

# Algorithmique et Programmation Avancée

Anh-Kiet DUONG

May 22, 2024

## 1 Introduce

My name: Anh-Kiet DUONG.

Program: Master (M1) CRYPTIS.

Source code: [github](#)

Initially, my group had 4 members. But I texted them and they never replied. They don't seem to be working on this project. Then I decided to work alone and they agreed to it. With experience after many times participating in **ICPC regional (ASEAN-Vietnam)**, the algorithm is not too complicated for me. So it is not impossible for me to complete this project alone.

## 2 Required work

### 2.1 Experiment with algorithms (a) and (b) provided with different images, and different parameters.

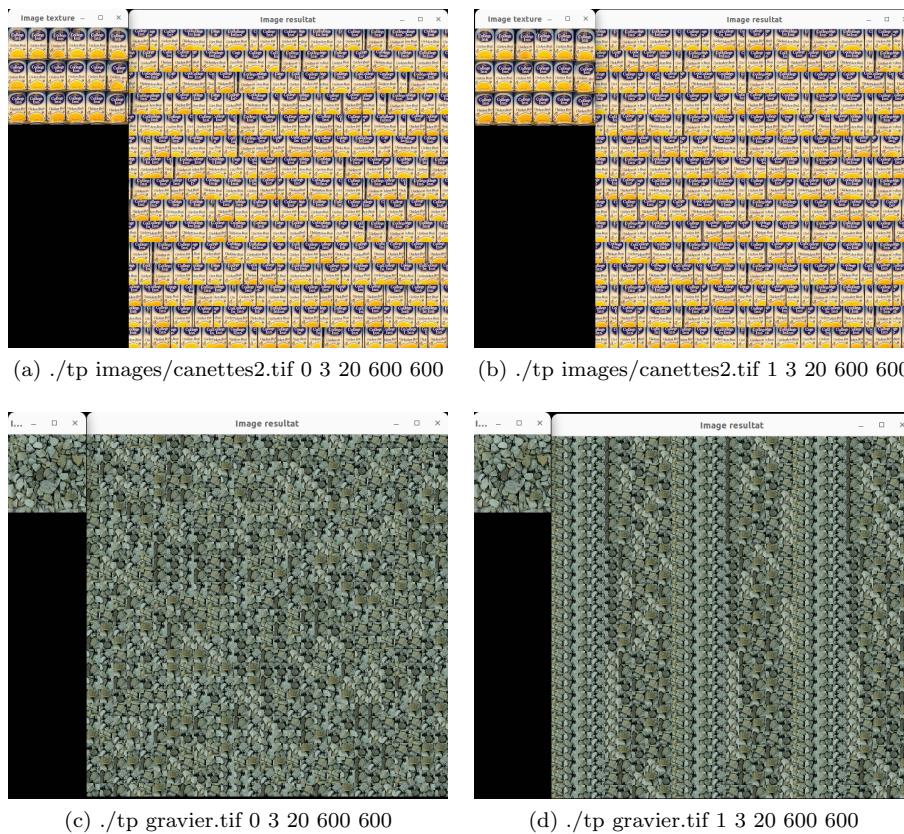


Figure 1: Experiment with algorithms (a) and (b) provided with different images.

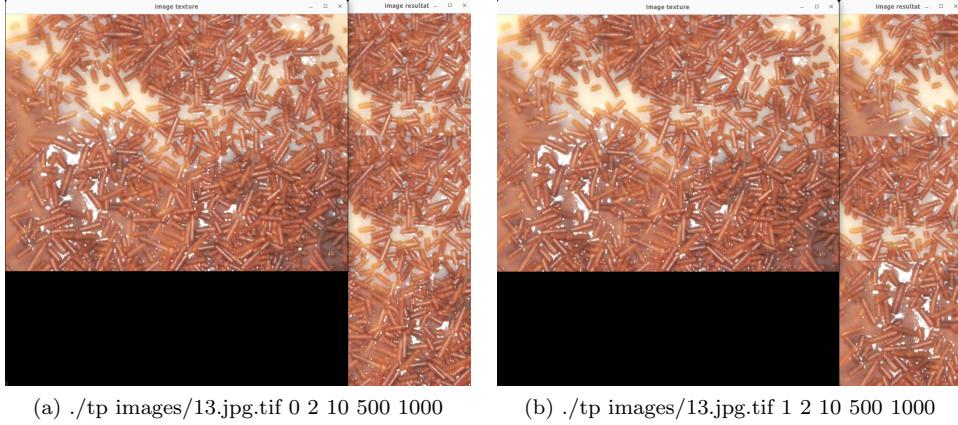


Figure 2: Experiment with algorithms (a) and (b) provided with different images, and different parameters.

