

**Anna TOMASZEWSKA**

ZACHODNIOPOMORSKI UNIWERSYTET TECHNOLOGICZNY, WYDZIAŁ INFORMATYKI,  
ul. Żołnierska 52, 71-210 Szczecin

## Wspomaganie sprzętowe do wyznaczenia statystyk obrazów naturalnych wyższego rzędu

Dr inż. Anna TOMASZEWSKA

Ukończyła studia na Wydziale Informatyki Politechniki Szczecińskiej w 2000 r. W 2003 r. obroniła pracę doktorską na tym samym wydziale. Obecnie pracuje jako adiunkt w Zakładzie Grafiki Komputerowej Instytutu Grafiki Komputerowej i Systemów Multimediałnych. Jej zainteresowania naukowe to przetwarzanie obrazów HDR, analiza percepcyjna oraz systemy graficzne czasu rzeczywistego.



e-mail: atomaszewska@wi.zut.edu.pl

### Streszczenie

Statystyki obrazów naturalnych, definiowanych jako nieprzetworzone obrazy rejestrowane przez człowieka, charakteryzują się dużą regularnością. Ich cechy wykorzystywane są w wielu aplikacjach grafiki komputerowej takich jak usuwanie szumu, czy kompresja. W artykule przedstawiono algorytm do szybkiego obliczenia statystyk wyższego rzędu na podstawie współczynników falek z wykorzystaniem programowalnego procesora graficznego. W rezultatach przedstawiono wyniki przyspieszenia uzyskanego przy wykorzystaniu GPU w porównaniu z implementacją na CPU.

**Słowa kluczowe:** statystyki obrazów naturalnych, system wizualny człowieka, programowalny procesor graficzny, GPGPU.

### GPU-based natural images statistics calculation

#### Abstract

A natural image is unprocessed reproduction of a natural scene observed by a human. The Human Visual System (HVS), during its evolution, has been adjusted to the information encoded in natural images. Computer images are interpreted best by a human when they fit natural image statistics that can model the information in natural images. The main requirement of such statistics is their striking regularity. It helps separate the information from noise, reconstruct information which is not available in an image, or only partially available. Other applications of statistics is compression, texture synthesis or finding distortion model in image like blur kernel. The statistics are translation and scale invariant, therefore a distribution of statistics does not depend on the object position in the image and on its size. In this paper there are presented higher order natural image statistics calculations based on GPU. The characteristic of the statistics is that they are independent of the scale and rotation transformations. Therefore, they are suitable for many graphic applications. To analyze images there is used statistics computed in the wavelet domain and there is considered the image contrast. The computation speedup is presented in the results. The paper is organized as follows: the overview of natural images statistics is introduced in Section 2. In Section 3 the GPU-based implementation is described. The obtained results are given in Section 4. Finally, there are presented the concluding remarks.

**Keywords:** natural image statistics, human visual system, graphics processing unit, GPGPU.

## 1. Wstęp

Definiowanie modeli opisujących naturalne sceny wymaga uwzględnienia szeregu różnych czynników takich jak pole widzenia, rozdzielczość obrazu, czy nieliniową charakterystykę sensorów i algorytmów postprocessingu. Istotny jest również rodzaj analizowanej sceny oraz warunki oświetlenia. Należy zauważyć, że sceny zewnętrzne różnią się od scen wewnętrznych, czy od portretów ludzi. Zdefiniowane pojedynczego modelu dla

dowolnego obrazu, bez wcześniejszej jego klasyfikacji, jest złożonym problemem [4]. Jednym ze sposobów jego rozwiązania jest zastosowanie statystyk obrazów naturalnych (NIS). Statystyki obrazów naturalnych to modele, które w syntetyczny sposób opisują charakterystyczne cechy obrazów naturalnych [1, 2, 4, 6, 7]. Obrazy to różne kombinacje wartości pikseli, których potencjalnie jest bardzo dużo. Okazuje się jednak, że w naturze występuje tylko niewielki podzbiór tych kombinacji, który bardzo dobrze wpisuje się w pewne modele statystyczne. Istnieje teoria, że ludzki układ wzrokowy, na drodze ewolucji, dopasował się właśnie do takich statystyk i najlepiej interpretuje informacje z obrazów, które do nich pasują.

