

Politechnika Poznańska
Wydział Elektroniki i Telekomunikacji
Katedra Telekomunikacji Multimedialnej i Mikroelektroniki

Autoreferat rozprawy doktorskiej

Modelowanie zaawansowanych koderów wizyjnych

Tomasz Grajek



Promotor: Prof. dr hab. inż. Marek Domański

Poznań, 2010

Spis treści

Spis treści	3
1. Wprowadzenie	5
1.1. Modelowanie koderów wizyjnych	5
1.2. Cele i teza pracy	8
1.3. Przegląd pracy	8
2. Zasadnicze osiągnięcia rozprawy	11
2.1. Wprowadzenie	11
2.2. Metodologia prowadzenia badań	12
2.3. Oryginalne modele kodera wizyjnego standardu MPEG-4 AVC/H.264	15
2.3.1. Wprowadzenie	15
2.3.2. Modelowanie średniej liczby bitów potrzebnych do reprezentacji skompresowanego obrazu	17
2.3.3. Modelowanie liczby bitów reprezentacji pojedynczego obrazu	19
2.3.4. Modelowanie łącznej liczby bitów potrzebnych do reprezentacji składowych luminancji i chrominancji pojedynczego obrazu	22
3. Podsumowanie i wnioski	25
3.1. Wnioski	25
3.2. Oryginalne osiągnięcia	26
4. Dorobek naukowy autora	29
4.1. Publikacje naukowe autora	29
4.2. Oryginalne osiągnięcia autora nie ujęte w rozprawie, a zastosowane w praktyce	30

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, AVS i VC-1 (ten ostatni jest związany z powszechnie znaną techniką WMV¹ 9.0). 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). Stosuje się je w systemach telekomunikacyjnych i teleinformatycznych nowych generacji, w telewizji cyfrowej, wideotelefonii, cyfrowych systemach nadzoru wizyjnego, transmisji obrazu w sieciach komputerowych itd.

W hybrydowych 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, liczbą szczegółów itd.). Szeroko stosowany adaptacyjny wybór trybów kodowania 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

¹ Windows Media Video

(tryb stałej prędkości bitowej – CBR¹). 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 – VBR²). Ponadto, opisane wyżej procesy muszą uwzględniać wielkość dopuszczalnego opóźnienia kodowania.

Znalezienie optymalnych wartości parametrów sterujących pracą kodera wymaga dużych nakładów obliczeniowych. Dlatego w praktycznych implementacjach stosuje się różnego rodzaju uproszczenia. Jednym z możliwych rozwiązań jest wykorzystanie modelu opisującego związek 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.

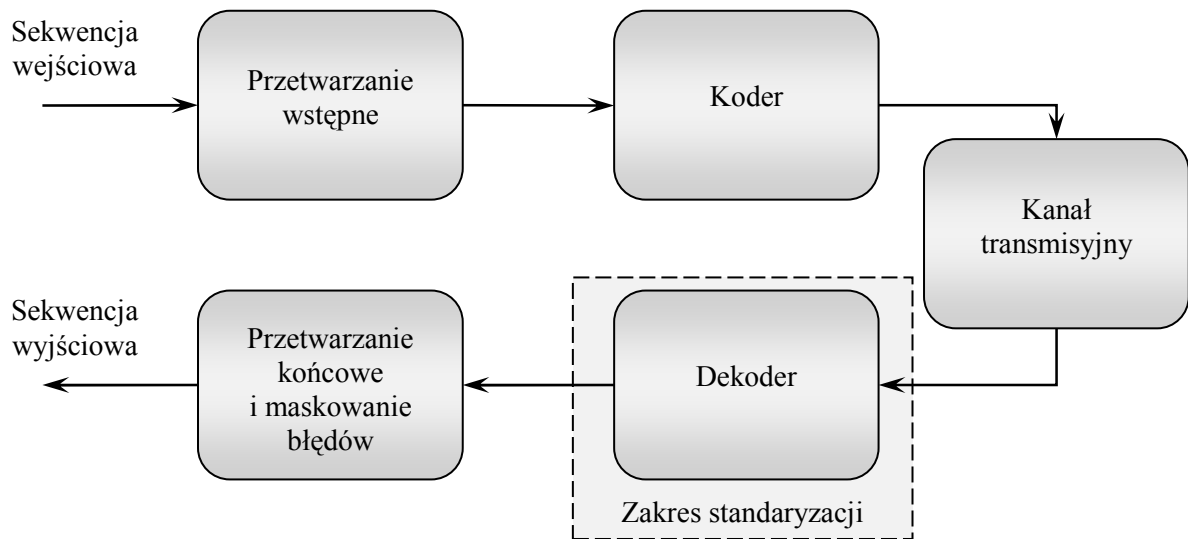
Sformułowanie choćby przybliżonych zależności ilościowych pomiędzy prędkością bitową danych na wyjściu kodera, parametrami sterującymi kodera i parametrami treści sekwencji ma olbrzymie znaczenie poznawcze. Posługując się takimi modelami można zbadać zachowanie się kodera w różnych stanach. Dla zaawansowanych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

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 parametr Q .

¹ Constant Bit Rate

² Variable Bit Rate

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. 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

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. 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.