## 2.2 Implement a variant of the algorithm (a). A random block permutation generator balances the number of occurrences of each block.

Done in **permuteur.cpp**. There are 3 algorithms that generate permute: dictionary order, heap recursive, and heap iterative. See the code for more.

Change the **testpermute.cpp** file to test more cases. See the code and README.md for more.

## 2.3 What do $e$ and $E$ represent?

In the equation:

$$E_{ij} = e_{ij} + \min(E_{i-1,j-1}, E_{i-1,j}, E_{i-1,j+1}) \quad (1)$$

Variable  $e_{ij}$  is the distance at point  $(i, j)$ . Calculate by  $e_{i,j} = |B_{i,j}^1 - B_{i,j}^2|$ .

Variable  $E$  is the total cost for the cut line.

## 2.4 What is the relationship between the cost of the optimal cut and the $E_{ij}$ ?

Variable  $E_{ij}$  is the minimum cost for the cut line end at point  $(i, j)$ , starting from  $i = 0$ . So the cost of the optimal cut is  $\min_{j \in [0, w]} (E_{hj})$ . Where  $h$  is the height of the intersection area,  $w$  is the width of the intersection area. The optimal cut will start from line 0 and end at the final line.

## 2.5 Clarify the concept of optimal cut on the following example

$e =$	$E =$	$E =$	$C =$																																																																	
<table border="1"> <tbody> <tr><td>1</td><td>3</td><td>2</td><td>1</td></tr> <tr><td>2</td><td>1</td><td>2</td><td>3</td></tr> <tr><td>1</td><td>3</td><td>4</td><td>2</td></tr> <tr><td>2</td><td>4</td><td>3</td><td>1</td></tr> <tr><td>4</td><td>3</td><td>1</td><td>2</td></tr> </tbody> </table>	1	3	2	1	2	1	2	3	1	3	4	2	2	4	3	1	4	3	1	2	<table border="1"> <tbody> <tr><td>1</td><td>3</td><td>2</td><td>1</td></tr> <tr><td>3</td><td>2</td><td>3</td><td>4</td></tr> <tr><td>3</td><td>5</td><td>6</td><td>5</td></tr> <tr><td>5</td><td>7</td><td>8</td><td>6</td></tr> <tr><td>9</td><td>8</td><td>7</td><td>8</td></tr> </tbody> </table>	1	3	2	1	3	2	3	4	3	5	6	5	5	7	8	6	9	8	7	8	<table border="1"> <tbody> <tr><td>1</td><td>3</td><td>2</td><td><b>1</b></td></tr> <tr><td>3</td><td>2</td><td><b>3</b></td><td>4</td></tr> <tr><td>3</td><td>5</td><td>6</td><td><b>5</b></td></tr> <tr><td>5</td><td>7</td><td>8</td><td><b>6</b></td></tr> <tr><td>9</td><td>8</td><td><b>7</b></td><td>8</td></tr> </tbody> </table>	1	3	2	<b>1</b>	3	2	<b>3</b>	4	3	5	6	<b>5</b>	5	7	8	<b>6</b>	9	8	<b>7</b>	8	<table border="1"> <tbody> <tr><td>3</td></tr> <tr><td>2</td></tr> <tr><td>3</td></tr> <tr><td>3</td></tr> <tr><td>2</td></tr> </tbody> </table>	3	2	3	3	2
1	3	2	1																																																																	
2	1	2	3																																																																	
1	3	4	2																																																																	
2	4	3	1																																																																	
4	3	1	2																																																																	
1	3	2	1																																																																	
3	2	3	4																																																																	
3	5	6	5																																																																	
5	7	8	6																																																																	
9	8	7	8																																																																	
1	3	2	<b>1</b>																																																																	
3	2	<b>3</b>	4																																																																	
3	5	6	<b>5</b>																																																																	
5	7	8	<b>6</b>																																																																	
9	8	<b>7</b>	8																																																																	
3																																																																				
2																																																																				
3																																																																				
3																																																																				
2																																																																				

And the minimum cost total is 7.

