

Politechnika Poznańska
Wydział Elektroniki i Telekomunikacji
Katedra Telekomunikacji Multimedialnej i Mikroelektroniki

Rozprawa doktorska

Modelowanie zaawansowanych koderów wizyjnych

Tomasz Grajek



Promotor: Prof. dr hab. inż. Marek Domański

Poznań, 2009

Spis treści

Spis treści	i
Streszczenie	vii
Abstract	viii
Spis ważniejszych skrótów i symboli	ix
Rozdział 1 Wprowadzenie	1
1.1. Modelowanie koderów wizyjnych.....	1
1.2. Cele i teza pracy.....	4
1.3. Metodologia prowadzenia badań	4
1.4. Przegląd pracy.....	8
Rozdział 2 Reprezentacja skompresowanych sygnałów wizyjnych	9
2.1. Wprowadzenie	9
2.2. Struktura strumienia bitowego na wyjściu kodera standardu MPEG-4 AVC/H.264	10
2.3. Warstwa kodowania obrazu – VCL	12
2.3.1. Makrobloki, plastry i obrazy	12
2.3.2. Predykcja wartości próbek obrazu	15
2.3.3. Kwantowanie.....	15
2.3.4. Zapis danych do strumienia bitowego.....	17
2.4. Podsumowanie	19
Rozdział 3 Przegląd rozwiązań w zakresie modelowania hybrydowych koderów wizyjnych	21
3.1. Wprowadzenie	21
3.2. Typowe modele wykorzystywane w implementacjach koderów standardu MPEG-4 AVC/H.264	22
3.3. Modyfikacje modeli z referencyjnej implementacji kodera MPEG-4 AVC/H.264	24
3.3.1. Model opisujący zależność między liczbą bitów a szerokością przedziałów kwantyzacji Q	24
3.3.2. Model aproksymujący wartość miary MAD	28
3.4. Model kodera sparymetryzowany procentem zerowych próbek transformaty	29
3.5. Model kodera w zależności od wartości parametru kwantyzacji QP	30
3.6. Podsumowanie	31

Rozdział 4 Modelowanie średniej liczby bitów potrzebnych do reprezentacji skompresowanego obrazu..... 33

4.1. Poszukiwany model	33
4.2. Wyznaczanie modelu na podstawie danych eksperymentalnych	36
4.3. Modelowanie kodera dla wartości Q z przedziału pierwszego (odcinek 1)	37
4.3.1. Wybór typu modelu.....	37
4.3.2. Model wielomianowy w przedziale pierwszym.....	39
4.3.3. Model potęgowy w przedziale pierwszym.....	51
4.3.4. Wnioski dotyczące modelu kodera dla wartości Q z przedziału pierwszego.....	55
4.4. Modelowanie kodera dla wartości Q z przedziału drugiego (odcinek 2)	56
4.4.1. Wybór typu modelu.....	56
4.4.2. Model hiperboliczny w przedziale drugim.....	59
4.4.3. Model potęgowy w przedziale drugim.....	65
4.4.4. Wnioski dotyczące modelu kodera dla wartości Q z przedziału drugiego.....	69
4.5. Modelowanie danych dla wartości Q z przedziału trzeciego (odcinek 3)	70
4.5.1. Model wielomianowy w przedziale trzecim	72
4.5.2. Model potęgowy w przedziale trzecim	84
4.5.3. Wnioski dotyczące modelu kodera dla wartości Q z przedziału trzeciego	87
4.6. Podsumowanie	88

Rozdział 5 Modelowanie liczby bitów reprezentacji pojedynczego obrazu. 91

5.1. Wprowadzenie	91
5.2. Modelowanie kodera dla wartości Q z przedziału pierwszego (odcinek 1)	92
5.2.1. Wykorzystywany model.....	92
5.2.2. Eksperymentalne wyznaczenie parametrów	93
5.3. Modelowanie kodera dla wartości Q z przedziału drugiego (odcinek 2)	97
5.4. Modelowanie kodera dla wartości Q z przedziału trzeciego (odcinek 3).....	98
5.5. Model z parametrem charakteryzującym złożoność treści obrazu	101
5.5.1. Miara charakterystyczna składowej luminancji obrazu	103
5.5.2. Miara charakterystyczna składowych luminancji i chrominancji	104
5.6. Podsumowanie	106

Rozdział 6 Modelowanie łącznej liczby bitów potrzebnych do reprezentacji składowych luminancji i chrominancji pojedynczego obrazu . 109

6.1. Wprowadzenie	109
6.2. Modelowanie kodera dla wartości Q z przedziału pierwszego (odcinek 1)	113
6.3. Modelowanie kodera dla wartości Q z przedziału drugiego (odcinek 2)	117
6.4. Modelowanie kodera dla wartości Q z przedziału trzeciego (odcinek 3).....	118
6.5. Model z parametrem charakteryzującym złożoność treści obrazu	121
6.6. Podsumowanie	123

<u>Rozdział 7 Podsumowanie</u>	125
7.1. Wnioski.....	125
7.2. Oryginalne osiągnięcia	126
Spis literatury	129
<u>Dodatek A Sekwencje testowe</u>	141
<u>Dodatek B Prędkości bitowe zakodowanych sekwencji wizyjnych</u>	149

Podziękowania

Chciałbym szczególnie podziękować mojemu promotorowi, prof. dr. hab. inż. Markowi Domańskiemu za poświęcony mi czas i cenne wskazówki, które pozwoliły na ukończenie tej rozprawy.

Dziękuję również wszystkim Koleżankom i Kolegom z Katedry Telekomunikacji Multimedialnej i Mikroelektroniki Politechniki Poznańskiej za wsparcie i pomocne uwagi.

Jestem wdzięczny moim rodzicom za ich bezwarunkową miłość i wsparcie we wszystkim co robię.

Na koniec, w sposób szczególny, dziękuję mojej żonie Justynie za cierpliwość i serce.

Streszczenie

Praca dotyczy modelowania zaawansowanych koderów wizyjnych, a w szczególności koderów wytwarzających strumienie binarne zgodne ze standardem MPEG-4 AVC/H.264, który jest coraz szerzej stosowany w systemach multimedialnych. W szczególności przedstawione w rozprawie modelowanie dotyczy koderów, których sprawność kompresji jest równa sprawności kompresji referencyjnej implementacji kodera wizyjnego standardu MPEG-4 AVC/H.264.

W rozdziale 2 przedstawiono sposób reprezentacji skompresowanych sekwencji wizyjnych na wyjściu kodera MPEG-4 AVC/H.264 zgodnego z profilem głównym.

W rozdziale 3 dokonano przeglądu literatury dotyczącej modelowania hybrydowych koderów wizyjnych ze szczególnym uwzględnieniem koderów wykorzystujących zaawansowane techniki kompresji.

Rozdziały 4 do 6 zawierają autorskie propozycje ilościowych modeli kodera wizyjnego MPEG-4 AVC/H.264. Zaproponowane modele mogą być wykorzystywane do ustalania szerokości przedziałów kwantyzacji Q dla zadanej liczby bitów w obrazach typu I, P i B. Ponadto, dokonano analizy dokładności zaproponowanych modeli dla wielu reprezentatywnych sekwencji wizyjnych o standardowej rozdzielczości telewizyjnej.

Zaproponowane modele wyrażają zależności między szerokością przedziałów kwantyzacji Q a:

- średnią liczbą bitów potrzebnych do zakodowania obrazu danego typu w obrębie jednego ujęcia,
- liczbą bitów potrzebnych do zakodowania pojedynczego obrazu danego typu,
- liczbą bitów potrzebnych do zakodowania skwantowanych wartości próbek transformaty błędu predykcji składowych luminancji i chrominancji obrazu danego typu.

Z przeprowadzonych badań wynika, że dokładność proponowanych modeli jest wystarczająca dla praktycznych zastosowań.

Abstract

Dissertation is dealing with modeling of advanced video encoders, mainly video encoders producing bitstreams that are compliant with MPEG-4 AVC/H.264 standard. This standard has growing applications in multimedia systems. In particular, the proposed modeling refers to encoders with compression performance being the same as that of the reference implementation of the video encoder of the standard MPEG-4 AVC/H.264.

Chapter 2 describes representations of compressed video sequences at the output of an MPEG-4 AVC/H.264 video encoder working in main profile.

Chapter 3 reviews modeling of hybrid video encoders with special emphasis on advanced video encoding techniques.

Chapters 4 to 6 include quantitative MPEG-4 AVC/H.264 encoder models proposed by the author. They can be used to set a value of the quantization step size Q for a given number of bits for an I- P- or B-frame. Moreover, accuracy analysis of the proposed models has been made based on many representative video sequences with standard television resolution.

Proposed models define the relationship between the quantization step size Q and:

- the average number of bits needed to encode frame of a given type within a single shot,
- the number of bits needed to encode a single frame of a given type,
- the number of bits needed to encode the quantized transformed prediction error coefficients of the luminance and chrominance.

The results lead to the conclusion, that the accuracy of the proposed models is sufficient for practical applications.

Spis ważniejszych skrótów i symboli

4CIF	4 x Common Intermediate Format	format sekwencji wizyjnej: rozdzielczość przestrzenna - 704x576 punktów; wybieranie kolejnoliniowe; schemat próbkowania - 4:2:0
AVC	Advanced Video Coding	standard kodowania sekwencji wizyjnych ISO/IEC MPEG-4 AVC oraz ITU-T H.264
AVS	Audio Video Standard	standard kodowania sekwencji wizyjnych (Chiny)
CABAC	Context-based Adaptive Binary Arithmetic Coding	kontekstowe, adaptacyjne kodowanie arytmetyczne
CAVLC	Context-based Adaptive Variable Length Coding	kontekstowe, adaptacyjne kodowanie o zmiennej długości słowa
CBR	Constant Bit Rate	stała prędkość bitowa (przepływność)
CIF	Common Intermediate Format	format sekwencji wizyjnej: rozdzielczość przestrzenna - 352x288 punktów; wybieranie kolejnoliniowe; schemat próbkowania - 4:2:0
DCT	Discrete Cosine Transform	dyskretna transformacja kosinusowa
FMO	Flexible Macroblock Ordering	konfigurowalna kolejność kodowania makrobloków w MPEG-4 AVC/H.264
FRExt	Fidelity Range Extensions	grupa profili w MPEG-4 AVC/H.264 o rozszerzonym zakresie dokładności
GOP	Group Of Pictures	grupa obrazów
IEC	International Electrotechnical Commission	Międzynarodowa Komisja Elektrotechniczna
ISO	International Organization for Standardization	Międzynarodowa Organizacja Standaryzacyjna (Normalizacyjna)
ITU-T	International Telecommunication Union Telecommunication Standardization Sector	Sektor Standaryzacji (Normalizacji) Telekomunikacji Międzynarodowej Unii Telekomunikacyjnej
JM	Joint Model	nazwa oprogramowania referencyjnego standardu MPEG-4 AVC/H.264

MAD	Mean Absolute Difference	średnia bezwzględnych wartości różnic
MPEG	Motion/Moving Pictures Experts Group	grupa robocza ISO/IEC zajmująca się rozwojem standardów kodowania dźwięku i obrazu zatwierdzona przez ISO grupa formatów zapisu danych zawierających obraz i dźwięk
NAL	Network Abstraction Layer	warstwa wyodrębniania sieci
Q	Quantization step size	szerokość przedziałów kwantyzacji
QP	Quantization Parameter	parametr kwantyzacji
RDO	Rate Distortion Optimization	optymalizacja „prędkość bitowa – zniekształcenie”
ROI	Region of Interest	region zainteresowania
SAD	Sum of Absolute Differences	suma bezwzględnych wartości różnic
SATD	Sum of Absolute Transformed Differences	suma bezwzględnych wartości próbek transformaty różnic
SEMPTE	Society of Motion Picture and Television Engineers	organizacja zajmująca się przygotowaniem standardów związanych z produkcją telewizyjną
VBR	Variable Bit Rate	zmienna prędkość bitowa (przepływność)
VC-1	Video Coding – 1	standard kodowania sekwencji wizyjnych opracowany przez Microsoft i formalnie przyjęty przez SMPTE jako SMPTE 421M
VCEG	Video Coding Experts Group	grupa robocza ITU-T zajmująca się rozwojem standardów kodowania dźwięku i obrazu
VCL	Video Coding Layer	warstwa kodowania obrazu
WMV	Windows Media Video	format plików dla skompresowanych sekwencji wizyjnych
$B_{\text{śred}}$	średnia liczba bitów reprezentacji obrazu danego typu w ramach pojedynczego ujęcia	
B_{obraz}	liczba bitów reprezentacji pojedynczego obrazu danego typu	
B_{YCbCr}	sumaryczna liczba bitów reprezentacji składowych luminancji i chrominancji obrazu danego typu	
$\varepsilon(x_i)$	względny błąd aproksymacji	

N_Y, N_{Cb}, N_{Cr}	liczba nienakładających się bloków o rozmiarze 4x4 punkty składowej odpowiednio Y, Cb i Cr
p.p.	punkty procentowe (różnica między dwiema wartościami podanymi w procentach)
$\sigma_i^Y, \sigma_i^{Cb}, \sigma_i^{Cr}$	odchylenie standardowe wartości próbek składowej odpowiednio Y, Cb i Cr obrazu w i -tym bloku o rozmiarze 4x4 punkty
$\overline{\sigma}_Y$	odchylenie standardowe wartości próbek składowej luminancji obrazu wyznaczone w blokach o rozmiarze 4x4 punkty uśrednione po wszystkich blokach w obrazie
$\overline{\sigma}_{YCbCr}$	suma uśrednionych odchyleń standardowych wartości próbek składowych Y, Cb i Cr obrazu wyznaczonych w blokach o rozmiarze 4x4 punkty

Rozdział 1

Wprowadzenie

1.1. Modelowanie koderów wizyjnych

Rozprawa dotyczy zaawansowanych technik kompresji obrazów ruchomych (sekwencji wizyjnych). Przez zaawansowane techniki kompresji sekwencji wizyjnych rozumie się najnowsze techniki rozwinięte po roku 2000, zwłaszcza techniki ujęte w najnowszych standardach takich jak MPEG-4 AVC/H.264 [AVC, Marp05, Oste04, Wieg03b], AVS [AVS] i VC-1 [VC1] (ten ostatni jest związany z powszechnie znaną techniką Windows Media Video (WMV) 9.0 [Lian04]). Techniki te wykorzystują nowe, często skomplikowane narzędzia kodowania, przez co osiągają lepszą sprawność kompresji w porównaniu z dotychczas powszechnie stosowanymi algorytmami (MPEG-1, MPEG-2, H.263) [Oste04, Ohm04, Raja04, Wieg03a]. Stosuje się je w systemach telekomunikacyjnych i teleinformatycznych nowych generacji, w telewizji cyfrowej, wideotelefonii, cyfrowych systemach nadzoru wizyjnego, transmisji obrazu w sieciach komputerowych itd.

Zaawansowane algorytmy kompresji sekwencji wizyjnych są bardzo złożone. Na przykład napisany w języku C program realizujący koder MPEG-4 AVC/H.264 w głównym profilu ma około 58 tysięcy linii kodu. Zaawansowane algorytmy kompresji charakteryzują się bardzo silnymi mechanizmami adaptacji algorytmu i jego parametrów w zależności od treści kodowanej sekwencji. W szczególności szeroko stosuje się adaptacyjny wybór trybów kodowania.

W koderach wizyjnych ilościowe zależności między jakością obrazu, szybkością bitową strumienia danych i parametrami sterującymi kodera zawsze silnie zależą od treści kodowanej

sekwencji (od charakteru i ilości ruchu, liczby szczegółów itd.). Wspomniana bardzo silna rola adaptacji w koderach zaawansowanych czyni ten związek jeszcze bardziej skomplikowanym.

Zaawansowane kodery sekwencji wizyjnych są regulowane przez wiele parametrów, z których część powinna być automatycznie wyznaczana przez procedurę sterowania. Algorytm sterowania dobiera parametry kodera tak, aby uzyskać maksymalną jakość kodowanej sekwencji wizyjnej przy założonej prędkości bitowej (tryb stałej prędkości bitowej – CBR ang. *constant bit rate*). W innym scenariuszu algorytm sterowania dąży do utrzymania stałej jakości i minimalizacji ogólnej liczby bitów reprezentujących sekwencję wizyjną (tryb stałej jakości przy zmiennej prędkości bitowej – VBR ang. *variable bit rate*). Ponadto, opisane wyżej procesy muszą uwzględniać wielkość dopuszczalnego opóźnienia.

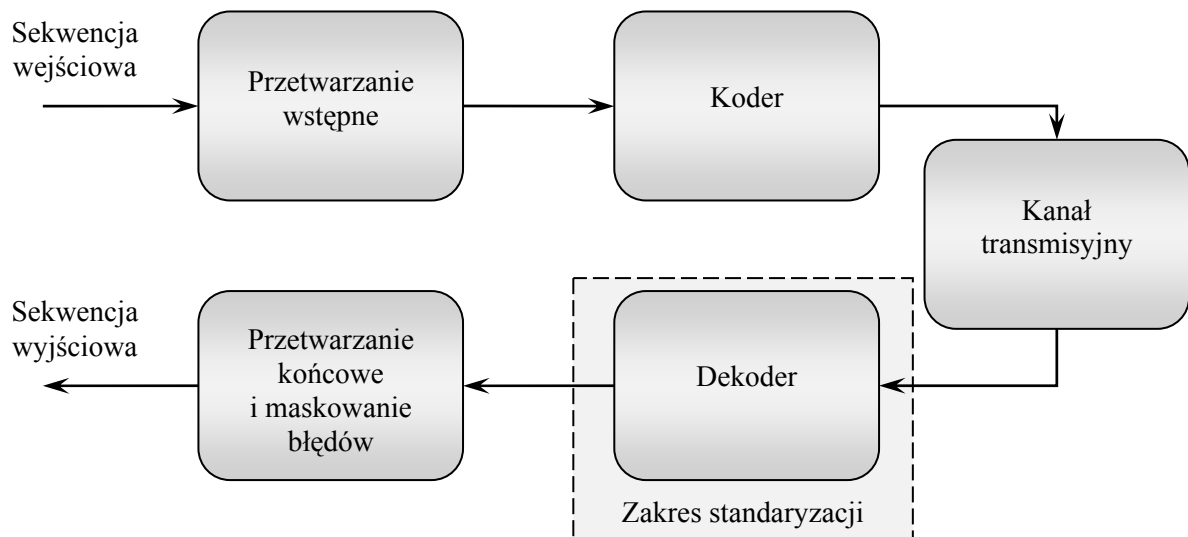
Znalezienie optymalnych wartości parametrów wymaga dużych nakładów obliczeniowych. Dlatego w praktycznych implementacjach stosuje się różnego rodzaju uproszczenia. Jednym z możliwych rozwiązań jest wykorzystanie modelu opisującego związku pomiędzy parametrami takiego kodera, prędkością bitową strumienia danych na wyjściu kodera, a parametrami treści kodowanej sekwencji wizyjnej. Wykorzystanie modelu pozwala na szybkie wyznaczenie parametrów sterujących w odpowiedzi na gwałtowną zmianę warunków transmisyjnych. W klasycznych zastosowaniach kompresji sygnałów wizyjnych szybkość transmisji była najczęściej stała lub rzadko modyfikowana (rozsiewcza telewizja cyfrowa naziemna, kablowa i satelitarna). Obecnie, w systemach internetowych i przy wykorzystaniu sieci bezprzewodowych i mobilnych, sytuacja się zmieniła. Kodery wizyjne muszą szybko reagować na zmiany w sieci i w kanałach transmisyjnych. Dlatego rośnie znaczenie praktyczne dobrych modeli koderów.

Sformułowanie choćby przybliżonych zależności ilościowych pomiędzy prędkością bitową danych na wyjściu kodera, parametrami sterującymi kodera i parametrami treści sekwencji ma olbrzymie znaczenie poznawcze. Posługując się takimi modelami można zbadać zachowanie się kodera w różnych stanach. Dla zaawansowanych technik kompresji dotychczas opracowane modele [Lee08a, Li03, Xie07, Yuan06] nie są wystarczające. Brakuje modeli analogicznych do tych, które zostały wcześniej opracowane dla starszych technik kompresji sygnałów wizyjnych [Bai01, Chia97, Ding96, He01d, Lin98, Łucz03b, Hang97].

Ponieważ zaawansowane kodery hybrydowe są bardzo złożone, to zbiór parametrów sterujących koderem jest duży. Większość tych parametrów dobiera się w celu uzyskania jak najlepszej relacji między jakością zakodowanej sekwencji wizyjnej a założoną prędkością bitową. Natomiast jedynym parametrem pozwalającym na zmianę prędkości bitowej

strumienia danych na wyjściu kodera w bardzo szerokim zakresie jest Q - parametr określający szerokość przedziałów kwantyzacji. Dlatego do modelowania hybrydowych koderów wizyjnych, wykorzystuje się głównie ten parametr Q [Bai01, Chia97, Doma02, Dong08, Kwon07, Lee08a, Li03, Li04, Lim07, Łucz02, Łucz03b, Ma05, Wieg03a, Xie07, Yuan06].

Międzynarodowe standardy kompresji sygnałów wizyjnych (np. MPEG-2, MPEG-4 (część 2), H.263, MPEG-4 AVC/H.264) definiują tylko proces dekodowania wraz z semantyką i składnią strumienia bitowego (rysunek 1.1). Tym samym dają dużą swobodę konstruktorom sprzętu i oprogramowania, którzy mogą stosunkowo dowolnie optymalizować kodery ze względu na ich szybkość działania i sprawność kompresji [Oste04, Rich03]. Swoboda projektowania dotyczy także procedur sterowania koderami. Takie procedury nie są objęte standardami i ciągle trwają badania dotyczące ich ulepszenia.



Rysunek 1.1. Zakres standaryzacji koderów wizyjnych (opracowanie własne na podstawie [Oste04])

Procedury sterowania często wykorzystują modele koderów. Modele są potrzebne także dla opisu i zrozumienia sposobu działania koderów. Modele związane są z poszczególnymi odmianami koderów.

Aby zaproponowane w rozprawie modele miały uniwersalny charakter i mogły być w prosty sposób adaptowane do różnych zaawansowanych technik kompresji, autor do badań wykorzystał oprogramowanie referencyjne standardu MPEG-4 AVC/H.264 w wersji JM 13.2 [AVCref]. Jest to oprogramowanie ogólnodostępne, w którym zaimplementowano wszystkie

narzędzia przewidziane przez standard. Ponadto, jest ono wykorzystywane jako odniesienie dla innych zaawansowanych kodeków wizyjnych w badaniach naukowych i testach.

1.2. Cele i teza pracy

Celem pracy jest zaproponowanie ilościowych modeli matematycznych koderów standardu MPEG-4 AVC/H.264, których sprawność kompresji reprezentowana zależnościami „prędkość bitowa – jakość obrazu zdekodowanego” jest w przybliżeniu równa sprawności implementacji referencyjnej koder standardu MPEG-4 AVC/H.264. Modele te mają opisywać związki pomiędzy prędkością bitową strumienia danych na wyjściu koder a szerokością przedziałów kwantyzacji Q .

Ponadto, celem pracy jest analiza dokładności modelowania koder standardu MPEG-4 AVC/H.264 za pomocą zaproponowanych modeli.

Teza pracy:

Możliwe jest opracowanie matematycznych modeli zaawansowanych koderów wizyjnych pozwalających na wyznaczenie liczby bitów generowanych przez koder w zależności od szerokości przedziałów kwantyzacji Q .

Modele zostaną zaproponowane na podstawie analizy danych eksperymentalnych uzyskanych z wykorzystaniem referencyjnej implementacji kodeka standardu MPEG-4 AVC/H.264 dla szerokiej klasy sekwencji wizyjnych o standardowej rozdzielczości telewizyjnej.

1.3. Metodologia prowadzenia badań

Proponowane modele zostały zbudowane w oparciu o analizę statystyczną danych eksperymentalnych zebranych dla zbioru sekwencji testowych o rozdzielczości przestrzennej 704x576 punktów i częstotliwościach obrazu 25 lub 30 Hz. Każda sekwencja testowa zawierała obrazy tylko z pojedynczego ujęcia, czyli w sekwencjach wizyjnych nie

występowały cięcia. Należy zaznaczyć, że kiedy rozpoczynano badania brakowało dobrych sekwencji testowych formatów większych niż 4CIF. Dlatego badania przeprowadzono na sekwencjach o standardowej rozdzielczości telewizyjnej, a dostępne sekwencje o wyższej rozdzielczości niż 4CIF zostały odpowiednio zdecydowane. Szczegółowe informacje o wykorzystanych sekwencjach wizyjnych zawiera Dodatek A. Sekwencje zostały zakodowane kolejno ze wszystkimi możliwymi szerokościami przedziałów kwantyzacji Q (jedno zakodowanie sekwencji dla jednej, stałej szerokości przedziałów kwantyzacji Q). Dla każdego takiego kodowania zebrano informacje dotyczące wielkości strumienia danych na wyjściu kodera. W ten sposób otrzymano krzywe eksperymentalne przedstawiające liczbę bitów potrzebnych do reprezentowania zarówno całych obrazów, jak również poszczególnych składników strumienia danych w funkcji szerokości przedziałów kwantyzacji Q . Dla takich krzywych, autor szukał prostych opisów matematycznych z możliwie małą liczbą parametrów.

Do znalezienia wzoru matematycznego i parametrów opisujących otrzymaną krzywą eksperymentalną autor zastosował metodę dopasowania funkcji do danych eksperymentalnych. W pierwszym kroku tej metody, spośród różnych klas funkcji (wykładniczych, logarytmicznych, potęgowych, itd.) wybiera się te, których wykresy są najbardziej podobne do badanej krzywej. Ponieważ takie dopasowanie może okazać się dobrym tylko z pozoru, to w drugim kroku, przydatność wybranego wzoru sprawdza się metodą rektyfikacji. Rektyfikacja polega na tym, że dla danych wartości x i y szukamy takich nowych X i Y postaci:

$$X = \varphi(x, y), \quad Y = \psi(x, y), \quad (1.1)$$

które są ze sobą powiązane w przybliżeniu liniowo, tzn. $Y = AX + B$.

Przykładowo, jeśli sprawdzana jest zależność $y = ax^b$, to dokonuje się przekształcenia $X = \log x$ oraz $Y = \log y$. Następnie sprawdza się, czy po wyliczeniu nowych X i Y dla wszystkich x i y oraz przedstawieniu ich na wykresie, otrzymano punkty układające się w przybliżeniu wzdłuż linii prostej. Wartości parametrów A i B można wyznaczyć np. metodą najmniejszych kwadratów [Bron07, Pres07, Rao09]. Do stwierdzenia, na ile otrzymane w ten sposób dane są zgodne z liniową zależnością autor wykorzystał współczynnik determinacji liniowej [Ciec07, Koro01]:

$$R^2 = \frac{\sum_i ((A \cdot X_i + B) - \bar{Y})^2}{\sum_i (Y_i - \bar{Y})^2} \quad (1.2)$$

gdzie:

- R^2 to współczynnik determinacji liniowej,
- Y_i to wartość y_i po rektyfikacji,
- X_i to wartość x_i po rektyfikacji,
- \bar{Y} to średnia wartość zbioru $\{Y_i\}$,
- A i B to parametry funkcji liniowej wyznaczone metodą najmniejszych kwadratów.

Współczynnik determinacji liniowej przyjmuje wartości z przedziału $<0; 1>$. Im wartość współczynnika jest bliższa jedności, tym lepsze jest dopasowanie badanego modelu do danych eksperymentalnych. Procedura poszukiwania i weryfikacji jest powtarzana do czasu znalezienia funkcji o możliwie małej liczbie parametrów i dobrze dopasowanej do danych eksperymentalnych. Takie podejście jest dokładnie opisane w literaturze [Bran99, Bron07] i powszechnie stosowane do rozwiązywania podobnych problemów.

Następnie, stosując minimalizację maksymalnego względnego błędu aproksymacji, autor wyznaczał wartości parametrów znalezionej funkcji opisującej dane eksperymentalne. Do minimalizacji zastosowano metodę quasi-newtonowską [Chon08, Rao09, Ravi06], a jako punkty startowe do procedury optymalizacji wykorzystano wartości parametrów wyznaczone metodą najmniejszych kwadratów dla zrektifikowanych danych.

Dla wcześniej wskazanego przykładu funkcji postaci $y = ax^b$, problem minimalizacji maksymalnego względnego błędu aproksymacji można zapisać jako:

$$\min_{a,b} \max_x \frac{|y_z(x) - y(x, a, b)|}{y_z(x)} \cdot 100\%, \quad (1.3)$$

gdzie:

- a, b to parametry modelu,
- $y_z(x)$ to wartość zmierzona,
- $y(x, a, b)$ to wartość oszacowana (modelem $y = ax^b$).

Dla tak wyznaczonych modeli i ich parametrów, autor określał maksymalny i/lub średni względny błąd aproksymacji danych eksperymentalnych (wzory 1.5 i 1.6). Względny błąd aproksymacji oznaczono jako:

$$\varepsilon(x_i) = \frac{|y_z(x_i) - y(x_i)|}{y_z(x_i)} \cdot 100\%, \quad (1.4)$$

gdzie:

- $\varepsilon(x_i)$ to względny błąd aproksymacji,
- $y_z(x_i)$ to wartość zmierzona,
- $y(x_i)$ to wartość oszacowana przy pomocy modelu.

Maksymalny względny błąd aproksymacji danych eksperymentalnych wyznaczano jako:

$$\max_{x_i} \varepsilon(x_i), \quad (1.5)$$

a średni względny błąd aproksymacji jako:

$$\frac{1}{n} \sum_{i=1}^n \varepsilon(x_i), \quad (1.6)$$

gdzie:

- n to liczebność zbioru $\{x_i\}$.

Oceny dokładności oszacowania (aproksymacji) danych eksperymentalnych przez badany model dokonywano na podstawie wartości średniego względnego błędu aproksymacji (wzór 1.6). W rozprawie dokładność aproksymacji danych przez badany model będzie nazywana w skrócie dokładnością modelu.

Do analizy danych eksperymentalnych i opracowania modeli matematycznych autor wykorzystał pakiet matematyczny Mathcad [Math] firmy PTC w wersji 14.0.

1.4. Przegląd pracy

Praca dotyczy modelowania zaawansowanych koderów wizyjnych, a w szczególności koderów wytwarzających strumienie binarne zgodne ze standardem MPEG-4 AVC/H.264, który jest coraz szerzej stosowany w systemach multimedialnych. W szczególności przedstawione w rozprawie modelowanie dotyczy koderów, których sprawność kompresji jest równa sprawności kompresji referencyjnej implementacji kodera wizyjnego standardu MPEG-4 AVC/H.264.

W rozdziale 2 przedstawiono sposób reprezentacji skompresowanych sekwencji wizyjnych na wyjściu kodera MPEG-4 AVC/H.264 zgodnego z profilem głównym.

W rozdziale 3 dokonano przeglądu literatury dotyczącej modelowania hybrydowych koderów wizyjnych ze szczególnym uwzględnieniem koderów wykorzystujących zaawansowane techniki kompresji.

Rozdziały 4 do 6 zawierają autorskie propozycje ilościowych modeli kodera wizyjnego standardu MPEG-4 AVC/H.264 wraz z opisem kolejnych etapów tworzenia tych modeli. Ponadto, umieszczono w nich wyniki analizy dokładności zaproponowanych modeli przeprowadzonej dla wielu reprezentatywnych sekwencji wizyjnych o standardowej rozdzielczości telewizyjnej. Należy zaznaczyć, że w rozdziale 4 zamieszczono opracowania badań wraz ze wszystkimi wynikami cząstkowymi. Natomiast w rozdziałach 5 i 6, ze względu na bogaty zbiór wyników, zamieszczono tylko zbiorcze opracowania rezultatów badań z pominięciem wyników cząstkowych.

Wyniki rozprawy podsumowano w rozdziale 7. Przedstawiono w nim główne wnioski z przeprowadzonych badań oraz zaprezentowano najważniejsze osiągnięcia. Wskazano także możliwe kierunki dalszych badań.

Rozdział 2

Reprezentacja skompresowanych sygnałów wizyjnych

2.1. Wprowadzenie

Nieskompresowane cyfrowe reprezentacje sekwencji wizyjnych wymagają ogromnych ilości danych, przez co ich przechowywanie na nośnikach danych i przesyłanie przez systemy transmisyjne o ograniczonym paśmie jest bardzo kosztowne, a często nawet niemożliwe. Z tego powodu przed wysłaniem lub zapisaniem sekwencji wizyjną poddaje się procesowi kompresji. Zmniejszenie ilości danych potrzebnych do reprezentowania sekwencji wizyjnej (kompresja) jest możliwe ze względu na to, że sekwencje wizyjne charakteryzują się dużą nadmiarowością statystyczną oraz zawierają informacje percepcyjnie nieistotne [Baxe94, Bovi05, Doma98, Ohm04, Skar98, Teka95]. Urządzenie lub oprogramowanie realizujące proces kompresji danych wizyjnych (kodowanie) nazywa się koderem, natomiast proces odwrotny (dekodowanie) realizowany jest przez dekoder. Łącznie oba te elementy nazywa się kodekiem sekwencji wizyjnych.

Dwa główne podejścia do kompresji cyfrowych sygnałów wizyjnych to kodowanie hybrydowe i kodowanie falkowe [Acha05, Bovi05]. W kodowaniu hybrydowym, najważniejsze techniki prowadzące do zredukowania strumienia bitowego, to predykcja międzyobrazowa z kompensacją ruchu, transformacja w blokach o określonym rozmiarze i kodowanie entropijne [Doma98, Rich02, Sull05]. Natomiast w kodowaniu falkowym,

zamiast transformacji w blokach, stosuje się analizę subpasmową przestrzenną i/lub czasową [Bovi05, Ohm04]. Mimo że kodowanie falkowe jest nadal rozwijane, to nie znalazło szerszego zastosowania w praktycznych aplikacjach. Wyjątkiem jest międzynarodowy standard JPEG2000 [JPEG2000, Penn93, Taub02] stosowany do kodowania obrazów nieruchomych oraz wewnątrzobrazowego kodowania obrazów ruchomych (np. w kinach cyfrowych). Obecnie wszystkie powszechnie stosowane standardy kompresji cyfrowych sekwencji wizyjnych bazują na kodowaniu hybrydowym [Oste04]. Dlatego w dalszej części praca dotyczy wyłącznie kodowania hybrydowego.

Po roku 2000 rozwinięto nowe techniki hybrydowego kodowania sekwencji wizyjnych, nazywane *zaawansowanymi technikami kompresji*, które ujęto w standardach MPEG-4 AVC/H.264 [AVC], AVS [AVS] i VC-1 [VC1]. Spośród tych standardów, MPEG-4 AVC/H.264 wykorzystuje najbardziej złożone narzędzia kompresji [Fan04, Kalv07, Lee08b, Wieg03b, Yu09]. Ponadto, oprogramowanie referencyjne tego standardu jest publicznie dostępne [AVCref]. Dlatego jest on obecnie powszechnie wykorzystywany jako punkt odniesienia w badaniach naukowych i testach. W dalszej części praca będzie dotyczyć modelowania kodera standardu MPEG-4 AVC/H.264.

2.2. Struktura strumienia bitowego na wyjściu kodera standardu MPEG-4 AVC/H.264

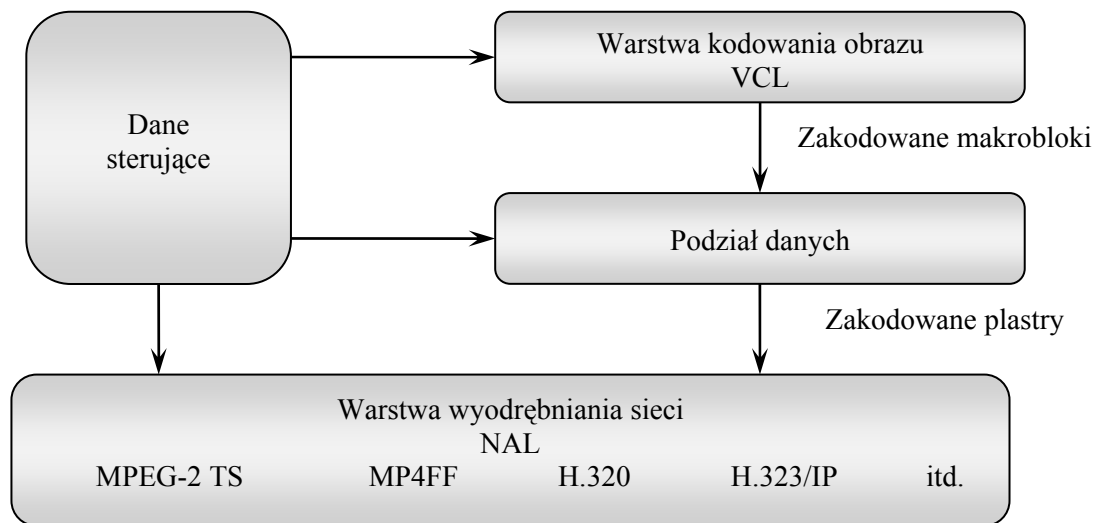
Jak zaznaczono we wstępie niniejszej pracy, zakres standaryzacji obejmuje strumień binarny, a dokładnie jego format i składnię oraz sposób dekodowania odebranych danych [Oste04]. Format strumienia określa z jakich elementów może składać się strumień. Natomiast składnia definiuje w jakiej kolejności mogą one występować w strumieniu, a w związku z tym, jak należy je interpretować. Żeby zapewnić poprawną współpracę koderów i dekoderów pracujących w różnych środowiskach, standard definiuje tzw. punkty zgodności dla koderów i dekoderów opisane przez profile i poziomy. Profile określają dozwoloną składnię strumienia bitowego, a poziomy ograniczają wartości istotnych parametrów (np. prędkość bitową strumienia danych) [Rich03].

Standard MPEG-4 AVC/H.264 definiuje 11 profili dedykowanych do różnych zastosowań [AVC, Marp05]. Wszystkie wyniki badań zaprezentowane w tej pracy dotyczą strumieni bitowych zgodnych z profilem głównym (ang. *Main*), ponieważ jest on najczęściej

stosowany, zwłaszcza w transmisji rozsyłkowej programów telewizyjnych [Marp06, Rich03, Sull05].

W odróżnieniu od wcześniejszych standardów w MPEG-4 AVC/H.264 wyraźnie rozróżniono dwie części (rysunek 2.1):

- warstwę kodowania obrazu VCL (ang. *Video Coding Layer*),
- warstwę wyodrębniania sieci NAL (ang. *Network Abstraction Layer*).



Rysunek 2.1. Struktura kodera MPEG-4 AVC/H.264 (opracowanie własne na podstawie [Wieg03b])

Za wytwarzanie efektywnej reprezentacji sekwencji wizyjnej odpowiada warstwa kodowania obrazu (VCL), a warstwa wyodrębniania sieci (NAL) tylko ją formatuje i zapewnia dodatkowe informacje potrzebne dla różnych nośników danych i systemów transportowych [Doma05]. Takie wyraźne rozróżnienie przez standard dwóch warstw: kodowania obrazu i wyodrębniania sieci, ma zapewnić łatwe dostosowanie strumienia do specyfiki poszczególnych nośników danych lub sieci telekomunikacyjnych (RTP/IP, format plików MP4, protokoły H.32x, strumień transportowy MPEG-2 itd.) [Rich03].

Ponieważ to w warstwie kodowania obrazu decyduje się, ile bitów będzie potrzebnych do reprezentowania treści sekwencji wizyjnej, dlatego w dalszej części autor skupi się wyłącznie na niej.

2.3. Warstwa kodowania obrazu – VCL

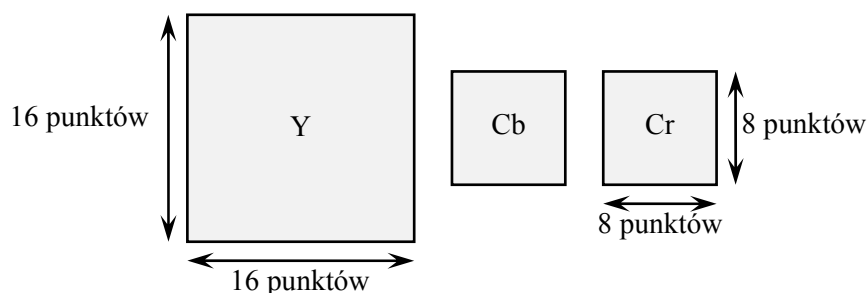
W warstwie kodowania obrazu (VCL) wykorzystuje się dobrze znany schemat blokowego, hybrydowego kodowania sekwencji wizyjnych [Bovi05, Doma98, Ghan99, Rich03], który był wykorzystywany także w wielu wcześniejszych standardach przyjętych przez ISO/IEC oraz ITU-T, takich jak H.261 [H261], H.263 [H263], MPEG-1 [MPEG1], MPEG-2 [MPEG2] i MPEG-4 (część 2) [MPEG4]. Przewaga MPEG-4 AVC/H.264 nad starszymi technikami kompresji wynika z wielu modyfikacji i ulepszeń wprowadzonych właśnie w tej warstwie. Pozwoliły one na osiągnięcie nawet 50% redukcji prędkości bitowej w stosunku do ciągle popularnego standardu MPEG-2 [Kama03, Lee08b, Raja04, Scha03, Sull05, Wieg03a].

Cyfrowa sekwencja wizyjna składa się z obrazów zarejestrowanych w kolejnych chwilach czasu. W praktyce obrazy te reprezentuje się w postaci 3 macierzy całkowitoliczbowych odpowiadających kolejnym składowym w przestrzeni kolorów YCbCr [Oste04, Rich02]. Składowa Y jest nazywana luminancją, natomiast składowe Cb i Cr to składowe różnicowe chrominancji. Składowe chrominancji zajmują zazwyczaj węższe pasmo częstotliwości przestrzennych niż luminancja. Ponadto, człowiek jest bardziej wrażliwy na zniekształcenia składowej Y niż składowych Cb i Cr. Dlatego przyjęto, że liczba próbek w każdej macierzy składowych chrominancji jest równa jednej czwartej liczby próbek z macierzy luminancji (odpowiada to zmniejszeniu rozmiaru każdej macierzy o połowę zarówno w pionie jak i poziomie). Taki sposób decymacji chrominancji określa się mianem schematu próbkowania 4:2:0 [Doma98, Rich02]. W dalszej części pracy zakłada się, że wszystkie sekwencje wizyjne będą reprezentowane właśnie w takiej postaci.

2.3.1. Makrobloki, plastry i obrazy

W procesie kodowania obraz dzieli się na bloki o rozmiarze 16x16 punktów luminancji i odpowiadające im bloki o rozmiarze 8x8 punktów obu składowych chrominancji. Taką strukturę nazywa się *makroblokiem* (rysunek 2.2). Makroblok jest podstawowym fragmentem obrazu, dla którego dokonuje się kompresji. Jeśli do zakodowania makrobloku wykorzystuje się informacje pochodzące wyłącznie z obecnie kodowanego obrazu, to takie kodowanie nazywa się *wewnątrzobrazowym*, a o makrobloku mówi się, że jest typu *Intra*. Natomiast, jeżeli przy kodowaniu korzysta się z informacji zawartych w innym obrazie lub obrazach, to wtedy o makrobloku mówi się, że jest typu *Inter*, a kodowanie nazywa się *międzyobrazowym*.

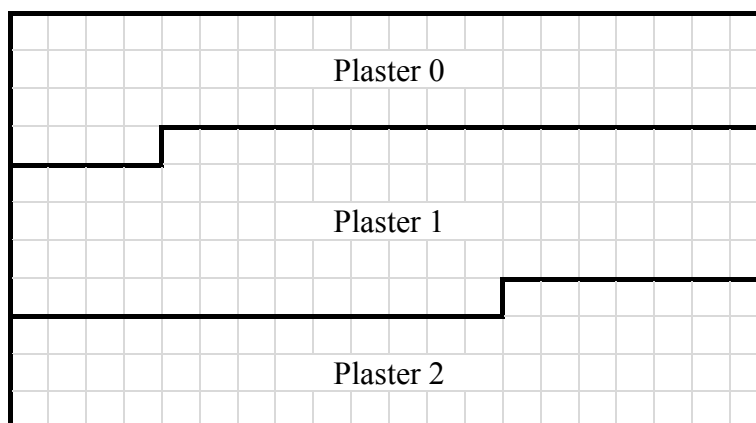
Obrazy, z których pochodzą informacje wykorzystywane przy kodowaniu makrobloków typu Inter, nazywa się *obrazami odniesienia*.



Rysunek 2.2. Struktura makrobloku w MPEG-4 AVC/H.264 (schemat próbkowania 4:2:0)

Makrobloki łączy się w struktury zwane *plastrami* (ang. *slice*), które reprezentują obszary obrazu zakodowane niezależnie od siebie. Każdy plaster składa się z grupy makrobloków kodowanych w kolejności liniowej tzn. od lewej do prawej i z góry na dół (rysunek 2.3). W MPEG-4 AVC/H.264 wyjątkiem od tej reguły jest technika FMO (ang. *Flexible Macroblock Ordering*) [Scha03, Weng03], która dopuszcza inne schematy kolejności kodowania makrobloków (a tym samym podziału na plastry), jednak jej użycie nie jest dozwolone w profilu głównym [AVC, Marp06, Oste04].

Zakodowane plastry, które zawierają wyłącznie makrobloki Intra nazywa się plastrami typu I. Jeśli w plastrach dopuszcza się użycie makrobloków Inter, które korzystają z informacji wyłącznie z wcześniejszych obrazów, to takie plastry nazywane są plastrami typu P. Natomiast, gdy można użyć informacji również z obrazów późniejszych od bieżącego, wtedy takie plastry są nazywane plastrami typu B.

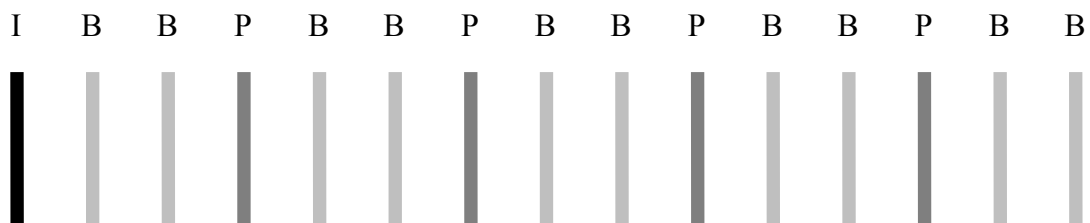


Rysunek 2.3. Podział obrazu na plastry (opracowanie własne na podstawie [Wieg03b])

Podział obrazu na plastry pozwala na redukcję strat jakości obrazu zdekodowanego ze strumienia z błędami transmisji [Weng03]. Niestety odbywa się to kosztem pogorszenia sprawności kompresji [Sze09, Zhao08]. Często jednak nie korzysta się z tego mechanizmu i wtedy jeden plaster obejmuje cały obraz. Wszystkie sekwencje wizyjne użyte do badań zostały zakodowane właśnie z takim założeniem. Dlatego w dalszej części pracy, aby na każdym kroku nie zaznaczać, że obraz został zakodowany jako jeden plaster, autor będzie się posługiwał pojęciami obrazów typu I, P i B.

