

Politechnika Poznańska  
Wydział Elektroniki i Telekomunikacji  
Katedra Telekomunikacji Multimedialnej i Mikroelektroniki

Streszczenie rozprawy doktorskiej

# **Algorytmy kompresji sekwencji wielowidokowych**

Krzysztof Klimaszewski

Promotor: prof. dr hab. inż. Marek Domański

Poznań 2012

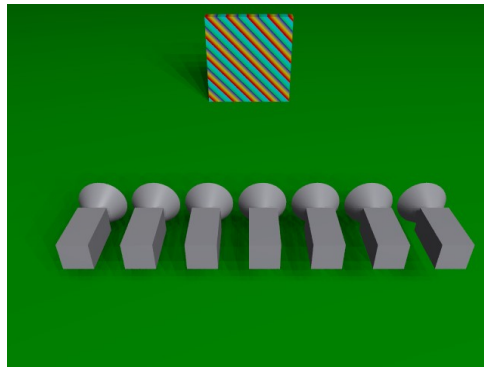
## Spis treści

1. Wprowadzenie.....	3
1.1. Zakres pracy.....	3
1.2. Założenia pracy.....	5
1.3. Cel i teza pracy.....	6
1.4. Przegląd pracy.....	6
2. Najważniejsze wyniki prezentowane w pracy.....	8
2.1. Przetwarzanie wstępne sekwencji wielowidokowych.....	8
2.2. Wpływ zniekształceń kodowania obrazów i głębi na jakość sekwencji syntezowanych.....	10
2.3. Wybór optymalnych par QP-QD dla oddzielnego kodowania obrazów i map głębi.....	15
2.4. Łączne kodowanie obrazu i głębi.....	24
3. Podsumowanie pracy.....	30
4. Główne osiągnięcia.....	32

# 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 (przykładowy układ kamer w takim systemie przedstawiony jest na rysunku 1.1). Sekwencje obrazów wielowidokowych są wykorzystywane w systemach wizji wielowidokowej drugiej generacji. W takich systemach, będących wciąż w początkowej fazie rozwoju, wykorzystuje się złożone metody prezentacji obrazu za pomocą wyświetlaczy autostereoskopowych. Wyświetlacze autostereoskopowe pozwalają na oglądanie scen trójwymiarowych bez zakładania specjalnych okularów.



Rysunek 1.1. Przykładowy sposób ustawienia kamer w systemie akwizycji sekwencji wielowidokowych z kamerami ustawionymi w linii.

Dużym wyzwaniem w systemach wizji wielowidokowej wykorzystujących wyświetlacze autostereoskopowe jest dostarczenie widoków w liczbie wystarczającej do poprawnego działania wyświetlaczy. Jakość wyświetlanego 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. Istotnym problemem badawczym jest zatem opracowanie metod efektywnej reprezentacji danych z systemów wielowidokowych. Opracowana reprezentacja powinna być możliwa do wykorzystania w systemach z wyświetlaczami autostereoskopowymi.

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*).

Przesyłanie obrazów z kamer w liczbie wystarczającej do poprawnego działania współcześnie wykorzystywanych wyświetlaczy autostereoskopowych wymaga jednak zbyt dużych prędkości bitowych. Wykorzystywane jest zatem inne rozwiązanie. Zamiast przesyłania kilkunastu lub kilkudziesięciu widoków, możliwe jest przesłanie jedynie kilku z nich wraz z odpowiadającymi tym widokom mapami głębi. Mapa głębi zawiera informacje o położeniu obiektów w scenie, a mówiąc ściślej – o ich odległości od kamery. Na podstawie obrazów z kamer wraz z przyporządkowanymi im mapami głębi możliwe jest wygenerowanie dodatkowych widoków syntetycznych, które mogą być wykorzystane do zapewnienia poprawnej pracy wyświetlacza. Wykorzystanie procesu syntezy widoków pozwala na znaczną redukcję liczby widoków koniecznych do przesłania. Obecnie rozważa się systemy, w których przesyłane są dwa lub trzy widoki wraz z mapami głębi. Na rysunku 1.2 przedstawiona jest przykładowa mapa głębi i odpowiadający jej obraz z sekwencji wielowidokowej.

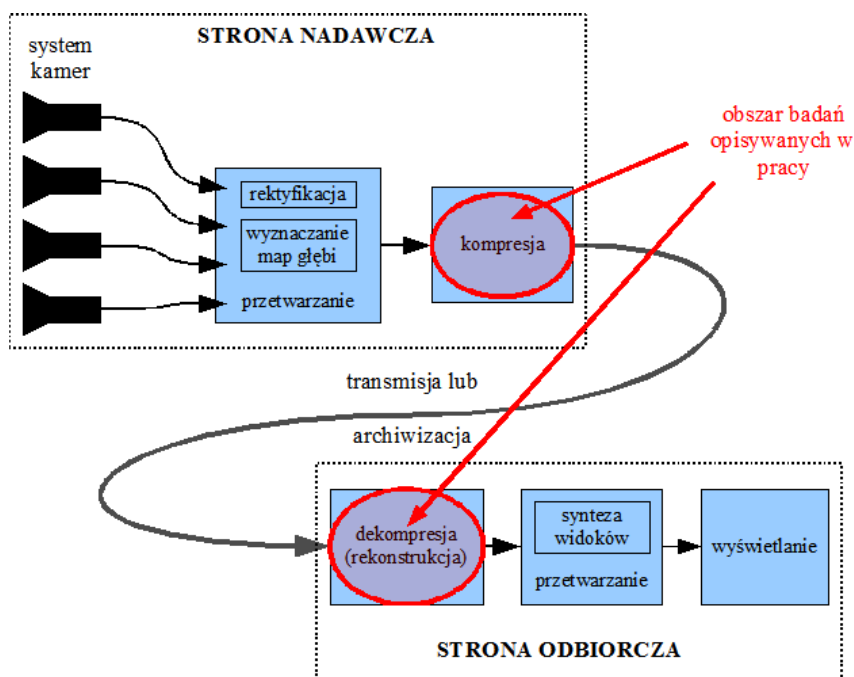


Rysunek 1.2. Przykładowy obraz (a) i mapa głębi (b) z sekwencji wielowidokowej.

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 ostatniego z wymienionych

zagadnień – wydajnych sposobów kompresji sekwencji wielowidokowych wraz z mapami głębi (patrz rysunek 1.3).



Rysunek 1.3. 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 możliwych zastosowań kodera przeznaczonego do kompresji sekwencji wielowidokowych (bez map głębi) do kompresji 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. 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, rozszerzający normę AVC o opis kodowania sekwencji wielowidokowych. W roku 2011 rozpoczęte zostały prace nad modyfikacją umożliwiającą kodowanie sekwencji wielowidokowych z mapami głębi (nazywanymi również sekwencjami trójwymiarowymi). W pracach tych, poza metodami opisywanymi w rozprawie, rozpatrywane są również nowe metody wykorzystujące złożone algorytmy predykcji, jak na przykład rzutowanie trójwymiarowe. Metody

predykcji tego typu nie są jednak przedmiotem prezentowanej rozprawy.

Badania prowadzone są na drodze eksperymentalnej, z wykorzystaniem dużego zestawu testowych sekwencji wielowidokowych z głębią, udostępnianych w ramach prac grupy MPEG (ang. *Moving Pictures Experts Group*). Dołożono starań, aby wykorzystywane w badaniach sekwencje były reprezentatywne dla ogółu dostępnych sekwencji, dzięki czemu uzyskane wyniki i wyciągane na podstawie wyników wnioski byłyby słuszne dla dowolnej sekwencji wielowidokowej z głębią.

### 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 metody kompresji wykorzystującej zależność pomiędzy sekwencjami wielowidokowymi i mapami głębi.* Opracowane techniki 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

Praca podzielona jest na rozdziały opisujące kolejne zagadnienia będące tematyką 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. Rozważone są w nim dwa sposoby pomiaru jakości widoków syntetycznych, różniące się pochodzeniem przyjętego w pomiarach obrazu odniesienia. Opisane są wady i zalety obydwu metod.

W **rozdziale 5** przedstawione są oryginalne 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. Najważniejsze wyniki prezentowane w pracy

### 2.1. Przetwarzanie wstępne sekwencji wielowidokowych

Istotny problem w przetwarzaniu sekwencji wielowidokowych stanowią niewielkie różnice pomiędzy egzemplarzami kamer, przejawiające się w zmienionych profilach kolorów dla obrazów z różnych kamer oraz niedokładności w geometrycznej orientacji kamer. Różnice te sprawiają, że zyski wynikające z predykcji międzywidokowej wykorzystywanej w koderze MVC są ograniczone. Jednym ze sposobów na zwiększenie efektywności kodowania sekwencji wielowidokowych jest przetwarzanie wstępne sekwencji podlegających kompresji.