## 2.6 Estimate the complexity of a naive recursive solution for computing the optimal cut.

Call  $new_h, new_w; h, w; h', w'$  respectively is height, width of new image; height, width of original image; height, width of intersection.  $w'$  is input,  $nb\_block = racineNOMBREBLOCS^2$ ,  $h' = block_h = h/nb\_block$ ,  $block_w = w/nb\_block$ .  $verti = new_h/(block_h - 2 \times w')$ ,  $hori = new_w/(block_w - 2 \times w')$ .

At each point, it calls to 3 other points while the height equals 0. For each  $i$  in the range of  $w'$ , we need to calculate  $E_{ih'}$ . For each  $j$  in the range of  $h'$ , we call to 3 next points. So it costs  $O(3^{h'})$  to calculate  $E_{ih'}$  for each point  $i$  in the range of  $w'$ . Totally it costs  $O(w' \times 3^{h'})$  in time complexity.

Here, space complexity has two ways of looking at it. According to programming languages, we usually don't use any variables when calling recursively (shortest way) so  $O(1)$ . However, when viewed at the assembly level, for each recursive call a next pointer, parameters, and return value are added to the stack. There is a maximum  $h'$  time stack recursive all the time so it is  $O(h')$  pointers in stack very easy stack overflow. So the space complexity is  $O(h')$ .

But we also need to calculate the cutline. There are two ways to calculate the cut line. One is to create a new array to save the cutline for each recursive call. This way increases the space complexity to  $O(h'^2)$ , and time is the same. The second way is, after getting min we call that point again to set the cutline. This way increases the time complexity to  $O(w \times 4^{h'})$ , and space is the same.

Combine with question 2.9 to get the total complexity.

## 2.7 Implement a recursive solution without redundant computation.

Done in file `raccordeur_recursif.cpp`.

## 2.8 Implement an iterative solution.

Done in file `raccordeur_iteratif.cpp`.

## 2.9 Additional question: what is the overall complexity of the algorithm (c) taking into account all the parameters?

Note: (time, space)

Compute between 2 blocks (recursive, naif, travel again):  $O(w' \times 4^{h'})$ ,  $O(h')$

Compute between 2 blocks (recursive, naif, new arrays):  $O(w' \times 3^{h'})$ ,  $O(h'^2)$

Compute between 2 blocks (recursive, memory):  $O(w' \times h')$ ,  $O(w' \times h')$

Compute between 2 blocks (iterative):  $O(w' \times h')$ ,  $O(w' \times h')$

For each algorithm, call  $O_1$ ,  $O_2$  respectively is the time, and space complexity for 2 blocks.

Compute all without cache:  $O(\text{verti} \times \text{hori} \times \text{nb\_block} \times O_1)$ ,  $O_2$

Compute all with cache cost:  $O(O_1 \times \min(\text{verti} \times \text{hori} \times \text{nb\_block}, \text{nb\_block}^2))$ ,  $O(O_2 + \text{nb\_block}^2)$

Compute all with cache cost, cut:  $O(O_1 \times \min(\text{verti} \times \text{hori} \times \text{nb\_block}, \text{nb\_block}^2))$ ,  $O(O_2 + \text{nb\_block}^2 \times h')$

## 3 Compare

### 3.1 Experiment with algorithm (c)

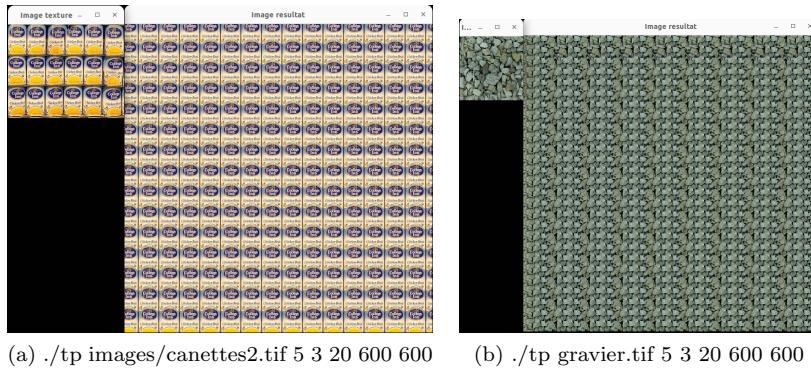


Figure 3: Experiment with algorithms (c) provided with different images.

