

**ŁĄCZNE WYKORZYSTANIE TECHNIKI RDOQ I METODY OPTIMALIZACJI KODÓW
HUFFMANA W KODERZE JPEG**
COMBINED USE OF THE RDOQ TECHNIQUE AND THE HUFFMAN CODE OPTIMIZATION METHOD IN
THE JPEG ENCODER

Jakub Stankowski, Artur Fojut, Damian Karwowski, Tomasz Grajek

Instytut Telekomunikacji Multimedialnej, Politechnika Poznańska, Poznań
{jakub.stankowski, damian.karwowski, tomasz.grajek}@put.poznan.pl

Streszczenie: Artykuł prezentuje autorskie badania dotyczące ulepszeń w kodowaniu JPEG. Zaproponowano połączenie techniki Rate-Distortion Optimized Quantization (RDOQ) z optymalizacją tablic kodów Huffmana w celu zwiększenia efektywności kompresji. Opracowano implementację tych metod i zintegrowano ją z autorskim kodekiem JPEG. Przeprowadzone eksperymenty na szerokim zestawie obrazów testowych wykazały poprawę efektywności kompresji od kilku do 20%, zależnie od ustawień kodera JPEG.

Abstract: The article presents original research on improvements to JPEG compression. The proposed approach combines the RDOQ technique with Huffman code table optimization to enhance coding efficiency. The methods were implemented and integrated into a custom JPEG codec. Experimental results on a diverse set of test images demonstrate compression gains ranging from several percent up to 20%, depending on the JPEG encoder settings.

Słowa kluczowe: Huffman (kodowanie), JPEG, kompresja obrazu, RDOQ, sterowanie koderem.

Keywords: Huffman (coding), image/video compression, image/video encoder control, JPEG, RDOQ.

1. WSTĘP

Minęły ponad trzy dekady od opracowania normy kompresji JPEG cyfrowych obrazów [8, 17]. Pomimo pojawienia się znacznie bardziej wydajnych technik kodowania obrazu, takich jak np. JPEG 2000 [9], JPEG-XR [11], JPEG-XL [7], HEVC Intra (HEIF) [6], czy wreszcie AV1 Intra (AVIF) [4], kodowanie JPEG nadal cieszy się ogromną popularnością i jest powszechnie stosowane w różnych aplikacjach i produktach rynkowych [5, 15, 19]. Wynika to głównie z jego prostoty, niskich wymagań sprzętowych i obliczeniowych oraz zgodności ze starszym oprogramowaniem i sprzętem.

Aby w pełni wykorzystać potencjał danej techniki kompresji, niezbędny jest odpowiedni algorytm sterowania koderem [12, 14]. Ze względu na nienormatywny charakter tego algorytmu (nie jest on opisany w normie danej techniki kompresji), istnieje możliwość zwiększenia efektywności techniki kompresji obrazu poprzez udoskonalenie mechanizmu sterowania koderem — nawet długo po

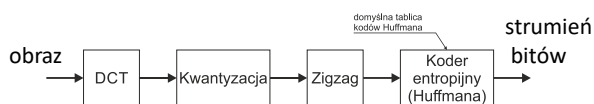
opracowaniu samej normy. Biorąc pod uwagę niesłabnącą popularność techniki JPEG, perspektywa poprawy współczynnika kompresji dzięki ulepszeniu algorytmu sterowania koderem wydaje się niezwykle atrakcyjna. Właśnie tego zagadnienia dotyczą badania, które przedstawiono w niniejszym artykule.

W literaturze można znaleźć różne podejścia do optymalizacji algorytmu sterowania w koderach obrazu statycznego i obrazu ruchomego, np. [2, 3, 13, 15, 18]. W przypadku techniki JPEG większość takich rozwiązań zakłada zdefiniowanie i przesłanie do dekodera obrazu nowych tablic kwantyzacji lub kodów entropijnych Huffmana albo zastosowanie wysokowydajnej zoptymalizowanej kwantyzacji (Rate-Distortion Optimized Quantization – RDOQ).

