

Henryk PALUS, Mariusz FRĄCKIEWICZ, Aron MANDRELLA  
Politechnika Śląska

## SZYBKĄ KWANTYZACJĄ BARWY POPRZEZ KLASTERYZACJĘ METODĄ K-ŚREDNICH NA WYBRANYCH PRÓBKACH OBRAZU

**Streszczenie.** Kwantyzacja barwy stała się ważną operacją w zadaniach przetwarzania obrazów barwnych i wizji komputerowej. Potrzebne są szybkie metody kwantyzacji, które jednocześnie generują skwantowane obrazy wysokiej jakości. W artykule przedstawiono metodę kwantyzacji barwy opartą na podpróbkiowaniu obrazu oryginalnego i klasteryzacji techniką K-średnich wybranych pikseli obrazu z deterministyczną inicjalizacją algorytmem Wu. Metoda ta znacznie przyspieszyła kwantyzację barwy bez zauważalnej utraty jakości obrazu, ocenianej za pomocą trzech wskaźników: MSE, HPSI i DSCSI.

## FAST COLOR QUANTIZATION USING K-MEANS CLUSTERING ON SELECTED IMAGE SAMPLES

**Summary.** Color quantization has become an important operation in color image processing and computer vision tasks. Fast quantization methods that simultaneously generate quantized high-quality images are required. This paper presents a color quantization method based on downsampling the original image and clustering of reduced number pixels using K-means technique with deterministic initialization by Wu's algorithm. This method significantly accelerated the color quantization without any noticeable loss of image quality, evaluated by three indices: MSE, HPSI and DSCSI.

### 1. Wprowadzenie

Kwantyzacja obrazów barwnych odgrywa pomocniczą, ale wciąż ważną rolę w takich zadaniach jak kompresja obrazów, segmentacja obrazów, wyszukiwanie obrazów itp. Operacja ta znacznie zmniejsza liczbę barw w obrazie, a jednocześnie powinna zachować podobieństwo skwantowanego obrazu do obrazu oryginalnego. Wśród metod kwantyzacji barw można wyróżnić dwie duże grupy, tj. metody podziałowe, np. *Median – Cut* (MC) [7] czy algorytm Wu [20] oraz metody klasteryzacji, np. metoda K-Średnich (*K – Means*, KM) [12]. Ta ostatnia metoda daje bardzo dobre wyniki, ale jej iteracyjność i złożoność obliczeniowa powoduje, że czas obliczeń jest długi. Wyniki uzyskane za pomocą KM zależą od inicjalizacji, tzn. od tego, w jaki sposób zostają ustalone początkowe środki klastrów. Ich losowy wybór, zwykle stosowany w KM, prowadzi do niepowtarzalnych wyników. Zastosowanie kwantyzacji podziałowej w postaci metod MC lub Wu może zapewnić deterministyczną inicjalizację dla metody KM [13].

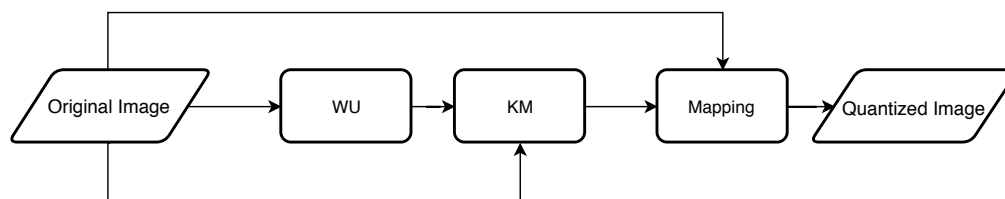
Wysoka złożoność obliczeniowa KM wynika z konieczności przeprowadzania

bardzo dużej liczby porównań między danymi wejściowymi, tj. barwami pikseli. Jest to szczególnie istotne w przypadku obrazów megapikselowych. Dlatego zastosowanie odpowiednich struktur danych oraz zmniejszenie liczby obliczanych odległości między danymi pozwala na skrócenie czasu obliczeń w metodzie KM [8]. Innym podejściem, szczególnie przydatnym w kwantyzacji barwy, może być klasteryzacja prowadzona na wybranej próbce obrazu zamiast na całym obrazie. Pozostaje do rozwiązania problem, w jaki sposób próbować obraz: losowo czy systematycznie? Kolejnym wyzwaniem jest określenie odpowiedniego rozmiaru takiej próbki, aby uzyskać jak najlepsze wyniki i krótszy czas obliczeń [3]. W swojej pracy Bejarano i inni zastosowali maksymalny rozmiar próbki wynoszący około 0,5 % wszystkich danych, a wyniki klasteryzacji były bardzo dobre i uzyskane w znacznie krótszym czasie. Zmniejszenie liczby pikseli zmniejsza liczbę barw przetwarzanego obrazu i można wtedy przeprowadzić kwantyzację barwy na kilku tysiącach barw, a nie na kilku milionach barw. Efektem powinna być niewielka utrata jakości obrazu i znacznie krótszy czas całej kwantyzacji.