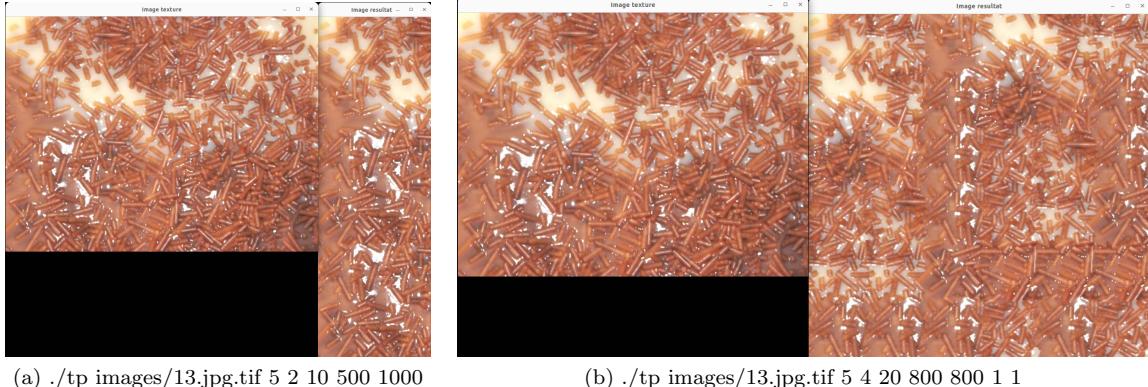


Figure 4: Experiment with algorithms (c) provided with different parameters

### 3.2 Time comparison of algorithm implementation versions (c)

	Test1	Test2	Test3	Test4	Test5	Test6
Recursivenaif,noroute	801ms±13.7ms	12.4s±102ms	>1min	>1min	>1min	>1min
Recursivenaif,route	756ms±4.37ms	3.37s±37.9ms	12.5s±73.6ms	>1min	>1min	>1min
Recursivememo	653ms±659μs	284ms±18.5ms	285ms±232μs	188ms±1.65ms	622ms±25.1ms	7.78s±50.1ms
Iterative	546ms±2.82ms	240ms±12.2ms	238ms±18.7ms	203ms±5.87ms	493ms±23.1ms	7.16s±377ms

Table 1: Time comparison of implementation versions of the algorithm (c). For more information about testcase parameters see [Compare algo.ipynb](#)

	Normal	Cache cost	Cache cut	Cache cost, cut
Algo 3, test 7	9.61 s ± 21.7 ms	6.57 s ± 8.15 ms	9.91 s ± 241 ms	7.48 s ± 155 ms
Algo 6, test 7	3.87 s ± 32.8 ms	2.65 s ± 32.2 ms	3.79 s ± 66.6 ms	3.03 s ± 104 ms
Algo 4, test 8	9.6 s ± 2.01 ms	611 ms ± 98.5 μs	10.3 s ± 85.6 ms	218 ms ± 1.84 ms
Algo 5, test 8	7.43 s ± 84.6 ms	542 ms ± 65.2 ms	7.83 s ± 40.1 ms	204 ms ± 8.04 ms

Table 2: Time comparison of with and without cache implementation versions of the algorithm (c). For more information about testcase parameters see [Compare cache.ipynb](#)

In the cache cut version, it's not much different than without cache, because we don't cache costs so we need to calculate the cost and coupe again for each pair of blocks. Only faster when the index of the best block is found, but slower at assigning a useless cut between pair of blocks (not best).

## 4 Advanced Optimize

### 4.1 Fix some bugs

Line 28<sup>th</sup> file **image\_4b.cpp** changed to

```
if(tif) // fix bug
```

Line 579<sup>th</sup> and 593<sup>th</sup> in new **textureur.cpp** (Line 433<sup>th</sup> and 439<sup>th</sup> in old **textureur.cpp**) changed from

```
Coupe_GD(racordeur, &table_blocs[ind], &table_blocs[res_bindex-&get(c-1, l)], coupe);
...
Coupe_HB(racordeur, &table_blocs[ind], &table_blocs[res_bindex-&get(c, l-1)], coupe);
```

to

```

Coupe_GD(racordeur, &table_blocs[res_bindex-&gt;get(c-1, l)], &table_blocs[ind], coupe);
...
Coupe_HB(racordeur, &table_blocs[res_bindex-&gt;get(c, l-1)], &table_blocs[ind], coupe);

```

Because the previous block is the front parameter.

