# Pontificia Universidad Javeriana Cali First Laboratory Report Parallel Programming Course

Submitted by: Katherine Camacho Calderón Juan José Hoyos Urcué

Submitted to the teacher: Roger Gomez Nieto

#### INTRODUCTION

This workshop is divided into two parts:

In the first part, a series of analyzes and modifications were carried out on a sequential algorithm which purpose is to obtain the quality of a video by the measure of a score. Initially the modifications were made on the sequential algorithm, then embarrassing parallel programming is used on the improved algorithm, and finally comparisons and conclusions are made on said algorithm in order to identify the improvements in its performance

The second part will analyze the times of three algorithms that were executed sequentially and then concurrently, which allowed identifying the differences in performance between these two forms of execution.

#### First part

In the first part of this laboratory, a series of analyzes and modifications were carried out on the FRIQUEE algorithm, which is written in such a way that its execution is given sequentially, such analyzes and modifications were given as follows:

- First, the algorithm time was analyzed, without any modification, and what was obtained was:

Mean of scores, video score(mean±std) and the total time were:

```
Command Window

La media de los scores es:
62.5041

El score final del video está entre:
56.5911 68.4171

Elapsed time is 8051.251270 seconds.
```

The time for each function was:

Function Name	Calls	Total Time	Self Time*	Total Time Plot (dark band = self time)
Complex DIIVINE feature	1204	7457.626 s	44.666 s	
Complex DIIVINE feature>RelativePhase HV	1204	107.931 s	31.476 s	Į.
Complex DIIVINE feature>Scale CW SSIM	1204	539.177 s	2.653 s	•
DIIVINE feature>WrapCauchyEstimate	28896	2155.029 s	2155.029 s	
DoGFeat	1204	105.180 s	0.150 s	L
MGGD ParaEstimate	9632	3322.639 s	232.059 s	
MGGD ParaEstimate>f	39806	3083.494 s	3083.494 s	
addpath	1507	61.058 s	0.943 s	I .
angle	14448	35.173 s	35.173 s	
applycform	1505	21.702 s	1.016 s	
ty(stringValue)  ~ischar(stringValue)	4515	0.004 s	0.004 s	
applycform>check_cform	1505	0.086 s	0.081 s	
applycform>check input image dimensions	1505	0.013 s	0.013 s	
axescheck	602	0.013 s	0.013 s	
binomialFilter	1204	0.038 s	0.012 s	
blanks	1808	0.405 s	0.405 s	
buildLpyr	4816	13.221 s	2.393 s	
buildSCFpyr	1204	223.543 s	21.157 s	T.

- Second, a series of modifications were made to the sequential algorithm, these modifications basically consisted of simplifying the number of mathematical operations, since there were repeated operations and taking out operations of some for loop, which did not depend on any value that was within it. Some of these modifications were:
  - In this image, the selected lines show operations that were previously repeated up to three times within this same function, so the simplification of the number of operations significantly reduced the execution time.

```
function [fo, f1] = f(b, r,a_,b_)
k = (r.^b);
t = (2/b);
r_b = sum(k);
ss = a_;

N = b_;
j = b * r_b;
pp = psi(t);
rr = log(0.5*j/N);
r_b_log = sum(ss.* k);
r_b_log2 = sum(ss.*ss.* k);
f0 = 1 + 2*pp/b - 2*r_b_log/r_b + (t) * rr;

q = (b^2);
A = -2*pp/q - 4*psi(1,t)/(b^3);
B = -2*(r_b_log2 * r_b - r_b_log^2/(r_b^2);
C = -2*rr/q + 2/q + 2*r_b_log/(j);

f1 = A + B + C;
```

• In this image, the selected lines show operations that were within the while loop and that were repeated unnecessarily many times, since their result was constant.Removing these operations from the while loop significantly improved the performance of the algorithm.

```
sinm = sin(AngleMatrix);
cosm = cos(AngleMatrix);
while (k < 1000)
    mul = ul;
    mu2 = u2;
    w = 1./(1 - (mul*cosm) - (mu2*sinm));
    num1 = w.*cosm;
    num2 = w.*sinm;
    u1 = sum(num1(:)) / sum(w(:));
    u2 = sum(num2(:))/sum(w(:));
    k = k + 1;
    if (abs(u1-mu1) < e)&&(abs(u2-mu2) < e)
        break;
end
end</pre>
```

• In this image the two lines selected were taken out from a for loop, since they were operations that did not depend on any value within it, their result within the for loop was always the same.