Aby umożliwić rozpoczęcie dekodowania sekwencji także od innego miejsca niż początek strumienia danych oraz zapewnić ochronę tego strumienia przed błędami transmisji, obrazy łączy się w niezależnie dekodowalne grupy (ang. *GOP – Group Of Pictures*). Na pierwszym miejscu w każdej grupie obrazów jest obraz typu I, a na kolejnych obrazy typu P, z których każdy może być poprzedzony jednym lub kilkoma obrazami typu B. Na rysunku 2.4 przedstawiono typową strukturę GOP dla zastosowań telewizyjnych, gdzie przyjmuje się, że punkty, od których można rozpocząć dekodowanie strumienia, mają być w odstępach około 0,5 sekundy. W literaturze strukturę grupy obrazów opisuje się przeważnie na jeden z dwóch poniższych sposobów (rysunek 2.4):

- pełny: IBBPBBPBBPBB,
- skrócony: 4P2B.



Rysunek 2.4. Typowa struktura grupy obrazów (GOP) dla zastosowań telewizyjnych

W przeprowadzonych badaniach obrazy typu I i P są obrazami *referencyjnymi*, tzn. mogą zostać użyte jako obrazy odniesienia dla innych obrazów w sekwencji wizyjnej, natomiast obrazy typu B nie są obrazami referencyjnymi (tzn. nie są odniesieniami dla innych obrazów) [Wieg03b].

2.3.2. Predykcja wartości próbek obrazu

W standardzie MPEG-4 AVC/H.264, poza praktycznie nie stosowanym trybem I_PCM, bezpośrednio nie przesyła się wartości próbek obrazu [AVC]. Dla wszystkich próbek składowych luminancji i obu chrominancji obrazu, makroblok po makrobloku, tworzony jest sygnał predykcji w przestrzeni lub czasie, a następnie oblicza się wartości różnic między uzyskanym sygnałem predykcji a sygnałem oryginalnym. Wyznaczony błąd predykcji jest poddawany transformacji całkowitoliczbowej w blokach o rozmiarze 4x4 punkty. Zastosowana transformacja jest przybliżeniem dyskretnej transformacji kosinusowej (DCT) [Baxe94, Bovi05, Doma98, Skar98] dobrze znanej z wcześniejszych standardów. Następnie otrzymane wartości próbek transformaty są kwantowane [Malv03]. Dopiero skwantowane wartości próbek transformaty błędu predykcji są kodowane koderem entropijnym i zapisywane do strumienia danych [Rich03, Sull04]. W makroblokach typu Intra predykcji wartości próbek dokonuje się na podstawie wcześniej zakodowanych makrobloków znajdujących się w tym samym obrazie. Natomiast w makroblokach typu Inter wykorzystuje się technikę predykcji z kompensacją ruchu. Celem tej techniki jest poszukanie w obrazie odniesienia, fragmentu obrazu jak najbardziej podobnego do obecnie kodowanego [Wedi03]. Informacja o odległości w przestrzeni, tego znalezionej fragmentu względem aktualnie kodowanego, umieszczana jest w strumieniu danych w postaci dwóch wartości, które łącznie nazywa się *wektorem ruchu*. Należy zaznaczyć, że predykcji z kompensacją ruchu można dokonywać dla całego makrobloku, jak również dla jego części (wtedy każda część makrobloku ma przypisany osobny wektor ruchu). Najmniejsza część makrobloku, dla której dokonuje się predykcji z kompensacją ruchu, ma rozmiar 4x4 punkty luminancji [Dzię05, Dzię06, Doma05, Rich03, Wieg03b].

2.3.3. Kwantowanie

Kwantowanie wartości próbek transformaty jest etapem, w którym dokonuje się zasadnicza kompresja danych, przy czym utrata informacji jest nieodwracalna [Doma98]. W standardzie MPEG-4 AVC/H.264, podobnie jak we wcześniejszych standardach kompresji sekwencji wizyjnych, zastosowano kwantyzator równomierny, tzn. o stałych wartościach szerokości przedziałów kwantyzacji Q [Malv03, Marp06]. W celu regulowania w szerokim zakresie wielkości strumienia danych na wyjściu kodera wizyjnego, zdefiniowano 52 szerokości przedziałów kwantyzacji Q . Jednak w strumieniu danych nie umieszcza się ich

wartości wprost, tylko poprzez podanie indeksu parametru kwantyzacji QP (ang. *Quantization Parameter*) dla składowej luminancji [Sull05, Wieg03b]. Szerokości przedziałów kwantyzacji dla składowej luminancji oraz odpowiadające im indeksy QP zebrano w tabeli 2.1.

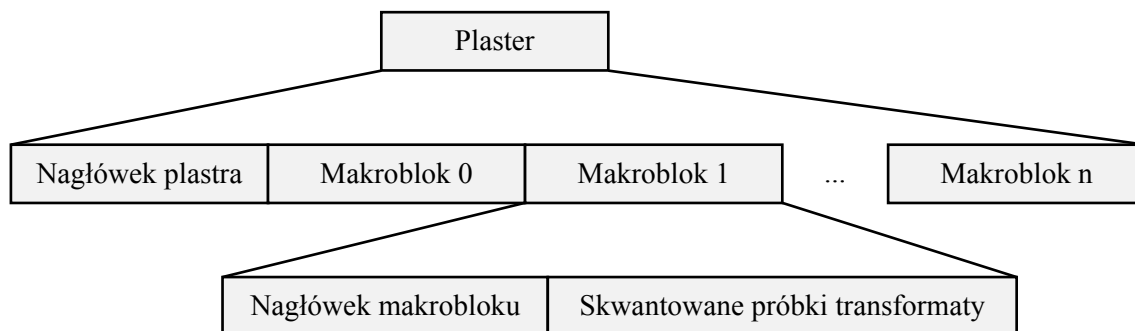
Tabela 2.1. Szerokości przedziałów kwantyzacji Q i opowiadające im wartości parametru kwantyzacji QP dla składowej luminancji (opracowanie własne na podstawie [Rich03])

QP	Q	QP	Q
0	0,625	26	13
1	0,6875	27	14
2	0,8125	28	16
3	0,875	29	18
4	1	30	20
5	1,125	31	22
6	1,25	32	26
7	1,375	33	28
8	1,625	34	32
9	1,75	35	36
10	2	36	40
11	2,25	37	44
12	2,5	38	52
13	2,75	39	56
14	3,25	40	64
15	3,5	41	72
16	4	42	80
17	4,5	43	88
18	5	44	104
19	5,5	45	112
20	6,5	46	128
21	7	47	144
22	8	48	160
23	9	49	176
24	10	50	208
25	11	51	224

Z powyższej tabeli wynika, że zwiększeniu wartości parametru kwantyzacji QP o 6 odpowiada dwukrotny wzrost szerokości przedziału kwantyzacji Q (zwiększenie QP o 1 daje około 12,5% wzrost wartości Q). Należy tu dodać, że wartość szerokości przedziałów kwantyzacji można ustalać indywidualnie nawet dla każdego makrobloku. Szczegółowe informacje dotyczące kwantowania wartości próbek transformaty błędu predykcji w standardzie MPEG-4 AVC/H.264 można znaleźć w [Malv03, Rich03].

2.3.4. Zapis danych do strumienia bitowego

Schemat zapisu danych do strumienia bitowego jest wspólny dla plastrów różnych typów. Na początku każdego plastra umieszcza się w strumieniu dane nagłówkowe (m. in. typ plastra, numer pierwszego makrobloku itd.). Po nagłówku plastra zapisuje się dane o poszczególnych makroblokach w kolejności ich kodowania. W każdym makrobloku również można wyróżnić informacje nagłówkowe (typ makrobloku, ewentualnie wektory ruchu itd.), po których w strumieniu bitowym umieszcza się zakodowane skwantowane wartości próbek transformaty błędu predykcji składowej luminancji i obu składowych chrominancji (rysunek 2.5).



Rysunek 2.5. Struktura strumienia bitowego

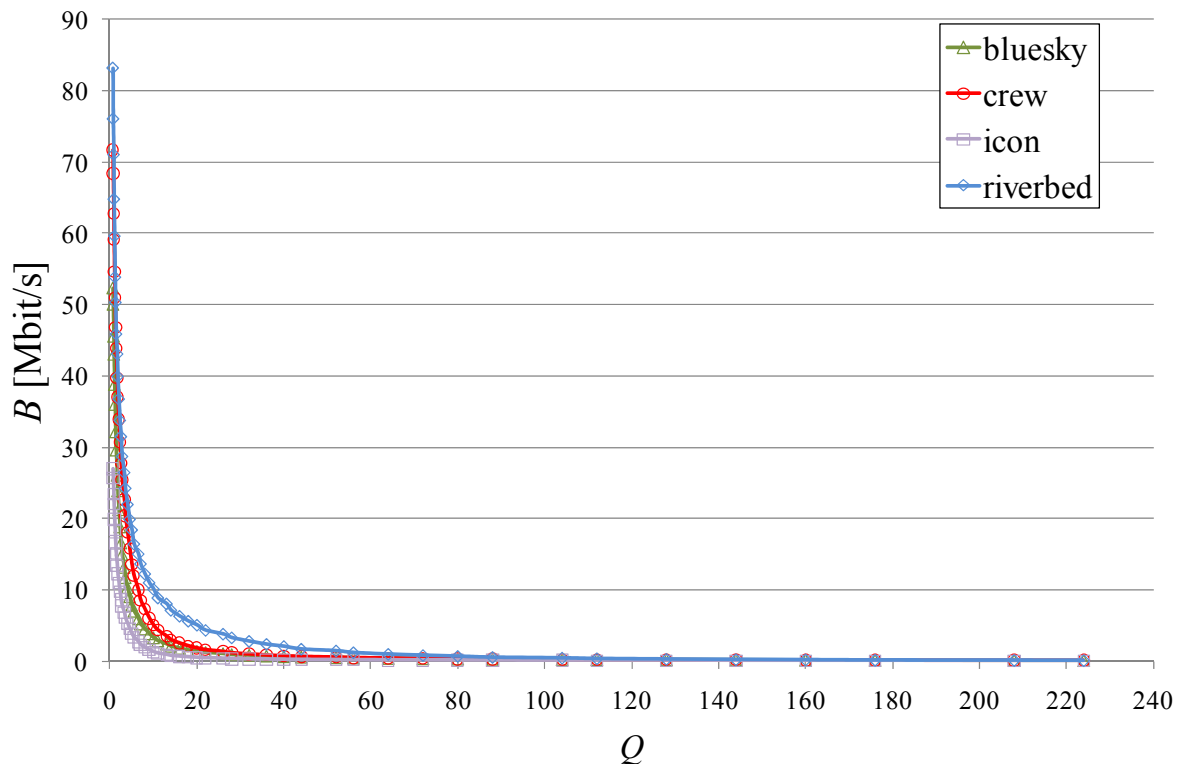
W profilu głównym standardu MPEG-4 AVC/H.264, do zapisu danych do strumienia bitowego można stosować jedną z dwóch technik kontekstowego, adaptacyjnego kodowania entropijnego:

- kodowanie o zmiennej długości słowa – CAVLC (ang. *Context-based Adaptive Variable Length Coding*) [AVC, Rich03, Sayo05, Wieg03b],
- kodowanie arytmetyczne – CABAC (ang. *Context-based Adaptive Binary Arithmetic Coding*) [AVC, Marp04, Salo07, Sayo05, Sull05].

W porównaniu z CAVLC, technika CABAC odznacza się lepszą sprawnością kompresji, jednakże złożoność procesu kodowania i dekodowania tą techniką jest dużo większa [Graj05, Marp03, Wieg03b]. Wszystkie badane sekwencje wizyjne zostały zakodowane przy użyciu kodera entropijnego CABAC, ze względu na lepszą sprawność kompresji.

Na wykresie 2.1 pokazano zależność prędkości bitowej na wyjściu kodera MPEG-4 AVC/H.264 (zgodnego z profilem głównym) od szerokości przedziałów kwantyzacji Q dla przykładowych sekwencji wizyjnych o rozdzielczości przestrzennej 704x576 punktów

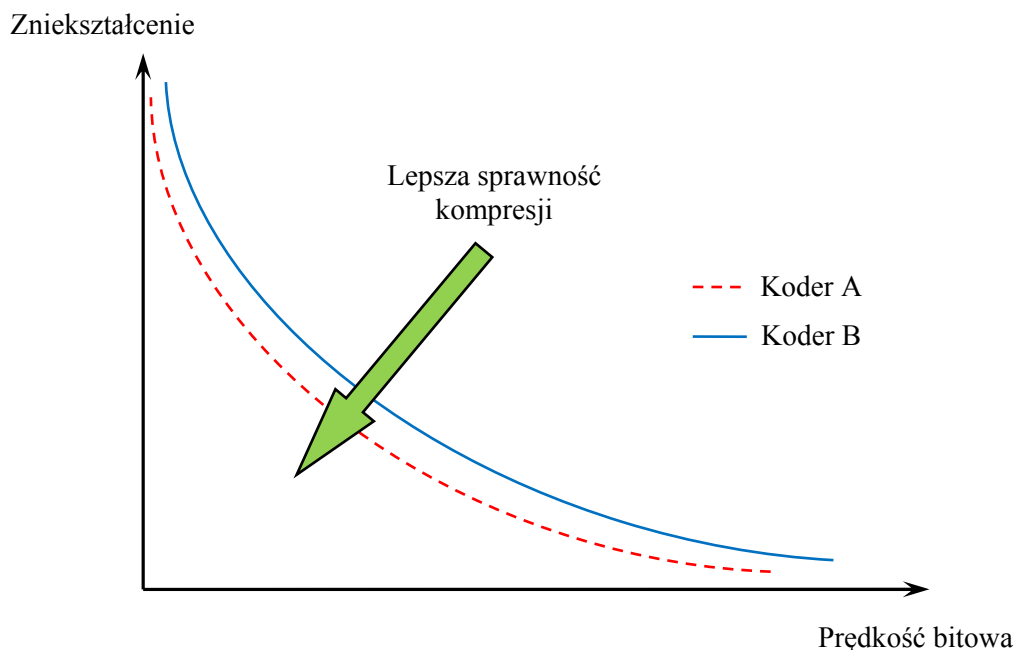
luminancji, struktury GOP typowej dla zastosowań telewizyjnych (4P2B) oraz kodowania CABAC.



Wykres 2.1. Prędkość bitowa strumienia danych B na wyjściu koderu MPEG-4 AVC/H.264 w zależności od szerokości przedziałów kwantyzacji Q dla 4 przykładowych sekwencji wizyjnych o rozmiarze 4CIF (struktura GOP – 4P2B)

Z wykresu 2.1 wynika, że zmiana szerokości przedziałów kwantyzacji Q pozwala na regulację prędkości bitowej strumienia danych w bardzo szerokim zakresie, tj. od dziesiątek megabitów do dziesiątek kilobitów na sekundę (sekwencje o rozdzielczości przestrzennej 4CIF). Z tego wykresu można również wywnioskować, że prędkość bitowa strumienia danych przy stałej wartości Q zależy od treści kodowanej sekwencji wizyjnej. Dokładne wartości prędkości bitowej strumienia danych na wyjściu koderu wizyjnego dla wszystkich badanych sekwencji wizyjnych zebrano w dodatku B.

Poszczególne wartości prędkości bitowej dla konkretnej kodowanej sekwencji odpowiadają różnej jakości obrazu zrekonstruowanego ze strumienia. Dla poszczególnych typów koderów i dla poszczególnych sekwencji można wykreślić linie „prędkość bitowa – zniekształcenie” (wykres 2.2).



Wykres 2.2. Przykładowe krzywe „prędkość bitowa – zniekształcenie” dla dwóch koderów o różnej sprawności kompresji

Jeżeli dla koder A wspomniana krzywa leży poniżej krzywej dla koder B, to mówi się, że koder A wykazuje dla danej sekwencji lepszą sprawność kompresji. Często sprawność kompresji koderów porównuje się za pomocą miar Bjontegaarda [Bjon01, Bjon08, Tan08].

2.4. Podsumowanie

W rozdziale została przedstawiona reprezentacja skompresowanych sekwencji wizyjnych na przykładzie standardu MPEG-4 AVC/H.264. Opisano także niektóre aspekty algorytmu hybrydowego kodowania sekwencji wizyjnych, potrzebne do zrozumienia podstawowych informacji dotyczących semantyki i składni strumienia bitowego na wyjściu zaawansowanego koder sekwencji wizyjnych. Na przykładzie kilku sekwencji wizyjnych pokazano również, że parametrem pozwalającym na regulację w bardzo szerokim zakresie prędkości bitowej strumienia danych na wyjściu koderu wizyjnego jest szerokość przedziałów kwantyzacji Q .

Rozdział 3

Przegląd rozwiązań w zakresie modelowania hybrydowych koderów wizyjnych

3.1. Wprowadzenie

Opracowanie dobrego modelu matematycznego koderów wizyjnych ma duże znaczenie zarówno poznawcze, jak i praktyczne. Znajomość modelu opisującego w zależności od treści sekwencji relację pomiędzy parametrami kodera a wyjściowym strumieniem bitowym pozwalałaby na przewidywanie zachowania się kodera w różnych warunkach. Ponadto, taki model mógłby zostać wykorzystany do ustalania wartości szerokości przedziałów kwantyzacji Q dla zadanej liczby bitów w algorytmie sterowania koderem. Dlatego w literaturze dotyczącej technik hybrydowego kodowania sekwencji wizyjnych modelowanie koderów wizyjnych jest stosunkowo często poruszonym tematem.

W kolejnych punktach zostaną przedstawione znane z literatury propozycje modelowania hybrydowych koderów wizyjnych, ze szczególnym uwzględnieniem modeli matematycznych dla koderów standardu MPEG-4 AVC/H.264.

3.2. Typowe modele wykorzystywane w implementacjach koderów standardu MPEG-4 AVC/H.264

Standardy ISO/IEC i zalecenia ITU-T dotyczące koderów wizyjnych nie określają procesu kodowania sekwencji wizyjnych (patrz punkt 1.1). Mimo to, rozwiązania zastosowane w referencyjnej implementacji koder standardu MPEG-4 AVC/H.264 [AVCref] są traktowane jako punkt odniesienia do badań dotyczących kompresji cyfrowych sekwencji wizyjnych. Opis tych rozwiązań można znaleźć w [Lim05], gdzie między innymi wskazano, jakie modele matematyczne wykorzystuje się do wyznaczenia szerokości przedziałów kwantyzacji Q dla zadanej liczby bitów.

Do modelowania zależności między liczbą bitów potrzebnych do reprezentacji skwantowanych wartości próbek transformaty błędu predykcji obrazu a szerokością przedziałów kwantyzacji Q w obrazach typu P zastosowano model zaproponowany w [Lee00]. Jest on zmodyfikowaną wersją modelu z [Chia97], który został wyprowadzony przy założeniu, że funkcję gęstości prawdopodobieństwa próbek transformaty można opisać rozkładem Laplace'a [Chia97, Vite79]. W zmodyfikowanej wersji modelu założono, że dla stałej wartości szerokości przedziałów kwantyzacji Q , liczba bitów potrzebnych do reprezentacji skwantowanych wartości próbek transformaty obrazu typu P jest wprost proporcjonalna do miary złożoności treści tego obrazu:

$$B - H = a \cdot \frac{MAD}{Q} + b \cdot \frac{MAD}{Q^2}, \quad (3.1)$$

gdzie:

- a i b to parametry modelu,
- B to liczba bitów reprezentacji obrazu typu P,
- H to liczba bitów reprezentujących nagłówki wraz z wektorami ruchu,
- Q to szerokość przedziałów kwantyzacji,
- MAD to przyjęta miara złożoności treści obrazu (wzór 3.2).

Występująca w tym modelu miara złożoności treści obrazu MAD (ang. *Mean Absolute Difference*) jest wyznaczoną dla składowej luminancji średnią bezwzględnych wartości różnic próbek obrazu oryginalnego i ich predykcji:

$$MAD = \frac{\sum_{x,y} |s_{x,y} - r_{x,y}|}{N}, \quad (3.2)$$

gdzie:

- $s_{x,y}$ to wartość próbki obrazu oryginalnego o współrzędnych (x,y) ,
- $r_{x,y}$ to wynik predykcji dla próbki obrazu o współrzędnych (x,y) ,
- x, y to współrzędne punktów obrazu, odpowiednio w poziomie i w pionie,
- N to liczba punktów składowej luminancji obrazu.

W referencyjnej implementacji kodera MPEG-4 AVC/H.264 zastosowano optymalizację „prędkość bitowa – zniekształcenie” – RDO (ang. *Rate Distortion Optimization*), która opiera się na metodzie mnożników Lagrange’a [Sull98, Wieg01, Wieg03a]. Optymalizacja ta jest wykorzystywana do wyboru trybu kodowania makrobloku dla ustalonej wartości szerokości przedziałów kwantyzacji Q . Aby wyznaczyć wartość Q (wzór 3.1), należy znać wartość miary MAD , czyli trzeba znać sygnał predykcji. Jednak sygnał predykcji jest dostępny dopiero po wykonaniu wspomnianej optymalizacji. Dlatego w referencyjnej implementacji kodera stosuje się jeszcze jeden model, w którym na podstawie wartości miary MAD dla wcześniej zakodowanego obrazu referencyjnego typu P, przewiduje się wartość miary MAD dla obecnie kodowanego obrazu referencyjnego typu P:

$$\widehat{MAD}_i = \alpha \cdot MAD_{i-1} + \beta, \quad (3.3)$$

gdzie:

- α i β to parametry modelu,
- \widehat{MAD}_i to przewidywana wartość miary MAD dla bieżącego obrazu typu P,
- MAD_{i-1} to wartość miary MAD dla wcześniej zakodowanego obrazu typu P.

Wartości parametrów, zarówno a i b , jak i α oraz β , wyznacza się na podstawie regresji liniowej odpowiednio dobranych danych uzyskanych dla wcześniej zakodowanych obrazów referencyjnych typu P [Lee00, Lim05]. Liczba bitów H potrzebnych do zakodowania nagłówek wraz z wektorami ruchu (wzór 3.1), jest traktowana jako składnik wolnozmienny [Lee00]. Dlatego, za wartość H dla aktualnie kodowanego obrazu typu P przyjmuje się liczbę bitów przypadających na nagłówki wraz z wektorami ruchu w poprzednio zakodowanym obrazie typu P.

Przedstawione modele (wzory: 3.1 i 3.3) pozwalają wyznaczyć szerokość przedziałów kwantyzacji Q dla bieżącego obrazu typu P na podstawie oszacowanej wielkości miary MAD . Takie same modele wykorzystuje się także do opisywania różnie zdefiniowanych fragmentów obrazu zwanych jednostkami bazowymi (ang. *basic unit*) [Li03]. Zmienne w opisanych wcześniej wzorach (3.1 - 3.3), zamiast do całego obrazu, odnoszą się do poszczególnych jednostek bazowych w bieżącym obrazie typu P lub jednostek bazowych na tej samej pozycji co modelowana jednostka we wcześniejszym obrazie typu P [Li04].

W referencyjnej implementacji kodera MPEG-4 AVC/H.264 modelowanie ogranicza się wyłącznie do obrazów typu P. Szerokość przedziałów kwantyzacji Q dla obrazów typu I ustala się na podstawie analizy wartości szerokości przedziałów kwantyzacji Q w obrazach typu P we wcześniej zakodowanej grupie obrazów GOP. Natomiast dla obrazów typu B analizuje się wartości Q z sąsiednich obrazów typu P [Li03, Li04, Lim05].

Opisane powyżej modele (wzory: 3.1 i 3.3) są szeroko stosowane w implementacjach zaawansowanych koderów wizyjnych oraz traktuje się je jako punkt odniesienia w badaniach dotyczących modelowania koderów wizyjnych. Należy podkreślić, że są one wykorzystywane tylko do modelowania obrazów typu P.

3.3. Modyfikacje modeli z referencyjnej implementacji kodera MPEG-4 AVC/H.264

3.3.1. Model opisujący zależność między liczbą bitów a szerokością przedziałów kwantyzacji Q

Model wykorzystywany do wyznaczania szerokości przedziałów kwantyzacji Q dla obrazów typu P (wzór 3.1) został opracowany na podstawie koderów hybrydowych wcześniejszych standardów. Ze względu na nowe skomplikowane narzędzia kompresji wykorzystane w standardzie MPEG-4 AVC/H.264 w literaturze można znaleźć prace pokazujące, że wspomniany model nie sprawdza się w przypadku koderów tego standardu. Na przykład w [Wan06] autorzy dostrzegają, że dla wielu sekwencji wizyjnych, liczba bitów potrzebnych do zakodowania skwantowanych wartości próbek transformaty składowych luminancji i chrominancji nie jest wprost proporcjonalna do wartości miary MAD . Dlatego zaproponowali oni, aby model ze wzoru 3.1 zastąpić modelem:

$$B - H = a + b \cdot \frac{MAD}{Q} + c \cdot \frac{MAD^2}{Q^2}, \quad (3.4)$$

gdzie:

- a , b i c to parametry modelu,
- B to całkowita liczba bitów reprezentacji obrazu typu P,
- H to liczba bitów przeznaczonych na zakodowanie nagłówków wraz z wektorami ruchu,
- Q to szerokość przedziałów kwantyzacji,
- MAD to miara złożoności treści obrazu (wzór 3.2).

Nowe modele do opisu kodera na poziomie jednostek bazowych zaproponowano w pracach [Xie04, Xie07, Xu04, Yuan05, Yuan06]. W [Xu04] zaproponowano model ze wzoru 3.4, ale z parametrem a równym 0. Wspomniany model ze wzoru 3.4 jeszcze bardziej uproszczono w [Xie04, Xie07], gdzie zaproponowano pozostawienie tylko członu zawierającego Q^2 w mianowniku oraz zastąpiono miarę MAD własną miarą złożoności treści obrazu. Ta zaproponowana miara złożoności treści obrazu została wyrażona jako:

$$T = (B - H) \cdot Q^2, \quad (3.5)$$

gdzie:

- T to zaproponowana miara złożoności treści obrazu,
- B to liczba bitów reprezentacji jednostki bazowej w obrazie typu P,
- H to liczba bitów przeznaczonych na zakodowanie nagłówków wraz z wektorami ruchu w jednostce bazowej,
- Q to szerokość przedziałów kwantyzacji.

Niestety wartość tej miary również trzeba przewidywać na podstawie wcześniej zakodowanych obrazów.

W [Yuan05, Yuan06] przedstawiono model kwadratowy postaci:

$$B_i - H_i = a_i \cdot Q_i^2 + b_i \cdot Q_i + c_i, \quad (3.6)$$

gdzie:

- a_i , b_i i c_i to parametry modelu dla i -tego makrobloku w obrazie typu P,
- B_i to liczba bitów reprezentacji i -tego makrobloku w obrazie typu P,

- H_i to liczba bitów potrzebnych do zakodowania nagłówków wraz z wektorami ruchu w i -tym makrobloku we wcześniejszym obrazie typu P,
- Q_i to szerokość przedziałów kwantyzacji w i -tym makrobloku.

Ta propozycja, w przeciwieństwie do wcześniejszych, nie wymaga znajomości miary złożoności obrazu. Jednak model ten jest wykorzystywany wyłącznie do modelowania kodera na poziomie jednostek bazowych w obrazach typu P. Do wyznaczenia wartości szerokości przedziałów kwantyzacji Q na poziomie obrazów wykorzystuje się model z referencyjnej implementacji kodera MPEG-4 AVC/H.264 (wzór 3.1).

Zależność między liczbą bitów reprezentacji obrazu a szerokością przedziałów kwantyzacji Q próbuje się także opisywać modelem postaci:

$$B = a \cdot \frac{MAD}{Q} + b, \quad (3.7)$$

gdzie:

- a i b to parametry modelu,
- B to liczba bitów reprezentacji obrazu,
- Q to szerokość przedziałów kwantyzacji,
- MAD to miara skomplikowania treści obrazu (wzór 3.2).

Propozycje takiego modelu można znaleźć w [Dong07a, Dong08, Jian06, Liu07]. Ponadto, w wymienionych pracach zauważono, że udział bitów potrzebnych do zakodowania nagłówków wraz z wektorami ruchu w całym strumieniu danych wzrósł w porównaniu do starszych standardów kompresji sekwencji wizyjnych. Jest to szczególnie widoczne dla dużych wartości szerokości przedziałów kwantyzacji Q (odpowiada to małym wartościom prędkości strumienia bitowego). Dlatego zaproponowany model (wzór 3.7) wyraża liczbę bitów reprezentacji całego obrazu typu P.

Model analogiczny do opisanego wzorem 3.7 został przedstawiony także w [Ma03, Ma05], gdzie jako miarę skomplikowania treści obrazu przyjęto sumę wartości bezwzględnych różnic próbek obrazu oryginalnego i ich predykcji *SAD* (ang. *Sum of Absolute Differences*):

$$SAD = \sum_{x,y} |s_{x,y} - r_{x,y}|, \quad (3.8)$$

gdzie:

- $s_{x,y}$ to wartość próbki obrazu oryginalnego o współrzędnych (x,y) ,
- $r_{x,y}$ to wynik predykcji dla próbki obrazu o współrzędnych (x,y) ,
- x, y to współrzędne punktów obrazu, odpowiednio w poziomie i w pionie,

Można również znaleźć prace, w których zamiast MAD wyznacza się sumę bezwzględnych wartości próbek transformaty błędu predykcji SATD (ang. *Sum of Absolute Transformed Differences*) [Kwon07].

Jednak, ze względu na stosowanie w koderze referencyjnym optymalizacji „prędkość bitowa – zniekształcenie” (RDO), wartości miar SAD oraz $SATD$ również nie są znane przed wyborem wartości Q .

Wyznaczając model kodera wizyjnego na podstawie analizy wartości próbek dyskretnej transformaty kosinusowej (DCT) [Bovi05, Baxe94, Doma98, Skar98] bardzo często zakłada się, że funkcja gęstości prawdopodobieństwa próbek transformaty jest dobrze aproksymowana rozkładem Laplace’a [Chia97, Ming07, Riba99, Wan06, Yuan06]. Można znaleźć też prace, w których do opisu funkcji gęstości prawdopodobieństwa wykorzystuje się rozkład Cauchy’ego [Kama04a, Kama05]. W pracach tych pokazuje się, że rozkład Cauchy’ego lepiej aproksymuje charakter źródła sygnału niż rozkład Laplace’a. Przyjmując nowe założenie o rozkładzie wartości próbek transformaty, autorzy wymienionych prac wyznaczyli teoretyczny model kodera wyrażający liczbę bitów reprezentacji składowych luminancji i chrominancji obrazu w funkcji szerokości przedziałów kwantyzacji Q postaci:

$$B = a \cdot Q^{-\alpha}, \quad (3.9)$$

gdzie:

- a i α to parametry modelu,
- B to liczba bitów reprezentacji skwantowanych wartości próbek transformaty,
- Q to szerokość przedziałów kwantyzacji.

Taki model rozkładu gęstości prawdopodobieństwa próbek transformaty wykorzystano także w [Lee08a] do modelowania obrazów typu I. Model opisany wzorem 3.9 został rozszerzony o zaproponowaną miarę złożoności treści obrazu wyznaczaną dla wartości próbek obrazu przed kompresją:

$$B = (a \cdot TC + b) \cdot Q^{-\alpha}, \quad (3.10)$$

gdzie:

- a , b i α to parametry modelu,
- B to liczba bitów reprezentacji obrazu typu I,
- TC to zaproponowana miara złożoności treści obrazu wyznaczona dla bloków o rozmiarze 5x5 punkty luminancji,
- Q to szerokość przedziałów kwantyzacji.

Model potęgowy (wzór 3.9), rozszerzony o wyraz wolny, zaproponowano także dla starszych standardów kodowania, np. w [Ding96].

3.3.2. Model aproksymujący wartość miary MAD

Do predykcji miary *MAD* w referencyjnej implementacji kodera standardu MPEG-4 AVC/H.264 stosuje się model liniowy (wzór 3.3), który wykorzystuje informacje tylko z poprzednich obrazów typu P. Przez to jest on bardzo niedokładny w przypadku sekwencji wizyjnych o dużej dynamice ruchu i skomplikowanej treści oraz na granicach ujęć [Liu07]. Obserwacja ta przyczyniła się do podjęcia badań mających na celu poprawienie dokładności oszacowania wartości miary *MAD*. W pracach [Liu06, Liu07, Yi04] zaproponowano usprawnienie modelu aproksymującego wartość miary *MAD* polegające na wykorzystaniu informacji o średniej bezwzględnej różnicy między wartościami próbek obrazu obecnie kodowanego a wartościami próbek we wcześniej zdekodowanym obrazie. Natomiast w [Ruol07] proponuje się, żeby przy modelowaniu na poziomie jednostek bazowych zamiast modelu liniowego użyć odpowiednio ważonej wartości miary *MAD* z odpowiednich jednostek bazowych z poprzedniego obrazu typu P.

Ponieważ parametry α oraz β modelu (wzór 3.3) wyznacza się na podstawie regresji liniowej danych uzyskanych dla wcześniej zakodowanych obrazów referencyjnych typu P [Lee00, Lim05], podjęto także badania sprawdzające, czy inny sposób doboru danych, wykorzystywanych do wyznaczania parametrów, może poprawić dokładność tego modelu [Dong07b, Dong08, Yu05].

3.4. Model kodera sparametryzowany procentem zerowych próbek transformaty

Całkiem inne podejście do modelowania koderów wizyjnych zostało przedstawione w serii publikacji [He01a, He01b, He01c, He01d, He02a, He02b, Kim01]. Autorzy proponują, aby hybrydowe kodery wizyjne modelować w tak zwanej przestrzeni ρ , gdzie ρ oznacza procentowy udział skwantowanych próbek transformaty o wartości zero we wszystkich skwantowanych próbkach transformaty. Ponieważ wartość ρ rośnie monotonicznie wraz ze wzrostem wartości Q , to można znaleźć jednoznaczne odwzorowanie wartości ρ na Q , a przez to liczbę bitów potrzebnych do reprezentacji skompresowanego obrazu wyrazić w funkcji ρ . Bazując na obserwacjach koderów zgodnych ze standardami MPEG-2 [MPEG2], H.263 [H263] i MPEG-4 [MPEG4] autorzy zaproponowali model liniowy postaci:

$$B = \theta \cdot (1 - \rho), \quad (3.11)$$

gdzie:

- θ to stała modelu,
- B to liczba bitów reprezentacji skwantowanych wartości próbek transformaty,
- ρ to procentowy udział skwantowanych próbek transformaty o wartości zero we wszystkich skwantowanych próbkach transformaty.

Znając wartość θ można wyznaczyć jaka powinna być wartość ρ dla zadanej liczby bitów. A następnie, na podstawie analizy procesu kwantowania w danym koderze, dobrać wartość Q tak, aby procentowy udział skwantowanych próbek transformaty o wartości zero we wszystkich skwantowanych próbkach transformaty wynosił ρ .

Obserwacje dotyczące liniowej zależności między liczbą bitów reprezentacji skwantowanych wartości próbek transformaty a ρ wykorzystano w pracach [Bobi04, Mila03, Lim07, Shin04] gdzie zaproponowano modele dla kodera standardu MPEG-4 AVC/H.264. Należy zaznaczyć, że modele te zostały wyrażone w zależności od parametru kwantyzacji QP . Model liniowy opisany wzorem 3.11 wykorzystano również w serii publikacji [Past06, Piet05a, Piet05b, Piet05c, Piet05d, Skar06] do kodowania sekwencji wizyjnych z tzw. regionami zainteresowania - ROI (ang. *region of interest*).

Bardzo podobne wnioski dotyczące zależności między liczbą bitów reprezentacji składowych luminancji i chrominancji a liczbą niezerowych próbek transformaty, wysunięte

na podstawie analizy histogramów wartości próbek transformaty, wykorzystano do budowy modeli dla wcześniejszych standardów: MPEG-2 i H.263 w serii publikacji [Łucz03a, Łucz04a, Łucz04b].

Ponieważ do wyznaczenia sygnału błędu predykcji w koderze referencyjnym należy wcześniej określić wartość Q (z powodu stosowania RDO), bezpośrednie zastosowanie liniowego modelu w przestrzeni ρ w koderze MPEG-4 AVC/H.264 jest utrudnione [He03, Lin08, Shin04]. Ponadto, nie sprawdza się on w przypadku modelowania obrazów typu B.

3.5. Model kodera w zależności od wartości parametru kwantyzacji QP

W standardzie MPEG-4 AVC/H.264 szerokość przedziałów kwantyzacji Q jest określana za pomocą parametru kwantyzacji QP , którego wartość umieszcza się w strumieniu bitowym (punkt 2.3.3). Ze względu na pośredni sposób określania wartości Q , w literaturze pojawiły się propozycje, w których wzory opisujące koder wyraża się w funkcji parametru kwantyzacji QP , a nie szerokości przedziałów kwantyzacji Q . Przykłady takich rozwiązań dla obrazów typu P można znaleźć w [Mila03, Ming07, Lim07, Shin04, Zhou08, Zhou09, Zhuo07]. Dodatkowo w pracach [Shin04, Lim07, Zhou08, Zhou09] zaproponowano modele dla obrazów typu I. Model wykładniczy (z wielomianem 2-go stopnia w wykładniku) dla zależności liczby bitów reprezentacji obrazu typu I od wartości QP przedstawiono w [Shin04]. Model wykładniczy rozszerzony o miarę złożoności treści obrazu typu I przedstawiono w [Zhou08, Zhou09]:

$$B = \alpha \cdot GradSOH \cdot e^{-\beta \cdot QP}, \quad (3.12)$$

gdzie:

- α i β to parametry modelu,
- B to liczba bitów reprezentacji obrazu typu I,
- $GradSOH$ to miara złożoności treści obrazu typu I,
- QP to parametr kwantyzacji.

Zaproponowana we wzorze 3.12 miara złożoności treści obrazu „GradSOH”, to iloczyn wartości Grad (wzór 3.13) i SOH (wzór 3.14) [Zhou09]:

$$\begin{aligned}
 Grad = & \frac{1}{W_Y \cdot K_Y} \sum_{i=0}^{W_Y-2} \sum_{j=0}^{K_Y-2} (|Y_{i,j} - Y_{i,j+1}| + |Y_{i,j} - Y_{i+1,j}|) \\
 & + \frac{1}{W_{Cb} \cdot K_{Cb}} \sum_{i=0}^{W_{Cb}-2} \sum_{j=0}^{K_{Cb}-2} (|Cb_{i,j} - Cb_{i,j+1}| + |Cb_{i,j} - Cb_{i+1,j}|), \\
 & + \frac{1}{W_{Cr} \cdot K_{Cr}} \sum_{i=0}^{W_{Cr}-2} \sum_{j=0}^{K_{Cr}-2} (|Cr_{i,j} - Cr_{i,j+1}| + |Cr_{i,j} - Cr_{i+1,j}|)
 \end{aligned} \tag{3.13}$$

$$SOH = \sum_{l=0}^{255} (\log_2 Hist_Y[l] + \log_2 Hist_Cb[l] + \log_2 Hist_Cr[l]), \tag{3.14}$$

gdzie:

- $Y_{i,j}, Cb_{i,j}, Cr_{i,j}$ to wartości próbek składowych odpowiednio Y, Cb i Cr obrazu na pozycji (i,j) ,
- W_Y, W_{Cb}, W_{Cr} to liczba wierszy próbek składowych odpowiednio Y, Cb i Cr obrazu,
- K_Y, K_{Cb}, K_{Cr} to liczba kolumn próbek składowych odpowiednio Y, Cb i Cr obrazu,
- $Hist_Y, Hist_Cb, Hist_Cr$ to histogramy składowych odpowiednio Y, Cb i Cr obrazu.

Natomiast w [Lim07] przedstawiono liniowy model zależności parametru QP od jakości zdekodowanego obrazu typu I (wyrażonej miarą PSNR [Bovi05, Doma98, Skar98]).

3.6. Podsumowanie

Analizując literaturę z zakresu modelowania hybrydowych koderów wizyjnych, można znaleźć szereg przykładów pokazujących, że dla wcześniejszych standardów kodowania sekwencji wizyjnych udało się opracować modele wyrażające liczbę bitów na wyjściu kodera w funkcji parametru kwantyzacji [Bai01, Chia97, Ding96, Hang97, He01d, Lee00, Lin98, Łucz01, Łucz02, Riba99]. Jednak ciągle brakuje dobrych modeli dla koderów standardu MPEG-4 AVC/H.264. Jednym z powodów jest to, że rozwiązania znane ze starszych technik

kodowania sekwencji wizyjnych, nie dają się bezpośrednio zaadaptować do koderów standardu MPEG-4 AVC/H.264 (głównie ze względu na wykorzystywanie w koderze optymalizacji „prędkość bitowa – zniekształcenie” (RDO)). W związku z tym trzeba stosować kolejne modele, które niestety wprowadzają dodatkowe błędy [Lee00, Li03, Liu07, Yi04]. Co więcej, duży stopień komplikacji zaawansowanych koderów wizyjnych, skutecznie utrudnia wyprowadzenie modeli matematycznych opierających się na analizie struktury kodera i wykorzystywanych w nim algorytmów. Większość proponowanych modeli dla takich koderów dotyczy wyłącznie obrazów typu P. Obecnie brakuje dokładnych modeli matematycznych opisujących obrazy pozostałych typów, a w szczególności obrazy typu B.

W następnych punktach autor przedstawi propozycje własnych modeli kodera wizyjnego standardu MPEG-4 AVC/H.264.

Rozdział 4

Modelowanie średniej liczby bitów potrzebnych do reprezentacji skompresowanego obrazu

4.1. Poszukiwany model

Rozdział przedstawia wyniki badań własnych, których celem było stworzenie ogólnego modelu matematycznego opisującego związku pomiędzy średnią liczbą bitów potrzebnych do reprezentacji skompresowanego obrazu w obrębie jednego ujęcia i szerokością przedziałów kwantyzacji Q . Poszukiwany model jest wyrażony zależnością:

$$B_{\text{śred}} = f(Q), \quad (4.1)$$

gdzie:

- $B_{\text{śred}}$ to średnia liczba bitów reprezentacji jednego obrazu pewnego typu (I, P lub B) w ramach jednego ujęcia,
- Q to szerokość przedziałów kwantyzacji.

Model $f(\cdot)$ zależy od treści kodowanej sekwencji wizyjnej, która zmienia się gwałtownie na granicach ujęć (w punktach cięcia). Co więcej, wartość $B_{\text{śred}}$ zależy od typu obrazu (I, P lub

B). Z tego powodu parametry modelu zostaną wyznaczone indywidualnie dla każdego typu obrazu w obrębie pojedynczego ujęcia.

Parametry modelu muszą być wyznaczone indywidualnie dla każdego ujęcia, gdyż przynajmniej część parametrów silnie zależy od treści sekwencji w każdym indywidualnym ujęciu. Im większa jest liczba parametrów modelu, tym więcej pomiarów należy wykonać dla estymacji tych parametrów. Dlatego model $f(\cdot)$ powinien być opisany możliwie małą liczbą parametrów.

Pomimo licznych prób i zastosowania różnych podejść do modelowania zaawansowanych koderów wizyjnych (patrz rozdział 3), nie udało się dotychczas opracować dla tych koderów dobrego modelu matematycznego (szczególnie dla obrazów typu I i B). Opracowanie takiego modelu jest trudne głównie ze względu na silną adaptację algorytmu kodowania i jego parametrów do treści kodowanej sekwencji wizyjnej. Dlatego autor podjął próbę opracowania modelu dla kodera jako całości, bez analizowania szczegółów bardzo skomplikowanego algorytmu kodowania. W tym celu zebrane zostały dane eksperymentalne dla szerokiej klasy sekwencji wizyjnych, a następnie starano się znaleźć model matematyczny do ich opisu.

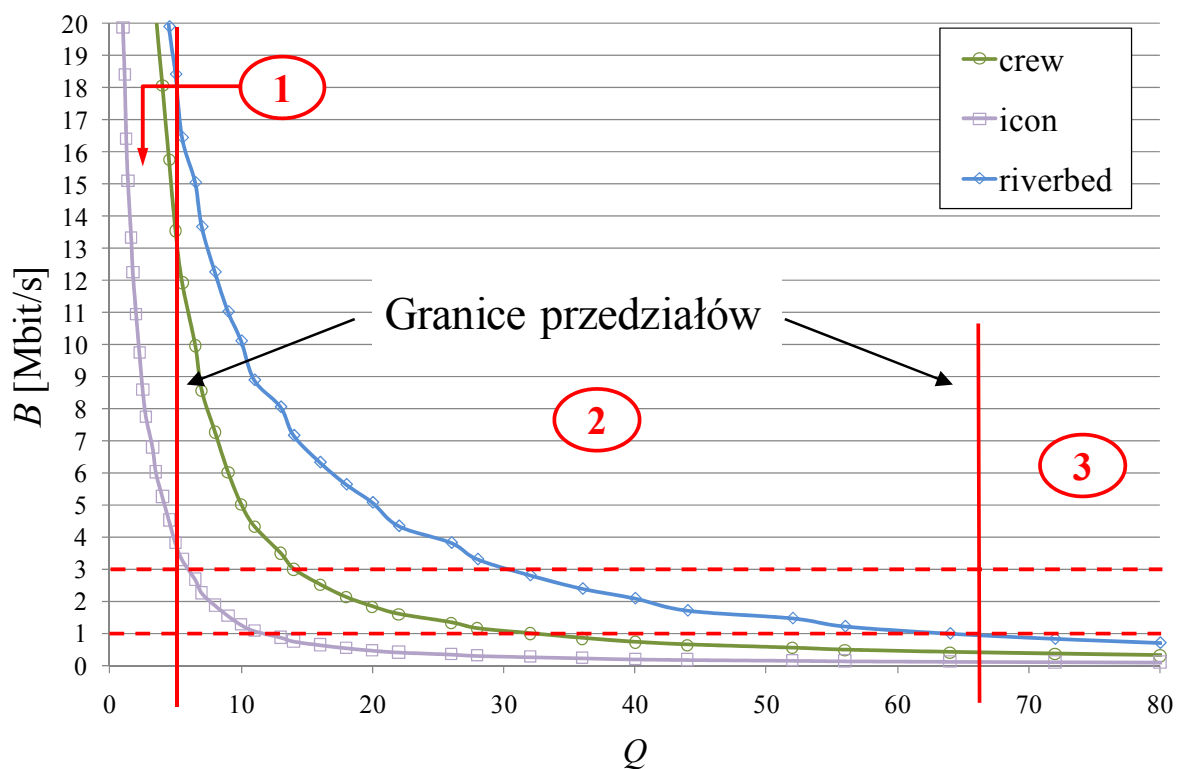
W literaturze dobrze znane są metody pozwalające na znalezienie funkcji, która przechodzi dokładnie przez punkty pomiarowe (np. interpolacja wielomianowa) [Bron07, Pres07]. Niestety zastosowanie takich metod w przypadku poszukiwania dopasowania do wielu punktów pomiarowych oznaczałoby konieczność wyznaczenia wartości wielu parametrów niezależnie dla każdego typu obrazu w ramach pojedynczego ujęcia. A to z kolei, mocno komplikowałoby, a może nawet czyniłoby niemożliwym, ewentualne zastosowanie takiego modelu w rozwiązaniach praktycznych. Dlatego autor zdecydował się na wykorzystanie metody dopasowania funkcji do danych eksperymentalnych przedstawionej w punkcie 1.3. Jest to takie samo podejście, które z sukcesem zastosowano dla koderów standardu MPEG-2 i H.263 w pracach [Doma02, Łucz01, Łucz03b].

