

Analiza wpływu dwuwymiarowych funkcji okien na liniową korelację nowoczesnych miar jakości obrazu z oceną subiektywną

Krzysztof Okarma

Streszczenie: W artykule zamieszczono analizę wpływu funkcji przesuwne okna dwuwymiarowego użytego w procedurze wyznaczania wartości dwóch nowoczesnych miar jakości obrazu (Multi-Scale SSIM oraz VIF) na ich korelację z subiektywną oceną jakości obrazów cyfrowych poddanych typowym zniekształceniom. Obrazy testowe wykorzystane w pracy są dostępne wraz z subiektywnymi ocenami DMOS jako tzw. baza LIVE zawierająca 982 obrazy poddane pięciu typom zniekształceń. Uzyskane dla nich wyniki przeanalizowano pod kątem korelacji liniowej dla dwuwymiarowych wersji kilku najbardziej typowych funkcji okien znanych z widmowej analizy sygnałów.

Słowa kluczowe: ocena jakości obrazów cyfrowych, okna czasowe 2-D

1. WPROWADZENIE

Użyteczność metod obiektywnej oceny jakości obrazów cyfrowych, w szczególności należących do grupy metod skalarnych, jest bardzo duża. Komputerowa analiza obrazu prowadząca do wyrażenia jego jakości za pomocą pojedynczej wartości liczbowej jest atrakcyjna, szczególnie z punktu widzenia różnorodnych metod optymalizacji algorytmów przetwarzania obrazów, a także ich stratnej kompresji. Obiektywne metody oceny jakości obrazów stanowią także podstawę do bardziej rozbudowanych technik oceny jakości sygnałów wideo.

2. NOWOCZESNE MIARY JAKOŚCI OBRAZÓW OPARTE NA PODOBIENSTWIE STRUKTURALNYM

Przez dziesiątki lat porównawcza ocena jakości obrazów oparta była na błędzie średniokwadratowym oraz podobnych do niego wskaźnikach [1], których korelacja z wrażeniami subiektywnymi była dyskusyjna. Biorąc to pod uwagę, zaproponowanie w 2002 roku [12] techniki oceny jakości opartej na iloczynie trzech czynników, znacznie lepiej odpowiadających sposobowi postrzegania typowych zniekształceń obrazu przez oko ludzkie, można uznać za swoistą rewolucję w dziedzinie przetwarzania i analizy obrazów. Twórcy tej miary, nazwanej uniwersalnym wskaźnikiem jakości obrazu (*Universal Image Quality Index*), zwrócili jednak uwagę na jej niestabilność dla ciemnych i płaskich fragmentów obrazu, co zaowocowało zdefiniowaniem tzw. podobieństwa

strukturalnego (*Structural SIMilarity* - *SSIM*) wyrażonego w zwartej formie jako [13]:

$$SSIM = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}, \quad (1)$$

gdzie średnie wartości jasności fragmentów obrazu oryginalnego i ocenianego, wariancje oraz kowariancja obliczane są jako:

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

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

$$\sigma_x^2 = \frac{1}{N^2 - 1} \sum_{i=1}^N \sum_{j=1}^N (x_{ij} - \mu_x)^2, \quad (4)$$

$$\sigma_y^2 = \frac{1}{N^2 - 1} \sum_{i=1}^N \sum_{j=1}^N (y_{ij} - \mu_y)^2, \quad (5)$$

$$\sigma_{xy} = \frac{1}{N^2 - 1} \sum_{i=1}^N \sum_{j=1}^N (x_{ij} - \mu_x)(y_{ij} - \mu_y). \quad (6)$$

Rolę zabezpieczającą przed dzieleniem przez zero pełnią współczynniki oznaczone jako C_1 oraz C_2 , które powinny być na tyle małymi liczbami, aby nie wprowadzać istotnego błędu jeśli wartości mianowników nie są bliskie zero. Proponowane przez twórców wskaźnika SSIM wartości wynoszą $C_1 = 0.01 \times L$ oraz $C_2 = 0.03 \times L$, gdzie L oznacza wartość maksymalnej jasności pikseli (typowo $L=255$ dla obrazów 8-bitowych).

Zakładając zerowe wartości tych współczynników, uzyskujemy definicję uniwersalnego wskaźnika jakości obrazu, z tą dodatkową różnicą, iż obie miary wykorzystują różne funkcje wagowe wewnątrz przesuwne okna do obliczania lokalnego wskaźnika jakości dla fragmentów obrazu zgodnie z zależnością (1).

