechanizmów predykcji – w procesie dekodowania tego makrobloku wykorzystuje się bezpośrednio wartości pochodzące z algorytmów predykcji. W tym przypadku również numer ramki odniesienia i wektor ruchu wyliczany jest na podstawie algorytmów predykcji. W tym trybie nie są też przesyłane żadne próbki transformaty błędu predykcji, zatem zdekodowany makroblok stanowi idealną kopię fragmentu jednej z ramek odniesienia. Koszt zakodowania jednego makrobloku typu *skip*, liczony jako średnia liczba bitów koniecznych do umieszczenia w strumieniu wynikowym, osiąga zwykle wartości zdecydowanie mniejsze niż 1 bit.

Zestaw danych kodujących informacje o makrobloku kodowanym międzyobrazowo jest większy niż dla kodowania wewnątrzobrazowego, jednak mimo tego może on wymagać mniejszej liczby bitów do zakodowania. Na zestaw ten składa się między innymi:

- znacznik trybu *skip*,
- typ makrobloku,
- sposób podziału makrobloku na bloki,
- oznaczenie wykorzystywanej ramki odniesienia,
- sygnał błędu predykcji wektora lub wektorów ruchu,
- informacja o tym, czy dla kolejnych bloków 8x8 punktów luminancji lub dla bloku chrominancji przesyłane są próbki transformaty sygnału błędu (pole CBP - *coded block pattern*),
- przesłanie skwantowanych próbek transformaty błędu predykcji dla makrobloku, zarówno dla luminancji jak i dla chrominancji.

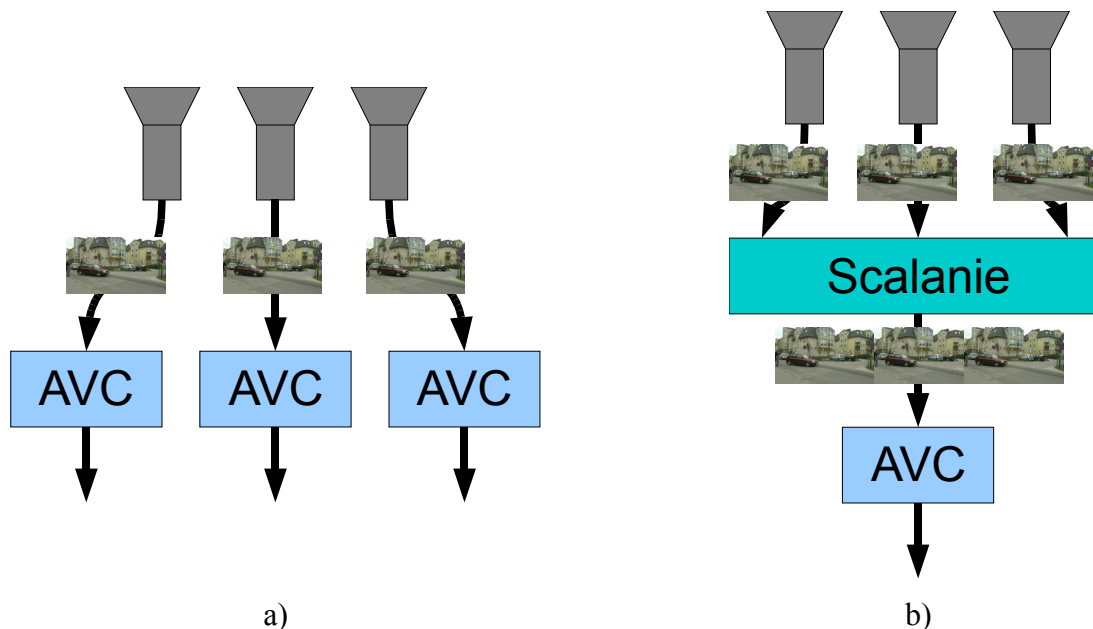
### 3.3. Metody kompresji sekwencji wielowidokowych

Najbardziej oczywistym i najprostszym sposobem kompresji sekwencji wielowidokowych jest użycie standardowego kodera sekwencji wizyjnych, przystosowanego do kompresji jednowidkowej sekwencji i zastosowanie go do kompresji każdego z widoków oddzielnie. Przykładowym rozwiązaniem tego typu jest użycie kodera zgodnego z normą AVC, w wersji przeznaczonej do kompresji sekwencji jednowidokowych.

Na rysunku 3.6 przedstawione są schematy możliwych sposobów kompresji sekwencji wielowidokowych z użyciem kodera sekwencji jednowidokowych AVC. Zastosowanie kodera jednowidkowego powoduje, że uzyskuje się, dla systemu z  $N$  kamerami,  $N$  niezależnych strumieni, zawierających informacje o pojedynczych widokach (rysunek 3.6a). Tego typu podejście jest dość powszechnie spotykane i zwykle nazywane jest kompresją równoczesną (ang. *simulcast*) [Tech\_09] [Merk\_09]. Ten termin będzie też używany w dalszej części pracy w odniesieniu do niezależnej kompresji każdego z widoków w systemie wielowidokowym.

Innym rozwiązaniem problemu kompresji jest scalenie, dla każdej chwili czasu, obrazów z  $N$  kamer do jednego obrazu o większej rozdzielczości i poddanie wynikowej sekwencji kompresji z użyciem kodera jednowidkowego (rysunek 3.6b). Warto zaznaczyć, że większość praktycznych implementacji koderów stawia ograniczenie na maksymalny rozmiar kompresowanego obrazu, zatem tą metodą nie da się skompresować sekwencji pochodzącej z bardzo wielu kamer przy

zachowaniu wysokiej rozdzielczości obrazów pojedynczego widoku. Metoda ta jest jednak niezwykle popularna w zastosowaniach, w których wymagane jest przesłanie jedynie dwu widoków – tego typu rozwiązanie funkcjonuje pod angielską nazwą kodowania „*side-by-side*”[Oster\_10].



Rysunek 3.6. Możliwe scenariusze kompresji sekwencji wielowidokowej z użyciem jednowidokowego kodera sekwencji wizyjnych AVC. Oddzielna kompresja sekwencji z kamer – kompresja równoczesna *simulcast* (a) oraz kompresja scalonego obrazu z wielu kamer (b).

Wymienione powyżej metody kompresji sekwencji wielowidokowych są proste w zastosowaniu i implementacji, gdyż wykorzystują istniejący, dobrze poznany koder AVC. Jest to oczywista zaleta, jednak metody te nie są w stanie wykorzystać charakterystycznych cech sekwencji wielowidokowych. Przez to ich zastosowanie nie pozwala na uzyskanie efektywnej kompresji sekwencji wielowidokowych.

W przypadku sekwencji wielowidokowych występuje bardzo duże podobieństwo pomiędzy obrazami z sąsiednich kamer, które można wykorzystać w procesie predykcji międzyobrazowej [Feck\_05][Kaup\_06]. Dzieje się tak, jeżeli kamery w systemach wielowidokowych są usytuowane dość blisko siebie. Ponieważ warunek ten jest spełniony dla wszystkich budowanych obecnie systemów, to pola widzenia kamer pokrywają się w bardzo dużym stopniu. Właśnie to zjawisko pokrywania się pól widzenia zostało wykorzystane podczas opracowywania specjalizowanej metody kompresji sekwencji wielowidokowych.

W opublikowanej w roku 2009 normie kompresji sekwencji wielowidokowych, stanowiącej aneks H do standardu AVC, wprowadzony został mechanizm predykcji międzywidokowej,



pozwalający wykorzystać duże podobieństwo widoków z sąsiednich kamer. W czasie prac nad standardem koder sekwencji wielowidokowych oznaczany był jako MVC i ten skrót będzie używany w dalszej części pracy jako oznaczenie kodera standardu AVC implementującego mechanizmy opisane w aneksie H.

### 3.4. Koder MVC

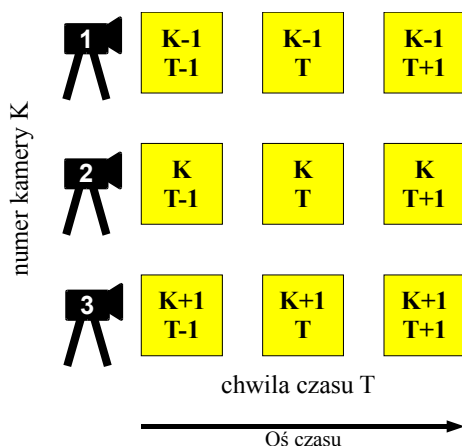
Opisany w poprzednim punkcie koder AVC projektowany był przede wszystkim do zastosowań, w których wykorzystywany jest pojedynczy widok z jednej kamery. Rozwój systemów wielowidokowych wymusił jednak opracowanie standardu kompresji dla sekwencji wielowidokowych. Jak zostało to wcześniej wspomniane, techniki łączenia obrazów z wielu kamer w jeden i poddawanie wynikowej mozaiki widoków kompresji z użyciem kodera jednowidokowego (jak na rysunku 3.6b), nie pozwalają wykorzystać wzmiankowanej już cechy sekwencji wielowidokowych, czyli dużego podobieństwa między widokami.

Od 2004 roku prace nad standaryzacją metod kompresji sekwencji wielowidokowych prowadzone są w ramach grupy MPEG (Moving Pictures Experts Group) [Pfis\_04]. W wyniku prac powstał dedykowany standard kompresji sekwencji wielowidokowych. Standard ten, nazywany MVC, jest rozszerzeniem normy AVC.

Podstawowym mechanizmem pozwalającym zwiększyć efektywność kodera MVC w porównaniu z AVC jest predykcja międzywidokowa, czyli wprowadzenie dodatkowego mechanizmu kompensacji rozbieżności (działającego na wzór kompensacji ruchu w AVC, a jedynie używającego innych ramek odniesienia). Jego działanie polega na umożliwieniu wykorzystania w predykcji międzyobrazowej dodatkowych obrazów odniesienia pochodzących z sąsiednich widoków – predykcja taka, wykorzystująca obrazy z sąsiednich kamer nazywana jest predykcją międzywidokową. Przykładowo dla obrazu z kamery o numerze  $K$  z chwili czasu  $T$  (patrz rysunek 3.7), poza zrekonstruowanymi obrazami z kamery  $K$ -tej z innych chwil czasu (na przykład  $T - 1$ ,  $T + 1$ ), w predykcji międzywidokowej można wykorzystać również zrekonstruowane obrazy z kamer sąsiednich, na przykład  $K + 1$  oraz  $K - 1$  z chwili czasu  $T$ .

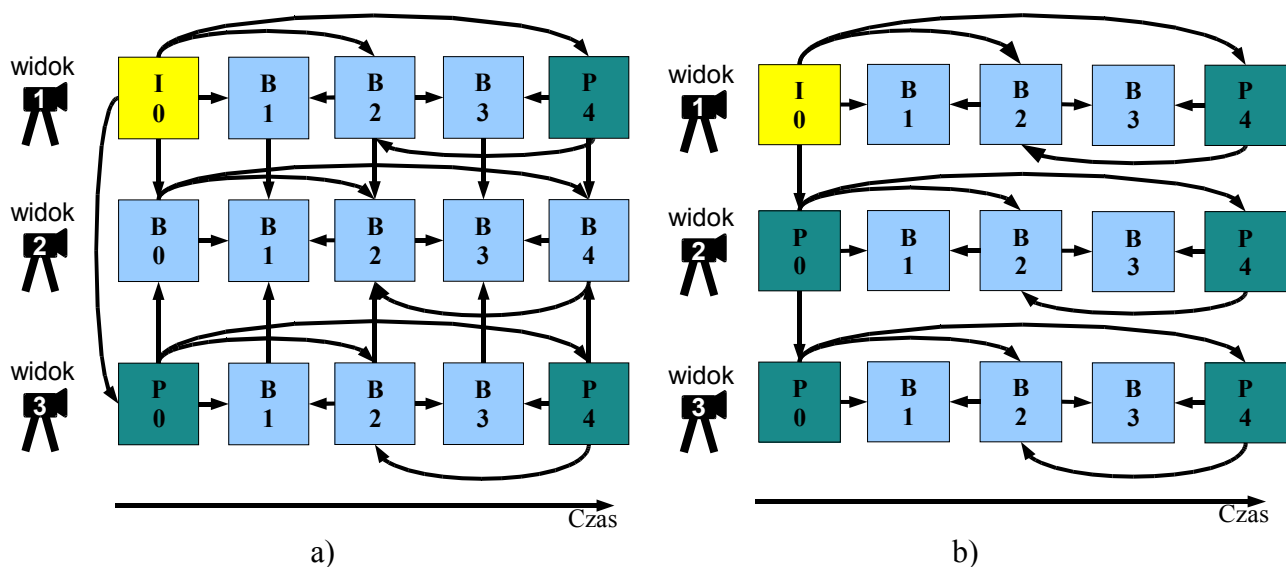
Taki mechanizm narzuca oczywiście konkretną kolejność kodowania widoków z kolejnych kamer, tak, aby dostępne były już wszystkie zrekonstruowane widoki potrzebne w procesie predykcji międzywidokowej. Przykładowe schematy wykorzystywania ramek odniesienia dla kodera MVC pokazane są na rysunku 3.8. Znaczenie strzałek jest takie samo jak na rysunku 3.5. Na rysunku przedstawione zostały dwa różne, przykładowe schematy. Pierwszy, oznaczony literą a, charakteryzuje się możliwością wykorzystania predykcji międzywidokowej dla każdego obrazu z

widoków z kamery 2 i 3. Schemat przedstawiony na rysunku 3.8b jest znacznie mniej złożony, odniesienia między widokami występują tylko dla pierwszej ramki w obrębie grupy obrazów (GOP). Dla obu schematów z rysunku 3.8, dzięki wykorzystaniu predykcji międzywidokowej, ze strumienia wyeliminowana jest większość ramek typu I, które zastąpione są dającymi znacznie mniejszą prędkość bitową ramkami typu P lub B. Dobór schematu wykorzystania ramek odniesienia jest kluczowym zagadnieniem dla kompresji z wykorzystaniem kodera MVC. Różne schematy wykorzystania ramek odniesienia rozważane są na przykład w pracach [Feck\_05] [Kaup\_06][Merk\_07][Flier\_07][Huo\_11].



Rysunek 3.7. Oznaczenie ramek w kodowaniu wielowidokowym.

Podobnie jak w przypadku kodowania jednowidokowego, również przy kodowaniu wielowidokowym dobór odpowiedniej struktury wpływa na właściwości strumienia wynikowego. Przykładowo - kodowanie według schematu z rysunku 3.8a daje nieznacznie większą efektywność kodowania niż kodowanie według schematu z rysunku 3.8b, jednak okupione jest to zwiększeniem złożoności obliczeniowej [Merk\_07]. Możliwe jest również wprowadzenie dodatkowych mechanizmów, pozwalających na zwiększenie efektywności kodowania, jak na przykład te przedstawione w [Klima\_07], [Kuo\_09], [Lee\_11].



Rysunek 3.8. Schematy wykorzystania ramek odniesienia dla kodera MVC w przypadku kodowania obrazów z 3 kamer. Schemat (a) przedstawia złożony schemat odniesień, natomiast schemat (b) przedstawia uproszczony schemat odniesień.

### 3.5. Efektywność mechanizmów wprowadzonych w MVC

Zastosowanie predykcji międzywidokowej pozwala na osiągnięcie większej efektywności kompresji w porównaniu z kodowaniem za pomocą kodera AVC bez mechanizmów kompresji sekwencji wielowidokowych. Kompresja sekwencji wielowidokowej złożonej z 8 lub więcej widoków przy użyciu kodera MVC daje prędkość bitową średnio o około 20% mniejszą przy zachowaniu podobnej jakości zrekonstruowanych obrazów w porównaniu z kodowaniem każdego widoku niezależnie (kodowanie równoczesne, ang. *simulcast*) [Chen\_09], przy czym dla niektórych sekwencji zysk wynosi mniej niż 10%. Maksymalne osiągane przez autorów prac [Merk\_07] [Mull\_11] zyski wynoszą około 40%. Warto podkreślić, że wyniki porównania w przypadku użycia do kodowania równoczesnego referencyjnego kodera AVC [Tour\_07] są lepsze – zysk z zastosowania kodera MVC wynosi powyżej 40% prędkości bitowej [Klim\_08]. Uzyskany wzrost efektywności silnie zależy od sekwencji. Na przykładzie sekwencji stereowizyjnych, badanych przez autorów [Merk\_09] widać, że zyski wynikające ze stosowania MVC mogą wahać się nawet od poniżej 10% do 50%.

### 3.6. Przetwarzanie wstępne sekwencji wielowidokowych

Jak zostało wspomniane w poprzednim punkcie, predykcja międzywidokowa wprowadzona do kodera MVC daje zyski sięgające od 10% do 50% wielkości strumienia bitowego w porównaniu z kodowaniem niezależnym. Redukcja strumienia jedynie o 10% może dziwić, gdyż widoki z

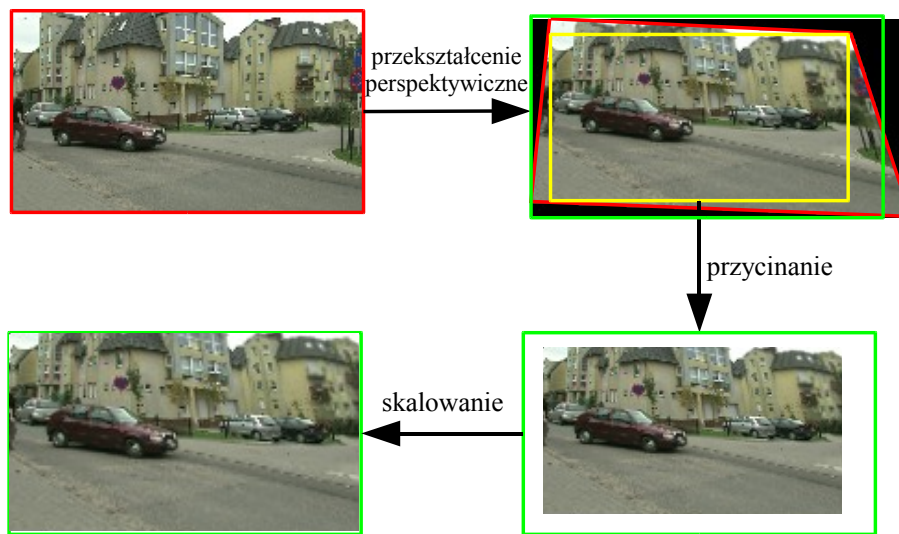
sąsiednich kamer są do siebie bardzo podobne – prezentują jedynie scenę z nieco innego punktu obserwacji. Mimo tego, obrazy z sąsiednich widoków rzadko wybierane są jako obrazy odniesienia dla kodowanych makrobloków (średnio w 17% przypadków), znacznie częściej wybierane są poprzednie obrazy z tej samej kamery (średnio 83% przypadków)[Feck\_05]. Obrazy z sąsiednich widoków i tej samej chwili czasu okazują się być zatem gorzej skorelowane z obrazem aktualnie kodowanym niż obraz z poprzedniej chwili z tej samej kamery. Istotnym problemem w przetwarzaniu sekwencji wielowidokowych jest bowiem istnienie słabo dostrzegalnych dla widza różnic między obrazami z sąsiednich kamer. Sprawiające najczęściej problemów różnice między widokami to różnice w histogramach kolorów dla obrazów z różnych kamer. Przyczynami tych różnic są: nierównomierne oświetlenie sceny oraz różnice w wewnętrznych cechach przetwornika każdej kamery. Inną ważną przyczyną powstawania różnic między widokami z sąsiednich kamer są niedokładności w geometrycznej orientacji kamer, mogące zmniejszać powierzchnię wspólnych obszarów widoków. Te dwie, na pierwszy rzut oka trudno dostrzegalne przez widza cechy sekwencji wielowidokowych sprawiają, że zyski wynikające z predykcji międzywidokowej są ograniczone. Poza tym, cechy te sprawiają problemy na późniejszych etapach przetwarzania sekwencji wielowidokowych, które będą opisane w dalszej części pracy.

Efekty związane z różnym odwzorowaniem kolorów w kamerach można zredukować poprzez zastosowanie procedury korekcji kolorów. Polega ona na przeprowadzeniu operacji wyrównywania histogramów dla wszystkich kamer w systemie wielowidokowym [Chen\_06][Yama\_07][Feck\_08][Hou\_09]. Korekcja kolorów nie tylko pozwala poprawić efektywność kompresji (patrz rysunek 3.10), ale również ułatwia dalsze przetwarzanie sekwencji.

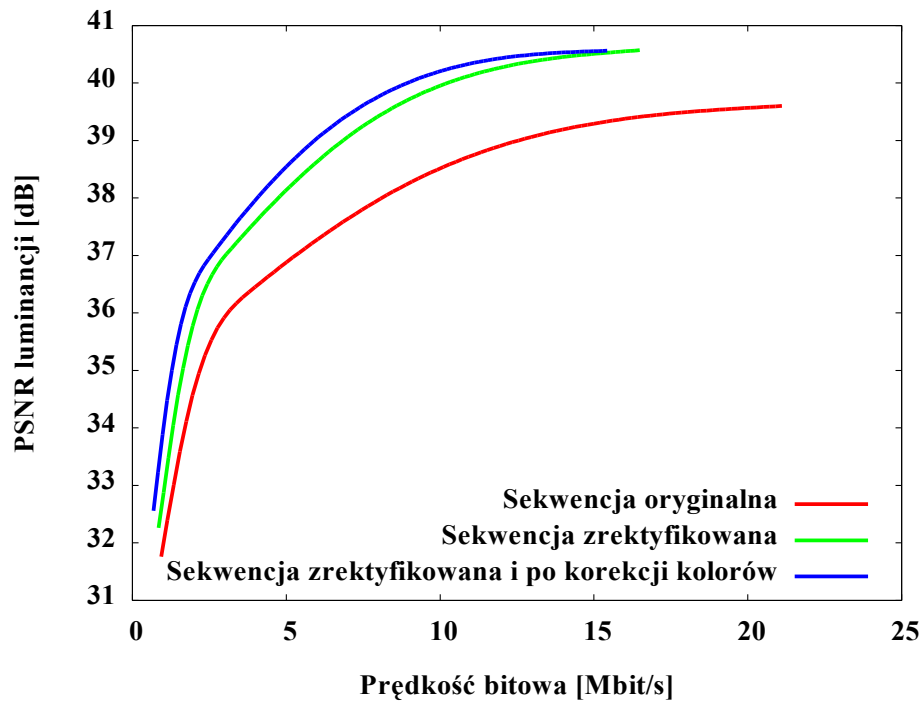
Problemy wynikające z nieidealnej geometrii systemu wielokamerowego rozwiązuje się zwykle przez rektyfikację epipolarną obrazów z wielu kamer [Stanko\_10]. Ten krok jest również pomocny przy dalszym przetwarzaniu sekwencji wielowidokowych jak również znakomicie ułatwia i przyspiesza proces syntezy widoków pośrednich. Proces rektyfikacji epipolarnej polega na poddaniu obrazu odpowiednio wyliczonemu przekształceniu perspektywicznemu [Fusie\_00] a następnie przycięciu obrazu tak, aby zawierał obszar bez powstających w wyniku procesu rektyfikacji marginesów. Przycięty obraz poddawany jest następnie skalowaniu do pierwotnych rozmiarów. Schemat procesu rektyfikacji pokazany jest na rysunku 3.9.

Wpływ rektyfikacji i korekcji kolorów na efektywność kompresji w sekwencjach wielowidokowych ilustrują przykładowe wykresy na rysunkach od 3.10 do 3.12, na których pokazana została zależność jakości zdekodowanych widoków (liczonej jako średnia jakość po wszystkich widokach) od całkowitej prędkości bitowej (dla wszystkich widoków) dla sekwencji

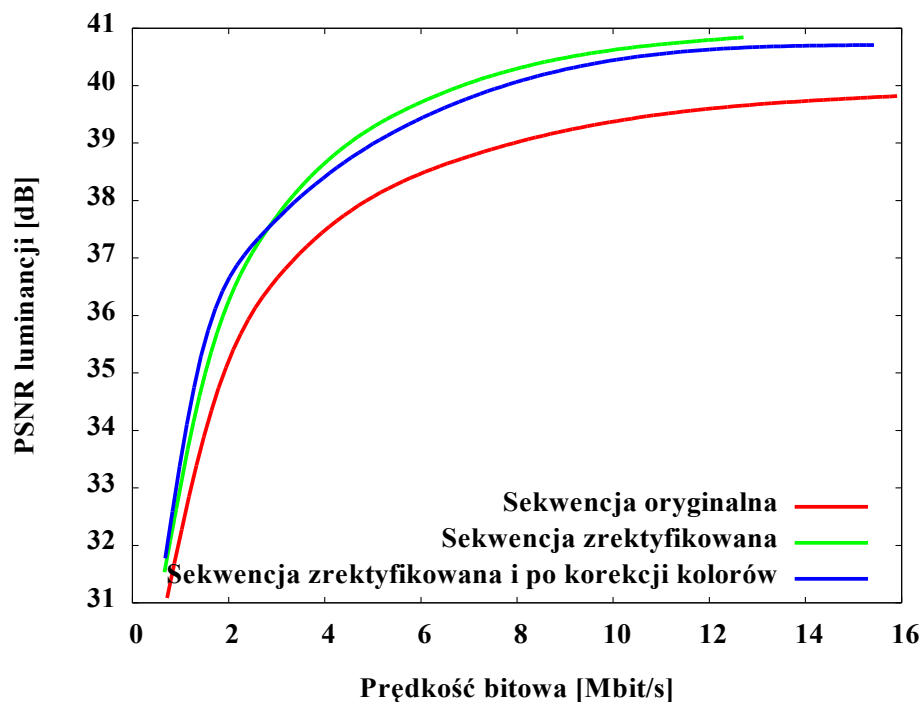
poddanych kolejnym etapom przetwarzania. Prezentowane wykresy utworzone zostały z wykorzystaniem wyników uzyskanych przez autora rozprawy na drodze badań eksperymentalnych. Przeprowadzenie eksperymentów było możliwe dzięki temu, że autor miał dostęp do sekwencji poddawanych kolejnym etapom przetwarzania. Tego typu częściowo przetworzone sekwencje nie są zwykle udostępniane przez autorów sekwencji. Szczegóły dotyczące eksperymentu podane są w Aneksie I. Podana na wykresach wartość PSNR liczona jest dla składowej luminancji względem luminancji nieskompresowanych obrazów sekwencji wejściowej kodera.



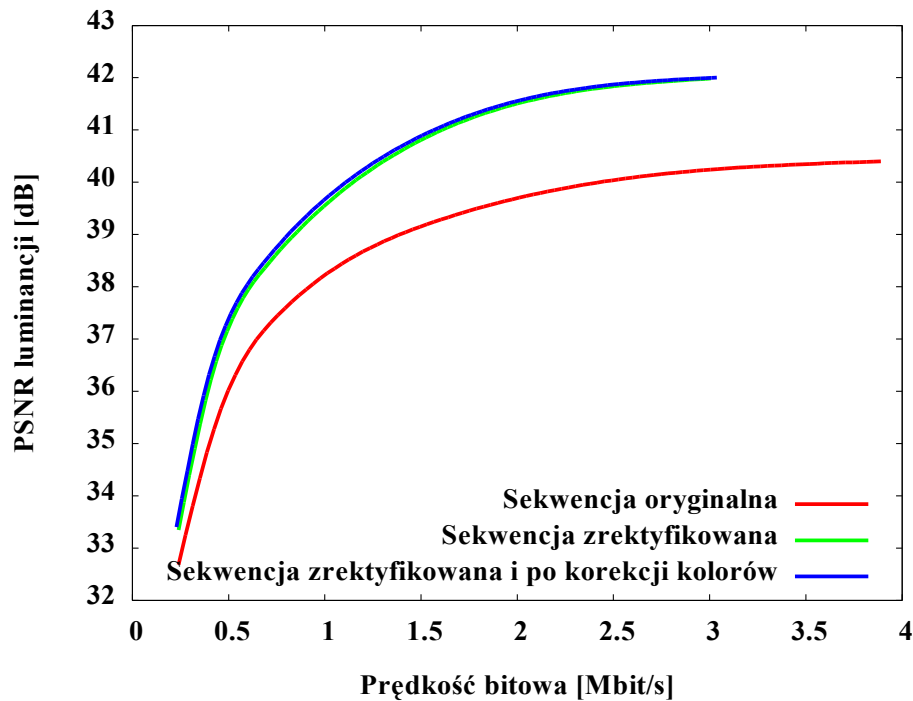
Rysunek 3.9. Schematyczne przedstawienie procesu rektyfikacji epipolarnej. Efekty przekształcenia perspektywicznego są znacznie przerysowane, aby uczynić je łatwo dostrzegalnymi.



Rysunek 3.10. Porównanie efektywności kompresji sekwencji Poznan Street na różnych etapach przetwarzania wstępnego. Kolor czerwony – sekwencja oryginalna, kolor zielony – sekwencja poddana procesowi rektyfikacji epipolarnej, kolor niebieski – sekwencja poddana procesowi rektyfikacji epipolarnej oraz procesowi korekcji kolorów. Prędkość bitowa podana jako sumaryczna wartość dla 3 widoków.



Rysunek 3.11. Porównanie efektywności kompresji sekwencji Poznan CarPark na różnych etapach przetwarzania wstępnego. Kolor czerwony – sekwencja oryginalna, kolor zielony – sekwencja poddana procesowi rektyfikacji epipolarnej, kolor niebieski – sekwencja poddana procesowi rektyfikacji epipolarnej oraz procesowi korekcji kolorów. Prędkość bitowa podana jako sumaryczna wartość dla 3 widoków.

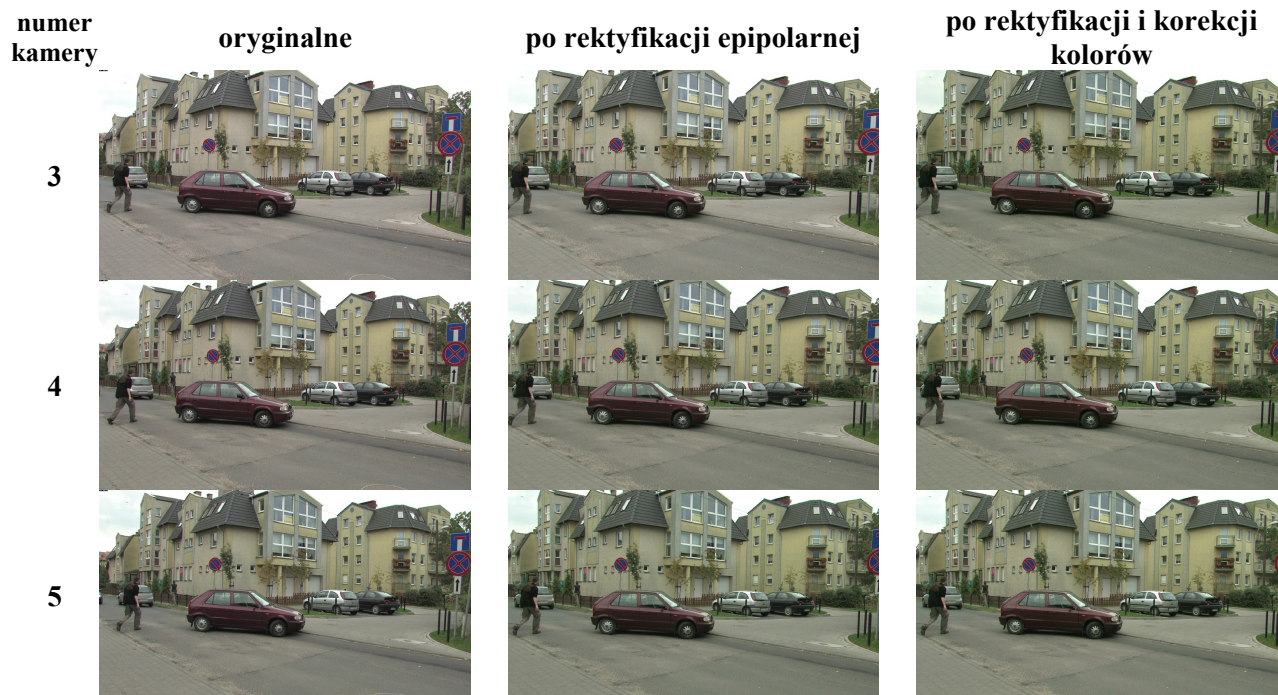


Rysunek 3.12. Porównanie efektywności kompresji sekwencji Book Arrival na różnych etapach przetwarzania wstępnego. Kolor czerwony – sekwencja oryginalna, kolor zielony – sekwencja poddana procesowi rektyfikacji epipolarnej, kolor niebieski – sekwencja poddana procesowi rektyfikacji epipolarnej oraz procesowi korekcji kolorów. Prędkość bitowa podana jako sumaryczna wartość dla 3 widoków.

Uwidocznione na rysunkach od 3.10 do 3.12 krzywe świadczą o wyraźnym wzroście efektywności kodowania sekwencji wielowidokowych w wyniku przeprowadzenia rektyfikacji sekwencji. Zysk wynosi około 2 dB dla większych prędkości bitowych i około 30 % prędkości bitowej dla mniejszych prędkości bitowych. Wpływ procesu korekcji kolorów na efektywność kodowania sekwencji nie jest jednoznaczny – w przypadku sekwencji CarPark dla prędkości bitowych powyżej 3 Mbit/s efektywność kompresji sekwencji po korekcji kolorów jest niższa o około 0,3 dB niż sekwencji bez korekcji kolorów. Poniżej tej prędkości daje się zaobserwować zysk rzędu 0,2 dB. W przypadku sekwencji Poznan Street dla całego przebadanego zakresu prędkości bitowych obserwowany jest zysk rzędu 0,3 dB wynikający z przeprowadzenia korekcji kolorów, jednak wyraźnie spada on wraz ze wzrostem prędkości bitowej. W przypadku sekwencji Book Arrival dla całego przebadanego zakresu prędkości bitowych obserwowany jest zysk rzędu 0,1 dB.

Porównanie efektywności kompresji dla sekwencji rektyfikowanej i oryginalnej może budzić pewne wątpliwości, ze względu na dużą ingerencję w materiał wejściowy koderów. W istocie jednak, dla przebadanych sekwencji proces rektyfikacji zawęży pole widzenia jedynie w niewielkim stopniu, trudnym do zauważenia dla widzów. Na rysunku 3.13 pokazane są całe obrazy dla 3

widoków sekwencji Poznan Street wybrane z sekwencji użytych w eksperymencie, poddane kolejnym etapom przetwarzania.



Rysunek 3.13. Porównanie obrazów z sekwencji oryginalnej (kolumna lewa), poddanej rektyfikacji epipolarnej (kolumna środkowa) i rektyfikacji epipolarnej wraz z korekcją kolorów (kolumna prawa).

Na rysunku 3.13 daje się zaobserwować jedynie niewielkie zawężenie pola widzenia w wyniku przeprowadzenia procesu rektyfikacji. Efekty działania procesu korekcji kolorów są bardzo trudne do zauważenia przez widzów, jednak powodują zauważalne zwiększenie efektywności kodowania. Podobnie trudne do zauważenia są transformacje w przypadku pozostałych przetestowanych sekwencji.

### 3.7. Podsumowanie

Kompresja sekwencji wielowidokowych może być realizowana na wiele sposobów, lecz spośród omówionych powyżej najnowocześniejszy jest sposób wykorzystujący koder MVC, czyli zmodyfikowany koder AVC, przystosowany do kompresji sekwencji wielowidokowych. Nowe narzędzia wprowadzone do MVC pozwalają na uzyskanie strumieni o mniejszej prędkości bitowej przy zachowanej tej samej jakości jaka jest uzyskiwana z wykorzystaniem niezależnego kodowania każdego widoku.

Wnioski z przedstawionych w rozdziale badań oraz podsumowanie dokonane w nim przeglądu stanu wiedzy można ująć poniższymi stwierdzeniami.



1. Kodowanie wielowidokowe z użyciem MVC pozwala na redukcję prędkości bitowej o około 20% do 40% względem kodowania równoczesnego (ang. *simulcast*) [Merk\_07] [Chen\_09][Mull\_11], przy czym redukcja liczona względem wyników uzyskanych względem kodowania równoczesnego z użyciem referencyjnego kodera AVC jest nieco większa [Klim\_08].
2. Zyski kodowania z użyciem MVC uzyskuje się dzięki wykorzystaniu predykcji międzywidokowej, pozwalającej na wyeliminowanie ze strumienia części powielanych między widokami informacji. Duże zyski uzyskuje się również przez eliminację ze strumienia większości obrazów typu I i zastąpienie ich obrazami kodowanymi międzyobrazowo z wykorzystaniem predykcji międzywidokowej.
3. Przetwarzanie wstępne sekwencji jest konieczne w celu przyspieszenia dalszych etapów przetwarzania oraz poprawy jakości obrazów wynikowych z dalszych etapów przetwarzania. Dodatkowo, rektyfikacja epipolarna sekwencji wielowidokowych sprawia, że kodowanie tak przetworzonej sekwencji jest bardziej efektywne. Poprawę efektywności kodowania może dać również przeprowadzenie korekcji kolorów w sekwencji wielowidokowej.
4. W przypadku przesyłania sekwencji złożonej z wielu widoków, wymaganych na przykład do poprawnej pracy wyświetlacza autostereoskopowego, redukcja strumienia oferowana przez MVC okazuje się niewystarczająca. Dla tego typu wyświetlaczy konieczne jest opracowanie innej metody przesyłania danych.



## 4. Kodowanie map głębi

### 4.1. Przegląd literatury

W dostępnych pozycjach literaturowych problem kompresji map głębi rozważany jest jako problem kompresji obrazów statycznych i jako problem kompresji sekwencji obrazów.

W większości przypadków autorzy zwracają szczególnie dużą uwagę na konieczność wiernego odwzorowania krawędzi w rekonstruowanej mapie głębi.

Wśród metod kompresji map głębi traktowanych jako obrazy statyczne, w literaturze opisywane są poniższe.

W pracach [Mait\_08][Mait\_10] autorzy proponują zastosowanie do kompresji map głębi dyskretnego przekształcenia falkowego dopasowanego do krawędzi występujących w mapach głębi. W metodzie tej przebieg krawędzi w mapie głębi przesyłany jest odrębnie. Liczba przesyłanych krawędzi (a przez to granic obszarów w mapie głębi) dopasowywana jest do dostępnej prędkości bitowej. Do kodowania przebiegu krawędzi autorzy proponują wykorzystanie korelacji pomiędzy położeniem krawędzi w mapie głębi i w odpowiadającym mu obrazie (w badaniach wykorzystane zostały materiały [Middlebury]).

Propozycja dopasowania wykorzystanej w kodowaniu mapy głębi transformaty do położenia krawędzi obiektów pojawia się również w [Shen\_10].

W pracy [Krish\_01] autorzy wykorzystują do kodowania map głębi koder JPEG2000 [JPEG2000]. Silnym zniekształceniom krawędzi w mapach głębi zapobiega w tym przypadku odpowiednia alokacja bitów w rejonach występowania krawędzi.

W pracy [Chai\_02] autorzy proponują kodowanie mapy głębi poprzez jej przybliżenie za pomocą siatki trójkątów. Rozłożenie węzłów siatki dopasowywane jest do przebiegu konturów w mapie głębi, przez co umożliwia dokładne odwzorowanie granic między obiektami bez wprowadzania zniekształceń.

Zbliżone metody zaprezentowane są w pracach [Morv\_05][Morv\_05a][Merk\_08], wykorzystujących podobne podejście co we wcześniejszych pracach [Tzov\_96][Willet\_02]. Proponowana metoda polega na przybliżeniu mapy głębi za pomocą bloków o jednolitej wartości głębi oraz za pomocą bloków zawierających granicę pomiędzy dwoma obszarami o różnej wartości głębi. Autorzy proponują również wprowadzenie bloków o gradientowo zmieniającej się wartości głębi.

W opisanych powyżej pracach mapy głębi traktowane są jako obrazy statyczne. Jednak ponieważ w zastosowaniach rozważanych w pracy mapy głębi mogą być rozumiane przede wszystkim jako sekwencje obrazów monochromatycznych, naturalnym wydaje się być zaadaptowanie do ich kompresji algorytmów kompresji sekwencji wizyjnych. Ten sposób daje nadzieję na osiągnięcie wyższej efektywności kodowania niż indywidualne kodowanie kolejnych obrazów sekwencji.

Wśród metod kodowania map głębi jako sekwencji obrazów w literaturze opisywane są poniższe metody.

Propozycja zastosowania mechanizmów kodera hybrydowego do kompresji sekwencji map głębi pojawia się w [Tzov\_96]. Autorzy postulują wykorzystanie kompensacji ruchu w kompresji sekwencji map głębi.

W dostępnych nowszych publikacjach dotyczących kodowania sekwencji map głębi opisywane są głównie metody wykorzystujące koder AVC. Autorzy skupiają się w nich na zagadnieniach dotyczących sposobów zachowania wyraźnych krawędzi w rekonstruowanych mapach głębi. Pojawiają się również prace, w których opisywane są sposoby optymalizacji procesu kompresji sekwencji map głębi pod kątem ich przeznaczenia w procesach realizowanych w wyświetlaczu.

W pracach [Kamo\_08][Kamo\_09][DeSil\_09][Fan\_11] opisane są metody wykorzystujące trójwymiarowy wektor ruchu do kompresji sekwencji map głębi. Poza współrzędnymi pionową i poziomą wektor ruchu dla mapy głębi ma w tych propozycjach dodatkową składową opisującą ruch obiektów w kierunku prostopadłym do płaszczyzny mapy głębi.

W pracy [Shen\_10a] autorzy proponują modyfikację kodera AVC wykorzystującą informacje o krawędziach do efektywnej predykcji głębi w trybie wewnątrzobrazowym.

W pracy [Zhang\_11] przedstawiona jest modyfikacja kodera AVC, polegająca na wprowadzeniu zmian w algorytmie sterowania. Dzięki wprowadzonym zmianom optymalizowana jest jakość widoku syntezy z użyciem kompresowanej mapy głębi, a nie samej mapy głębi. Podejście takie odzwierciedla rzeczywisty sens przesyłania map głębi w systemach wielowidokowych jako informacji pobocznej, nie prezentowanej widzowi, a służącej jedynie do przeprowadzenia procesu syntezy widoków wirtualnych. Podobne prace, których autorzy zwracają uwagę na konieczność innego od dotychczasowego podejścia do zagadnienia optymalizacji procesu kompresji mapy głębi, udokumentowane są w [Kim\_09][Kim\_10][Cheu\_10][Vale\_12].

W pracach [Oh\_09][Oh\_11] opisane są prace nad filtrami rekonstruującymi krawędzie w mapie głębi poddanej kompresji za pomocą kodera transformatowego. W pracy [Liu\_10] opisana jest metoda wykorzystania w procesie filtrowania mapy głębi również informacji o luminancji odpowiadającego obszaru.

Podobne podejście prezentowane jest w pracy [Jung\_12], przy czym w tym wypadku stosowane są binarne wzorce dla odpowiadających sobie rejonów w mapie głębi i luminancji.

W pracy [Oh\_09a], bazującej na metodzie filtrowania mapy głębi, poruszone jest również zagadnienie interpolacji mapy głębi. Podobne prace przedstawione są na przykład w [Zanutt\_09], jednak w tym przypadku dotyczą one kompresji statycznych map głębi. Wśród innych badań nad interpolacją map głębi wymienić można te opisane w [Schw\_12].

Kompresja sekwencji map głębi działająca w oparciu na dekompozycję na obszary zawierające krawędzie i pozostałe obszary opisana jest w [Kim\_07]. Autorzy używają przedstawionej przez siebie reprezentacji map głębi i koderu AVC do ostatecznego zakodowania sekwencji.

Opisane powyżej prace bazują w większości na kompresji sekwencji z użyciem koderu AVC. W ostatnim czasie pojawiać zaczynają się prace nad opracowaniem podobnie działających algorytmów dla nowo rozwijanych algorytmów kompresji sekwencji wizyjnych [Jager\_12].

Skompresowane mapy głębi mogą być przesyłane jako dane poboczne razem z obrazami z kamer, na przykład zgodnie z normą MPEG [MPEG-C\_3][Bour\_06]. Mapy głębi mogą być wtedy kompresowane z użyciem standardowych koderów sekwencji wizyjnych. Kodowanie obrazów i map głębi odbywa się w takim przypadku niezależnie, z użyciem tego samego koderu, a strumienie wyjściowe z koderów są wysyłane do odbiorców.

## 4.2. Szczególne właściwości mapy głębi

Z dokonanego przeglądu literatury jak też i z własnych obserwacji wysnuć można wnioski dotyczące szczególnych właściwości map głębi. Właściwości te odróżniają w sposób wyraźny mapy głębi od obrazów pochodzących z kamer.

Mapy głębi, rozważane jako obrazy monochromatyczne, mogą być kodowane za pomocą tych samych narzędzi, które używane są do kodowania obrazów z kamer. Należy jednak zwrócić uwagę na szczególne cechy map głębi, odróżniające je od obrazów. Różnice te z pewnością wpływają na efektywność kompresji.

Na rysunku 4.1 przedstawiony jest przykładowy obraz i odpowiadająca mu mapa głębi. Widoczne są charakterystyczne cechy mapy głębi:

- tylko jedna składowa (obraz monochromatyczny),
- liczne obszary o stałej wartości bądź z jednostajnie zmieniającą się wartością,
- ostre, wyraźne krawędzie.



Rysunek 4.1. Przykładowy obraz (a) i mapa głębi (b) dla sekwencji wielowidokowej.

W badaniach prezentowanych w pracy do kodowania map głębi wykorzystywany jest koder MVC, będący koderem hybrydowym. Fakt dominacji w mapie głębi gładkich powierzchni powoduje, że dla większości makrobloków w wyniku transformacji dokonywanej dla bloku mapy głębi nie będą pojawiać się duże wartości składowych wysokoczęstotliwościowych. Z drugiej strony, występowanie kontrastowych krawędzi będzie powodowało znaczne zwiększenie zawartości składowych wysokoczęstotliwościowych dla tych bloków, przez które przechodzi krawędź obiektu. Wobec tego można oczekiwać, że kodowanie optymalne z punktu widzenia kodowania obrazów z kamer nie będzie równie dobrze dostosowane do kompresji map głębi. Jednak mimo wykorzystania nieoptymalnej metody kompresji, występowanie w mapie głębi dużych obszarów o jednolitej wartości (a zatem możliwych do bardzo efektywnego zakodowania) może sprawiać, że strumienie zawierające informacje o obrazach map głębi będą mniejsze niż strumienie dla odpowiadających im obrazów z kamer zakodowanych z takimi samymi parametrami kodowania.

Dodatkową cechą odróżniającą mapę głębi od obrazu jest fakt, że ewentualna niewielka zmiana wartości jasności punktu obrazu nie wpływa w sposób znaczący na subiektywną jakość obrazu zrekonstruowanego [Yang\_05]. Jednak taka sama niewielka zmiana wartości próbek mapy głębi może (ale nie musi) doprowadzić do bardzo dokuczliwych zmian jakości obrazu syntezy z użyciem tak zniekształconej mapy głębi. Związane jest to ze sposobem działania procesu syntezy, w którym do tworzenia obrazu wirtualnego wykorzystywane są wprawdzie konkretne wartości próbek mapy głębi, lecz kolejne próbki mapy głębi nie są tak samo istotne [Vale\_12].

Opisane cechy map głębi świadczą o tym, że kompresja map głębi z wykorzystaniem kodera MVC będzie prowadziła do powstawania strumieni o mniejszej prędkości bitowej niż kompresja odpowiadających tym mapom obrazów.

### 4.3. Ocena jakości map głębi

W procesie kodowania sekwencji wielowidokowych z mapami głębi ważnym aspektem jest dobór odpowiedniej proporcji pomiędzy liczbą bitów służących reprezentacji obrazu i służących reprezentacji głębi. Celowe jest osiągnięcie możliwie najmniejszej sumarycznej liczby bitów koniecznych do reprezentacji sekwencji, a przez to zmniejszenie prędkości bitowej wymaganej do przesłania zakodowanej treści. W przypadku sekwencji jednowidokowych problem ten sprowadza się do uzyskania jak najlepszej jakości rekonstruowanego obrazu (jakości rozumianej najczęściej jako wartość stosunku sygnału użytecznego do szumu – PSNR) przy ograniczonej od góry możliwej do zaakceptowania liczby bitów (prędkości bitowej transmisji). Sterowanie koderem w celu osiągnięcia optymalnej jakości przy zachowaniu ograniczonej liczby bitów reprezentacji daje się zrealizować na drodze odpowiedniego doboru wartości indeksu parametru kwantyzacji, określanego jako QP.

W przypadku sekwencji wielowidokowych, reprezentowanych zarówno przez obrazy jak i odpowiadające obrazom mapy głębi, konieczne jest określenie nieco innych kryteriów doboru parametrów procesu kodowania. Podstawą nie może być tu, jak w przypadku kodowania jednowidokowego, proste obliczenie miary PSNR. Konieczność uwzględnienia w budżecie bitów reprezentacji dwóch rodzajów danych o zupełnie różnym charakterze i przeznaczeniu wyklucza możliwość prostego wnioskowania o jakości materiału z systemu wielowidokowego jedynie na podstawie miary PSNR dla zrekonstruowanego obrazu.

Miara PSNR dla zrekonstruowanych map głębi nie daje pełnej wiedzy o jakości map głębi dla całości systemu wielowidokowego [Hewa\_09]. O prawdziwości takiej tezy świadczą również wyniki zaprezentowane w [Leon\_08]. Obiektywne miary jakości liczone dla map głębi nie są dobrą miarą jakości zdekodowanych sekwencji wielowidokowych z dwóch ważnych powodów:

- po pierwsze, obraz głębi nie jest nigdy prezentowany użytkownikom wprost, a jedynie używany jest jako informacja dodatkowa, niezbędna na pewnych etapach przetwarzania danych,
- po drugie, przynajmniej przy obecnym stanie technologii, oryginalne, nieskompresowane mapy głębi nie zawsze są dokładne i nie zapewniają uzyskania bardzo wysokiej jakości wyników syntezy (świadczą o tym przytoczone w dalszej części pracy wyniki badań). Uzyskanie całkowicie poprawnych map głębi – czyli takich, które zawierają informacje całkowicie zgodne z rzeczywistą strukturą sceny – możliwe jest obecnie jedynie dla sekwencji syntetycznych.

Wobec tego, używanie dostępnych obecnie map głębi dla sekwencji naturalnych jako wzorca

może prowadzić do błędnych wniosków, gdyż można oczekiwać, że wraz z postępem badań, jakość dostępnych map głębi będzie coraz lepsza.

W procesie oceny jakości map głębi należy rozważyć sposób, w jaki zniekształcenia wynikające ze stratnego kodowania map głębi i obrazów wpływają na jakość syntezy widoku. Duży wpływ na jakość syntetycznego widoku będą miały nie tylko efekty związane z rozmywaniem i zniekształcaniem krawędzi w mapach głębi, ale również zmiany wartości próbek mapy głębi. Sprawdzenia wymaga zatem wpływ tych zniekształceń na końcowy wynik procesu, w którym mapy głębi są wykorzystywane. Należy zatem badać, jaki wpływ mają wprowadzane zniekształcenia na widok syntetyczny.

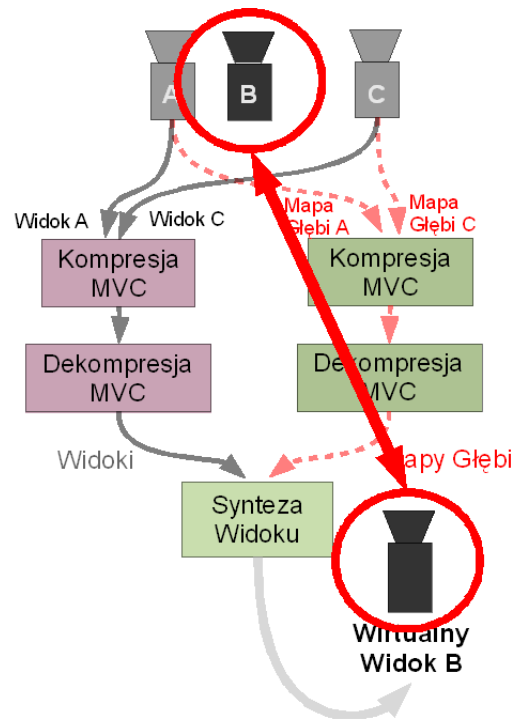
Charakterystyczną cechą procesu syntezy jest to, że nie wszystkie błędy lub zniekształcenia w mapach głębi mają jednakowy wpływ na jakość syntezy widoku. Zależnie od sceny, zniekształcenia pewnych fragmentów mapy głębi mogą się w ogóle nie ujawniać w syntezy obrazie, a zniekształcenia innych fragmentów mapy głębi mogą spowodować drastyczny spadek jakości obrazu syntezy. Wobec tego jakości rekonstrukcji map głębi nie można rozpatrywać w oderwaniu od jej konkretnego zastosowania w systemach wielowidokowych [Kim\_09][Kim\_10][Zhang\_11]. Mapa głębi nie może być traktowana w procesie szacowania jakości rekonstruowanych danych na równi z obrazem – jej jakość nie powinna być oceniana poprzez porównanie z oryginalną mapą głębi.

Obecnie najpowszechniej wykorzystywaną techniką oceny jakości zrekonstruowanych danych w systemach wielowidokowych jest pomiar PSNR widoków syntetycznych [MPEG\_CfP] – patrz wzór (3.1). Zaletą takiej metody jest jej prostota implementacyjna i możliwość wyznaczenia jakości jedynie na podstawie znajomości obrazu oryginalnego i zniekształconego. Do najważniejszych wad takiej metody można zaliczyć niepełną korelację uzyskiwanych za jej pomocą wyników z wynikami ocen subiektywnych [Wang\_09]. W przypadku sekwencji wielowidokowych problemem jest również wybór obrazu odniesienia – obrazu, z którym porównywany jest wytworzony na podstawie zrekonstruowanych danych widok syntetyczny. Rozpatrywane są w tym kontekście z reguły dwie możliwości.

Pierwszą metodą jest wykorzystanie jako obrazu odniesienia obrazu z rzeczywistej kamery umieszczonej w miejscu, z którego widok jest syntezy. Schemat takiej metody określania jakości rekonstruowanych danych przedstawiony jest na rysunku 4.2. W tym przypadku kompresji podlegają obrazy z kamer A oraz C wraz z odpowiadającymi im mapami głębi. Kompresja obrazów dokonywana jest za pomocą kodera (tutaj kodera MVC). Kompresja map głębi dokonywana jest oddzielnie, również za pomocą kodera MVC. Ocena jakości jest dokonywana na podstawie porównania miarą PSNR syntezy widoku z obrazem z kamery B, umieszczonej w dokładnie



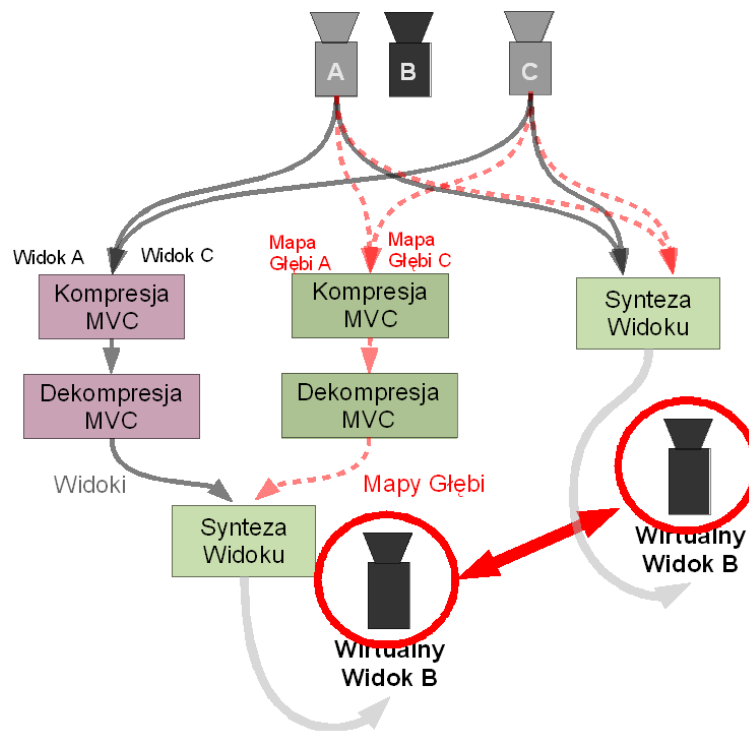
tym samym miejscu, z którego widok jest syntezywany.



Rysunek 4.2. Ocena jakości sekwencji wielowidkowej z głębią po kompresji – porównanie z rzeczywistą kamerą. Linia ciągłą oznaczony został przepływ danych o obrazie, linią przerywaną – danych o mapach głębi.

W ten sposób porównania dokonuje się z rzeczywistym, niezniekształconym widokiem, a więc takim, jakiego oczekuje odbiorca – uzyskane wyniki będą zatem lepiej odzwierciedlać wyniki uzyskane przy badaniu subiektywnej jakości [Nur\_10]. Niestety, porównywanie z widokiem z rzeczywistej kamery ma bardzo poważną wadę – można je przeprowadzić tylko wtedy, gdy dostępny jest widok z rzeczywistej kamery oznaczonej na rysunku jako B. W obecnie rozważanych scenariuszach zastosowań sekwencji wielowidkowych widok ten może nie być dostępny. Dodatkową wadą opisywanej metody oceny jakości jest duży wpływ algorytmu syntezy widoku na końcowy wynik.

Drugą metodą oceny jakości map głębi, której wyniki nie są zakłócane wpływem dodatkowych, poza efektami związanymi z kompresją, czynników jest natomiast porównanie otrzymanego z użyciem zdekompresowanych danych widoku wirtualnego z widokiem syntetycznym wyliczonym na podstawie nieskompresowanych danych wejściowych. Schemat takiej metody pomiaru przedstawiony jest na rysunku 4.3.



Rysunek 4.3. Ocena jakości sekwencji wielowidkowej z mapami głębi po kompresji – porównanie z widokiem wirtualnym wyznaczonym na podstawie nieskompresowanych danych.

W przypadku z rysunku 4.3 kompresji z użyciem kodera MVC podlegają obrazy z kamer A oraz C. Odpowiadające im mapy głębi kodowane są w podobny sposób, za pomocą kodera MVC. Syntezowany z użyciem zrekonstruowanych danych widok wirtualny porównywany jest z widokiem wirtualnym wygenerowanym na podstawie nieskompresowanych danych (ścieżka po prawej stronie rysunku 4.3). W ten sposób wynik pomiaru jakości w mniejszym stopniu zależy od cech algorytmu syntezy użytego w systemie. Wadą tej metody jest jednak fakt, że uzyskiwane za jej pomocą wyniki nieco gorzej odzwierciedlają subiektywne oceny materiałów wynikowych niż metoda opisana wcześniej, używająca widoku rzeczywistego jako widoku odniesienia.

Mimo wymienionej wady metoda ta jest powszechnie wykorzystywana [El-Y\_10] [MPEG\_CfP], co zawdzięcza również temu, że do jej użycia nie jest konieczne dostarczenie rzeczywistego widoku z kamery oznaczanej na rysunkach 4.2 oraz 4.3 jako B. Jak to już zostało wspomniane, w istocie kamery takiej może w systemie wielokamerowym nie być, bądź to ze względu na ograniczenie kosztów, bądź to ze względu na fizyczne ograniczenia systemu. Kamery nie da się przecież umieścić w dowolnym miejscu, choćby ze względu na jej rozmiary.

W dalszej części pracy wyniki podawane będą w postaci uzyskanej za pomocą drugiej metody, a tam, gdzie będzie to możliwe, dodatkowo podane zostaną wyniki uzyskane pierwszą metodą.

## **5. Bezpośrednia metoda kodowania obrazów i map głębi**

Przesłanie do odbiorcy zarówno kolorowego obrazu jak i mapy głębi wymaga odpowiedniego podziału dostępnej prędkości bitowej pomiędzy te dwa rodzaje danych. W rozdziale 4 zaprezentowany został przegląd metod kompresji map głębi dla systemów wielowidokowych. W dużej części bazują one na algorytmach kompresji obrazu nieruchomego, a te, które dostosowane są do kompresji map głębi bazują na koderze AVC z modyfikacjami. Najprostszym sposobem przesłania sekwencji wielowidokowej i map głębi jest jednak zastosowanie do kodowania zarówno obrazów jak i map głębi kodera MVC. Podejście wykorzystujące do kompresji głębi niezmodyfikowany koder jest najprostszym z możliwych i jest stosowane w działających systemach [Akar\_07]. Wadą takiej metody kodowania jest brak możliwości wykorzystania specyficznych właściwości kodowanych map głębi, niezaprzeczną zaletą jest jednak fakt wykorzystywania istniejącego, dobrze poznanego kodera do kodowania zarówno obrazów jak i map głębi. W wypadku wykorzystania kodera MVC do kodowania zarówno obrazów jak i map głębi istotne jest jednak takie dobranie parametrów kodowania, aby przy założonej maksymalnej prędkości bitowej uzyskać jak najwyższą jakość obrazów syntezowanych na podstawie zdekodowanych danych. Konieczne jest zatem opracowanie metody doboru parametrów kodowania obrazów i map głębi. Najwygodniejszą metodą jest uzależnienie parametrów kodowania map głębi od parametrów kodowania obrazów. W dalszej części rozdziału prezentowane jest autorskie ujęcie problemu doboru parametrów kodowania map głębi w bezpośrednim kodowaniu obrazów i map głębi.

### **5.1. Wpływ zniekształceń kodowania obrazów i głębi na jakość sekwencji syntezowanych**

W celu określenia zasad dobierania parametrów kodowania należy zbadać, jak kompresja obrazów i map głębi z wykorzystaniem kodera MVC wpływa na jakość obrazów syntezowanych z ich użyciem

#### **5.1.1. Przegląd literatury**

Problem wpływu kodowania sekwencji na jakość syntezowanych widoków poruszany jest w pracach [Morv\_07a][Morv\_09], jednak przedstawione tam wyniki uzyskane zostały przy bardzo

silnie ograniczających założeniach – wykorzystane zostało jedynie kodowanie wewnątrzobrazowe. Co więcej, badania oparte są na analizie wyników jedynie dla dwóch sekwencji o bardzo zbliżonym charakterze. Podobnie ograniczone są badania, których wyniki zaprezentowane są w pracy [Mull\_09]. Problem wpływu kompresji na jakość generowanych widoków syntetycznych poruszany jest przy okazji przeglądu różnych metod transmisji sekwencji stereoskopowych w opracowaniu [Tech\_09] oraz w [Merk\_09a], jednak nie podlega on głębszej analizie.

Nieco inne podejście prezentują autorzy pracy [El-Y\_10], gdyż rozważają oni jedynie wpływ kompresji map głębi na jakość widoku syntezy. Pomijają w ten sposób ważny aspekt wzajemnej zależności wpływu zniekształceń obrazów i map głębi na końcowy efekt syntezy.

Analiza uwzględniająca wzajemne zależności pomiędzy jakością obrazu i mapy głębi przedstawiona jest w pracy [Cheu\_11], jednak jej autorzy skupiają się na opracowaniu algorytmu dobierającego parametry kodowania w sposób adaptacyjny w zależności od wybranego widoku syntetycznego.

W literaturze znanej autorowi rozprawy brak jest wyników eksperymentalnych badań dotyczących wpływu kompresji obrazu i map głębi na jakość syntezy obrazu, szczególnie dla zdecydowanej mapy głębi.

### **5.1.2. Cel badań**

Celem prezentowanych badań jest przedstawienie ilościowych wniosków dla analizy zagadnienia wpływu kompresji obrazów i map głębi na jakość syntezy widoku. Ważnym celem było również stwierdzenie, czy możliwe jest uzyskiwanie dobrej jakości widoków syntetycznych w przypadku, gdy wykorzystane były mapy głębi o zmniejszonej rozdzielczości. Należało również zbadać, jak zmniejszenie rozdzielczości mapy głębi wpływa na efektywność jej kodowania przy uwzględnieniu jakości generowanych widoków wirtualnych.

Badania, których wyniki są prezentowane w tym rozdziale zapoczątkowane były pracami udokumentowanymi w [Klim\_09][Klim\_09a][Klim\_09b][Doma\_10].

### **5.1.3. Metodologia badań**

Badania polegały na kodowaniu obrazów i map głębi sekwencji testowych z różnymi wartościami parametru kwantyzacji. Wykorzystane zostały po dwa widoki i odpowiadające im mapy głębi z każdej przebadanej sekwencji. Mapy głębi kodowane były w pełnej rozdzielczości oraz z użyciem decymacji do  $\frac{1}{2}$  i do  $\frac{1}{4}$  pierwotnych rozmiarów (zarówno w pionie jak i w poziomie). W celu zmiany prędkości bitowych zmieniana była wartość parametru kwantyzacji. W

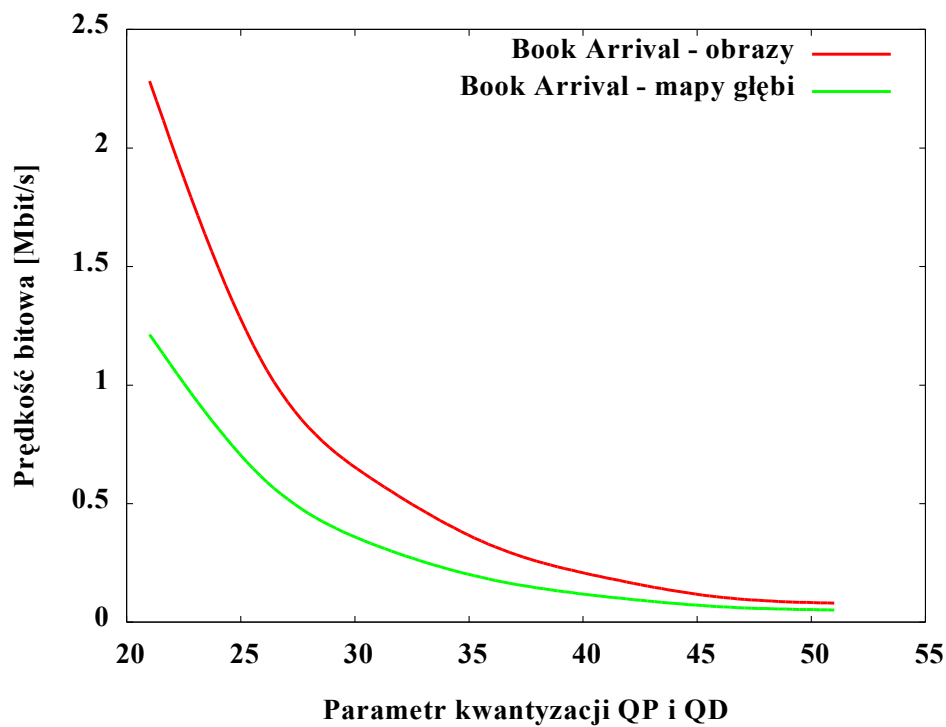
przypadku parametru kwantyzacji w kodowaniu obrazów w dalszej części tekstu używany będzie skrót QP, a w przypadku indeksu parametru kwantyzacji w kodowaniu mapy głębi w dalszej części tekstu używany będzie skrót QD. Szczegółowy opis warunków przeprowadzonego eksperymentu zawarty jest w Aneksie II.

## **5.1.4. Wyniki eksperymentu**

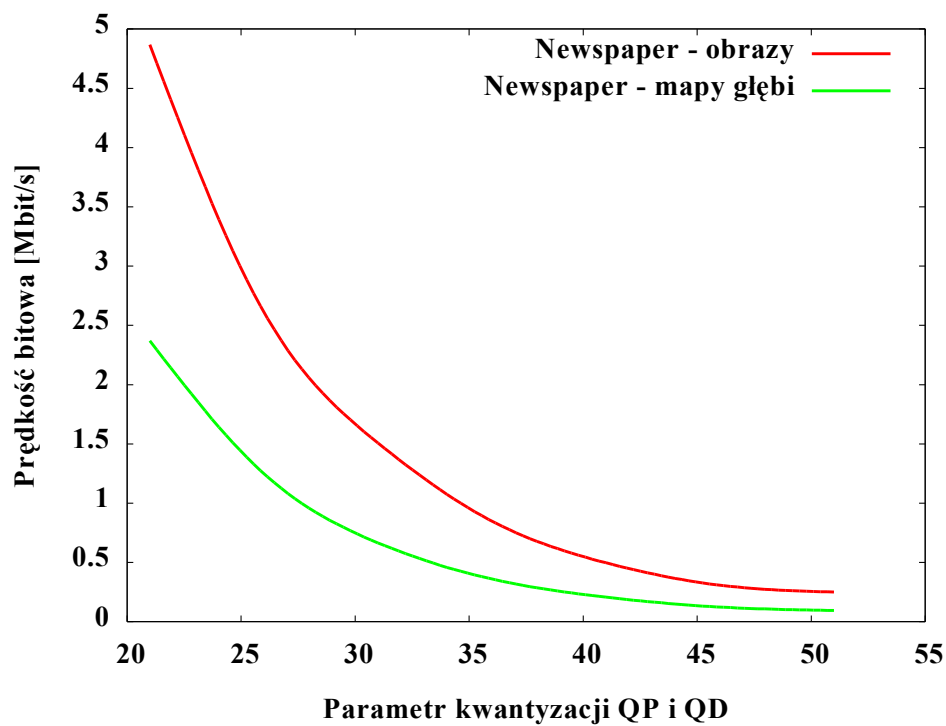
Wyniki przedstawione zostały w formie wykresów dla każdej z trzech użytych w eksperymencie sekwencji. Jako pierwsze przedstawione są wyniki wstępnego porównania wydajności kompresji obrazów i map głębi (punkt 5.1.4.1), a w dalszej części prezentowane są wyniki pomiaru jakości widoku syntezowanego z wykorzystaniem skompresowanych danych przy wykorzystaniu mapy głębi w pełnej rozdzielczości (punkt 5.1.4.2) i przy wykorzystaniu mapy głębi o rozdzielczości zredukowanej w procesie decymacji (punkt 5.1.4.3).

### **5.1.4.1. Efektywność kompresji obrazów i map głębi**

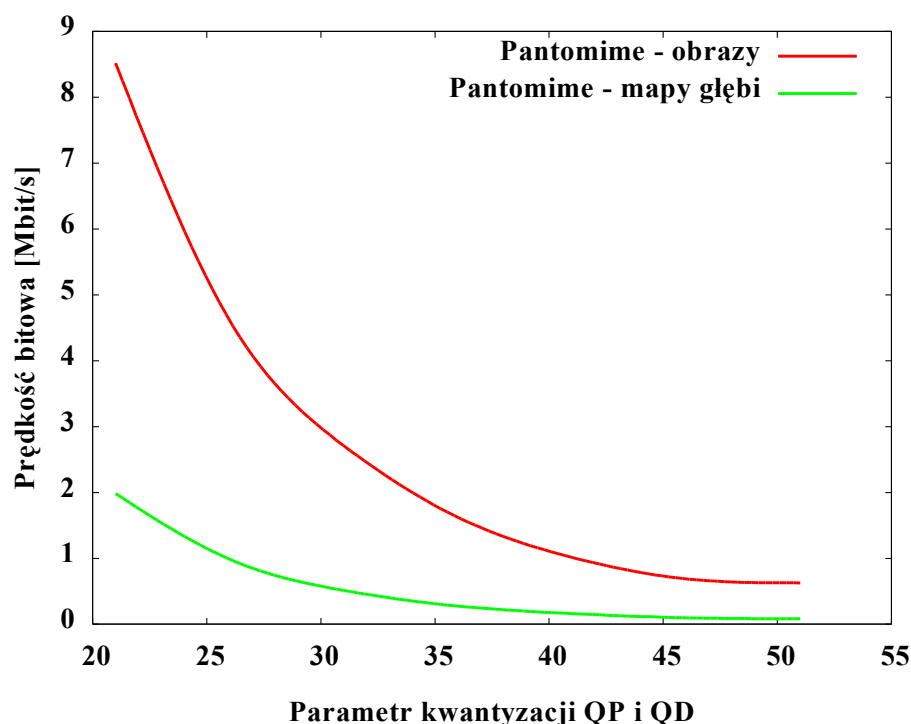
W celu oszacowania wydajności kodowania map głębi w porównaniu z kodowaniem obrazów, przeprowadzony został eksperyment polegający na kompresowaniu sekwencji wielowidkowej złożonej z obrazów oraz na kompresowaniu, z identycznymi ustawieniami kodera, sekwencji map głębi w pełnej rozdzielczości. użytym koderem był koder MVC. Wykorzystano trzy testowe sekwencje wielowidkowe z mapami głębi: Book Arrival [Feld\_08], Newspaper [Ho\_08] oraz Pantomime [Tani\_08]. Uzyskane w ten sposób wyniki zostały porównane, a wyniki przedstawione na wykresach. Podane na rysunkach 5.1, 5.2 oraz 5.3 wartości miary jakości PSNR uzyskane zostały przez wyliczenie wartości średniej miar jakości dla wszystkich ramek z obydwu badanych widoków. Wartości prędkości bitowych przedstawione zostały w megabitach na sekundę i przedstawiają one wartości sumaryczne dla obu widoków. Wykresy przedstawiają wykresy efektywności kompresji obrazów i map głębi dla trzech przetestowanych sekwencji. Przedstawione zostały na nich krzywe obrazujące zależność wymaganej prędkości bitowej od indeksu parametru kwantyzacji wykorzystanego przy kompresji. Pokazane są oddzielne krzywe dla kompresji obrazów i map głębi.



Rysunek 5.1. Zależność prędkości bitowej od indeksu parametru kwantyzacji dla sekwencji Book Arrival – zarówno dla obrazów jak i map głębi.



Rysunek 5.2. Zależność prędkości bitowej od indeksu parametru kwantyzacji dla sekwencji Newspaper – zarówno dla obrazów jak i map głębi.

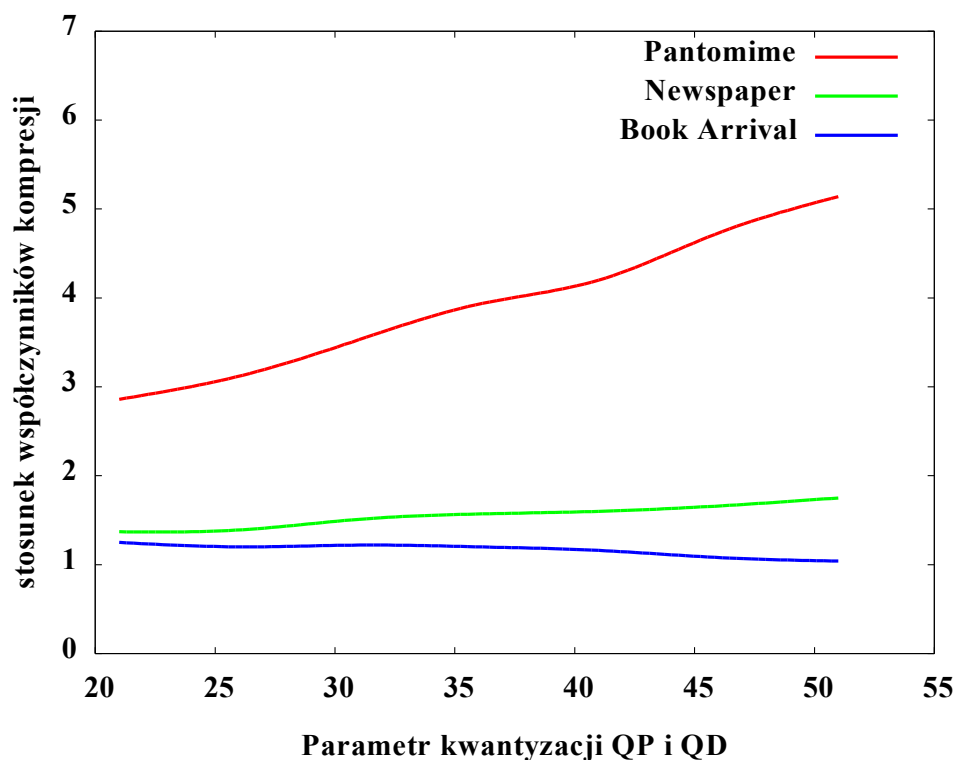


Rysunek 5.3. Zależność prędkości bitowej od indeksu parametru kwantyzacji dla sekwencji Pantomime – zarówno dla obrazów jak i map głębi.

Wyniki eksperymentu dowodzą słuszności założeń mówiących o większej podatności na kompresję map głębi niż obrazów. Wynika to z występowania dużej ilości gładkich powierzchni w obrazach map głębi, które dają się efektywnie zakodować. W tabeli 5.1 przedstawione jest przykładowe porównanie stopnia kompresji danych dla sekwencji wykorzystanych w eksperymencie, a na rysunku 5.4 przedstawiona jest zależność stosunku współczynników kompresji od wartości parametrów QP i QD dla trzech przebadanych sekwencji.

Tabela 5.1. Porównanie efektywności kodowania dla przykładowych sekwencji.

rodzaj danych		Prędkość bitowa dla 2 widoków [Mbit/s]		współczynnik kompresji	stosunek współczynników kompresji
		nieskompresowane	skompresowane (QP = 31)		
Book arrival	obrazy	314,64	0,59	533	1,23
	mapy głębi	209,76	0,32	656	
Newspaper	obrazy	566,23	1,51	375	1,50
	mapy głębi	377,49	0,67	563	
Pantomime	obrazy	867,34	2,71	320	3,54
	mapy głębi	578,22	0,51	1134	



Rysunek 5.4. Zależność stosunku współczynników kompresji od parametrów QP i QD.

Dla wszystkich trzech sekwencji redukcja prędkości bitowej dla głębi jest większa niż dla obrazów kodowanych z użyciem tej samej wartości parametru kwantyzacji. Jak widać na rysunku 5.4, dla wszystkich przebadanych sekwencji stosunek współczynnika kompresji głębi do współczynnika kompresji obrazów jest większy niż 1. Oznacza to, że mapy głębi są, bardziej niż obrazy, podatne na kodowanie z użyciem kodera MVC. Wyraźnie większe wartości współczynnika kompresji głębi dla sekwencji Pantomime związane są z niewielką liczbą obiektów w tej sekwencji, a co za tym idzie – mało złożoną mapą głębi, podatną na kompresję.

#### 5.1.4.2. Badanie jakości syntezy ze zdekodowanych danych

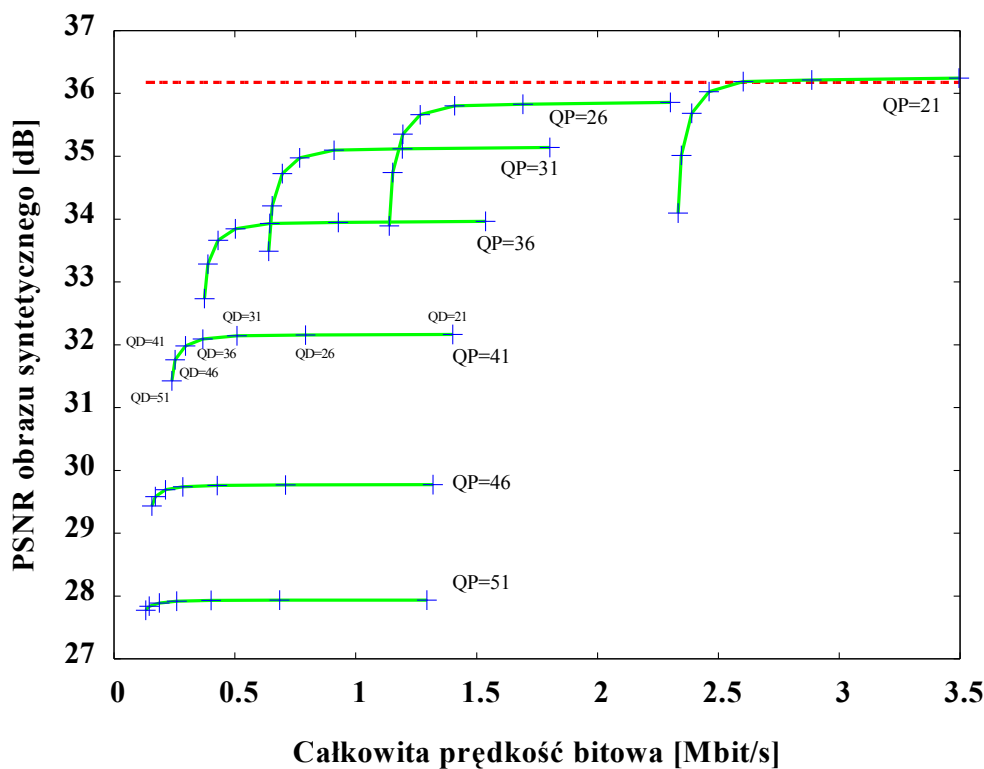
Kolejnym etapem badań było zbadanie, w jaki sposób zmiany parametrów kodowania mapy głębi i obrazu wpływają na jakość obrazu syntezy. Na poniższych wykresach przedstawione zostały wyniki pomiaru jakości widoków wirtualnych syntezy z zrekonstruowanych obrazów i map głębi, zgodnie z tabelą 7.5 (Aneks II).

Przedstawione krzywe odzwierciedlają jakość obrazu wirtualnego (mierzoną jako PSNR luminancji w porównaniu z rzeczywistym obrazem z kamery) w zależności od sumarycznej

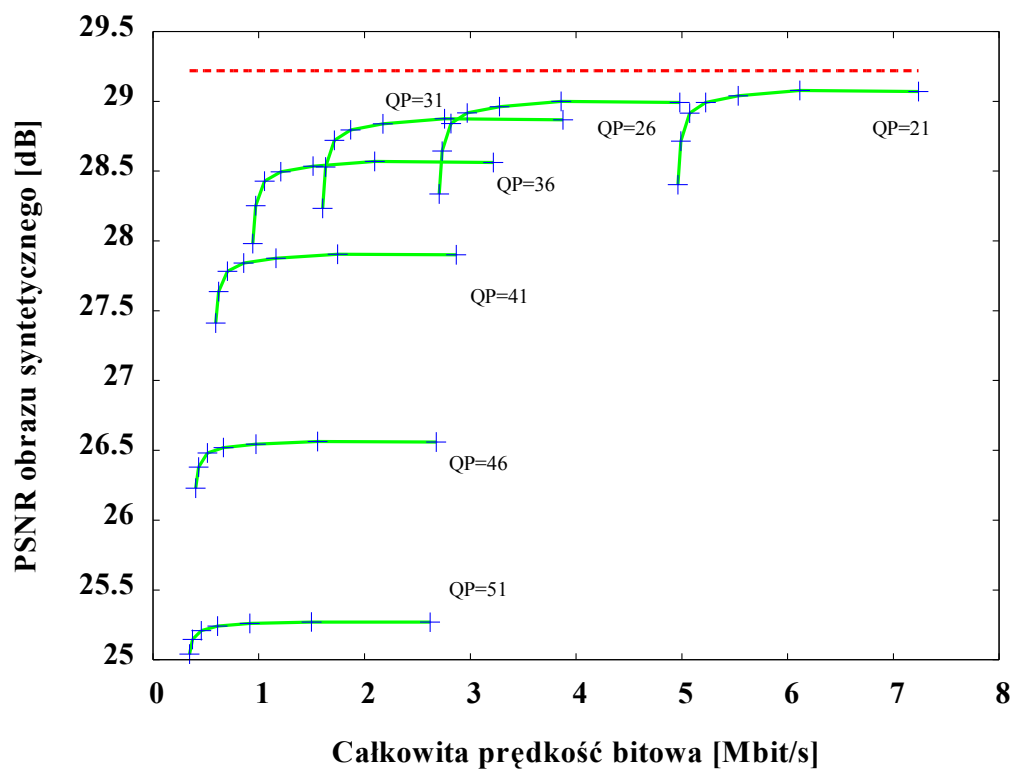


prędkości bitowej potrzebnej do przesłania dwóch sekwencji obrazów i dwóch sekwencji map głębi.

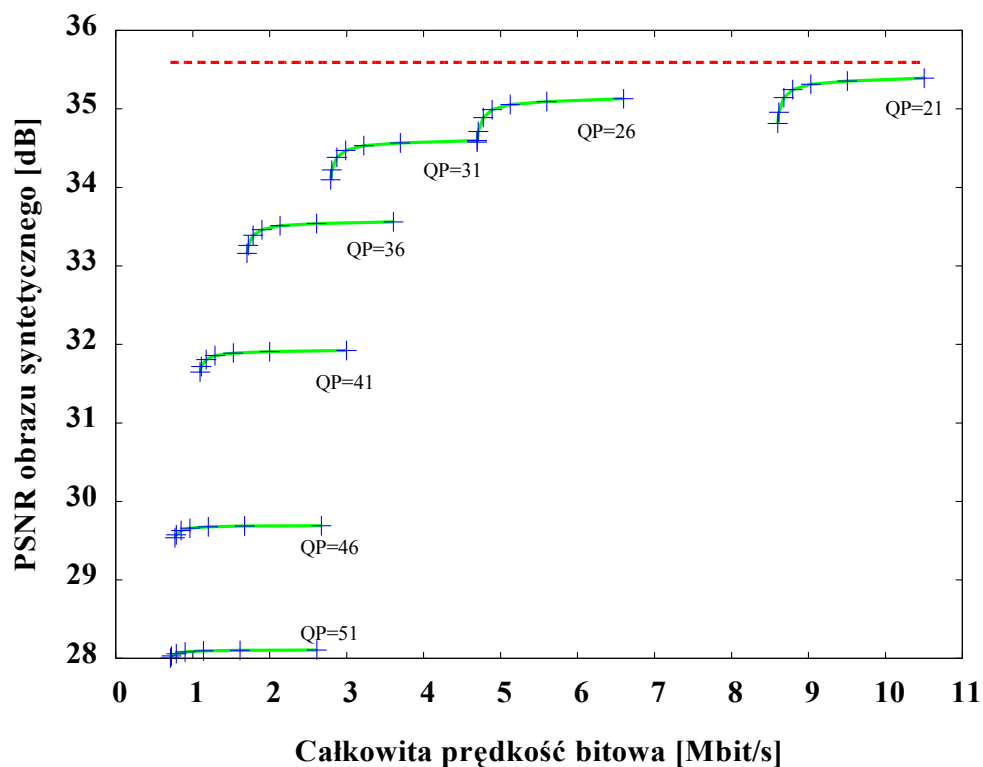
Na wykresach, za pomocą krzyży zaznaczone są kolejne punkty pomiarowe. **Czerwona, przerywana linia** wyznacza poziom wartości PSNR uzyskanej dla widoku syntezowanego z wykorzystaniem nieskompresowanych obrazów i map głębi. Dodatkowo na wykresach, przy kolejnych krzywych, umieszczone są etykiety z wartością indeksu parametru kwantyzacji wykorzystanego przy kompresji obrazów (QP) dla danej krzywej. Kolejne punkty pomiarowe położone na takiej krzywej odpowiadają kombinacji tej wartości QP kolejno ze wszystkimi wartościami indeksu parametru kwantyzacji wykorzystanymi przy kompresji map głębi (QD), wymienionymi w tabeli 7.4 (Aneks II).



Rysunek 5.5. Zależność jakości widoku wirtualnego od całkowitej prędkości bitowej dla sekwencji Book Arrival – opis w tekście



Rysunek 5.6. Zależność jakości widoku wirtualnego od całkowitej prędkości bitowej dla sekwencji Newspaper – opis w tekście



Rysunek 5.7. Zależność jakości widoku wirtualnego od całkowitej prędkości bitowej dla sekwencji Pantomime – opis w tekście

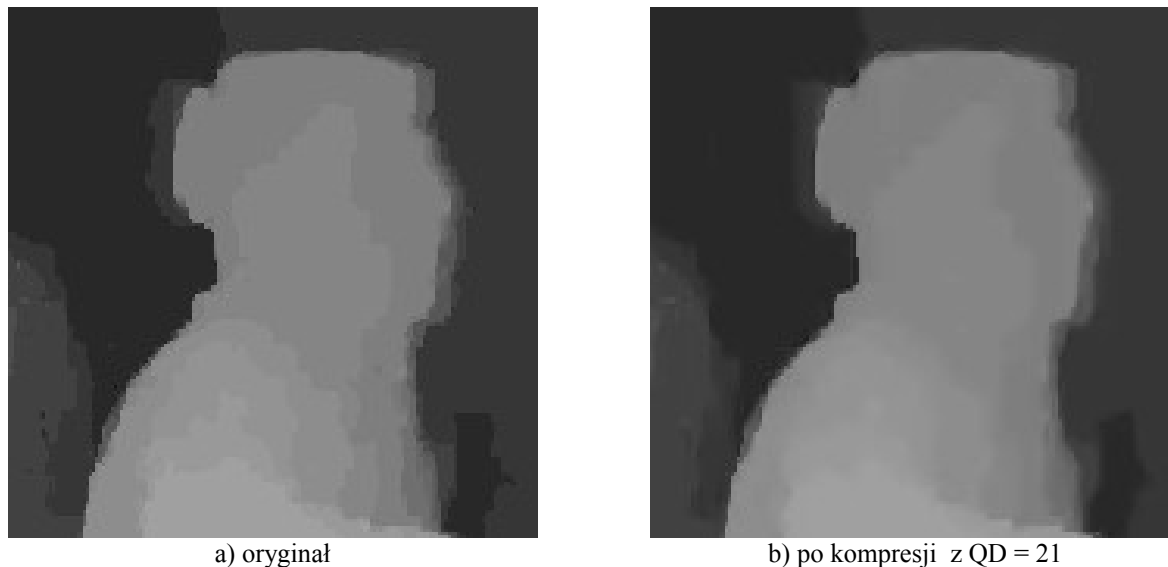
Otrzymane wyniki potwierdzają użyteczność niezmodyfikowanego kodera MVC do kompresji map głębi. Zniekształcenia powstające w wyniku kompresji obrazów map głębi nie dyskwalifikują zrekonstruowanych map głębi z użycia w procesie syntezy widoków wirtualnych.

Z przedstawionych wykresów wynika, że dla większości przebadanych punktów pomiarowych zmiana indeksu QD ma znacznie mniejszy wpływ na jakość syntezy widoku wirtualnego niż zmiana indeksu QP. Dla sekwencji Book Arrival maksymalna zmiana jakości dla całego zakresu zmian QD wynosi 2 dB, a dla przebadanego zakresu zmian QP (od QP = 21 do QP = 51) wynosi 8 dB. Dla sekwencji Newspaper wartości te wynoszą odpowiednio 0,8 dB i 3,7 dB, a dla sekwencji Pantomime – 0,6 dB i 7,4 dB.

Wynika stąd wniosek, że zniekształcenia w mapach głębi mają mniejszy wpływ na jakość syntezy obrazu niż zniekształcenia w obrazach z kamer. Można przy tym zauważyć, że zmiana wartości indeksu QD w sposób monotoniczny wpływa na jakość kodowanej mapy głębi w mierze PSNR. Dla punktów pomiarowych, dla których wartości indeksu QP przyjmowały większe wartości (QP = 46, QP = 51), wpływ wartości indeksu QD na jakość generowanej sekwencji jest bardzo niewielki – zmiany wynoszą około 0,1 dB. Wpływ ten jest nieco większy dla punktów pomiarowych z niskimi wartościami indeksu QP (QP = 21), jednak, co warto podkreślić, staje się on wyraźny dla wartości QD powyżej pewnej granicznej wartości. Na podstawie przedstawionych wyników można określić granicę jako QD = 31.