Rozprawa jest wynikiem przemyśleń, analiz, badań, jak również doświadczeń zdobytych przez doktoranta podczas realizacji wielu projektów badawczo-rozwojowych wykonywanych na zlecenia dużych firm zarówno polskich, jak i zagranicznych.

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 kodera standardu MPEG-4 AVC/H.264. Modele te mają opisywać związki pomiędzy prędkością bitową strumienia danych na wyjściu kodera i szerokością przedziałów kwantyzacji Q .

Ponadto, celem pracy jest **analiza dokładności modelowania kodera 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 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.

1.3. 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 rozprawy 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 modelowania 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 zreferowano badania eksperymentalne 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.

2. Zasadnicze osiągnięcia rozprawy

2.1. Wprowadzenie

Opracowanie modeli matematycznych koderów wizyjnych ma duże znaczenie zarówno poznawcze, jak i praktyczne. Znajomość modelu opisującego relację pomiędzy parametrami kodera a wielkością zakodowanego strumienia pozwala na przewidywanie zachowania się kodera w różnych warunkach. Ponadto, taki model może zostać wykorzystany w algorytmie sterowania koderem do ustalania wartości szerokości przedziałów kwantyzacji Q dla zadanej liczby bitów. Dlatego w literaturze dotyczącej technik hybrydowego kodowania sekwencji wizyjnych modelowanie koderów wizyjnych jest stosunkowo często poruszonym tematem.

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. 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” (RDO¹)). 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

¹ Rate Distortion Optimization

typu P. Obecnie brakuje dokładnych modeli matematycznych opisujących obrazy pozostałych typów, a w szczególności obrazy typu B.

Powyższe argumenty skłoniły autora do podjęcia badań nad modelowaniem koderów wizyjnych, co zaowocowało zaproponowaniem własnych modeli kodera wizyjnego standardu MPEG-4 AVC/H.264.

2.2. Metodologia prowadzenia badań

Tworzenie modelu autor podzielił 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.

Proponowane modele zostały zbudowane w oparciu o analizę statystyczną danych eksperymentalnych zebranych dla zbioru sekwencji testowych o rozdzielczości przestrzennej 704x576 punktów (4CIF¹) i częstotliwościach obrazu 25 lub 30 Hz. Zbiór ten obejmuje sekwencje testowe wykorzystywane do badań porównawczych prowadzonych w ramach działalności grup ekspertów MPEG² (działa z ramienia ISO³) i VCEG⁴ (działa z ramienia ITU-T⁵). Wspomniane grupy ekspertów stosują te sekwencje do testowania koderów wizyjnych i podczas tworzenia nowych standardów kompresji sekwencji wizyjnych. Wybrane sekwencje są bardzo zróżnicowane pod względem treści oraz charakteru i dynamiki ruchu. Każda sekwencja testowa zawierała obrazy tylko z pojedynczego ujęcia, czyli w sekwencjach wizyjnych nie występowały cięcia. Należy zaznaczyć, że kiedy rozpoczynano badania nie było ustalonego zbioru reprezentatywnych sekwencji testowych wysokiej rozdzielczości (HD⁶). Dlatego badania przeprowadzono na sekwencjach o standardowej rozdzielczości telewizyjnej, a dostępne sekwencje o wyższej rozdzielczości niż 4CIF zostały odpowiednio zdecydowane.

¹ 4×Common Intermediate Format

² Motion/Moving Pictures Experts Group

³ International Organization for Standardization

⁴ Video Coding Experts Group

⁵ International Telecommunication Union - Telecommunication Standardization Sector

⁶ High Definition

Sekwencje zostały zakodowane kolejno ze wszystkimi dozwolonymi szerokościami przedziałów kwantyzacji Q (jedno zakodowanie sekwencji dla jednej, stałej szerokości przedziałów kwantyzacji Q). Do kodowania przyjęto strukturę grupy obrazów GOP¹ typową dla zastosowań telewizyjnych (4P2B²). 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). 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. Przydatność wybranego wzoru sprawdza się metodą rektyfikacji. Rektyfikacja polega na tym, że przy założeniu istnienia zależności funkcyjnej między wielkościami x i y , znajdujemy nowe wielkości X i Y postaci:

$$X = \varphi(x, y), \quad Y = \psi(x, y), \quad (2.1)$$

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

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. Do stwierdzenia, na ile otrzymane w ten sposób dane są zgodne z liniową zależnością autor wykorzystał współczynnik determinacji liniowej. Procedura poszukiwania i weryfikacji jest powtarzana do czasu znalezienia funkcji o możliwie małej liczbie parametrów i dobrze dopasowanej do danych

¹ Group of Pictures

² 4P2B – grupa obrazów składająca się z kolejnych obrazów typu: I, B, B, P, B, B, P, B, B, P, B, B, P, B, B

eksperymentalnych. Takie podejście jest dokładnie opisane w literaturze 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, a jako punkty startowe do procedury optymalizacji wykorzystano wartości parametrów wyznaczone metodą najmniejszych kwadratów dla zredukowanych danych.