## 4.2 Stored cut to a new array

For a recursive algorithm, there are two ways to save the path. One is to travel again, the other is to create a new array to save the way each time you travel. The first way is slower but saves memory. The second way is faster but takes more memory

## 4.3 Cache cost

Modified "texture" to be able to save the distance of a pair of blocks. Because every time we want to add a block, we have to go through all the blocks to calculate the cost with the previous block. There will be 2 computations, one **vertically** and one **horizontally** (ignored if there is no neighbor block). To minimize these computations we will create two 2-dimensional arrays of size  $[nb\_bloc \times nb\_bloc]$ . One to save cost vertically and one horizontally. All will be initialized to  $-1$ , then if  $A[i, j] == -1$ , we need to calculate and save it in  $A[i, j]$ . Next time we need to use it, just get  $A[i, j]$ .

## 4.4 Cache cut

Same as cache cost 4.3. We create two 3d-arrays (one **vertically** and one **horizontally**) to save cut between two blocks.

# 5 Conclusion

It can be seen that the recursive naif algorithm runs for a very long time, and can only be executed with inputs with many blocks (so that the exponent is small). The iterative algorithm runs fast, but the more blocks it has, the slower it will be because it has to compare pairs of blocks many times (no cache). The permutation algorithm will not run if there are too many blocks (not enough memory). We can expand otherwise without remembering the permutations, but generating a new one each time next() is called. And with cache the algorithm runs significantly faster.

## References

- [1] R. P. Brent and H. T. Kung, “A regular layout for parallel adders,” *IEEE transactions on Computers*, vol. 31, no. 03, pp. 260–264, 1982.
- [2] P. M. Kogge and H. S. Stone, “A parallel algorithm for the efficient solution of a general class of recurrence equations,” *IEEE transactions on computers*, vol. 100, no. 8, pp. 786–793, 1973.
- [3] R. Rivest, “Rfc1321: The md5 message-digest algorithm,” 1992.
- [4] T. Nguyen-Quang, T.-D. H. Nguyen, T.-L. Nguyen-Ho, A.-K. Duong, N. Hoang-Xuan, V.-T. Nguyen-Truong, H.-D. Nguyen, and M.-T. Tran, “Hcmus at mediaeval 2020: Image-text fusion for automatic news-images re-matching,” 2020.
- [5] A.-K. Duong, H.-L. Nguyen, and T.-T. Truong, “Large margin cotangent loss for deep similarity learning,” in *2022 International Conference on Advanced Computing and Analytics (ACOMPA)*, pp. 40–47, IEEE, 2022.
- [6] A.-K. Duong, H.-L. Nguyen, and T.-T. Truong, “Enhanced face authentication with separate loss functions,” *arXiv preprint arXiv:2302.11427*, 2023.
- [7] M. S. A. Toofanee, S. Dowlut, M. Hamroun, K. Tamine, V. Petit, A. K. Duong, and D. Sauveron, “Dfu-siam a novel diabetic foot ulcer classification with deep learning,” *IEEE Access*, 2023.
- [8] M. S. A. Toofanee, S. Dowlut, M. Hamroun, K. Tamine, A. K. Duong, V. Petit, and D. Sauveron, “Dfu-helper: An innovative framework for longitudinal diabetic foot ulcer diseases evaluation using deep learning,” *Applied Sciences*, vol. 13, no. 18, p. 10310, 2023.
- [9] M. S. A. Toofanee, M. Hamroun, S. Dowlut, K. Tamine, V. Petit, A. K. Duong, and D. Sauveron, “Federated learning: Centralized and p2p for a siamese deep learning model for diabetes foot ulcer classification,” *Applied Sciences*, vol. 13, no. 23, p. 12776, 2023.
- [10] A.-K. Duong and V.-H. Huynh, “Bilinear cnns model and test time augmentation for screening viral and covid-19 pneumonia,” *SSICT2020*, p. 112, 2020.