Celem niniejszej pracy jest eksperymentalna ocena wpływu jednoczesnego zastosowania wspomnianych mechanizmów, czyli techniki RDOQ oraz zmiany tablic kodów Huffmana na sprawność kompresji kodera JPEG. Przedstawione prace badawcze obejmują programową realizację zaproponowanych mechanizmów, ich włączenie do struktury oprogramowania kodera JPEG, oraz wykonanie serii eksperymentów kodowania obrazów testowych.

2. MOŻLIWOŚCI OPTIMALIZACJI KODOWANIA W TECHNICIE JPEG

W najprostszym przypadku przedstawianym w wielu opracowaniach (np. [12]), koder JPEG składa się z czterech bloków funkcjonalnych: dyskretnej transformacji kosinusowej (DCT), kwantyzacji otrzymanych próbek widma, uporządkowania zygzakowatego współczynników transformaty po kwantyzacji oraz kodera entropijnego właściwych danych (Rys. 1.). Dodatkowo można jeszcze uwzględnić krok konwersji przestrzeni barw (np. z RGB do YCbCr), który dla czytelności pominięto w prezentowanych poniżej schematach. Kodowanie entropijne danych w JPEG zazwyczaj utożsamiane jest z kodowaniem Huffmana, ponieważ możliwy, alternatywny wariant wykorzystujący kodowanie arytmetyczne danych jest praktycznie niespotykany (z powodu obowiązujących jeszcze do niedawna ograniczeń patentowych).



Rys. 1. – Schemat blokowy prostego kodera JPEG (bez optymalizacji)

Nawet tak prosta technika kompresji pozwala twórcom kodera obrazu na zastosowanie różnych algorytmów sterujących jego pracą. W kolejnych podpunktach przedstawiono dwie autorskie techniki pozwalające na znaczną poprawę sprawności kompresji z zachowaniem pełnej zgodności z normą [8] i gwarantujące pełną kompatybilność z dowolnym dekodującym JPEG.

2.1. Tablice kodów Huffmana

Technika JPEG definiuje zestaw domyślnych tablic kodów Huffmana. Jednak wiele formatów wymiany danych (w tym najpopularniejszy JFIF [10]) zakłada, że koder obrazu może wysłać własny zestaw tablic kodów Huffmana. Te indywidualnie wyznaczone tablice kodów powinny być lepiej dopasowane do charakterystyki kodowanego obrazu (charakteru jego treści). Tablice te przesyłane są w zwartej, pośredniej formie, a następnie są interpretowane przez dekodera obrazu i konwertowane na pełne tablice kodów.

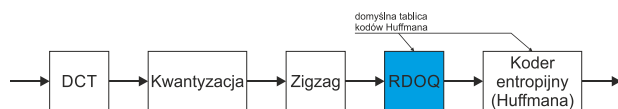
W celu zastosowania tej techniki konieczne jest uzupełnienie kodera obrazu o mechanizm analizy częstości występowania poszczególnych symboli danych (oznaczających poszczególne pary R-L (Run-Lenght) [12]) oraz dodatkowo o mechanizm wyznaczania tablicy kodów Huffmana, działający na podstawie wiedzy o prawdopodobieństwach wystąpienia poszczególnych symboli danych. Schemat blokowy takiego kodera został przedstawiony na Rys. 2.



Rys. 2. – Schemat blokowy kodera JPEG z optymalizacją tablic Huffmana

2.2. Rate-Distortion Optimized Quantization

Rate-Distortion Optimized Quantization (RDOQ) jest nienormatywną techniką pozwalającą na istotne zwiększenie wydajności kompresji danych poprzez zastosowanie dodatkowej analizy skwantowanych współczynników transformaty, wykonanej jeszcze przed kodowaniem entropijnym. RDOQ modyfikuje wartości skwantowanych współczynników DCT biorąc pod uwagę zarówno koszt bitowy ich zakodowania, jak i również ich wpływ na poziom zniekształceń zdekodowanego obrazu (Rys. 3.). Bardziej wyczerpujący opis techniki RDOQ można znaleźć w pracach [3, 16]. W szczególności zaś adaptacja techniki RDOQ do zastosowania w koderze JPEG została szczegółowo opisana w [3].