Dobranie dokładnego modelu dla całego dozwolonego zakresu wartości szerokości przedziałów kwantyzacji Q , które udało się zrealizować dla starszych technik kodowania sekwencji wizyjnych [Doma02, Łucz02, Łucz03b], w przypadku MPEG-4 AVC/H.264, okazało się nieskuteczne (duże błędy aproksymacji), przez co nie zostało zaprezentowane w rozprawie. Z powodu wspomnianych dużych błędów aproksymacji, autor zdecydował się na podzielenie danych eksperymentalnych na 3 przedziały (odcinki) jak na rysunku 4.1. Granice przedziałów zostały dobrane tak (tabela 4.1), aby środkowy odcinek obejmował wartości Q odpowiadające użytecznemu zakresowi prędkości bitowej sekwencji wizyjnych

o rozmiarze 4CIF, tj. od 1 do 3 Mbit/s dla badanych sekwencji testowych. Należy zaznaczyć, że dla poszczególnych sekwencji wizyjnych, użytecznemu zakresowi prędkości bitowych strumienia danych odpowiada różny zakres wartości Q (patrz Dodatek B). Dlatego granice drugiego przedziału ustalono, z jednej strony jako najmniejszą wartość Q odpowiadającą prędkości 3 Mbit/s, a z drugiej strony jako największą wartość Q odpowiadającą prędkości 1 Mbit/s (rysunek 4.1).

Tabela 4.1. Granice przedziałów (odcinków)

Przedział (odcinek)	Zakres Q	Zakres QP
1	$\langle 0,625; 4,5 \rangle$	$\langle 0;17 \rangle$
2	$\langle 5; 64 \rangle$	$\langle 18;40 \rangle$
3	$\langle 72; 224 \rangle$	$\langle 41;51 \rangle$



Wykres 4.1. Przyjęte granice przedziałów na przykładzie 3 sekwencji wizyjnych. Dla przejrzystości wykresu obcięto zakres wartości B i Q

Następnie w każdym przedziale autor dobierał odpowiedni model matematyczny niezależnie od pozostałych przedziałów. Podobny podział wartości Q został przyjęty do badań w innych pracach, na przykład w [Boyc04, Liu07, Ma05, Wan06, Xie07, Yu05, Yuan05, Yuan06].

Jednak autorzy wymienionych prac ograniczali się do analizy kodera tylko dla wartości Q należących do przedziału drugiego.

Część wyników, zaprezentowanych w tym rozdziale, została już opublikowana przez autora w pracach [Graj09a, Graj09b].

4.2. Wyznaczanie modelu na podstawie danych eksperymentalnych

Tworzenie modelu można podzielić na dwa główne etapy:

- zaplanowanie i przeprowadzenie eksperymentów, łącznie z zebraniem odpowiednich danych eksperymentalnych,
- opracowanie zebranych danych eksperymentalnych, tj. wyznaczenie modelu i ocena jego dokładności.

Do przeprowadzenia badań autor wybrał reprezentatywną grupę 21 sekwencji testowych o rozdzielczości przestrzennej 704×576 punktów luminancji i rozdzielczości czasowej 25 lub 30 obrazów na sekundę (patrz Dodatek A). Każda sekwencja składała się ze 199 obrazów należących do pojedynczego ujęcia. Aby zebrać dane eksperymentalne do budowy modelu, wszystkie testowe sekwencje wizyjne zostały zakodowane ze wszystkimi możliwymi wartościami szerokości przedziałów kwantyzacji Q (jedno kodowanie sekwencji dla jednej, stałej szerokości przedziałów kwantyzacji Q). Do kodowania przyjęto strukturę grupy obrazów GOP typową dla zastosowań telewizyjnych (4P2B, rysunek 2.4). Dodatkowo, żeby dla danej szerokości przedziałów kwantyzacji Q i konkretnych ustawień pozostałych parametrów kodera, kodowanie powodowało możliwie najmniejsze straty jakości kompresowanej sekwencji, włączono optymalizację „prędkość bitowa – zniekształcenie” (RDO) [Sull98, Wieg01, Wieg03a]. Dla każdego takiego kodowania zebrano informacje dotyczące średniej liczby bitów potrzebnych do reprezentacji skompresowanego obrazu danego typu.

Do znalezienia wzoru matematycznego opisującego związku pomiędzy średnią liczbą bitów potrzebnych do reprezentacji skompresowanego obrazu w obrębie jednego ujęcia a szerokością przedziałów kwantyzacji Q , autor zastosował metodę dopasowania funkcji do danych eksperymentalnych przedstawioną w punkcie 1.3. Parametry modelu zostały

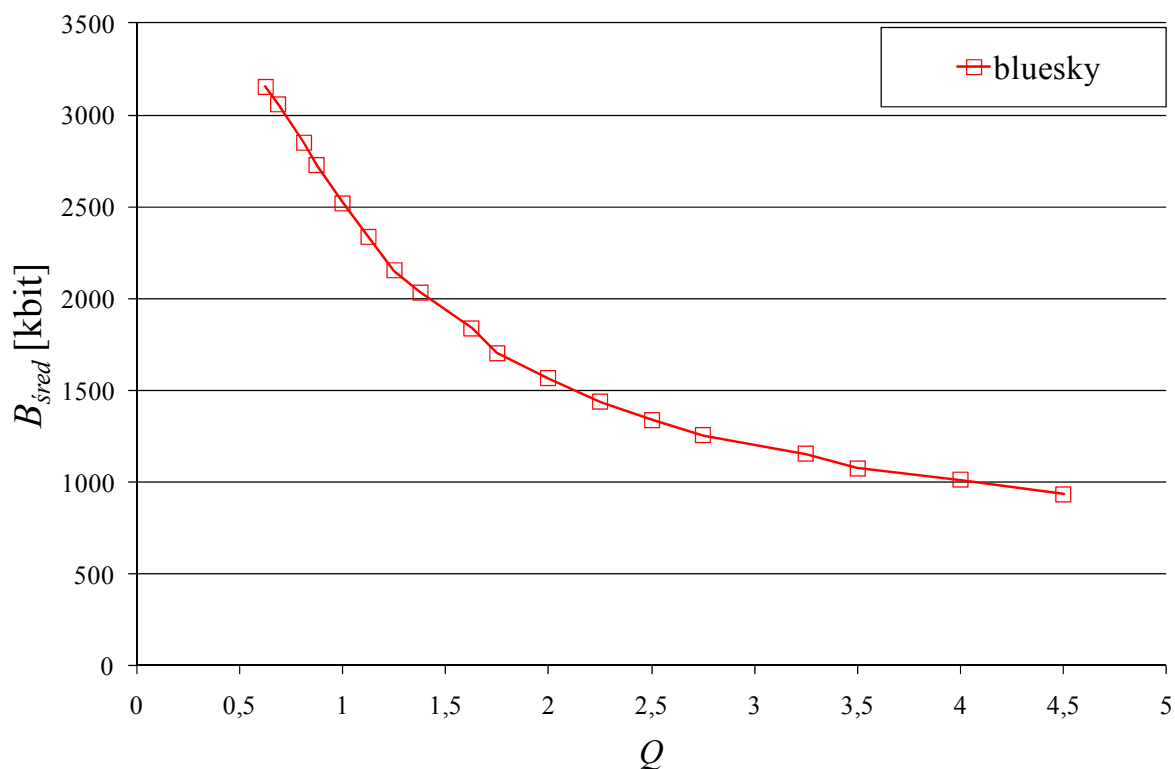
wyznaczone poprzez minimalizację maksymalnego względnego błędu aproksymacji metodą quasi-newtonowską [Chon08, Rao09, Ravi06].

W następnych punktach zostaną przedstawione szczegóły dotyczące drugiego etapu tworzenia modelu, czyli wyznaczenia modelu i oceny jego dokładności, przeprowadzone niezależnie dla danych z 3 przedziałów wartości Q .

4.3. Modelowanie kodera dla wartości Q z przedziału pierwszego (odcinek 1)

4.3.1. Wybór typu modelu

W pierwszym kroku metody dopasowania funkcji do danych eksperymentalnych (punkt 1.3), na podstawie wykresów tych danych, wnioskuje się o charakterze badanej zależności. Na wykresie 4.2 pokazano zależność średniej liczby bitów ($B_{\text{śred}}$) potrzebnej do reprezentacji skompresowanego obrazu typu I w zależności od szerokości przedziałów kwantyzacji Q z pierwszego przedziału dla przykładowej sekwencji wizyjnej. Wykresy dla pozostałych sekwencji testowych i typów kodowania obrazu są bardzo podobne do tego przykładowego, dlatego nie zostały umieszczone w rozprawie. Wnioskując o charakterze badanej zależności właśnie na podstawie wykresów takich jak 4.2, można domniemywać, że jest to krzywa, którą można opisać za pomocą funkcji wielomianowej, potęgowej lub wykładniczej.



Wykres 4.2. Średnia liczba bitów potrzebnych do reprezentacji skompresowanego obrazu typu I w sekwencji *bluesky* (odcinek 1)

Do rozstrzygnięcia o tym, którą ze wspomnianych funkcji użyć do modelowania danych, wykorzystano analizę współczynnika determinacji [Ciec07, Koro01]. W tabeli 4.2 podano wartości współczynników determinacji dla wspomnianych klas funkcji uśrednione po wszystkich badanych sekwencjach wizyjnych niezależnie dla obrazów typu I, P i B. Im wartość współczynnika determinacji jest bliższa jedności, tym model jest lepiej dopasowany do danych eksperymentalnych. Należy zaznaczyć, że zastosowanie wielomianów kolejnych wyższych rzędów nieznacznie poprawiało dokładność modelu, ale niestety zwiększało liczbę jego parametrów.

Tabela 4.2. Wartości współczynnika determinacji dla sprawdzanych klas funkcji uśrednione po badanych sekwencjach wizyjnych (odcinek 1)

Typ obrazu	Typ modelu		
	wielomianowy 3-go stopnia	potęgowy	wykładniczy
I	0,9985	0,9905	0,9542
P	0,9944	0,9911	0,9529
B	0,9934	0,9798	0,9719

Z analizy tabeli 4.2 wynika, że dla modelu wielomianowego 3-go stopnia i modelu potęgowego otrzymano bardzo zbliżone wyniki. Jednak trudno na tej podstawie wnioskować, który z nich będzie lepszy. Należy zauważyć, że model wielomianowy 3-go stopnia ma cztery parametry, a model potęgowy dwa. Dlatego, na tym etapie, autor odrzucił model wykładniczy, a do dalszej analizy pozostawił modele: wielomianowy 3-go stopnia i potęgowy. Dokładność aproksymacji oraz wartości parametrów tych dwóch modeli, wyznaczone niezależnie dla obrazów typu I, P i B, zostaną przedstawione w następnych punktach.

4.3.2. Model wielomianowy w przedziale pierwszym

4.3.2.1. Model dla obrazów typu I

Pierwszym modelem wybranym do aproksymacji danych z przedziału pierwszego jest funkcja wielomianowa 3-go stopnia postaci:

$$B_{\text{śred}}(Q, a, b, c, d) = a \cdot Q^3 + b \cdot Q^2 + c \cdot Q + d, \quad (4.2)$$

gdzie:

- a, b, c, d to parametry modelu,
- Q to szerokość przedziałów kwantyzacji,
- $B_{\text{śred}}$ to średnia liczba bitów reprezentacji obrazu typu I.

Parametry modelu zostały wyznaczone poprzez minimalizację maksymalnego względnego błędu aproksymacji metodą quasi-newtonowską (punkt 1.3):

$$\min_{a,b,c,d} \max_Q \frac{|B_x(Q) - B_{\text{śred}}(Q, a, b, c, d)|}{B_x(Q)} \cdot 100\%, \quad (4.3)$$

gdzie:

- a, b, c, d to parametry modelu,
- $B_x(Q)$ to zmierzona średnia liczba bitów reprezentacji obrazu typu I,
- $B_{\text{śred}}(Q, a, b, c, d)$ to wartość oszacowania średniej liczby bitów reprezentacji obrazu typu I modelem wielomianowym 3-go stopnia z 4 parametrami.

W tabeli 4.3 podano wartości parametrów modelu, maksymalny i średni względny błąd aproksymacji dla obrazów typu I dla wszystkich badanych sekwencji wizyjnych.

Tabela 4.3. Wartości parametrów modelu wielomianowego oraz maksymalny i średni względny błąd aproksymacji dla obrazów typu I (odcinek 1)

Nazwa sekwencji	Obrazy typu I					
	<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>	względny błąd dopasowania [%]	
					maksymalny	średni
basket	-74 498	788 528	-2 935 569	5 096 611	1,99	1,27
bluesky	-69 789	733 279	-2 695 014	4 547 544	2,71	1,53
bus	-66 619	674 935	-2 397 860	3 971 240	3,91	1,67
cheer	-69 375	721 548	-2 655 498	4 653 161	2,61	1,20
city	-78 329	799 567	-2 847 477	4 676 680	3,27	2,00
crew	-50 803	526 043	-1 978 779	3 529 029	4,18	1,73
flow	-77 447	832 782	-3 207 002	5 911 343	3,02	1,35
football	-34 142	374 119	-1 525 524	2 893 859	3,32	1,26
harbour	-82 754	844 293	-2 999 724	4 864 715	3,32	2,09
ice	-25 211	298 283	-1 298 037	2 443 776	2,63	1,34
icon	-25 043	271 134	-1 052 321	1 847 473	2,59	1,61
pedestrian	-35 054	386 966	-1 540 643	2 709 255	2,87	1,16
riverbed	-71 372	725 097	-2 594 278	4 265 139	4,43	1,80
rushhour	-37 240	412 600	-1 622 628	2 693 731	3,11	1,28
soccer	-62 793	644 583	-2 331 367	3 878 668	3,73	1,56
station2	-56 751	588 875	-2 167 480	3 633 529	3,64	2,16
stefan	-60 310	622 593	-2 278 977	3 878 070	3,54	1,45
sunflower	-54 073	565 861	-2 087 727	3 480 029	3,67	2,10
tractor	-65 339	676 456	-2 471 666	4 184 482	3,13	1,92
universal	-36 701	385 577	-1 449 329	2 465 382	4,09	1,60
warner	-38 582	413 650	-1 593 961	2 741 470	3,66	1,33

Z tabeli 4.3 wynika, że dokładność zaproponowanego modelu jest wysoka (średni względny błąd dopasowania dla większości badanych sekwencji wizyjnych poniżej 2%).

Biorąc pod uwagę możliwości zastosowania zaproponowanego modelu do sterowania koderem wizyjnym, autor podjął dalsze badania mające na celu uproszczenie modelu poprzez redukcję liczby jego parametrów. Analizując wartości współczynników korelacji liniowej [Ciec07, Koro01] dla poszczególnych par parametrów *a*, *b*, *c* i *d* okazało się, że są one bliskie jedności (patrz tabela 4.4).

Tabela 4.4. Wartości współczynnika korelacji liniowej dla poszczególnych par parametrów modelu wielomianowego dla obrazów typu I (odcinek 1)

Parametry modelu	a	b	c	d
a	1	-0,997	0,9813	-0,9376
b	-0,997	1	-0,993	0,9597
c	0,9813	-0,993	1	-0,985
d	-0,9376	0,9597	-0,985	1

Dlatego, wykorzystując regresję liniową do wyznaczenia zależności między kolejnymi parami parametrów [Ciec07, Koro01, Pres07], został opracowany model wielomianowy tylko z jednym parametrem d postaci:

$$B_{\text{śred}}(Q, d) = (a_1 d + a_2) \cdot Q^3 + (b_1 d + b_2) \cdot Q^2 + (c_1 d + c_2) \cdot Q + d, \quad (4.4)$$

gdzie:

- d to parametr modelu,
- $B_{\text{śred}}$ to średnia liczba bitów reprezentacji obrazu typu I,
- Q to szerokość przedziałów kwantyzacji,
- a_1, a_2, b_1, b_2, c_1 i c_2 to uniwersalne stałe modelu dla obrazów typu I.

Parametr d modelu został wyznaczony, analogicznie jak wcześniej, poprzez minimalizację maksymalnego względnego błędu aproksymacji.

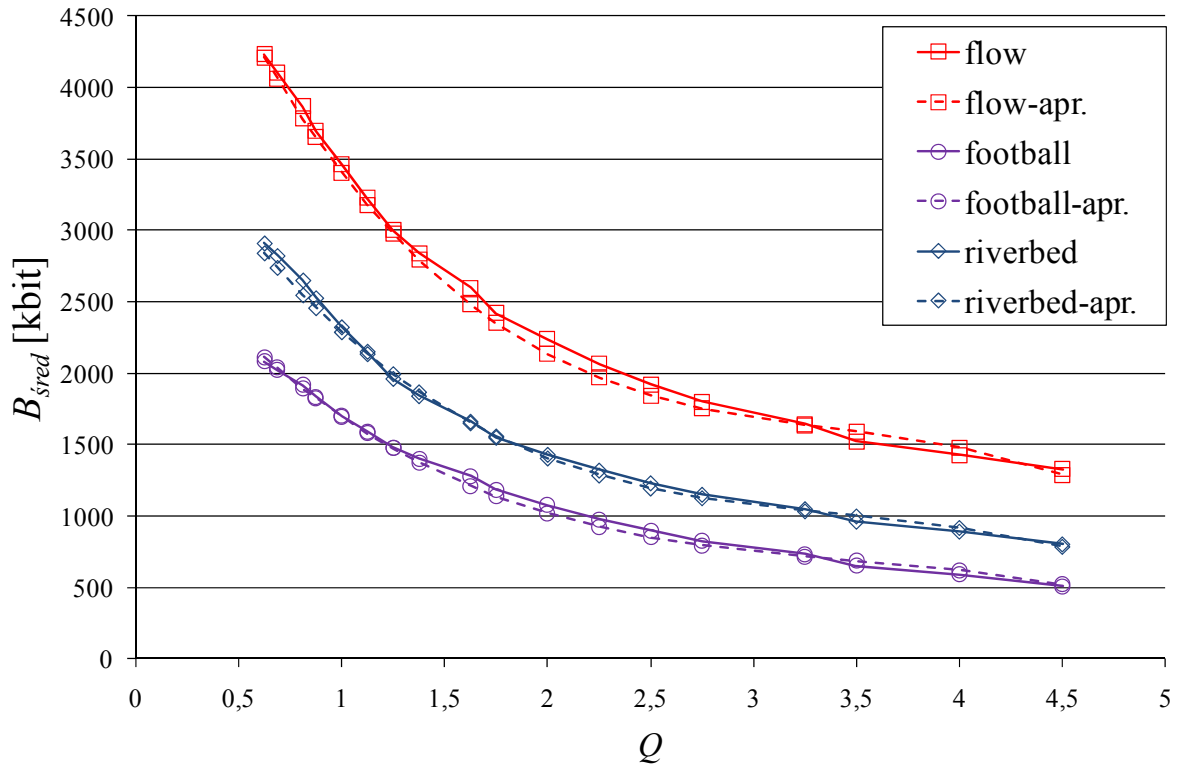
W tabeli 4.5 podano wartości stałych dla obrazów typu I, natomiast w tabeli 4.6 zebrano wartości parametru d , maksymalny i średni względny błąd aproksymacji dla wszystkich badanych sekwencji wizyjnych. Ponadto, na wykresie 4.3 przedstawiono krzywe eksperymentalne i ich oszacowania dla 3 przykładowych sekwencji wizyjnych.

Tabela 4.5. Wartości stałych uproszczonego modelu wielomianowego dla obrazów typu I (odcinek 1)

Stała	Wartość
$a1$	-0,017
$a2$	6385,8485
$b1$	0,1702
$b2$	-41606,227
$c1$	-0,5875
$c2$	1451,112

Tabela 4.6. Wartości parametru uproszczonego modelu wielomianowego oraz maksymalny i średni względny błąd aproksymacji dla obrazów typu I (odcinek 1)

Nazwa sekwencji	Obrazy typu I		
	d	względny błąd dopasowania [%]	
		maksymalny	średni
basket	5 125 019	3,51	1,27
bluesky	4 514 954	4,40	1,40
bus	4 021 171	5,10	2,50
cheer	4 851 385	3,97	2,00
city	4 576 801	4,03	1,93
crew	3 670 807	4,33	2,34
flow	6 077 910	4,65	2,47
football	3 060 834	5,44	2,90
harbour	4 632 519	3,16	1,62
ice	2 424 888	7,62	4,51
icon	2 118 733	13,51	7,93
pedestrian	2 654 235	4,93	3,12
riverbed	4 109 293	3,86	2,02
rushhour	2 438 183	9,32	4,74
soccer	3 835 366	3,40	1,49
station2	3 533 381	3,77	2,13
stefan	3 900 147	3,17	1,51
sunflower	3 412 873	4,35	2,03
tractor	4 167 942	3,21	1,91
universal	2 565 351	6,20	2,94
warner	2 689 461	3,74	2,45



Wykres 4.3. Krzywe eksperymentalne (linia ciągła) i oszacowane uproszczonym modelem wielomianowym (linia przerywana) dla obrazów typu I dla 3 przykładowych sekwencji wizyjnych (odcinek 1)

Dokładność zaproponowanego modelu wielomianowego 3-go stopnia z czterema parametrami, dla obrazów typu I, jest wysoka (średni względny błąd aproksymacji dla większości sekwencji wizyjnych poniżej 2%). Uproszczenie modelu, dokonane poprzez zredukowanie liczby parametrów, pogorszyło dokładność oszacowania dla obrazów typu I średnio o około 1 p.p. (punkt procentowy). Wyjątkiem jest sekwencja *icon*, z błędem dopasowania większym o 6 p.p.

4.3.2.2. Model dla obrazów typu P

Dokładność oraz wartości parametrów modelu wielomianowego (wzór 4.2) wyznaczone dla obrazów typu P, dla wszystkich sekwencji testowych zebrano w tabeli 4.7. Wynika z niej, że średni względny błąd aproksymacji dla większości sekwencji wizyjnych jest poniżej 4%.

Tabela 4.7. Wartości parametrów modelu wielomianowego oraz maksymalny i średni względny błąd aproksymacji dla obrazów typu P (odcinek 1)

Nazwa sekwencji	Obrazy typu P					
	<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>	względny błąd dopasowania [%]	
					maksymalny	średni
basket	-88 948	860 837	-2 876 191	4 266 012	8,16	4,07
bluesky	-46 036	476 693	-1 735 736	2 686 322	5,84	2,42
bus	-66 938	657 986	-2 258 494	3 393 741	8,37	4,00
cheer	-73 554	716 627	-2 443 791	3 841 036	7,07	3,55
city	-97 086	936 801	-3 108 750	4 404 399	10,34	4,99
crew	-65 961	658 230	-2 325 196	3 697 993	6,71	3,02
flow	-60 046	599 998	-2 126 656	3 481 153	5,91	2,70
football	-46 430	485 528	-1 825 632	3 008 844	6,14	2,47
harbour	-106 067	1 019 269	-3 349 150	4 799 357	8,63	4,26
ice	-39 662	430 684	-1 656 273	2 540 525	8,10	2,66
icon	-37 511	377 729	-1 318 052	1 890 589	9,24	3,60
pedestrian	-55 163	559 989	-1 958 832	2 705 624	12,21	4,33
riverbed	-93 358	914 183	-3 078 187	4 559 223	7,15	3,25
rushhour	-53 255	552 252	-1 979 320	2 782 869	10,57	3,51
soccer	-79 616	780 935	-2 641 688	3 817 822	8,95	4,02
station2	-53 368	550 676	-1 971 936	2 715 523	16,86	4,84
stefan	-60 782	601 161	-2 104 837	3 305 491	7,20	3,24
sunflower	-59 434	605 262	-2 105 358	2 799 166	14,74	4,96
tractor	-81 953	799 164	-2 708 771	3 998 390	8,53	3,81
universal	-46 018	463 829	-1 635 401	2 469 542	7,27	3,11
warner	-50 048	510 054	-1 825 870	2 779 037	6,93	2,72

Podobnie jak dla obrazów typu I, między poszczególnymi parami parametrów *a*, *b*, *c* i *d* modelu dla obrazów typu P istnieją silne zależności liniowe (patrz tabela 4.8).

Tabela 4.8. Wartości współczynnika korelacji liniowej dla poszczególnych par parametrów modelu wielomianowego dla obrazów typu P (odcinek 1)

Parametry modelu	<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>
<i>a</i>	1	-0,999	0,9937	-0,9632
<i>b</i>	-0,999	1	-0,9974	0,9648
<i>c</i>	0,9937	-0,9974	1	-0,9743
<i>d</i>	-0,9632	0,9648	-0,9743	1

Dlatego dla obrazów typu P, autor także zdecydował się na uproszczenie modelu do jednego parametru. Stałe uproszczonego modelu wielomianowego (wzór 4.4) zostały umieszczone w tabeli 4.9. Natomiast wartości parametru *d*, maksymalny i średni względny błąd

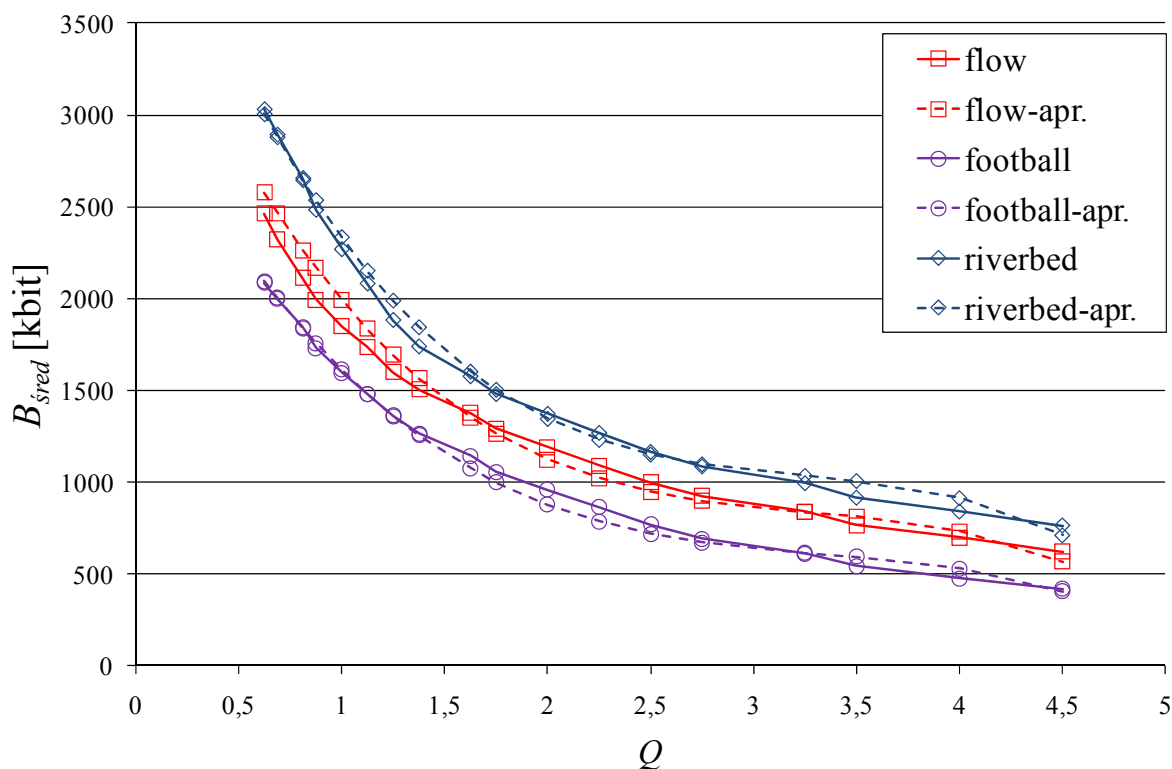
aproxymacji dla wszystkich badanych sekwencji wizyjnych zebrano w tabeli 4.10. Ponadto, na wykresie 4.4 przedstawiono krzywe eksperymentalne i oszacowane uproszczonym modelem wielomianowym dla 3 przykładowych sekwencji wizyjnych.

Tabela 4.9. Wartości stałych uproszczonego modelu wielomianowego dla obrazów typu P (odcinek 1)

Stała	Wartość
<i>a1</i>	-0,0241
<i>a2</i>	15595,2752
<i>b1</i>	0,2198
<i>b2</i>	-86498,6983
<i>c1</i>	-0,6742
<i>c2</i>	5666,28

Tabela 4.10. Wartości parametru uproszczonego modelu wielomianowego oraz maksymalny i średni względny błąd aproxymacji dla obrazów typu P (odcinek 1)

Nazwa sekwencji	Obrazy typu P		
	<i>d</i>	względny błąd dopasowania [%]	
		maksymalny	średni
basket	4 378 794	8,45	4,09
bluesky	2 890 723	8,76	4,39
bus	3 512 385	7,72	4,16
cheer	4 270 940	11,00	5,55
city	4 031 623	9,53	4,66
crew	3 931 911	7,58	3,92
flow	3 951 239	11,15	5,44
football	3 219 691	9,19	3,82
harbour	4 646 116	8,35	3,87
ice	2 329 092	11,93	8,75
icon	2 166 496	20,90	11,62
pedestrian	2 395 037	10,51	4,54
riverbed	4 601 786	9,15	3,53
rushhour	2 402 695	12,25	6,54
soccer	3 664 507	8,33	3,74
station2	2 160 822	19,57	10,71
stefan	3 550 574	7,81	4,55
sunflower	2 279 884	16,80	7,09
tractor	3 877 690	7,83	3,94
universal	2 739 873	12,96	6,45
warner	2 908 738	7,82	3,54



Wykres 4.4. Krzywe eksperymentalne (linia ciągła) i oszacowane uproszczonym modelem wielomianowym (linia przerywana) dla obrazów typu P dla 3 przykładowych sekwencji wizyjnych (odcinek 1)

Dokładność oszacowania danych zaproponowanym modelem wielomianowym 3-go stopnia z czterema parametrami dla obrazów typu P jest wysoka (średni względny błąd aproksymacji poniżej 4% dla większości sekwencji testowych). Uproszczenie modelu, dokonane poprzez zredukowanie liczby parametrów, pogorszyło dokładność oszacowania dla obrazów typu P średnio o około 2 punkty procentowe. Wyjątek stanowią sekwencje *icon* i *station2*, dla których średni względny błąd aproksymacji dla uproszczonego modelu wielomianowego jest wyższy o około 6 p.p.

4.3.2.3. Model dla obrazów typu B

Dokładność oraz wartości parametrów modelu wielomianowego 3-go stopnia z czterema parametrami (wzór 4.2), wyznaczone dla obrazów typu B, dla wszystkich sekwencji testowych zebrano w tabeli 4.11. Podobnie, jak dla obrazów typu P, średni względny błąd aproksymacji dla większości sekwencji wizyjnych jest poniżej 4%.

Tabela 4.11. Wartości parametrów modelu wielomianowego oraz maksymalny i średni względny błąd aproksymacji dla obrazów typu B (odcinek 1)

Nazwa sekwencji	Obrazy typu B					
	<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>	względny błąd dopasowania [%]	
					maksymalny	średni
basket	-49 878	511 357	-1 879 951	3 127 561	5,27	2,29
bluesky	-44 698	465 386	-1 679 883	2 406 080	7,87	3,83
bus	-37 858	397 433	-1 511 620	2 530 802	5,04	2,03
cheer	-42 019	436 743	-1 658 562	2 934 648	4,86	2,18
city	-46 704	491 374	-1 874 949	3 213 243	3,78	2,24
crew	-52 819	542 036	-2 002 965	3 306 357	5,63	2,38
flow	-49 836	508 591	-1 858 463	3 057 536	5,46	2,31
football	-40 523	430 114	-1 649 704	2 688 629	5,97	2,29
harbour	-62 908	631 635	-2 246 885	3 636 150	5,23	3,26
ice	-34 542	399 899	-1 598 233	2 360 474	6,11	3,38
icon	-27 234	281 226	-1 015 066	1 497 862	6,91	3,36
pedestrian	-40 266	446 507	-1 702 444	2 408 824	7,86	4,97
riverbed	-88 289	858 267	-2 873 811	4 256 696	7,63	3,81
rushhour	-50 532	529 592	-1 896 075	2 560 980	10,94	5,08
soccer	-49 039	509 310	-1 880 965	3 016 784	4,62	2,72
station2	-30 172	348 627	-1 418 069	2 127 330	11,19	6,45
stefan	-39 007	411 708	-1 559 165	2 599 103	4,85	2,17
sunflower	-37 185	412 548	-1 562 797	2 143 622	8,81	5,77
tractor	-44 839	482 381	-1 844 222	2 932 444	4,89	2,73
universal	-37 155	384 265	-1 393 603	2 084 314	6,99	3,38
warner	-45 193	463 306	-1 677 214	2 563 872	6,54	3,23

Analogicznie jak dla obrazów typów I i P, między poszczególnymi parami parametrów *a*, *b*, *c* i *d* modelu istnieją silne zależności liniowe (patrz tabela 4.12). Z tego powodu, również dla obrazów typu B, autor zdecydował się na uproszczenie modelu do tylko jednego parametru – *d*.

Tabela 4.12. Wartości współczynnika korelacji liniowej dla poszczególnych par parametrów modelu wielomianowego dla obrazów typu B (odcinek 1)

Parametry modelu	<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>
<i>a</i>	1	-0,996	0,9788	-0,8914
<i>b</i>	-0,996	1	-0,9912	0,8913
<i>c</i>	0,9788	-0,9912	1	-0,9213
<i>d</i>	-0,8914	0,8913	-0,9213	1

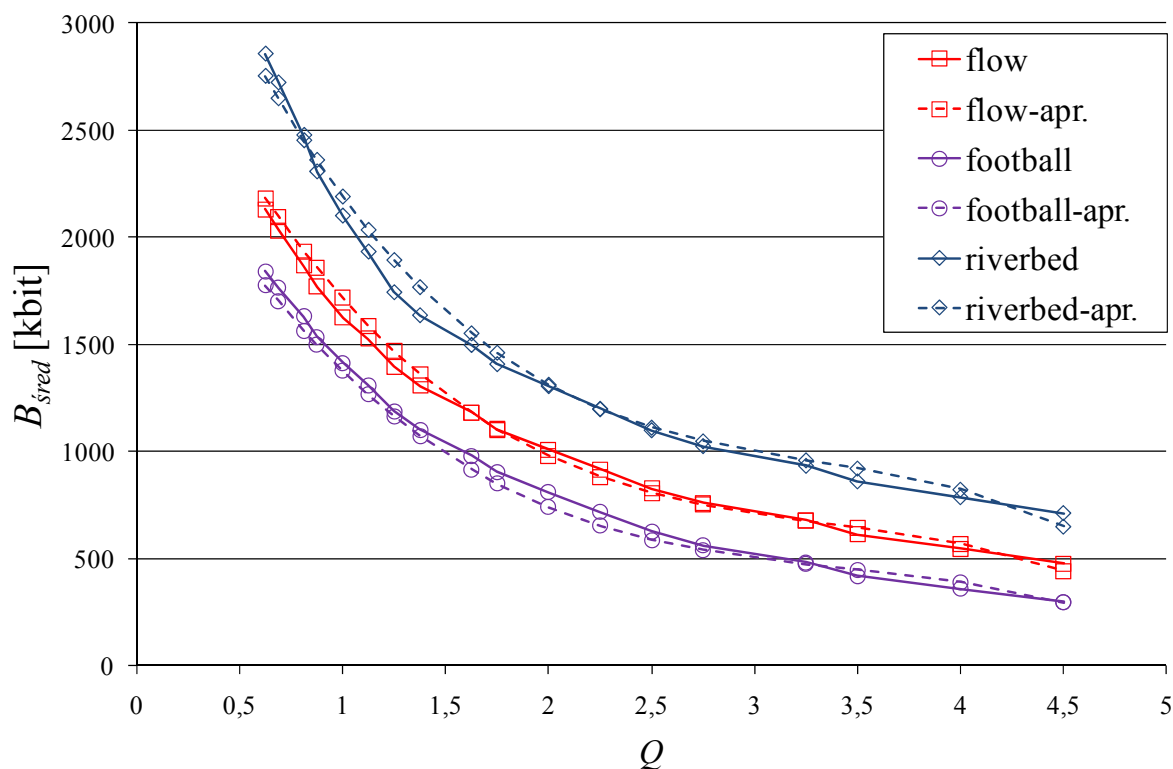
Stałe uproszczonego modelu wielomianowego (wzór 4.4) zostały umieszczone w tabeli 4.13. Natomiast wartości parametru d , maksymalny i średni względny błąd aproksymacji dla wszystkich badanych sekwencji wizyjnych zebrano w tabeli 4.14. Dodatkowo, na wykresie 4.5 przedstawiono krzywe eksperymentalne i ich oszacowania dla 3 przykładowych sekwencji wizyjnych.

Tabela 4.13. Wartości stałych uproszczonego modelu wielomianowego dla obrazów typu B (odcinek 1)

Stała	Wartość
$a1$	-0,0188
$a2$	4732,1641
$b1$	0,1729
$b2$	9943,7422
$c1$	-0,5618
$c2$	-230247,155

Tabela 4.14. Wartości parametru uproszczonego modelu wielomianowego oraz maksymalny i średni względny błąd aproksymacji dla obrazów typu B (odcinek 1)

Nazwa sekwencji	Obrazy typu B		
	d	względny błąd dopasowania [%]	
		maksymalny	średni
basket	3 350 846	7,37	3,51
bluesky	2 312 136	8,00	2,92
bus	2 780 435	7,34	4,26
cheer	3 324 337	9,95	4,99
city	3 368 607	6,46	3,10
crew	3 341 142	6,66	3,12
flow	3 262 700	7,12	3,53
football	2 690 822	8,93	4,49
harbour	3 747 318	7,38	3,41
ice	1 980 431	16,51	11,58
icon	1 959 864	31,31	18,88
pedestrian	2 001 551	13,14	7,72
riverbed	4 061 325	8,68	4,05
rushhour	2 152 026	13,62	6,60
soccer	3 062 766	5,84	2,92
station2	1 816 207	24,81	18,45
stefan	2 886 283	9,83	4,51
sunflower	1 834 127	19,20	10,93
tractor	2 764 190	7,50	4,65
universal	2 303 471	12,83	6,75
warner	2 636 918	6,03	3,24



Wykres 4.5. Krzywe eksperymentalne (linia ciągła) i oszacowane uproszczonym modelem wielomianowym (linia przerywana) dla obrazów typu B dla 3 przykładowych sekwencji wizyjnych (odcinek 1)

Dokładność zaproponowanego modelu wielomianowego 3-go stopnia z czterema parametrami dla obrazów typu B jest również wysoka (średni względny błąd aproksymacji poniżej 4%). Uproszczenie modelu, dokonane poprzez zredukowanie liczby parametrów, pogorszyło dokładność oszacowania dla badanego typu obrazów średnio o około 3 punkty procentowe. Wyjątek stanowią sekwencje *ice*, *icon*, *station2* i *sunflower*, dla których średni względny błąd aproksymacji dla uproszczonego modelu jest wyższy odpowiednio o około 8, 15, 12,5 i 6 p.p.

4.3.2.4. Weryfikacja dokładności uproszczonego modelu wielomianowego w przedziale pierwszym

Dla uproszczonego modelu wielomianowego wyznaczono wartości uniwersalnych stałych niezależnie dla obrazów typu I, P i B, które zebrano w tabeli 4.15. Stałe te zostały wyznaczone dla szerokiej klasy sekwencji wizyjnych.

Tabela 4.15. Wartości stałych uproszczonego modelu wielomianowego (odcinek 1)

Stała	Typ obrazu		
	I	P	B
$a1$	-0,017	-0,0241	-0,0188
$a2$	6385,8485	15595,2752	4732,1641
$b1$	0,1702	0,2198	0,1729
$b2$	-41606,227	-86498,6983	9943,7422
$c1$	-0,5875	-0,6742	-0,5618
$c2$	1451,112	5666,28	-230247,155

Aby zweryfikować dokładność uproszczonego modelu wielomianowego przeprowadzono dodatkowy eksperyment z wykorzystaniem zbioru kontrolnego, który składał się z 8 sekwencji wizyjnych formatu 4CIF. Sekwencje te nie występowały w zbiorze testowym użytym do wyznaczenia wartości uniwersalnych stałych. Dodatkowy eksperyment polegał na wyznaczeniu wartości parametru d oraz maksymalnego i średniego względnego błędu aproksymacji danych eksperymentalnych otrzymanych dla kontrolnego zbioru sekwencji wizyjnych z wykorzystaniem uniwersalnych stałych z tabeli 4.15. Wyniki tego eksperymentu zebrano w tabeli 4.16.

Tabela 4.16. Wartości parametru modelu oraz maksymalny i średni względny błąd aproksymacji danych z wykorzystaniem uproszczonego modelu wielomianowego dla obrazów typu I, P i B dla kontrolnego zbioru sekwencji wizyjnych (odcinek 1)

Nazwa sekwencji	Obrazy typu I			Obrazy typu P			Obrazy typu B		
	$d \cdot 10^3$	względny błąd dopasowania [%]		$d \cdot 10^3$	względny błąd dopasowania [%]		$d \cdot 10^3$	względny błąd dopasowania [%]	
		maks.	średni		maks.	średni		maks.	średni
carpark	3 879	3.46	2.17	2 322	16.17	10.85	1 855	32.09	25.52
night	4 363	4.52	2.10	3 852	7.46	4.11	3 219	6.87	2.86
outdorr	3 681	3.74	2.18	2 419	16.81	10.09	1 862	32.47	26.16
raven	2 988	5.37	2.51	2 854	9.64	5.96	2 349	13.30	7.53
shuttle	2 274	8.87	4.50	2 144	14.51	6.76	1 843	17.78	8.34
spincalendar	5 371	2.76	1.62	3 713	11.83	5.16	3 130	13.05	7.91
toys	3 574	4.08	2.64	2 683	16.11	11.00	2 202	19.96	14.68
vintagecar	3 561	5.32	2.42	3 280	7.93	4.00	2 343	15.17	9.55

Wyniki eksperymentu potwierdzają, że zaproponowany uproszczony model wielomianowy z zestawem uniwersalnych stałych dobrze aproksymuje dane eksperymentalne uzyskane dla sekwencji spoza zbioru testowego dla obrazów typu I, P i B. Średni względny błąd aproksymacji dla badanego modelu wynosi odpowiednio około 3, 7 i 9%. Wyjątkiem są

sekwencje *carpark* i *outdoor*, dla których dla obrazów typu B otrzymano średni względny błąd aproksymacji na poziomie 26%.

4.3.2.5. Podsumowanie dotyczące modelu wielomianowego w przedziale pierwszym

Dla danych eksperymentalnych otrzymanych dla wartości Q z przedziału pierwszego, autor zaproponował model wielomianowy 3-go stopnia z jednym parametrem d wyrażający średnią liczbę bitów reprezentacji skompresowanego obrazu danego typu w zależności od wartości Q . Dokładność uproszczonego modelu wielomianowego została zweryfikowana na dodatkowym zbiorze sekwencji wizyjnych. Zaproponowanym modelem można aproksymować dane dla szerokiej klasy sekwencji wizyjnych ze średnim względnym błędem dopasowania około 3, 6 i 7% odpowiednio dla obrazów typu I, P i B.

4.3.3. Model potęgowy w przedziale pierwszym

4.3.3.1. Model dla obrazów typu I

Drugim modelem wybranym do opisu danych eksperymentalnych otrzymanych dla wartości Q z przedziału pierwszego jest funkcja potęgowa postaci:

$$B_{\text{śred}}(Q, a, b) = a \cdot Q^b, \quad (4.5)$$

gdzie:

- a, b to parametry modelu,
- Q to szerokość przedziałów kwantyzacji,
- $B_{\text{śred}}$ to średnia liczba bitów reprezentacji obrazu typu I.

Parametry modelu potęgowego zostały wyznaczone poprzez minimalizację maksymalnego względnego błędu aproksymacji:

$$\min_{a,b} \max_Q \frac{|B_x(Q) - B_{\text{śred}}(Q, a, b)|}{B_x(Q)} \cdot 100\%, \quad (4.6)$$

gdzie:

- a, b to parametry modelu,

- $B_x(Q)$ to zmierzona średnia liczba bitów reprezentacji obrazów typu I,
- $B_{\text{śred}}(Q,a,b)$ to wartość oszacowania średniej liczby bitów reprezentacji obrazu typu I modelem potęgowym z 2 parametrami.

W tabeli 4.17 podano wartości parametrów modelu potęgowego, maksymalny i średni względny błąd aproksymacji dla obrazów typu I, dla wszystkich sekwencji testowych.

Tabela 4.17. Wartości parametrów modelu potęgowego oraz maksymalny i średni względny błąd aproksymacji dla obrazów typu I (odcinek 1)

Nazwa sekwencji	Obrazy typu I			
	a	b	względny błąd dopasowania [%]	
			maksymalny	średni
basket	2 775 466	-0,6016	3,47	2,28
bluesky	2 433 689	-0,6243	3,44	2,05
bus	2 110 163	-0,6154	3,34	1,76
cheer	2 561 774	-0,5893	3,56	2,08
city	2 446 658	-0,6094	3,26	1,79
crew	1 956 423	-0,6650	6,49	3,44
flow	3 329 032	-0,5938	3,96	2,24
football	1 633 669	-0,7117	9,69	5,45
harbour	2 503 583	-0,6199	3,35	1,91
ice	1 320 889	-0,7982	12,13	6,89
icon	1 001 750	-0,7023	5,91	3,03
pedestrian	1 436 462	-0,7466	8,56	4,77
riverbed	2 230 582	-0,6504	4,05	1,93
rushhour	1 352 001	-0,8132	9,03	5,20
soccer	2 042 944	-0,6443	4,07	2,01
station2	1 907 634	-0,6753	5,18	2,61
stefan	2 075 082	-0,6393	4,19	2,07
sunflower	1 845 153	-0,6821	3,78	1,80
tractor	2 256 600	-0,6407	3,51	1,51
universal	1 303 092	-0,7054	6,29	3,34
warner	1 446 695	-0,7347	7,54	4,11

Wynika z niej, że dokładność zaproponowanego modelu dla obrazów typu I jest wysoka (średni względny błąd aproksymacji dla większości sekwencji testowych poniżej 4%). Wyjątkiem są sekwencje *football*, *ice* i *rushhour*, dla których średni względny błąd aproksymacji wynosi odpowiednio 5,45, 6,89 i 5,20%.