Proponowana w artykule metoda kwantyzacji barwy ma kilka zalet. Przede wszystkim zapewnia jednocześnie wysoką jakość skwantowanego obrazu i krótki czas kwantyzacji. Metoda ta ogranicza się do etapu generowania palety barw (inicjalizacja, klasteryzacja) z danych obrazu poddanego podpróbkiowaniu metodą interpolacji NNI (ang. *Nearest Neighbor Interpolation*). Wskaźniki jakości skwantowanych obrazów porównane ze wskaźnikami dla innej metody opartej na ograniczonej próbce pikseli (algorytm oparty na tzw. *rdzeniach*) wskazują na przewagę proponowanej metody.

## 2. Próbki obrazu w procesie kwantyzacji barwy

Podstawą teoretyczną do rozwiązania naszego problemu jest interpolacja obrazu. Operacja ta występuje na niektórych etapach przetwarzania obrazu, takich jak zmiana rozmiaru obrazu (ang. *resampling*), czy jego rektyfikacja. Problem interpolacji obrazu został szeroko opisany w pracy [16]. Powszechnie stosowane techniki interpolacji nieadaptacyjnej to interpolacja metodą najbliższych sąsiadów (NNI), interpolacja biliniowa i interpolacja bikubiczna. NNI jest najprostszą metodą interpolacji, która po prostu odwzorowuje najbliższe sąsiednie piksele. Oznacza to, że w przypadku obrazów barwnych w wyniku takiej interpolacji, nie są tworzone nowe barwy. Dlatego podpróbkiowanie obrazu może jedynie zmniejszyć liczbę barw, ale ich wartości pozostaną niezmienione. W 2002 r. Papamarkos i in. przedstawili koncepcję podpróbkiowania obrazu za pomocą techniki skanowania fraktalnego [14] opartą na krzywej Hilberta wypełniającej przestrzeń (algorytm Adaptacyjnej Redukcji Barw). W 2014 r. Szilagy i współpracownicy [17] poprzedzili kwantyzację barwy opartą na metodzie KM specjalnym przetwarzaniem wstępnym polegającym na statycznej kwantyzacji barw na około 140 000 barw oraz obliczeniu histogramu barw z odrzuceniem najrzadziej występujących barw. Kluczowym parametrem przetwarzania wstępnego jest ustalenie jaka część pikseli obrazu będzie włączana do procesu klasteryzacji. Rezygnacja z 2-5% pikseli oryginalnego obrazu zmniejszyła liczbę barw do 500-5000 barw i skróciła czas kwantyzacji obrazu (współczynnik przyspieszenia 2-3) bez utraty szczegółów obrazu. W pracy Valenzueli, Celebiego i Schaefera [19] zaproponowano przyspieszoną wersję metody KM wykorzystującą podpróbkiowanie (decymację) obrazu wejściowego, niedeterministyczny ale efektywny algorytm inicjalizacji KM++ oraz wydajną konstrukcję zbioru tzw. rdzeni



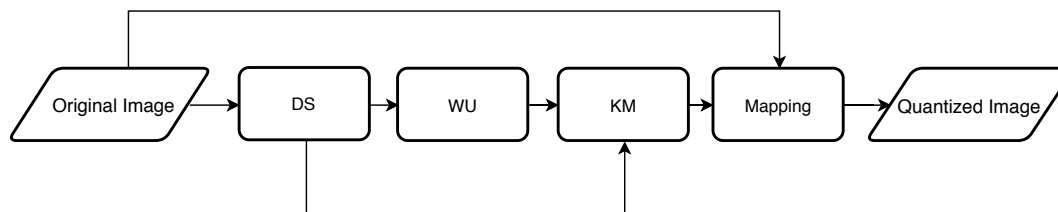
Rys. 1. Ogólna idea metody kwantyzacji barwy Wu+KM

(ang. *coresets*). Autorzy użyli nazwy *CoreMeans* (CM) dla swojej metody kwantyzacji barwy. Decymacja jest prostym przekształceniem obrazu polegającym na zachowaniu każdej  $n$ -tej próbki i odrzuceniu pozostałych. W artykule rozważano dwie wartości współczynnika decymacji  $df$  oznaczone jako CM2 i CM4. Algorytm KM++ jest znany z literatury [2]. Zbiór rdzeni został skonstruowany na podstawie podpróbkowanego obrazu oryginalnego oraz na podstawie początkowego zestawu środków klastrów pochodzących z algorytmu KM++. Algorytm Lucica i in, [11] został zaadaptowany do konstrukcji zestawu rdzeni. Przetestowano cztery wartości ułamków rdzeniowych:  $cf=1/2, 3/8, 1/4, 1/8$ . W tej pracy na czterech obrazach przetestowano metodę CM omówioną powyżej. CM była znacznie szybsza niż metoda KM i zwróciła obrazy skwantowane o jakości zbliżonej do wyników KM. Szczegółowe porównanie wyników CM z wynikami proponowanej przez nas metody przedstawiono w rozdziale 4.