Ciekawą obserwacją, którą można poczynić na podstawie przedstawionych wykresów jest to, że w niektórych przypadkach możliwe jest poprawienie jakości widoku syntezy, czyli uzyskanie wyższej wartości miary PSNR niż przy użyciu oryginalnych, nieskompresowanych map głębi i obrazów. Efekt ten widoczny jest dla sekwencji Book Arrival (rysunek 5.5). Zastosowanie kompresji z niewielką wartością parametru kwantyzacji QD (na przykład QD = 21) powoduje wygładzenie map głębi i usuwa z nich drobne elementy powodujące błędy w syntezy. W ten sposób kompresja może wpłynąć na poprawę osiągniętych w przypadku używania map głębi o niskiej jakości. Dodatkowo kompresja powoduje wprowadzenie pośrednich wartości punktów w mapach głębi, przez co jakość syntezy w niektórych fragmentach mapy głębi jest poprawiana [Stank\_10]. Efekt ten jest niepożądany w przypadku krawędzi obiektów, gdyż powoduje ich rozmywanie na obrazie syntezy, jednak poprawia jakość syntezy dla wnętrza dużych obiektów. Ponieważ znakomita większość próbek mapy głębi należy zazwyczaj do spójnych obszarów odpowiadających pojedynczym obiektom w scenie, obserwowana jest poprawa jakości syntezy spowodowana przez rozmycie granic obszarów o różnych wartościach głębi. Efekt rozmywania krawędzi poprzez wprowadzanie wartości pośrednich jest też wykorzystywany w narzędziach do poprawy jakości map głębi, jak na przykład w narzędziu prezentowanym w [Stank\_10]. Przykładowy fragment

oryginalnego obrazu mapy głębi i obrazu po kompresji z indeksem QD = 21 pokazany jest na rysunku 5.8.



Rysunek 5.8. Fragment obrazu mapy głębi: a) – obraz oryginalny, b) – obraz po kompresji z indeksem kwantyzacji wynoszącym 21 (widoczne rozmycie powodowane przez kompresję).

Wpływ podziału całkowitej prędkości bitowej pomiędzy część przeznaczoną do transmisji map głębi a część przeznaczoną do transmisji obrazów przedstawiony został na kolejnych wykresach, od 5.9 do 5.11. Uwidoczniona jest na nich zależność jakości widoku wirtualnego w zależności od stosunku prędkości bitowej przeznaczonej na transmisję skompresowanych sekwencji obrazów map głębi do całkowitej prędkości bitowej przeznaczonej na transmisję skompresowanych sekwencji obrazów i map głębi, wyznaczonego według wzoru (5.1).

$$S = \frac{P_D}{P_T + P_D} \quad (5.1)$$

gdzie:

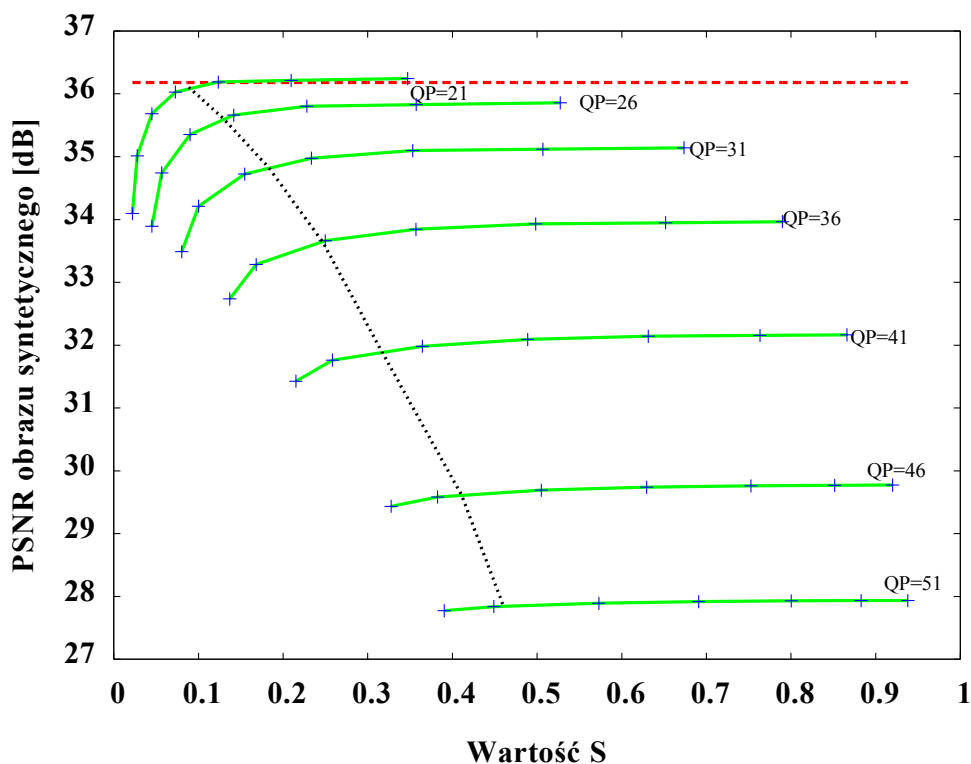
$S$  – stosunek prędkości bitowej map głębi do całkowitej prędkości bitowej,

$P_D$  – prędkość bitowa potrzebna do przesłania map głębi,

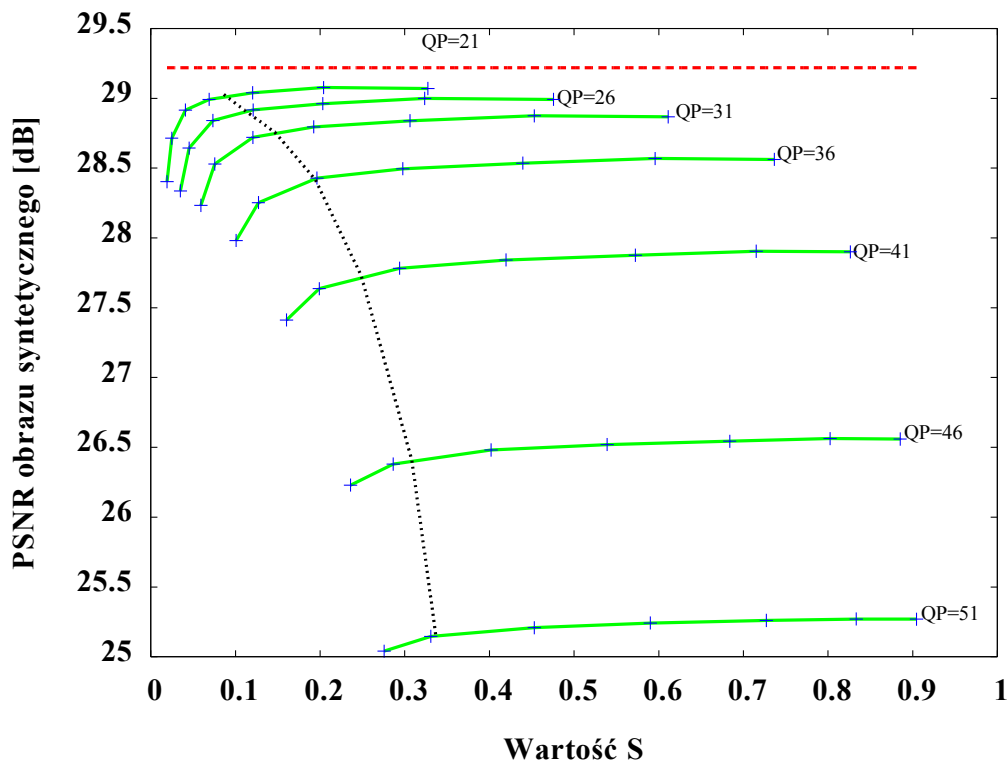
$P_T$  – prędkość bitowa potrzebna do przesłania obrazów.

Na wykresach, za pomocą krzyży, zaznaczone są kolejne punkty pomiarowe. **Przerzywana czerwona pozioma linia** wyznacza wartość miary jakości PSNR uzyskanej dla widoku syntetycznego uzyskanego z wykorzystaniem nieskompresowanych obrazów i map głębi. Dodatkowo na wykresach, przy kolejnych krzywych, umieszczone są etykiety z wartością indeksu parametru

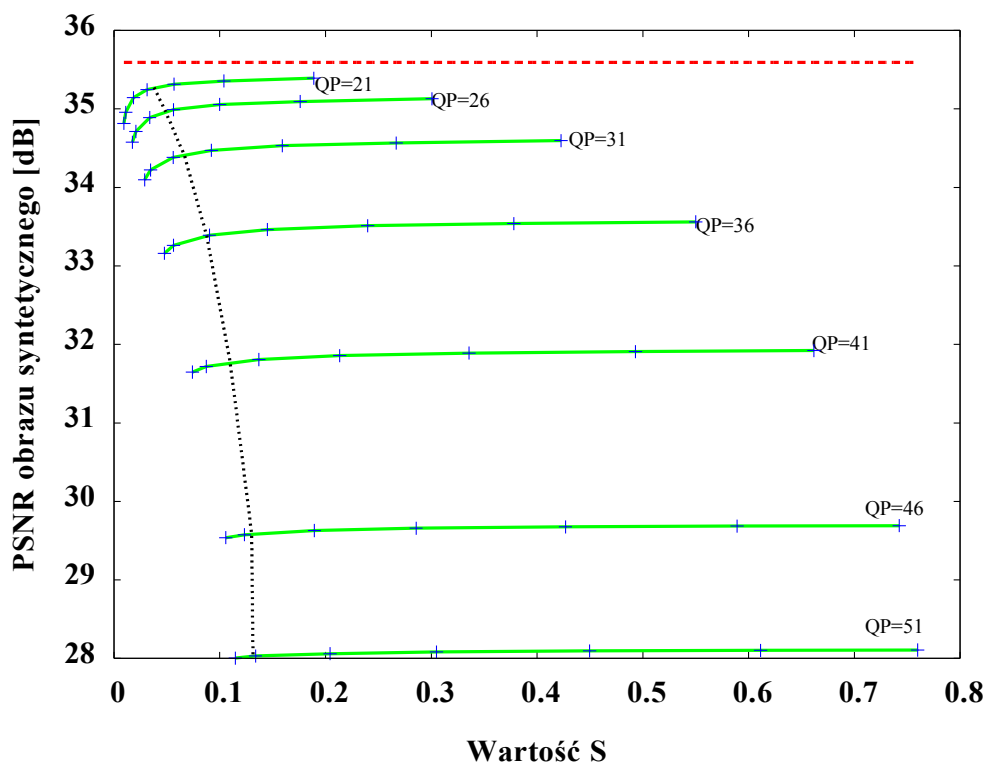
kwantyzacji wykorzystanego przy kompresji obrazów (QP) dla danej krzywej. Tak jak poprzednio, kolejne punkty pomiarowe położone na takiej krzywej odpowiadają kombinacji tej wartości QP ze wszystkimi wymienionymi w tabeli 7.4 (Aneks II) wartościami indeksu parametru kwantyzacji wykorzystanego przy kompresji obrazów map głębi (QD). Czarna przerywana linia wyznacza granicę dla zakresu wartości indeksu QD, które nie powodują pogorszenia maksymalnej dla danego QP jakości o więcej niż o 0,5 dB.



Rysunek 5.9. Zależność jakości widoku wirtualnego od stosunku prędkości bitowej danych o mapie głębi do całkowitej prędkości bitowej dla sekwencji Book Arrival – opis w tekście.



Rysunek 5.10. Zależność jakości widoku wirtualnego od stosunku prędkości bitowej danych o mapie głębi do całkowitej prędkości bitowej dla sekwencji Newspaper – opis w tekście.

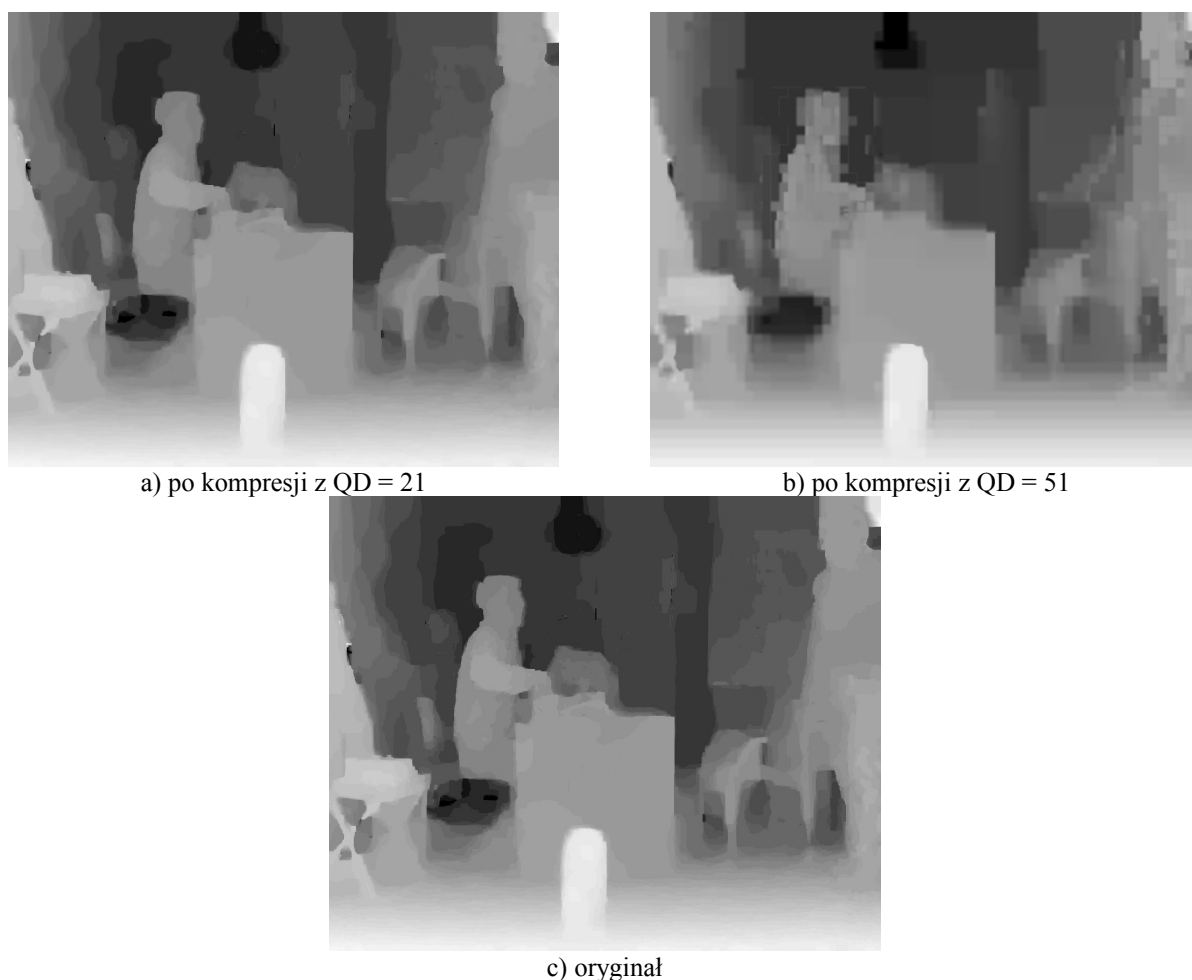


Rysunek 5.11. Zależność jakości widoku wirtualnego od stosunku prędkości bitowej danych o mapie głębi do całkowitej prędkości bitowej dla sekwencji Pantomime – opis w tekście.

Przykładowe, skompresowane obrazy mapy głębi z sekwencji Book Arrival zastosowane w eksperymencie pokazane są na rysunku 5.12.

Zależności przedstawione na wykresach od 5.9 do 5.11 pozwalają wysnuć następujące wnioski, słuszne dla przebadanego szerokiego zakresu wartości QP i QD:

- prędkość bitowa konieczna do transmisji danych o głębi może być wyraźnie mniejsza niż prędkość bitowa konieczna do transmisji danych o obrazie – dla większości przebadanych przypadków, dla  $S$  mniejszych od 0,5, jakość widoku syntezywanego pozostaje na poziomie nie różniącym się od maksymalnego dla danego QP o więcej niż 0,5 dB,
- dla większości punktów pomiarowych informacje o obrazie zajmują większą część strumienia ( $S < 0,5$ ),
- redukcja stosunku  $S$  do wartości 0,2 a nawet 0,1 nie powoduje znaczącego obniżenia jakości generowanego widoku wirtualnego dla wysokiej jakości przesyłanego obrazu ( $QP = 21$ ).



Rysunek 5.12. Porównanie obrazów map głębi skompresowanych z użyciem skrajnych indeksów parametru kwantyzacji z oryginalnym obrazem mapy głębi dla sekwencji Book Arrival. (a) – obraz skompresowany z indeksem QD = 21, (b) – obraz skompresowany z indeksem QD = 51, (c) – obraz oryginalny, bez kompresji.

Na podstawie uzyskanych wyników można wyznaczyć zależność pozwalającą określić obszar, dla którego wpływ zmian stosunku  $S$  na jakość widoku wirtualnego jest mały. Przy założeniu, że „mały wpływ” oznaczać będzie spadek jakości o mniej niż 0,5 dB PSNR, można oszacować, że wartość indeksu QD powinna być nie większa niż wynikająca ze wzoru, wyznaczonego przy użyciu regresji liniowej dla punktów odpowiadających spadkowi jakości o 0,5 dB dla wszystkich przebadanych sekwencji:

$$QD_{\max} = 49 - 0,005 \cdot (76 - QP)^2 \quad (5.2)$$

gdzie:

$QD_{\max}$  – maksymalna wartość indeksu QD nie powodująca znaczącego spadku jakości,

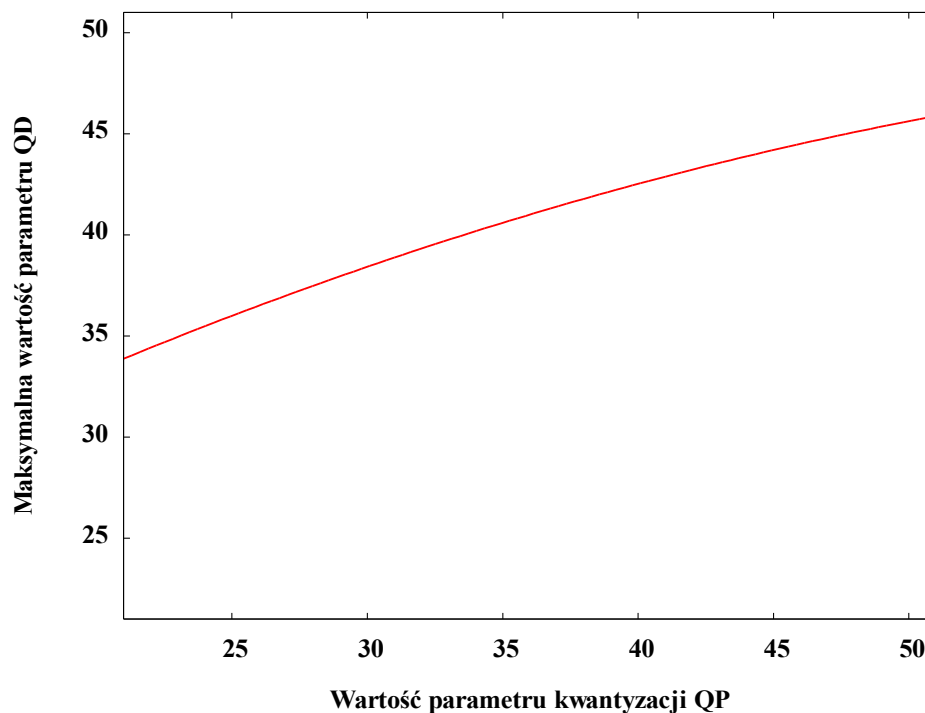
QP – wartość indeksu parametru kwantyzacji dla kodowania obrazu.

Wzór (5.2) umożliwi dobór wartości QD powodujący spadek jakości widoku syntezywanego o mniej niż 0,5 dB dla wszystkich przebadanych wartości QP.

Zastosowanie wzoru pozwala wyznaczyć krzywe dzielące płaszczyznę wykresów na dwa obszary, jak przedstawione to zostało na rysunkach od 5.9 do 5.11. Pierwszy obszar, o zbyt wysokich wartościach parametru QD, dla których obserwowany jest znaczący spadek jakości (po lewej stronie różowej krzywej na wykresach) oraz drugi obszar, w którym jakość obrazu syntezywanego jest w dużym stopniu niezależna od wyboru wartości QD (na prawo od różowej krzywej na wykresach).

Przebieg krzywej opisanej wzorem (5.2) w zależności od wartości QP ukazany jest na rysunku 5.13. Aby nie powodować znacznego pogorszenia jakości syntezywanych widoków, zastosowane wartości indeksu QD powinny mieć wartość mniejszą niż wartość wyznaczana przez czerwoną linię na wykresie.



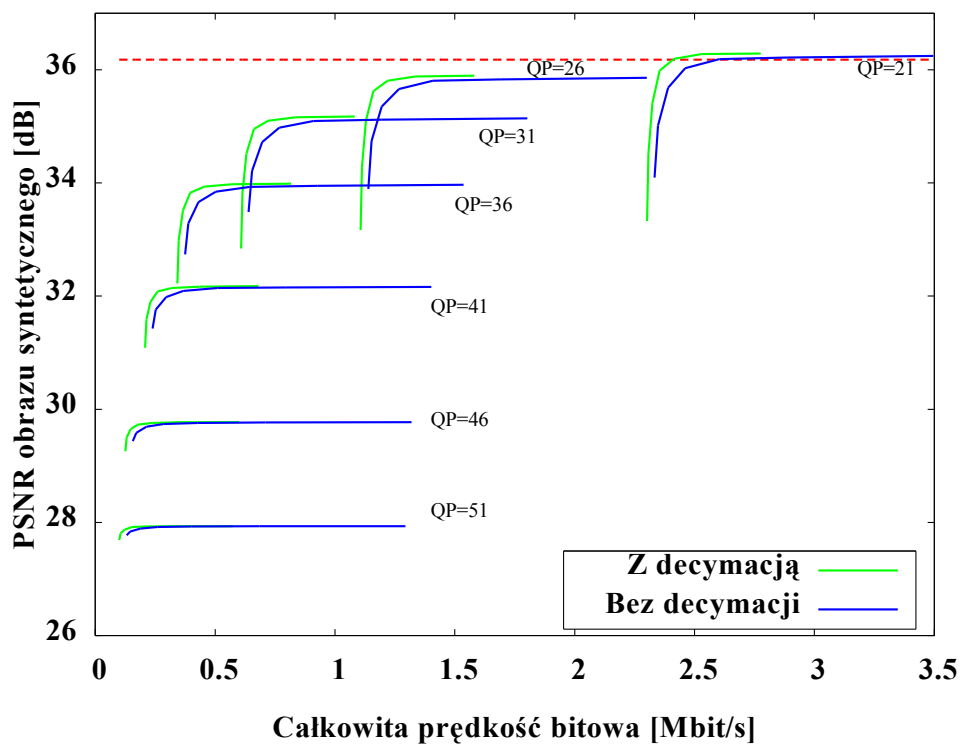


Rysunek 5.13. Wartości maksymalne parametru kwantyzacji QD w zależności od QP.

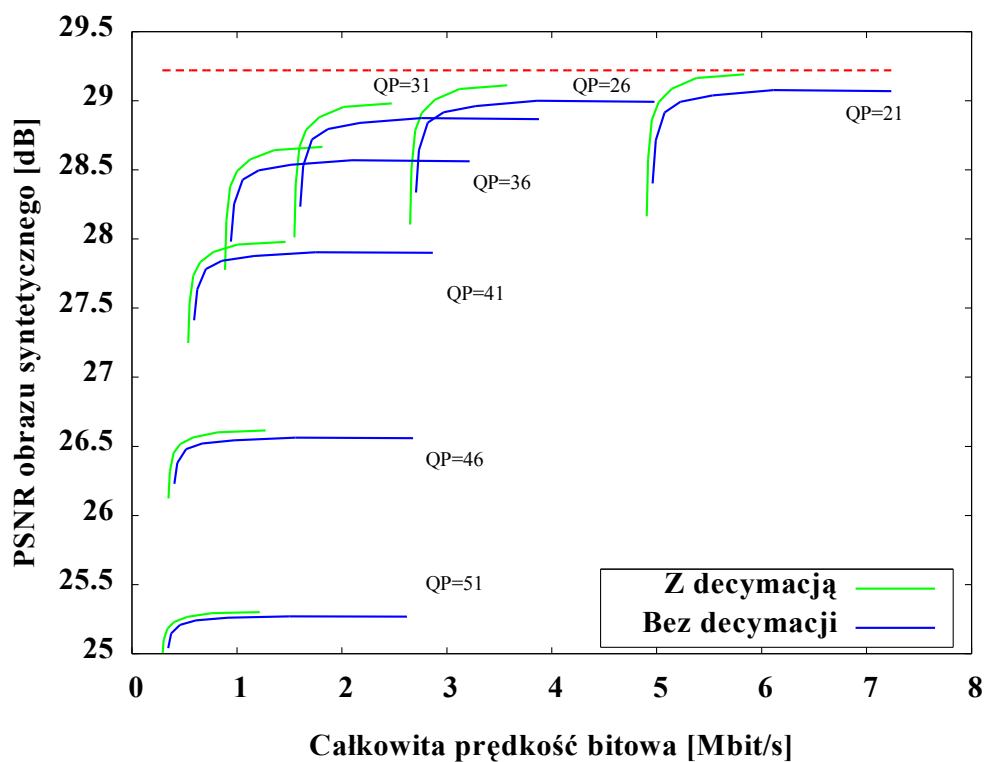
Ciekawą obserwacją wynikającą z wykresu 5.13 jest fakt, że dla bardzo dużych wartości parametru kwantyzacji QP (powyżej 43), parametr kwantyzacji QD musi być niższy niż QP. Jest to spowodowane tym, że przy zastosowaniu bardzo dużych wartości QD całkowitemu zanikowi lub znacznym zniekształceniom ulegają krawędzie w mapie głębi. Ponieważ krawędzie w mapie głębi niosą bardzo istotną część informacji dla algorytmu syntezy widoków wirtualnych, w takich wypadkach jakość widoków syntezy spada.

### 5.1.4.3. Badanie jakości syntezy ze zdecydowanych map głębi

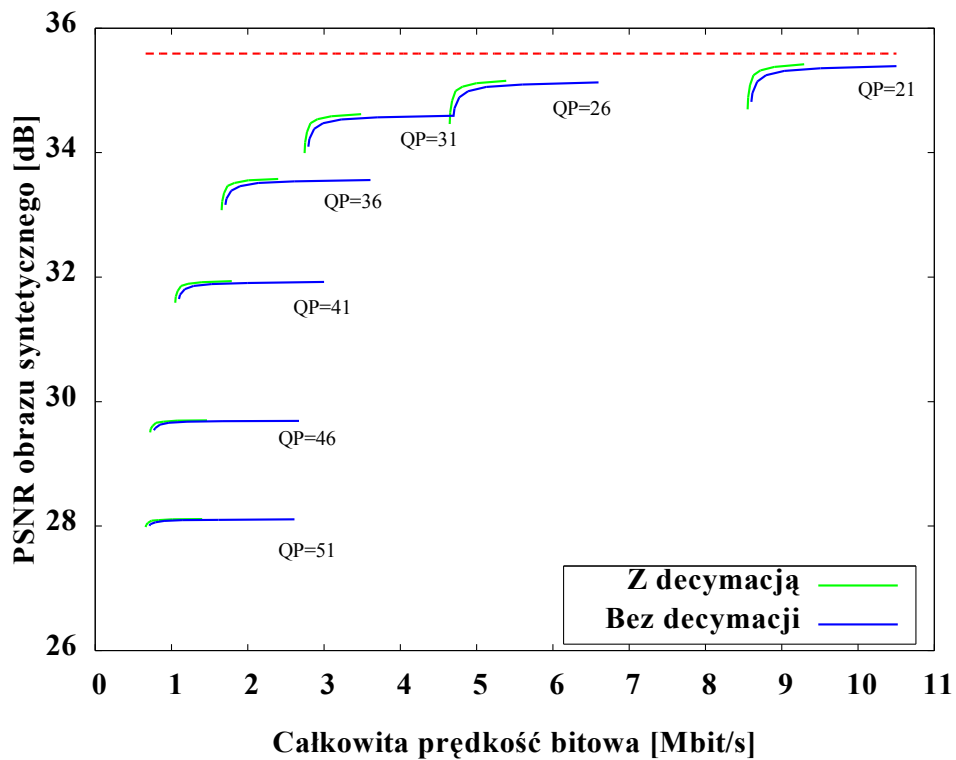
Kolejnym etapem badań było użycie map głębi o zmniejszonej rozdzielczości. Wyniki przedstawione są na poniższych wykresach. Linia **niebieska** przedstawiona obrazuje wyniki dla niezdecymowanej mapy głębi, (jak na wcześniejszych wykresach od 5.5 do 5.11). Linia **zielona** odpowiada scenariuszowi z wykorzystaniem map głębi zdecydowanych do połowy oryginalnych rozmiarów (pionowego i poziomego). **Przerywana czerwona** linia odpowiada wartości miary PSNR dla widoku syntezy z użyciem nieskompresowanych danych.



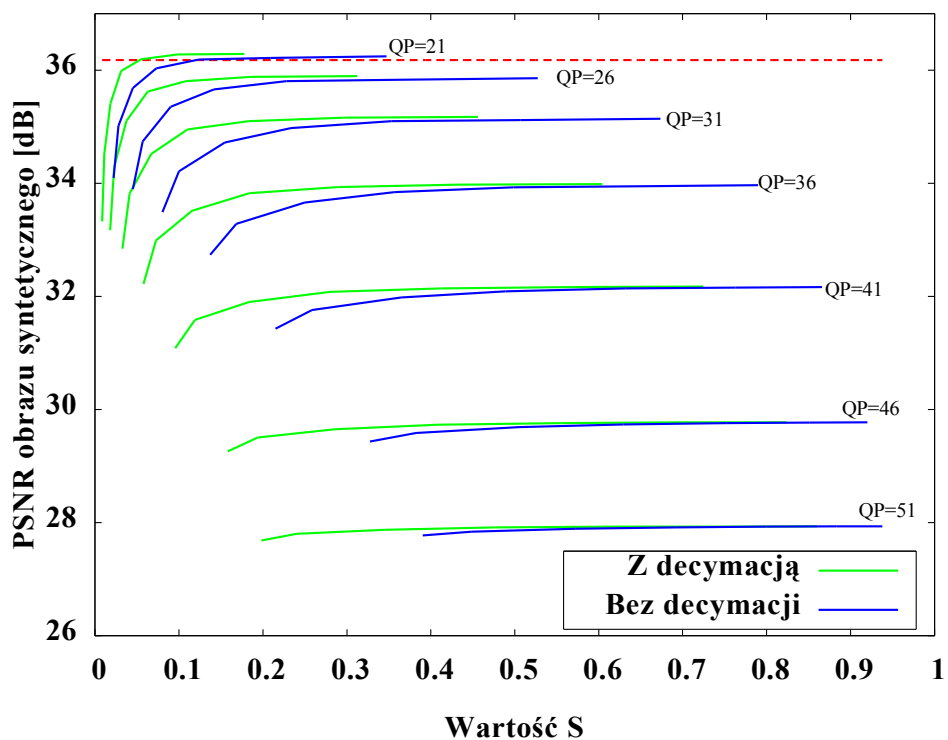
Rys. 5.14. Zależność jakości widoku wirtualnego od prędkości bitowej dla sekwencji Book Arrival ze zdecydowanymi do połowy wymiarów mapami głębi.



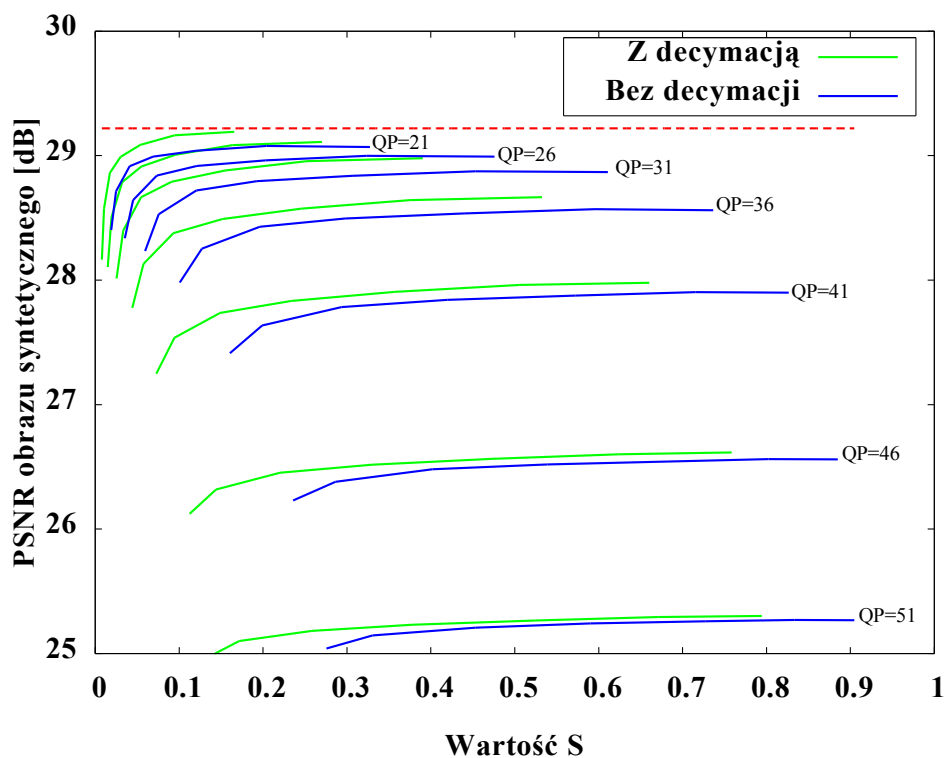
Rys. 5.15. Zależność jakości widoku wirtualnego od prędkości bitowej dla sekwencji Newspaper ze zdecydowanymi do połowy wymiarów mapami głębi.



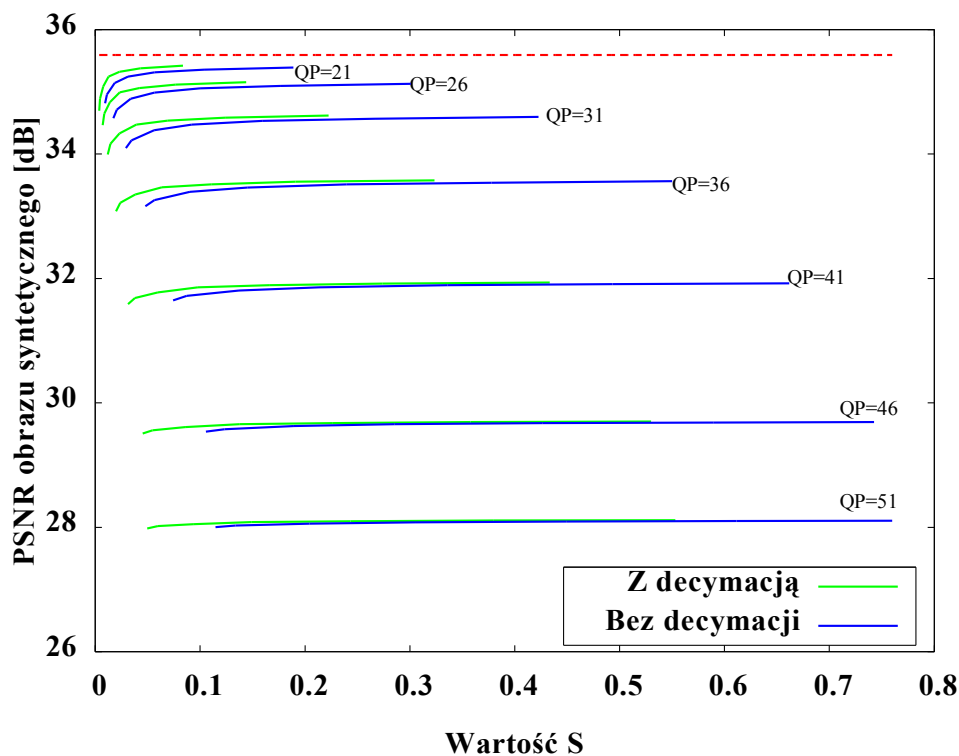
Rys. 5.16. Zależność jakości widoku wirtualnego od prędkości bitowej dla sekwencji Pantomime ze zdecydowanymi do połowy wymiarów mapami głębi.



Rysunek 5.17. Zależność jakości widoku wirtualnego od stosunku prędkości bitowej danych o mapie głębi do całkowitej prędkości bitowej dla sekwencji Book Arrival ze zdecydowanymi do połowy wymiarów mapami głębi.



Rysunek 5.18. Zależność jakości widoku wirtualnego od stosunku prędkości bitowej danych o mapie głębi do całkowitej prędkości bitowej dla sekwencji Newspaper ze zdecymowanymi do połowy wymiarów mapami głębi.



Rysunek 5.19. Zależność jakości widoku wirtualnego od stosunku prędkości bitowej danych o mapie głębi do całkowitej prędkości bitowej dla sekwencji Pantomime ze zdecymowanymi do połowy wymiarów mapami głębi.

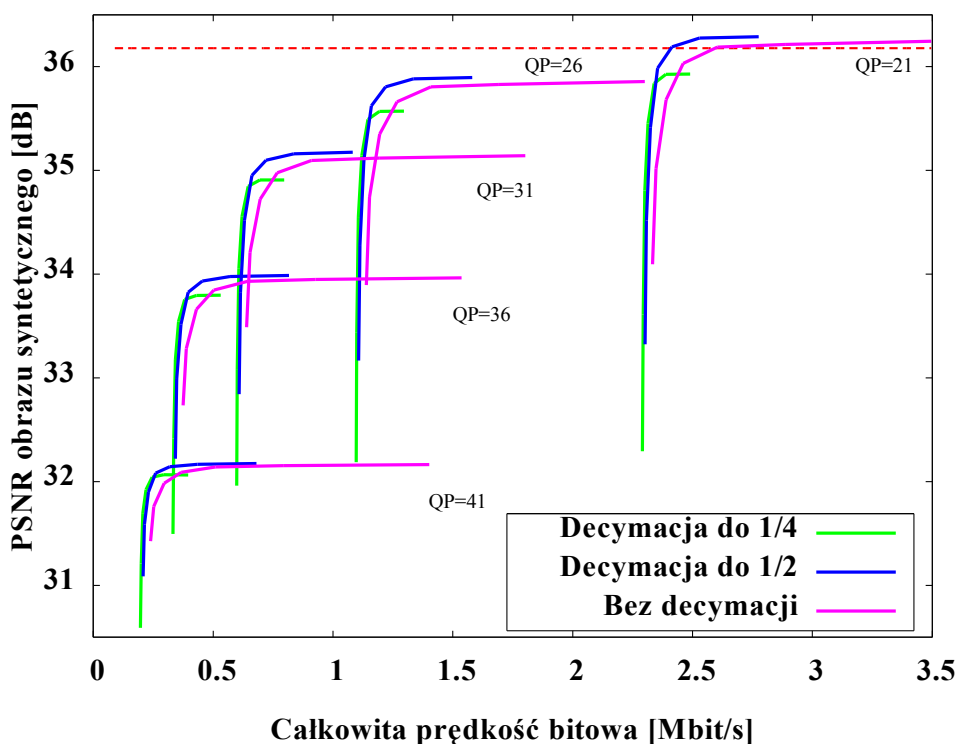
Przedstawione wyniki pokazują, że zastosowanie decymacji map głębi do połowy wymiarów (zarówno wysokości jak i szerokości) pozwala zwiększyć efektywność kompresji przy zachowaniu tej samej jakości generowanego widoku wirtualnego. Dla tej samej prędkości bitowej, jakość uzyskana przy wykorzystaniu zdecydowanych danych jest wyższa nawet o 0,2 dB niż dla przypadku z mapą głębi w pełnej rozdzielczości. Efekt ten jest tym wyraźniej widoczny, im mniejsze prędkości bitowe są rozpatrywane. Dodatkowo w przypadku zastosowania decymacji mapy głębi osiągane są wyższe maksymalne wartości miary jakości dla tych samych wartości QP w porównaniu z przypadkiem, gdy obraz mapy głębi nie jest poddawany decymacji. Warto zaznaczyć, że maksimum jakości w przypadku wykorzystania zdecydowanych map głębi występuje dla niższych prędkości bitowych niż maksimum dla przypadku z mapami głębi nie poddanymi decymacji. Należy jednak zaznaczyć, że efekt ten jest obserwowany głównie dla map głębi wyznaczonych automatycznie i jest znacznie mniej wyraźny w przypadku map głębi wysokiej jakości wyznaczonych w sposób półautomatyczny z użyciem wprowadzanych przez operatora danych pomocniczych (patrz rysunek 5.22 - linia [niebieska](#)).

Wyjaśnienie zjawiska zwiększenia jakości dla zdecydowanych map głębi jest podobne jak dla efektu zwiększenia jakości w wyniku przeprowadzenia kompresji stratnej. Decymacja i następnie przywrócenie rozdzielczości pierwotnej obrazu mapy głębi poprzez powielenie wartości punktu obrazu (wykorzystany algorytm decymacji i rekonstrukcji przedstawiony jest w Aneksie II) prowadzi do wygładzenia, ujednolicenia wartości jasności w obrazie mapy głębi na gładkich powierzchniach. Dodatkowo, proces ten nie wiąże się ze szkodliwym efektem rozmycia obrazu mapy głębi na krawędziach obiektów, a jedynie z niewielkim (jednopunktowym) przesunięciem tej krawędzi.

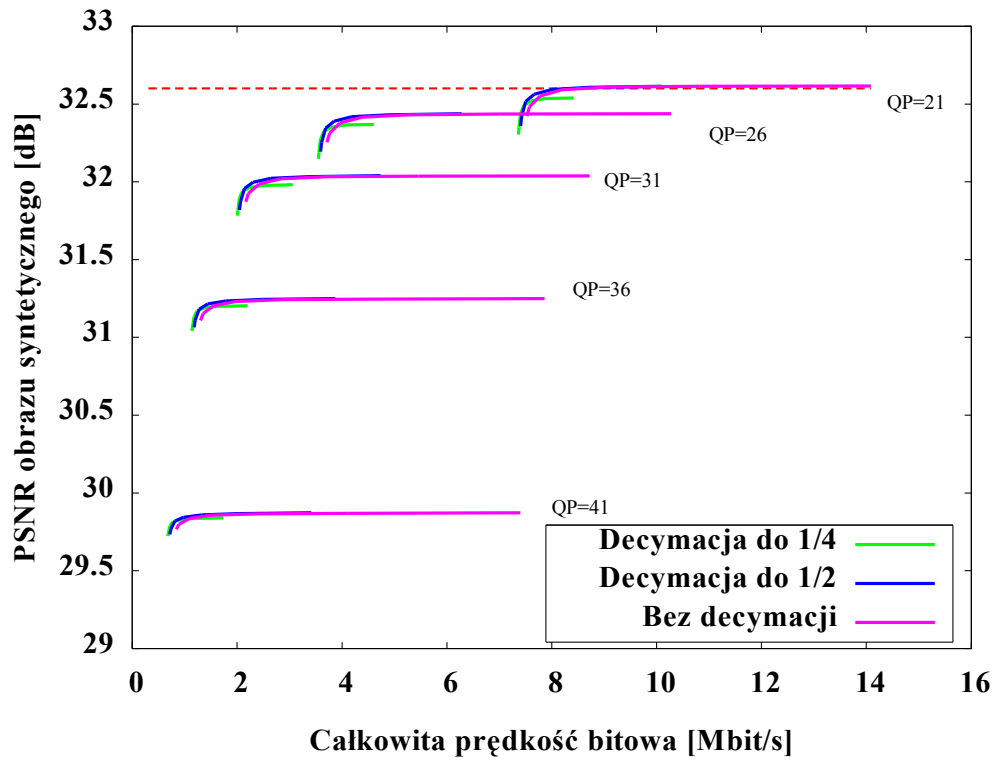
Przebadany został również przypadek ze zwiększonym stopniem decymacji. Mapy głębi zmniejszane były do jednej czwartej pierwotnych wymiarów. Przykładowe wyniki badań z wykorzystaniem mapy głębi zdecydowanej do jednej czwartej wysokości i szerokości przedstawione są na wykresach od 5.20 do 5.25. [Przerywana czerwona](#) linia odpowiada wartości miary jakości dla widoku wirtualnego wyliczonego na podstawie nieskompresowanych danych wejściowych, linie [zielone](#) odpowiadają jakości dla widoku wirtualnego stworzonego z użyciem map głębi zdecydowanych do  $\frac{1}{4}$  szerokości i wysokości. Dla porównania, linia [różowa](#) i [niebieska](#) przedstawiają wyniki, odpowiednio, dla mapy głębi niezdecymowanej i mapy głębi zdecydowanej do połowy rozmiarów.

Wyniki uzyskane w przypadku wykorzystania mapy głębi zdecydowanej do jednej czwartej rozmiarów są nieco inne niż dla decymacji do połowy rozmiarów. Nie obserwuje się zwiększonych wartości miary jakości w całym zakresie prędkości bitowych (jak na rysunkach od 5.14 do 5.19),

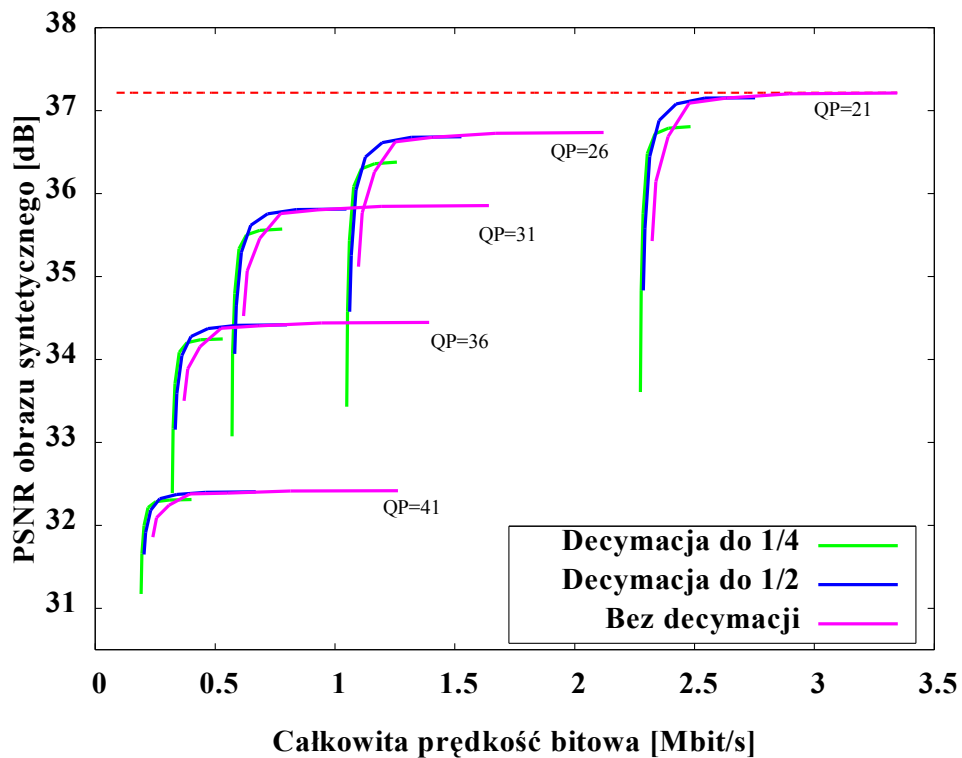
lecz jedynie dla mniejszych prędkości bitowych. Dla większych prędkości bitowych mapy głębi zdecydowane do  $\frac{1}{4}$  wysokości i szerokości powodują uzyskiwanie gorszych wyników syntezy widoku wirtualnego niż dla oryginalnych map głębi. Przy zwiększaniu indeksu kwantyzacji głębi QD i utrzymaniu indeksu kwantyzacji dla obrazu QP na stałym poziomie (a zatem – przy poruszaniu się wzdłuż jednej krzywej oznaczonej na wykresach jedną wartością QP w kierunku mniejszych wartości prędkości bitowych lub mniejszego stosunku prędkości) dochodzi jednak do zmiany kolejności krzywych – wyższą jakość uzyskuje się dla obrazu mapy głębi zdecydowanego do  $\frac{1}{4}$  rozmiarów niż dla obrazu o oryginalnej rozdzielczości i niż dla obrazu zdecydowanego do  $\frac{1}{2}$  rozmiarów. Na podstawie uzyskanych wyników można założyć, że zamiana kolejności zachodzi dla kompresji z parametrem QD większym od 36. Lepszą jakość widoku syntezowanego osiąga się wtedy w przypadku wykorzystania mapy głębi z decymacją do  $\frac{1}{4}$  rozmiarów luminancji.



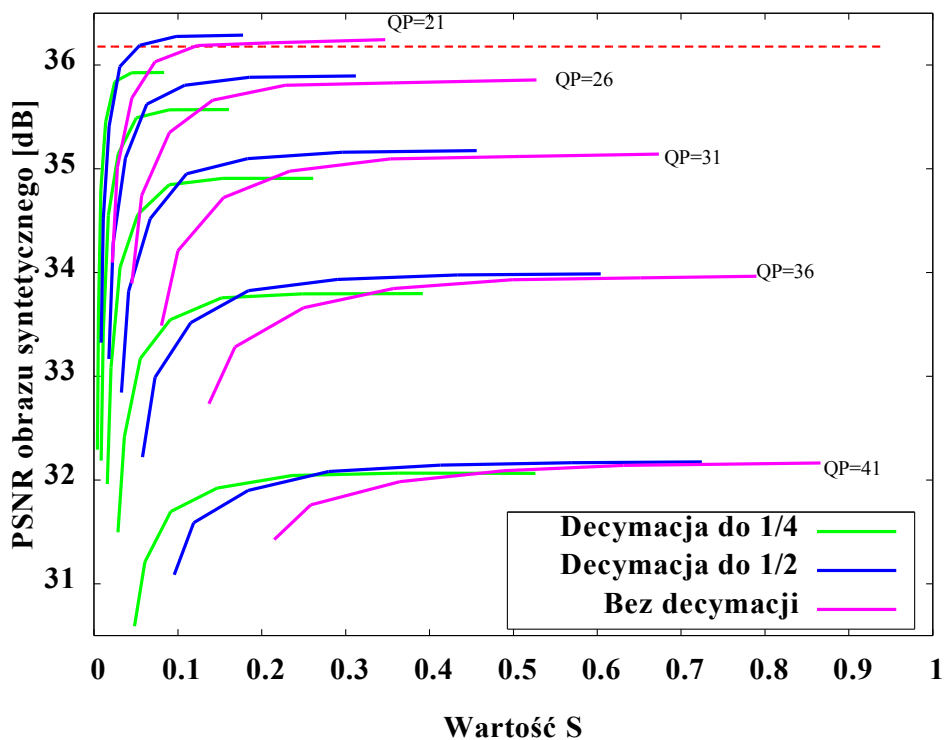
Rys. 5.20. Zależność jakości widoku wirtualnego od prędkości bitowej dla sekwencji Book Arrival z mapą głębi zdecydowaną do  $\frac{1}{4}$  rozmiarów.



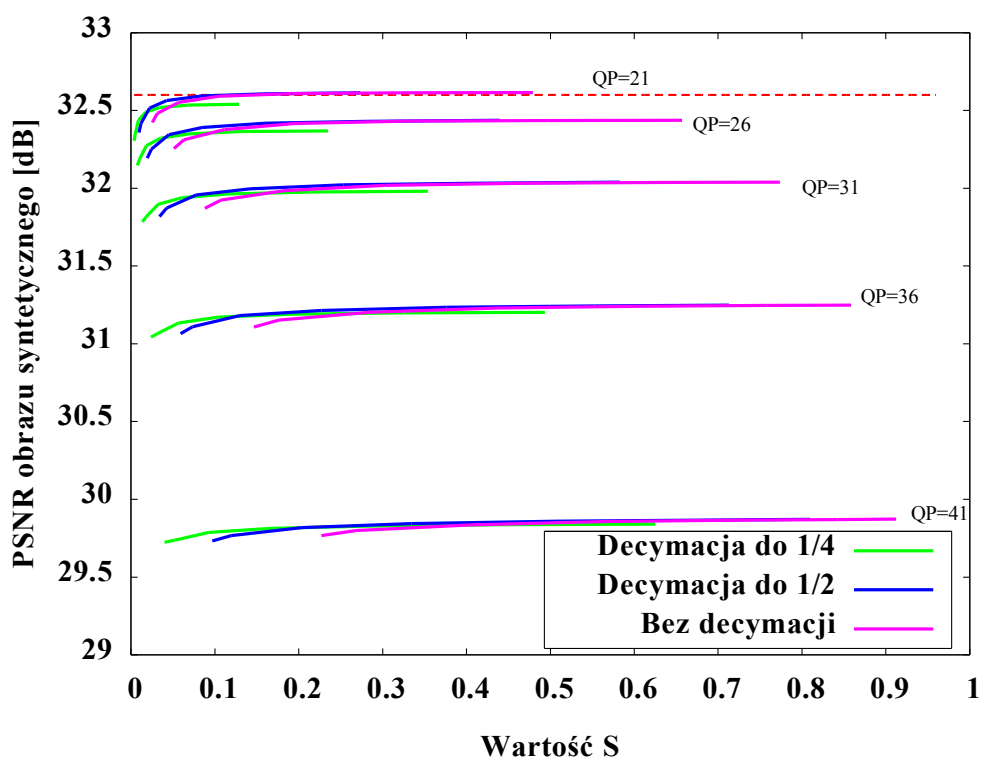
Rys. 5.21. Zależność jakości widoku wirtualnego od prędkości bitowej dla sekwencji Dog z mapą głębi zdecymowaną do  $\frac{1}{4}$  rozmiarów



Rys. 5.22. Zależność jakości widoku wirtualnego od prędkości bitowej dla sekwencji Book Arrival z mapą głębi wysokiej jakości zdecymowaną do  $\frac{1}{4}$  rozmiarów

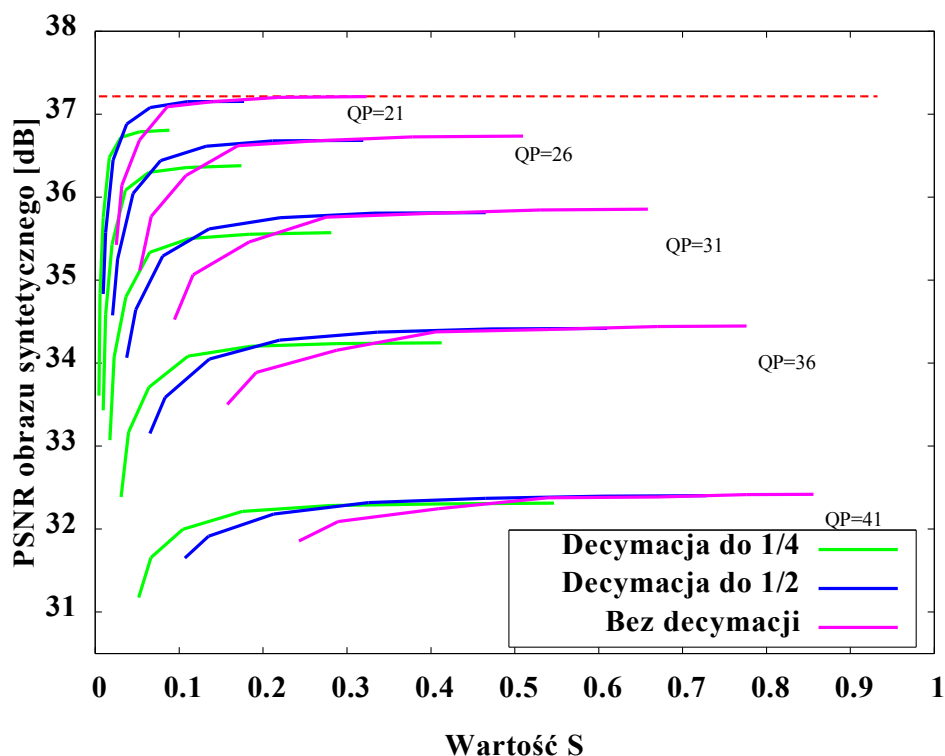


Rysunek 5.23. Zależność jakości widoku wirtualnego od stosunku prędkości bitowej danych o mapie głębi do całkowitej prędkości bitowej dla sekwencji Book Arrival z mapą głębi zdecyimowaną do ¼ rozmiarów.



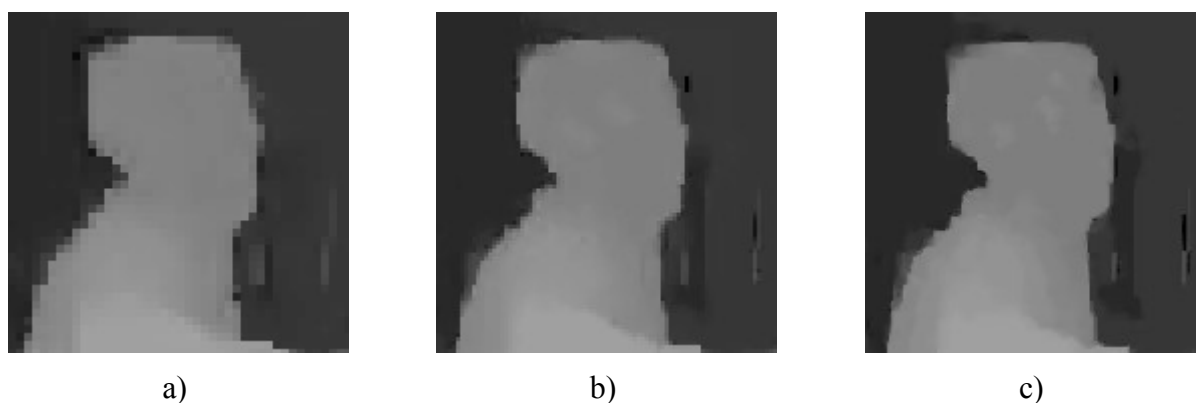
Rysunek 5.24. Zależność jakości widoku wirtualnego od stosunku prędkości bitowej danych o mapie głębi do całkowitej prędkości bitowej dla sekwencji Dog z mapą głębi zdecyimowaną do ¼ rozmiarów.



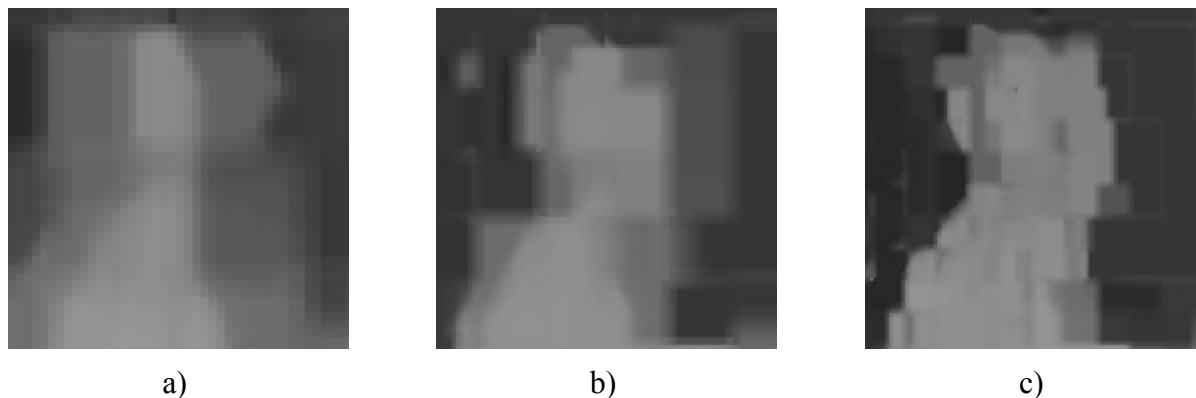


Rysunek 5.25. Zależność jakości widoku wirtualnego od stosunku prędkości bitowej danych o mapie głębi do całkowitej prędkości bitowej dla sekwencji Book Arrival z mapą głębi wysokiej jakości zdecydowaną do  $\frac{1}{4}$  rozmiarów.

Na rysunkach 5.26 oraz 5.27 pokazane są przykładowe fragmenty mapy głębi zrekonstruowanej po kompresji i po przywróceniu oryginalnej rozdzielczości dla decymacji do połowy i jednej czwartej rozmiarów przy wartości współczynnika QD = 21 oraz QD = 46.



Rysunek 5.26. Fragment mapy głębi po przywróceniu oryginalnej rozdzielczości dla trzech przypadków:  
 (a) – mapa zrekonstruowana po decymacji do  $\frac{1}{4}$  wysokości i szerokości przy kompresji z QD = 21,  
 (b) - mapa zrekonstruowana po decymacji do połowy wymiarów przy kompresji z QD = 21,  
 (c) – mapa zrekonstruowana bez decymacji przy kompresji z QD = 21.



Rysunek 5.27. Fragment mapy głębi po przywróceniu oryginalnej rozdzielczości dla trzech przypadków:

- (a) - mapa zrekonstruowana po decymacji do  $\frac{1}{4}$  wysokości i szerokości przy kompresji z  $QD = 46$ ,
- (b) - mapa zrekonstruowana po decymacji do połowy wymiarów przy kompresji z  $QD = 46$ ,
- (c) – mapa zrekonstruowana bez decymacji przy kompresji z  $QD = 46$ .

Na rysunkach 5.26 oraz 5.27 widoczny jest wpływ decymacji mapy głębi na jej postać po rekonstrukcji. W przypadku kompresji z  $QD = 21$  różnice pomiędzy mapą głębi w pełnej rozdzielczości i po decymacji do połowy wymiarów (rysunek 5.26b i 5.26c) są trudne do zauważenia. Natomiast dla bardzo silnej kompresji ( $QD = 46$ ), widoczny jest efekt związany z wygładzeniem jednolitych powierzchni w zdecydowanych obrazach jak również zniekształcenia o innym charakterze w porównaniu z obrazem skompresowanym z oryginalną rozdzielczością. Należy przypomnieć, że dla silnej kompresji ( $QD > 36$ ) lepszą jakość syntezy uzyskuje się decymując obrazy mapy głębi do jednej czwartej szerokości i wysokości przed poddaniem ich kompresji.

### 5.1.5. Podsumowanie i wnioski

Przedstawione rozważania i wyniki badań można podsumować następująco.

1. Dla pomiarów w systemach wielowidokowych wykorzystujących mapy głębi duże znaczenie i wpływ na wyniki ma wybór metody pomiaru jakości.
  - a. Pomiar jakości zrekonstruowanych map głębi poprzez obliczenie miary PSNR zdekompresowanej mapy głębi nie jest odpowiedni, gdyż nie uwzględnia specyfiki wykorzystania danych z mapy głębi, czyli używania ich jedynie w procesie syntezy widoków. Pomiar PSNR obrazu syntezy z wykorzystaniem zrekonstruowanych map głębi daje lepszy obraz jakości sekwencji wielowidokowej z głębią, gdyż uwzględnia specyfikę wykorzystania danych z mapy głębi.
  - b. Odrębną kwestią jest wybór obrazu odniesienia do obliczeń miary PSNR obrazów wirtualnych. Może nim być:

- widok syntezowany z użyciem nieskompresowanych danych oryginalnych,
- widok z rzeczywistej kamery umieszczonej w tym samym miejscu.

W tym drugim przypadku otrzymane wyniki mogą być zaburzone, gdyż na wynik wpływają cechy algorytmu syntezy. Dodatkową wadą drugiej możliwości jest fakt, że nie zawsze widok z rzeczywistej kamery jest możliwy do uzyskania. Niezaprzeczalną zaletą drugiej metody jest jednak uzyskiwanie wyników bardziej odpowiadających ocenom subiektywnym testowanego materiału [Klim\_09b].

2. Proces kompresji map głębi z wykorzystaniem kodera MVC daje w wyniku strumienie o mniejszych prędkościach bitowych niż proces kompresji obrazów z kamer przy identycznych ustawieniach kodera. Dla takich samych parametrów kodowania obrazów i map głębi większą część strumienia bitowego w transmisji sekwencji wielowidokowych z głębią zużyta będzie na transmisję danych o obrazach.
3. W systemach wielowidokowych wykorzystujących do swego działania mapy głębi (a zatem przeprowadzające na jakimś etapie syntezy widoków), jakość obrazów ma znacznie większy wpływ na jakość obrazu syntezowanego niż jakość dekodowanych map głębi (jakość map głębi w tym przypadku liczona w mierze PSNR względem oryginalnych map, podawanych na wejście kodera). Z tego względu, w większości przypadków, należy przyjąć, że priorytetem w kompresji sekwencji wielowidokowych z głębią powinno być zachowanie jak najwyższej jakości zrekonstruowanych obrazów.
4. Na podstawie otrzymanych wyników opracowany został wzór (5.2) pozwalający na określenie maksymalnych wartości parametrów QD nie powodujących znacznego obniżenia (o więcej niż 0,5 dB) jakości widoków syntetycznych generowanych z użyciem skompresowanych sekwencji.
5. Dla zastosowanych w eksperymencie sekwencji wielowidokowych z głębią wyznaczaną w pełni automatycznie, poprawę efektywności kompresji uzyskać można dla wszystkich prędkości bitowych poprzez zastosowanie decymacji map głębi do połowy wysokości i szerokości. Prowadzi to do wyraźnego zmniejszenia prędkości bitowej wymaganej do przesłania danych o głębi, dając takie same lub lepsze wyniki, mierzone jako wzrost jakości obrazu wirtualnego (rysunek 5.23). W przypadku map głębi o wysokiej jakości poprawę dzięki tej metodzie otrzymuje się tylko dla małych prędkości bitowych – poprawa obserwowana jest dla wartości  $QD > 31$  (rysunek 5.25).
6. Przy zastosowaniu redukcji rozmiarów obrazów map głębi do  $\frac{1}{4}$  wysokości i szerokości poprawę efektywności kodowania wobec przypadku wykorzystania map głębi o większych rozdzielczościach uzyskuje się jedynie dla wysokich wartości indeksu QD ( $QD = 46$  i

więcej), czyli dla silnie skompresowanych danych o głębi. Dla mniejszych wartości QD obserwowana jest strata jakości.

7. Jak wynika z badań, możliwe jest wytworzenie wielu strumieni danych sekwencji wielowidokowych z głębią o tej samej prędkości bitowej, a różniących się jakością, ze względu na różne proporcje pomiędzy prędkościami bitowymi przeznaczonymi na transmisję danych o obrazie i o mapie głębi. Pożądane jest zatem ustalenie sposobu wyznaczania proporcji skutkującej uzyskaniem maksymalnej możliwej jakości dla zadanej prędkości bitowej.

## 5.2. Wybór optymalnych par QP-QD

Opisany w punkcie 5.1 eksperyment pozwolił określić maksymalne wartości indeksu parametru kwantyzacji przy kompresji map głębi w pełnej rozdzielczości (wzór (5.2)). Odrębną kwestią pozostaje jednak wyznaczenie optymalnej, czyli dającej najlepsze wyniki, wartości indeksu QD dla konkretnych wartości QP. Jak zostało już wskazane, istnieje wiele kombinacji parametrów QP i QD dających bardzo zbliżone prędkości bitowe strumieni wynikowych, jednak różniących się dość znacznie jakością generowanych widoków wirtualnych.

Celem kolejnego etapu badań jest zatem opracowanie metody automatycznego doboru parametru QD przy zadanej wartości parametru QP. Dobór powinien odbywać się w sposób uwzględniający zależność maksymalnej (dla danej prędkości bitowej) jakości obrazów syntezowanych od prędkości bitowej. Otrzymana w ten sposób funkcja  $QD = f(QP)$  pozwoliłaby określić, jaką wartość powinien przyjmować indeks QD dla uzyskania maksymalnej jakości widoku syntezowanego przy danej prędkości bitowej. Przebadanie tego zagadnienia jest istotne dla prac nad kompresją sekwencji wielowidokowych z głębią, gdyż pozwala na automatyczne wyznaczenie odpowiednich wartości parametru QD w trakcie kodowania.

W toku badań należy zatem kolejno:

- wyznaczyć zależność jakości obrazów syntezowanych od całkowitej prędkości bitowej (dla obrazów i map głębi),
- wyznaczyć pary wartości QP-QD dające maksymalną jakość dla danej prędkości bitowej,
- wyznaczyć wzór pozwalający na automatyczne wyznaczenie wartości QD.

### 5.2.1. Metodologia badań

W celu wyznaczenia optymalnych par indeksów QP-QD przeprowadzone zostały wszechstronne badania na dużej próbie sekwencji wielowidokowych. Badania polegały na kompresji obrazów i kompresji map głębi a następnie przeprowadzeniu procesu syntezy widoku wirtualnego z wykorzystaniem zrekonstruowanych danych. Jakość uzyskanego widoku wirtualnego była mierzona za pomocą miary PSNR dla luminancji. Jako obrazy odniesienia do pomiarów wykorzystywane były obrazy wirtualne uzyskane z nieskompresowanych danych, oraz, jeśli tylko było to możliwe, dodatkowo wykonywany był pomiar przez porównanie z widokiem z rzeczywistej kamery umieszczonej w tym samym miejscu, z którego widok był syntezowany. W tabeli 5.2 podane są rozważane w trakcie eksperymentu scenariusze. Szczegółowy opis przebiegu

eksperymentu wraz z opisem danych testowych zamieszczony jest w Aneksie III. Ze względu na chęć zachowania jak największej ilości szczegółów w rekonstruowanych mapach głębi oraz ograniczenie dostępnych stopni swobody w konfiguracji eksperymentu, nie rozważano tu przypadków ze zdecydowaną mapą głębi.

Wątpliwości może budzić podana w tabeli 5.2 długość sekwencji testowej wynosząca jedynie 16 ramek. Wybór tak małej liczby ramek podyktowany był względami praktycznymi – wyraźnie krótszym niż dla sekwencji składającej się ze 100 ramek czasem obliczeń i mniejszym zapotrzebowaniem na przestrzeń dyskową w trakcie przeprowadzania eksperymentów, co jest bardzo istotne jeśli weźmie się pod uwagę liczbę przebadanych par parametrów QP-QD oraz dużą rozdzielczość sekwencji testowych (1920x1088 punktów). Uzyskane dla tak krótkich sekwencji wyniki nie odbiegają od wyników uzyskiwanych dla dłuższych fragmentów tej samej sekwencji, co zostało pokazane przez porównanie wyników dla tej samej sekwencji z użyciem 16 i następnie 100 ramek (scenariusze 5 i 6 z tabeli 5.2). Przebadany został również wpływ odległości między kamerami, z których pochodzą kompresowane sekwencje na jakość sekwencji wirtualnej. Badanie tej zależności przeprowadzone zostało przez porównanie wyników dla pary scenariuszy 4 i 5 oraz dla pary scenariuszy 7 i 8.

Tabela 5.2. Scenariusze testowe wykorzystane w eksperymencie

Numer scenariusza	Nazwa sekwencji	rozdzielczość	Liczba ramek na sekundę	Numery użytych kamer	Numer kamery syntezowanej	Liczba użytych ramek
1	Book Arrival	1024 x 768	16,67	8 i 10	9	100
2	Newspaper	1024 x 768	30	4 i 6	5	100
3	Undo Dancer	1920 x 1088	25	1 i 3	2	16
4	Poznan Hall 2	1920 x 1088	25	5 i 7	6	16
5	Poznan Hall 2	1920 x 1088	25	6 i 7	w połowie między 6 i 7	16
6	Poznan Hall 2	1920 x 1088	25	6 i 7	w połowie między 6 i 7	100
7	Poznan Street	1920 x 1088	25	3 i 5	4	16
8	Poznan Street	1920 x 1088	25	3 i 4	w połowie między 3 i 4	16

## 5.2.2. Wyniki eksperymentu

Uzyskane wyniki eksperymentu przedstawione są w kolejnych punktach.

W punkcie 5.2.2.1 przedstawione są wyniki badań nad jakością sekwencji wielowidokowych z głębią kompresowanych z użyciem parametrów QP i QD o wartości w zakresie od 10 do 51.

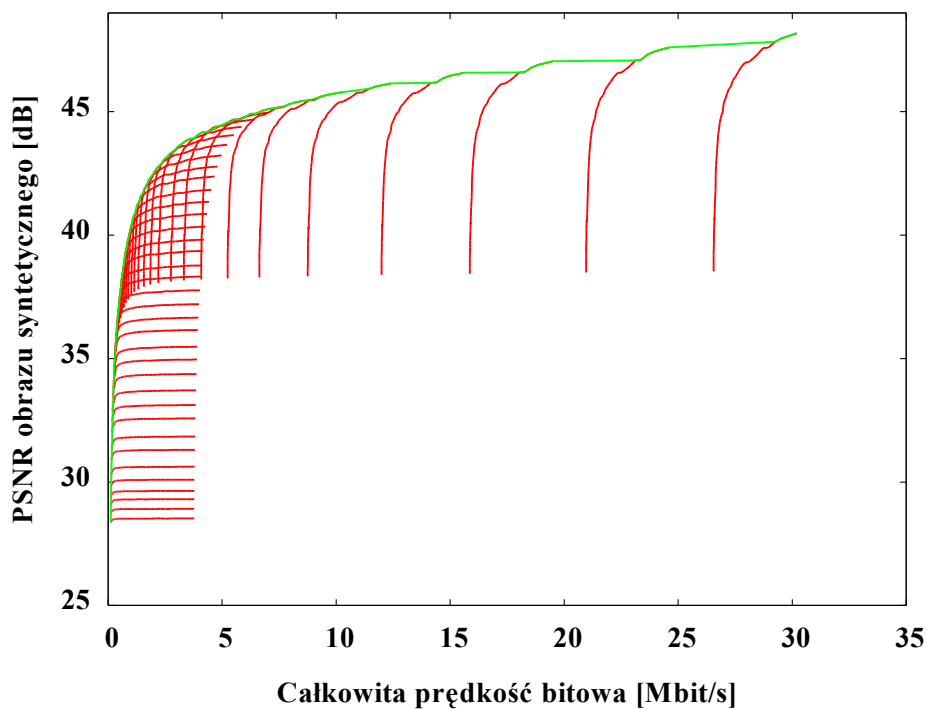
W punkcie 5.2.2.2 przedstawione są wyniki badań nad doбором par parametrów QP i QD, dla których uzyskiwana jest maksymalna jakość przy określonej prędkości bitowej. Przedstawione są analizy parametrów strumieni bitowych uzyskiwanych dla dobranych par parametrów.

W punkcie 5.2.2.3 przedstawione są wyniki prac zmierzających do określenia zależności pomiędzy wartościami parametrów QP i QD.

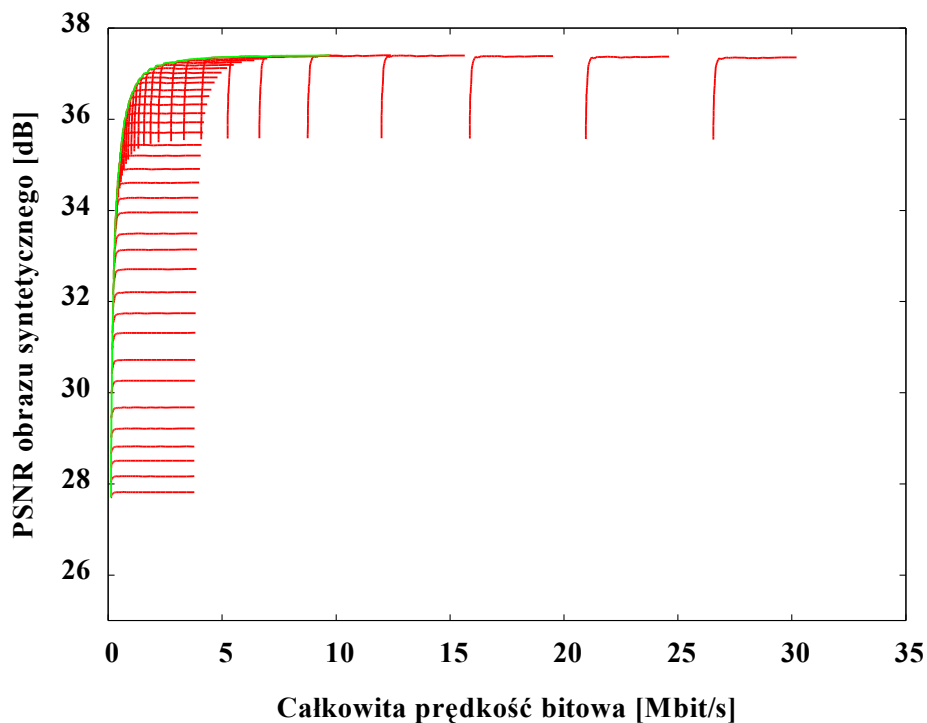
### **5.2.2.1. Efektywność kompresji w zależności od wyboru par QP-QD**

Na przedstawionych poniżej wykresach zaprezentowane są wykresy jakości widoku syntetycznego (PSNR luminancji w porównaniu z widokiem syntezy z nieskompresowanych danych i w porównaniu z widokiem z rzeczywistej kamery) w funkcji całkowitej prędkości bitowej przeznaczonej na transmisję danych dwóch widoków i dwóch map głębi. Badania przeprowadzone były dla różnych scenariuszy testowych, wymienionych w tabeli 5.2. Kolejne **czerwone linie** na wykresach odpowiadają wynikom uzyskanym dla stałej wartości indeksu QP i zmieniającej się wartości indeksu QD. Utworzona w ten sposób rodzina krzywych dla wszystkich przebadanych wartości QP umożliwia stwierdzenie, że dla tej samej prędkości bitowej możliwe jest uzyskanie różnej jakości syntezy widoku. Dodatkowo zatem na wykresie umieszczona została **zielona linia** łącząca punkty pomiarowe o maksymalnej jakości dla danej prędkości bitowej. Dla tych scenariuszy, w których było to możliwe, przedstawione są dwa wykresy – na pierwszym prezentowane są wyniki pomiaru jakości uzyskane przez porównanie uzyskanego obrazu syntetycznego do widoku syntezy z użyciem nieskompresowanych danych, a na drugim – przez porównanie do widoku z rzeczywistej kamery. Dla pozostałych scenariuszy prezentowany jest jeden wykres, na którym prezentowane są wyniki pomiaru jakości uzyskane przez porównanie do widoku syntezy z użyciem nieskompresowanych danych.

## Scenariusz 1



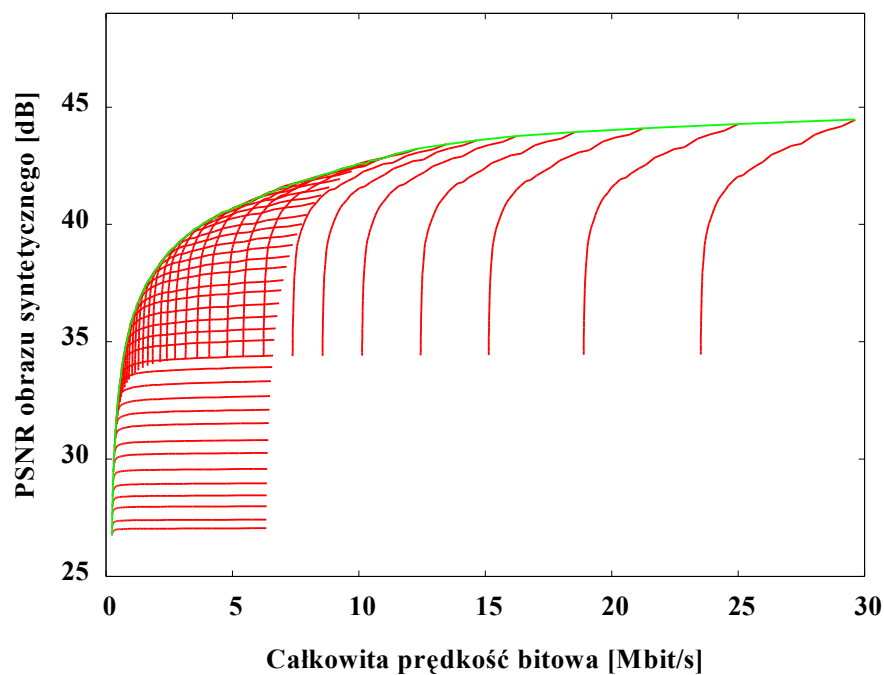
Rysunek 5.28. PSNR luminancji widoku syntezowanego w odniesieniu do widoku syntezowanego z użyciem nieskompresowanych danych dla sekwencji Book Arrival wraz z linią wyznaczającą maksymalną jakość dla danej prędkości bitowej.



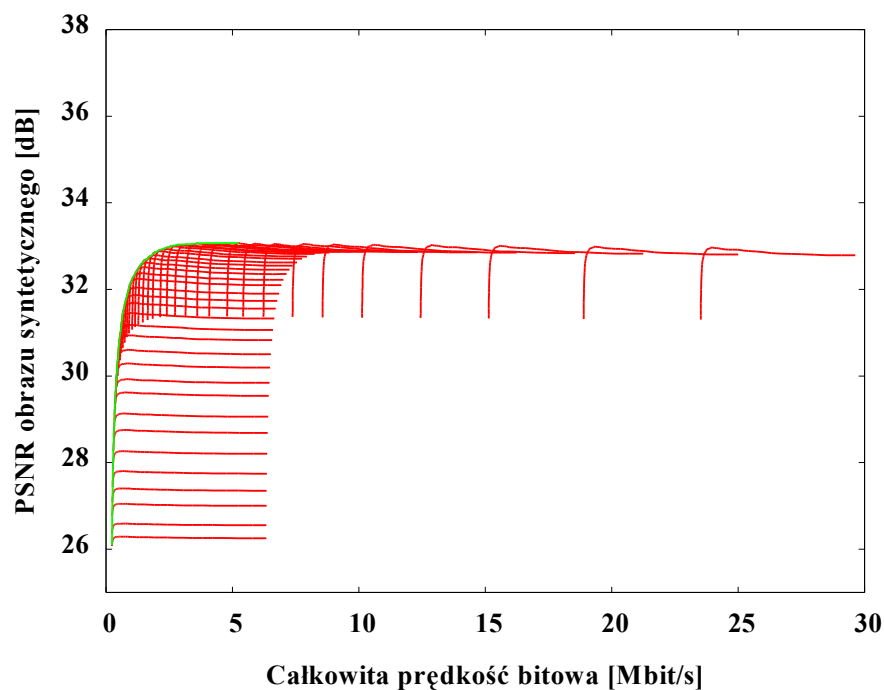
Rysunek 5.29. PSNR luminancji widoku syntezowanego w odniesieniu do rzeczywistego widoku z kamery dla sekwencji Book Arrival wraz z linią wyznaczającą maksymalną jakość dla danej prędkości bitowej.



## Scenariusz 2

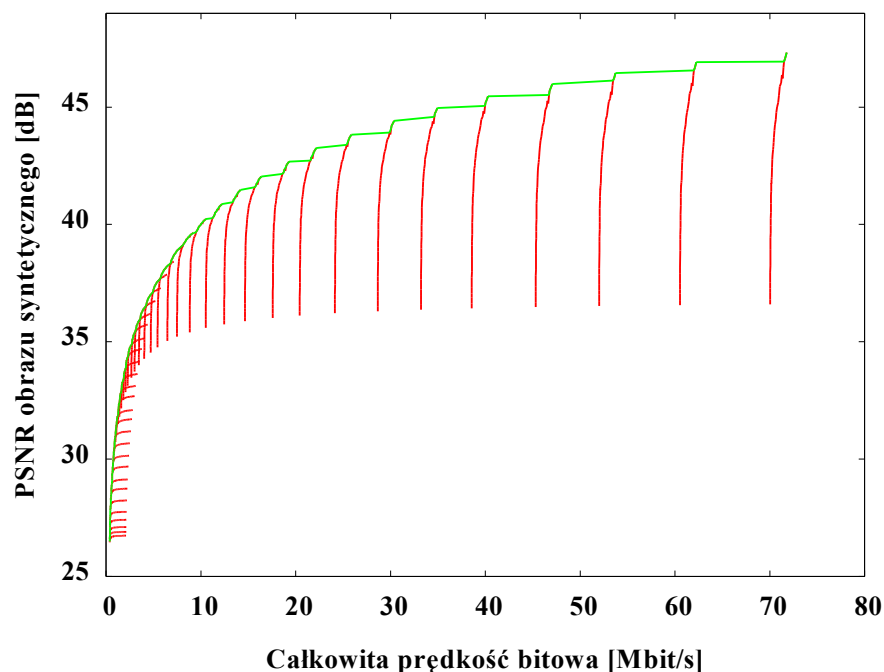


Rysunek 5.30. PSNR luminancji widoku syntezowanego w odniesieniu do widoku syntezowanego z użyciem nieskompresowanych danych dla sekwencji Newspaper wraz z linią wyznaczającą maksymalną jakość dla danej prędkości bitowej.

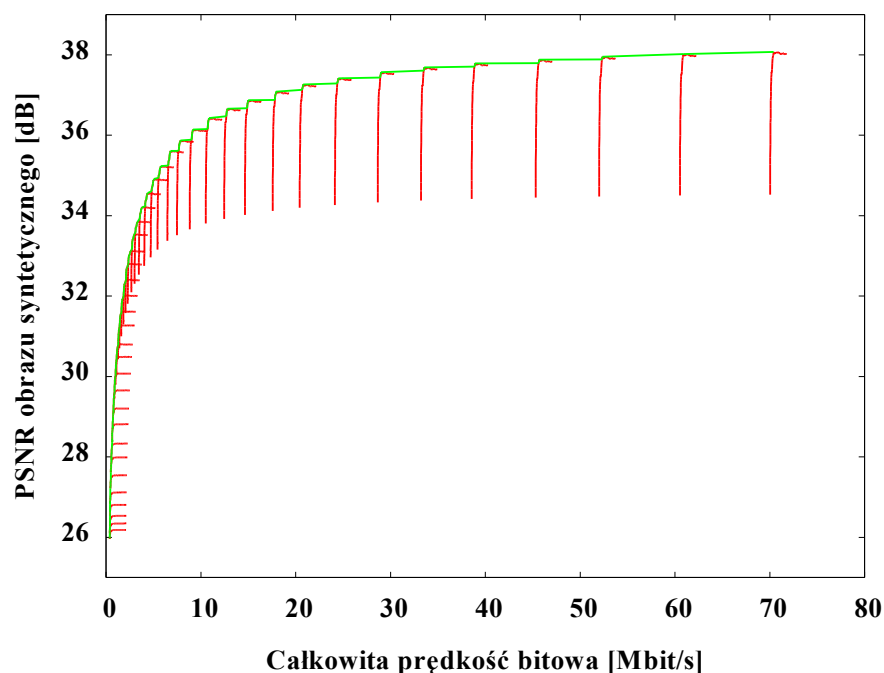


Rysunek 5.31. PSNR luminancji widoku syntezowanego w odniesieniu do rzeczywistego widoku z kamery dla sekwencji Newspaper wraz z linią wyznaczającą maksymalną jakość dla danej prędkości bitowej.

### Scenariusz 3

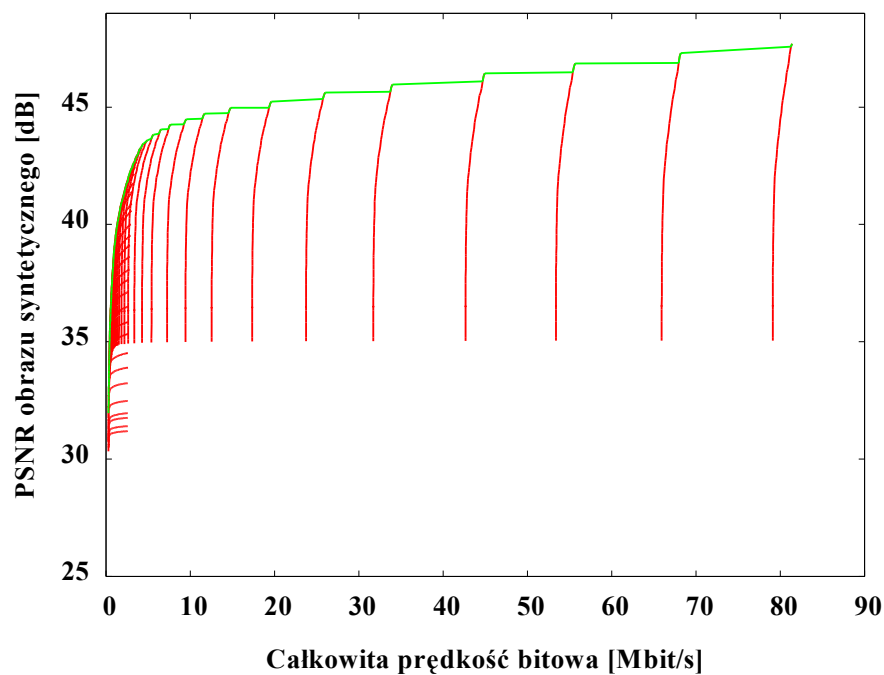


Rysunek 5.32. PSNR luminancji widoku syntezowanego w odniesieniu do widoku syntezowanego z użyciem nieskompresowanych danych dla sekwencji Undo Dancer wraz z linią wyznaczającą maksymalną jakość dla danej prędkości bitowej.

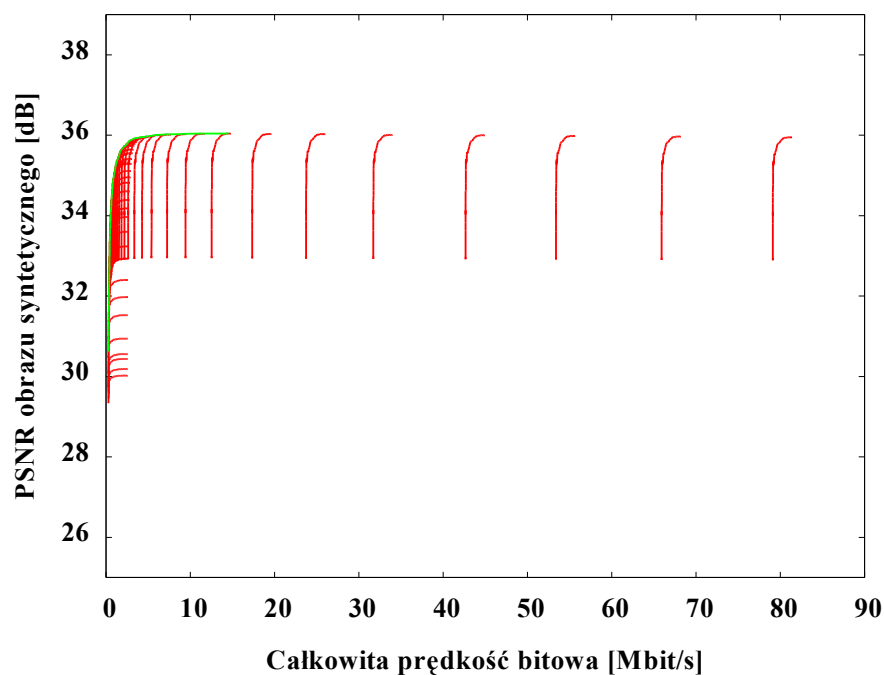


Rysunek 5.33. PSNR luminancji widoku syntezowanego w odniesieniu do rzeczywistego widoku z kamery dla sekwencji Undo Dancer wraz z linią wyznaczającą maksymalną jakość dla danej prędkości bitowej.

