Przemysław MAZUREK, **Krzysztof OKARMA**
ZACHODNIOPOMORSKI UNIWERSYTET TECHNOLOGICZNY W SZCZECINIE, ul. 26. Kwietnia 10, 71-126 Szczecin

An efficient estimation of the Structural Similarity index using the GPGPU programming techniques

PhD Eng. Przemysław MAZUREK

Assistant professor in the Department of Signal Processing and Multimedia Engineering at the Faculty of Electrical Engineering, West-Pomeranian University of Technology in Szczecin. Author of more then 80 papers related to the digital signal processing, estimation of object kinematics, biosignals acquisition and processing.

e-mail: przemyslaw.mazurek@zut.edu.pl

PhD Eng. Krzysztof OKARMA

Assistant professor in the Chair of Signal Processing and Multimedia Engineering of West Pomeranian Technological University in Szczecin. He was graduated in the Electrical Faculty (1999) and Informatics Faculty 2001) of the Technological University. PhD degree he has got from the Electrical Faculty (2003). Currently he is a chief of part-time studies at the Faculty. An author of more than 80 papers concerning the processing and analysis of pictures and signals.

e-mail: okarma@zut.edu.pl

Abstract

In the paper a fast implementation of the SSIM index is presented. Because of specific features of the CUDA architecture, the 8×8 pixel sliding window is used. In order to speed up the computations, the vertical passes are limited to the 32-pixels wide fragments, so four pixels margins should be left for each fragment. The verification of the proposed estimation is performed for the LIVE database with images corrupted by five common types of distortions and their subjective evaluations.

Keywords: Image Quality Assessment, GPGPU Programming, Structural Similarity.

Wydajna estymacja wskaźnika podobieństwa strukturalnego z wykorzystaniem technik programowania układów GPGPU

Streszczenie

W artykule zaprezentowano wydajną technikę implementacji nowoczesnej metody oceny jakości obrazu znanej jako podobieństwo strukturalne (SSIM). Uwzględniając specyficzne uwarunkowania architektury CUDA, obliczenia wykonano przy użyciu okna przesuwnego o rozmiarze 8×8 pikseli, podobnie jak we wcześniejszym wariancie tego wskaźnika określanym jako uniwersalny wskaźnik jakości obrazu (UIQI). W celu przyspieszenia obliczeń, przebiegi pionowe zostały ograniczone do fragmentów obrazu o szerokości 32 pikseli, co przy tym rozmiarze okna wymaga pozostawienia czteropikselowych marginesów z obu stron. Estymowana wartość globalna wskaźnika SSIM jest obliczana jako średnia z wartości lokalnych obliczanych dla każdego fragmentu obrazu. Praktyczna weryfikacja dokładności proponowanej metody została przeprowadzona z wykorzystaniem obrazów ze znanej bazy LIVE Image Quality Assessment Database Release 2 zawierającej obrazy poddane pięciu typowym rodzajom zniekształceń wraz z ich ocenami subiektywnymi (wartościami DMOS).

Słowa kluczowe: ocena jakości obrazów cyfrowych, programowanie układów GPGPU, podobieństwo strukturalne.

1. Introduction

A fast and reliable image quality assessment is one of the most relevant aspects of the contemporary image processing and analysis. Starting from the very classical methods [1], mainly based on the Mean Squared Error (MSE) with their poor correlation with the Human Visual System (HVS), to some modern ones, it has always been a tool needed by the image processing society, not only just for the determination the quality of images. A good image quality metric is an essential element for the assessment of some new algorithms related e.g. to image filtering or lossy compression.

There are many image quality assessment methods, which are very specialised and can detect only one or two types of contaminations e.g. JPEG block artefacts [2] or Gaussian blur [3]. Some less popular techniques used for the image quality assessment are wavelets, Singular Value Decomposition [4] or transforms. Some of them can work as the no-reference (blind) methods [5] which do not require any information related to the original image. Nevertheless, currently the full reference methods seem to be much more universal in comparison to the "blind" approach.

2. The Structural Similarity

A very popular approach to the full reference image quality assessment is the Structural Similarity index proposed in 2004 by Wang and Bovik [6]. The idea of the method is based on the usage of the sliding window in order to calculate the quality map of the image (based on the local quality values) taking into account the three common types of contaminations: the loss of correlation, mean distortion and variance distortion. The local SSIM index is calculated as:

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

where, the constants C_1 and C_2 are chosen in the way that they do not introduce significant changes of the results. The values suggested by the authors of the paper [6] are $C_1 = (0.01 \times L)^2$ and $C_2 = (0.03 \times L)^2$, where L denotes the number of the grayscale levels in the image. The remaining elements in the formula (1) are calculated as:

$$\mu_x = \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N} x_{ij}, \qquad \mu_y = \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N} y_{ij}, \tag{2}$$

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

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

where M and N denote the resolution of the image. The recommended sliding window's size is 11×11 pixels and the Gaussian window should be applied. The choice of the Gaussian window and its size is somehow arbitrary, it is treated as the default value, but the usage of 8×8 pixels rectangular window, as primarily proposed, is also allowed.

