

Krzysztof OKARMA

ZACHODNIOPOMORSKI UNIWERSYTET TECHNOLOGICZNY W SZCZECINIE

Optimalizacja korelacji szybkiej metody estymacji jakości obrazów opartej na podobieństwie strukturalnym z oceną subiektywną

Dr inż. Krzysztof OKARMA

Adiunkt w Katedrze Przetwarzania Sygnałów i Inżynierii Multimedialnej Zachodniopomorskiego Uniwersytetu Technologicznego w Szczecinie. Absolwent Wydziału Elektrycznego (1999) i Informatyki (2001) Politechniki Szczecińskiej, stopień doktora nauk technicznych uzyskał w 2003 r. na Wydziale Elektrycznym, gdzie od 2008 r. pełni funkcję Kierownika Studiów Niestacjonarnych. Autor ponad 60 publikacji z zakresu cyfrowego przetwarzania oraz analizy obrazów i sygnałów.



e-mail: okarma@zut.edu.pl

Streszczenie

W pracy przedstawiono analizę wyników estymacji jakości obrazów przy użyciu metody Monte Carlo w połączeniu z podobieństwem strukturalnym (SSIM) pod kątem korelacji z oceną subiektywną. Jako platformy testowej użyto bazy obrazów poddanych pięciu typowym rodzajom zniekształceń wraz z ich ocenami subiektywnymi udostępnianej przez Laboratorium Inżynierii Obrazu i Wideo (LIVE) Uniwersytetu w Teksasie. Uzyskane wyniki korelacji liniowej oraz z zastosowaniem funkcji logistycznej potwierdzają wysoką przydatność zaproponowanej szybkiej metody estymacji jakości obrazu.

Słowa kluczowe: ocena jakości obrazów cyfrowych, podobieństwo strukturalne, metoda Monte Carlo.

Optimisation of the correlation of the fast image quality estimation based on the Structural Similarity with the subjective evaluation

Abstract

The paper presents the analysis of the results of the image quality estimation when using the Monte Carlo method applied to the Structural Similarity index [5] in the aspect of correlation with the subjective evaluation. As a test platform, the LIVE database (available at the Laboratory for Image & Video Engineering of the University of Texas at Austin) [7] containing the Differential Mean Opinion Score (DMOS) values for almost 1000 images corrupted by five common types of distortions has been used. After the short introduction containing the analysis of some modern image quality assessment techniques, the idea of the image quality fast estimation with use of the Monte Carlo Structural Similarity Index [6] is described. The results of the Pearson's correlation of the DMOS values and MC-SSIM for the individual types of distortions and for the whole database, excluding the original images, are presented in Tables 1-2. Tables 3-4 illustrate the optimised results obtained after the nonlinear mapping with use of the logistic function (6). The obtained results of the linear correlation, as well as those obtained when using the nonlinear mapping by the optimised logistic function, confirm the usefulness of the proposed fast image quality estimation technique. An interesting fact is not only the preservation of the high correlation coefficients even for the small number of pixels used for the quality estimation but also a slight increase in them which may be an interesting stimulus for further research.

Keywords: digital image quality assessment, Structural Similarity, Monte Carlo method.

1. Wstęp

Automatyczna ocena jakości obrazów cyfrowych stanowi obecnie jedną z najdynamiczniej rozwijających się gałęzi komputerowej analizy obrazów. Przyczyną takiego stanu rzeczy jest duża przydatność takich uniwersalnych miar jakości obrazów zarówno do obiektywnej oceny efektywności technik stratnej kompresji obrazów, jak również algorytmów ich przetwarzania, w szczegól-

ności filtracji, a także badanie korelacji takich wskaźników np. ze współczynnikami poprawności klasyfikacji w zadaniach związanych z rozpoznawaniem wzorców.