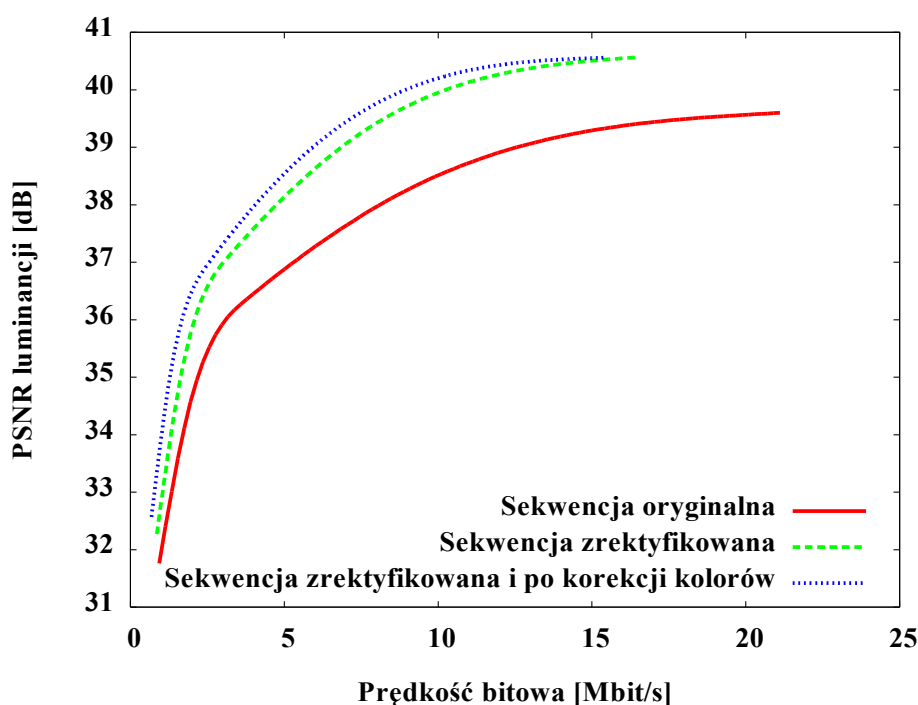
Efekty związane z różnym odwzorowaniem kolorów w kamerach można zredukować poprzez zastosowanie procedury korekcji kolorów. Polegać może ona, w najprostszej wersji, na przeprowadzeniu operacji wyrównywania histogramów dla wszystkich kamer w systemie wielowidokowym. Korekcja kolorów nie tylko pozwala poprawić efektywność kompresji, lecz również ułatwia przetwarzanie sekwencji w kolejnych etapach, jak na przykład synteza widoków.

Problemy wynikające z nieidealnej geometrii systemu wielokamerowego rozwiązuje się zwykle przez rektyfikację epipolarną obrazów z wielu kamer. Rektyfikacja epipolarna polega na zastosowaniu przekształcenia perspektywicznego do modyfikacji obrazów z kamer. Wynikiem przekształcenia są obrazy w postaci bardzo zbliżonej do tej, jaką można by uzyskać z systemu idealnie ustawionych kamer. Odpowiednie przekształcenie wyliczane jest na podstawie danych zebranych na etapie kalibracji systemu wielokamerowego. Więcej informacji na temat rektyfikacji można odnaleźć w artykule [18]. Rektyfikacja jest pomocna przy dalszym przetwarzaniu sekwencji wielowidokowych jak również znakomicie ułatwia i przyspiesza proces syntezy widoków pośrednich.

Zagadnieniem rozpatrywanym w badaniach była kwestia wpływu przetwarzania wstępnego (rektyfikacji i korekcji kolorów) na efektywność kompresji sekwencji wielowidokowych. Przeprowadzone zostały badania eksperymentalne z wykorzystaniem sekwencji poddanych kolejnym etapom przetwarzania wstępnego. Wpływ rektyfikacji i korekcji kolorów na efektywność kompresji w sekwencjach wielowidokowych ilustruje przykładowy wykres na rysunku 2.1, na którym pokazana została zależność jakości zdekodowanych widoków (liczonej jako średnia jakość ze wszystkich widoków) od całkowitej prędkości bitowej (dla wszystkich widoków). Wykresy dla



innych sekwencji testowych odnaleźć można w pracy. Prezentowany wykres utworzony został z wykorzystaniem wyników uzyskanych przez autora rozprawy na drodze badań eksperymentalnych. Przeprowadzenie eksperymentów było możliwe dzięki dostępowi do sekwencji poddawanych kolejnym etapom przetwarzania. Tego typu częściowo przetworzone sekwencje nie są zwykle udostępniane przez autorów sekwencji. Jednak autor rozprawy brał udział w tworzeniu zestawu sekwencji testowych, zgłoszonych i przyjętych do prac w grupie MPEG [10], dzięki czemu miał możliwość przeprowadzenia opisywanych tu eksperymentów. Podana na wykresach wartość PSNR (ang. *Peak Signal to Noise Ratio*) liczona jest przez porównanie składowej luminancji obrazów poddanych procesowi kompresji ze składową luminancji obrazów nieskompresowanych.



Rysunek 2.1. 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.

Uzyskane wyniki (przykładowe pokazane są na rysunku 2.1) świadczą o wyraźnym wzroście efektywności kodowania sekwencji wielowidokowych w wyniku przeprowadzenia rektyfikacji sekwencji. Zysk wynosi około 2 dB PSNR. Dla dolnego zakresu przebadanych prędkości bitowych przekłada się to na redukcję prędkości bitowej o około 30%. W prezentowanym na wykresie 2.1 przypadku, dla całego przebadanego zakresu prędkości bitowych obserwowany jest dodatkowy zysk rzędu 0,3 dB PSNR wynikający z przeprowadzenia korekcji kolorów. Zysk ten wyraźnie

jednak spada wraz ze wzrostem prędkości bitowej. Dla pozostałych przebadanych sekwencji również daje się zaobserwować niewielki zysk wynikający z przeprowadzenia korekcji kolorów. Szczegółowe wyniki dotyczące pozostałych sekwencji przedstawione są w pracy.

Na podstawie otrzymanych wyników można stwierdzić, że rektyfikacja epipolarna sekwencji wielowidokowych sprawia, że kodowanie tak przetworzonej sekwencji jest bardziej efektywne. Niewielką poprawę efektywności kodowania może dać również przeprowadzenie korekcji kolorów w sekwencji wielowidokowej. Niezależnie od wpływu na efektywność kompresji, przeprowadzenie tych dwóch operacji na etapie przetwarzania wstępnego znacznie ułatwia przeprowadzanie kolejnych etapów przetwarzania, czyli syntezy widoków.

## **2.2. Wpływ zniekształceń kodowania obrazów i głębi na jakość sekwencji synteżowanych**

Celem badań było ustalenie wpływu stopnia kompresji obrazów i map głębi na jakość synteżowanego widoku i przydatności zdekompresowanych danych w procesie syntezy widoków wirtualnych. Wnioskiem z badań miało być określenie, jak silnie można kompresować mapy głębi przy założonej, stałej jakości obrazów, aby nie doprowadzać do nadmiernego zmniejszenia jakości obrazów synteżowanych na podstawie rekonstruowanych danych. Kolejnym celem było przebadanie wpływu decymacji mapy głębi na jakość synteżowanego widoku i na efektywność kodowania mapy głębi. Ze względu na fakt, że nieznanym jest dokładny model teoretyczny zjawisk zachodzących w procesach wykorzystywanych w czasie przetwarzania sekwencji wielowidokowych z głębią, konieczne było przeprowadzenie badań eksperymentalnych.

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.

Jednym z podstawowych problemów w opisywanych badaniach jest podział prędkości bitowej strumienia na dwie części – przeznaczoną na transmisję danych o obrazie i drugiej, przeznaczoną na transmisję danych o mapach głębi.

Definiowany jest w tym celu współczynnik S:

$$S = \frac{P_D}{P_T + P_D} \quad (2.1)$$

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.