## Scenariusz 4

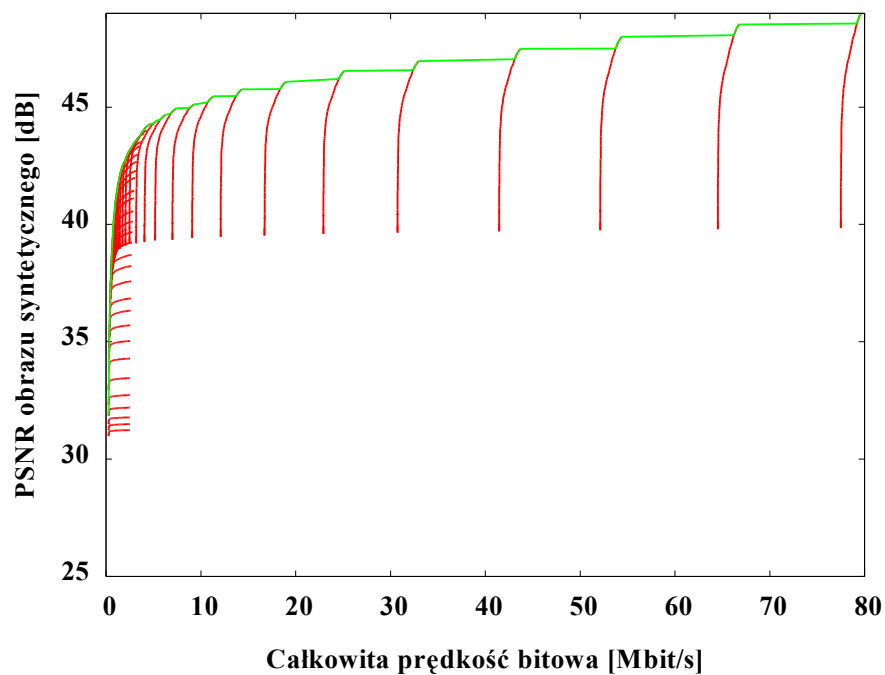


Rysunek 5.34. PSNR luminancji widoku syntezowanego w odniesieniu do widoku syntezowanego z użyciem nieskompresowanych danych dla sekwencji Poznan Hall 2 wraz z linią wyznaczającą maksymalną jakość dla danej prędkości bitowej.



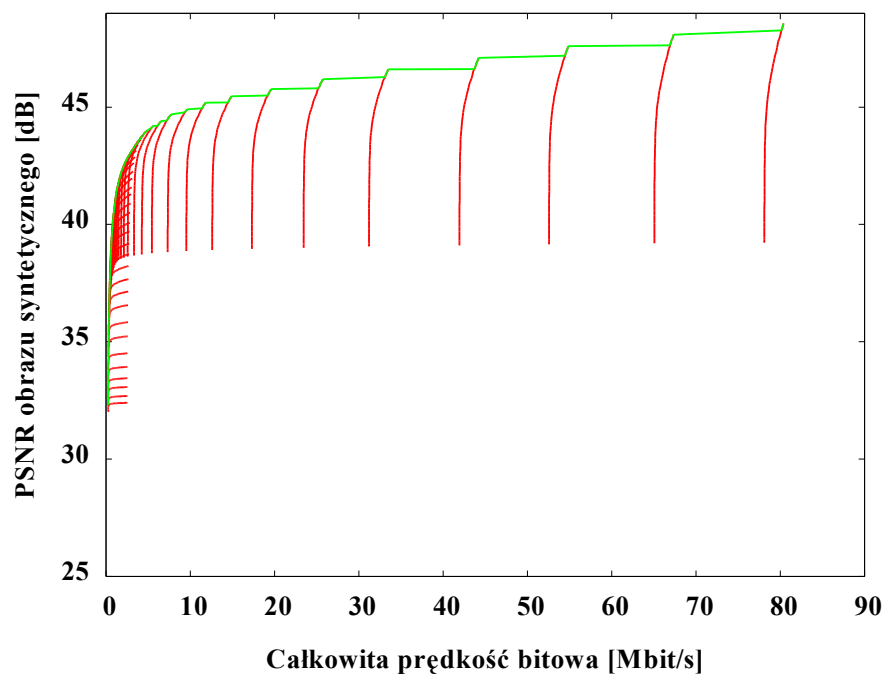
Rysunek 5.35. PSNR luminancji widoku syntezowanego w odniesieniu do rzeczywistego widoku z kamery dla sekwencji Poznan Hall 2 wraz z linią wyznaczającą maksymalną jakość dla danej prędkości bitowej. Zwiększenie prędkości bitowej powyżej 15 Mbit/s nie powoduje już zwiększenia jakości.

## Scenariusz 5



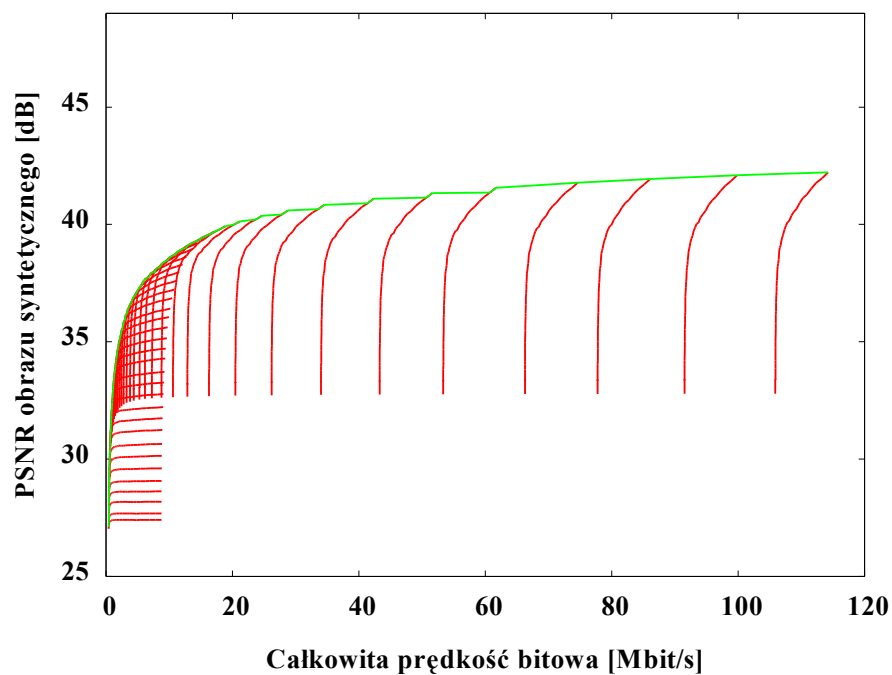
Rysunek 5.36. PSNR luminancji widoku syntezywanego w odniesieniu do widoku syntezywanego z użyciem nieskompresowanych danych dla sekwencji Poznan Hall 2 wraz z linią wyznaczającą maksymalną jakość dla danej prędkości bitowej.

## Scenariusz 6

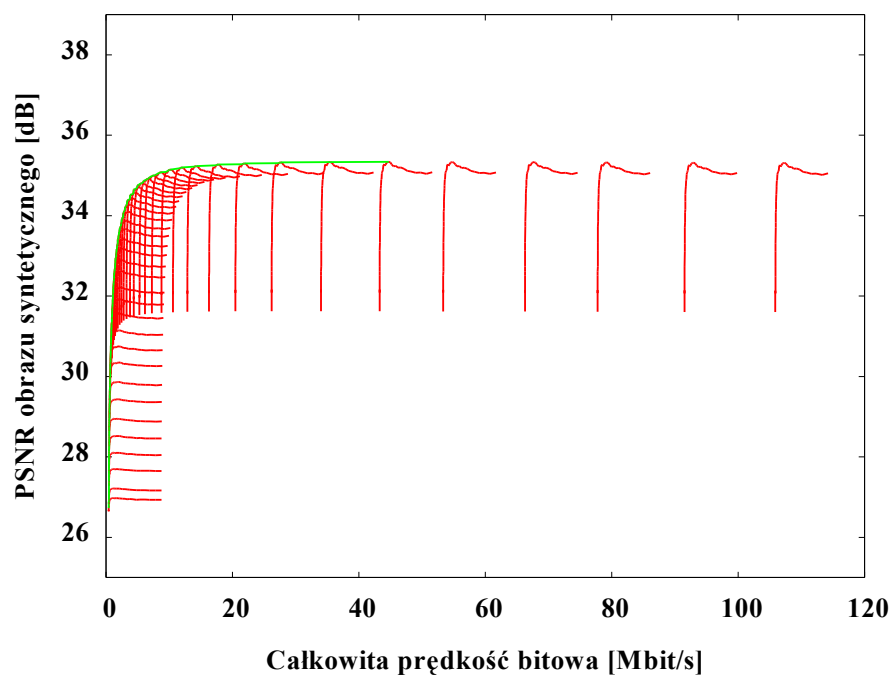


Rysunek 5.37. PSNR luminancji widoku syntezowanego w odniesieniu do widoku syntezowanego z użyciem nieskompresowanych danych dla sekwencji Poznan Hall 2 (100 ramek) wraz z linią wyznaczającą maksymalną dla danej prędkości bitowej jakość.

## Scenariusz 7

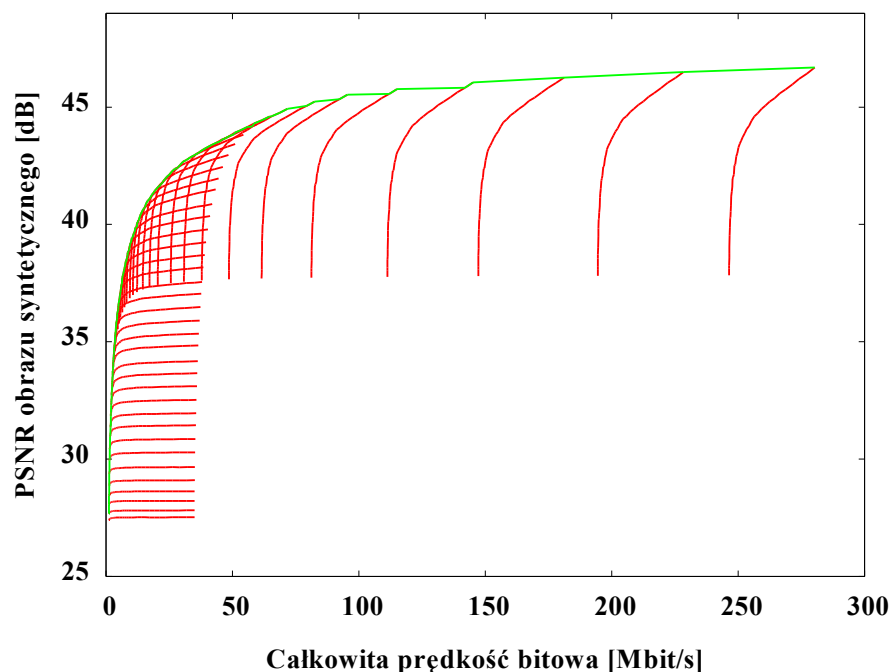


Rysunek 5.38. PSNR luminancji widoku syntezowanego w odniesieniu do widoku syntezowanego z użyciem nieskompresowanych danych dla sekwencji Poznan Street wraz z linią wyznaczającą maksymalną jakość dla danej prędkości bitowej.



Rysunek 5.39. PSNR luminancji widoku syntezowanego w odniesieniu do rzeczywistego widoku z kamery dla sekwencji Poznan Street wraz z linią wyznaczającą maksymalną jakość dla danej prędkości bitowej.

## Scenariusz 8



Rysunek 5.40. PSNR luminancji widoku syntezowanego w odniesieniu do widoku syntezywanego z użyciem nieskompresowanych danych dla sekwencji Poznan Street wraz z linią wyznaczającą maksymalną jakość dla danej prędkości bitowej.

W przedstawionych wynikach dostrzec można pewną prawidłowość, powtarzającą się dla każdego z przebadanych scenariuszy. W przypadku pomiaru jakości z wykorzystaniem jako odniesienia widoku syntezywanego z użyciem nieskompresowanych danych, jakość maksymalna widoku syntezywanego (zielona linia) rośnie monotonicznie wraz ze zwiększaniem prędkości bitowej. W przypadku pomiaru, w którym jako odniesienie użyty został widok z rzeczywistej kamery dostrzegalne jest inne zjawisko – maksymalną jakość uzyskuje się dla pewnej wartości prędkości bitowej, wyraźnie mniejszej od maksymalnej prędkości bitowej w danym scenariuszu. Na wykresach wartość ta odpowiada końcowi zielonej krzywej od strony wyższych prędkości bitowych. Zwiększanie prędkości bitowej powyżej tej charakterystycznej wartości nie prowadzi do poprawy jakości. Na przykład dla wyników pokazanych na rysunku 5.39 maksymalną jakość uzyskuje się dla przepływności około 45 Mbit/s. Zwiększanie prędkości bitowej ponad tę wartość nie powoduje już poprawy jakości syntezywanego widoku. Daje się zatem zaobserwować nasycenie wartości miary jakości w opisywanym przypadku. Ograniczeniem w tym wypadku jest zapewne proces syntezy widoku, który mimo otrzymywania na wejściu danych coraz bliższych oryginalnym nie jest w stanie wygenerować widoku wirtualnego lepszej jakości. Drugim czynnikiem ograniczającym jest niewystarczająca dokładność wyznaczania map głębi (nie dotyczy to

scenariusza 3, w którym wykorzystana jest sekwencja z dokładną mapą głębi).

W przypadku porównania z widokiem z rzeczywistej kamery daje się zaobserwować ciekawe zjawisko – dla każdego ustalonego indeksu QP (jedna z czerwonych krzywych na wykresie) istnieje optymalna wartość indeksu QD, dla której PSNR luminancji przyjmuje wartość maksymalną. Najwyraźniej widoczne jest to na powyższych wykresach dla scenariuszy 2, 3 oraz 7 – rysunki 5.31, 5.33, 5.39. Zachowanie takie powodowane jest przez opisane już w punkcie 5.1.4.2 zjawisko – kompresja wprowadza do mapy głębi zniekształcenia, które mogą sprawić, że obraz syntezowany z użyciem zrekonstruowanych danych będzie bardziej przypominał obraz z rzeczywistej kamery niż w przypadku wykorzystania oryginalnej mapy głębi. Dodatkowo, spowodowane przez zniekształcenia wprowadzane przez kompresję, wygładzenie obrazu mapy głębi poprawia tu jakość syntezowanego widoku.

Co ciekawe, globalne maksimum dla danej prędkości bitowej (czyli punkt leżący na zielonej linii) nie zawsze jest jednocześnie maksimum jakości możliwej do uzyskania dla takiego samego QP. Znaczący to, że maksima linii czerwonych nie zawsze leżą na zielonej linii optymalnej jakości na powyższych wykresach. Jest to widoczne dla małych prędkości bitowych (dobrą ilustracją są tu na przykład rysunki 5.28, 5.39).

W przypadku porównania do widoku syntezowanego z użyciem nieskompresowanych danych nie obserwuje się zjawiska nasycania wartości miary jakości – jakość zmienia się w sposób monotoniczny w całym zakresie indeksów QD. Jakikolwiek zniekształcenie wprowadzone do obrazu mapy głębi sprawia tutaj, że wynikowy obraz syntetyczny różni się od obrazu odniesienia.

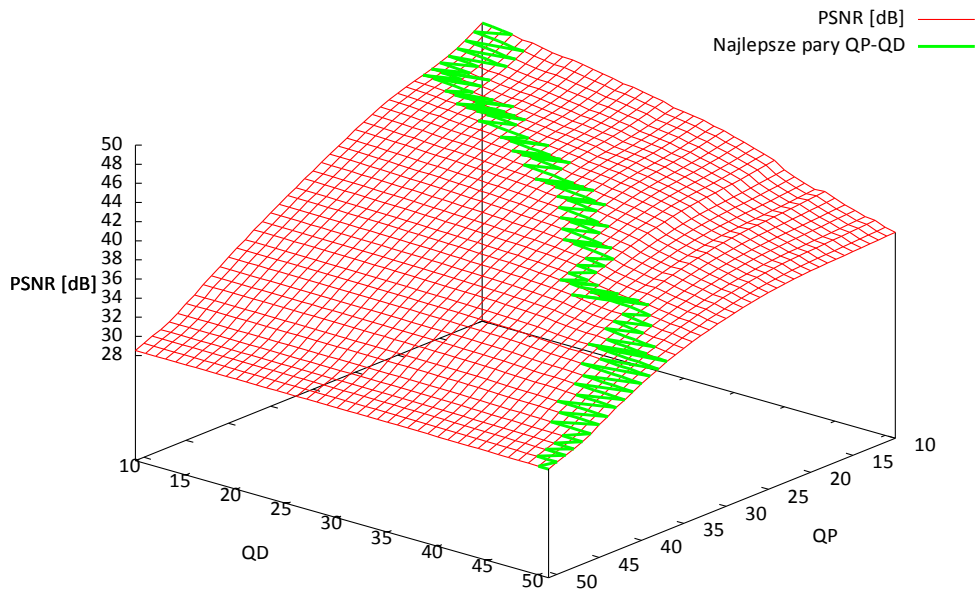
Jak już było to wspomniane, scenariusze 5 i 6 różnią się jedynie liczbą klatek użytych w eksperymencie. Te dwa scenariusze służyły sprawdzeniu, czy możliwe jest wykorzystanie krótkich sekwencji (o długości pojedynczej grupy obrazów – GOP) do określenia optymalnych par indeksów QP-QD. Wykorzystanie krótkich sekwencji jest uzasadnione znacznym skróceniem czasu obliczeń i zmniejszeniem wymaganej przestrzeni dyskowej, jak również daje nadzieję na możliwość praktycznego wykorzystania wyników prowadzonych badań. Porównanie wyników dla scenariuszy 5 oraz 6 pozwala stwierdzić, że wyniki są bardzo podobne – krzywe zachowują takie same cechy (kształt i ułożenie) w obu przypadkach. Również wartości miary jakości są podobne (dla większości punktów różnica jakości wynosi mniej niż 0,5 dB przy różnicy prędkości bitowej o mniej niż 5%). Pozwala to zakładać, że na podstawie krótkiego fragmentu sekwencji można będzie z wystarczającą dokładnością oszacować charakterystyki kodowania dla dłuższych fragmentów sekwencji.



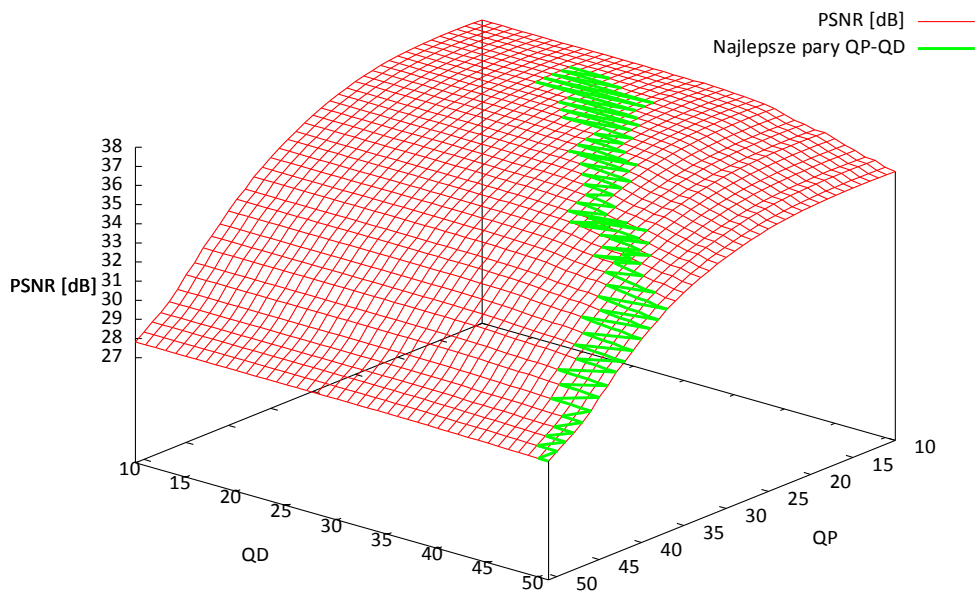
### 5.2.2.2. Optymalne pary QP-QD

Po wyznaczeniu przebiegu krzywej maksymalnej dla danej prędkości bitowej jakości (zielona linia na wykresach od 5.28 do 5.40), kolejnym krokiem jest ustalenie, jakim parom indeksów QP-QD odpowiadają kolejne punkty wyznaczające tę krzywą. Zauważyć należy, że dla każdej pary wartości parametrów QP i QD otrzymywany jest strumień o określonej prędkości bitowej. Wiele różnych par QP-QD daje strumienie o zbliżonej prędkości bitowej, lecz różniących się jakością widoków syntetycznych otrzymywanych ze zrekonstruowanych danych. Należy zatem wyznaczyć takie pary QP-QD, które dają najlepszą jakość obrazów syntezy dla danej prędkości bitowej.

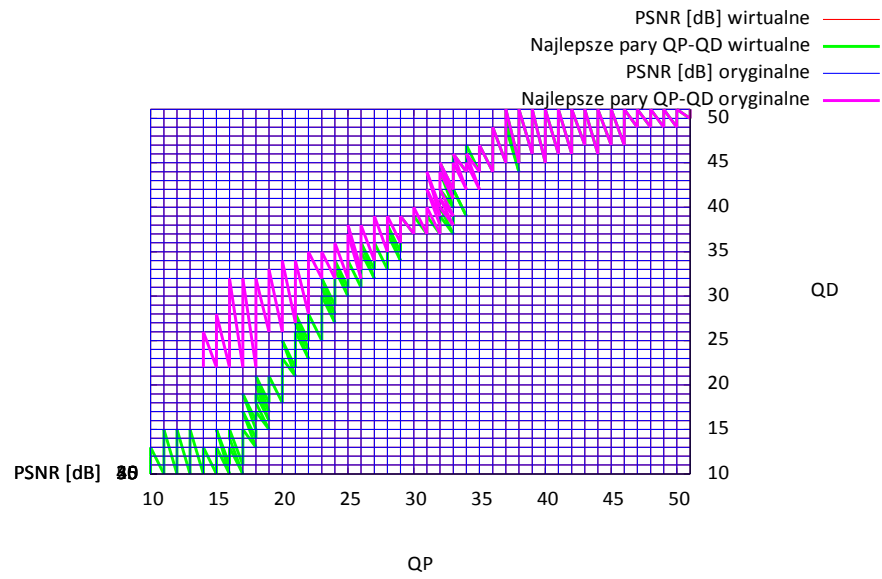
Analiza wyników eksperymentu pozwala przygotować odpowiednie dane. Na poniższych rysunkach (od 5.41 do 5.61) prezentowane są wykresy miary jakości w zależności od wartości parametrów QP i QD wraz z zaznaczoną linią łączącą pary QP-QD dające maksymalną jakość dla danej prędkości bitowej. Poruszanie się wzdłuż linii, widocznej na rysunkach poniżej, w kierunku rosnących wartości miary jakości odpowiada poruszaniu się wzdłuż zielonej krzywej w kierunku rosnącej prędkości bitowej na rysunkach od 5.28 do 5.40. Na rysunkach każdy węzeł siatki odpowiada jednemu punktowi pomiarowemu. Wyniki dla kolejnych badanych scenariuszy przedstawione są na kolejnych stronach.



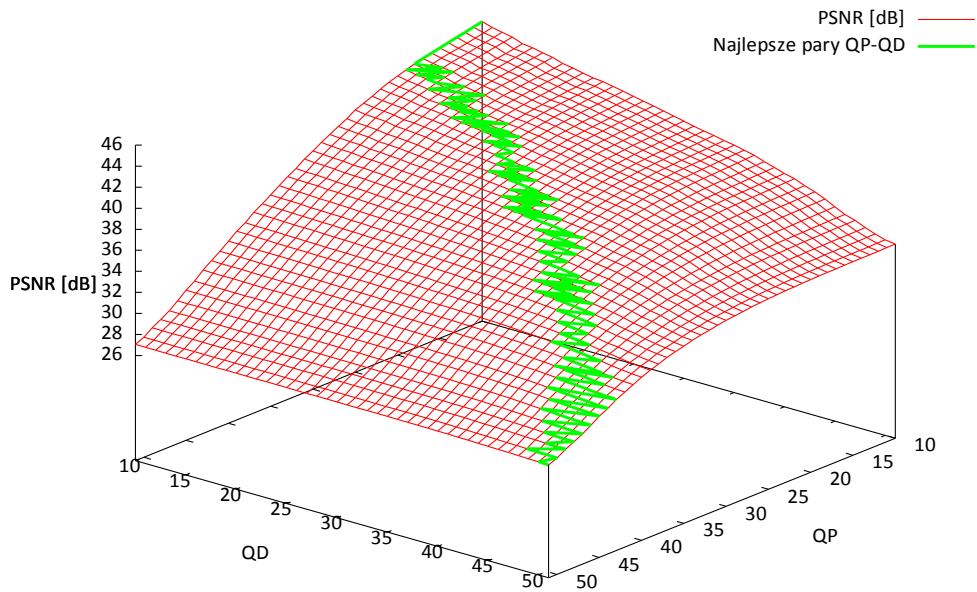
Rysunek 5.41. Zależność jakości widoku wirtualnego od wartości indeksów QP i QD dla sekwencji Book Arrival (scenariusz 1). Zaznaczona linia łącząca pary optymalne. Jakość mierzona miarą PSNR z odniesieniem do widoku wirtualnego wygenerowanego z nieskompresowanych danych.



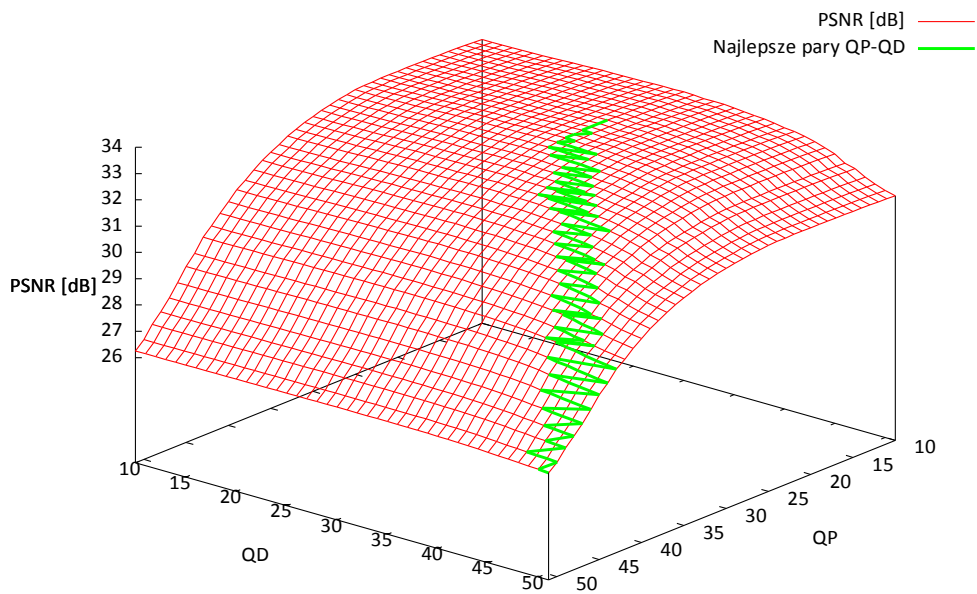
Rysunek 5.42. Zależność jakości widoku wirtualnego od wartości indeksów QP i QD dla sekwencji Book Arrival (scenariusz 1) wraz z linią łączącą optymalne pary. Jakość mierzona miarą PSNR z odniesieniem do widoku z rzeczywistej kamery.



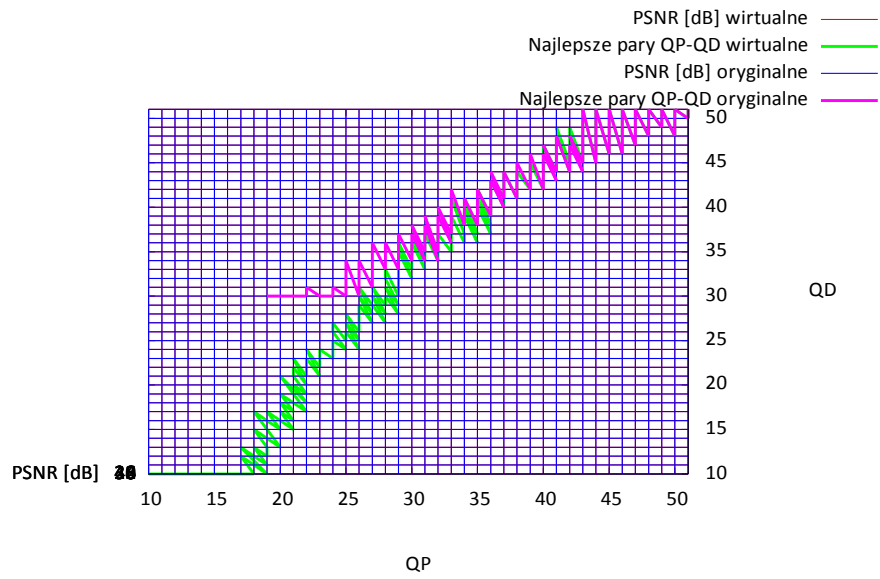
Rysunek 5.43. Linie łączące punkty pomiarowe odpowiadające optymalnym parom indeksów QP-QD dla sekwencji Book Arrival (scenariusz 1). Porównanie danych z rysunków 5.41 (wirtualne) i 5.42 (oryginalne).



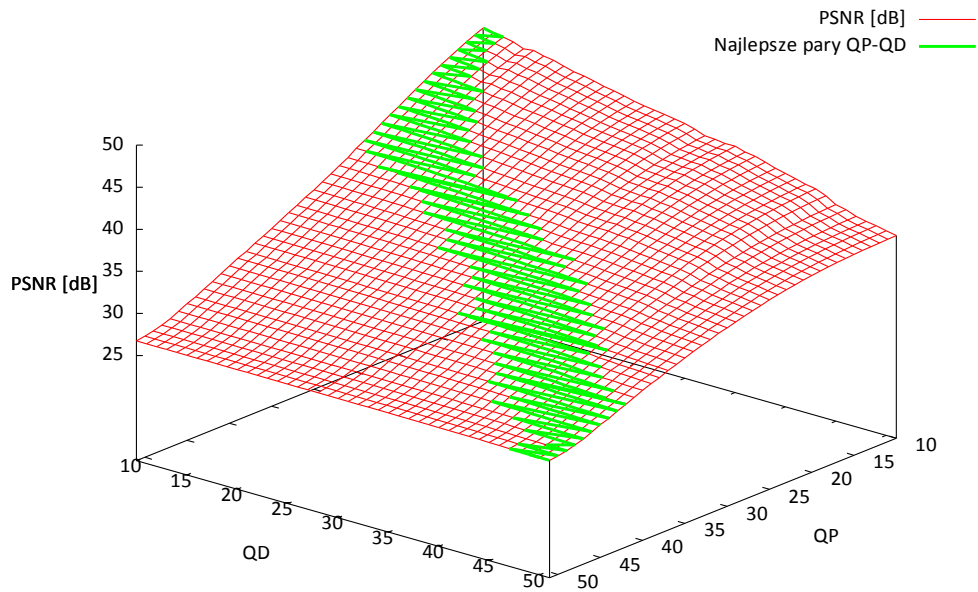
Rysunek 5.44. Zależność jakości widoku wirtualnego od wartości indeksów QP i QD dla sekwencji Newspaper (scenariusz 2). Zaznaczona linia łącząca optymalne pary. Jakość mierzona miarą PSNR z odniesieniem do widoku wirtualnego wygenerowanego z nieskompresowanych danych.



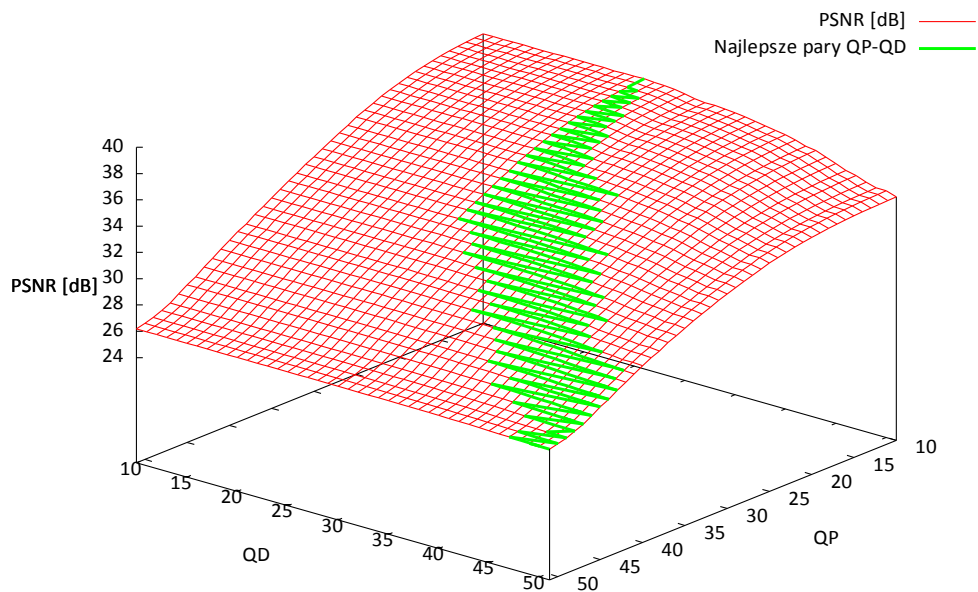
Rysunek 5.45. Zależność jakości widoku wirtualnego od wartości indeksów QP i QD dla sekwencji Newspaper (scenariusz 2). Jakość mierzona miarą PSNR z odniesieniem do widoku z rzeczywistej kamery.



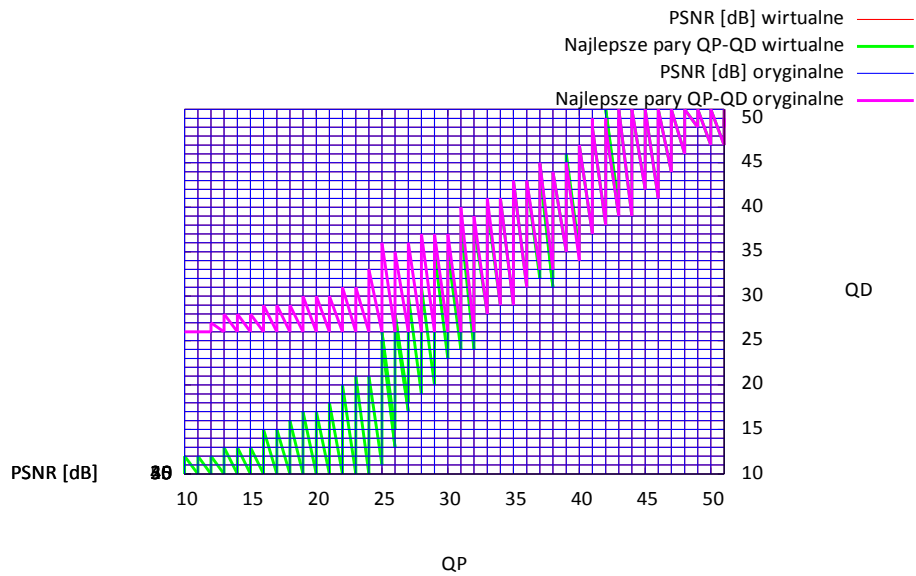
Rysunek 5.46. Linie łączące punkty pomiarowe odpowiadające optymalnym parom indeksów QP-QD dla sekwencji Newspaper (scenariusz 2). Porównanie danych z rysunków 5.44 (wirtualne) i 5.45 (oryginalne).



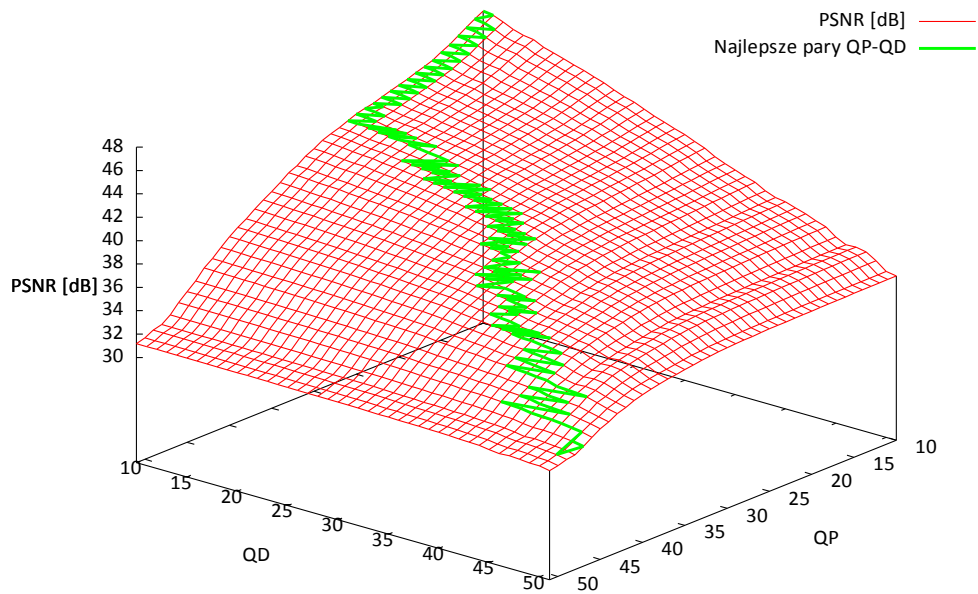
Rysunek 5.47. Zależność jakości widoku wirtualnego od wartości indeksów QP i QD dla sekwencji Undo Dancer (scenariusz 3). Zaznaczona linia łącząca optymalne pary. Jakość mierzona miarą PSNR z odniesieniem do widoku wirtualnego wygenerowanego z nieskompresowanych danych.



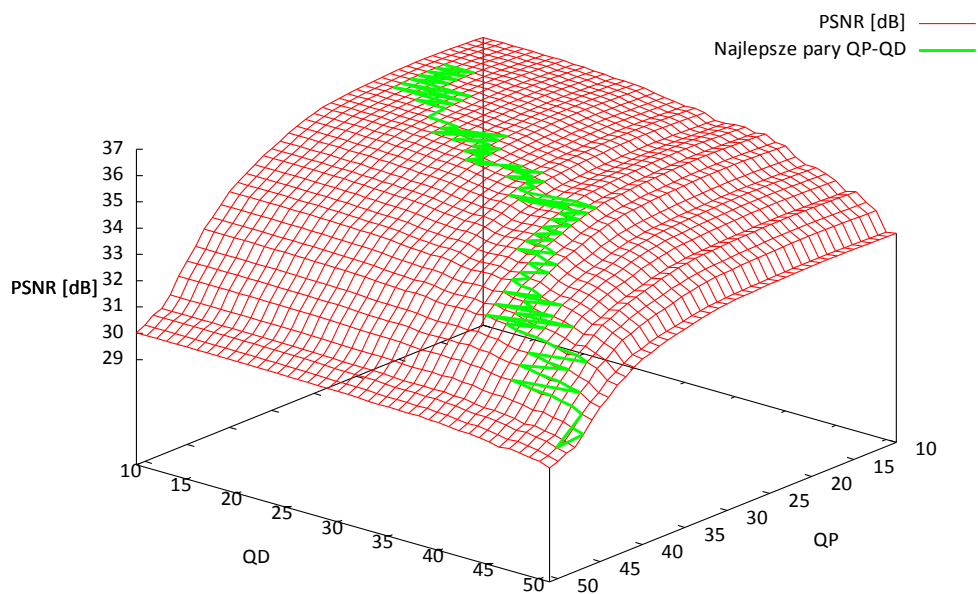
Rysunek 5.48. Zależność jakości widoku wirtualnego od wartości indeksów QP i QD dla sekwencji Undo Dancer (scenariusz 3). Zaznaczona linia łącząca optymalne pary. Jakość mierzona miarą PSNR z odniesieniem do widoku z rzeczywistej kamery.



Rysunek 5.49. Linie łączące punkty pomiarowe odpowiadające optymalnym parom indeksów QP-QD dla sekwencji Undo Dancer (scenariusz 3). Porównanie danych z rysunków 5.47 (wirtualne) i 5.48 (oryginalne).

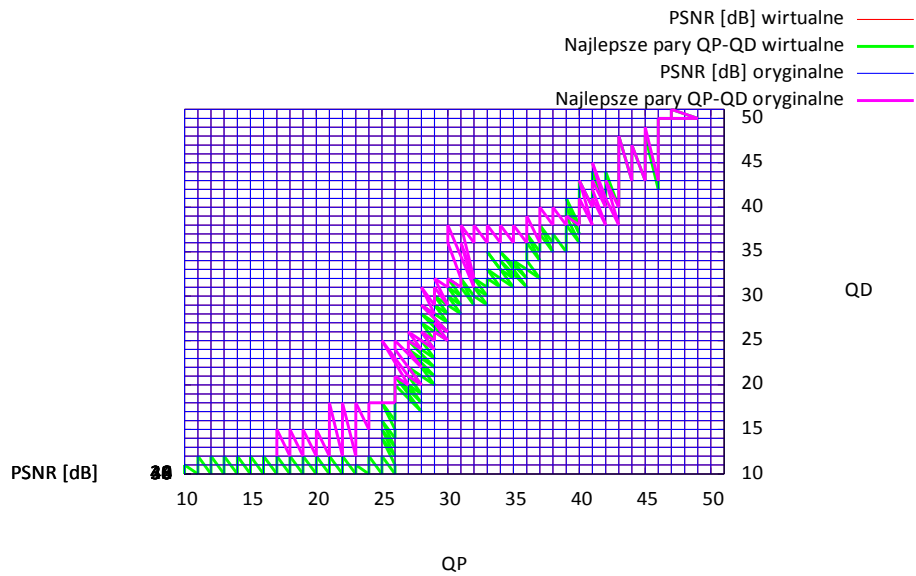


Rysunek 5.50. Zależność jakości widoku wirtualnego od wartości indeksów QP i QD dla sekwencji Poznan Hall 2 (scenariusz 4). Zaznaczona linia łącząca optymalne pary. Jakość mierzona miarą PSNR z odniesieniem do widoku wirtualnego wygenerowanego z nieskompresowanych danych.

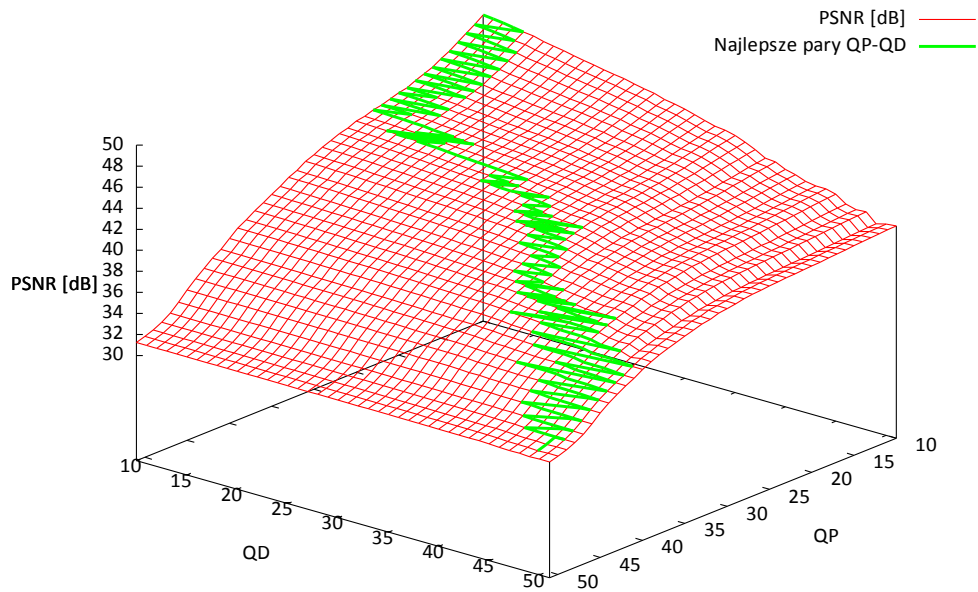


Rysunek 5.51. Zależność jakości widoku wirtualnego od wartości indeksów QP i QD dla sekwencji Poznan Hall 2 (scenariusz 4). Zaznaczona linia łącząca optymalne pary. Jakość mierzona miarą PSNR z odniesieniem do widoku z rzeczywistej kamery.

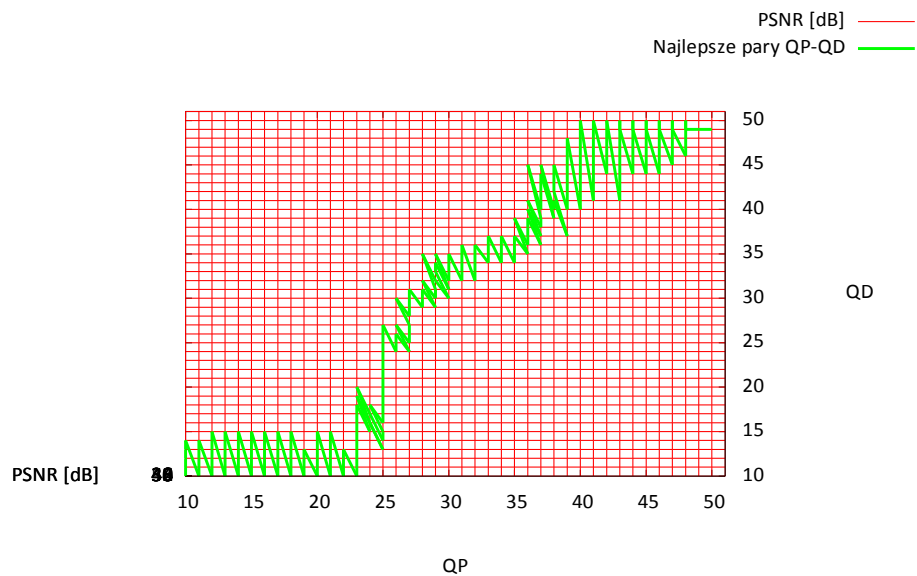




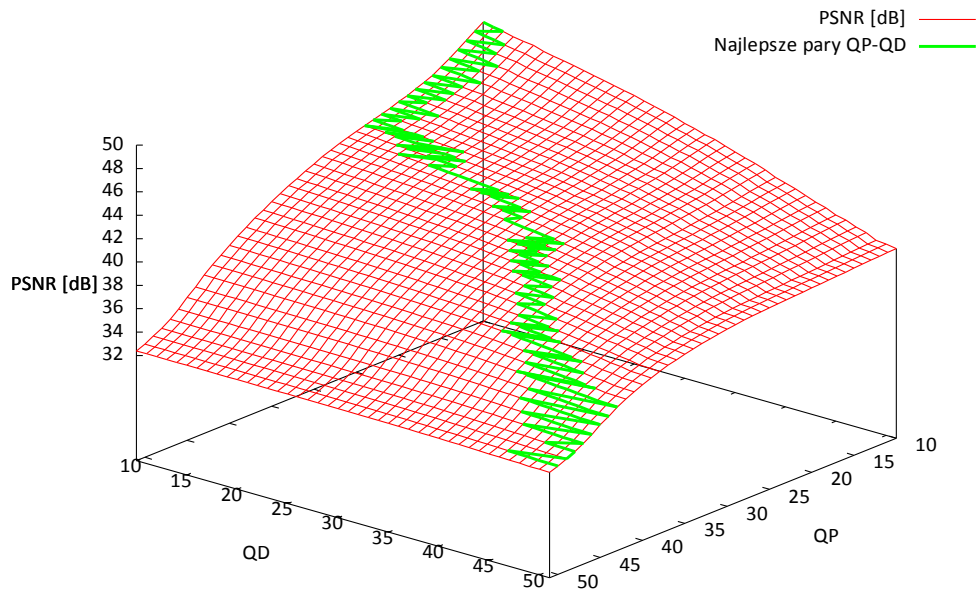
Rysunek 5.52. Linie łączące punkty pomiarowe odpowiadające optymalnym parom indeksów QP-QD dla sekwencji Poznan Hall 2 (scenariusz 4). Porównanie danych z rysunków 5.50 (wirtualne) i 5.51 (oryginalne).



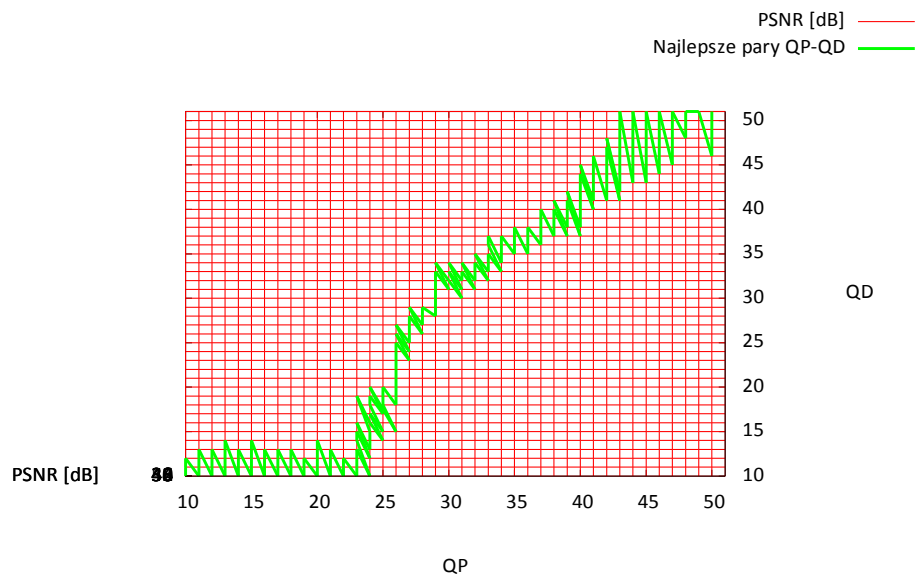
Rysunek 5.53. Zależność jakości widoku wirtualnego od wartości indeksów QP i QD dla sekwencji Poznan Hall 2 (scenariusz 5). Zaznaczona linia łącząca optymalne pary. Jakość mierzona miarą PSNR z odniesieniem do widoku wirtualnego wygenerowanego z nieskompresowanych danych.



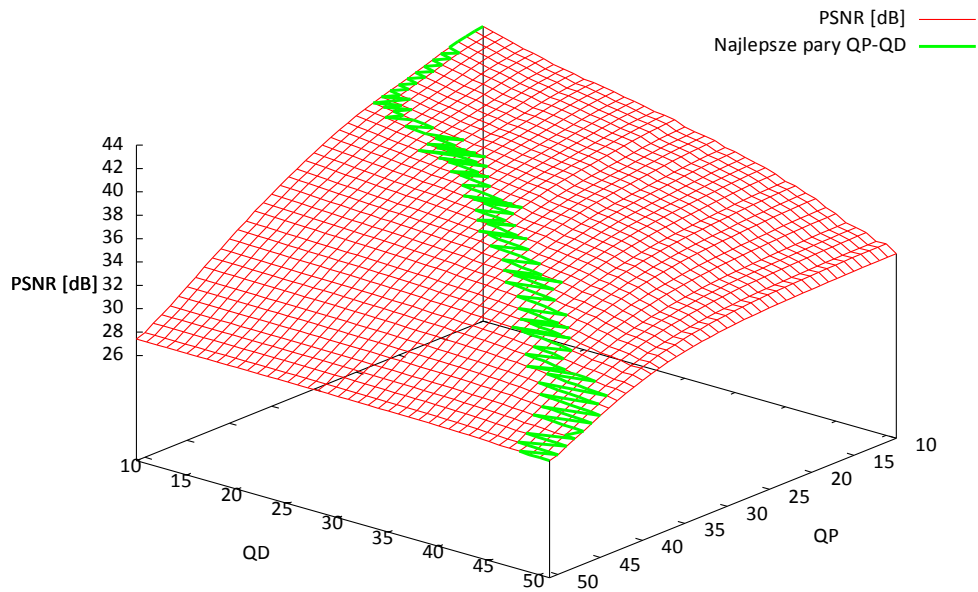
Rysunek 5.54. Linie łączące punkty pomiarowe odpowiadające optymalnym parom indeksów QP-QD dla sekwencji Poznan Hall 2 (scenariusz 5). Z wykorzystaniem miary jakości liczonej w odniesieniu do widoku wirtualnego.



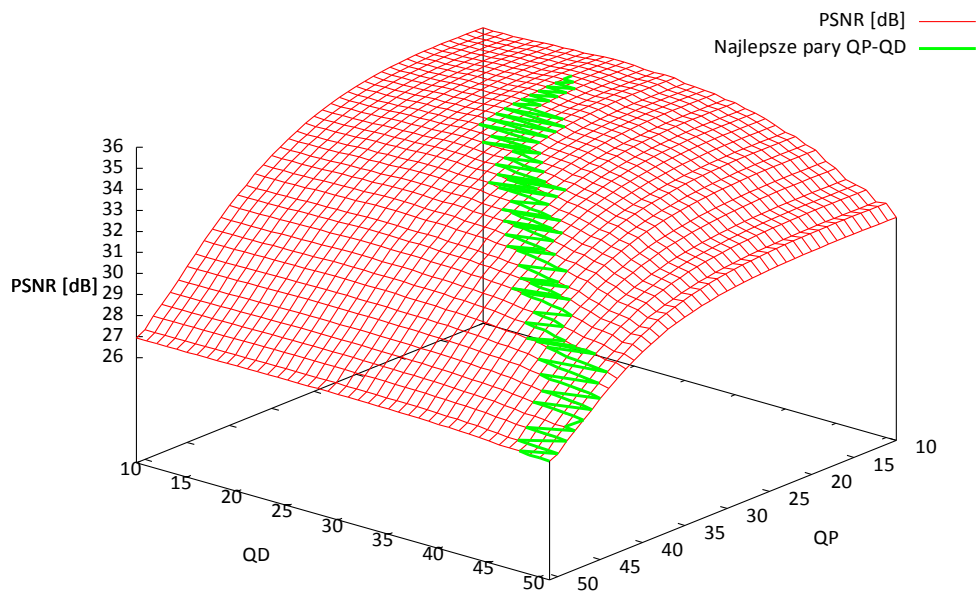
Rysunek 5.55. Zależność jakości widoku wirtualnego od wartości indeksów QP i QD dla sekwencji Poznan Hall 2 (scenariusz 6). Zaznaczona linia łącząca optymalne pary. Jakość mierzona miarą PSNR z odniesieniem do widoku wirtualnego wygenerowanego z nieskompresowanych danych.



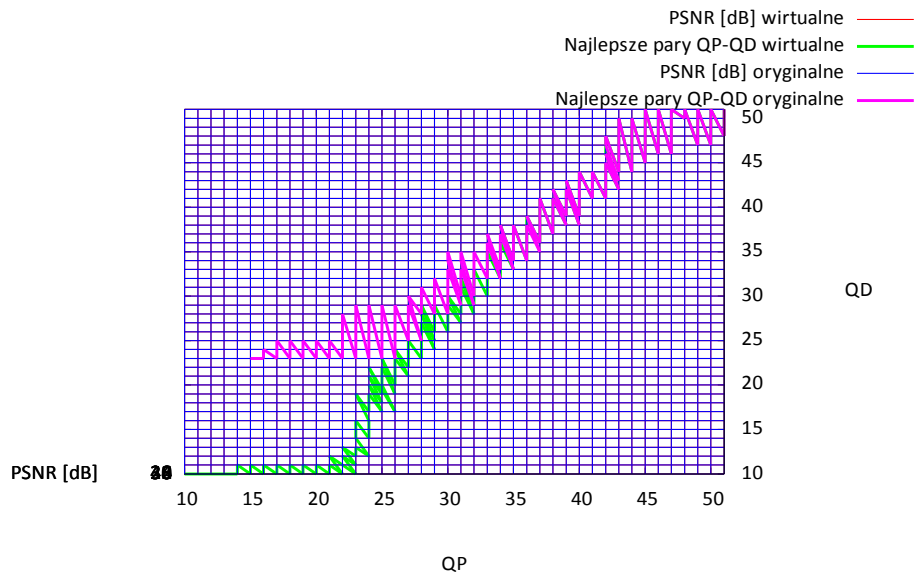
Rysunek 5.56. Linie łączące punkty pomiarowe odpowiadające optymalnym parom indeksów QP-QD dla sekwencji Poznan Hall 2 (scenariusz 6). Z wykorzystaniem miary jakości liczonej w odniesieniu do widoku wirtualnego.



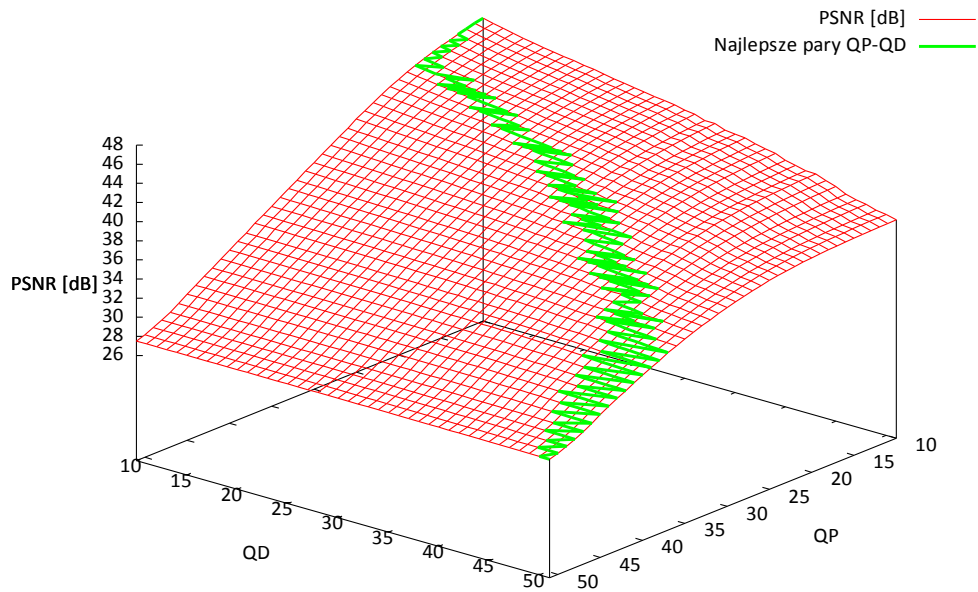
Rysunek 5.57. Zależność jakości widoku wirtualnego od wartości indeksów QP i QD dla sekwencji Poznan Street (scenariusz 7). Zaznaczona linia łącząca optymalne pary. Jakość mierzona miarą PSNR z odniesieniem do widoku wirtualnego wygenerowanego z nieskompresowanych danych.



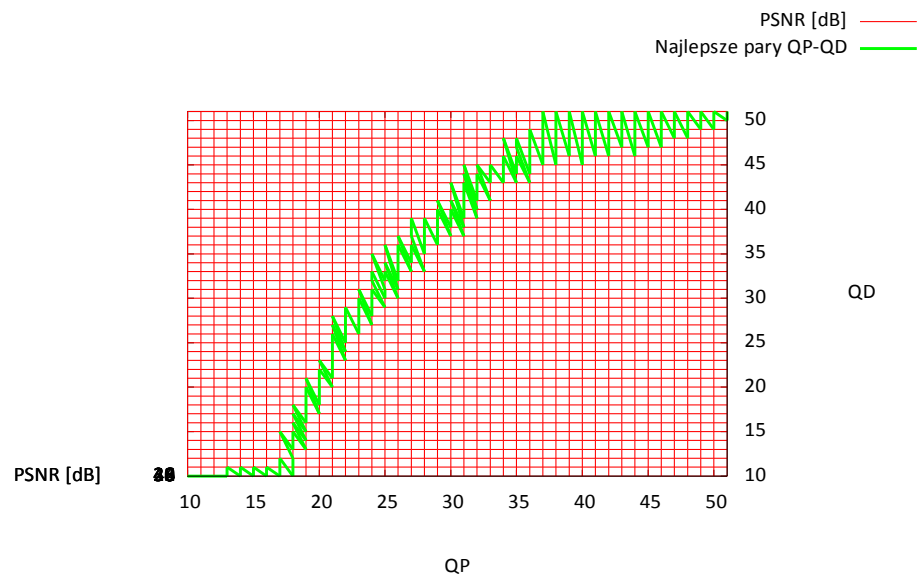
Rysunek 5.58. Zależność jakości widoku wirtualnego od wartości indeksów QP i QD dla sekwencji Poznan Street (scenariusz 7). Zaznaczona linia łącząca optymalne pary. Jakość mierzona miarą PSNR z odniesieniem do widoku z rzeczywistej kamery.



Rysunek 5.59. Linie łączące punkty pomiarowe odpowiadające optymalnym parom indeksów QP-QD dla sekwencji Poznan Street (scenariusz 7). Porównanie danych z rysunków 5.57 (wirtualne) i 5.58 (oryginalne).



Rysunek 5.60. Zależność jakości widoku wirtualnego od wartości indeksów QP i QD dla sekwencji Poznan Street (scenariusz 8). Zaznaczona linia łącząca optymalne pary. Jakość mierzona miarą PSNR z odniesieniem do widoku wirtualnego wygenerowanego z nieskompresowanych danych.



Rysunek 5.61. Linie łączące punkty pomiarowe odpowiadające optymalnym parom indeksów QP-QD dla sekwencji Poznan Street (scenariusz 8). Z wykorzystaniem miary jakości liczonej w odniesieniu do widoku wirtualnego.

Podsumowując przedstawione wyniki należy zwrócić uwagę na następujące wnioski.

Wyniki dla scenariuszy 4 i 5 są bardzo podobne. Również wyniki dla scenariuszy 7 i 8 cechują się dużym podobieństwem. W obu przypadkach większe wartości miary jakości uzyskuje się dla scenariusza, w którym kamery ustawione są bliżej siebie (kompresji poddawane są sekwencje z kamer o numerach różniących się o 1, czyli z sąsiednich kamer). Wynika z tego, że:

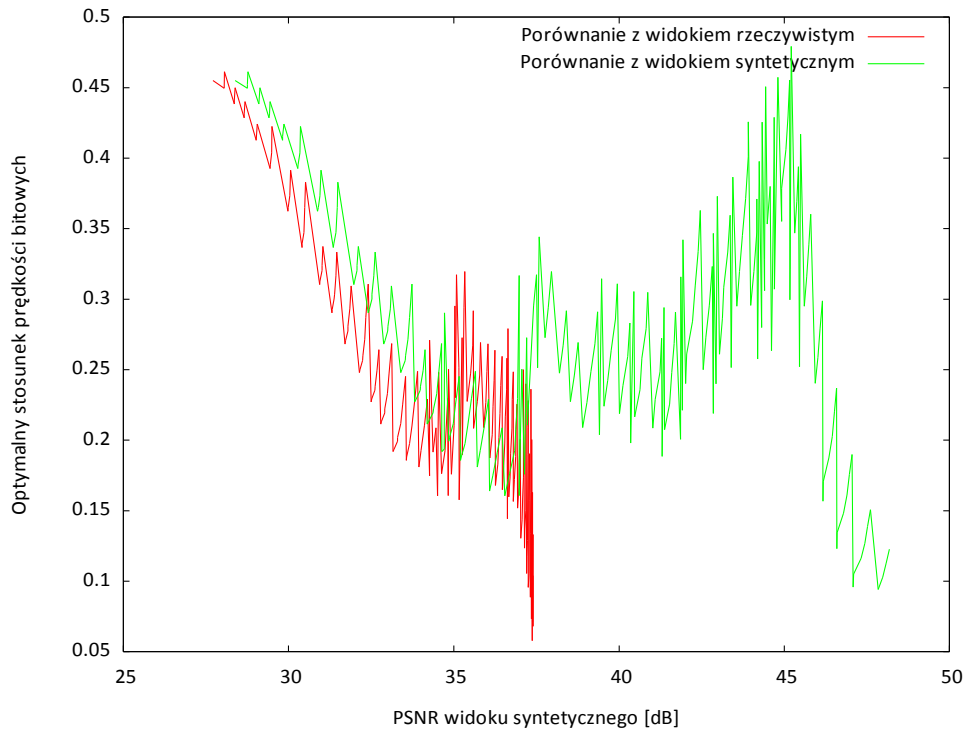
- im mniejsza odległość widoku syntetycznego od rzeczywistych kamer, z których pochodzą dane wejściowe, tym wyższa jakość widoku wirtualnego. Jest to dość łatwe do uzasadnienia, bowiem im mniejsze są odległości, tym mniejsze przesunięcia musi wprowadzić algorytm syntezy do widoku wirtualnego, a wobec tego jakiegokolwiek błędy czy zniekształcenia ujawniają się w znacznie mniejszym stopniu,
- wybór optymalnych par indeksów QP-QD jest uzależniony od treści sekwencji w znacznie większym stopniu niż od odległości między kamerami rzeczywistymi a wirtualną.

Wyniki dla sekwencji syntetycznej (Undo Dancer – scenariusz 3) odbiegają nieco od pozostałych. Warto w tym miejscu przypomnieć, że dla tej sekwencji dostępna jest idealna, pozbawiona błędów mapa głębi. Wobec tego zrozumiałym jest, że wybór indeksu QD ma znacznie większy wpływ na jakość widoku wirtualnego w porównaniu do innych rozpatrywanych scenariuszy. Większy wpływ wartości indeksu QD dobrze widać na rysunku 5.49 – optymalne pary QP-QD charakteryzują się większą zmiennością indeksu QD dla stałej wartości QP.

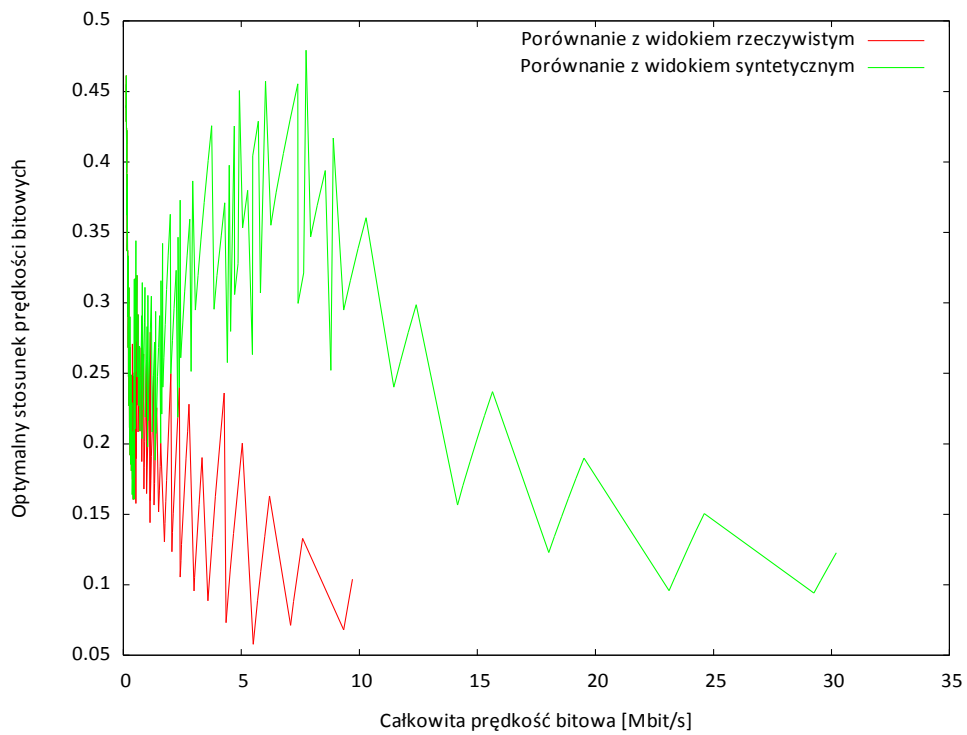
Wykresy prezentujące optymalne pary QP-QD dla wszystkich przebadanych scenariuszy mają wspólną cechę – zwiększanie jakości obrazów syntetycznych uzyskiwanych ze zrekonstruowanych danych polega w przeważającej większości przypadków na zmniejszeniu wartości indeksu QD przy utrzymaniu tego samego indeksu QP lub też na zmniejszeniu wartości indeksu QP przy jednoczesnym zmniejszeniu wartości indeksu QD. Bardzo rzadko zwiększenie jakości wiąże się ze zmianą wartości indeksów w inny sposób (na przykład zmniejszenie wartości QP przy stałym QD).

Zaskakujący może być fakt, że wartości miary PSNR uzyskiwane dla sekwencji „Undo Dancer” nie odbiegają od wartości uzyskiwanych dla innych sekwencji, i to pomimo korzystania z idealnych map głębi. Oznacza to, że obecnie stosowane metody nie zapewniają uzyskania poprawnego obrazu syntezy nawet z użyciem bezbłędnych map głębi.

Ciekawa obserwacja wynika z analizy stosunku prędkości bitowych poświęconych transmisji danych o obrazach map głębi do całości strumienia. Na rysunkach od 5.62 do 5.69 przedstawione są wykresy obrazujące tę wartość dla uzyskiwanych w eksperymentach wartości prędkości bitowych i wartości miary jakości PSNR.



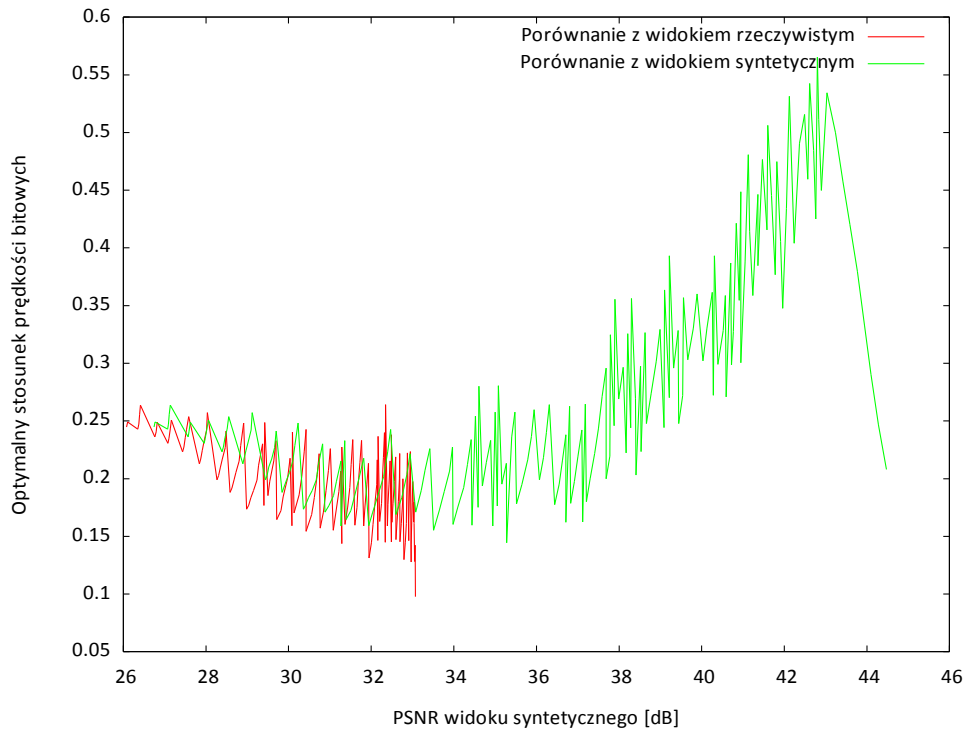
a)



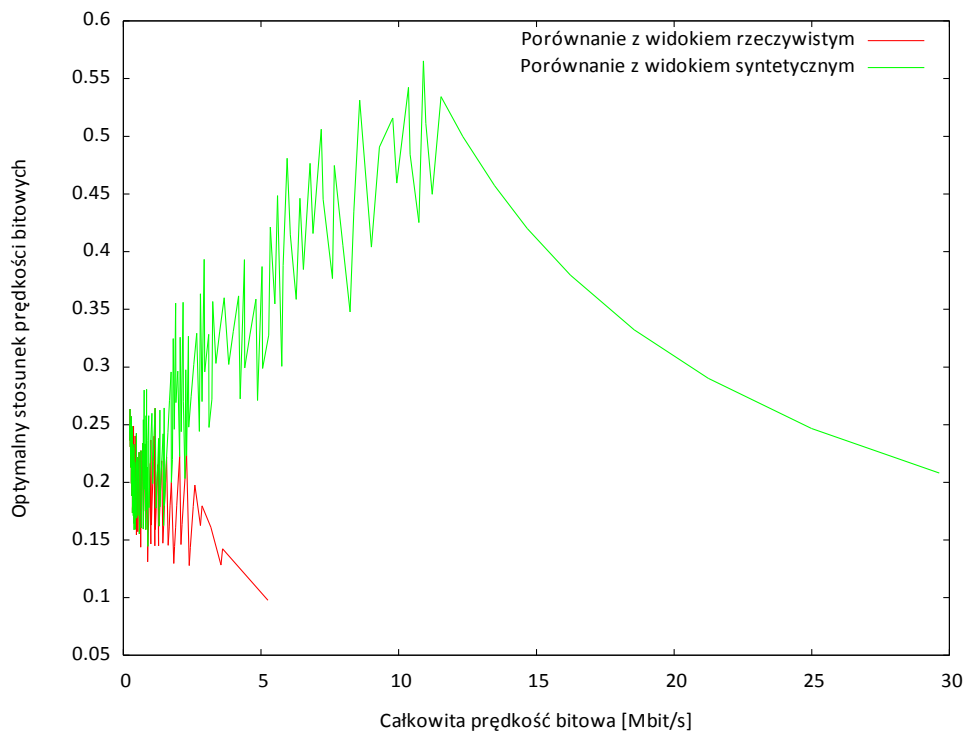
b)

Rysunek 5.62. Optymalny stosunek prędkości bitowej dla głębi w stosunku do sumarycznej prędkości bitowej w zależności od miary jakości widoku syntetycznego (a) i od całkowitej prędkości bitowej (b). Sekwencja Book Arrival (scenariusz 1).



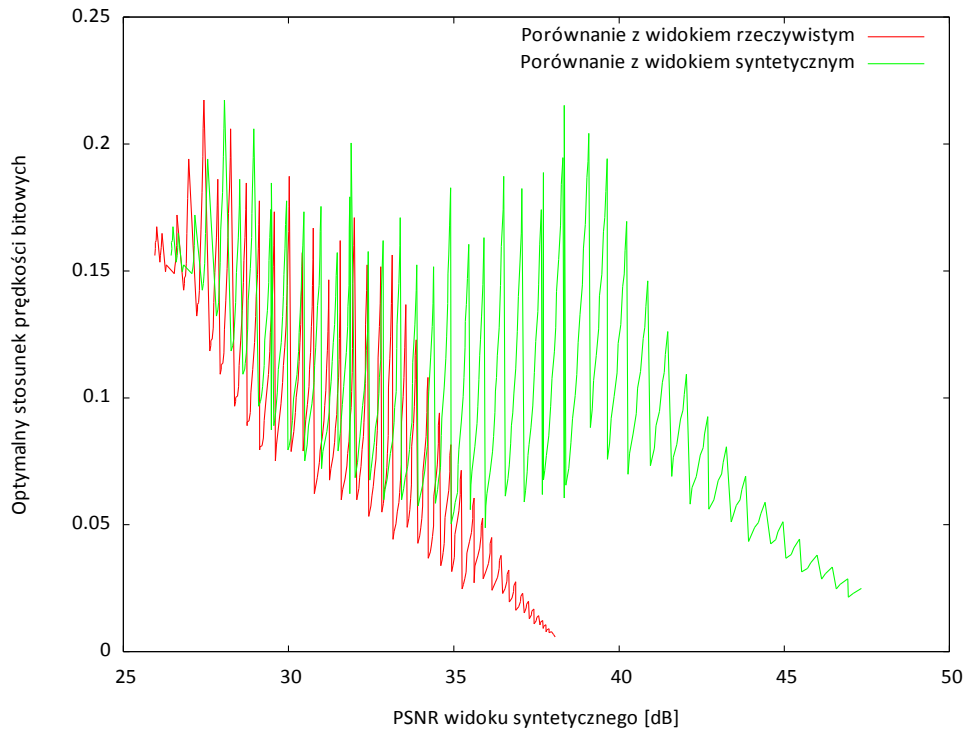


a)

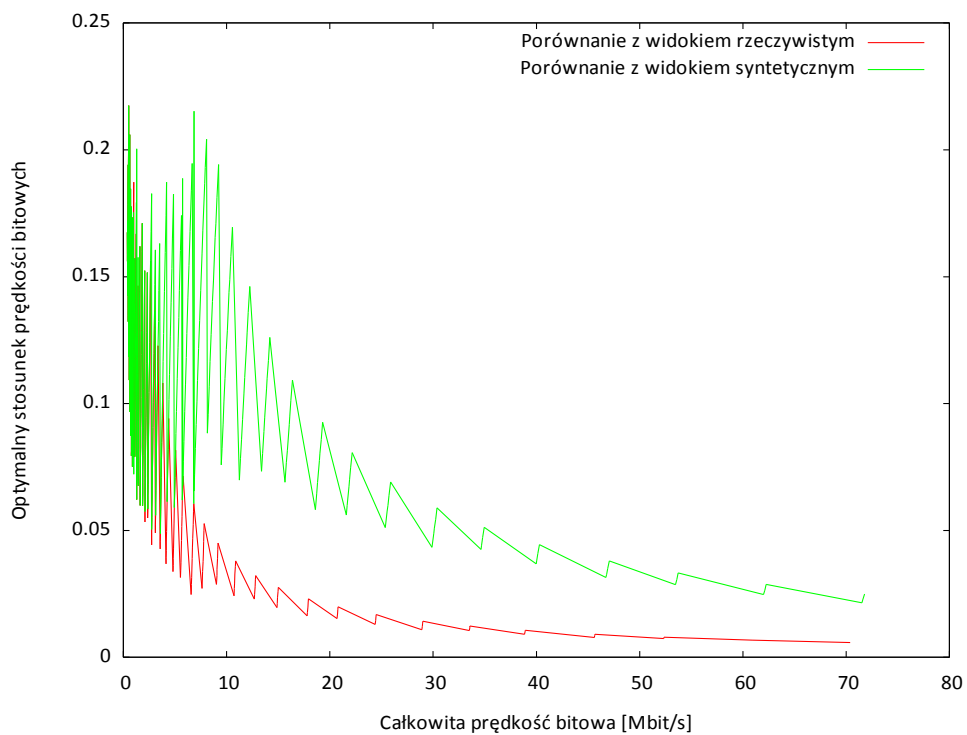


b)

Rysunek 5.63. Optymalny stosunek prędkości bitowej dla głębi w stosunku do sumarycznej prędkości bitowej w zależności od miary jakości widoku syntetycznego (a) i od całkowitej prędkości bitowej (b). Sekwencja Newspaper (scenariusz 2).

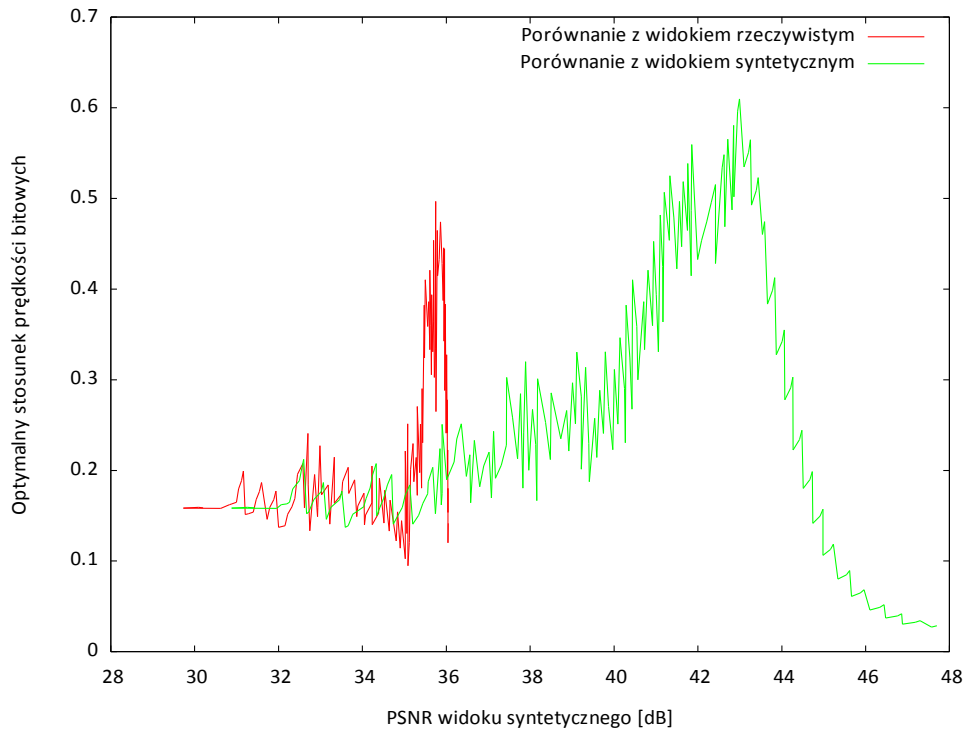


a)

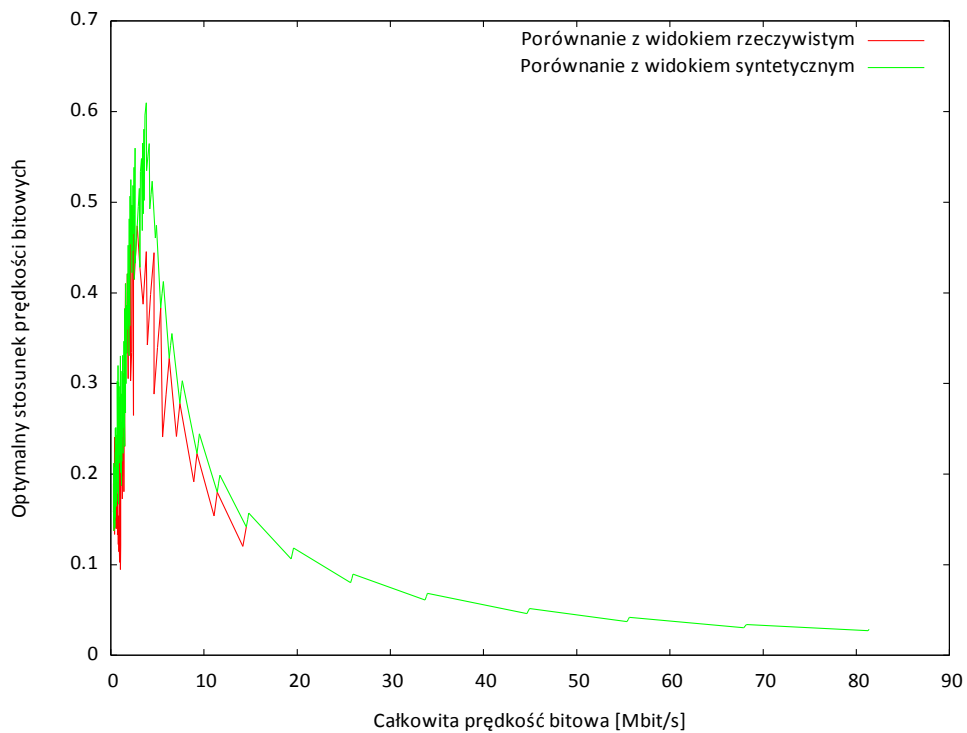


b)

Rysunek 5.64. Optymalny stosunek prędkości bitowej dla głębi w stosunku do sumarycznej prędkości bitowej w zależności od miary jakości widoku syntetycznego (a) i od całkowitej prędkości bitowej (b). Sekwencja Undo Dancer (scenariusz 3).

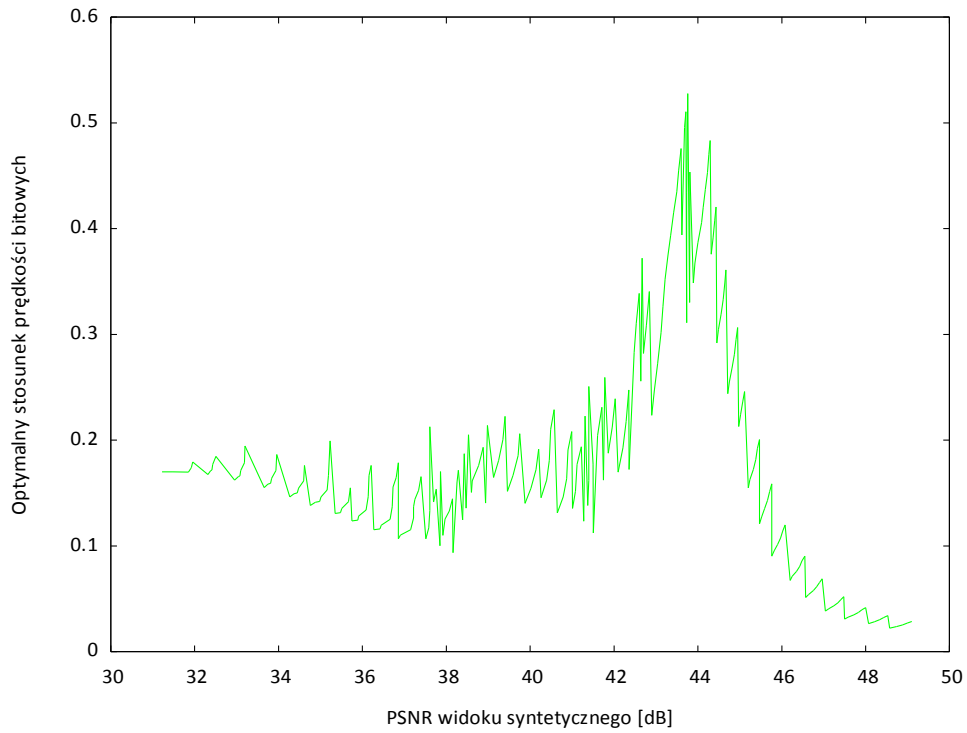


a)

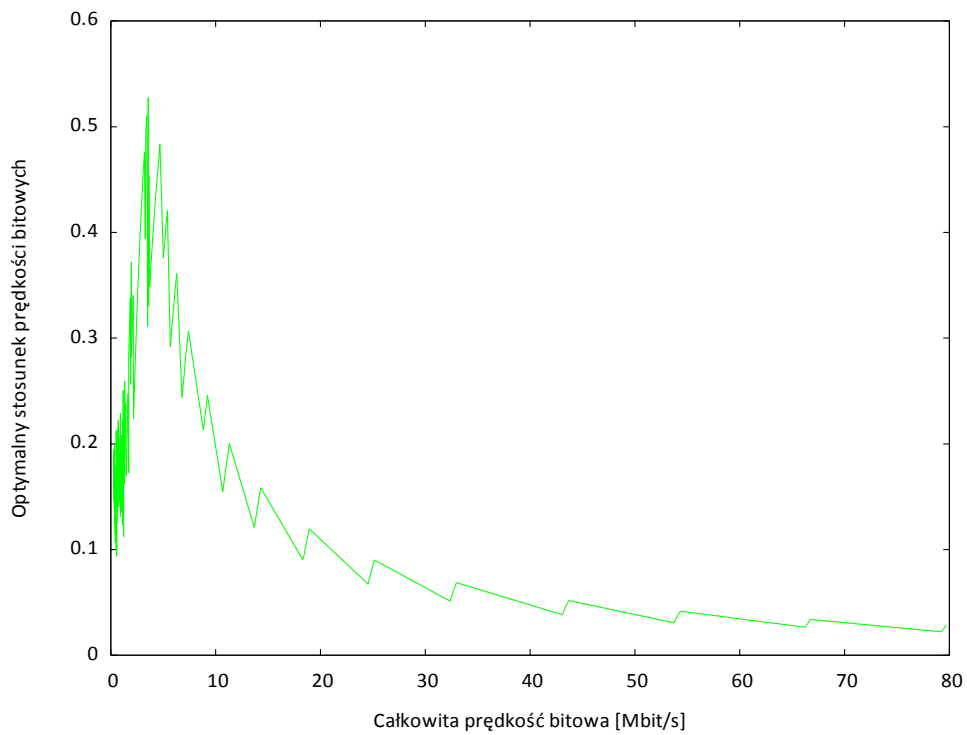


b)

Rysunek 5.65. Optymalny stosunek prędkości bitowej dla głębi w stosunku do sumarycznej prędkości bitowej w zależności od miary jakości widoku syntetycznego (a) i od całkowitej prędkości bitowej (b). Sekwencja Poznan Hall 2 (scenariusz 4).