2. Techniki obiektywnej oceny jakości obrazów cyfrowych

Klasyczne podejście do obiektywnej oceny jakości obrazów cyfrowych oparte jest na zastosowaniu błędu średniokwadratowego i miar na nim opartych [1, 2]. Przykładami takich miar mogą być szczytowy stosunek sygnału do szumu (ang. *Peak Signal to Noise Ratio*), znormalizowany błąd średniokwadratowy (*Normalised Mean Squared Error*), czy też średni błąd bez-względny (*Mean Absolute Error*). Wspólną wadą wszystkich tych miar jest ich słaba korelacja z wynikami oceny subiektywnej dokonywanej przez obserwatorów. Ponadto warto zwrócić uwagę na często różny zakres dynamiczny poszczególnych wskaźników, co także można uznać za czynnik utrudniający interpretację uzyskiwanych wyników, a także na fakt uzyskiwania czasem diametralnie różnych wartości dla różnych obrazów poddawanych podobnym zniekształceniom.

Zdecydowana większość spośród tradycyjnych metod oceny jakości obrazów cyfrowych należy do grupy metod porównawczych bazujących na założeniu dokładnej znajomości obrazu oryginalnego. Trudno to jednak uznać za znaczącą wadę ze względu na małą ilość metod tzw. "ślepej" (nie wymagającej znajomości obrazu bez zniekształceń) oceny jakości [3], które są z reguły mało uniwersalne, ograniczając się do wrażliwości na jeden lub dwa rodzaje zniekształceń.

Osobnym zagadnieniem, nie rozwiązany dotąd w sposób jednoznaczny i kompletny, pozostaje technika oceny jakości obrazów kolorowych. Zdecydowana większość, także nowoczesnych, metod oceny jakości obrazów definiowana jest dla obrazów w skali szarości, co często stawia pod znakiem zapytania celowość ich stosowania, szczególnie dla oceny efektów działania metod przetwarzania obrazów dedykowanych dla obrazów o pełnym zakresie barw. Przykładem mogą być algorytmy wektorowej filtracji medianowej, czy też wektorowej filtracji kierunkowej.

Pomimo tych wad wielu autorów publikacji wciąż używa metod klasycznych, także w zastosowaniach związanych z oceną jakości obrazów kolorowych, często ograniczając analizę jakości do luminancji, czasem uzupełniając wyniki o wartości znormalizowanej różnicy barw (*Normalised Colour Difference*) zdefiniowanej w przestrzeni CIE L*a*b*. Uzasadnieniem takiego stanu rzeczy może być mała złożoność obliczeniowa klasycznych miar jakości, jak również zdecydowanie większa czułość oka ludzkiego na zmiany luminancji niż chrominancji, co w pewnym stopniu może stanowić usprawiedliwienie w niektórych zastosowaniach. Biorąc pod uwagę porównawczy charakter analizy wyników dla szybkiej estymacji jakości przedstawionej w niniejszym artykule, również ograniczono się do analizy obrazów przekonwertowanych do skali szarości, pozostawiając aspekty związane z szybką oceną jakości obrazów kolorowych jako pole do dalszych badań.

Znaczący postęp w zakresie opracowania nowych metod oceny jakości obrazów nastąpił na początku obecnej dekady, czego przykładem może być tzw. uniwersalny wskaźnik jakości obrazu (*Universal Image Quality Index*) zaproponowany w roku 2002 [4] zdefiniowany jako iloczyn trzech czynników:

$$Q = \frac{2\mu_x\mu_y}{\mu_x^2 + \mu_y^2} \cdot \frac{2\sigma_x\sigma_y}{\sigma_x^2 + \sigma_y^2} \cdot \frac{\sigma_{xy}}{\sigma_x\sigma_y} = \frac{4\mu_x\mu_y\sigma_{xy}}{(\mu_x^2 + \mu_y^2)(\sigma_x^2 + \sigma_y^2)}, \quad (1)$$

odpowiadających trzem typowym rodzajom zniekształceń (zmiana jasności, utrata kontrastu oraz zmiana struktury obrazu), gdzie poszczególne elementy zdefiniowane są następująco:

$$\mu_x = \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N x_{ij}, \quad \mu_y = \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N y_{ij}, \quad (2)$$

$$\sigma_x^2 = \frac{1}{N^2 - 1} \sum_{i=1}^N \sum_{j=1}^N (x_{ij} - \mu_x)^2, \quad \sigma_y^2 = \frac{1}{N^2 - 1} \sum_{i=1}^N \sum_{j=1}^N (y_{ij} - \mu_y)^2, \quad (3)$$

$$\sigma_{xy} = \frac{1}{N^2 - 1} \sum_{i=1}^N \sum_{j=1}^N (x_{ij} - \mu_x)(y_{ij} - \mu_y). \quad (4)$$