Z powodu dużej dokładności modelu i związanej z tym trudności w rozróżnieniu na wykresie krzywych eksperymentalnych od ich oszacowań oraz ograniczonych rozmiarów rozprawy, autor nie zamieścił w pracy wykresów dla modelu potęgowego.

4.3.3.2. Model dla obrazów typu P

Dokładność oraz wartości parametrów modelu potęgowego (wzór 4.5) wyznaczone dla obrazów typu P, dla wszystkich sekwencji testowych zebrano w tabeli 4.18.

Tabela 4.18. Wartości parametrów modelu potęgowego oraz maksymalny i średni względny błąd aproksymacji dla obrazów typu P (odcinek 1)

Nazwa sekwencji	Obrazy typu P			
	a	b	względny błąd dopasowania [%]	
			maksymalny	średni
basket	2 095 865	-0,7055	2,79	1,43
bluesky	1 308 554	-0,8260	6,65	3,81
bus	1 674 260	-0,7781	4,62	2,64
cheer	1 994 563	-0,6738	3,40	1,92
city	2 062 391	-0,7952	4,20	2,29
crew	1 901 186	-0,7421	5,33	3,09
flow	1 847 663	-0,6948	4,28	2,50
football	1 547 865	-0,8142	8,82	5,22
harbour	2 293 226	-0,7223	3,77	1,61
ice	1 170 736	-1,0494	14,56	9,01
icon	862 583	-0,9011	6,13	3,74
pedestrian	1 163 690	-1,0180	8,35	5,12
riverbed	2 234 666	-0,7041	2,63	1,27
rushhour	1 198 202	-1,0353	9,93	6,05
soccer	1 802 225	-0,7848	3,29	1,83
station2	1 136 246	-1,2203	16,11	9,97
stefan	1 690 542	-0,7648	6,02	3,49
sunflower	1 146 599	-1,1201	9,48	5,83
tractor	1 951 542	-0,7830	4,92	2,36
universal	1 193 729	-0,8120	5,50	3,23
warner	1 334 996	-0,8376	6,74	4,10

Analiza danych z tej tabeli prowadzi do wniosku, że średni względny błąd aproksymacji dla większości sekwencji wizyjnych jest poniżej 6%. Wyjątkiem są sekwencje *ice*, *rushhour* i *station2*, dla których średni względny błąd aproksymacji jest wyższy i wynosi odpowiednio 9,01, 6,05 i 9,97%.

4.3.3.3. Model dla obrazów typu B

Dokładność oraz wartości parametrów modelu potęgowego z dwoma parametrami (wzór 4.5), wyznaczone dla obrazów typu B, dla wszystkich sekwencji testowych zebrano w tabeli 4.19.

Tabela 4.19. Wartości parametrów modelu potęgowego oraz maksymalny i średni względny błąd aproksymacji dla obrazów typu B (odcinek 1)

Nazwa sekwencji	Obrazy typu B			
	a	b	względny błąd dopasowania [%]	
			maksymalny	średni
basket	1 646 744	-0,7452	6,48	3,84
bluesky	1 076 704	-1,0072	10,44	6,08
bus	1 309 796	-0,8264	9,83	5,97
cheer	1 622 550	-0,7206	7,42	4,47
city	1 703 359	-0,7738	8,83	5,28
crew	1 721 883	-0,7920	8,24	4,92
flow	1 593 163	-0,7570	6,70	4,00
football	1 344 176	-0,9195	12,37	7,61
harbour	1 913 007	-0,7201	5,22	2,87
ice	1 002 736	-1,2620	19,22	12,60
icon	693 402	-0,9407	9,53	5,78
pedestrian	968 775	-1,2306	16,69	10,22
riverbed	2 088 319	-0,7036	2,40	1,16
rushhour	1 054 943	-1,1516	12,40	7,47
soccer	1 511 054	-0,8075	7,68	4,58
station2	958 401	-1,6460	35,36	22,28
stefan	1 354 283	-0,7877	7,83	4,36
sunflower	889 163	-1,4409	21,22	12,71
tractor	1 403 780	-0,8987	10,20	6,09
universal	983 641	-0,9250	9,42	5,73
warner	1 244 235	-0,8867	9,43	5,67

Średni względny błąd aproksymacji dla większości sekwencji wizyjnych jest poniżej 6%. Wyjątkiem są sekwencje *ice*, *pedestrian* i *sunflower* ze średnim względnym błędem aproksymacji powyżej 10% oraz *station2* ze średnim względnym błędem powyżej 22%.

4.3.3.4. Podsumowanie dotyczące modelu potęgowego w przedziale pierwszym

Dla wartości Q z przedziału pierwszego autor przebadął model potęgowy z dwoma parametrami wyrażający średnią liczbę bitów reprezentacji skompresowanego obrazu danego typu w zależności od wartości Q . Średni względny błąd aproksymacji tego modelu dla większości badanych sekwencji wizyjnych jest poniżej 4, 6 i 6% odpowiednio dla obrazów typu I, P i B.

4.3.4. Wnioski dotyczące modelu kodera dla wartości Q z przedziału pierwszego

Autor zaproponował i przebadął dwa modele wyrażające średnią liczbę bitów potrzebnych do reprezentacji skompresowanego obrazu danego typu dla danych eksperymentalnych otrzymanych dla wartości Q z przedziału pierwszego (odcinek 1). Modele te zostały wyrażone za pomocą funkcji wielomianowej 3-go stopnia oraz potęgowej. O wysokiej dokładności modeli świadczy, uśredniony po wszystkich sekwencjach testowych, średni względny błąd aproksymacji przedstawiony w tabeli 4.20. Dodatkowo, we wspomnianej tabeli, umieszczono wartości odchylenia standardowego dla tego uśrednionego błędu aproksymacji.

Tabela 4.20. Średni względny błąd aproksymacji dla zaproponowanych modeli w pierwszym przedziale wartości Q uśredniony po wszystkich sekwencjach testowych oraz wartości odchylenia standardowego dla tego błędu

Typ obrazu	Typ modelu					
	wielomianowy 3-go stopnia (4 parametry)		wielomianowy 3-go stopnia (1 parametr)		potęgowy (2 parametry)	
	błąd średni [%]	odchylenie standardowe	błąd średni [%]	odchylenie standardowe	błąd średni [%]	odchylenie standardowe
I	1,59	0,32	2,63	1,52	2,97	1,49
P	3,57	0,80	5,47	2,33	3,83	2,35
B	3,33	1,27	6,36	4,76	6,84	4,54

Z analizy danych z tabeli 4.20 wynika, że model wielomianowy 3-go stopnia z czterema parametrami lepiej aproksymuje dane eksperymentalne niż model potęgowy, który ma dwa parametry. Można także stwierdzić, że uproszczenie modelu wielomianowego, dokonane poprzez zredukowanie liczby parametrów, pogorszyło jego dokładność średnio o około 1, 2 i 3 p.p. odpowiednio dla obrazów typu I, P i B. Dokładność takiego uproszczonego modelu wielomianowego jest porównywalna z modelem potęgowym. Ze względu na mniejszą liczbę parametrów uproszczonego modelu wielomianowego autor proponuje przyjęcie tego modelu do opisu danych otrzymanych dla wartości Q z przedziału pierwszego. Należy zaznaczyć, że dokładność aproksymacji modelami: wielomianowym z jednym parametrem i potęgowym sekwencji *ice*, *icon*, *station2* oraz *sunflower* była gorsza niż pozostałych sekwencji, szczególnie dla obrazów typu B (z tego powodu, wartości odchylenia standardowych dla obrazów typu B są wyższe od pozostałych). Dla uproszczonego modelu wielomianowego, głównym powodem pogorszenia dokładności dopasowania jest właśnie dokonana redukcja

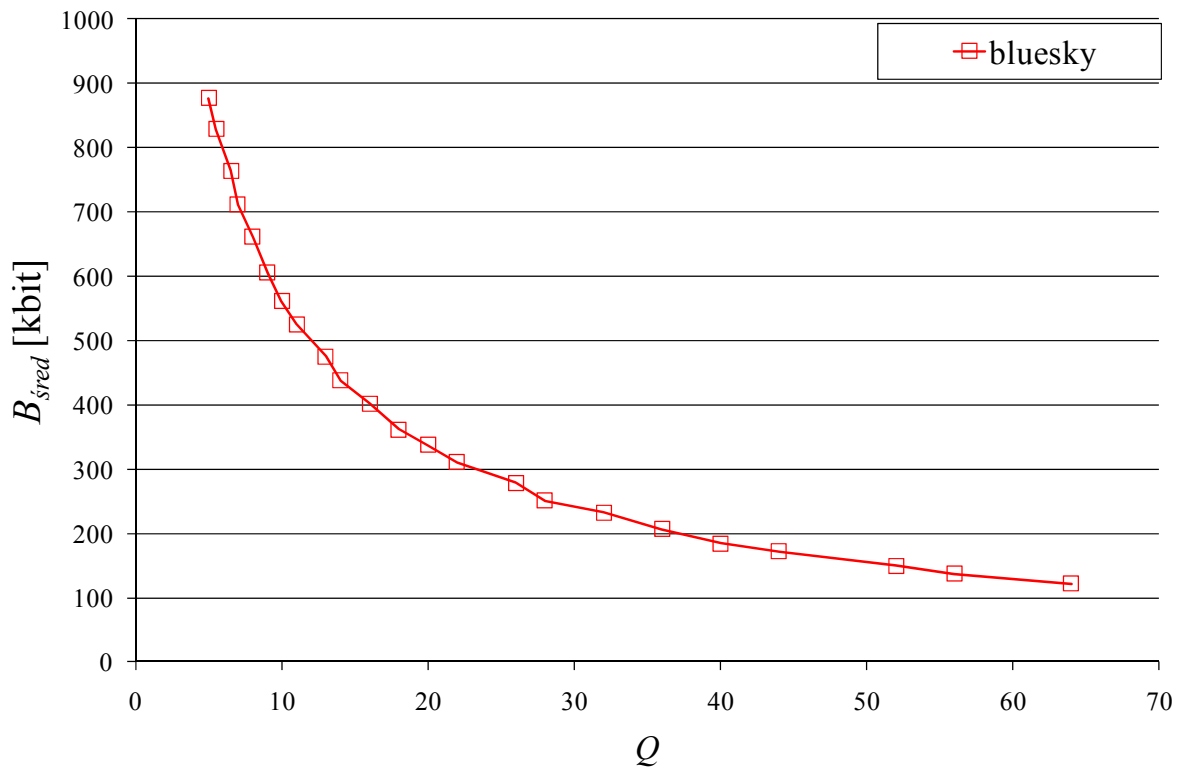
parametrów. Natomiast w przypadku modelu potęgowego, charakter zależności między liczbą bitów reprezentacji obrazów typu B a szerokością przedziałów kwantyzacji Q dla tych sekwencji, wyraźnie odbiega od zależności potęgowej.

4.4. Modelowanie kodera dla wartości Q z przedziału drugiego (odcinek 2)

4.4.1. Wybór typu modelu

Przedział drugi, ze względu na przyjęte granice (tabela 4.1), obejmuje wartości Q , które odpowiadają zakresowi prędkości bitowej strumienia danych na wyjściu kodera wizyjnego od 1 do 3 Mbit/s (szczegóły podano w punkcie 4.1). Z punktu widzenia ewentualnej możliwości wykorzystania zaproponowanych modeli w zastosowaniach praktycznych, jest to najbardziej interesujący przedział do badań, co zauważono również, m. in. w [Ma05, Wan06, Xie07, Yuan05].

Analogicznie jak dla danych z pierwszego przedziału wartości Q , o charakterze badanej zależności wnioskowano na podstawie wykresów przedstawiających średnią liczbę bitów potrzebnych do reprezentacji skompresowanego obrazu danego typu $B_{\text{śred}}$ w funkcji szerokości przedziałów kwantyzacji Q . Na wykresie 4.6 pokazano taką zależność dla przykładowej sekwencji wizyjnej dla obrazu typu I. Ze względu na podobny charakter badanej zależności dla pozostałych sekwencji testowych, jak i typów kodowania obrazu, w pracy ograniczono się do pokazania wybranego przykładu. Na podstawie analizy wykresów dla wszystkich sekwencji testowych wywnioskowano, że badaną zależność można opisać za pomocą funkcji hiperbolicznej, potęgowej lub wykładniczej.

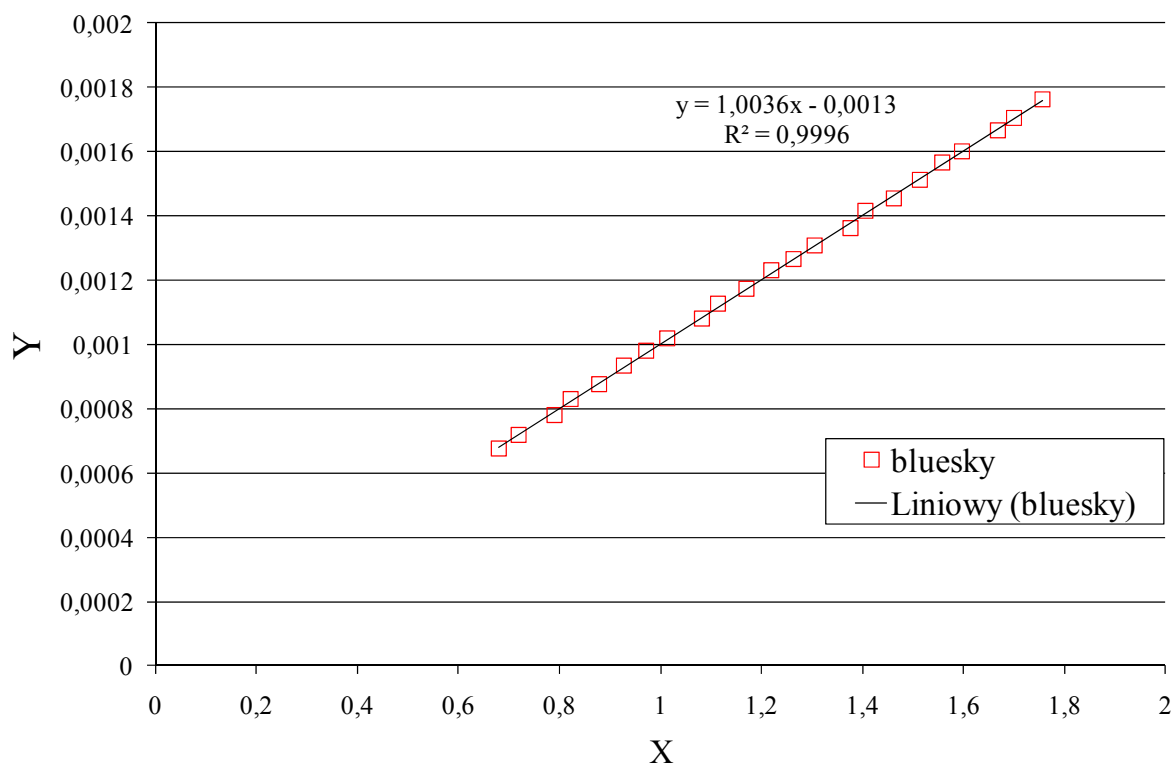


Wykres 4.6. Średnia liczba bitów potrzebnych do reprezentacji skompresowanego obrazu typu I w sekwencji *bluesky* (odcinek 2)

W celu weryfikacji przydatności modelu hiperbolicznego autor zastosował wcześniej wspomnianą metodę rektyfikacji (punkt 1.3), wprowadzając nowe zmienne według [Bron07]:

$$X = \log Q, \quad Y = \log\left(\frac{a}{B_{\text{śred}}} - c\right), \quad (4.7)$$

Na wykresie 4.7 przedstawiono wyniki zastosowania rektyfikacji dla przykładowych danych z wykresu 4.6 i parametrów modelu wyznaczonych metodą średniej. Dodatkowo na wspomnianym wykresie naniesiono krzywą regresji liniowej z podaną wartością współczynnika determinacji liniowej – R^2 . Wiedząc, że zmienne ze wzoru 4.7 są powiązane liniowo, tzn. $Y = AX + B$ można wyznaczyć poszukiwany wzór na $B_{\text{śred}}$.



Wykres 4.7. Średnia liczba bitów potrzebnych do reprezentacji skompresowanego obrazu typu I w sekwencji *bluesky* (odcinek 2)

Podobnie jak dla danych z przedziału pierwszego, w tabeli 4.21 podano wartości współczynników determinacji dla wybranych klas funkcji, uśrednione po wszystkich badanych sekwencjach wizyjnych niezależnie dla obrazów typu I, P i B.

Tabela 4.21. Wartości współczynnika determinacji dla sprawdzanych klas funkcji uśrednione po badanych sekwencjach wizyjnych (odcinek 2)

Typ obrazu	Typ modelu		
	hiperboliczny	potęgowy	wykładniczy
I	0,9992	0,9961	0,8919
P	0,9976	0,9958	0,8604
B	0,9876	0,9883	0,8359

Analiza wartości współczynników determinacji z tabeli 4.21 pozwoliła na wykluczenie z dalszych badań modelu wykładniczego. Natomiast niewielkie różnice między modelem hiperbolicznym a potęgowym spowodowały, że zostały one pozostawione do dalszych badań.

W kolejnych punktach zostaną przedstawione wartości parametrów oraz dokładność tych dwóch modeli wyznaczone niezależnie dla obrazów typu I, P i B.

4.4.2. Model hiperboliczny w przedziale drugim

4.4.2.1. Model dla obrazów typu I

Pierwszym modelem wybranym do aproksymacji danych z przedziału drugiego jest funkcja hiperboliczna wyznaczona na podstawie przekształcenia wzorów 4.7, postaci:

$$B_{\acute{s}red}(Q, a, b, c) = \frac{a}{Q^b + c}, \quad (4.8)$$

gdzie:

- a, b, c to parametry modelu,
- Q to szerokość przedziałów kwantyzacji,
- $B_{\acute{s}red}$ to średnia liczba bitów reprezentacji obrazu typu I.

Parametry modelu zostały wyznaczone poprzez minimalizację maksymalnego względnego błędu aproksymacji metodą optymalizacji quasi-newtonowską (punkt 1.3):

$$\min_{a,b,c} \max_Q \frac{|B_x(Q) - B_{\acute{s}red}(Q, a, b, c)|}{B_x(Q)} \cdot 100\%, \quad (4.9)$$

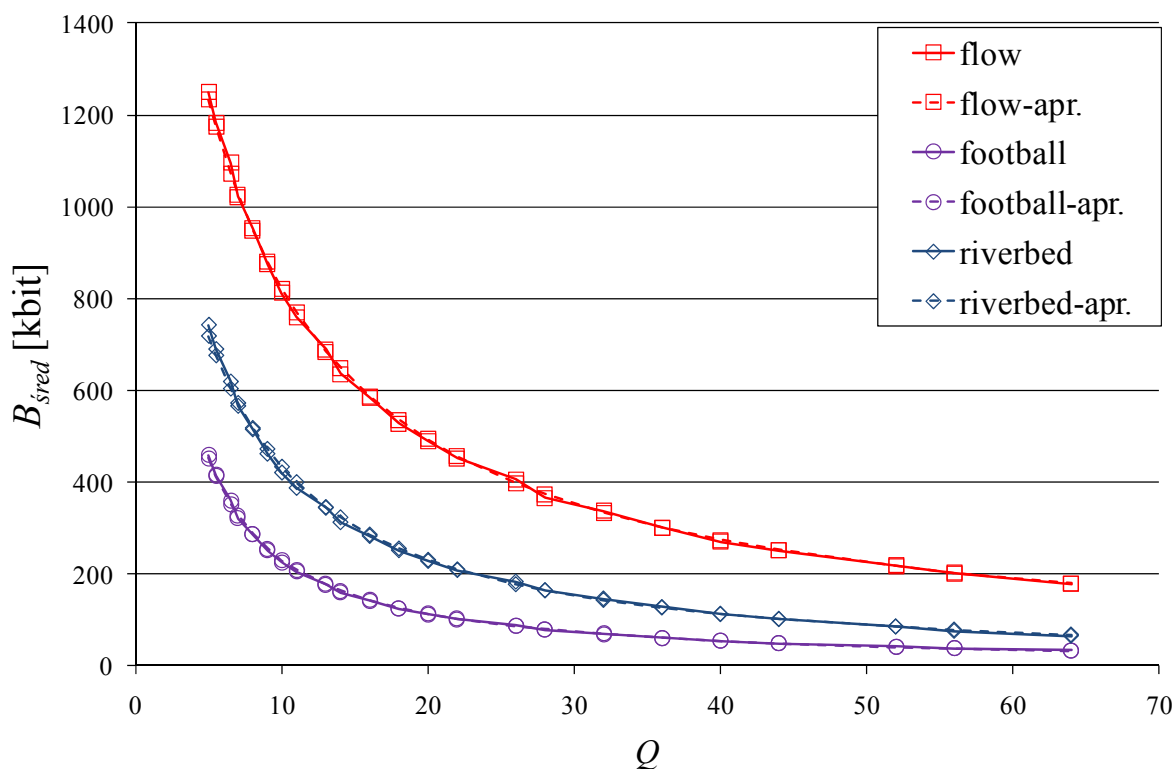
gdzie:

- a, b, c to parametry modelu,
- $B_x(Q)$ to zmierzona średnia liczba bitów reprezentacji obrazu typu I,
- $B_{\acute{s}red}(Q, a, b, c)$ to wartość oszacowania średniej liczby bitów reprezentacji obrazu typu I modelem hiperbolicznym.

Wartości parametrów modelu hiperbolicznego, maksymalny i średni względny błąd aproksymacji danych pochodzących z obrazów typu I zebrano w tabeli 4.22. Natomiast przykładowe krzywe eksperymentalne i ich oszacowania dla kilku sekwencji wizyjnych pokazano na wykresie 4.8.

Tabela 4.22. Wartości parametrów modelu hiperbolicznego oraz maksymalny i średni względny błąd aproksymacji dla obrazów typu I (odcinek 2)

Nazwa sekwencji	Obrazy typu I				
	<i>a</i>	<i>b</i>	<i>c</i>	względny błąd dopasowania [%]	
				maksymalny	średni
basket	7 245 655	0,9532	2,3934	2,30	1,27
bluesky	7 524 591	0,9718	3,8350	2,21	1,14
bus	6 577 059	1,0431	3,4849	3,04	1,16
cheer	5 982 781	0,9325	1,7947	2,24	0,90
city	11 864 487	1,1866	7,0595	3,49	1,41
crew	2 946 375	1,0188	-0,5114	3,48	1,69
flow	12 121 877	0,9948	4,8679	2,12	1,08
football	2 869 546	1,0680	0,7567	2,25	1,36
harbour	6 596 947	1,0058	2,4142	2,88	1,00
ice	665 107	0,7640	-1,3983	3,42	1,82
icon	4 068 992	1,2881	5,6750	3,45	1,27
pedestrian	2 020 047	0,9778	0,3726	3,10	1,13
riverbed	8 394 833	1,1465	5,3560	3,36	1,77
rushhour	1 321 316	0,8563	-0,0488	2,19	1,08
soccer	13 710 977	1,3492	11,7213	4,14	1,68
station2	6 587 251	1,2274	3,1877	5,42	2,58
stefan	3 646 003	0,8768	1,0981	1,98	0,90
sunflower	3 775 271	0,9478	1,8114	3,10	1,20
tractor	6 682 803	1,0540	3,4500	3,24	1,34
universal	2 680 466	1,0511	1,5645	2,58	1,26
warner	1 131 976	0,8319	-1,1089	3,73	1,56



Wykres 4.8. Krzywe eksperymentalne (linia ciągła) i oszacowane modelem hiperbolicznym (linia przerywana) dla obrazów typu I dla 3 przykładowych sekwencji wizyjnych (odcinek 2)

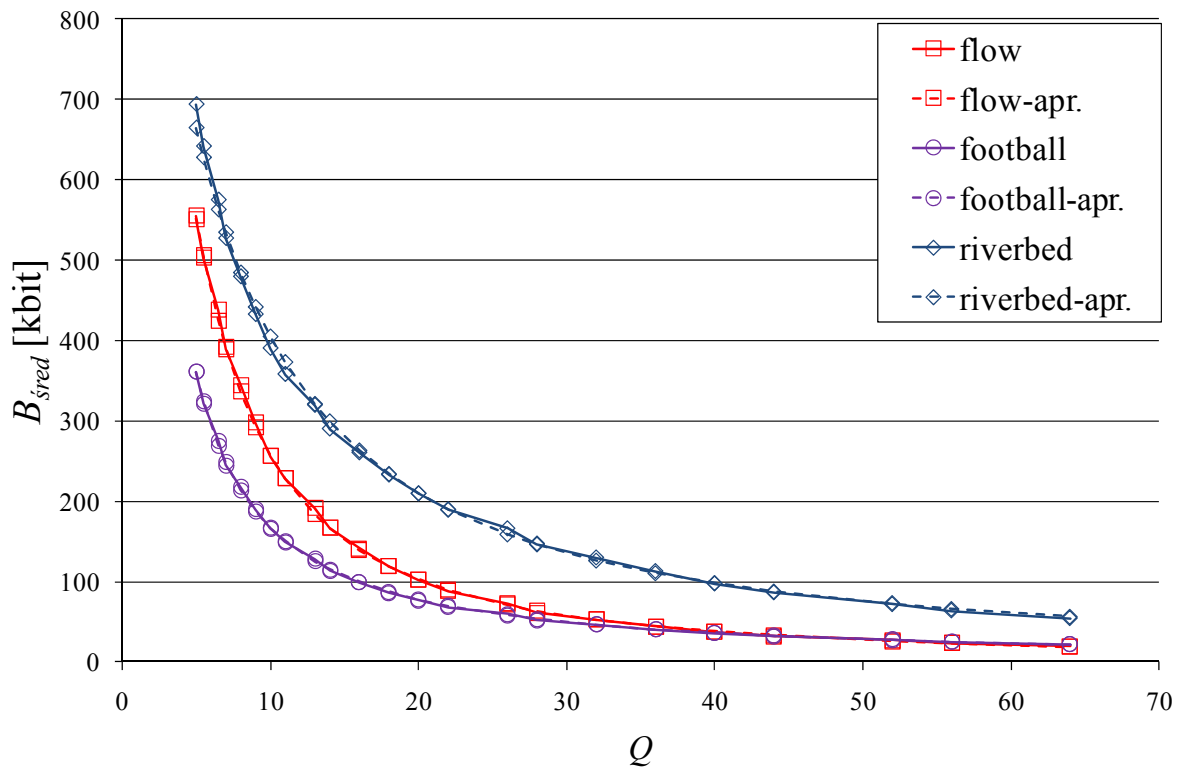
Dokładność zaproponowanego modelu hiperbolicznego z trzema parametrami dla obrazów typu I jest wysoka (średni względny błąd dopasowania poniżej 2% dla większości sekwencji testowych). Wyjątkiem jest sekwencja *station2* ze średnim względnym błędem aproksymacji równym 2,58%.

4.4.2.2. Model dla obrazów typu P

Dokładność oraz wartości parametrów modelu hiperbolicznego (wzór 4.8) wyznaczone dla obrazów typu P, dla wszystkich sekwencji testowych zebrano w tabeli 4.23. Ponadto, na wykresie 4.9 pokazano krzywe eksperymentalne i ich oszacowania dla kilku przykładowych sekwencji wizyjnych.

Tabela 4.23. Wartości parametrów modelu hiperbolicznego oraz maksymalny i średni względny błąd aproksymacji dla obrazów typu P (odcinek 2)

Nazwa sekwencji	Obrazy typu P				
	<i>a</i>	<i>b</i>	<i>c</i>	względny błąd dopasowania [%]	
				maksymalny	średni
basket	6 885 564	1,2511	3,1832	2,54	1,46
bluesky	3 168 695	1,3052	2,0144	3,20	1,45
bus	3 051 751	1,1993	-0,1092	3,24	1,59
cheer	4 513 529	1,0721	1,5742	2,37	1,15
city	7 285 272	1,6048	-0,2078	11,56	6,97
crew	3 722 714	1,2297	-0,2510	3,21	1,96
flow	9 538 615	1,4900	6,1873	4,14	1,71
football	2 063 494	1,0960	-0,1247	2,66	1,27
harbour	14 262 441	1,4589	10,5790	5,10	1,93
ice	473 704	0,8808	-1,5527	3,33	1,41
icon	487 512	1,0074	-2,7714	18,31	8,28
pedestrian	1 022 954	1,0170	0,0052	3,78	1,59
riverbed	10 608 263	1,2473	8,5117	4,22	2,06
rushhour	1 399 726	1,1416	0,8177	2,63	1,39
soccer	3 435 942	1,2235	-0,1501	5,59	2,72
station2	265 727	0,9596	-2,3208	8,98	5,19
stefan	5 305 155	1,3651	2,4789	3,92	1,81
sunflower	1 246 436	1,2468	0,0377	4,75	2,03
tractor	3 978 821	1,2085	0,7476	3,73	1,61
universal	1 822 635	1,1086	0,3814	2,63	1,49
warner	1 043 015	0,9989	-1,6600	6,71	2,79



Wykres 4.9. Krzywe eksperymentalne (linia ciągła) i oszacowane modelem hiperbolicznym (linia przerywana) dla obrazów typu P dla 3 przykładowych sekwencji wizyjnych (odcinek 2)

Średni względny błąd dopasowania zaproponowanego modelu hiperbolicznego z trzema parametrami dla obrazów typu P, dla większości sekwencji testowych jest poniżej 2,8%. Wyjątek stanowią sekwencje *city*, *icon* i *station2* ze średnim względnym błędem aproksymacji równym odpowiednio 6,97, 8,28 i 5,19%.

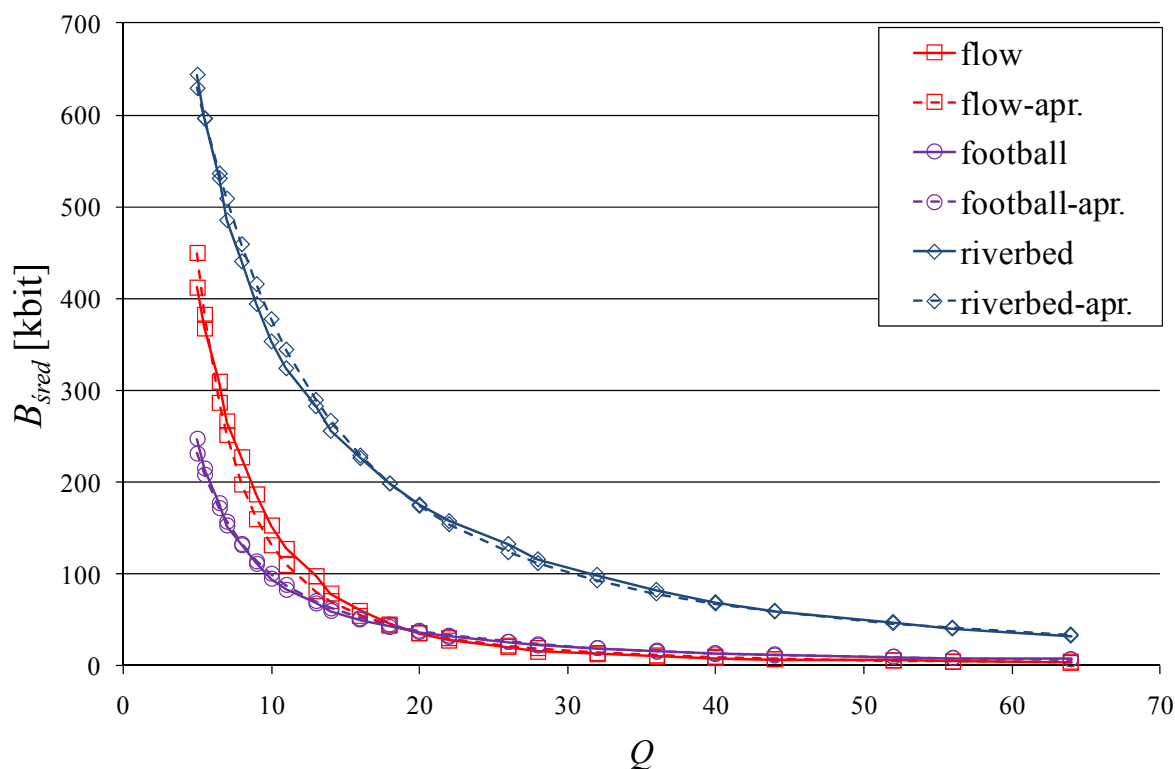
4.4.2.3. Model dla obrazów typu B

Analogicznie jak dla obrazów typu I i P, w tabeli 4.24 zebrano wartości parametrów modelu hiperbolicznego (wzór 4.8) oraz maksymalny i średni względny błąd aproksymacji wyznaczone dla obrazów typu B, dla wszystkich sekwencji testowych. Ponadto, krzywe eksperymentalne i ich oszacowania dla kilku przykładowych sekwencji wizyjnych pokazano na wykresie 4.10.

Średni względny błąd dopasowania zaproponowanego modelu hiperbolicznego z trzema parametrami dla obrazów typu B, dla większości sekwencji testowych jest poniżej 11,3%. Wyjątek stanowią sekwencje *bluesky* i *city* ze średnim względnym błędem aproksymacji powyżej 20% oraz *station2*, ze średnią wielkością błędu aproksymacji na poziomie 15%.

Tabela 4.24. Wartości parametrów modelu hiperbolicznego oraz maksymalny i średni względny błąd aproksymacji dla obrazów typu B (odcinek 2)

Nazwa sekwencji	Obrazy typu B				
	<i>a</i>	<i>b</i>	<i>c</i>	względny błąd dopasowania [%]	
				maksymalny	średni
basket	11 098 741	1,6546	14,7570	13,15	7,76
bluesky	3 664 704	1,8275	15,5224	40,01	24,54
bus	5 075 941	1,6291	4,8714	6,88	3,69
cheer	41 997 474	1,9422	90,9492	17,42	9,81
city	3 046 591	1,7394	-10,7340	35,86	21,74
crew	4 867 580	1,5944	-1,8604	6,91	3,87
flow	11 382 569	1,9186	3,4167	17,38	10,24
football	3 531 057	1,4943	4,1595	6,17	3,66
harbour	52 261 195	2,2573	48,9720	11,67	6,52
ice	357 166	1,0583	-2,3336	7,18	3,34
icon	548 729	1,3457	-4,3813	16,89	9,67
pedestrian	1 004 195	1,2711	1,4769	2,47	1,18
riverbed	21 967 401	1,5483	22,8337	6,92	3,38
rushhour	5 596 236	1,8763	20,9500	4,90	2,11
soccer	1 554 545	1,2648	-3,7167	11,81	6,34
station2	117 237	1,1243	-2,4703	26,06	15,30
stefan	15 047 532	2,0236	26,1983	16,34	11,22
sunflower	1 110 611	1,6019	5,8485	14,73	8,29
tractor	2 314 322	1,4103	-1,6901	4,94	2,62
universal	2 397 284	1,5108	2,3553	6,63	3,51
warner	1 293 738	1,2998	-3,3237	8,90	5,35



Wykres 4.10. Krzywe eksperymentalne (linia ciągła) i oszacowane modelem hiperbolicznym (linia przerywana) dla obrazów typu B dla 3 przykładowych sekwencji wizyjnych (odcinek 2)

4.4.2.4. Podsumowanie dotyczące modelu hiperbolicznego w przedziale drugim

W punkcie 4.4.2 autor zaproponował i przebadął model hiperboliczny z trzema parametrami wyrażający średnią liczbę bitów reprezentacji skompresowanego obrazu danego typu w zależności od wartości Q (wzór 4.8). Średni względny błąd aproksymacji tego modelu, dla większości badanych sekwencji wizyjnych, jest poniżej 2, 2,8 i 11,3% odpowiednio dla obrazów typu I, P i B.

4.4.3. Model potęgowy w przedziale drugim

4.4.3.1. Model dla obrazów typu I

Drugim modelem wybranym do aproksymacji danych z przedziału drugiego jest, tak jak dla przedziału pierwszego, funkcja potęgowa postaci:

$$B_{\text{śred}}(Q, a, b) = a \cdot Q^b, \quad (4.10)$$

gdzie:

- a, b to parametry modelu,
- Q to szerokość przedziałów kwantyzacji,
- $B_{\text{śred}}$ to średnia liczba bitów reprezentacji obrazu typu I.

Parametry modelu potęgowego zostały wyznaczone poprzez minimalizację maksymalnego względnego błędu aproksymacji (patrz wzór 4.6). W tabeli 4.25 podano wartości parametrów modelu potęgowego, maksymalny i średni względny błąd aproksymacji dla obrazów typu I dla wszystkich sekwencji testowych.

Tabela 4.25. Wartości parametrów modelu potęgowego oraz maksymalny i średni względny błąd aproksymacji dla obrazów typu I (odcinek 2)

Nazwa sekwencji	Obrazy typu I			
	a	b	względny błąd dopasowania [%]	
			maksymalny	średni
basket	3 937 581	-0,8089	5,92	3,16
bluesky	3 223 345	-0,7722	6,75	3,67
bus	3 299 249	-0,8778	6,95	3,62
cheer	3 667 937	-0,8120	4,60	2,35
city	4 440 551	-0,9480	9,79	5,43
crew	3 546 964	-1,0727	2,94	1,59
flow	4 550 611	-0,7627	6,77	3,82
football	2 390 930	-1,0204	2,65	1,12
harbour	3 833 647	-0,8759	5,96	3,12
ice	1 344 365	-0,9391	6,89	3,74
icon	1 885 045	-1,0890	8,68	4,97
pedestrian	1 881 149	-0,9615	3,39	1,30
riverbed	3 701 077	-0,9487	8,46	4,10
rushhour	1 344 611	-0,8604	2,08	1,09
soccer	4 239 195	-1,0574	12,22	7,24
station2	3 847 139	-1,0947	9,93	5,35
stefan	2 622 494	-0,7977	3,17	1,53
sunflower	2 420 340	-0,8435	5,55	2,56
tractor	3 511 909	-0,9021	6,77	3,31
universal	1 902 046	-0,9656	4,03	1,75
warner	1 795 732	-0,9440	5,44	2,79

Z tabeli 4.25 wynika, że dokładność zaproponowanego modelu dla obrazów typu I jest wysoka (średni względny błąd dopasowania dla większości sekwencji testowych poniżej 4%).

Wyjątkiem są sekwencje *city*, *icon*, *soccer* i *station2*, dla których średni względny błąd dopasowania wynosi odpowiednio 5,43, 4,97, 7,24 i 5,35%.

Ponieważ na wykresach trudno rozróżnić krzywe eksperymentalne od ich oszacowań (z powodu dużej dokładności modelu), autor nie zamieścił w pracy wykresów dla modelu potęgowego.

4.4.3.2. Model dla obrazów typu P

Dokładność oraz wartości parametrów modelu potęgowego (wzór 4.10) wyznaczone dla obrazów typu P, dla wszystkich sekwencji testowych zebrano w tabeli 4.26.

Tabela 4.26. Wartości parametrów modelu potęgowego oraz maksymalny i średni względny błąd aproksymacji dla obrazów typu P (odcinek 2)

Nazwa sekwencji	Obrazy typu P			
	a	b	względny błąd dopasowania [%]	
			maksymalny	średni
basket	4 059 107	-1,1138	6,97	3,62
bluesky	2 390 512	-1,2340	6,12	2,54
bus	3 123 352	-1,2046	3,12	1,67
cheer	3 158 991	-0,9807	3,84	1,84
city	7 383 579	-1,6083	11,66	6,96
crew	3 972 508	-1,2508	3,15	1,76
flow	4 837 077	-1,2996	10,04	5,52
football	2 114 914	-1,1025	2,66	1,33
harbour	5 700 715	-1,2329	13,56	7,55
ice	870 125	-1,0263	6,33	3,59
icon	1 432 676	-1,3078	6,78	3,61
pedestrian	1 009 984	-1,0136	3,91	1,53
riverbed	3 851 692	-1,0005	10,90	5,64
rushhour	1 196 708	-1,1016	3,28	1,55
soccer	3 564 299	-1,2329	5,89	2,67
station2	688 147	-1,2061	19,09	11,23
stefan	3 569 241	-1,2558	7,22	4,18
sunflower	1 238 129	-1,2449	4,77	2,03
tractor	3 522 079	-1,1747	3,20	1,81
universal	1 709 597	-1,0926	2,78	1,58
warner	1 730 650	-1,1167	7,60	4,47

Średni względny błąd aproksymacji modelu dla obrazów typu P dla większości sekwencji wizyjnych jest poniżej 6%. Wyjątkiem są sekwencje *city*, *harbour* i *station2*, dla których średni względny błąd dopasowania wynosi odpowiednio 6,96, 7,55 i 11,23%.

4.4.3.3. Model dla obrazów typu B

Analogicznie, jak dla obrazów typu I i P, w tabeli 4.27 zebrano wartości parametrów modelu potęgowego (wzór 4.10) oraz maksymalny i średni względny błąd aproksymacji wyznaczone dla obrazów typu B, dla wartości Q z drugiego przedziału, dla wszystkich sekwencji testowych.

Tabela 4.27. Wartości parametrów modelu potęgowego oraz maksymalny i średni względny błąd aproksymacji dla obrazów typu B (odcinek 2)

Nazwa sekwencji	Obrazy typu B			
	a	b	względny błąd dopasowania [%]	
			maksymalny	średni
basket	4 957 726	-1,4567	9,67	5,41
bluesky	2 621 058	-1,7413	37,86	19,11
bus	3 788 303	-1,5569	6,92	3,18
cheer	4 165 001	-1,3225	10,86	6,35
city	8 305 436	-2,0478	36,72	23,00
crew	5 798 170	-1,6421	7,64	4,12
flow	10 432 747	-1,9008	18,77	10,43
football	2 318 625	-1,3868	3,68	1,48
harbour	12 280 096	-1,8288	28,36	17,38
ice	680 465	-1,2204	10,09	5,99
icon	1 124 870	-1,5373	24,44	15,39
pedestrian	798 263	-1,2090	3,93	1,86
riverbed	5 184 845	-1,1892	18,79	10,54
rushhour	2 110 169	-1,6036	14,09	8,36
soccer	3 098 495	-1,4418	17,44	10,25
station2	236 453	-1,3185	32,98	19,98
stefan	7 585 333	-1,8300	19,30	10,86
sunflower	666 502	-1,4664	9,42	5,43
tractor	3 014 779	-1,4892	6,07	2,89
universal	1 995 663	-1,4638	5,10	2,35
warner	2 221 858	-1,4373	13,51	8,00

Średni względny błąd aproksymacji dla obrazów typu B jest poniżej 11% dla większości sekwencji wizyjnych. Wyjątkiem są sekwencje *bluesky*, *harbour*, *icon* i *station2* ze średnim względnym błędem dopasowania powyżej 15% oraz *city* ze średnim względnym błędem na poziomie 23%.

4.4.3.4. Podsumowanie dotyczące modelu potęgowego w przedziale drugim

Także i w tym przedziale, autor przebadął model potęgowy wyrażający średnią liczbę bitów reprezentacji skompresowanego obrazu danego typu w zależności od wartości Q . Średni względny błąd aproksymacji tego modelu, dla większości badanych sekwencji wizyjnych, jest poniżej 4, 6 i 11% odpowiednio dla obrazów typu I, P i B.

4.4.4. Wnioski dotyczące modelu kodera dla wartości Q z przedziału drugiego

Dla danych eksperymentalnych otrzymanych dla wartości Q z drugiego przedziału (odcinek 2) autor zaproponował i przebadął dwa modele: hiperboliczny i potęgowy. W tabeli 4.28 zebrano, uśredniony po wszystkich sekwencjach testowych, średni względny błąd aproksymacji wraz z odchyleniem standardowym wyznaczone niezależnie dla obrazów typu I, P i B.

Tabela 4.28. Średni względny błąd aproksymacji dla zaproponowanych modeli w drugim przedziale wartości Q uśredniony po wszystkich sekwencjach testowych oraz wartości odchylenia standardowego dla tego błędu

Typ obrazu	Typ modelu			
	hiperboliczny (3 parametry)		potęgowy (2 parametry)	
	błąd średni [%]	odchylenie standardowe	błąd średni [%]	odchylenie standardowe
I	1,36	0,39	3,22	1,61
P	2,47	1,93	3,65	2,55
B	7,82	6,23	9,16	6,44

Na podstawie informacji zawartych w tabeli 4.28 można stwierdzić, że model hiperboliczny z trzema parametrami lepiej dopasowuje się do badanych danych, niż model potęgowy z dwoma parametrami. Jednak, zarówno model hiperboliczny, jak i potęgowy, gorzej radzą sobie z dopasowaniem do danych z obrazów typu P i B dla sekwencji *bluesky*, *city* i *station2* od pozostałych sekwencji. Miało to istotny wpływ na wartości odpowiednich średnich i odchylen standardowych ujętych w tabeli podsumowującej dokładność badanych modeli.

Ponieważ wartości Q z drugiego przedziału obejmują użyteczny zakres prędkości strumienia danych na wyjściu kodera wizyjnego, dlatego autor proponuje przyjęcie do

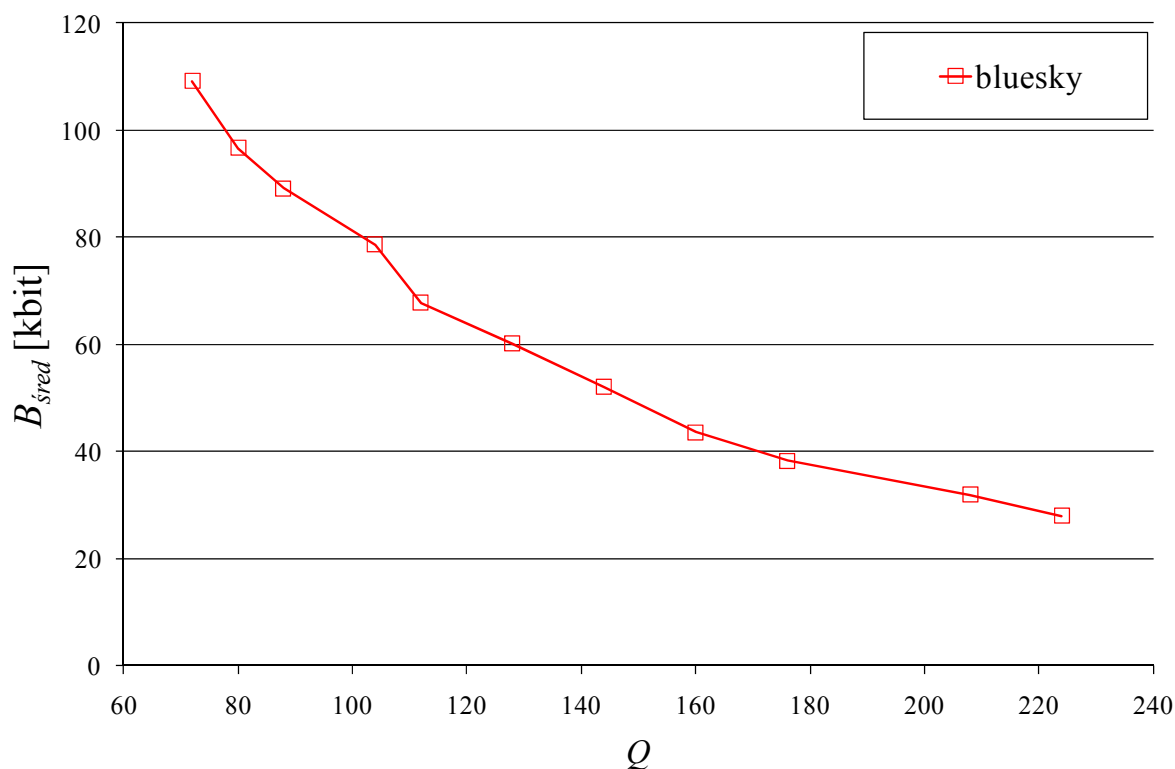
modelowania danych dla tego odcinka modelu hiperbolicznego, mimo że ma on o jeden parametr więcej niż model potęgowy.

