

GmImgProc

Alexandra Olteanu

SCPD

Alexandru Ștefănescu

SCPD

Image processing for OCR and content extraction

Implies a set of costly operations like:

- Binarization – the resulted image is formed by two classes separated by a threshold t : foreground and background
- Noise detection
- Clustering for creating meta-entities (ex: characters, lines)
- Detection and/or classification of page components: line, table, character, frame, image, etc
- Skew Detection and correction
- Geometrical distortions correction
- And the list may continue

For this project our aim was to implement methods for image binarization and to try different solutions to improve our results and to remove noise:

- Otsu 1D
- Otsu 2D
- Niblack
- Color Uniformization

Input (grayscale)

recorded we were again married on August 28, 1905, in Brooklyn.

"The witnesses were William Moore and a Mr. Toy. In September, a few days previous to her brother's marriage, Mrs. Carlton had some words with her parents, and she and her brother left home, the brother occupying my room at 198 Sands street, until his marriage. After one week at the Hotel St. George Mrs. Carlton went to board in East New York. She remained there until November, when she joined me in Washington. She visited Brooklyn in February, remaining one month.

"In May Mrs. Carlton, in putting on a felt slipper with a rubber sole, stamped on the floor, and a needle sticking in the matting ran through the sole into her foot and broke off. It was extracted. Two days later she went with me to the Zoological Gardens. Her foot paining her, she removed her shoe and stocking and walked barefooted for two

GENERAL BLACKMAR

Assumes Headship of the
Grand Army of Republic.

WILL ISSUE GENERAL ORDER

New Commander-in-Chief Expresses
Great Grief at Death of
Predecessor.

...ing ... to ... we were again married on August 28, 1907, in Brooklyn.

"The witnesses were William Moore and a Mr. Toy. In September, a few days previous to her brother's marriage, Mrs. Carlton had some words with her parents, and she and her brother left home, the brother occupying my room at 198 Sands street, until his marriage. After one week at the Hotel St. George Mrs. Carlton went to board in East New York. She remained there until November, when she joined me in Washington. She visited Brooklyn in February, remaining one month.

"In May Mrs. Carlton, in putting on a felt slipper with a rubber sole, stamped on the floor, and a needle sticking in the matting ran through the sole into her foot and broke off. It was extracted. Two days later she went with me to the Zoological Gardens. Her foot, painful, she removed her shoe and stocking and walked barefooted for two hours. Returning home, the foot ...

GENERAL BLACKMAR

Assumes Headship of the Grand Army of Republic.

WILL ISSUE GENERAL ORDER

New Commander-in-Chief Expresses Great Grief at Death of Predecessor.

recorded we were again married on August 23, 1902, in Brooklyn. The witnesses were William Moore and a Mr. Toy. In September, a few days previous to her brother's marriage, Mrs. Carlton had some words with her parents, and she and her brother left home, the brother occupying my room at 198 Sands street, until his marriage. After one week at the Hotel St. George Mrs. Carlton went to board in East New York. She remained there until November, when she joined me in Washington. She visited Brooklyn in February, remaining one month.

"In May Mrs. Carlton, in putting on a felt slipper with a rubber sole, stamped on the floor, and a needle sticking in the matting ran through the sole into her foot and broke off. It was extracted. Two days later she went with me to the Zoological Gardens. Her foot paining her, she removed her shoe and stocking and walked barefooted for two

GENERAL BLACKMAR

Assumes Headship of the
Grand Army of Republic.

WILL ISSUE GENERAL ORDER

New Commander-in-Chief Expresses
Great Grief at Death of
Predecessor.

Color Uniformization (intermediate grayscale)

recorded we were again married on August 28, 1903, in Brooklyn.

"The witnesses were William Moore and a Mr. Toy. In September, a few days previous to her brother's marriage, Mrs. Carlton had some words with her parents, and she and her brother left home, the brother occupying my room at 198 Sands street, until his marriage. After one week at the Hotel St. George Mrs. Carlton went to board in East New York. She remained there until November, when she joined me in Washington. She visited Brooklyn in February, remaining one month.

"In May Mrs. Carlton, in putting on a felt slipper with a rubber sole, stamped on the floor, and a needle sticking in the matting ran through the sole into her foot and broke off. It was extracted. Two days later she went with me to the Zoological Gardens. Her foot paining her, she removed her shoe and stocking and walked barefooted for two

GENERAL BLACKMAR

Assumes Headship of the
Grand Army of Republic.

WILL ISSUE GENERAL ORDER

New Commander-in-Chief Expresses
Great Grief at Death of
Predecessor.

Color Uniformization + Otsu 1D

recorded we were again married on August 23, 1902, in Brooklyn.

"The witnesses were William Moore and a Mr. Toy. In September, a few days previous to her brother's marriage, Mrs. Carlton had some words with her parents, and she and her brother left home, the brother occupying my room at 198 Sands street, until his marriage. After one week at the Hotel St. George Mrs. Carlton went to board in East New York. She remained there until November, when she joined me in Washington. She visited Brooklyn in February, remaining one month.

"In May Mrs. Carlton, in putting on a felt slipper with a rubber sole, stamped on the floor, and a needle sticking in the matting ran through the sole into her foot and broke off. It was extracted. Two days later she went with me to the Zoological Gardens. Her foot paining her, she removed her shoe and stockings and walked barefooted for two

GENERAL BLACKMAR

Assumes Headship of the
Grand Army of Republic.

WILL ISSUE GENERAL ORDER

New Commander-in-Chief Expresses
Great Grief at Death of
Predecessor.