Dla tak wyznaczonych modeli i ich parametrów, autor określał maksymalny i/lub średni względny błąd aproksymacji danych eksperymentalnych (wzory 2.3 i 2.4). Względny błąd aproksymacji oznaczono jako:

$$\varepsilon(x_i) = \frac{|y_z(x_i) - y(x_i)|}{y_z(x_i)} \cdot 100\%, \quad (2.2)$$

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:

$$\max_{x_i} \varepsilon(x_i), \quad (2.3)$$

a średni względny błąd aproksymacji jako:

$$\frac{1}{n} \sum_{i=1}^n \varepsilon(x_i), \quad (2.4)$$

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 2.4). W rozprawie dokładność aproksymacji danych przez badany model jest nazywana w skrócie dokładnością modelu.

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

2.3. Oryginalne modele koder a wizyjnego standardu MPEG-4 AVC/H.264

2.3.1. Wprowadzenie

Rozprawa przedstawia wyniki badań własnych, których celem było stworzenie modeli matematycznych opisujących związki pomiędzy liczbą bitów wytwarzanych przez koder i szerokością przedziałów kwantyzacji Q , co ogólnie można opisać jako:

$$B = f(Q), \quad (2.5)$$

gdzie:

- B to liczba bitów produkowanych przez koder,
- Q to szerokość przedziałów kwantyzacji.

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

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.

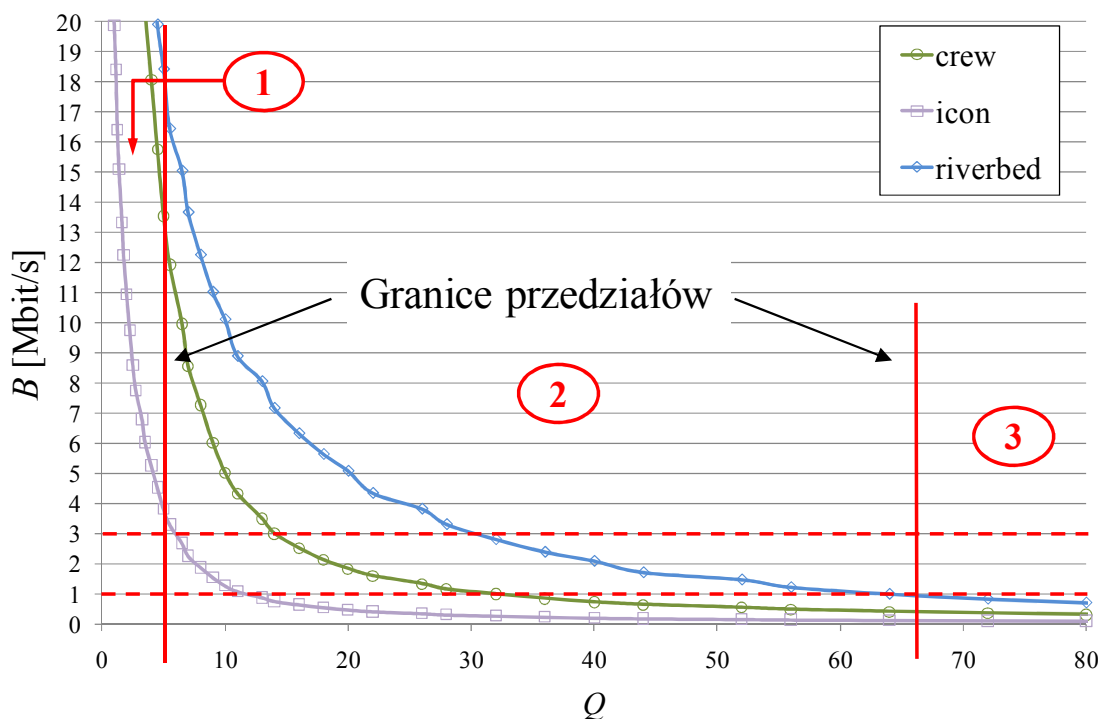
W literaturze dobrze znane są metody pozwalające na znalezienie funkcji, która przechodzi dokładnie przez punkty pomiarowe (np. interpolacja wielomianowa). Niestety zastosowanie takich metod w przypadku poszukiwania dopasowania do wielu punktów pomiarowych oznaczałoby konieczność wyznaczenia wartości 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 poprzednim punkcie. Jest to takie samo podejście, które z sukcesem zastosowano dla koderów standardu MPEG-2 i H.263.

Niestety nie powiodły się próby stworzenia dokładnych modeli dla dowolnej wartości szerokości przedziałów kwantyzacji Q . Dlatego zaproponowane modele określono niezależnie dla przyjętych 3 przedziałów wartości Q . Granice przedziałów zostały dobrane tak (tabela 2.1), aby środkowy przedział obejmował wartości Q odpowiadające użytecznemu zakresowi prędkości bitowej sekwencji wizyjnych formatu 4CIF, tj. od 1 do 3 Mbit/s dla badanych sekwencji testowych (rysunek 2.1).