Kwantyzacja barwy wraz z podpróbkowaniem obrazu znalazła również zastosowanie w sztuce *pixel art* [6], która przypomina ręczną pracę artystów tworzących piksele. Obrazy tego typu, powstałe w wyniku automatycznej konwersji barwnego obrazu wejściowego na pikselowy (ikoniczny) obraz wyjściowy, wykorzystują niską rozdzielczość przestrzenną i małą paletę barw. Opisana procedura konwersji składa się z dwóch części: algorytmu SLIC (ang. *Simple Linear Iterative Clustering*), który segmentuje obraz wejściowy na superpiksele oraz algorytmu klasteryzacji rozmytej MCDA (ang. *Mass Constrained Deterministic Annealing*). Praktycznym zastosowaniem kwantyzacji barwy i podpróbkowania obrazu, jest bezprzewodowa sieć kamer z zasilaniem bateryjnym [1]. Sieć ta została wykorzystana do monitorowania ptasich gniazd. Obie operacje znacznie zmniejszyły ilość danych przesyłanych między kamerami a serwerem i jednocześnie zmniejszyły zużycie energii. Podejście to działa jak algorytm stratnej kompresji obrazu ograniczony koniecznością zachowania odpowiedniej jakości obrazu. Zdaniem autorów najkorzystniejszą metodą kwantyzacji barwy była kwantyzacja przy użyciu stałej palety (8, 16, 32 barw) oraz algorytmu interpolacji biliniowej do podpróbkowania obrazu.

### 3. Opis proponowanej metody

Głównym zadaniem w stosunku do proponowanej metody była ocena wpływu generowania palety barw z próbki pikseli zamiast z całego obrazu na czas obliczeń i jakość kwantyzacji. W 2010 roku Palus i Frąckiewicz zastosowali algorytm podziałowy Wu jako efektywną metodę inicjalizacji KM [13]. Na Rys. 1 przedstawiono ideę kwantyzacji barwy, gdzie oryginalny obraz o pełnej rozdzielczości jest punktem wyjścia dla wszystkich procesów, tj. inicjalizacji (algorytm Wu), klasteryzacji (KM) i końcowego mapowania pikseli. Złożoność czasowa tej metody kwantyzacji wynika głównie ze zło-



Rys. 2. Idea zaproponowanej metody kwantyzacji barwy

Tabela 1

Wartości MSE dla kwantyzacji na 32 barwy

| $cf, df$ | NNI    | CM2 | CM4 | NNI  | CM2 | CM4 | NNI     | CM2 | CM4 | NNI   | CM2 | CM4 |
|----------|--------|-----|-----|------|-----|-----|---------|-----|-----|-------|-----|-----|
|          | Baboon |     |     | Lena |     |     | Peppers |     |     | Pills |     |     |
| 1        | 378    | 378 | 378 | 126  | 120 | 120 | 234     | 233 | 233 | 204   | 203 | 203 |
| 1/2      | 379    | 383 | 385 | 124  | 121 | 123 | 233     | 232 | 236 | 210   | 205 | 206 |
| 1/4      | 376    | 382 | 395 | 123  | 122 | 124 | 233     | 234 | 240 | 208   | 205 | 208 |
| 1/8      | 381    | 385 | 398 | 123  | 122 | 125 | 237     | 235 | 244 | 206   | 203 | 208 |
| 1/16     | 392    | N/A | N/A | 121  | N/A | N/A | 238     | N/A | N/A | 207   | N/A | N/A |
| 1/32     | 388    | N/A | N/A | 123  | N/A | N/A | 238     | N/A | N/A | 207   | N/A | N/A |
| 1/64     | 384    | N/A | N/A | 123  | N/A | N/A | 237     | N/A | N/A | 207   | N/A | N/A |
| 1/128    | 412    | N/A | N/A | 122  | N/A | N/A | 242     | N/A | N/A | 207   | N/A | N/A |
| 1/256    | 400    | N/A | N/A | 132  | N/A | N/A | 249     | N/A | N/A | 212   | N/A | N/A |

żoności metody KM i wynosi  $O(df \cdot n \cdot k \cdot d \cdot i)$  [4], gdzie:  $n$  jest liczbą pikseli,  $k$  jest liczbą oczekiwanych barw,  $d$  jest liczbą składowych barw, a  $i$  jest liczbą iteracji. Na Rys. 2 pokazano ideę nowej metody opartej na wstępnym podpróbkowaniu obrazu (blok DS).