Otsu 1D

Otsu method is based on a very simple idea: find the threshold that minimizes the weighted within-class variance, defined as a weighted sum of variances of the two classes.

$$\sigma_w^2(t) = \omega_1(t)\sigma_1^2(t) + \omega_2(t)\sigma_2^2(t)$$

- ω_i are the probabilities of the two classes separated by a threshold t
- σ_i^2 are variances of these classes

This can be demonstrated to be the same as maximizing the between-class variance.

$$\sigma_b^2(t) = \sigma^2 - \sigma_w^2(t) = \omega_1(t)\omega_2(t) [\mu_1(t) - \mu_2(t)]^2$$

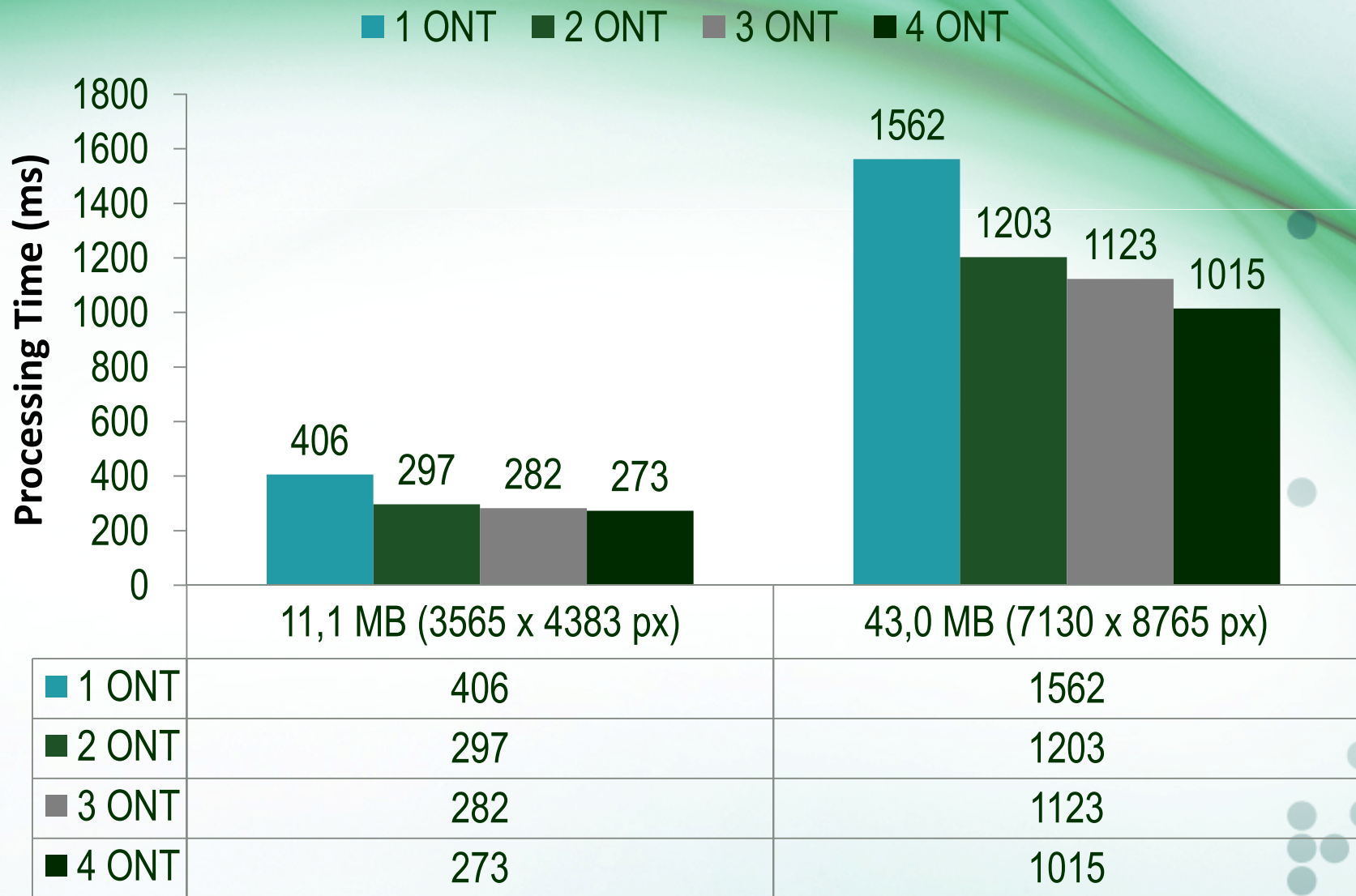
where μ_i represent the class means

Implementation: we simply compute the class means and their probabilities

Advantages:

This method is fast because operates directly on the gray level histogram. It provides very good results when the numbers of pixels in each class are close to each other.