Ze względu na potencjalną niestabilność wyrażenia (1) zwłaszcza dla ciemnych obszarów obrazu o jednolitym kolorze spowodowaną małą wartością jednego z mianowników zastosowane zostały dodatkowe stałe współczynniki zapobiegające dzieleniu przez liczby bliskie zeru. W efekcie uzyskana została, przedstawiona w 2004 roku [5], zależność określająca podobieństwo strukturalne (*Structural Similarity*) dwóch fragmentów obrazów:

$$SSIM = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}, \quad (5)$$

przy czym wartości stałych proponowane przez twórców tej miary wynoszą $C_1 = 0.01 \cdot L$ oraz $C_2 = 0.03 \cdot L$, gdzie L oznacza zakres dynamiczny dla wartości pikseli (typowo $L = 255$ dla obrazów 8-bitowych).

Efektem zastosowania zależności (8) dla fragmentów obrazu określonych przez przesuwne okno o wymiarach $N \times N$ pikseli jest mapa jakości obrazu o rozdzielczości mniejszej od obrazu o $N-1$ wierszy oraz $N-1$ kolumn. Zalecanym przez twórców wskaźnika SSIM oknem jest dwuwymiarowe okno Gaussa o rozmiarze 11×11 pikseli, przy czym w szczególnym przypadku stosowania wskaźnika (1) stosuje się okno prostokątne dla $N = 8$.

3. Zastosowanie metody Monte Carlo do estymacji jakości obrazów cyfrowych

Metoda Monte Carlo, jako jedna z najbardziej powszechnych metod statystycznych mających na celu redukcję ilości analizowanych punktów może być z powodzeniem zastosowana do celów estymacji jakości obrazów cyfrowych. Jak łatwo przewidzieć, losowe ograniczenie ilości analizowanych pikseli obrazu, ma istotny wpływ na dokładność estymacji w stosunku do wyników uzyskanych po analizie całego obrazu, co może być zaobserwowane zarówno przy użyciu tradycyjnych miar jakości, ja również podobieństwa strukturalnego. Warunkiem uzyskania stosunkowo małych błędów względnych przy niewielkiej liczbie analizowanych próbek są dobre właściwości statystyczne generatora liczb pseudolosowych użytego do losowania próbek obrazu poddawanych analizie przy użyciu określonej metody oceny jakości [6].

Oszacowanie błędów względnych estymacji określonej miary jakości, w szczególności SSIM, nie pozwala na dokonanie pełnej analizy efektywności zastosowania metody Monte Carlo chociażby ze względu na fakt, iż aktualnie nie jest znana żadna "doskonała" metoda oceny jakości, która byłaby idealnie skorelowana ze sposobem postrzegania obrazu przez oko ludzkie. Wszystkie znane porównawcze metody oceny jakości obrazów są w zasadzie estymatorami stopnia podobieństwa obrazu oryginalnego i zniekształconego, dlatego też bardziej celowe jest odniesienie uzyskanych wyników do rezultatów oceny subiektywnej, aniżeli do wyników uzyskiwanych zblizoną metodą.

Wiarygodne określenie korelacji wyników obiektywnej oceny jakości obrazów z oceną subiektywną wymaga wykorzystania bazy danych zawierającej możliwie dużą ilość obrazów z różnymi rodzajami zniekształceń oraz ich ocen subiektywnych dokonanych

przez możliwie największą obserwatorów. Niewątpliwie najlepszą i najczęściej wykorzystywaną taką bazę stanowi tzw. *LIVE Image Quality Database* [7] udostępniona przez Uniwersytet w Teksasie, która zawiera 29 obrazów z pięcioma rodzajami zniekształceń (z różną intensywnością bądź różnym współczynnikiem kompresji): kompresją stratną JPEG oraz JPEG2000, rozmyciem Gaussa, szumem białym dla składowych RGB oraz z błędami bitowymi w strumieniu JPEG2000 po transmisji przez symulowany kanał z zanikami Rayleigha. W efekcie siedmiokrotnej oceny każdego obrazu przez kilkanaście do kilkudziesięciu osób uzyskano 982 średnie oceny różnicowe (*Differential Mean Opinion Scores*) znormalizowane do zakresu 0–100 (w tym 203 oceny dla 29 obrazów oryginalnych).