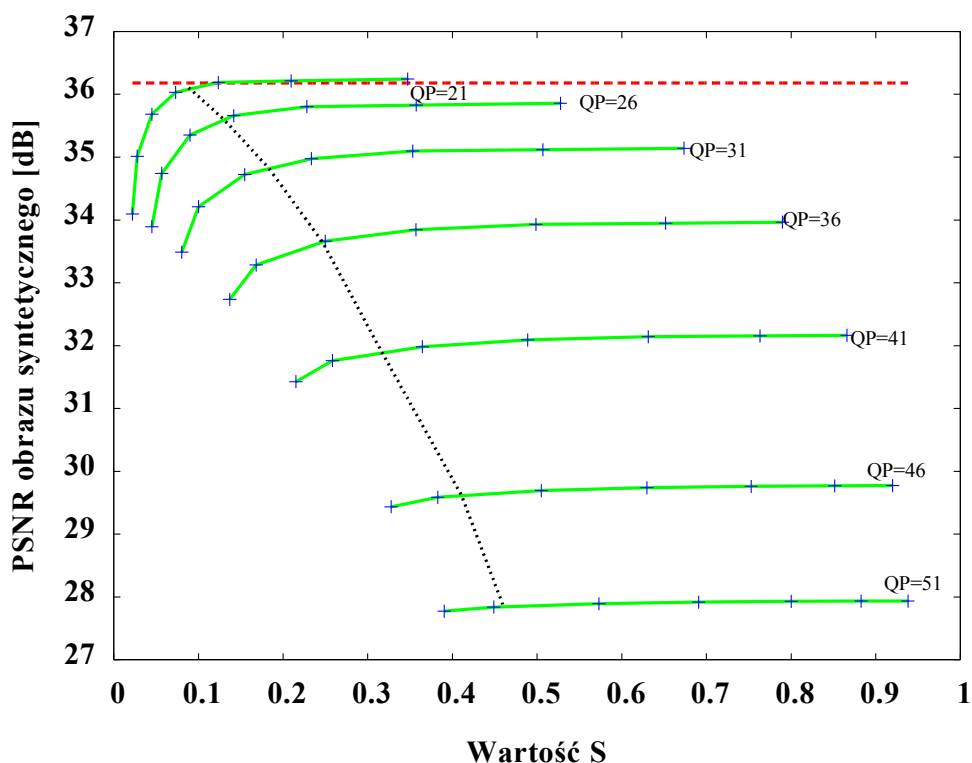
Na rysunku 2.2 zaprezentowany został przykładowy wykres jakości obrazu syntezowanego w zależności od wartości współczynnika S. Na wykresie, przy kolejnych krzywych, umieszczone są etykiety z wartością indeksu parametru kwantyzacji wykorzystanego przy kompresji obrazów (QP) dla danej krzywej. Kolejne punkty pomiarowe położone na takiej krzywej odpowiadają kombinacji tej wartości QP kolejno ze wszystkimi wartościami indeksu parametru kwantyzacji wykorzystanymi przy kompresji map głębi (QD). Czerwona przerywana linia wyznacza w tym wypadku poziom jakości uzyskiwany dla nieskompresowanych danych, natomiast czarna przerywana linia łączy punkty, dla których obserwowany jest spadek jakości o około 0,5 dB względem maksimum dla danej wartości QP. Przebieg tej linii wyznaczony jest z użyciem **autorskiego, opracowanego w trakcie badań wzoru (2.2)**, słusznego dla wszystkich przebadanych pod tym kątem sekwencji. Wzór wyznaczony został z wykorzystaniem regresji wielomianowej na podstawie wyników dla różnych sekwencji testowych. Wzór pozwala na określenie maksymalnej wartości współczynnika QD (a zatem minimalnej prędkości bitowej dla głębi), która nie powoduje znaczącego spadku jakości widoków syntezowanych. Za progową wartość spadku jakości przyjęto wartość 0,5 dB.

$$QD_{\max} = 49 - 0,005 \cdot (76 - QP)^2 \quad (2.2)$$

gdzie:

$QD_{\max}$  – maksymalna wartość indeksu QD nie powodująca znaczącego (ponad 0,5 dB) spadku jakości,

QP – wartość indeksu parametru kwantyzacji dla kodowania obrazu.



Rysunek 2.2. Zależność jakości widoku syntetycznego (wirtualnego) od wartości współczynnika S – stosunku prędkości bitowej danych o mapie głębi do całkowitej prędkości bitowej – dla sekwencji Book Arrival.

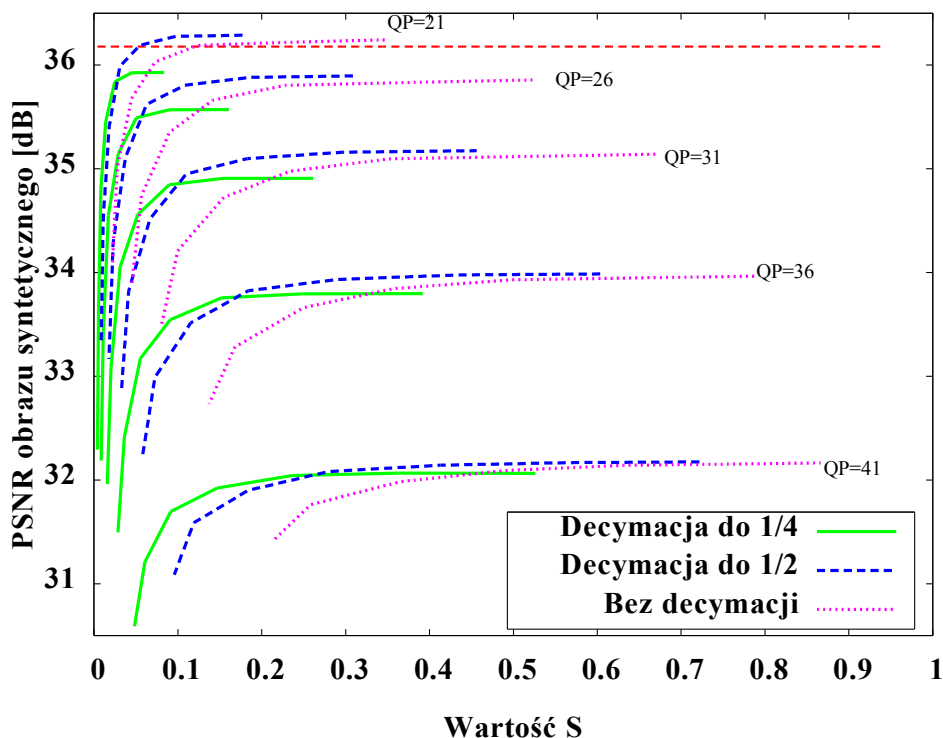
Na podstawie wyników przeprowadzonych badań (przykładowe wyniki pokazane są na rysunku 2.2) można stwierdzić, że jakość obrazu syntezowanego ze zdekompresowanych danych może, w pewnych przypadkach, przewyższać jakość obrazu syntezowanego z wykorzystaniem oryginalnych obrazów i map głębi. Zjawisko to wyjaśnione jest w pracy na podstawie obserwacji dotyczącej usuwania drobnych błędów w oryginalnej mapie głębi na drodze wygładzania map głębi. Z podobnym wygładzeniem wiąże się przeprowadzenie procesu kompresji stratnej z wykorzystaniem kodera MVC.

Na rysunku 2.3 przedstawione są z kolei przykładowe wyniki uzyskane z wykorzystaniem map głębi o zmniejszonej rozdzielczości.

Przedstawione wyniki pokazują, że zastosowanie decymacji map głębi do połowy wymiarów (zarówno wysokości jak i szerokości) – niebieskie linie na wykresie, pozwala zwiększyć efektywność kompresji przy zachowaniu tej samej jakości generowanego widoku wirtualnego. Dla tej samej prędkości bitowej, jakość uzyskana przy wykorzystaniu zdecymowanych 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 osiągnęte są wyższe maksymalne wartości miary jakości dla tych samych wartości QP w porównaniu z przypadkiem, gdy obraz mapy głębi nie jest poddawany decymacji. Wyjaśnienie zjawiska zwiększenia jakości dla zdecydowanych map głębi jest podobne jak dla efektu zwiększenia jakości w wyniku przeprowadzenia kompresji stratnej. Decymacja a następnie przywrócenie rozdzielczości pierwotnej obrazu mapy głębi prowadzi do wygładzenia, ujednolicenia wartości 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.

Wyniki uzyskane w przypadku wykorzystania mapy głębi zdecydowanej do jednej czwartej rozmiarów – zielone linie, są nieco inne niż dla decymacji do połowy rozmiarów. Nie obserwuje się poprawy jakości w całym zakresie prędkości bitowych, 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 mniejszego stosunku prędkości) dochodzi jednak do przecięcia krzywych jakości – 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 dla kompresji z parametrem QD większym od 36 lepszą jakość widoku syntezy osiąga się w przypadku wykorzystania mapy głębi z decymacją do  $\frac{1}{4}$  rozmiarów luminancji.



Rysunek 2.3. 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.

Wnioski opracowane na podstawie przedstawionych w pracy wyników można podsumować następująco.

