

Politechnika Poznańska
Wydział Elektroniki i Telekomunikacji
Katedra Telekomunikacji Multimedialnej i Mikroelektroniki

Autoreferat rozprawy doktorskiej

**Transkodowanie strumieni wizyjnych
wykorzystujące zaawansowane
techniki kompresji**

Jarosław Marek

Promotor: prof. dr hab. inż. Marek Domański

Poznań, 2009

Spis treści autoreferatu*

1	Wprowadzenie	5
1.1.	Transkodowanie	5
1.2.	Cele i teza pracy	6
1.3.	Przegląd pracy	7
2	Oryginalne wyniki pracy:	
	Analiza transkodera kaskadowego	8
2.1.	Wprowadzenie	8
2.2.	Metodologia badań	8
2.3.	Najważniejsze wnioski	10
3	Zasadnicze oryginalne wyniki pracy:	
	Algorytmy transkodowania homogenicznego	12
3.1.	Wprowadzenie	12
3.2.	Algorytmy transkodowania	12
3.2.1.	Oryginalne strategie redukcji prędkości bitowej strumienia wizyjnego	12
3.2.2.	Modyfikacja próbek transformaty niewpływająca na przewidywane wartości próbek sąsiednich bloków	14
3.2.3.	Ograniczanie utraty subiektywnej jakości sekwencji wizyjnej	17
3.3.	Badania zaproponowanych technik transkodowania z wykorzystaniem obiektywnej miary jakości	17
3.4.	Badania zaproponowanych technik transkodowania z wykorzystaniem subiektywnej metody oceny jakości	20
3.5.	Złożoność zaproponowanych rozwiązań	21
4	Podsumowanie	22
4.1.	Wnioski	22
4.2.	Oryginalne osiągnięcia rozprawy	23
5	Dorobek naukowy autora	25
5.1.	Publikacje wyników rozprawy	25
5.2.	Uzyskane przez autora i nie ujęte w rozprawie oryginalne wyniki naukowe	25

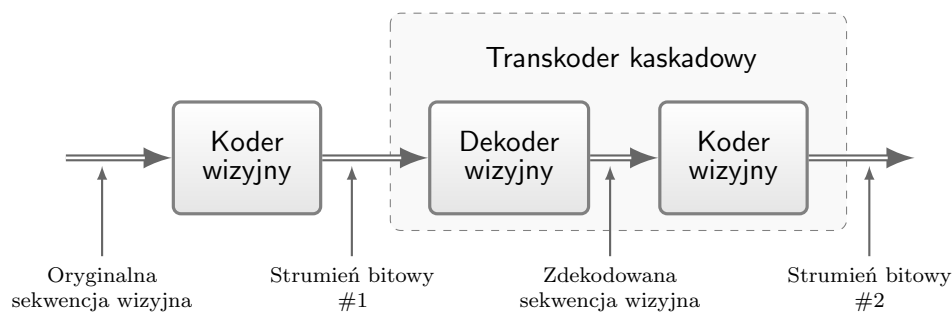
* Układ rozdziałów autoreferatu nie odpowiada układowi rozdziałów rozprawy.

1 Wprowadzenie

1.1. Transkodowanie

Adaptacja skompresowanego strumienia wizyjnego do możliwości sieci transmisyjnej oraz terminali abonenckich może być realizowana na dwa sposoby. Pierwszy polega na wykorzystaniu skalowalnych strumieni bitowych. Rozwiązanie to dotychczas nie upowszechniło się. Inne podejście, mające obecnie duże znaczenie praktyczne, polega na wykorzystaniu efektywnych narzędzi pozwalających na zmianę parametrów zakodowanego strumienia wizyjnego. Proces taki nazywa się *transkodowaniem* i jest realizowany przez *transkoder*. Jeżeli strumienie bitowe wejściowy i wyjściowy zgodne są z takim samym standardem/technologią kodowania to transkodowanie nazywane jest *homogenicznym*. Bardzo duże znaczenie praktyczne ma transkodowanie homogeniczne służące zmniejszeniu szybkości bitowej strumienia danych. Rozprawa dotyczy wyłącznie takiego transkodowania.

W najprostszej postaci transkoder jest kaskadowym zestawieniem dekodera i kodera wizyjnego, jak to zostało przedstawione na rysunku 1.1. Takie rozwiązanie jest skrótowo nazywane *transkoderem kaskadowym*.



Rysunek 1.1: Transkoder kaskadowy

Ze względu na istotne znaczenie praktyczne transkodowania sekwencji wizyjnych, zainteresowanie tą tematyką wykazują duże firmy. Przykładem mogą być przedsiębiorstwa projektujące i produkujące scalone układy graficzne. Firma NVIDIA w roku 2007 zaprezentowała architekturę przetwarzania równoległego danych o nazwie CUDA (ang. *Compute Unified Device Architecture*), wykorzystującą układy scalone tego producenta. Jednym z pierwszych publicznie pokazanych rozwiązań wykorzystujących tę technologię był właśnie transkoder sekwencji wizyjnych*. Również firma ATI przedstawiła implementację transkodera wykorzystującą produkowane scalone układy graficzne†.

Problem transkodowania jest również dostrzegany przez polskie firmy. Autor rozprawy brał udział w projekcie badawczo-rozwojowym realizowanym przez Politechnikę Poznańską dla dużej firmy zajmującej się projektowaniem i wytwarzaniem dekodów telewizji cyfrowej.

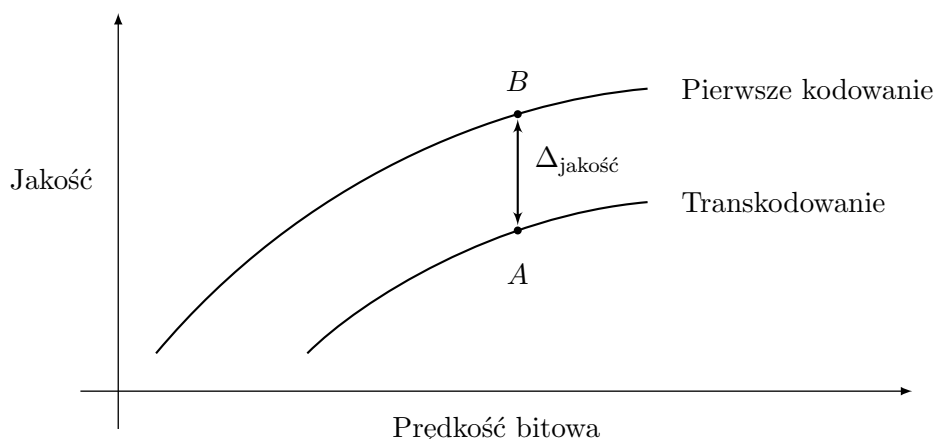
*Zaprezentowany na konferencji NVISION 2008, San Jose, CA.

†Zaprezentowana na konferencji SIGGRAPH 2008, Los Angeles, CA.

Celem projektu było zaprojektowanie i implementacja transkodera, który wykorzystywałby zaawansowane techniki kompresji.

Wykorzystanie transkodera do zmniejszenia prędkości strumienia bitowego prowadzi zwykle do pogorszenia relacji pomiędzy prędkością strumienia a jakością zdekodowanego sygnału. Oznacza to, że dla takiej samej docelowej prędkości bitowej wyższą jakość uzyskuje się w przypadku kodowania oryginalnej sekwencji, niż transkodując strumień bitowy o większej początkowej prędkości. Analogicznie, taką samą jakość zdekodowanych obrazów możemy uzyskać przy mniejszym strumieniu bitowym, gdy kodowaniu poddawana jest oryginalna sekwencja niż redukując inny strumień.

Na rysunku 1.2 zilustrowano różnicę jakości obrazu przy takiej samej prędkości bitowej dwóch strumieni wizyjnych ($\Delta_{\text{jakość}}$). Jeden ze strumieni uzyskany jest w wyniku kodowania materiału źródłowego, a drugi został otrzymany w wyniku transkodowania. W rozprawie takie różnice jakości nazywane są *stratami jakości na skutek transkodowania*, bądź w skrócie *stratami transkodowania*.



Rysunek 1.2: Strata jakości ($\Delta_{\text{jakość}}$) obrazu spowodowane transkodowaniem sekwencji wizyjnej

Ze względu na fakt, że transkodowanie kaskadowe wymaga znacznych nakładów obliczeniowych, dotychczasowe prace badawcze koncentrują się na zmniejszeniu złożoności tego procesu kosztem akceptowalnego zwiększenia strat transkodowania. Możliwe jest także odmienne zdefiniowanie problemu. Celem badań zaprezentowanych w pracy jest takie zmniejszanie prędkości bitowej strumieni wizyjnych skompresowanych z wykorzystaniem zaawansowanych technik kompresji, aby straty jakości były mniejsze niż w przypadku wykorzystania transkodera kaskadowego.

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

1.2. Cele i teza pracy

Głównym celem pracy jest *uzyskanie takich algorytmów transkodowania strumieni wizyjnych skompresowanych z wykorzystaniem najnowszych, zaawansowanych technik kompresji, które zapewniają mniejsze straty transkodowania niż transkoder kaskadowy.*

Zaawansowane techniki kompresji zdefiniowane są m.in w standardach VC-1, AVS oraz MPEG-4 AVC/H.264. Ponieważ ostatni z wymienionych standardów jest najbardziej rozpowszechniony oraz publicznie dostępna jest jego referencyjna implementacja, został on wybrany do badań.

Dotychczas opisane w literaturze prace dotyczące transkodowania wykorzystującego zaawansowane techniki kompresji skupione były w głównej mierze na problemie zmniejszenia złożoności obliczeniowej. W odróżnieniu od tych prac, celem rozprawy jest *uzyskanie lepszej jakości sekwencji wynikowej niż to jest możliwe w przypadku wykorzystania transkodera kaskadowego*.

Dodatkowym celem jest *uzyskanie transkodera o złożoności mniejszej od transkodera kaskadowego*.

Teza pracy jest następująca:

Możliwe jest stworzenie transkodera sekwencji wizyjnych wykorzystującego zaawansowane techniki kompresji, który zapewni wyższą jakość transkodowanej sekwencji niż transkoder kaskadowy. Możliwe jest jednocześnie zmniejszenie nakładu obliczeniowego w stosunku do transkodera kaskadowego.

W celu sprawdzenia zaproponowanych rozwiązań założono stworzenie odpowiednich implementacji. Przyjęto, że podobnie jak w literaturze, wyniki badań porównywane będą z wynikami uzyskanymi dla przypadku transkodowania kaskadowego wykorzystującego oprogramowanie referencyjne MPEG-4 AVC/H.264.

1.3. Przegląd pracy

Praca dotyczy transkodowania strumieni binarnych reprezentujących obraz ruchomy zakodowany zgodnie ze standardem MPEG-4 AVC/H.264. W szczególności rozprawa koncentruje się na transkodowaniu mającym na celu redukcję prędkości transmisji.

Rozdział 1 wprowadza w tematykę rozprawy, zawiera cele i tezę pracy, a także jej przegląd.

Rozdział 2 wyjaśnia zagadnienia hybrydowego kodowania sekwencji wizyjnych oraz prezentuje istotne zaawansowane techniki kompresji zdefiniowane w standardzie MPEG-4 AVC/H.264. Ponadto znajduje się w nim przegląd literatury dotyczącej transkodowania skompresowanych strumieni wizyjnych. Oprócz tego przedstawiono w nim zagadnienie pomiaru jakości sekwencji wizyjnych.

Rozdział 3 zawiera prezentację wyników badań strat jakości transkodera kaskadowego, które zostały przeprowadzone przez autora tej pracy.