Otsu 1D parallelization using OpenMP



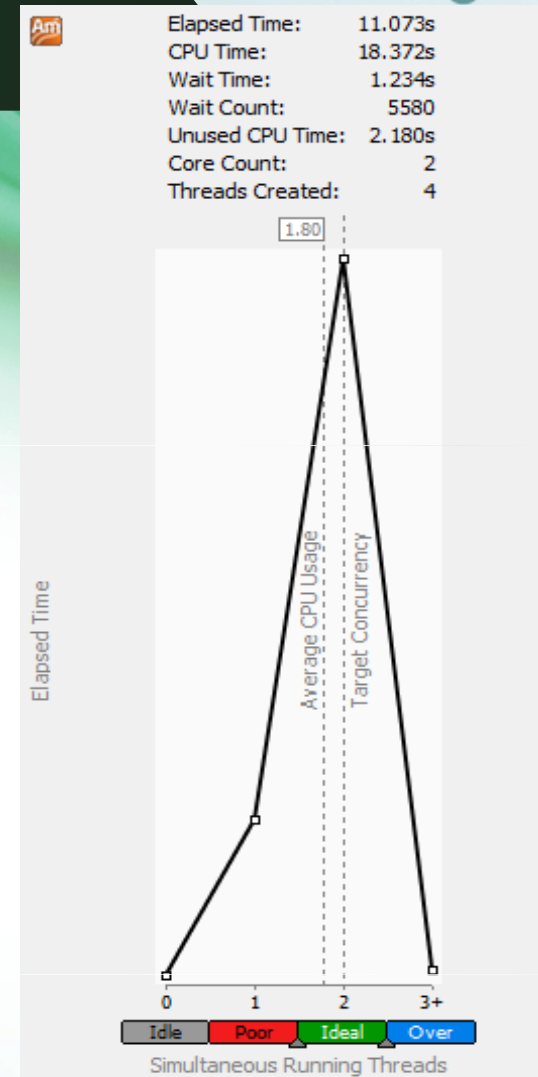
ONT = OMP_NUM_THREADS

Observations

- ✓ granularity => good
- ✓ fast serial algorithm => low speedup
- ✓ load balance => good

Profiling results:

- ✓ CPU time utilization => very good
- ✓ hot spot => histogram computation



recorded we were again married on August 28, 1897, in Brooklyn.

"The witnesses were William Moore and a Mr. Toy. In September, a few days previous to her brother's marriage, Mrs. Carlton had some words with her parents, and she and her brother left home, the brother occupying my room at 198 Sands street, until his marriage. After one week at the Hotel St. George Mrs. Carlton went to board in East New York. She remained there until November, when she joined me in Washington. She visited Brooklyn in February, remaining one month.

"In May Mrs. Carlton, in putting on a felt slipper with a rubber sole, stamped on the floor, and a needle sticking in the matting ran through the sole into her foot and broke off. It was extracted. Two days later she went with me to the Zoological Gardens. Her foot being better, she removed her shoe and stocking and walked barefooted for two

GENERAL BLACKMAR

Assumes Headship of the Grand Army of Republic.

WILL ISSUE GENERAL ORDER

New Commander-in-Chief Expresses Great Grief at Death of Predecessor.



recorded we were again married in August 28, 1897, in Brooklyn.

"The witnesses were William Moore and a Mr. Toy. In September, a few days previous to her brother's marriage, Mrs. Carlton had some words with her parents, and she and her brother left home, the brother occupying my room at 198 Sands street, until his marriage. After one week at the Hotel St. George Mrs. Carlton went to board in East New York. She remained there until November, when she joined me in Washington. She visited Brooklyn in February, remaining one month.

"In May Mrs. Carlton, in putting on a felt slipper with a rubber sole, stamped on the floor, and a needle sticking in the matting ran through the sole into her foot and broke off. It was extracted. Two days later she went with me to the Zoological Gardens. Her foot being better, she removed her shoe and stocking and walked barefooted for two

GENERAL BLACKMAR

Assumes Headship of the Grand Army of Republic.

WILL ISSUE GENERAL ORDER

New Commander-in-Chief Expresses Great Grief at Death of Predecessor.



of about five months, through a letter of introduction from Mr. Laughlin. We were married secretly the next month. Owing to a failure to have the marriage recorded we were again married on August 28, 1897, in Brooklyn.

"The witnesses were William Moore and a Mr. Toy. In September, a few days previous to her brother's marriage, Mrs. Carlton had some words with her parents, and she and her brother left home, the brother occupying my room at 198 Sands street, until his marriage. After one week at the Hotel St. George Mrs. Carlton went to board in East New York. She remained there until November, when she joined me in Washington. She visited Brooklyn in February, remaining one month.

"In May Mrs. Carlton, in putting on a felt slipper with a rubber sole, stamped on the floor, and a needle sticking in the matting ran through the sole into her foot and broke off. It was extracted. Two days later she went with me to the Zoological Gardens. Her foot being better, she removed her shoe and stocking and walked barefooted for two

KING SUCCEEDS GENERAL BLACKMAR

Assumes Headship of the Grand Army of Republic.

WILL ISSUE GENERAL ORDER

New Commander-in-Chief Expresses Great Grief at Death of Predecessor.