3. The Idea of the Proposed Approach

Taking into account the fact that actually there is no ideal objective image quality assessment method and all the more or less universal metrics should be treated as some quality estimators, in some applications the speed of the quality index calculations may be a much more crucial element than its accuracy. One of the most attractive solutions for such purposes seems to be the usage of the GPGPU computational power. Nevertheless, an optimal, or even sub-optimal, utilisation of the modern architectures such as the CUDA, require specific representation of data and some specific programming tricks in order to minimise the computation time, memory usage etc.

In our approach, the estimation of the SSIM index is proposed, which is based on the fast parallel computations of some of the local SSIM values leading to the partial quality map of the image. The size of the blocks depend on the capabilities of the specific graphic card and the size of the sliding window has been set to 8×8 pixels, assuming the rectangular window, similarly as in the Universal Image Quality Index [7] - the "predecessor" of the SSIM index. The abandonment of the use of the 11×11 pixels Gaussian window has been caused by its size (8 pixels is much more convenient size for some fast implementations) and the necessity of some additional multiplications for the non-rectangular window.

For the testing purposes the NVIDIA CUDA computing platform has been used with the G80 core (GeForce 8800 GTS: 128 stream processors, 650MHz core clock, 1625MHz shader clock, 1944MHz memory data rate, 256-bit memory interface). The CUDA gives the ability of writing programs using the modified C language what is a great advantage over the shader language. The cost is that the obtained code is not as fast as in an optimal shader based implementation. The organization of the thread block has been the 1×32 pixels and a single thread has been assigned to the single pixel column.

Thread blocks are organized horizontally and for each block the image column is processed using a pipeline approach. The results obtained for the four left and right boundary threads are not used in the further averaging, assuming the usage of the 8×8 pixels sliding window, since these local SSIM values are computed without the access to the next or previous 32 pixels wide blocks.

The most relevant advantage of such approach is the single memory fetch for each row, what is especially important due to the main bottleneck of the fast image and video processing systems, which is related to the device memory transfers. During the computations of the sums (further used during the variance calculations) the fast shared memory is used and the threads accesses are synchronized in order to fulfil the coalescence requirements.

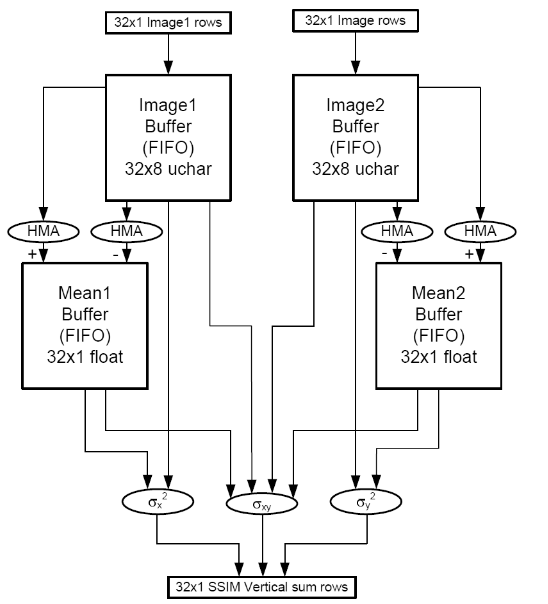

Fig. 1. Organisation of the calculations
Rys. 1. Organizacja obliczeń

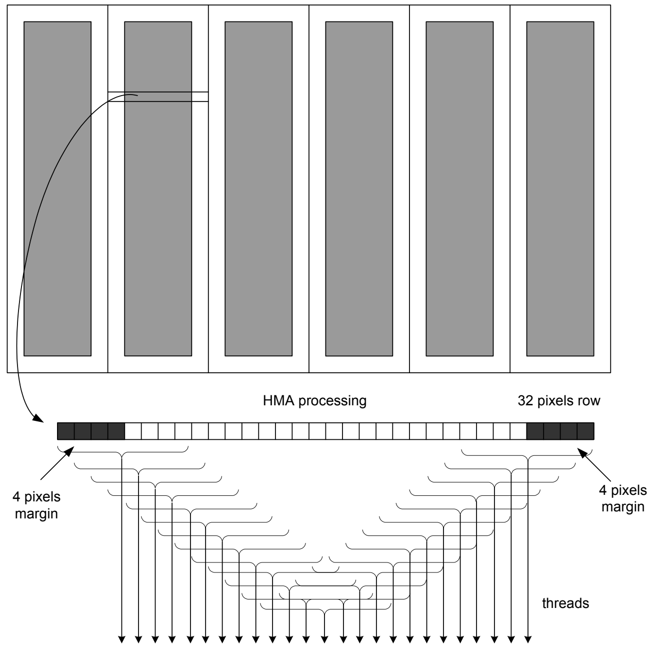

Fig. 2. The idea of the HMA processing
Rys. 2. Zasada przetwarzania HMA

The horizontal operations (for calculations of the mean values) are performed without the pipelines so these operations are not computationally optimal. Nevertheless, they do not require the device memory transfers, which would take much more time than these mathematical operations. Finally, the SSIM values calculated for all columns are returned as the float numbers and aggregated by the host processor.