• In this image, the selected line shows an operation that was previously performed twice, which is expensive because it is within a function that can be called many times.

```
function indices = pyrBandIndices(pind,band)
tamanlo = size(pind,l);
if ((band > tamanio) || (band < 1))
    error(sprintf('BAND_NUM must be between 1 and number of pyramid bands (%d).', ...
    tamanio));
end

if (size(pind,2) ~= 2)
    error('INDICES must be an Nx2 matrix indicating the size of the pyramid subbands');
end

ind = 1;
for l=1:band-1
    ind = ind + prod(pind(l,:));
end

indices = ind:ind+prod(pind(band,:))-1;</pre>
```

• In this image, the selected lines show two operations that were repeated twice within a double for loop, so saving them in a variable to avoid doing it twice, improves the performance of the algorithm.

```
for scale=1:Nsc-1
      for orien=1:Nor
          nband = (scale-1)*Nor+orien+1; % except the ll
                  if(prod(band-nband-1) ~=0)
                          continue;
          aux_c = pyrBand(pyro, pind, nband);
          aux = abs(aux_c);
[Nsy,Nsx] = size(aux);
          prnt = parent & (nband <= Nband-(Nsc-1)*Nor); % has the subband a parent?
          define that only the finest scale has a parent
BL = zeros(Nsy,Nsx,1 + prnt);
          BL(:,:,1) = aux;
                auxp = pyrBand(pyro, pind, nband+Nor);
                    if nband>Nor+1, % resample 2x2 the parent if not in the high-pass oriented subbands.
auxp = real(imenlarge2(auxp)); % this was uncommented
               auxp = abs(imresize(auxp,2)); %
               % end
% fprintf('parent band and size is %d %d %d \n',nband+Nor,Nsy,Nsx)
               BL(:,:,2) = auxp(1:Nsy,1:Nsx);
           y = BL;
[nv,nh, nb] = size(y);
```

- Third, a report was made of the current time to perceive the improvements in the performance of the algorithm and this was obtained:

The total time of the algorithm was:

```
La media de los scores es:
62.5041

El score final del video está entre:
56.5911 68.4171

Elapsed time is 4403.962525 seconds.
```

The time for each of the functions was:

Function Name	Calls	Total Time	Self Time*	Total Time Plot (dark band = self time)
Complex DIIVINE feature	1204	3808.329 s	46.194 s	
Complex DIIVINE feature>RelativePhase HV	1204	110.120 s	31.815 s	I
Complex DIIVINE feature>Scale CW SSIM	1204	544.192 s	3.104 s	
DIIVINE feature>WrapCauchyEstimate	28896	569.175 s	569.175 s	
DoGFeat	1204	105.152 s	0.152 s	1
MGGD ParaEstimate	9632	1239.951 s	261.697 s	
MGGD ParaEstimate>f	39806	971.151 s	971.151 s	
addpath	1507	61.255 s	0.940 s	1
angle	14448	35.348 s	35.348 s	Ī
applycform	1505	21.637 s	1.026 s	
ty(stringValue)  ~ischar(stringValue)	4515	0.004 s	0.004 s	
applycform>check cform	1505	0.087 s	0.083 s	
applycform>check input image dimensions	1505	0.014 s	0.014 s	
axescheck	602	0.013 s	0.013 s	
binomialFilter	1204	0.038 s	0.012 s	
<u>blanks</u>	1808	0.404 s	0.404 s	
<u>buildLpyr</u>	4816	13.263 s	2.380 s	
buildSCFpyr	1204	224.182 s	21.287 s	