4. Procedura testowa

W celu weryfikacji przydatności zaproponowanej w artykule metody przyspieszenia obliczeń opartej na metodzie Monte Carlo przeprowadzono po kilkanaście losowań ustalonej liczby próbek z obrazów, na podstawie których wyznaczono wartości wskaźnika podobieństwa strukturalnego SSIM. Jako generatora liczb pseudolosowych użyto wbudowanej funkcji *rand* z pakietu MATLAB. Wyniki uzyskane dla ustalonej ilości wylosowanych pikseli zostały następnie uśrednione, po czym zostały wyznaczone współczynniki korelacji liniowej Pearsona dla poszczególnych pięciu rodzajów zniekształceń oraz dla zestawu wszystkich zniekształconych obrazów. Warto zwrócić uwagę, iż liczba wylosowanych punktów została potraktowana jako wartość stała niezależna od rozdzielczości ocenianego obrazu, a nie jako procent ilości pikseli.

Takie podejście zostało podcytkowane faktem, iż celem badań nie było określenie dokładnego przebiegu błędu czy też korelacji z wartościami DMOS w funkcji ilości analizowanych próbek, a jedynie weryfikacja przydatności zaproponowanej metody i przybliżona ocena zmiany korelacji spowodowanej zmniejszeniem ilości analizowanych danych.

Pomimo, iż spośród 29 oryginalnych obrazów w bazie LIVE wyróżnić można obraz o siedmiu różnych rozdzielczościach, można uznać, iż ich rozmiary są zbliżone (najmniejszy obraz ma rozdzielczość 634×438 a największy 768×512 pikseli), szczególnie iż zaproponowana metoda służyć powinna do radykalnego obniżenia ilości analizowanych próbek obrazu i znaczącego przyspieszenia obliczeń (użycie nawet 10 000 próbek oznacza prawie 30-krotne zmniejszenie ich liczby dla obrazów o najniższej rozdzielczości).

Wszystkie obliczenia przeprowadzono dla obrazów przekonwertowanych do skali szarości: jako luminancja Y zgodnie z "televizyjnymi" modelami barw YUV oraz YIQ, a także jako średnia arytmetyczna z trzech wartości dla kanałów RGB. Współczynniki korelacji liniowej uzyskane dla poszczególnych rodzajów zniekształceń oraz różnych ilości losowo wybranych pikseli zostały przedstawione odpowiednio w tabelach 1 oraz 2.

Warto podkreślić, iż współczynniki te były obliczane bez uwzględniania wskaźników SSIM i ocen obrazów oryginalnych, które stanowią ponad 20% wyników w bazie LIVE, co powoduje, iż wyniki te są niższe aniżeli obliczane dla całej bazy prezentowane w publikacjach innych autorów. Jednakże takie podejście można uznać za w pełni uzasadnione ze względu na przewidywalność wyników uzyskiwanych dla obrazów oryginalnych oraz ich znaczący wpływ na wartości poszczególnych współczynników korelacji liniowej obliczanych z uwzględnieniem tych obrazów.

Ze względu na nieliniowość charakterystyki wzroku ludzkiego oraz nieliniowy charakter zależności postrzeganej jakości obrazu od ilości i rodzaju zniekształceń, zgodnie z wytycznymi organizacji VQEG [8] zastosowano dodatkowo mapowanie nieliniowe wartości wskaźnika SSIM za pomocą funkcji logistycznej postaci:

$$SSIM2 = \beta_1 \cdot \left(\frac{1}{2} - \frac{1}{1 + \exp(\beta_2 \cdot (SSIM - \beta_3))} \right) + \beta_4 \cdot SSIM + \beta_5. \quad (6)$$