Cechą charakterystyczną NIS jest ich niezależność od skali oraz rotacji [8], dzięki czemu dobrze opisują obrazy bez względu na wykonane na nich przekształcenia. Dzięki swoim cechom, statystyki obrazów naturalnych są wykorzystywane w algorytmach grafiki komputerowej takich jak usuwanie szumu [8], usuwanie nadmiarowej informacji [8], czy transformacja stylów malarskich [7], stanowiąc często składową większego systemu analizy i przetwarzania obrazów. Implementacja statystyk wykorzystująca architekturę procesora graficznego w dużym stopniu przyspieszy czas ich wyznaczania. W artykule przedstawiono klasyfikację statystyk obrazów naturalnych oraz sprzętową implementację algorytmu wyznaczania statystyk wyższego rzędu.

W rozdziale 2 przedstawiono przegląd literatury związanej z obrazami naturalnymi oraz krótką charakterystykę statystyk pierwszego, drugiego oraz wyższych rzędów. Implementację sprzętową statystyk przedstawiono w rozdziale 3. Rezultaty przeprowadzonych testów wydajnościowych zaprezentowano w rozdziale 4.

## 2. Statystyki obrazów naturalnych

Wygląd oraz parametry statystyki obrazów naturalnych zależą od jej rzędu. Rozważanymi najczęściej w literaturze oraz wykorzystywanymi w obliczeniach są statystyki pierwszego, drugiego oraz wyższego rzędu [4].

Cechą statystyk pierwszego rzędu wykorzystywanie wyłącznie wartościach poszczególnych pikseli, bez uwzględnienia wzajemnych relacji pomiędzy otoczeniem. Statystyką taką opisuje histogram wartości pikseli, niezależny od translacji oraz skali obrazu, niestety nie nadający się do identyfikacji obrazów naturalnych. Dodanie dowolnie wygenerowanego szumu do histogramu zakłóca czystość sygnału, którego nie można w żaden sposób odseparować.

W przeciwieństwie do nich, statystyki drugiego rzędu uwzględniają wzajemne zależności pomiędzy pikselami, co wzbogaca statystykę w dodatkowe informacje. Z praktycznego punktu widzenia, statystyka drugiego rzędu opisuje autokorelację pomiędzy parami pikseli, gdzie intensywność dwóch pikseli o określonych pozycjach opisana jest za pomocą wektora. Transformacja autokorelacji w dziedzinie Fouriera tworzy tzw. power domain, gdzie odległość pomiędzy dwoma pikselami opisana jest za pomocą częstotliwości, natomiast różnica intensywności za pomocą amplitudy. Wadą takiego podejścia, ze względu na uśrednione wartości amplitudy oraz fazy, jest utrata informacji o lokalizacji szczegółów w obrazie.

Rozwinięcie statystyk drugiego rzędu stanowią statystyki wyższego rzędu, uwzględniające relacje pomiędzy wartościami pikseli z trzech lub większej liczby pozycji. Podejście umożliwia detekcję krawędzi oraz informacji o obiektach zawartych w obrazie. Obliczenia wykonywane są w dziedzinie falek, dzięki

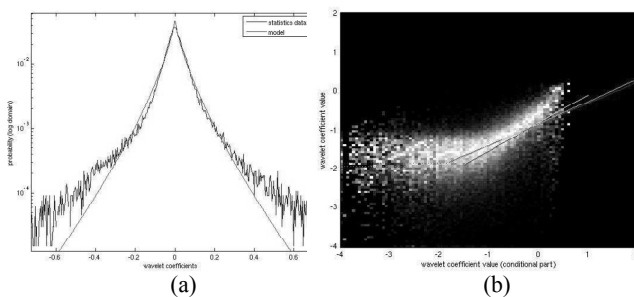
czemu statystyki wyższego rzędu zachowują zarówno częstotliwościowe jak i przestrzenne informacje zawarte w obrazie. Wśród statystyk wyższego rzędu wyróżnia się dwa podstawowe modele: marginal oraz joint wavelet.

Pierwszy model należący do statystyk wyższego rzędu (model marginal) zakłada niezależność współczynników falek wewnątrz tego samego przedziału oraz identyczność ich rozkładu [3, 5]. Obliczenie takiej statystyki wymaga transformacji obrazu do przestrzeni współczynników falek, wyboru zakresu dla którego liczone są współczynniki a następnie sortowania wyznaczonych współczynników i ich normalizacji. Rozkład współczynników falek (Rysunek 1a) dobrze wpisuje się w uogólniony rozkład Laplace'a (Równanie 1) [4].