Stosując okno przesuwne o rozmiarze $N \times N$ pikseli, dla którejkolwiek z obu miar, uzyskać można tzw. mapę jakości obrazu o rozdzielczości mniejszej od rozdzielczości ocenianego obrazu o $N-1$ wierszy i $N-1$ kolumn. Jako skalarny wskaźnik jakości dla całego obrazu przyjmuje się wartość uśrednioną z całej mapy jakości.

W przypadku definicji uniwersalnego wskaźnika jakości obrazu używano okna o rozmiarze 8×8 pikseli bez żadnej funkcji wagowej, co w teorii sygnałów odpowiada

dwuwymiarowemu oknu prostokątnemu. Z kolei definicja podobieństwa strukturalnego określa okno przesuwne o rozmiarze 11×11 pikseli jako spróbkowaną dwuwymiarową funkcję Gaussa. Prosty eksperyment z wykorzystaniem obrazu bez ciemnych plam oraz bez obszarów o stałej intensywności w obrębie całego okna pozwala stwierdzić zauważalne różnice w wartościach uzyskiwanych dla obu miar spowodowane wyłącznie wpływem użytej funkcji okna. W wypadku wskaźnika podobieństwa strukturalnego (SSIM) wpływ zarówno kształtu, jak również wielkości funkcji okna został określony, co zostało przedstawione w poprzednich publikacjach [6-8].

Rozwinięcie wskaźnika SSIM, określane jako Multi-Scale SSIM, oparte jest na idei analizy obrazu w różnej skali na zasadzie piramidy dwójkowej. Trzy elementy występujące w iloczynie we wskaźniku SSIM (zniekształcenie luminancji – l , utrata kontrastu – c oraz zaburzenie struktury – s) są ważone przy użyciu wykładników określonych osobno dla każdej skali. Definicja tego wskaźnika może być wyrażona jako:

$$MS-SSIM = (l_M)^{\alpha_M} \cdot \prod_{j=1}^M [(c_j)^{\beta_j} \cdot (s_j)^{\gamma_j}] \quad (7)$$

gdzie M oznacza najwyższy współczynnik skali uzyskany po $M-1$ iteracji składających się z operacji filtracji dolnoprzepustowej oraz dwukrotnego zmniejszenia rozdzielczości obrazu [14].

Kolejnym interesującym pomysłem opartym na podobnych założeniach jest wykorzystanie wskaźnika wierności informacji wizualnej (*Visual Information Fidelity* - VIF) opartego na dekompozycji falkowej lub obliczanego w dziedzinie płaszczyzny obrazu dla pikseli znajdujących się wewnątrz okna. Warto zauważyć, że wersja działająca w dziedzinie pikseli (VIFp) jest szybsza, a zamiast podpasm dekompozycji falkowej są używane współczynniki skali obrazu. Ogólną definicję wskaźnika VIF można przedstawić jako [10,11]:

$$VIF = \frac{\sum_{j=0}^S \sum_{i=0}^{M_j} I(c_{i,j}; f_{i,j})}{\sum_{j=0}^S \sum_{i=0}^{M_j} I(c_{i,j}; e_{i,j})} \quad (8)$$

gdzie S oznacza liczbę podpasm, M_j oznacza liczbę bloków dla j -tego podpasma, natomiast $I(a;b)$ oznacza współczynnik wzajemnej wymiany informacji (*mutual information*) pomiędzy a oraz b . Mianownik oraz licznik wyrażenia (8) reprezentują odpowiednio informację uzyskaną z obrazu odniesienia i ocenianego. Współczynnik wzajemnej wymiany informacji obliczany jest na podstawie percepcji (e) dwóch odpowiadających sobie bloków obrazu referencyjnego (c) z szumem addytywnym n oraz percepcji (f) bloku obrazu zniekształconego.

W obu przedstawionych metodach domyślną funkcją okna wykorzystywaną w obliczeniach jest dwuwymiarowe okno Gaussa o rozmiarze 11×11 pikseli, podobnie jak ma to miejsce dla wskaźnika SSIM. Celowe jest zatem zbadanie wpływu kształtu funkcji okna na uzyskiwane

wyniki, w szczególności na ich korelację z subiektywnymi ocenami jakości obrazów, także dla miar jakości stanowiących rozwinięcie idei podobieństwa strukturalnego.