Algorytm kwantyzacji Wu jako typowa metoda podziałowa korzysta ze statystycznej eksploracji barw występujących w oryginalnym obrazie. Jest on oparty na ortogonalnym, dwuczęściowym podziale sześcianu RGB na  $k$  "pudełek", przy czym na każdym etapie "pudełko" o największej wariancji barwy jest dzielone wzdłuż osi, która minimalizuje sumę wariancji po obu stronach. Autor algorytmu skutecznie rozwiązał problem jak obliczać statystyki barw (średnie i wariancje) dla tych "pudełek". Średnie barwy w ostatecznych "pudełkach" stają się elementami palety barw. Inne szczegóły dotyczące algorytmu można znaleźć w pracy [18].

Wszystkie algorytmy stosowane w proponowanej metodzie zostały zaimplementowane w języku Visual C++, a eksperymenty obliczeniowe przeprowadzono na komputerze PC z systemem operacyjnym Windows, procesorem Intel Core i7 920@2,67GHz i 8 GB pamięci RAM.

#### 4. Wyniki eksperymentów obliczeniowych

W niniejszej pracy dla każdego obrazu testowego przeprowadzono kwantyzację na 32, 64, 128 i 256 barw, obliczono dla obrazów skwantowanych wskaźniki jakości MSE, DSCSI [10] i HPSI [15] oraz zmierzono czasy pracy procesora. Generowanie palety barw miało miejsce na wszystkich pikselach obrazu (współczynnik podpróbkowania 1/1) oraz na ośmiu podzbiorach pikseli uzyskanych metodą NNI z oryginalnych obrazów, gdzie współczynnik podpróbkowania wynosił odpowiednio: 1/2, 1/4, 1/8, 1/16, 1/32, 1/64, 1/128 i 1/256. W ten sposób dla każdego obrazu testowego uzyskano po dziewięć wyników kwantyzacji dla każdej z 4 liczb barw.

Tabela 2

Wartości MSE dla kwantyzacji na 256 barw

| <i>cf, df</i> | NNI    | CM2 | CM4 | NNI  | CM2 | CM4 | NNI     | CM2 | CM4 | NNI   | CM2 | CM4 |
|---------------|--------|-----|-----|------|-----|-----|---------|-----|-----|-------|-----|-----|
|               | Baboon |     |     | Lena |     |     | Peppers |     |     | Pills |     |     |
| 1             | 99     | 97  | 97  | 31   | 30  | 30  | 54      | 54  | 54  | 42    | 41  | 41  |
| 1/2           | 100    | 101 | 107 | 31   | 31  | 33  | 54      | 56  | 59  | 42    | 42  | 44  |
| 1/4           | 100    | 103 | 109 | 31   | 32  | 34  | 54      | 57  | 61  | 42    | 43  | 45  |
| 1/8           | 101    | 105 | 114 | 31   | 33  | 35  | 55      | 58  | 63  | 42    | 44  | 47  |
| 1/16          | 102    | N/A | N/A | 32   | N/A | N/A | 56      | N/A | N/A | 42    | N/A | N/A |
| 1/32          | 105    | N/A | N/A | 32   | N/A | N/A | 57      | N/A | N/A | 43    | N/A | N/A |
| 1/64          | 108    | N/A | N/A | 34   | N/A | N/A | 59      | N/A | N/A | 45    | N/A | N/A |
| 1/128         | 113    | N/A | N/A | 35   | N/A | N/A | 64      | N/A | N/A | 47    | N/A | N/A |
| 1/256         | 119    | N/A | N/A | 38   | N/A | N/A | 68      | N/A | N/A | 50    | N/A | N/A |

Tabela 3

Czasy przetwarzania dla 32 barw

| <i>cf, df</i> | NNI    | CM2  | CM4 | NNI  | CM2  | CM4 | NNI     | CM2  | CM4 | NNI   | CM2  | CM4 |
|---------------|--------|------|-----|------|------|-----|---------|------|-----|-------|------|-----|
|               | Baboon |      |     | Lena |      |     | Peppers |      |     | Pills |      |     |
| 1             | 345    | 443  | 443 | 124  | 477  | 477 | 154     | 474  | 474 | 828   | 647  | 647 |
| 1/2           | 155    | 1370 | 116 | 62   | 1499 | 121 | 79      | 1327 | 102 | 134   | 3201 | 233 |
| 1/4           | 201    | 696  | 55  | 40   | 841  | 64  | 50      | 684  | 52  | 49    | 1634 | 122 |
| 1/8           | 45     | 368  | 57  | 20   | 400  | 31  | 30      | 356  | 28  | 52    | 825  | 68  |
| 1/16          | 15     | N/A  | N/A | 21   | N/A  | N/A | 12      | N/A  | N/A | 16    | N/A  | N/A |
| 1/32          | 9      | N/A  | N/A | 13   | N/A  | N/A | 4       | N/A  | N/A | 11    | N/A  | N/A |
| 1/64          | 6      | N/A  | N/A | 4    | N/A  | N/A | 2       | N/A  | N/A | 5     | N/A  | N/A |
| 1/128         | 2      | N/A  | N/A | 4    | N/A  | N/A | 1       | N/A  | N/A | 3     | N/A  | N/A |
| 1/256         | 2      | N/A  | N/A | 2    | N/A  | N/A | 1       | N/A  | N/A | 3     | N/A  | N/A |