Rozdział 4 służy przedstawieniu technik pozwalających na redukcję prędkości bitowej strumienia wizyjnego. Zostały w nim przedstawione i udowodnione przez autora warunki modyfikacji próbek transformaty w strumieniu bitowym zgodnym ze standardem MPEG-4 AVC/H.264, których spełnienie zapewnia, że nie ulegną zmianie przewidywane wartości próbek sąsiednich bloków. Ponadto w rozdziale znajdują się wyniki przeprowadzonych przez doktoranta badań wpływu modyfikacji próbek transformaty na jakość sekwencji wizyjnej.

Rozdział 5 przedstawia zaproponowane przez autora algorytmy transkodowania strumieni bitowych wykorzystujących zaawansowane techniki kompresji. Ponadto prezentuje wyniki badań eksperymentalnych przeprowadzonych z wykorzystaniem przygotowanych przez autora tej pracy modeli zaproponowanych algorytmów transkodowania. Wyniki porównano z rezultatami uzyskanymi dla kodowania oryginalnych sekwencji wizyjnych oraz transkodowania kaskadowego opartego na wykorzystaniu referencyjnego oprogramowania standardu MPEG-4 AVC/H.264.

Rozdział 6 podsumowuje pracę. Zaprezentowane w nim są najważniejsze wnioski zaprezentowane w tej pracy oraz oryginalne osiągnięcia. Wskazany jest również kierunek dalszych badań.

2 Oryginalne wyniki pracy: Analiza transkodera kaskadowego

2.1. Wprowadzenie

Kwantyzacja jest procesem stratnym, a więc nieodwracalnym. Ponowna kwantyzacja już raz skwantowanego sygnału wprowadza dodatkowe zniekształcenia. Z tego powodu transkodowana sekwencja wizyjna, nawet przy założeniu takiej samej prędkości bitowej strumienia binarnego, charakteryzuje się na ogół niższą jakością obrazu niż sekwencja wizyjna otrzymana w wyniku procesu kodowania sekwencji oryginalnej. Fragmentaryczne wyniki dotyczące strat jakości spowodowanych wykorzystaniem transkodera kaskadowego odnaleźć można w literaturze. Dotyczą one jednakże strumieni wizyjnych uzyskanych dzięki wykorzystaniu klasycznych technik kompresji. W literaturze brak kompleksowych wyników badań nad stratą jakości wynikającą z transkodowania strumieni wizyjnych powstałych w wyniku wykorzystania zaawansowanych technik kompresji. Stosowanie nowych narzędzi kompresji ma wpływ na zniekształcenia wynikające z ponownej kwantyzacji sygnału. W pracy zaprezentowane zostały własne, oryginalne wyniki badań pokazujące jaki wpływ na jakość zdekodowanej sekwencji ma zastosowanie transkodera kaskadowego.

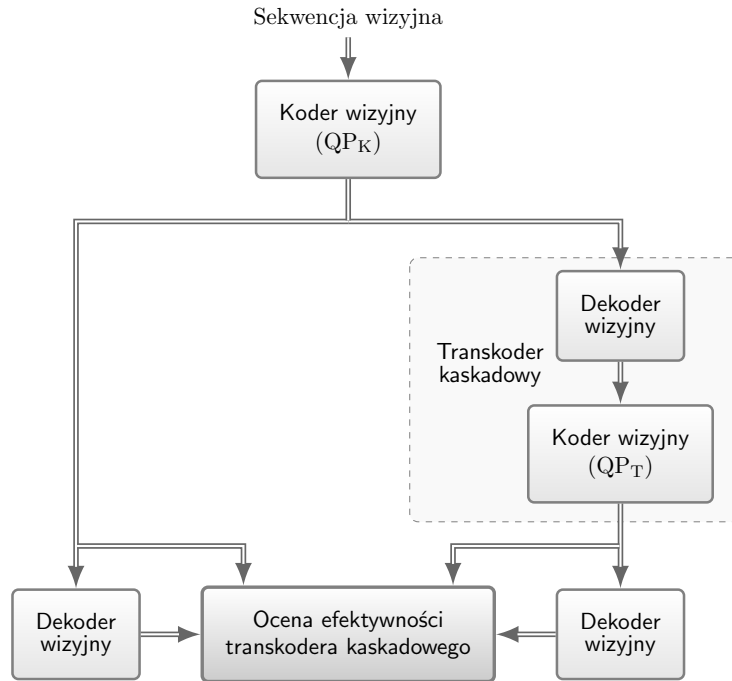
2.2. Metodologia badań

Do badań użyte zostały cztery powszechnie wykorzystywane sekwencje testowe: CITY, CREW HARBOUR i SOCCER. Sposób przeprowadzania testów zilustrowany został na rysunku 2.1.

Badania przeprowadzono używając referencyjnego oprogramowania standardu MPEG-4 AVC/H.264. Eksperymenty zostały zrealizowane z wykorzystaniem stałych wartości parametrów kwantyzacji w obrazach tych samych typów. Takie podejście ma na celu uniezależnienie uzyskiwanych wyników badań od wpływu mechanizmu sterowania prędkością bitową strumienia wizyjnego.

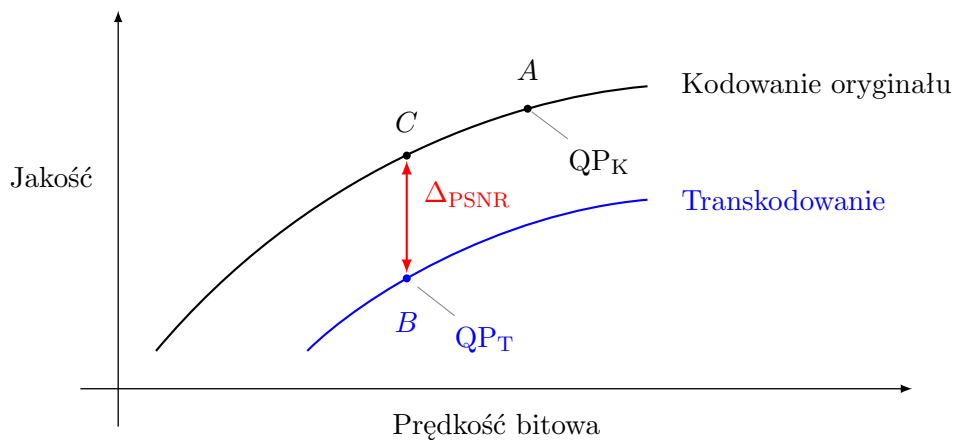
Dla uzyskania wysokiej jakości obrazu odzyskanego ze strumienia w narzędziach włączone były mechanizmy optymalizacji „prędkość bitowa-zniekształcenie” (RD).

Konfiguracje koderów wizyjnych: wykorzystywanego do kodowania oryginalnych sekwencji oraz użytego w transkoderze były zawsze zgodne. Wyjątkiem była jedynie wartość parametru kwantyzacji (QP). W pierwszej kolejności oryginalne sekwencje testowe zostały poddane kodowaniu z zestawem wartości indeksu QP_K . Otrzymane w ten sposób strumienie binarne były dekodowane i w rezultacie możliwe stało się wykreślenie krzywych „prędkość bitowa-jakość” dla kodowania oryginalnych sekwencji wizyjnych. Aby zbadać, jak zmienia się jakość sekwencji w wyniku transkodowania z wykorzystaniem transkodera kaskadowego, ze zbioru otrzymanych wcześniej strumieni binarnych wybrano te, dla których indeks kwantyzatora wynosił odpowiednio $QP_K \in \{25, 28, 31, 34, 37, 40\}$. Następnie wybrane strumienie poddane zostały transkodowaniu z nowymi indeksami



Rysunek 2.1: Badania efektywności transkodera kaskadowego

kwantyzatora: $QP_T = QP_K + n$, $n = 0, 1, 2, \dots$. Każdy z tak otrzymanych strumieni binarnych był dekodowany. Dzięki temu możliwe stało się wykreślenie odpowiednich krzywych „prędkość bitowa-jakość” dla transkodowania.



Rysunek 2.2: Strata jakości (Δ_{PSNR}) spowodowana transkodowaniem kaskadowym

W pracy przeanalizowane zostały straty jakości obrazu spowodowane transkodowaniem. Zostały one zilustrowane na rysunku 2.2. Kodując sekwencję oryginalną z wartością indeksu QP_K otrzymujemy punkt A na krzywej pierwszego kodowania. W wyniku transkodowania kaskadowego otrzymanego strumienia bitowego z nową wartością parametru kwantyzacji QP_T otrzymujemy nowy strumień binarny, którego parametry umiejscawiają go w punkcie B na krzywej transkodowania. Znając prędkość bitową tak otrzymanego strumienia możemy na krzywej „jakość-prędkość bitowa” kodowania oryginalnej sekwencji wskazać punkt C odpowiadający takiej samej prędkości bitowej. Różnica odległości punktów B i C odpowiada wielkości strat jakości powstałych na skutek transkodowania

kaskadowego Δ_{PSNR} . Takie straty można przedstawić na przykład w funkcji różnicy wartości parametrów kwantyzacji w strumieniu wyjściowym i wejściowym $\Delta_{\text{QP}} = \text{QP}_T - \text{QP}_K$. Te straty można także przedstawić w funkcji stosunku prędkości bitowej po transkodowaniu (punkt B) do prędkości bitowej przed transkodowaniem (punkt A).

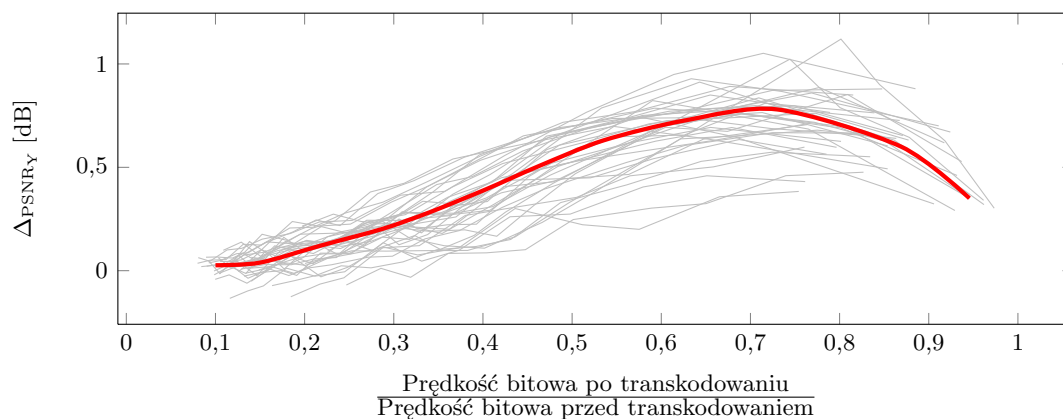
Badania strat transkodowania strumieni bitowych, w których występują obrazy I, P oraz B, przeprowadzone zostały dla wielu kombinacji parametrów kodowania wyszczególnionych poniżej:

- Filtr deblokujący: włączony, wyłączony;
- Struktura grupy obrazów (GOP): I-BBB-P-BBB-P-BBB-P-BBB (3P3B), I-BB-P-BB-P-BB-P-BB-P-BB (4P2B), I-BB-P-BB-P-BB-P-BB-P-BB-P-BB (5P2B);
- Relacja pomiędzy wartościami parametrów kwantyzacji w obrazach typu I (QP^I), P (QP^P) oraz B (QP^B): $\text{QP}^I = \text{QP}^P = \text{QP}^B$, $\text{QP}^I = \text{QP}^P = \text{QP}^B - 2$, $\text{QP}^I = \text{QP}^P = \text{QP}^B - 6$, $\text{QP}^I + 2 = \text{QP}^P = \text{QP}^B - 2$.

Straty jakości obrazu powstające na skutek transkodowania kaskadowego strumieni wizyjnych składających się wyłącznie z obrazów typu I przebadane zostały dla przypadków transkodowania strumieni uzyskanych z wykorzystaniem filtra usuwającego efekt blokowy, a także bez wykorzystania tego narzędzia.