3. DWUWYMIAROWE FUNKCJE OKIEN

Zasadniczym celem stosowania okien czasowych w widmowej analizie sygnałów jednowymiarowych jest redukcja niekorzystnego efektu rozmycia widmowej gęstości mocy. Efekt ten jest spowodowany skończonym czasem obserwacji dyskretnego ciągu próbek sygnału, co powoduje jego gwałtowne obcięcie równoznaczne ze stosowaniem okna prostokątnego. Do celów analizy widmowej sygnałów, a także projektowania filtrów cyfrowych o skończonej odpowiedzi impulsowej, zaproponowano szereg okien czasowych o kształtach powodujących łagodniejsze "obcinanie" ciągu obserwowanych próbek.

Za najprostsze okno, wyłączając prostokątne, uznać można okno Bartletta, mające kształt symetrycznego trójkąta. Zakładając, iż wszystkie okna dane są w postaci znormalizowanej tzn. mają wartość równą 1 na środku przedziału obserwacji, dla nieparzystej ilości próbek będzie się ono charakteryzowało wartościami równymi 0 dla próbek skrajnych oraz 1 dla próbki środkowej.

Do najpopularniejszych okien czasowych, powszechnie stosowanych w przetwarzaniu sygnałów jednowymiarowych, zaliczyć można okna z rodziny kosinusoidalnych: Hamminga, Hannę oraz Blackmana [2] określone odpowiednio następującymi wzorami (w postaci ciągłej zakładając przedział obserwacji, w którym zdefiniowana jest funkcja okna jako $[-T; T]$):

$$w_{Hm}(t) = 0.54 + 0.46 \cdot \cos\left(\frac{\pi \cdot t}{T}\right) \quad (9)$$

$$w_{Hn}(t) = 0.5 + 0.5 \cdot \cos\left(\frac{\pi \cdot t}{T}\right) \quad (10)$$

$$w_{Bl}(t) = 0.42 + 0.5 \cdot \cos\left(\frac{\pi \cdot t}{T}\right) + 0.08 \cdot \cos\left(\frac{2\pi \cdot t}{T}\right) \quad (11)$$

Ponadto, dla celów porównawczych, w niniejszej pracy użyto okien Parzena (rzadziej określane jako okno de la Valle-Poussina [2]) oraz Nuttalla [4], również zaliczanego do grupy okien kosinusoidalnych, zdefiniowanych odpowiednio jako:

$$w_P(t) = \begin{cases} 1 - 6 \cdot \left(\frac{t}{T}\right)^2 \cdot \left(1 - \frac{|t|}{T}\right) & \text{dla } 0 \leq t \leq \frac{T}{2} \\ 2 \cdot \left(1 - \frac{|t|}{T}\right)^3 & \text{dla } \frac{T}{2} \leq t \leq T \end{cases} \quad (12)$$

$$w_{Nut}(t) = 0.3635819 + 0.4891775 \cdot \cos\left(\frac{\pi \cdot t}{T}\right) + 0.1365995 \cdot \cos\left(\frac{2\pi \cdot t}{T}\right) + 0.0106411 \cdot \cos\left(\frac{3\pi \cdot t}{T}\right) \quad (13)$$

Tabela 1.

Współczynniki korelacji liniowej wartości wskaźnika MS-SSIM uzyskane dla różnych funkcji okien z oceną subiektywną (wartościami DMOS)

Zniekształ- Funkcja cenie okna	JPEG	JPEG 2000	Szum biały	Rozmycie Gausa	Symulowa ny kanał z zanikami Rayleigha	Wszystkie
Gausa	0,6995	0,8155	0,7489	0,7023	0,7046	0,4762
prostokątna	0,6976	0,8230	0,7534	0,7032	0,6961	0,4942
Bartletta	0,6990	0,8178	0,7505	0,6987	0,7004	0,4841
Hamminga	0,6995	0,8181	0,7508	0,6997	0,7001	0,4842
Hanna	0,6992	0,8168	0,7500	0,6997	0,7017	0,4818
Blackmana	0,6994	0,8158	0,7493	0,7013	0,7036	0,4784
Nuttalla	0,6991	0,8157	0,7510	0,7018	0,7038	0,4791
Parzena	0,6994	0,8158	0,7493	0,7014	0,7037	0,4781

Przekształcenie postaci okien jednowymiarowych do postaci dwuwymiarowej zostało zrealizowane w postaci iloczynu wartości funkcji okna w kierunku poziomym i pionowym. W celu analizy wpływu wyłącznie kształtu okna na uzyskiwane wyniki przyjęto stały rozmiar okna przesuwonego równy 11×11 pikseli zgodnie z zaleceniami twórców wskaźnika SSIM.