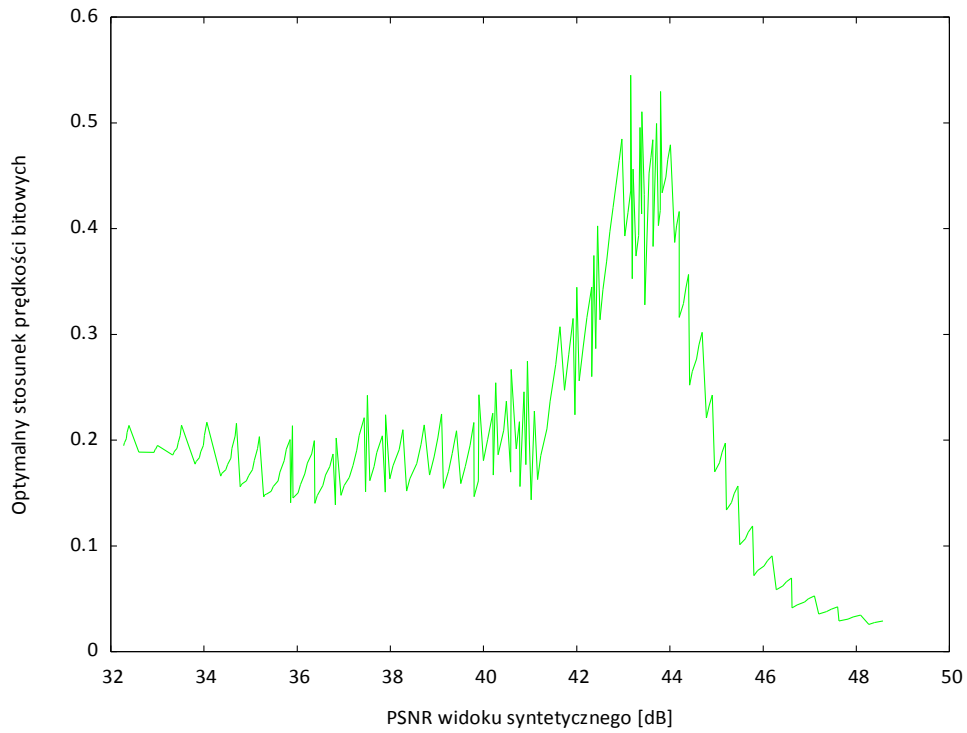


a)

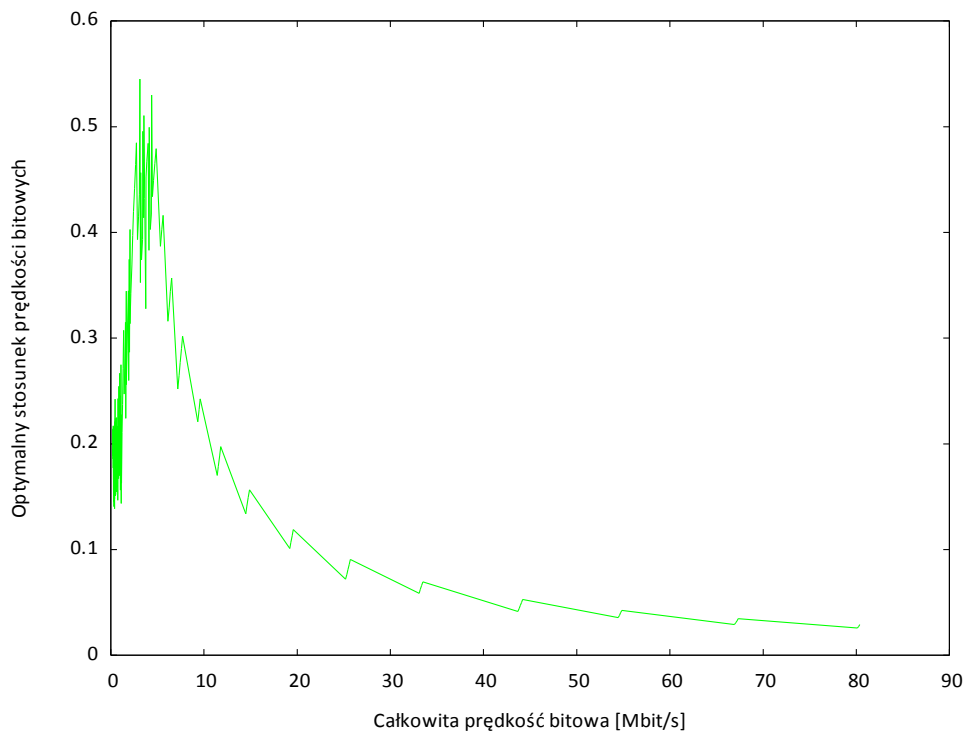


b)

Rysunek 5.66. Optymalny stosunek prędkości bitowej dla głębi w stosunku do sumarycznej prędkości bitowej w zależności od miary jakości widoku syntetycznego (a) i od całkowitej prędkości bitowej (b). Sekwencja Poznan Hall 2 (scenariusz 5).

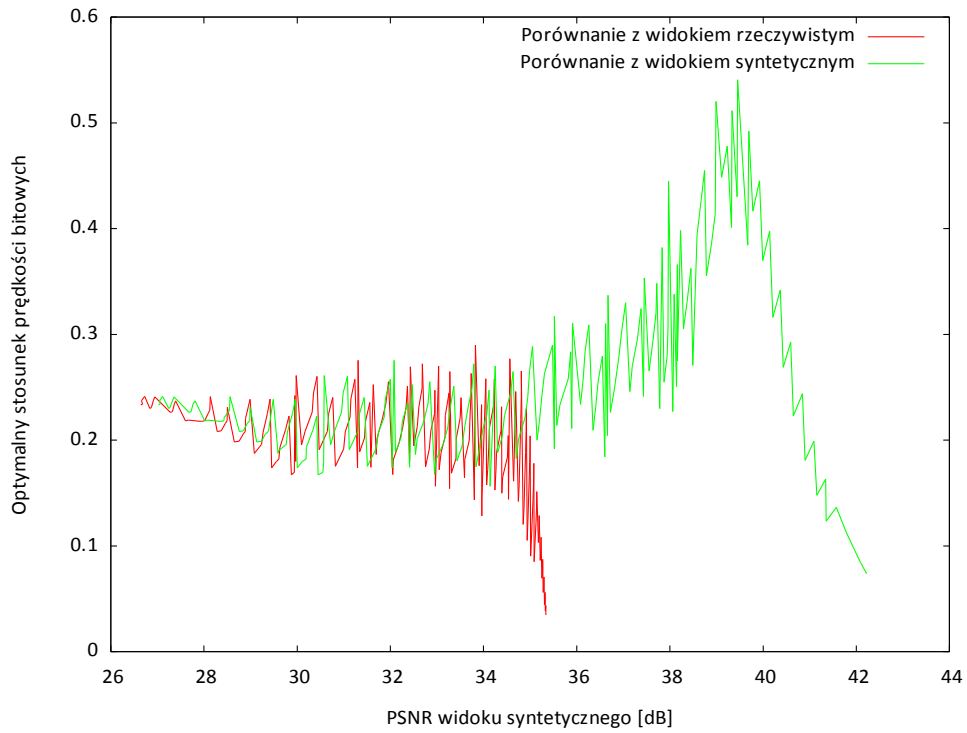


a)

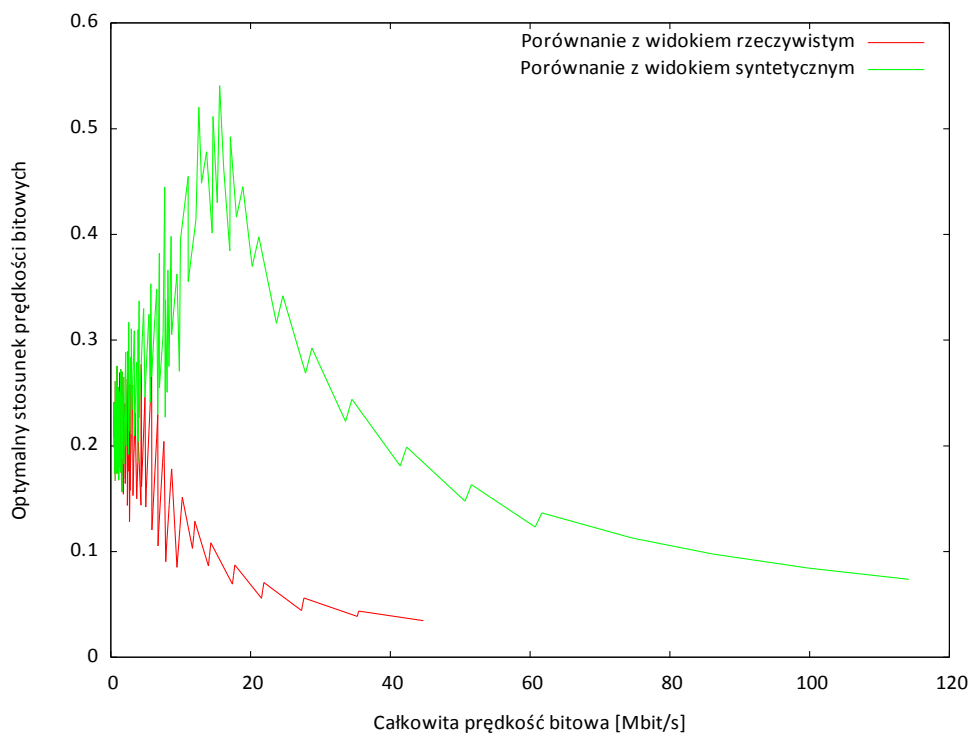


b)

Rysunek 5.67. Optymalny stosunek prędkości bitowej dla głębi w stosunku do sumarycznej prędkości bitowej w zależności od miary jakości widoku syntetycznego (a) i od całkowitej prędkości bitowej (b). Sekwencja Poznan Hall 2 (scenariusz 6).

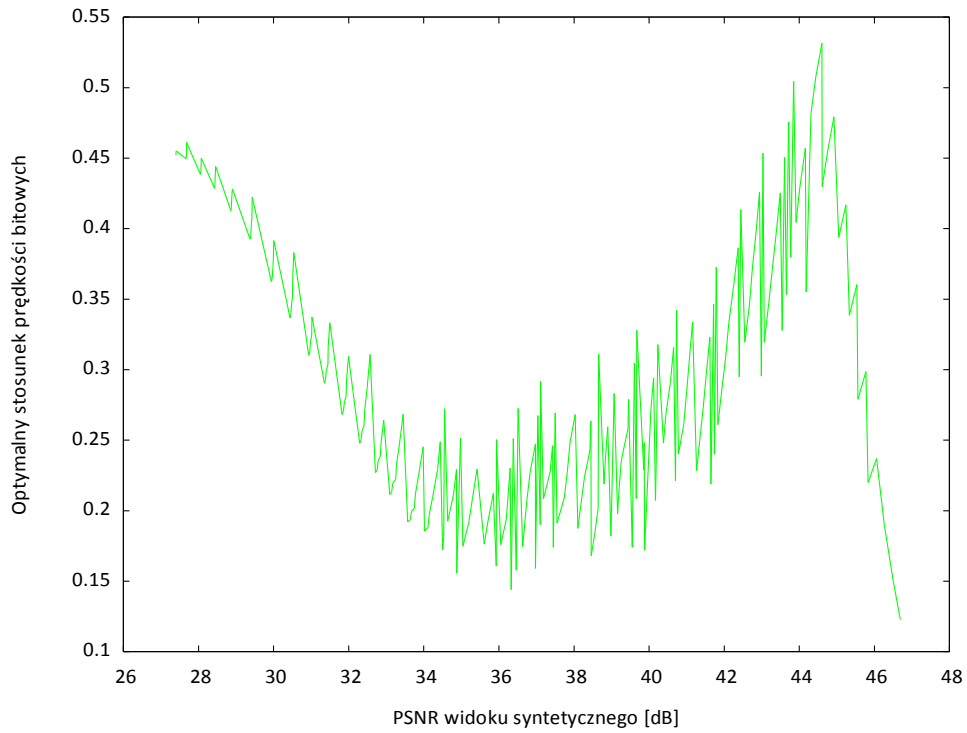


a)

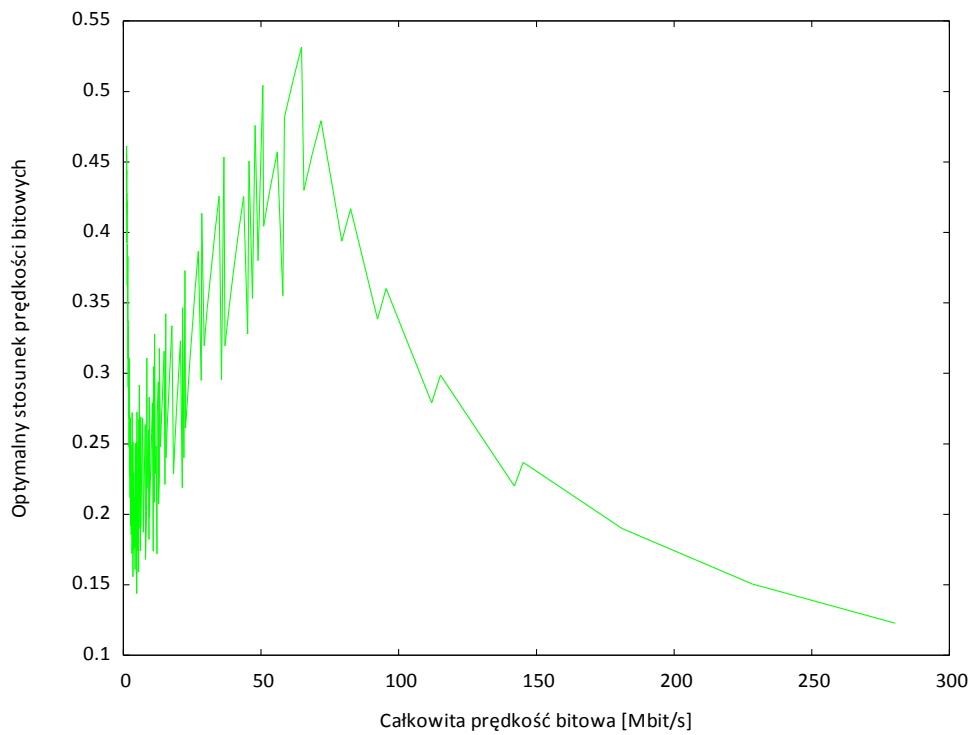


b)

Rysunek 5.68. Optymalny stosunek prędkości bitowej dla głębi w stosunku do sumarycznej prędkości bitowej w zależności od miary jakości widoku syntetycznego (a) i od całkowitej prędkości bitowej (b). Sekwencja Poznan Street (scenariusz 7).



a)



b)

Rysunek 5.69. Optymalny stosunek prędkości bitowej dla głębi w stosunku do sumarycznej prędkości bitowej w zależności od miary jakości widoku syntetycznego (a) i od całkowitej prędkości bitowej (b). Sekwencja Poznan Street (scenariusz 8).

Na wykresach od 5.62 do 5.69 widać, że, co może być nieco zaskakujące, dla uzyskania najlepszej jakości dla założonej prędkości bitowej syntezy widoku wirtualnego może okazać się konieczne poświęcenie więcej niż połowy prędkości bitowej na transmisję map głębi. Efekt ten jest widoczny dla przypadków, w których PSNR liczony jest na drodze porównania z widokiem syntezy z użyciem nieskompresowanych danych. Dla przypadków, dla których obliczenia miary jakości dokonuje się przez porównanie uzyskanego widoku z widokiem z rzeczywistej kamery, nie dochodzi do takich sytuacji. Jednak nawet wtedy stosunek prędkości bitowych dochodzi do zaskakujących wartości bliskich 50% całkowitej prędkości bitowej.

Zjawisko takie jest zaskakujące, ponieważ mapy głębi zdają się nieść znacznie mniej informacji niż obrazy, a do tego, jak zostało to pokazane w punkcie 5.1.4.1, ich kompresja jest bardziej efektywna niż kompresja obrazów. Mimo tego, do uzyskania dobrej jakości obrazu syntetycznego konieczne okazuje się zapewnienie odpowiednio dobrej jakości rekonstruowanego obrazu mapy głębi. O ile da się wytłumaczyć duże wartości stosunku prędkości bitowych dla małej całkowitej prędkości bitowej (w takim wypadku zarówno strumienie obrazów jak i map głębi są małe i w dużej mierze składają się z informacji sterujących), to znacznie bardziej zaskakujące jest pojawianie się wyraźnego maksimum tej wartości dla stosunkowo dużych wartości całkowitej prędkości bitowej. Wytłumaczeniem w tym wypadku może być pewne ciekawe zjawisko, zaobserwowane w trakcie prowadzonych prac. Okazuje się bowiem, że o jakości obrazu syntezy w ogólności (w większości przypadków) decyduje jakość obrazu użytego do przeprowadzenia procesu syntezy. Jakość map głębi okazuje się mieć znacznie mniejsze znaczenie (jest to szczególnie widoczne w badaniach subiektywnych) dla ostatecznej oceny jakości obrazu syntetycznego. Dla dużych prędkości bitowych uzyskiwane są dobre jakości obrazu, co pociąga za sobą dobre jakości obrazu syntetycznego. Jednak wraz z rosnącą jakością obrazu, coraz wyraźniej w obrazie syntezy widoczne stają się błędy wynikające z zakłóceń w mapach głębi. Zwiększanie jakości obrazu przy zachowaniu jakości map głębi nie daje od tego momentu tak dużego przyrostu jakości obrazu syntezy. Z tego powodu konieczne staje się zwiększanie prędkości bitowej poświęcanej transmisji map głębi, aby móc poprawić jakość obrazu syntezy.

### **5.2.2.3. Wyznaczenie zależności $QD = f(QP)$**

Z zaprezentowanych w punkcie 5.1.4.2 wyników można wysnuć wnioski dotyczące zależności optymalnej wartości indeksu QD od wartości indeksu QP. Wyniki uzyskane dla dwu używanych w pracy sposobów oceny jakości widoku syntetycznego należy rozpatrywać oddzielnie, gdyż charakter zależności  $QD = f(QP)$  jest dla nich nieco odmienny. Zależnie od wybranego sposobu



oceny jakości należy więc użyć innego, odpowiedniego dla danej metody, wzoru.

W celu wyznaczenia krzywych przybliżających zależność optymalnych indeksów QD od QP wyznaczone zostały, metodą regresji liniowej, funkcje drugiego stopnia przybliżające uzyskane w eksperymencie wartości dla par parametrów QP-QD. Dla każdego scenariusza wyznaczona została oddzielna krzywa, której przebieg odzwierciedla przybliżone położenie optymalnych par na płaszczyźnie parametrów QP oraz QD. Wykresy wraz z odpowiednimi wzorami przedstawione zostały na rysunkach poniżej, oddzielnie dla przypadków, dla których jakość mierzona była przez porównanie otrzymanego widoku z widokiem syntezowanym z nieskompresowanych danych i oddzielnie dla porównania z widokiem z rzeczywistej kamery.

### Pomiar jakości przez porównanie z widokiem syntetycznym

Jako pierwsze zbadane zostały pary wyznaczone dla przypadków, gdy miara jakości liczona była w odniesieniu do widoku syntezowanego z użyciem nieskompresowanych danych. Na wykresach z rysunków od 5.70 do 5.77 przedstawione są uzyskane wyniki. Dla każdego scenariusza podana została również wartość współczynnika determinacji  $R^2$ , którego wartość świadczy o jakości przybliżenia danych pomiarowych przez krzywą regresji [Ciec\_07][Koro\_09]. Współczynnik ten wyznaczany jest z poniższego wzoru (5.3):

$$R^2 = \frac{\sum_i (\widehat{QD} - \overline{QD})^2}{\sum_i (QD_i - \overline{QD})^2}, \quad (5.3)$$

gdzie:

$i$  – liczba par QP-QD,

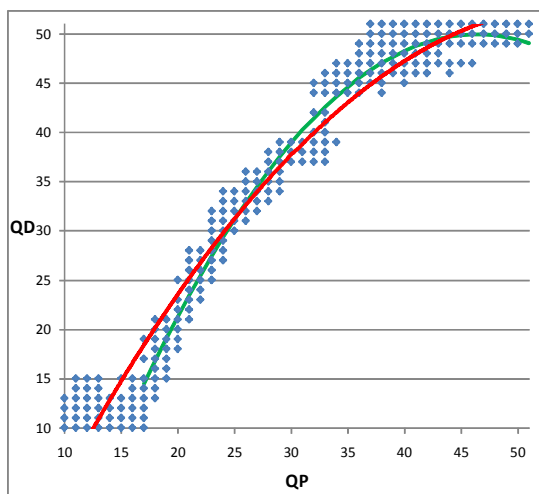
$\widehat{QD}$  - wartość QD wyliczona z pomocą wyznaczonego wzoru  $QD = f(QP)$ ,

$\overline{QD}$  - wartość średnia QD.

Im wartość tego współczynnika bliższa jest jedności, tym lepsze jest dopasowanie krzywej regresji do uzyskanych punktów eksperymentalnych.

Oznaczenia na rysunkach od 5.70 do 5.77 są następujące:

- linia A – czerwona – wyznacza przebieg krzywej wyznaczonej za pomocą regresji liniowej z całości danych (QP z zakresu od 10 do 51),
- linia B – zielona – wyznacza przebieg krzywej wyznaczonej za pomocą regresji liniowej z części danych (QP z zakresu od 17 do 51),
- niebieskie punkty – punkty oznaczające optymalne pary QP-QD.



dla linii A:

$$QD = -0,0231 \cdot QP^2 + 2,5745 \cdot QP - 18,728$$

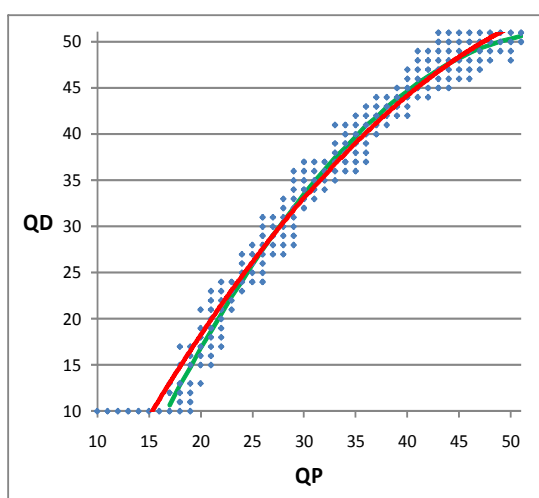
$$R^2 = 0,9532$$

dla linii B:

$$QD = -0,041 \cdot QP^2 + 3,8024 \cdot QP - 38,281$$

$$R^2 = 0,9652$$

Rysunek 5.70. Przybliżona zależność  $QD=f(QP)$  dla optymalnych par w scenariuszu 1 (sekwencja Book Arrival).



dla linii A:

$$QD = -0,01855 \cdot QP^2 + 2,4103 \cdot QP - 22,5911$$

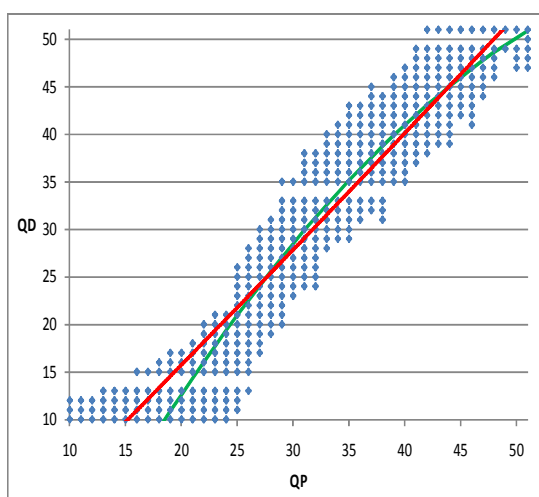
$$R^2 = 0,965$$

dla linii B:

$$QD = -0,02765 \cdot QP^2 + 3,0548 \cdot QP - 33,2524$$

$$R^2 = 0,9734$$

Rysunek 5.71. Przybliżona zależność  $QD=f(QP)$  dla optymalnych par w scenariuszu 2 (sekwencja Newspaper).



dla linii A:

$$QD = -0,0008 \cdot QP^2 + 1,1722 \cdot QP - 7,9863$$

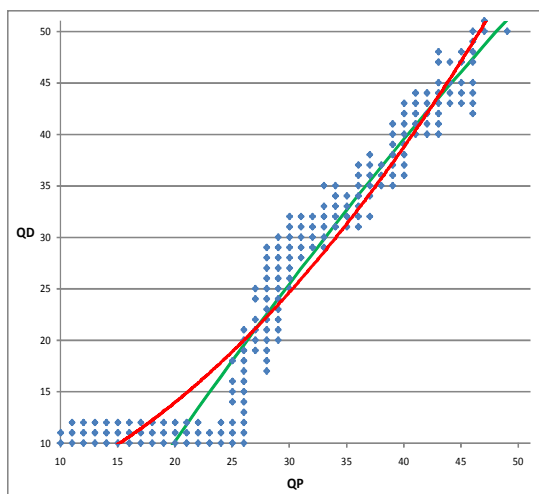
$$R^2 = 0,8864$$

dla linii B:

$$QD = -0,01662 \cdot QP^2 + 2,4151 \cdot QP - 29,049$$

$$R^2 = 0,8892$$

Rysunek 5.72. Przybliżona zależność  $QD=f(QP)$  dla optymalnych par w scenariuszu 3 (sekwencja Undo Dancer).



dla linii A:

$$QD = 0,0171 \cdot QP^2 + 0,2197 \cdot QP + 2,7581$$

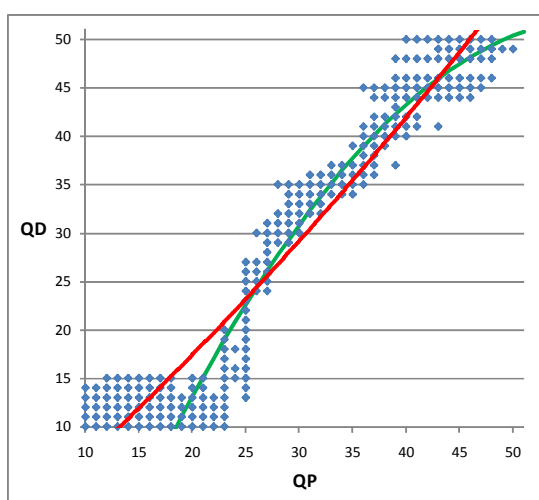
$$R^2 = 0,9222$$

dla linii B:

$$QD = -0,006 \cdot QP^2 + 1,8289 \cdot QP - 23,934$$

$$R^2 = 0,9289$$

Rysunek 5.73. Przybliżona zależność  $QD=f(QP)$  dla optymalnych par w scenariuszu 4 (sekwencja Poznan Hall 2).



dla linii A:

$$QD = 0,0049 \cdot QP^2 + 0,9328 \cdot QP - 3,1613$$

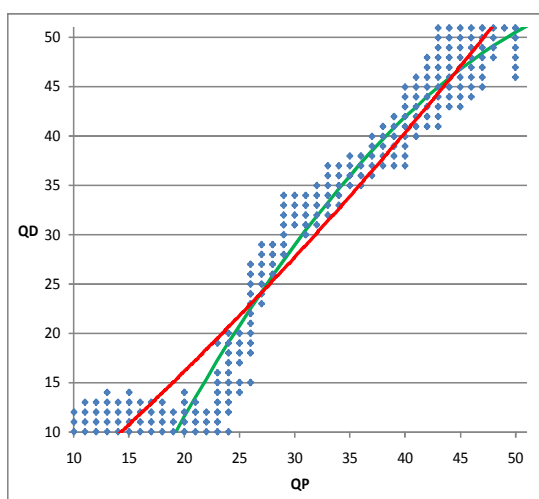
$$R^2 = 0,913$$

dla linii B:

$$QD = -0,02676 \cdot QP^2 + 3,1163 \cdot QP - 38,5598$$

$$R^2 = 0,9341$$

Rysunek 5.74. Przybliżona zależność  $QD=f(QP)$  dla optymalnych par w scenariuszu 5 (sekwencja Poznan Hall 2).



dla linii A:

$$QD = 0,0053 \cdot QP^2 + 0,8945 \cdot QP - 3,8828$$

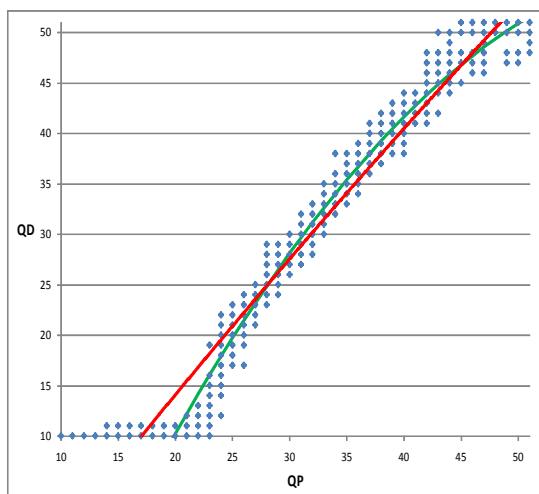
$$R^2 = 0,931$$

dla linii B:

$$QD = -0,0219 \cdot QP^2 + 2,8353 \cdot QP - 36,364$$

$$R^2 = 0,9483$$

Rysunek 5.75. Przybliżona zależność  $QD=f(QP)$  dla optymalnych par w scenariuszu 6 (sekwencja Poznan Hall 2).



dla linii A:

$$QD = 0,0027 \cdot QP^2 + 1,4809 \cdot QP - 14,415$$

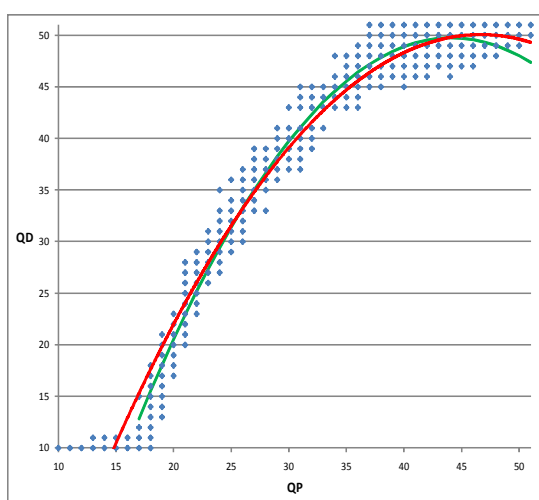
$$R^2 = 0,9483$$

dla linii B:

$$QD = -0,02143 \cdot QP^2 + 2,8526 \cdot QP - 38,1798$$

$$R^2 = 0,9670$$

Rysunek 5.76. Przybliżona zależność  $QD=f(QP)$  dla optymalnych par w scenariuszu 7 (sekwencja Poznan Street).



dla linii A:

$$QD = -0,0395 \cdot QP^2 + 3,6837 \cdot QP - 35,843$$

$$R^2 = 0,9565$$

dla linii B:

$$QD = -0,05023 \cdot QP^2 + 4,4333 \cdot QP - 48,0555$$

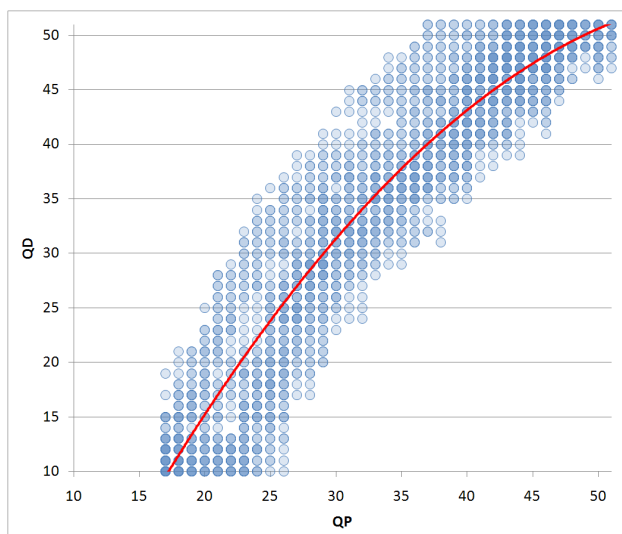
$$R^2 = 0,96579$$

Rysunek 5.77. Przybliżona zależność  $QD=f(QP)$  dla optymalnych par w scenariuszu 8 (sekwencja Poznan Street).

Przedstawione na powyższych wykresach czerwone krzywe regresji mają różny przebieg dla rozpatrzonych scenariuszy. Daje się jednak zauważyć, że wyraźnie zaznacza się pewna prawidłowość dotycząca optymalnych par QP-QD dla wartości parametrów QP większych od około 17 (zielona krzywa). Dla wszystkich scenariuszy obserwuje się w tym wypadku bardzo podobne rozłożenie punktów odpowiadających optymalnym parom QP-QD. Wydaje się zatem celowe wyznaczenie uogólnionej krzywej, za pomocą której można przybliżyć optymalne wartości indeksu QD dla każdego z przebadanych scenariuszy. Krzywa taka może być wyznaczona poprzez wyznaczenie krzywej regresji dla wszystkich par QP-QD dla wszystkich sekwencji.

Przebieg uogólnionej krzywej regresji dla przedziału wartości indeksu QP od 17 do 51 zaprezentowany został wraz z punktami odpowiadającymi optymalnym parom QP-QD dla przebadanych sekwencji na rysunku 5.78. Nasycenie barwy punktu odpowiada częstości jego występowania w wynikach dla rozpatrzonych scenariuszy. Ponieważ rozpatrzono 8 scenariuszy,

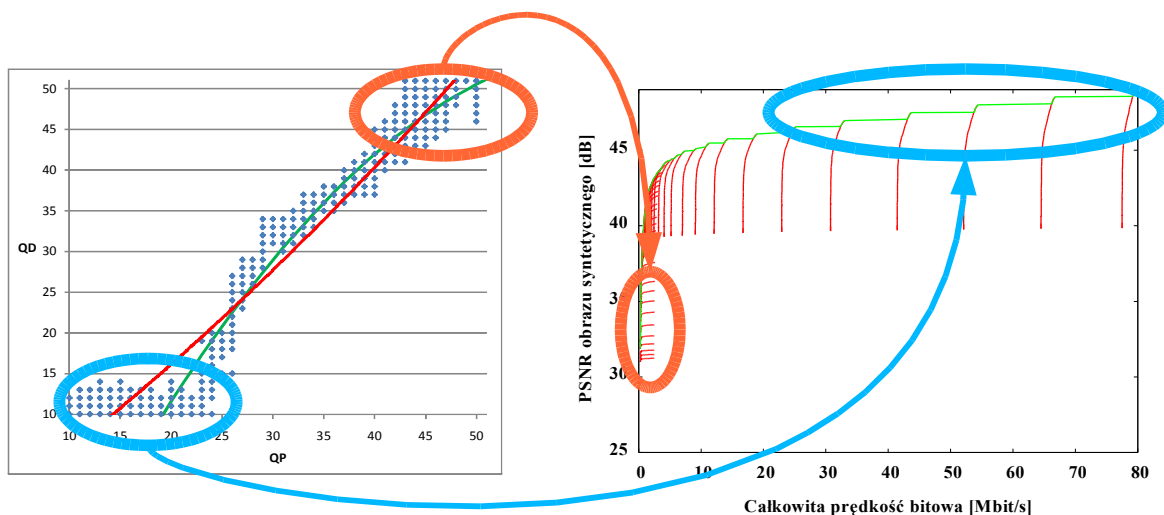
każdy punkt może wystąpić maksymalnie 8 razy i tylokrotne powtórzenie odpowiada największemu nasyceniu koloru punktu.



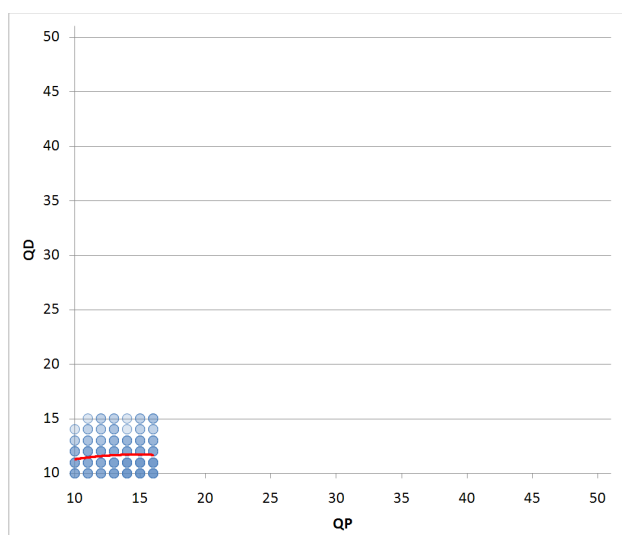
dla czerwonej linii:  
$$QD = -0,0216 \cdot QP^2 + 2,6872 \cdot QP - 29,876$$
$$R^2 = 0,8521$$

Rysunek 5.78. Optymalne pary indeksów QP-QD dla wszystkich rozpatrywanych scenariuszy (dane łączne) w przedziale QP od 17 do 51.

Dla wartości indeksu QP z przedziału od 10 do 16 dla wszystkich scenariuszy obserwuje się podobne ułożenie punktów odpowiadających optymalnym parom QP-QD. Dla tego zakresu obserwuje się utrzymywanie wartości optymalnego indeksu QD na stałym poziomie zbliżonym do wartości minimalnych z rozpatrywanego zbioru. Warto w tym miejscu podkreślić, że temu przedziałowi wartości indeksu QP odpowiadają prędkości bitowe rzędu powyżej 10 Mbit/s potrzebne do transmisji 2 widoków wraz z mapami głębi (a zatem o ograniczonym praktycznym znaczeniu), jak pokazane zostało to na rysunku 5.79.



Rysunek 5.79. Położenie punktów odpowiadających tym samym optymalnym parom QP-QD na wykresach prezentowanych w pracy.



dla czerwonej linii:  
 $QD = -0,0199 \cdot QP^2 + 0,5854 \cdot QP + 7,4142$   
 $R^2 = 0,0091$

Rysunek 5.80. Optymalne pary indeksów QP-QD dla wszystkich rozpatrywanych scenariuszy (dane łączne) w przedziale QP od 10 do 16.

Jak widać na rysunku 5.80, dla wartości QP mniejszych niż 17, krzywa regresji przyjmuje wartości, które po zaokrągleniu w dół wynoszą 11. Warto zwrócić uwagę, że uzyskana w tym przypadku wartość współczynnika determinacji  $R^2$  jest bardzo mała, i świadczy o tym, że wyznaczona zależność słabo przybliża punkty uzyskane w eksperymencie. Uprawnione wydaje się być twierdzenie o możliwości zastosowania w zakresie wartości indeksu QP z przedziału od 10 do 16 stałej wartości QD równej 11.

W przypadkach, gdy miara jakości PSNR wyznaczana jest przez porównanie z obrazem syntezowanym z nieskompresowanych danych wartość QD w zależności od QP określona może

być zatem poniższą zależnością:

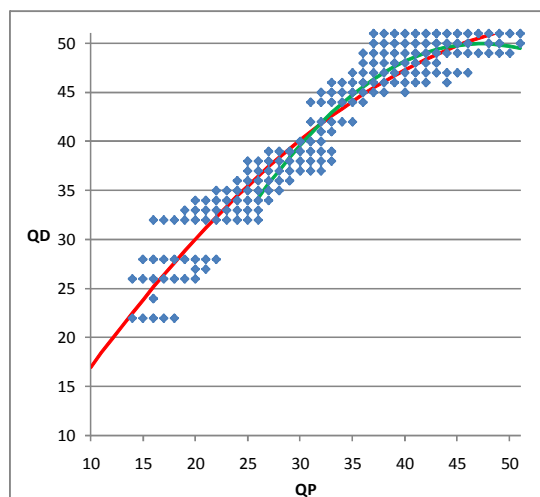
$$QD = \begin{cases} [-0,0216 \cdot QP^2 + 2,6872 \cdot QP - 29,876 + 0,5] & \text{dla } QP > 16 \\ 11 & \text{dla } QP \leq 16 \end{cases} \quad (5.4)$$

## Pomiar jakości przez porównanie z widokiem z rzeczywistej kamery

W drugiej kolejności przebadane zostały optymalne pary wartości QP-QD dla przypadków, gdy jakość generowanego widoku wyznaczana była na podstawie porównania z widokiem z rzeczywistej kamery. Na wykresach z rysunków od 5.81 do 5.85 przedstawione są wyniki.

Oznaczenia na rysunkach od 5.81 do 5.85 są następujące:

- **linia A** – czerwona – wyznacza przebieg krzywej wyznaczonej za pomocą regresji liniowej z całości danych (QP z zakresu od 10 do 51),
- **linia B** – zielona – wyznacza przebieg krzywej wyznaczonej za pomocą regresji liniowej z części danych (QP z zakresu od 27 do 51),
- **niebieskie** punkty – punkty oznaczające optymalne pary QP-QD.



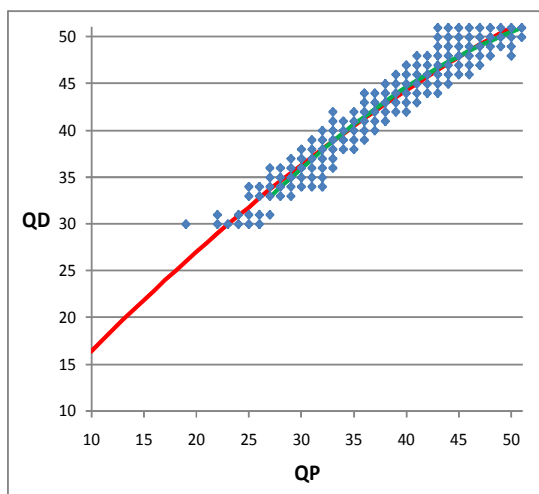
dla linii A:

$$QD = -0,0149 \cdot QP^2 + 1,75365 \cdot QP - 0,9295 \\ R^2 = 0,9151$$

dla linii B:

$$QD = -0,0349 \cdot QP^2 + 3,2937 \cdot QP - 27,789 \\ R^2 = 0,8564$$

Rysunek 5.81. Przybliżona zależność  $QD=f(QP)$  dla optymalnych par w scenariuszu 1 (sekwencja Book Arrival).



dla linii A:

$$QD = -0,00651 \cdot QP^2 + 1,2568 \cdot QP - 4,4915$$

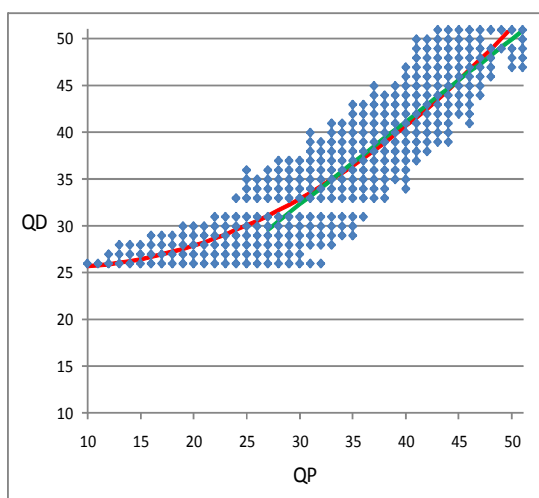
$$R^2 = 0,923$$

dla linii B:

$$QD = -0,0137 \cdot QP^2 + 1,8226 \cdot QP - 6,3907$$

$$R^2 = 0,9034$$

Rysunek 5.82. Przybliżona zależność  $QD=f(QP)$  dla optymalnych par w scenariuszu 2 (sekwencja Newspaper).



dla linii A:

$$QD = 0,01389 \cdot QP^2 - 0,1967 \cdot QP + 26,329$$

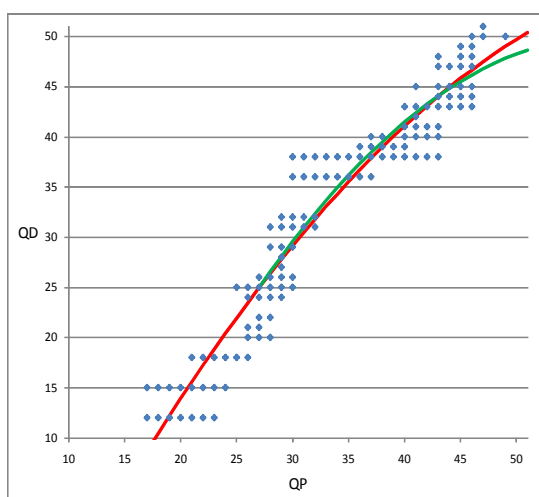
$$R^2 = 0,7923$$

dla linii B:

$$QD = -0,00047 \cdot QP^2 + 0,9215 \cdot QP + 5,1195$$

$$R^2 = 0,6994$$

Rysunek 5.83. Przybliżona zależność  $QD=f(QP)$  dla optymalnych par w scenariuszu 3 (sekwencja Undo Dancer).



dla linii A:

$$QD = -0,0165 \cdot QP^2 + 2,3459 \cdot QP - 26,441$$

$$R^2 = 0,9235$$

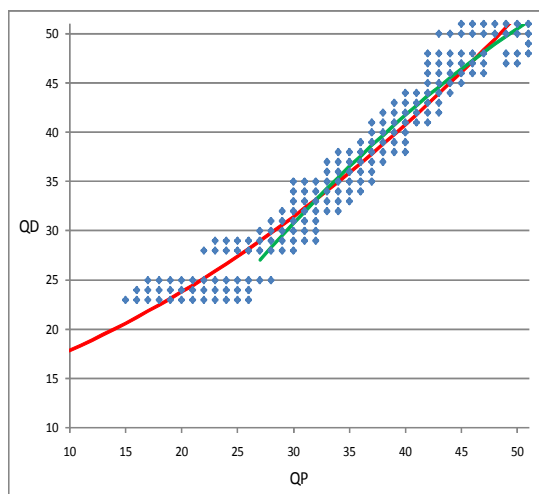
dla linii B:

$$QD = -0,0257 \cdot QP^2 + 2,9877 \cdot QP - 36,955$$

$$R^2 = 0,8631$$

Rysunek 5.84. Przybliżona zależność  $QD=f(QP)$  dla optymalnych par w scenariuszu 4 (sekwencja Poznan Hall 2).





dla linii A:

$$QD = 0,00857 \cdot QP^2 + 0,339 \cdot QP + 13,5755$$

$$R^2 = 0,9338$$

dla linii B:

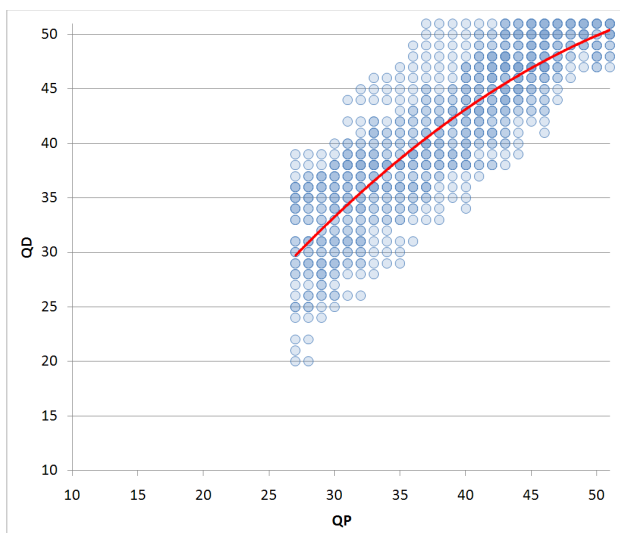
$$QD = -0,011 \cdot QP^2 + 1,8688 \cdot QP - 15,399$$

$$R^2 = 0,9132$$

Rysunek 5.85. Przybliżona zależność  $QD=f(QP)$  dla optymalnych par w scenariuszu 7 (sekwencja Poznan Street).

Przedstawione na powyższych wykresach czerwone krzywe regresji mają różny przebieg dla rozpatrzonych scenariuszy. Pewne podobieństwo jest zauważalne w przypadku optymalnych par QP-QD dla wartości indeksów QP większych od około 26 (zielona krzywa). Dla wszystkich scenariuszy obserwuje się w tym wypadku podobne rozłożenie punktów odpowiadających optymalnym parom QP-QD i zbliżone wartości współczynników regresji. Można zatem podjąć próby wyznaczenia uogólnionej krzywej dla tego przedziału.

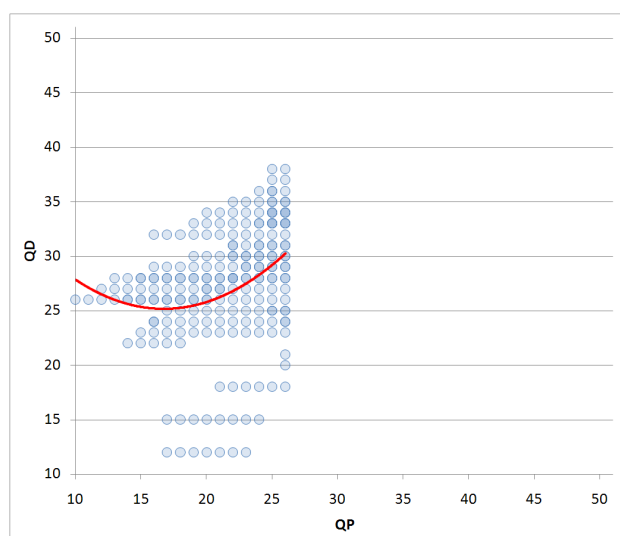
Przebieg uogólnionej krzywej regresji dla przedziału wartości indeksu QP od 27 do 51 zaprezentowany został wraz z punktami odpowiadającymi optymalnym parom QP-QD dla przebadanych sekwencji na rysunku 5.86. Nasycenie koloru punktu odpowiada częstości jego występowania w wynikach dla rozpatrzonych scenariuszy. Ponieważ rozpatrzono 5 scenariuszy (tylko dla nich możliwe było porównanie do widoku z rzeczywistej kamery), każdy punkt może wystąpić maksymalnie 5 razy i tyłokrotne powtórzenie odpowiada największemu nasyceniu koloru punktu. Wyznaczona krzywa regresji dość dobrze przybliży rozłożenie uzyskanych w eksperymentach punktów – wartość współczynnika  $R^2$  około 0,7 – nieco mniejsza niż w przypadku indywidualnego rozpatrywania scenariuszy i mniejsza niż uogólniona krzywa dla przypadku porównywania otrzymanych obrazów syntetycznych z obrazem uzyskanym poprzez syntezę z oryginalnych danych.



dla czerwonej linii:  
 $QD = -0,0155 \cdot QP^2 + 2,073 \cdot QP - 14,885$   
 $R^2 = 0,6936$

Rysunek 5.86. Optymalne pary indeksów QP-QD dla wszystkich rozpatrywanych scenariuszy (dane łączne) w przedziale QP od 27 do 51.

Dla wartości indeksu QP z przedziału od 10 do 26 dla wszystkich scenariuszy obserwuje się różne zachowanie punktów odpowiadających optymalnym parom QP-QD. Zbiorcze zestawienie optymalnych par QP-QD przedstawione jest dla tego przedziału na rysunku 5.87.



dla czerwonej linii:  
 $QD = 0,0585 \cdot QP^2 - 1,9579 \cdot QP + 41,562$   
 $R^2 = 0,1037$

Rysunek 5.87. Optymalne pary indeksów QP-QD dla wszystkich rozpatrywanych scenariuszy (dane łączne) w przedziale QP od 10 do 26.

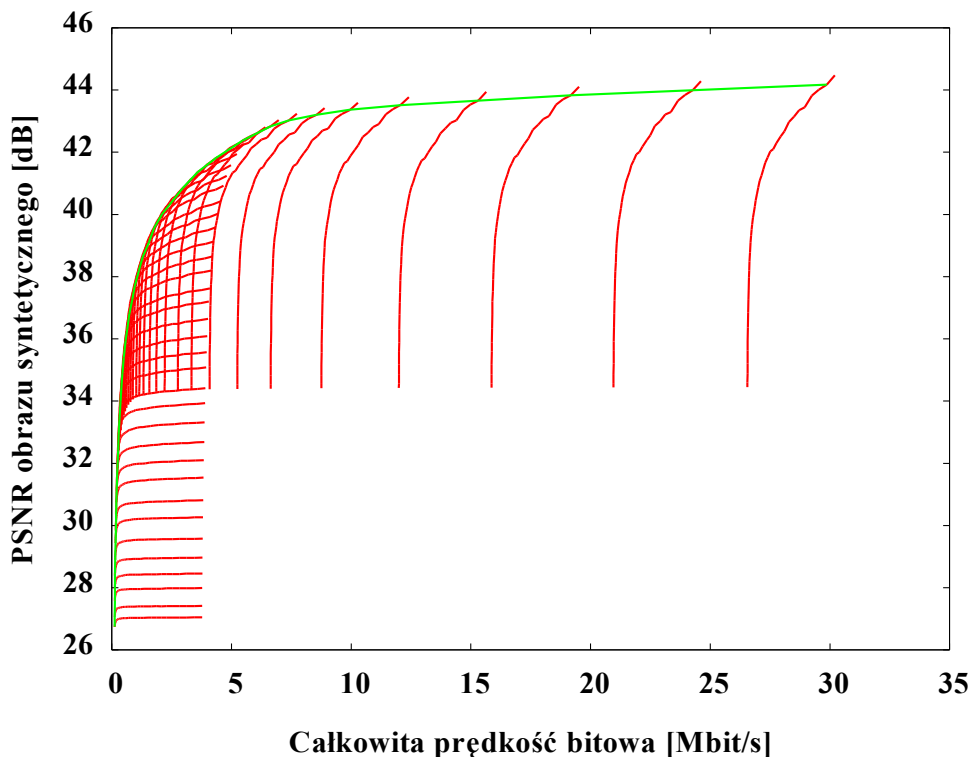
Jak widać na rysunku 5.87, dla wartości QP mniejszych niż 27, nie da się skutecznie przewidzieć wartości optymalnego indeksu QD na podstawie znajomości indeksu QP. Uzyskana w tym przypadku krzywa regresji bardzo słabo (współczynnik  $R^2$  osiąga niską wartość) przybliża rozłożenie punktów pomiarowych. Ostatecznie zatem, biorąc pod uwagę duży rozrzut punktów,

wzór zależności  $QD = f(QP)$  przyjmować może poniższą postać:

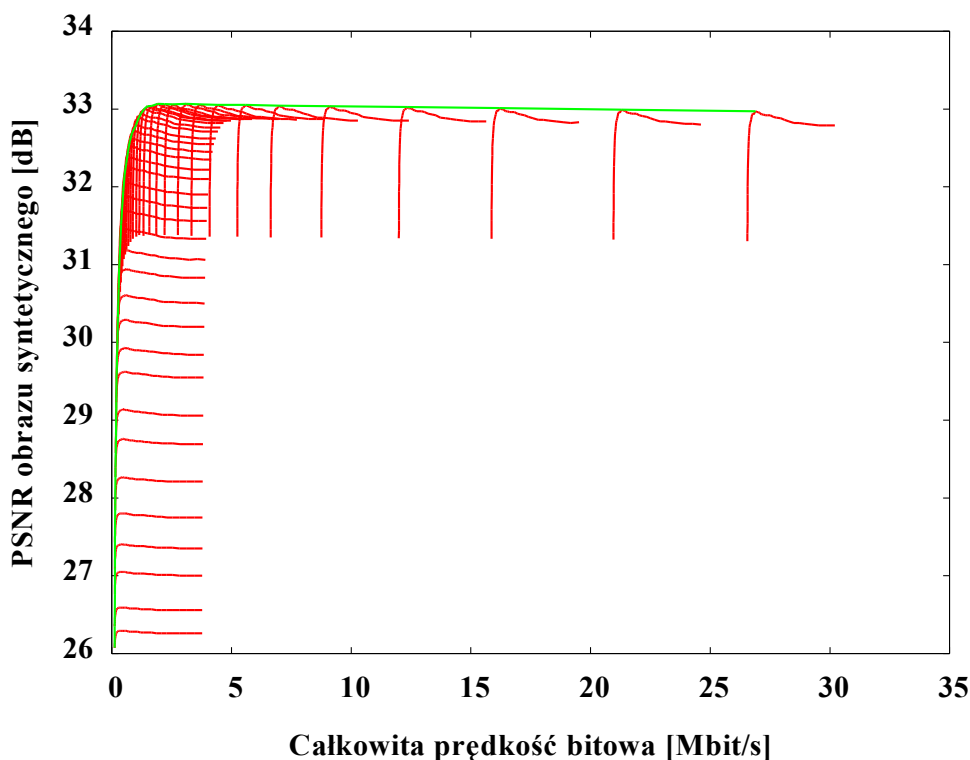
$$QD = -0,0155 \cdot QP^2 + 2,073 \cdot QP - 14,885 \quad (5.5)$$

### 5.2.3. Weryfikacja opracowanych wzorów

W celu zweryfikowania przydatności opracowanych wzorów przeprowadzone zostało badanie z wykorzystaniem sekwencji spoza zbioru scenariuszy, z których pochodziły dane do wyznaczenia zależności  $QD = f(QP)$ . Jako sekwencję spoza zbioru wybrano Champagne Tower [Tani\_08]. Dokonano podobnych obliczeń co dla scenariuszy 1 do 8. Przebadano fragment sekwencji o długości 100 ramek. Uzyskane wyniki, wraz z wynikami uzyskanymi przy wykorzystaniu zaproponowanych wzorów przedstawione są na rysunkach 5.88 i 5.89. Linie czerwone przedstawiają wyniki dla wszystkich par z przebadanego zakresu, a linia zielona – krzywą uzyskaną dla par, w których  $QD$  określone jest za pomocą odpowiedniego wzoru.



Rysunek 5.88. Jakość widoku syntezy w odniesieniu do widoku syntezy z nieskompresowanych danych dla sekwencji Champagne Tower wraz z wynikami uzyskanymi przy użyciu proponowanego wzoru (5.4).



Rysunek 5.89. Miara jakości widoku syntezowanego w odniesieniu do rzeczywistego widoku z kamery dla sekwencji Champagne Tower wraz z wynikami uzyskanymi przy użyciu proponowanego wzoru (5.5).

Wyniki zaprezentowane na rysunkach 5.88 oraz 5.89 świadczą o tym, że proponowane wzory (5.4) oraz (5.5) dobrze nadają się do automatycznego wyznaczania wartości parametru QD na podstawie parametru QP. W prezentowanym przykładzie jakość uzyskiwana w przypadku wykorzystania wzoru dobrze przybliża przebieg linii wyznaczającej najlepszą jakość dla danej prędkości bitowej, przy czym największe różnice obserwuje się dla dużych prędkości bitowych (czyli dla niskich wartości parametru QP), dla których współczynnik korelacji opracowanego wzoru z danymi eksperymentalnymi był niski. Warto jednak zaznaczyć, że w całym zakresie różnica między uzyskaną jakością a jakością maksymalną wynosi mniej niż 0,5 dB.

#### 5.2.4. Podsumowanie i wnioski

W punkcie 5.2 przedstawiono wyniki badań dotyczących doboru wartości parametrów QP i QD, pozwalających na uzyskanie jak najwyższej jakości widoków syntezowanych z użyciem zdekompresowanych danych. Przebadany został również udział map głębi w całkowitej prędkości bitowej dla takich par QP-QD. Wyznaczone zostały wzory pozwalające na określenie wartości parametru QD, dla których uzyskuje się najlepszą jakość obrazów syntezowanych dla zadanej całkowitej prędkości bitowej przy ustalonej wartości parametru QP.

Wnioski z przedstawionych badań można podsumować następująco.

1. Przedstawione w rozdziale wyniki pozwalają stwierdzić, że uzyskanie maksymalnej możliwej dla zadanej prędkości bitowej jakości widoku syntezywanego w systemie wielowidokowym wymaga odpowiedniego dobrania parametrów kodowania. Uzyskanie takiej jakości możliwe jest jedynie poprzez odpowiednie dobranie proporcji pomiędzy prędkością bitową przeznaczaną na transmisję danych o mapach głębi i o obrazie. Przeprowadzone doświadczenia pokazują, że optymalny stosunek prędkości bitowych przeznaczanych na dwa wymienione typy danych dla pewnych wartości prędkości bitowych i pewnych scenariuszy (scenariusze 4, 5, 6, 8 – rysunki 5.65, 5.66, 5.67, 5.69) może osiągać wartość 1:1, co jest wynikiem zaskakującym. Według pierwotnych oczekiwań prędkość bitowa danych o mapie głębi powinna zajmować mniej niż połowę całkowitej prędkości bitowej. Warto podkreślić, że w większości punktów pomiarowych i dla pozostałych rozpatrzonych scenariuszy stosunek prędkości bitowej danych o mapie głębi do całkowitej prędkości bitowej w istocie oscyluje w granicach 20% - 30% i mniej, zatem jest zgodny z oczekiwaniami.
2. Istotnym wnioskiem wynikającym z przeprowadzonych badań jest z pewnością obserwacja, że możliwe jest wyznaczenie przybliżonej zależności optymalnej wartości indeksu QD w zależności od wartości indeksu QP (funkcji  $QD = f(QP)$ ). Podobna zależność była wyznaczona w punkcie 5.1.4.2, jednak warto podkreślić różnice w założeniach do obu badań.
  - a. W punkcie 5.1.4.2 rozważana była kompresja map głębi uzyskanych w pełni automatycznie, zatem charakteryzujących się niższą jakością. Założeniem przy wyznaczaniu wzoru 5.2 było uzyskanie krzywej uniwersalnej dla przebadanych sekwencji, która pozwala wyznaczyć pary QP-QD dające widok syntetyczny o jakości, która jest maksymalnie o 0,5 dB niższa od wartości maksymalnej dla danej wartości parametru QP. Nie była tu brana pod uwagę najlepsza możliwa do uzyskania dla danej prędkości bitowej jakość, lecz jedynie jakość dla konkretnej wartości QP.
  - b. W punkcie 5.2.2.3 rozważana jest kompresja map głębi wyznaczonych w sposób półautomatyczny, a zatem korzystający z wstępnych informacji dostarczonych przez operatora. Wyznaczone w ten sposób mapy głębi odznaczają się zdecydowanie wyższą jakością obrazu syntezywanego. Porównanie wykresów 5.5 z 5.29 oraz 5.6 z 5.31 pokazuje, że odpowiednio dla sekwencji Book Arrival różnica jakości

wynikająca z zastosowania nowych map głębi wynosi około 1 dB dla krzywej, dla której  $QP = 21$ , a dla sekwencji Newspaper – około 4 dB dla krzywej dla tego samego  $QP = 21$ . Dodatkowo, w punkcie 5.2.2.3 brana była pod uwagę maksymalna możliwa do uzyskania jakość widoku syntetycznego dla danej całkowitej prędkości bitowej.

3. Opracowane zostały wzory (5.4) i (5.5) służące do wyznaczania wartości parametru QD na podstawie wartości parametru QP. Wybór jednego z dwóch podanych wzorów dokonywany jest zależnie od przyjętego sposobu wyznaczania miary jakości PSNR.
4. Ważnym wynikiem badań jest ukazanie możliwości wyznaczenia dokładnej, odpowiedniej dla danej sekwencji, zależności  $QD = f(QP)$  na podstawie krótkiego fragmentu sekwencji i zastosowanie uzyskanych w ten sposób wyników dla całej sekwencji.

### 5.3. Podsumowanie

W rozdziale 5 przedstawione zostały wyniki badań nad oddzielną kompresją sekwencji wielowidokowych z głębią, które można podsumować w poniższy sposób.

1. Przeprowadzone zostało porównanie efektywności kompresji obrazów i map głębi z użyciem kodera AVC. Kodowanie map głębi jest bardziej efektywne (osiąga się większe współczynniki kompresji) niż obrazów.
2. Przebadany został wpływ zniekształceń powstających w procesie kodowania sekwencji na jakość widoków syntezowanych z użyciem dekompresowanych danych.
3. Opracowany został wzór (5.2) pozwalający na rozgraniczenie zakresów wartości parametru QP, dla których jakość widoków syntezowanych jest silnie wrażliwa lub mało wrażliwa na zmiany wartości parametru QD.
4. Przebadana została efektywność kodowania sekwencji wielowidokowych w przypadku użycia zdecydowanych map głębi. Uzyskane wyniki świadczą o użyteczności decymacji map głębi w celu zwiększenia efektywności kodowania sekwencji wielowidokowych z głębią.
5. Przebadane zostały metody automatycznego doboru wartości parametru QD na podstawie parametru QP. Opracowane zostały wzory (5.4) i (5.5) opisujące zależność  $QD = f(QP)$ . Wybór jednego z dwóch wzorów uzależniony jest od przyjętego sposobu oceny jakości. W przypadku gdy jakość sekwencji mierzona jest przez porównanie widoku syntezowanego z widokiem syntezowanym z nieskompresowanych danych, należy użyć wzoru (5.4), a gdy porównanie jest dokonywane z widokiem z rzeczywistej kamery, należy użyć wzoru (5.5).





## 6. Łączne kodowanie obrazu i głębi

### 6.1. Wprowadzenie

Oprócz wymienionych już we wcześniejszych rozdziałach właściwości obrazów map głębi, obrazy te mają również inne cechy, które mogą być pomocne przy zwiększaniu efektywności kompresji sekwencji wielowidokowych. Bardzo istotną cechą jest to, że pomiędzy informacjami, jakie zawarte są w obrazie i mapie głębi dostrzegalne są pewne podobieństwa co do ogólnych kształtów obiektów, czyli zachodzi korelacja pomiędzy mapami głębi i obrazami. W istocie bowiem obraz i odpowiadająca mu mapa głębi prezentują tę samą scenę, rejestrowaną z tego samego punktu obserwacji, a jedynie zawierają informacje o różnych właściwościach sceny. W związku z tym, dla prawidłowo wyznaczonych map głębi, położenie obiektów na obrazie i na mapie głębi jest zawsze takie samo. Wobec tego krawędzie obiektów widocznych na obrazie i na mapie głębi powinny być położone w tych samych miejscach. Dodatkowo można też oczekiwać zauważalnej korelacji pomiędzy wektorami ruchu wyliczonymi przez koder przy kodowaniu obrazów i map głębi. Ta ostatnia cecha wynika z tego, że obiekty, które można zidentyfikować zarówno na obrazie jak i na mapie głębi, poruszają się dokładnie w ten sam sposób. Wobec tego, algorytm sterowania w koderze powinien wybrać podobny wektor ruchu bądź rozbieżności tak dla kodowanego obrazu jak i dla map głębi.

Wnioski płynące z powyższej analizy jak też z przeprowadzonych badań pozwalają sądzić, że efektywnym rozwiązaniem zagadnienia kompresji sekwencji z danymi o głębi z użyciem kodera MVC mogłoby okazać się zastosowanie pewnej formy kodowania łącznego obrazów i map głębi. Pozwoliłoby to uniknąć wielokrotnego przesyłania tych samych informacji zarówno w strumieniu zawierającym dane o obrazie jak i w strumieniu zawierającym dane o mapie głębi. Dodatkowo, dzięki wykorzystaniu informacji zawartych w jednym ze strumieni można usprawnić kodowanie informacji zawartych w drugim strumieniu poprzez wykorzystanie istniejących zależności i podobieństw łączących dane o obrazie i o głębi sceny.

### 6.2. Przegląd literatury

W literaturze opisanych jest wiele metod wspólnej kompresji obrazów i map głębi. Część opracowanych metod charakteryzuje się osiąganiem większej efektywności kodowania, natomiast inne umożliwiają redukcję złożoności obliczeniowej procesu kodowania. Przegląd literatury dotyczącej zagadnienia wspólnego kodowania sekwencji obrazów i map głębi przedstawiony jest

poniżej.

Wykorzystanie do zwiększenia efektywności kodowania sekwencji z głębią korelacji pomiędzy wektorami ruchu dla obrazu i głębi jest jedną z opisywanych w literaturze możliwości. W pracy [Grew\_04] wykorzystano ten mechanizm do kodowania sekwencji z głębią za pomocą kodera MPEG-2. Metoda wykorzystująca w kompresji z użyciem AVC wspólne wektory ruchu dla obrazu i mapy głębi opisana jest w pracy [Dari\_09]. W pracy [Oh\_06] autorzy opisują algorytm wykorzystujący wektory ruchu sekwencji obrazów do kodowania sekwencji map głębi, skupiając się na uzyskanej redukcji złożoności procesu kodowania. Również na redukcji złożoności procesu kompresji sekwencji wielowidokowych skupiają się autorzy pracy [Lin\_11]. Redukcję złożoności uzyskują dzięki usprawnieniu procesu wyboru trybów kodowania obrazu na podstawie analizy mapy głębi. W pracach [Tao\_09][Zhang\_10a] autorzy opisują z kolei propozycję wykorzystania mechanizmu kodera skalowalnego do kodowania wektorów ruchu dla map głębi (warstwy rozszerzającej) na podstawie wektorów ruchu obrazu (warstwy podstawowej). Wspólne wykorzystanie wektorów ruchu jest mechanizmem wykorzystywanym także w pracy [Seo\_10].

Inną możliwością jest zastosowanie złożonej predykcji wektorów ruchu dla makrobloków na podstawie wektorów ruchu z sąsiednich widoków. Metody wykorzystania map głębi do predykcji wektorów ruchu dla kodowanego obrazu opisane są w pracach [Konie\_10a][Konie\_10b][Konie\_11].

Innym podejściem do zagadnienia wspólnej kompresji jest wykorzystanie w procesie kompresji syntezy obrazów. Badania wykorzystujące obraz syntezy w procesie predykcji należą do często opisywanych w literaturze. Użycie w trakcie kompresji algorytmów syntezy widok aktualnie kodowany i wykorzystanie go w procesie złożonej predykcji aktualnie kodowanego obrazu opisane jest z kolei w pracy [Doma\_12]. Podobne podejścia opisane są również w pracach [Yea\_07][Yama\_07][Shimi\_07][Morv\_07][Morv\_07b][Ekme\_07], jednak ich autorzy nie badają problemu efektywnej transmisji map głębi, niezbędnych w opisywanym zastosowaniu. Użycie podobnego mechanizmu do podniesienia efektywności kompresji zarówno obrazów jak i map głębi opisane jest w pracy [Na\_08].

Podobny sposób wykorzystania map głębi proponują autorzy pracy [Zhang\_10], w której opisany jest również sposób kompresji mapy głębi wykorzystujący predykcję i kodowanie słownikowe.

Rodzajem wspólnego kodowania obrazów i głębi jest również metoda opisana w [Mait\_10][Mait\_08], w której wykorzystana jest znajomość położenia konturów w obrazach do optymalizacji procesu kodowania mapy głębi poprzez odpowiednie dopasowanie wykorzystanej transformaty.

### 6.3. Kompresja głębi jako czwartej składowej obrazu

Inną możliwością wspólniej kompresji sekwencji wielowidokowych z głębią, nie opisywaną dotychczas w literaturze znanej autorowi rozprawy, jest przesyłanie mapy głębi jako dodatkowej, czwartej składowej D, przesyłanej wraz z trzema składowymi obrazu: Y, C<sub>B</sub> oraz C<sub>R</sub>.

Na podstawie przedstawionych w wcześniejszych rozdziałach wyników można wnioskować, że dla większości przypadków wierne zachowanie map głębi nie ma tak dużego znaczenia dla uzyskiwania wysokiej jakości syntezowanych widoków jak jakość obrazów. Można zatem stwierdzić, iż mapę głębi można traktować jako dodatkowy komponent, będący obok luminancji i dwóch chrominancji, czwartą składową kodowanej sekwencji wielowidokowej. Składową taką można przesłać, na wzór chrominancji, jako kolejny komponent dodatkowy. W ten sposób można uniknąć konieczności przesyłania większości informacji sterujących, które musiałyby zostać umieszczone w strumieniu zakodowanej sekwencji map głębi w przypadku jej oddzielnego kodowania.

Celem badań opisanych w dalszej części rozdziału jest opracowanie i przebadanie metod wspólnej kompresji sekwencji obrazów i map głębi wykorzystujących reprezentację czterokomponentową w różnych konfiguracjach.

Przedstawione w tym rozdziale wyniki otrzymane zostały w trakcie badań stanowiących rozwinięcie prac przedstawionych w [Klim\_10].

### 6.4. Metodologia badań

W celu zbadania efektywności kodowania sekwencji wielowidokowych w reprezentacji czterokomponentowej z głębią jako czwartym komponentem, przygotowano modyfikacje referencyjnego oprogramowania do kodowania sekwencji wielowidokowych MVC [JMVC] w wersji 6. W wyniku wprowadzenia modyfikacji powstał oryginalny, autorski kodek (para koder i dekoder sekwencji) przeznaczony do kodowania sekwencji wielowidokowych z głębią w formie sekwencji czterokomponentowych. Wprowadzone zmiany były dość poważne i polegały na dodaniu do kodera i dekodera możliwości obsługi sekwencji czterokomponentowych oraz na dodaniu procedur kodowania danych zawartych w dodatkowym komponencie głębi. Na podstawie wcześniejszych badań poczynione zostało dodatkowe założenie: dla kodowania czterokomponentowego, dekompresowane obrazy z kamer (składowe Y, C<sub>B</sub>, C<sub>R</sub>) muszą być identyczne z obrazami dekompresowanymi w przypadku zakodowania samych obrazów z kamer z użyciem takich samych parametrów. Założenie takie wynika z opisywanego w rozdziale 5

dominującego wpływu jakości obrazów na jakość syntezy obrazów.

W badaniach nad łączną kompresją sekwencji obrazów i map głębi z wykorzystaniem reprezentacji czterokomponentowej większość uwagi skupiono na metodach kodowania wewnątrzobrazowego. Cechami kodowania wewnątrzobrazowego, sugerującymi skupienie się na tych metodach są: relatywnie mała złożoność obliczeniowa (kodowanie makrobloku w trybie wewnątrzobrazowym trwa wielokrotnie krócej niż kodowanie tego samego makrobloku w trybie międzyobrazowym) oraz prostota koncepcyjna i implementacyjna przy zachowaniu dużej wartości praktycznej jak też poznawczej. W dalszej części pracy podane są również uwagi na temat rozszerzenia proponowanej metody kompresji łącznej na kodowanie międzyobrazowe.

W badaniach rozważane było jedynie kodowanie wewnątrzobrazowe. Rozpatrzone zostały następujące sposoby kompresji sekwencji czterokomponentowych.

- Głębia jako składowa o zmniejszonej rozdzielczości, kompresowana wspólnie ze składowymi chrominancji (punkt 6.5). W tym przypadku mapa głębi była decymowana, aby uzyskać jej rozmiary równe połowie wysokości i szerokości obrazu luminancji – dokładnie takie, jakie są rozmiary chrominancji w wykorzystywanym w badaniach schemacie 4:2:0. Mapa głębi podlegała następnie takim samym operacjom jak chrominancje: dla każdego makrobloku wykorzystywany kierunek predykcji oraz dane o obecności próbek transformaty błędu predykcji były wspólne dla chrominancji i dla map głębi. Aby spełnić założenie o nie zmienionym obrazie z kamer po dekodowaniu, wartość mapy głębi nie miała wpływu na wybór kierunku predykcji.
- Głębia jako składowa o pełnej rozdzielczości, kompresowana wspólnie z luminancją (punkt 6.6.2). W tym przypadku mapa głębi była kodowana w oryginalnych rozmiarach. Jej kodowanie przebiegało natomiast w taki sam sposób jak luminancji. Takie same, wybrane dla kolejnych makrobloków, podziały makrobloku na bloki i kierunki predykcji były wykorzystywane do kodowania luminancji i mapy głębi. Dla mapy głębi przesyłane były oddzielne dane o niezerowych współczynnikach skwantowanych próbek transformaty błędu predykcji – oddzielne pole CBP dla głębi, nazwane przez autora rozprawy DCBP (od angielskiego *Depth Coded Block Pattern*). Zgodnie z założeniami, mapa głębi nie była brana pod uwagę w procesie wyboru kierunku predykcji ani podziału makrobloku na bloki.
- Głębia jako składowa o pełnej rozdzielczości, kompresowana z użyciem oddzielnych trybów, przy czym kodowanie trybów wykorzystuje korelację między głębią i luminancją (punkt 6.6.3). W tym przypadku mapa głębi również była kodowana w oryginalnych

rozmiarach. Kodowanie składowej głębi dla kolejnych makrobloków przebiegało na takich samych zasadach, jak kodowanie luminancji. Podział na bloki i kierunki predykcji w tych blokach wyznaczany był jednak indywidualnie dla mapy głębi, a jedynie kodowany w strumieniu wynikowym razem z tymi parametrami dla luminancji. Oddzielnie przesyłane było również pole DCBP, zawierające informacje o istnieniu niezerowych skwantowanych współczynników transformaty sygnału błędu predykcji dla głębi.

Szczegółowe informacje na temat wykorzystanych sekwencji i przebiegu eksperymentów zawarte są w Aneksie IV.

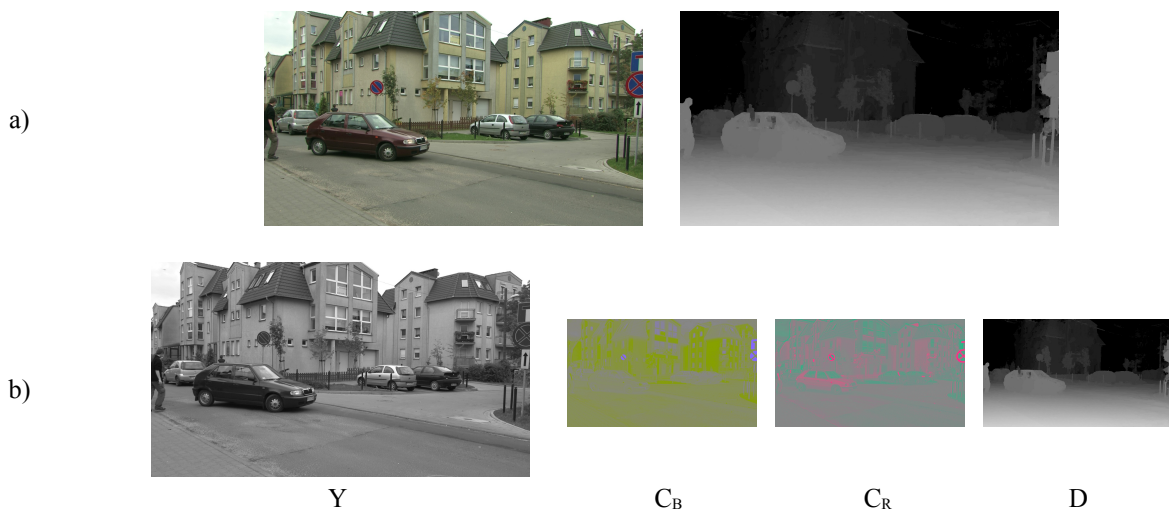
Prezentowane poniżej wyniki uporządkowane są w następujący sposób:

- w punkcie 6.5 opisana jest metoda, w której mapa głębi kodowana jest w zmniejszonej rozdzielczości,
- w punkcie 6.6 opisane są badania nad korelacją map głębi i luminancji oraz metody wykorzystujące kodowanie mapy głębi w pełnej rozdzielczości, w tym:
  - w punkcie 6.6.1 przedstawione są wyniki badań nad korelacją między trybami kodowania map głębi i luminancji,
  - w punkcie 6.6.2 przedstawione są wyniki dla metody, w której mapa głębi kodowana jest z wykorzystaniem takich samych parametrów jak luminancja w danym makrobloku,
  - w punkcie 6.6.3 przedstawione są wyniki dla metody, w której mapa głębi kodowana jest z wykorzystaniem oddzielnego zestawu parametrów.

## **6.5. Mapy głębi o zmniejszonej rozdzielczości**

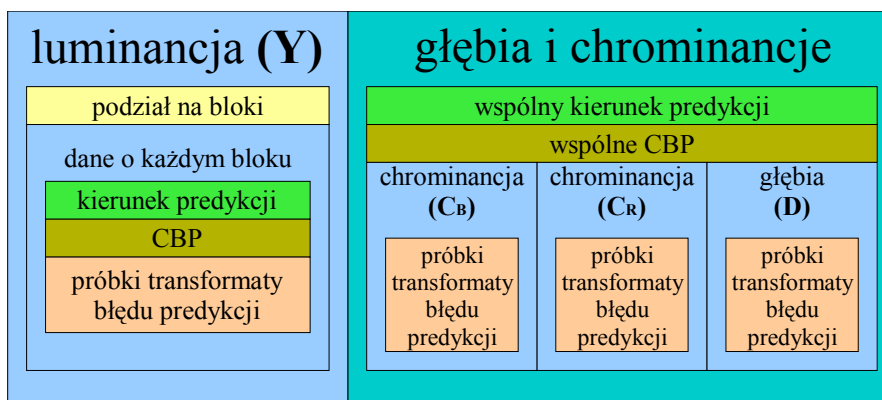
W pierwszym etapie badań wykorzystano reprezentację czterokomponentową, w której wykorzystywano obraz mapy głębi o zredukowanej rozdzielczości. Redukcja rozdzielczości dokonana została na podobnej zasadzie co dla chrominancji, poprzez decymację obrazu do połowy rozmiarów w obu kierunkach. Taka reprezentacja pozwala na wykorzystanie zjawiska opisanego w rozdziale 5.1.4.3, polegającego na poprawie efektywności kodowania sekwencji wielowidokowych w wyniku decymacji obrazów map głębi. Schematyczne przedstawienie pojedynczej, czterokomponentowej ramki sekwencji wielowidokowej widnieje na rysunku 6.1 poniżej. Warto w tym miejscu zwrócić uwagę na to, że w odróżnieniu od składowych chrominancji, w składowej głębi daje się zaobserwować dużą zawartość informacji – wyraźnie rozróżnialne są poszczególne

elementy sceny, szczególnie te leżące na pierwszym planie.



Rysunek 6.1. Pojedyncza ramka jednego widoku w sekwencji wielowidokowej z głębią; a) – obraz i głębia, b) – ta sama ramka po poddaniu dekompozycji na 4 komponenty: luminancji Y o pełnej rozdzielczości oraz zredukowanej rozdzielczości chrominancje C<sub>B</sub> i C<sub>R</sub> oraz głębię D.

Wstępne badania przeprowadzono dla scenariusza, w którym każdy makroblok kodowany był w sposób przedstawiony na rysunku 6.2. Blok luminancji kodowany był w niezmienny sposób, taki jak w koderze MVC, z wyznaczeniem sposobu podziału, kierunku predykcji i znaczników występowania niezerowych próbek transformaty błędu predykcji w blokach. Bloki chrominancji i blok mapy głębi kodowane były w ten sam sposób co standardowe kodowanie chrominancji w koderze MVC – wyznaczany był wspólny dla tych bloków kierunek predykcji i z jego użyciem wyznaczany był sygnał predykcji. W strumieniu umieszczane były zdecydowane próbki transformaty błędu predykcji.



Rysunek 6.2. Schematyczne przedstawienie sposobu kodowania pojedynczego makrobloku dla kodowania czterokomponentowego – mapa głębi o zmniejszonej rozdzielczości.

Niestety już wstępne badania wykazały, że wykorzystanie zdecydowanych obrazów map głębi w kodowaniu czterokomponentowym nie dość, że nie wykazuje opisanych wcześniej cech związanych z decymacją map głębi, to, co gorsza, prowadzi do zmniejszenia efektywności kodowania w porównaniu ze scenariuszem oddzielnego kodowania obrazów i map głębi.

Powody zmniejszenia efektywności kodowania można zawrzeć w następujących punktach.

1. W eksperymencie decymacja i kodowanie mapy głębi nie poprawia jakości syntezy obrazu. Powoduje natomiast usunięcie zbyt wielu szczegółów z pierwotnej mapy głębi, nie dając jednocześnie poprawy wynikającej z wygładzenia mapy głębi.
2. Kodowanie mapy głębi jako dodatkowej składowej, która nie jest brana pod uwagę w procesie sterowania i wyboru trybów, powoduje wybieranie nieoptymalnych z punktu widzenia kodowania mapy głębi trybów kodowania.
3. Predykcja składowej mapy głębi prowadzona jest w ten sam sposób jak predykcja chrominancji. Takie postępowanie daje gorsze efekty niż przy kodowaniu oddzielnym, ze względu na uboższą paletę dostępnych trybów predykcji w porównaniu z trybami dostępnymi dla składowej luminancji. O ile takie postępowanie ze składowymi chrominancjami nie powoduje widocznego obniżenia postrzeganej jakości obrazów, to nie sprawdza się dla mapy głębi.

## **6.6. Mapy głębi w pełnej rozdzielczości**

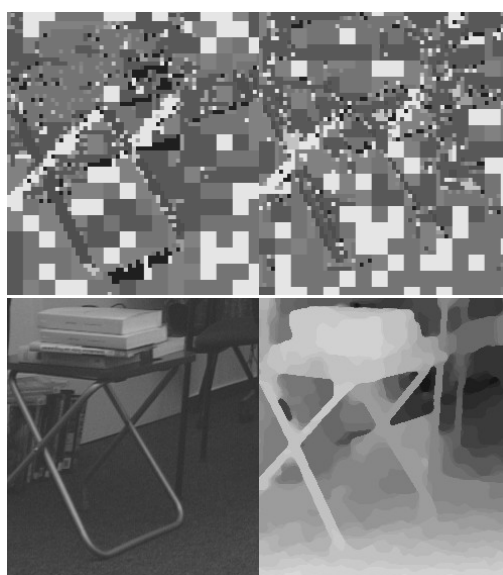
W związku z powyższymi wnioskami, kolejne próby były prowadzone z obrazami mapy głębi o pełnej rozdzielczości, tak, jak pokazane to zostało na rysunku 6.3. Zaletą takiego podejścia jest to, że możliwe jest zachowanie drobnych szczegółów w mapie głębi. Warto tu wyjaśnić, że wraz z postępem badań nad zagadnieniem wyznaczania map głębi dla sceny można mieć nadzieję na uzyskiwanie map głębi coraz wierniej oddających skomplikowaną trójwymiarową strukturę sceny. Wobec tego należy się liczyć z tym, że coraz ważniejsze może stawać się zachowanie mapy głębi w postaci jak najbliższej oryginałowi, bez utraty drobnych szczegółów, co wykluczać może zastosowanie zdecydowanych map głębi.



Rysunek 6.3. Pojedyncza ramka jednego widoku w sekwencji wielowidokowej z głębią. Ramka po poddaniu dekompozycji na 4 komponenty: luminancji Y i głębi D o pełnej rozdzielczości oraz zredukowanej rozdzielczości chrominancje  $C_B$  i  $C_R$ .

### 6.6.1. Korelacja dla wewnątrzobrazowego kodowania luminancji i głębi

W przypadku kodowania map głębi z tą samą rozdzielczością co luminancja, możliwe jest proste wykorzystanie podobieństw między luminancją i głębią. Istniejące podobieństwo można wykorzystać w celu zwiększenia efektywności wspólnej kompresji tych danych w porównaniu z oddzielną kompresją obrazów i map głębi. Na rysunku 6.4 pokazane zostały fragmenty jednego z obrazów sekwencji Book Arrival wraz z obrazem ilustrującym podział makrobloku na bloki i kierunki predykcji dla obrazu luminancji i dla mapy głębi. Różne odcienie szarości na ilustracji odpowiadają różnym wybranym kierunkom predykcji. Wielkość jednolitych obszarów świadczy o wybranym sposobie podziału makrobloku.



Rysunek 6.4. Tryby kodowania makrobloków dla niezależnego kodowania obrazu i mapy głębi. U góry – schemat podziałów i kierunków predykcji (różne odcienie szarości), na dole – odpowiadające fragmenty obrazu – po lewej – składowa luminancji, po prawej – składowa mapy głębi.

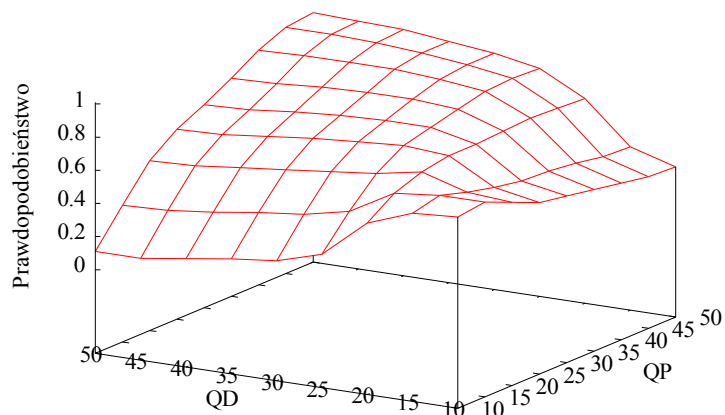


Na rysunku 6.4 można zaobserwować, że pomiędzy sposobami podziałów makrobloków i kierunkami predykcji dla składowych luminancji i głębi istnieje pewna korelacja. Jest to szczególnie widoczne w okolicach krawędzi obiektów. Co istotne – krawędzie te muszą być widoczne zarówno na obrazie mapy głębi jak i na obrazie luminancji. Warunek ten jest spełniony dla widocznego na rysunku taboretu – zarówno na obrazie przedstawiającym wybór trybów kodowania dla składowej luminancji jak i dla składowej głębi wyraźnie można zaobserwować ułożenie nóg taboretu. Rejony te kodowane są z użyciem tych samych kierunków predykcji użytych w blokach. Również sposób podziału na bloki cechuje się zauważalnym podobieństwem.

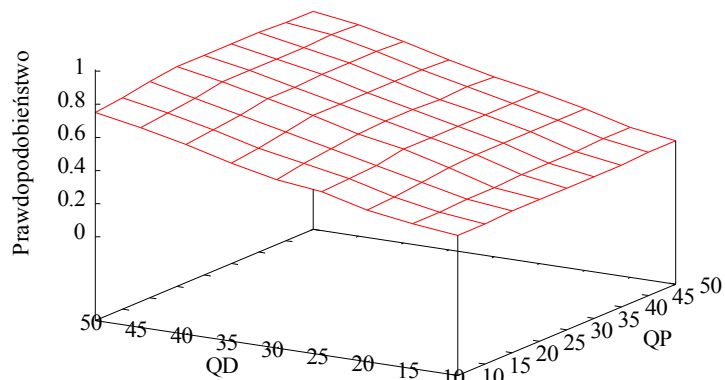
Przeprowadzone zostały zatem badania mające na celu określenie stopnia podobieństwa trybów kodowania wybieranych dla luminancji i głębi. Celem tych badań było ustalenie, czy możliwe jest wykorzystanie korelacji między danymi do efektywnego ich zakodowania w strumieniu.

### **6.6.1.1. Podziały makrobloku i kierunki predykcji dla luminancji i głębi**

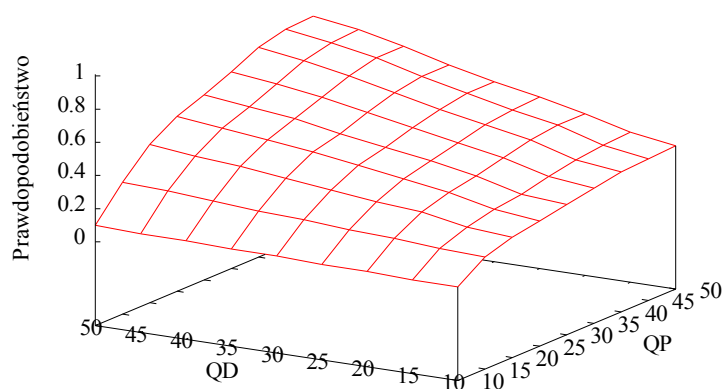
Na rysunkach poniżej zaprezentowane zostały wykresy prawdopodobieństw wyboru takiego samego sposobu podziału makrobloku dla kodowania luminancji i mapy głębi, takiego samego kierunku predykcji oraz prawdopodobieństwa, że dla luminancji i dla mapy głębi wartości te są takie same. W trakcie kodowania możliwy był wybór jednego z dwóch trybów podziału makrobloku – na 1 blok 16x16 punktów lub na 16 bloków o wymiarach 4x4 punkty. Dostępne były wszystkie możliwe kierunki predykcji (4 dla bloków 16x16 oraz 8 dla bloków 4x4 punkty). Przedstawione zostały wyniki dla przykładowych sekwencji: Book Arrival, Newspaper i Poznan Street. Wyniki uzyskane zostały na drodze oddzielnej, niezależnej analizy każdego bloku o rozmiarach 4x4 punkty, zatem nie biorą one pod uwagę przynależności danego bloku do określonego makrobloku.



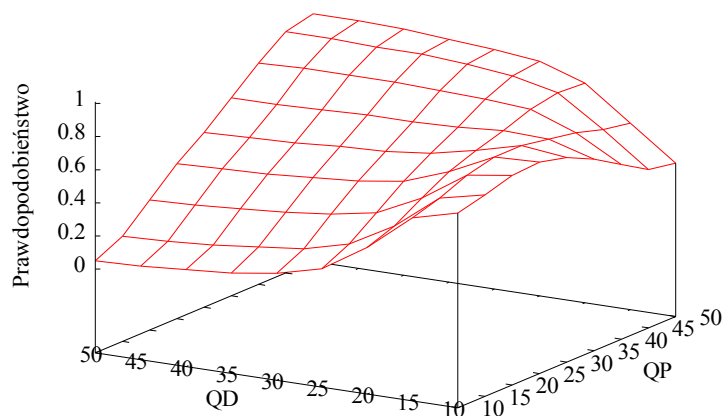
Rysunek 6.5. Sekwencja Book Arrival – prawdopodobieństwo wykorzystania tego samego trybu podziału makrobloku dla kodowania luminancji i mapy głębi w zależności od wartości parametru kwantyzacji dla luminancji QP i dla mapy głębi QD.