1. W systemach wielowidokowych wykorzystujących do swego działania mapy głębi (a zatem przeprowadzające na jakimś etapie syntezę 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.
2. Na podstawie otrzymanych wyników opracowany został wzór (2.2) 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.
3. 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 2.3).

4. 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 dużych wartości indeksu QD ( $QD = 46$  i więcej), czyli dla silnie skompresowanych map głębi. Dla mniejszych wartości QD obserwowana jest strata jakości w porównaniu z przypadkiem wykorzystania mapy głębi w pełnej rozdzielczości.

### **2.3. Wybór optymalnych par QP-QD dla oddzielnego kodowania obrazów i map głębi**

Celem kolejnego etapu badań było opracowanie metody automatycznego doboru parametru QD przy zadanej wartości parametru QP. Dobór powinien odbywać się w sposób umożliwiający uzyskanie optymalnej jakości obrazów syntezy. Pod pojęciem „optymalna jakość” rozumiana jest w pracy maksymalna jakość uzyskiwana przy danej prędkości bitowej. Otrzymana w wyniku badań funkcja  $QD = f(QP)$  pozwoliłaby określić, jaką wartość powinien przyjmować indeks QD dla uzyskania maksymalnej jakości widoku syntezy przy danej prędkości bitowej. Przebadanie tego zagadnienia jest istotne dla prac nad kompresją sekwencji wielowidokowych z głębią, gdyż pozwala na automatyczne wyznaczanie odpowiednich wartości parametru QD w trakcie kodowania. Możliwe jest wtedy regulowanie prędkości bitowej wynikowego strumienia za pomocą pojedynczego parametru – parametru QP.

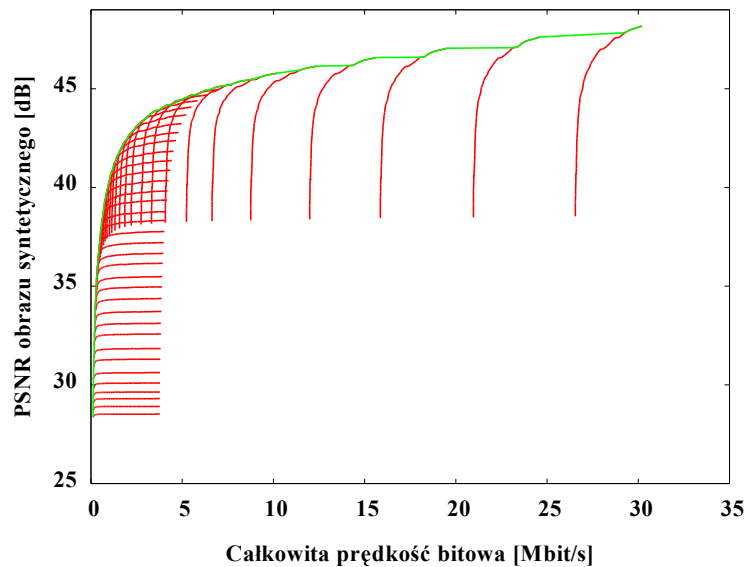
W toku badań należało zatem kolejno:

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

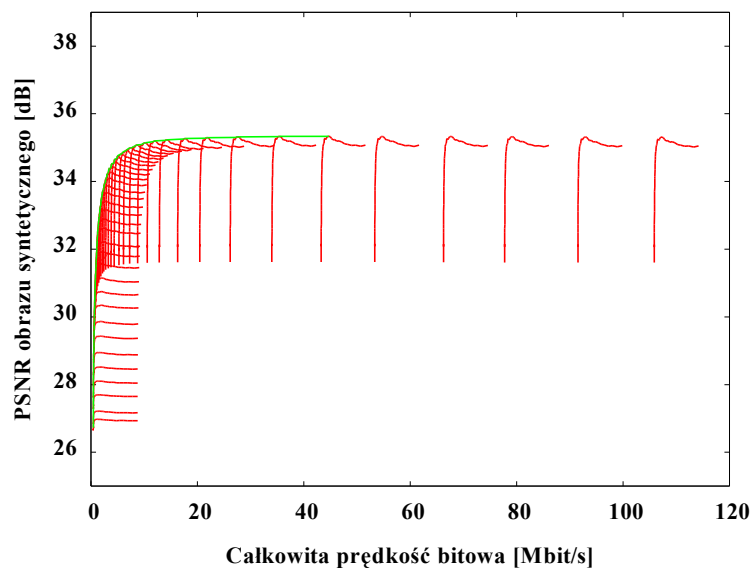
Uzyskane przykładowe wyniki przedstawione są na kolejnych rysunkach. Wyniki uzyskane

dla innych przypadków odnaleźć można w pracy.

Dla wszystkich sekwencji wykorzystanych w badaniach wyznaczone zostały następnie pary parametrów QP i QD, dla których jakość syntezowanych widoków osiągała wartość maksymalną dla danej prędkości bitowej. Wartości te odpowiadają punktom leżącym na zielonej linii z wykresów 2.4 oraz 2.5. Wyznaczone pary QP i QD zostały zaznaczone na rysunkach 2.6 oraz 2.7.

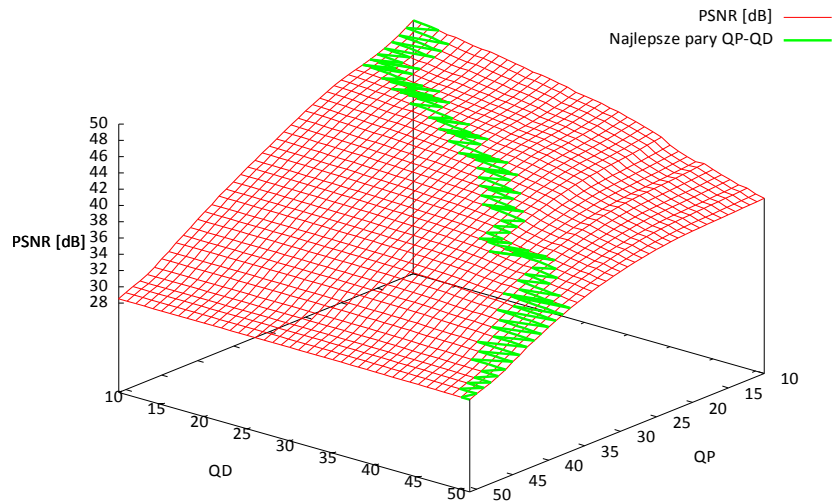


Rysunek 2.4. PSNR luminancji widoku syntezowanego w odniesieniu do widoku syntezowanego z użyciem nieskompresowanych danych dla sekwencji Book Arrival wraz z linią wyznaczającą maksymalną jakość dla danej prędkości bitowej.

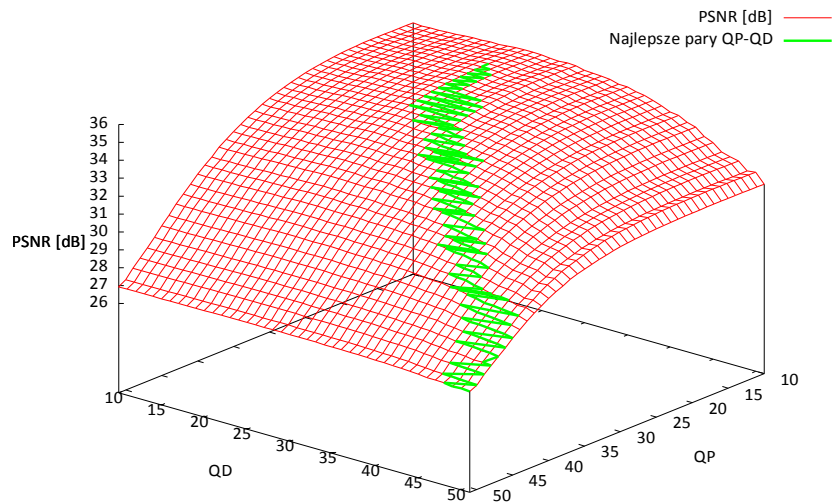


Rysunek 2.5. PSNR luminancji widoku syntezowanego w odniesieniu do rzeczywistego widoku z kamery dla sekwencji Poznan Street wraz z linią wyznaczającą maksymalną jakość dla danej prędkości bitowej.





Rysunek 2.6. Zależność jakości widoku wirtualnego od wartości indeksów QP i QD dla sekwencji Book Arrival (scenariusz 1). Zaznaczona linia łącząca pary optymalne. Jakość mierzona przez porównanie z widokiem wirtualnym wygenerowanym z nieskompresowanych danych.