2.3. Najważniejsze wnioski

Najważniejszym wnioskiem płynącym z badań przedstawionych w rozprawie jest stwierdzenie, że krzywa przedstawiająca zależność strat jakości obrazu spowodowanych transkodowaniem od redukcji strumienia bitowego posiada pewien typowy kształt. Badania wykazały, że jest on niezależny od treści sekwencji i parametrów kodowania. Rysunek 2.3 przedstawia uśrednioną zależność straty jakości obrazu spowodowanej transkodowaniem na tle uzyskanych wyników badań.



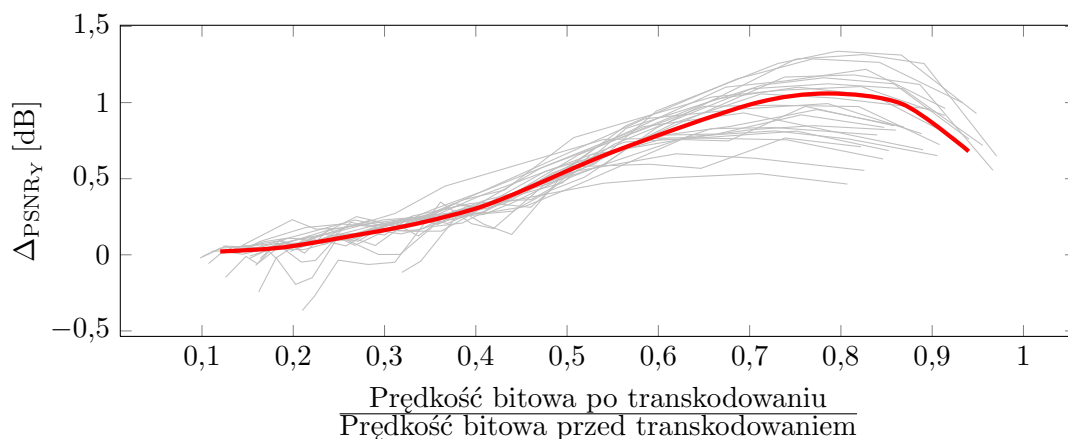
Rysunek 2.3: Uśredniona krzywa strat transkodowania kaskadowego strumieni bitowych zgodnych ze standardem MPEG-4 AVC/H.264, w których występują obrazy typu I, P oraz B, na tle wyników badań eksperymentalnych

Dla wszystkich przebadanych przypadków stwierdzić można, że największe straty jakości powstające na skutek transkodowania kaskadowego występują, gdy strumień bitowy zmniejszany jest o około 30%. Po przekroczeniu tej wartości redukcji, dalsze jej zwiększanie prowadzi do systematycznego zmniejszania się strat transkodowania. Są to właściwości analogiczne do znanych z literatury właściwości transkodowania strumieni bitowych uzyskanych dzięki wykorzystaniu klasycznych technik kompresji, takich jak MPEG-2.

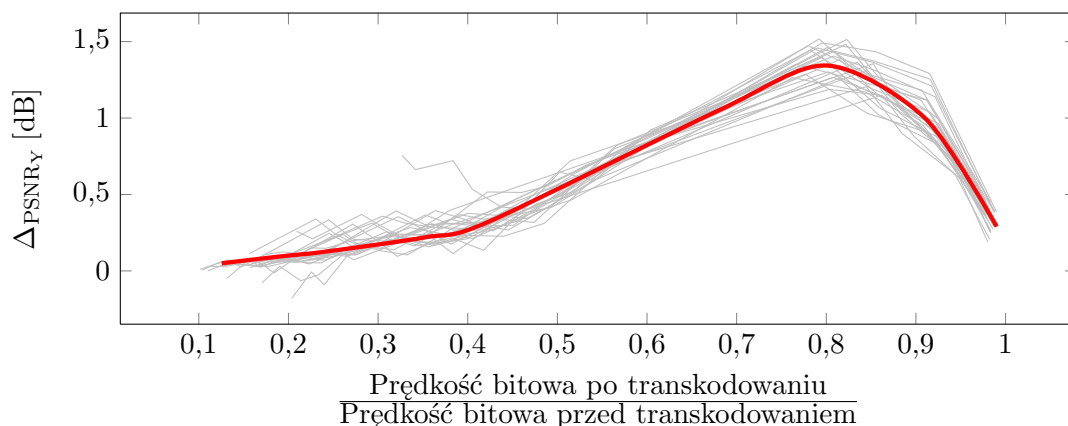
Eksperymenty wykazały niewielki wpływ struktury grupy obrazów i relacji pomiędzy wartościami parametrów kwantyzacji w obrazach I, P, B na wielkość strat transkodowania.

Większy wpływ na wielkość tych strat ma wykorzystanie filtra usuwającego efekt blokowy. Ten wyraźny wpływ zaobserwowano tylko wtedy, gdy różnica pomiędzy parametrami kwantyzacji w strumieniach wejściowym i wyjściowym nie przekracza wartości 3.

Powyższe wyniki uzyskane zostały dla typowych strumieni wizyjnych złożonych z obrazów typów I, P oraz B. W rozprawie przedstawione zostały również rezultaty uzyskane dla strumieni wizyjnych zbudowanych wyłącznie z obrazów kodowanych wewnątrzobrazowo. Wykazano, że w takim przypadku największe straty transkodowania przypadają na około 20 % redukcję strumienia bitowego. Wyniki zaprezentowane w rozprawie wskazują, że wykorzystywanie mechanizmu filtracji usuwającej efekt blokowy na etapie kodowania i transkodowania sygnału wizyjnego ma wpływ na przebieg strat transkodowania kaskadowego w funkcji redukcji prędkości bitowej. Zostało to przedstawione na rysunku 2.4.



(a) Kodowanie z wykorzystaniem filtra usuwającego efekt blokowy



(b) Kodowanie bez wykorzystania filtra usuwającego efekt blokowy

Rysunek 2.4: Uśrednione krzywe strat transkodowania kaskadowego strumieni bitowych złożonych wyłącznie z obrazów typu I. W tle umieszczone zostały wyniki badań eksperymentalnych

Przedstawione w pracy wyniki badań wskazują, że transkodowanie kaskadowe, podczas którego wykorzystywany jest taki sam kwantyzator jak w procesie kodowania, prowadzi do utraty jakości skompresowanej sekwencji wizyjnej. Strata taka może osiągać znaczące wartości i przekraczać nawet 0,8 dB.

Badania pokazały, że niezależnie od treści sekwencji wizyjnej i parametrów kodowania, **największe straty jakości na skutek transkodowania kaskadowego występują przy 10–40 % redukcji strumienia bitowego.**

3 Zasadnicze oryginalne wyniki pracy:

Algorytmy transkodowania homogenicznego

3.1. Wprowadzenie

Na podstawie wyników badań zaprezentowanych w rozprawie, a omówionych w poprzednim rozdziale, stwierdzić należy, że największe straty jakości spowodowane transkodowaniem powstają w wyniku stosunkowo niewielkich redukcji strumienia bitowego. Silniejsza redukcja prędkości bitowej wiąże się z poszerzeniem przedziałów kwantyzacji. W efekcie zniekształcenia obrazu spowodowane silną kwantyzacją stają się relatywnie duże w stosunku do zniekształceń spowodowanych rekwantyzacją. Wynikiem są niewielkie straty jakości spowodowane transkodowaniem. Celowym jest więc opracowanie mechanizmu transkodowania, który zapewniłby mniejsze straty jakości niż transkoder kaskadowy dla niewielkich redukcji strumienia bitowego.

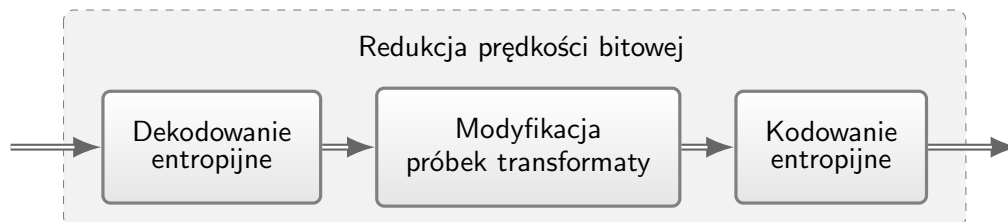
Autor przebadiał wiele scenariuszy transkodowania homogenicznego. Ze względu na obszerność eksperymentów, w rozprawie znalazła się jedynie ich reprezentatywna część. Na podstawie obserwacji otrzymanych wyników badań oraz dogłębnej analizy standardu MPEG-4 AVC/H.264 i rozwiązań prezentowanych w literaturze, zaproponowane zostały nowe, oryginalne algorytmy transkodowania homogenicznego z ograniczonymi stratami jakości.

3.2. Algorytmy transkodowania

3.2.1. Oryginalne strategie redukcji prędkości bitowej strumienia wizyjnego

Autor proponuje by transkodowanie przebiegało przy otwartej pętli sprzężenia zwrotnego. W pracy proponuje się, szczegółowo uzasadniając, aby w procesie transkodowania zachować typy obrazów, tryby kodowania makrobloków, submakrobloków i bloków, wektory ruchu wraz z indeksami odniesienia. W zaproponowanym algorytmie (rysunek 3.1) modyfikacji podlegają wyłącznie skwantowane wartości próbek transformaty. W pierwszej kolejności dekodery entropijny dekoduje strumień bitowy. W następnym kroku modyfikowane są wartości próbek transformaty. Ostatnim etapem jest sformowanie nowego strumienia bitowego.

Jednym z powodów wysokiej efektywności kompresji uzyskiwanej dla strumieni wizyjnych zgodnych ze standardem MPEG-4 AVC/H.264 jest wykorzystywanie kodowania kontekstowego i predykcji. Stosowanie mechanizmu predykcji powoduje, że modyfikacja niewielkiego fragmentu obrazu może mieć znaczący wpływ na jego pozostałą część, a nawet



Rysunek 3.1: Schemat zaproponowanego transkodera homogenicznego

na kolejne obrazy w sekwencji wizyjnej. W związku z powyższym, pojawiający się w wyniku modyfikacji próbek transformaty błąd rekwantyzacji może się propagować i kumulować zarówno wewnątrz obrazu, jak i pomiędzy obrazami sekwencji. Zjawisko to prowadzi może do znacznej degradacji jakości obrazu. Dlatego też bardzo ważne jest by mechanizmy transkodowania uwzględniały fakt propagacji i kumulacji błędów.

W rozprawie autor zaprezentował **oryginalne strategie redukcji prędkości bitowej strumienia wizyjnego**. Powstały one na podstawie wyników przeprowadzonych rozległych badań. Ich stosowanie pozwala w znaczącym stopniu ograniczyć zjawisko propagacji błędów rekwantyzacji pomiędzy obrazami.