4.5. Modelowanie danych dla wartości Q z przedziału trzeciego (odcinek 3)

Przedział trzeci obejmuje największe wartości szerokości przedziałów Q (tabela 4.1). Dla sekwencji wizyjnych o rozdzielczości przestrzennej i czasowej, odpowiednio 4CIF i 25 lub 30 obrazów na sekundę oraz struktury grupy obrazów (GOP) typowej dla zastosowań telewizyjnych, odpowiada on prędkościom bitowym strumienia danych na wyjściu kodera wizyjnego poniżej 1 Mbit/s.

Podobnie jak dla danych z przedziału pierwszego i drugiego, o charakterze badanej zależności wnioskowano na podstawie wykresów przedstawiających średnią liczbę bitów $B_{\text{śred}}$ potrzebnych do reprezentacji skompresowanego obrazu danego typu w funkcji szerokości przedziałów kwantyzacji Q . Na wykresie 4.11 pokazano taką zależność dla przykładowej sekwencji wizyjnej dla obrazu typu I. Ze względu na podobny charakter badanej zależności dla pozostałych sekwencji testowych, jak i typów kodowania obrazu, również dla wartości Q z przedziału 3 ograniczono się do pokazania wybranego, przykładowego wykresu. Na podstawie analizy wykresów wywnioskowano, że badaną zależność można próbować opisać za pomocą funkcji wielomianowej, potęgowej lub wykładniczej.

Zgodnie z przyjętą metodologią badań (punkt 1.3), także i dla tego przedziału wartości Q , do rozstrzygnięcia o tym, którą z wymienionych funkcji użyć do modelowania danych, wykorzystano analizę współczynnika determinacji. W tabeli 4.29 podano wartości współczynników determinacji dla wymienionych modeli, uśrednione po wszystkich badanych sekwencjach wizyjnych niezależnie dla obrazów typu I, P i B. Należy zaznaczyć, że zastosowanie wielomianów kolejnych wyższych rzędów nieznacznie poprawiało dokładność modelu, ale zwiększało liczbę jego parametrów.



Wykres 4.11. Średnia liczba bitów potrzebnych do reprezentacji skompresowanego obrazu typu I w sekwencji *bluesky* (odcinek 3)

Tabela 4.29. Wartości współczynnika determinacji dla sprawdzanych klas funkcji uśrednione po badanych sekwencjach wizyjnych

Typ obrazu	Funkcja		
	wielomian 3-go stopnia	potęgowa	wykładnicza
I	0,9973	0,9954	0,9666
P	0,9958	0,9909	0,9603
B	0,9962	0,9791	0,972

Podobnie jak w przedziale 1, dla modelu wielomianowego 3-go stopnia i modelu potęgowego otrzymano bardzo zbliżone wyniki. Jednak trudno na tej podstawie wnioskować, który z nich należy wykorzystać do modelowania danych, ponieważ model wielomianowy 3-go stopnia ma cztery parametry, a model potęgowy dwa. Dlatego, na tym etapie, autor odrzucił model wykładniczy, a do dalszej analizy pozostawił modele: wielomianowy 3-go stopnia i potęgowy. Dokładność oraz wartości parametrów tych dwóch modeli, wyznaczone niezależnie dla obrazów typu I, P i B, zostaną przedstawione w następnych punktach.

4.5.1. Model wielomianowy w przedziale trzecim

4.5.1.1. Model dla obrazów typu I

Pierwszym modelem wybranym do aproksymacji danych z przedziału trzeciego jest funkcja wielomianowa 3-go stopnia postaci:

$$B_{\text{śred}}(Q, a, b, c, d) = a \cdot Q^3 + b \cdot Q^2 + c \cdot Q + d, \quad (4.11)$$

gdzie:

- a, b, c, d to parametry modelu,
- Q to szerokość przedziałów kwantyzacji,
- $B_{\text{śred}}$ to średnia liczba bitów reprezentacji obrazów typu I.

Parametry modelu zostały wyznaczone poprzez minimalizację maksymalnego względnego błędu aproksymacji:

$$\min_{a,b,c,d} \max_Q \frac{|B_x(Q) - B_{\text{śred}}(Q, a, b, c, d)|}{B_x(Q)} \cdot 100\%, \quad (4.12)$$

gdzie:

- a, b, c, d to parametry modelu,
- $B_x(Q)$ to zmierzona średnia liczba bitów reprezentacji obrazu typu I,
- $B_{\text{śred}}(Q, a, b, c, d)$ to wartość oszacowania średniej liczby bitów reprezentacji obrazu typu I modelem wielomianowym 3-go stopnia z 4 parametrami.

W tabeli 4.30 podano wartości parametrów modelu, maksymalny i średni względny błąd dopasowania dla obrazów typu I, dla wszystkich badanych sekwencji. Wynika z niej, że średni względny błąd dopasowania dla większości sekwencji wizyjnych jest poniżej 2%.

Biorąc pod uwagę możliwości zastosowania zaproponowanego modelu do sterowania koderem wizyjnym, także i dla tego modelu autor podjął dalsze badania mające na celu uproszczenie modelu poprzez redukcję liczby jego parametrów. Analizując wartości współczynników korelacji liniowej dla poszczególnych par parametrów a, b, c i d (tabela 4.31) okazało się, że są one bliskie jedności.

Tabela 4.30. Wartości parametrów modelu wielomianowego oraz maksymalny i średni względny błąd aproksymacji dla obrazów typu I (odcinek 3)

Nazwa sekwencji	Obrazy typu I					
	$a \cdot 10^3$	b	c	d	względny błąd dopasowania [%]	
					maksymalny	średni
basket	-22,2131	13,5528	-2966,59	264708	2,67	1,69
bluesky	-14,1705	9,5056	-2328,27	231420	3,23	2,00
bus	-15,5017	9,1431	-1928,33	166562	2,06	1,48
cheer	-14,1074	9,3267	-2216,16	224650	3,35	1,93
city	-23,3143	13,2056	-2594,53	191887	2,69	1,93
crew	-8,5199	5,0249	-1034,87	87853	3,01	1,78
flow	-23,7399	15,5108	-3645,56	349652	2,98	1,97
football	-6,2820	3,7125	-771,98	69111	2,45	1,91
harbour	-18,5741	11,1027	-2365,36	201494	1,94	1,08
ice	-3,5995	2,1831	-492,29	51684	1,71	1,21
icon	-4,5134	2,5326	-491,83	39983	2,42	1,56
pedestrian	-6,7870	4,0298	-838,82	71101	2,16	1,21
riverbed	-16,2187	9,3986	-1907,16	150395	1,68	1,38
rushhour	-6,5493	3,8741	-821,92	74221	2,12	1,20
soccer	-12,4510	7,0738	-1390,80	106604	2,51	1,87
station2	-11,1205	6,3459	-1245,59	91476	3,10	2,36
stefan	-11,4723	7,5848	-1810,42	178185	2,50	1,51
sunflower	-10,1481	6,3738	-1422,08	134858	3,79	2,25
tractor	-15,9104	9,5110	-1992,41	167221	3,25	1,85
universal	-5,8819	3,4956	-725,14	65642	3,52	1,96
warner	-3,7966	2,5555	-621,76	67079	3,22	1,96

Tabela 4.31. Wartości współczynnika korelacji liniowej dla poszczególnych par parametrów modelu wielomianowego dla obrazów typu I (odcinek 3)

Parametry modelu	a	b	c	d
a	1	-0,9919	0,9615	-0,8888
b	-0,9919	1	-0,9886	0,9396
c	0,9615	-0,9886	1	-0,9802
d	-0,8888	0,9396	-0,9802	1

Dlatego, wykorzystując liniowe zależności między parametrami, opracowano model wielomianowy z jednym parametrem, postaci:

$$B_{\text{śred}}(Q, d) = (a_1 d + a_2) \cdot Q^3 + (b_1 d + b_2) \cdot Q^2 + (c_1 d + c_2) \cdot Q + d, \quad (4.13)$$

gdzie:

- d to parametr modelu,
- $B_{\text{śred}}$ to średnia liczba bitów reprezentacji obrazu typu I,
- Q to szerokość przedziałów kwantyzacji,
- a_1, a_2, b_1, b_2, c_1 i c_2 to uniwersalne stałe modelu dla obrazów typu I.

Parametr d modelu został wyznaczony tak jak wcześniej, poprzez minimalizację maksymalnego względnego błędu aproksymacji. W tabeli 4.32 podano wartości stałych uproszczonego modelu wielomianowego dla obrazów typu I, dla odcinka 3.

Tabela 4.32. Wartości stałych uproszczonego modelu wielomianowego dla obrazów typu I (odcinek 3)

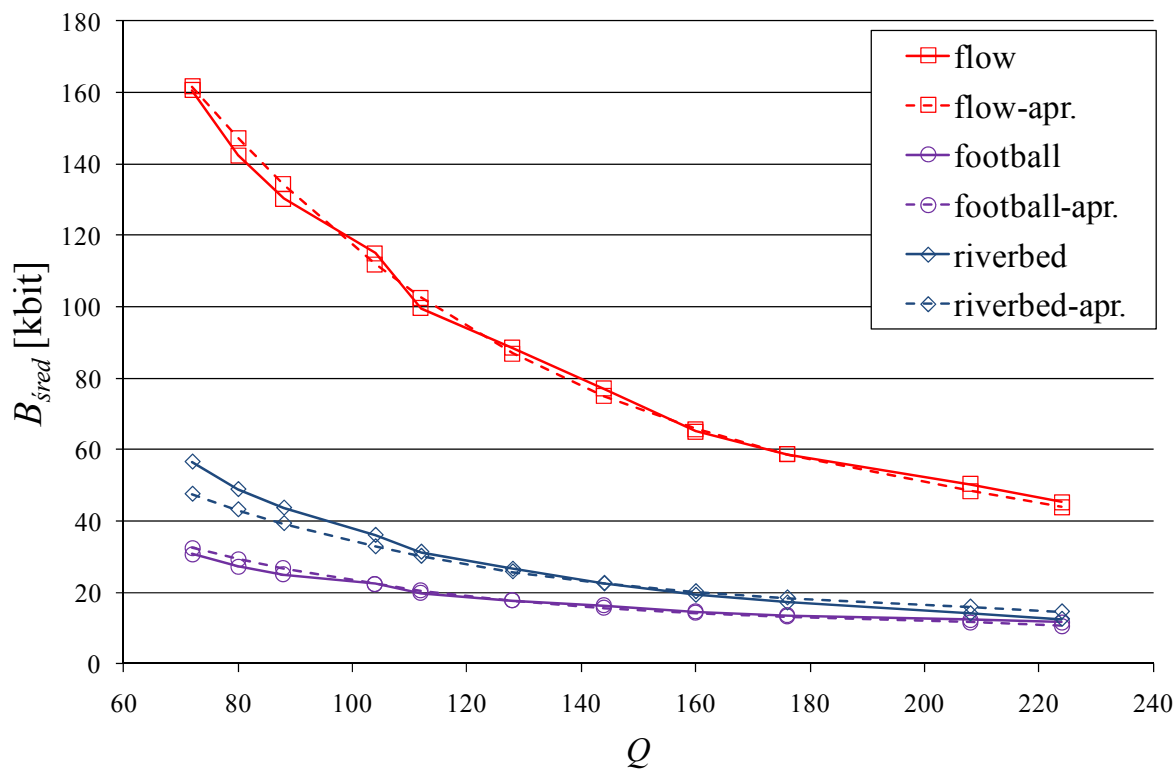
Stała	Wartość
a_1	-0,0000000692
a_2	-0,0023148451
b_1	0,0000451012
b_2	0,9871857978
c_1	-0,0105840255
c_2	-97,5048684104

W kolejnej tabeli (4.33), zebrano wartości parametru d , maksymalny i średni względny błąd aproksymacji dla wszystkich badanych sekwencji wizyjnych. Ponadto, na wykresie 4.12 przedstawiono krzywe eksperymentalne i ich oszacowania uproszczonym modelem wielomianowym dla 3 przykładowych sekwencji wizyjnych.

Dokładność zaproponowanego modelu wielomianowego 3-go stopnia z czterema parametrami, dla obrazów typu I, dla wartości Q z trzeciego przedziału jest wysoka (średni względny błąd dopasowania poniżej 2% dla większości sekwencji wizyjnych). Uproszczenie modelu, dokonane poprzez zredukowanie liczby parametrów, pogorszyło jego dokładność dla tego typu obrazów średnio o około 4 p.p., za wyjątkiem sekwencji *city*, *riverbed* i *station2* z błędem dopasowania większym odpowiednio o około 13, 7 i 16 p.p.

Tabela 4.33. Wartości parametru uproszczonego modelu wielomianowego oraz maksymalny i średni względny błąd aproksymacji dla obrazów typu I (odcinek 3)

Nazwa sekwencji	Obrazy typu I		
	d	względny błąd dopasowania [%]	
		maksymalny	średni
basket	257 881	2,65	1,37
bluesky	244 716	6,35	2,88
bus	159 826	2,43	1,43
cheer	278 078	12,96	7,27
city	113 800	27,22	15,00
crew	90 034	6,13	2,93
flow	368 807	3,41	2,25
football	79 064	7,89	4,28
harbour	175 458	8,78	4,72
ice	70 996	9,54	6,52
icon	41 292	5,88	3,18
pedestrian	69 123	4,66	2,67
riverbed	112 739	16,01	8,70
rushhour	81 732	3,80	2,00
soccer	84 007	11,44	6,58
station2	54 633	30,79	18,70
stefan	196 383	3,10	1,77
sunflower	159 802	12,68	6,79
tractor	152 504	4,17	2,10
universal	81 327	14,98	8,23
warner	91 593	11,38	6,96



Wykres 4.12. Krzywe eksperymentalne (linia ciągła) i oszacowane uproszczonym modelem wielomianowym (linia przerywana) dla obrazów typu I dla 3 przykładowych sekwencji wizyjnych (odcinek 3)

4.5.1.2. Model dla obrazów typu P

Wartości parametrów modelu wielomianowego (wzór 4.11) wyznaczone dla obrazów typu P oraz dokładność tego modelu dla wszystkich sekwencji testowych zebrano w tabeli 4.34. Średni względny błąd aproksymacji dla większości sekwencji wizyjnych jest poniżej 2,5% za wyjątkiem sekwencji *harbour* (3,11%) i *stefan* (4,66%).

Tabela 4.34. Wartości parametrów modelu wielomianowego oraz maksymalny i średni względny błąd aproksymacji dla obrazów typu P (odcinek 3)

Nazwa sekwencji	Obrazy typu P					
	$a \cdot 10^3$	b	c	d	względny błąd dopasowania [%]	
					maksymalny	średni
basket	-8,7402	5,0242	-1012,28	81856	3,00	1,79
bluesky	-3,4281	1,9180	-372,51	30126	3,19	2,07
bus	-3,7801	2,2594	-481,81	42883	2,09	1,71
cheer	-8,7272	5,4116	-1199,75	106958	3,61	2,23
city	-3,0609	1,6726	-308,45	23745	2,78	2,11
crew	-5,2892	2,9738	-592,67	48356	2,90	2,17
flow	-4,7468	2,6943	-535,58	43523	2,89	1,93
football	-4,8265	2,7956	-568,62	47384	2,93	1,63
harbour	-9,9406	5,5620	-1069,99	76202	4,53	3,11
ice	-2,1363	1,2349	-263,15	24892	2,16	1,53
icon	-1,2904	0,7304	-147,76	13162	2,28	1,39
pedestrian	-3,0220	1,7239	-347,82	29690	2,08	1,43
riverbed	-16,0000	9,0540	-1778,14	132300	3,08	2,18
rushhour	-2,5668	1,4968	-310,50	25801	3,11	2,08
soccer	-5,1491	2,9197	-586,64	48636	2,57	1,67
station2	0,0598	0,0766	-52,82	8465	3,39	2,44
stefan	-6,0755	3,2424	-596,96	43330	6,92	4,66
sunflower	-1,6671	0,9005	-168,11	14335	3,68	2,05
tractor	-5,0090	2,9241	-622,81	54407	3,37	2,40
universal	-3,9005	2,1860	-432,35	37028	2,68	1,80
warner	-3,1785	1,8668	-395,62	36216	3,13	1,94

Podobnie jak dla obrazów typu I, między poszczególnymi parami wyznaczonych parametrów a , b , c i d istnieją silne zależności liniowe (patrz tabela 4.35). Dlatego, dla obrazów typu P, autor także podjął decyzję o uproszczeniu modelu do tylko jednego parametru.

Tabela 4.35. Wartości współczynnika korelacji liniowej dla poszczególnych par parametrów modelu wielomianowego dla obrazów typu P (odcinek 3)

Parametry modelu	a	b	c	d
a	1	-0,9982	0,9891	-0,9594
b	-0,9982	1	-0,9961	0,9741
c	0,9891	-0,9961	1	-0,99
d	-0,9594	0,9741	-0,99	1

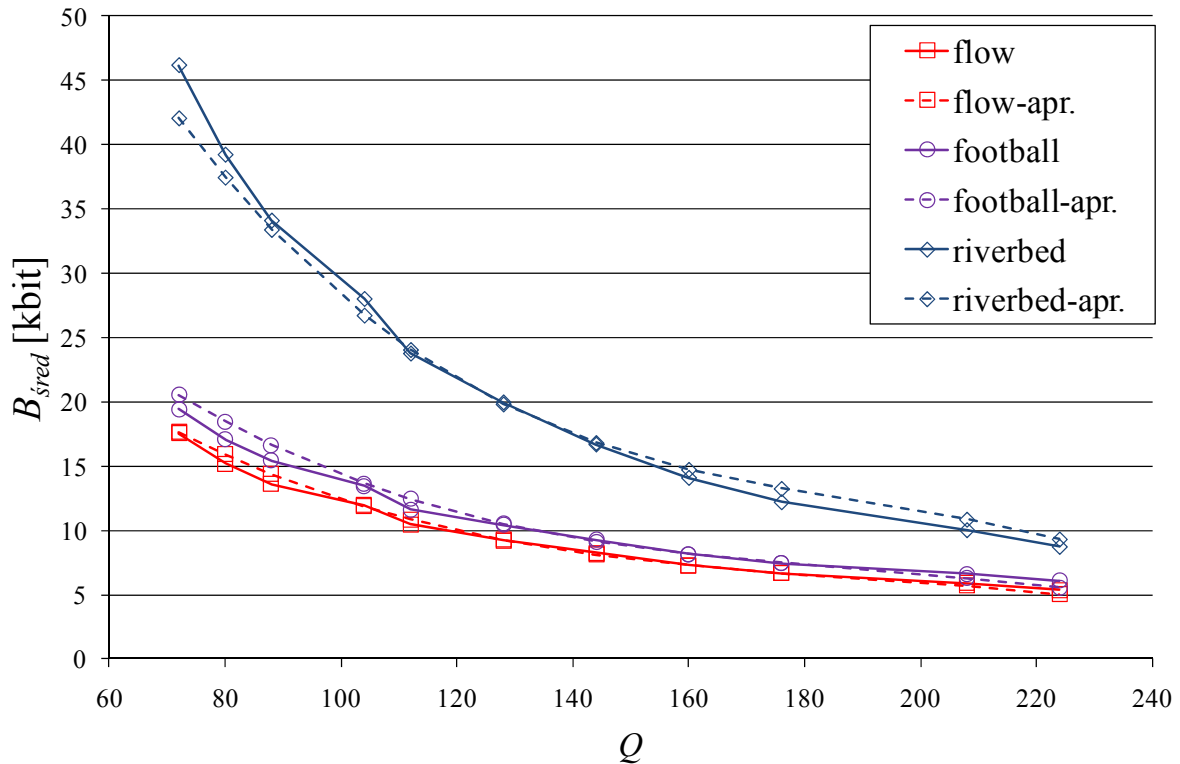
W tabeli 4.36. zostały umieszczone stałe uproszczonego modelu wielomianowego (wzór 4.13) dla obrazów typu P. Natomiast wartości parametru d , maksymalny i średni względny błąd aproksymacji dla wszystkich badanych sekwencji wizyjnych zebrano w tabeli 4.37. Ponadto, na wykresie 4.13 przedstawiono, dla 3 przykładowych sekwencji wizyjnych, krzywe eksperymentalne i ich aproksymacje uproszczonym modelem wielomianowym.

Tabela 4.36. Wartości stałych uproszczonego modelu wielomianowego dla obrazów typu P (odcinek 3)

Stała	Wartość
$a1$	-0,0000001138
$a2$	0,0003129255
$b1$	0,0000654241
$b2$	-0,2024468200
$c1$	-0,0131349116
$c2$	39,9207555922

Tabela 4.37. Wartości parametru uproszczonego modelu wielomianowego oraz maksymalny i średni względny błąd aproksymacji dla obrazów typu P (odcinek 3)

Nazwa sekwencji	Obrazy typu P		
	d	względny błąd dopasowania [%]	
		maksymalny	średni
basket	93 378	10,80	5,40
bluesky	29 558	7,10	3,51
bus	50 513	6,16	3,80
cheer	132 214	7,40	5,10
city	21 199	11,13	6,11
crew	49 951	4,08	2,44
flow	44 719	5,50	2,46
football	53 004	8,23	4,16
harbour	46 732	27,88	15,20
ice	28 674	3,73	2,39
icon	7 422	23,41	13,32
pedestrian	30 779	4,01	1,93
riverbed	114 354	8,86	4,75
rushhour	16 019	27,53	16,21
soccer	53 706	6,36	3,74
station2	-3 586	86,83	52,44
stefan	34 674	10,05	3,22
sunflower	12 108	5,38	3,00
tractor	58 969	8,13	4,25
universal	45 567	14,60	8,38
warner	45 161	10,29	5,98



Wykres 4.13. Krzywe eksperymentalne (linia ciągła) i oszacowane uproszczonym modelem wielomianowym (linia przerywana) dla obrazów typu P, dla 3 przykładowych sekwencji wizyjnych (odcinek 3)

Dla zaproponowanego modelu wielomianowego 3-go stopnia z czterema parametrami, dla obrazów typu P z trzeciego przedziału wartości Q , średni względny błąd dopasowania dla większości sekwencji wizyjnych jest poniżej 2,44%. Uproszczenie modelu dokonane poprzez zredukowanie liczby parametrów pogorszyło dokładność oszacowania tego typu obrazów średnio o około 6 p.p., z wyjątkiem sekwencji *harbour* (12 p.p.), *icon* (12 p.p.), *rushhour* (14 p.p.) i *station2* (50 p.p.).

4.5.1.3. Model dla obrazów typu B

Dokładność oraz wartości parametrów modelu wielomianowego z czterema parametrami (wzór 4.11) wyznaczone dla obrazów typu B, dla wszystkich sekwencji testowych, zebrano w tabeli 4.38. Średni względny błąd aproksymacji dla większości sekwencji wizyjnych jest poniżej 5%, za wyjątkiem sekwencji *city* (9,76%), *station2* (7,26%) i *tractor* (8,81%).

Tabela 4.38. Wartości parametrów modelu wielomianowego oraz maksymalny i średni względny błąd aproksymacji dla obrazów typu B (odcinek 3)

Nazwa sekwencji	Obrazy typu B					
	$a \cdot 10^3$	b	c	d	względny błąd dopasowania [%]	
					maksymalny	średni
basket	-3,6479	1,9550	-364,60	26198	4,07	2,67
bluesky	-0,5926	0,3057	-56,41	5530	2,10	1,18
bus	-1,8555	0,9937	-186,04	14015	3,34	2,28
cheer	-5,3266	2,9754	-573,02	40144	7,98	4,08
city	1,3998	-0,5885	59,85	766	17,40	9,76
crew	-1,8007	1,0578	-217,65	16396	5,72	3,06
flow	-1,3402	0,6800	-126,75	10198	3,33	1,94
football	-1,4584	0,8366	-178,33	15525	4,03	2,13
harbour	-1,8870	1,0173	-187,97	12731	8,81	4,03
ice	-0,9639	0,5514	-114,08	9912	2,07	1,47
icon	-0,6253	0,3561	-71,24	5779	2,65	1,57
pedestrian	-1,1474	0,6546	-135,25	11285	3,16	2,14
riverbed	-13,2122	7,2500	-1355,41	90346	10,36	5,08
rushhour	-0,8330	0,4625	-88,22	6156	8,39	4,40
soccer	-2,0992	1,1977	-242,91	19983	2,82	1,94
station2	-0,8306	0,4423	-77,42	4644	11,64	7,26
stefan	-0,8931	0,5144	-105,06	8535	4,20	2,98
sunflower	-0,1898	0,1069	-24,08	2656	3,29	2,11
tractor	-0,6611	0,5938	-163,55	14678	13,77	8,81
universal	-0,7208	0,4829	-114,81	10168	4,61	2,72
warner	-1,0742	0,6940	-160,76	13987	4,40	2,58

Analogicznie jak dla obrazów poprzednich typów, między poszczególnymi parami parametrów a , b , c i d istnieją silne zależności liniowe (patrz tabela 4.39). Z tego powodu, również dla obrazów typu B, autor uprościł model wielomianowy do tylko jednego parametru. Stałe uproszczonego modelu wielomianowego (wzór 4.13) zostały umieszczone w tabeli 4.40.

Tabela 4.39. Wartości współczynnika korelacji liniowej dla poszczególnych par parametrów modelu wielomianowego dla obrazów typu B (odcinek 3)

Parametry modelu	a	b	c	d
a	1	-0,999	0,9951	-0,9876
b	-0,999	1	-0,9985	0,9931
c	0,9951	-0,9985	1	-0,9977
d	-0,9876	0,9931	-0,9977	1

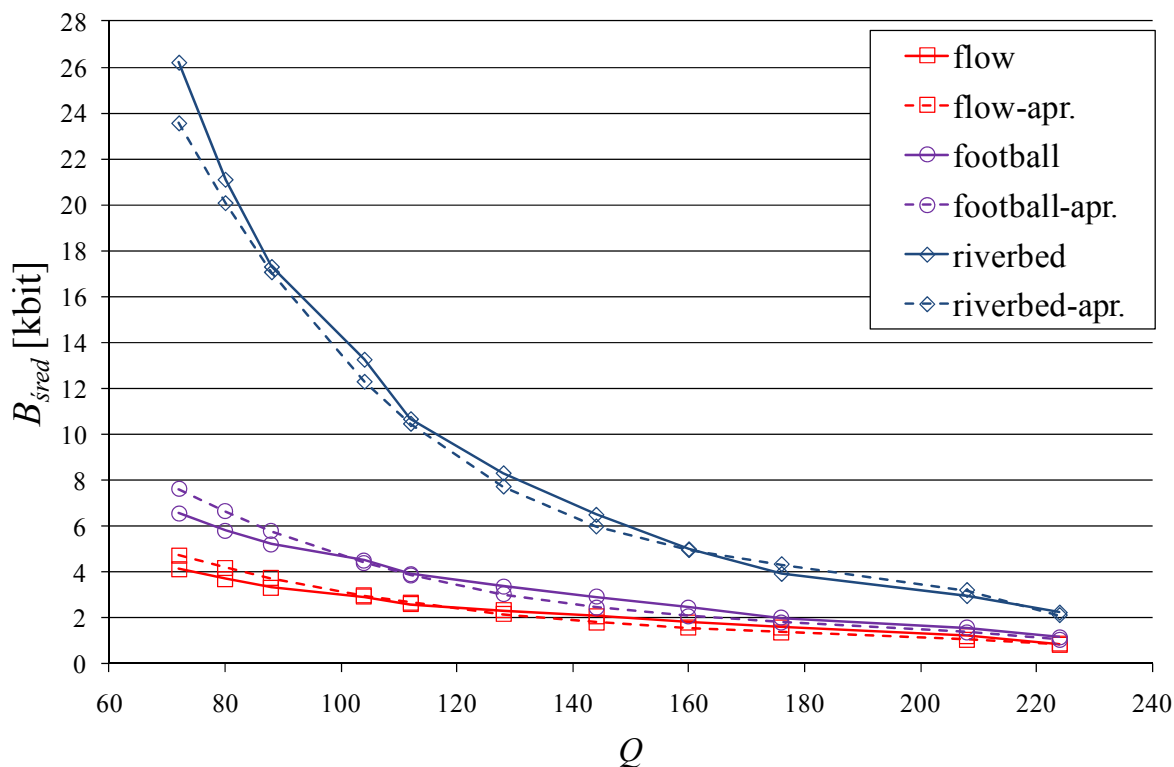
Tabela 4.40. Wartości stałych uproszczonego modelu wielomianowego dla obrazów typu B
(odcinek 3)

Stała	Wartość
<i>a1</i>	-0,0000001504
<i>a2</i>	0,0005353337
<i>b1</i>	0,0000821261
<i>b2</i>	-0,2524531645
<i>c1</i>	-0,0152342836
<i>c2</i>	32,5775452826

W tabeli 4.41 zebrano wartości parametru *d*, maksymalny i średni względny błąd aproksymacji dla wszystkich badanych sekwencji wizyjnych. Ponadto, na wykresie 4.14 przedstawiono krzywe eksperymentalne i ich oszacowania uproszczonym modelem wielomianowym dla 3 przykładowych sekwencji wizyjnych.

Tabela 4.41. Wartości parametru uproszczonego modelu wielomianowego oraz maksymalny i średni względny błąd aproksymacji dla obrazów typu B (odcinek 3)

Nazwa sekwencji	Obrazy typu B		
	<i>d</i>	względny błąd dopasowania [%]	
		maksymalny	średni
basket	35 595	24,24	14,98
bluesky	10 831	46,37	27,89
bus	19 625	28,49	17,28
cheer	37 762	9,21	5,76
city	-3 036	83,63	51,00
crew	6 213	48,27	29,04
flow	12 779	13,86	9,83
football	23 384	15,88	10,58
harbour	3 914	40,81	20,27
ice	14 792	24,59	14,34
icon	2 550	14,73	8,88
pedestrian	13 375	8,50	5,83
riverbed	81 885	10,10	6,06
rushhour	-5 393	112,00	61,04
soccer	36 660	39,03	21,96
station2	-8 077	189,72	138,04
stefan	7 059	6,25	3,77
sunflower	-986	31,99	16,69
tractor	-254	79,00	50,72
universal	5 575	34,23	20,66
warner	13 182	18,18	12,89



Wykres 4.14. Krzywe eksperymentalne (linia ciągła) i oszacowane (linia przerywana) dla obrazów typu B dla 3 przykładowych sekwencji wizyjnych (odcinek 3)

Wybierając model wielomianowy 3-go stopnia z czterema parametrami do modelowania danych dla obrazów typu B, uzyskano średni względny błąd dopasowania dla większości sekwencji testowych poniżej 5%. Uproszczenie modelu, dokonane poprzez zredukowanie liczby parametrów, pogorszyło dokładność oszacowania badanego typu obrazów średnio o około 23 punkty procentowe. Wyjątek stanowią sekwencje *city*, *rushhour*, *station2* i *tractor*, dla których średni względny błąd aproksymacji dla uproszczonego modelu jest wyższy odpowiednio o około 41, 57, 131 i 42 p.p.

4.5.1.4. Weryfikacja dokładności uproszczonego modelu wielomianowego w przedziale trzecim

Analogicznie jak dla modelu z przedziału pierwszego, w celu weryfikacji dokładności zaproponowanego uproszczonego modelu wielomianowego z przedziału trzeciego, przeprowadzono dodatkowe badania z wykorzystaniem kontrolnego zbioru sekwencji wizyjnych. Badania te polegały na wyznaczeniu wartości parametru d oraz maksymalnego i średniego względnego błędu aproksymacji danych eksperymentalnych otrzymanych dla

kontrolnego zbioru sekwencji wizyjnych z wykorzystaniem uniwersalnych stałych zebranych w tabeli 4.42.

Tabela 4.42. Wartości stałych uproszczonego modelu wielomianowego (odcinek 3)

Stała	Typ obrazu		
	I	P	B
a_1	-0,0000000692	-0,0000001138	-0,0000001504
a_2	-0,0023148451	0,0003129255	0,0005353337
b_1	0,0000451012	0,0000654241	0,0000821261
b_2	0,9871857978	-0,2024468200	-0,2524531645
c_1	-0,0105840255	-0,0131349116	-0,0152342836
c_2	-97,5048684104	39,9207555922	32,5775452826

W tabeli 4.43 umieszczono wartości parametru d oraz maksymalnego i średniego względnego błędu aproksymacji uzyskane dla kontrolnego zbioru sekwencji wizyjnych..

Tabela 4.43. Wartości parametru modelu oraz maksymalny i średni względny błąd aproksymacji danych z wykorzystaniem uproszczonego modelu wielomianowego dla obrazów typu I, P i B dla kontrolnego zbioru sekwencji wizyjnych (odcinek 3)

Nazwa sekwencji	Obrazy typu I			Obrazy typu P			Obrazy typu B		
	$d \cdot 10^2$	względny błąd dopasowania [%]		$d \cdot 10^2$	względny błąd dopasowania [%]		$d \cdot 10^2$	względny błąd dopasowania [%]	
		maks.	średni		maks.	średni		maks.	średni
carpark	2 579	2,65	1,37	934	10,80	5,40	356	24,24	14,98
night	2 447	6,35	2,88	296	7,10	3,51	108	46,37	27,89
outdorr	1 598	2,43	1,43	505	6,16	3,80	196	28,49	17,28
raven	2 781	12,96	7,27	1 322	7,41	5,10	378	9,21	5,76
shuttle	1 138	27,22	15,00	212	11,13	6,11	0,10	100,40	42,53
spincalendar	900	6,13	2,93	500	4,08	2,44	62	48,27	29,04
toys	3 688	3,41	2,25	447	5,50	2,46	128	13,86	9,83
vintagecar	791	7,89	4,28	530	8,23	4,16	234	15,88	10,58

Wyniki eksperymentu potwierdzają, że zaproponowany uproszczony model wielomianowy z zestawem uniwersalnych stałych dobrze aproksymuje dane eksperymentalne uzyskane dla sekwencji spoza zbioru testowego dla obrazów typu I, P i B. Średni względny błąd aproksymacji badanego modelu wynosi odpowiednio około 5, 6 i 19%. Wyjątkiem są sekwencje *night*, *shuttle* i *spincalendar*, dla których średni względny błąd aproksymacji dla obrazów typu B wynosi odpowiednio około 28, 42 i 29%.

4.5.1.5. Podsumowanie dotyczące modelu wielomianowego w przedziale trzecim

Dla danych eksperymentalnych otrzymanych dla wartości Q z przedziału trzeciego autor zaproponował model wielomianowy 3-go stopnia z jednym parametrem d wyrażający średnią liczbę bitów reprezentacji skompresowanego obrazu danego typu w zależności od wartości Q . Do modelu wyznaczono także wartości uniwersalnych stałych niezależnie dla obrazów typu I, P i B. Ponadto, dokładność uproszczonego modelu wielomianowego została zweryfikowana na dodatkowym zbiorze sekwencji wizyjnych. Zaproponowanym modelem można przybliżyć dane uzyskane dla szerokiej klasy sekwencji wizyjnych ze średnim względnym błędem dopasowania około 6, 8,5 i 28% odpowiednio dla obrazów typu I, P i B.

4.5.2. Model potęgowy w przedziale trzecim

4.5.2.1. Model dla obrazów typu I

Drugim modelem wybranym do aproksymacji danych z trzeciego przedziału wartości Q , podobnie jak dla dwóch pierwszych przedziałów, jest funkcja potęgowa postaci:

$$B_{\text{śred}}(Q, a, b) = a \cdot Q^b, \quad (4.14)$$

gdzie:

- a, b to parametry modelu,
- Q to szerokość przedziałów kwantyzacji,
- $B_{\text{śred}}$ to średnia liczba bitów reprezentacji obrazu typu I.

Parametry modelu potęgowego, tak jak dla poprzednich przedziałów, zostały wyznaczone poprzez minimalizację maksymalnego względnego błędu aproksymacji. W tabeli 4.44 podano wartości parametrów modelu potęgowego, maksymalny i średni względny błąd aproksymacji dla obrazów typu I, dla wszystkich sekwencji testowych. Wynika z niej, że dokładność zaproponowanego modelu dla obrazów typu I jest wysoka (średni względny błąd dopasowania dla większości sekwencji testowych poniżej 3%).

Podobnie, jak dla dwóch pierwszych przedziałów, w pracy nie zamieszczono wykresów dla modelu potęgowego.

Tabela 4.44. Wartości parametrów modelu potęgowego oraz maksymalny i średni względny błąd aproksymacji dla obrazów typu I (odcinek 3)

Nazwa sekwencji	Obrazy typu I			
	a	b	względny błąd dopasowania [%]	
			maksymalny	średni
basket	15 691 286	-1,1443	3,29	2,00
bluesky	19 035 250	-1,1948	5,78	3,33
bus	8 613 433	-1,1211	2,61	1,59
cheer	6 138 430	-0,9384	4,09	2,09
city	47 926 074	-1,5350	2,63	1,85
crew	1 973 212	-0,9403	4,90	2,96
flow	20 024 940	-1,1205	4,33	2,64
football	1 157 968	-0,8552	3,08	2,12
harbour	17 597 885	-1,2462	3,31	1,72
ice	821 657	-0,8009	2,34	1,48
icon	548 436	-0,8350	5,24	2,81
pedestrian	2 319 882	-1,0265	3,80	2,21
riverbed	16 549 830	-1,3254	2,67	1,35
rushhour	1 979 027	-0,9577	2,47	1,02
soccer	6 264 564	-1,1940	4,07	2,51
station2	15 206 879	-1,4552	4,87	2,80
stefan	9 965 070	-1,1085	4,38	2,48
sunflower	3 024 721	-0,9093	4,15	2,34
tractor	7 382 940	-1,1004	4,23	2,38
universal	694 358	-0,7468	5,82	3,05
warner	1 142 917	-0,8179	3,90	1,66

4.5.2.2. Model dla obrazów typu P

Wartości parametrów oraz dokładność modelu potęgowego (wzór 4.14) wykorzystanego do aproksymacji danych z przedziału trzeciego dla obrazów typu P zebrano w tabeli 4.45. Średni względny błąd aproksymacji dla większości badanych sekwencji wizyjnych jest poniżej 3,3%. Wyjątkiem są sekwencje *city* i *station2*, dla których średni względny błąd dopasowania wynosi odpowiednio 4,73 i 9,05%.

Tabela 4.45. Wartości parametrów modelu potęgowego oraz maksymalny i średni względny błąd aproksymacji dla obrazów typu P (odcinek 3)

Nazwa sekwencji	Obrazy typu P			
	a	b	względny błąd dopasowania [%]	
			maksymalny	średni
basket	3 889 235	-1,1234	3,04	2,03
bluesky	493 209	-0,8767	5,19	3,13
bus	1 540 615	-1,0291	2,99	1,62
cheer	9 404 629	-1,2345	5,07	3,06
city	183 232	-0,7175	7,85	4,73
crew	1 983 974	-1,0837	3,16	1,94
flow	1 397 745	-1,0309	3,16	1,99
football	1 508 841	-1,0238	3,41	1,95
harbour	22 254 005	-1,5885	5,13	2,81
ice	683 042	-0,9460	2,59	1,78
icon	206 012	-0,8349	2,34	1,20
pedestrian	652 418	-0,9283	2,63	1,54
riverbed	24 036 035	-1,4612	3,09	1,76
rushhour	1 935 509	-1,2179	3,78	1,89
soccer	1 711 327	-1,0448	2,93	1,18
station2	1 958 843	-1,3561	14,39	9,05
stefan	1 511 571	-1,0847	6,83	3,13
sunflower	70 215	-0,5736	4,95	3,29
tractor	4 765 863	-1,2388	6,16	3,15
universal	601 094	-0,8551	3,23	1,85
warner	859 176	-0,9260	3,31	1,84

4.5.2.3. Model dla obrazów typu B

Podobnie jak dla obrazów typu I i P, dokładność oraz wartości parametrów modelu potęgowego z dwoma parametrami (wzór 4.14), wyznaczone dla obrazów typu B, dla wszystkich sekwencji testowych, zebrano w tabeli 4.46. Z analizy danych zawartych w tabeli wynika, że średni względny błąd aproksymacji dla większości sekwencji wizyjnych jest poniżej 8%. Wyjątkiem są sekwencje *crew*, *flow* i *tractor* ze średnim względnym błędem dopasowania powyżej 8,5% oraz *city* i *station2* ze średnim względnym błędem aproksymacji powyżej 16%.

Tabela 4.46. Wartości parametrów modelu potęgowego oraz maksymalny i średni względny błąd aproksymacji dla obrazów typu B (odcinek 3)

Nazwa sekwencji	Obrazy typu B			
	a	b	względny błąd dopasowania [%]	
			maksymalny	średni
basket	3 457 419	-1,3971	6,52	3,27
bluesky	22 305	-0,4852	3,12	1,78
bus	595 532	-1,1139	3,26	1,82
cheer	57 331 127	-1,9591	8,33	4,40
city	9 129 442	-1,8556	30,81	18,39
crew	53 937 465	-2,1153	12,70	8,56
flow	1 155 214	-1,3057	16,00	9,98
football	3 268 705	-1,4414	13,46	7,86
harbour	6 501 047	-1,7348	6,93	2,96
ice	477 676	-1,0994	3,16	2,10
icon	192 970	-1,0442	3,45	2,21
pedestrian	1 923 773	-1,3968	7,96	5,76
riverbed	241 370 000	-2,1280	7,16	3,73
rushhour	4 139 452	-1,7874	5,04	2,62
soccer	1 139 496	-1,1557	2,77	1,71
station2	909 548	-1,6379	28,15	16,76
stefan	687 720	-1,2418	4,92	2,58
sunflower	71 817	-0,9108	8,94	5,44
tractor	291 680 000	-2,4949	21,93	11,18
universal	9 635 303	-1,7837	11,82	7,07
warner	13 709 294	-1,7939	11,73	7,41

4.5.2.4. Podsumowanie dotyczące modelu potęgowego w przedziale trzecim

Analogicznie jak w dwóch wcześniejszych przedziałach, autor przebadął model potęgowy wyrażający średnią liczbę bitów reprezentacji skompresowanego obrazu danego typu w zależności od wartości Q . Średni względny błąd aproksymacji tego modelu, dla większości badanych sekwencji wizyjnych, jest poniżej 3, 3,3 i 16% odpowiednio dla obrazów typu I, P i B.

4.5.3. Wnioski dotyczące modelu kodera dla wartości Q z przedziału trzeciego

Dla danych eksperymentalnych z trzeciego przedziału wartości Q (odcinek 3), autor przebadął dwa modele dopasowując je do średniej liczby bitów potrzebnych do reprezentacji skompresowanego obrazu danego typu. Te modele to funkcja wielomianowa i potęgowa.

O dużej dokładności tych modeli świadczy uśredniony po wszystkich sekwencjach testowych, średni względny błąd dopasowania zebrany w tabeli 4.47. Dodatkowo, we wspomnianej tabeli, umieszczono wartości odchylenia standardowego dla uśrednionego błędu aproksymacji.

Tabela 4.47. Średni względny błąd aproksymacji dla zaproponowanych modeli w trzecim przedziale wartości Q uśredniony po wszystkich sekwencjach testowych oraz wartości odchylenia standardowego dla tego błędu

Typ obrazu	Typ modelu					
	wielomianowy 3-go stopnia (4 parametry)		wielomianowy 3-go stopnia (1 parametr)		potęgowy (2 parametry)	
	błąd średni [%]	odchylenie standardowe	błąd średni [%]	odchylenie standardowe	błąd średni [%]	odchylenie standardowe
I	1,72	0,36	5,54	4,48	2,21	0,61
P	2,11	0,70	7,99	11,00	2,61	1,71
B	3,53	2,38	26,07	30,17	6,08	4,78

Analiza danych z tabeli 4.47 prowadzi do wniosku, że model wielomianowy 3-go stopnia z czterema parametrami lepiej dopasowuje się do danych eksperymentalnych niż model potęgowy, który ma dwa parametry. Uproszczenie modelu wielomianowego, dokonane poprzez zredukowanie liczby parametrów, pogorszyło dokładność oszacowania danych średnio o około 4, 6 i 23 p.p. odpowiednio dla obrazów typu I, P i B. Mimo że dokładność takiego uproszczonego modelu wielomianowego jest niższa w porównaniu do modelu potęgowego, to do modelowania danych otrzymanych dla wartości Q z przedziału trzeciego autor wybrał ten model (ze względu na mniejszą liczbę parametrów).

4.6. Podsumowanie

Rozdział 4 zawiera autorskie propozycje modeli dopasowanych do średniej liczby bitów potrzebnych do reprezentowania skompresowanego obrazu danego typu. Wspomniana średnia liczba bitów jest traktowana jako cecha charakterystyczna obrazów danego typu w ramach pojedynczego ujęcia. Dlatego modele i ich parametry, zostały wyznaczone indywidualnie dla każdego typu obrazu. Ze względu na nieudane próby stworzenia modelu dla całego dozwolonego zakresu szerokości przedziałów kwantyzacji Q , zaproponowano modele niezależnie dla przyjętych przedziałów wartości Q (punkt 4.1).

W przedziale pierwszym, ze względu na mniejszą liczbę parametrów przy porównywalnej dokładności badanych modeli, wybrano model wielomianowy 3-go stopnia z 1 parametrem. Średni względny błąd aproksymacji dla większości sekwencji testowych jest poniżej 3% dla obrazów typu I, 6% dla obrazów typu P i 7% dla obrazów typu B. W przedziale drugim, gdzie szczególnie istotna jest dokładność aproksymacji danych, wybrano model hiperboliczny z 3 parametrami, ze średnim względnym błędem dopasowania dla większości sekwencji testowych poniżej 2, 3 i 8%, odpowiednio dla obrazów typu I, P i B. W przedziale trzecim wybrano uproszczony model wielomianowy ze względu na mniejszą liczbę parametrów. Średni względny błąd aproksymacji dla tego modelu dla większości sekwencji wizyjnych jest poniżej 6, 9% i 28% odpowiednio dla obrazów typu I, P i B.