Tabela 2.1. Granice przedziałów

Przedział	Zakres Q	Zakres QP
1	<0,625; 4,5>	<0;17>
2	<5; 64>	<18;40>
3	<72; 224>	<41;51>



Wykres 2.1. Przyjęte granice przedziałów na przykładzie 3 sekwencji wizyjnych. Dla przejrzystości wykresu obcięto zakres wartości B i Q

Należy zaznaczyć, że dla poszczególnych sekwencji wizyjnych, użytecznemu zakresowi prędkości bitowych strumienia danych odpowiada różny zakres wartości Q . 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 2.1).

W strumieniu danych nie umieszcza się wartości szerokości przedziałów kwantyzacji Q wprost, tylko poprzez podanie indeksu parametru kwantyzacji QP^1 dla składowej luminancji. Dlatego w tabeli 2.1 obok zakresów wartości Q podano odpowiadające im zakresy indeksu QP .

Podobny podział wartości Q został przyjęty do badań w wielu innych pracach. Jednak autorzy tych prac ograniczali się do analizy kodera tylko dla wartości Q należących do przedziału drugiego.

2.3.2. Modelowanie średniej liczby bitów potrzebnych do reprezentacji skompresowanego obrazu

Rozdział 4 rozprawy zawiera autorskie propozycje modeli 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 dla każdego z trzech przedziałów Q .

W przedziale pierwszym autor zaproponował model wielomianowy 3-go stopnia z 1 parametrem i zestawem uniwersalnych stałych:

$$B_{\text{śred}}(Q, d) = (a_1d + a_2) \cdot Q^3 + (b_1d + b_2) \cdot Q^2 + (c_1d + c_2) \cdot Q + d, \quad (2.6)$$

gdzie:

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

¹ Quantization Parameter

Dla zaproponowanego modelu (wzór 2.6) 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.

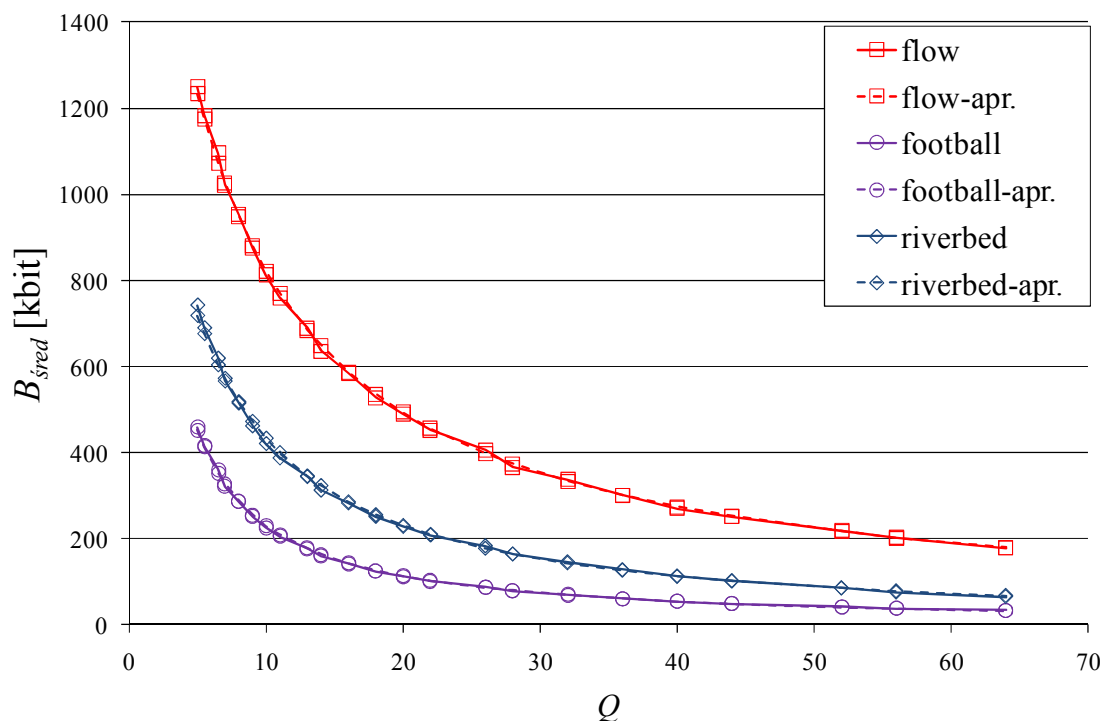
W przedziale drugim, gdzie szczególnie istotna jest dokładność aproksymacji danych, wybrano model hiperboliczny z 3 parametrami:

$$B_{\text{śred}}(Q, a, b, c) = \frac{a}{Q^b + c}, \quad (2.7)$$

gdzie:

- a, b, c to parametry modelu,
- Q to szerokość przedziałów kwantyzacji,
- $B_{\text{śred}}$ to średnia liczba bitów reprezentacji obrazu danego typu (I, P lub B).

Średni względny błąd dopasowania dla większości sekwencji testowych jest poniżej 2, 3 i 8%, odpowiednio dla obrazów typu I, P i B. Na wykresie 2.2. pokazano przykładowe krzywe eksperymentalne i ich oszacowania modelem hiperbolicznym dla kilku sekwencji wizyjnych dla przedziału drugiego, dla obrazów typu I.