Pierwszy test został przeprowadzony na czterech obrazach: Baboon, Lena, Peppers i Pills pokazanych na Rys. 3, aby porównać wyniki z metodą opartą na próbkowaniu rdzeniowym [19]. Pierwsze trzy obrazy mają rozdzielczość  $512 \times 512$  pikseli, a obraz Pills -  $800 \times 519$  pikseli.

W Tabelach 1-4 przedstawiono porównanie średnich wartości MSE oraz czasów kwantyzacji dla algorytmu opartego na rdzeniach oraz dla proponowanej nowej metody oznaczonej jako NNI w przypadku odpowiednio 32 i 256 barw.

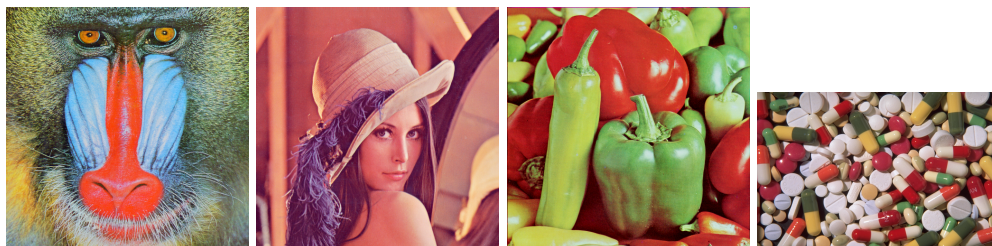
Na Rys. 4 pokazano uzyskane przyspieszenia obliczeń w zależności od wartości współczynnika podpróbkowania dla  $k = 256$ , gdzie przewaga proponowanej metody jest znacząca.