Rysunek 2.7. Zależność jakości widoku wirtualnego od wartości indeksów QP i QD dla sekwencji Poznan Street (scenariusz 7). Zaznaczona linia łącząca optymalne pary. Jakość mierzona przez porównanie z widokiem z rzeczywistej kamery.

W celu wyznaczenia krzywych przybliżających zależność optymalnych indeksów QD od QP wyznaczone zostały, metodą regresji, funkcje drugiego stopnia przybliżające uzyskane w eksperymencie wartości dla par parametrów QP-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. Niebieskie punkty na wykresach 2.8, 2.9, 2.11 i 2.12 mają współrzędne odpowiadające wartościom parametrów QP i QD dających maksymalną jakość przy danej prędkości bitowej. Im większe nasycenie koloru, tym częściej w wynikach dla różnych rozważanych scenariuszy dana para QP-QD dawała najlepsze wyniki dla danej prędkości bitowej. Podane są również wartości współczynnika determinacji  $R^2$ , którego wartość świadczy o jakości przybliżenia danych pomiarowych przez krzywą regresji. Współczynnik ten wyznaczany jest z poniższego wzoru (2.3):

$$R^2 = \frac{\sum_i (\widehat{QD} - \overline{QD})^2}{\sum_i (QD_i - \overline{QD})^2}, \quad (2.3)$$

gdzie:

$i$  – liczba par QP-QD,

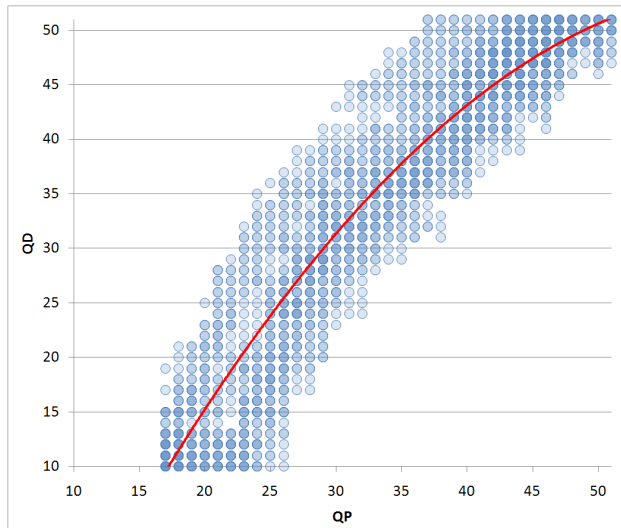
$\widehat{QD}$  - wartość QD wyliczona z pomocą wyznaczonego wzoru  $QD = f(QP)$ ,

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

Krzywe regresji wyznaczane są w rozpatrywanych przypadkach dla dwóch przedziałów. Podział na dwa przedziały umożliwia uzyskanie krzywych o zadowalającej wartości współczynnika  $R^2$  dla zakresu częściej wykorzystywanych – wyższych – wartości parametru QP. Wartości współczynnika  $R^2$  dla przedziału niższych wartości parametru QP są znacząco niższe, co świadczy o słabej jakości przybliżenia punktów pomiarowych przez wyznaczoną krzywą. Krzywe obliczone dla całych zakresów przebadanych wartości QP stanowiłyby wyraźnie gorsze przybliżenie badanej zależności niż krzywe wyznaczone dla 2 odrębnych przedziałów.

Dla obydwu rozpatrywanych sposobów mierzenia jakości obrazu syntezowanego granica przedziałów wyznaczana była oddzielnie.

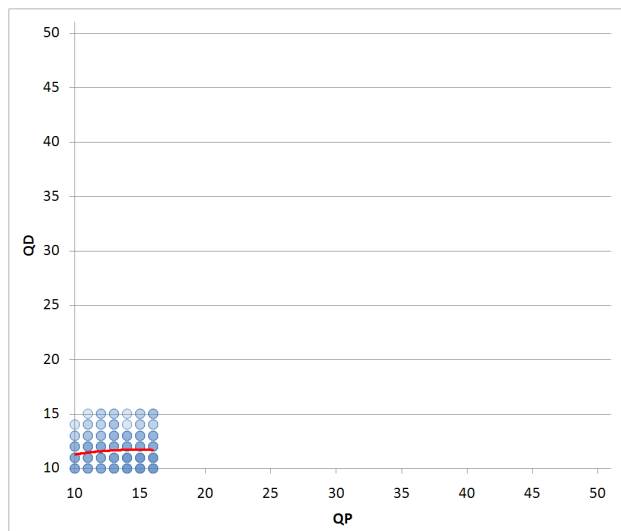
## Wyniki dla porównania z widokiem syntezowanym z nieskompresowanych danych



dla czerwonej linii:

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

Rysunek 2.8. Optymalne pary indeksów QP-QD dla wszystkich rozpatrywanych scenariuszy (dane łączne) w przedziale QP od 17 do 51. Dokładny opis w tekście.

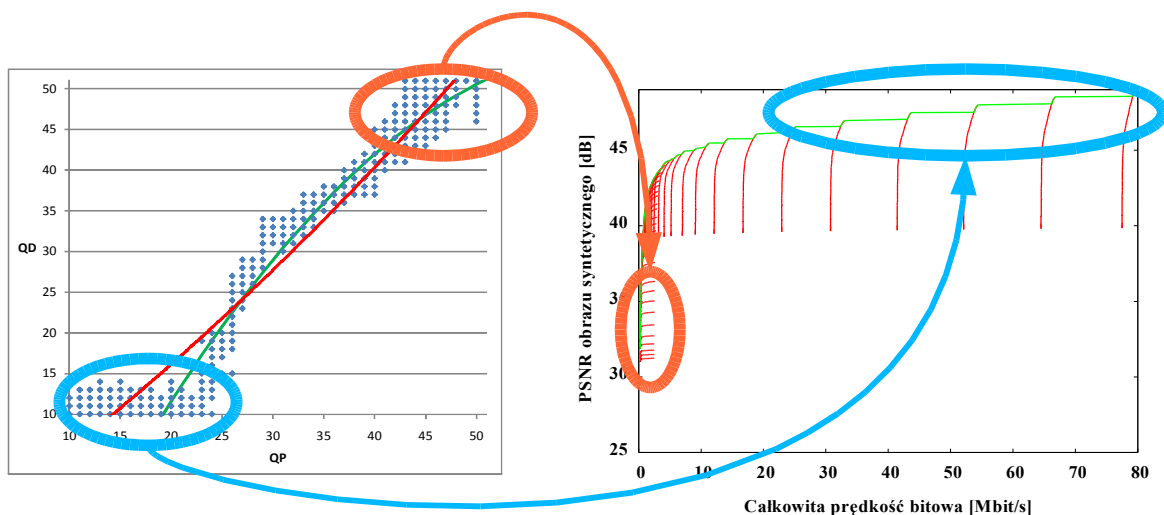


dla czerwonej linii:

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

Rysunek 2.9. Optymalne pary indeksów QP-QD dla wszystkich rozpatrywanych scenariuszy (dane łączne) w przedziale QP od 10 do 16. Dokładny opis w tekście.

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), co pokazane zostało to na rysunku 2.10.



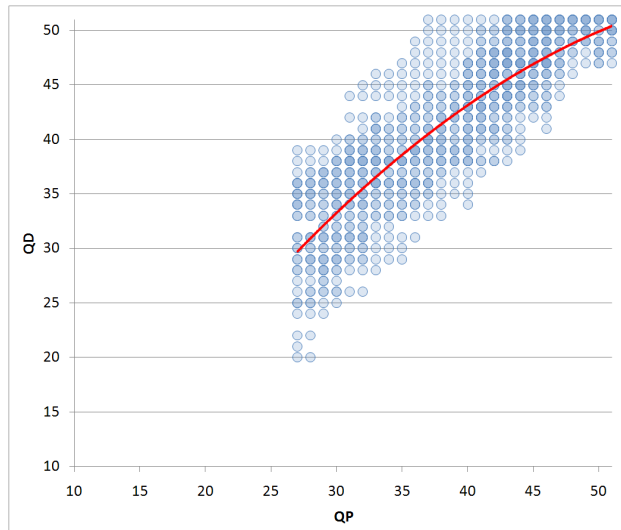
Rysunek 2.10. Położenie punktów odpowiadających tym samym optymalnym parom QP-QD na wykresach prezentowanych w pracy.