Each block of 32 threads uses four blocks of shared memory, two of them are used for the buffering of images in order to prevent the unnecessary memory accesses (multiple reading of the same pixel's value). For the reduction of the number of the move operations the software based FIFO queuing has been applied. Calculations of the mean value for each 8×8 pixels block are performed using two additional 32×1 bytes buffers containing the vertical mean values. The Horizontal Moving Average (HMA) calculations have been applied in parallel mode, so each thread can use the same fragments of the memory independently. The overall organisation of the calculations is presented in Fig. 1 and the idea of the HMA filter processing is illustrated in Fig. 2.

4. Results

As the test platform, some images from the well known LIVE Image Quality Database Release 2 [8] have been used in their greyscale versions. Such a database contains nearly 1000 images with five types of distortions: JPEG and JPEG2000 compression, contamination by the white noise, Gaussian blurred images and some of the images transmitted over the simulated fast fading Rayleigh channel. Omitting about 200 original images, also present in the database, the SSIM index has been calculated using the original method and our estimation. Finally, the estimation error has been computed as the mean absolute difference for five types of distortions as well as for the whole database (without the original images). Additionally, the Pearson linear correlation coefficients of the obtained SSIM estimates and the exact values with the Differential Mean Opinion Score (DMOS) values derived from the LIVE database have been calculated (Table 3).

The mean absolute differences between the exact SSIM values obtained using the 8×8 pixels rectangular window and our implementation are illustrated in Table 1, while the absolute differences between the effects of using 11×11 pixels Gaussian window and the 8×8 rectangular one are shown in Table 2.

Tab. 1. Mean absolute differences between the SSIM index for the 8×8 pixels rectangular sliding window and our estimation
Tab. 1. Średnie różnice bezwzględne pomiędzy wartościami wskaźnika SSIM dla przesuwnego okna prostokątnego o rozmiarze 8×8 pikseli a wynikami proponowanej estymacji

Distortion type	JPEG 2000	JPEG	White noise	Gaussia n blur	Fast fading Rayleigh channel	**All**
Mean Absolute Difference	0.0014	0.0012	0.0017	0.0014	0.0019	**0.0015**

Tab. 2. Mean absolute differences between the SSIM index for the 8×8 pixels rectangular sliding window and 11×11 Gaussian one
Tab. 2. Średnie różnice bezwzględne pomiędzy wartościami wskaźnika SSIM dla przesuwnego okna prostokątnego o rozmiarze 8×8 pikseli a wynikami dla okna Gaussa o rozmiarze 11×11 pikseli

Distortion type	JPEG 2000	JPEG	White noise	Gaussia n blur	Fast fading Rayleigh channel	**All**
Mean Absolute Difference	0.0136	0.0172	0.0238	0.0177	0.0158	**0.0175**

Tab. 3. The linear correlation coefficients between the SSIM index values and the subjective evaluation (DMOS values)
Tab. 3. Współczynniki korelacji liniowej wskaźnika SSIM z ocenami subiektywnymi (wartościami DMOS)

Distortion / window	JPEG 2000	JPEG	White noise	Gaussian blur	Fast fading Rayleigh channel	**All**
8×8 rectangular	0.9034	0.8464	0.9664	0.8875	0.9047	**0.7486**
11×11 Gaussian	0.8977	0.8507	0.9644	0.8494	0.9011	**0.7367**
Our estimation	0.9015	0.8457	0.9666	0.8870	0.9034	**0.7471**