Drugi test został przeprowadzony na 24 obrazach z zestawu obrazów Kodaka [9],

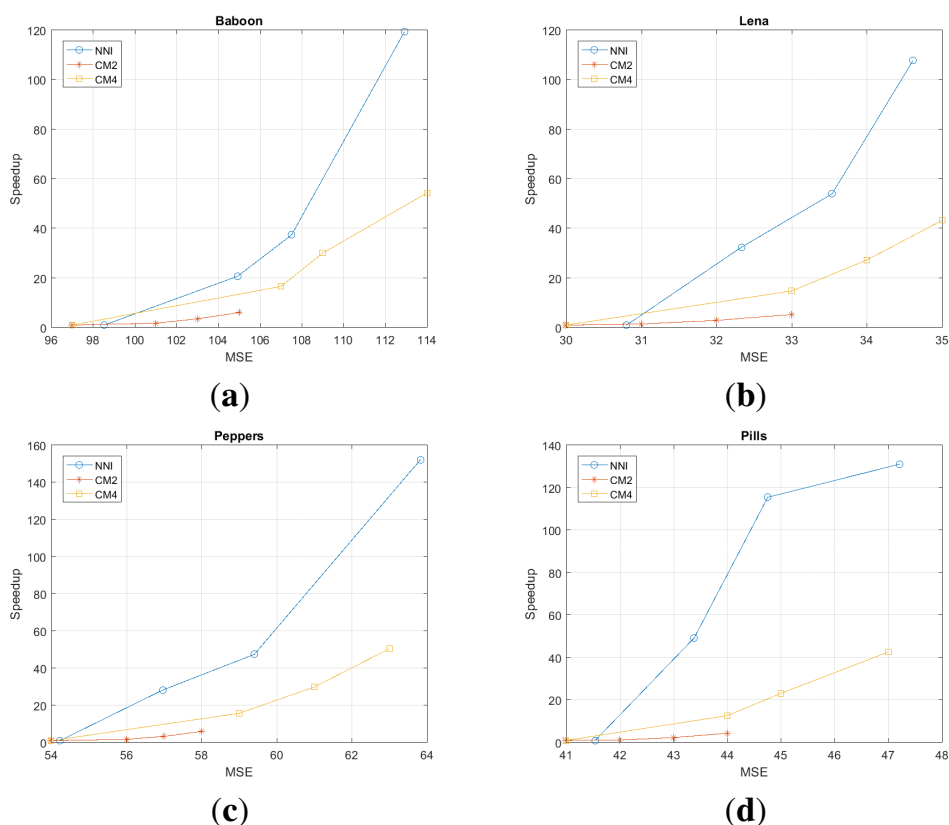
Tabela 4

Czasy przetwarzania dla 256 barw

| <i>cf, df</i> | NNI    | CM2  | CM4  | NNI  | CM2  | CM4  | NNI     | CM2  | CM4  | NNI   | CM2  | CM4  |
|---------------|--------|------|------|------|------|------|---------|------|------|-------|------|------|
|               | Baboon |      |      | Lena |      |      | Peppers |      |      | Pills |      |      |
| 1             | 1260   | 3152 | 3152 | 1124 | 3026 | 3026 | 1390    | 2929 | 2929 | 3223  | 4606 | 4606 |
| 1/2           | 526    | 1796 | 190  | 339  | 2137 | 205  | 564     | 1673 | 186  | 875   | 3830 | 364  |
| 1/4           | 368    | 898  | 105  | 281  | 1047 | 111  | 572     | 890  | 98   | 612   | 1989 | 199  |
| 1/8           | 209    | 522  | 58   | 143  | 575  | 70   | 144     | 490  | 58   | 308   | 1046 | 108  |
| 1/16          | 93     | N/A  | N/A  | 55   | N/A  | N/A  | 113     | N/A  | N/A  | 243   | N/A  | N/A  |
| 1/32          | 61     | N/A  | N/A  | 35   | N/A  | N/A  | 49      | N/A  | N/A  | 65    | N/A  | N/A  |
| 1/64          | 34     | N/A  | N/A  | 21   | N/A  | N/A  | 29      | N/A  | N/A  | 28    | N/A  | N/A  |
| 1/128         | 11     | N/A  | N/A  | 10   | N/A  | N/A  | 9       | N/A  | N/A  | 25    | N/A  | N/A  |
| 1/256         | 5      | N/A  | N/A  | 4    | N/A  | N/A  | 4       | N/A  | N/A  | 7     | N/A  | N/A  |



Rys. 3. Obrazy testowe: (a) Baboon, (b) Lena, (c) Peppers, (d) Pills



Rys. 4. Przyspieszenie vs. MSE dla 256 barw

gdzie są obrazy o rozdzielczości  $768 \times 512$  lub  $512 \times 768$  pikseli pokazane na Rys. 5. Ponadto w zestawie tym dostępne są również obrazy wysokiej rozdzielczości  $3072 \times 2048$  lub  $2048 \times 3072$  pikseli.

Na Rys. 6-7 przedstawiono wykresy pudełkowe uzyskane dla wszystkich obrazów z zestawu obrazów Kodaka i dla  $k=32, 64, 128$  i  $256$ . Na Rys. 6 przedstawiono wykresy dla klasycznego wskaźnika jakości MSE, natomiast Rys. 7 zawiera oceny uzyskane dla nowego wskaźnika jakości HPSI. Jak widać, spadek jakości obrazu związany z próbkowaniem do poziomu  $1/128$  jest niewielki, natomiast zmiana czasu kwantyzacji jest znacząca, co pokazano w Tabeli 5. Taka sytuacja występuje dla trzech badanych wskaźników jakości obrazu.

Maksymalne przyspieszenie uzyskane przy kwantyzacji 24 obrazów wyniosło 357 w przypadku współczynnika próbkowania równego  $1/256$  i palety 256 barw. W tym przypadku jednak wartości wskaźników jakości uległy wyraźnemu pogorszeniu. Z drugiej strony, takiego spadku jakości nie zaobserwowano dla współczynnika próbkowania