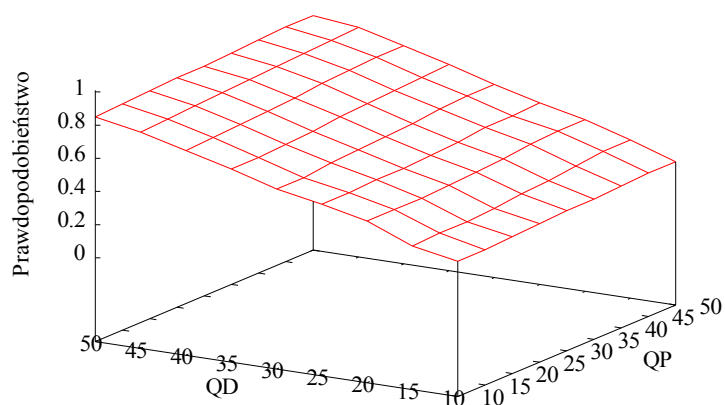
Rysunek 6.6. Sekwencja Book Arrival – prawdopodobieństwo wykorzystania tego samego kierunku predykcji dla kodowania luminancji i mapy głębi w zależności od wartości parametru kwantyzacji dla luminancji QP i dla mapy głębi QD.



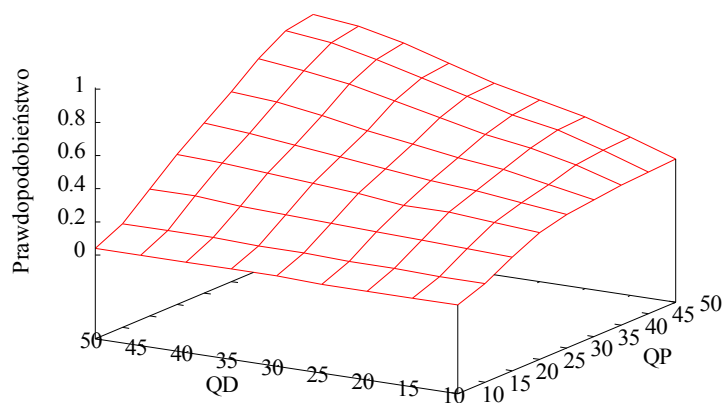
Rysunek 6.7. Sekwencja Book Arrival – prawdopodobieństwo wykorzystania tego samego trybu podziału makrobloku i tego samego kierunku predykcji dla kodowania luminancji i mapy głębi w zależności od wartości parametru kwantyzacji dla luminancji QP i dla mapy głębi QD.



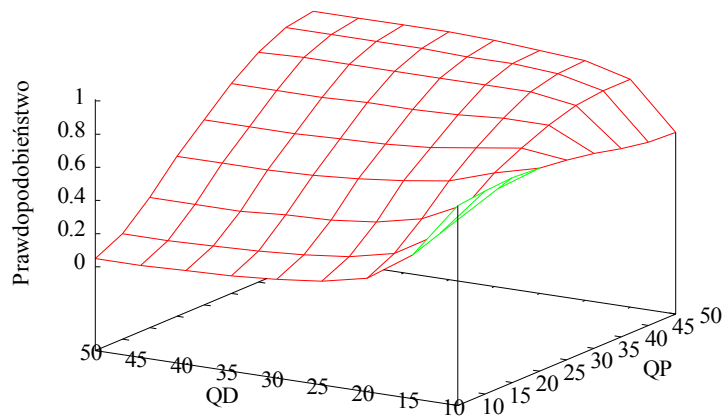
Rysunek 6.8. Sekwencja Newspaper – prawdopodobieństwo wykorzystania tego samego trybu podziału makrobloku dla kodowania luminancji i mapy głębi w zależności od wartości parametru kwantyzacji dla luminancji QP i dla mapy głębi QD.



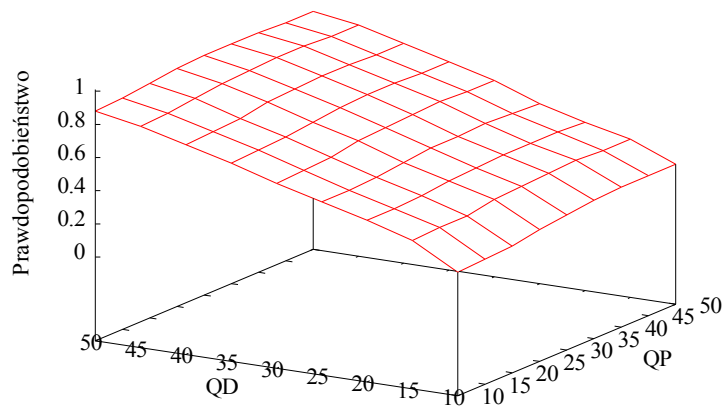
Rysunek 6.9. Sekwencja Newspaper – prawdopodobieństwo wykorzystania tego samego kierunku predykcji dla kodowania luminancji i mapy głębi w zależności od wartości parametru kwantyzacji dla luminancji QP i dla mapy głębi QD.



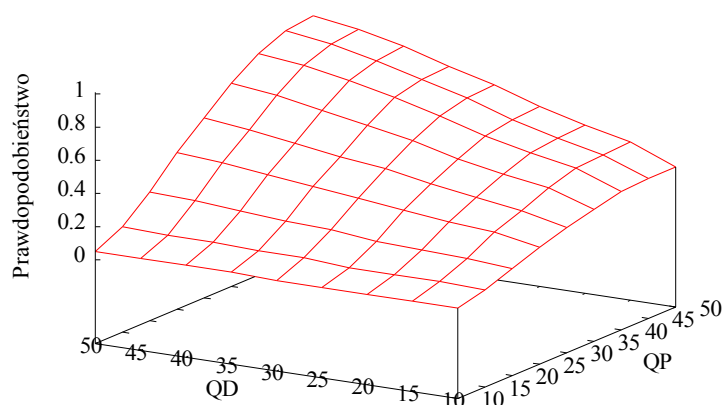
Rysunek 6.10. Sekwencja Newspaper – prawdopodobieństwo wykorzystania tego samego trybu podziału makrobloku i tego samego kierunku predykcji dla kodowania luminancji i mapy głębi w zależności od wartości parametru kwantyzacji dla luminancji QP i dla mapy głębi QD.



Rysunek 6.11. Sekwencja Poznan Street – prawdopodobieństwo wykorzystania tego samego trybu podziału makrobloku dla kodowania luminancji i mapy głębi w zależności od wartości parametru kwantyzacji dla luminancji QP i dla mapy głębi QD.



Rysunek 6.12. Sekwencja Poznan Street – prawdopodobieństwo wykorzystania tego samego kierunku predykcji dla kodowania luminancji i mapy głębi w zależności od wartości parametru kwantyzacji dla luminancji QP i dla mapy głębi QD.



Rysunek 6.13. Sekwencja Poznan Street – prawdopodobieństwo wykorzystania tego samego trybu podziału makrobloku i tego samego kierunku predykcji dla kodowania luminancji i mapy głębi w zależności od wartości parametru kwantyzacji dla luminancji QP i dla mapy głębi QD.

Na wykresach z rysunków 6.5 - 6.13 widać, że dla większych wartości parametrów QP i QD (wartości większe niż 35) występują bardzo duże wartości prawdopodobieństwa wyboru tego samego trybu podziału makrobloku i trybu predykcji. Prawdopodobieństwa te osiągają maksymalną wartość powyżej 0,9, niezależnie od badanej sekwencji. Pozwala to oczekiwać, że zastosowanie wspólnego kodowania informacji dotyczących sposobu podziału makrobloku oraz kierunku predykcji dla kodowania luminancji i mapy głębi przynieść może zwiększenie efektywności kompresji. Poprawa efektywności powinna być największa dla zakresu wysokich wartości parametrów QP i QD. Jest to korzystne, gdyż w tym zakresie dane dotyczące podziału makrobloku i kierunku predykcji zajmują większą część strumienia niż dla przypadków kodowania z niskimi wartościami QP i QD, można zatem liczyć na pewien zysk wynikający z zastosowania prezentowanej metody.

### 6.6.1.2. Podział na bloki dla luminancji i głębi dla kodowania wewnątrzobrazowego

Przeprowadzone zostały badania polegające na określeniu dokładnych zależności pomiędzy wyborem optymalnych trybów dla obrazów i głębi. Ich wyniki przedstawione są na kolejnych rysunkach. Zaprezentowane są jedynie wykresy dla wybranych par indeksów QP-QD, jednak wybór dokonany został tak, aby ukazać istotne zależności dostrzeżone w czasie prowadzenia badań.

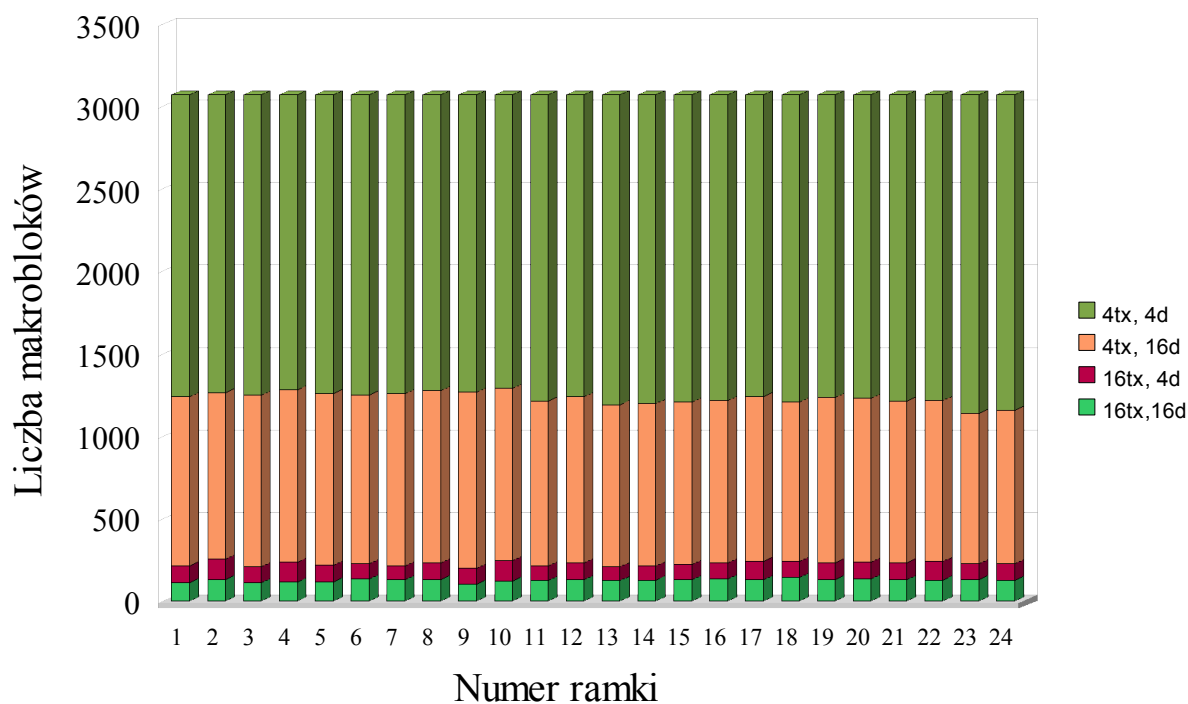
Warto podkreślić, że dla wszystkich przebadanych sekwencji wyniki prezentują się bardzo podobnie, co pozwala na wyciągnięcie ogólnych wniosków dotyczących wspólnej kompresji sekwencji obrazów i głębi.

Jako pierwsze, na rysunkach od 6.14 do 6.19 zostaną zaprezentowane wykresy obrazujące zależności pomiędzy wyborem dla danego makrobloku trybu podziału na bloki przy kodowaniu obrazu (tx) i głębi (d). Przebadane zostały oddzielnie przypadki dla włączonego i wyłączzonego trybu podziału na bloki 8x8 punktów. Przedstawione są wyniki dla 24 ramek, po 12 z każdego z dwóch kodowanych widoków. Ramki pierwszego widoku oznaczone są numerami od 1 do 12, a widoku drugiego – numerami od 13 do 24. W opisie wykresów numery odpowiadają rozmiarom bloków, i tak na przykład opis „**4tx, 16d**” oznacza makrobloki, których składowa luminancji kodowana była z użyciem podziału na bloki 4x4 punkty, a składowa głębi – w jednym bloku o rozmiarach 16x16 punktów.

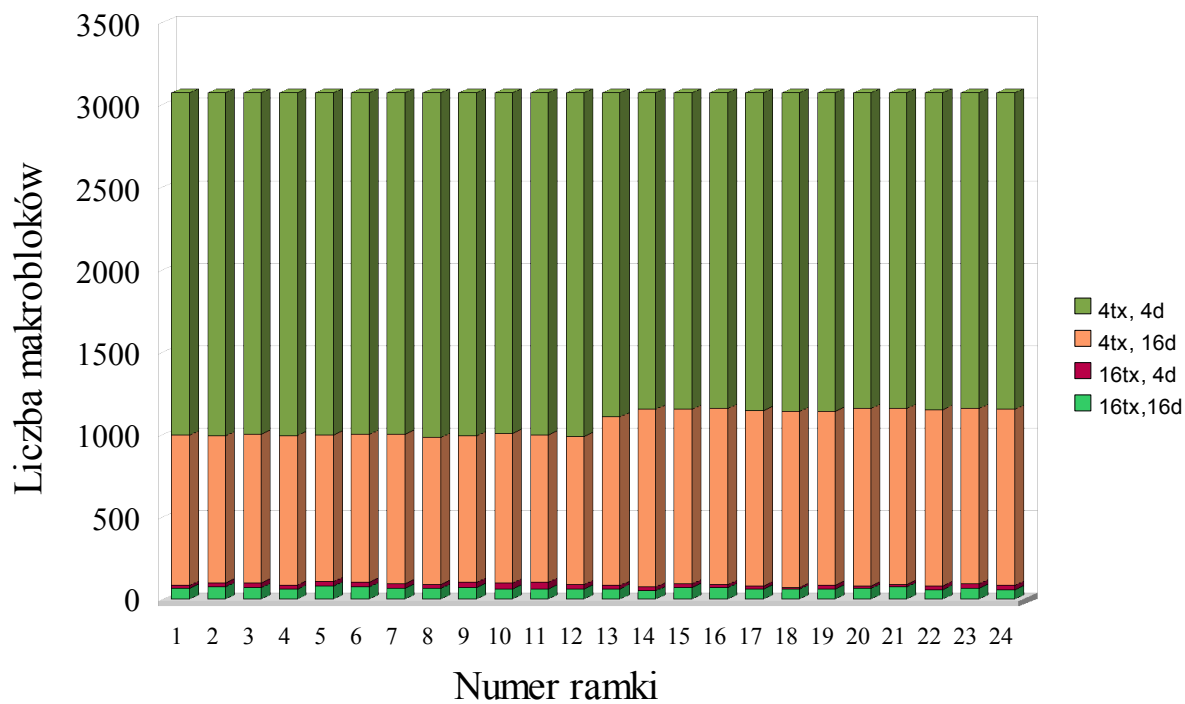
Na rysunkach od 6.14 do 6.19 przedstawione są zestawienia podziałów makrobloków dla kompresji luminancji i głębi dla niskiej wartości parametrów  $QP=10$  i  $QD=10$  dla trzech wybranych sekwencji testowych. Na rysunku 6.20 przedstawione jest zestawienie liczby makrobloków o takich samych podziałach dla luminancji i głębi dla przypadków z wyłączoną możliwością wyboru podziału na bloki o rozmiarach 8x8 punktów. Z kolei zestawienie z rysunku 6.21 sporządzone zostało dla przypadków, w których możliwe było dokonywanie podziału makrobloków na bloki 8x8 punktów.

Wyniki zaprezentowane na rysunkach od 6.14 do 6.19 świadczą o tym, że w kodowaniu z bardzo niskimi wartościami (w prezentowanym przykładzie użyto wartości 10) indeksów QP i QD najczęściej wykorzystywany jest podział na bloki 4x4 punkty, zarówno w kodowaniu głębi jak i obrazu. W przypadku głębi znacznie częściej niż w kodowaniu obrazu wykorzystywany jest tryb 16x16 punktów. Bardzo rzadko natomiast w kompresji głębi wykorzystywany jest podział na bloki 8x8 punktów.

Jak wynika z wykresów z rysunku 6.20, około 2/3 makrobloków przy kodowaniu z wykorzystaniem wartości  $QP = 10$  i  $QD = 10$  i wyłączonej możliwości wyboru podziału na bloki 8x8 kodowanych jest z tym samym podziałem dla składowej luminancji i składowej głębi. Wyniki z rysunku 6.21 świadczą natomiast o tym, że dla wartości  $QP = 10$  i  $QD = 10$  i włączonej możliwości wyboru podziału na bloki 8x8 punktów, około połowa makrobloków kodowana jest z takim samym podziałem dla luminancji i głębi.

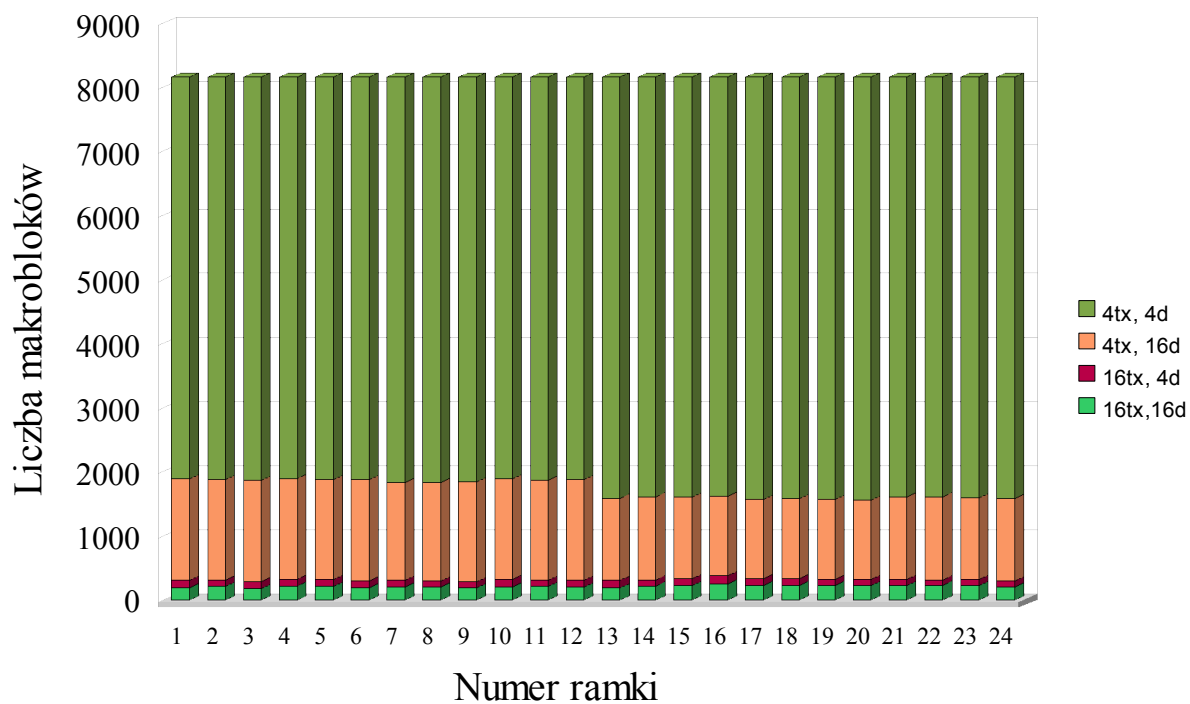


Rysunek 6.14. Rozmiary bloków dla kodowania sekwencji Book Arrival.  
 Parametry kodowania: QP = 10, QD = 10. Podział na bloki 8x8 punktów wyłączony.

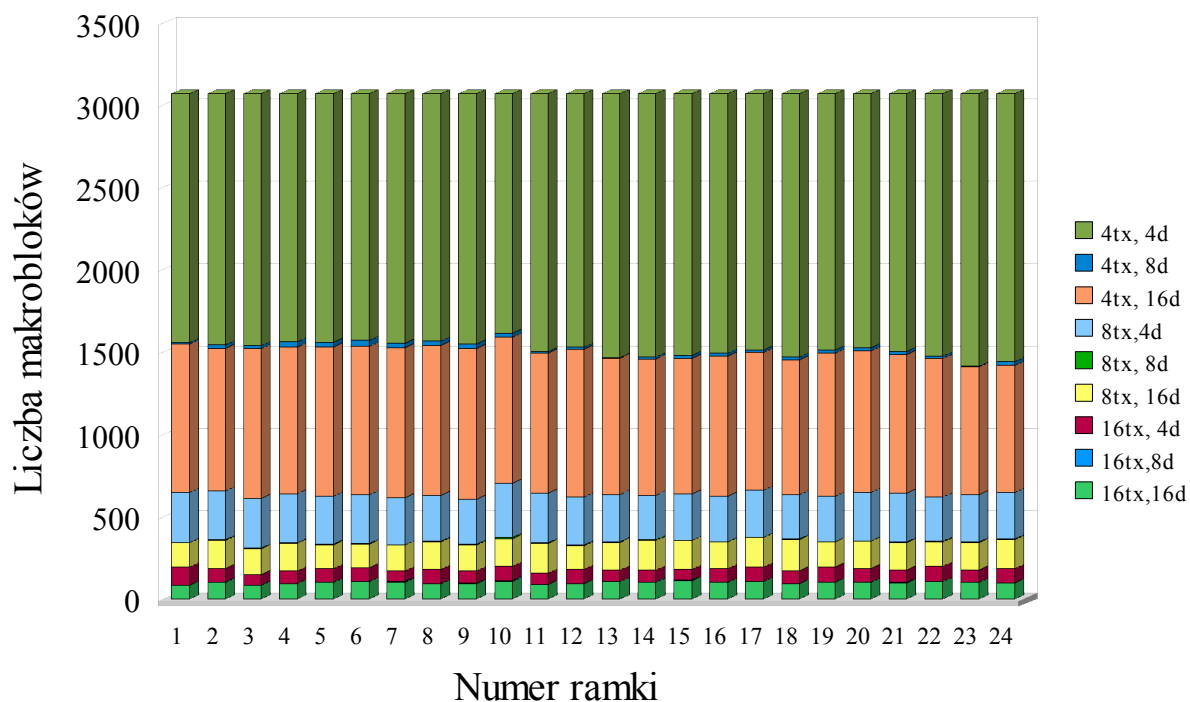


Rysunek 6.15. Rozmiary bloków dla kodowania sekwencji Newspaper.  
 Parametry kodowania: QP = 10, QD = 10. Podział na bloki 8x8 punktów wyłączony.

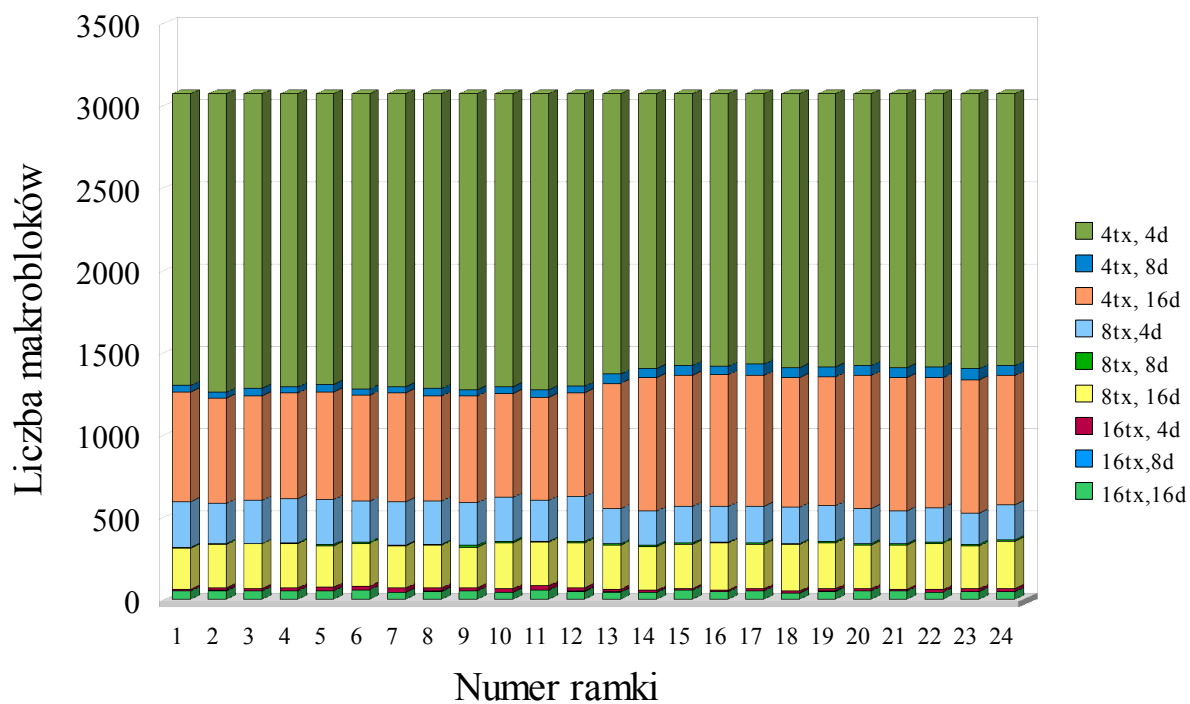




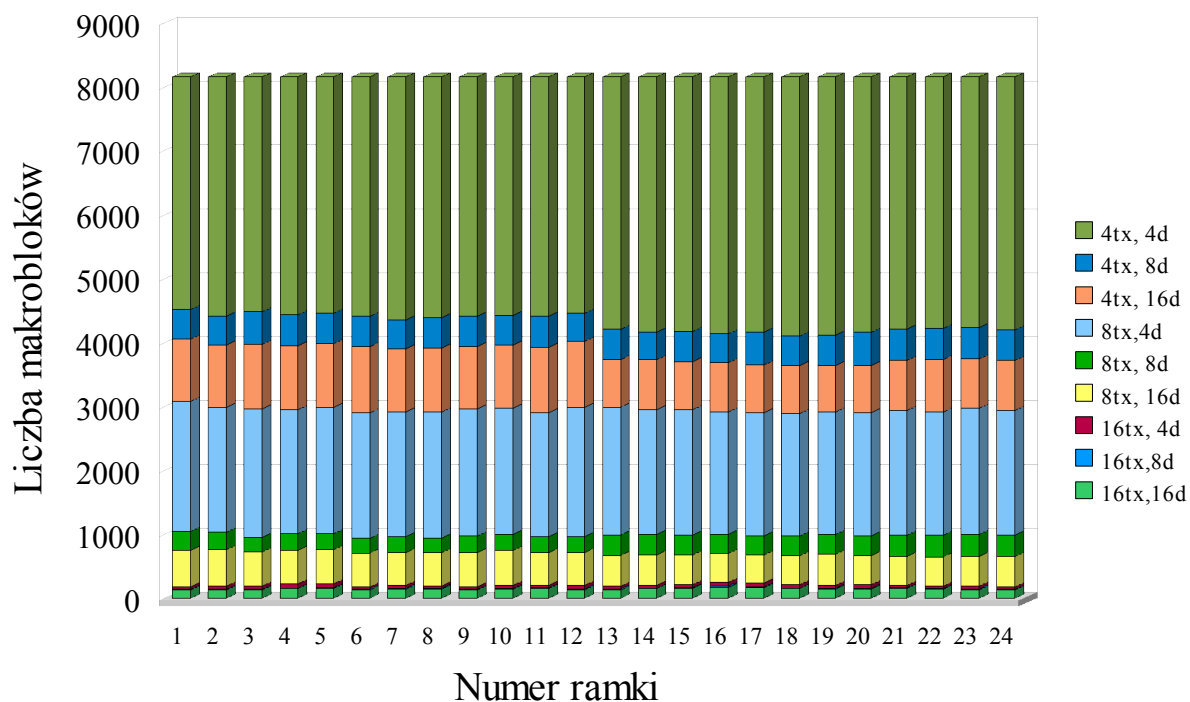
Rysunek 6.16. Rozmiary bloków dla kodowania sekwencji Poznan Street.  
 Parametry kodowania: QP = 10, QD = 10. Podział na bloki 8x8 punktów wyłączony.



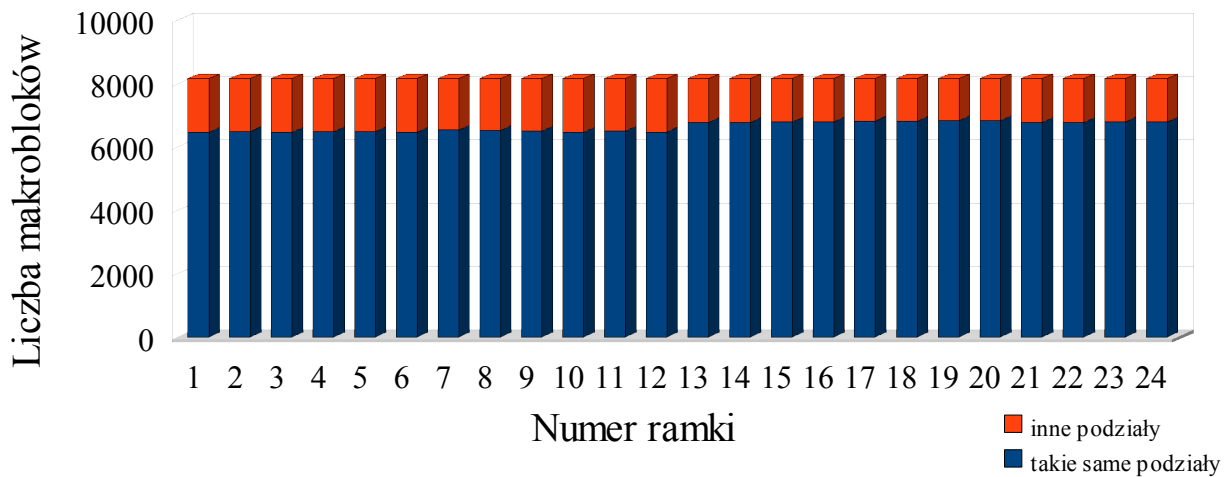
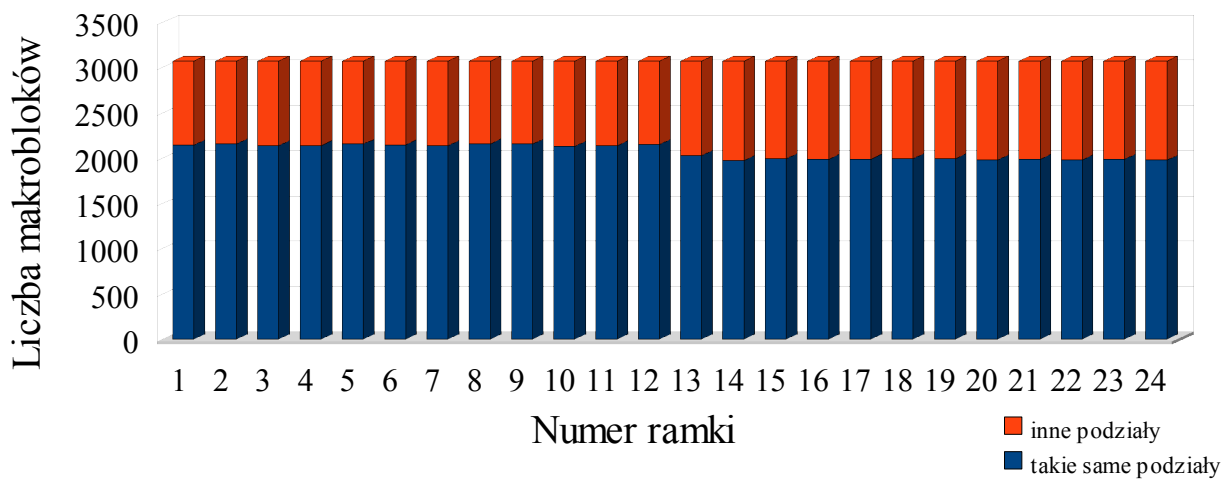
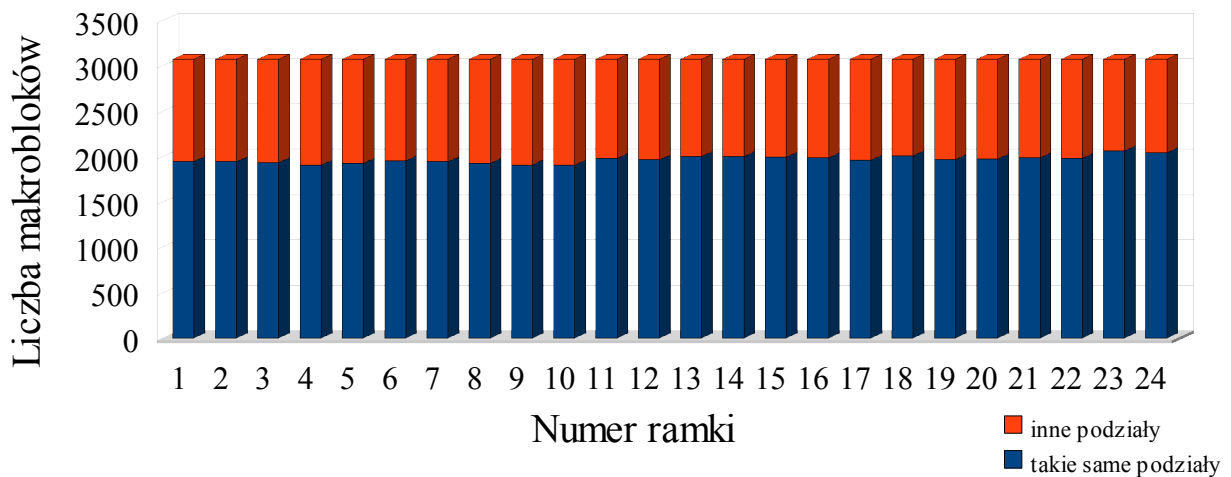
Rysunek 6.17. Rozmiary bloków dla kodowania sekwencji Book Arrival.  
 Parametry kodowania: QP = 10, QD = 10. Podział na bloki 8x8 punktów włączony.



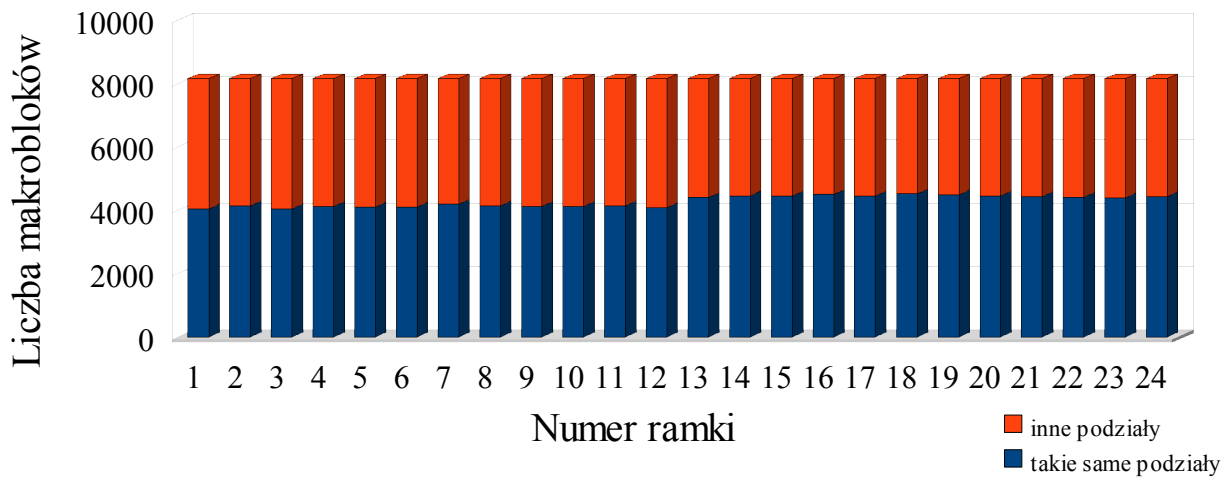
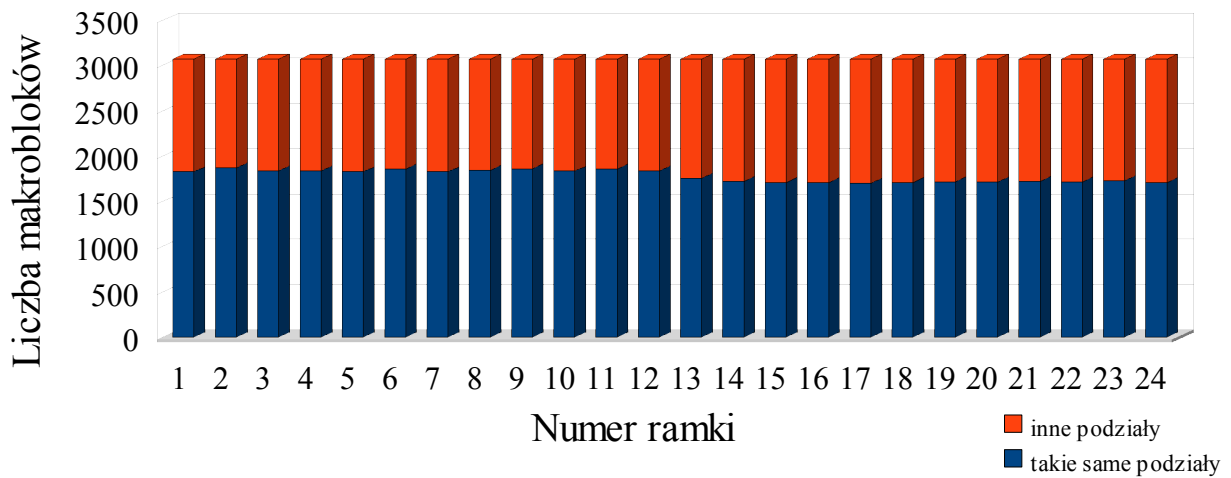
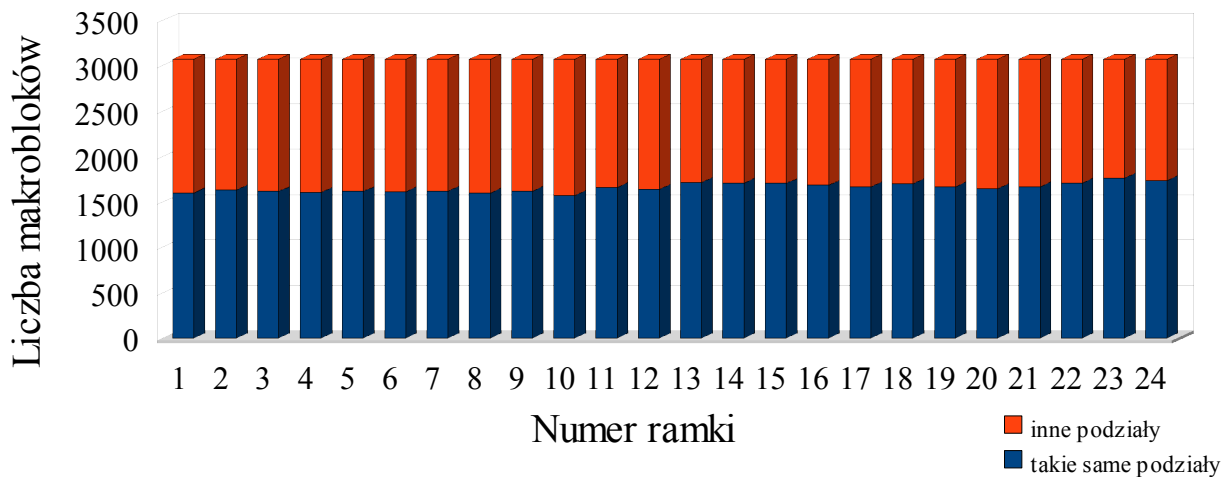
Rysunek 6.18. Rozmiary bloków dla kodowania sekwencji Newspaper.  
 Parametry kodowania: QP = 10, QD = 10. Podział na bloki 8x8 punktów włączony.



Rysunek 6.19. Rozmiary bloków dla kodowania sekwencji Poznan Street.  
 Parametry kodowania: QP = 10, QD = 10. Podział na bloki 8x8 punktów włączony.



Rysunek 6.20. Wykorzystane przy kodowaniu makrobloków podziały dla sekwencji (od góry do dołu): Book Arrival, Newspaper, Poznan Street. Parametry kodowania: QP = 10, QD = 10. Podział na bloki 8x8 punktów wyłączony.



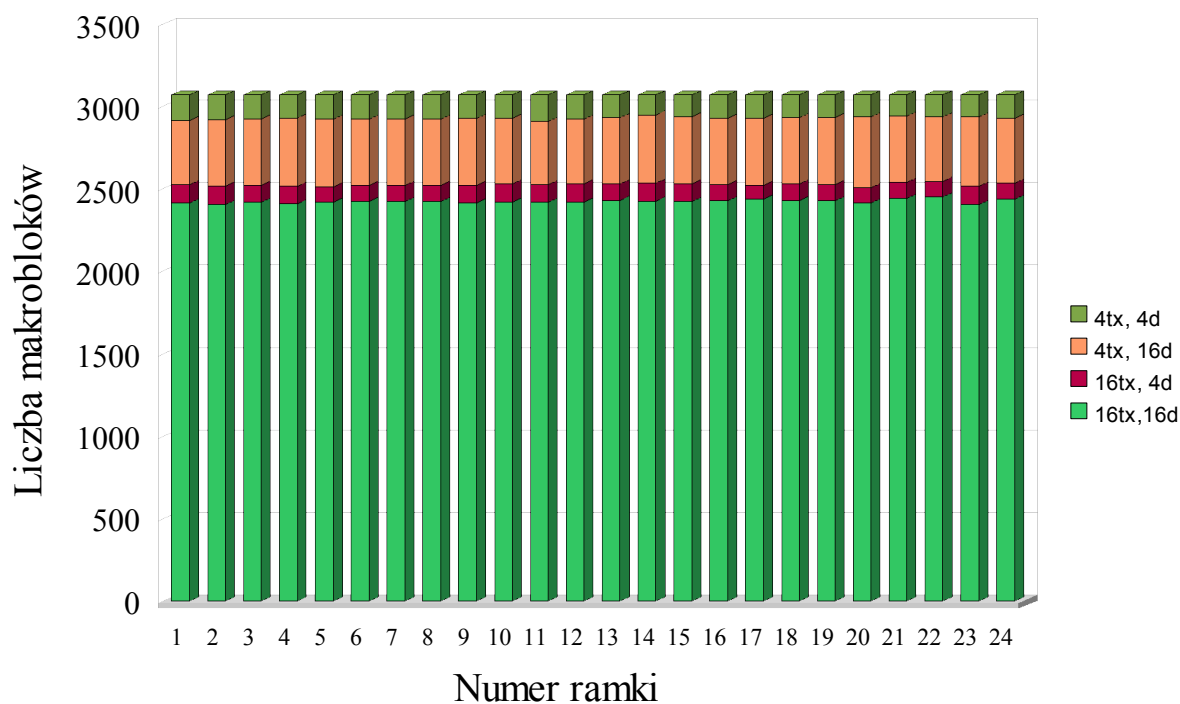
Rysunek 6.21. Wykorzystane przy kodowaniu makrobloków podziały dla sekwencji (od góry do dołu): Book Arrival, Newspaper, Poznan Street. Parametry kodowania: QP = 10, QD = 10. Podział na bloki 8x8 punktów włączony.

Na rysunkach od 6.22 do 6.27 przedstawione są zestawienia podziałów makrobloków dla kompresji luminancji i głębi dla wysokiej wartości parametrów  $QP=40$  i  $QD=40$  dla trzech wybranych sekwencji testowych. Na rysunku 6.28 przedstawione jest zestawienie liczby makrobloków o takich samych podziałach dla luminancji i głębi dla przypadków z wyłączoną możliwością wyboru podziału na bloki o rozmiarach  $8 \times 8$  punktów dla kompresji z parametrami  $QP=40$  i  $QD=40$ . Z kolei zestawienie z rysunku 6.29 sporządzone zostało dla przypadków, w których możliwe było dokonywanie podziału makrobloków na bloki  $8 \times 8$  punktów.

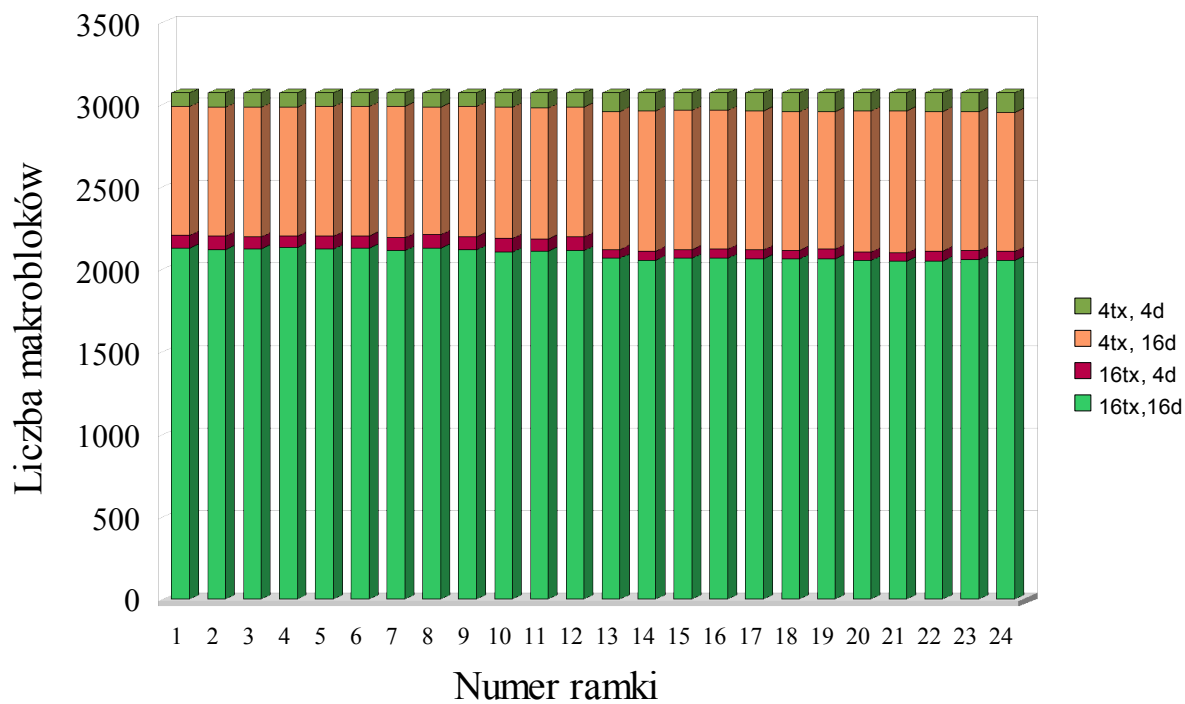
Wyniki zaprezentowane na rysunkach od 6.22 do 6.27 świadczą o tym, że w kodowaniu z bardzo wysokimi wartościami (dla przedstawionych przykładowych wyników użyto wartości 40) indeksów  $QP$  i  $QD$  najczęściej wykorzystywany jest podział na bloki  $16 \times 16$  punktów, zarówno w kodowaniu głębi jak i obrazu. W przypadku kodowania głębi znacznie częściej niż w kodowaniu obrazu wykorzystywany jest tryb  $16 \times 16$  punktów. Bardzo rzadko natomiast w kompresji głębi wykorzystywany jest podział na bloki  $8 \times 8$  punktów. Dwie ostatnie obserwacje są identyczne jak dla przypadku pokazanego wcześniej, dla którego  $QP = 10$  i  $QD = 10$ , można zatem założyć, że podobne zachowanie obserwowane będzie dla większości przypadków, dla których wartości  $QP$  i  $QD$  są takie same.

Wyniki z rysunku 6.28 świadczą o tym, że około  $3/4$  makrobloków przy kodowaniu z wykorzystaniem wartości  $QP = 40$  i  $QD = 40$  i wyłączonej możliwości wyboru podziału na bloki  $8 \times 8$  kodowanych jest z tym samym podziałem na bloki dla składowej luminancji i składowej głębi. Wyniki z rysunku 6.29 świadczą natomiast o tym, że dla wartości  $QP = 40$  i  $QD = 40$  i włączonej możliwości wyboru podziału na bloki  $8 \times 8$  punktów, około  $2/3$  makrobloków kodowanych jest z takim samym podziałem dla luminancji i głębi. Widoczny jest zatem wyraźny wzrost liczby makrobloków korzystających z tego samego sposobu podziału bloków luminancji i mapy głębi wobec przypadku  $QP = 10$ ,  $QD = 10$  (rysunki 6.20 oraz 6.21).

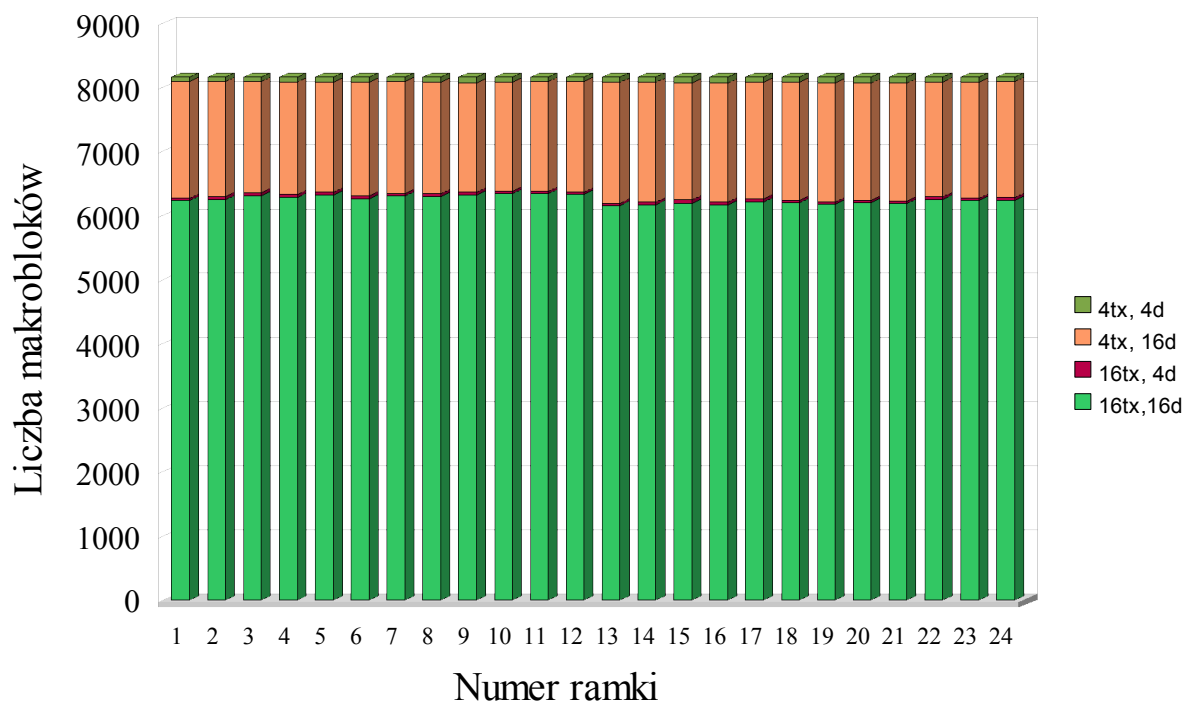
Jak widać z wykresów z rysunków od 6.22 do 6.27, w większości przypadków makrobloki o takim samym podziale dla składowej luminancji i głębi to makrobloki wykorzystujące kodowanie obydwu składowych w blokach  $16 \times 16$  punktów.



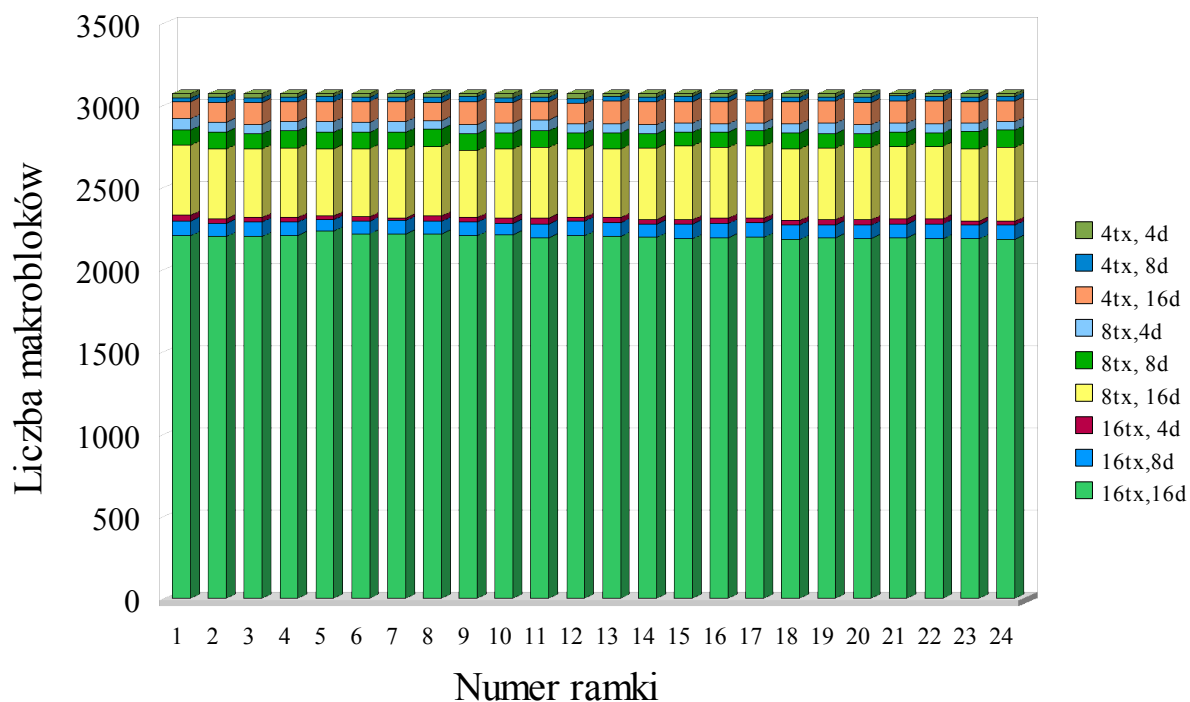
Rysunek 6.22. Rozmiary bloków dla kodowania sekwencji Book Arrival.  
 Parametry kodowania: QP = 40, QD = 40. Podział na bloki 8x8 punktów wyłączony.



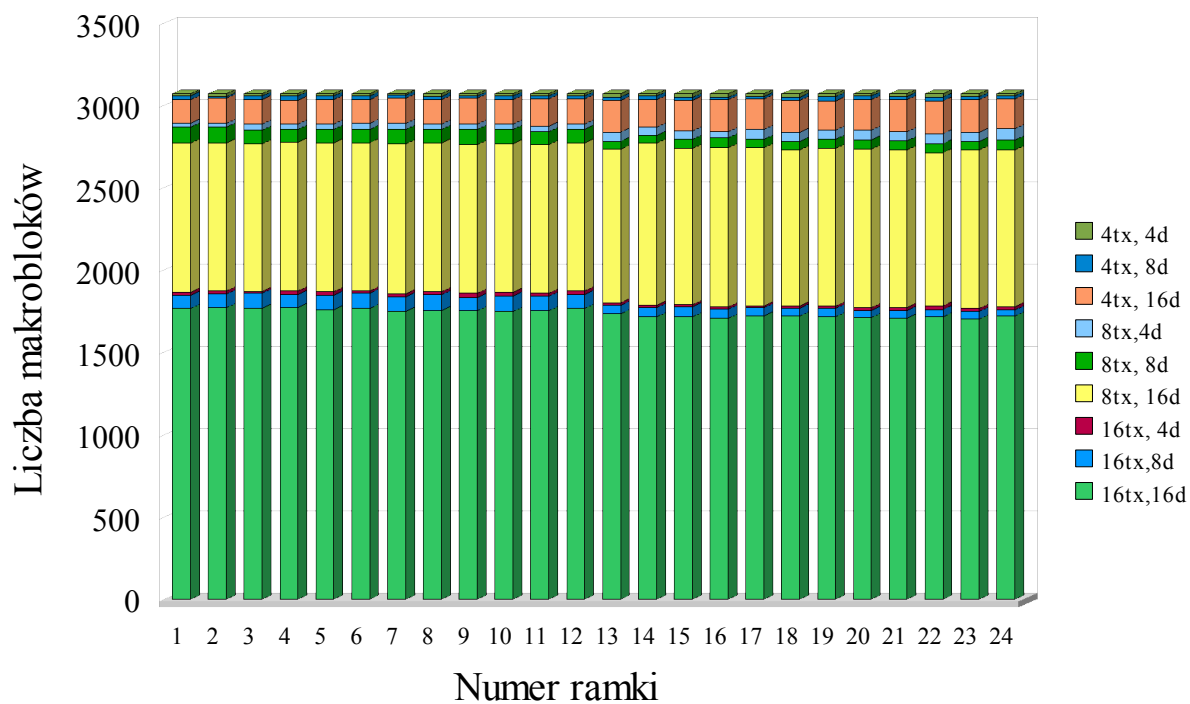
Rysunek 6.23. Rozmiary bloków dla kodowania sekwencji Newspaper.  
 Parametry kodowania: QP = 40, QD = 40. Podział na bloki 8x8 punktów wyłączony.



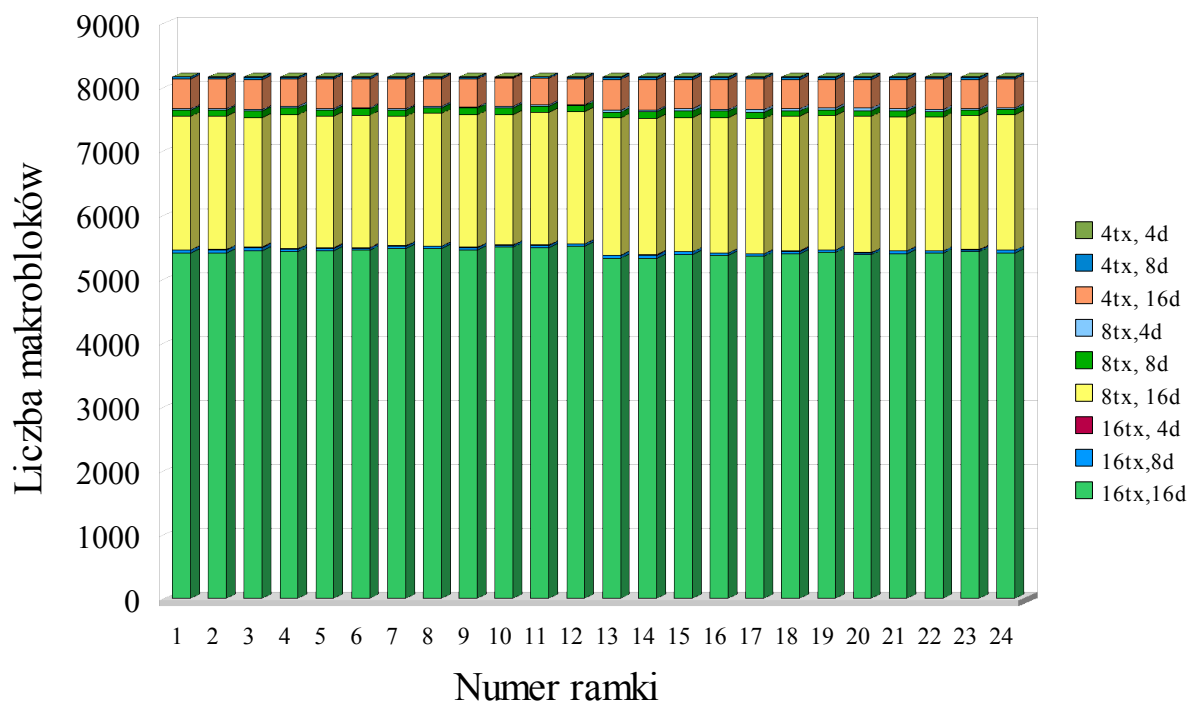
Rysunek 6.24. Rozmiary bloków dla kodowania sekwencji Poznan Street.  
 Parametry kodowania: QP = 40, QD = 40. Podział na bloki 8x8 punktów wyłączony.



Rysunek 6.25. Rozmiary bloków dla kodowania sekwencji Book Arrival.  
 Parametry kodowania: QP = 40, QD = 40. Podział na bloki 8x8 punktów włączony.

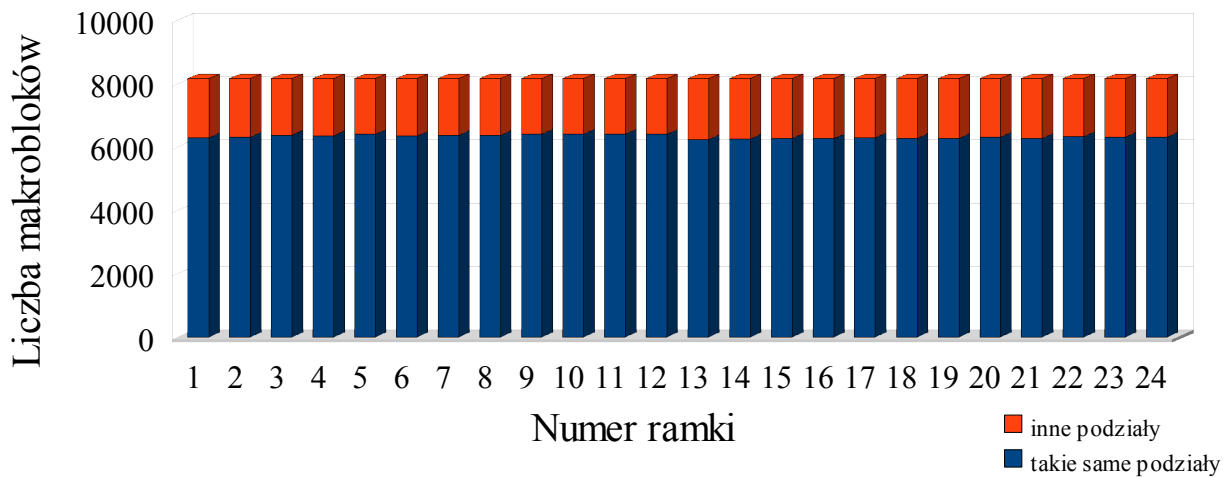
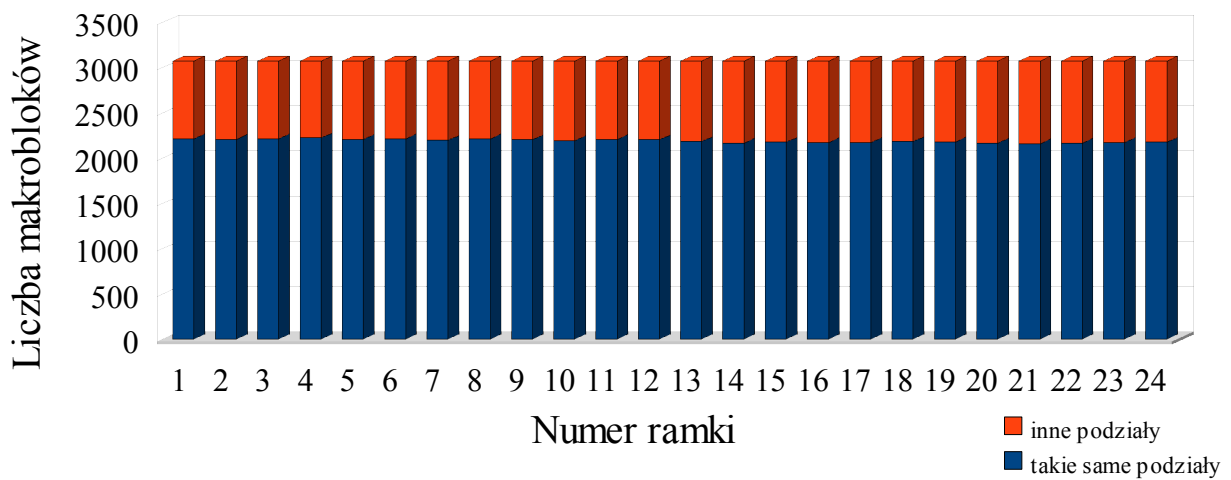
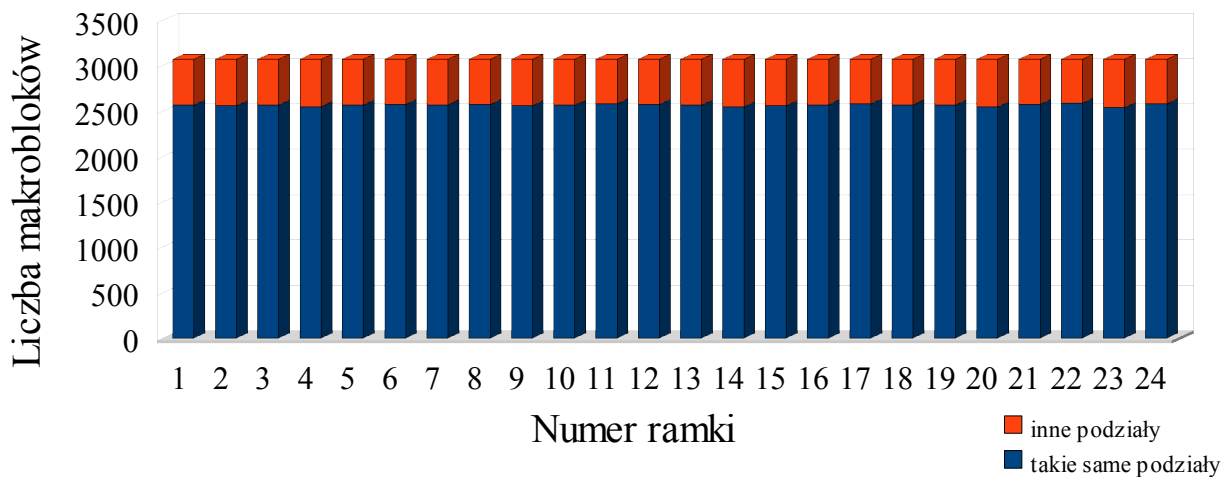


Rysunek 6.26. Rozmiary bloków dla kodowania sekwencji Newspaper.  
 Parametry kodowania: QP = 40, QD = 40. Podział na bloki 8x8 punktów włączony.

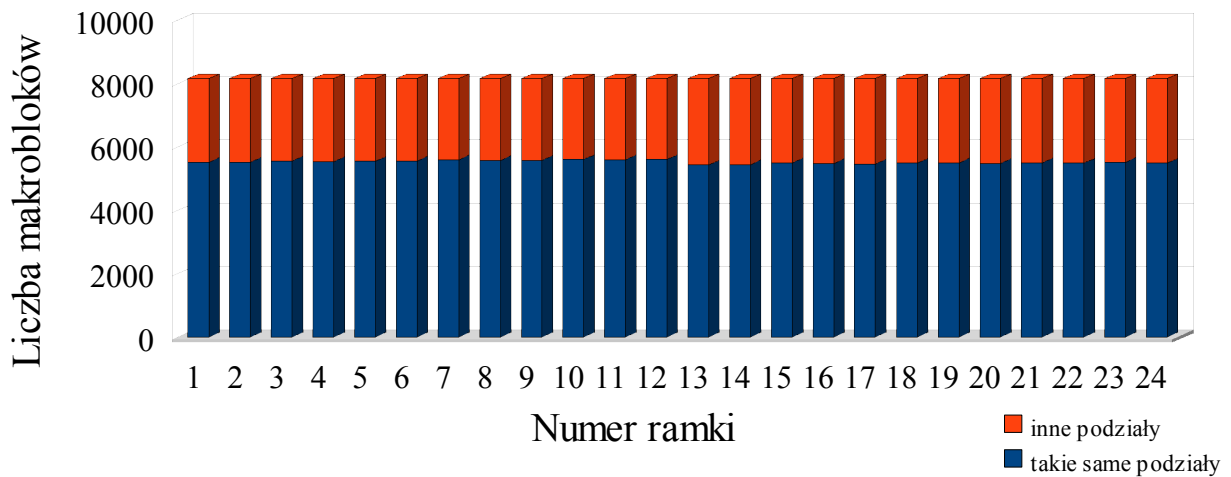
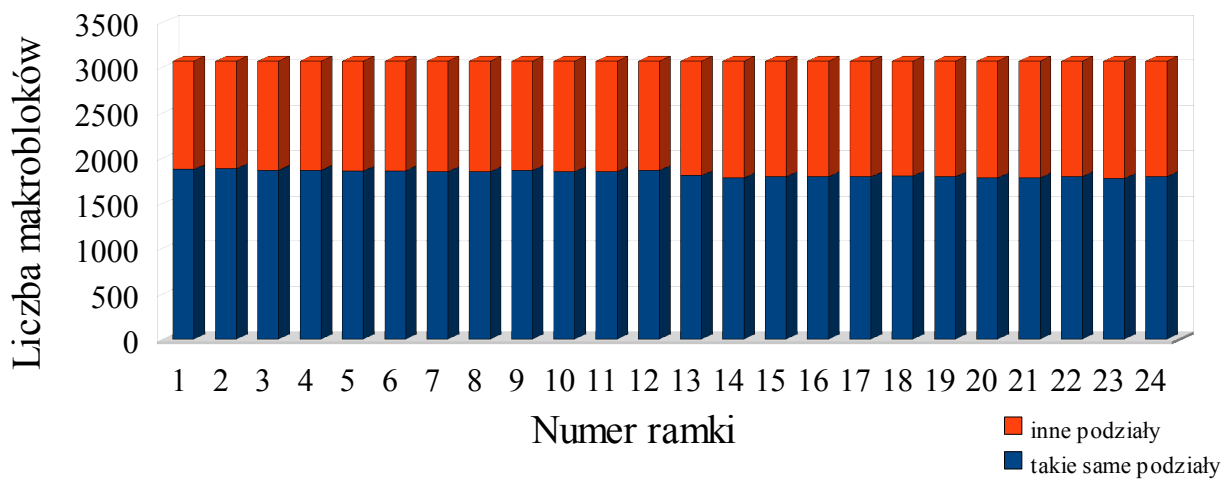
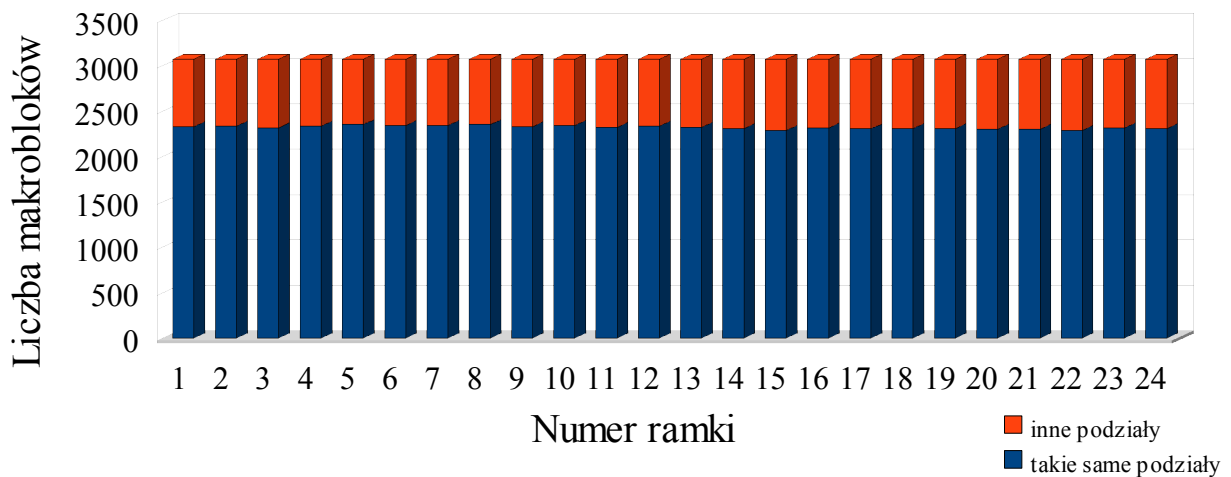


Rysunek 6.27. Rozmiary bloków dla kodowania sekwencji Poznan Street.  
 Parametry kodowania: QP = 40, QD = 40. Podział na bloki 8x8 punktów włączony.





Rysunek 6.28. Wykorzystane przy kodowaniu makrobloków podziały dla sekwencji (od góry do dołu): Book Arrival, Newspaper, Poznan Street.. Parametry kodowania: QP = 40, QD = 40. Podział na bloki 8x8 punktów wyłączony.



Rysunek 6.29. Wykorzystane przy kodowaniu makrobloków podziały dla sekwencji (od góry do dołu): Book Arrival, Newspaper, Poznan Street. Parametry kodowania: QP = 40, QD = 40. Podział na bloki 8x8 punktów włączony.

### **6.6.1.3. Kierunki predykcji dla luminancji i głębi w kodowaniu wewnątrzobrazowym**

W trakcie badań sprawdzony został również wybór kierunków predykcji dla bloków luminancji i głębi w obrębie tego samego makrobloku. Odpowiednie wykresy przedstawione są na rysunkach od 6.30 do 6.39. W przypadkach, w których występowało duże podobieństwo wyników dla wszystkich przebadanych sekwencji, ograniczono się do prezentacji wyników tylko dla jednej sekwencji.

Na rysunkach 6.30 oraz 6.31 zebrane zostały dane uzyskane podczas kodowania z parametrami  $QP=10$  i  $QD=10$  z wyłączoną możliwością wyboru podziału na bloki  $8 \times 8$  punktów. W zestawieniu podsumowano kierunki predykcji w blokach, dla których luminancja i głębia kodowana była z użyciem tych samych podziałów. W zestawieniu biorą zatem udział bloki, które na wykresie z rysunku 6.20 oznaczone są kolorem niebieskim.

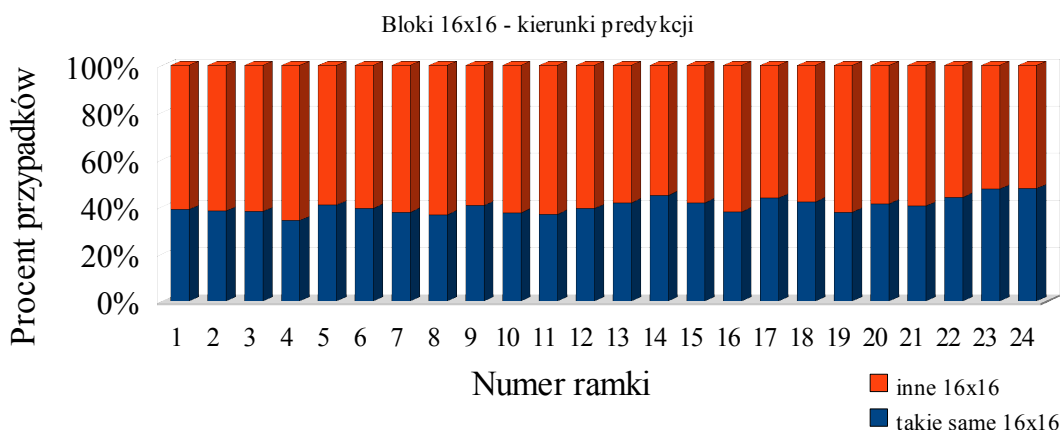
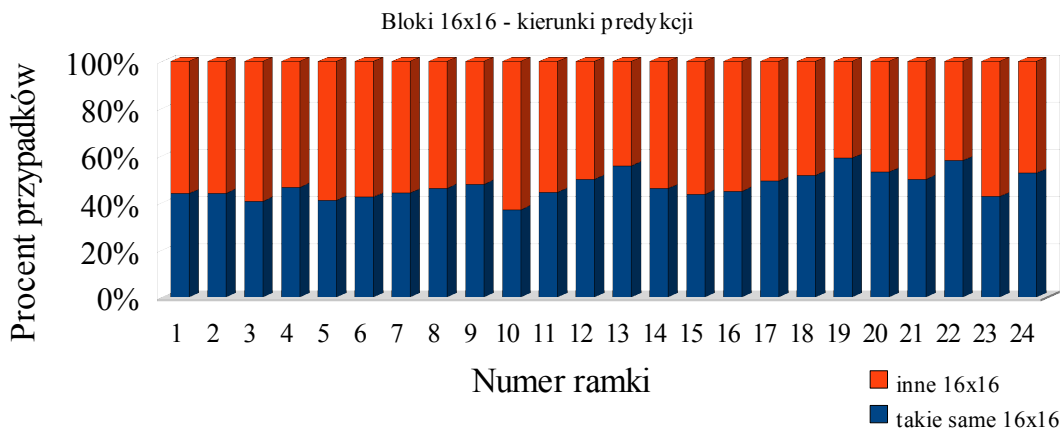
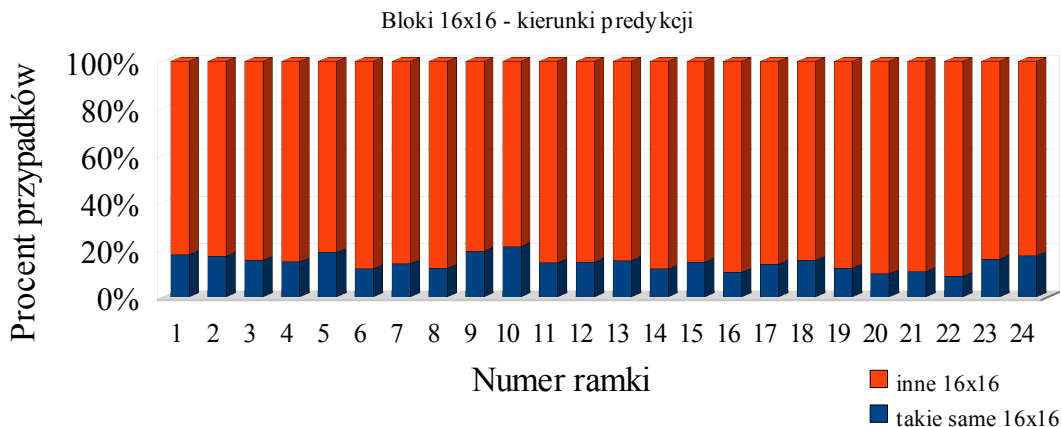
Na rysunkach 6.32, 6.33, 6.34 zebrane zostały dane uzyskane podczas kodowania z parametrami  $QP=10$  i  $QD=10$  z dopuszczoną możliwością wyboru podziału na bloki  $8 \times 8$  punktów. W zestawieniu podsumowano kierunki predykcji z bloków, dla których luminancja i głębia kodowana była z użyciem tych samych podziałów, czyli te, które na wykresie z rysunku 6.21 oznaczone są kolorem niebieskim.

Na rysunkach 6.35 oraz 6.36 zebrane zostały dane uzyskane podczas kodowania z parametrami  $QP=40$  i  $QD=40$  z wyłączoną możliwością wyboru podziału na bloki  $8 \times 8$  punktów. W zestawieniu podsumowano kierunki predykcji z bloków, dla których luminancja i głębia kodowana była z użyciem tych samych podziałów. W zestawieniu biorą zatem udział bloki, które na wykresie z rysunku 6.28 oznaczone są kolorem niebieskim.

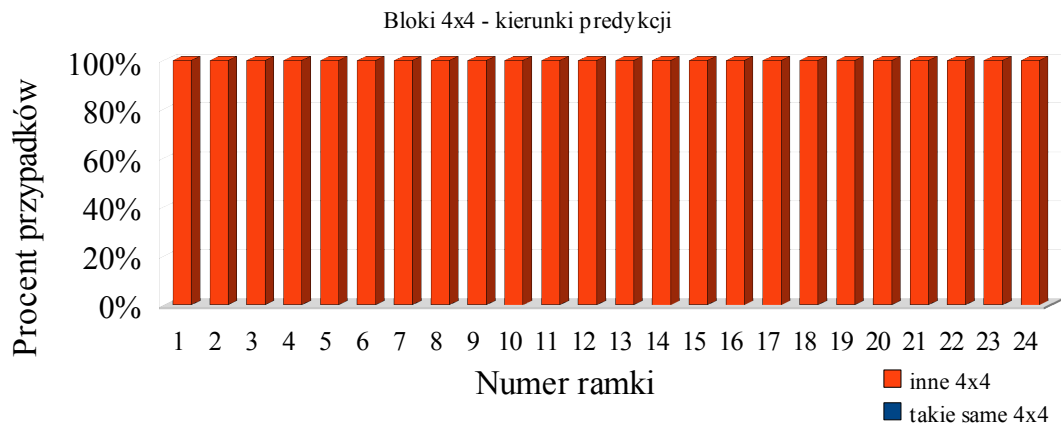
Na rysunkach 6.37, 6.38 oraz 6.39 zebrane zostały dane uzyskane podczas kodowania z parametrami  $QP=40$  i  $QD=40$  dla konfiguracji, w której możliwy był wybór podziału na bloki  $8 \times 8$  punktów. W zestawieniu podsumowano kierunki predykcji z bloków, dla których luminancja i głębia kodowana była z użyciem tych samych podziałów, czyli te, które na wykresie z rysunku 6.29 oznaczone są kolorem niebieskim.

Bardzo istotną obserwacją nasuwającą się na podstawie wyników z rysunków 6.31, 6.33, 6.34, 6.36, 6.38 oraz 6.39 jest praktyczny brak korelacji pomiędzy kierunkami predykcji w przypadku kodowania bloków głębi i luminancji z podziałem na bloki  $8 \times 8$  lub  $4 \times 4$  punkty. W przypadku wykorzystania tych podziałów nie można zatem liczyć na wykorzystanie korelacji pomiędzy kierunkami predykcji dla luminancji i głębi.

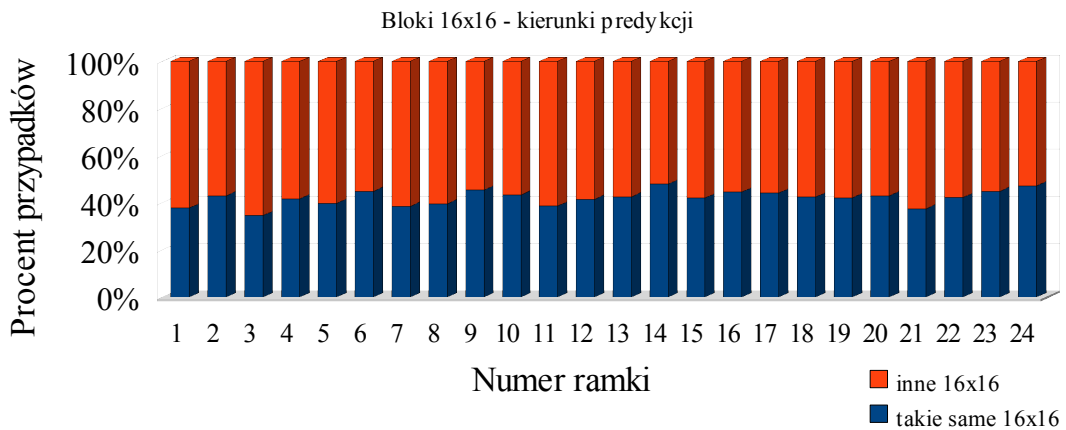
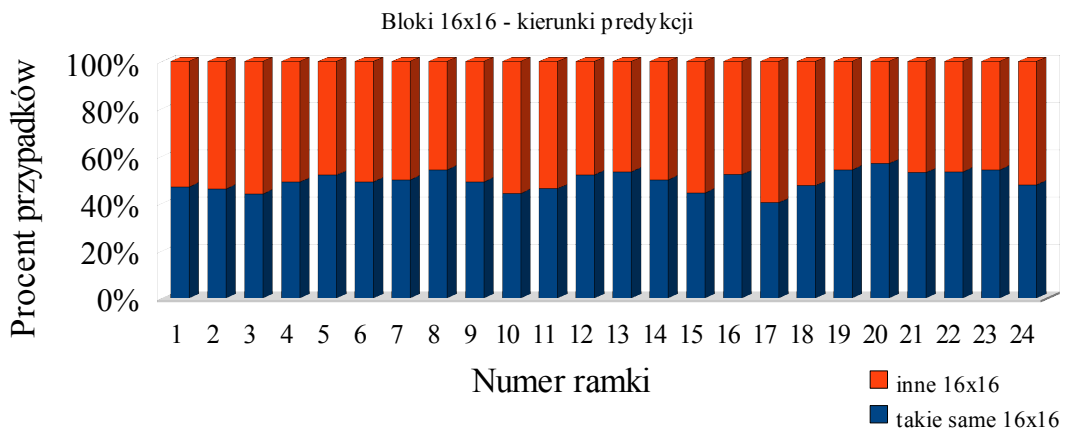
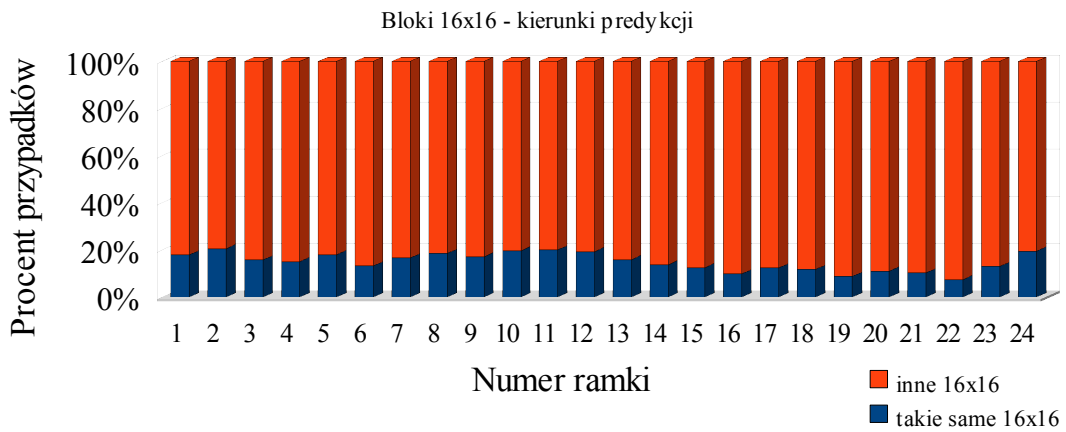
Dla podziału 16x16 punktów odsetek przypadków wykorzystania tego samego kierunku predykcji silnie zależy od badanej sekwencji i wynosi mniej niż 50% w większości przypadków, niezależnie od wybranej wartości QP i QD, choć dla QP = 40 i QD = 40 udział makrobloków wykorzystujących ten sam kierunek predykcji jest zauważalnie większy.



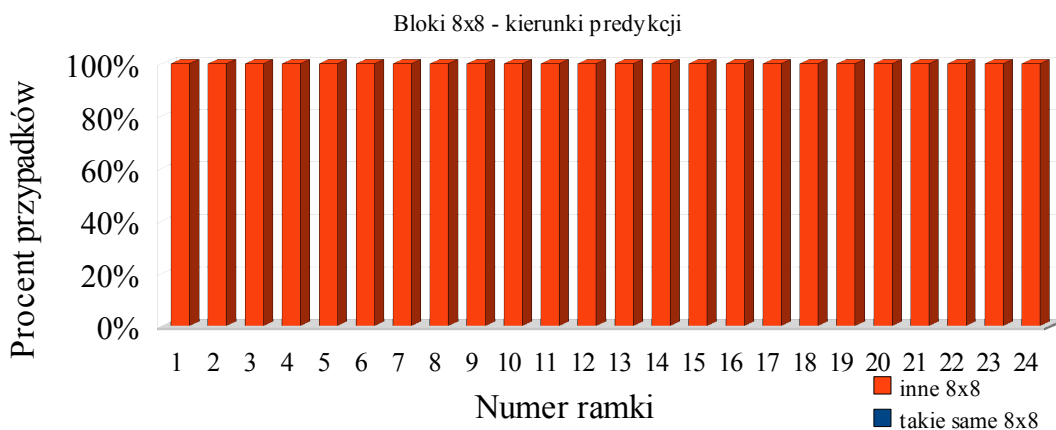
Rysunek 6.30. Porównanie wykorzystanych przy kodowaniu luminancji i głębi kierunków predykcji dla przypadków, gdy w kodowaniu obydwu komponentów w makrobloku wykorzystano bloki 16x16 (od góry do dołu): Book Arrival, Newspaper, Poznan Street. Parametry kodowania: QP = 10, QD = 10. Podział na bloki 8x8 punktów wyłączony.



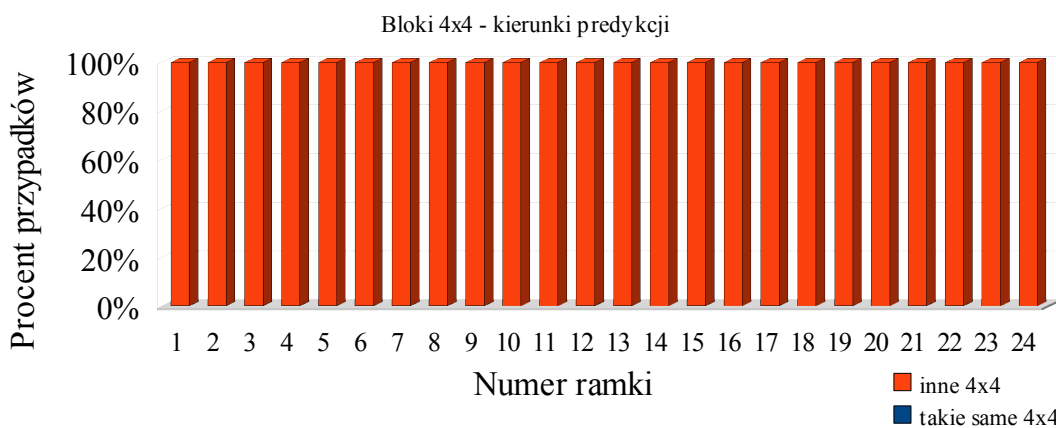
Rysunek 6.31. Porównanie wykorzystanych przy kodowaniu luminancji i głębi kierunków predykcji dla przypadków, gdy w kodowaniu obydwu komponentów w makrobloku wykorzystano bloki 4x4, sekwencja Newspaper. Parametry kodowania: QP = 10, QD = 10. Podział na bloki 8x8 punktów wyłączony.



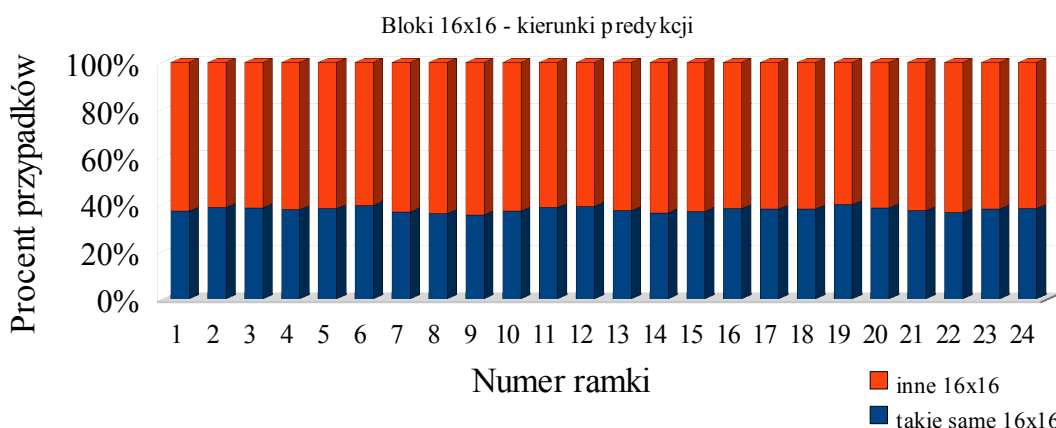
Rysunek 6.32. Porównanie wykorzystanych przy kodowaniu luminancji i głębi kierunków predykcji dla przypadków, gdy w kodowaniu obydwu komponentów w makrobloku wykorzystano bloki 16x16 (od góry do dołu): Book Arrival, Newspaper, Poznan Street. Parametry kodowania: QP = 10, QD = 10. Podział na bloki 8x8 punktów włączony.