Otsu 2D

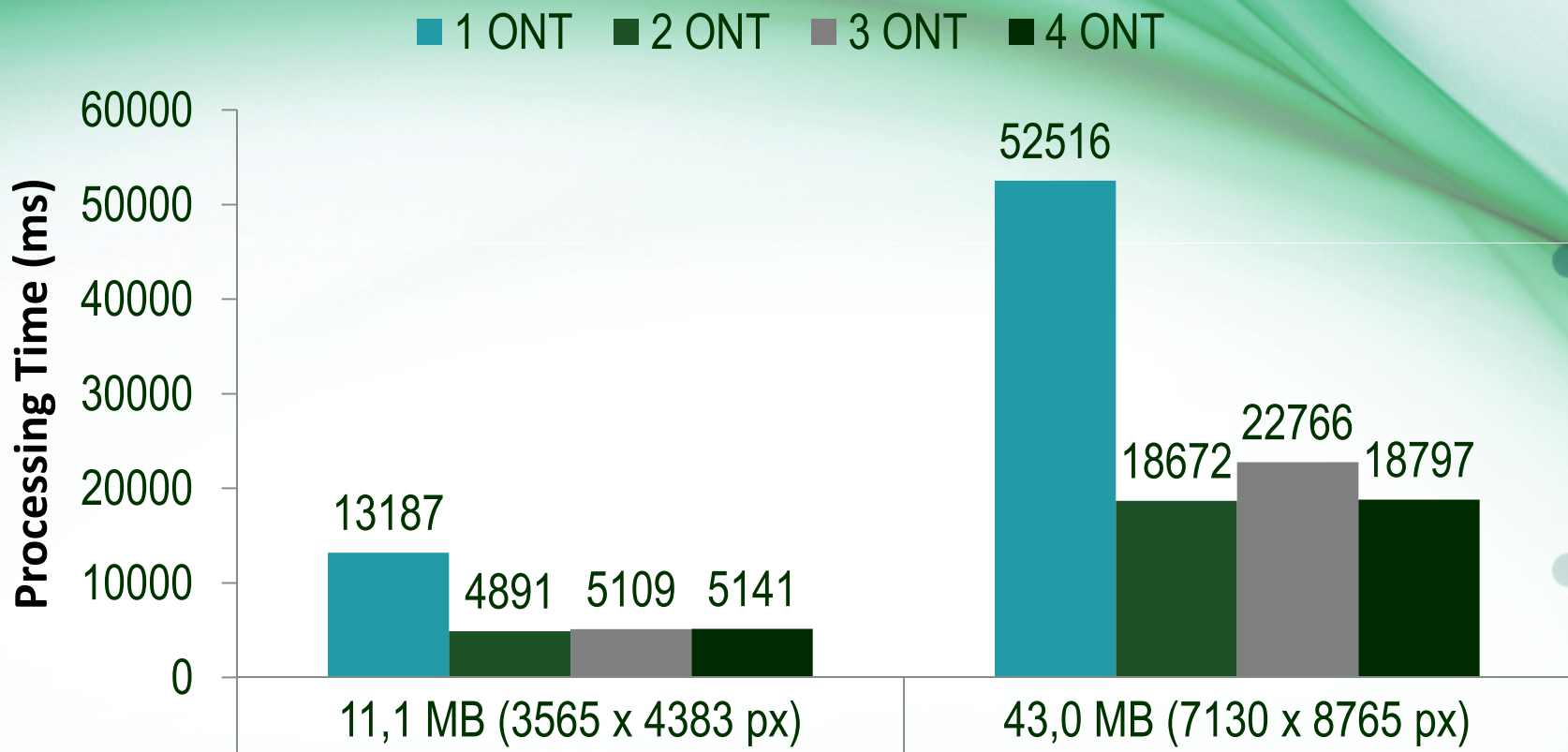
- considers both the gray value of a pixel and the average gray value of its neighborhood
- pixels are partitioned into two classes by a threshold pair (s, t) , where s is the gray level of a pixel and t is its local average gray level
- Otsu 2D method is based on the same principles as classic Otsu
- it uses the two-dimensional histogram
- by using the trace of the between-class variance matrix as the measurement of between-class variance and maximizing it, we select the corresponding threshold (s, t) for this value.

Implementation: we use a fast recursive algorithm to implement Otsu 2D method.

Advantages:

- more robust to noise than Otsu 1D
- provides very good results when the numbers of pixels in each class are close to each other

Otsu 2D parallelization using OpenMP



	11,1 MB (3565 x 4383 px)	43,0 MB (7130 x 8765 px)
■ 1 ONT	13187	52516
■ 2 ONT	4891	18672
■ 3 ONT	5109	22766
■ 4 ONT	5141	18797

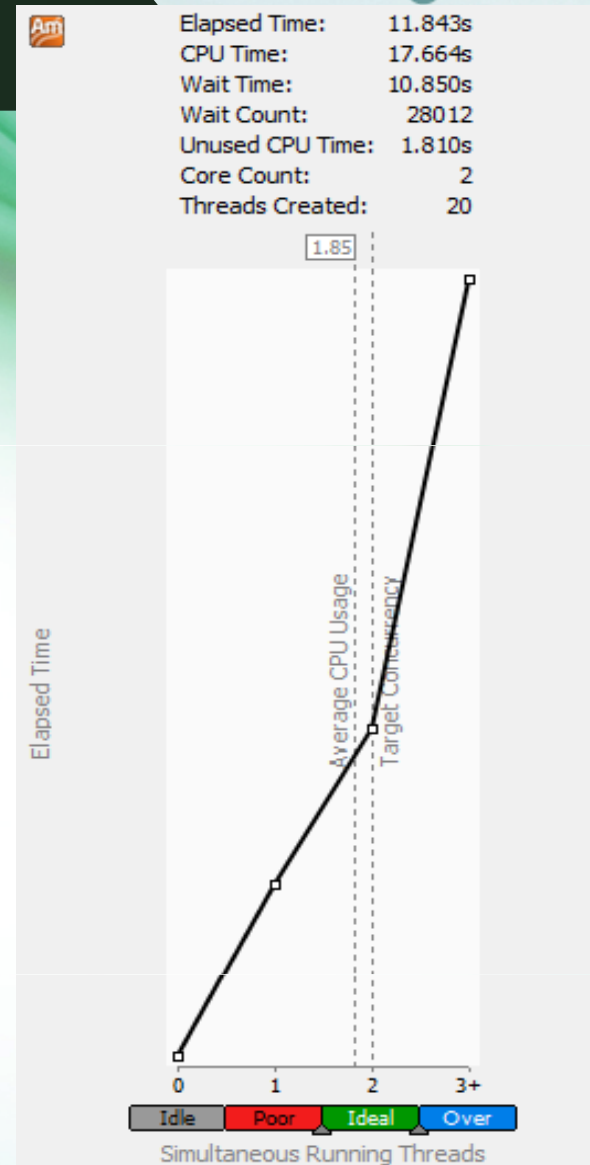
ONT = OMP_NUM_THREADS

Observations

- ✓ in OpenMP 2.5 and in earlier versions the nested parallelism can be exploit only by using nested parallel regions
- ✓ granularity => good
- ✓ atomic constructs introduce synchronization overhead
- ✓ load imbalance – conditional instructions

Profiling results:

- ✓ hot spot => local average gray level computation



recorded we were again married on August 25, 1902, in Brooklyn.