Rys. 5. Zestaw obrazów Kodaka

Tabela 5

Czasy przetwarzania (w ms) dla zestawu obrazów Kodaka

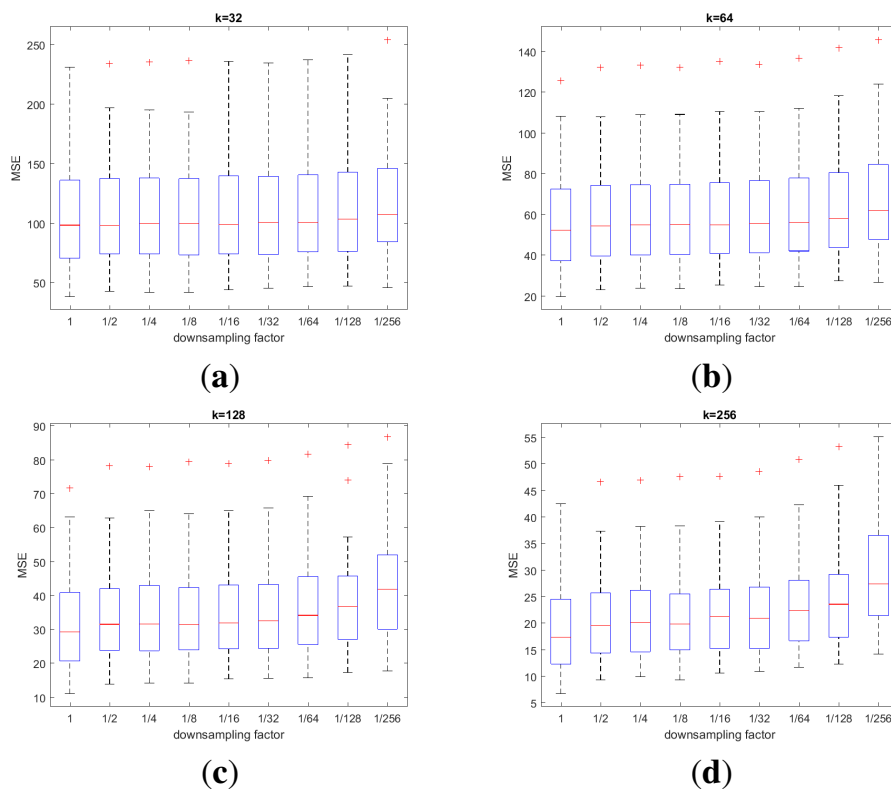
| $df$  | $k$ |     |      |      |
|-------|-----|-----|------|------|
|       | 32  | 64  | 128  | 256  |
| 1     | 361 | 650 | 1465 | 2500 |
| 1/2   | 161 | 366 | 643  | 1106 |
| 1/4   | 80  | 177 | 347  | 565  |
| 1/8   | 46  | 76  | 165  | 298  |
| 1/16  | 22  | 44  | 75   | 134  |
| 1/32  | 11  | 21  | 37   | 68   |
| 1/64  | 6   | 11  | 19   | 32   |
| 1/128 | 3   | 5   | 7    | 15   |
| 1/256 | 2   | 3   | 4    | 7    |

równego 1/128, gdzie maksymalne przyspieszenie wyniosło 167.

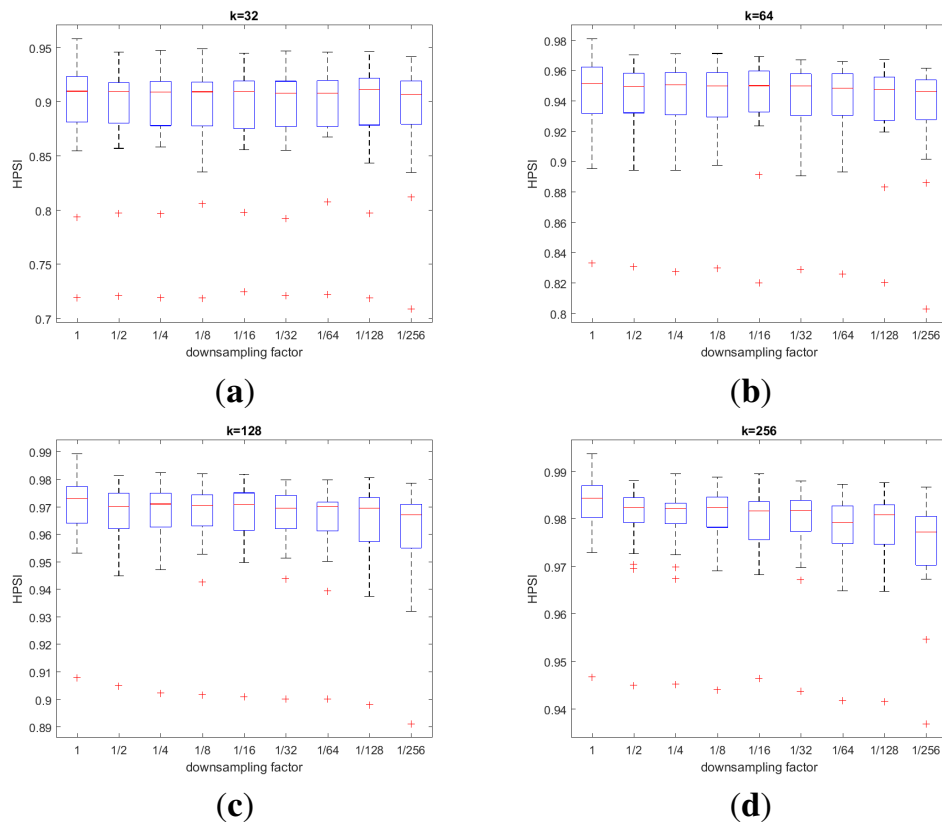
Inne szczegóły dotyczące opracowanej metody i dodatkowe wyniki badań można znaleźć w publikacji [5].