Pierwsza z zaproponowanych strategii transkodowania przeznaczona jest do redukcji typowych strumieni bitowych, w których występują obrazy typów I, P oraz B, a w obrazach kodowanych międzyobrazowo dopuszcza się wykorzystanie technik kodowania wewnątrzobrazowego. Strategia ta zakłada, że do momentu, w którym zostanie osiągnięta zamierzona redukcja strumienia bitowego, wykonywane będą kolejno następujące kroki:

1. Modyfikacja obrazów typu B:
 - usuwanie próbek transformaty o amplitudzie równej 1 zarówno w makroblokach kodowanych międzyobrazowo jak i wewnątrzobrazowo,
 - zwiększenie o 6 wartość indeksu QP^B w makroblokach kodowanych międzyobrazowo, które nie stanowią odniesienia do kodowania makrobloków typu Intra,
 - usuwanie próbek transformaty o wartości bezwzględnej równej 1 w makroblokach kodowanych międzyobrazowo, w których zmieniono wartość parametru kwantyzacji, a także modyfikacja wartości próbek transformaty o wartość nie większą niż 2 w makroblokach kodowanych wewnątrzobrazowo;
2. W obrazach typu P, rozpoczynając od ostatniego w strukturze GOP, kolejno w kierunku jej początku:
 - usuwanie próbek transformaty o wartości bezwzględnej równej 1 w makroblokach kodowanych międzyobrazowo jak i wewnątrzobrazowo,
 - zwiększenie o 6 wartość indeksu QP^P w makroblokach kodowanych międzyobrazowo, które nie stanowią odniesienia do kodowania makrobloków typu Intra,
 - usuwanie próbek transformaty o wartości bezwzględnej równej 1 w makroblokach kodowanych międzyobrazowo, w których zmieniono wartość parametru kwantyzacji, a także modyfikacja wartości próbek transformaty o wartość nie większą niż 2 w makroblokach kodowanych wewnątrzobrazowo;
3. Modyfikacja próbek transformaty w obrazach typu I o wartość nie większą niż 2.

Przy wykorzystaniu przedstawionej powyżej strategii redukcja prędkości bitowej możliwa jest w pewnym ograniczonym zakresie. W wyniku przeprowadzonej analizy autor rozprawy

stwierdził, że zakres ten może być większy, gdy w obrazach typów P oraz B nie występują makrobloki typu Intra. W rezultacie przeprowadzonych badań zaproponowana została odmienna procedura redukcji strumieni bitowych przeznaczona do transkodowania takich specyficznych strumieni. Znaleźć ją można w pełnym tekście rozprawy.

Przedstawione w pracy procedury transkodowania opisują kolejność modyfikacji obrazów w pojedynczej grupie obrazów. Określają także amplitudy próbek mogących podlegać modyfikacji. Jednak próbek o takiej samej amplitudzie może być wiele, a wybór właściwej kolejności i sposobu modyfikacji może mieć bardzo duży wpływ na wartość relacji pomiędzy prędkością bitową strumienia i jakością zdekodowanego obrazu.

3.2.2. Modyfikacja próbek transformaty niewpływająca na przewidywane wartości próbek sąsiednich bloków

Należy zauważyć, że szczególnie w przypadku kodowania wewnątrzobrazowego zjawisko kumulacji błędów kwantyzacji może prowadzić do znaczącej degradacji jakości obrazu ze względu na istniejące bardzo długie drogi propagacji błędów. Możliwe są jednak takie modyfikacje próbek transformaty, które zapewniają, że powstający w ich wyniku błąd rekwantyzacji nie będzie propagował się wewnątrz obrazu. W pracy autor **sformułował i udowodnił** twierdzenia dotyczące takich modyfikacji próbek transformaty. Twierdzenia te wskazują jaka zmiana wartości próbek transformaty nie wpływa na wartości w bloku o rozmiarze 4×4 :

- punktów znajdujących się w jego ostatniej kolumnie,
- punktów należących do jego ostatniego wiersza,
- punktu leżącego w ostatniej kolumnie jego ostatniego wiersza.

W rozprawie autor wskazał, które z trzech powyższych przypadków powinny być uwzględniane w zależności od trybów kodowania sąsiednich bloków.

Poniżej przedstawione zostaną przykładowe rozważania dotyczące modyfikacji próbek transformaty niewpływającej na wartości punktów znajdujących się w ostatniej kolumnie bloku. W pełnym tekście rozprawy znaleźć można analogiczne wywody dla pozostałych przypadków.

Przekształcenie odwrotne definiowane przez standard MPEG-4 AVC/H.264 dane jest równaniem:

$$\mathbf{X} = \mathbf{C}_i^T (\mathbf{Y} \otimes \mathbf{K}_i) \mathbf{C}_i, \quad (3.1)$$

gdzie:

$$\mathbf{C}_i = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1/2 & -1/2 & -1 \\ 1 & -1 & -1 & 1 \\ 1/2 & -1 & 1 & -1/2 \end{bmatrix}, \quad (3.2)$$

$$\mathbf{K}_i = \begin{bmatrix} a^2 & ab & a^2 & ab \\ ab & b^2 & ab & b^2 \\ a^2 & ab & a^2 & ab \\ ab & b^2 & ab & b^2 \end{bmatrix}. \quad (3.3)$$

Macierz \mathbf{Y} otrzymywana jest w wyniku dekwantyzacji próbek transformaty. Proces ten jest połączony z mnożeniem Hadamarda (poelementowym) przez macierz skalującą \mathbf{K}_i . Zdefiniujmy więc nową macierz:

$$\mathbf{T} = \mathbf{Y} \otimes \mathbf{K}_i, \quad (3.4)$$

o elementach:

$$\mathbf{T} = [t_{ij}]. \quad (3.5)$$

Równanie (3.1) przyjmie wtedy następującą postać:

$$\mathbf{X} = \mathbf{C}_i^T \cdot \mathbf{T} \cdot \mathbf{C}_i, \quad (3.6)$$

gdzie:

$$\mathbf{X} = [x_{ij}]. \quad (3.7)$$

Macierz \mathbf{X} możemy zapisać też jako kombinację liniową funkcji bazowych:

$$\mathbf{X} = t_{00}\mathbf{A}_{00} + t_{01}\mathbf{A}_{01} + \dots + t_{33}\mathbf{A}_{33} = \sum_{ij} t_{ij}\mathbf{A}_{ij}, \quad (3.8)$$

gdzie \mathbf{A}_{ij} są macierzami bazowymi przekształcenia odwrotnego definiowanego przez standard MPEG-4 AVC/H.264. Zostały one zaprezentowane w tabeli 3.1. Zdefiniujemy dodatkowo pomocniczą macierz \mathbf{W} o postaci:

$$\mathbf{W} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \quad (3.9)$$

$i \setminus j$	0	1	2	3
0	$\begin{bmatrix} a^2 & a^2 & a^2 & a^2 \\ a^2 & a^2 & a^2 & a^2 \\ a^2 & a^2 & a^2 & a^2 \\ a^2 & a^2 & a^2 & a^2 \end{bmatrix}$	$\begin{bmatrix} ab & \frac{ab}{2} & -\frac{ab}{2} & -ab \\ ab & \frac{ab}{2} & -\frac{ab}{2} & -ab \\ ab & \frac{ab}{2} & -\frac{ab}{2} & -ab \\ ab & \frac{ab}{2} & -\frac{ab}{2} & -ab \end{bmatrix}$	$\begin{bmatrix} a^2 & -a^2 & -a^2 & a^2 \\ a^2 & -a^2 & -a^2 & a^2 \\ a^2 & -a^2 & -a^2 & a^2 \\ a^2 & -a^2 & -a^2 & a^2 \end{bmatrix}$	$\begin{bmatrix} \frac{ab}{2} & -ab & ab & -\frac{ab}{2} \\ \frac{ab}{2} & -ab & ab & -\frac{ab}{2} \\ \frac{ab}{2} & -ab & ab & -\frac{ab}{2} \\ \frac{ab}{2} & -ab & ab & -\frac{ab}{2} \end{bmatrix}$
1	$\begin{bmatrix} ab & ab & ab & ab \\ \frac{ab}{2} & \frac{ab}{2} & \frac{ab}{2} & \frac{ab}{2} \\ -\frac{ab}{2} & -\frac{ab}{2} & -\frac{ab}{2} & -\frac{ab}{2} \\ -ab & -ab & -ab & -ab \end{bmatrix}$	$\begin{bmatrix} b^2 & \frac{b^2}{2} & -\frac{b^2}{2} & -b^2 \\ \frac{b^2}{2} & \frac{b^2}{4} & -\frac{b^2}{4} & -\frac{b^2}{2} \\ -\frac{b^2}{2} & -\frac{b^2}{4} & \frac{b^2}{4} & \frac{b^2}{2} \\ -b^2 & -\frac{b^2}{2} & \frac{b^2}{2} & b^2 \end{bmatrix}$	$\begin{bmatrix} ab & -ab & -ab & ab \\ \frac{ab}{2} & -\frac{ab}{2} & -\frac{ab}{2} & \frac{ab}{2} \\ -\frac{ab}{2} & \frac{ab}{2} & \frac{ab}{2} & -\frac{ab}{2} \\ -ab & ab & ab & -ab \end{bmatrix}$	$\begin{bmatrix} \frac{b^2}{2} & -b^2 & b^2 & -\frac{b^2}{2} \\ \frac{b^2}{4} & -\frac{b^2}{2} & \frac{b^2}{2} & -\frac{b^2}{4} \\ -\frac{b^2}{4} & \frac{b^2}{2} & -\frac{b^2}{2} & \frac{b^2}{4} \\ -\frac{b^2}{2} & b^2 & -b^2 & \frac{b^2}{2} \end{bmatrix}$
2	$\begin{bmatrix} a^2 & a^2 & a^2 & a^2 \\ -a^2 & -a^2 & -a^2 & -a^2 \\ -a^2 & -a^2 & -a^2 & -a^2 \\ a^2 & a^2 & a^2 & a^2 \end{bmatrix}$	$\begin{bmatrix} ab & \frac{ab}{2} & -\frac{ab}{2} & -ab \\ -ab & -\frac{ab}{2} & \frac{ab}{2} & ab \\ -ab & -\frac{ab}{2} & \frac{ab}{2} & ab \\ ab & \frac{ab}{2} & -\frac{ab}{2} & -ab \end{bmatrix}$	$\begin{bmatrix} a^2 & -a^2 & -a^2 & a^2 \\ -a^2 & a^2 & a^2 & -a^2 \\ -a^2 & a^2 & a^2 & -a^2 \\ a^2 & -a^2 & -a^2 & a^2 \end{bmatrix}$	$\begin{bmatrix} \frac{ab}{2} & -ab & ab & -\frac{ab}{2} \\ -\frac{ab}{2} & ab & -ab & \frac{ab}{2} \\ -\frac{ab}{2} & ab & -ab & \frac{ab}{2} \\ \frac{ab}{2} & -ab & ab & -\frac{ab}{2} \end{bmatrix}$
3	$\begin{bmatrix} \frac{ab}{2} & \frac{ab}{2} & \frac{ab}{2} & \frac{ab}{2} \\ -ab & -ab & -ab & -ab \\ ab & ab & ab & ab \\ -\frac{ab}{2} & -\frac{ab}{2} & -\frac{ab}{2} & -\frac{ab}{2} \end{bmatrix}$	$\begin{bmatrix} \frac{b^2}{2} & \frac{b^2}{4} & -\frac{b^2}{4} & -\frac{b^2}{2} \\ -b^2 & -\frac{b^2}{2} & \frac{b^2}{2} & b^2 \\ b^2 & \frac{b^2}{2} & -\frac{b^2}{2} & -b^2 \\ -\frac{b^2}{2} & -\frac{b^2}{4} & \frac{b^2}{4} & \frac{b^2}{2} \end{bmatrix}$	$\begin{bmatrix} \frac{ab}{2} & -\frac{ab}{2} & -\frac{ab}{2} & \frac{ab}{2} \\ -ab & ab & ab & -ab \\ ab & -ab & -ab & ab \\ -\frac{ab}{2} & \frac{ab}{2} & \frac{ab}{2} & -\frac{ab}{2} \end{bmatrix}$	$\begin{bmatrix} \frac{b^2}{4} & -\frac{b^2}{2} & \frac{b^2}{2} & -\frac{b^2}{4} \\ -\frac{b^2}{2} & b^2 & -b^2 & \frac{b^2}{2} \\ \frac{b^2}{2} & -b^2 & b^2 & -\frac{b^2}{2} \\ -\frac{b^2}{4} & \frac{b^2}{2} & -\frac{b^2}{2} & \frac{b^2}{4} \end{bmatrix}$

Tabela 3.1: Macierze \mathbf{A}_{ij} tworzące bazę odwrotnego przekształcenia zdefiniowanego w standardzie MPEG-4 AVC/H.264, gdzie: $a = \frac{1}{2}$, $b = \sqrt{\frac{2}{5}}$

Wprowadźmy także dwie pomocnicze równości. Ostatnie kolumny macierzy bazowych odwrotnego przekształcenia zdefiniowanego w standardzie MPEG-4 AVC/H.264 są sobie równe, gdy:

$$\mathbf{A}_{i0}\mathbf{W} = \mathbf{A}_{i2}\mathbf{W}, \quad (3.10a)$$

$$\mathbf{A}_{i1}\mathbf{W} = 2\mathbf{A}_{i3}\mathbf{W}, \quad (3.10b)$$

dla $i = 0, 1, 2, 3$. Powyższe zależności wynikają z analizy tabeli 3.1.