Rys. 3. – Schemat blokowy kodera JPEG z RDOQ

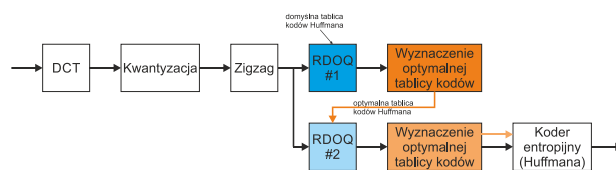
2.3. Łączne zastosowanie obydwu technik

Możliwe jest efektywne połączenie obu wyżej wymienionych technik, czyli jednocześnie zastosowanie RDOQ oraz optymalizacji tablic kodów Huffmana. W tym przypadku optymalne tablice kodów Huffmana są wyznaczane na podstawie skwantowanych współczynników DCT, wyznaczonych przez algorytm RDOQ (Rys. 4.). Warte zainteresowania jest też analiza zależności pomiędzy RDOQ a optymalizacją Huffmana i sprawdzenie, czy poprawa sprawności kompresji wnoszona przez każdą z technik jest niezależna (addytywna).



Rys. 4. – Schemat blokowy kodera JPEG z RDOQ i optymalizacją tablic kodów Huffmana

Kolejnym pytaniem jest, czy warto realizować wiele przebiegów wspomnianej optymalizacji? Możliwe jest, że RDOQ wykorzystujące zoptymalizowane tablice kodów Huffmana lepiej dopasuje się do krzywej prędkość bitowa-zniekształcenie (Rate-Distortion – RD). W tym celu zaproponowano dwuprzebiegową optymalizację kodowania, gdzie pierwszy przebieg wykorzystany jest do wyliczenia tablic kodów Huffmana, które to tablice są następnie wykorzystywane w optymalizacji RD przez drugi przebieg RDOQ (Rys. 5.). W ten sposób finalny krok RDOQ pracuje na tablicach słów kodowych Huffmana, które są lepiej dopasowane do charakterystyki kodowanego obrazu.



Rys. 5. – Schemat blokowy kodera JPEG z dwuprzebiegową optymalizacją RDOQ i tablic kodów Huffmana

3. WERYFIKACJA EKSPERYMENTALNA

Weryfikację przedstawionych w punkcie 2 koncepcji zrealizowano w dwóch krokach. W pierwszym przygotowana została implementacja kodera JPEG, a następnie przeprowadzone zostały eksperymenty mające na celu ocenę wpływu przyjętych rozwiązań na sprawność kompresji danych.

3.1. Oprogramowanie

Autorska implementacja kodera JPEG zawierająca zarówno technikę RDOQ, jak i mechanizm wyznaczania optymalnych tablic kodów Huffmana została przygotowana na bazie opracowanego przez autorów oprogramowania JOptEnc. Wspomniane oprogramowanie jest udostępnione na licencji BSD w repozytorium: <https://github.com/jstankowski/joptenc>.

Istniejąca wcześniej wersja JOptEnc zawierała już implementację techniki RDOQ, natomiast na potrzeby opisywanych w tej pracy badań zaimplementowano algorytm wyznaczania optymalnych tablic kodów Huffmana oraz rozwiązano wszystkie problemy, jakie pojawiły się w związku z interakcją pomiędzy dwiema technikami.

Między innymi takie, jak konieczność wyznaczenia kosztu bitowego zakodowania symbolu danych, który nie występuje w zoptymalizowanej tablicy kodów Huffmana. Zmodyfikowana wersja oprogramowania JOptEnc jest dostępna w wymienionym powyżej repozytorium.

3.2. Metodyka badań

W badaniach skupiono się na ocenie efektywności zaproponowanych rozwiązań, w kontekście kompresji sekwencji obrazów (obrazy ruchome), a nie kompresji obrazów statycznych. W przypadku takiego scenariusza koder JPEG koduje niezależnie od siebie kolejne obrazy sekwencji wizyjnej.