$$P_c(c;s,p) = \frac{\exp\left(-\frac{|c|}{s}\right)^p}{Z(s,p)}, \quad (1)$$

gdzie parametry  $s$  i  $p$  dobierane są tak aby prawdopodobieństwo dopasowania modelu oraz obrazu było jak największe. Wykładnik  $p=2$  odpowiada rozkładowi Gaussa, natomiast dla  $p=1$  funkcja przypomina rozkład Laplace'a. Uogólniając, małe wartości wykładnika  $p$  odpowiadają rozkładowi o dużej koncentracji wartości w pobliżu zera. Zaobserwowano, że wartości parametru  $p$  zawierają się najczęściej w zbiorze [0.4, 0.8]. Współczynnik  $s$  zmienia się monotonnie wraz z poziomem skali funkcji bazowej, zwiększając wariancję parametrów o niskiej częstotliwości [4].

Stałą normalizacji  $Z(s,p)$  wyraża zależność  $Z(s,p) = 2 \frac{s}{p} \Gamma \frac{1}{p}$ .



Rys. 1. Statystyki wyższego rzędu: model marginal (a), model joint (b) dla obrazu Lena

Fig. 1. Higher order statistics for Lena image: marginal model (a) and joint model (b)

Kształt modelu marginal zależy od dynamiki oraz „naturalności” obrazu [5]. Dla obrazów o dużej liczbie kontrastów statystyka charakteryzuje się długimi „ogonami” oraz ostrym pikiem dla współczynników równych 0. Jest to typowe dla obrazów naturalnych, gdzie pik odpowiada obszarom płaskim – zawartym pomiędzy krawędziami (obszary przewidywalne, nadmiarowość informacji), „ogony” natomiast krawędziom w obrazie. Obrazy sztucznie wygenerowane, mają zazwyczaj dużo mniejsze „ogony” oraz więcej obszarów płaskich. Niskie częstotliwości są zwykle ignorowane przy interpretacji obrazów ponieważ ich widoczność jest ograniczona ostrością widzenia. Układ wzrokowy człowieka najczulszy jest na średnie częstotliwości. Natomiast brak wysokich częstotliwości powoduje, że obraz jest nieostry i ma mało szczegółów.

Relacje pomiędzy współczynnikami falek obliczonymi dla różnych poziomów skali w postaci prawdopodobieństwa warunkowego opisuje obliczany z wykorzystaniem falek model joint. Zależności występujące w tej statystyce wskazują na tendencje współczynników o dużych wartościach do występowania koło siebie w ramach tego samego poziomu oraz dotyczą odpowiadających sobie współczynników na różnych poziomach dla różnych orientacji [3]. Histogram rozkładu warunkowego (równanie 2) (rysunek 1b) obliczany jest poprzez

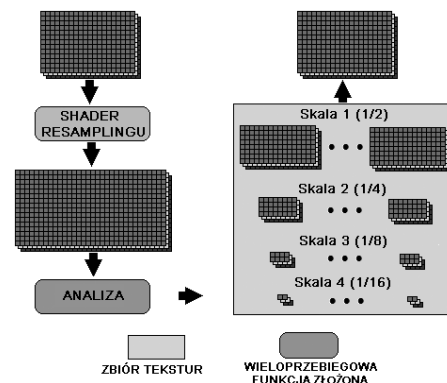
podzielenie każdej kolumny histogramu rozkładu łącznego przez odpowiadającą jej wartość histogramu warunkowego obliczonego dla danej skali.

$$P(C|P) = \frac{P(C \cap P)}{P(P)}, \quad (2)$$

gdzie  $P(C \cap P)$  oznacza prawdopodobieństwo że dwa piksele na odpowiadających sobie pozycjach przyjmują takie same wartości współczynników,  $P(P)$  natomiast jest zwykłym histogramem (prawdopodobieństwo, że piksel na określonej pozycji przyjmuje daną wartość).

### 3. Sprzętowa implementacja algorytmu

Implementację algorytmu wykonano w API OpenGL. Jest to uniwersalny interfejs dostępu do funkcji platformy sprzętowej, umożliwiający tworzenie grafiki 3D. Zawiera on rozszerzenia, które umożliwiają wykorzystanie dodatkowych funkcji sprzętowych. OpenGL pozwala na wykorzystanie bibliotek pomocniczych, między innymi takich jak GLEW czy GLUT, które są wykorzystywane w stworzonej aplikacji. Programy dla procesorów wierzchołków (vertex shader) oraz fragmentów (fragment shader) napisane zostały w języku GLSL. Język ten pomaga w definiowaniu własnych przekształceń geometrycznych wierzchołków i operacji na fragmentach. Aplikacja została napisana w języku C++. Kompilatorem, przy pomocy którego napisano program jest G++ w wersji 4.2.3 (Ubuntu 4.2.3.-2ubuntu7).