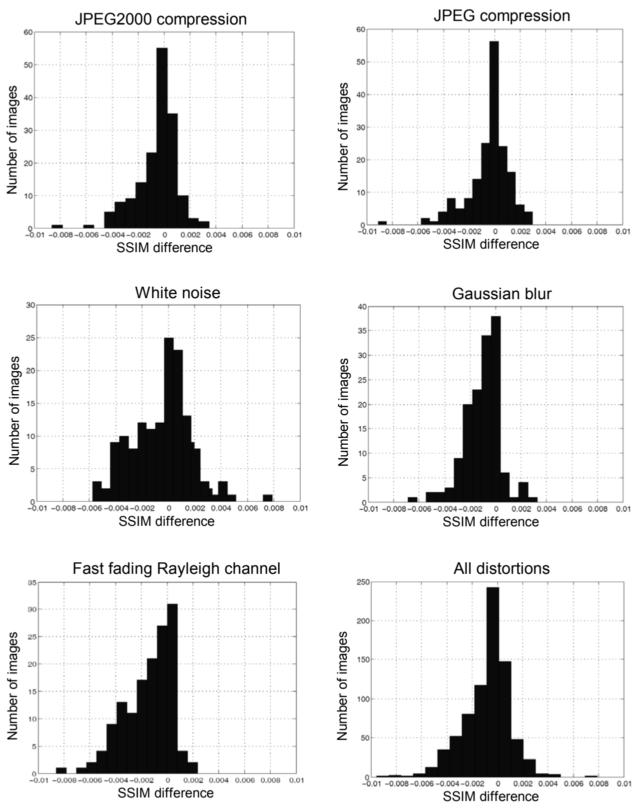

Fig. 3. Distributions of the differences between the SSIM values and their estimations for various types of distortions
Rys. 3. Rozkłady różnic pomiędzy wartościami wskaźnika SSIM a ich estymacjami dla różnych rodzajów zniekształceń

Analysing the obtained results it can be easily noticed that the differences between the results obtained using the proposed technique and the exact SSIM values are much smaller than the differences caused by the change of the type and size of the sliding window. The linear correlation coefficients with the DMOS values obtained for the SSIM values for both types of sliding windows as well as for our method are similar, so the approximation errors of the method presented in the paper can be considered as the acceptable ones.

Analysing the histograms presented in Fig. 3 it can be noticed that the obtained SSIM values are underestimated for most of the images, especially for the Gaussian blurred ones and those transmitted over the simulated fast fading Rayleigh channel. The images without any distortions (about 20% of the database) have not been considered since their estimated values are always equal to 1 (there are no differences in any fragment of both images).

5. Conclusions

Proposed method can be treated as an efficient solution, especially as the implementation of the fast estimation of the image quality algorithm for the video quality assessment purposes. Such an algorithm can be based on the Video SSIM – one of the further extensions of the SSIM index. The application of the modified quality estimation based on the proposed approach, as well as on the statistical method [9], for the fast approximation of the Video SSIM index also for the High Definition video streams, is planned as one of the directions of our future work.

The speed of 56-86 Mpix/s and the GPU time about 0.95 ms per image obtained during the calculations of the results presented in this paper, may be very promising for the real-time video quality estimation purposes. Comparing the execution times of our estimation to the usage of the standard Matlab implementation of the SSIM index (using the Intel Q6600 2.4 GHz processor), the average performance of the algorithm is about 150 times faster.

6. References

[1] Eskicioglu A.: Quality Measurement for Monochrome Compressed Images in the Past 25 Years. Proc. of IEEE International Conference on Acoustics Speech and Signal Processing, pp. 1907–1910, Istanbul, Turkey, 2000.
[2] Meesters L., Martens J.-B.: A Single-Ended Blockiness Measure for JPEG-Coded Images. Signal Processing vol. 82 no. 3, pp. 369–387, 2002.
[3] Marziliano P., Dufaux F., Winkler S., Ebrahimi T.: A No-Reference Perceptual Blur Metric. Proceedings of IEEE International Conference on Image Processing, pp. 57–60, Rochester, USA, 2002.
[4] Shnayderman A., Gusev A., Eskicioglu A.: An SVD-Based Gray-Scale Image Quality Measure for Local and Global Assessment. IEEE Transactions on Image Processing vol. 15 no. 2, pp. 422–429, 2006.
[5] Li X.: Blind Image Quality Assessment. Proceedings of IEEE Int. Conference on Image Processing, pp. 449–452, 2002.
[6] Wang Z., Bovik A.C., Sheikh H., Simoncelli E.: Image Quality Assessment: From Error Measurement to Structural Similarity. IEEE Trans. on Image Processing vol. 13 no. 4, pp. 600–612, 2004.
[7] Wang Z., Bovik A.C.: A Universal Image Quality Index. IEEE Signal Processing Letters vol. 9 no. 3, pp. 81–84, 2002.
[8] Sheikh H.R., Wang Z., Cormack L., Bovik A.C.: LIVE Image Quality Assessment Database Release 2. http://live.ece.utexas.edu/research/quality
[9] Okarma K., Lech P.: A Statistical Reduced-Reference Approach to Digital Image Quality Assessment. Lecture Notes in Computer Science vol. 5337, pp. 43–54, Springer-Verlag, Berlin Heidelberg, 2009.

otrzymano / received: 07.05.2010
przyjęto do druku / accepted: 04.06.2010 *artykuł recenzowany*