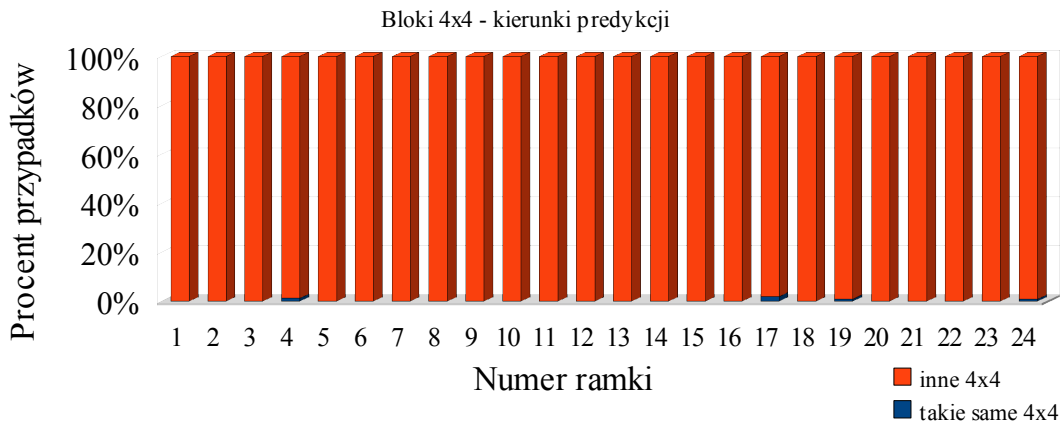
Rysunek 6.33. Porównanie wykorzystanych przy kodowaniu luminancji i głębi kierunków predykcji dla przypadków, gdy w kodowaniu obydwu komponentów w makrobloku wykorzystano bloki 8x8, sekwencja Newspaper. Parametry kodowania: QP = 10, QD = 10. Podział na bloki 8x8 punktów włączony.



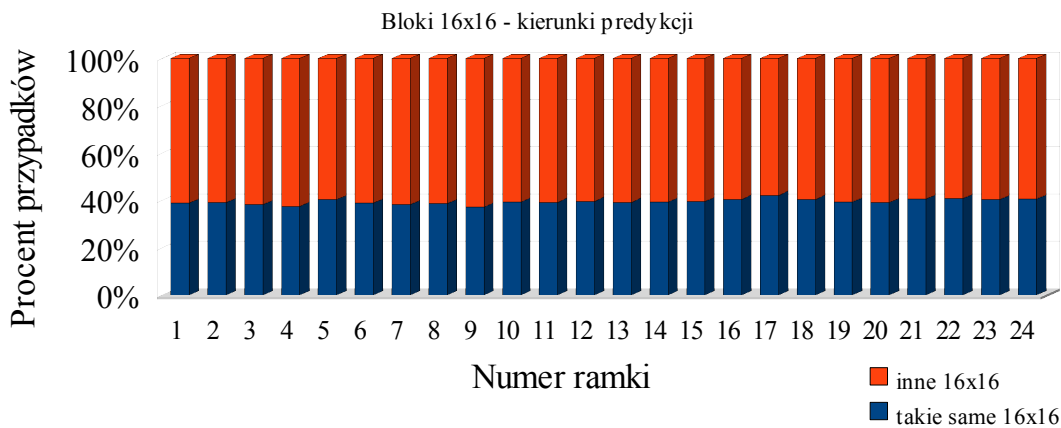
Rysunek 6.34. Porównanie wykorzystanych przy kodowaniu luminancji i głębi kierunków predykcji dla przypadków, gdy w kodowaniu obydwu komponentów w makrobloku wykorzystano bloki 4x4, sekwencja Newspaper. Parametry kodowania: QP = 10, QD = 10. Podział na bloki 8x8 punktów włączony.



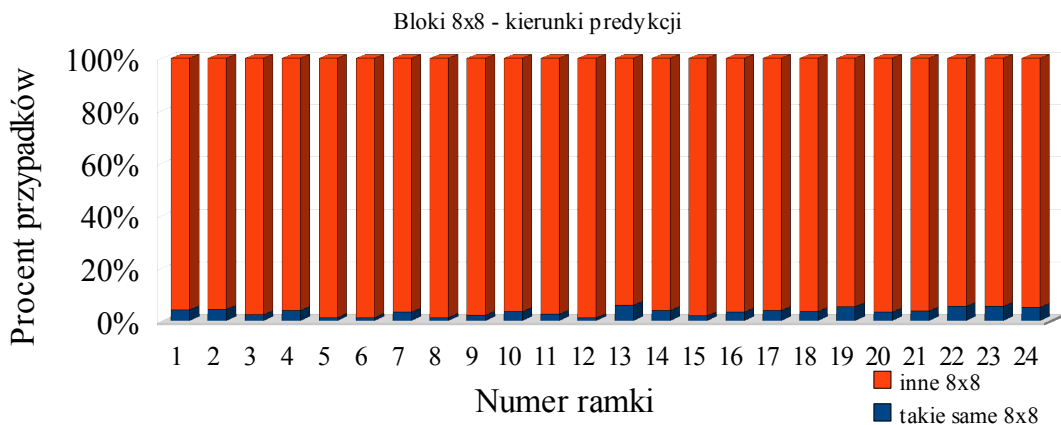
Rysunek 6.35. Porównanie wykorzystanych przy kodowaniu luminancji i głębi kierunków predykcji dla przypadków, gdy w kodowaniu obydwu komponentów w makrobloku wykorzystano bloki 16x16, sekwencja Newspaper. Parametry kodowania: QP = 40, QD = 40. Podział na bloki 8x8 punktów wyłączony.



Rysunek 6.36. Porównanie wykorzystanych przy kodowaniu luminancji i głębi kierunków predykcji dla przypadków, gdy w kodowaniu obydwu komponentów w makrobloku wykorzystano bloki 4x4, sekwencja Newspaper. Parametry kodowania: QP = 40, QD = 40. Podział na bloki 8x8 punktów wyłączony.

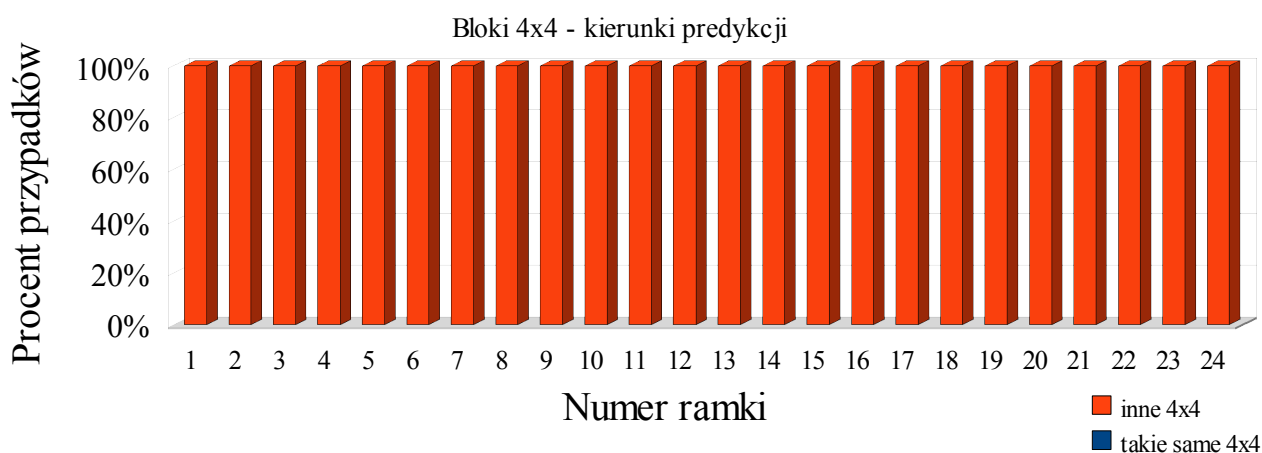


Rysunek 6.37. Porównanie wykorzystanych przy kodowaniu luminancji i głębi kierunków predykcji dla przypadków, gdy w kodowaniu obydwu komponentów w makrobloku wykorzystano bloki 16x16, sekwencja Newspaper. Parametry kodowania: QP = 40, QD = 40. Podział na bloki 8x8 punktów włączony.



Rysunek 6.38. Porównanie wykorzystanych przy kodowaniu luminancji i głębi kierunków predykcji dla przypadków, gdy w kodowaniu obydwu komponentów w makrobloku wykorzystano bloki 8x8, sekwencja Newspaper. Parametry kodowania: QP = 40, QD = 40. Podział na bloki 8x8 punktów włączony.





Rysunek 6.39. Porównanie wykorzystanych przy kodowaniu luminancji i głębi kierunków predykcji dla przypadków, gdy w kodowaniu obydwu komponentów w makrobloku wykorzystano bloki 4x4, sekwencja Newspaper. Parametry kodowania: QP = 40, QD = 40. Podział na bloki 8x8 punktów włączony.

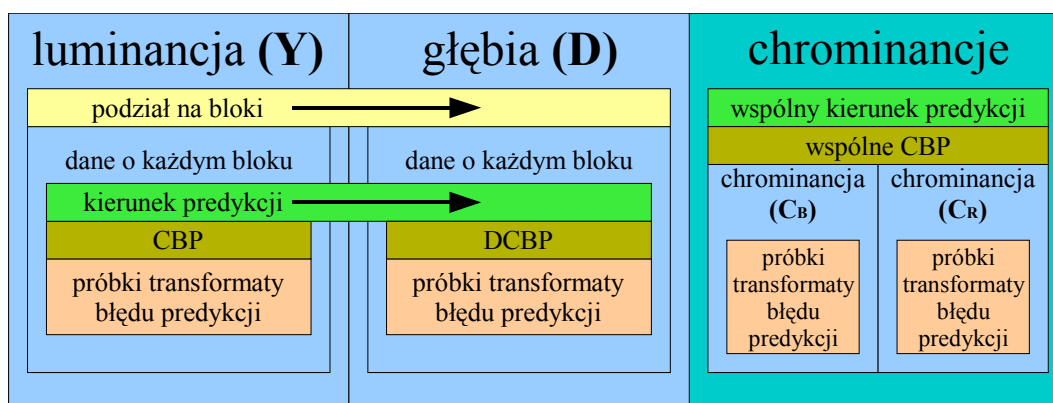
Wyniki przedstawione na rysunkach od 6.14 do 6.39 świadczą o możliwości wykorzystania istniejącej w pewnych warunkach korelacji pomiędzy sposobami kodowania bloków luminancji i głębi w makrobloku do zwiększenia efektywności kompresji sekwencji wielowidokowych z głębią. Dostrzegalny jest praktyczny brak korelacji kierunków predykcji dla podziałów makrobloku na bloki 8x8 (rysunki 6.33 i 6.38) oraz 4x4 punkty (rysunki 6.31, 6.34, 6.36 i 6.39).

### 6.6.2. Wewnątrzobrazowe kodowanie głębi z wykorzystaniem trybów dla luminancji

Najprostszym sposobem wykorzystania korelacji między obrazami i mapami głębi jest wykorzystanie do ich kodowania tych samych podziałów makrobloku i tych samych kierunków predykcji. Przeprowadzone zostały badania eksperymentalne z wykorzystaniem takiej metody kodowania wspólnego głębi i obrazów.

Dla map głębi kodowanych w pełnej rozdzielczości wprowadzona została modyfikacja do sposobu kodowania makrobloku – rysunek 6.40. Sposób podziału na bloki w obrębie makrobloku oraz kierunek predykcji wewnątrz bloków dla głębi są takie same, jak te, które zostały wybrane dla luminancji. Wybór tych parametrów dokonywany jest jedynie na podstawie składowej luminancji.

Odrębnie przesyłane jest natomiast pole DCBP dla mapy głębi, będące odpowiednikiem pola CBP dla luminancji. Zapisane są w nim informacje o tym, które spośród bloków o rozmiarach 8x8 punktów w obrębie makrobloku zawierają niezerowe próbki skwantowanej transformaty błędu predykcji.



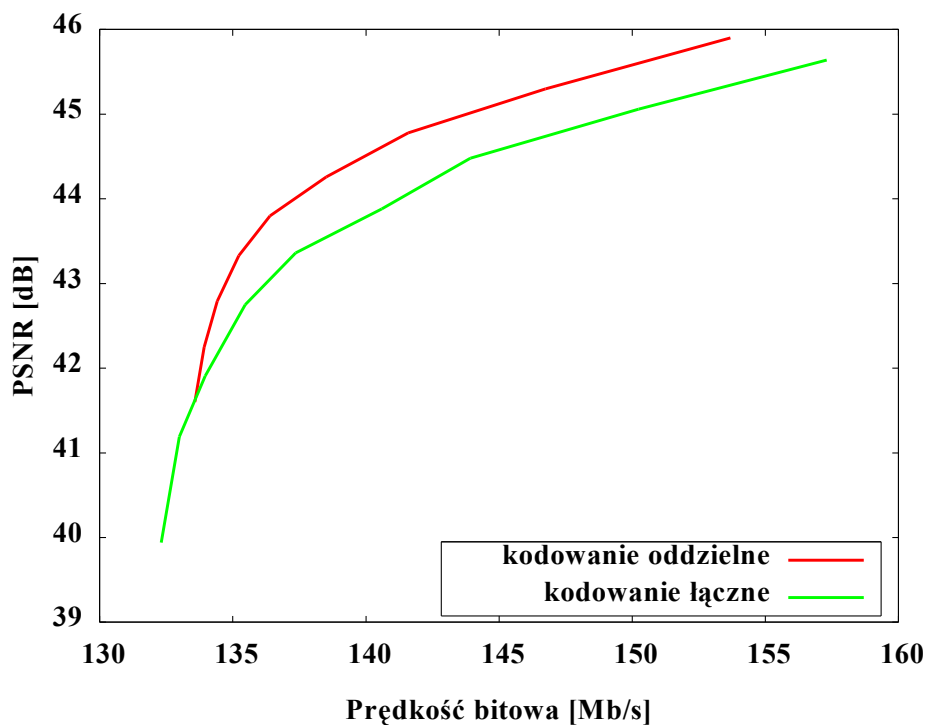
Rysunek 6.40. Schematyczne przedstawienie sposobu kodowania pojedynczego makrobloku dla kodowania czterekomponentowego – kodowanie z wykorzystaniem parametrów kodowania luminancji.

Wyniki eksperymentu wykorzystujące odpowiednio zmodyfikowane oprogramowanie przedstawione są na poniższych wykresach. Na rysunkach od 6.41 do 6.52 przedstawiono przykładowe wyniki dla wybranych sekwencji.

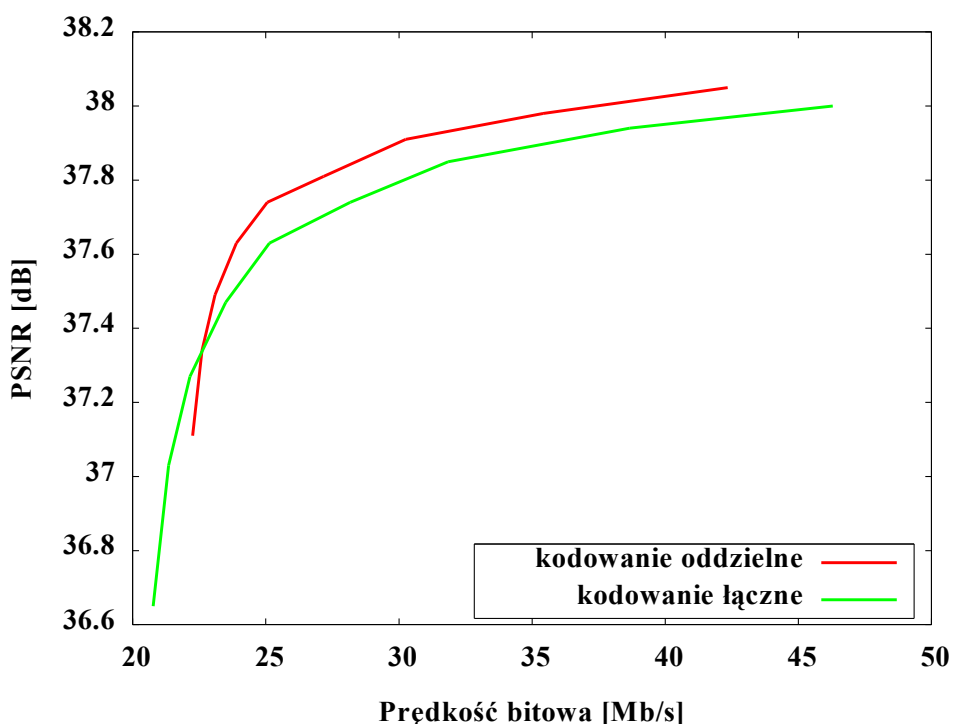
Dla poniższych przypadków przedstawiono przykładowe wyniki, uzyskane dla 3 różnych wartości ustalonego parametru, wynoszących: 16, 34 i 40. Przedstawiono wyniki dla wybranych sekwencji testowych. Zaprezentowane wyniki wykorzystują jako miarę jakości miarę PSNR i używają jako obrazu odniesienia obrazu syntezowanego z użyciem nieskompresowanych danych o obrazie i o głębi. Przedstawiona na wykresach prędkość bitowa jest całkowitą prędkością bitową dla 2 widoków i 2 map głębi.

### 6.6.2.1. Wyniki dla stałego QP

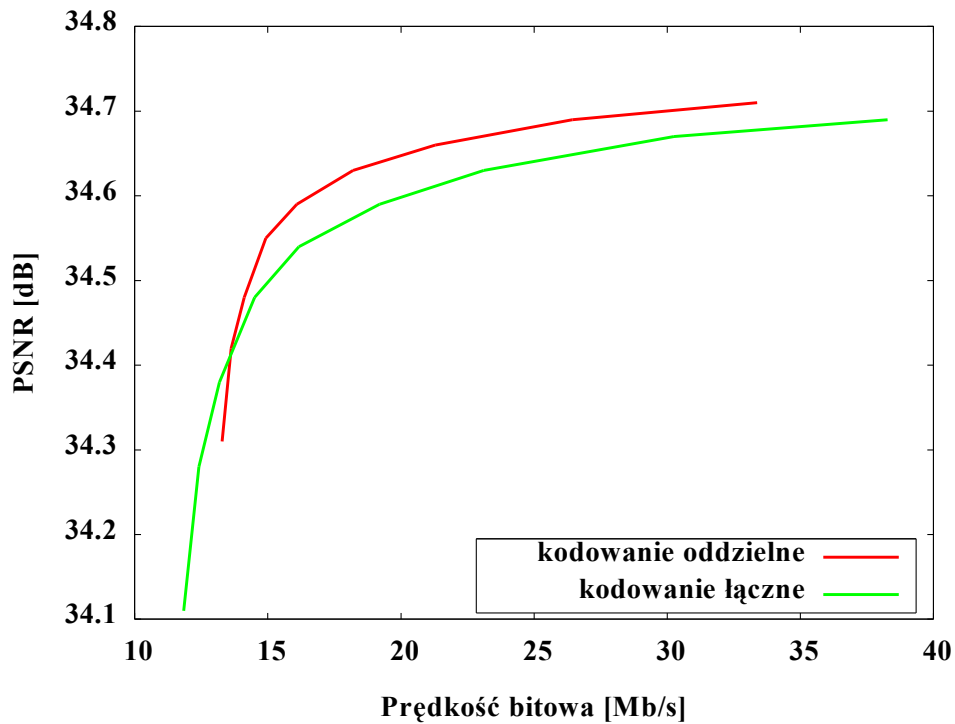
1. Kompresja z wykorzystaniem podziału makrobloku na bloki o rozmiarach 16x16, 8x8 lub 4x4 punkty dla stałego QP i zmieniającego się indeksu QD. Indeks QD przyjmował kolejne wartości z listy: 16, 19, 22, 25, 28, 31, 34, 37, 40 (rysunki od 6.41 do 6.46).
2. Kompresja z wykorzystaniem podziału makrobloku na bloki o rozmiarach 16x16 lub 4x4 punkty dla stałego QP i zmieniającego się indeksu QD. Indeks QD przyjmował kolejne wartości z listy: 16, 19, 22, 25, 28, 31, 34, 37, 40 (rysunki od 6.47 do 6.52).



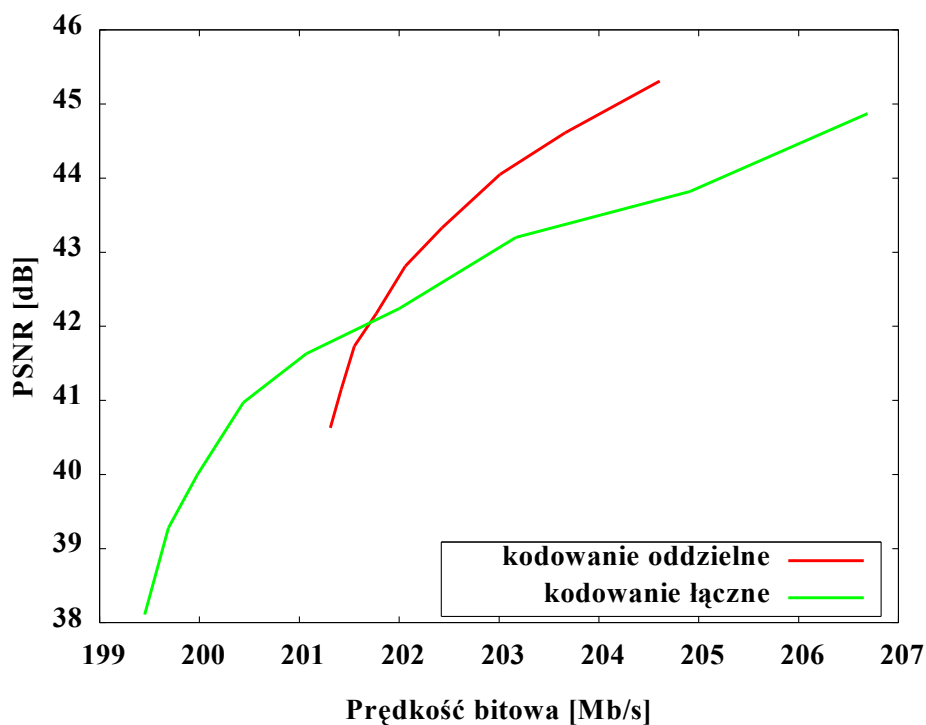
Rysunek 6.41. Porównanie efektywności kompresji kodowania czterokomponentowego do oddzielnego kodowania obrazów i map głębi. Kodowanie z możliwym podziałem na bloki 16x16, 8x8 i 4x4. Sekwencja Poznan Street, QP = 16, QD zmieniane w zakresie od 16 do 40.



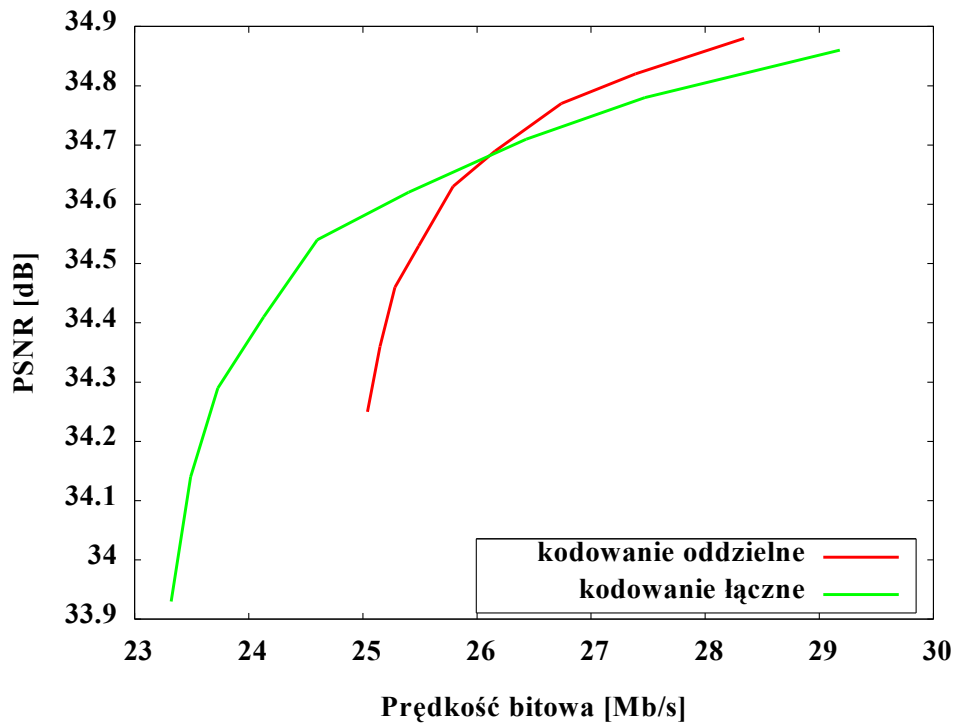
Rysunek 6.42. Porównanie efektywności kompresji kodowania czterokomponentowego do oddzielnego kodowania obrazów i map głębi. Kodowanie z możliwym podziałem na bloki 16x16, 8x8 i 4x4. Sekwencja Poznan Street, QP = 34, QD zmieniane w zakresie od 16 do 40.



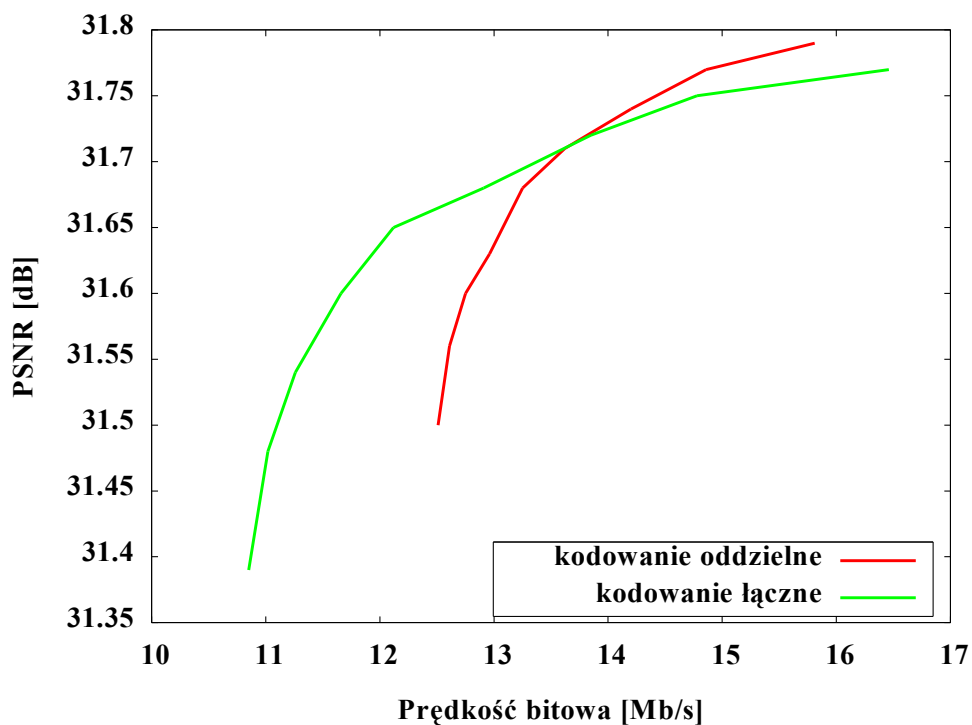
Rysunek 6.43. Porównanie efektywności kompresji kodowania czterokomponentowego do oddzielnego kodowania obrazów i map głębi. Kodowanie z możliwym podziałem na bloki 16x16, 8x8 i 4x4. Sekwencja Poznan Street, QP = 40, QD zmieniane w zakresie od 16 do 40.



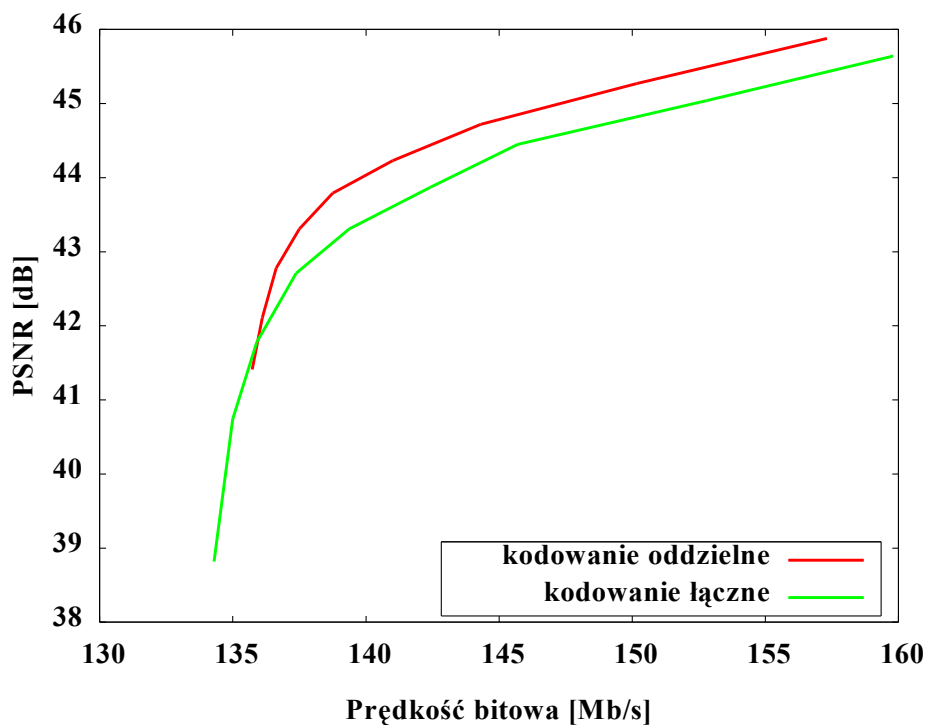
Rysunek 6.44. Porównanie efektywności kompresji kodowania czterokomponentowego do oddzielnego kodowania obrazów i map głębi. Kodowanie z możliwym podziałem na bloki 16x16, 8x8 i 4x4. Sekwencja Undo Dancer, QP = 16, QD zmieniane w zakresie od 16 do 40.



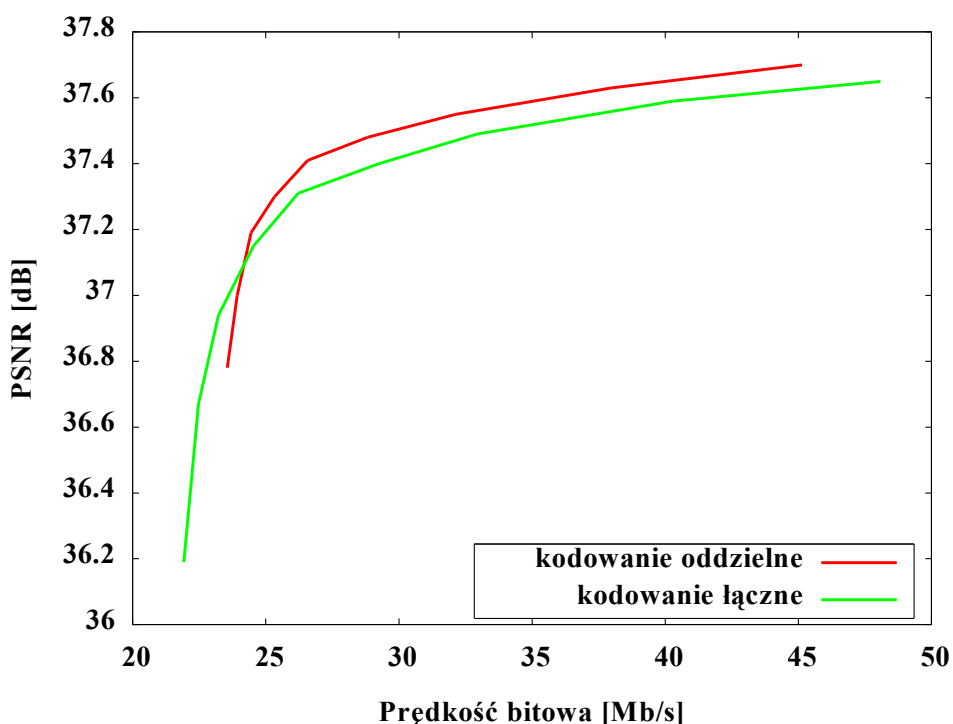
Rysunek 6.45. Porównanie efektywności kompresji kodowania czterokomponentowego do oddzielnego kodowania obrazów i map głębi. Kodowanie z możliwym podziałem na bloki 16x16, 8x8 i 4x4. Sekwencja Undo Dancer, QP = 34, QD zmieniane w zakresie od 16 do 40.



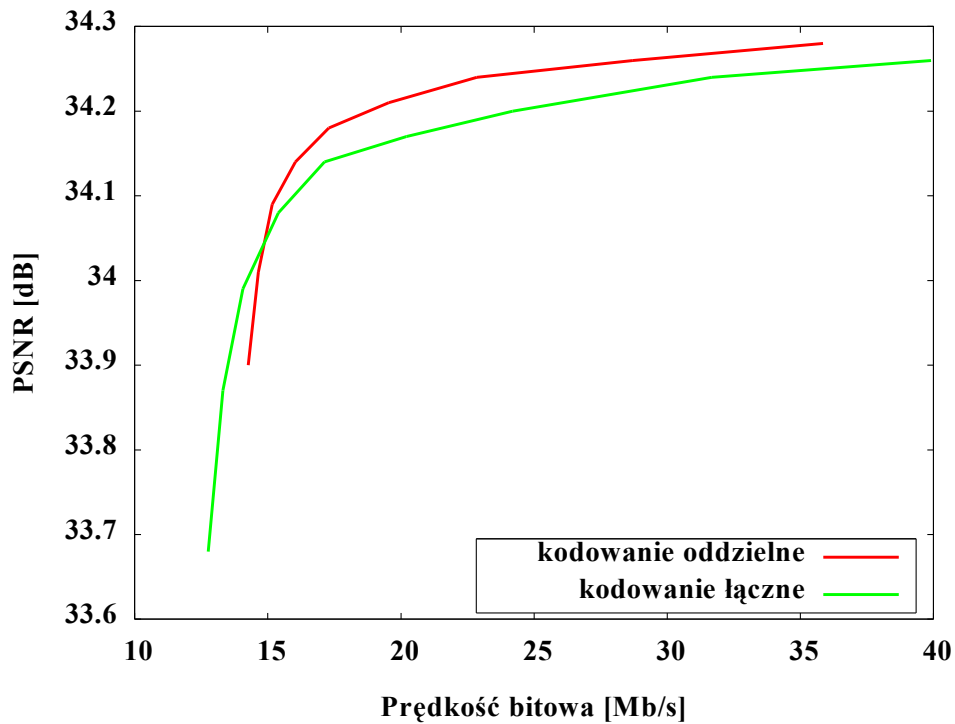
Rysunek 6.46. Porównanie efektywności kompresji kodowania czterokomponentowego do oddzielnego kodowania obrazów i map głębi. Kodowanie z możliwym podziałem na bloki 16x16, 8x8 i 4x4. Sekwencja Undo Dancer, QP = 40, QD zmieniane w zakresie od 16 do 40.



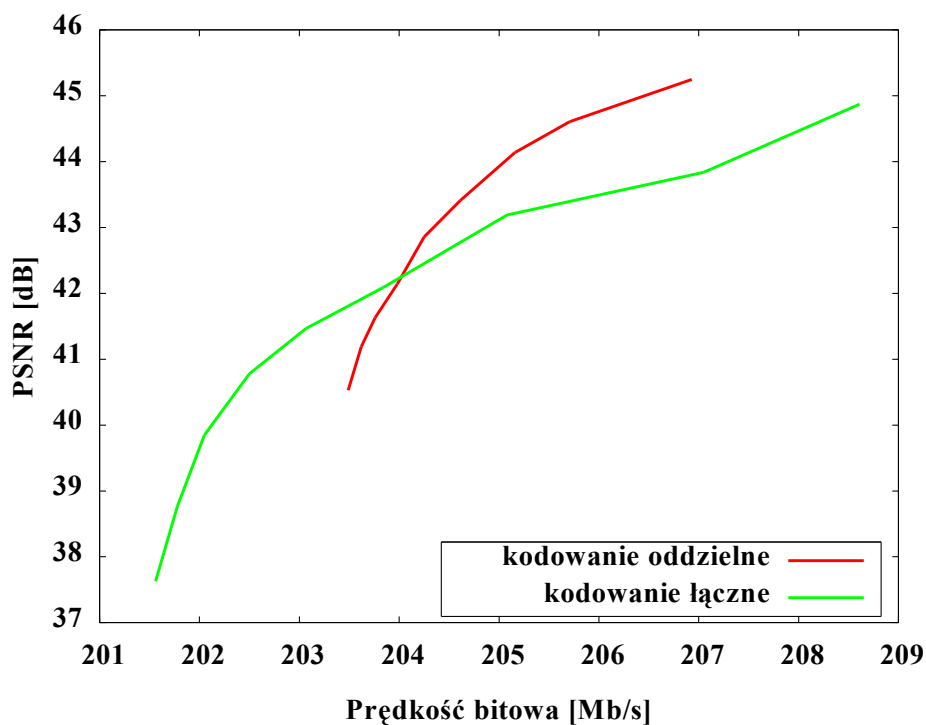
Rysunek 6.47. Porównanie efektywności kompresji kodowania czterokomponentowego do oddzielnego kodowania obrazów i map głębi. Kodowanie z możliwym podziałem na bloki 16x16 i 4x4. Sekwencja Poznan Street, QP = 16, QD zmieniane w zakresie od 16 do 40.



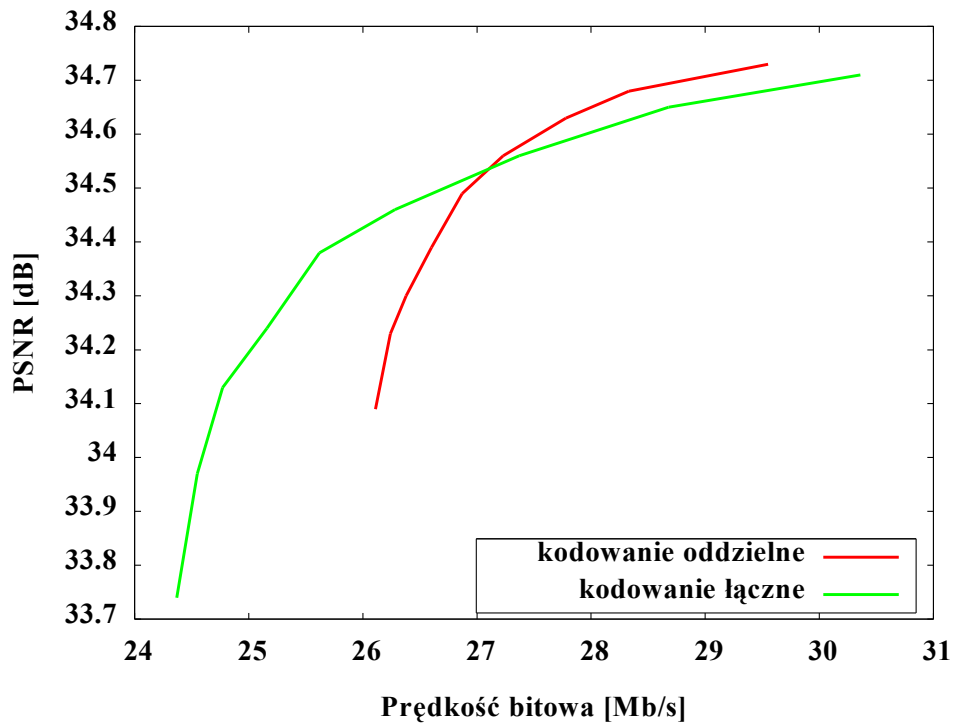
Rysunek 6.48. Porównanie efektywności kompresji kodowania czterokomponentowego do oddzielnego kodowania obrazów i map głębi. Kodowanie z możliwym podziałem na bloki 16x16 i 4x4. Sekwencja Poznan Street, QP = 34, QD zmieniane w zakresie od 16 do 40.



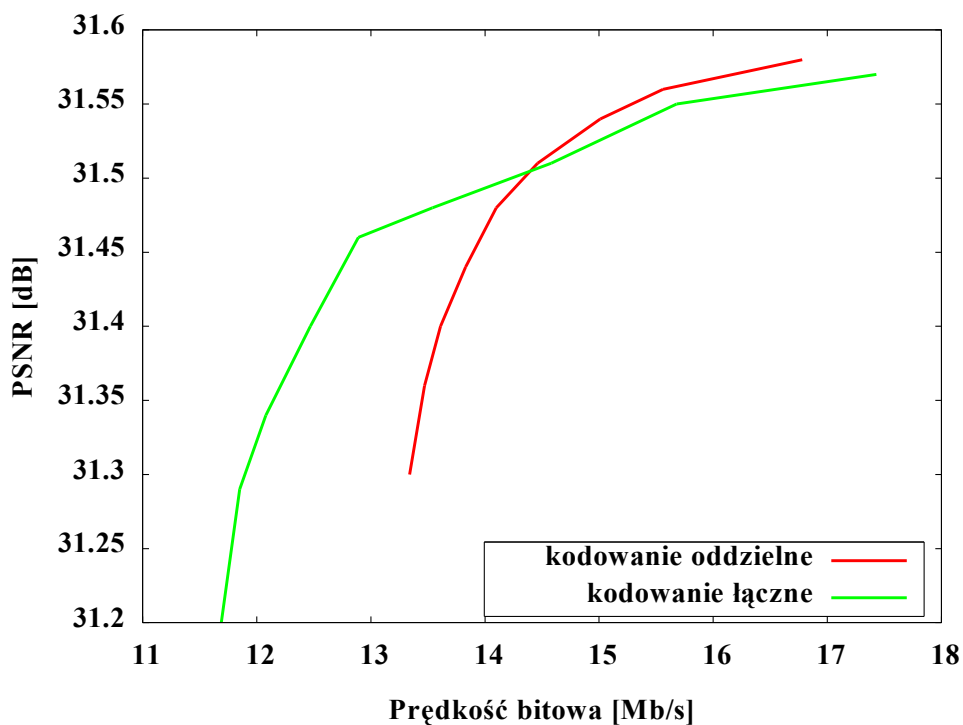
Rysunek 6.49. Porównanie efektywności kompresji kodowania czterokomponentowego do oddzielnego kodowania obrazów i map głębi. Kodowanie z możliwym podziałem na bloki 16x16 i 4x4. Sekwencja Poznan Street, QP = 40, QD zmieniane w zakresie od 16 do 40.



Rysunek 6.50. Porównanie efektywności kompresji kodowania czterokomponentowego do oddzielnego kodowania obrazów i map głębi. Kodowanie z możliwym podziałem na bloki 16x16 i 4x4. Sekwencja Undo Dancer, QP = 16, QD zmieniane w zakresie od 16 do 40.



Rysunek 6.51. Porównanie efektywności kompresji kodowania czterokomponentowego do oddzielnego kodowania obrazów i map głębi. Kodowanie z możliwym podziałem na bloki 16x16 i 4x4. Sekwencja Undo Dancer, QP = 34, QD zmieniane w zakresie od 16 do 40.



Rysunek 6.52. Porównanie efektywności kompresji kodowania czterokomponentowego do oddzielnego kodowania obrazów i map głębi. Kodowanie z możliwym podziałem na bloki 16x16 i 4x4. Sekwencja Undo Dancer, QP = 40, QD zmieniane w zakresie od 16 do 40.



Na wykresach z rysunków od 6.41 do 6.52 przedstawiono wybrane wyniki uzyskane w przypadku zastosowania stałej wartości indeksu QP i zmieniania wartości indeksu QD. Dla każdego przypadku przedstawionego na oddzielnym wykresie jakość obrazów jest zatem stała, a zmianie ulega jedynie jakość i prędkość bitowa map głębi.

Porównanie metod kompresji dla przypadków z włączonym (rysunki od 6.41 do 6.46) i wyłączonym (rysunki od 6.47 do 6.52) podziałem na bloki 8x8 punktów daje podobne wyniki i na ich podstawie poczynić można bardzo zbliżone obserwacje.

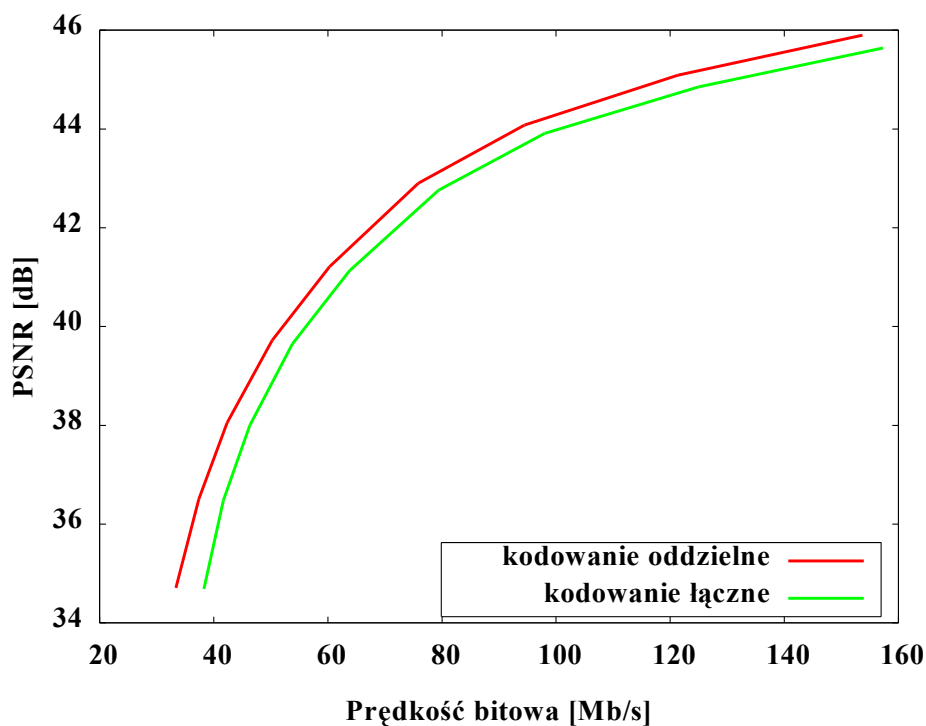
Wyniki świadczą o tym, że proponowana metoda kompresji czterokomponentowej powoduje w większości przebadanych w tej konfiguracji przypadków spadek efektywności kodowania. Spadek jakości dla najwyższych otrzymywanych prędkości bitowych wynosi od ponad 1 dB do mniej niż 0,1 dB dla przypadków odpowiednio QP = 16 oraz QP = 40.

Dla części przebadanych sekwencji (Poznan Street i Undo Dancer) obserwowany jest jednak wzrost efektywności kodowania w przypadkach, w których wykorzystane są duże wartości indeksu QD (powyżej 30). Wzrost efektywności mierzony jako spadek prędkości bitowej przy zachowaniu tej samej jakości wynosi w tych przypadkach nawet od 4% do 5% prędkości bitowej w porównaniu z oddzielnym kodowaniem sekwencji obrazów i głębi. Badana metoda daje zatem w wypadku tych sekwencji poprawę efektywności kodowania dla niskich prędkości bitowych.

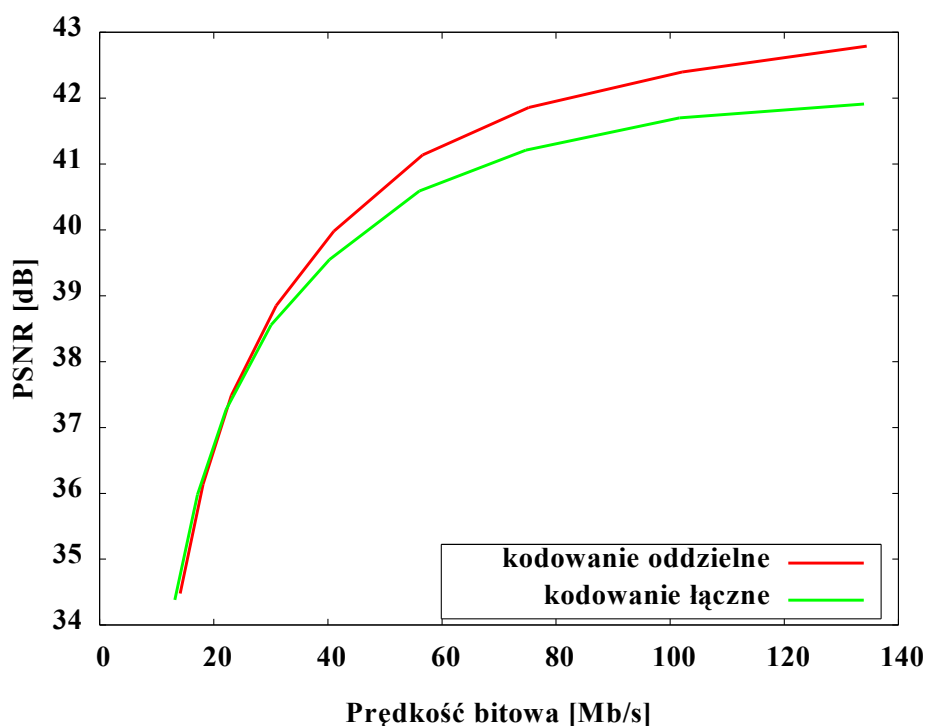
### **6.6.2.2. Wyniki dla stałego QD**

W następnej kolejności przedstawione są wyniki dla następujących przypadków (rysunki od 6.53 do 6.64):

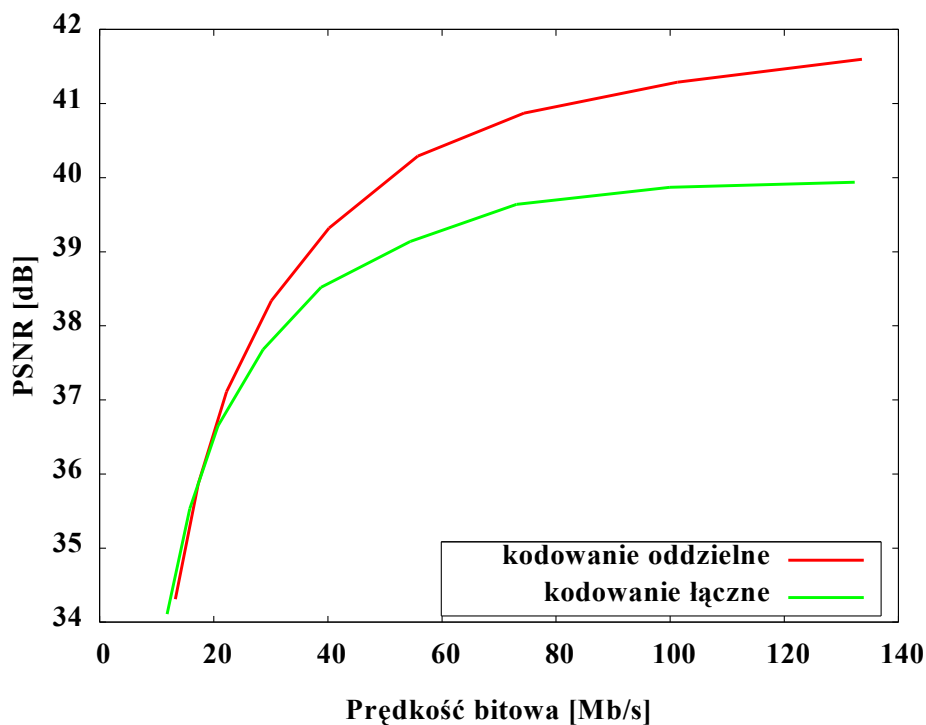
1. Kompresja z wykorzystaniem podziału makrobloku na bloki o rozmiarach 16x16, 8x8 lub 4x4 punkty dla stałego QD i zmieniającego się indeksu QP. Indeks QP przyjmował kolejne wartości z listy: 16, 19, 22, 25, 28, 31, 34, 37, 40 (rysunki od 6.53 do 6.58).
2. Kompresja z wykorzystaniem podziału makrobloku na bloki o rozmiarach 16x16 lub 4x4 punkty dla stałego QD i zmieniającego się indeksu QP. Indeks QP przyjmował kolejne wartości z listy: 16, 19, 22, 25, 28, 31, 34, 37, 40 (rysunki od 6.59 do 6.64).



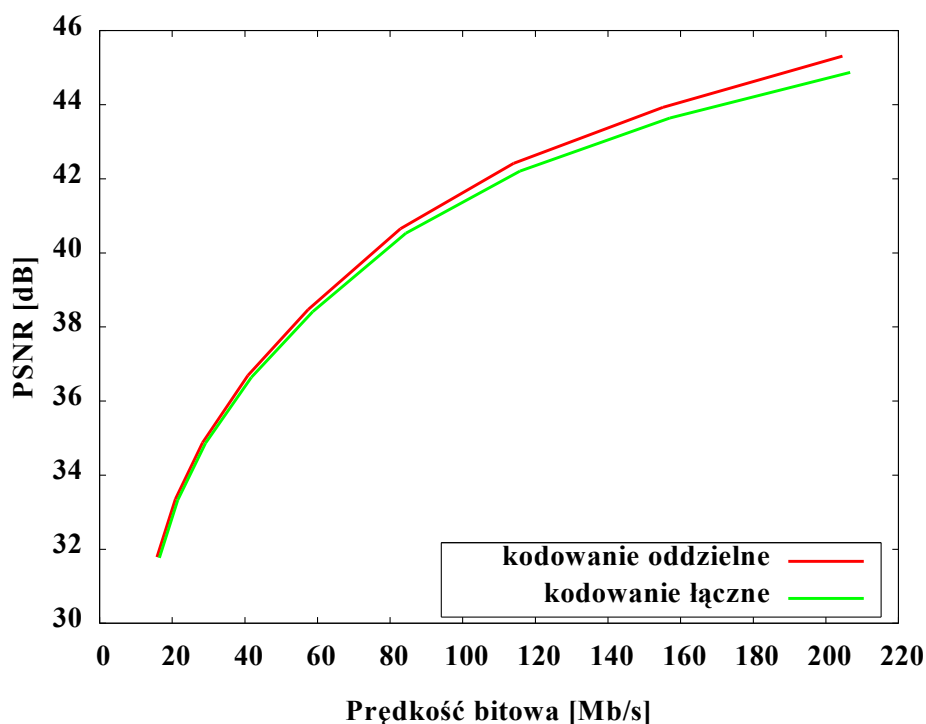
Rysunek 6.53. Porównanie efektywności kompresji kodowania czterokomponentowego do oddzielnego kodowania obrazów i map głębi. Kodowanie z możliwym podziałem na bloki 16x16, 8x8 i 4x4. Sekwencja Poznan Street, QD= 16, QP zmieniane w zakresie od 16 do 40.



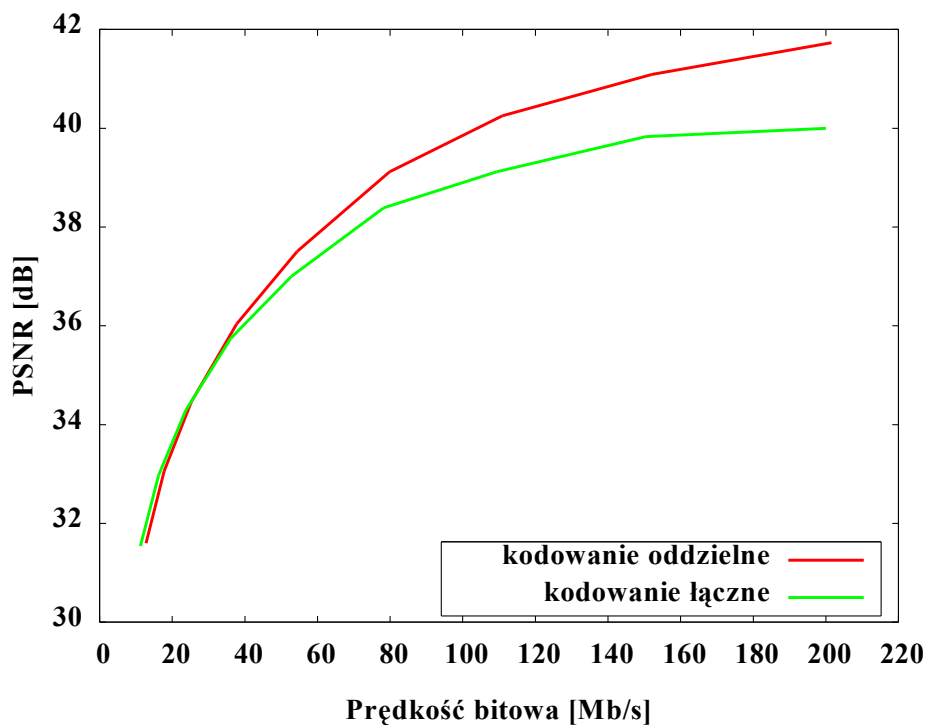
Rysunek 6.54. Porównanie efektywności kompresji kodowania czterokomponentowego do oddzielnego kodowania obrazów i map głębi. Kodowanie z możliwym podziałem na bloki 16x16, 8x8 i 4x4. Sekwencja Poznan Street, QD= 34, QP zmieniane w zakresie od 16 do 40.



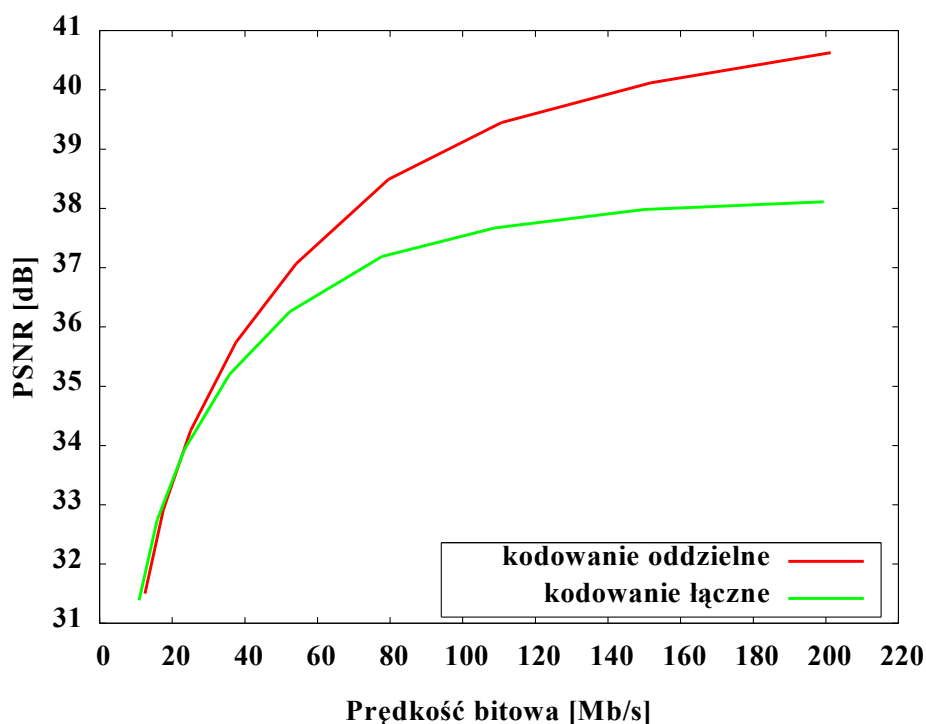
Rysunek 6.55. Porównanie efektywności kompresji kodowania czterokomponentowego do oddzielnego kodowania obrazów i map głębi. Kodowanie z możliwym podziałem na bloki 16x16, 8x8 i 4x4. Sekwencja Poznan Street, QD= 40, QP zmieniane w zakresie od 16 do 40.



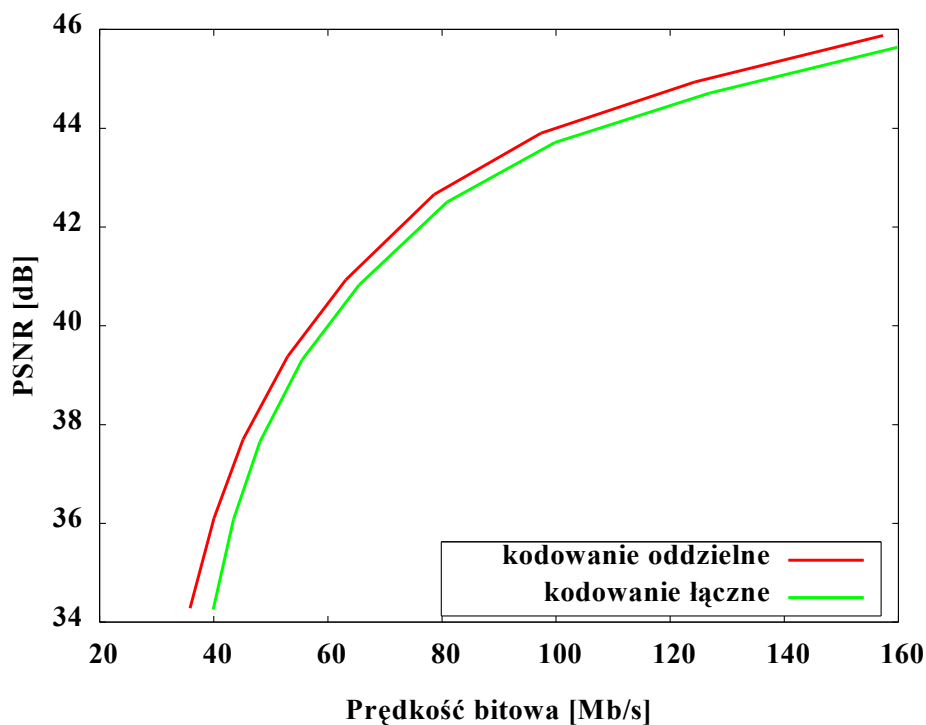
Rysunek 6.56. Porównanie efektywności kompresji kodowania czterokomponentowego do oddzielnego kodowania obrazów i map głębi. Kodowanie z możliwym podziałem na bloki 16x16, 8x8 i 4x4. Sekwencja Undo Dancer, QD= 16, QP zmieniane w zakresie od 16 do 40.



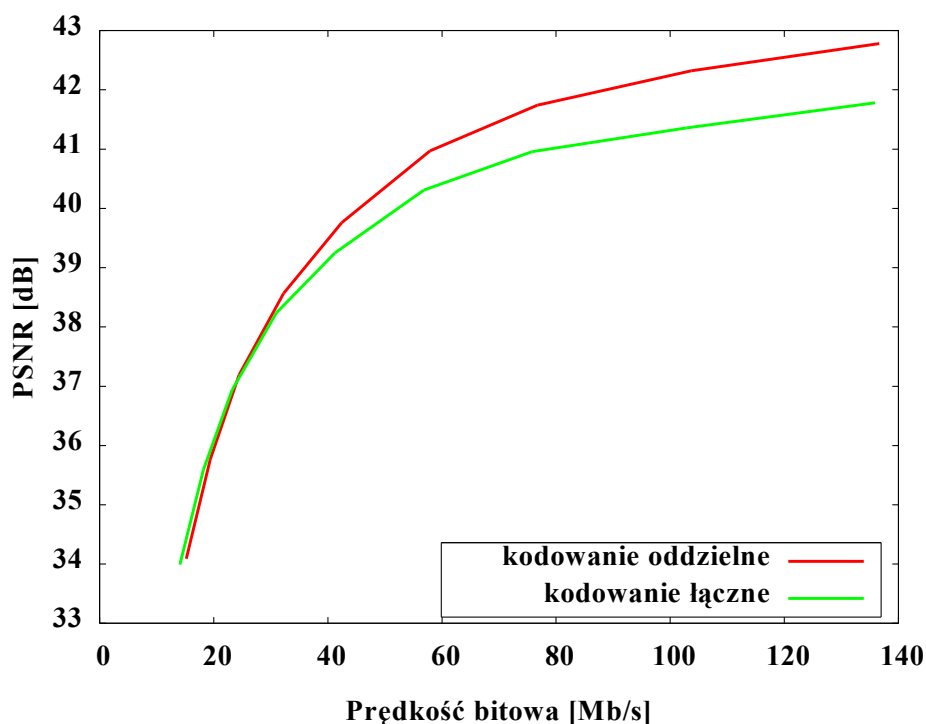
Rysunek 6.57. Porównanie efektywności kompresji kodowania czterokomponentowego do oddzielnego kodowania obrazów i map głębi. Kodowanie z możliwym podziałem na bloki 16x16, 8x8 i 4x4. Sekwencja Undo Dancer, QD= 34, QP zmieniane w zakresie od 16 do 40.



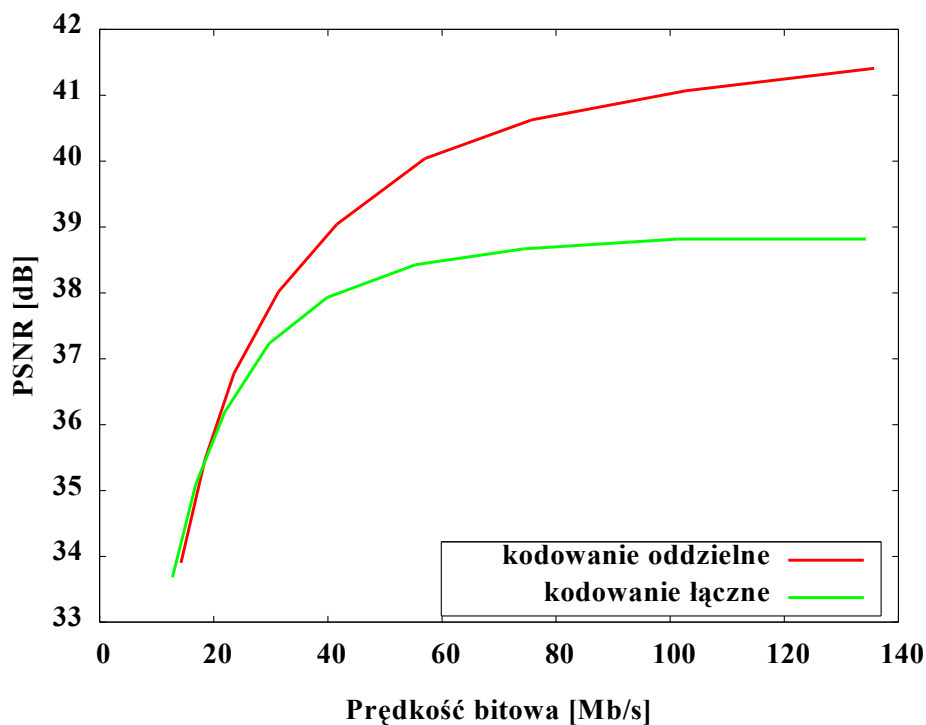
Rysunek 6.58. Porównanie efektywności kompresji kodowania czterokomponentowego do oddzielnego kodowania obrazów i map głębi. Kodowanie z możliwym podziałem na bloki 16x16, 8x8 i 4x4. Sekwencja Undo Dancer, QD= 40, QP zmieniane w zakresie od 16 do 40.



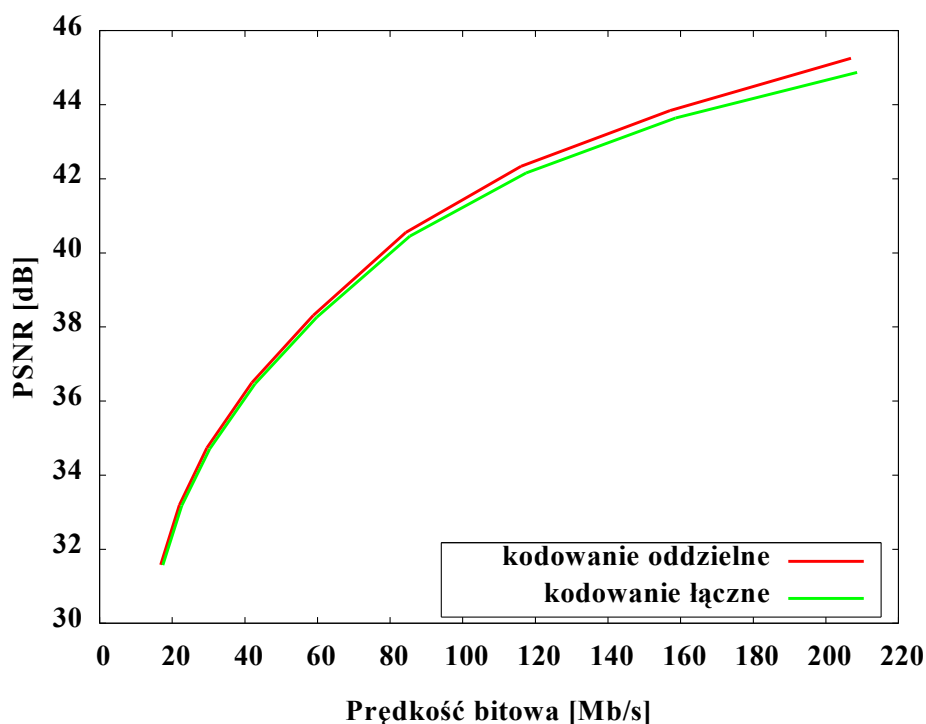
Rysunek 6.59. Porównanie efektywności kompresji kodowania czterokomponentowego do oddzielnego kodowania obrazów i map głębi. Kodowanie z możliwym podziałem na bloki 16x16 i 4x4. Sekwencja Poznan Street, QD= 16, QP zmieniane w zakresie od 16 do 40.



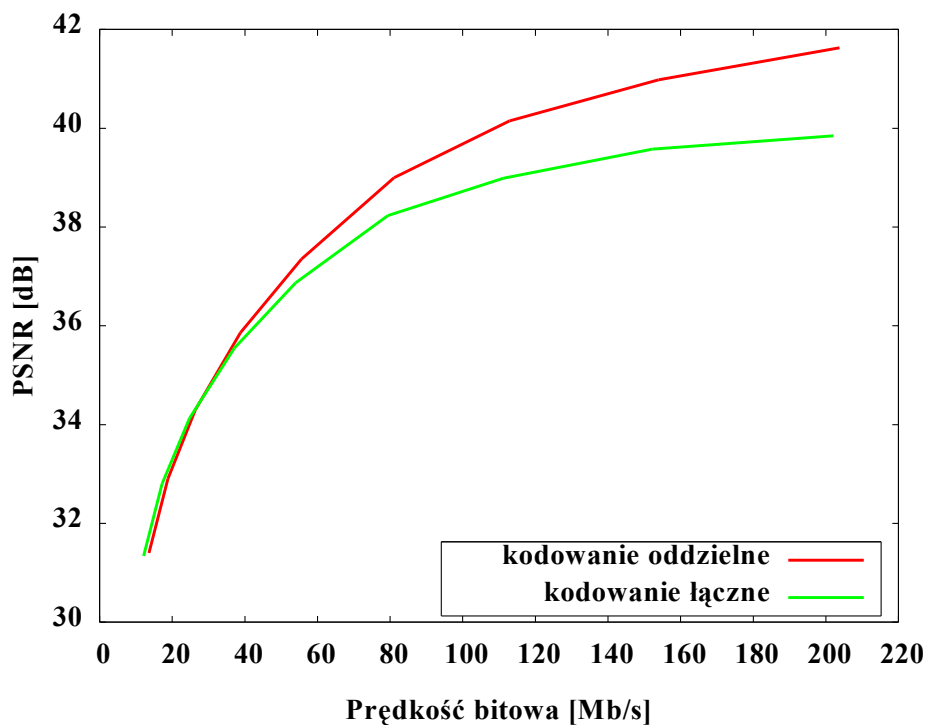
Rysunek 6.60. Porównanie efektywności kompresji kodowania czterokomponentowego do oddzielnego kodowania obrazów i map głębi. Kodowanie z możliwym podziałem na bloki 16x16 i 4x4. Sekwencja Poznan Street, QD= 34, QP zmieniane w zakresie od 16 do 40.



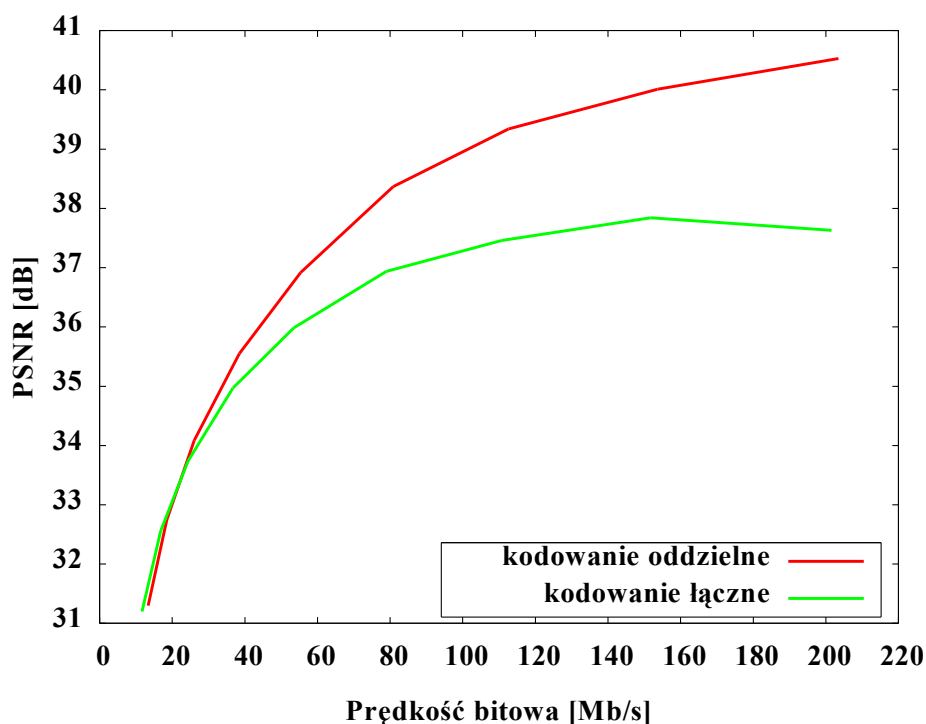
Rysunek 6.61. Porównanie efektywności kompresji kodowania czterokomponentowego do oddzielnego kodowania obrazów i map głębi. Kodowanie z możliwym podziałem na bloki 16x16 i 4x4. Sekwencja Poznan Street, QD= 40, QP zmieniane w zakresie od 16 do 40.



Rysunek 6.62. Porównanie efektywności kompresji kodowania czterokomponentowego do oddzielnego kodowania obrazów i map głębi. Kodowanie z możliwym podziałem na bloki 16x16 i 4x4. Sekwencja Undo Dancer, QD= 16, QP zmieniane w zakresie od 16 do 40.



Rysunek 6.63. Porównanie efektywności kompresji kodowania czterokomponentowego do oddzielnego kodowania obrazów i map głębi. Kodowanie z możliwym podziałem na bloki 16x16 i 4x4. Sekwencja Undo Dancer, QD= 34, QP zmieniane w zakresie od 16 do 40.



Rysunek 6.64. Porównanie efektywności kompresji kodowania czterokomponentowego do oddzielnego kodowania obrazów i map głębi. Kodowanie z możliwym podziałem na bloki 16x16 i 4x4. Sekwencja Undo Dancer, QD= 40, QP zmieniane w zakresie od 16 do 40.

Na wykresach z rysunków od 6.53 do 6.64 przedstawiono przykładowe wyniki uzyskane w przypadku zastosowania stałej wartości indeksu QD i zmieniania wartości indeksu QP. Dla każdego przypadku przedstawionego na oddzielnym wykresie jakość i prędkość bitowa obrazów jest zatem zmienna. Niewielkiej zmianie ulega również jakość i prędkość bitowa map głębi, ze względu na uzależnienie wykorzystywanych do jej kodowania trybów od trybów dla luminancji.

Spadek jakości dla najwyższych otrzymywanych prędkości bitowych wynosi, w zależności od sekwencji, od 3 do 4 dB w przypadku użycia wartości indeksu QD równej 40. Obserwowany spadek jakości jest mniejszy dla mniejszych wartości indeksu QD i dla  $QD = 16$  wynosi już tylko około 1 dB.

W tym przypadku tak duża, czterodecybelowa różnica jakości może być wyjaśniona na podstawie obserwacji poczynionych w ramach badań opisanych w punkcie 5.2.2.2: w przypadku przeprowadzania syntezy widoku na podstawie mapy głębi o niskiej jakości, nie można przekroczyć pewnej maksymalnej jakości widoku syntezy, nawet jeśli do syntezy użyte zostaną obrazy bardzo wysokiej jakości. Efekt ograniczania jakości widoku syntezy jest wyraźnie widoczny na wykresach z rysunków 6.61, 6.63 oraz 6.64.

Dla części spośród badanych sekwencji (Book Arrival, Poznan Street i Undo Dancer) obserwowany jest jednak pewien wzrost efektywności kodowania dla przypadków, w których wykorzystane są duże wartości indeksu QP (powyżej 34) dla QD wynoszących 34 i 40. Wzrost efektywności mierzony jako spadek prędkości bitowej przy zachowaniu tej samej jakości wynosi w tych przypadkach nawet do 4% prędkości bitowej przy oddzielnym kodowaniu sekwencji obrazów i głębi. Jest to zatem bardzo podobny wynik do tego, jaki uzyskany był w konfiguracjach ze stałą wartością QP.

Porównanie metod kompresji dla przypadków z włączonym i wyłączonym podziałem na bloki 8x8 punktów daje podobne wyniki i na ich podstawie poczynić można bardzo zbliżone obserwacje.

### **6.6.2.3. Podsumowanie**

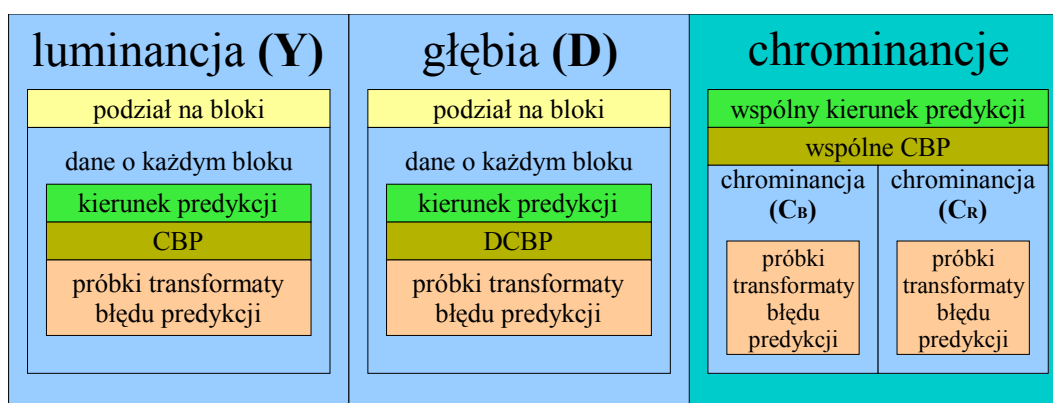
Wyniki przedstawione na wykresach z rysunków od 6.41 do 6.64 pokazują, że wykorzystanie łącznego kodowania sekwencji czterokomponentowych pozwala na uzyskanie dla części przebadanych przypadków zauważalnie lepszej efektywności kodowania (spadek prędkości bitowej do 4%) dla wysokich wartości indeksów kwantyzacji QP oraz QD (powyżej 30). W przypadku użycia mniejszych wartości tych indeksów przedstawiona metoda kodowania prowadzi do obniżenia efektywności kodowania, rozumianej jako stosunek jakości widoku syntezy do zrekonstruowanych obrazów do prędkości bitowej skompresowanego strumienia.



### 6.6.3. Wewnątrzobrazowe kodowanie głębi z możliwością wyboru odrębnych trybów

W dalszej kolejności przebadany został nieco bardziej złożony sposób kodowania komponentu głębi, dający swobodę wyboru trybu kodowania informacji o głębi. Modyfikacja taka wprowadzona była po poczynieniu obserwacji, że obniżenie efektywności kodowania opisane w punkcie 6.6.2 wiąże się z wykorzystywaniem do kodowania mapy głębi identycznych trybów kodowania makrobloków co dla kompresji obrazu. Oznacza to, że obraz głębi kodowany jest w nieoptymalny dla tej składowej sposób. Wypróbowany został zatem sposób kodowania informacji o makrobloku dla sekwencji czterokomponentowej, schematycznie przedstawiony na rysunku 6.65. W tym przypadku dla luminancji i głębi informacje o sposobie podziału na bloki przesyłane są oddzielne. Oddzielnie są również przesyłane informacje na temat wybranych dla poszczególnych bloków kierunków predykcji. Zarówno bloki luminancji jak i głębi kodowane są w pełnej rozdzielczości. Chrominancje i luminancja kodowane są w standardowy dla koderów MVC sposób. Koder działający dokładnie w sposób przedstawiono na rysunku nie został jednak przetestowany, gdyż nie daje on możliwości wykorzystania korelacji między mapami głębi a obrazem.

Sposobem na wykorzystanie korelacji jest w takim wypadku wspólne kodowanie parametrów kodowania, to znaczy sposobu podziału na bloki i kierunku predykcji.

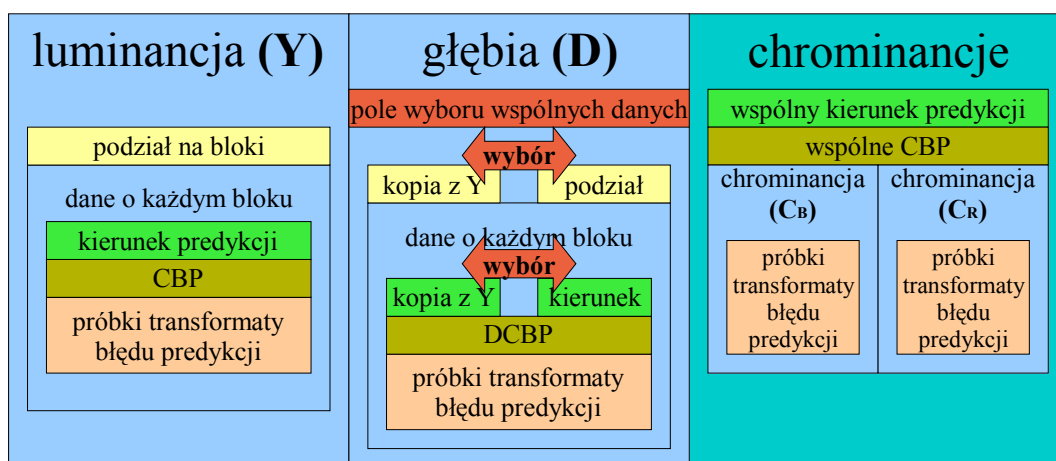


Rysunek 6.65. Schematyczne przedstawienie sposobu kodowania pojedynczego makrobloku dla kodowania czterokomponentowego – oddzielne parametry dla mapy głębi.

W celu ustalenia sposobu efektywnego zakodowania w strumieniu danych dotyczących głębi (czyli sposobu podziału na bloki, kierunków predykcji, pola DCBP), przeprowadzona została analiza badań, których wyniki przedstawione zostały w punkcie 6.6.1.

Przedstawiona dalej metoda ma na celu wykorzystanie korelacji pomiędzy wyborem sposobu

podziału makrobloku na bloki oraz kierunków predykcji dla luminancji i mapy głębi do zwiększenia efektywności kodowania czterokomponentowych sekwencji wielowidokowych. Wykorzystanie korelacji jest realizowane poprzez ograniczenie przesyłania oddzielnych danych dotyczących sposobu podziału na bloki i wybranych kierunków predykcji dla głębi jedynie do sytuacji, gdy parametry te są inne niż dla luminancji w tym samym makrobloku. Wiąże się z tym jednak konieczność przesłania dodatkowego pola określającego, czy dane o wyborze podziału bloku i o kierunku predykcji są wspólne dla kodowanej luminancji i głębi, czy też nie. Proponowana struktura danych dotyczących pojedynczego makrobloku w tak zmienionej formie zaprezentowana została na rysunku 6.66. Dane przesyłane dla luminancji i chrominancji nie ulegają zmianie, natomiast wśród danych dla głębi pojawia się dodatkowe pole, oznaczające, która część parametrów kodowania jest wspólna dla luminancji i mapy głębi. Jeśli sposób podziału na bloki dla głębi jest inny niż dla luminancji, w strumieniu umieszczana jest informacja o podziale wybranym dla mapy głębi. W przypadku, gdy pole wyboru wspólnych danych wskazuje na to, że podział na bloki dla głębi jest taki sam, jak dla luminancji, nie są przesyłane dane o sposobie podziału dla mapy głębi. W trakcie dekodowania używany jest wtedy ten sam sposób podziału na bloki co dla luminancji. Podobnie – jeśli pole wyboru wspólnych danych wskazuje na to, że nie ma konieczności przesyłania dla głębi odrębnych kierunków predykcji wewnątrzobrazowej, nie są one przesyłane w strumieniu, a podczas dekodowania wykorzystywane są takie same dane jakie zostały przesłane dla luminancji. W przeciwnym razie przesyłane jest odrębne pole niosące informacje o kierunkach predykcji dla mapy głębi.



Rysunek 6.66. Schematyczne przedstawienie sposobu kodowania pojedynczego makrobloku dla kodowania czterokomponentowego – oddzielne parametry dla mapy głębi, kodowane wspólnie z parametrami dla luminancji.

Odrębnego omówienia wymaga organizacja pola wyboru wspólnych danych, uwidocznionego na rysunku 6.66. W polu tym zawarta jest informacja o tym, które pola są kopiowane z danych o luminancji, a które przesyłane są oddzielnie dla mapy głębi. Zaproponowany sposób kodowania danych w tym polu przedstawiony jest w tabeli 6.1 poniżej. Wykorzystany został kod o zmiennej długości.

Tabela 6.1. Liczba bitów oraz kody przeznaczone do kodowania informacji w polu wyboru wspólnych danych

Rodzaj danych dla kodowanej mapy głębi		długość kodu	kod
Tryb podziału bloku danych	Kierunek predykcji	[bity]	
Taki sam jak dla luminancji	Taki sam jak dla luminancji	1	1
Taki sam jak dla luminancji	Inny niż dla luminancji (dosłany)	2	01
Inny niż dla luminancji (dosłany)	Taki sam jak dla luminancji	3	001
Inny niż dla luminancji (dosłany)	Inny niż dla luminancji (dosłany)	3	000

Zaprezentowane w tabeli 6.1 przyporządkowanie kodów pozwala na zmaksymalizowanie zysków kodowania mapy głębi wynikających z powtórnego użycia części danych wykorzystywanych przy kodowaniu luminancji. Najczęściej występującej dla dużych wartości parametrów QP i QD sytuacji przyporządkowany jest najkrótszy kod. Kolejne możliwe sytuacje mają przyporządkowane kody o większej długości, przy czym długość kodu jest największa dla najrzadziej występujących przypadków. Warto zauważyć, że najkrótszy kod jest przyporządkowany sytuacji pozwalającej na zaoszczędzenie największej liczby bitów strumienia, gdyż nie są dla niej przesyłane żadne odrębne parametry dla głębi.

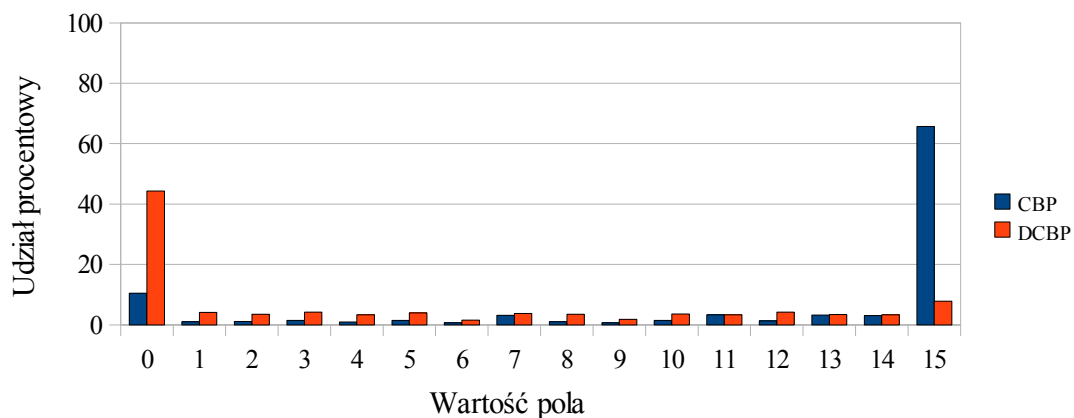
W przypadku, gdy luminancja kodowana jest z podziałem na bloki 4x4 punkty lub 8x8 punktów, niezwykle rzadko występuje przypadek, gdy kierunki predykcji dla luminancji i głębi są takie same. Wobec tego, w takim przypadku sygnalizowane jest jedynie, czy podział na bloki dla luminancji i głębi jest taki sam, czy inny. Kierunki predykcji dla składowej głębi wysyłane są w takim przypadku zawsze oddzielnie. Do sygnalizacji wykorzystane są wtedy jedynie dwa najkrótsze kody z tabeli 6.1.

Dodatkową zmianą, wprowadzoną do kodera czterokomponentowego, jest odmienne kodowanie pola DCBP. Pole CBP dla luminancji kodowane jest z wykorzystaniem ściśle przyporządkowanych kodów, których długość zależy od częstości występowania danego kodu w strumieniu. Kodowane są w nim zarówno informacje o luminancji, jak i o chrominancji. Przypadkom częściej występującym przyporządkowane są kody krótsze. Możliwych jest 48 różnych wartości dla pola CBP – 16 możliwości dla luminancji i 3 możliwości dla chrominancji w dowolnej kombinacji.