Rys. 2. Schemat sprzętowej implementacji statystyk obrazów naturalnych wyższego rzędu

Fig. 2. The schema of a hardware implementation of higher order natural images statistics

Wykonywane w aplikacji operacje macierzowe wykorzystują obiekt ramki bufora (ang. *Frame Buffer Object - FBO*). Umożliwia on ustawienie tekstury jako obiektu wyjściowego procesora graficznego zamiast standardowego bufora okna, co w połączeniu z programowalnymi jednostkami cieniowania wierzchołków i fragmentów, umożliwia wykorzystanie procesora GPU do wykonywania różnorodnych obliczeń arytmetycznych. Jednostki cieniowania wierzchołków uruchamiane są jednocześnie dla poszczególnych wierzchołków transformując ich położenie w wirtualnej przestrzeni 3D na współrzędne 2D na ekranie. Wyznaczenie koloru piksela realizowane jest przez jednostki cieniowania pikseli.

Wszelkie obrazy w pamięci GPU występują jako tekstury i kodowane są w postaci 16 lub 32-bitowych liczb zmiennoprzecinkowych. Standardowo rozszerzenie FBO pozwala na użycie tylko jednej tekstury wyjściowej. W związku z tym, że aplikacja nie mogła być ograniczona tylko do jednej tekstury wyjściowej, wykorzystano rozszerzenie Multiple Render Targets (MRT). Pozwala ono zdefiniować kilka tekstur na wyjściu programu fragmentów, dzięki czemu można napisać programy GPU, generujące kilka obrazów jednocześnie. Przechowywanie danych w pamięci karty graficznej pozwala na szybsze odwołanie się do

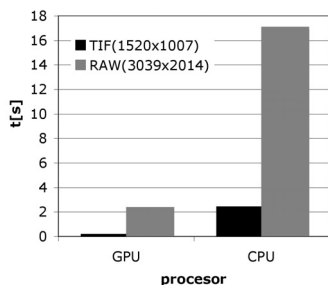
nich przez GPU, a także odciąża wykorzystanie magistrali PCI-E, jednocześnie zmniejszając w znacznym stopniu użycie CPU.

Schemat algorytmu przedstawiono na Rys. 2. W pierwszej kolejności obraz wejściowy powiększany jest do rozmiarów najbliższej potęgi liczby dwa. Powiększenie następuje przez resampling tekstury wejściowej, interpolowanej dwuliniowo, do tekstury o większych wymiarach. Operacja ta wykonywana jest przez shader resamplingu. Kolejnym krokiem jest transformacja falkowa (Rys. 4). Głównymi operacjami transformacji są filtracje kolumnowe i wierszowe. W metodzie filtracji kolumnowej, założono że po każdym filtrowaniu następuje redukcja, która usuwa parzyste elementy kolumn (tzn. parzyste wiersze) obrazu wyjściowego, dlatego też nie są one wyznaczane. Okno próbkowania obrazu wejściowego przesuwa się o dwa piksele na każde przesunięcie o jeden piksel wyznaczanego elementu wyjściowego. Obrazy wyjściowe filtracji kolumnowej stanowią obrazy wejściowe filtracji wierszowej. Filtracja po wierszach różni się od poprzedniej, orientacją okna próbkowania oraz dzięki rozszerzeniu MRT równoległą realizacją dla 4 obrazów (wszystkich gałęzi jednocześnie: RGBA). W kolejnym etapie wyznaczane są współczynniki falek i zapisywane do tekstury wyjściowej.

## 4. Rezultaty

Oprogramowanie przetestowano na komputerze PC wyposażonym w kartę z procesorem graficznym ATI Radeon HD 5470, procesor Intel Core i5 2.5GHz, 4GB pamięci RAM, system operacyjny Windows 7 Professional x64.

Do przeprowadzenia badań wykorzystano 2 obrazy w różnych formatach: szesnastobitowym formacie TIF o rozdzielczości 1520x1007 oraz dwunastobitowym formacie RAW o 4-krotnie większej rozdzielczości 3039x2014. W ramach badań przeprowadzono porównanie wydajności obliczeń zaprojektowanych pod kątem realizacji na GPU (rys. 3). Zastosowanie wspomaganie sprzętowego umożliwiło 12.5-krotne przyspieszenia obliczeń w porównaniu z implementacją programową.