W celu potwierdzenia prawdziwości równań (3.10) przeanalizujemy równość (3.10a) dla $i = 0$:

$$\mathbf{A}_{00}\mathbf{W} = \mathbf{A}_{02}\mathbf{W}. \quad (3.11)$$

Podstawiając odpowiednie wartości otrzymujemy:

$$\begin{bmatrix} a^2 & a^2 & a^2 & a^2 \\ a^2 & a^2 & a^2 & a^2 \\ a^2 & a^2 & a^2 & a^2 \\ a^2 & a^2 & a^2 & a^2 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} a^2 & -a^2 & -a^2 & a^2 \\ a^2 & -a^2 & -a^2 & a^2 \\ a^2 & -a^2 & -a^2 & a^2 \\ a^2 & -a^2 & -a^2 & a^2 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad (3.12)$$

$$\begin{bmatrix} 0 & 0 & 0 & a^2 \\ 0 & 0 & 0 & a^2 \\ 0 & 0 & 0 & a^2 \\ 0 & 0 & 0 & a^2 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & a^2 \\ 0 & 0 & 0 & a^2 \\ 0 & 0 & 0 & a^2 \\ 0 & 0 & 0 & a^2 \end{bmatrix}. \quad (3.13)$$

Dowodzi to prawdziwości równania (3.10a) dla $i = 0$. W taki sam sposób można udowodnić prawdziwość tego równania dla $i = 1, 2, 3$, oraz równania (3.10b) dla $i = 0, 1, 2, 3$.

Określmy teraz zależność ostatniej kolumny macierzy \mathbf{X} od wartości próbek t_{ij} transformy \mathbf{T} . W tym celu pomnożmy obie strony równania (3.8) prawostronnie przez macierz \mathbf{W} :

$$\mathbf{XW} = \left(\sum_{ij} t_{ij} \mathbf{A}_{ij} \right) \mathbf{W}. \quad (3.14)$$

Twierdzenie. Równanie (3.14) jest prawdziwe dla $i = 0, 1, 2, 3$, gdy:

$$\begin{cases} t_{i0} + t_{i2} = \text{const} \\ t_{i1} + \frac{1}{2}t_{i3} = \text{const} \end{cases}. \quad (3.15)$$

Dowiedźmy teraz, iż modyfikacja elementów macierzy \mathbf{T} zgodnie z powyższymi warunkami zapewni, że nie zmieniają się wartości w ostatniej kolumnie macierzy \mathbf{X} .

Dowód. Równanie (3.14) możemy zapisać w następujący sposób:

$$\mathbf{XW} = \left(\sum_i t_{i0} \mathbf{A}_{i0} + \sum_i t_{i1} \mathbf{A}_{i1} + \sum_i t_{i2} \mathbf{A}_{i2} + \sum_i t_{i3} \mathbf{A}_{i3} \right) \mathbf{W}. \quad (3.16)$$

Powyższe możemy zapisać jako:

$$\mathbf{XW} = \left(\sum_i t_{i0} \mathbf{A}_{i0} \mathbf{W} + \sum_i t_{i2} \mathbf{A}_{i2} \mathbf{W} \right) + \left(\sum_i t_{i1} \mathbf{A}_{i1} \mathbf{W} + \sum_i t_{i3} \mathbf{A}_{i3} \mathbf{W} \right). \quad (3.17)$$

Korzystając z równań (3.10) wykonujemy podstawienia za macierze \mathbf{A}_{i2} oraz \mathbf{A}_{i3} :

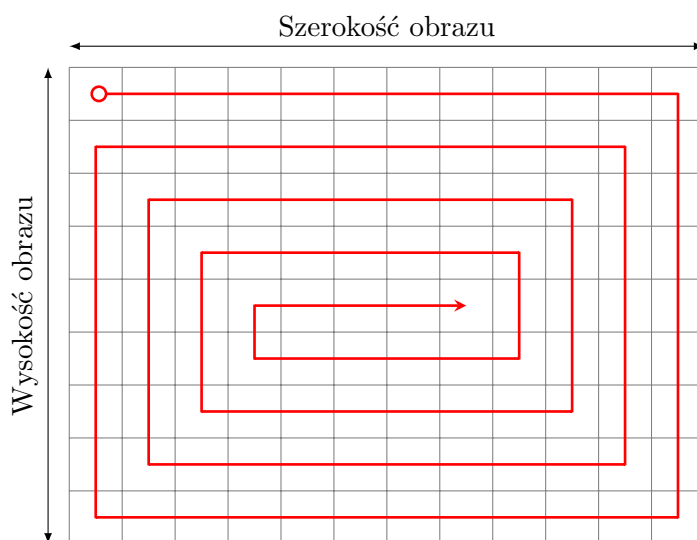
$$\mathbf{XW} = \left(\sum_i t_{i0} \mathbf{A}_{i0} \mathbf{W} + \sum_i t_{i2} \mathbf{A}_{i0} \mathbf{W} \right) + \left(\sum_i t_{i1} \mathbf{A}_{i1} \mathbf{W} + \sum_i \frac{1}{2} t_{i3} \mathbf{A}_{i1} \mathbf{W} \right), \quad (3.18)$$

$$\mathbf{XW} = \sum_i (t_{i0} + t_{i2}) \mathbf{A}_{i0} \mathbf{W} + \sum_i \left(t_{i1} + \frac{1}{2} t_{i3} \right) \mathbf{A}_{i1} \mathbf{W}. \quad (3.19)$$

Macierze \mathbf{A}_{ij} i \mathbf{W} mają stałe wartości elementów. W związku z tym, modyfikacja elementów t_{ij} macierzy \mathbf{T} spełniająca warunki (3.15) zapewnia, że nie ulegnie zmianie wartość iloczynu \mathbf{XW} . Oznacza to, że nie zmieniają się wartości elementów w ostatniej kolumnie macierzy \mathbf{X} , czyli nie ulegną zmianie przewidywane wartości punktów sąsiedniego bloku.

3.2.3. Ograniczanie utraty subiektywnej jakości sekwencji wizyjnej

Uwaga widza zwykle skoncentrowana jest na pewnym fragmencie obrazu, co powoduje, że zniekształcenia w innych jego częściach często pozostają niezauważone. Położenie obszaru zainteresowania zależy od treści, ale jest to też sprawa indywidualna. Z różnych badań eksperymentalnych wynika, że dla szerokiej klasy sekwencji obszar zainteresowania znajduje się w pobliżu geometrycznego środka obrazu. W związku z tym, w rozprawie proponuje się aby makrobloki modyfikowane były w kolejności zgodnej z uporządkowaniem spiralnym przedstawionym na rysunku 3.2. Dzięki temu w pierwszej kolejności jakość będzie ulegała pogorszeniu w zewnętrznej części obrazu. Ma to na celu podwyższenie subiektywnie postrzeganej jakości obrazu.



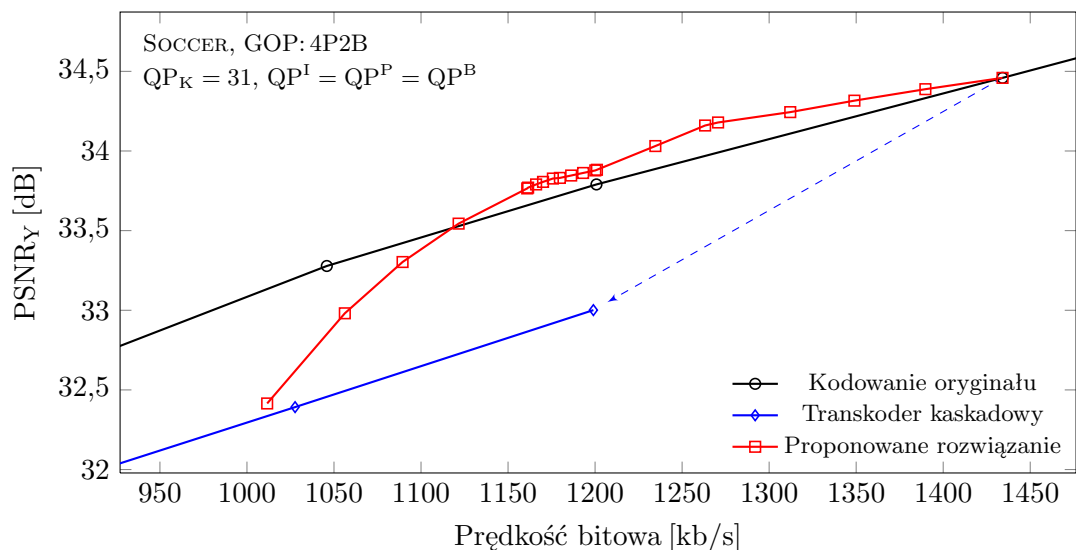
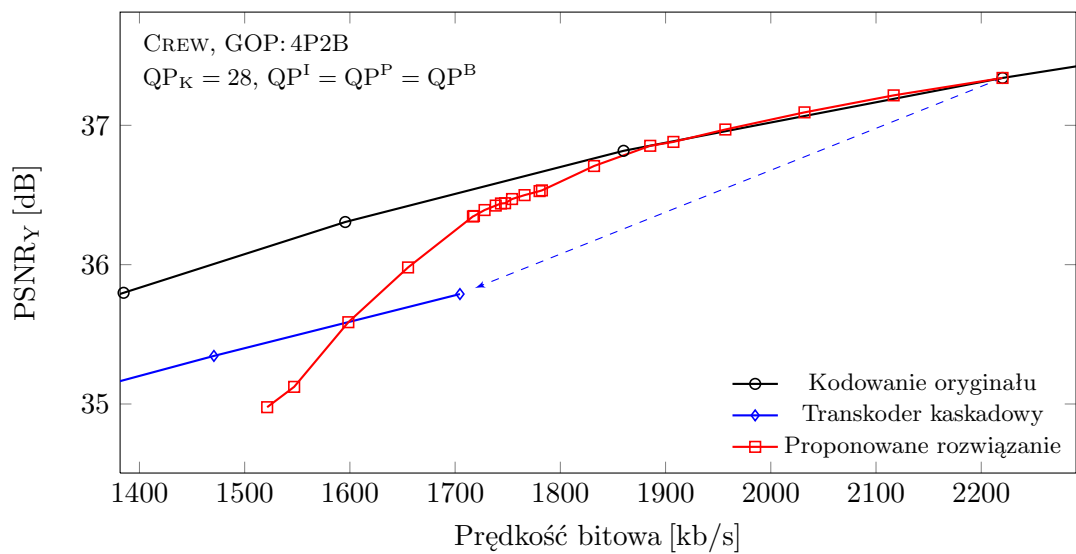
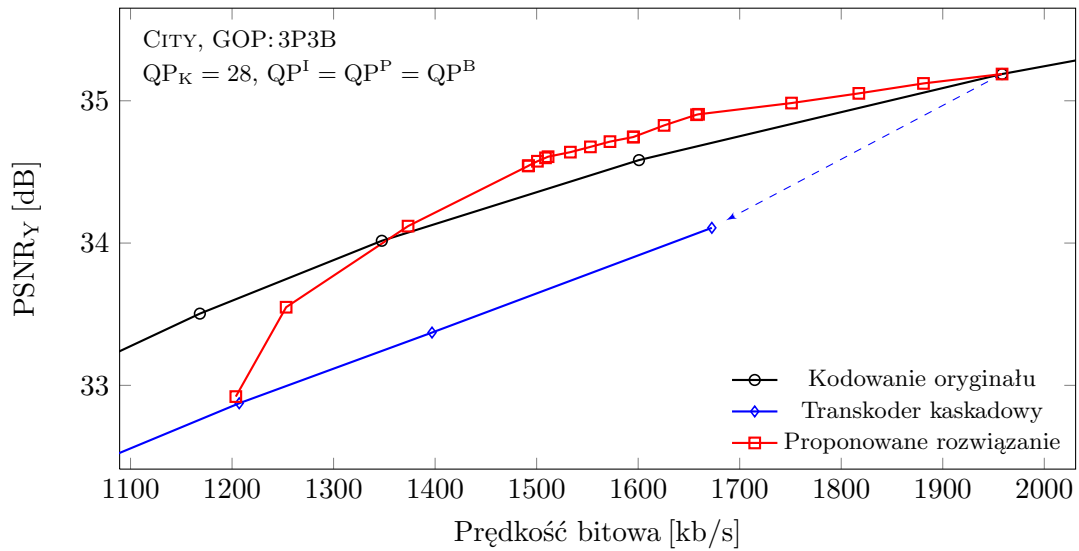
Rysunek 3.2: Spiralne uporządkowanie makrobloków