Wykres 2.2. Krzywe eksperymentalne (linia ciągła) i oszacowane modelem hiperbolicznym (linia przerywana) dla obrazów typu I dla 3 przykładowych sekwencji wizyjnych (przedział 2)

W przedziale trzecim autor także zaproponował uproszczony model wielomianowy (wzór 2.6). 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.

Dodatkowo, dla uproszczonego modelu wielomianowego (wzór 2.6) w przedziałach pierwszym i trzecim autor przeprowadził eksperymenty z wykorzystaniem dodatkowego zbioru kontrolnego, który składał się z 8 sekwencji wizyjnych formatu 4CIF. Sekwencje te nie występowały w zbiorze testowym użytym do wyznaczenia wartości uniwersalnych stałych. Wyniki eksperymentu potwierdziły, że zaproponowany uproszczony model wielomianowy z zestawem uniwersalnych stałych dobrze aproksymuje dane eksperymentalne uzyskane także dla sekwencji spoza zbioru testowego dla obrazów typu I, P i B.

Zaproponowane modele określają średnią liczbę bitów reprezentacji obrazów w ramach pojedynczego ujęcia. Z tego względu mogą 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.

2.3.3. Modelowanie liczby bitów reprezentacji pojedynczego obrazu

W poprzednim punkcie 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 mogą zostać wykorzystane 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.

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 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 przez predykcję wartości parametrów z poprzedniego obrazu tego samego typu, będą określane mianem „predykcji parametrów”.

W tabeli 2.2 podsumowano wartości średniego względnego błędu aproksymacji uśrednionego po wszystkich obrazach poszczególnych typów.

Tabela 2.2. 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
	Przedział 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
	Przedział 2					
hiperboliczny(3)	1,30	11,77	3,11	11,68	7,60	21,95
	Przedział 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

Uzyskane wyniki potwierdzają, że modele zaproponowane do określania średniej liczby bitów reprezentacji obrazu danego typu 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. Uproszczenie modelu wielomianowego do jednego parametru pogarsza dokładność aproksymacji danych odpowiednio o około 2, 5 i 19 punktów

procentowych dla obrazów typu I, P i B dla pierwszego przedziału i 4, 12 oraz 32 punkty procentowe dla trzeciego przedziału.

Dokonano także weryfikacji dokładności uproszczonego modelu wielomianowego z wykorzystaniem kontrolnego zbioru sekwencji wizyjnych. Wyniki eksperymentu potwierdziły, że zaproponowany uproszczony model wielomianowy z zestawem uniwersalnych stałych dobrze aproksymuje dane eksperymentalne uzyskane także dla sekwencji spoza zbioru testowego dla obrazów typu I, P i B.

Dodatkowo dla obrazów typu I autor zaproponował miary charakteryzujące złożoność treści obrazów tego typu, których wykorzystanie pozwala zwiększyć dokładność modelu hiperbolicznego w przedziale drugim o około 7 punktów procentowych przy predykcji wartości parametrów z poprzedniego obrazu typu I należącego do tego samego ujęcia.

Zaproponowane miary złożoności treści obrazu, to:

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

$$\overline{\sigma_Y} = \frac{1}{N_Y} \sum_{i=1}^{N_Y} \sigma_i^Y, \quad (2.8)$$

gdzie:

- $\overline{\sigma_Y}$ to odchylenie standardowe wartości próbek składowej luminancji obrazu wyznaczone w blokach o rozmiarze 4×4 punkty uśrednione po wszystkich blokach w obrazie,
- σ_i^Y to odchylenie standardowe wartości próbek składowej luminancji obrazu w i -tym bloku o rozmiarze 4×4 punkty,
- N_Y to liczba nienakładających się bloków luminancji o rozmiarze 4×4 punkty w obrazie.

2. suma uśrednionych odchyleń standardowych wartości próbek składowych luminancji (Y) i chrominancji (Cb i Cr) obrazu wyznaczonych w blokach o rozmiarze 4×4 punkty:

$$\overline{\sigma_{YCbCr}} = \frac{1}{N_Y} \sum_{i=1}^{N_Y} \sigma_i^Y + \frac{1}{N_{Cb}} \sum_{i=1}^{N_{Cb}} \sigma_i^{Cb} + \frac{1}{N_{Cr}} \sum_{i=1}^{N_{Cr}} \sigma_i^{Cr}, \quad (2.9)$$

gdzie:

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

Szczegółowe wyniki dotyczące wykorzystania miar charakteryzujących złożoność treści obrazów typu I przedstawiono w punkcie 5.5 rozprawy.