"The witnesses were William Moore and a Mr. Toy. In September, a few days previous to her brother's marriage, Mrs. Carlton had some words with her parents, and she and her brother left home, the brother occupying my room at 138 Sands street, until his marriage. After one week at the Hotel St. George Mrs. Carlton went to board in East New York. She remained there until November, when she joined me in Washington. She visited Brooklyn in February, remaining one month.

"In May Mrs. Carlton, in putting on a felt slipper with a rubber sole, stamped on the floor, and a needle sticking in the matting ran through the sole into her foot and broke off. It was extracted. Two days later she went with me to the Zoological Gardens. Her foot paining her, she removed her shoe and stockings and walked barefooted for two

GENERAL BLACKMAR

Assumes Headship of the Grand Army of Republic.

WILL ISSUE GENERAL ORDER

New Commander-in-Chief Expresses Great Grief at Death of Predecessor.



...the witnesses were William Moore and a Mr. Toy. In September, a few days previous to her brother's marriage, Mrs. Carlton had some words with her parents, and she and her brother left home, the brother occupying my room at 138 Sands street, until his marriage. After one week at the Hotel St. George Mrs. Carlton went to board in East New York. She remained there until November, when she joined me in Washington. She visited Brooklyn in February, remaining one month.

"In May Mrs. Carlton, in putting on a felt slipper with a rubber sole, stamped on the floor, and a needle sticking in the matting ran through the sole into her foot and broke off. It was extracted. Two days later she went with me to the Zoological Gardens. Her foot paining her, she removed her shoe and stockings and walked barefooted for two

KING SUCCEEDS GENERAL BLACKMAR

Assumes Headship of the Grand Army of Republic.

WILL ISSUE GENERAL ORDER

New Commander-in-Chief Expresses Great Grief at Death of Predecessor.



...the witnesses were William Moore and a Mr. Toy. In September, a few days previous to her brother's marriage, Mrs. Carlton had some words with her parents, and she and her brother left home, the brother occupying my room at 138 Sands street, until his marriage. After one week at the Hotel St. George Mrs. Carlton went to board in East New York. She remained there until November, when she joined me in Washington. She visited Brooklyn in February, remaining one month.

"In May Mrs. Carlton, in putting on a felt slipper with a rubber sole, stamped on the floor, and a needle sticking in the matting ran through the sole into her foot and broke off. It was extracted. Two days later she went with me to the Zoological Gardens. Her foot paining her, she removed her shoe and stockings and walked barefooted for two

KING SUCCEEDS GENERAL BLACKMAR

Assumes Headship of the Grand Army of Republic.

WILL ISSUE GENERAL ORDER

New Commander-in-Chief Expresses Great Grief at Death of Predecessor.

Otsu 2D tuned with OpenCL

CPU: Intel Core 2 Duo, T7500, 2.2 GHz, 1.99 GB RAM - Display Mobile Intel(R) 965

	OpenCL (seconds)			Serial (seconds)		
Image size	11 MB	44 MB	62 MB	11 MB	44 MB	62 MB
Transform data into CL structures	0.438	1.735	1.765			
Create CL context, buffers etc.	0.5	0.421	0.453			
Runtime	0.953	3.094	3.141	2.937	11.672	11.859
Copy computed data from CL structures into app structures	0.156	0.641	0.719			
TOTAL time	2.781	9.359	9.547	3.344	13.297	13.453

Helped by Iulian Milas 😊

Niblack

Niblack's method is a local threshold binarization method which involves computing for each pixel in the grayscale image the mean and standard deviation of the colors of the neighboring pixels in an area (window) of predefined size. The threshold for determining if the pixel will be black or white is computed by the formula:

$$T = \text{mean} + k * \text{stdev}$$

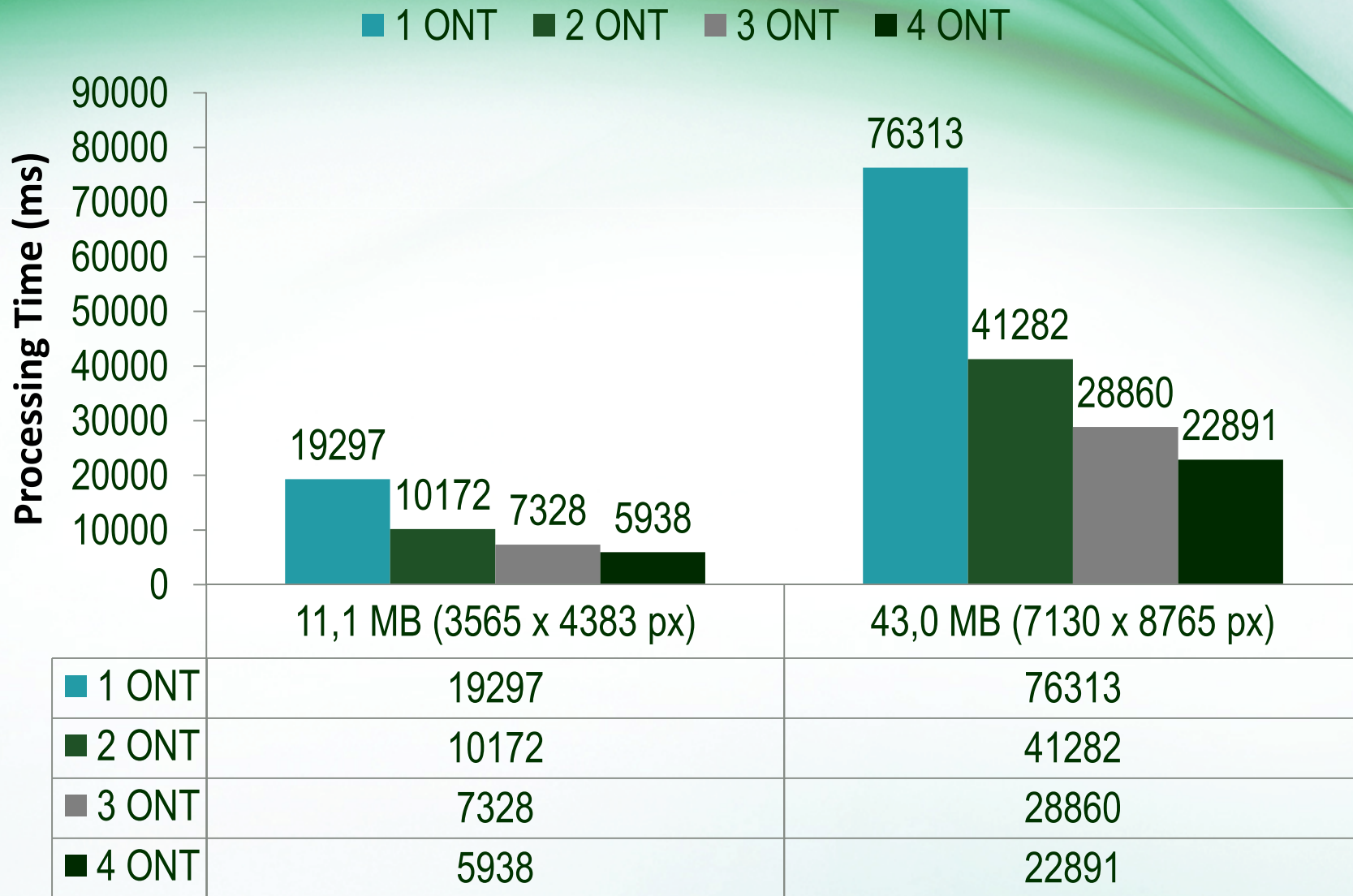
where k is a constant, preselected coefficient.