Zaproponowane modele określają średnią liczbę bitów reprezentacji obrazów w ramach pojedynczego ujęcia. Z tego względu mogłyby one posłużyć do oszacowania liczby bitów potrzebnej do reprezentacji całej grupy obrazów GOP o strukturze typowej dla zastosowań telewizyjnych.

Autorowi nie są znane żadne pozycje literaturowe, w których dokonano modelowania średniej liczby bitów potrzebnej do reprezentowania skompresowanego obrazu danego typu w ramach pojedynczego ujęcia.

Rozdział 5

Modelowanie liczby bitów reprezentacji pojedynczego obrazu

5.1. Wprowadzenie

W poprzednim rozdziale przedstawiono nowe modele wyrażające, liczoną w ramach jednego ujęcia, średnią liczbę bitów potrzebnych do reprezentacji obrazu określonego typu w zależności od szerokości przedziałów kwantyzacji Q . Duża dokładność zaproponowanych modeli skłoniła autora do przeprowadzenia kolejnych badań mających na celu sprawdzenie, czy takie same modele można wykorzystać do opisywania liczby bitów reprezentujących pojedynczy obraz danego typu. Dodatkową motywacją do podjęcia takich badań była obserwacja dotycząca dużego podobieństwa obrazów w ramach jednego ujęcia. Jeśli model dokładnie opisuje średnią liczbę bitów reprezentacji obrazu wyznaczoną na podstawie obrazów o podobnej treści, to można domniemywać, że sprawdzi się też w przypadku osobnego modelowania każdego z tych obrazów.

Rozdział przedstawia wyniki badań własnych, których celem było sprawdzenie, czy modele matematyczne zaproponowane w rozdziale 4 można zastosować do opisywania zależności między liczbą bitów potrzebnych do reprezentacji obrazu danego typu a szerokością przedziałów kwantyzacji Q . Ogólnie taki model można opisać:

$$B_{obraz} = f(Q), \quad (5.1)$$

gdzie:

- B_{obraz} to liczba bitów reprezentacji obrazu pewnego typu (I, P lub B),
- Q to szerokość przedziałów kwantyzacji.

Ze względu na duży zbiór danych eksperymentalnych w pracy nie zamieszczono wyników cząstkowych dotyczących poszczególnych obrazów (analizowano ponad 5,5 tysiąca obrazów). Wyniki badań przedstawiono poprzez wyznaczenie wartości kilku wybranych statystyk opisujących uzyskane średnie względne błędy aproksymacji. Te wybrane statystyki, to wartości: maksymalna, minimalna i średnia oraz odchylenie standardowe [Ciec07, Koro01] wyznaczone niezależnie dla obrazów typu I, P i B. Dodatkowo autor wyznaczał wartości wymienionych statystyk dla sytuacji, w której wartości parametrów modelu dla bieżącego obrazu są wyznaczone na podstawie wartości parametrów z poprzedniego obrazu danego typu w ramach tego samego ujęcia. Wartości statystyk uzyskane dla modelu z parametrami wyznaczonymi indywidualnie dla każdego obrazu będą określane mianem „dopasowania wprost”, a uzyskane przy predykcji wartości parametrów z poprzedniego obrazu tego samego typu, będą określane mianem „predykcji parametrów”.

W kolejnych punktach zostaną przedstawione wyniki badań uzyskane niezależnie dla każdego z trzech przedziałów wartości Q (tabela 4.1).

5.2. Modelowanie kodera dla wartości Q z przedziału pierwszego (odcinek 1)

5.2.1. Wykorzystywany model

W rozdziale 4 dane eksperymentalne otrzymane dla wartości Q z pierwszego przedziału zostały opisane modelem wielomianowym 3-go stopnia z 4 parametrami, który następnie został uproszczony do tylko jednego parametru d . Dlatego do opisu zależności między liczbą bitów reprezentacji pojedynczego obrazu danego typu a szerokością przedziałów kwantyzacji Q dla danych otrzymanych w opisywanym przedziale również zaproponowano zastosowanie modelu wielomianowego 3-go stopnia z 4 parametrami postaci:

$$B_{obraz}(Q, a, b, c, d) = a \cdot Q^3 + b \cdot Q^2 + c \cdot Q + d, \quad (5.2)$$

gdzie:

- a, b, c, d to parametry modelu,
- Q to szerokość przedziałów kwantyzacji,
- B_{obraz} to liczba bitów reprezentacji obrazu.

Parametry modelu wyznaczono minimalizując maksymalny względny błąd aproksymacji metodą quasi-newtonowska (patrz punkt 1.3) niezależnie dla każdego obrazu:

$$\min_{a,b,c,d} \max_Q \frac{|B_x(Q) - B_{obraz}(Q, a, b, c, d)|}{B_x(Q)} \cdot 100\%, \quad (5.3)$$

gdzie:

- a, b, c, d to parametry modelu,
- $B_x(Q)$ to zmierzona liczba bitów reprezentacji obrazu,
- $B_{obraz}(Q, a, b, c, d)$ to wartość oszacowania liczby bitów reprezentacji obrazu modelem wielomianowym 3-go stopnia z 4 parametrami.

Następnie, mając wyznaczone wartości parametrów a, b, c i d , wyznaczono współczynniki korelacji liniowej dla poszczególnych par parametrów niezależnie dla obrazów typu I, P i B. Na podstawie tak wyznaczonych wartości współczynników korelacji autor podejmował decyzję o ewentualnym uproszczeniu modelu do jednego parametru.

5.2.2. Eksperymentalne wyznaczenie parametrów

W celu zebrania danych eksperymentalnych autor przeprowadził badania analogiczne do tych opisanych w punkcie 4.2. Jednak tym razem, podczas każdego kodowania, zbierano informacje dotyczące liczby bitów potrzebnych do reprezentacji pojedynczych obrazów. Dla tak zebranych danych wyznaczono parametry modelu wielomianowego (wzór 5.2) stosując minimalizację maksymalnego względnego błędu aproksymacji (wzór 5.3). Następnie wyliczono wartości odpowiednich współczynników korelacji. Wartości tych współczynników dla poszczególnych par parametrów wyznaczone niezależnie dla obrazów typu I, P i B zebrano w tabeli 5.1.

Tabela 5.1. Wartości współczynnika korelacji liniowej dla poszczególnych par parametrów modelu wielomianowego dla obrazów typu I, P i B (odcinek 1)

Parametry modelu	a	b	c	d
Obrazy typu I				
a	1	-0,9966	0,9785	-0,9293
b	-0,9966	1	-0,9919	0,954
c	0,9785	-0,9919	1	-0,9827
d	-0,9293	0,954	-0,9827	1
Obrazy typu P				
a	1	-0,9984	0,9899	-0,9544
b	-0,9984	1	-0,9959	0,9602
c	0,9899	-0,9959	1	-0,9736
d	-0,9544	0,9602	-0,9736	1
Obrazy typu B				
a	1	-0,9955	0,9738	-0,8817
b	-0,9955	1	-0,989	0,8884
c	0,9738	-0,989	1	-0,9247
d	-0,8817	0,8884	-0,9247	1

Ponieważ parametry są silnie skorelowane, to autor zdecydował o uproszczeniu modelu do jednego parametru d :

$$B_{obraz}(Q, d) = (a_1d + a_2) \cdot Q^3 + (b_1d + b_2) \cdot Q^2 + (c_1d + c_2) \cdot Q + d, \quad (5.4)$$

gdzie:

- d to parametr modelu,
- B_{obraz} to liczba bitów reprezentacji obrazu danego typu,
- Q to szerokość przedziałów kwantyzacji,
- a_1, a_2, b_1, b_2, c_1 i c_2 to uniwersalne stałe modelu wyznaczone niezależnie dla obrazów różnych typów.

W tabeli 5.2 zebrano wartości uniwersalnych stałych modelu dla obrazów typu I, P i B.

Tabela 5.2. Wartości stałych uproszczonego modelu wielomianowego dla obrazów typu I, P i B
(odcinek 1)

Stała	Wartości stałych modelu		
	Obrazy typu I	Obrazy typu P	Obrazy typu B
<i>a1</i>	-0,0165	-0,0239	-0,0185
<i>a2</i>	4584,8013	14911,3634	3963,4159
<i>b1</i>	0,1661	0,2199	0,1726
<i>b2</i>	-26281,1108	-86757,4069	10671,5321
<i>c1</i>	-0,5793	-0,6821	-0,5698
<i>c2</i>	-29441,8191	32090,8272	-208040,6235

Ze względu na duży zbiór danych eksperymentalnych w pracy nie zamieszczono wyników cząstkowych dotyczących poszczególnych obrazów. Ograniczono się do podania wybranych statystyk opisujących średni względny błąd aproksymacji dla modelu wielomianowego oraz uproszczonego modelu wielomianowego. Wybrane statystyki opisujące średni względny błąd aproksymacji, to wartości: maksymalna, minimalna i średnia oraz odchylenie standardowe [Ciec07, Koro01] wyznaczone niezależnie dla obrazów typu I, P i B. Statystyki te zebrano w tabeli 5.3. W tabeli tej pokazano także wartości wymienionych statystyk średniego względnego błędu aproksymacji dla sytuacji, w której wartości parametrów modelu w aktualnie kodowanym obrazie uzyskuje się z wcześniejszego obrazu tego samego typu w ramach pojedynczego ujęcia (predykcja parametrów).

Tabela 5.3. Statystyki średniego względnego błędu aproksymacji danych eksperymentalnych wyznaczone dla bieżącego obrazu i przy predykcji parametrów z poprzedniego obrazu danego typu dla odcinka 1 (cyfra w nawiasie oznacza liczbę parametrów modelu)

Model	Dopasowanie wprost				Predykcja parametrów			
	Średni względny błąd dopasowania [%]							
	maksymalny	minimalny	średni	odchylenie standardowe	maksymalny	minimalny	średni	odchylenie standardowe
Obrazy typu I								
wielomianowy(4)	2,27	0,80	1,52	0,36	148,51	1,08	7,23	12,99
wielomianowy(1)	70,02	1,16	3,84	7,98	172,68	1,33	8,76	15,63
Obrazy typu P								
wielomianowy(4)	9,11	1,77	3,67	1,00	433,5	1,92	6,19	14,47
wielomianowy(1)	152,16	2,79	7,93	14,50	468,95	2,91	9,78	20,63
Obrazy typu B								
wielomianowy(4)	11,59	1,60	3,29	1,39	407,4	1,63	7,17	11,34
wielomianowy(1)	197,15	2,43	9,29	20,00	445,15	2,50	12,26	22,62

Z analizy danych z tabeli 5.3 można wywnioskować, że zarówno model wielomianowy z czterema parametrami, jak i uproszczony model wielomianowy z jednym parametrem dobrze aproksymują liczbę bitów potrzebnych do reprezentacji pojedynczego obrazu danego typu. Jednak zdarzają się obrazy, które trudno opisać uproszczonym modelem wielomianowym. Obrazy te pochodzą z sekwencji *icon*. Natomiast przy predykcji parametrów z wcześniejszego obrazu danego typu należącego do tego samego ujęcia dokładność aproksymacji danych pogarsza się średnio o około 3 – 6 punktów procentowych. Również i przy takim wyznaczaniu parametrów może się zdarzyć, że natrafi się na obrazy, w których wspomniany prosty mechanizm predykcji wartości parametrów nie pozwoli uzyskać satysfakcjonujących rezultatów (również sekwencja *icon*).

Dodatkowo dokonano weryfikacji dokładności uproszczonego modelu wielomianowego z wykorzystaniem kontrolnego zbioru sekwencji wizyjnych. Przeprowadzono badania dla zbioru 8 sekwencji wizyjnych formatu 4CIF, które nie występowały w zbiorze testowym użytym do wyznaczenia wartości uniwersalnych stałych. Podczas badań wyznaczono wartości parametru d oraz maksymalne i średnie względne błędy aproksymacji danych eksperymentalnych otrzymanych dla kontrolnego zbioru sekwencji wizyjnych z wykorzystaniem uniwersalnych stałych z tabeli 5.2. Otrzymane w ten sposób wyniki zebrano w tabeli 5.4.

Tabela 5.4. Statystyki średniego względnego błędu aproksymacji danych z wykorzystaniem uproszczonego modelu wielomianowego dla obrazów typu I, P i B dla kontrolnego zbioru sekwencji wizyjnych (odcinek 1)

Model	Średni względny błąd aproksymacji [%]			
	maksymalny	minimalny	średni	odchylenie standardowe
	Obrazy typu I			
wielomianowy(1)	7,21	1,50	2,80	1,33
	Obrazy typu P			
wielomianowy(1)	19,19	2,77	8,03	3,44
	Obrazy typu B			
wielomianowy(1)	34,06	2,81	14,09	9,10

Zaproponowany uproszczony model wielomianowy z zestawem uniwersalnych stałych dobrze aproksymuje dane eksperymentalne uzyskane dla sekwencji wizyjnych spoza zbioru testowego. Świadczą o tym statystyki zebrane w tabeli 5.4. Średni względny błąd aproksymacji badanego modelu dla obrazów typu I, P i B wynosi odpowiednio w przybliżeniu 3, 8 i 14%.

5.3. Modelowanie kodera dla wartości Q z przedziału drugiego (odcinek 2)

W poprzednim rozdziale, do opisu danych otrzymanych dla drugiego przedziału wartości Q zaproponowano model hiperboliczny z 3 parametrami. Zastosowanie tego modelu do opisu zależności między liczbą bitów reprezentacji pojedynczego obrazu danego typu a szerokością przedziałów kwantyzacji Q można przedstawić w postaci:

$$B_{obraz}(Q, a, b, c) = \frac{a}{Q^b + c}, \quad (5.5)$$

gdzie:

- a, b, c to parametry modelu,
- Q to szerokość przedziałów kwantyzacji,
- B_{obraz} to liczba bitów reprezentacji obrazu.

Parametry modelu zostały wyznaczone tak jak poprzednio, poprzez minimalizację maksymalnego względnego błędu aproksymacji niezależnie dla każdego obrazu. Statystyki średniego względnego błędu aproksymacji, analogiczne do wyznaczonych dla wartości Q z przedziału pierwszego, zebrano w tabeli 5.5.

Tabela 5.5. Statystyki średniego względnego błędu aproksymacji danych eksperymentalnych wyznaczone dla bieżącego obrazu i przy predykcji parametrów z poprzedniego obrazu danego typu (odcinek2)

Typ obrazu	Dopasowanie wprost				Predykcja parametrów			
	Średni względny błąd dopasowania [%]							
	maksymalny	minimalny	średni	odchylenie standardowe	maksymalny	minimalny	średni	odchylenie standardowe
I	7,36	0,81	1,30	0,55	515,89	0,77	11,77	34,22
P	35,13	0,92	3,11	3,03	825,23	0,98	11,68	29,57
B	55,12	0,89	7,60	6,77	1585,37	1,27	21,95	44,05

Na podstawie wyników zamieszczonych w tabeli 5.5, można wywnioskować, że zaproponowany model hiperboliczny dobrze aproksymuje liczbę bitów potrzebnych do

reprezentacji obrazów typu I, P i B dla większości sekwencji wizyjnych. Dokładność modelu przy zastosowaniu predykcji wartości parametrów z wcześniejszego obrazu tego samego typu w ramach pojedynczego ujęcia jest niższa średnio o około 11 – 15 punktów procentowych. Największe wartości względnego błędu aproksymacji otrzymano dla sekwencji *icon*, co wskazuje na to, że treść tej sekwencji zmienia się stosunkowo szybko z obrazu na obraz, mimo że obrazy należą do jednego ujęcia.

5.4. Modelowanie kodera dla wartości Q z przedziału trzeciego (odcinek 3)

W rozdziale 4 dane eksperymentalne otrzymane dla wartości Q z trzeciego przedziału zostały opisane modelem wielomianowym 3-go stopnia z czterema parametrami. Po przeprowadzeniu dodatkowych badań, model ten został uproszczony do tylko jednego parametru d . Dlatego do opisu zależności między liczbą bitów reprezentacji pojedynczego obrazu danego typu a szerokością przedziałów kwantyzacji Q dla danych otrzymanych w opisywanym przedziale również zaproponowano model wielomianowy 3-go stopnia z czterema parametrami postaci:

$$B_{obraz}(Q, a, b, c, d) = a \cdot Q^3 + b \cdot Q^2 + c \cdot Q + d, \quad (5.6)$$

gdzie:

- a, b, c, d to parametry modelu,
- Q to szerokość przedziałów kwantyzacji,
- B_{obraz} to liczba bitów reprezentacji obrazu.

Jest to taki sam model jaki przyjęto do opisywania danych w pierwszym przedziale wartości Q . Dlatego, aby wyznaczyć wartości parametrów modelu, przeprowadzono badania analogiczne do tych opisanych w punkcie 5.2.2. W wyniku tych badań otrzymano wartości parametrów oraz błędów aproksymacji danych eksperymentalnych dla każdego obrazu. Analiza wartości współczynników korelacji dla poszczególnych par parametrów przeprowadzona niezależnie dla obrazów typu I, P i B pokazała, że między parametrami istnieją silne zależności liniowe (tabela 5.6). Dlatego również w tym przedziale wartości Q

autor zdecydował się na uproszczenie modelu do jednego parametru d (wzór 5.4). W tabeli 5.7 umieszczono wartości uniwersalnych stałych wyznaczonych niezależnie dla obrazów typu I, P i B.

Tabela 5.6. Wartości współczynnika korelacji liniowej dla poszczególnych par parametrów modelu wielomianowego dla obrazów typu I, P i B (odcinek 3)

Parametry modelu	a	b	c	d
Obrazy typu I				
a	1	-0,9911	0,9582	-0,8822
b	-0,9911	1	-0,9877	0,9363
c	0,9582	-0,9877	1	-0,9793
d	-0,8822	0,9363	-0,9793	1
Obrazy typu P				
a	1	-0,9952	0,974	-0,918
b	-0,9952	1	-0,9913	0,9508
c	0,974	-0,9913	1	-0,9825
d	-0,918	0,9508	-0,9825	1
Obrazy typu B				
a	1	-0,9977	0,9887	-0,9686
b	-0,9977	1	-0,9964	0,9817
c	0,9887	-0,9964	1	-0,9935
d	-0,9686	0,9817	-0,9935	1

Tabela 5.7. Wartości stałych modelu dla obrazów typu I, P i B (odcinek 3)

Stała	Wartości stałych modelu		
	Obrazy typu I	Obrazy typu P	Obrazy typu B
$a1$	-0,0000000706	-0,0000001254	-0,0000001487
$a2$	-0,0021123221	0,0008503749	0,0005083291
$b1$	0,0000457856	0,0000701522	0,0000810765
$b2$	0,8899953917	-0,4216033093	-0,2354180923
$c1$	-0,0106746354	-0,0136141864	-0,0150593010
$c2$	-84,6368623339	62,1364053028	29,7421559401

Ze względu na dużą liczbę eksperymentów również i dla tego przedziału wartości Q podano wybrane statystyki średniego względnego błędu aproksymacji (dopasowanie wprost) oraz statystyki tego błędu, gdy wartości parametrów modelu ustala się na podstawie wcześniejszego obrazu tego samego typu w ramach jednego ujęcia (predykcja parametrów). Uzyskane wyniki zebrano w tabeli 5.8.

Tabela 5.8. Statystyki średniego względnego błędu aproksymacji danych eksperymentalnych wyznaczone dla bieżącego obrazu i przy predykcji parametrów z poprzedniego obrazu danego typu dla odcinka 3 (cyfra w nawiasie oznacza liczbę parametrów modelu)

Model	Dopasowanie wprost				Predykcja parametrów			
	Średni względny błąd dopasowania [%]							
	maksymalny	minimalny	średni	odchylenie standardowe	maksymalny	minimalny	średni	odchylenie standardowe
Obrazy typu I								
wielomianowy(4)	4,40	0,73	1,62	0,49	529,85	0,95	10,53	33,28
wielomianowy(1)	38,80	1,15	6,06	4,97	530,05	1,36	13,03	33,13
Obrazy typu P								
wielomianowy(4)	34,52	0,88	3,20	2,57	809,94	1,67	15,91	42,53
wielomianowy(1)	380,36	1,57	14,70	31,38	824,27	2,24	22,35	48,85
Obrazy typu B								
wielomianowy(4)	42,25	0,00	6,54	4,62	1591,01	0,00	28,61	62,80
wielomianowy(1)	164,15	2,78	38,32	35,64	1460,92	3,64	50,17	62,79

Z danych przedstawionych w tabeli 5.8 można wnioskować, że dokładność modelu wielomianowego z 4 parametrami jest wysoka, mimo że dla niektórych obrazów typu B błąd aproksymacji wynosi około 40%. Uproszczenie modelu dokonane poprzez redukcję liczby jego parametrów pogarsza dokładność modelu średnio o około 4, 12 i 32% odpowiednio dla obrazów typu I, P i B. Natomiast przy predykcji parametrów z wcześniejszego obrazu danego typu należącego do tego samego ujęcia dokładność aproksymacji dla modelu wielomianowego z 4 parametrami pogarsza się odpowiednio o 9, 13 i 22 p.p. dla obrazów typu I, P i B. Natomiast dla uproszczonego modelu wielomianowego dokładność aproksymacji danych jest niższa o około 7, 8 i 12 p.p. Średnia wartość błędu aproksymacji została zawyżona przez sekwencję *icon*, dla której prosta predykcja wartości parametrów okazała się nieskuteczna.

Analogicznie jak dla danych z przedziału pierwszego dokonano weryfikacji dokładności uproszczonego modelu wielomianowego z wykorzystaniem kontrolnego zbioru sekwencji wizyjnych. Eksperyment, tak jak wcześniej, polegał na wyznaczeniu wartości parametru *d* oraz maksymalnego i średniego względnego błędu aproksymacji danych eksperymentalnych otrzymanych dla kontrolnego zbioru sekwencji wizyjnych z wykorzystaniem uniwersalnych stałych zebranych w tabeli 5.7. Wyniki eksperymentu przedstawiono w tabeli 5.9.

Tabela 5.9. Statystyki średniego względnego błędu aproksymacji danych z wykorzystaniem uproszczonego modelu wielomianowego dla obrazów typu I, P i B dla kontrolnego zbioru sekwencji wizyjnych (odcinek 3)

Model	Średni względny błąd aproksymacji [%]			
	maksymalny	minimalny	średni	odchylenie standardowe
	Obrazy typu I			
wielomianowy(1)	38,80	1,25	5,37	5,35
	Obrazy typu P			
wielomianowy(1)	94,04	1,86	6,68	6,02
	Obrazy typu B			
wielomianowy(1)	310,67	2,81	28,69	27,23

Podobnie jak dla danych z przedziału pierwszego, zaproponowany uproszczony model wielomianowy z zestawem uniwersalnych stałych dobrze aproksymuje dane eksperymentalne uzyskane dla sekwencji wizyjnych spoza zbioru testowego dla obrazów typu I i P (średni względny błąd aproksymacji wynosi odpowiednio około 5 i 7%). Natomiast dla obrazów typu B średni względny błąd dopasowania wynosi 29% i jest on porównywalny do błędu uzyskanego dla zbioru sekwencji testowych.

5.5. Model z parametrem charakteryzującym złożoność treści obrazu

Mając wyznaczone modele wyrażające liczbę bitów reprezentacji obrazu, autor podjął badania, których celem było sprawdzenie, czy można wyznaczyć miarę charakteryzującą złożoność treści modelowanego obrazu powiązaną z liczbą bitów reprezentacji tego obrazu. Z literatury znane są propozycje różnych takich miar [Kwon07, Li03, Lee08a, Ma05, Zhou09]. Niestety część z proponowanych miar (np. *MAD*, *SAD*) jest wyznaczana na podstawie sygnału predykcji, a nie tylko wartości oryginalnych próbek obrazu, dlatego zostały one pominięte w dalszych badaniach.

Poszukując dobrego parametru charakteryzującego złożoność treści obrazu sprawdzano, czy badany parametr jest w prosty sposób powiązany z liczbą bitów reprezentacji tego obrazu. W tym celu wyznaczano wartość współczynnika korelacji między wybraną miarą złożoności treści obrazu a liczbą bitów reprezentacji tego obrazu dla wszystkich obrazów danego typu.

Sprawdzano także jak wartość współczynnika korelacji zmienia się wraz ze zmianą szerokości przedziałów kwantyzacji Q . Autor zaproponował i przebadął dwie własne miary złożoności treści obrazu. Dodatkowo porównał je z jedną z miar dostępnych w literaturze. Przebadane miary, to:

1. odchylenie standardowe wartości próbek składowej luminancji obrazu wyznaczone w blokach o rozmiarze 4x4 uśrednione po takich blokach w całym obrazie:

$$\overline{\sigma_Y} = \frac{1}{N_Y} \sum_{i=1}^{N_Y} \sigma_i^Y, \quad (5.7)$$

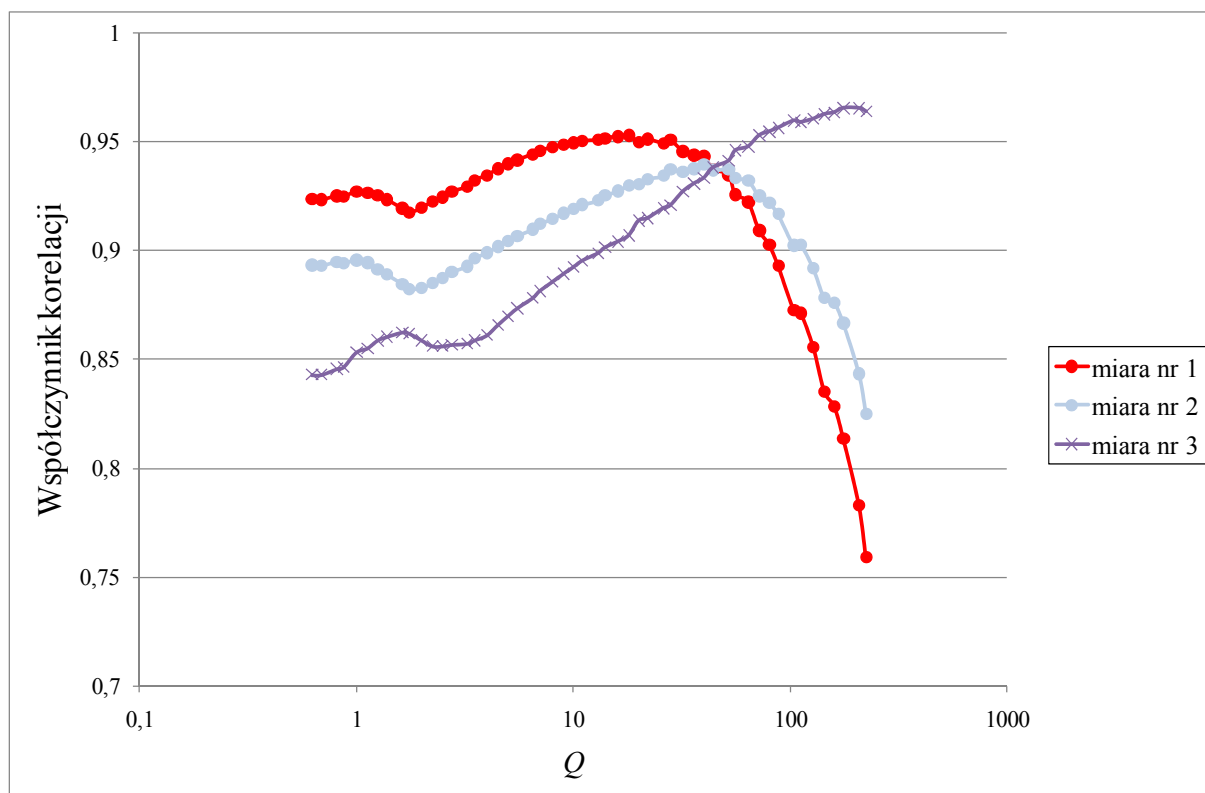
gdzie:

- $\overline{\sigma_Y}$ to odchylenie standardowe wartości próbek składowej luminancji obrazu wyznaczone w blokach o rozmiarze 4x4 punkty uśrednione po wszystkich blokach w obrazie,
 - σ_i^Y to odchylenie standardowe wartości próbek składowej luminancji obrazu w i -tym bloku o rozmiarze 4x4 punkty,
 - N_Y to liczba nienakładających się bloków luminancji o rozmiarze 4x4 punkty w obrazie,
2. odchylenie standardowe wartości próbek składowej luminancji obrazu wyznaczone w blokach o rozmiarze 16x16 uśrednione po wszystkich blokach w całym obrazie (miara określona analogicznie do wzoru 5.7, ale dla bloków o rozmiarze 16x16 punktów),
 3. wartość miary „Grad·SOH” [Zhou09] (punkt 3.5, wzory 3.13 i 3.14.).

Badania przeprowadzono dla trzech typów obrazów i całego zakresu szerokości przedziałów kwantyzacji Q . Ponieważ zadowalające wyniki uzyskano tylko dla obrazów typu I, dlatego w dalszej części przedstawiono wyniki tylko dla tych obrazów.

Na wykresie 5.1 pokazano zależność współczynnika korelacji wyznaczonego dla wspomnianych miar złożoności treści obrazu w zależności od szerokości przedziałów kwantyzacji Q dla obrazów typu I (wykres w skali logarytmicznej). Miara nr 1 oznacza pierwszą miarę z listy, miara nr 2 – drugą, a miara nr 3 – trzecią. Z analizy wykresu wynika, że dla prawie całego zakresu Q najwyższą wartość współczynnika korelacji otrzymano dla pierwszej miary z powyższej listy, czyli odchylenia standardowego wartości próbek

składowej luminancji obrazu wyznaczonego w blokach o rozmiarze 4x4 punkty uśrednione po wszystkich takich blokach w obrazie. Dlatego właśnie ta miara została wybrana do dalszych badań.



Wykres 5.1. Współczynnik korelacji miary złożoności treści obrazu i liczby bitów reprezentacji obrazu typu I dla całego zakresu wartości Q

Ponieważ drugi przedział wartości Q odpowiada użytecznemu zakresowi prędkości strumienia bitowego, dlatego dalsze badania przeprowadzono tylko dla tego odcinka.

5.5.1. Miara charakterystyczna składowej luminancji obrazu

Wyznaczoną miarę złożoności treści obrazu wykorzystano do modelowania danych otrzymanych dla obrazów typu I dla wartości Q z przedziału drugiego. Autor zaproponował rozszerzenie modelu hiperbolicznego ze wzoru 5.5 o parametr charakteryzujący złożoność treści obrazu $\overline{\sigma}_Y$:

$$B_{obraz}(Q, a, b, c) = \overline{\sigma}_Y \cdot \frac{a}{Q^b + c}, \quad (5.8)$$

gdzie:

- a, b, c to parametry modelu,
- Q to szerokość przedziałów kwantyzacji,
- B_{obraz} to liczba bitów reprezentacji obrazu,
- $\bar{\sigma}_Y$ to zaproponowana przez autora miara złożoności treści obrazu (wzór 5.7).

Dla tak określonego modelu i miary złożoności treści obrazu wyznaczono parametry modelu minimalizując maksymalny względny błąd aproksymacji. W tabeli 5.10 zebrano statystyki średniego względnego błędu aproksymacji danych z wykorzystaniem zaproponowanego rozszerzonego modelu hiperbolicznego. Dodatkowo w tabeli tej pokazano statystyki średniego względnego błędu aproksymacji dla sytuacji, w której wartości parametrów modelu w aktualnie kodowanym obrazie uzyskuje się z wcześniejszego obrazu typu I w obrębie jednego ujęcia.

Tabela 5.10. Statystyki średniego względnego błędu aproksymacji danych eksperymentalnych wyznaczone dla bieżącego obrazu typu I i przy predykcji parametrów z poprzedniego obrazu typu I dla rozszerzonego modelu hiperbolicznego

Model	Dopasowanie wprost				Predykcja parametrów			
	Średni względny błąd dopasowania [%]							
	maksymalny	minimalny	średni	odchylenie standardowe	maksymalny	minimalny	średni	odchylenie standardowe
hiperboliczny rozszerzony (3 parametry + $\bar{\sigma}_Y$)	9,84	0,78	1,34	0,73	213,27	0,90	4,48	13,86

Porównując dane z tabel 5.5 i 5.10 można wywnioskować, że dodanie parametru charakteryzującego złożoność treści obrazu poprawia dokładności modelu średnio o około 7 p.p. w przypadku, gdy wartości parametrów modelu dla bieżącego obrazu ustala się na podstawie wartości tych parametrów z wcześniejszego obrazu typu I z tego samego ujęcia.

5.5.2. Miara charakterystyczna składowych luminancji i chrominancji

Zaproponowana w poprzednim punkcie miara złożoności treści obrazu została wyznaczona tylko dla składowej luminancji. Z tego powodu autor postanowił sprawdzić, czy

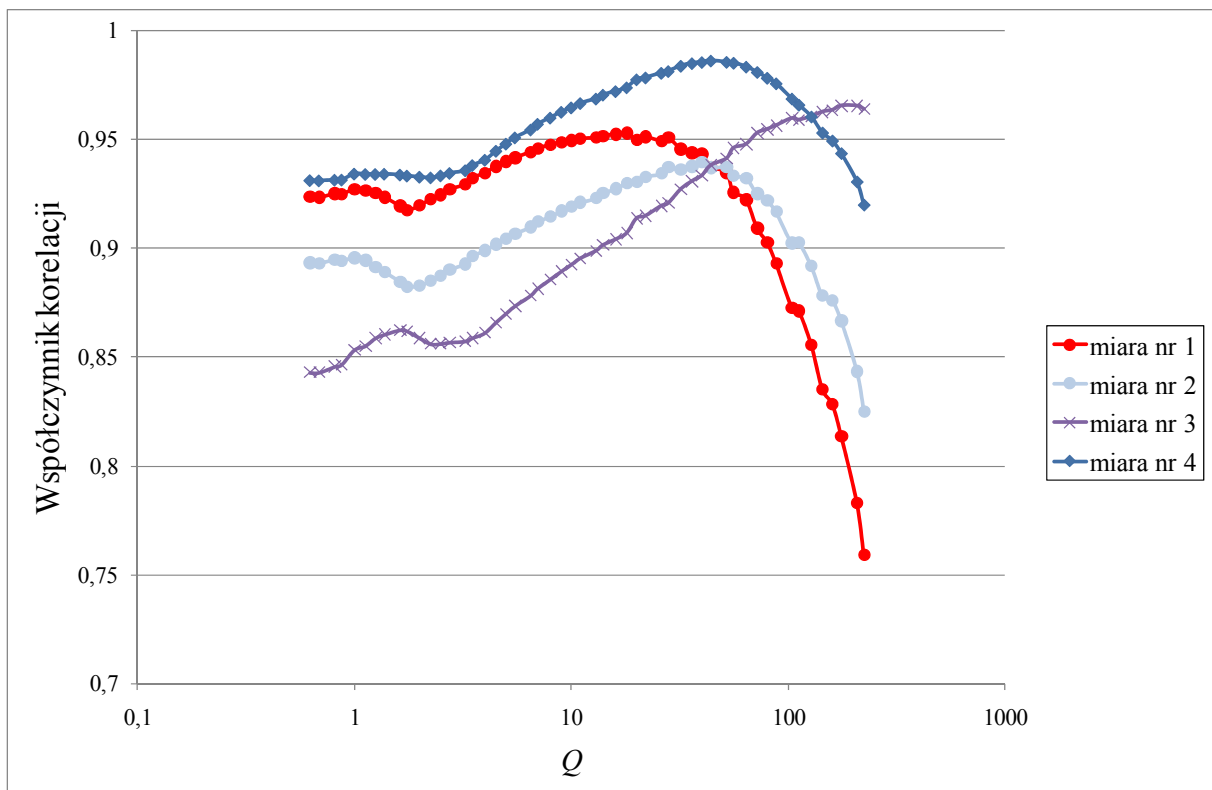
na dokładność modelu wpłynie wyznaczenie tej miary także dla składowych chrominancji. Dlatego zaproponowano nową miarę wyznaczoną jako:

$$\overline{\sigma_{YCbCr}} = \frac{1}{N_Y} \sum_{i=1}^{N_Y} \sigma_i^Y + \frac{1}{N_{Cb}} \sum_{i=1}^{N_{Cb}} \sigma_i^{Cb} + \frac{1}{N_{Cr}} \sum_{i=1}^{N_{Cr}} \sigma_i^{Cr}, \quad (5.9)$$

gdzie:

- $\overline{\sigma_{YCbCr}}$ to suma uśrednionych odchyłeń standardowych wartości próbek składowych Y, Cb i Cr obrazu wyznaczonych w blokach o rozmiarze 4x4 punkty,
- $\sigma_i^Y, \sigma_i^{Cb}, \sigma_i^{Cr}$ to odchylenie standardowe wartości próbek składowej odpowiednio Y, Cb i Cr obrazu w i -tym bloku o rozmiarze 4x4 punkty,
- N_Y, N_{Cb}, N_{Cb} to liczba nienakładających się bloków o rozmiarze 4x4 punkty składowej odpowiednio Y, Cb i Cr.

Na wykresie 5.2 przedstawiono zależność współczynnika korelacji wyznaczonego dla nowej miary oraz poprzednich miar złożoności treści obrazu w zależności od szerokości przedziałów kwantyzacji Q . Na wykresie miara $\overline{\sigma_{YCbCr}}$ została określona jako „miara nr 4”.



Wykres 5.2. Współczynnik korelacji miary złożoności treści obrazu i liczby bitów reprezentacji obrazu dla całego zakresu wartości Q

Ponieważ dla nowej miary uzyskano wyższe wartości współczynnika korelacji niż dla wcześniejszej miary, to autor przeprowadził badania dla rozszerzonego modelu hiperbolicznego z wykorzystaniem zaproponowanej miary (czyli miara $\overline{\sigma_Y}$ została zastąpiona miarą $\overline{\sigma_{YCbCr}}$). Wartości parametrów oraz dokładność rozszerzonego modelu hiperbolicznego z miarą $\overline{\sigma_{YCbCr}}$ wyznaczono, podobnie jak wcześniej, minimalizując maksymalny względny błąd aproksymacji. W tabeli 5.11 zebrano statystyki średniego względnego błędu aproksymacji danych z wykorzystaniem zaproponowanej miary złożoności treści obrazu oraz przy predykcji wartości parametrów z wcześniejszego obrazu typu I w obrębie tego samego ujęcia.

Tabela 5.11. Statystyki średniego względnego błędu aproksymacji danych eksperymentalnych wyznaczone dla bieżącego obrazu typu I i przy predykcji parametrów z poprzedniego obrazu typu I dla rozszerzonego modelu hiperbolicznego

Model	Dopasowanie wprost				Predykcja parametrów			
	Średni względny błąd dopasowania [%]							
	maksymalny	minimalny	średni	odchylenie standardowe	maksymalny	minimalny	średni	odchylenie standardowe
hiperboliczny rozszerzony (3 parametry + $\overline{\sigma_{YCbCr}}$)	7,38	0,78	1,32	0,60	184,20	0,87	4,64	12,72

Z analizy danych w tabeli 5.11 można wywnioskować, że podobnie jak dla pierwszej miary o około 7 p.p. zmniejszył się błąd oszacowania danych w przypadku, gdy wartości parametrów modelu dla bieżącego obrazu ustala się na wartości tych parametrów z wcześniejszego obrazu typu I.

5.6. Podsumowanie

W rozdziale przedstawiono wyniki badań własnych dotyczących modelowania liczby bitów reprezentacji obrazu danego typu. W tabeli 5.12 podsumowano wartości średniego względnego błędu aproksymacji uśrednionego po wszystkich obrazach poszczególnych

typów. Podano także wartość tak samo obliczonego średniego względnego błędu aproksymacji dla sytuacji, w której parametry modelu dla bieżącego obrazu przewidywano na podstawie wcześniejszego obrazu tego samego typu w tym samym ujęciu.

Tabela 5.12. Średni względny błąd dopasowania wyznaczony dla bieżącego obrazu i przy predykcji parametrów z poprzedniego obrazu danego typu uśredniony po obrazach danego typu (cyfra w nawiasie oznacza liczbę parametrów modelu)

Model	Obrazy typu I		Obrazy typu P		Obrazy typu B	
	Wprost	Predykcja	Wprost	Predykcja	Wprost	Predykcja
ODCINEK 1						
wielomianowy(4)	1,52	7,23	3,67	6,19	1,39	7,17
wielomianowy(1)	3,84	8,76	8,27	9,78	20,00	12,26
ODCINEK 2						
hiperboliczny(3)	1,30	11,77	3,11	11,68	7,60	21,95
ODCINEK 3						
wielomianowy(4)	1,62	10,53	3,20	15,91	6,54	28,61
wielomianowy(1)	6,06	13,03	14,70	22,35	38,32	50,17

Z tabeli 5.12 wynika, że modele zaproponowane do określania średniej liczby bitów reprezentacji obrazu danego typu (rozdział 4) można zastosować do opisywania liczby bitów reprezentacji pojedynczego obrazu. Ponadto, wartości parametrów modelu dla bieżącego obrazu można ustalać na podstawie wcześniejszego obrazu danego typu w tym samym ujęciu (wyjątek stanowi sekwencja *icon*).

Uproszczenie modelu wielomianowego do jednego parametru pogarsza dokładność aproksymacji danych odpowiednio o około 2, 5 i 19% dla obrazów typu I, P i B dla pierwszego przedziału i 4, 12 oraz 32% dla trzeciego przedziału.

Dodatkowo dla obrazów typu I autor zaproponował własne parametry charakteryzujące złożoność treści obrazów tego typu, których wykorzystanie pozwala zwiększyć dokładność modelu przy predykcji wartości parametrów z poprzedniego obrazu typu I należącego do tego samego ujęcia.

W literaturze dotyczącej modelowania hybrydowych koderów wizyjnych można znaleźć propozycje modelowania liczby bitów reprezentacji pojedynczego obrazu danego typu. Natomiast nie są dostępne implementacje takich rozwiązań, przez co nie można przeprowadzić badań porównawczych.