## 5. Podsumowanie

W niniejszej pracy wprowadzono ulepszoną metodę kwantyzacji barwy, która znacznie przyspieszyła kwantyzację (niekiedy kilkaset razy), nie obniżając przy tym znacząco jakości otrzymanego obrazu. Należy zaznaczyć, że do oceny skwantowanego obrazu użyto nie tylko wskaźnika jakości MSE, ale również wskaźników DSCSI i HPSI, silnie związanych z percepcją wzrokową obserwatora. Usprawnienie działania metody kwantyzacji polegało na zastosowaniu podpróbki obrazu wejściowego przez interpolację typu NNI. W przyszłości można analizować przydatność tego podejścia do innych, jeszcze bardziej czasochłonnych technik klasteryzacji, takich jak *Fuzzy C-Means*



Rys. 6. Wartości MSE dla zestawu obrazów Kodaka z różnymi współczynnikami podpróbkiwania



Rys. 7. Wartości HPSI dla zestawu obrazów Kodaka z różnymi współczynnikami podpróbkiwania



(FCM), czy *K-Harmonic Means* (KHM).

### Podziękowania

Praca powstała przy wsparciu finansowym Ministerstwa Edukacji i Nauki dla Politechniki Śląskiej w ramach grantów badawczych: 02/010/BK-19/0143 oraz 02/070/BK\_22/0035.

### LITERATURA

1. Ahn, J., Park, J., Park, D., Paek, J., and Ko, J. Convolutional neural network-based classification system design with compressed wireless sensor network images. *PLoS one* 13, 5 (2018), e0196251.
2. Arthur, D., and Vassilvitskii, S. k-means++: The advantages of careful seeding. In *Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms* (2007), Society for Industrial and Applied Mathematics, pp. 1027–1035.
3. Bejarano, J., Bose, K., Brannan, T., Thomas, A., Adragani, K., Neerchal, N. K., and Ostrouchov, G. Sampling within k-means algorithm to cluster large datasets. *UMBC Student Collection, University of Maryland, Baltimore County* (2011).
4. Celebi, M. E. *Partitional Clustering Algorithms*. Springer, 2014.
5. Frackiewicz, M., Mandrella, A., and Palus, H. Fast color quantization by k-means clustering combined with image sampling. *Symmetry* 11, 8 (2019), 963.
6. Gerstner, T., DeCarlo, D., Alexa, M., Finkelstein, A., Gingold, Y., and Nealen, A. Pixelated image abstraction with integrated user constraints. *Computers & Graphics* 37, 5 (2013), 333–347.
7. Heckbert, P. Color image quantization for frame buffer display. *ACM SIG-GRAPH Computer Graphics* 16, 3 (1982), 297–307.
8. Kanungo, T., Mount, D. M., Netanyahu, N. S., Piatko, C. D., Silverman, R., and Wu, A. Y. An efficient k-means clustering algorithm: Analysis and implementation. *IEEE Transactions on Pattern Analysis & Machine Intelligence* 7 (2002), 881–892.
9. Kodak images. <http://r0k.us/graphics/kodak/>. Accessed in July, 2022.
10. Lee, D., and Plataniotis, K. N. Towards a full-reference quality assessment for color images using directional statistics. *Image Processing, IEEE Transactions on* 24, 11 (2015), 3950–3965.
11. Lucic, M., Bachem, O., and Krause, A. Strong coresets for hard and soft Bregman clustering with applications to exponential family mixtures. *arXiv preprint arXiv:1508.05243* (2015).

12. MacQueen, J. Some methods for classification and analysis of multivariate observations. In *Proceedings of the 5th Berkeley Symposium on Mathematics, Statistics, and Probabilities, Berkeley, USA* (1967), pp. 281–297.
13. Palus, H., and Frackiewicz, M. New approach for initialization of k-means technique applied to color quantization. In *Information Technology (ICIT), 2010 2nd International Conference on* (2010), IEEE, pp. 205–209.
14. Papamarkos, N., Atsalakis, A. E., and Strouthopoulos, C. P. Adaptive color reduction. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* 32, 1 (2002), 44–56.
15. Reisenhofer, R., Bosse, S., Kutyniok, G., and Wiegand, T. A Haar wavelet-based perceptual similarity index for image quality assessment. *arXiv preprint arXiv:1607.06140* (2016).
16. Sinha, P. K. *Image acquisition and preprocessing for machine vision systems*. SPIE, 2012.
17. Szilágyi, L., Dénesi, G., and Szilágyi, S. M. Fast color reduction using approximative c-means clustering models. In *2014 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)* (2014), IEEE, pp. 194–201.
18. Timmerman, B. Wu’s Color Quantizer. <https://gist.github.com/bert/1192520>. Accessed in July, 2022.
19. Valenzuela, G., Celebi, M. E., and Schaefer, G. Color quantization using coresets sampling. In *2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC)* (2018), IEEE, pp. 2096–2101.
20. Wu, X. Efficient statistical computations for optimal color quantization. In *Graphic Gems II*, J. Arvo, Ed. Academic Press, New York, USA, 1991, pp. 126–133.