Tab. 1. Współczynniki korelacji liniowej uzyskanych wyników z oceną subiektywną (wartościami DMOS) dla luminancji Y

Tab. 1. Linear correlation coefficients of the obtained results with the subjective evaluation (DMOS values) for the Y luminance

Zniekształcenie / ilość pikseli	JPEG 2000	JPEG	Szum biały	Rozmycie Gaussa	Symulowany kanał z zanikami Rayleigha	Wszystkie
N=50	0,8825	0,8592	0,9625	0,8405	0,8938	0,7350
N=100	0,9021	0,8507	0,9639	0,8611	0,9052	0,7455
N=200	0,8963	0,8489	0,9659	0,8447	0,9018	0,7416
N=500	0,9014	0,8533	0,9639	0,8520	0,9037	0,7414
N=1000	0,8965	0,8483	0,9641	0,8496	0,9005	0,7365
N=5000	0,8969	0,8510	0,9642	0,8476	0,9006	0,7364
N=10000	0,8980	0,8503	0,9644	0,8489	0,9009	0,7363
Cały obraz	0,8974	0,8503	0,9644	0,8486	0,9008	0,7364

Tab. 2. Współczynniki korelacji liniowej uzyskanych wyników z oceną subiektywną (wartościami DMOS) dla średniej z kanałów RGB

Tab. 2. Linear correlation coefficients of the obtained results with the subjective evaluation (DMOS values) for the RGB average

Zniekształcenie / ilość pikseli	JPEG 2000	JPEG	Szum biały	Rozmycie Gaussa	Symulowany kanał z zanikami Rayleigha	Wszystkie
N=50	0,8808	0,8594	0,9667	0,8359	0,8929	0,7553
N=100	0,9025	0,8548	0,9672	0,8594	0,9056	0,7662
N=200	0,8969	0,8529	0,9695	0,8412	0,9018	0,7627
N=500	0,9022	0,8561	0,9677	0,8497	0,9038	0,7624
N=1000	0,8967	0,8509	0,9678	0,8471	0,9009	0,7575
N=5000	0,8972	0,8538	0,9680	0,8447	0,9007	0,7575
N=10000	0,8984	0,8533	0,9682	0,8461	0,9009	0,7576
Cały obraz	0,8977	0,8532	0,9681	0,8457	0,9008	0,7575

Współczynniki korelacji Pearsona obliczone dla wartości wskaźnika SSIM2 uzyskanych w wyniku zastosowania mapowania nieliniowego po optymalizacji współczynników $\beta_1 - \beta_5$ z wartościami ocen subiektywnych DMOS przedstawione zostały w tabelach 3 oraz 4.

5. Analiza uzyskanych wyników

Jak łatwo zaobserwować, nawet poważne ograniczenie liczby analizowanych pikseli pozwala zachować wysoką korelację uzyskanych wyników z ocenami subiektywnymi, co można uznać za istotną zaletę rozważanej metody. Oczywiście dalsze ograniczanie ilości analizowanych próbek prowadziłyby do zauważalnego pogorszenia zgodności i stabilności uzyskiwanych wyników, szczególnie dla zniekształceń o charakterze lokalnym. Z tego też względu zalecana liczba analizowanych próbek nie powinna być traktowana jako określony procent ze wszystkich pikseli obrazu, szczególnie przy ocenie jakości obrazów o niższych rozdzielczościach, gdyż mogłoby to prowadzić do nadmiernego ograniczenia ilości analizowanych próbek.

Zastosowanie optymalizacji korelacji z użyciem funkcji logistycznej pozwoliło na jej znaczące podwyższenie, w szczególności dla oceny uniwersalnej dotyczącej wszystkich pięciu analizowanych rodzajów zniekształceń obrazu.

Rozważana w artykule metoda szybkiej estymacji jakości obrazów cyfrowych może być szczególnie przydatna w systemach o stosunkowo niskiej wydajności, w szczególności w aplikacjach, gdzie niska złożoność obliczeniowa jest elementem kluczowym, znacznie istotniejszym od dokładności estymacji jakości obrazu.