4. ZASTOSOWANIE FUNKCJI OKIEN DO OBLICZENIA WARTOŚCI WSKAŹNIKÓW JAKOŚCI OBRAZU

W celu zbadania wpływu zastosowanej funkcji okna na zgodność wskaźników Multi-Scale SSIM oraz VIF z oceną subiektywną przeprowadzono obliczenia obu wartości z wykorzystaniem ośmiu różnych funkcji okien o rozmiarze 11×11 pikseli dla 982 obrazów stanowiących tzw. bazę LIVE [9]. Jest to powszechnie stosowana baza 29 obrazów poddanych pięciu rodzajom zniekształceń o różnych intensywnościach wraz z uśrednionymi ocenami subiektywnymi w postaci wartości DMOS (Differential Mean Opinion Scores) przeskalowanych do zakresu <0; 100>. Ze względu na fakt, iż 203 spośród ocen dotyczą obrazów oryginalnych, dla których wartości obu rozważanych miar jakości są z założenia równe 1, zostały one wyłączone z analizy, aby nie powodowały zawyżenia współczynników korelacji, pomimo iż w wielu publikacjach prezentowane są wyniki analiz dla całej bazy. Po obliczeniu wartości wskaźników jakości z użyciem różnych funkcji okien dla poszczególnych rodzajów zniekształceń, jak również dla wszystkich zniekształconych obrazów dostępnych w bazie LIVE, wyznaczono współczynniki korelacji Pearsona pomiędzy wartościami wskaźników MS-SSIM oraz DMOS zebrane w Tabeli 1, a także analogiczne współczynniki korelacji pomiędzy wartościami VIF oraz DMOS (Tabela 2). W obu tabelach pogrubieniem czcionki zaznaczono wskaźniki korelacji wyższe aniżeli uzyskane dla standardowego okna Gaussa, przy czym wyniki najlepsze

zaznaczono dodatkowo pochyleniem. Obliczenia wykonano dla składowej luminancyjnej występującej w "telewizyjnych" modelach barw YUV/YIQ.

Analizując uzyskane wyniki można zauważyć, iż istnieje zależność pomiędzy kształtem funkcji okna a korelacją uzyskiwanych wyników z subiektywnymi ocenami jakości zniekształconych obrazów.

W wypadku obu rozpatrywanych miar jakości można zauważyć, iż współczynnik korelacji dla całej bazy jest znacznie niższy aniżeli dla każdego z pięciu rodzajów zniekształceń osobno, co jest naturalne ze względu na różnorodność typów zniekształceń obrazu (zaszumienie, kompresja, rozmycie, błędy bitowe).

Dla wskaźnika Multi-Scale SSIM w zdecydowanie największej liczbie przypadków najefektywniejsze okazało się okno prostokątne, co ma tę zaletę, iż umożliwia dodatkowe zmniejszenie złożoności obliczeniowej procedury oceny jakości. W wypadku miary VIF uzyskano wyższe wskaźniki korelacji, jednak pomimo najlepszych ogólnych wyników uzyskanych dla okna prostokątnego, dla poszczególnych rodzajów zniekształceń nieco lepsze okazywały się inne rodzaje okien (np. okno Nuttalla w wypadku kompresji).

5. PODSUMOWANIE I WNIOSKI

Metody oceny jakości rozważane w niniejszym artykule należą do grupy metod obiektywnych wymagających pełnej znajomości obrazu oryginalnego (porównawczych, czyli tzw. *full-reference*). Nie należy jednak traktować tego jako znaczącej wady ze względu na dużą uniwersalność takiego podejścia. Aktualnie znane są również metody tzw. "ślepej" (*no-reference*) oceny jakości obrazów [3], nie wykorzystujące w ogóle informacji o obrazach oryginalnych, w przeciwieństwie do niektórych typu *reduced-reference* (wymagających częściowej informacji o oryginalne) [5]. Są one zwykle specjalizowane, pozwalając estymować wielkość zniekształceń obecnych w obrazie tylko dla określonego ich typu, najczęściej rozmycia lub artefaktów blokowych charakterystycznych dla kompresji JPEG.

Tabela 2.