Jako materiał testowy do badań wykorzystano 17 sekwencji testowych używanych w ramach prac grupy MPEG. Wszystkie sekwencje miały rozdzielczość przestrzenną FullHD i zawierały co najmniej 100 obrazów. Sekwencje znacząco różniły się treścią, ilością detali, stopniem zaszumienia, itd. Pozwoliło to na przetestowanie wspomnianych technik na szerokim zakresie treści i zebranie dużego zestawu danych eksperymentalnych.

Jako metodę porównania sprawności kompresji przyjęto dobrze w kompresji znaną miarę Bjontegaarda [1], a dokładniej metrykę BD-rate. Wartość BD-rate wyznaczana była dla dwóch punktów pracy kodera odpowiadających: średniej jakości (wyższy stopień kompresji) – z widocznymi, acz nieuciążliwymi zniekształceniami oraz wysokiej jakości (niższy stopień kompresji) – bez widocznych zniekształceń zakodowanej treści. W technice JPEG jakością zakodowanych obrazów steruje się poprzez regulację parametru Q (Quality), który ma bezpośredni wpływ na zawartość tablic kwantyzacji [8]. W związku z tym przyjęto, że średnia jakość odpowiada zakresowi $Q \in \{50, 55, 60, 65\}$ a wysoka jakość zakresowi $Q \in \{80, 85, 90, 95\}$. W obu przypadkach użyto czterech wartości Q , gdyż tylu punktów pomiarowych wymaga obliczenie miary BD-rate.

Motywacją do przebadania dwóch różnych punktów pracy kodera był fakt występowania wpływu stopnia (siły) kwantyzacji danych na liczbę przesyłanych niezerowych współczynników próbek DCT. Zakres średniej jakości odpowiada małej liczbie przesyłanych współczynników DCT, zaś zakres wysokiej jakości – dużej liczbie przesyłanych współczynników DCT.

3.3. Scenariusze testowe

W ramach weryfikacji wpływu techniki RDOQ i metody zoptymalizowanych tablic kodów Huffmana na sprawność kompresji danych zdefiniowano pięć scenariuszy badawczych:

1. Odniesienia – w tym przypadku zastosowano prosty koder JPEG (przedstawiony na Rys. 1.)
2. OHT – koder JPEG z mechanizmem wyznaczania optymalnych tablic kodów Huffmana (przedstawiony na Rys. 2.)
3. RDOQ – koder JPEG, który wykorzystuje techniki adaptacyjnej (zoptymalizowanej) kwantyzacji RDOQ (przedstawiony na Rys. 3.)

4. RDOQ+OHT – koder JPEG używający obydwu technik optymalizacji (przedstawiony na Rys. 4.)
5. Dwuprzebiegowe RDOQ+OHT – rozbudowany wariant z pkt 4. (przedstawiony na Rys. 5.)

4. WYNIKI EKSPERYMENTALNE

Wyniki eksperymentalne zostały zebrane w tabelach 1 i 2, odpowiednio dla wyższego i niższego stopnia kompresji danych.

Tab. 1. Wyniki dla wyższego stopnia kompresji (średniej jakości) $Q \in \{50, 55, 60, 65\}$. Ujemne wartości oznaczają poprawę efektywności kodowania obrazów.

Nazwa sekwencji	Poprawa sprawności kompresji (DB-rate) dla danego scenariusza [dB]			
	OHT	RDOQ	RDOQ + OHT	dwuprzebiegowe RDOQ + OHT
Poznan_Street	-5,41%	-10,16%	-14,43%	-14,93%
BasketballDrive	-6,69%	-11,03%	-17,10%	-17,71%
blue_sky	-1,79%	-6,56%	-8,33%	-8,76%
BQTerrace	-3,31%	-11,60%	-14,27%	-14,97%
Cactus	-2,81%	-10,49%	-12,96%	-13,34%
Kimono1	-6,47%	-8,89%	-14,95%	-15,12%
ParkScene	-3,38%	-9,55%	-12,59%	-12,89%
pedestrian_area	-7,94%	-8,39%	-15,91%	-15,91%
Riverbed	-4,66%	-8,19%	-12,57%	-12,71%
rush_hour	-10,46%	-9,01%	-19,15%	-19,25%
station2	-6,25%	-9,08%	-14,55%	-14,88%
Sunflower	-4,13%	-7,13%	-11,41%	-11,75%
Tennis	-9,61%	-9,22%	-18,30%	-18,72%
toys_calendar	-4,87%	-10,12%	-14,50%	-14,83%
Tractor	-4,07%	-7,81%	-11,59%	-11,95%
vintage_car	-6,68%	-9,40%	-15,33%	-15,84%
walking_couple	-4,84%	-10,10%	-14,37%	-14,66%
Średnia	-5,49%	-9,22%	-14,25%	-14,60%