Jak widać na rysunku 2.9, dla wartości QP mniejszych niż 17, krzywa regresji przyjmuje wartości, które po zaokrągleniu w dół wynoszą dla całego tego zakresu 11. Uzyskana w tym przedziale wartość współczynnika determinacji  $R^2$  jest jednak bardzo mała i świadczy o tym, że wyznaczona zależność słabo przybliża punkty uzyskane w eksperymencie. Warto jednak podkreślić, że rozrzut wartości jest niewielki – występują jedynie wartości z przedziału od 10 do 15. 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 syntezy z nieskompresowanych danych **wartość QD w zależności od QP określona może być zatem poniższą zależnością, proponowaną przez autora pracy do wykorzystania w automatycznym doborze parametru QD:**

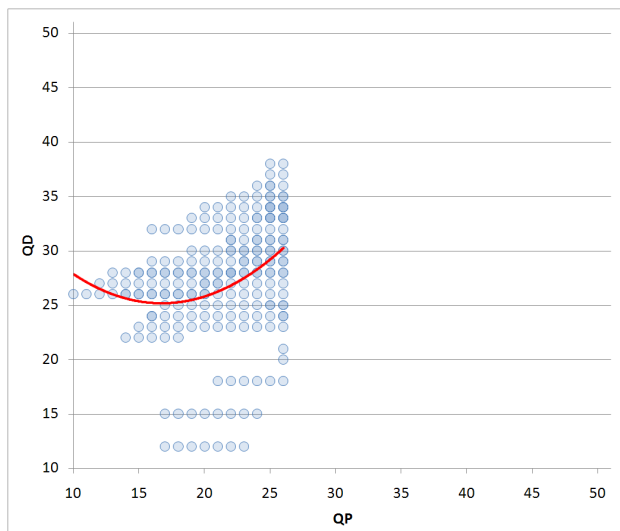
$$\text{QD} = \begin{cases} \lfloor -0,0216 \cdot \text{QP}^2 + 2,6872 \cdot \text{QP} - 29,876 + 0,5 \rfloor & \text{dla } \text{QP} > 16 \\ 11 & \text{dla } \text{QP} \leq 16 \end{cases} \quad (2.4)$$

## Wyniki dla porównania z widokiem z rzeczywistej kamery



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

Rysunek 2.11. Optymalne pary indeksów QP-QD dla wszystkich rozpatrywanych scenariuszy (dane łączne) w przedziale QP od 27 do 51. Dokładny opis w tekście.



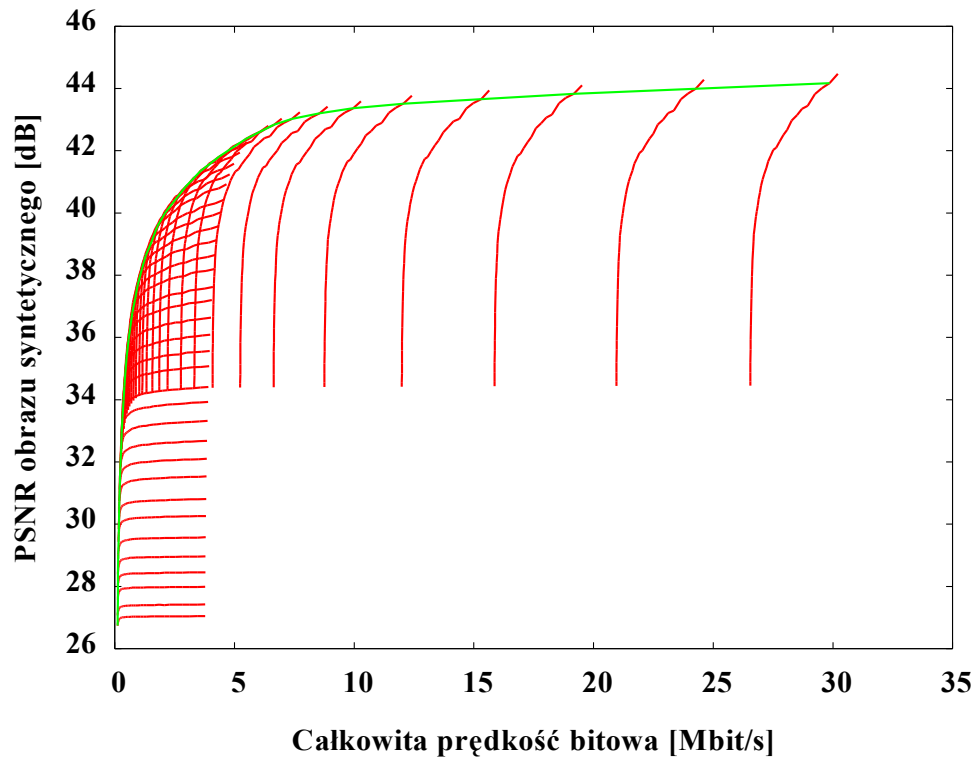
dla czerwonej linii:  
 $QD = 0,0585 \cdot QP^2 - 1,9579 \cdot QP + 41,562$   
 $R^2 = 0,1037$

Rysunek 2.12. Optymalne pary indeksów QP-QD dla wszystkich rozpatrywanych scenariuszy (dane łączne) w przedziale QP od 10 do 26. Dokładny opis w tekście.

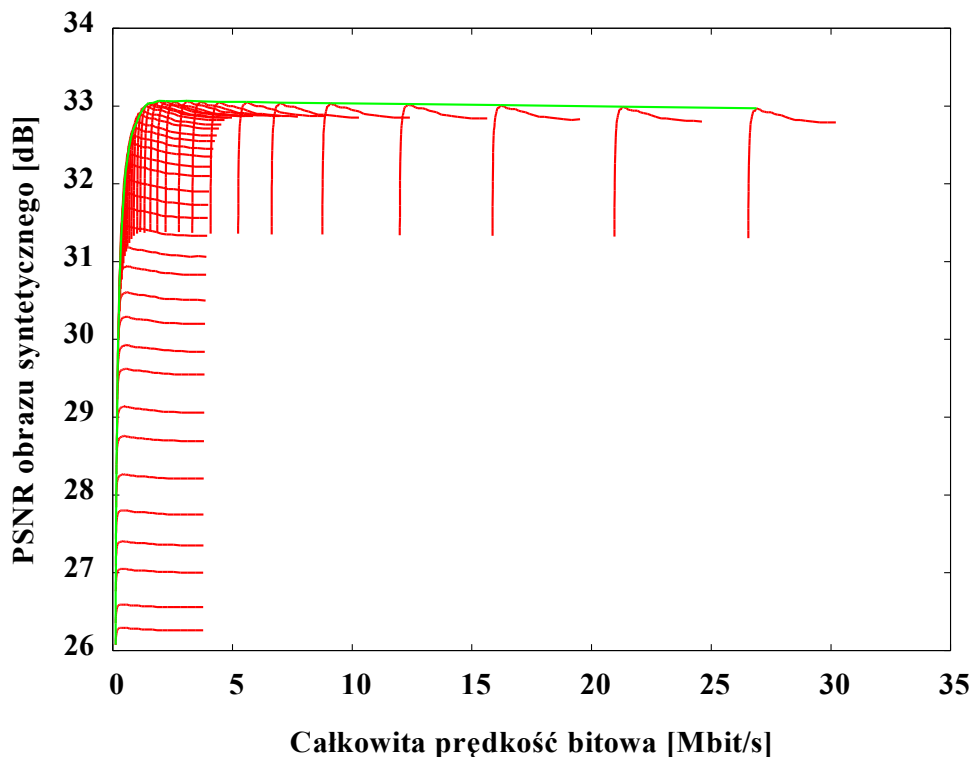
Jak widać na rysunku 2.12, dla wartości QP mniejszych niż 27, nie da się skutecznie przewidzieć wartości optymalnego indeksu QD na podstawie znajomości indeksu QP. Uzyskana w tym przypadku krzywa regresji bardzo słabo przybliża rozłożenie punktów pomiarowych (współczynnik  $R^2$  osiąga niską wartość). Ostatecznie zatem, biorąc pod uwagę duży rozrzut punktów, **wzór proponowanej do wykorzystania przez autora pracy zależności  $QD = f(QP)$  przyjmować może poniższą, uproszczoną postać:**

$$QD = -0,0155 \cdot QP^2 + 2,073 \cdot QP - 14,885 \quad . \quad (2.5)$$

Przedstawione wzory (2.4) oraz (2.5) zostały wykorzystane do określenia par QP-QD dla próbnej sekwencji, nie wchodzącej w skład zbioru sekwencji wykorzystanych w procesie opracowywania wzoru. Uzyskane wyniki świadczą o tym, że z pomocą wyznaczonych wzorów daje się uzyskać wyniki bliskie maksymalnym, co pokazane zostało na rysunkach 2.13 oraz 2.14.



Rysunek 2.13. Jakość widoku syntezywanego w odniesieniu do widoku syntezywanego z nieskompresowanych danych dla sekwencji Champagne Tower (czerwone linie) wraz z wynikami uzyskanymi przy doborze QD z wykorzystaniem proponowanego wzoru (2.4) (zielona linia).