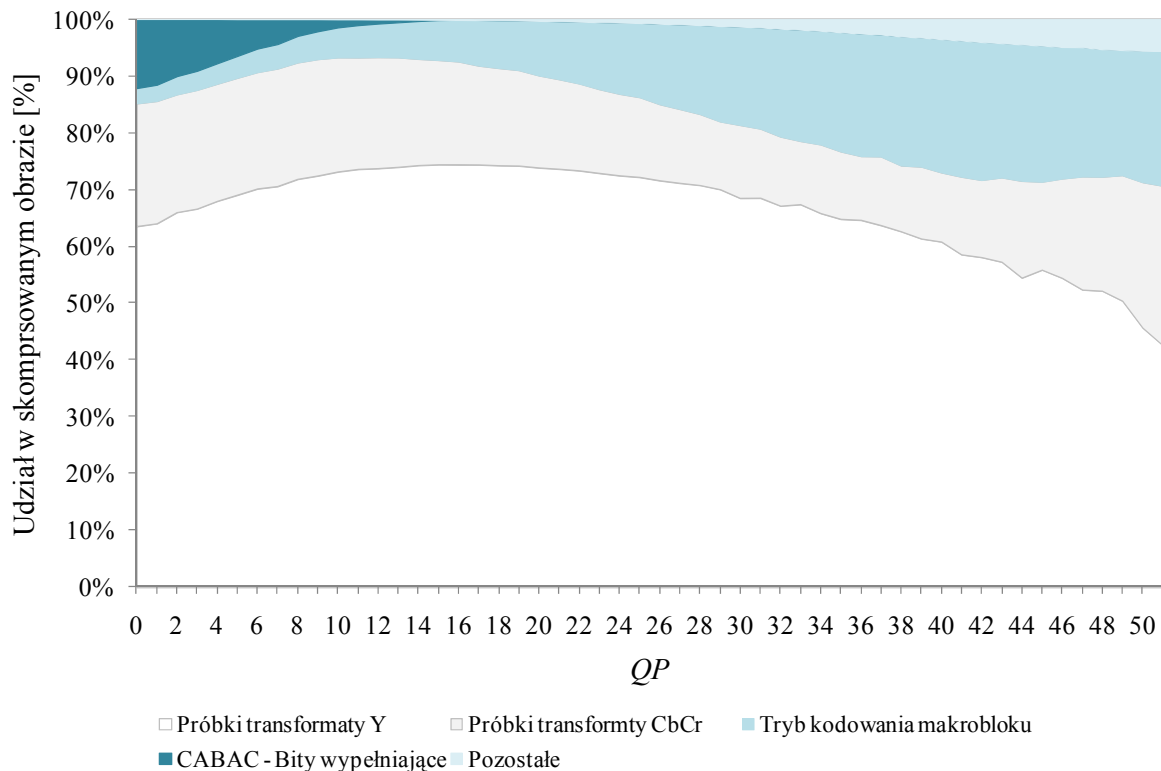
Rozdział 6

Modelowanie łącznej liczby bitów potrzebnych do reprezentacji składowych luminancji i chrominancji pojedynczego obrazu

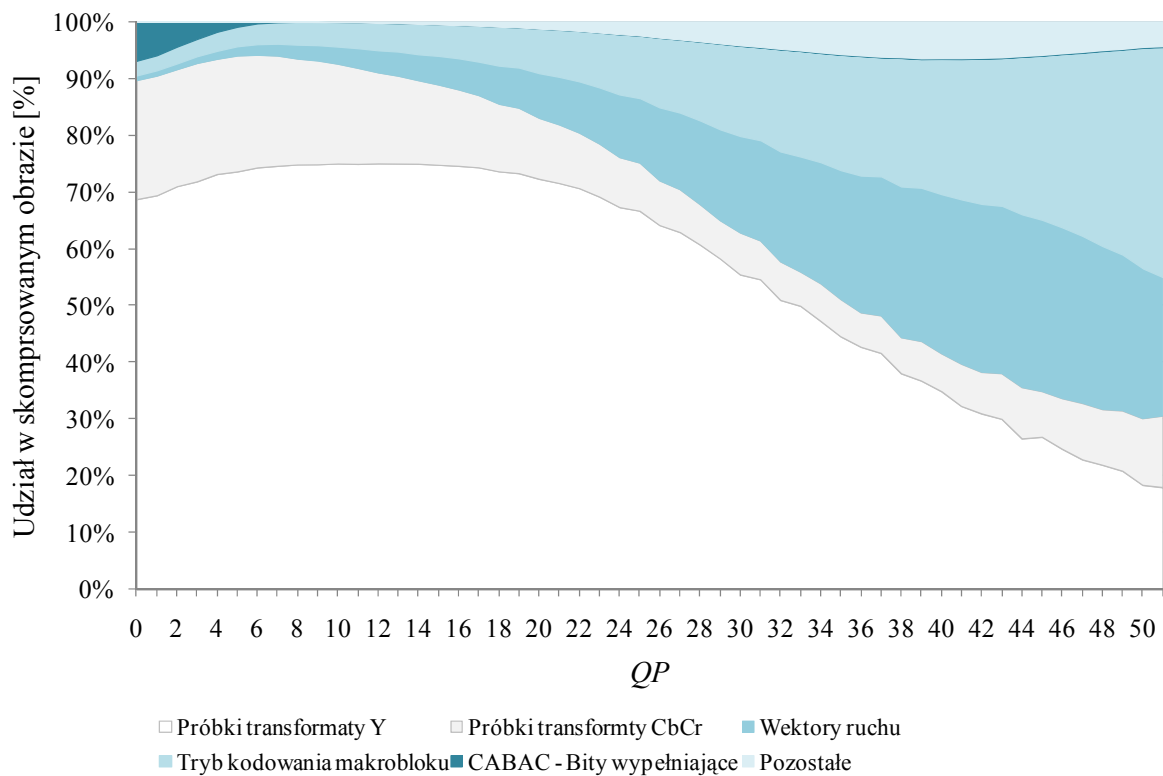
6.1. Wprowadzenie

W rozdziałach czwartym i piątym zaproponowano modele określające średnią liczbę bitów potrzebnych do reprezentacji obrazu danego typu w ramach pojedynczego ujęcia oraz liczbę bitów reprezentacji obrazu danego typu w zależności od szerokości przedziałów kwantyzacji Q . Takie podejście nie wymagało głębokiego wnikania w strukturę strumienia bitowego. Mając opracowane wyżej wspomniane modele, autor podjął dalsze badania mające na celu analizę strumienia bitowego na niższym poziomie niż obrazy. Aby zebrać informacje dotyczące liczby bitów potrzebnych do reprezentacji poszczególnych składników strumienia danych przeprowadzono badania analogiczne do tych opisanych w punkcie 4.2. Jednak tym razem, podczas każdego kodowania, zbierano informacje dotyczące liczby bitów potrzebnych do reprezentacji poszczególnych składników strumienia bitowego. Na wykresach od 6.1 do 6.3 przedstawiono procentowy udział wybranych typów danych w obrazach typu I, P i B

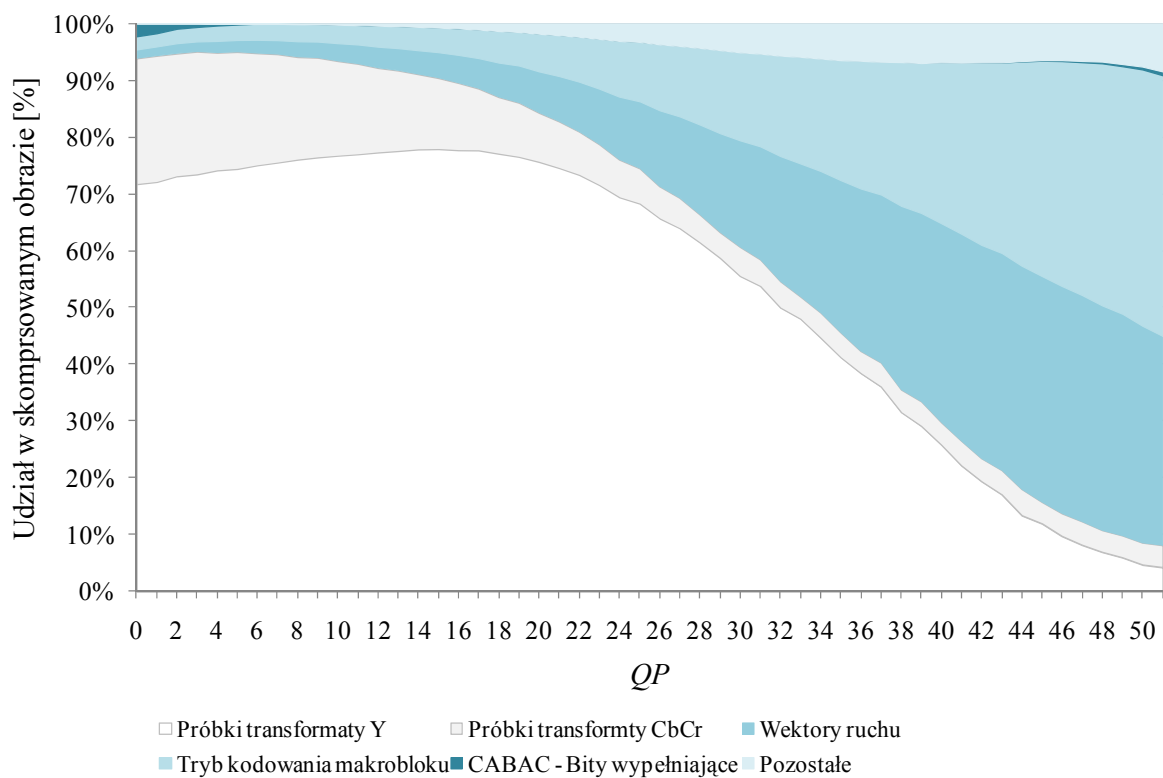
(uśredniony po wszystkich badanych sekwencjach testowych) w zależności od wartości parametru kwantyzacji QP (wybrano QP ze względu na większą przejrzystość wykresów niż dla Q). Relację między wartościami Q a QP opisano w punkcie 2.3.3. Z analizy tych wykresów wynika, że na reprezentację skwantowanych wartości próbek transformaty błędu predykcji składowych luminancji i chrominancji w obrazie typu I potrzeba od 71 do 93 % liczby bitów reprezentujących cały obraz. W obrazach typu P jest to od 30 do 94%, a w obrazach typu B od 8 do 95%. Należy zaznaczyć, że dla małych wartości szerokości przedziałów kwantyzacji Q (co odpowiada dużym prędkościom strumienia bitowego) we wszystkich typach kodowania obrazów od kilku do nawet kilkunastu procent bitów reprezentacji obrazu mogą zajmować dane określone w standardzie jako *bity wypełniające* [AVC], które nie przenoszą żadnej informacji. Takie dane występują w strumieniu tylko w przypadku kodowania danych koderem entropijnym CABAC.



Wykres 6.1. Procentowy udział poszczególnych typów danych w skompresowanych obrazach typu I w zależności od wartości indeksu QP uśredniony po badanych sekwencjach wizyjnych.

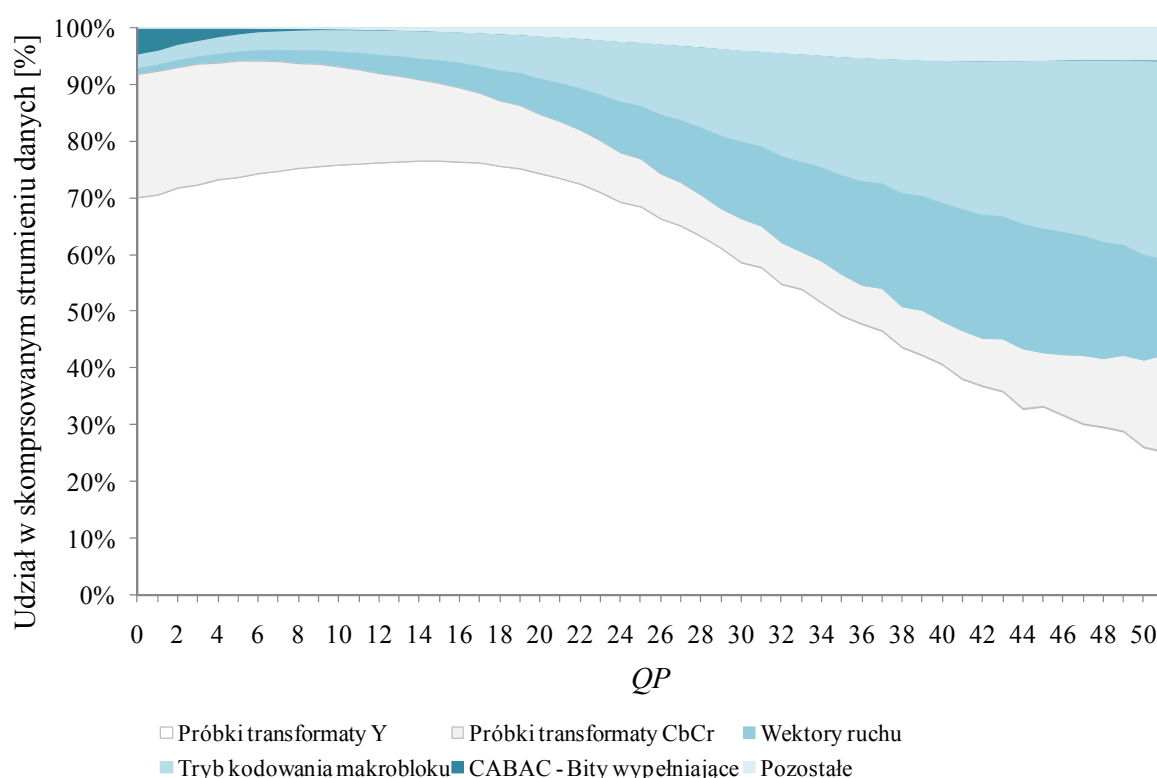


Wykres 6.2. Procentowy udział poszczególnych typów danych w skompresowanych obrazach typu P w zależności od wartości indeksu QP uśredniony po badanych sekwencjach wizyjnych.



Wykres 6.3. Procentowy udział poszczególnych typów danych w skompresowanych obrazach typu B w zależności od wartości indeksu QP uśredniony po badanych sekwencjach wizyjnych.

Na kolejnym wykresie (6.4) pokazano, uśredniony po badanych sekwencjach, udział wybranych typów danych w całkowitym strumieniu bitowym reprezentującym skompresowaną sekwencję wizyjną zakodowaną ze strukturą grupy obrazów GOP – 4P2B. Z analizy wykresu wynika, że do reprezentacji skwantowanych wartości próbek transformaty błędu predykcji składowych luminancji i chrominancji potrzeba od 41 do 94% całkowitej liczby bitów reprezentujących skompresowaną sekwencję wizyjną. Ponadto, stosunkowo dużą część całkowitego strumienia bitowego zajmują dane określające tryb kodowania makrobloku oraz wektory ruchu (łącznie od 4 do 50 %). Im większa szerokość przedziałów kwantyzacji Q , tym większy udział danych nagłówkowych w całkowitym strumieniu danych.



Wykres 6.4. Procentowy udział poszczególnych typów danych w skompresowanym strumieniu bitowym zgodnym z profilem głównym MPEG-4 AVC/H.264 w zależności od wartości indeksu QP uśredniony po badanych sekwencjach wizyjnych. Struktura GOP – 4P2B (rysunek 2.4)

Na podstawie obserwacji dotyczących udziału poszczególnych typów danych w skompresowanym strumieniu bitowym autor podjął dalsze badania mające na celu opisanie zależności między sumaryczną liczbą bitów potrzebnych do reprezentacji skwantowanych wartości próbek transformaty błędu predykcji składowych luminancji i chrominancji a szerokością przedziałów kwantyzacji Q . Także w referencyjnej implementacji kodera standardu MPEG-4 AVC/H.264 [AVCref] modeluje się wspomnianą zależność, ale tylko dla

obrazów typu P [Li03, Li04]. Autor postanowił sprawdzić, czy modele zaproponowane w poprzednich rozdziałach można wykorzystać również do opisanie zależności między sumaryczną liczbą bitów potrzebnych do reprezentacji skwantowanych wartości próbek transformaty błędu predykcji składowych luminancji i chrominancji a szerokością przedziałów kwantyzacji Q .

W kolejnych punktach zostaną przedstawione wyniki badań dotyczące modelowania opisanej powyżej zależności w trzech przedziałów wartości Q . Podobnie jak w rozdziale 5, ze względu na duży zbiór danych eksperymentalnych w pracy nie zamieszczono wyników cząstkowych dotyczących poszczególnych obrazów. Wyniki badań przedstawiono poprzez wyznaczenie wartości kilku wybranych statystyk opisujących uzyskane średnie względne błędy aproksymacji danych. Te wybrane statystyki, to wartości: maksymalna, minimalna i średnia oraz odchylenie standardowe [Ciec07, Koro01] wyznaczone niezależnie dla obrazów typu I, P i B.

6.2. Modelowanie kodera dla wartości Q z przedziału pierwszego (odcinek 1)

Do opisu zależności między sumaryczną liczbą bitów potrzebnych do reprezentacji składowych luminancji i chrominancji pojedynczego obrazu danego typu a szerokością przedziałów kwantyzacji Q autor zastosował opisany w rozdziałach 4 i 5 model wielomianowy 3-go stopnia z 4 parametrami postaci:

$$B_{YCbCr}(Q, a, b, c, d) = a \cdot Q^3 + b \cdot Q^2 + c \cdot Q + d, \quad (6.1)$$

gdzie:

- a, b, c, d to parametry modelu,
- Q to szerokość przedziałów kwantyzacji,
- B_{YCbCr} to łączna liczba bitów reprezentacji składowych luminancji i chrominancji obrazu.

Parametry modelu zostały wyznaczone poprzez minimalizację maksymalnego względnego błędu aproksymacji metodą quasi-newtonowską (patrz punkt 1.3) niezależnie dla każdego obrazu:

$$\min_{a,b,c,d} \max_Q \frac{|B_x(Q) - B_{YCbCr}(Q, a, b, c, d)|}{B_x(Q)} \cdot 100\%, \quad (6.2)$$

gdzie:

- a, b, c, d to parametry modelu,
- $B_x(Q)$ to zmierzona liczba bitów reprezentacji obrazu,
- $B_{YCbCr}(Q, a, b, c, d)$ to wartość oszacowania liczby bitów reprezentacji składowych luminancji i chrominancji obrazu modelem wielomianowym 3-go stopnia z 4 parametrami.

Następnie analizując wartości współczynnika korelacji dla poszczególnych par parametrów, niezależnie dla obrazów typu I, P i B (tabela 6.1), autor zdecydował o uproszczeniu modelu do tylko jednego parametru d . Wartości stałych modelu dla obrazów typu I, P i B przedstawiono w tabeli 6.2.

Tabela 6.1. Wartości współczynnika korelacji liniowej dla poszczególnych par parametrów modelu wielomianowego dla obrazów typu I, P i B (odcinek 1)

Parametry modelu	a	b	c	d
Obrazy typu I				
a	1	-0,994	0,9518	-0,7235
b	-0,994	1	-0,9764	0,7515
c	0,9518	-0,9764	1	-0,8504
d	-0,7235	0,7515	-0,8504	1
Obrazy typu P				
a	1	-0,9931	0,9465	-0,6365
b	-0,9931	1	-0,9741	0,6783
c	0,9465	-0,9741	1	-0,8074
d	-0,6365	0,6783	-0,8074	1
Obrazy typu B				
a	1	-0,9919	0,9468	-0,6502
b	-0,9919	1	-0,9738	0,6755
c	0,9468	-0,9738	1	-0,801
d	-0,6502	0,6755	-0,801	1

Tabela 6.2. Wartości stałych uproszczonego modelu wielomianowego dla obrazów typu I, P i B
(odcinek 1)

Stała	Wartości stałych modelu		
	Obrazy typu I	Obrazy typu P	Obrazy typu B
$a1$	-0,006985	-0,010149	-0,009389
$a2$	-13965,624568	-18937,709740	-17898,253143
$b1$	0,075699	0,105315	0,097796
$b2$	155972,586489	195741,649750	189803,730193
$c1$	-0,329964	-0,427363	-0,410114
$c2$	-552877,647924	-600594,366407	-590877,381858

Podobnie jak w rozdziale 5, ze względu na obszerny zbiór danych eksperymentalnych, w pracy nie umieszczono wyników cząstkowych. Do opisu uzyskanych wyników, analogicznie jak w poprzednim rozdziale, wyznaczono wartości wybranych statystyk opisujących średni względny błąd aproksymacji dla modelu wielomianowego 3-go stopnia z 4 parametrami oraz uproszczonego modelu wielomianowego. Te wybrane statystyki opisujące średni względny błąd aproksymacji, to wartości: maksymalna, minimalna i średnia oraz odchylenie standardowe wyznaczone niezależnie dla obrazów typu I, P i B. Wymienione statystyki zebrano w tabeli 6.3. Dodatkowo w tabeli tej podano wartości takich samych statystyk w przypadku wykorzystania modelu z kodera referencyjnego (wzór 3.1), ale tylko dla obrazów typu P (patrz punkt 3.2).

Tabela 6.3. Statystyki średniego względnego błędu aproksymacji danych eksperymentalnych wyznaczone niezależnie dla obrazów typu I, P i B (odcinek 1)

Model	Dopasowanie wprost			
	Średni względny błąd dopasowania [%]			
	maksymalny	minimalny	średni	odchylenie standardowe
	Obrazy typu I			
wielomianowy(4)	2,49	0,64	0,98	0,29
wielomianowy(1)	102,82	0,90	4,58	11,61
	Obrazy typu P			
wielomianowy(4)	10,30	1,45	3,00	1,51
wielomianowy(1)	167,21	1,83	8,21	16,23
referencyjny(2)	29,60	1,58	8,63	5,42
	Obrazy typu B			
wielomianowy(4)	15,36	1,42	3,38	2,04
wielomianowy(1)	199,62	2,01	10,09	20,33

Dane z tabeli 6.3 pozwalają stwierdzić, że model wielomianowy z czterema parametrami oraz uproszczony model wielomianowy z jednym parametrem dobrze aproksymują liczbę bitów potrzebnych do reprezentacji składowych luminancji i chrominancji pojedynczego obrazu danego typu. Jednak zdarzają się obrazy, które trudno opisać uproszczonym modelem wielomianowym. Należy podkreślić, że dla obrazów typu P zaproponowany model wielomianowy z czterema parametrami ma wyższą dokładność w porównaniu z modelem z kodera referencyjnego (o około 5,6 p.p.).

Analogicznie jak w rozdziałach 4 i 5 dokonano weryfikacji dokładności uproszczonego modelu wielomianowego z wykorzystaniem kontrolnego zbioru sekwencji wizyjnych (dodatek A). Eksperyment, tak jak wcześniej, polegał na wyznaczeniu wartości parametru d oraz maksymalnego i średniego względnego błędu aproksymacji danych eksperymentalnych otrzymanych dla kontrolnego zbioru sekwencji wizyjnych z wykorzystaniem uniwersalnych stałych zebranych w tabeli 6.2. Wyniki eksperymentu przedstawiono w tabeli 6.4.

Tabela 6.4. Statystyki średniego względnego błędu aproksymacji danych z wykorzystaniem uproszczonego modelu wielomianowego dla obrazów typu I, P i B dla kontrolnego zbioru sekwencji wizyjnych (odcinek 1)

Model	Średni względny błąd aproksymacji [%]			
	maksymalny	minimalny	średni	odchylenie standardowe
Obrazy typu I				
wielomianowy(1)	8,02	0,94	2,98	1,66
Obrazy typu P				
wielomianowy(1)	22,61	1,88	9,53	4,99
Obrazy typu B				
wielomianowy(1)	38,74	2,23	15,96	10,11

Na podstawie porównania wyników zamieszczonych w tabelach 6.3 i 6.4 można stwierdzić, że zaproponowany uproszczony model wielomianowy z zestawem uniwersalnych stałych dobrze aproksymuje dane eksperymentalne uzyskane dla sekwencji wizyjnych spoza zbioru testowego dla obrazów typu I i P (średni względny błąd aproksymacji wynosi odpowiednio około 3 i 10%). Natomiast dla obrazów typu B średni względny błąd dopasowania wynosi 16%. Otrzymane wyniki są porównywalne do tych uzyskanych dla zbioru sekwencji testowych.

6.3. Modelowanie kodera dla wartości Q z przedziału drugiego (odcinek 2)

Do opisu zależności między sumaryczną liczbą bitów potrzebnych do reprezentacji składowych luminancji i chrominancji pojedynczego obrazu danego typu a szerokością przedziałów kwantyzacji Q , podobnie jak w poprzednich rozdziałach dla danych otrzymanych dla wartości Q z przedziału drugiego, autor zaproponował model hiperboliczny postaci:

$$B_{B_YCbCr}(Q, a, b, c) = \frac{a}{Q^b + c}, \quad (6.3)$$

gdzie:

- a, b, c to parametry modelu,
- Q to szerokość przedziałów kwantyzacji,
- B_{B_YCbCr} łączna liczba bitów reprezentacji składowych luminancji i chrominancji obrazu.

Parametry modelu zostały wyznaczone poprzez minimalizację maksymalnego względnego błędu aproksymacji niezależnie dla każdego obrazu. Statystyki średniego względnego błędu aproksymacji, analogiczne do wyznaczonych dla wartości Q z przedziału pierwszego, zebrano w tabeli 6.5. W tabeli tej podano również odpowiednie statystyki dla modelu z kodera referencyjnego.

Tabela 6.5. Statystyki średniego względnego błędu aproksymacji danych eksperymentalnych wyznaczone niezależnie dla obrazów typu I, P i B (odcinek 2)

Typ obrazu	Średni względny błąd dopasowania [%]			
	maksymalny	minimalny	średni	odchylenie standardowe
I	6,92	0,86	1,56	0,71
P	72,00	0,98	4,27	4,82
P- model referencyjny	101,10	2,45	15,87	9,43
B	95,29	1,34	13,61	16,62

Na podstawie analizy danych w tabeli 6.5, można wywnioskować, że zaproponowany model hiperboliczny dobrze aproksymuje sumaryczną liczbę bitów potrzebnych do

reprezentacji składowych luminancji i chrominancji pojedynczego obrazu danego typu. Średni błąd aproksymacji uśredniony po wszystkich obrazach danego typu wynosi 1,56, 4,27 i 13,61% odpowiednio dla obrazów typu I, P i B. Dla danych otrzymanych dla wartości Q z przedziału drugiego, zaproponowany model hiperboliczny jest lepszy od modelu z referencyjnej implementacji kodera standardu MPEG-4 AVC/H.264 (średnio o około 11,6 p.p.).

6.4. Modelowanie kodera dla wartości Q z przedziału trzeciego (odcinek 3)

Tak jak w poprzednich rozdziałach, do opisu danych eksperymentalnych otrzymanych dla wartości Q z trzeciego przedziału wykorzystano model wielomianowy 3-go stopnia z 4 parametrami:

$$B_{YCbCr}(Q, a, b, c, d) = a \cdot Q^3 + b \cdot Q^2 + c \cdot Q + d, \quad (6.4)$$

gdzie:

- a, b, c, d to parametry modelu,
- Q to szerokość przedziałów kwantyzacji,
- B_{YCbCr} łączna liczba bitów reprezentacji składowych luminancji i chrominancji obrazu.

Parametry modelu wyznaczono metodą minimalizacji maksymalnego względnego błędu aproksymacji. Analiza wartości współczynników korelacji dla poszczególnych par parametrów przeprowadzona niezależnie dla obrazów typu I, P i B pokazała, że między parametrami a, b, c i d istnieją silne zależności liniowe (tabela 6.6). Dlatego również w tym przedziale wartości Q zdecydowano o uproszczeniu modelu do jednego parametru d . Wartości uniwersalnych stałych wyznaczonych niezależnie dla obrazów typu I, P i B zebrano w tabeli 6.7.

Tabela 6.6. Wartości współczynnika korelacji liniowej dla poszczególnych par parametrów modelu wielomianowego dla obrazów typu I, P i B (odcinek 3)

Parametry modelu	<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>
Obrazy typu I				
<i>a</i>	1	-0,9971	0,9849	-0,9447
<i>b</i>	-0,9971	1	-0,9948	0,9623
<i>c</i>	0,9849	-0,9948	1	-0,9829
<i>d</i>	-0,9447	0,9623	-0,9829	1
Obrazy typu P				
<i>a</i>	1	-0,9953	0,9759	-0,929
<i>b</i>	-0,9953	1	-0,9924	0,9593
<i>c</i>	0,9759	-0,9924	1	-0,9862
<i>d</i>	-0,929	0,9593	-0,9862	1
Obrazy typu B				
<i>a</i>	1	-0,9994	0,9967	-0,9877
<i>b</i>	-0,9994	1	-0,9988	0,9922
<i>c</i>	0,9967	-0,9988	1	-0,9969
<i>d</i>	-0,9877	0,9922	-0,9969	1

Tabela 6.7. Wartości stałych uproszczonego modelu wielomianowego dla obrazów typu I, P i B (odcinek 3)

Stała	Wartości stałych modelu		
	Obrazy typu I	Obrazy typu P	Obrazy typu B
<i>a1</i>	-0,0000001014	-0,0000001279	-0,0000001714
<i>a2</i>	0,0006251313	0,0000760935	-0,0000018960
<i>b1</i>	0,0000576381	0,0000717416	0,0000915728
<i>b2</i>	-0,1837242900	-0,0403197083	0,0003716539
<i>c1</i>	-0,0117646052	-0,0138780758	-0,0163526482
<i>c2</i>	10,9954879341	7,0014386213	0,0758203813

Tak jak wcześniej, ze względu na obszerny materiał testowy w pracy nie pokazano wyników cząstkowych. Na ich podstawie wyznaczono wybrane statystyki opisujące średni względny błąd aproksymacji, które zebrano w tabeli 6.8.

Tabela 6.8. Statystyki średniego względnego błędu aproksymacji danych eksperymentalnych wyznaczone niezależnie dla obrazów typu I, P i B (odcinek 3)

Model	Dopasowanie wprost			
	Średni względny błąd dopasowania [%]			
	maksymalny	minimalny	średni	odchylenie standardowe
Obrazy typu I				
wielomianowy(4)	7,22	0,60	1,86	0,70
wielomianowy(1)	27,16	1,23	8,33	6,33
Obrazy typu P				
wielomianowy(4)	583,79	0,81	10,98	33,68
wielomianowy(1)	15443,92	2,43	121,09	903,64
referencyjny(2)	99,16	1,26	11,04	12,42
Obrazy typu B				
wielomianowy(4)	1875,40	0,00	140,98	204,92
wielomianowy(1)	864,64	11,12	237,94	246,41

Z analizy danych z tabeli 6.8 wynika, że zaproponowany model wielomianowy dobrze aproksymuje dane eksperymentalne dla obrazów typu I i P. Natomiast, ze względu na duże błędy aproksymacji, nie można go wykorzystać do modelowania liczby bitów reprezentacji składowych luminancji i chrominancji w obrazach typu B. Główna trudność w aproksymacji danych otrzymanych dla tego przedziału wartości Q , wynika z silnej kwantyzacji wartości próbek transformaty, przez co do reprezentacji składowych luminancji i chrominancji wykorzystuje się niewielką liczbę bitów (szczególnie dla obrazów typu B).

Dla danych otrzymanych dla przedziału trzeciego również dokonano weryfikacji dokładności uproszczonego modelu wielomianowego z wykorzystaniem kontrolnego zbioru sekwencji wizyjnych (dodatek A). Eksperyment polegał na wyznaczeniu wartości parametru d oraz maksymalnego i średniego względnego błędu aproksymacji danych otrzymanych dla kontrolnego zbioru sekwencji wizyjnych z wykorzystaniem uniwersalnych stałych zebranych w tabeli 6.7. Wyniki eksperymentu umieszczono w tabeli 6.9.

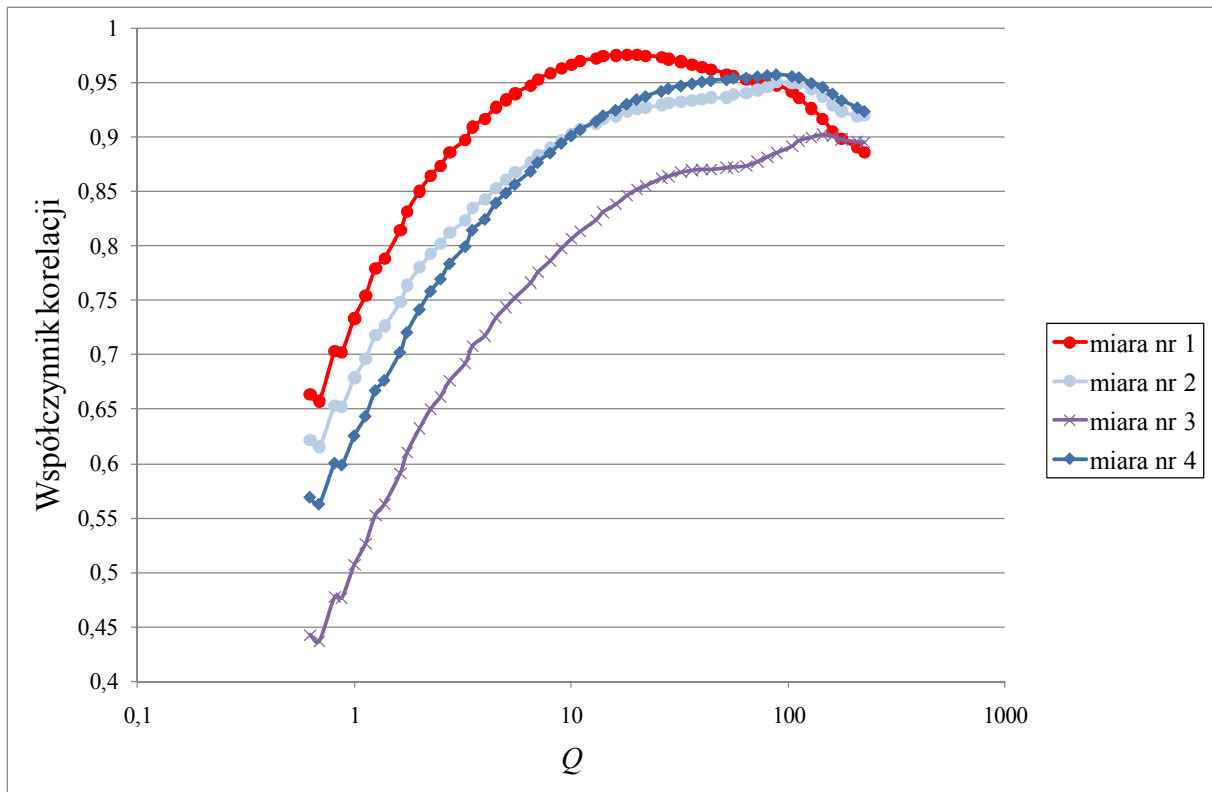
Tabela 6.9. Statystyki średniego względnego błędu aproksymacji danych z wykorzystaniem uproszczonego modelu wielomianowego dla obrazów typu I, P i B dla kontrolnego zbioru sekwencji wizyjnych (odcinek 3)

Model	Średni względny błąd aproksymacji [%]			
	maksymalny	minimalny	średni	odchylenie standardowe
	Obrazy typu I			
wielomianowy(1)	26,67	1,81	6,92	6,81
	Obrazy typu P			
wielomianowy(1)	122,32	2,43	16,46	11,66
	Obrazy typu B			
wielomianowy(1)	808,91	11,12	179,94	196,27

Z wyników zamieszczonych w tabeli 6.9 wynika, że średni względny błąd aproksymacji danych dla kontrolnego zbioru sekwencji jest niższy od błędu uzyskanego dla testowego zbioru sekwencji na którym wyznaczono wartości stałych. Mimo to uproszczony model wielomianowy nadal generuje duży błąd aproksymacji dla obrazów typu B.

6.5. Model z parametrem charakteryzującym złożoność treści obrazu

W punkcie 5.5 autor przebadiał kilka miar charakteryzujących złożoność treści obrazów typu I. Miary te zostały wyznaczone na podstawie wartości próbek obrazu przed kodowaniem. Mając wyznaczone wartości wspomnianych miar, autor postanowił sprawdzić, czy można je wykorzystać przy modelowaniu sumarycznej liczby bitów potrzebnych do reprezentacji składowych luminancji i chrominancji obrazu. Na wykresie 6.5 przedstawiono zależność współczynnika korelacji wyznaczonego dla wspomnianych miar złożoności treści obrazu a sumaryczną liczbą bitów potrzebnych do reprezentacji składowych luminancji i chrominancji obrazu typu I w zależności od szerokości przedziałów kwantyzacji Q .



Wykres 6.5. Współczynnik korelacji miary złożoności treści obrazu i liczby bitów reprezentacji składowych luminancji i chrominancji obrazu dla całego zakresu wartości Q (przedstawione miary opisano w punkcie 5.5)

Z wykresu 6.5 można odczytać, że spośród badanych miar najwyższe wartości współczynnika korelacji w drugim przedziale wartości Q uzyskano dla miary nr 1, czyli $\overline{\sigma_Y}$ (wzór 5.7). Dlatego autor zaproponował rozszerzenie modelu hiperbolicznego o parametr charakteryzujący złożoność treści obrazu typu I:

$$B_{YCbCr}(Q, a, b, c) = \overline{\sigma_Y} \cdot \frac{a}{Q^b + c}, \quad (6.5)$$

gdzie:

- a, b, c to parametry modelu,
- Q to szerokość przedziałów kwantyzacji,
- B_{YCbCr} łączna liczba bitów reprezentacji składowych luminancji i chrominancji obrazu,
- $\overline{\sigma_Y}$ to zaproponowana miara złożoności treści obrazu (wzór 5.7).

Dla tak określonego modelu i miary złożoności obrazu autor wyznaczył parametry tego modelu minimalizując maksymalny względny błąd aproksymacji. W tabeli 6.10 umieszczono

statystyki średniego względnego błędu aproksymacji danych z wykorzystaniem zaproponowanego modelu.

Tabela 6.10. Statystyki średniego względnego błędu aproksymacji danych eksperymentalnych dla rozszerzonego modelu hiperbolicznego

Model	Średni względny błąd dopasowania [%]			
	maksymalny	minimalny	średni	odchylenie standardowe
hiperboliczny rozszerzony (3 parametry + $\overline{\sigma_Y}$)	7,38	0,78	1,32	0,60

Porównanie danych z tabel 6.5 i 6.10 pozwala stwierdzić, że zastosowanie parametru charakteryzującego złożoność treści obrazu typu I nieznacznie poprawia dokładność aproksymacji danych eksperymentalnych otrzymanych dla wartości Q z przedziału drugiego (średnio o około 0,24 p.p.).

6.6. Podsumowanie

Do modelowania liczby bitów potrzebnych do reprezentacji składowych luminancji i chrominancji autor zaproponował wykorzystanie takich samych modeli jak w rozdziałach 4 i 5. Podsumowanie dokładności modeli z tego rozdziału umieszczono w tabeli 6.11. W tabeli tej nie podano dokładności modelu zastosowanego w referencyjnej implementacji kodera dla obrazów typu I i B, ponieważ model ten wykorzystuje się tylko do opisu obrazów typu P [Li03, Li04]. Z analizy danych z tabeli 6.11 można wywnioskować, że zaproponowane modele dobrze aproksymują dane eksperymentalne z wyjątkiem danych dla obrazów typu B w trzecim przedziale wartości Q . Z tabeli tej wynika również, że dla zaproponowanych modeli uzyskano niższe wartości błędów aproksymacji danych w porównaniu z modelem referencyjnym. Uproszczenie modelu wielomianowego w trzecim przedziale wartości Q

spowodowało znaczny wzrost wartości błędu aproksymacji, szczególnie dla obrazów typu P i B.

Tabela 6.11. Średni względny błąd dopasowania uśredniony po obrazach danego typu

Model	Typ obrazu		
	I	P	B
	Odcinek 1		
propozycja(4)	0,98	3,00	3,38
propozycja(1)	4,58	8,21	10,09
referencyjny(2)	x	8,63	x
	Odcinek 2		
propozycja(3)	1,56	4,27	13,61
referencyjny(2)	x	15,87	x
	Odcinek 3		
propozycja(4)	1,86	10,98	140,98
propozycja(1)	8,33	121,09	237,94
referencyjny(2)	x	11,04	x

Poza modelem zaproponowanym w referencyjnej implementacji kodera standardu MPEG-4 AVC/H.264 w literaturze można znaleźć inne propozycje modelowania liczby bitów reprezentacji składowych luminancji i chrominancji obrazu. Niestety nie są dostępne implementacje takich rozwiązań, przez co nie można przeprowadzić badań porównawczych.

Rozdział 7

Podsumowanie

7.1. Wnioski

Celem badań przedstawionych w rozprawie było zaproponowanie ilościowych modeli matematycznych zaawansowanych koderów wizyjnych, wyrażających liczbę bitów generowanych przez koder w funkcji szerokości przedziałów kwantyzacji Q .

Na przykładzie kodera standardu MPEG-4 AVC/H.264 autor pokazał, że można wyznaczyć proste modele matematyczne opisujące wspomnianą zależność, co potwierdza tezę rozprawy.

Modele zostały zaproponowane na podstawie analizy danych eksperymentalnych uzyskanych z wykorzystaniem referencyjnej implementacji kodeka standardu MPEG-4 AVC/H.264 dla szerokiej klasy sekwencji wizyjnych o standardowej rozdzielczości telewizyjnej.

Dalsze badania powinny dotyczyć możliwości zastosowania zaproponowanych modeli w algorytmie sterowania zaawansowanego kodera wizyjnego. Autor zaproponował modele dla trzech typów obrazów, a obecnie modeluje się przeważnie tylko obrazy typu P.

Przy zastosowaniu zaawansowanych koderów wizyjnych mamy do czynienia ze znacznym udziałem danych nagłówkowych w całkowitym strumieniu danych. Dlatego dalsze prace mogłyby dotyczyć modelowania także takich danych strumienia bitowego.

Modele przedstawione w rozprawie zostały zaproponowane dla danych eksperymentalnych uzyskanych przy wykorzystaniu kodera standardu MPEG-4 AVC/H.264 zgodnego z profilem głównym. Obecnie, coraz bardziej upowszechnia się telewizja wysokiej

jakości. Dlatego dalsze badania powinny dotyczyć modelowania zaawansowanych koderów wizyjnych na potrzeby takich zastosowań, np. koderów profilu wysokiego lub profili rozszerzonego zakresu dokładności (FRExt) [Marp05, Sull04].

7.2. Oryginalne osiągnięcia

Zasadniczym, oryginalnym, naukowym osiągnięciem rozprawy jest przygotowanie modeli określających liczbę bitów produkowanych przez koder w zależności od szerokości przedziałów kwantyzacji Q . Modele te zostały wyznaczone z wykorzystaniem bogatego zbioru danych uzyskanych w wyniku przeprowadzenia ponad 3000 eksperymentów kodowania sekwencji wizyjnych. Dokładność zaproponowanych modeli została zweryfikowana eksperymentalnie.

W szczególności w pracy zaproponowano oryginalne, globalne modele strumienia bitowego wyrażające zależności między szerokością przedziałów kwantyzacji Q a:

- średnią liczbą bitów potrzebnych do reprezentacji obrazu danego typu w ramach jednego ujęcia,
- liczbą bitów potrzebnych do reprezentacji pojedynczego obrazu danego typu,
- łączną liczbą bitów potrzebnych do reprezentacji skwantowanych wartości próbek transformaty błędu predykcji składowych luminancji i chrominancji obrazu danego typu.

Modele pozwalające na wyznaczenie średniej liczby bitów reprezentacji obrazu danego typu w ramach jednego ujęcia nie są znane z literatury.

Zaproponowano także nowy globalny model strumienia bitowego z parametrem charakteryzującym złożoność treści obrazów typu I, który opisuje:

- liczbę bitów potrzebnych do zakodowania pojedynczego obrazu danego typu lub
- sumaryczną liczbę bitów potrzebnych do zakodowania składowych luminancji i chrominancji pojedynczego obrazu danego typu w funkcji szerokości przedziałów kwantyzacji Q .

Zbadano wykorzystanie następujących parametrów charakteryzujących złożoność treści obrazów typu I w sekwencji wizyjnej:

- odchylenie standardowe dla wartości próbek składowej luminancji wyznaczone w blokach o rozmiarze 4x4 uśrednione po całym obrazie,
- suma odchyleń standardowych wartości próbek składowych luminancji i chrominancji wyznaczonych w blokach o rozmiarze 4x4 uśrednionych po całym obrazie.

Zaproponowany globalny model strumienia bitowego z parametrem charakteryzującym złożoność treści obrazów typu I nie jest znany z literatury.

Przeprowadzono analizę dokładności zaproponowanych modeli. Dodatkowo zbadano dokładność modelu wyrażającego liczbę bitów reprezentacji pojedynczego obrazu w zależności od szerokości przedziałów kwantyzacji Q przy predykcji parametrów tego modelu z poprzedniego obrazu danego typu należącego do tego samego ujęcia.

Porównano również dokładność zaproponowanego modelu z modelem opisanym w części informacyjnej standardu MPEG-4 AVC/H.264, który wykorzystuje się do modelowania liczby bitów reprezentacji składowych luminancji i chrominancji obrazów typu P. Dla użytecznego zakresu prędkości transmisji model zaproponowany przez autora jest bardziej dokładny od tego wykorzystanego w referencyjnej implementacji kodera o około 11,5 punktów procentowych.

Spis literatury

- [Acha05] Acharya, T., Ray, A.K.: Image processing: principles and applications, John Wiley & Sons, Inc. (2005).
- [AVC] ISO/IEC 14496-10 (MPEG-4AVC) / ITU-T Rec. H.264: Advanced Video Coding for Generic Audiovisual Services, International Organization for Standardization / International Electrotechnical Commission, International Telecommunication Union – Telecommunication (2007).
- [AVCref] Suhring, K.: H.264/AVC software coordination, <http://iphome.hhi.de/suehring/tml/>
- [AVS] Audio Video Coding Standard Workgroup of China (AVS), The Standards of People's Republic of China GB/T 20090.2-2006, Information Technology – Advanced Coding of Audio and Video – Part 2: Video (2006).
- [Bai01] Bai, J.F., Feng, C., Liao, Q.M., Lin, X.G., Zhang, X.H.: Accurate estimation of R-D characteristics for rate control in real-time video encoding, IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP, vol. 3, pp. 1381 – 1384, Salt Lake City, USA (2001).
- [Baxe94] Baxes, G.A.: Digital image processing: principles and applications, John Wiley & Sons, Inc. (1994).
- [Bjon01] Bjontegaard, G.: Calculation of average PSNR differences between RD-curves, ITU-T Video Coding Experts Group (VCEG) (ITU-T SG16 Q.6), document VCEG-M33, Austin, Texas, USA (2001).
- [Bjon08] Bjontegaard, G.: Improvements of the BD-PSNR model, ITU-T Video Coding Experts Group (VCEG) (ITU-T SG16 Q.6), document VCEG-AI11, Berlin, Germany (2008).
- [Bobi04] Bobinski, P., Skarbek, W.: Analysis of RD models for coding efficiency in H.264 standard, 5th International Workshop on Image Analysis for Multimedia Interactive Services, WIAMIS, Lisboa, Portugal (2004).
- [Bovi05] Bovik, A. (redaktor): Handbook of image and video processing, Academic Press, wydanie 2 (2005).
- [Boyc04] Boyce, J.M.: Weighted prediction in the H.264/MPEG AVC video coding standard, IEEE International Symposium on Circuits and Systems, ISCAS, vol. 3, pp. 789 – 792, Vancouver, Canada (2004).
- [Bran99]] Brandt, S.: Analiza danych, metody statystyczne i obliczeniowe, Wydawnictwo Naukowe PWN, wydanie 2, Warszawa (1999).
- [Bron07] Bronsztejn, I.N., Siemiendajew, K.A., Musiol, G., Muhlig, H.: Nowoczesne kompendium matematyki, Wydawnictwo Naukowe PWN, wydanie 1, Warszawa (2007).

- [Chia97] Chiang, T., Zhang, Y.-Q.: A new rate control scheme using quadratic rate distortion model, *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 7, pp. 246 – 250 (1997).
- [Chon08] Chong, E.K.P., Żak, S.H.: An introduction to optimization, John Wiley & Sons, Inc., wydanie 3, New Jersey (2008).
- [Ciec07] Cieciura, M., Zacharski, J.: Metody probabilistyczne w ujęciu praktycznym, Vizja Press&IT, wydanie 1, Warszawa (2007).
- [Ding96] Ding, W., Liu, B.: Rate control of MPEG video coding and recording by rate-quantization modeling, *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 6, no. 2, pp. 12 – 20 (1996).
- [Doma98] Domański, M.: Zaawansowane techniki kompresji obrazów i sekwencji wizyjnych, Wydawnictwo Politechniki Poznańskiej, Poznań (1998).
- [Doma02] Domański, M., Łuczak, A.: Simple global model of an MPEG-2 bitstream, *IEEE International Conference on Multimedia and Expo, ICME*, pp. 645 – 648, Lausanne, Switzerland (2002).
- [Doma05] Domański, M., Grajek, T., Marek, J.: Zaawansowana kompresja cyfrowych sygnałów wizyjnych – standard AVC/H.264, *Systemy Alarmowe*, nr 2, rok wyd. XIV, vol. 75, str. 14 – 21 (2005).
- [Dong07a] Dong, J., Ling, N.: Enhanced linear R-Q model based rate control for H.264/AVC using context-adaptive parameter estimation, *IEEE Workshop on Signal Processing Systems, SiPS*, pp. 698 – 703, Shanghai, China (2007).
- [Dong07b] Dong, J., Ling, N.: On model parameter estimation for H.264/AVC rate control, *IEEE International Symposium on Circuits and Systems, ISCAS*, pp. 289 – 292, New Orleans, USA (2007).
- [Dong08] Dong, J., Ling, N.: A model parameter and MAD prediction scheme for H.264 macroblock layer rate control, *IEEE International Symposium on Circuits and Systems, ISCAS*, pp. 628 – 631, Seattle, USA (2008).
- [Dzię05] Dzięcielewski, T., Grajek, T., Marek, J.: Eksperymentalna analiza częstotliwości wyboru trybów zaawansowanego kodowania wizyjnego, *X Poznańskie Warsztaty Telekomunikacyjne*, str. 115 – 120, Poznań (2005).
- [Dzię06] Dzięcielewski, T., Grajek, T., Marek, J.: Eksperymentalna analiza efektywności trybów kodowania międzyobrazowego w zaawansowanym koderze wizyjnym, *Krajowa Konferencja Radiokomunikacji, Radiofonii i Telewizji*, str. 219 – 222, Poznań (2006).
- [Fan04] Fan, L., Ma, S., Wu, F.: Overview of AVS video standard. *IEEE International Conference on Multimedia and Expo, ICME*, vol. 1, pp. 423 – 426, Taipei, Taiwan (2004).
- [Ghan99] Ghanbari, M.: Video coding – an introduction to standard codecs, Institution Electrical Engineers, Stevenage, UK (1999).