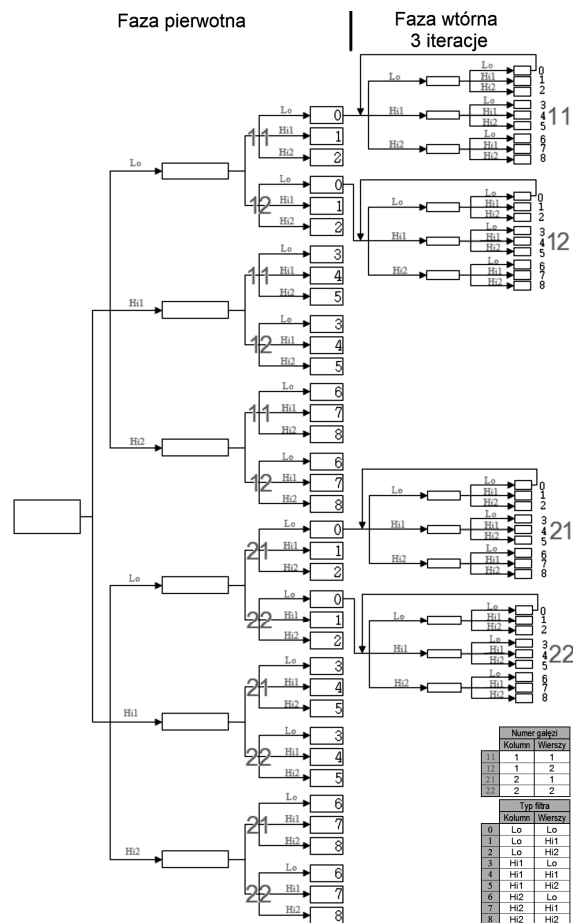


Rys. 3. Porównanie wydajności obliczeń wykonywanych na CPU i GPU  
Fig. 3. Results of efficiency for CPU and GPU based computations

## 5. Podsumowanie

W artykule przedstawiono klasyfikację statystyk obrazów naturalnych. Podczas obliczeń główny nacisk położono na statystyki wyższego rzędu: modele marginal i joint. Zaprezentowano sprzętową implementację algorytmu wyznaczania współczynników statystyk obrazów naturalnych wyższego rzędu. Wykorzystanie architektury procesora graficznego do obliczeń ogólnego przeznaczenia umożliwiło przyspieszenie pracy algorytmu o 12,5 razy w porównaniu z implementacją wykorzystującą wyłącznie CPU.

Algorytmy wykorzystujące NIS stanowią często składowe większych systemów wykorzystywanych w przetwarzaniu obrazów. Skrócenie czasu obliczeń każdego z etapów systemu stanowi istotne znaczenie dla wydajności całego systemu.



Rys. 4. Schemat transformaty falkowej  
Fig. 4. The schema of a wavelet transformation

Praca została wykonana w ramach realizacji grantu MNiSW nr N516 193537 (Polska).

## 6. Literatura

- [1] Van der Schaaf A.: Natural Images Statistics and Visual Processing, PhD thesis, Rijksuniversitet Groningen, The Netherlands, 1998.
- [2] Reinhard E., Shirley P., Ashikhim M., Troscianko T.: Second Order Image Statistics in Computer Graphics, In Proc. APGV, 2004.
- [3] Simoncelli E. P.: Modeling the Joint Statistics of Images in the Wavelet Domain, In Proc. SPIE, 44th Annual Meeting, vol. 3813, pp. 188-195, 1999.
- [4] Simoncelli E. P.: Statistical Modeling of Photographic Images. Handbook of Video and Image Processing, 2nd edition ed. Alan Bovik, Academic Press, 2005.
- [5] Buccigrossi R. W., Simoncelli E. P.: Image Compression via Joint Statistical Characterization in Wavelet Domain, In Proc. IEEE Transaction on Image Processing, vol. 8, iss. 12, pp. 1688-1701, 1999.
- [6] Ruderman D. L.: The statistics of Natural Images, Computation in Neural Systems 5, pp. 517-548, 1994.
- [7] Bae Soonmin and Durand Frédo: Statistical Analysis and Transfer of Pictorial Styles, In proceedings of 2004 MIT Student Oxygen Workshop, 2004.
- [8] Lyu S., Simoncelli E.P.: Statistically and Perceptually Motivated Nonlinear Image Representation, Human Vision and Electronic Imaging XII, Proc. Of SPIE-IS&T Electronic Imaging, SPIE, 2007.