Współczynniki korelacji liniowej wartości wskaźnika VIFp uzyskane dla różnych funkcji okien z oceną subiektywną (wartościami DMOS)

Zniekształ- Funkcja cenie okna	JPEG	JPEG 2000	Szum biały	Rozmycie Gausa	Symulowa ny kanał z zanikami Rayleigha	Wszystkie
Gausa	0,9232	0,9285	0,8708	0,9484	0,9578	0,7528
prostokątna	0,9098	0,9142	0,8721	0,9391	0,9492	0,7614
Bartletta	0,9236	0,9291	0,8706	0,9494	0,9579	0,7515
Hamminga	0,9232	0,9286	0,8709	0,9488	0,9578	0,7529
Hanna	0,9265	0,9319	0,8696	0,9488	0,9588	0,7464
Blackmana	0,9299	0,9354	0,8676	0,9466	0,9592	0,7391
Nuttalla	0,9312	0,9364	0,8661	0,9451	0,9594	0,7355
Parzena	0,9305	0,9358	0,8668	0,9459	0,9594	0,7373

Przedstawione wyniki badania korelacji liniowej nowoczesnych porównawczych miar jakości obrazów z ocenami subiektywnymi w zestawieniu z uzyskanymi uprzednio wynikami [7] dotyczącymi zależności pomiędzy optymalną wielkością okna dla wskaźnika SSIM a rozdzielczością ocenianego obrazu pozwolić mogą na opracowanie kolejnych modyfikacji znanych wskaźników jakości obrazów charakteryzujących się jeszcze lepszym dopasowaniem do ludzkiej percepcji obrazów cyfrowych.

LITERATURA

- [1] Eskicioglu A., „Quality Measurement for Monochrome Compressed Images in the Past 25 Years” *Proceedings of the IEEE International Conference on Acoustics Speech and Signal Processing*, Istanbul, Turkey 2000, ss. 1907-1910.
- [2] Harris F.J., “On the Use of Windows for Harmonic Analysis with Discrete Fourier Transform”. *Proceedings of the IEEE*, vol. 66, no 1, January 1978, ss. 51–83.
- [3] Li X., „Blind Image Quality Assessment” *Proceedings of the IEEE International Conference on Image Processing*, Rochester, USA, 2002, ss. 449-452.
- [4] Nuttall A.H., “Some Windows with Very Good Sidelobe Behavior” *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 29, no. 1, February 1981, ss. 84–91.
- [5] Okarma K., Lech P., „A Statistical Reduced-Reference Approach to Digital Image Quality Assessment”, *Lecture Notes in Computer Science*, vol. 5337, 2009, ss. 43-54.
- [6] Okarma K., „Colour Image Quality Assessment using Structural Similarity Index and Singular Value Decomposition”, *Lecture Notes in Computer Science*, vol. 5101, 2008, ss. 55-65.
- [7] Okarma K., „Two-Dimensional Windowing in the Structural Similarity Index for the Colour Image Quality Assessment”, *Lecture Notes in Computer Science*, vol. 5702, 2009, ss. 501-508.
- [8] Okarma K., „Optymalizacja korelacji szybkiej metody estymacji jakości obrazów opartej na podobieństwie strukturalnym z oceną subiektywną”, *Pomiary Automatyka Kontrola*, vol. 55, nr 7, ss. 444-446, 2009.
- [9] Sheikh H.R., Wang Z., Cormack L., Bovik A.C., LIVE Image Quality Assessment Database Release 2. <http://live.ece.utexas.edu/research/quality>.
- [10] Sheikh H.R., Bovik A.C., de Veciana G., „An Information Fidelity Criterion for Image Quality Assessment Using Natural Scene Statistics”, *IEEE Transactions on Image Processing*, vol. 14, no. 12, 2005, ss. 2117-2128.
- [11] Sheikh H.R., Bovik A.C., „Image Information and Visual Quality”. *IEEE Transactions on Image Processing*, vol. 15, no. 2, 2006, ss. 430-444
- [12] Wang, Z., Bovik, A., „A Universal Image Quality Index”, *IEEE Signal Processing Letters*, vol. 9, no. 3, 2002, ss. 81–84.
- [13] Wang, Z., Bovik, A., Sheikh, H., Simoncelli, E., „Image Quality Assessment: From Error Measurement to Structural Similarity”, *IEEE Transactions on Image Processing*, vol. 13, no. 4, 2004, ss. 600–612.
- [14] Wang, Z., Simoncelli, E., Bovik, A., „Multi-Scale Structural Similarity for Image Quality Assessment”, *Proceedings of the 37th IEEE Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, California, USA, 2003.



dr inż. Krzysztof Okarma
Zachodniopomorski Uniwersytet
Technologiczny w Szczecinie
Wydział Elektryczny
Katedra Przetwarzania Sygnałów
i Inżynierii Multimedialnej
ul. 26. Kwietnia 10
71-126 Szczecin
tel.: 91 449 53 11
e-mail: okarma@zut.edu.pl