Ludzki system widzenia wykazuje większą tolerancję na zniekształcenia wysokoczęstotliwościowych niż niskoczęstotliwościowych składników sygnału wizyjnego. Dlatego też proponowane w rozprawie schematy transkodowania zakładają, że najpierw modyfikowane są próbki transformaty odpowiadające za najwyższe częstotliwości w obrazie, a następnie w razie konieczności zmieniane mogą być wartości próbek transformaty odpowiadające coraz niższym częstotliwościom.

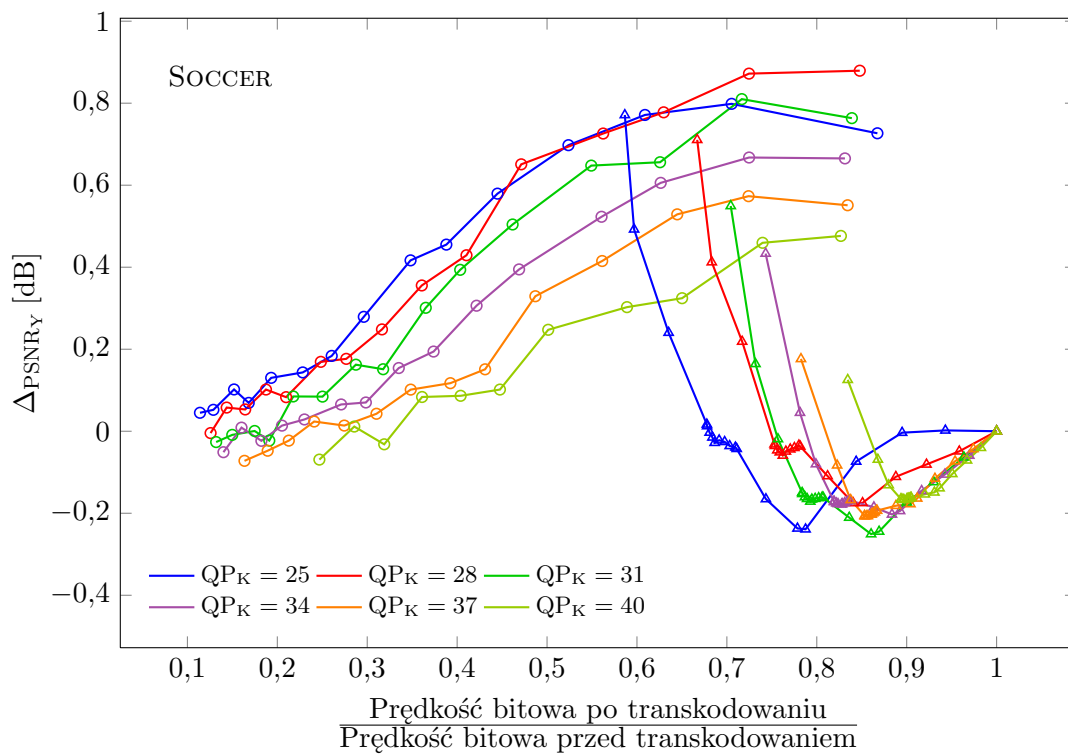
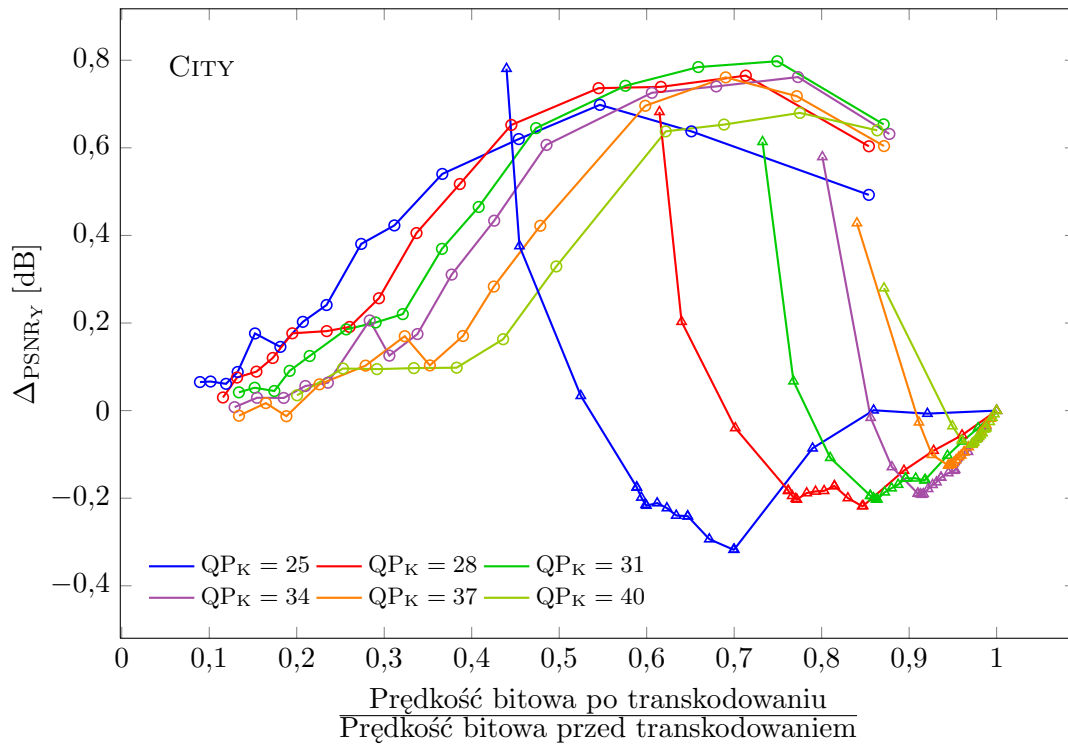
3.3. Badania zaproponowanych technik transkodowania z wykorzystaniem obiektywnej miary jakości

W rozprawie zastosowano miarę PSNR jako obiektywną miarę jakości. Jest to miara powszechnie stosowana i dzięki temu uzyskane wyniki można porównywać z wynikami badań innych autorów.

Na rysunku 3.3 przedstawiono przykładowe jakości kodowanych i transkodowanych sekwencji wizyjnych, natomiast rysunek 3.4 zawiera przykładowe porównanie strat transkodowania kaskadowego i transkodowania z wykorzystaniem jednego z zaproponowanych schematów transkodowania. Jest to znikomy fragment wyników przeprowadzonych doświadczeń. Dodatek do rozprawy, prezentujący wyłącznie wybrane rezultaty badań, zawiera



Rysunek 3.3: Jakość sekwencji wizyjnych uzyskanych w wyniku zakodowania oryginalnej sekwencji oraz transkodowania strumieni bitowych (włączony filtr usuwający efekt blokowy, dozwolone makrobloki Intra w obrazach P i B)



Rysunek 3.4: Straty jakości dla transkodera kaskadowego (\circ) i proponowanej metody redukcji strumienia bitowego (\triangle) w odniesieniu do jakości otrzymanej w wyniku jednokrotnej kompresji sekwencji oryginalnej (struktura GOP: 3P3B, $QP^I = QP^P = QP^B$, włączony filtr usuwający efekt blokowy, dozwolone makrobloki Intra w obrazach P i B)

ponad 150 stron podobnych wykresów. Dokumentuje on właściwości przedstawionych w zasadniczej części pracy technik transkodowania.

Na podstawie wyników przeprowadzonych doświadczeń stwierdzić można, że **zaproponowane metody transkodowania pozwalają uzyskać znacząco mniejsze straty jakości obrazu niż transkodowanie kaskadowe**. Te mniejsze straty występują w ograniczonym zakresie redukcji strumienia binarnego, przypadającym jednakże na obszar największych strat jakości obrazu spowodowanych transkodowaniem kaskadowym. **Zysk jakości wynikający z zastosowania zaproponowanych przez autora rozprawy technik transkodowania może sięgać nawet 1,2 dB**.

Analiza uzyskanych wyników doświadczeń pozwala na stwierdzenie, że wykorzystanie zaproponowanych metod transkodowania pozwala na bardzo efektywną realizację niewielkich redukcji strumienia bitowego. Jest to niemożliwe do uzyskania w przypadku transkodowania kaskadowego, gdyż takie transkodowanie z niezmienną wartością parametru kwantyzacji może redukować prędkość bitową strumienia nawet o 23 %. Zastosowanie większej niż w strumieniu wejściowym wartości parametru kwantyzacji powoduje dużą nadmiarowość strumienia bitowego czyli znaczące zwiększenie strat transkodowania.

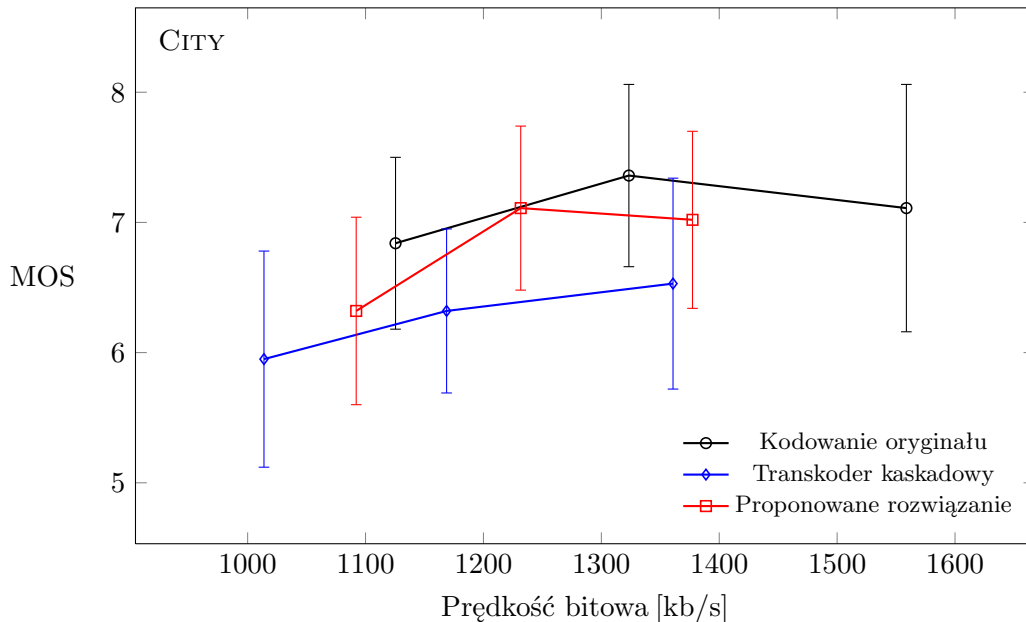
Dodatkowo **zaproponowane techniki transkodowania umożliwiają płynną redukcję prędkości strumienia bitowego, tak zwaną skalowalność drobnoziarnistą**.