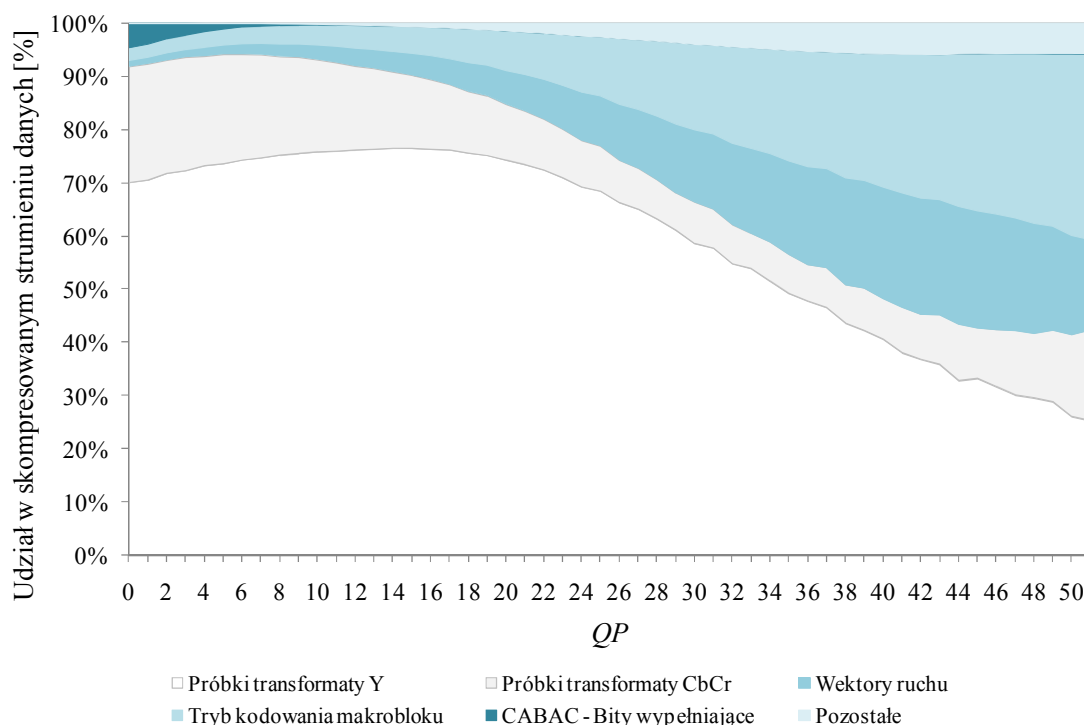
W literaturze dotyczącej modelowania hybrydowych koderów wizyjnych można znaleźć propozycje modelowania liczby bitów reprezentacji pojedynczego obrazu danego typu. Natomiast nie są dostępne implementacje takich rozwiązań. Dlatego trudno byłoby przeprowadzić eksperymentalne badania porównawcze.

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

Na wykresie (2.3) pokazano, uśredniony po badanych sekwencjach, udział wybranych typów danych w całkowitym strumieniu bitowym reprezentującym skompresowaną sekwencję wizyjną zakodowaną ze strukturą grupy obrazów GOP – 4P2B¹ (wybrano QP ze względu na większą przejrzystość wykresu niż dla Q). 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

¹ 4P2B – grupa obrazów składająca się z kolejnych obrazów typu: I, B, B, P, B, B, P, B, B, P, B, B, P, B, B

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 2.3. 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 (IBBPBBPBBPBBPBB)

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 . Autor postanowił sprawdzić, czy modele zaproponowane w poprzednich punktach można wykorzystać również do opisanie wspomnianej zależności. Także w referencyjnej implementacji kodera standardu MPEG-4 AVC/H.264 modeluje się wspomnianą zależność, ale tylko dla obrazów typu P.

Do modelowania liczby bitów potrzebnych do reprezentacji składowych luminancji i chrominancji autor zaproponował wykorzystanie takich samych modeli jak opisanych w rozdziałach 4 i 5 rozprawy. Podsumowanie uzyskanych wyników

umieszczono w tabeli 2.3. 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. Z analizy danych z tabeli 2.3 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. Należy zaznaczyć, że oszacowanie liczby bitów w przedziale trzecim jest trudniejsze niż w przedziałach pierwszym i drugim, ponieważ dla dużych wartości Q koder produkuje małą liczbę bitów (zwłaszcza dla obrazów typu P i B). Dlatego nawet niewielka pomyłka w oszacowaniu liczby bitów powoduje duże względne błędy aproksymacji.

Tabela 2.3. Średni względny błąd dopasowania uśredniony po obrazach danego typu

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

Analogicznie jak w dwóch poprzednich punktach dokonano weryfikacji dokładności uproszczonego modelu wielomianowego z zestawem uniwersalnych stałych z wykorzystaniem kontrolnego zbioru sekwencji wizyjnych.

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. Dlatego trudno byłoby przeprowadzić eksperymentalne badania porównawcze.

3. Podsumowanie i wnioski

3.1. 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 osobne modele dla trzech typów obrazów, a obecnie modeluje się przeważnie tylko obrazy typu P.

Przy kodowaniu sekwencji zaawansowanymi koderami wizyjnymi mamy do czynienia ze znacznym udziałem danych nagłówkowych w całkowitym strumieniu danych. Dlatego dalsze prace mogłyby dotyczyć modelowania także innych elementów składni 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).

3.2. Oryginalne osiągnięcia

Zasadniczym, oryginalnym, naukowym osiągnięciem rozprawy jest zaproponowanie 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 **nowe globalne modele strumienia bitowego** wyrażające zależności między szerokością przedziałów kwantyzacji Q a:

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

Istotnym elementem rozprawy jest analiza dokładności zaproponowanych modeli w różnych zakresach wartości szerokości przedziałów kwantyzacji Q .