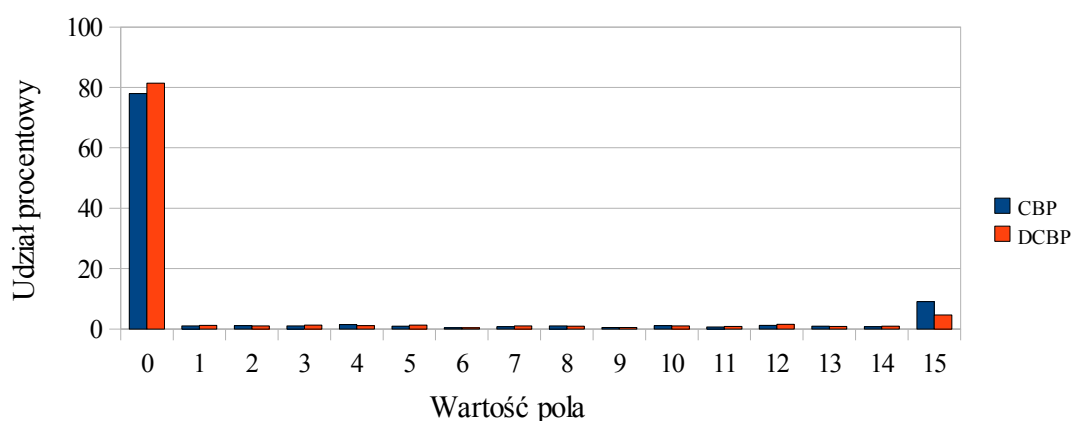
Jedynka w zapisie binarnym na jednej z 4 pozycji przeznaczonych dla luminancji oznacza obecność niezerowych próbek transformaty błędu predykcji dla odpowiadającego bloku 8x8 punktów. W tabeli 6.2 zestawione zostały kody przyporządkowane kolejnym wartościom pola CBP dla kodowania luminancji w trybie wewnątrzobrazowym. Warty odnotowania jest fakt, że najkrótsze trzy kody przyporządkowane są sytuacji, gdy wszystkie bloki 8x8 punktów zawierają niezerowe próbki transformaty błędu predykcji (są to kody o wartościach 15, 31, 47). Przeprowadzone badania wykazały, że istotnie w przypadku kodowania luminancji są to najczęściej wybierane kody. Jednak w przypadku kodowania głębi rozkład prawdopodobieństwa wystąpienia wartości DCBP jest zupełnie inny. Prezentują to rysunki od 6.67 do 6.70, na których przedstawiona jest częstość wystąpienia poszczególnych czterobitowych wartości (dwa najstarsze bity kodów, na których kodowane są dane chrominancji, są pomijane) otrzymanych w trakcie eksperymentów prowadzonych na sekwencjach testowych. Widać wyraźnie, że dla kodowania komponentu głębi jako wartość pola DCBP najczęściej wybierana jest wartość 0. Druga w kolejności wybierana jest wartość 15. Pozostałe możliwe wartości wybierane są ze znacznie niższym prawdopodobieństwem, zbliżonym dla wszystkich czternastu pozostałych możliwości.

Tabela 6.2. Sposób kodowania pola CBP dla luminancji i chrominancji w koderze MVC. Opracowano na podstawie [AVC].

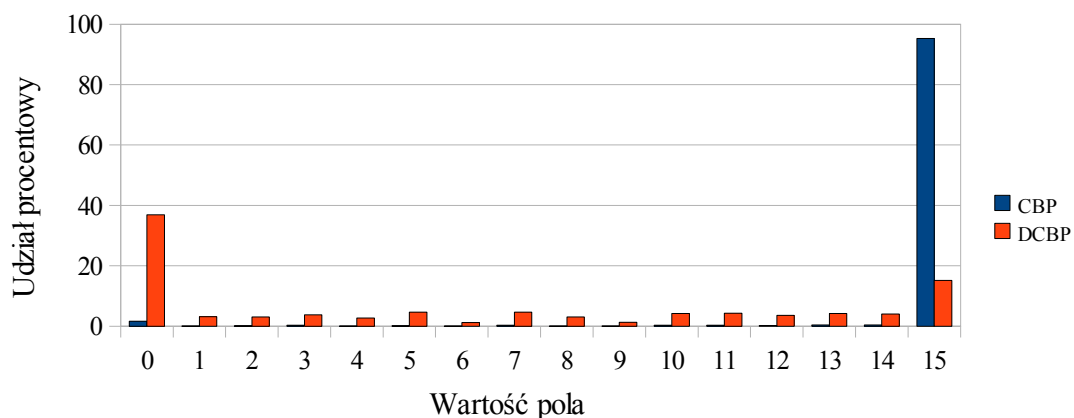
wartość CBP	chrominancje	luminancja	przyporządkowany kod	długość kodu [bity]	kod (binarnie)
0	00	0000	3	5	00100
1	00	0001	29	9	000011110
2	00	0010	30	9	000011111
3	00	0011	17	9	000010010
4	00	0100	31	11	00000100000
5	00	0101	18	9	000010011
6	00	0110	37	11	00000100110
7	00	0111	8	7	0001001
8	00	1000	32	11	00000100001
9	00	1001	38	11	00000100111
10	00	1010	19	9	000010100
11	00	1011	9	7	0001010
12	00	1100	20	9	000010101
13	00	1101	10	7	0001011
14	00	1110	11	7	0001100
15	00	1111	2	3	011
16	01	0000	16	9	000010001
17	01	0001	33	11	00000100010
18	01	0010	34	11	00000100011
19	01	0011	21	9	000010110
20	01	0100	35	11	00000100100
21	01	0101	22	9	000010111
22	01	0110	39	11	00000101000
23	01	0111	4	5	00101
24	01	1000	36	11	00000100101
25	01	1001	40	11	00000101001
26	01	1010	23	9	000011000
27	01	1011	5	5	00110
28	01	1100	24	9	000011001
29	01	1101	6	5	00111
30	01	1110	7	7	0001000
31	01	1111	1	3	010
32	10	0000	41	11	00000101010
33	10	0001	42	11	00000101011
34	10	0010	43	11	00000101100
35	10	0011	25	9	000011010
36	10	0100	44	11	00000101101
37	10	0101	26	9	000011011
38	10	0110	46	11	00000101111
39	10	0111	12	7	0001101
40	10	1000	45	11	00000101110
41	10	1001	47	11	00000110000
42	10	1010	27	9	000011100
43	10	1011	13	7	0001110
44	10	1100	28	9	000011101
45	10	1101	14	7	0001111
46	10	1110	15	9	000010000
47	10	1111	0	1	1



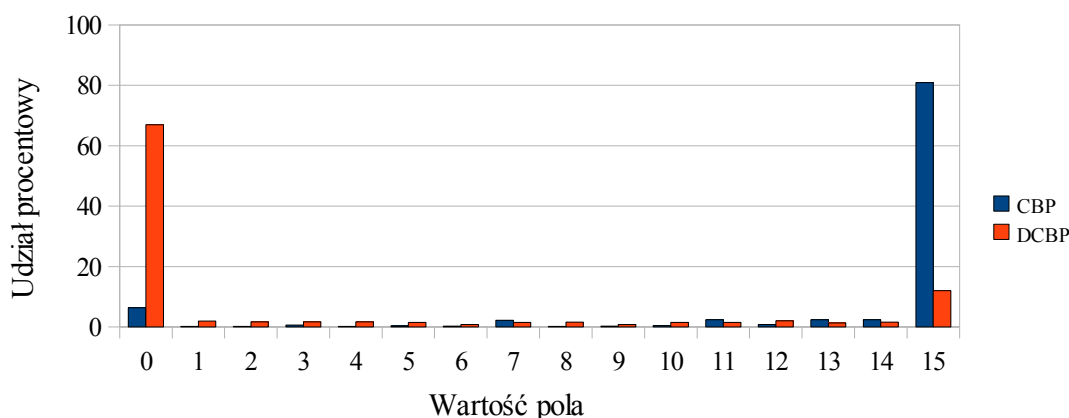
Rysunek 6.67. Udział procentowy bloków kodowanych z użyciem określonej wartości pola CBP i DCBP. Sekwencja Book Arrival, QP = 20, QD = 20, bloki 8x8 punktów włączone.



Rysunek 6.68. Udział procentowy bloków kodowanych z użyciem określonej wartości pola CBP i DCBP. Sekwencja Book Arrival, QP = 40, QD = 30, bloki 8x8 punktów wyłączone.



Rysunek 6.69. Udział procentowy bloków kodowanych z użyciem określonej wartości pola CBP i DCBP. Sekwencja Newspaper, QP = 10, QD = 10, bloki 8x8 punktów włączone.



Rysunek 6.70. Udział procentowy bloków kodowanych z użyciem określonej wartości pola CBP i DCBP. Sekwencja Poznan Street, QP = 20, QD = 20, bloki 8x8 punktów wyłączone.

W wyniku analizy otrzymanych wyników zaproponowany został inny sposób przyporządkowania kodów wartościom z pola DCBP. Sposób ten uwzględnia zmienione statystyki występowania wartości oraz uwzględnia brak konieczności kodowania danych dla chrominancji. Proponowany sposób kodowania zaprezentowany został w tabeli 6.3.

Tabela 6.3. Proponowany sposób kodowania pola DCBP dla głębi.

wartość CBP	głębia	przyporządkowany kod	długość kodu [bity]	kod (binarnie)
0	0000	0	1	1
1	0001	2	3	011
2	0010	3	5	00100
3	0011	15	5	00001
4	0100	4	5	00101
5	0101	14	7	0001111
6	0110	13	7	0001110
7	0111	5	5	00110
8	1000	6	5	00111
9	1001	12	7	0001101
10	1010	11	7	0001100
11	1011	7	7	0001000
12	1100	10	7	0001011
13	1101	8	7	0001001
14	1110	9	7	0001010
15	1111	1	3	010

Warto przytoczyć w tym miejscu wyniki kodowania pól DCBP z użyciem zmodyfikowanego i niezmodyfikowanego sposobu przyporządkowania słów kodowych. Wyniki podane są dla tych samych przypadków, które przedstawione zostały na rysunkach od 6.67 do 6.70. Zamieszczone w tabeli 6.4 wyniki ukazują średnią liczbę bitów potrzebnych do zakodowania pola DCBP z użyciem standardowej metody kodowania pola CBP i proponowanej metody przyporządkowania kodów.

Tabela 6.4. Porównanie sposobów kodowania pola DCBP

Użyty przykład	Średnia liczba bitów przypadająca na jedno pole DCBP	
	Metoda używana dla CBP	Proponowana metoda
Przykład z rysunku 6.67	6,68	3,63
Przykład z rysunku 6.68	5,45	1,86
Przykład z rysunku 6.69	6,44	4,02
Przykład z rysunku 6.70	5,58	2,48

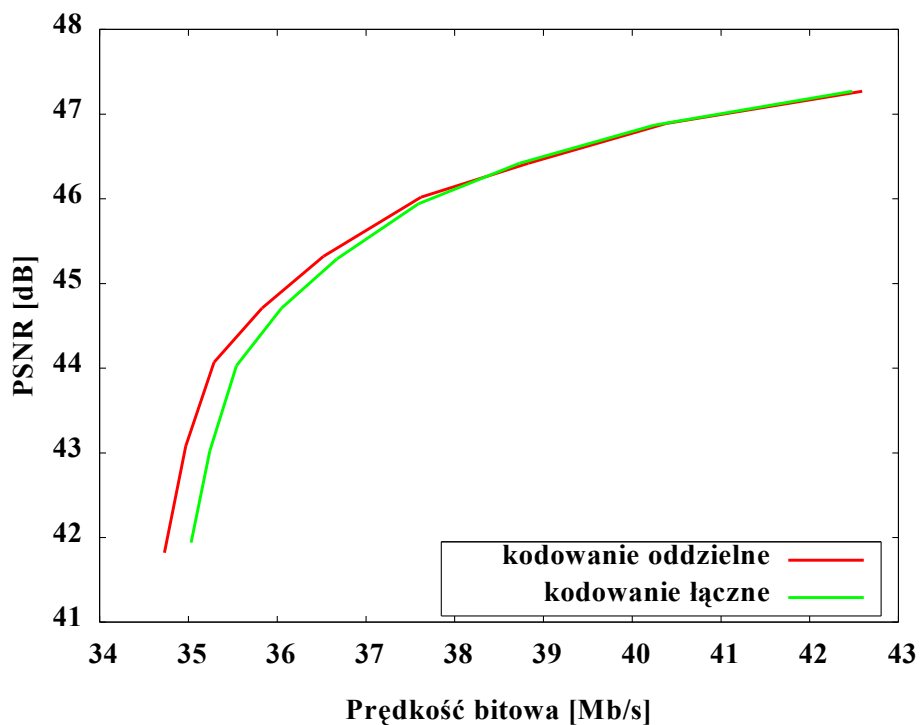
W celu sprawdzenia efektywności kodowania sekwencji czterokomponentowych w opisany sposób, przeprowadzono kodowania sekwencji testowych. Wyniki uzyskane dla 3 różnych wartości stałego parametru: 16, 34 i 40 przedstawione są w punktach 6.6.3.1 oraz 6.6.3.2. Na rysunkach przedstawiono wyniki jedynie dla wybranych sekwencji. Wykorzystano miarę jakości PSNR, dla której jako obraz odniesienia wykorzystano obraz syntezowany z użyciem nieskompresowanych obrazów i głębi. Podane prędkości bitowe są wartościami sumarycznymi dla 2 widoków i 2 map głębi.

### 6.6.3.1. Wyniki dla stałego QP

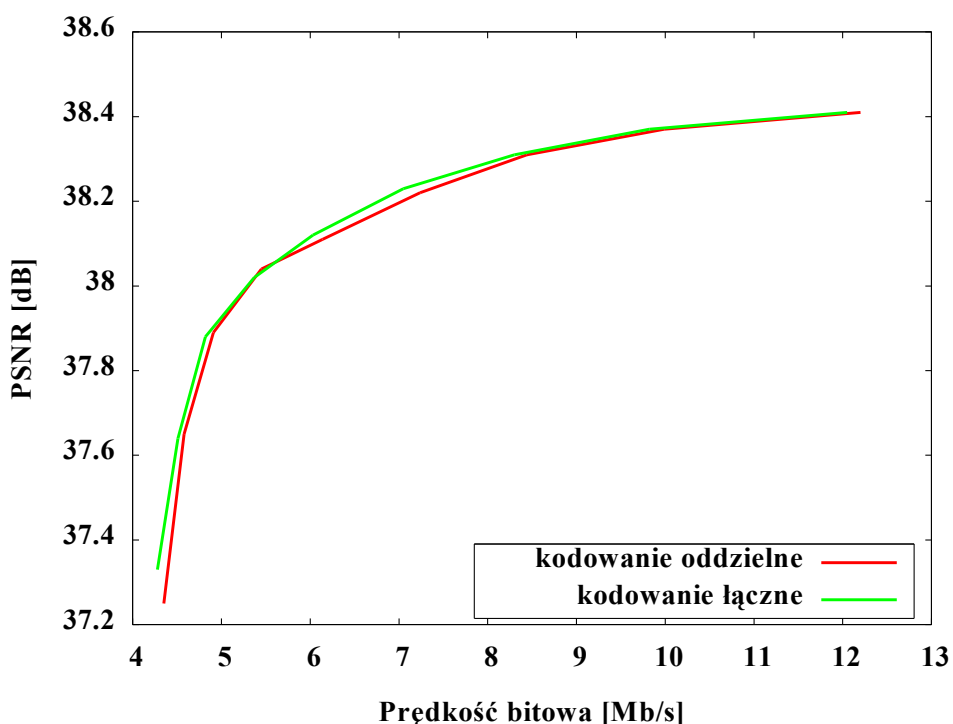
Wyniki eksperymentu dla stałej wartości QP przedstawione są na wykresach od 6.71 do 6.88.

1. Kompresja z wykorzystaniem podziału makrobloku na bloki o rozmiarach 16x16, 8x8 lub 4x4 punkty dla stałego QP i zmieniającego się indeksu QD. Indeks QD przyjmował kolejne wartości z listy: 16, 19, 22, 25, 28, 31, 34, 37, 40 (rysunki od 6.71 do 6.79).
2. Kompresja z wykorzystaniem podziału makrobloku na bloki o rozmiarach 16x16 lub 4x4 punkty dla stałego QP i zmieniającego się indeksu QD. Indeks QD przyjmował kolejne wartości z listy: 16, 19, 22, 25, 28, 31, 34, 37, 40 (rysunki od 6.80 do 6.88).

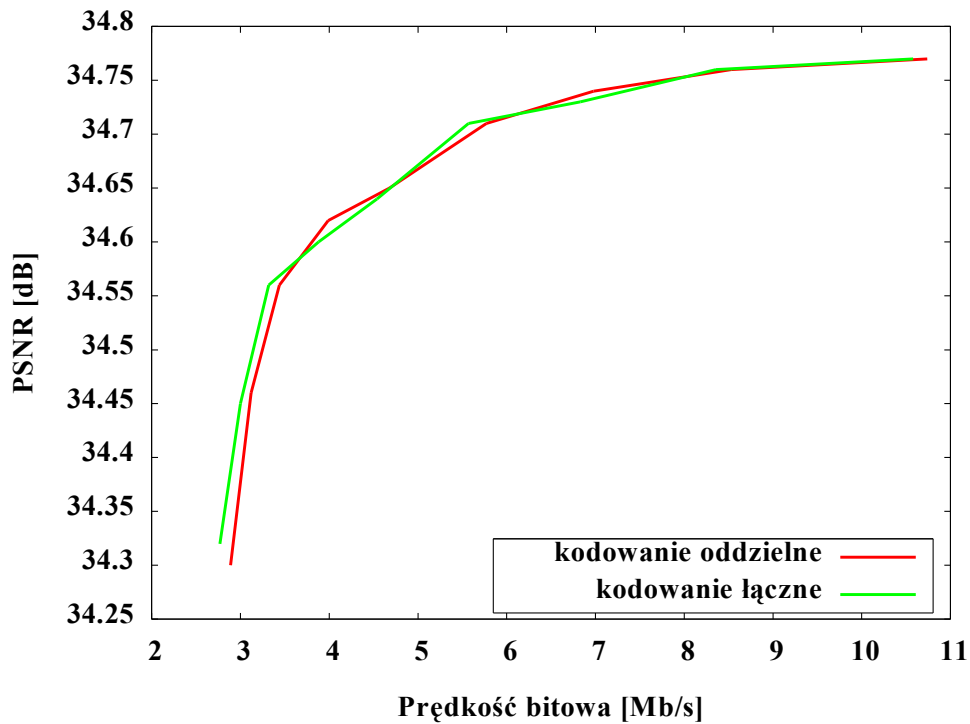




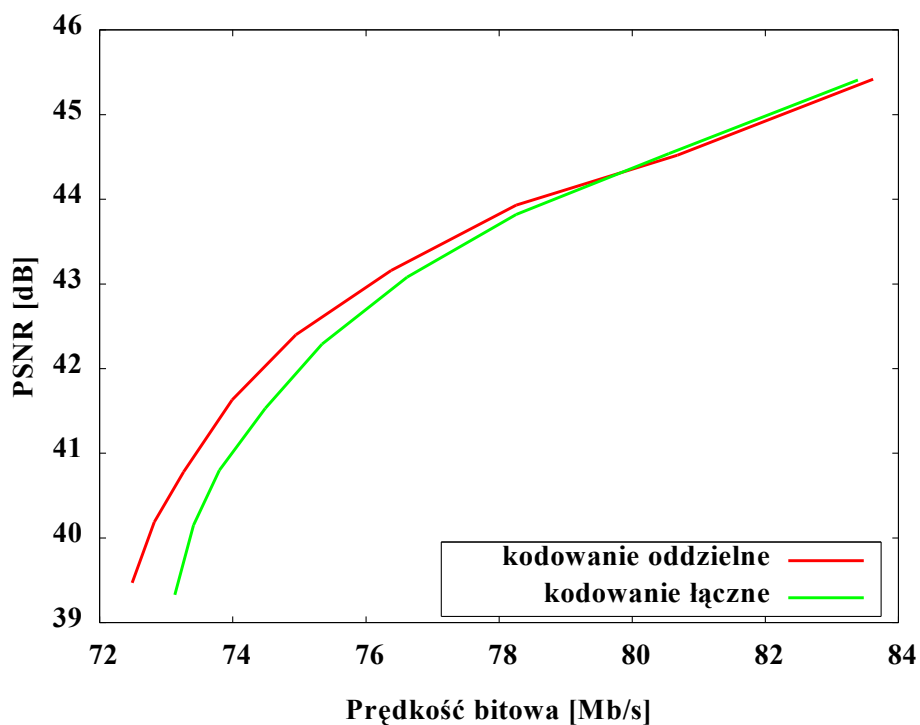
Rysunek 6.71. Porównanie efektywności kompresji kodowania czterokomponentowego do oddzielnego kodowania obrazów i map głębi. Kodowanie z możliwym podziałem na bloki 16x16, 8x8 i 4x4. Sekwencja Book Arrival, QP = 16, QD zmieniane w zakresie od 16 do 40.



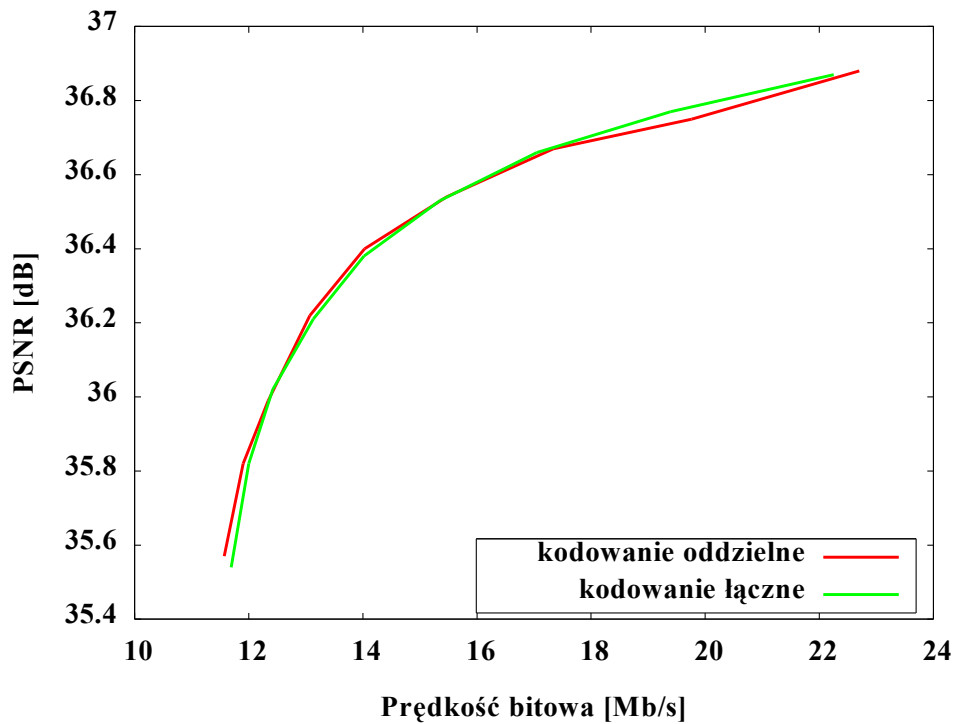
Rysunek 6.72. Porównanie efektywności kompresji kodowania czterokomponentowego do oddzielnego kodowania obrazów i map głębi. Kodowanie z możliwym podziałem na bloki 16x16, 8x8 i 4x4. Sekwencja Book Arrival, QP = 34, QD zmieniane w zakresie od 16 do 40.



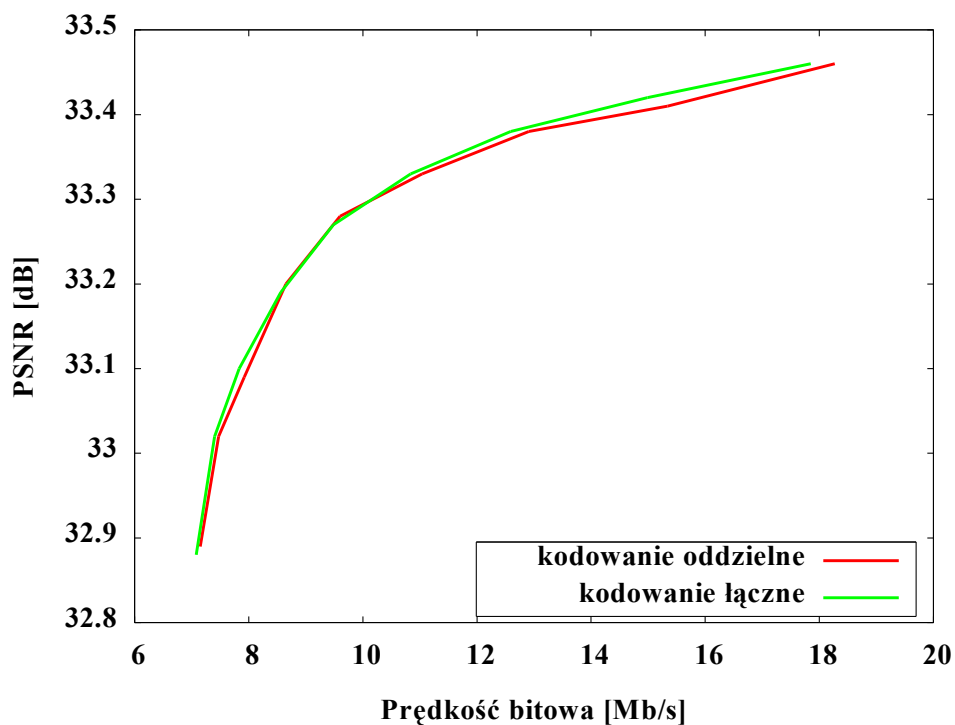
Rysunek 6.73. Porównanie efektywności kompresji kodowania czterokomponentowego do oddzielnego kodowania obrazów i map głębi. Kodowanie z możliwym podziałem na bloki 16x16, 8x8 i 4x4. Sekwencja Book Arrival, QP = 40, QD zmieniane w zakresie od 16 do 40.



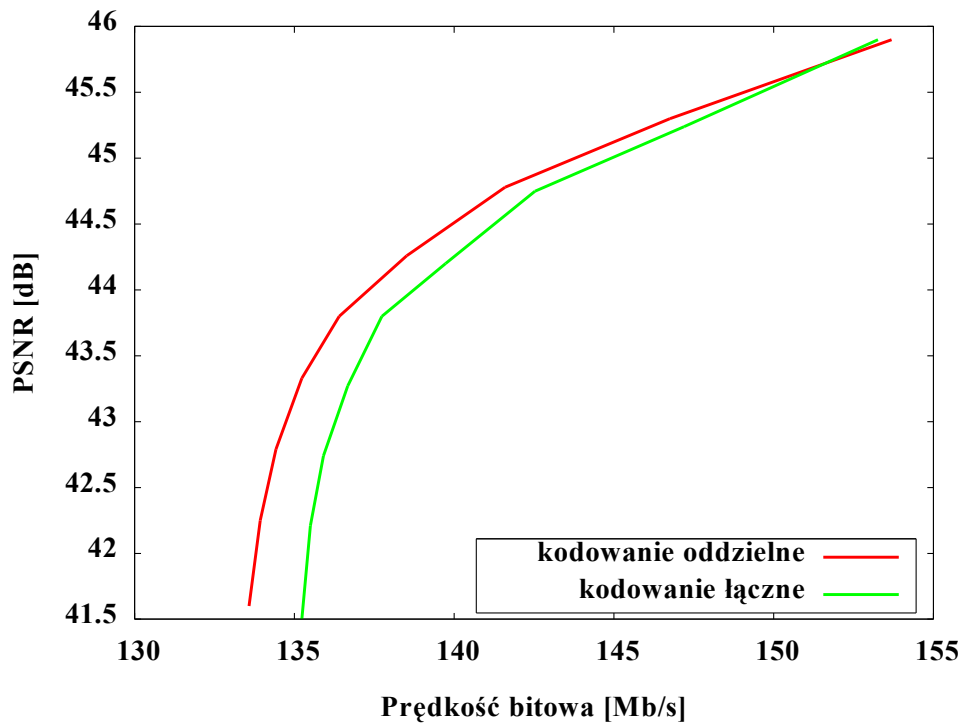
Rysunek 6.74. Porównanie efektywności kompresji kodowania czterokomponentowego do oddzielnego kodowania obrazów i map głębi. Kodowanie z możliwym podziałem na bloki 16x16, 8x8 i 4x4. Sekwencja Newspaper, QP = 16, QD zmieniane w zakresie od 16 do 40.



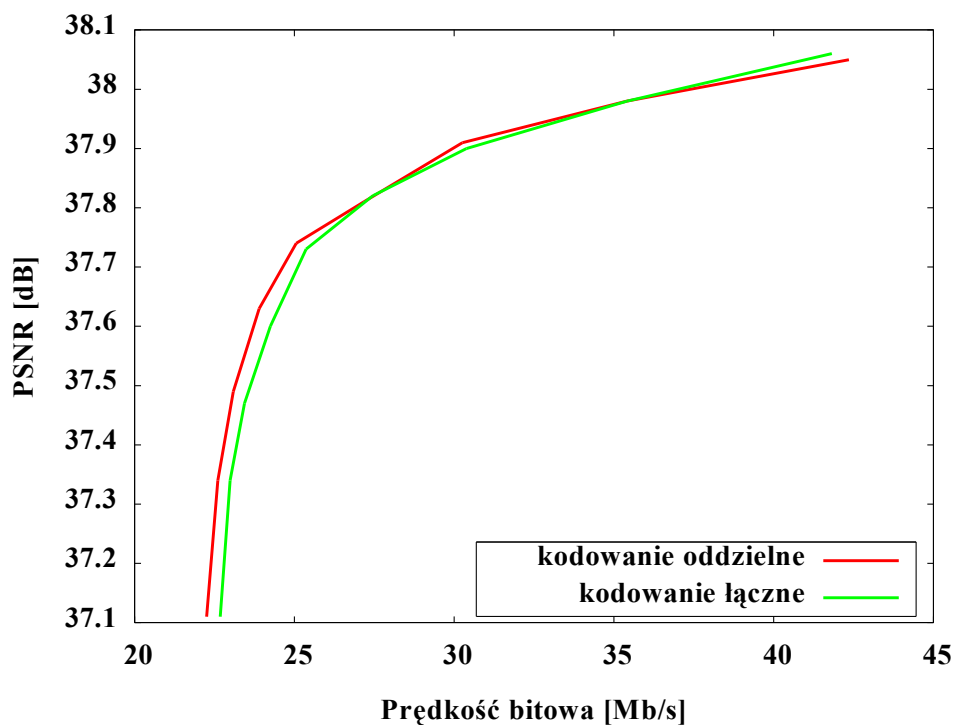
Rysunek 6.75. Porównanie efektywności kompresji kodowania czterokomponentowego do oddzielnego kodowania obrazów i map głębi. Kodowanie z możliwym podziałem na bloki 16x16, 8x8 i 4x4. Sekwencja Newspaper, QP = 34, QD zmieniane w zakresie od 16 do 40.



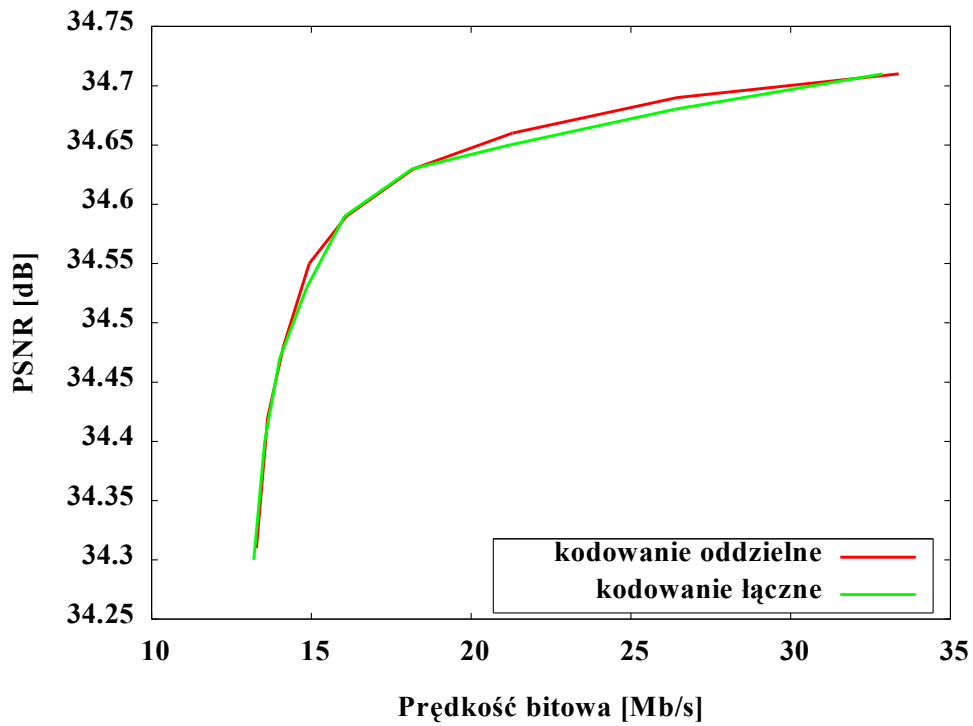
Rysunek 6.76. Porównanie efektywności kompresji kodowania czterokomponentowego do oddzielnego kodowania obrazów i map głębi. Kodowanie z możliwym podziałem na bloki 16x16, 8x8 i 4x4. Sekwencja Newspaper, QP = 40, QD zmieniane w zakresie od 16 do 40.



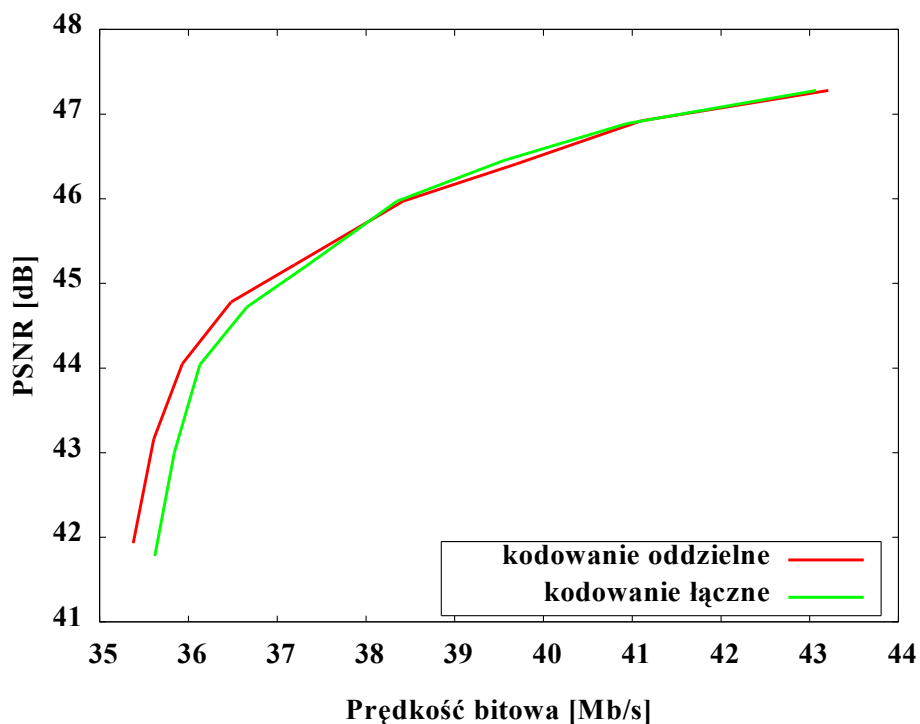
Rysunek 6.77. Porównanie efektywności kompresji kodowania czterokomponentowego do oddzielnego kodowania obrazów i map głębi. Kodowanie z możliwym podziałem na bloki 16x16, 8x8 i 4x4. Sekwencja Poznan Street, QP = 16, QD zmieniane w zakresie od 16 do 40.



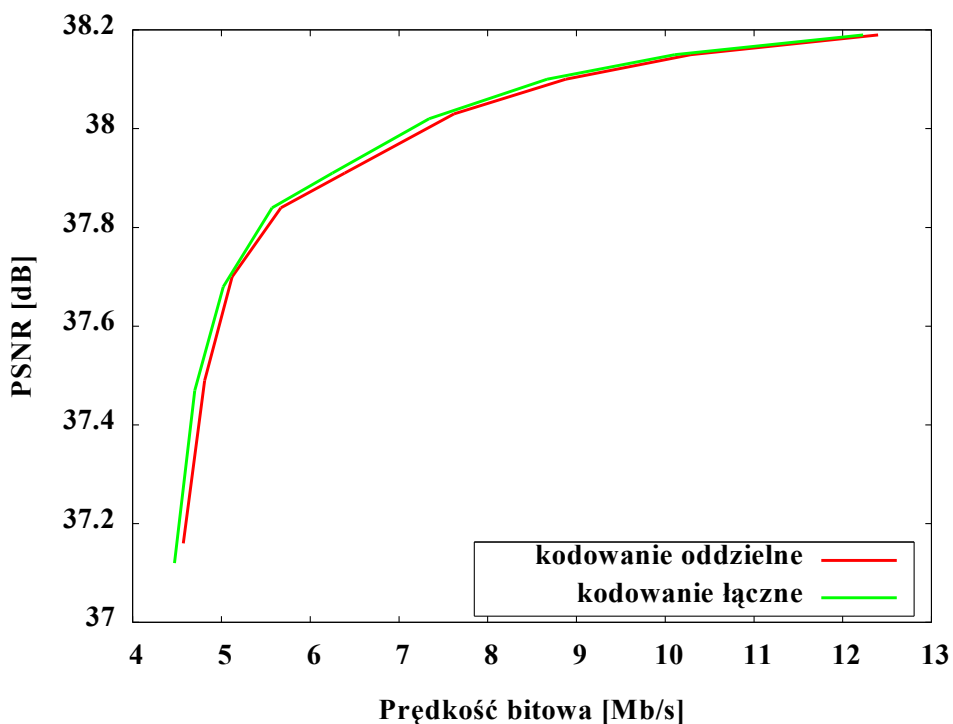
Rysunek 6.78. Porównanie efektywności kompresji kodowania czterokomponentowego do oddzielnego kodowania obrazów i map głębi. Kodowanie z możliwym podziałem na bloki 16x16, 8x8 i 4x4. Sekwencja Poznan Street, QP = 34, QD zmieniane w zakresie od 16 do 40.



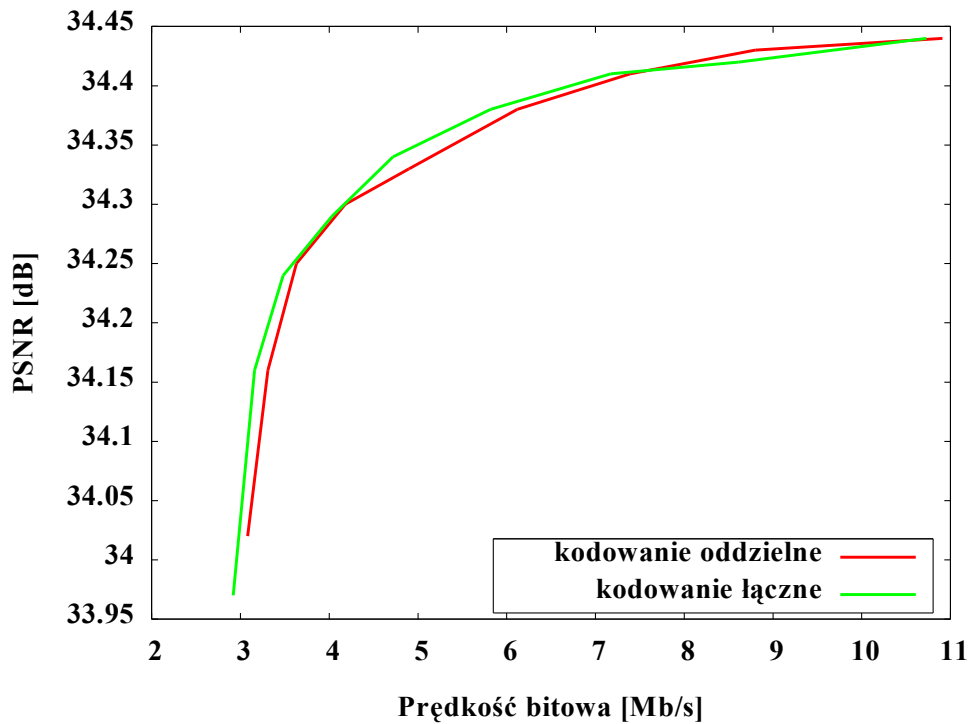
Rysunek 6.79. Porównanie efektywności kompresji kodowania czterokomponentowego do oddzielnego kodowania obrazów i map głębi. Kodowanie z możliwym podziałem na bloki 16x16, 8x8 i 4x4. Sekwencja Poznan Street, QP = 40, QD zmieniane w zakresie od 16 do 40.



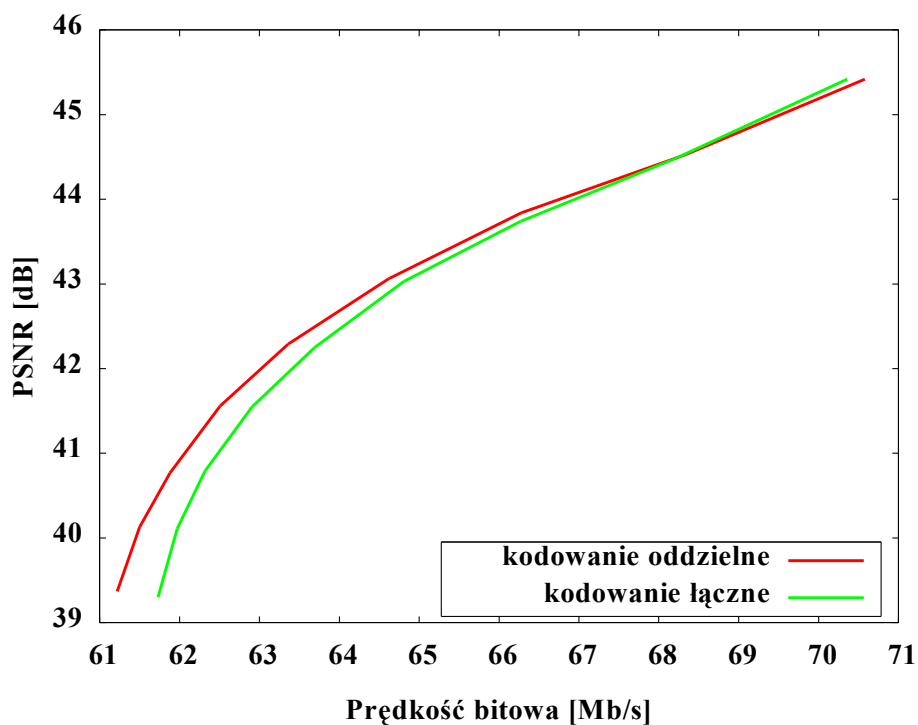
Rysunek 6.80. Porównanie efektywności kompresji kodowania czterokomponentowego do oddzielnego kodowania obrazów i map głębi. Kodowanie z możliwym podziałem na bloki 16x16 i 4x4. Sekwencja Book Arrival, QP = 16, QD zmieniane w zakresie od 16 do 40.



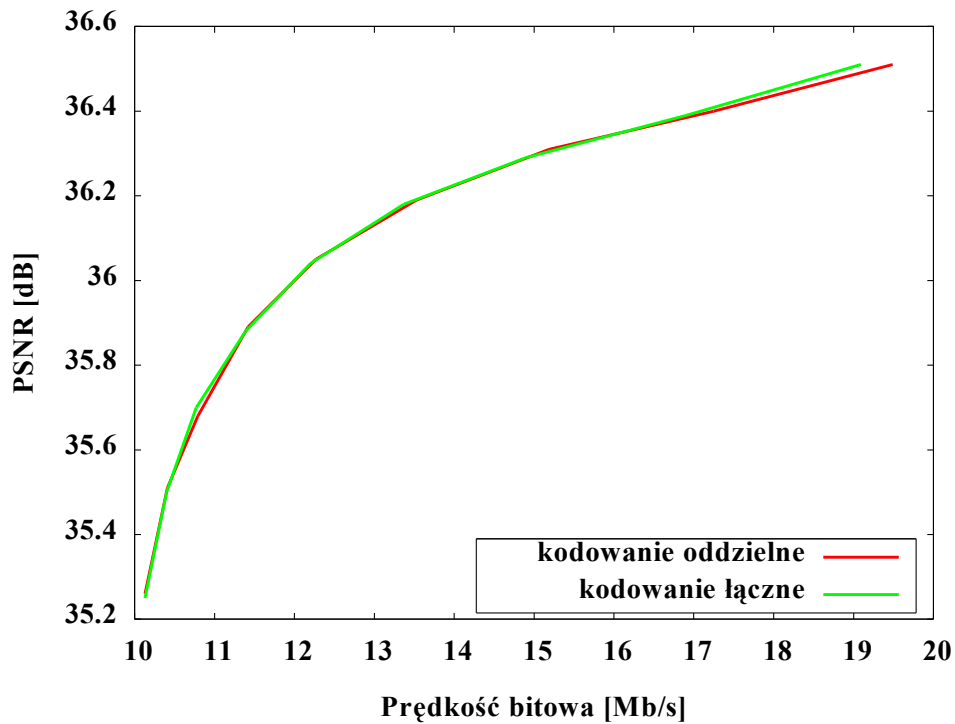
Rysunek 6.81. Porównanie efektywności kompresji kodowania czterokomponentowego do oddzielnego kodowania obrazów i map głębi. Kodowanie z możliwym podziałem na bloki 16x16 i 4x4. Sekwencja Book Arrival, QP = 34, QD zmieniane w zakresie od 16 do 40.



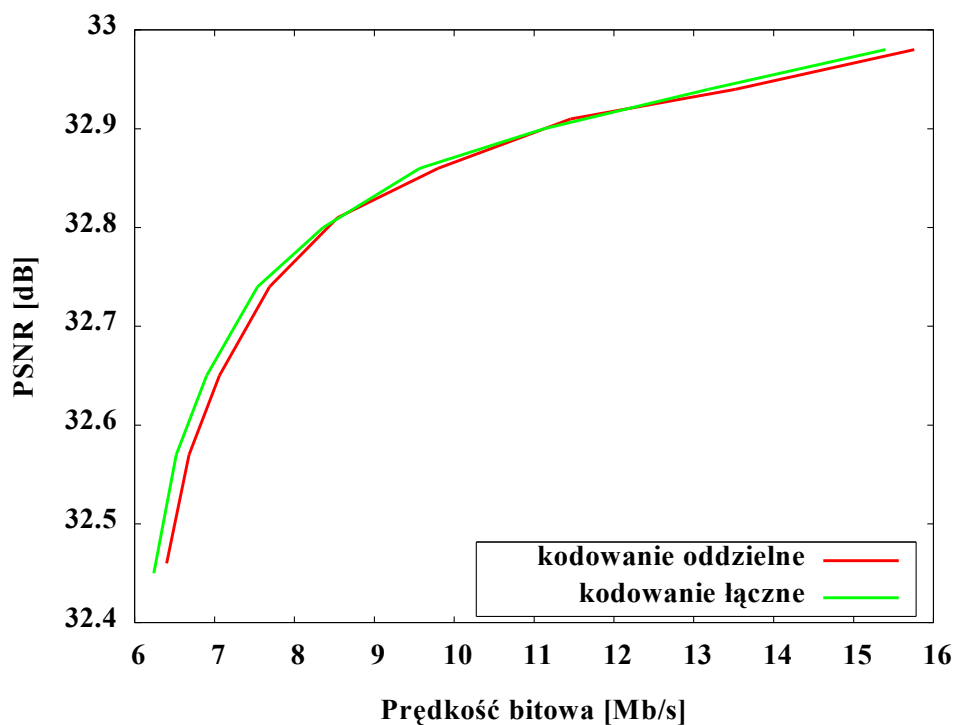
Rysunek 6.82. Porównanie efektywności kompresji kodowania czterokomponentowego do oddzielnego kodowania obrazów i map głębi. Kodowanie z możliwym podziałem na bloki 16x16 i 4x4. Sekwencja Book Arrival, QP = 40, QD zmieniane w zakresie od 16 do 40.



Rysunek 6.83. Porównanie efektywności kompresji kodowania czterokomponentowego do oddzielnego kodowania obrazów i map głębi. Kodowanie z możliwym podziałem na bloki 16x16 i 4x4. Sekwencja Newspaper, QP = 16, QD zmieniane w zakresie od 16 do 40.

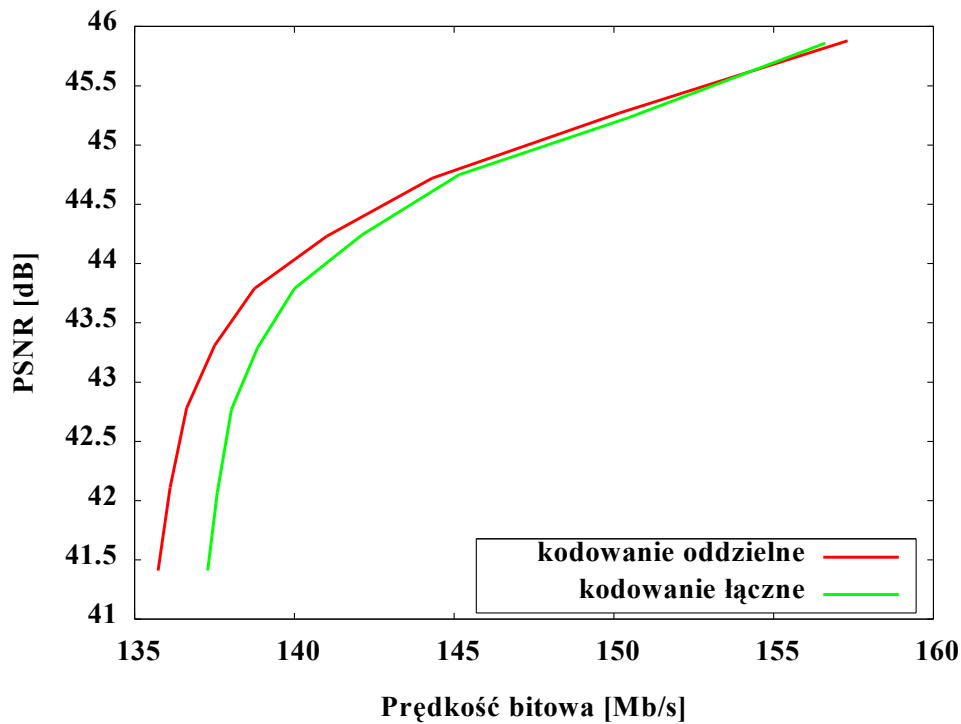


Rysunek 6.84. Porównanie efektywności kompresji kodowania czterokomponentowego do oddzielnego kodowania obrazów i map głębi. Kodowanie z możliwym podziałem na bloki 16x16 i 4x4. Sekwencja Newspaper, QP = 34, QD zmieniane w zakresie od 16 do 40.

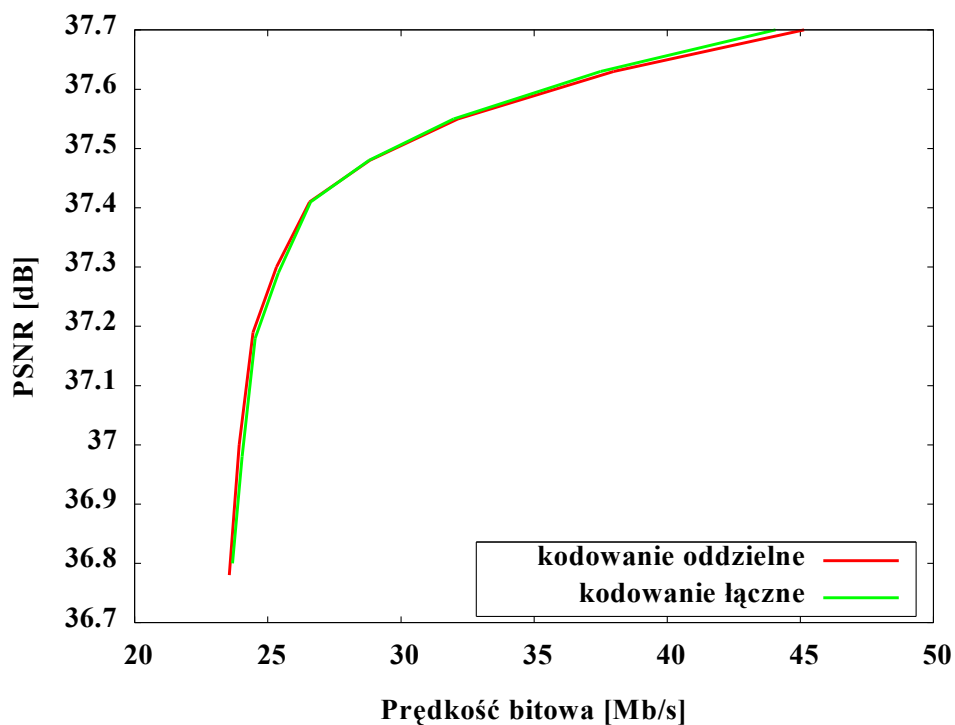


Rysunek 6.85. Porównanie efektywności kompresji kodowania czterokomponentowego do oddzielnego kodowania obrazów i map głębi. Kodowanie z możliwym podziałem na bloki 16x16 i 4x4. Sekwencja Newspaper, QP = 40, QD zmieniane w zakresie od 16 do 40.

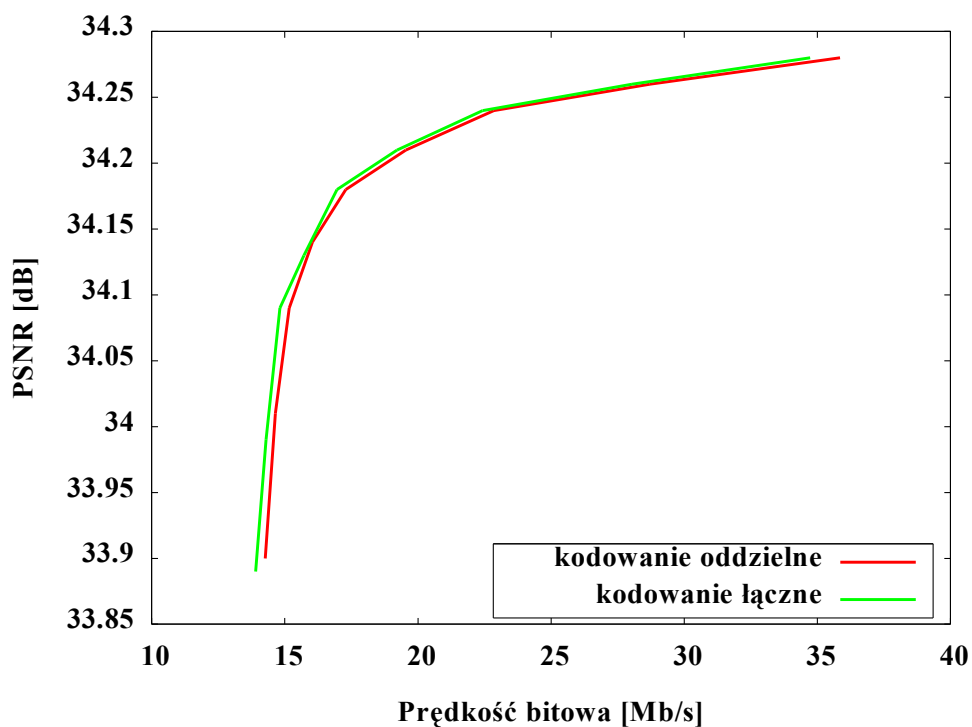




Rysunek 6.86. Porównanie efektywności kompresji kodowania czterokomponentowego do oddzielnego kodowania obrazów i map głębi. Kodowanie z możliwym podziałem na bloki 16x16 i 4x4. Sekwencja Poznan Street, QP = 16, QD zmieniane w zakresie od 16 do 40.



Rysunek 6.87. Porównanie efektywności kompresji kodowania czterokomponentowego do oddzielnego kodowania obrazów i map głębi. Kodowanie z możliwym podziałem na bloki 16x16 i 4x4. Sekwencja Poznan Street, QP = 34, QD zmieniane w zakresie od 16 do 40.



Rysunek 6.88. Porównanie efektywności kompresji kodowania czterokomponentowego do oddzielnego kodowania obrazów i map głębi. Kodowanie z możliwym podziałem na bloki 16x16 i 4x4. Sekwencja Poznan Street, QP = 40, QD zmieniane w zakresie od 16 do 40.

Na wykresach z rysunków od 6.71 do 6.88 przedstawiono wyniki uzyskane w przypadku zastosowania stałej wartości indeksu QP i zmieniania wartości indeksu QD. Dla każdego przypadku przedstawionego na oddzielnym wykresie jakość obrazów jest zatem stała, a zmianie ulega jedynie jakość i prędkość bitowa map głębi.

W przypadkach, gdy wartość indeksu QP wynosiła 16, dla wszystkich sekwencji obserwowana była strata rzędu 1% prędkości strumienia bitowego związana z użyciem kodera czterokomponentowego w porównaniu do oddzielnego kodowania map głębi i obrazów. Strata obserwowana była niezależnie od tego, czy podział na bloki 8x8 był włączony, czy nie.

W przypadkach, gdy użyta została wartość QP równa 34, obserwowany był zysk wynoszący około 2% dla sekwencji Book Arrival.

W przypadkach, gdy użyta została wartość QP równa 40, obserwowany był zysk dla wszystkich przebadanych sekwencji. Zysk wyniósł do 4% dla sekwencji Book Arrival. W przypadku, gdy podział na bloki 8x8 punktów był wyłączony, obserwowany zysk był o około 1 punkt procentowy większy niż dla przypadków, gdy podział na te bloki był włączony. Wytłumaczeniem tego zjawiska jest fakt, że w przypadku, gdy podział na bloki 8x8 jest wyłączony, zachodzi większe prawdopodobieństwo, że bloki luminancji i głębi kodowane będą z użyciem bloku o tym samym

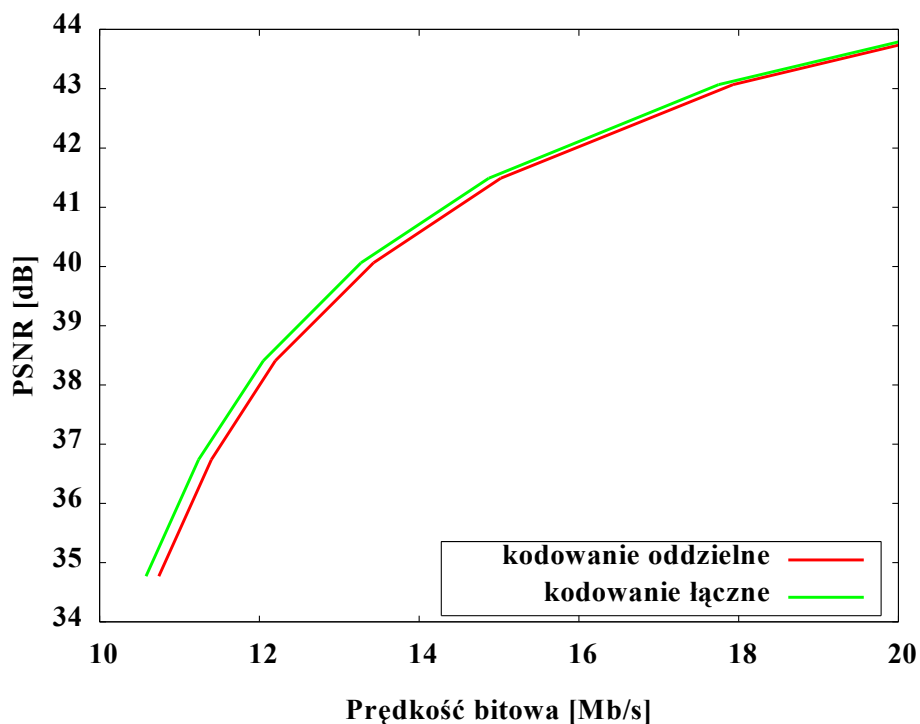
rozmiarze – jak widać to na rysunkach 6.28 i 6.29 - na rysunku 6.29 liczba bloków kodowanych z użyciem tego samego podziału dla obrazu i głębi jest o około 10% mniejsza.

Wyniki świadczą o tym, że proponowana metoda kompresji czterekomponentowej może dawać, dla części przebadanych w tej konfiguracji przypadków, wzrost efektywności kodowania. Wzrost efektywności kodowania obserwowany jest w przypadkach, gdy obraz i głębia kodowane są z większymi wartościami parametrów QP i QD. Wzrost jest wyraźniejszy dla przypadków, w których podział na bloki o wymiarach 8x8 punktów pozostaje wyłączony.

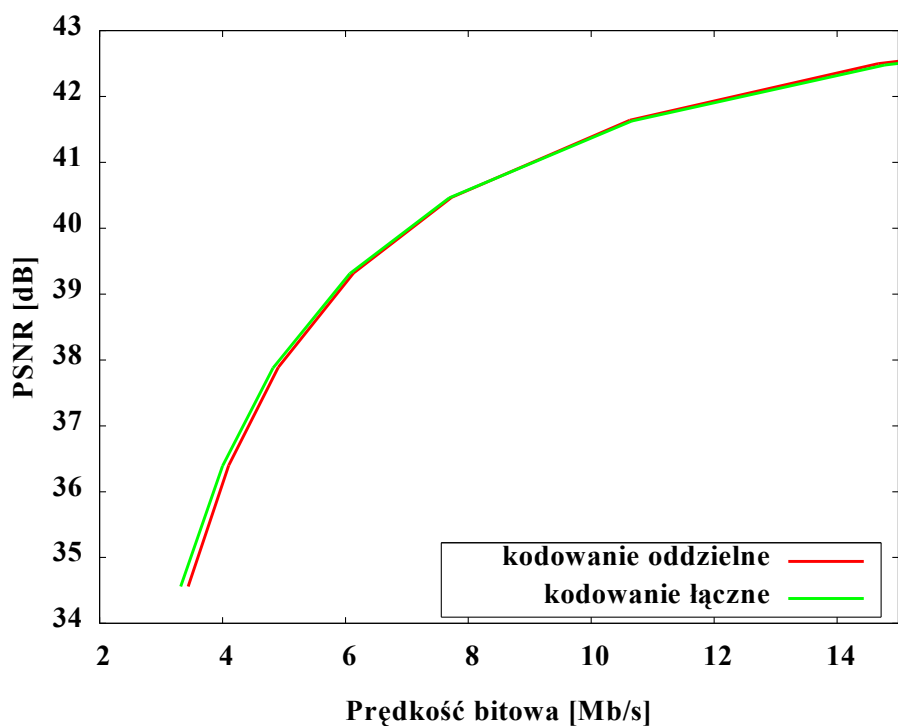
### **6.6.3.2. Wyniki dla stałego QD**

W następnej kolejności wyniki przedstawione są w innym ujęciu (rysunki od 6.89 do 6.106) oraz w nieco zawężonej skali – pokazane są wyniki dla małych prędkości bitowych :

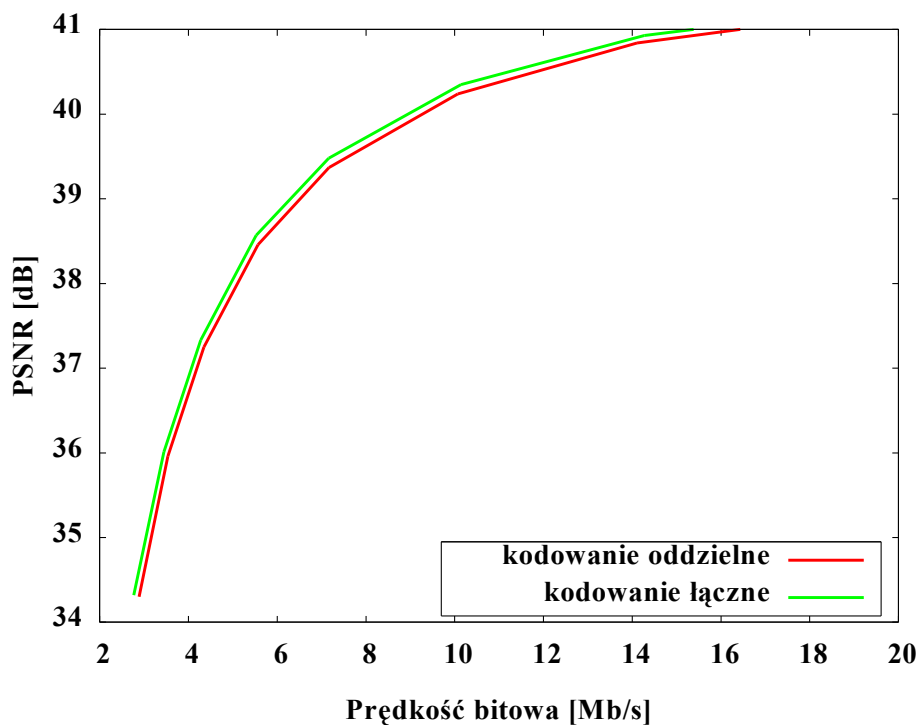
1. Kompresja z wykorzystaniem podziału makrobloku na bloki o rozmiarach 16x16, 8x8 lub 4x4 punkty dla stałego QD i zmieniającego się indeksu QP. Indeks QP przyjmował kolejne wartości z listy: 16, 19, 22, 25, 28, 31, 34, 37, 40 (rysunki od 6.89 do 6.97).
2. Kompresja z wykorzystaniem podziału makrobloku na bloki o rozmiarach 16x16 lub 4x4 punkty dla stałego QD i zmieniającego się indeksu QP. Indeks QP przyjmował kolejne wartości z listy: 16, 19, 22, 25, 28, 31, 34, 37, 40 (rysunki od 6.98 do 6.106).



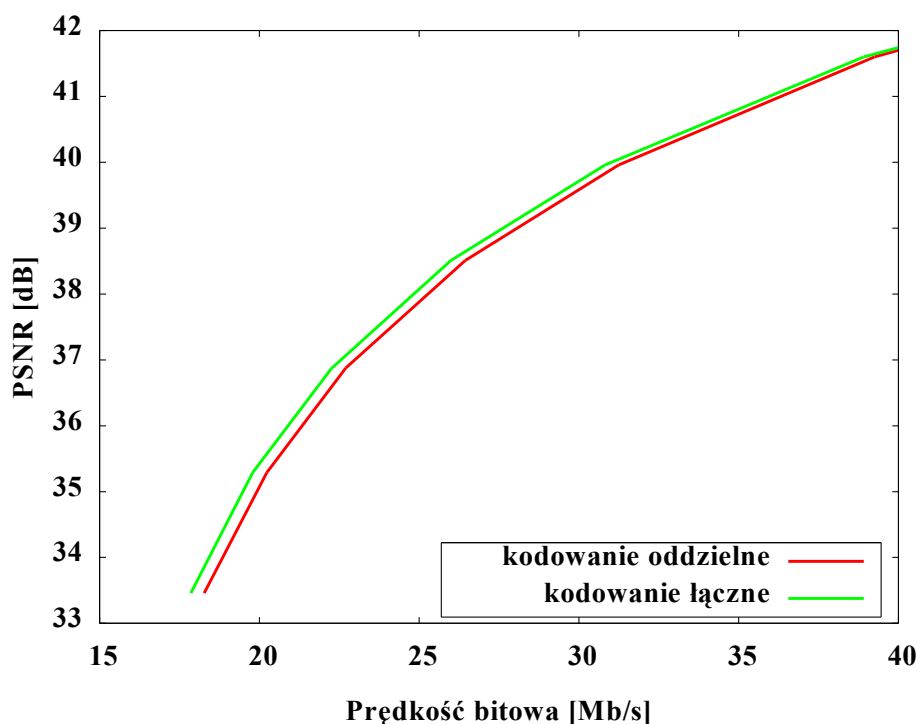
Rysunek 6.89. Porównanie efektywności kompresji kodowania czterokomponentowego do oddzielnego kodowania obrazów i map głębi. Kodowanie z możliwym podziałem na bloki 16x16, 8x8 i 4x4. Sekwencja Book Arrival, QD = 16, QP zmieniane w zakresie od 16 do 40.



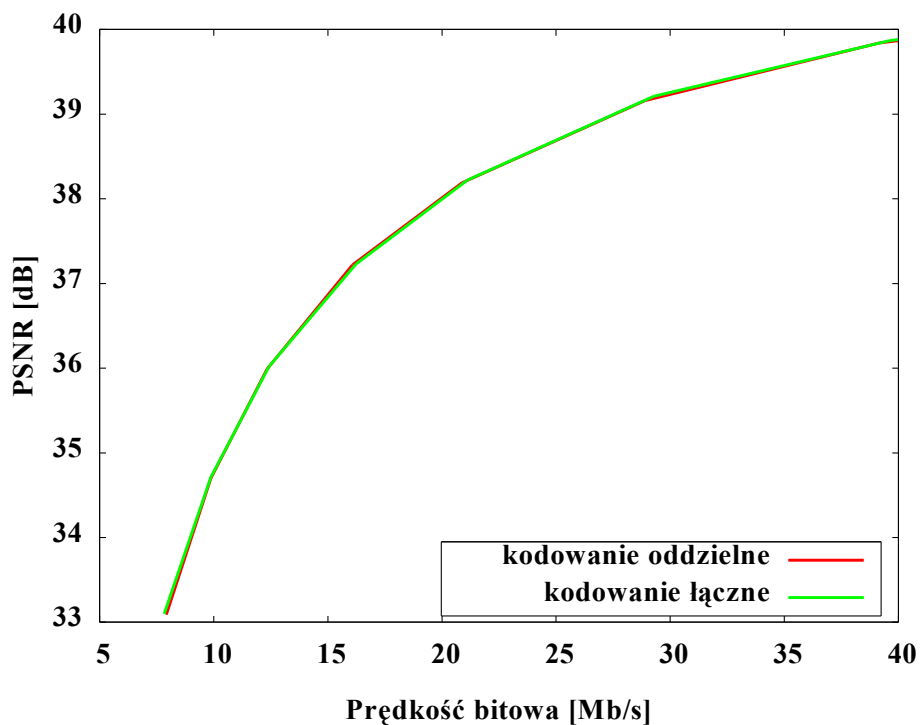
Rysunek 6.90. Porównanie efektywności kompresji kodowania czterokomponentowego do oddzielnego kodowania obrazów i map głębi. Kodowanie z możliwym podziałem na bloki 16x16, 8x8 i 4x4. Sekwencja Book Arrival, QD = 34, QP zmieniane w zakresie od 16 do 40.



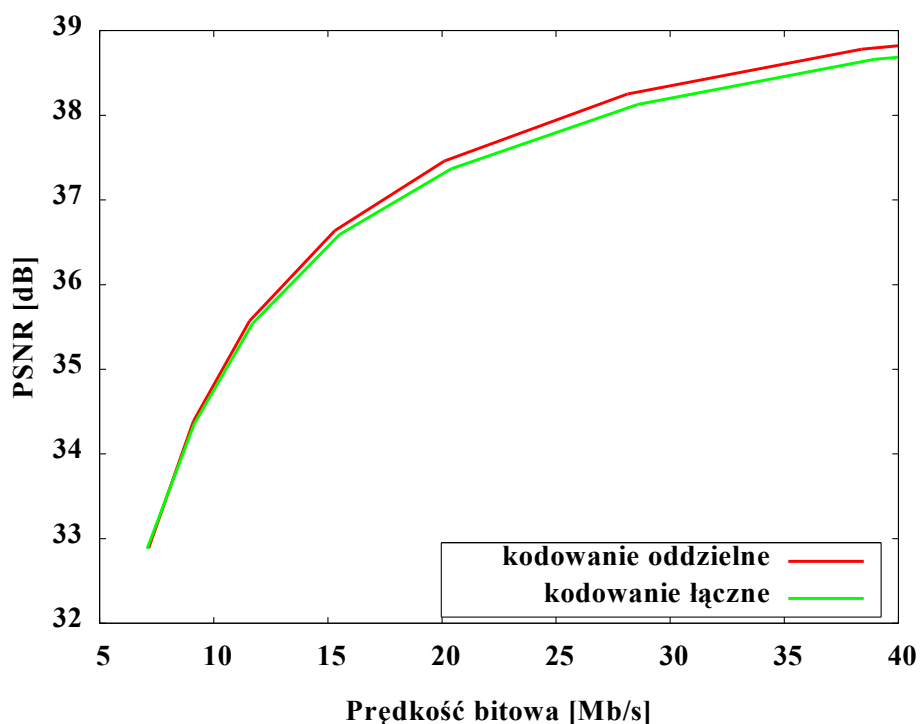
Rysunek 6.91. Porównanie efektywności kompresji kodowania czterokomponentowego do oddzielnego kodowania obrazów i map głębi. Kodowanie z możliwym podziałem na bloki 16x16, 8x8 i 4x4. Sekwencja Book Arrival, QD = 40, QP zmieniane w zakresie od 16 do 40.



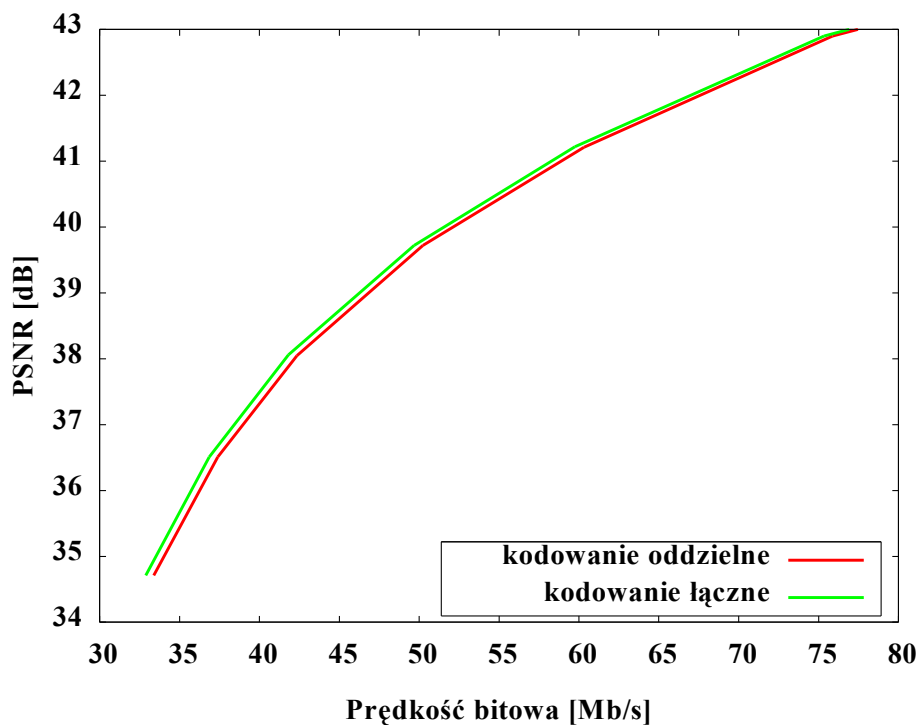
Rysunek 6.92. Porównanie efektywności kompresji kodowania czterokomponentowego do oddzielnego kodowania obrazów i map głębi. Kodowanie z możliwym podziałem na bloki 16x16, 8x8 i 4x4. Sekwencja Newspaper, QD = 16, QP zmieniane w zakresie od 16 do 40.



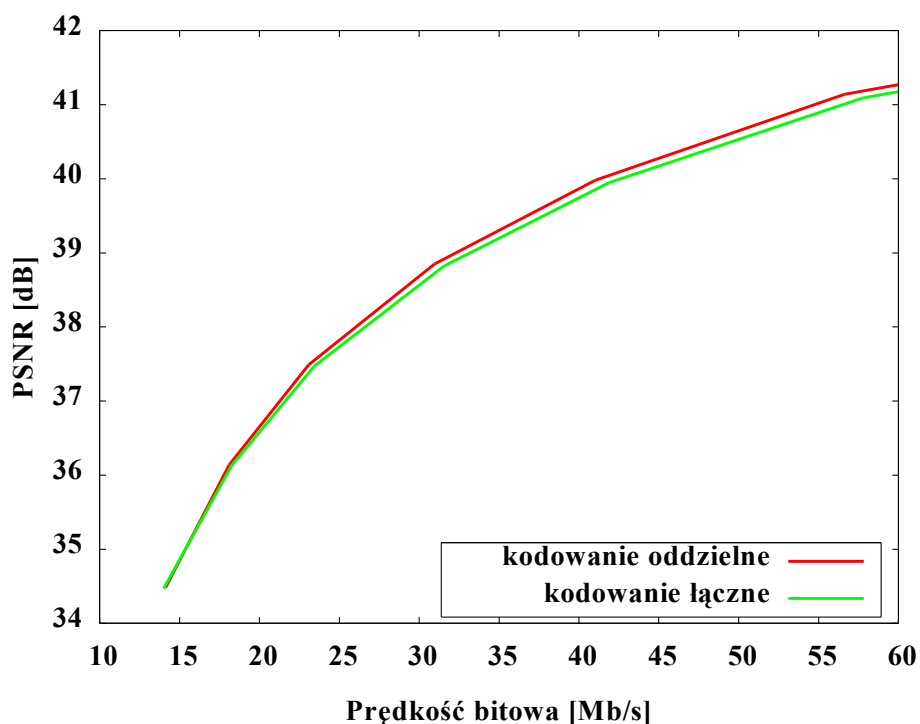
Rysunek 6.93. Porównanie efektywności kompresji kodowania czterokomponentowego do oddzielnego kodowania obrazów i map głębi. Kodowanie z możliwym podziałem na bloki 16x16, 8x8 i 4x4. Sekwencja Newspaper, QD = 34, QP zmieniane w zakresie od 16 do 40.



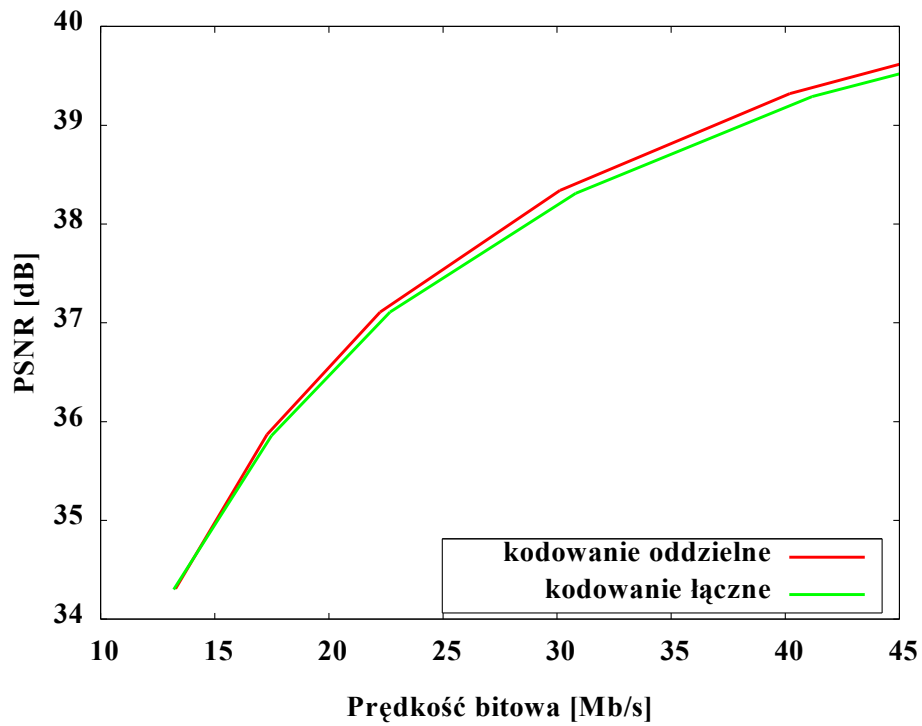
Rysunek 6.94. Porównanie efektywności kompresji kodowania czterokomponentowego do oddzielnego kodowania obrazów i map głębi. Kodowanie z możliwym podziałem na bloki 16x16, 8x8 i 4x4. Sekwencja Newspaper, QD = 40, QP zmieniane w zakresie od 16 do 40.



Rysunek 6.95. Porównanie efektywności kompresji kodowania czterokomponentowego do oddzielnego kodowania obrazów i map głębi. Kodowanie z możliwym podziałem na bloki 16x16, 8x8 i 4x4. Sekwencja Poznan Street, QD = 16, QP zmieniane w zakresie od 16 do 40.

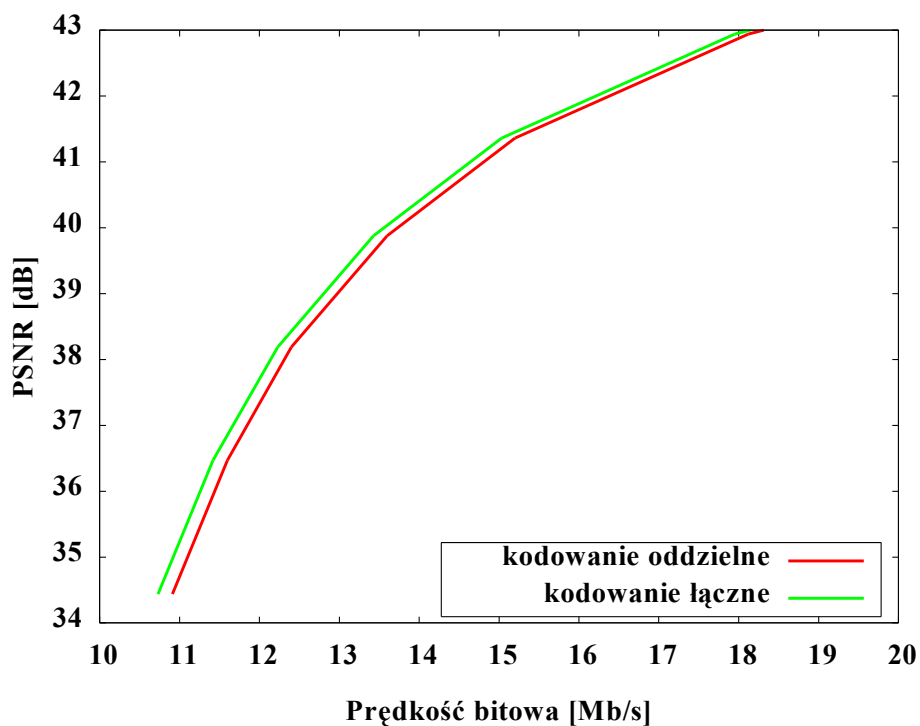


Rysunek 6.96. Porównanie efektywności kompresji kodowania czterokomponentowego do oddzielnego kodowania obrazów i map głębi. Kodowanie z możliwym podziałem na bloki 16x16, 8x8 i 4x4. Sekwencja Poznan Street, QD = 34, QP zmieniane w zakresie od 16 do 40.

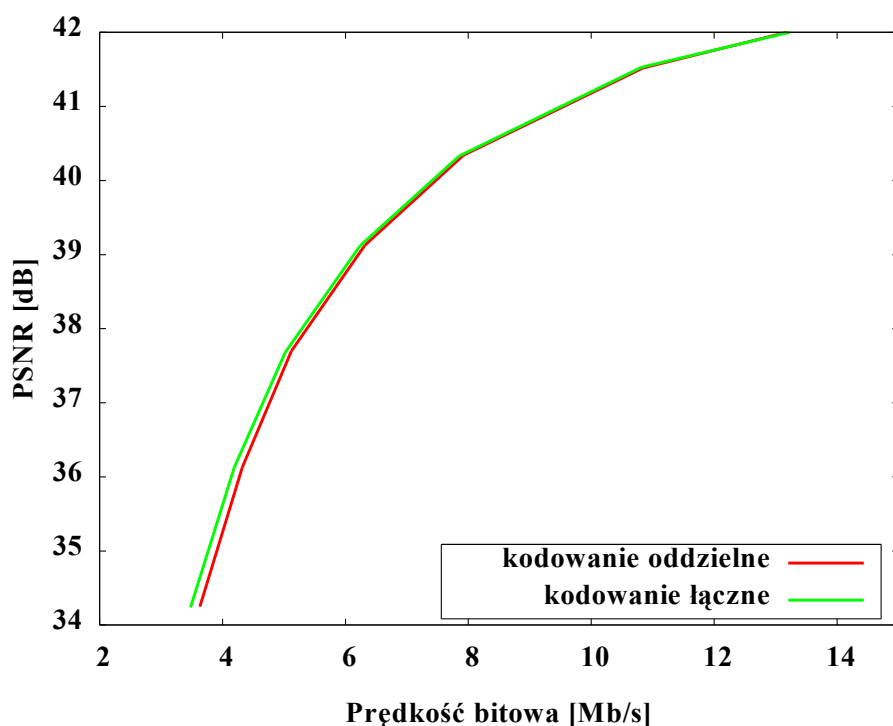


Rysunek 6.97. Porównanie efektywności kompresji kodowania czterokomponentowego do oddzielnego kodowania obrazów i map głębi. Kodowanie z możliwym podziałem na bloki 16x16, 8x8 i 4x4. Sekwencja Poznan Street, QD = 40, QP zmieniane w zakresie od 16 do 40.

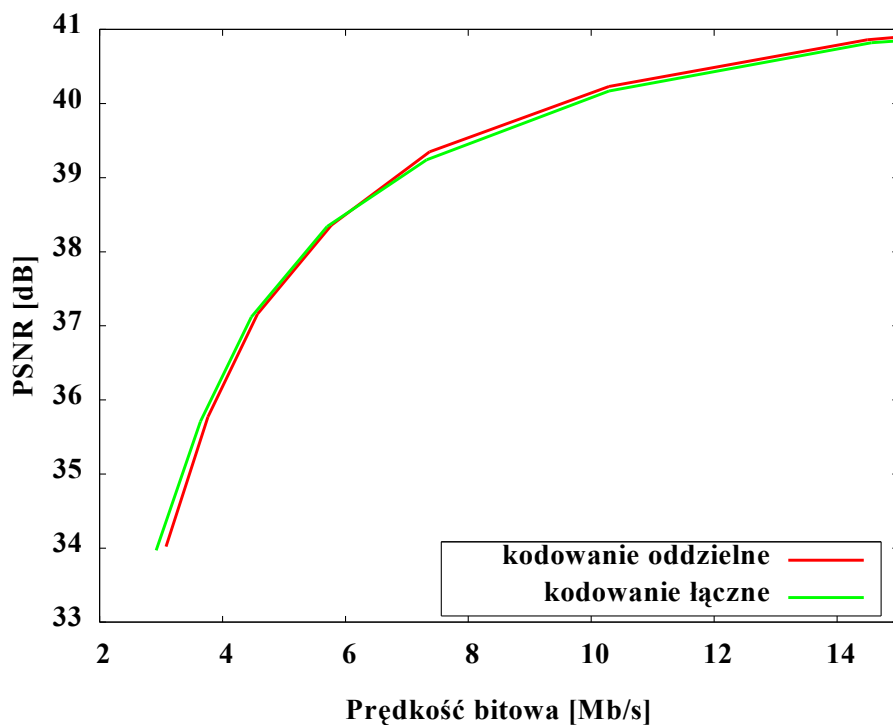




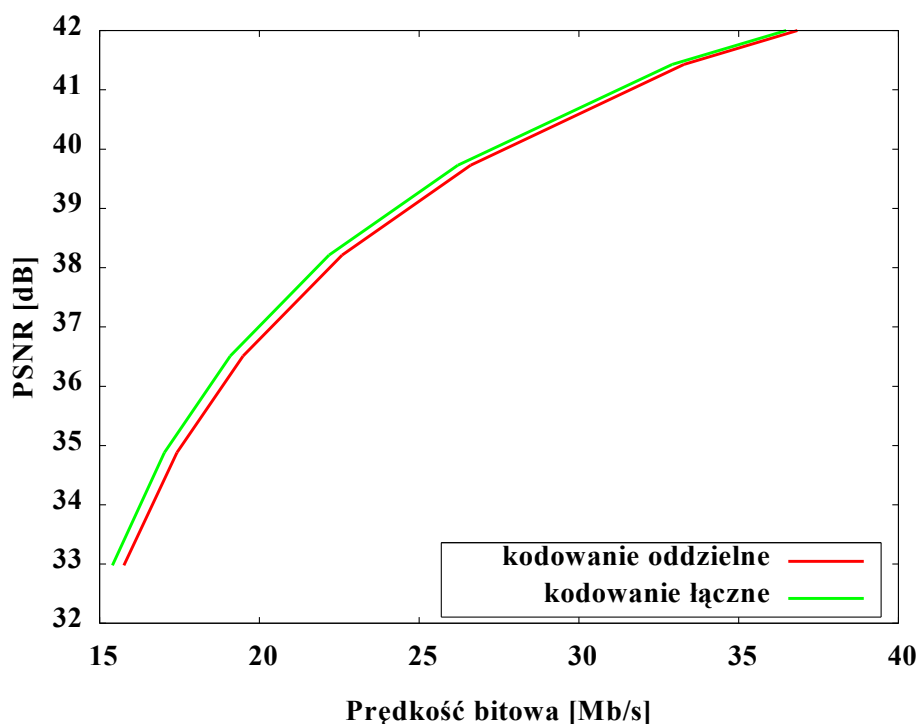
Rysunek 6.98. Porównanie efektywności kompresji kodowania czterokomponentowego do oddzielnego kodowania obrazów i map głębi. Kodowanie z możliwym podziałem na bloki 16x16 i 4x4. Sekwencja Book Arrival, QD = 16, QP zmieniane w zakresie od 16 do 40.



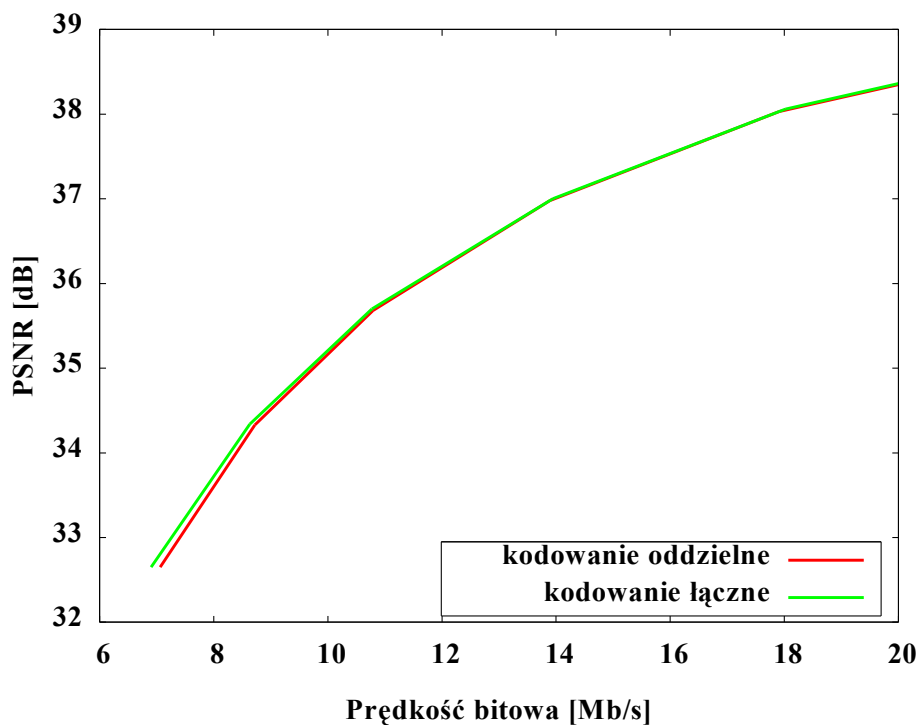
Rysunek 6.99. Porównanie efektywności kompresji kodowania czterokomponentowego do oddzielnego kodowania obrazów i map głębi. Kodowanie z możliwym podziałem na bloki 16x16 i 4x4. Sekwencja Book Arrival, QD = 34, QP zmieniane w zakresie od 16 do 40.