3.4. Badania zaproponowanych technik transkodowania z wykorzystaniem subiektywnej metody oceny jakości

W celu potwierdzenia właściwości zaproponowanych algorytmów transkodowania określonych na podstawie badań obiektywnej miary jakości zdekodowanych sekwencji, przeprowadzone zostały badania jakości subiektywnej. Testy prowadzone były zgodnie ze wskazaniami zawartymi w zaleceniu ITU-R BT.500-11. Do ich przeprowadzenia wykorzystane zostało oprogramowanie przygotowane przez autora rozprawy.

W trakcie badań 19 obserwatorom zaprezentowano w losowej kolejności 60 sekwencji o długości 10s każda. Do badań wykorzystany został profesjonalny monitor *EIZO ColorEdge CG240W*, skalibrowany wcześniej przy użyciu kalibratora *Eye-One Display 2*. Testy prowadzone były z udziałem każdego z obserwatorów niezależnie. Szczegółowy opis przeprowadzonych badań znajduje się w pełnym tekście rozprawy. Przykładowe uśrednione wyniki opinii obserwatorów (ang. *Mean Opinion Score* – MOS) oraz przedziały ufności uzyskane zgodnie z zaleceniem ITU-R BT.500 znajdują się na rysunku 3.5.

Przedstawione wyniki nie pozwalają na statystyczną ocenę, czy widzowie byli w stanie rozróżnić jakość sekwencji otrzymanych w wyniku zakodowania oryginalnego materiału wizyjnego oraz sekwencji otrzymanych w wyniku transkodowania z wykorzystaniem badanych metod, gdyż uzyskane przedziały ufności nie są rozłączne. Zawężenie tych przedziałów wymagałoby przeprowadzenia eksperymentu ze znacząco większą grupą obserwatorów. Jest to zadanie niezwykle trudne, ze względu na potrzebę znalezienia odpowiedniej liczby osób chętnych do wzięcia udziału w eksperymencie, jak i z powodu bardzo długiego czasu testów. Dlatego do oceny jakości subiektywnej, wykorzystano uśrednione oceny obserwatorów. Jest to zgodne z rozwiązaniami stosowanymi przez innych autorów, a także z metodologią stosowaną przez grupę standaryzacyjną MPEG. Na podstawie uśrednionych wartości ocen obserwatorów można stwierdzić, że zaledwie cztery z dwudziestu jeden sekwencji uzyskanych w wyniku transkodowania z wykorzystaniem zaproponowanych algorytmów cechowało się niższą wartością wskaźnika MOS, niż interpolowane wartości wskaźnika MOS dla sekwencji transkodowanych kaskadowo. Należy zauważyć, że trzy z tych przypadków mają miejsce dla największych przebadanych redukcji strumienia bitowego, a w czwartym przypadku różnica jest pomijalnie mała.



Rysunek 3.5: Wyniki uśrednionych ocen subiektywnej jakości sekwencji wizyjnych wraz z przedziałami ufności 95 %. Prędkość bitowa strumienia bitowego poddawanego transkodowaniu wynosiła 1559 kb/s (struktura GOP: 3P3B, $QP^I = QP^P = QP^B - 2$, włączony filtr usuwający efekt blokowy, brak makrobloków Intra w obrazach P i B)

Powyższe rezultaty **nie stoją w sprzeczności z wynikami badań przeprowadzonych z wykorzystaniem obiektywnej miary jakości**. W ograniczonym przedziale redukcji prędkości bitowej zaproponowane algorytmy pozwalają na uzyskanie mniejszych strat jakości spowodowanych transkodowaniem niż transkodowanie kaskadowe.

3.5. Złożoność zaproponowanych rozwiązań

W pracy położono główny nacisk na to, aby proponowane mechanizmy redukcji prędkości bitowej zapewniały możliwie najmniejszą stratę jakości sekwencji spowodowaną transkodowaniem. Dodatkowo, realizacja zaproponowanych metod transkodowania wymaga niewielkich nakładów obliczeniowych.

Autor dysponował wyłącznie niezoptymalizowanymi implementacjami zaproponowanych rozwiązań. W związku z tym, nie było możliwe porównanie czasów transkodowania z wykorzystaniem zaproponowanych algorytmów, z czasami transkodowania kaskadowego. Możliwe jest jednak oszacowanie ich wzajemnej relacji. W rozprawie dla uproszczenia przyjęto, że nakłady obliczeniowe są wprost proporcjonalne do czasów obliczeń.

Przeprowadzona w pracy analiza prowadzi do wniosku, że **realizacja zaproponowanych algorytmów redukcji strumienia bitowego wymaga dwa rzędy wielkości mniej nakładów obliczeniowych niż transkodowanie kaskadowe**.

Dodatkowo, zastosowanie zaproponowanych procedur redukcji strumienia bitowego powoduje ograniczenie wymaganej pamięci danych w stosunku do transkodowania kaskadowego. Ze względu na znaczące uproszczenie algorytmiczne, zmniejszeniu ulega również zapotrzebowanie na pamięć programu.

4 Podsumowanie

4.1. Wnioski

Badania przedstawione w rozprawie koncentrowały się na stworzeniu nowych technik transkodowania homogenicznego sekwencji wizyjnych kodowanych z wykorzystaniem zaawansowanych technik kompresji. Uzyskane *wyniki badań potwierdzają tezę pracy*, to znaczy **zostały zaproponowane procedury transkodowania sekwencji wizyjnych wykorzystujące zaawansowane techniki kompresji, zapewniające wyższą jakość transkodowanej sekwencji niż transkodowanie kaskadowe**. Równocześnie, **realizacja zaproponowanych algorytmów transkodowania wymaga mniejszych nakładów obliczeniowych niż transkodowanie kaskadowe**.

Poniżej przedstawione zostaną najważniejsze wnioski rozprawy.

- Możliwe jest transkodowanie strumieni wizyjnych z otwartą pętlą sprzężenia zwrotnego, zapewniające mniejsze straty transkodowania niż w przypadku transkodowania kaskadowego.
- Realizacja zaproponowanych algorytmów redukcji strumienia bitowego wymaga około stukrotnie mniejszych nakładów obliczeniowych niż transkodowanie kaskadowe.
- Zakres redukcji prędkości transmisji dla algorytmów transkodowania zaproponowanych w rozprawie zależy w dużym stopniu od treści sekwencji wizyjnej. Największa redukcja, w przypadku niemodyfikowania obrazów typu I, wyniosła, w przebadanych przypadkach, prawie 70 %, a najmniejsza 6 %. Modyfikacja obrazów typu I, przeprowadzona zgodnie z procedurami przedstawionymi w rozprawie, w przebadanych przypadkach, pozwoliła na zwiększenie zakresu redukcji prędkości transmisji o maksymalnie 8 punktów procentowych.
- Zakres redukcji prędkości transmisji oferowany przez zaproponowane algorytmy jest uzależniony od występowania makrobloków Intra w obrazach P i B. Kiedy takie makrobloki nie występują, zakres redukcji może być nawet o 26 punktów procentowych większy.
- Możliwe jest ograniczenie błędu dryftu poprzez modyfikację obrazów typu P w kolejności odwrotnej do kolejności ich kodowania.
- Po spełnieniu warunków przedstawionych w pracy, możliwa jest taka modyfikacja próbek transformaty, która nie zmienia przewidywanych wartości próbek obrazu w sąsiednich blokach kodowanych wewnątrzobrazowo.
- Liczba zmian parametru kwantyzacji może być ograniczona poprzez wykorzystanie faktu, że w części makrobloków wszystkie próbki transformaty mają zerową wartość.
- Straty transkodowania kaskadowego strumieni wizyjnych zawierających wyłącznie obrazy typu I zależą od wykorzystania filtra usuwającego efekt blokowy. W zależności od wykorzystania tego narzędzia, możliwe jest wskazanie typowych zależności strat transkodowania w funkcji redukcji strumienia bitowego. Największe straty występują przy około 20 % redukcji prędkości transmisji strumienia bitowego.

- Zależność strat transkodowania kaskadowego strumieni wizyjnych zawierających obrazy typu P i B od redukcji prędkości bitowej ma typowy charakter. Największe straty występują przy 30 % redukcji strumienia bitowego i maleją wraz ze zwiększaniem stopnia redukcji.
- Badania przeprowadzone z wykorzystaniem referencyjnego oprogramowania standardu MPEG-4 AVC/H.264 wykazały, że w zakresie wartości parametru kwantyzacji 20–51, reprezentacja próbek transformaty stanowić może od 30 do 93 % całości strumienia bitowego.
- Znacząca większość niezerowych próbek transformaty w strumieniu bitowym ma amplitudę równą jeden. W obrazach P i B takie próbki stanowią około 80 % liczby niezerowych próbek, niezależnie od wartości parametru kwantyzacji.

4.2. Oryginalne osiągnięcia rozprawy

Głównym osiągnięciem autora jest **opracowanie technik transkodowania homogenicznego strumieni wizyjnych zgodnych ze standardem MPEG-4 AVC/H.264**. Autor pracy zaimplementował zaproponowane algorytmy. Pozwoliło to na przeprowadzenie eksperymentalnej analizy ich właściwości. W rozprawie przedstawiono dwa algorytmy, które umożliwiają redukcję prędkości bitowej. Wybór najlepszego algorytmu uzależniony jest od występowania makrobloków Intra w obrazach typu P i typu B. Oba zaprezentowane rozwiązania pozwalają na uzyskanie mniejszych strat jakości sekwencji niż transkodowanie kaskadowe. Różnica może wynosić nawet 1,2 dB. Jednocześnie, realizacja zaproponowanych algorytmów redukcji strumienia bitowego wymaga dwa rzędy wielkości mniej nakładów obliczeniowych niż transkodowanie kaskadowe. Zaproponowane metody redukcji strumienia binarnego umożliwiają realizację skalowalności drobnoziarnistej strumieni wizyjnych.

Istotnym osiągnięciem przedstawionym w pracy jest **analiza właściwości transkodowania kaskadowego strumieni wizyjnych zgodnych ze standardem MPEG-4 AVC/H.264**. Wynikiem przeprowadzonych przez autora pracy badań jest stwierdzenie, że zależność strat takiego transkodowania od redukcji prędkości bitowej strumienia wizyjnego ma charakter stały, niezależny od treści sekwencji i parametrów kodowania. Dla strumieni wizyjnych, w których występują wyłącznie obrazy typu I, największe straty transkodowania kaskadowego występują przy 20 % redukcji prędkości bitowej. W przypadku transkodowania strumieni w których występują dodatkowo obrazy typu P i typu B, największa wartość strat ma miejsce przy około 30 % redukcji strumienia bitowego. W obu przypadkach zwiększenie redukcji strumieni bitowych ponad wskazane powyżej wartości powoduje stopniowe zmniejszanie strat transkodowania. Podobne właściwości charakteryzują transkodowanie kaskadowe strumieni wizyjnych zakodowanych z wykorzystaniem klasycznych technik kompresji.

Poniżej przedstawiono *oryginalne osiągnięcia autora pracy, wspierające główne osiągnięcia*.