Ad. 1. Autor określił dokładność zaproponowanych modeli dla 3 zakresów wartości Q dla obrazów typu I, P i B. **Modele pozwalające na wyznaczenie średniej liczby bitów reprezentacji obrazu danego typu w ramach jednego ujęcia nie są znane z literatury.**

Ad. 2. W rozprawie określono dokładność zaproponowanych modeli w dwóch wariantach:

- wartości parametrów modeli zostały wyznaczone indywidualnie dla każdego obrazu,
- wartości parametrów modeli dla bieżącego obrazu danego typu zostały wyznaczone na podstawie wartości parametrów z wcześniejszego obrazu tego samego typu.

Dodatkowo w rozprawie zaproponowano **globalny model strumienia bitowego z parametrem (miarą) charakteryzującym złożoność treści obrazów typu I**.

Zbadano dokładność modeli wykorzystujących następujące miary charakteryzujące złożoność treści obrazów typu I w sekwencji wizyjnej:

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

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

Ad. 3. Autor określił dokładność zaproponowanych modeli dla 3 zakresów wartości Q dla obrazów typu I, P i B.

W rozprawie 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 modelu zastosowanego w referencyjnej implementacji kodera średnio o około 11,5 punktów procentowych.** Należy podkreślić, że model z referencyjnej implementacji kodera standardu MPEG-4 AVC/H.264 jest

szeroko stosowany w praktycznych implementacjach zaawansowanych koderów wizyjnych. Uzyskane wyniki wskazują na potencjalne możliwości praktycznych zastosowań opracowanych modeli.

Dodatkowo w rozprawie zaproponowano **model strumienia bitowego z parametrem (miarą) charakteryzującym złożoność treści obrazów typu I**. Zbadano dokładność modelu wykorzystującego sumę uśrednionych odchyleń standardowych wartości próbek składowych luminancji i chrominancji wyznaczonych w blokach o rozmiarze 4×4 punkty jako miarę charakteryzującą złożoność treści obrazów typu I w sekwencji wizyjnej.

Zaproponowane modele mają duże znaczenie poznawcze, gdyż dotychczas w literaturze nie opisano globalnych modeli analogicznych do tych przedstawionych w rozprawie. Modele te opisują charakter zależności między liczbą bitów produkowanych na wyjściu zaawansowanego koderu wizyjnego i szerokością przedziałów kwantyzacji Q , która jest podstawową wielkością sterującą koderem. Opracowane modele pozwalają wyznaczać przybliżoną wielkość strumienia generowanego przez zaawansowany koder wizyjny przy różnym charakterze treści kodowanego materiału wizyjnego. Jest to bardzo istotne z punktu widzenia przyszłych zastosowań zaproponowanych modeli w procedurach sterowania zaawansowanymi koderami wizyjnymi.

4. Dorobek naukowy autora

4.1. Publikacje naukowe autora

Konferencje międzynarodowe o światowym zasięgu i surowym procesie recenzowania:

1. T. Grajek, M. Domański: “*A Global Model of AVC/H.264 Video Encoders*”, 27th Picture Coding Symposium, PCS, 6-8 maja 2009, str. 1-4, Chicago, Illinois, USA (2009), także: IEEE Xplore Digital Library (<http://ieeexplore.ieee.org>).

Wydawnictwa z tzw. „listy filadelfijskiej”:

2. T. Grajek, M. Domański: “*A Simple Quantitative Model of AVC/H.264 Video Coders*” w L. Bolc, J.L. Kulikowski, and K. Wojciechowski (eds.): *Computer Vision and Graphics, International Conference, ICCVG 2008, Lecture Notes in Computer Science - LNCS 5337*, pp. 174-183, Springer-Verlag (2009).
3. T. Grajek, M. Domański: “*Single Frame Rate-Quantization Model for MPEG-4 AVC/H.264 Video Encoders*” w L. Bolc, R. Tadeusiewicz, and L.J. Chmielewski (eds.): *Computer Vision and Graphics, Proc. ICCVG 2010, Lecture Notes in Computer Science, Springer-Verlag (2010)* – zgłoszony, dwie pozytywne recenzje.

Konferencje międzynarodowe:

4. M. Domański, T. Grajek, K. Klimaszewski, M. Kurc, O. Stankiewicz, J. Stankowski, K. Wegner, „*Poznań Multiview Video Test Sequences and Camera Parameters*”, ISO/IEC JTC1/SC29/WG11 MPEG 2009/M17050, Xian, Chiny, (2009).

Czasopisma krajowe:

5. M. Domański, T. Grajek, J. Marek, „*Zaawansowana kompresja cyfrowych sygnałów wizyjnych - standard AVC/H.264*”, Systemy Alarmowe, nr 2, marzec-kwiecień 2005, rok wyd. XIV, vol. 75, str. 14-21, Warszawa (2005).
6. T. Grajek „*Model ilościowy kodera wizyjnego standardu MPEG-4 AVC/H.264 z parametrem charakteryzującym treść sekwencji*”, Przegląd Telekomunikacyjny, również: Krajowa Konferencja Radiokomunikacji, Radiofonii i Telewizji, KKRRiT, 16-18 czerwca 2010, Kraków (2010) – przyjęty do druku po pozytywnych recenzjach.