Implementation: We simply slide the window onto each pixel and recompute the mean and standard deviation.

Assessment:

- does not involve a unique threshold for the entire image and thus is adaptable for local particularities
- had worst results in our tests, despite being considered the best binarization method for text images
- is much slower than other methods.

Niblack parallelization using OpenMP



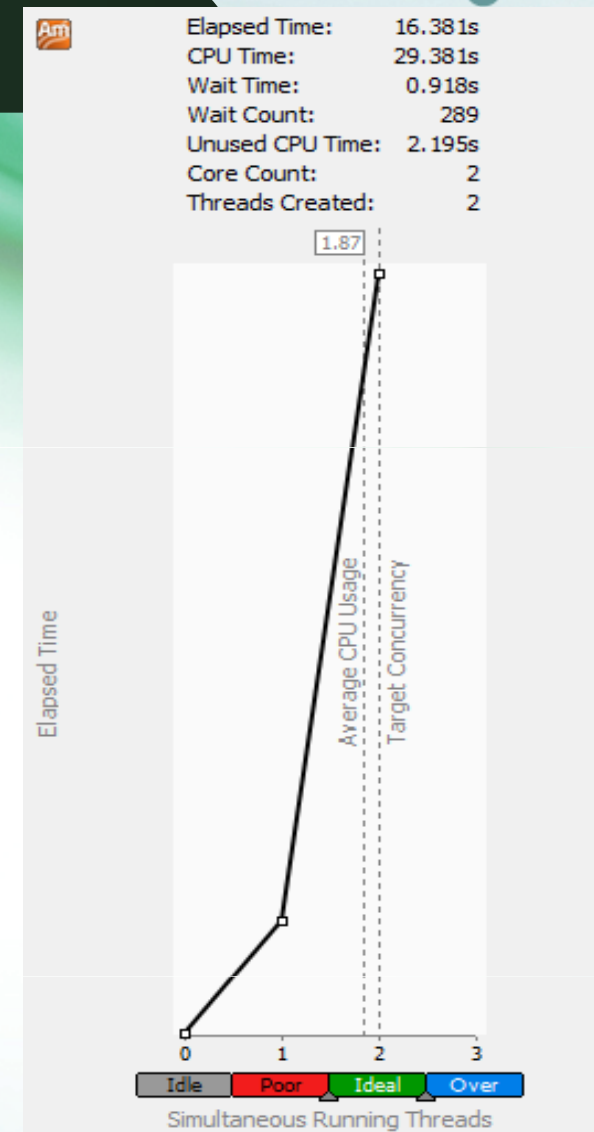
ONT = OMP_NUM_THREADS

Observations

- ✓ granularity => good
- ✓ trivial to parallelize
- ✓ each iteration is completely independent on the rest
- ✓ load balancing easy to achieve
- ✓ CPU wait time practically zero (except for I/O)

Profiling results:

- ✓ hot spot => computing mean and stdev



recorded we were again married on August 23, 1917, in Brooklyn.

"The witnesses were William Moore and a Mr. Toy. In September, a few days previous to her brother's marriage, Mrs. Carlton had some words with her parents, and she and her brother left home, the brother occupying my room at 198 Sands street, until his marriage. After one week at the Hotel St. George Mrs. Carlton went to board in East New York. She remained there until November, when she joined me in Washington. She visited Brooklyn in February, remaining one month.

"In May Mrs. Carlton, in putting on a felt slipper with a rubber sole, stamped on the floor, and a needle sticking in the matting ran through the sole into her foot and broke off. It was extracted. Two days later she went with me to the Zoological Gardens. Her foot paining her, she removed her shoe and stockings and walked barefooted for two

GENERAL BLACKMAR

Assumes Headship of the Grand Army of Republic.

WILL ISSUE GENERAL ORDER

New Commander-in-Chief Expresses Great Grief at Death of Predecessor.



recorded we were again married on August 23, 1917, in Brooklyn.