Tab. 3. Optymalizowane współczynniki korelacji liniowej uzyskanych wyników z oceną subiektywną (wartościami DMOS) po mapowaniu nieliniowym z użyciem funkcji logistycznej dla luminancji Y

Tab. 3. Optimised linear correlation coefficients of the obtained results with the subjective evaluation (DMOS values) after the nonlinear mapping using the logistic function for the Y luminance

Zniekształcenie / ilość pikseli	JPEG 2000	JPEG	Szum biały	Rozmycie Gaussa	Symulowany kanał z zanikami Rayleigha	Wszystkie
N=50	0,9297	0,9281	0,9772	0,8695	0,9389	0,8608
N=100	0,9391	0,9335	0,9782	0,8829	0,9473	0,8684
N=200	0,9326	0,9238	0,9810	0,8704	0,9432	0,8626
N=500	0,9371	0,9301	0,9784	0,8729	0,9445	0,8646
N=1000	0,9344	0,9294	0,9791	0,8722	0,9432	0,8621
N=5000	0,9357	0,9295	0,9791	0,8724	0,9441	0,8630
N=10000	0,9366	0,9294	0,9777	0,8734	0,9440	0,8631
Cały obraz	0,9364	0,9294	0,9651	0,8734	0,9449	0,8632

Tab. 4. Optymalizowane współczynniki korelacji liniowej uzyskanych wyników z oceną subiektywną (wartościami DMOS) po mapowaniu nieliniowym z użyciem funkcji logistycznej dla średniej z kanałów RGB

Tab. 4. Optimised linear correlation coefficients of the obtained results with the subjective evaluation (DMOS values) after the nonlinear mapping using the logistic function for the RGB average

Zniekształcenie / ilość pikseli	JPEG 2000	JPEG	Szum biały	Rozmycie Gaussa	Symulowany kanał z zanikami Rayleigha	Wszystkie
N=50	0,9285	0,9278	0,9785	0,8647	0,9350	0,8730
N=100	0,9383	0,9342	0,9788	0,8805	0,9453	0,8811
N=200	0,9316	0,9250	0,9813	0,8668	0,9414	0,8756
N=500	0,9370	0,9304	0,9792	0,8705	0,9427	0,8777
N=1000	0,9337	0,9295	0,9798	0,8692	0,9410	0,8751
N=5000	0,9350	0,9296	0,9798	0,8693	0,9420	0,8761
N=10000	0,9362	0,9297	0,9801	0,8703	0,9417	0,8763
Cały obraz	0,9357	0,9296	0,9701	0,8702	0,9425	0,8445

6. Literatura

- [1] A. Eskicioglu, P. Fisher, S. Chen: Image Quality Measures and Their Performance. IEEE Transactions on Communications vol. 43 no. 12, pp. 2959–2965, 1995.
- [2] A. Eskicioglu: Quality Measurement for Monochrome Compressed Images in the Past 25 Years. Proceedings of IEEE International Conference on Acoustics Speech and Signal Processing, pp. 1907–1910, Istanbul, Turkey, 2000.
- [3] X. Li: Blind Image Quality Assessment. Proceedings of IEEE International Conference on Image Processing, pp. 449–452, 2002.
- [4] Z. Wang, A.C. Bovik: A Universal Image Quality Index. IEEE Signal Processing Letters vol. 9 no. 3, pp. 81–84, 2002.
- [5] Z. Wang, A.C. Bovik, H. Sheikh, E. Simoncelli: Image Quality Assessment: From Error Measurement to Structural Similarity. IEEE Transactions on Image Processing vol. 13 no. 4, pp. 600–612, 2004.
- [6] K. Okarma, P. Lech: Monte Carlo Based Algorithm for Fast Preliminary Video Analysis. Lecture Notes in Computer Science vol. 5101, pp. 790–799, Springer-Verlag, Heidelberg, 2008.
- [7] H.R. Sheikh, Z. Wang, L. Cormack, A.C. Bovik: LIVE Image Quality Assessment Database Release 2. <http://live.ece.utexas.edu/research/quality>
- [8] Video Quality Experts Group: Final Report on the Validation of Objective Models of Video Quality Assessment. August 2003. <http://www.vqeg.org>