Konferencje krajowe:

7. T. Grajek, D. Karwowski, „*Złożoność obliczeniowa i efektywność kodowania entropijnego w standardzie H.264/AVC*”, Krajowa Konferencja Radiokomunikacji, Radiofonii i Telewizji, KKRRiT, 15-17 czerwca 2005, str. 377-380, Kraków (2005).
8. T. Dzieścielewski, T. Grajek, J. Marek, „*Eksperymentalna analiza częstotliwości wyboru trybów zaawansowanego kodowania wizyjnego*”, X Poznańskie Warsztaty Telekomunikacyjne, 8-9 grudnia 2005, str. 115-120, Poznań (2005).
9. T. Dzieścielewski, T. Grajek, J. Marek, „*Eksperymentalna analiza efektywności trybów kodowania międzyobrazowego w zaawansowanym koderze wizyjnym*”, Krajowa Konferencja Radiokomunikacji, Radiofonii i Telewizji, KKRRiT, 7-9 czerwca 2006, str. 219-222, Poznań (2006).

Publikacje zgłoszone, oczekujące na recenzje:

10. T. Grajek, M. Domański: “*New Model of MPEG-4 AVC/H.264 Video Encoders*”, The International Conference on Image Processing, ICIP, 26-29 września 2010, Hong Kong (2010).

4.2. Oryginalne osiągnięcia autora nie ujęte w rozprawie, a zastosowane w praktyce

1. *Koder skalowanych strumieni wizyjnych zgodnych z MPEG-4 AVC/H.264 - optymalizacja referencyjnego oprogramowania pozwalająca na kodowanie skalowalnych strumieni wizyjnych wysokich rozdzielczości.*
Praca zespołowa pod kierunkiem prof. dr. hab. inż. M. Domańskiego.

Udział własny:

- Badania nad statystykami wyboru trybów kodowania makrobloków.

- Opracowanie algorytmu wyboru trybów kodowania makrobloków w skalowalnym koderze wizyjnym.
- Autorska implementacja rozwiązań pozwalających na poprawę szybkości kodowania skalowalnych strumieni wizyjnych na komputerach osobistych.

Wyniki prac zostały wykorzystane w oprogramowaniu kodera zakupionego przez Mitsubishi Electric Information Technology Centre Europe, UK.

2. *Dekoder skalowanych strumieni wizyjnych zgodnych z MPEG-4 AVC/H.264* - optymalizacja referencyjnego oprogramowania pozwalająca na dekodowanie skalowalnych strumieni wizyjnych wysokich rozdzielczości.
Praca zespołowa pod kierunkiem prof. dr. hab. inż. M. Domańskiego.

Udział własny:

- Opracowanie i implementacja autorskich rozwiązań pozwalających na znaczną poprawę szybkości dekodowania skalowalnych strumieni wizyjnych na komputerach osobistych.

Opracowany dekoderek został zakupiony przez Mitsubishi Electric Information Technology Centre Europe, UK.

3. *Transkoder sygnałów wizyjnych standardu MPEG-2 przetwarzanych do formatu MPEG-4 AVC/H.264* - autorska implementacja transkodera heterogenicznego.
Praca zespołowa pod kierunkiem prof. dr. hab. inż. M. Domańskiego.

Udział własny:

- Badania nad sterowaniem przepływnością w transkoderze wizyjnym.
- Prace nad integracją dekodera MPEG-2 i kodera MPEG-4 AVC/H.264.

Wyniki prac badawczych zostały wykorzystane w transkoderze zakupionym przez Advanced Digital Broadcast Polska.

4. *Koder wizyjny MPEG-4 AVC/H.264 PP* - autorska implementacja i optymalizacja kodera wizyjnego standardu MPEG-4 AVC/H.264.
Praca zespołowa pod kierunkiem prof. dr. hab. inż. M. Domańskiego.

Udział własny:

- Prace badawcze nad algorytmami sterowania przepływnością kodera.
- Opracowanie autorskiego algorytmu sterowania przepływnością kodera w trybie stałej przepływności bitowej.
- Autorska implementacja modułów kodera hybrydowego standardu H.264/AVC w języku ANSI C oraz assembler realizujących kodowanie binarne.
- Oryginalna implementacja modułów kodera hybrydowego standardu H.264/AVC w języku ANSI C realizujących formater strumienia binarnego.
- Optymalizacja wydajnościowa powyższych modułów kodera.

Koder zakupiony i wdrożony przez Advanced Digital Broadcast Polska.

5. *Kodek sygnałów wizyjnych* – autorska implementacja i optymalizacja kodeka wizyjnego zintegrowanego w systemie telewizji interaktywnej iTVP.
Praca zespołowa pod kierunkiem prof. dr. hab. inż. M. Domańskiego.

Udział własny:

- Opracowanie oryginalnych modułów realizujących proces kodowania i dekodowania binarnego.
- Opracowanie wydajnych modułów realizujących formatowanie i parsowanie strumienia binarnego w autorskim kodeku czasu rzeczywistego.

Kodek zakupiony przez Telewizję Polską S.A.