"The witnesses were William Moore and a Mr. Toy. In September, a few days previous to her brother's marriage, Mrs. Carlton had some words with her parents, and she and her brother left home, the brother occupying my room at 198 Sands street, until his marriage. After one week at the Hotel St. George Mrs. Carlton went to board in East New York. She remained there until November, when she joined me in Washington. She visited Brooklyn in February, remaining one month.

"In May Mrs. Carlton, in putting on a felt slipper with a rubber sole, stamped on the floor, and a needle sticking in the matting ran through the sole into her foot and broke off. It was extracted. Two days later she went with me to the Zoological Gardens. Her foot paining her, she removed her shoe and stockings and walked barefooted for two

GENERAL BLACKMAR

Assumes Headship of the Grand Army of Republic.

WILL ISSUE GENERAL ORDER

New Commander-in-Chief Expresses Great Grief at Death of Predecessor.

Image color uniformization

- is a **preprocessing technique** for grayscale images before the binarization step.
- creates an intermediate grayscale image which binarized with global methods like Otsu 1D or Otsu 2D produces better results than without the preprocessing
- the main idea is to “diffuse” color differences weighted by distances between neighboring pixels
- each pixel in the intermediate image is computed using the following sum, for the pixels in a neighboring area (window) of predefined size:

$$\sigma(x, y) = \sum_{(x', y') \in W} \frac{\text{color}(x', y') - \text{color}(x, y)}{\sqrt{(x - x')^2 + (y - y')^2}}$$

for every pixel of coordinated (x', y') located in the window W centered in (x, y)

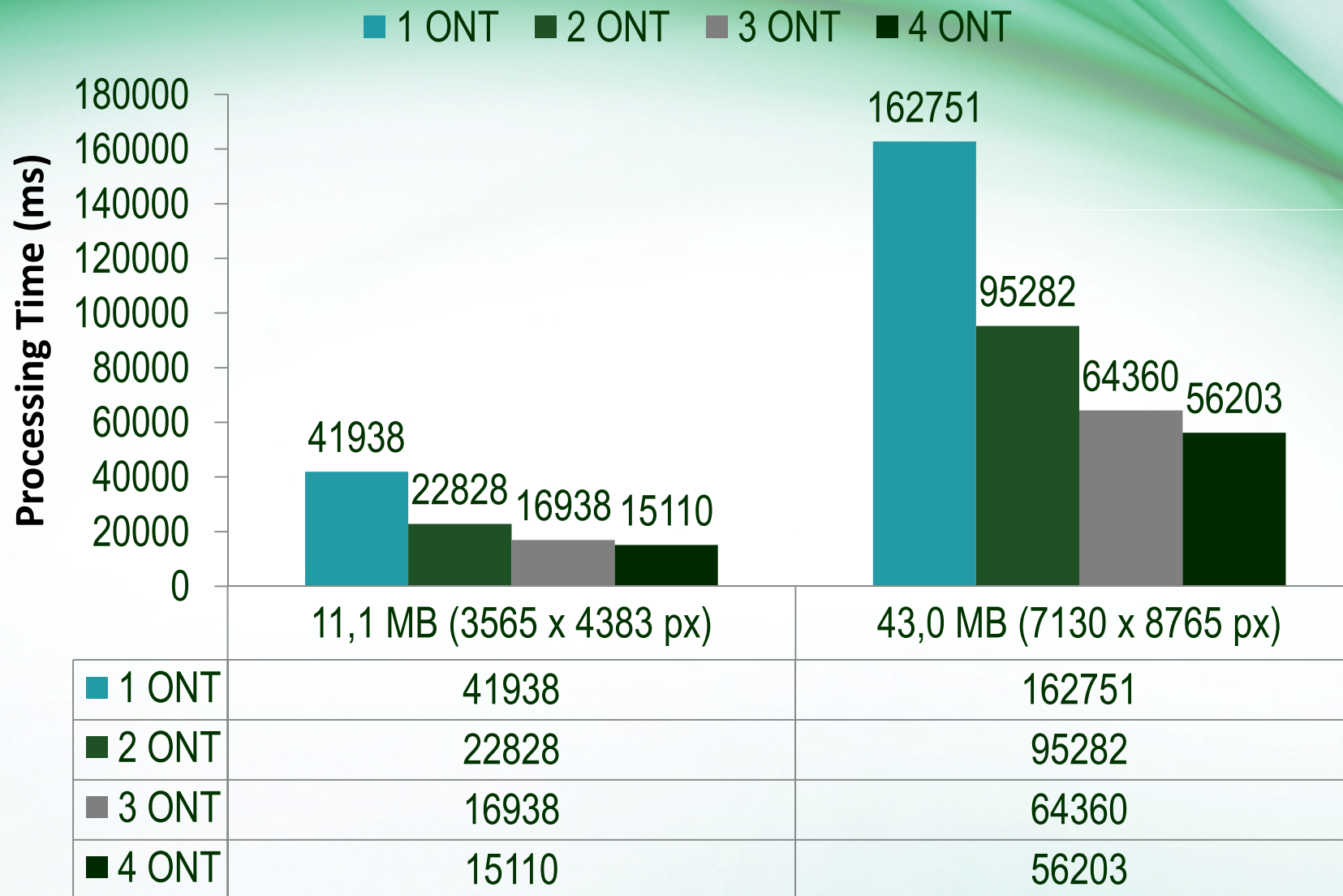
- in the end, the sums are scaled to 0 – 255 to correspond to a level of gray.

Implementation: We simply slide the window onto each pixel and recompute the mean and standard deviation.

Assessment:

- obtained a **great improvement in the image binarized with Otsu 1D and Otsu 2D using this preprocessing step** as compared to binarizing without it.
- similar to Niblack’s method, it is very slow.