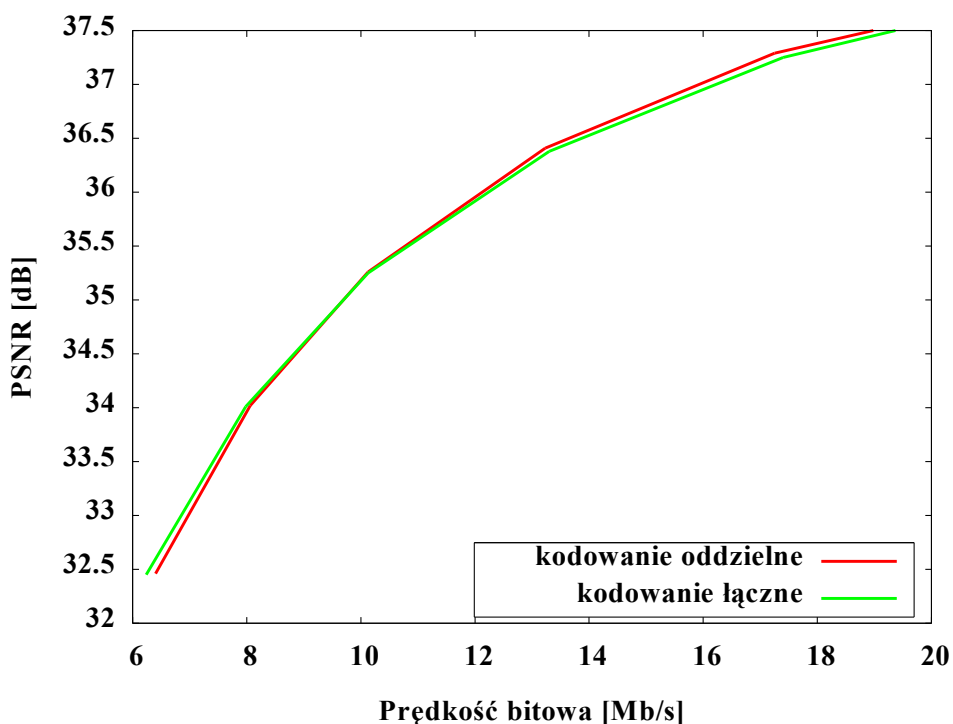
Rysunek 6.100. Porównanie efektywności kompresji kodowania czterokomponentowego do oddzielnego kodowania obrazów i map głębi. Kodowanie z możliwym podziałem na bloki 16x16 i 4x4. Sekwencja Book Arrival, QD = 40, QP zmieniane w zakresie od 16 do 40.



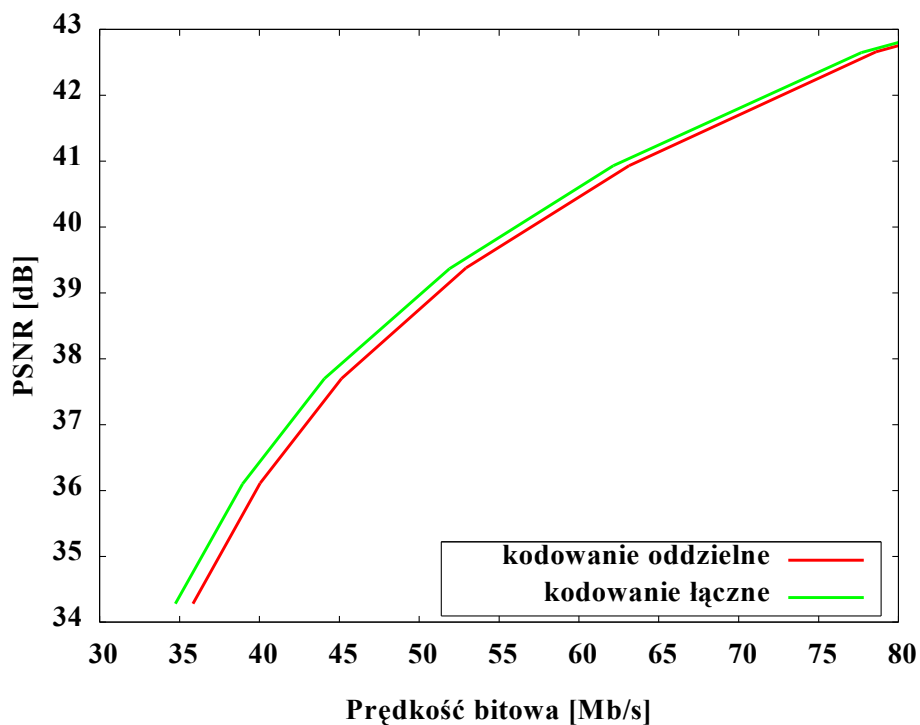
Rysunek 6.101. Porównanie efektywności kompresji kodowania czterokomponentowego do oddzielnego kodowania obrazów i map głębi. Kodowanie z możliwym podziałem na bloki 16x16 i 4x4. Sekwencja Newspaper, QD = 16, QP zmieniane w zakresie od 16 do 40.



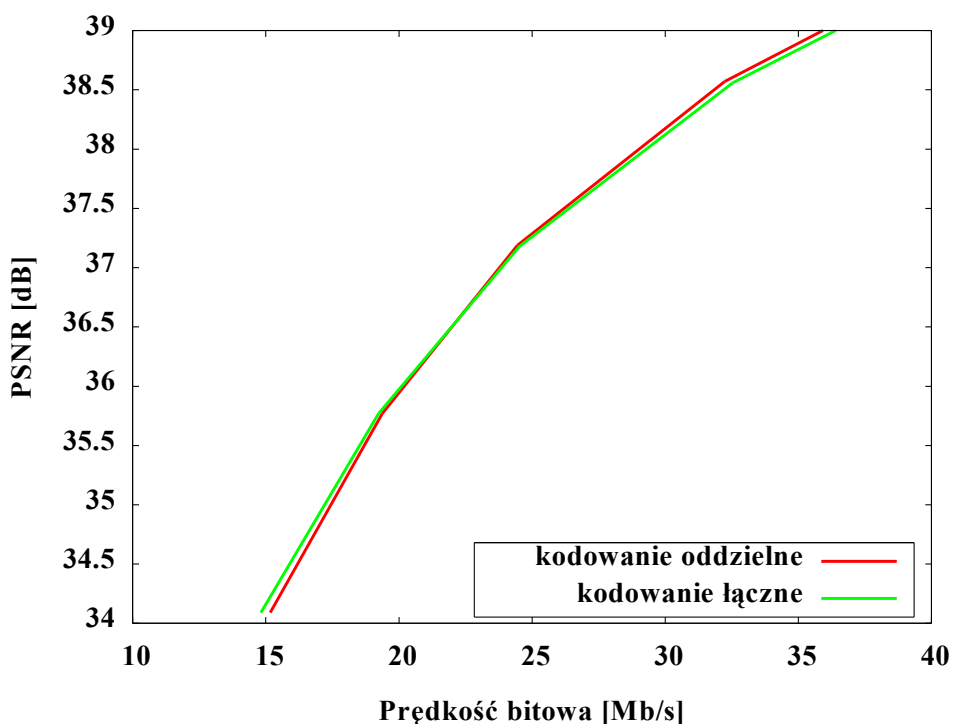
Rysunek 6.102. Porównanie efektywności kompresji kodowania czterokomponentowego do oddzielnego kodowania obrazów i map głębi. Kodowanie z możliwym podziałem na bloki 16x16 i 4x4. Sekwencja Newspaper, QD = 34, QP zmieniane w zakresie od 16 do 40.



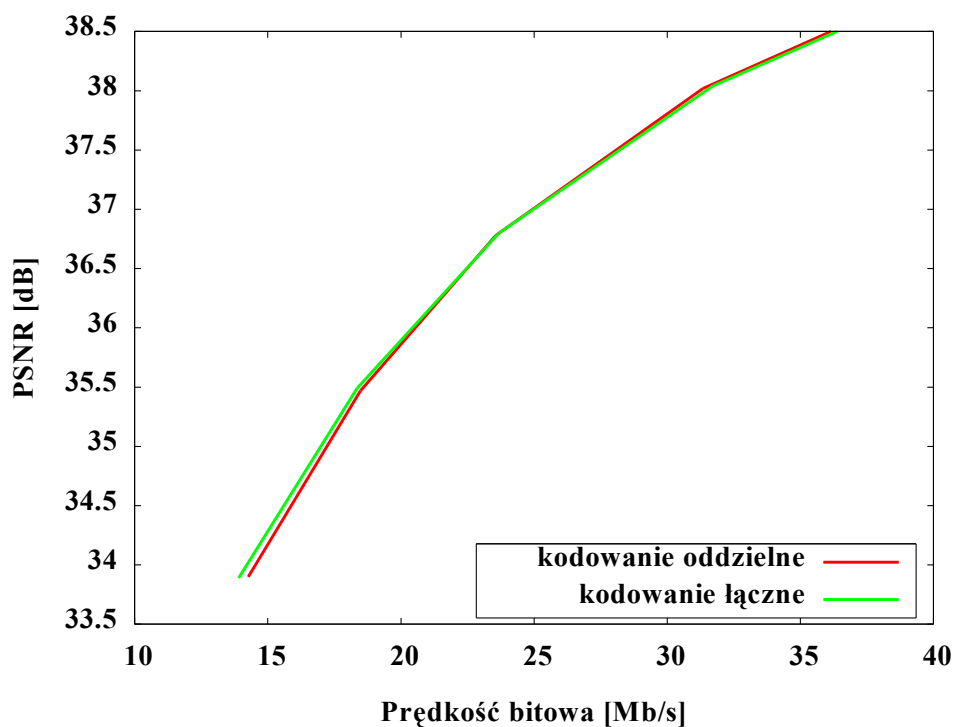
Rysunek 6.103. Porównanie efektywności kompresji kodowania czterokomponentowego do oddzielnego kodowania obrazów i map głębi. Kodowanie z możliwym podziałem na bloki 16x16 i 4x4. Sekwencja Newspaper, QD = 40, QP zmieniane w zakresie od 16 do 40.



Rysunek 6.104. Porównanie efektywności kompresji kodowania czterokomponentowego do oddzielnego kodowania obrazów i map głębi. Kodowanie z możliwym podziałem na bloki 16x16 i 4x4. Sekwencja Poznan Street, QD = 16, QP zmieniane w zakresie od 16 do 40.



Rysunek 6.105. Porównanie efektywności kompresji kodowania czterokomponentowego do oddzielnego kodowania obrazów i map głębi. Kodowanie z możliwym podziałem na bloki 16x16 i 4x4. Sekwencja Poznan Street, QD = 34, QP zmieniane w zakresie od 16 do 40.



Rysunek 6.106. Porównanie efektywności kompresji kodowania czterekomponentowego do oddzielnego kodowania obrazów i map głębi. Kodowanie z możliwym podziałem na bloki 16x16 i 4x4. Sekwencja Poznan Street, QD = 40, QP zmieniane w zakresie od 16 do 40.

Na wykresach z rysunków od 6.89 do 6.106 przedstawiono wyniki uzyskane w przypadku zastosowania stałej wartości indeksu QD i zmiennej wartości indeksu QP. Dla każdego przypadku przedstawionego na oddzielnym wykresie jakość i prędkość bitowa obrazów jest zatem zmienna, niewielkiej zmianie ulega również jakość i prędkość bitowa map głębi.

Wyniki świadczą o tym, że proponowana metoda kompresji czterekomponentowej daje w przypadku prawie wszystkich przebadanych w tej konfiguracji przypadków redukcję prędkości bitowej nawet do 3% dla zakresu mniejszych prędkości bitowych pokazanych na wykresach (parametr QP > 30). Dla części przypadków (na przykład z rysunków 6.94, 6.96, 6.97) zastosowanie proponowanej metody wiąże się jednak z pogorszeniem efektywności kompresji dla zakresu większych prędkości bitowych – strumienie mają wtedy prędkości bitowe o ponad 1% większe.

W przypadku, gdy wyłączona jest możliwość wyboru kodowania z podziałem na bloki 8x8 punktów. W tym przypadku koder czterekomponentowy w proponowanej postaci osiąga dla wszystkich sekwencji dla wartości QP większych niż 30 wyniki nie gorsze niż otrzymywane w przypadku oddzielnego kodowania obrazów i głębi. Obserwowana jest redukcja prędkości strumienia bitowego nawet o około 3% dla zakresu mniejszych prędkości bitowych (na przykład

rysunki 6.98, 6.101, 6.104).

### **6.6.3.3. Podsumowanie**

Wyniki przedstawione na wykresach z rysunków od 6.71 do 6.106 pokazują, że umożliwienie indywidualnego wyboru sposobu kodowania bloków głębi powoduje poprawę efektywności kompresji sekwencji czterokomponentowych w porównaniu do przypadku wcześniej zaprezentowanego (rysunki od 6.41 do 6.64), w którym składowa głębi była kodowana tak samo, jak składowa luminancji.

W ponad połowie przedstawionych przypadków kodowanie sekwencji czterokomponentowych z wykorzystaniem schematu kodowania makrobloku zaprezentowanego na rysunku 6.66 charakteryzuje się nie gorszą efektywnością w porównaniu z oddzielnym kodowaniem sekwencji obrazów i sekwencji map głębi. Dla pomiarów realizowanych przy niskich przepływnościach ( $QP < 30$  na rysunkach od 6.71 do 6.88,  $QD < 30$  na rysunkach od 6.89 do 6.106), w 9 przypadkach proponowany koder osiąga gorsze wyniki, w 4 zbliżone a w 23 – lepsze niż dla oddzielnego kodowania obrazu i map głębi. Zwiększenie efektywności powodowane jest przez wspólne użycie części wysyłanych danych dla kodowania obrazów i głębi. Zyski metody są jednak ograniczane poprzez konieczność zasygnalizowania w strumieniu, które dane są wspólnie używane przy kodowaniu obrazu i głębi, a które są przesyłane oddzielnie. Warto jednak podkreślić, że proponowana metoda nie zwiększa złożoności obliczeniowej procesu kodowania ani dekodowania.

### **6.6.4. Wspólne kodowanie wewnątrzobrazowe - podsumowanie**

Porównanie wyników kodowania uzyskanych z użyciem koderów realizujących kodowanie według schematów z rysunków 6.40 oraz 6.66 pokazuje, że ten drugi koder daje nieco lepsze wyniki. Uzyskiwane zyski dla małych prędkości bitowych są porównywalne w przypadku obu schematów, natomiast drugi schemat charakteryzuje się mniejszym obniżeniem efektywności kompresji dla dużych prędkości bitowych strumienia. W przypadku pierwszego schematu, obniżenie jakości dla dużych przepływności może być znaczne, jak pokazane to zostało na wykresach z rysunków 6.55, 6.58, 6.61 oraz 6.64. Wyraźnie widoczne jest na nich obniżenie jakości syntezy widoku dla większych prędkości bitowych, nawet o 4 dB względem przypadku oddzielnego kodowania obrazu i głębi. Obniżenie to powodowane jest niską jakością rekonstruowanych obrazów map głębi, które kodowane były z użyciem trybów wyznaczonych dla luminancji w każdym makrobloku. Kodowanie map głębi odbywało się zatem w sposób nieoptymalny z punktu widzenia jakości rekonstruowanej mapy głębi. Porównanie z wykresami z

rysunków, odpowiednio 6.94, 6.97, 6.100, 6.103, 6.106, pokazuje przewagę kodowania z użyciem schematu umożliwiającego wybór odrębnych sposobów kodowania bloków luminancji i głębi w tym samym makrobloku. W tym przypadku nie zachodzi redukcja jakości spowodowana zaniżoną jakością dekodowanych map głębi. Zachowane są natomiast zyski dla niskich prędkości bitowych.

## 6.7. Predykcja międzyobrazowa w kodowaniu czterokomponentowym

W czasie prac nad koderem czterokomponentowym większość uwagi skupiona była na metodach kodowania wewnątrzobrazowego, ze względu na duże potencjalne zyski wynikające z poprawienia tego sposobu kodowania sekwencji. Warto jednak pamiętać, że wysoka efektywność współczesnych koderów hybrydowych wynika głównie z zastosowania w nich zaawansowanych metod predykcji międzyobrazowej. W przypadku kodowania wielowidokowego z użyciem kodera MVC, kodowanie międzyobrazowe wykorzystuje mechanizmy predykcji międzyobrazowej w obrębie jednego widoku (predykcja wewnątrzwidokowa) jak i z wykorzystaniem obrazów z różnych widoków (predykcja międzywidokowa).

Możliwe jest oczywiście dostosowanie kodera czterokomponentowego do działania w trybie kodowania międzyobrazowego na analogicznej zasadzie jak zostało to zaprezentowane dla kodowania wewnątrzobrazowego w punkcie 6.6.2. W rozwiązaniu tym składowa głębi kodowana byłaby z wykorzystaniem tych samych wektorów ruchu i ramek odniesienia, co luminancja. Ten sposób kodowania jest jednak mało efektywny i nie był szerzej badany.

Konieczne jest zatem przesłanie odrębnych wektorów ruchu i numerów ramek odniesienia dla składowej głębi. Co równie istotne, konieczne jest przesłanie odrębnego sposobu podziału na bloki dla składowej głębi – w kodowaniu międzyobrazowym koder MVC oferuje znacznie więcej sposobów podziału na bloki niż w kodowaniu wewnątrzobrazowym (szczegóły można odnaleźć na przykład w książkach [Doma\_11] i [Rich\_10]). Konieczne jest również przesłanie typu makrobloku dla bloku głębi.

Liczba stopni swobody jest zatem znacznie większa niż w kodowaniu wewnątrzobrazowym. Dodatkowym problemem występującym w przypadku kompresji międzyobrazowej jest występowanie makrobloków typu *skip*. Jak zostało to już wspomniane, tego typu bloki charakteryzują się niezwykle wysoką efektywnością kompresji. Niestety, ich efektywne wykorzystanie w przypadku, gdy kodowana jest sekwencja czterokomponentowa jest utrudnione. Dzieje się tak ze względu na fakt, że tryb *skip* oznacza pominięcie wszystkich informacji o makrobloku, łącznie z głębią. Ponieważ założeniem przyświecającym badaniom było zachowanie

takiej samej postaci rekonstruowanego obrazu co dla kodowania oddzielnego, decyzja o tym, które makrobloki kodowane mogą być jako *skip*, zależy od treści składowej luminancji. Tymczasem całkowite pominięcie w strumieniu danych dla makrobloku może skutkować dużym zniekształceniem rekonstruowanej składowej mapy głębi dla tego makrobloku. Niestety, w sytuacji gdy wybranym sposobem zakodowania tylko jednej ze składowych – luminancji bądź głębi, jest tryb *skip*, nie może on zostać wykorzystany ze względu na konieczność przesłania danych dla składowej, dla której optymalne jest kodowanie w innym trybie. Wskutek tego, liczba makrobloków typu *skip* ulega w kodowaniu czterokomponentowym zmniejszeniu. Dodatkowym problemem jest w tym przypadku sposób kodowania informacji o makroblokach *skip* w wykorzystanym w pracach kodowaniu UVLC. Sposób ten polega na wykorzystaniu kodowania „*run-length*”, czyli zapisywaniu w strumieniu informacji o tym, ile makrobloków poprzedzających aktualnie odczytywany ze strumienia makroblok zakodowanych zostało jako makrobloki typu *skip*. Sposób ten jest bardzo wydajny dla przypadków, gdy całe obszary obrazu kodowane są z użyciem bloków *skip*, jednak działa słabo w przypadkach, gdy bloki *skip* rozproszone są w małych grupach. Niestety, kodowanie czterokomponentowe skutkuje zmniejszeniem liczby makrobloków *skip* i skróceniem ciągów następujących po sobie bloków tego typu. Problem ten zobrazowany jest na rysunku 6.107. W 2 rzędach przedstawione są na nim optymalne sposoby kodowania bloków luminancji i głębi. Dla każdego bloku określona jest też wartość oznaczająca liczbę poprzedzających dany blok bloków *skip*. Zaznaczone są sytuacje uniemożliwiające zastosowanie bloku *skip*. Na rysunku 6.108 przedstawiony jest wynikowy sposób kodowania kolejnych makrobloków. Widoczny jest spadek liczby makrobloków typu *skip* w porównaniu z trybami wybieranymi dla oddzielnego kodowania obrazów i głębi.



### Kodowanie optymalne dla luminancji

I16x16 0	P16x16 0	P8x16 0	Skip	Skip	I4x4 2	Skip	P8x8 1	P8x8 0
P16x16 0	I4x4 0	P8x16 0	Skip	I4x4 1	Skip	Skip	P8x8 2	P8x8 0

### Kodowanie optymalne dla głębi

Rysunek 6.107. Porównanie optymalnych trybów kodowania bloków luminancji i głębi z uwzględnieniem kodowania międzyobrazowego. Na czerwono zaznaczone sytuacje uniemożliwiające wykorzystanie bloków *skip*.

### Kodowanie łączne

I16x16 P16x16 0	P16x16 I4x4 0	P8x16 P8x16 0	Skip	Skip I4x4 1	I4x4 Skip 0	Skip	P8x8 P8x8 1	P8x8 P8x8 0
-----------------------	---------------------	---------------------	------	-------------------	-------------------	------	-------------------	-------------------

Rysunek 6.108. Tryby kodowania wybierane dla wspólnego kodowania luminancji i głębi. Widoczna mniejsza w porównaniu z rysunkiem 6.107 liczba bloków *skip*.

## 6.8. Podsumowanie i wnioski

Przedstawiona została tu metoda kompresji sekwencji wielowidokowych z głębią polegająca na stworzeniu czterokomponentowej sekwencji wielowidokowej, zawierającej poza trzema składowymi opisującymi obraz sceny (luminancja i dwie chrominancje), składową głębi, niosącą informację o trójwymiarowej strukturze sceny. Zaletą tego typu sekwencji jest brak konieczności synchronizacji danych o obrazie i głębi, które w dotychczas stosowanych rozwiązaniach kodowane były zupełnie niezależnie w odrębnych przebiegach kodera.

Opracowany został, na podstawie referencyjnego kodera MVC, koder dostosowany do kompresji sekwencji czterokomponentowych, wykorzystujący zależności między składowymi luminancji i głębi. Jego możliwości zostały następnie przetestowane z użyciem najnowszych dostępnych w czasie prowadzenia badań sekwencji wielowidokowych z głębią. Przeprowadzone badania podsumować można następująco.

1. W przypadku kompresji łącznej sekwencji obrazów i map głębi, podobnie jak w przypadku kompresji oddzielnej, znacznie większy wpływ na jakość syntezowanych widoków ma jakość obrazów.
2. Kodowanie zdecydowanego dwukrotnie komponentu głębi w sposób stosowany dla

chrominancji jest nieefektywny i nie powinien być stosowany.

3. Kodowanie mapy głębi powinno odbywać się w pełnej rozdzielczości, aby nie doprowadzać do utraty istotnych szczegółów, obecnych w mapach głębi wysokiej jakości uzyskiwanych z użyciem najnowszych technik.
4. Kodowanie wewnątrzobrazowe bloków mapy głębi z użyciem tych samych trybów, które są używane do kodowania bloków luminancji w makrobloku prowadzi do pogorszenia jakości obrazów syntezy uzyskiwanych dla małych wartości parametrów kwantyzacji QD. Dla dużych wartości parametrów QD możliwe jest uzyskanie za pomocą tej metody wyników lepszych niż dla oddzielnego kodowania obrazów i map głębi.
5. Kodowanie wewnątrzobrazowe bloków mapy głębi w sposób umożliwiający indywidualny dobór sposobu kodowania pozwala na uzyskanie wyższej efektywności niż dla oddzielnego kodowania obrazów i map głębi przy zachowaniu takiej samej złożoności obliczeniowej procesu kodowania i dekodowania.
6. Efektywność kodowania czterokomponentowego w dużym stopniu zależy od właściwości sekwencji oraz jakości dostępnych dla niej map głębi.
7. Statystyki związane z częstością występowania możliwych wartości pola CBP i DCBP są inne dla składowych obrazu i dla składowej mapy głębi. Możliwe jest zatem zwiększenie efektywności kompresji głębi poprzez odpowiednie dostosowanie kodów entropijnych używanych do kodowania wartości tego pola.

## 7. Podsumowanie

### 7.1. Podsumowanie pracy

Zgodnie z postawionymi celami, w pracy przebadane zostały możliwości wykorzystania kodera MVC do kompresji sekwencji wielowidokowych z głębią.

Cele pracy osiągnięte zostały przez przebadanie wpływu wyboru wartości podstawowych parametrów sterujących procesem kompresji na jakość rekonstruowanych danych i na prędkości bitowe strumieni zakodowanych sekwencji wielowidokowych z mapami głębi. Na podstawie otrzymanych wyników *określone zostały sposoby doboru parametrów kompresji* pod kątem uzyskiwania jak najlepszej efektywności kompresji w założonych warunkach. Przedstawiony został sposób pomiaru jakości sekwencji wielowidokowych z głębią, polegający na badaniu jakości widoków syntezowanych z użyciem zrekonstruowanych danych. Sposób ten został użyty do oceny jakości sekwencji przy zastosowaniu oddzielnej kompresji obrazów i map głębi za pomocą standardowego kodera MVC. Na podstawie tak uzyskanych wyników *opracowane zostały wzory pozwalające na uzyskanie jak najlepszej jakości rekonstruowanych sekwencji wielowidokowych z głębią przy danej prędkości bitowej poprzez odpowiedni dobór parametrów kodowania*.

Kolejnym celem pracy było opracowanie i przebadanie takich modyfikacji kodera MVC, które umożliwiałyby wykorzystanie zależności pomiędzy danymi w systemach wielowidokowych. Cel pracy osiągnięto poprzez *zapropozowanie, zaimplementowanie i przetestowanie kodera i dekodera sekwencji wielowidokowych z głębią w formie kodera i dekodera sekwencji czterokomponentowych*. Sekwencje czterokomponentowe są sekwencjami obrazów składających się z czterech komponentów – luminancji, dwóch chrominancji oraz składowej mapy głębi. Reprezentacja w postaci sekwencji czterokomponentowych pozwala na zagregowanie całości danych koniecznych do poprawnej pracy współczesnych systemów wizji wielowidokowej do jednego zbioru, eliminując konieczność ich synchronizacji. W zaproponowanym koderze wprowadzono modyfikacje umożliwiające bardziej efektywną kompresję sekwencji wielowidokowych z głębią. Najważniejsze modyfikacje to *wspólne kodowanie kierunków predykcji i podziałów bloków w kodowaniu wewnątrzobrazowym, zoptymalizowane kodowanie danych dotyczących obecności niezerowych próbek transformaty błędu predykcji dla bloków mapy głębi oraz szereg drobnych zmian dostosowujących koder do pracy z sekwencjami czterokomponentowymi*. Działanie kodera zostało przetestowane przy użyciu sekwencji testowych dostępnych w ramach badań nad nowymi

standardami kompresji sekwencji wielowidokowych.

Przedstawione w pracy wyniki, będące rezultatami przeprowadzonych badań, pozwalają stwierdzić, że dowiedziona została teza pracy, która brzmi:

*„Istnieją proste reguły, umożliwiające ustalenie proporcji pomiędzy prędkościami bitowymi obrazów i głębi, pozwalające na maksymalizację efektywności kompresji sekwencji wielowidokowych z mapami głębi przeprowadzanej przy użyciu kodera MVC. Mapy głębi mogą być w procesie kompresji traktowane jako dodatkowa składowa barwna, co pozwala na proste wykorzystanie zależności pomiędzy obrazami i odpowiadającymi im mapami głębi.”.*

Przedstawione w pracy reguły dotyczą sposobu takiego przeprowadzenia procesu kodowania, dla którego osiągnięte wyniki będą odznaczały się maksymalną efektywnością przy danej prędkości bitowej. Kluczowy jest w tym wypadku odpowiedni dobór par parametrów kwantyzacji QP i QD.

Zaproponowany i przebadany został sposób kompresji sekwencji wielowidokowych z głębią wykorzystujący zależności między obrazem a mapą głębi. Przedstawiony w pracy sposób wykorzystania zależności pomiędzy danymi o luminancji i głębi pozwala na zwiększenie efektywności kompresji w przypadku małych prędkości bitowych. W przypadku strumieni o małej prędkości bitowej używane są duże wartości parametrów kwantyzacji. Użycie dużych wartości parametru kwantyzacji, szczególnie – parametru kwantyzacji dla luminancji, powoduje, że wybierane dla bloków luminancji i głębi sposoby kodowania wewnątrzobrazowego ulegają daleko idącej unifikacji. To z kolei pozwala na użycie tych samych podziałów i kierunków predykcji w kompresji wewnątrzobrazowej dla bloków luminancji oraz głębi i uzyskanie mniejszych prędkości bitowych generowanego strumienia.

Przedstawione w pracy wyniki badań pozwalają stwierdzić, że osiągnięte zostały wszystkie cele pracy.

## **7.2. Główne osiągnięcia**

Główne, autorskie osiągnięcia pracy podzielić można na dwa obszary.

1. W ramach badań nad metodami doboru parametrów kodowania sekwencji wielowidokowych z mapami głębi:
  - Wyznaczono maksymalne wartości indeksów kwantyzacji dla kodowania głębi w zależności od wartości parametru kwantyzacji dla kodowania obrazów (punkt 5.1.4.2);
  - Określono wpływ decymacji map głębi na efektywność procesu kodowania sekwencji wielowidokowych z mapami głębi (punkt 5.1.4.3);

- Opracowano metodę doboru parametrów kodowania map głębi dających najlepsze efekty procesu syntezy widoków dla danych parametrów kodowania obrazu (punkt 5.2.2.3).

Przedstawione wyniki mają duże znaczenie praktyczne. Szczególnie użyteczne są wzory (5.4) oraz (5.5), pozwalające na automatyczne wyznaczenie parametrów kodowania map głębi dla bezpośredniego kodowania koderem MVC sekwencji wielowidokowych z mapami głębi. W dostępnej literaturze problem doboru właściwych proporcji prędkości bitowych dla obrazu głębi jest dostrzegany, jednak w żadnej publikacji nie są podane wzory pozwalające na automatyczny dobór parametrów kodowania.

2. W ramach badań nad wspólną kompresją sekwencji obrazów i sekwencji map głębi:
  - opracowano oryginalną metodę wspólnej kompresji sekwencji wielowidokowych z głębią i przebadano różne jej warianty (punkty 6.6.3 oraz 6.6.2);
  - przebadano zależności pomiędzy trybami kodowania współleżących makrobloków dla kompresji obrazów i map głębi w trybie kodowania wewnątrzobrazowego (punkt 6.6.1);
  - zaproponowano i przebadano sposób scalenia obrazów i map głębi do postaci sekwencji czterokomponentowej (punkt 6.3);
  - zaproponowano modyfikacje koderów zwiększające efektywność kodowania map głębi (punkt 6.6.3).

Badania nad metodami wspólnej kompresji sekwencji wielowidokowych z głębią, wykorzystującymi zależności istniejące pomiędzy tymi danymi mają obecnie dość duże znaczenie. Zaprezentowane wyniki mogą stanowić punkt wyjścia do opracowania metod kodowania tego typu sekwencji w ramach prowadzonych obecnie prac nad koderem sekwencji trójwymiarowych [3DV\_WD].

Dodatkowe osiągnięcia i cele opisane w rozprawie przedstawione są poniżej.

1. Przebadano wpływ przetwarzania wstępnego, polegającego na dokonaniu rektyfikacji epipolarnej i korekcji kolorów, na efektywność kompresji sekwencji wielowidokowych. Przedstawione wyniki mają duży walor poznawczy. Są one niemożliwe do otrzymania bez, uzyskanego przez autora rozprawy, dostępu do sekwencji wielowidokowych w postaci w jakiej poddawane są one kolejnym etapom przetwarzania.

### 7.3. **Możliwe kierunki dalszych badań**

Podsumowując przedstawioną pracę należy stwierdzić, że najbardziej prawdopodobnym scenariuszem rozwoju telewizji trójwymiarowej w najbliższej przyszłości jest kontynuacja prac nad systemami autostereoskopowymi.

Warto w tym miejscu zaznaczyć, że niezwykle interesującym rozwiązaniem problemu prezentacji danych telewizji trójwymiarowej, którego rozwój możliwy jest dzięki postępującemu rozwojowi technologii wyświetlaczy o ultra wysokiej rozdzielczości, jest wykorzystanie technologii obrazowania zintegrowanego (ang. *integral imaging*) [Taka\_08][Nunes\_12]. Technologia ta stanowi rozwinięcie technologii wyświetlaczy autostereoskopowych i można liczyć na jej rozpowszechnienie w perspektywie najbliższych lat.

W kontekście wykorzystania dostępnych obecnie wyświetlaczy autostereoskopowych, duże znaczenie mają prace nad zapewnieniem efektywnej kompresji sekwencji wielowidokowych z głębią, składających się z 2 lub 3 widoków [MPEG\_CfP]. Wśród proponowanych metod kodowania tego typu sekwencji szczególnie wyróżniają się sposoby wykorzystujące najnowsze rozwiązania w dziedzinie kodowania sekwencji wizyjnych, proponowane do umieszczenia w nowym koderze HEVC (*High Efficiency Video Coder*) [HEVC][Doma\_11a][Doma\_12a]. Propozycje charakteryzują się znaczną odmiennością przyjętych koncepcji w porównaniu z aktualnie używanym standardem AVC. Prace nad nowym standardem trwają i należy oczekiwać, że ostateczna jego wersja będzie oferować wyraźnie lepszą efektywność niż standard AVC [Wegn\_10][Doma\_11b]. Można zatem przypuszczać, że opracowywane w przeciągu następnych kilku lat metody kompresji sekwencji wielowidokowych z głębią bazować będą na tej nowej technologii.

Jakiegokolwiek dalsze badania powinny zatem skupiać się na odpowiednim dostosowaniu kodera według powstającego standardu. Należy zbadać, czy wnioski z prezentowanej pracy mają również zastosowanie w jego przypadku.

Celowym wydaje się być wykorzystanie w dalszych badaniach sposobów kodowania, które nie wiązałyby ze sobą tak ściśle kodowania komponentów luminancji i głębi. Obecnie najbardziej perspektywiczne wydają się metody realizujące kodowanie tych komponentów niezależnie, umożliwiające jednak wzajemne korzystanie z zawartych w nich informacji na etapie predykcji. Rozwiązania takie powinny być dobrze dopasowane do koncepcji kodera następnej generacji HEVC, charakteryzującego się znacznie bardziej rozbudowanymi mechanizmami predykcji w porównaniu z koderem wykorzystywanym w pracy, bazującym w pełni na koderze AVC.

## **Aneksy**





## Aneks I

Dla badań nad wpływem przetwarzania wstępnego sekwencji wielowidokowych na efektywność ich kompresji wykorzystano sekwencje Poznan Street [Doma\_09], Poznan CarPark [Doma\_09] i Book Arrival [Feld\_08]. Szczegółowe dane wymienione są w tabeli 7.1.

Tabela 7.1. Sekwencje testowe wykorzystane w eksperymencie.

Nazwa sekwencji	rozdzielczość	Liczba ramek na sekundę	Liczba użytych ramek	Liczba użytych widoków
Book Arrival	1024 x 768	16,67	100	3
Poznan Street	1920 x 1088	25	100	3
Poznan CarPark	1920 x 1088	25	100	3

Sekwencje wejściowe obrazów były w trakcie eksperymentu poddawane kolejnym procesom przetwarzania. Najpierw kodowana była sekwencja nieprzetworzona, potem sekwencja poddana rektyfikacji, a na końcu – sekwencja poddana rektyfikacji i korekcji kolorów. Proces rektyfikacji epipolarnej przeprowadzany był według metody opisanej w [Stanko\_10] a korekcja kolorów według metody opisanej w [Stanko\_10a][Wegn\_10a]. Do kodowania sekwencji wykorzystany został koder MVC w wersji JMVC 8.5 [JMVC] z domyślnymi ustawieniami, podanymi w tabeli 7.3 (aneks II). Kodowaniu podlegały po 3 widoki z każdej sekwencji, po 100 ramek z każdego widoku.



## Aneks II

W celu określenia wpływu zniekształceń kodowania, występujących w zrekonstruowanych mapach głębi i w obrazach, na jakość syntezowanych widoków, przeprowadzone zostały badania na grupie sekwencji wielowidokowych: Book Arrival [Feld\_08], Newspaper [Ho\_08], Pantomime [Tani\_08]. W eksperymencie wykorzystane zostały po dwa widoki z trzech sekwencji wielowidokowych, których najważniejsze parametry podane zostały w tabeli 7.2. Każda sekwencja składa się z obrazów i map głębi o tej samej rozdzielczości co obrazy.

Tabela 7.2. Sekwencje testowe wykorzystane w eksperymencie.

Nazwa sekwencji	rozdzielczość	Liczba ramek na sekundę	Numery użytych kamer	Numer kamery syntezywanej	Liczba użytych ramek
Book Arrival	1024 x 768	16,67	7 i 10	8	100
Newspaper	1024 x 768	30	3 i 6	4	100
Pantomime	1280 x 960	29,41	38 i 41	39	100

W trakcie eksperymentu, kompresji podlegały widoki z dwu kamer oraz odpowiadające im mapy głębi. Wykorzystane zostały mapy głębi wyznaczone w sposób w pełni automatyczny za pomocą oprogramowania DERS [Tani\_08b]. Do kompresji obrazów wykorzystany został koder MVC w domyślnej konfiguracji (patrz tabela 7.3). Podobnie, do kompresji map głębi wykorzystany został ten sam koder MVC w domyślnej konfiguracji.

Tabela 7.3. Domyślna konfiguracja kodera MVC – najważniejsze parametry.

Parametr	Wartość
Długość GOP	16 ramek, hierarchiczne ramki B
Okres występowania ramek typu I	16
Tryby kodowania makrobloku	Wszystkie dostępne
Koder entropijny	CABAC
Maksymalna liczba ramek odniesienia	2

Zdekodowane dane wykorzystane zostały w procesie syntezy, przeprowadzonym z użyciem oprogramowania referencyjnego do syntezy obrazów VSRS [Tani\_08b]. Użyte opcje programu do syntezy widoku były identyczne z opcjami dla danej sekwencji wykorzystywanymi w pracach nad standardem kompresji sekwencji wielowidokowych przez członków grupy ekspertów MPEG

(*Moving Pictures Experts Group*). Przykładowe pliki konfiguracyjne oprogramowania syntezy są udostępniane wraz z sekwencjami.

Zrekonstruowane sekwencje obrazów i sekwencje map głębi użyto w procesie syntezy i wygenerowano widok wirtualny dla każdej możliwej do utworzenia pary indeksów QP oraz QD z tabeli 7.4. Otrzymany widok wirtualny porównany został z rzeczywistym widokiem z kamery umiejscowionej dokładnie w tym samym miejscu, którego koordynaty ustawione zostały w oprogramowaniu syntezy. Porównanie widoków przeprowadzone zostało z użyciem miary PSNR (*Peak Signal to Noise Ratio*) liczonej dla luminancji, a wynik przedstawiony został w mierze decybelowej, według wzoru (3.1).

W trakcie badań kompresowano sekwencje z dwu kamer z indeksami parametru kwantyzacji wymienionymi w tabeli 7.4. Zakres wartości indeksów jest szeroki, aby wyniki eksperymentu dawały pełny obraz badanej zależności. Wykluczone zostały jedynie niskie wartości poniżej 21, gdyż ich zastosowanie powoduje generowanie strumieni o bardzo dużej prędkości bitowej. Im większa wartość indeksu, tym większy rozmiar kroku kwantyzacji oraz, co za tym idzie, tym niższa ogólna wierność obrazu rekonstruowanego wobec oryginalnego. Jednocześnie z obniżaniem jakości maleje jednak prędkość bitowa generowanego strumienia danych.

Tabela 7.4. Wartości indeksów QP i QD wykorzystane w badaniach.

	<b>Parametr kwantyzacji dla kompresji obrazu QP</b>	<b>Parametr kwantyzacji dla kompresji mapy głębi QD</b>
Wartości wykorzystane w kompresji	21, 26, 31, 36, 41, 46, 51	21, 26, 31, 36, 41, 46, 51

W badaniach nad wykorzystaniem zdecydowanych map głębi wykorzystane zostały te same trzy sekwencje, co w poprzednim przypadku, a dodatkowo wykorzystano sekwencję Dog [Tani\_08] oraz sekwencję Book Arrival z nowszymi mapami głębi o lepszej jakości, wyznaczonymi półautomatycznie (tabela 7.5). W tym przypadku, kompresji poddane zostały mapy głębi zdecydowane (zarówno w pionie jak i w poziomie) do połowy i do jednej czwartej pierwotnego rozmiaru.

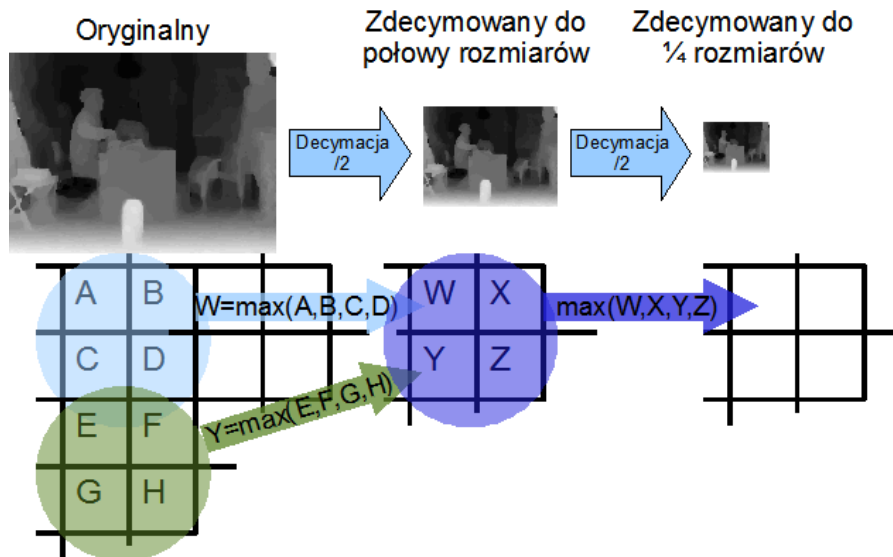
Tabela 7.5. Sekwencje testowe wykorzystane w eksperymencie.

Nazwa sekwencji	rozdzielczość	Liczba ramek na sekundę	Numery użytych kamer	Numer kamery syntezywanej	Liczba użytych ramek
Book Arrival	1024 x 768	16,67	7 i 10	8	100
Book Arrival z mapami głębi wysokiej jakości	1024 x 768	16,67	8 i 10	9	100
Newspaper	1024 x 768	30	3 i 6	4	100
Pantomime	1280 x 960	29,41	38 i 41	39	100
Dog	1280 x 960	29,41	38 i 41	39	100

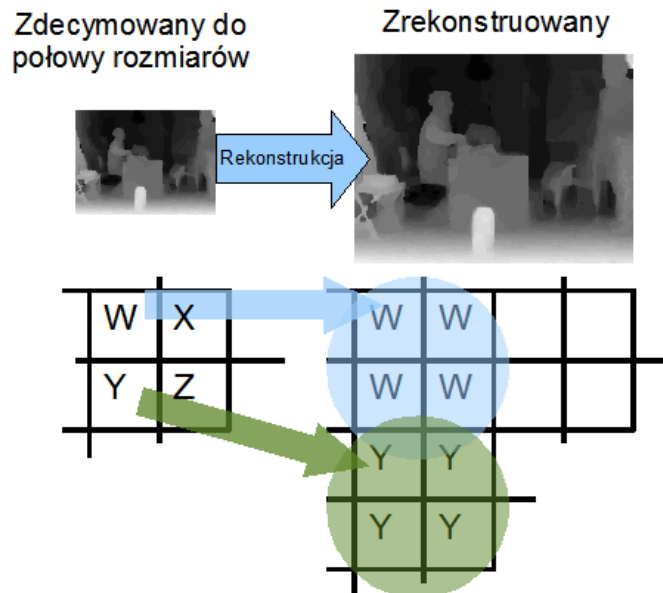
Decymację przeprowadzono z użyciem nieliniowego przetwarzania wartości punktów obrazu. Przy decymacji do połowy szerokości i wysokości, wartość próbki mapy głębi była równa największej wartości czterech sąsiadujących ze sobą (patrz rysunek 7.1) próbek oryginalnej mapy głębi (filtr maksymalny). Dla decymacji czterokrotnej (do  $\frac{1}{4}$  wysokości i szerokości), wartość próbki w wynikowej mapie głębi równa była maksymalnej wartości spośród 16 sąsiednich próbek oryginalnej mapy głębi. Uzasadnieniem dla takiego sposobu zmniejszenia rozmiaru mapy głębi jest obserwacja, że mapa głębi, wykorzystywana w syntezie widoków wirtualnych, służy do określania sposobu, w jaki obiekty przesłaniają się w widoku wirtualnym. Dla widza pierwszorzędne znaczenie mają z reguły obiekty znajdujące się na pierwszym planie. Opisany sposób decymacji pozwala na zachowanie poprawnej wartości dla mapy głębi dla obiektów pierwszego planu, zapobiegając silnym zniekształceniom w obrazie wirtualnym. Opisany sposób decymacji okazał się dawać najlepsze wyniki spośród rozpatrywanych prostych metod, nie wymagających złożonych obliczeń. Innymi rozpatrywanymi metodami było wybieranie wartości średniej, minimalnej i mediany.

Na etapie syntezy wykorzystano mapy głębi pełnej rozdzielczości, uzyskane przez powielenie sąsiednich punktów zrekonstruowanej mapy głębi, odpowiednio, cztero- lub szesnastokrotnie (patrz rysunek 7.2). Powielenie wartości próbki oznacza użycie wartości najbliższej dostępnej próbki (ang. *nearest neighbor*), jednak, jak wykazały badania, dla używanych w eksperymencie stopni decymacji, algorytm ten nie okazuje się być gorszy od znacznie bardziej złożonych obliczeniowo [Wild\_10].

Schematyczne przedstawienie algorytmu decymacji przedstawione jest na rysunku 7.1. Na rysunku 7.2 przedstawiony jest schemat rekonstrukcji mapy głębi do pierwotnej rozdzielczości.



Rysunek 7.1. Schemat nieliniowej decymacji zastosowanej w eksperymencie do obrazów map głębi.



Rysunek 7.2. Schemat rekonstrukcji obrazu do pierwotnej rozdzielczości po decymacji.

## Aneks III

W trakcie eksperymentu dotyczącego wyznaczenia optymalnych par parametrów QP i QD kompresji podlegały widoki z dwu kamer oraz odpowiadające im mapy głębi. Do kompresji obrazów wykorzystany został koder MVC w domyślnej konfiguracji (patrz tabela 7.3 z Aneksu II). Podobnie, do kompresji map głębi wykorzystany został koder MVC w domyślnej konfiguracji.

Zdekompresowane dane wykorzystane zostały w procesie syntezy, przeprowadzonym z użyciem oprogramowania VSRS [Tani\_08b]. Użyte opcje programu do syntezy widoku pokrywały się z opcjami optymalnymi dla danej sekwencji, które zostały dostarczone razem z sekwencjami wykorzystywanymi w pracach nad standardem kompresji sekwencji wielowidokowych przez członków grupy ekspertów MPEG.

W trakcie badań kompresowano sekwencje z dwu kamer z indeksami parametru kwantyzacji wymienionymi w tabeli 7.6. Warto zaznaczyć, że w porównaniu do eksperymentu opisanego w Aneksie II, liczba punktów pomiarowych jest wielokrotnie większa i obejmuje szerszy zakres wartości. Wzrost liczby wartości parametru kwantyzacji z 7 do 42 oznacza, że liczba przebadanych par wartości QP-QD wzrosła z 49 do 1764. Taka liczba punktów pomiarowych pozwala dokładnie określić optymalne parametry kodowania map głębi i obrazów, jednak sprawia, że badania są niezwykle czasochłonne.

Tabela 7.6. Wartości indeksów QP i QD wykorzystane w eksperymencie

	<b>Parametr kwantyzacji dla kompresji obrazów QP</b>	<b>Parametr kwantyzacji dla kompresji mapy głębi QD</b>
Wartości wykorzystane w kompresji	od 10 do 51 z krokiem co 1	od 10 do 51 z krokiem co 1

Uzyskane w eksperymencie zrekonstruowane sekwencje obrazów i sekwencje map głębi użyto w procesie syntezy i wygenerowano widok wirtualny dla każdej możliwej do utworzenia pary indeksów QP-QD.

W eksperymencie wykorzystane zostały po dwa widoki z sekwencji wielowidokowych Book Arrival [Feld\_08], Newspaper [Ho\_08], Undo Dancer [Rusa\_11], Poznan Hall 2 [Doma\_09] i Poznan Street [Doma\_09] w różnych scenariuszach, których najważniejsze parametry podane zostały w tabeli 5.2 w tekście pracy. Każda sekwencja składa się z obrazów i map głębi. Sekwencje map głębi miały tę samą rozdzielczość co odpowiadające im obrazy.

W przypadku tego eksperymentu, wybór numerów kamer, z których pochodziły sekwencje był

podyktowany dostępnością map głębi wysokiej jakości, czyli wyznaczanych w sposób półautomatyczny.

Warto zwrócić również uwagę, że sekwencja rozważana w scenariuszu 3, Undo Dancer, jest sekwencją syntetyczną i dostępne są dla niej mapy głębi o bardzo wysokiej jakości.



## Aneks IV

Badania dotyczące wspólnego kodowania obrazów i map głębi sekwencji wielowidokowych jako sekwencji czterokomponentowych przeprowadzone zostały z wykorzystaniem wybranych sekwencji testowych o różnym charakterze.

W trakcie eksperymentów wykorzystano sekwencje testowe składające się z widoków z dwóch kamer wraz z odpowiadającymi im mapami głębi. Użyte w badaniach sekwencje to: Book Arrival [Feld\_08], Newspaper [Ho\_08], Undo Dancer [Rusa\_11] oraz Poznan Street [Doma\_09]. Wykorzystano ustawienia koderów podane w tabeli 7.7. Długość sekwencji użytych w testach jest niewielka w związku z koniecznością przeprowadzenia bardzo licznych kodowań, jednak na podstawie wniosków z punktu 5.2.4 należy oczekiwać, że będą one bardzo podobne do wyników uzyskiwanych w przypadku użycia dłuższych fragmentów sekwencji testowych. W badaniach wykorzystano kodowanie symboli kodami o zmiennej długości CAVLC. Podyktowane było to chęcią uproszczenia procesu implementacji poprzez wyeliminowanie konieczności opracowania oddzielnych tak zwanych kontekstów koderów CABAC [Karw\_08][Doma\_11] dla nowych danych w strumieniu wynikowym.

Szacowanie efektywności kodowania było przeprowadzone na podstawie pomiaru jakości obiektywnej widoku syntezy z obrazów zrekonstruowanych z danych zakodowanych autorskim, zmodyfikowanym koderem umożliwiającym kodowanie sekwencji czterokomponentowych. Uzyskane dla zmodyfikowanego koderów wyniki były następnie porównywane z wynikami uzyskanymi dla obrazu syntezy z obrazów zrekonstruowanych z danych zakodowanych za pomocą niezmodyfikowanego koderów MVC w tej samej wersji, przy czym sekwencje obrazów kodowane były oddzielnie od sekwencji map głębi. Taki sposób porównania pozwolił na oszacowanie zysków wynikających ze wspólnego kodowania obrazów i głębi w koderze wielowidokowym opartym na technologii MVC.

Tabela 7.7. Konfiguracja koderów MVC użyta w eksperymentach z kodowaniem sekwencji czterokomponentowych

<b>Parametr</b>	<b>Wartość</b>
Długość GOP	12 ramek
Okres występowania ramek typu I	1 (użyte są wyłącznie ramki typu I)
Liczba kodowanych ramek	12
Koder entropijny	CAVLC



## Bibliografia

- [3DV\_WD] Suzuki T., Hannuksela M.M., Chen Y., *WD on MVC extensions for inclusion of depth maps*, ISO/IEC JTC1/SC29/WG11 Doc. N12351, Genewa, Szwajcaria, grudzień 2011.
- [Aggo\_10] Aggoun A., *3D Holographic video content capture, manipulation and display technologies*, 9th Euro-American Workshop on Information Optics (WIO), 2010
- [AVC] ISO/IEC 14496-10, International standard, *Information technology – Coding of audio-visual objects – Part 10: Advanced Video Coding*, 2010, również: ITU-T Rec. H.264, 2010.
- [Akar\_07] Akar G., Teklap A.M., Fehn C., Civanlar M.R., *Transport methods in 3DTV - a survey*, IEEE Transactions on Circuits and Systems for Video Technology Vol. 17, Nr 11, str. 1622 - 1630, 2007.
- [Anders\_05] Anderson D., Herman H., Kelly A., *Experimental characterization of commercial flash lidar devices*, International Conference of Sensing and Technology, 2005.
- [Bjont\_01] Bjøntegaard G., *Calculation of average PSNR differences between RD-curves*, VCEG Doc. VCEG-M33, Austin, Teksas, kwiecień 2001.
- [Bour\_06] Bourge A., Gobert J., Bruls F., *MPEG-C part 3: enabling the introduction of video plus depth contents*, Content Generation and Coding for 3D-television Workshop, Eindhoven, 2006.
- [BT500] ITU-R Rec. BT.500-11, *Methodology for the subjective assessment of the quality of television pictures*, 2002.
- [Chai\_02] Chai B.-B., Sethuraman S., Sawhney H.S., *A depth map representation for real-time transmission and view-based rendering of a dynamic 3d scene*, International Symposium on 3D Data Processing Visualization and Transmission, 2002.
- [Chen\_06] Chen Y., Cai C., Liu J., *YUV correction for multi-view video compression*, Proceedings of International Conference on Pattern Recognition (ICPR), str. 734 - 737, 2006.
- [Chen\_09] Chen Y., Hannuksela M.M., Zhu L., Hallapuro A., Gabbouj M., Li H., *Coding techniques in multiview video coding and joint multiview video model*, Picture Coding Symposium (PCS), 2009.
- [Chen\_87] Chen C., Kak A., *Modeling and calibration of a structured light scanner for 3-D robot vision*, Proceedings of IEEE International Conference on Robotics and Automation, str. 807 - 815, 1987.
- [Cheu\_10] Cheung G., Kubota A., Ortega A., *Sparse representation of depth maps for efficient transform coding*, Picture Coding Symposium (PCS), 2010.
- [Cheu\_11] Cheung G., Velisavljević V., Ortega A., *On dependent bit allocation for multiview image coding with depth-image-based rendering*, IEEE Transactions on Image Processing Vol. 20, Nr 11, str. 3179-3194, 2011.
- [Ciec\_07] Cieciora M., Zacharski J., *Metody probabilistyczne w ujęciu praktycznym*, wyd. Vizja Press&IT, Warszawa 2007.

- [Cigla\_09] Cigla C., Alatan A.A., *Temporally consistent dense depth map estimation via belief propagation*, 3DTV Conference: The True Vision - Capture, Transmission and Display of 3D Video, 2009.
- [Cyg\_02] Cyganek B., *Komputerowe przetwarzanie obrazów trójwymiarowych*, wyd. EXIT, Warszawa 2002.
- [Dari\_09] Daribo I., Tillier C., Pesquet-Popescu B., *Motion vector sharing and bit rate allocation for 3D video-plus-depth coding*, EURASIP Journal on Advances in Signal Processing Vol. 2009 (Special Issue on 3DTV), Article ID 258920, 2009.
- [DeSil\_09] De Silva D.V.S.X., Fernando W.A.C., Yasakethu S.L.P., *Object based coding of the depth maps for 3D video coding*, IEEE Transactions on Consumer Electronics Vol. 55, Nr 3, str. 1699-1706, 2009.
- [Doma\_09] Domański M., Grajek T., Klimaszewski K., Kurc M., Stankiewicz O., Stankowski J., Wegner K., *Poznań multiview video test sequences and camera parameters*, ISO/IEC JTC1/SC29/WG11 MPEG Doc. M17050, Xian, Chiny, październik 2009.
- [Doma\_09a] Domański M., Klimaszewski K., Konieczny J., Kurc M., Łuczak A., Stankiewicz O., Wegner K., *An experimental free-view television system*, Image Processing and Communications Challenges, wyd. EXIT, Warszawa 2009.
- [Doma\_10] Domański M., Klimaszewski K., Stankiewicz O., Stankowski J., Wegner K., *Efficient transmission of 3D video using MPEG-4 AVC/H.264 compression technology*, Computer Communication Networks and Telecommunications, Proceedings of Third International Workshop, FMN 2010, Lecture Notes in Computer Science Vol. 6157/2010, str. 145 - 156, Springer-Verlag, Berlin, 2010.
- [Doma\_10a] Domański M., *Telewizja trójwymiarowa - stan badań i perspektywy rozwoju*, Przegląd Telekomunikacyjny, str. 223 - 228, Nr 6/2010.
- [Doma\_11] Domański M., *Obraz cyfrowy*, wyd. Wydawnictwo Komunikacji i Łączności, Warszawa 2010.
- [Doma\_11a] Domański M., Grajek T., Karwowski D., Klimaszewski K., Konieczny J., Kurc M., Łuczak A., Ratajczak R., Siast J., Stankiewicz O., Stankowski J., Wegner K., *Multiview HEVC – experimental results*, JCT-VC ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11, Doc. JCTVC-G582 oraz MPEG Doc. M22147, Genewa, Szwajcaria, listopad 2011.
- [Doma\_11b] Domański M., Grajek T., Karwowski D., Klimaszewski K., Konieczny J., Kurc M., Łuczak A., Ratajczak R., Siast J., Stankiewicz O., Stankowski J., Wegner K., *Technical description of Poznan University of Technology proposal for call on 3d video coding technology*, ISO/IEC JTC1/SC29/WG11, MPEG Doc. M22697, Genewa, Szwajcaria, listopad 2011.

- [Doma\_12] Domański M., Grajek T., Karwowski D., Konieczny J., Kurc M., Łuczak A., Ratajczak R., Siast J., Stankiewicz O., Stankowski J., Wegner K., *Coding of multiple video+depth using HEVC technology and reduced representations of side views and depth maps*, IEEE Picture Coding Symposium (PCS), 2012.
- [Doma\_12a] Domański M., Grajek T., Karwowski D., Klimaszewski K., Konieczny J., Kurc M., Łuczak A., Ratajczak R., Siast J., Stankiewicz O., Stankowski J., Wegner K., *New coding technology for 3D video with depth maps as proposed for standardization within MPEG*, International Conference on Systems, Signals and Image Processing (IWSSIP), 2012.
- [Dose\_12] Doser I.T., Hassenpflug F., Zwing R., *The Glasses Are Half Empty: Turn Your Sight to Autostereoscopic 3-D*, IEEE Consumer Electronics Magazine Vol. 1, Nr 1, str. 49-60, 2012.
- [Ekme\_07] Ekmekcioglu E., Worrall S., Kondoz A.M., *Multi-view video coding via virtual view generation*, Picture Coding Symposium (PCS), 2007.
- [El-Y\_10] El-Yamany N. A., Ugur K., Hannuksela M. M., Gabbouj M., *Evaluation of depth compression and view synthesis distortions in multiview-video-plus-depth coding systems*, 3DTV-Conference: The True Vision - Capture, Transmission and Display of 3D Video (3DTV-CON), 2010.
- [Fan\_11] Fan Y.-C., Wu S.-F., Lin B.-L., *Three-dimensional depth map motion estimation and compensation for 3D video compression*, IEEE Transactions on Magnetics, Vol. 47, Nr 3, str. 691-695, 2011.
- [Feck\_05] Fecker U., Kaup A., *Statistical analysis of multi-reference block matching for dynamic light field coding*, Proceedings of 10th International Fall Workshop Vision, Modeling, and Visualization, 2005.
- [Feck\_08] Fecker U., Barkowsky M., Kaup A., *Histogram-based prefiltering for luminance and chrominance compensation of multiview video*, IEEE Transactions on Circuits and Systems for Video Technology Vol. 18, Nr 9, str. 1258-1267, 2008.
- [Feld\_08] Feldmann I., Müller M., Zilly F., Tanger R., Müller K., Smolic A., Kauff P., Wiegand T., *HHI Test Material for 3D Video*, ISO/IEC JTC1/SC29/WG11 MPEG Doc. M15413, Archamps, Francja, kwiecień 2008.
- [Flier\_07] Flierl M., Mavlankar A., Girod B., *Motion and disparity compensated coding for multiview video*, IEEE Transactions on Circuits and Systems for Video Technology Vol. 17, Nr 11, str. 1474-1484, 2007.
- [Forst\_06] Forster F., *A high-resolution and high accuracy real-time 3D sensor based on structured light*, Proceedings of Third International Symposium on 3D Data Processing, Visualization, and Transmission, str. 208 - 215, 2006.
- [Free\_10] Freedman B., Shpunt A., Machline M., Arieli Y., *Depth mapping using projected patterns*, Wniosek patentowy US 2010/0118123 A1, opublikowany 13 maja 2010.

- [Fuji\_06] Fuji T., Mori K., Takeda K., Mase K., Tanimoto M., Suenaga Y., *Multipoint measuring system for video and sound: 100-camera and microphone system*, Proceedings of IEEE International Conference on Multimedia Expo, 2006.
- [Fusie\_00] Fusiello A., Trucco E., *A compact algorithm for rectification of stereo pairs*, Machine Vision and Applications, Vol. 12, Nr 1, str. 16 - 22, wyd. Springer, 2000.
- [Gord\_06] Gordon I., Lowe D.G., *What and where: 3D object recognition with accurate pose*, Toward category-level object recognition, str. 67-82, wyd. Springer, 2006.
- [Grew\_04] Grewatsch S., Müller E., *Sharing of motion vectors in 3D video coding*, International Conference on Image Processing (ICIP), Vol. 5, str. 3271-3274, 2004.
- [Han\_03] Han F., Zhu S.C., *Bayesian reconstruction of 3D shapes and scenes from a single image*, IEEE International Workshop on Higher-Level Knowledge in 3D Modeling and Motion Analysis, str 12 - 20, 2003.
- [HEVC] Bros B., Han W.-J., Sullivan G., Ohm J.-R., Wiegand T., *High Efficiency Video Coding (HEVC) text specification draft 6*, JCT-VC ITU-T SG16 WP3 Doc. JCTVC-H1003, USA, San Jose, luty 2012.
- [Hewa\_09] Hewage C., Worrall S., Dogan S., Villette S., Kondo A., *Quality evaluation of color plus depth map-based stereoscopic video*, IEEE Journal of Selected Topics in Signal Processing Vol. 3, Nr 2, str. 304-318, 2009.
- [Ho\_08] Ho Y.-S., Lee E.-K., Lee C., *Multiview video test sequence and camera parameters*, ISO/IECJTC1/SC29/WG11 MPEG Doc. M15419, Archamps, Francja, kwiecień 2008.
- [Hoi\_05] Hoiem D., Efros A., *Geometric context from a single image*, International Conference on Computer Vision (ICCV), 2005.
- [Hou\_09] Hou L., Au O.C., Fan X., Ma M., *A novel ray-space based color correction algorithm for multi-view video*, Proceedings of IEEE International Symposium on Circuits and Systems (ISCAS), 2009.
- [Huo\_11] Huo J., Chang Y., Ma Y., *Efficient prediction structure for key pictures in multiview video coding*, Symposium on Photonics and Optoelectronics (SOPO), 2011.
- [Ince\_07] Ince S., Martinian E., Yea S., Vetro A., *Depth estimation for view synthesis in multiview video coding*, 3DTV Conference, 2007.
- [Jager\_12] Jäger F., Ballé J., *Median trilateral loop filter for depth map video coding*, Picture Coding Symposium (PCS), 2012.
- [JMVC] cvs: garcon.ient.rwth-aachen.de/cvs/jvt, ostatnio pobrane w kwietniu 2011.
- [JPEG2000] ISO/IEC 15444-1, International standard, *Information technology - JPEG 2000 image coding system: Core coding system*, 2004.
- [Jung\_12] Jung S., *Enhancement of image and depth map using adaptive joint trilateral filter*, IEEE Transactions on Circuits and Systems for Video Technology, art. przyjęty do publikacji.

- [Kamo\_08] Kamolrat B., Fernando W.A.C., Mrak M., *3D motion estimation for depth information compression in 3D-TV applications*, Electronic Letters Vol. 44, Nr 21, str. 1244-1245, 2008.
- [Kamo\_09] Kamolrat B., Fernando W.A.C., Mrak M., Kondoz A., *3D motion estimation for depth image coding in 3D video coding*, IEEE Transactions on Consumer Electronics Vol. 55, Nr 2, str. 824-830, 2009.
- [Kang\_11] Kang Y.-S., Ho Y.-S., *Disparity map generation for color image using TOF depth camera*, 3DTV Conference: The True Vision - Capture, Transmission and Display of 3D Video (3DTV-CON), 2011.
- [Karw\_08] Karwowski D., *Advanced Adaptation Algorithms of Arithmetic Coding in Hybrid Video Compression*, rozprawa doktorska, Politechnika Poznańska, Wydział Elektroniki i Telekomunikacji, Poznań, 2008.
- [Kauff\_07] Kauff P., Atzpadin N., Fehn C., Müller M., Schreer O., Smolic A., Tanger R., *Depth map creation and image-based rendering for advanced 3DTV services providing interoperability and scalability*, Signal Processing: Image Communication Vol. 22, Nr 2, str. 217–234, 2007.
- [Kaup\_06] Kaup A., Fecker U., *Analysis of multi-reference block matching for multi-view video coding*, 7th Workshop Digital Broadcasting, Erlangen, Niemcy, wrzesień 2006.
- [Kawa\_04] Kawada R., *KDDI multiview video sequences for MPEG 3DAV use*, ISO/IEC JTC1/SC29/WG11 MPEG Doc. M10533, Monachium, Niemcy, marzec 2004.
- [Kim\_07] Kim S.Y., Ho Y.S., *Mesh-based depth coding for 3D video using hierarchical decomposition of depth maps*, IEEE International Conference on Image Processing (ICIP), 2007.
- [Kim\_09] Kim W.-S., Ortega A., Lai P., Tian D., Gomila C., *Depthmap distortion analysis for view rendering and depth coding*, Proceedings of IEEE International Conference on Image Processing (ICIP), 2009.
- [Kim\_10] Kim W.-S., Ortega A., Lai P.L., Tian D., Gomila C., *Depth map coding with distortion estimation of rendered view*, SPIE Visual Information Processing and Communication, Proceedings of SPIE Vol. 7543, 75430B, 2010.
- [Kinect] <http://www.xbox.com/pl-PL/Kinect>, ostatni dostęp 20.04.2012.
- [Klim\_08] Klimaszewski K., *Ocena efektywności kompresji sekwencji wielowidokowych*, Przegląd Telekomunikacyjny, str. 581-584, Nr 4/2008.
- [Klim\_09] Klimaszewski K., Wegner K., *Wpływ kompresji obrazów i map głębi na syntezę widoków w systemie wielowidokowym*, Przegląd Telekomunikacyjny, str. 380-383, Nr 6/2009.
- [Klim\_09a] Klimaszewski K., Wegner K., Domański M., *Influence of views and depth compression onto quality of synthesized views*, ISO/IEC JTC1/SC29/WG11 MPEG Doc. M16758, Londyn, Wielka Brytania, czerwiec 2009.
- [Klim\_09b] Klimaszewski K., Wegner K., Domański M., *Distortions of synthesized views caused by compression of views and depth maps*, 3DTV Conference: The True Vision - Capture, Transmission and Display of 3D Video, 2009.

- [Klim\_10] Klimaszewski K., Wegner K., Domański M., *Joint intra coding of video and depth maps*, International Conference on Signals and Electronic Systems (ICSSES), 2010.
- [Klima\_07] Klimaszewski K., *Disparity prediction in multiview coding*, Poznań University of Technology Academic Journals - Electrical Engineering Nr 54, str. 89-96, 2007.
- [Kolmo\_02] Kolmogorov V., Zabih R., *Multi-camera scene reconstruction via graph cuts*, Proceedings of European Conference on Computer Vision (ECCV) – part 3, str. 82-96, 2002.
- [Konie\_10a] Konieczny J., Domański M., *Depth-based inter-view prediction of motion vectors for improved multiview video coding*, 3DTV Conference: The True Vision - Capture, Transmission and Display of 3D Video (3DTV-CON), 2010.
- [Konie\_10b] Konieczny J., Domański M., *Inter-view direct mode for multiview video coding*, ISO/IEC JTC1/SC29/WG11, MPEG Doc. M17800, Genewa, Szwajcaria, lipiec 2010.
- [Konie\_11] Konieczny J., Domański M., *Extended inter-view direct mode for multiview video coding*, IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2011.
- [Koro\_09] Koronacki J., Mielniczuk J., *Statystyka dla studentów kierunków technicznych i przyrodniczych*, wyd. WNT, Warszawa 2009.
- [Krish\_01] Krishnamurthy R., Chai B.-B., Tao H., Sethuraman S., *Compression and transmission of depth maps for image based rendering*, Proceedings of International Conference on Image Processing (ICIP) Vol.3 str. 828-831, 2001.
- [Kubo\_07] Kubota A., Smolic A., Magnor M., Tanimoto M., Chen T., Zhang C., *Multiview Imaging and 3DTV*, IEEE Signal Processing Magazine, Vol. 24, Nr 6, str. 10-21, 2007
- [Kuo\_09] Kuo T.-Y., Yeh C.-K., Tsai H.-Y., *A novel method for global disparity vector estimation in multiview video coding*, IEEE International Symposium on Circuits and Systems (ISCAS), 2009.
- [Lange\_01] Lange R., Seitz P., *Solid-state time-of-flight range camera*, IEEE Journal of Quantum Electronics Vol. 37, Nr 3, str. 390-397, 2001.
- [Lavo\_04] Lavoie P., Ionescu D., Petriu E.M., *3D object model recovery from 2D images using structured light*, IEEE Transactions on Instrumentation and Measurement Vol. 53, Nr 2, str. 437-443, 2004.
- [Lee\_08] Lee E.-K., Kim S.-Y., Jung Y.-K., Ho Y.-S., *High-resolution depth map generation by applying stereo matching based on initial depth information*, 3DTV Conference: The True Vision - Capture, Transmission and Display of 3D Video (3DTV-CON), 2008.
- [Lee\_11] Lee J.Y., Wey H., Park D.-S., Kim C.-Y., *Temporal and inter-view skip modes for multi-view video coding*, 3DTV Conference: The True Vision - Capture, Transmission and Display of 3D Video (3DTV-CON), 2011.
- [LeeSB\_08] Lee S.-B., Oh K.-J., Ho Y.-S., *Segment-based multi-view depth map estimation using belief propagation from dense multi-view video*, 3DTV Conference: The True Vision - Capture, Transmission and Display of 3D Video (3DTV-CON), 2008.



- [Leon\_08] Leon G., Kalva H., Furht B., *3D video quality evaluation with depth quality variations*, 3DTV Conference: The True Vision - Capture, Transmission and Display of 3D Video (3DTV-CON), 2008.
- [Li\_06] Li P., Farin D., Gunnewiek R.K., de With P.H.N., *On creating depth maps from monoscopic video using structure from motion*, Content Generation and Coding for 3D-television Workshop, Eindhoven, 2006.
- [Li\_09] Li Z., Xie X., Liu X., *An efficient 2D to 3D video conversion method based on skeleton line tracking*, 3DTV Conference: The True Vision - Capture, Transmission and Display of 3D Video (3DTV-CON), 2009.
- [Li\_11] Li X., Wang Y., Zhao D., Jiang T., Zhang N., *Joint just noticeable difference model based on depth perception for stereoscopic images*, IEEE Visual Communications and Image Processing Conference (VCIP), Tajwan, Tainan, 2011.
- [Lin\_11] Lin Y.-H., Wu J.-L., *A depth information based fast mode decision algorithm for color plus depth-map 3d videos*, IEEE Transactions on Broadcasting Vol. 57, Nr 2, str. 542-550, 2011.
- [LiT\_09] Li T., Ji X., Dai Q., *Depth map recovery for multi-view using belief propagation*, 3DTV Conference: The True Vision - Capture, Transmission and Display of 3D Video (3DTV-CON), 2009.
- [Liu\_10] Liu S., Lai P., Tian D., Gomila C., Chen C., *Joint trilateral filtering for depth map compression*, Visual Communications and Image Processing (VCIP), 2010.
- [Mait\_08] Maitre M., Do M.N., *Joint encoding of the depth image based representation using shape-adaptive wavelets*, International Conference on Image Processing (ICIP), 2008.
- [Mait\_10] Maitre M., Do M.N., *Depth and depth-color coding using shape-adaptive wavelets*, Journal of Visual Communication and Image Representation, Vol. 21, Nr 5-6, str. 513-522, 2010.
- [Merk\_07] Merkle P., Smolic A., Müller K., Wiegand T., *Efficient prediction structures for multiview video coding*, IEEE Transactions on Circuits and Systems for Video Technology Vol. 17, Nr 11, str. 1461-1473, 2007.
- [Merk\_08] Merkle P., Morvan Y., Smolic A., Farin D., Müller K. de With P.H.N. Wiegand T., *The effect of depth compression on multiview rendering quality*, 3DTV Conference: The True Vision - Capture, Transmission and Display of 3D Video (3DTV-CON), 2008.
- [Merk\_09] Merkle P., Brust H., Dix K., Müller K., Wiegand T., *Stereo video compression for mobile 3D services*, 3DTV Conference: The True Vision - Capture, Transmission and Display of 3D Video (3DTV-CON), 2009.
- [Merk\_09a] Merkle P., Wang Y., Müller K., Smolic A., Wiegand T., *Video plus depth compression for mobile 3D services*, 3DTV Conference: The True Vision - Capture, Transmission and Display of 3D Video (3DTV-CON), 2009.
- [Middlebury] Middlebury Stereo Vision Page, <http://vision.middlebury.edu/stereo/> ostatni dostęp 11.06.2012.

- [Mori\_09] Mori Y., Fukushima N., Yendo T., Fujii T., Tanimoto M., *View generation with 3D warping using depth information for FTV*, Signal Processing, Image Communications, vol. 24, Special Issue on Advances in Three-Dimensional Television and Video, Nr 1–2, str. 65-72, 2009.
- [Morv\_05] Morvan Y., Farin D., de With P.H.N., *Coding depth images with piecewise linear functions for multi-view synthesis*, European Signal Processing Conference (EUSIPCO), 2005.
- [Morv\_05a] Morvan Y., Farin D., de With P.H.N., *Novel coding technique for depth images using quadtree decomposition and plane approximation*, Visual Communications and Image Processing (VCIP), 2005.
- [Morv\_07] Morvan Y., Farin D., de With P.H.N., *Incorporating depth-image based view-prediction into H.264 for multiview-image coding*, IEEE International Conference on Image Processing (ICIP), 2007.
- [Morv\_07a] Morvan Y., Farin D., de With P.H.N., *Joint depth/texture bit-allocation for multi-view video compression*, Picture Coding Symposium (PCS), 2007.
- [Morv\_07b] Morvan Y., Farin D., de With P.H.N., *Multiview depth-image compression using an extended H.264 encoder*, International Conference on Advanced Concepts for Intelligent Vision Systems (ACIVS), 2007.
- [Morv\_09] Morvan Y., *Acquisition, Compression and Rendering of Depth and Texture for Multi-View Video*, praca doktorska, Technische Universiteit Eindhoven, 2009.
- [MPEG\_CfP] *Call for Proposals on 3D Video Coding Technology*, ISO/IEC JTC1/SC29/WG11 MPEG Doc. N12036, Genewa, Szwajcaria, marzec 2011.
- [MPEG-C\_3] ISO/IEC 23002-3 International standard, *Auxiliary Video Data Representations, MPEG-C Part 3*, 2006.
- [Mull\_08] Müller K., Smolic A., Dix K., Merkle P., Kauff P., Wiegand T., *View synthesis for advanced 3D video systems*, EURASIP Journal on Image Video Processing, vol. 2008, Special Issue on 3D Image and Video Processing, 2008
- [Mull\_09] Müller K., Smolic A., Dix K., Merkle P., Wiegand T., *Coding and intermediate view synthesis of multiview video plus depth*, IEEE International Conference on Image Processing (ICIP), 2009.
- [Mull\_11] Müller K., Merkle P., Wiegand T., *3-D video representation using depth maps*, Proceedings of the IEEE, str. 643-656, Nr 4, 2010.
- [Na\_08] Na S.T., Oh K.-J., Ho Y.-S., *Joint coding of multi-view video and corresponding depth map*, IEEE International Conference on Image Processing (ICIP), 2008.
- [Nunes\_12] Nunes P., Soares L.D., Aggoun A., *Recent advances in 3d holoscopic video*, ISO/IEC JTC1/SC29/WG11 MPEG Doc. M25146, Genewa, Szwajcaria, maj 2012.
- [Nur\_10] Nur G., Dogan S., Arachchi H.K., Kondoz A.M., *Impact of depth map spatial resolution on 3D video quality and depth perception*, 3DTV-Conference: The True Vision - Capture, Transmission and Display of 3D Video (3DTV-CON), 2010.

- [Oggi\_03] Oggier T., Lehmann M., Kaufmann R., Schweizer M., Richter M., Metzler P., Lang G., Lustenberger F., Blanc N., *An all-solid-state optical range camera for 3d real-time imaging with sub-centimeter depth resolution (SwissRanger)*, SPIE Conference on Optical System Design, Proceedings of SPIE, Vol. 5249, str. 534-545, 2003.
- [Oh\_06] Oh H., Ho Y.-S., *H.264-based depth map sequence coding using motion information of corresponding texture video*, Pacific Rim Symposium on Advances in Image and Video Technology (PSIVT), 2006.
- [Oh\_09] Oh K.-J., Yea S., Vetro A., Ho Y.-S., *Depth reconstruction filter for depth coding*, IET Electronic Letters Vol. 45, Nr 6, str. 305-306, 2009.
- [Oh\_09a] Oh K.-J., Yea S., Vetro A., Ho Y.-S., *Depth reconstruction filter and down/upsampling for depth coding in 3D video*, IEEE Signal Processing Letters Vol. 16, Nr 9, str. 747-750, 2009.
- [Oh\_11] Oh K.-J., Vetro A., Ho Y.-S., *Depth coding using a boundary reconstruction filter for 3-D video systems*, IEEE Transactions on Circuits and Systems for Video Technology Vol. 21, Nr 3, str. 350-359, 2011.
- [Oster\_10] Ostermann J., *3D information coding*, Picture Coding Symposium (PCS), 2010.
- [Park\_95] Park D., Kim K., Lee C., Son J., Lee Y., *Lenticular stereoscopic imaging and displaying techniques with no special glasses*, Proceedings of International Conference on Image Processing ICIP, Vol. 3, str. 137-140, 1995.
- [Pfis\_04] Pfister H., Vetro A., Sun H., *3D TV System*, ISO/IEC JTC1/SC29/WG11 MPEG Doc. M10624, Monachium, Niemcy, marzec 2004.
- [Przel\_05] Przelaskowski A., *Kompresja danych: podstawy, metody bezstratne, kodery obrazów*, wyd. BTC, Warszawa 2005.
- [Red\_02] Redert A., de Beeck M.O., Fehn C., Ijsselsteijn W., Pollefeys M., Van Gool L., Ofek E., Sexton I., Surman P., *Advanced three-dimensional television system technologies*, Proceedings of First International Symposium on 3D Data Processing Visualization and Transmission, str. 313-319, 2002.
- [Rich\_10] Richardson I.E., *The H.264 advanced video compression standard second edition*, wyd. John Wiley & Sons, 2010.
- [Rusa\_11] Rusanovskyy D., Aflaki P., Hannuksela M. M., *Undo Dancer 3DV sequence for purposes of 3DV standardization*, ISO/IEC JTC1/SC29/WG11 MPEG Doc. M20028, Genewa, Szwajcaria, marzec 2011.
- [Rusa\_12] Rusanovskyy D., Tian D., Hannuksela M.M., Vetro A., *Compression efficiency of 3DV-ATM in MVC-compatible and AVC-compatible coding configurations*, ISO/IEC JTC1/SC29/WG11 MPEG Doc. M24913, Genewa, Szwajcaria, maj 2012.
- [Sax\_07] Saxena A., Sun M., Ng A.Y., *Learning 3-D scene structure from a single still image*, International Conference on Computer Vision (ICCV) workshop on 3D Representation for Recognition, 2007.

- [Schar\_03] Scharstein D., Szeliski R., *High-accuracy stereo depth maps using structured light*, Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition Vol. 1, str. 195-202, 2003.
- [Schw\_12] Schwarz S., Olsson R., Sjoström M., Tourancheau S., *Adaptive depth filtering for HEVC 3D video coding*, Picture Coding Symposium, 2012.
- [Senoh\_07] Senoh T., Yamamoto K., Oi R., Mishina T., Okui M., *Consideration of depth format*, ISO/IEC JTC1/SC29/WG11 MPEG Doc. M15047, Antalya, Turcja, styczeń 2007.
- [Seo\_10] Seo J., Park D., Wey H.-C., Lee S., Sohn K., *Motion information sharing mode for depth video coding*, 3DTV-Conference: The True Vision - Capture, Transmission and Display of 3D Video (3DTV-CON), 2010.
- [Sext\_99] Sexton I., Surman P., *Stereoscopic and autostereoscopic display systems*, IEEE Signal processing magazine Vol. 16, Nr 3, str. 85-99, 1999.
- [Shen\_10] Shen G., Kim W.-S., Narang S.K., Ortega A., Lee J., Wey H., *Edge-adaptive transforms for efficient depth map coding*, Picture Coding Symposium (PCS), 2010.
- [Shen\_10a] Shen G., Kim W.-S., Ortega A., Lee J., Wey H., *Edge-aware intra prediction for depth-map coding*, International Conference on Image Processing (ICIP), 2010.
- [Shimi\_07] Shimizu S., Kitahara M., Kimata H., Kamikura K., Yashima Y., *View scalable multiview video coding using 3-D warping with depth map*, IEEE Transactions on Circuits and Systems for Video Technology Vol. 17, Nr 11, str. 1485-1495, 2007.
- [Skar\_09] Skarbek W., Nowakowski A., *Filter-less gray patterns detection in 3D modeling by structured light*, Photonics Applications in Astronomy, Communications, Industry, and High-Energy Physics Experiments, Proceedings of SPIE, Vol 7502, 750203, 2009.
- [Smis\_11] Smisek J., Jancosek M., Pajdla T., *3D with Kinect*, IEEE International Conference on Computer Vision Workshops, 2011.
- [Smol\_06] Smolic A., Müller K., Merkle P., Fehn C., Kauff P., Eisert P., Wiegand T., *3D video and free viewpoint video - technologies, applications and MPEG standards*, IEEE International Conference on Multimedia and Expo (ICME), 2006.
- [Souri\_10] Sourimant G., *Depth maps estimation and use for 3DTV*, Raport techniczny INRIA 0379.
- [Stank\_08] Stankiewicz O., Wegner K., *Depth map estimation software version 3*, ISO/IEC JTC1/SC29/WG11 MPEG Doc. M15540, Hanower, Niemcy, lipiec 2008.
- [Stank\_10] Stankiewicz O., Domanski M., Wegner K., *Stereoscopic depth refinement by mid-level hypothesis*, IEEE International Conference on Multimedia and Expo (ICME), 2010.
- [Stanko\_10] Stankowski J., Klimaszewski K., *Application of epipolar rectification algorithm in 3D Television*, Advances in Intelligent and Soft Computing Vol. 84, Image Processing and Communications Challenges 2, str. 345 - 352, wyd. Springer, 2010.

- [Stanko\_10a] Stankowski J., Klimaszewski K., Stankiewicz O., Wegner K., Domański M., *Preprocessing methods used for Poznan 3D/FTV test sequences*, ISO/IEC JTC1/SC29/WG11 MPEG Doc. M17174, Kyoto, Japonia, 2010.
- [Stre\_02] Strecha C., Van Gool L., *PDE-based multi-view depth estimation*, Proceedings of First International Symposium on 3D Data Processing Visualization and Transmission, str. 416-425, 2002.
- [Suarez\_12] Cancino-Suárez S., Klimaszewski K., Stankiewicz O., Domański M., *Enhancement of stereoscopic depth estimation by the use of motion information*, IEEE Southeastern Symposium on System Theory, 2012.
- [Sun\_03] Sun J., Zheng N.-N., Shum H.-Y., *Stereo matching using belief propagation*, IEEE Transactions on Pattern Analysis and Machine Intelligence Vol. 25, Nr 7, str. 787-800, 2003.
- [Taka\_08] Takahashi H., Kureyama N., Chikayama M., Yamada K., *Flatbed-type three-dimensional display system as a tool for cooperation working*, International Conference on Innovative Computing Information and Control (ICICIC), 2008.
- [Tani\_05] Tanimoto M., Fuji T., Senoh T., Aoki T., Sugihara Y., *Test sequences with different camera arrangements for call for proposals on multiview video coding*, ISO/IEC JTC1/SC29/WG11 MPEG Doc. M12338, Poznań, Polska, lipiec 2005.
- [Tani\_06] Tanimoto M., *Overview of free viewpoint television*, Signal Processing: Image Communications Vol. 21, Nr 6, str. 454 - 461, 2006.
- [Tani\_08] Tanimoto M., Fujii T., Fukushima N., *1D parallel test sequences for MPEG-FTV*, ISO/IEC JTC1/SC29/WG11 MPEG Doc. M15378, Archamps, Francja, kwiecień 2008.
- [Tani\_08b] Tanimoto M., Fujii T., Suzuki K., Fukushima N., Mori Y., *Reference softwares for depth estimation and view synthesis*, ISO/IEC JTC1/SC29/WG11 MPEG Doc. M15377, Archamps, Francja, kwiecień 2008.
- [Tani\_09] Tanimoto M., *Overview of FTV (free-viewpoint television)*, IEEE International Conference on Multimedia and Expo (ICME), 2009.
- [Tani\_11] Tanimoto M., Panahpour Tehrani M., Fuji T., Yendo T., *Free-viewpoint TV*, IEEE Signal Processing Magazine, Vol. 28, Nr 1, str. 67-76, 2011.
- [Tani\_12] Tanimoto M., Panahpour Tehrani M., Fujii T., Yendo T., *FTV for 3-D spatial communication*, Proceedings of the IEEE, Vol. 100, Nr 4, str. 905 - 917, 2012.
- [Tao\_09] Tao S., Chen Y., Hannuksela M.M., Wang Y.-K., Gabbouj M., Li H., *Joint texture and depth map video coding based on the scalable extension of H.264/AVC*, IEEE International Symposium on Circuits and Systems (ISCAS), str. 2353 - 2356, 2009.
- [Tech\_09] Tech G., Smolic A., Brust H., Merkle P., Dix K., Wang Y., Müller K., Wiegand T., *Optimization and comparison of coding algorithms for mobile 3DTV*, 3DTV Conference: The True Vision - Capture, Transmission and Display of 3D Video (3DTV-CON), 2009.
- [Tour\_07] Tourapis A.M., Leontaris A., Sühring K., Sullivan G., *H.264/MPEG-4 AVC Reference Software Manual*, ITU-T ISO/IEC JTC1 Doc. JVT-X072, 2007.

- [Tzov\_96] Tzovaraz D., Grammalidis N., Strintzis M., *Disparity field and depth map coding for multiview image sequence compression*, Proceedings of International Conference on Image Processing (ICIP) Vol. 2, str. 887-890, 1996.
- [Vale\_12] Valenzise G., Cheung G., Galvao R., Cagnazzo M., Pesquet-Popescu B., Ortega A., *Motion prediction of depth video for depth-image-based rendering using don't care regions*, Picture Coding Symposium (PCS), 2012.
- [Vetr\_05] Vetro A., McGuire M., Matusik W., Behrens A., Jinho Lee, Pfister H., *Multiview Video Test Sequences from MERL*, ISO/IEC JTC1/SC29/WG11 MPEG Doc. M12077, Busan, Korea, kwiecień 2005.
- [Wang\_09] Wang Z., Bovik A.C., *Mean squared error: Love it or leave it? A new look at Signal Fidelity Measures*, IEEE Signal Processing Magazine Vol. 26, Nr 1, str. 98-117, 2009.
- [Wegn\_10] Wegner K., Stankiewicz O., Klimaszewski K., Domański M., *Comparison of multiview compression performance using MPEG-4 MVC and prospective HVC technology*, ISO/IEC JTC1/SC29/WG11 MPEG Doc. M17913, Genewa, Szwajcaria, lipiec 2010.
- [Wegn\_10a] Wegner K., Klimaszewski K., Stankiewicz O., Stankowski J., Domański M., *Przygotowanie wielowidokowych sekwencji wizyjnych dla badań nad telewizją trójwymiarową*, Przegląd Telekomunikacyjny, str. 304-307, Nr 6/2010.
- [Wild\_10] Wildeboer M.O., Yendo T., Tehrani M.P., Fujii T., Tanimoto M., *Color based depth up-sampling for depth compression*, Picture Coding Symposium (PCS), 2010.
- [Willet\_02] Willet R., Nowak R., *Platelets for multiscale analysis in photon-limited imaging*, IEEE International Conference on Image Processing (ICIP), 2002.
- [Yama\_07] Yamamoto K., Kitahara M., Kimata H., Yendo T., Fujii T., Tanimoto M., Shimizu S., Kamikura K., Yashima Y., *Multiview video coding using view interpolation and color correction*, IEEE Transactions on Circuits and Systems for Video Technology Vol. 17, Nr 11, str. 1436-1449, 2007.
- [Yang\_05] Yang X.K., Lin W., Lu Z.K., Ong E.P., Yao S.S., *Just noticeable distortion model and its applications in video coding*, Signal Processing: Image Communication, Vol. 20, Nr 7, str. 662 - 680, 2005.
- [Yea\_07] Yea S., Vetro A., *RD-optimized view synthesis prediction for multiview video coding*, IEEE International Conference on Image Processing (ICIP), 2007.
- [Zabu\_08] Zabulis X., Grammalidis N., Bastanlar Y., Yilmaz E., Cetin Y.Y., *3D scene reconstruction based on robust camera motion estimation and space sweeping for a cultural heritage virtual tour system*, 3DTV Conference: The True Vision - Capture, Transmission and Display of 3D Video (3DTV-CON), 2008.
- [Zanutt\_09] Zanuttigh P., *Compression of depth information for 3D rendering*, 3DTV Conference: The True Vision - Capture, Transmission and Display of 3D Video (3DTV-CON), 2009.

- [Zhang\_10] Zhang Q., An P., Zuo Y., Zhang Z., *A novel depth coding scheme for multiview video plus depth compression*, International Conference on Audio Language and Image Processing (ICALIP), 2010.
- [Zhang\_10a] Zhang J., Hannuksela M.M., Li H., *Joint multiview video plus depth coding*, IEEE International Conference on Image Processing (ICIP), 2010.
- [Zhang\_11] Zhang Q., An P., Zhang Y., Zhang Z., *Efficient rendering distortion estimation for depth map compression*, IEEE International Conference on Image Processing (ICIP), 2011.
- [Zhu\_08] Zhu J., Wang L., Yang R., Davis J., *Fusion of time-of-flight depth and stereo for high accuracy depth maps*, IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2008.
- [Zitn\_04] Zitnick C.L., Kang S.B., Uyttendaele M., Winder S., Szeliski R., *High-quality video view interpolation using a layered representation*, Proceedings of ACM SIGGRAPH Vol. 23, Nr. 3, str. 600-608, 2004.