Rysunek 2.14. Miara jakości widoku syntezowanego w odniesieniu do rzeczywistego widoku z kamery dla sekwencji Champagne Tower (czerwone linie) wraz z wynikami uzyskanymi przy doborze QD z wykorzystaniem proponowanego wzoru (2.5) (zielona linia).

Istotnym wnioskiem wynikającym z przeprowadzonych badań jest z pewnością obserwacja, że możliwe jest wyznaczenie przybliżonej zależności optymalnej wartości indeksu QD od wartości indeksu QP (czyli funkcji  $QD = f(QP)$ ).

Wzory (2.4) i (2.5) służą do wyznaczania wartości parametru QD na podstawie wartości parametru QP. W tym przypadku brana była pod uwagę najlepsza możliwa do uzyskania jakość widoku syntetycznego dla danej całkowitej prędkości bitowej. Wybór jednego z dwóch podanych wzorów dokonywany jest zależnie od przyjętego sposobu wyznaczania miary jakości PSNR.

Z kolei założeniem przy wyznaczaniu wzoru (2.2), różniącego się od wzoru (2.4), było uzyskanie krzywej uniwersalnej dla przebadanych sekwencji, która pozwala wyznaczyć pary parametrów QP-QD dające widok syntetyczny o jakości, która jest maksymalnie o 0,5 dB niższa od wartości maksymalnej dla danej wartości parametru QP.

Przedstawione w pracy wzory wyznaczone zostały na podstawie empirycznych badań na zestawie sekwencji wielowidokowych. Użyte sekwencje pochodzą ze zbioru wykorzystywanego w pracach grupy MPEG. Dołożono starań, aby wykorzystane sekwencje były reprezentatywne dla

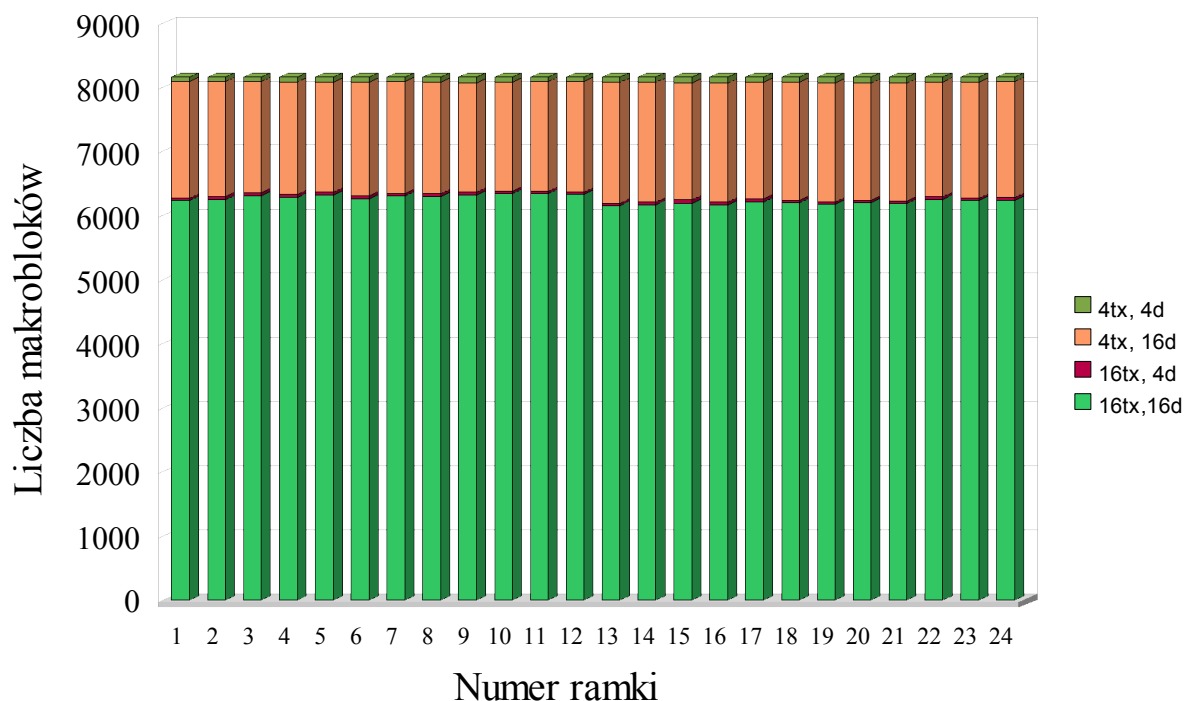
dostępnych obecnie sekwencji wielowidokowych. Zbiór przebadanych sekwencji zawiera zarówno sekwencje naturalne jak i sekwencję syntetyczną.

## 2.4. Łączne kodowanie obrazu i głębi

Analiza wyników przedstawionych w pracy doprowadziła do postawienia hipotezy, 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. Celem kolejnego etapu badań było zatem zweryfikowanie postawionej hipotezy oraz, w przypadku jej potwierdzenia, zaproponowanie metody efektywnego łącznego kodowania obrazów i odpowiadających im map głębi.

Kodowanie łączne pozwala 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. Przykładowo, zachodzi duże podobieństwo pomiędzy rozmiarami bloków wybranych w procesie kompresji obrazów i map głębi. Przedstawione jest to na przykładowym rysunku 2.15, na którym oznaczono wybrany rozmiar bloku w kompresji makrobloku obrazu jako 4tx (bloki o rozmiarach 4x4 punkty obrazu) oraz 16tx (blok o rozmiarach 16x16 punktów obrazu), natomiast dla głębi – odpowiednio 4d i 16d. Liczne wykresy dla innych sekwencji i innych parametrów kodowania prezentowane są w pracy.





Rysunek 2.15. Rozmiary bloków dla kodowania sekwencji Poznan Street.  
 Parametry kodowania: QP = 40, QD = 40. Podział na bloki 8x8 punktów wyłączony.

W badaniach nad łącznym kodowaniem obrazów i map głębi zaproponowana i przebadana została **oryginalna, autorska reprezentacja czterokomponentowa**, nie opisywana dotychczas w literaturze znanej autorowi rozprawy. Mapa głębi jest w tej reprezentacji przesyłana jako czwarta składowa D, przesyłana wraz z trzema składowymi obrazu: Y, C<sub>B</sub> oraz C<sub>R</sub>.

W celu zbadania efektywności kodowania sekwencji wielowidokowych w reprezentacji czterokomponentowej z głębią jako czwartym komponentem, przygotowane zostały modyfikacje referencyjnego oprogramowania do kodowania sekwencji wielowidokowych MVC w wersji 6. W wyniku wprowadzenia modyfikacji **powstał oryginalny, autorski kodek** (para koder i dekodek sekwencji) przeznaczony do kodowania sekwencji wielowidokowych z głębią w formie sekwencji czterokomponentowych. Wprowadzone zmiany były znaczne i polegały na dodaniu do koderów i dekodek możliwości obsługi sekwencji czterokomponentowych oraz na dodaniu procedur kodowania danych zawartych w dodatkowej składowej głębi. Wprowadzone zostały również modyfikacje w wykorzystywanych w koderze MVC tablicach kodowania, dostosowujące je do kompresji map 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 pracy 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.

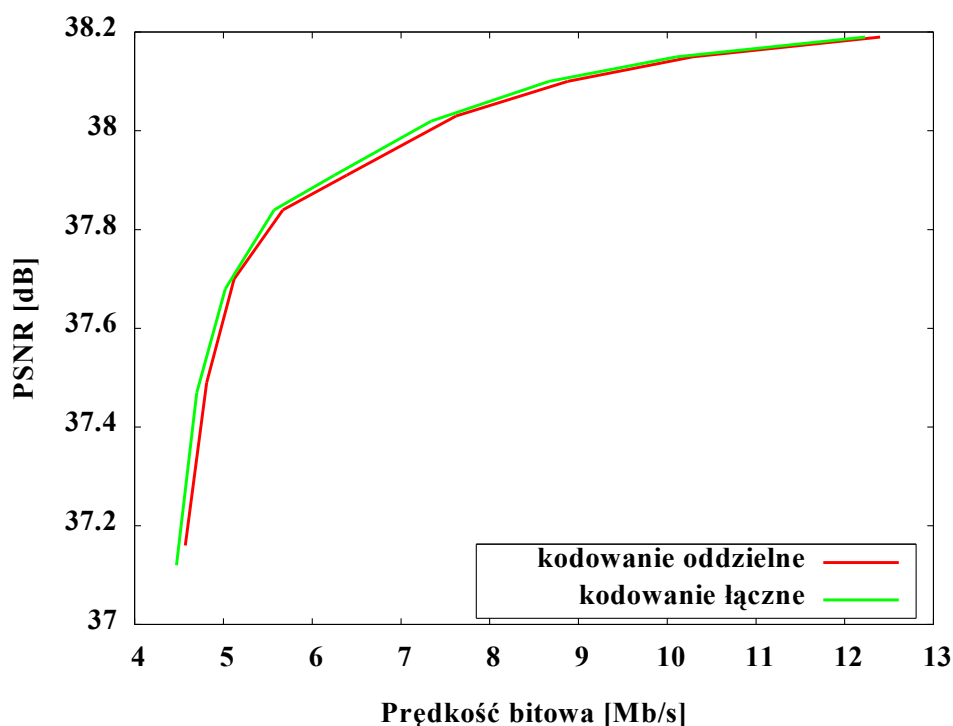
W badaniach rozważane było jedynie kodowanie wewnątrzobrazowe. Rozpatrzone zostały następujące sposoby kompresji sekwencji czterokomponentowych.