Parallelization of Color Uniformization algorithm with OpenMP



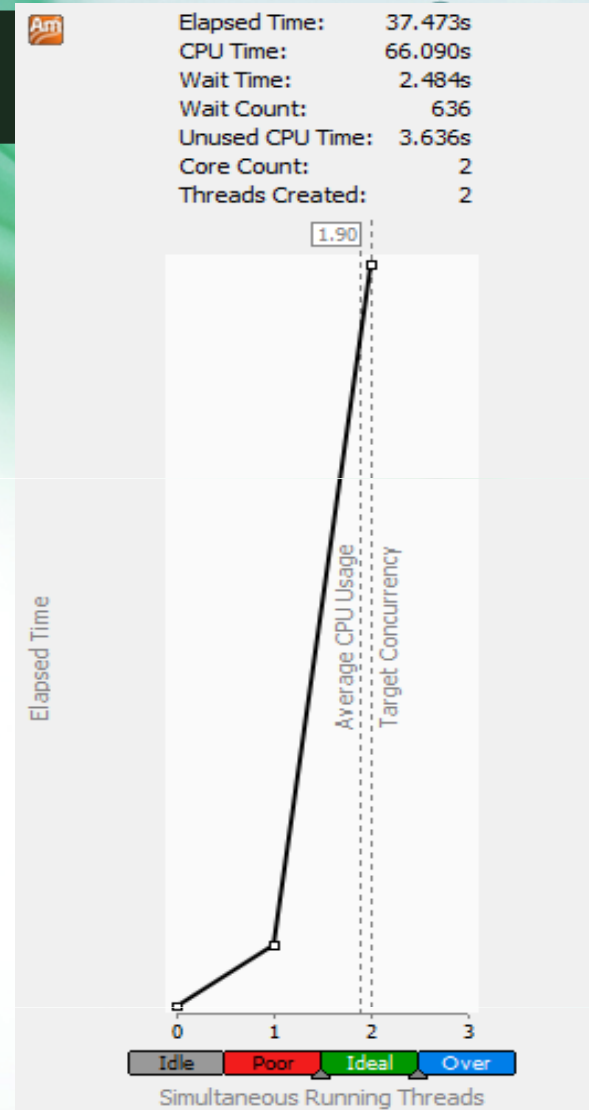
ONT = OMP_NUM_THREADS

Observations

- ✓ similar to Niblack's method
- ✓ granularity => good
- ✓ each iteration is completely independent on the rest
- ✓ CPU wait time practically zero (except for I/O)

Profiling results:

- ✓ hot spot => computing the sum



recorded we were again married on August 23, 1911, in Brooklyn. The witnesses were William Moore and a Mr. Toy. In September, a few days previous to her brother's marriage, Mrs. Carlton had some words with her parents, and she and her brother left home, the brother occupying my room at 118 Sands street, until his marriage. After one week at the Hotel St. George Mrs. Carlton went to board in East New York. She remained there until November, when she joined me in Washington. She visited Brooklyn in February, remaining one month.

"In May Mrs. Carlton, in putting on a felt slipper with a rubber sole, stamped on the floor, and a needle sticking in the matting ran through the sole into her foot and broke off. It was extracted. Two days later she went with me to the Zoological Gardens. Her foot paining her, she removed her shoe and stockings and waited here for two

GENERAL BLACKMAR

Assumes Headship of the Grand Army of Republic.

WILL ISSUE GENERAL ORDER

New Commander-in-Chief Expresses Great Grief at Death of Predecessor.



recorded we were again married on August 23, 1911, in Brooklyn. The witnesses were William Moore and a Mr. Toy. In September, a few days previous to her brother's marriage, Mrs. Carlton had some words with her parents, and she and her brother left home, the brother occupying my room at 118 Sands street, until his marriage. After one week at the Hotel St. George Mrs. Carlton went to board in East New York. She remained there until November, when she joined me in Washington. She visited Brooklyn in February, remaining one month.

"In May Mrs. Carlton, in putting on a felt slipper with a rubber sole, stamped on the floor, and a needle sticking in the matting ran through the sole into her foot and broke off. It was extracted. Two days later she went with me to the Zoological Gardens. Her foot paining her, she removed her shoe and stockings and waited here for two

GENERAL BLACKMAR

Assumes Headship of the Grand Army of Republic.

WILL ISSUE GENERAL ORDER

New Commander-in-Chief Expresses Great Grief at Death of Predecessor.



recorded we were again married on August 23, 1911, in Brooklyn. The witnesses were William Moore and a Mr. Toy. In September, a few days previous to her brother's marriage, Mrs. Carlton had some words with her parents, and she and her brother left home, the brother occupying my room at 118 Sands street, until his marriage. After one week at the Hotel St. George Mrs. Carlton went to board in East New York. She remained there until November, when she joined me in Washington. She visited Brooklyn in February, remaining one month.

"In May Mrs. Carlton, in putting on a felt slipper with a rubber sole, stamped on the floor, and a needle sticking in the matting ran through the sole into her foot and broke off. It was extracted. Two days later she went with me to the Zoological Gardens. Her foot paining her, she removed her shoe and stockings and waited here for two

GENERAL BLACKMAR

Assumes Headship of the Grand Army of Republic.

WILL ISSUE GENERAL ORDER

New Commander-in-Chief Expresses Great Grief at Death of Predecessor.

Conclusions (what we've learned!)

OpenMP:

- reducing false sharing => making use of private data as much as possible
- parallelize at the highest level possible
- make parallel regions as large as possible to reduce parallelization overhead
- use **CRITICAL** sections for fine-grained locking
- the work should be distributed evenly across the threads
- **NOWAIT** clause => to eliminate redundant or unnecessary barriers
- we also compiled with Intel C++ but we obtained similar results (but we hated it)
- profiling was made on a two core system

OpenCL:

- very fast
- does not support bidimensional arrays
- setting up the OpenCL kernel for running is very complicated

Q&A

..and autographs