Tab. 2. Wyniki dla niższego stopnia kompresji (wysokiej jakości) $Q \in \{80, 85, 90, 95\}$. Ujemne wartości oznaczają poprawę efektywności kodowania obrazów.

Nazwa sekwencji	Poprawa sprawności kompresji (DB-rate) dla danego scenariusza [dB]			
	OHT	RDOQ	RDOQ + OHT	dwuprzebiegowe RDOQ + OHT
Poznan_Street	-0,72%	-7,69%	-8,23%	-8,76%
BasketballDrive	-2,73%	-10,10%	-11,58%	-12,61%
blue_sky	-2,45%	-9,35%	-11,57%	-12,44%
BQTerrace	-2,56%	-8,49%	-10,07%	-11,38%
Cactus	-1,26%	-9,64%	-10,26%	-11,19%
Kimono1	-1,41%	-11,87%	-12,67%	-13,47%
ParkScene	-0,89%	-8,95%	-9,61%	-10,37%
pedestrian_area	-1,60%	-11,05%	-11,93%	-12,70%
riverbed	-1,47%	-8,30%	-9,49%	-10,01%
rush_hour	-2,93%	-13,31%	-15,27%	-16,10%
station2	-1,04%	-9,84%	-10,56%	-11,19%
sunflower	-2,48%	-10,60%	-12,82%	-13,62%
tennis	-2,23%	-13,08%	-14,56%	-15,42%
toys_calendar	-1,49%	-10,95%	-11,71%	-12,55%
tractor	-1,55%	-8,96%	-10,21%	-10,92%
vintage_car	-1,61%	-9,03%	-10,05%	-10,75%
walking_couple	-1,51%	-9,50%	-10,48%	-11,20%
Średnia	-1,76%	-10,04%	-11,24%	-12,04%

Z przedstawionych danych jasno wynika, że zysk z zastosowania OHT jest znacznie wyższy dla sytuacji, gdy przesyłana jest niewielka liczba niezerowych współczynników DCT i wyraźnie maleje dla punktu pracy ze słabiej skwantowanymi współczynnikami DCT.

Jeśli chodzi o RDOQ to średnie zyski z zastosowania tej techniki są bardzo podobne w obu analizowanych zakresach parametru jakości Q .

Co istotne, można zauważyć, że łączne użycie obu technik (scenariusz RDOQ+OHT) wiąże się ze wzrostem sprawności kompresji zbliżonym do sumy zysków wynikających z niezależnego stosowania każdej z technik z osobna.

Zastosowanie dwuprzebiegowej optymalizacji pozwala na uzyskanie dodatkowego zwiększenia sprawności kompresji o około 1 punkt procentowy.

5. PODSUMOWANIE

Zastosowanie wybranych metod optymalizacji kodowania danych w technice JPEG umożliwia istotne zwiększenie efektywności kompresji przy zachowaniu pełnej zgodności ze standardem ISO/IEC 10918-1. Oznacza to, że zakodowany strumień danych może być poprawnie zdekodowany przez dowolny standardowy dekodery JPEG, co ma kluczowe znaczenie w kontekście praktycznego wykorzystania uzyskanych rezultatów.