- Głębia jako składowa o zmniejszonej rozdzielczości, kompresowana wspólnie ze składowymi chrominancji. 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 niezmienionym 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. 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. W tym przypadku mapa głębi również była kodowana w oryginalnych rozmiarach.

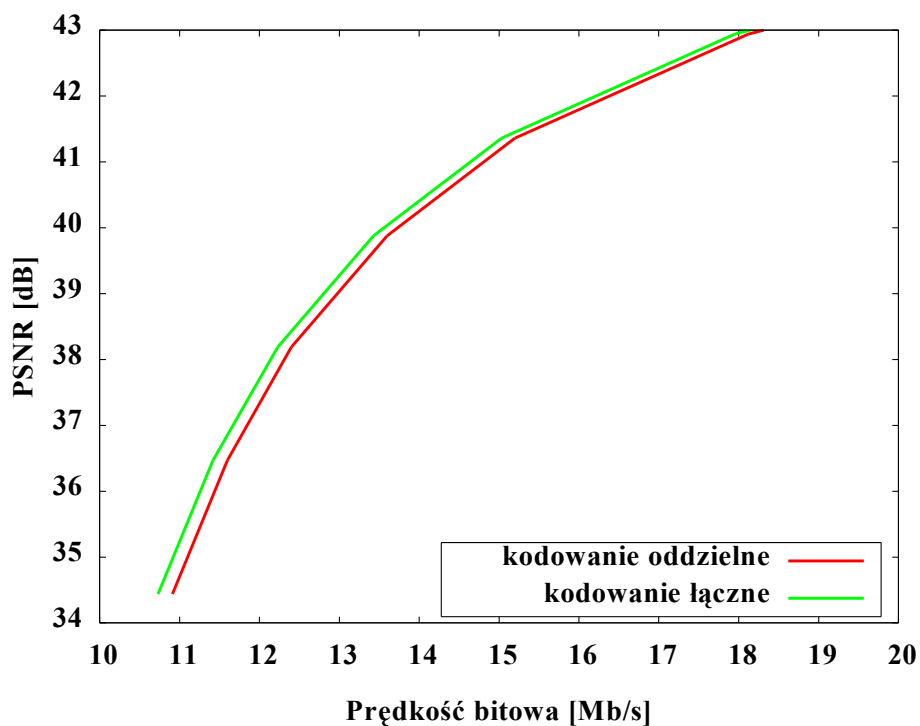
Kodowanie składowej głębi dla kolejnych makrobloków przebiegało na takich samych zasadach, jak kodowanie luminancji. Podział na bloki i kierunki predykcji w tych blokach wyznaczany był jednak indywidualnie dla mapy głębi, a jedynie kodowany w strumieniu wynikowym razem z tymi parametrami dla luminancji.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

Wyniki przeprowadzonych badań świadczą o tym, że najlepsze wyniki daje trzecia metoda, w której mapa głębi kodowana jest w pełnej rozdzielczości z możliwością wyboru innych trybów niż dla kodowania obrazu. Wyniki uzyskiwane z wykorzystaniem tej metody prezentowane są na przedstawianych dalej wykresach.

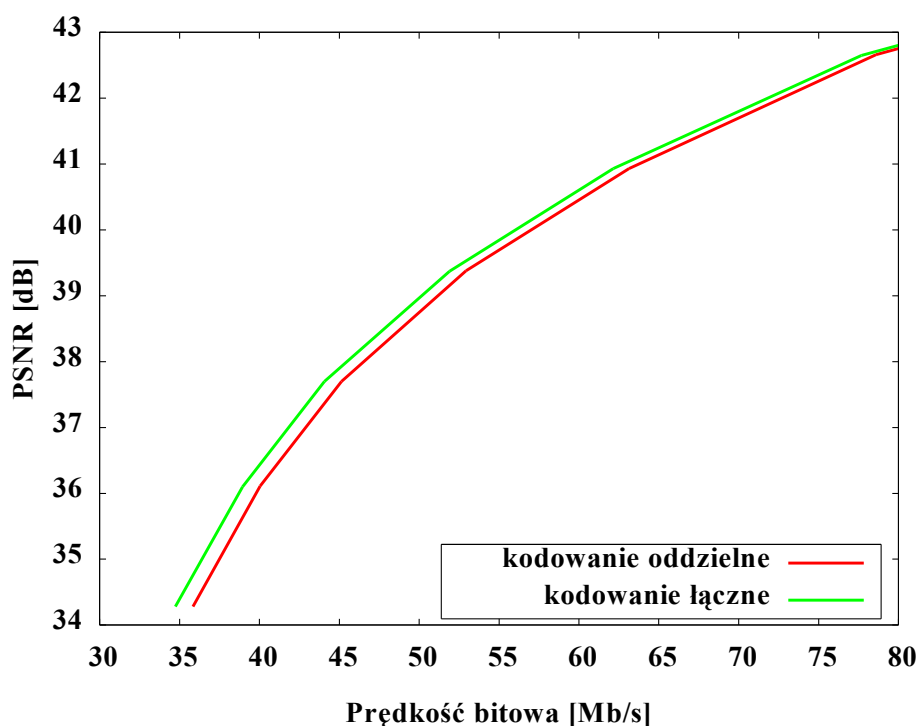
Przykładowe wyniki prezentowane są na rysunkach 2.16, 2.17 oraz 2.18. Więcej wykresów, dla innych sekwencji i innych ustawień kodera, znaleźć można w pracy.



Rysunek 2.16. 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 2.17. 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 2.18. 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.

Zyski wynikające z opracowanej metody sięgają mogą trzyprocentowej redukcji prędkości strumienia bitowego w porównaniu z niezależnym kodowaniem obrazów i map głębi. Warto też podkreślić, że zastosowanie metod kodowania wspólnego obrazów i map głębi wykorzystanych w stworzonym kodeku nie wiąże się ze wzrostem nakładów obliczeniowych.

Wadą metody jest występujące dla części przebadanych sekwencji zwiększenie prędkości bitowej w przypadku zastosowania małych wartości parametru kwantyzacji. Wyniki uzyskiwane z wykorzystaniem zmodyfikowanego kodera zależą zatem od własności sekwencji i jakości dostępnych map głębi.

### 3. 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 zaproponowanie, 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, rzędu 35 i więcej, 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.

## 4. 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;
  - Określono wpływ decymacji map głębi na efektywność procesu kodowania sekwencji wielowidokowych z mapami głębi;
  - Opracowano metodę doboru parametrów kodowania map głębi dających najlepsze efekty procesu syntezy widoków dla danych parametrów kodowania obrazu.

Przedstawione wyniki mają **duże znaczenie praktyczne**. Szczególnie użyteczne są wzory (2.4) oraz (2.5), 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;
  - przebadano zależności pomiędzy trybami kodowania współzależnych makrobloków dla kompresji obrazów i map głębi w trybie kodowania wewnątrzobrazowego;
  - zaproponowano i przebadano sposób scalenia obrazów i map głębi do postaci sekwencji czterokomponentowej;
  - zaproponowano modyfikacje kodera zwiększające efektywność kodowania map głębi.

Badania nad metodami wspólnej kompresji sekwencji wielowidokowych z głębią, wykorzystującymi zależności istniejące pomiędzy tymi składowymi mają obecnie duże znaczenie. Zaprezentowane wyniki mogą stanowić punkt wyjścia do opracowania metod kodowania tego typu sekwencji w ramach prowadzonych obecnie prac nad koderem



sekwencji trójwymiarowych.

Za dodatkowe osiągnięcia i cele opisane w rozprawie można uznać przebadanie wpływu 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.