- Analiza eksperymentalna zaproponowanych algorytmów, w tym przede wszystkim:
 - porównanie strat transkodowania z wykorzystaniem zaproponowanych algorytmów i strat transkodowania kaskadowego,
 - wyznaczenie zakresu redukcji strumienia bitowego uzyskanej dzięki wykorzystaniu proponowanych algorytmów,
 - wskazanie zakresu redukcji strumienia bitowego, w którym straty jakości na skutek transkodowania z wykorzystaniem zaproponowanych algorytmów są nie większe niż straty transkodowania kaskadowego.
- Przeprowadzenie badań eksperymentalnych wskazujących jak modyfikacja różnej liczby próbek transformaty w różnej liczbie makrobloków tworzących obraz wpływa

na jakość transkodowanej sekwencji przy określonym stopniu redukcji strumienia bitowego.

- Zbadanie jak liczba zmodyfikowanych obrazów w strukturze GOP wpływa na jakość transkodowanej sekwencji.
- Wskazanie i udowodnienie warunków modyfikacji próbek transformaty, których spełnienie zapewnia, że nie zmieniają się przewidywane wartości próbek sąsiednich bloków kodowanych w trybie Intra.
- Zaproponowanie sposobu minimalizacji dryftu powstającego na skutek transkodowania z otwartą pętlą sprzężenia zwrotnego.
- Propozycja procedury redukcji liczby zmian wartości parametru kwantyzacji w obrazie, polegająca na wykorzystaniu faktu, że w części makrobloków wszystkie próbki transformaty mają zerową wartość. W wyniku działania zaproponowanej procedury nie zmieniają się wartości próbek w zdekodowanym obrazie.
- Przeprowadzenie badań zaproponowanych algorytmów transkodowania z wykorzystaniem subiektywnej metody oceny jakości.
- Porównanie nakładów obliczeniowych wymaganych do realizacji zaproponowanych algorytmów i do wykonania transkodowania kaskadowego.
- Zbadanie udziału w strumieniu bitowym informacji o próbkach transformaty, a także wskazanie, jaka część niezerowych próbek ma amplitudę równą jeden.
- Przedstawienie aktualnego stanu techniki w dziedzinie transkodowania strumieni wizyjnych wykorzystującego zaawansowane techniki kompresji.
- Implementacja aplikacji służącej do prowadzenia testów subiektywnej oceny jakości sekwencji wizyjnych.

Rezultaty zaprezentowane w pracy powstały w wyniku ogromnej pracy eksperymentalnej, na którą składało się wykonanie blisko **250 tysięcy kodowań i transkodowań**. Autor **zakodował i transkodował łącznie około 263 godzin sekwencji wizyjnych** o rozdzielczości przestrzennej 704×576 punkty i rozdzielczości czasowej 30 Hz. Na trud powstania tej pracy złożyły się również przetworzenie i analiza otrzymanych wyników.

Tak obszerna praca eksperymentalna została przeprowadzona w celu oceny i udokumentowania powtarzalności analizowanych w pracy właściwości transkodowania kaskadowego, a także transkodowania polegającego na modyfikacji próbek transformaty. Dzięki dużej liczbie przeprowadzonych testów potwierdzona została uniwersalność zaproponowanych algorytmów. Podkreślenia wymaga fakt, że ocena jakości obrazu wykonywana była zawsze na podstawie sekwencji zdekodowanych z wykorzystaniem referencyjnego dekodera standardu MPEG-4 AVC/H.264. Pozwoliło to na potwierdzenie poprawności przygotowanych implementacji zaproponowanych algorytmów transkodowania.

5 Dorobek naukowy autora

5.1. Publikacje wyników rozprawy

Konferencje międzynarodowe

- Marek Domański, Jarosław Marek. Fine grain scalability of bitrate using AVC/H.264 bitstream truncation. *Picture Coding Symposium PCS'09*, Chicago, 2009.
- Jarosław Marek. Homogeneous Video Transcoding of H.264-AVC Intra Coded Frames *Lecture Notes In Computer Science*, 5337:198–207, 2008
International Conference on Computer Vision and Graphics ICCVG'08, Warszawa, 2008.
- Jarosław Marek, Marek Domański. Direct transcoding of AVC/H.264 bitstreams for bitrate reduction with optimized video quality losses. *The 2009 International Conference on Image Processing, Computer Vision, and Pattern Recognition IPCV'09*, Las Vegas, 2009.

Czasopisma krajowe

- Jarosław Marek. Transkodowanie sekwencji wizyjnych wykorzystujące zaawansowane techniki kompresji. *Przegląd Telekomunikacyjny*, 4:565–568, 2008
Krajowa Konferencja Radiokomunikacji, Radiofonii i Telewizji KKRRiT'08, Wrocław, 2008.

Artykuł w czasopiśmie

W przygotowaniu:

- Low-complexity MPEG-4 AVC/H.264 bitstreams transrating with reduced video quality losses.

5.2. Uzyskane przez autora i nie ujęte w rozprawie oryginalne wyniki naukowe

1. **Kodek prawie-bezstratny danych geologicznych** — przygotowanie prawie-bezstratnego kodeka danych geologicznych.

Praca pod kierunkiem prof. dr. hab. inż. M. Domańskiego.

Udział własny w pracach naukowych:

- *Badania nad charakterem kompresowanego sygnału, istniejącą w nim redundancją i sposobem jej eliminacji.*
- *Opracowanie technik prawie-bezstratnej kompresji danych geologicznych.*
- *Zaprojektowanie kodera i dekodera danych geologicznych.*
- *Przygotowanie programowych modeli zaproponowanych kodera i dekodera danych geologicznych w środowisku Matlab.*

Udział własny w pracach wdrożeniowych:

- *Implementacja przygotowanych rozwiązań w języku ANSI C wraz z optymalizacją algorytmiczną i programową.*

Wynik: **wdrożenie** w firmie Halliburton – Exton Technology Center, USA.

2. **Kodek stratny danych geologicznych** — przygotowanie kodeka danych geologicznych o bardzo wysokim stopniu kompresji.

Praca pod kierunkiem prof. dr. hab. inż. M. Domańskiego.

Udział własny w pracach naukowych:

- *Badania nad sposobem eliminacji redundancji istniejącej w danych geologicznych.*
- *Opracowanie technik silnie stratnej kompresji danych geologicznych.*
- *Zaprojektowanie kodera i dekodera danych geologicznych.*

Udział własny w pracach wdrożeniowych:

- *Przygotowanie programowych modeli zaproponowanych kodera i dekodera danych geologicznych w środowisku Matlab.*
- *Implementacja przygotowanych rozwiązań w języku ANSI C wraz z optymalizacją algorytmiczną i programową.*

Wynik: **wdrożenie** w firmie Halliburton – Exton Technology Center, USA.

3. **Dekoder skalowanych strumieni wizyjnych zgodnych z MPEG-4 AVC/H.264** — optymalizacja referencyjnego oprogramowania pozwalająca na dekodowanie skalowalnych strumieni wizyjnych wysokich rozdzielczości w czasie rzeczywistym.

Praca zespołowa pod kierunkiem prof. dr. hab. inż. M. Domańskiego.

Udział własny w pracach naukowo-wdrożeniowych:

- *Opracowanie zoptymalizowanych algorytmów dekodowania skalowalnych strumieni binarnych, oraz ich efektywna implementacja w języku C++.*

Wynik: dekodek **zakupiony** przez Mitsubishi Electric Information Technology Centre Europe, UK.

4. **Transkoder sygnałów wizyjnych standardu MPEG-2 przetwarzanych do formatu H.264/AVC** — autorska koncepcja i implementacja transkodera heterogenicznego.

Praca zespołowa pod kierunkiem prof. dr. hab. inż. M. Domańskiego.

Udział własny w badaniach naukowych:

- *Badania nad sterowaniem przepływnością w heterogenicznym transkoderze wizyjnym.*
- *Opracowanie modelu relacji pomiędzy prędkością bitową strumieni wejściowego zgodnego ze standardem MPEG-2 i wyjściowego zgodnego ze standardem MPEG-4 AVC/H.264 w heterogenicznym transkoderze wizyjnym.*

Udział własny w pracach wdrożeniowych:

- *Implementacja algorytmu sterowania prędkością bitową w heterogenicznym transkoderze wizyjnym.*
- *Implementacja szybkich algorytmów interpolacji obrazu.*

Wynik: efekty prac badawczych i wdrożeniowych **zakupione** przez Advanced Digital Broadcast Polska.

5. **Koder wizyjny MPEG-4 AVC/H.264 PP** — autorska implementacja szybkiego koderza wizyjnego MPEG-4 AVC/H.264.

Praca zespołowa pod kierunkiem prof. dr. hab. inż. M. Domańskiego.

Udział własny w pracach badawczych:

- *Opracowanie efektywnego, adaptacyjnego algorytmu wyboru typu kodowanej warstwy w koderze hybrydowym standardu MPEG-4 AVC/H.264.*
- *Zaprojektowanie szybkich implementacji modułów transformacji, kwantyzacji, interpolacji, wyznaczania miary błędu dla predykcji międzyobrazowej, filtru deblokującego.*

Udział własny w pracach wdrożeniowych:

- *Implementacja modułów transformacji, kwantyzacji, interpolacji, wyznaczania miary błędu dla predykcji międzyobrazowej, filtru deblokującego, adaptacyjnego wyboru typu kodowanej warstwy w szybkim koderze standardu MPEG-4 AVC/H.264 z wykorzystaniem języków programowania ANSI C oraz assembler oraz z uwzględnieniem możliwości przetwarzania równoległego procesorów klasy Pentium 4.*

Wynik: koder **zakupiony i wdrożony** w Advanced Digital Broadcast Polska.

6. **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. Praca zrealizowana dla TVP S.A.

Udział własny w pracach naukowych:

- *Opracowanie oryginalnych modułów realizujących proces kodowania międzyobrazowego z predykcją dwukierunkową w autorskim koderze wizyjnym czasu rzeczywistego.*
- *Opracowanie wydajnych modułów realizujących funkcje wyznaczania miary błędu.*
- *Opracowanie modułów w efektywny sposób korzystających z wektorowego przetwarzania współczesnych procesorów.*
- *Opracowanie wydajnych modułów realizujących proces kodowania międzyobrazowego z predykcją dwukierunkową w autorskim koderze wizyjnym czasu rzeczywistego.*
- *Zaprojektowanie oryginalnego algorytmu regulacji prędkości transmisji.*

Udział własny w pracach wdrożeniowych:

- *Implementacja przygotowanych na etapie badawczym elementów kodeka wizyjnego z wykorzystaniem języków ANSI C, C++ i assembler.*
- *Modyfikacja kodeka opakowanego w filtry DMO tak, aby zapewnić mu funkcjonalność równoważną aplikacji konsolowej.*

Wynik: efekty pracy **zakupiła** Telewizja Polska S.A.

7. **Publikacje będące wynikiem innych badań**

Wyniki prac w materiałach konferencyjnych:

- Tomasz Dzieścielewski, Tomasz Grajek, Jarosław Marek. Eksperymentalna analiza efektywności trybów kodowania międzyobrazowego w zaawansowanym koderze wizyjnym. *Krajowa Konferencja Radiokomunikacji, Radiofonii i Telewizji KKRRiT'06*, str. 219–222, Poznań, 2006.
- Tomasz Dzieścielewski, Tomasz Grajek, Jarosław Marek. Eksperymentalna analiza częstotliwości wyboru trybów zaawansowanego kodowania wizyjnego. *X Poznańskie Warsztaty Telekomunikacyjne PWT'05*, str. 115–120, Poznań, 2005.

Artykuł przeglądowy w czasopiśmie:

- Marek Domański, Tomasz Grajek, Jarosław Marek. Zaawansowana kompresja cyfrowych sygnałów wizyjnych — standard AVC/H.264. *Systemy Alarmowe*, nr 2, rok wyd. XIV, vol. 75, str. 14-21, marzec-kwiecień 2005.