W pracy zaproponowano połączenie dwóch rozwiązań: optymalizacji procesu kwantyzacji współczynników DCT z wykorzystaniem algorytmu sterowania RDOQ oraz adaptacyjnej modyfikacji tablic kodów Huffmana na podstawie rzeczywistego rozkładu statystycznego danych. Zaprojektowane warianty integracji tych technik w ramach kodera JPEG prowadzą do zauważalnej poprawy współczynnika kompresji.

Efektywność zaproponowanych rozwiązań została oceniona eksperymentalnie na szerokim zbiorze sekwencji testowych, przy różnych konfiguracjach kodera i poziomach jakości obrazu. Szczegółowe wyniki przedstawiono w Tabelach 1 i 2, a analiza danych wskazuje na wyraźne korzyści wynikające z zastosowania proponowanych metod – redukcja prędkości bitowej nawet o 20%.

LITERATURA

- [1] Bjøntegaard G. 2001. „Calculation of average PSNR differences between RD curves”, *ITU-T SG16/Q6*, dokument VCEG-M33.
- [2] Canh T., Xu M., Jeon B. 2018. „Rate-Distortion Optimized Quantization: A Deep Learning Approach”, *IEEE High Performance Extreme Computing Conference*.
- [3] Grajek T., Stankowski J. 2022. „Rate-Distortion Optimized Quantization in Motion JPEG”, *30th International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision*.
- [4] Guyon Y., Barnes L., Chang W.-T. (edytorzy). 2021. „AV1 Image File Format (AVIF)”, *Alliance for Open Media*, <https://aomediacodec.github.io/av1-avif/>.
- [5] Hudson G., Léger A., Niss B., Sebestyén I., Vaaben J. 2018. „JPEG-1 standard 25 years: past, present, and future reasons for a success”, *Journal of Electronic Imaging*, 27 (4) : 040901.
- [6] ISO. 2017. „Information technology — High efficiency coding and media delivery in heterogeneous environments — Part 12: Image File Format”, ISO/IEC 23008-12.
- [7] ISO. 2022. „Information technology - JPEG XL image coding system”, ISO/IEC IS 18181.
- [8] ISO/ITU. 1992. „Information Technology — Digital Compression and Coding of Continuous-Tone Still Images — Requirements and Guidelines”, ISO 10918-1, także jako ITU T.81.
- [9] ISO/ITU. 2019. „Information technology - JPEG 2000 image coding system”, ISO/IEC IS 15444 Edition 4, także jako ITU-T Rec. T.800, Edition 06/2019.
- [10] ITU. 2011. „Information technology – Digital compression and coding of continuous-tone still images: JPEG File Interchange Format (JFIF)”. ITU-T Rec. T.871.
- [11] ITU. 2020. „Information technology - JPEG XR image coding system”, ITU-T Rec. T.832.
- [12] Karwowski D. 2019. „Zrozumieć kompresję obrazu”. ISBN 978-83-953420-0-4.
- [13] Luo X., Talebi H., Yang F., Elad M., Milanfar P. 2021 „The Rate-Distortion-Accuracy Tradeoff: JPEG Case Study”, *Data Compression Conference*.
- [14] Richardson I. 2002. „Video Codec Design: Developing Image and Video Compression Systems”, Wiley.
- [15] Richter T. 2016. „JPEG on STEROIDS: Common optimization techniques for JPEG image compression”, *IEEE International Conference on Image Processing (ICIP)*, 61 – 65.
- [16] Stankowski, J., Korzeniewski, C., Domański, M., Grajek, T. 2015. „Rate-distortion optimized quantization in HEVC: performance limitations”, *31st Picture Coding Symposium*, 85-89.
- [17] Wallace G.K. 1992. „The JPEG still picture compression standard”, *IEEE Transactions on Consumer Electronics*, 38 (1) : xviii – xxxiv.
- [18] Wang Q., Liu P., Zhang L., Cheng F., Qiu J., Zhang X. 2022. „Rate-distortion optimal evolutionary algorithm for JPEG quantization with multiple rates”, *Knowledge-Based Systems*, 244: 108500.
- [19] W3Techs. 2025. „Usage statistics of JPEG for websites”, <https://w3techs.com/technologies/details/im-jpeg>.