- [Graj05] Grajek, T., Karwowski, D.: Złożoność obliczeniowa i efektywność kodowania entropijnego w standardzie H.264/AVC, Krajowa Konferencja Radiokomunikacji, Radiofonii i Telewizji, str. 377 – 380, Kraków (2005).
- [Graj09a] Grajek, T., Domański, M.: A global model of AVC/H.264 video encoders, 27th Picture Coding Symposium, PCS, Chicago, Illinois, USA (2009).
także: IEEE Xplore Digital Library, <http://ieeexplore.ieee.org>.
- [Graj09b] Grajek, T., Domański, M.: A simple quantitative model of AVC/H.264 video coders, Computer Vision and Graphics International Conference, ICCVG 2008, Lecture Notes in Computer Science – LNCS 5337, pp. 174 – 183, Springer-Verlag, Berlin(2009).
- [H261] ITU-T Rec. H.261: Video codec for audiovisual services at p×64 kbit/s, International Telecommunication Union – Telecommunication, wersja 2 (1993).
- [H263] ITU-T Rec. H.263: Video Coding for Low Bit Rate Communication, International Telecommunication Union – Telecommunication, wersja 3 (2000).
- [Hang97] Hang, H.M., Chen, J.J.: Source model for transform video coder and its application – part I: fundamental theory, IEEE Transactions on Circuits and Systems for Video Technology, vol. 7, no. 2, pp. 287 – 298 (1997).
- [He01a] He, Z., Kim, Y.K., Mitra, S.K.: Low-delay rate control for DCT video coding via ρ -domain source modeling, IEEE Transactions on Circuits and Systems for Video Technology, vol. 11, no. 8, pp. 928 – 940 (2001).
- [He01b] He, Z., Kim, Y.K., Mitra, S.K.: ρ -domain source modeling and rate control for video coding and transmission, IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP, vol. 3, pp. 1773 – 1776, Salt Lake City, USA (2001).
- [He01c] He, Z., Mitra, S.K.: A unified rate-distortion analysis framework for transform coding, IEEE Transactions on Circuits and Systems for Video Technology, vol. 11, no. 12, pp. 1221 – 1236 (2001).
- [He01d] He, Z., ρ -domain rate-distortion analysis and rate control for visual coding and communication, Doctoral Dissertation, University of California, Santa Barbara, USA (2001).
- [He02a] He, Z., Mitra, S.K.: A linear source model and a unified rate control algorithm for DCT video coding, IEEE Transactions on Circuits and Systems for Video Technology, vol. 12, no. 11, pp. 970 – 982 (2002).
- [He02b] He, Z., Mitra, S.K.: Optimum bit allocation and accurate rate control for video coding via ρ -domain source modeling, IEEE Transactions on Circuits and Systems for Video Technology, vol. 12, no. 10, pp. 840 – 849 (2002).

- [He03] He, Z., Chen, T.: Linear rate control for JVT video coding, International Conference on Information Technology: Research and Education, ITRE, pp. 65 – 68, Newark, USA (2003).
- [Jian06] Jiang, M., Ling, N.: Bit allocation scheme for low-delay H.264/AVC rate control, IEEE International Conference on Image Processing, ICIP, pp. 2501 – 2504, Atlanta, USA (2006).
- [JPEG2000] ISO/IEC 15444-1 (JPEG2000) / ITU-T Rec. T.800: Information Technology – JPEG2000 image coding system, International Organization for Standardization / International Electrotechnical Commission, International Telecommunication Union – Telecommunication (2000).
- [JSVM] JSVM software manual, JSVM (Joint Scalable Video Model) software for the Scalable Video Coding (SVC) project of the Joint Video Team (JVT) of the ISO/IEC MPEG and the ITU-T VCEG (2007).
- [Kalv07] Kalva, H., Lee, J.-B.: The vc-1 video coding standard, IEEE MultiMedia, vol. 14, no. 4, pp. 88 – 91 (2007).
- [Kama03] Kamaci, N., Altunbasak, Y.: Performance comparison of the emerging H.264 video coding standard with the existing standards, International Conference on Multimedia and Expo, ICME, vol. 1, pp. 345 – 348, Baltimore, USA (2003).
- [Kama04a] Kamaci, N., Altunbasak, Y.: An analysis of the DCT coefficient with the H.264 video coder, IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP, vol. 3, pp. 177 – 180, Montreal, Canada (2004).
- [Kama04b] Kamaci, N., Altunbasak, Y.: ρ -domain rate-distortion optimal rate control for DCT-based video coders, IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP, vol. 3, pp. 149 – 152, Montreal, Canada (2004).
- [Kama05] Kamaci, N., Altunbasak, Y., Mersereau, R.M.: Frame bit allocation for the H.264 AVC video coder via Cauchy-density-based rate and distortion models, IEEE Transactions on Circuits and Systems for Video Technology, vol. 15, no. 8, pp. 994 – 1006 (2005).
- [Kim01] Kim, Y.K., He, Z., Mitra, S.K.: A novel linear source model and a unified rate control algorithm for H.263/MPEG-2/MPEG-4, IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP. vol. 3, pp. 1777 – 1780, Salt Lake City, USA (2001).
- [Koro01] Koronacki, J., Mielniczuk, J.: Statystyka dla studentów kierunków technicznych i przyrodniczych, Wydawnictwa Naukowo-Techniczne, wydanie 1, Warszawa (2001).

- [Kwon07] Kwon, D.-K., Shen, M.-Y., Kuo, C.-C. J.: Rate control for H.264 video with enhanced rate and distortion models, *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 17, no. 5, pp. 517 – 529 (2007).
- [Lee00] Lee, H.J., Chiang, T., Zhang, Y.Q.: Scalable rate control for MPEG-4 video, *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 10, no. 6, pp. 878 – 894 (2000).
- [Lee08a] Lee, G.L., Lin, H.Y., Wang, M.J.: Textural complexity-based rate control algorithm, *IEEE International Conference on Multimedia and Expo, ICME*, pp. 1109 – 1112, Hannover, Germany (2008).
- [Lee08b] Lee, J.B., Kalva, H.: *The VC-1 and H.264 video compression standards for broadband video services*, Springer Science+Business Media, LLC, New York, USA (2008).
- [Li03] Li, Z.G., Pan, F., Lim, K.P., Feng, G., Lin, X., Rahardja, S.: Adaptive basic unit layer rate control for JVT, Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG (ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q.6), document JVT-G012r1, Pattaya, Thailand (2003).
- [Li04] Li, Z.G., Pan, F., Lim, K.P., Rahardja, S.: Adaptive rate control for H.264, *IEEE International Conference on Image Processing, ICIP*, pp. 745 – 748, Singapore (2004).
- [Lian04] Liang, J., Lee, M.C., Ribas-Corbera, J.: Windows Media Video 9.0: overview and applications, *Signal Processing: Image Communication*, vol. 19, no. 9, pp. 851 – 876 (2004).
- [Lim05] Lim, K.P., Sullivan, G.J., Wiegand, T.: Text description of Joint Model reference encoding methods and decoding concealment methods, Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG (ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q.6), document JVT-O079, Busan, S.Korea (2005).
- [Lim07] Lim, S.C., Na, H.R., Lee, Y.L.: Rate control based on linear regression for H.264/MPEG-4 AVC, *Signal Processing: Image Communication*, vol. 22, no. 1, pp. 39 – 58, Elsevier (2007).
- [Lin98] Lin L.J., Ortega, A.: Bit-rate control using piecewise approximated rate-distortion characteristics, *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 8, no. 4, pp. 446 – 459 (1998).
- [Lin08] Lin, G., Zheng, S., Hu, J.: A two-stage ρ -domain rate control scheme for H.264 encoder, *IEEE International Conference on Multimedia and Expo, ICME*, pp. 713 – 716, Hannover, Germany (2008).
- [Liu06] Liu, Y., Li, Z.G., Soh, Y.C.: Adaptive MAD prediction and refined R-Q model for H.264/AVC rate control, *IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP*, pp. II – 905 – 908, Toulouse, France (2006).

- [Liu07] Liu, Y., Li, Z.G., Soh, Y.C.: A novel rate control scheme for low delay video communication of H.264/AVC standard, *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 17, no. 1, pp. 68 – 78 (2007).
- [Łucz01] Łuczak, A.: Modelowanie strumieni binarnych generowanych przez koder obrazu MPEG-2, *Krajowa Konferencja Radiokomunikacji, Radiofonii i Telewizji, KKRRiT*, str. 7.4-1 – 7.4-3, Poznań (2001).
- [Łucz02] Łuczak, A., Domański, M.: Modeling of bitstreams produced by hybrid video coders, *IEEE International Symposium on Consumer Electronics, ISCE*, pp. E69 – E73, Erfurt, Germany (2002).
- [Łucz03a] Łuczak, A., Domański, M.: MPEG video bitrate control using coder modeling, *Picture Coding Symposium, PCS*, pp. 93 – 98, Saint Malo, France (2003).
- [Łucz03b] Łuczak, A.: Digital video bitrate control using coder models, *Rozprawa doktorska*, Poznań (2003).
- [Łucz04a] Łuczak, A., Domański, M.: Cosine transform analysis with applications to MPEG-2 video coder control, *International Workshop on Systems, Signals and Image Processing, IWSSIP*, pp. 279 – 282, Poznań (2004).
- [Łucz04b] Łuczak, A., Domański, M.: Hybrid video coder modeling using histograms of transform coefficients, *5th International Workshop on Image Analysis for Multimedia Interactive Services, WIAMIS*, Lisboa, Portugal (2004).
- [Ma03] Ma, S., Gao, W., Wu, F., Lu, Y.: Rate control for JVT video coding scheme with HRD considerations, *IEEE International Conference on Image Processing, ICIP*, vol. 3, pp. 793 – 796, Barcelona, Spain (2003).
- [Ma05] Ma, S., Gao, W.: Rate-distortion analysis for H.264/AVC video coding and its application to rate control, *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 15, no. 12, pp. 1533 – 1544 (2005).
- [Malv03] Malvar, H., Hallapuro, A., Karczewicz, M., Kerofsky, L.: Low-complexity transform and quantization in H.264/AVC, *IEEE Transaction on Circuits and Systems for Video Technology*, vol. 13, no. 7, pp. 598 – 603 (2003).
- [Marp03] Marpe, D., Schwarz, H., Wiegand, T.: Context-based adaptive binary arithmetic coding in the H.264/AVC video compression standard, *IEEE Transaction on Circuits and Systems for Video Technology*, vol. 13, no. 7, pp. 620 – 636 (2003).
- [Marp04] Marpe, D.: Adaptive context-based and tree-based algorithms for image coding and denoising, *Doctoral Dissertation*, Universität Rostock, Rostock, Germany (2004).
- [Marp05] Marpe, D., Gordon, S., Wiegand, T.: H.264/MPEG4-AVC fidelity range extensions: tools, profiles, performance, and application areas, *IEEE International Conference on Image Processing, ICIP*, vol. 1, pp. 593 – 596, Genova, Italy, (2005).

- [Marp06] Marpe, D., Wiegand, T., Sullivan, G.J.: The H.264/MPEG4 Advanced Video Coding standard and its applications, IEEE Communications Magazine, vol. 44, no. 8, pp. 134 – 143 (2006).
- [Math] Pakiet matematyczny – Mathcad wersja 14.0, <http://www.ptc.com/products/mathcad>.
- [Mila03] Milani, S., Celetto, L., Mian, G.A.: A rate control algorithm for the H.264 encoder, Sixth Baiona Workshop on Signal Processing in Communications, pp. 390 – 396 Baiona, Spain, (2003).
- [Ming07] Ming, Y.; Yun, X.; Fen; G., Shuting, C.: Improvements on MB-layer rate control scheme for H.264 video using complexity estimation, Eighth ACIS International Conference on Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing, SNPD, vol. 2, pp. 280 – 284, Qingdao, China (2007).
- [MPEG1] ISO/IEC 11172-2 (MPEG-1): Information technology – Coding of moving pictures and associated audio for digital storage media at up to about 1,5Mbits/sec – Part2: Video, International Organization for Standardization / International Electrotechnical Commission (1993).
- [MPEG2] ISO/IEC 13818-2 (MPEG-2Video) / ITU-T Rec. H.262: Generic Coding of Moving Pictures and Associated Audio – Part 2: Video, International Organization for Standardization / International Electrotechnical Commission, International Telecommunication Union – Telecommunication (1994).
- [MPEG4] ISO/IEC 14496-2 (MPEG-4Visual): Information technology – Coding of audiovisual objects – Part2: Visual, International Organization for Standardization / International Electrotechnical Commission (2003).
- [Ohm04] Ohm, J.R.: Multimedia communication technology. Representation, transmission and identification of multimedia signals, Springer-Verlag, Berlin (2004).
- [Oste04] Ostermann, J., Bormans, J., List, P., Marpe, D., Narroschke, M., Pereira, F., Stockhammer, T., Wedi, T.: Video coding with H.264/AVC: tools, performance and complexity, IEEE Circuits and Systems Magazine, vol. 4, no. 1, pp. 7 – 28, (2004).
- [Past06] Pastuszak, G., Buchowicz, A., Skarbek, W.: Rate-control techniques in the latest visual compression standards – a survey, New Trends in Audio and Video, NTIAV, pp. 83 – 90, Białystok (2006).
- [Penn93] Pennebaker, W.B., Mitchell, J.L.: JPEG: still image data compression standard (Digital Multimedia Standards), Kluwer Academic Publishers (1993).

- [Piet05a] Pietrowcew, A., Buchowicz, A., Skarbak, W.: Algorytm kontroli przepływności dla kodowania sekwencji video z regionami zainteresowania, Krajowa Konferencja Radiokomunikacji, Radiofonii i Telewizji, KKRRiT, str. 235 – 238, Kraków (2005).
- [Piet05b] Pietrowcew, A., Buchowicz, A., Skarbak, W.: Bit-rate control algorithm based on local image complexity for video coding with ROI, IEEE Conference on Advanced Video and Signal Based Surveillance, AVSS, pp. 582 – 587, Como, Italy (2005).
- [Piet05c] Pietrowcew, A., Buchowicz, A., Skarbak, W.: Bit-rate control algorithm for ROI enabled video coding, 11th International Conference on Computer Analysis of Images and Patterns, CAIP2005, Lecture Notes in Computer Science –LNCS 3691, pp. 514 – 521, Springer-Verlag, Berlin (2005).
- [Piet05d] Pietrowcew, A., Buchowicz, A., Skarbak, W.: Bit-rate control for video coding with ROI, 6th International Workshop on Image Analysis for Multimedia Interactive Services, WIAMIS, Montreux, Switzerland (2005)
- [Pres07] Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P.: Numerical recipes: the art of scientific computing, Cambridge University Press, wydanie 3 (2007).
- [Raja04] Raja, G., Mirza, J.M.: Performance comparison of advanced video coding H.264 standard with baseline H.263 and H.263+ standards, IEEE International Symposium on Communications and Information Technology, ISCIT, vol. 2, pp. 743 – 746, Sapporo, Japan (2004).
- [Rao09] Rao, S.S.: Engineering optimization: theory and practice, John Wiley & Sons, Inc., wydanie 4 (2009).
- [Ravi06] Ravindran, A., Ragsdell, K.M., Reklaitis, G.V.: Engineering optimization: methods and applications, John Wiley & Sons, Inc., wydanie 2, New Jersey (2006).
- [Riba99] Ribas-Corbera J., Lei, S.: Rate control in DCT video coding for low- delay communications, IEEE Transactions on Circuits and Systems for Video Technology, vol. 9, no. 1, pp. 172 – 185 (1999).
- [Rich02] Richardson, I.E.G.: Video codec design: developing image and video compression systems, John Wiley & Sons, Inc., Chichester (2002).
- [Rich03] Richardson, I.E.G.: H.264 and MPEG-4 video compression, video coding for next-generation multimedia, John Wiley & Sons, Inc., Chichester (2003).
- [Ruol07] Ruolin, R., Ruimin, H., Xuqing X.: A rate control scheme based on MAD weighted model for H.264/AVC, International Conference on Wireless Communications, Networking and Mobile Computing, WiCom, pp. 2903 – 2906, Shanghai, China (2007).
- [Salo07] Salomon, D.: Data compression: the complete reference, Springer-Verlag, wydanie 4 (2007).

- [Sayo05] Sayood, K.: Introduction to data compression, Academic Press, wydanie 3 (2005).
- [Scha03] Schafer, R., Wiegand, T., Schwarz, H.: The emerging H.264/AVC standard, European Broadcasting Union Technical Review (2003).
- [Shin04] Shin, I.H., Lee, Y.L., Park HW.: Rate control using linear rate- ρ model for H.264, Signal Processing: Image Communication, vol. 19, no. 4, pp. 341 – 352, Elsevier (2004).
- [Skar98] Skarbek, W. (redaktor): Multimedia. Algorytmy i standardy kompresji. Akademicka Oficyna Wydawnicza PLJ, Warszawa (1998).
- [Skar06] Skarbek, W., Buchowicz, A., Pietrowcew, A., Pereira, F.: Bit-rate control for compression of video with ROI, Computer Vision and Graphics International Conference, ICCVG 2004, Computational Imaging and Vision , vol. 32, pp. 228 – 239, Springer Netherlands (2006).
- [Sull98] Sullivan, G.J., Wiegand, T.: Rate-distortion optimization for video compression, IEEE Signal Processing Magazine, vol. 15, no. 6, pp. 74 – 90 (1998).
- [Sull04] Sullivan, G.J., Topiwala, P.N., Luthra, A.: The H.264/AVC advanced video coding standard: overview and introduction to the fidelity range extensions, Applications of Digital Image Processing XXVII. Edited by Tescher, A.G., Proceedings of the SPIE, Volume 5558, pp. 454-474 (2004).
- [Sull05] Sullivan, G.J., Wiegand, T.: Video compression – from concepts to the H.264/AVC standard. Proceedings of the IEEE, vol. 93, no. 1, pp. 18 – 31 (2005).
- [Sun06] Sun, S., Reichel, J.: AHG report on spatial scalability resampling, Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG (ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q.6), document JVT-R06r1, Bangkok, Thailand (2006).
- [Sze09] Sze, V., Chandrakasan, A.P.: A high throughput CABAC algorithm using syntax element partitioning,” IEEE International Conference on Image Processing, ICIP, pp 773 – 776 (2009).
- [Tan08] Tan, T.K., Sullivan, G., Wedi, T.: Recommended simulation common conditions for coding efficiency experiments, ITU-T Video Coding Experts Group (VCEG) (ITU-T SG16 Q.6), document VCEG-AJ10, San Diego, USA (2008).
- [Taub02] Taubman, D.S., Marcellin, M.W.: JPEG2000: Image compression fundamentals, standards and practice, Kluwer Academic Publishers, Boston (2002).
- [Teka95] Tekalp, A.M.: Digital video processing, Prentice Hall PTR (1995).

- [VC1] SMPTE 421M, Proposed SMPTE Standard for Television: VC-1 Compressed Video Bitstream Format and Decoding Process, SMPTE Technology Committee C24 on Video Compression Technology (2005).
- [Vite79] Viterbi, A., Omura, J.: A new rate control scheme using a new rate-distortion model, in Principle of Digital Communication and Coding. McGraw-Hill, New York (1979).
- [Wan06] Wan, S., Izquierdo, E.: A new rate distortion model for rate control of video coding, 7th International Workshop on Image Analysis for Multimedia Interactive Services, WIAMIS, Incheon, S.Korea (2006).
- [Wedi03] Wedi, T., Musmann, H.G.: Motion- and aliasing- compensated prediction for hybrid video coding, IEEE Transaction on Circuits and Systems for Video Technology, vol. 13, no. 7, pp. 577 – 586 (2003).
- [Weng03] Wenger, S.: H.264/AVC over IP,” IEEE Transaction on Circuits and Systems for Video Technology, vol. 13, no. 7, pp. 645 – 656 (2003).
- [Wieg01] Wiegand, T., Girod, B.: Lagrange multiplier selection in hybrid video coder control, IEEE International Conference on Image Processing, ICIP, vol. 3, pp. 542 – 545, Thessaloniki, Greece (2001).
- [Wieg03a] Wiegand, T., Schwarz, H., Joch, A., Kossentini, F., Sullivan, G.J.: Rate-constrained coder control and comparison of video coding standards, IEEE Transaction on Circuits and Systems for Video Technology, vol. 13, no. 7, pp. 688 – 703 (2003).
- [Wieg03b] Wiegand, T., Sullivan, G.J., Bjontegaard, G., Luthra, A.K.: Overview of the H.264/AVC video standard coding, IEEE Transaction on Circuits and Systems for Video Technology, vol. 13, no. 7, pp. 560 – 576 (2003).
- [Xie04] Xie, B., Zeng, W.: An improved rate-quantization model for rate control in real-time video coding, IEEE International Conference on Image Processing, ICIP, pp. 2765 – 2768, Singapore (2004)..
- [Xie07] Xie, Z., Liu, J.: Simple and efficient rate control for H.264/AVC, International Symposium on Intelligent Signal Processing and Communication Systems, ISPACS, pp. 862 – 865, Xiamen, China (2007).
- [Xu04] Xu, J., He, Y.: A novel rate control for H.264, IEEE International Symposium on Circuits and Systems, ISCAS, vol. 3, pp. 809 – 812, Vancouver, Canada (2004).
- [Yi04] Yi, X., Ling, N.: Rate control using enhanced frame complexity measure for H.264 video, IEEE Workshop on Signal Processing Systems, SIPS, pp. 263 – 268, Austin, Texas, USA (2004)
- [Yu05] Yu, H., Lin, Z., Pan, F.: An improved rate control algorithm for H.264, IEEE International Symposium on Circuits and Systems, ISCAS, pp. 312 – 315, Kobe, Japan (2005).

- [Yu09] Yu, L., Chen, S., Wang, J.: Overview of AVS-video coding standards, *Signal Processing: Image Communication*, vol. 24, no. 4, pp. 247 – 262, Elsevier (2009).
- [Yuan05] Yuan, W., Lin, S., Zhang, Y., Luo, H.: Optimum bit allocation and rate control for H.264/AVC, Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG (ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q.6), document JVT-O016, Busan, S.Korea (2005).
- [Yuan06] Yuan, W., Lin, S., Zhang, Y., Yuan, W., Luo, H.: Optimum bit allocation and rate control for H.264/AVC, *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 16, no. 6, pp. 705 - 715 (2006).
- [Zhao08] Zhao, J., Segall, A.: New results using entropy slices for parallel decoding, ITU-T Video Coding Experts Group (VCEG) (ITU-T SG16 Q.6), document VCEG-AI32, Berlin, Germany (2008).
- [Zhou08] Zhou, Y., Sun, Y., Feng, Z., Sun, S.: New rate-complexity-quantization modeling and efficient rate control for H.264/AVC, *IEEE International Conference on Multimedia and Expo, ICME*, pp. 717 - 720, Hannover, Germany (2008).
- [Zhou09] Zhou, Y., Sun, Y., Feng, Z., Sun, S.: New rate-distortion modeling and efficient rate control for H.264/AVC video coding, *Signal Processing: Image Communication*, vol. 24, no. 5, pp. 345 – 356, Elsevier (2009).
- [Zhuo07] Zhuo, L., Gao, X., Wang, Z., Feng, D.D., Shen, L.: A novel rate-quality model based H.264/AVC frame layer rate control method, *6th International Conference on Information, Communications & Signal Processing, ICICS*, pp. 1 - 5, Singapore (2007).

Dodatek A

Sekwencje testowe

Do przeprowadzenia badań wykorzystano 29 sekwencji wizyjnych o schemacie próbkowania 4:2:0 i rozdzielczości przestrzennej 704x576 punktów luminancji oraz częstotliwości obrazu wynoszącej 25 lub 30 Hz. Wszystkie stosowane sekwencje zostały zarejestrowane z wybieraniem kolejnoliniowym. Sekwencje te zostały podzielone na dwa zbiory: testowy (21 sekwencji) i kontrolny (8 sekwencji).

W badaniach wykorzystano sekwencje wyselekcjonowane jako sekwencje testowe do badań porównawczych prowadzonych w ramach działalności grup ekspertów MPEG (działa z ramienia ISO) i VCEG (działa z ramienia ITU-T). Wspomniane grupy ekspertów stosują te sekwencje w badaniach koderów wizyjnych i ogólnie podczas tworzenia nowych standardów kompresji sekwencji wizyjnych. Wybrane sekwencje są bardzo zróżnicowane pod względem treści oraz charakteru i dynamiki ruchu. Każda sekwencja przedstawia pojedyncze ujęcie, czyli nie zawiera cięć.

Oryginalnie sekwencje *bluesky*, *carpark*, *outdoor*, *pedestrian*, *riverbed*, *rushhour*, *station2*, *sunflower* i *tractor* są rozmiaru 1920x1080 punktów luminancji, a *night*, *raven*, *shuttle* oraz *spincalendar* 1280x720. Dlatego na potrzeby badań zostały przeskalowane do rozmiaru 4CIF. W tym celu wykorzystano oprogramowanie *DownConvertStatic* zawarte w referencyjnej implementacji kodeka skalowalnego SVC (ang. *Scalable Video Coding*) wersja 9.16 [JSVM], które do przeskalowania obrazów wykorzystuje filtrację opartą na funkcji *sinc* [Sun06].

Poniżej przedstawiono po jednym przykładowym kadrze z każdej sekwencji wizyjnej wykorzystanej do badań.

Zbiór sekwencji testowych



basket



bluesky



bus



cheer



city



crew



flow



football



harbour



ice



icon



pedestrian



riverbed



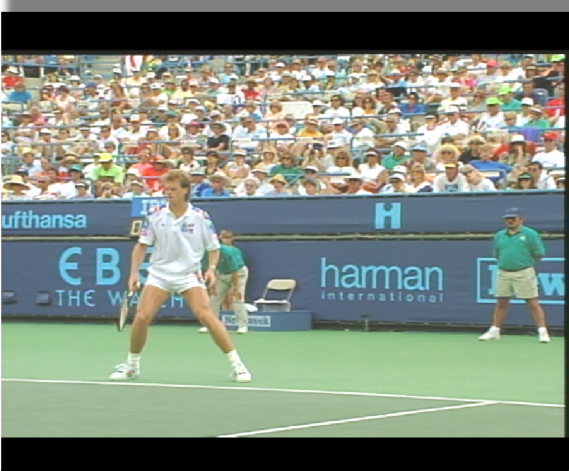
rushhour



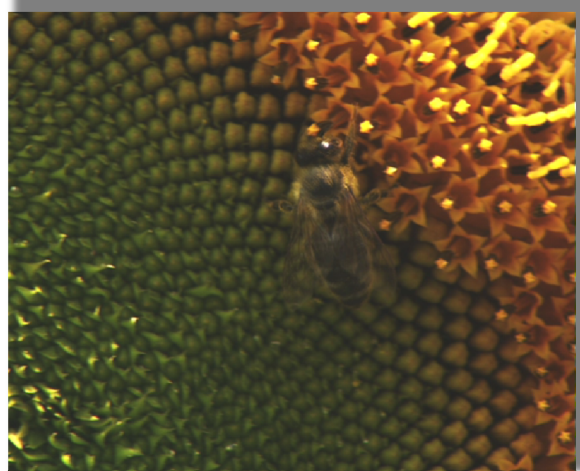
soccer



station2



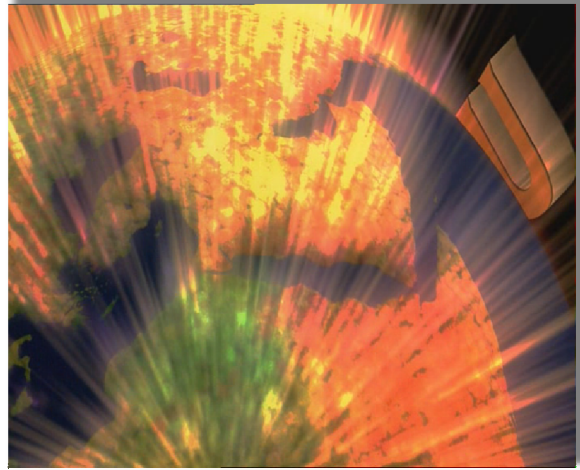
stefan



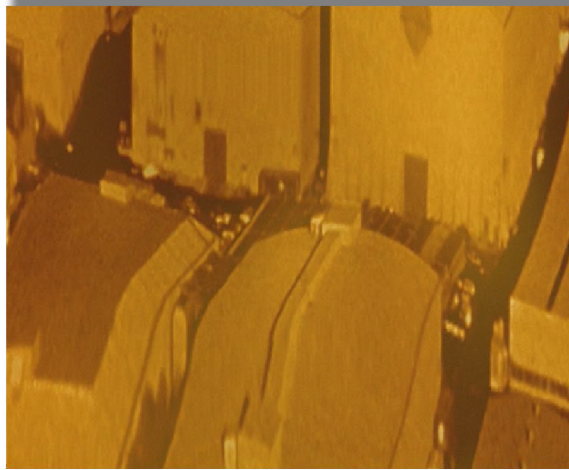
sunflower



tractor



universal



warner

Zbiór sekwencji kontrolnych



carpark



night



outdoor



raven



shuttle



spincalendar



toys



vintagecar

Dodatek B

Prędkości bitowe zakodowanych sekwencji wizyjnych

W tabelach B.1 do B.4 przedstawiono prędkość strumienia bitowego badanych testowych sekwencji wizyjnych o rozdzielczości 4CIF (704x576) dla wszystkich szerokości przedziałów kwantyzacji Q uzyskanych z wykorzystaniem kodera referencyjnego MPEG-4 AVC/H.264 w wersji 13.2 [AVCref]. Dla każdej sekwencji wizyjnej kolorem czerwonym, zaznaczono zakres prędkości bitowych strumienia danych pomiędzy 1 a 3 Mbit/s. Na podstawie analizy tych tabeli ustalono podział wartości Q na trzy przedziały.

Tabela B.1. Prędkość bitowa strumienia danych na wyjściu kodera dla badanych sekwencji wizyjnych w zależności od Q . Dodatkowo podano wartość indeksu QP

QP	Q	<i>basket</i>	<i>bluesky</i>	<i>bus</i>	<i>cheer</i>	<i>city</i>	<i>crew</i>
0	0,625	61 918 558	52 259 056	49 332 811	58 643 919	74 682 907	71 686 953
1	0,6875	58 887 350	49 961 223	47 071 458	55 813 759	71 027 083	68 395 178
2	0,8125	54 097 790	45 498 760	43 476 281	51 449 512	65 751 949	62 806 585
3	0,875	50 872 118	42 974 607	40 825 051	48 564 159	61 289 694	59 171 506
4	1	46 777 293	38 754 112	37 562 699	45 080 902	56 410 514	54 595 270
5	1,125	43 572 071	35 953 815	34 915 679	42 072 367	52 255 646	50 942 219
6	1,25	40 012 627	32 124 443	31 932 386	39 029 273	48 027 355	46 765 834
7	1,375	37 535 595	29 569 177	29 795 559	36 803 618	44 980 319	43 830 784
8	1,625	34 132 074	26 053 333	26 905 918	33 742 718	40 820 375	39 731 929
9	1,75	31 924 499	23 877 120	25 026 850	31 662 616	38 073 109	37 048 859
10	2	29 352 875	21 238 237	22 777 743	29 208 706	34 933 626	33 887 288
11	2,25	26 782 958	18 828 444	20 617 443	26 737 819	31 726 000	30 727 827
12	2,5	24 359 739	16 582 105	18 535 103	24 353 377	28 708 983	27 717 458
13	2,75	22 535 717	14 978 886	16 957 930	22 523 967	26 380 041	25 399 096
14	3,25	20 375 340	13 170 031	15 085 195	20 331 787	23 592 269	22 663 837
15	3,5	18 557 874	11 726 513	13 515 828	18 499 001	21 228 650	20 361 091
16	4	16 808 583	10 358 164	11 986 351	16 731 415	18 940 758	18 077 749
17	4,5	14 974 727	9 068 009	10 436 137	14 928 915	16 523 115	15 768 396
18	5	13 274 374	7 848 321	9 018 296	13 263 994	14 221 122	13 543 877
19	5,5	12 074 136	7 029 221	8 024 016	12 117 176	12 542 687	11 928 592
20	6,5	10 523 655	5 974 267	6 848 526	10 637 921	10 403 676	9 966 883
21	7	9 380 022	5 238 245	5 997 563	9 566 530	8 745 524	8 553 233
22	8	8 325 443	4 545 220	5 251 023	8 575 850	7 251 630	7 270 486
23	9	7 228 177	3 878 002	4 485 331	7 548 524	5 752 336	6 013 628
24	10	6 301 000	3 307 895	3 866 384	6 670 103	4 563 887	5 010 607
25	11	5 618 659	2 893 956	3 419 237	6 046 052	3 706 734	4 319 510
26	13	4 762 383	2 429 038	2 864 474	5 216 804	2 836 027	3 485 682
27	14	4 185 378	2 126 191	2 501 463	4 650 851	2 338 279	2 993 645
28	16	3 614 124	1 835 403	2 144 422	4 077 961	1 900 053	2 508 269
29	18	3 095 640	1 576 609	1 829 145	3 538 860	1 554 927	2 111 481
30	20	2 703 775	1 411 596	1 588 506	3 158 284	1 305 958	1 822 505
31	22	2 389 575	1 253 241	1 401 088	2 809 685	1 135 183	1 586 767
32	26	2 021 965	1 085 018	1 176 253	2 407 675	942 144	1 322 501
33	28	1 755 222	954 404	1 023 312	2 098 226	813 151	1 138 052
34	32	1 534 788	865 318	890 029	1 859 207	701 678	982 038
35	36	1 312 049	759 050	759 087	1 597 401	599 615	835 762
36	40	1 131 139	668 542	660 151	1 380 036	519 370	713 386
37	44	1 009 492	613 662	587 886	1 245 056	461 637	632 183
38	52	851 740	533 664	501 171	1 051 077	394 413	529 879
39	56	754 391	486 854	443 433	940 421	347 637	468 550
40	64	649 556	429 172	384 867	808 291	301 043	398 668
41	72	564 979	385 126	336 359	707 520	262 713	345 771
42	80	490 649	342 181	294 145	610 923	230 300	298 154
43	88	437 180	313 898	265 049	546 796	207 763	264 023
44	104	375 801	280 851	229 799	473 467	178 447	226 479
45	112	325 379	246 590	203 756	400 123	155 649	193 656
46	128	284 821	222 818	180 174	346 927	133 666	166 997
47	144	247 685	199 029	158 183	299 662	112 102	144 245
48	160	215 178	174 315	139 765	252 349	94 531	123 289
49	176	190 974	157 893	125 089	221 171	83 126	107 990
50	208	163 438	140 343	108 592	188 288	71 117	92 914
51	224	143 564	125 808	95 605	164 743	63 395	81 173

Tabela B.2. Prędkość bitowa strumienia danych na wyjściu kodera dla badanych sekwencji wizerunkowych w zależności od Q . Dodatkowo podano wartość indeksu QP (c.d.)

QP	Q	<i>flow</i>	<i>football</i>	<i>harbour</i>	<i>ice</i>	<i>icon</i>
0	0,625	59 190 533	48 106 306	83 310 926	47 278 173	26 953 987
1	0,6875	56 384 921	46 099 864	79 078 029	45 159 497	25 675 666
2	0,8125	51 816 969	42 699 046	72 070 688	41 241 177	23 412 781
3	0,875	49 108 723	40 242 319	67 049 860	38 706 533	22 014 974
4	1	45 380 789	37 070 782	61 612 287	34 740 914	19 872 132
5	1,125	42 564 867	34 404 082	57 086 872	31 914 533	18 420 200
6	1,25	39 090 423	31 394 767	52 385 209	28 088 273	16 427 187
7	1,375	36 674 133	29 173 495	49 137 454	25 551 596	15 107 595
8	1,625	33 341 735	26 141 849	44 949 636	22 117 337	13 332 002
9	1,75	31 145 497	24 140 670	42 182 694	20 000 205	12 264 918
10	2	28 634 948	21 719 985	39 006 361	17 211 434	10 955 720
11	2,25	26 092 187	19 399 210	35 827 248	14 466 885	9 749 818
12	2,5	23 741 524	17 122 244	32 795 357	11 639 618	8 587 713
13	2,75	21 945 937	15 381 792	30 450 406	9 698 352	7 746 109
14	3,25	19 780 712	13 400 937	27 568 436	8 023 783	6 798 371
15	3,5	17 955 128	11 742 753	25 204 542	6 819 858	6 026 670
16	4	16 235 681	10 175 988	22 863 758	5 816 844	5 256 753
17	4,5	14 411 668	8 656 829	20 402 569	4 941 495	4 524 639
18	5	12 703 629	7 305 638	18 121 397	4 194 134	3 807 227
19	5,5	11 526 913	6 430 003	16 507 178	3 714 517	3 295 714
20	6,5	9 965 801	5 402 573	14 390 458	3 129 941	2 678 545
21	7	8 797 543	4 733 468	12 827 486	2 725 850	2 255 510
22	8	7 736 302	4 133 556	11 354 650	2 380 456	1 875 729
23	9	6 616 998	3 558 552	9 799 473	2 031 886	1 545 705
24	10	5 662 102	3 073 476	8 460 931	1 748 133	1 265 891
25	11	4 971 636	2 731 303	7 516 476	1 553 161	1 075 961
26	13	4 116 579	2 307 777	6 214 410	1 312 642	873 986
27	14	3 530 487	2 026 904	5 347 184	1 161 852	750 978
28	16	2 972 764	1 743 776	4 446 017	1 013 549	636 590
29	18	2 477 722	1 497 810	3 653 258	884 383	539 695
30	20	2 125 227	1 314 640	3 025 311	784 033	466 080
31	22	1 843 417	1 164 574	2 570 183	707 917	409 217
32	26	1 543 082	987 250	2 083 225	615 320	343 553
33	28	1 322 340	857 054	1 764 680	550 347	300 160
34	32	1 158 257	751 215	1 484 462	493 618	261 091
35	36	992 032	645 327	1 244 234	437 313	225 521
36	40	860 215	554 844	1 048 567	390 797	193 881
37	44	777 303	501 095	910 133	358 062	174 641
38	52	666 882	423 901	754 462	316 152	149 928
39	56	599 402	380 330	645 303	288 798	134 585
40	64	525 316	328 862	542 718	258 408	118 686
41	72	468 092	292 578	455 369	233 563	105 753
42	80	413 126	258 073	384 572	209 651	94 405
43	88	374 954	233 494	332 401	192 687	85 227
44	104	330 666	203 834	273 612	171 513	74 585
45	112	287 877	177 536	233 315	154 238	66 718
46	128	256 348	156 779	196 878	137 729	59 321
47	144	225 254	139 073	165 198	124 765	54 016
48	160	193 064	120 910	139 260	110 743	48 811
49	176	173 895	107 085	120 153	100 392	44 864
50	208	148 131	91 685	100 132	87 506	40 307
51	224	129 519	80 618	84 941	78 080	37 209

Tabela B.3. Prędkość bitowa strumienia danych na wyjściu kodera dla badanych sekwencji wizyjnych w zależności od Q . Dodatkowo podano wartość indeksu QP (c.d.)

QP	Q	<i>pedestrian</i>	<i>riverbed</i>	<i>rushhour</i>	<i>soccer</i>	<i>station2</i>
0	0,625	47 273 916	83 224 382	50 092 935	66 583 837	49 673 527
1	0,6875	44 901 093	76 096 620	47 717 785	63 409 330	47 139 326
2	0,8125	40 584 911	71 076 397	43 228 055	58 175 606	42 454 220
3	0,875	37 984 609	64 830 254	40 417 986	54 495 179	39 775 956
4	1	33 752 679	59 639 355	35 805 055	49 932 435	35 176 208
5	1,125	31 045 158	53 881 545	32 949 782	46 193 023	32 269 749
6	1,25	27 186 186	50 342 041	28 690 907	42 042 835	28 199 790
7	1,375	24 619 860	45 911 890	25 930 123	39 253 784	25 366 636
8	1,625	21 227 231	43 114 313	22 244 029	35 323 495	21 683 629
9	1,75	19 251 776	39 960 987	20 109 311	32 859 078	19 489 328
10	2	16 655 099	36 801 423	17 306 242	29 970 771	16 652 861
11	2,25	14 260 716	33 795 816	14 841 430	27 186 972	14 102 980
12	2,5	11 876 923	31 521 688	12 525 430	24 462 754	11 558 869
13	2,75	10 221 549	28 749 929	10 945 772	22 458 962	9 748 615
14	3,25	8 581 199	26 489 439	9 290 481	20 068 457	7 872 608
15	3,5	7 324 991	24 272 300	8 035 053	18 043 177	6 244 421
16	4	6 243 511	22 012 233	6 877 975	16 136 527	4 928 529
17	4,5	5 359 796	19 925 144	5 909 317	14 145 791	3 895 604
18	5	4 614 521	18 440 808	5 080 274	12 255 490	3 078 808
19	5,5	4 129 218	16 463 843	4 524 467	10 864 226	2 616 889
20	6,5	3 539 192	15 061 581	3 860 376	9 180 759	2 111 180
21	7	3 160 454	13 683 395	3 419 936	7 871 034	1 813 259
22	8	2 808 759	12 272 258	3 012 519	6 741 472	1 560 065
23	9	2 472 573	11 037 613	2 616 537	5 612 825	1 341 080
24	10	2 177 905	10 125 683	2 276 478	4 723 755	1 162 492
25	11	1 962 411	8 905 464	2 014 877	4 037 445	1 024 012
26	13	1 696 348	8 066 791	1 696 256	3 334 577	879 065
27	14	1 522 451	7 178 014	1 490 318	2 876 673	781 359
28	16	1 335 215	6 339 717	1 275 379	2 430 379	687 587
29	18	1 172 500	5 648 211	1 093 381	2 101 710	602 254
30	20	1 042 080	5 086 694	946 855	1 825 840	536 358
31	22	936 141	4 352 697	826 678	1 631 954	483 417
32	26	806 407	3 830 181	696 396	1 381 459	421 770
33	28	715 649	3 317 444	601 921	1 201 311	374 847
34	32	631 950	2 818 017	519 298	1 060 974	332 950
35	36	549 649	2 397 497	444 422	912 678	292 438
36	40	481 188	2 096 401	382 108	796 936	258 226
37	44	431 335	1 716 649	338 105	709 804	232 804
38	52	369 706	1 475 920	289 041	598 677	200 002
39	56	331 142	1 223 954	254 905	532 931	176 814
40	64	288 189	998 994	220 098	463 429	151 914
41	72	253 412	836 530	191 844	401 922	128 767
42	80	223 471	709 130	167 325	352 847	109 344
43	88	201 822	563 882	148 046	314 611	94 843
44	104	175 397	468 544	127 852	271 553	79 781
45	112	154 761	381 175	111 207	236 373	68 154
46	128	136 357	310 379	96 206	207 951	57 785
47	144	121 102	253 059	83 957	183 795	48 486
48	160	106 595	212 024	72 710	161 146	40 139
49	176	95 135	169 048	64 418	144 192	34 915
50	208	82 266	141 061	55 151	124 587	28 831
51	224	73 204	117 080	49 311	111 958	25 313

Tabela B.4. Prędkość bitowa strumienia danych na wyjściu kodera dla badanych sekwencji wizyjnych w zależności od Q . Dodatkowo podano wartość indeksu QP (c.d.)

QP	Q	<i>stefan</i>	<i>sunflower</i>	<i>tractor</i>	<i>universal</i>	<i>warner</i>
0	0,625	50 098 405	48 070 254	66 363 522	37 064 911	44 343 044
1	0,6875	47 738 473	45 347 041	63 352 210	35 399 996	42 471 321
2	0,8125	43 811 257	40 693 608	58 000 359	32 322 966	38 885 745
3	0,875	41 355 458	37 997 099	54 628 591	30 439 155	36 703 671
4	1	38 072 506	33 285 338	49 626 984	27 543 965	33 292 628
5	1,125	35 552 599	30 503 873	46 046 882	25 615 314	30 999 709
6	1,25	32 605 396	26 279 802	41 790 601	22 992 945	27 867 364
7	1,375	30 542 258	23 529 536	38 765 908	21 258 396	25 864 877
8	1,625	27 619 411	19 907 227	34 687 274	18 899 566	23 018 778
9	1,75	25 695 072	17 823 619	32 149 896	17 467 110	21 310 554
10	2	23 433 434	15 021 839	29 066 950	15 683 980	19 206 224
11	2,25	21 166 570	12 473 046	26 171 108	14 023 326	17 225 907
12	2,5	19 002 096	10 229 380	23 263 643	12 401 384	15 293 556
13	2,75	17 406 530	8 749 276	21 184 006	11 220 074	13 864 055
14	3,25	15 575 895	7 351 547	18 846 807	9 889 740	12 201 277
15	3,5	14 043 137	6 269 433	16 805 709	8 798 712	10 825 245
16	4	12 579 835	5 371 143	14 817 151	7 699 332	9 437 688
17	4,5	11 069 612	4 568 812	12 893 493	6 675 739	8 115 098
18	5	9 749 696	3 925 854	11 258 277	5 705 906	6 871 779
19	5,5	8 823 095	3 498 218	9 997 761	5 045 765	6 015 097
20	6,5	7 626 664	2 973 324	8 488 957	4 255 001	5 001 782
21	7	6 750 401	2 624 014	7 434 585	3 718 372	4 290 641
22	8	5 911 407	2 324 489	6 492 697	3 241 421	3 651 805
23	9	5 060 555	2 023 055	5 576 183	2 782 332	3 050 275
24	10	4 321 053	1 777 046	4 757 401	2 392 136	2 568 450
25	11	3 805 087	1 567 707	4 195 898	2 120 005	2 236 331
26	13	3 058 483	1 353 678	3 479 838	1 781 757	1 835 136
27	14	2 636 126	1 188 416	3 046 520	1 566 758	1 587 881
28	16	2 181 693	1 040 680	2 645 496	1 348 358	1 343 248
29	18	1 794 668	908 212	2 282 945	1 158 752	1 141 560
30	20	1 525 852	818 522	2 022 709	1 025 953	1 003 174
31	22	1 300 234	724 412	1 789 398	902 609	885 202
32	26	1 077 204	628 497	1 534 268	767 298	751 731
33	28	928 909	549 229	1 347 665	664 531	660 458
34	32	799 193	497 579	1 179 621	584 519	589 438
35	36	686 687	431 465	1 021 815	503 310	517 266
36	40	581 117	378 336	875 492	431 467	454 510
37	44	528 283	344 328	781 555	387 949	416 032
38	52	443 109	295 335	661 955	330 530	361 841
39	56	398 121	269 711	584 490	296 252	332 097
40	64	349 858	236 407	500 932	256 193	293 230
41	72	307 058	211 124	439 822	228 779	264 197
42	80	266 036	190 062	384 802	200 897	233 544
43	88	236 296	175 194	346 601	182 841	213 413
44	104	207 081	158 908	295 493	162 265	188 619
45	112	184 684	140 178	246 970	140 686	162 960
46	128	161 678	128 544	209 772	124 195	143 452
47	144	140 950	117 998	177 017	111 921	126 486
48	160	123 210	105 630	146 560	97 593	108 919
49	176	107 749	98 556	126 835	88 844	96 443
50	208	94 676	90 433	106 470	78 379	83 363
51	224	83 703	83 473	93 300	71 949	74 408