## Relevant optimizations:

### **Function: Complex DIVINE feature**

7,457.626 ------ 100% 3,808.329 -----x

x = 51.06%. It's optimization was 48.94%

### **Function: MGGD ParaEstimate**

3,322.639 ------100% 1,239.951 -----x

x = 37.32%. It's optimization was 62.68%

#### **Function: MGGD ParaEstimate>f**

3,083.494 ------100% 971.171 -----x

x = 31.5% .It's optimization was 68.5%

- Fourth, data decomposition was performed, so each core worked with a unique data set, a frame for each core, which significantly improved performance. The following line was used to perform data decomposition:

```
(base) juan2203@Juan3ose:-/Downloads/FRIQUEE_Release/data/videoS ffmpeg -i G802.mp4 Pictures%d.jpg

ffmpeg version 3.4.6-8ubuntu8.18.64.1 Copyright (c) 2008-2019 the Ffmpeg developers

built with gcc 7 (Ubuntu 7.3.0-16ubuntu3)

configuration: --prefix=/usr --extra-version=8ubuntu8.18.64.1 --toolchain=hardened --libdir=/usr/lib/x86_64-linux-gnu --incdir=/usr/include/

x86_64-linux-gnu --enable-lidsable-stripping --enable-avresample --enable-avisynth --enable-gnutls --enable-ladspa --enable-libsss --enable-libbluray --enable-libssb --enable-libcdio --enable-libflite --enable-libpenting --enable-libpred --enable-libpred --enable-libpred --enable-libpred --enable-libpred --enable-libpred --enable-libpred --enable-libspee --enable-libpred --enable-libspee --enable-libspee --enable-libspee --enable-libspee --enable-libspee --enable-libved -
```

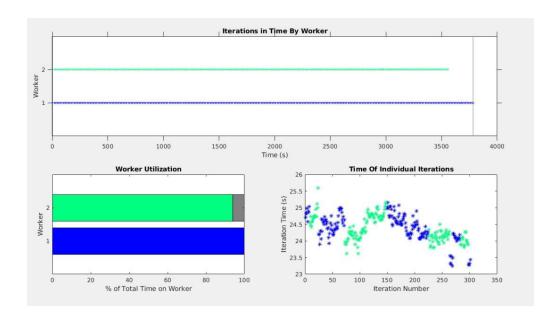
- Fifth, after having evaluated some for loops, we select the for loop that receive all the frames of the video, since at this point it is possible to divide the information in such a way that each core can return a value without relying on other core to finish performing a task. The selected for loop was the following:

```
parfor i= 1:n_frames
     % Read an image
    % Par.tic;
    name = 'Pictures';
    number = num2str(i);
    ext = '.jpg';
    frame name = strcat(name, number, ext);
    img = imread(strcat('data/video/',frame_name));
    % Extract FRIQUEE-ALL features of this image
    testFriqueeFeats = extractFRIQUEEFeatures(img);
     Load a learned model
    % load('data/friqueeLearnedModel.mat');
    % Scale the features of the test image accordingly.
    % The minimum and the range are computed on features of all the images of
    % LIVE Challenge Database
    testFriqueeALL = testFeatNormalize(testFriqueeFeats.friqueeALL, friqueeLearnedModel.trainDataMinVals, fri
    qualityScore = sympredict (0, double(testFriqueeALL), friqueeLearnedModel.trainModel, '-b 1 -q');
    scores(i) = qualityScore;
     % p(i) = Par.toc;
```

- Sixth, the ParTicToc tool was implemented to evaluate the operation of the parfor and the following was obtained:

The total time was:

```
Elapsed time is 3783.603234 seconds.
```



- Finally, the average of each frame was obtained and the range in which the video quality was found was calculated, the way in which this calculation was performed was:

```
media = mean(scores);
desviacion = std(scores);
ans_menos = media-desviacion;
ans_mas = media+desviacion;
res = [ans_menos,ans_mas];
disp(res);
toc;
profile viewer;
```

Where scores is an array that stores the score obtained in each one of the video frames.

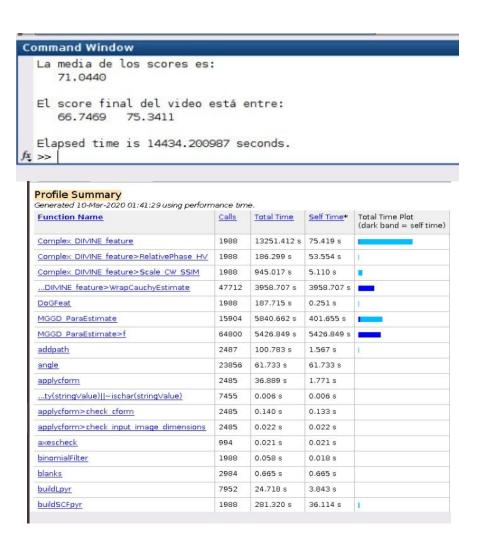
The range in which the score is found is:

```
La media de los scores es:
62.5041

El score final del video está entre:
56.5911 68.4171
```

It was subsequently tested with another video and the following results were obtained:

- The algorithm time was analyzed, without any modification, and what was obtained was:



- The algorithm time was analyzed after having made a series of modifications to the sequential algorithm, and the results were:

```
Command Window

>> example
La media de los scores es:
71.0440

El score final del video está entre:
66.7469 75.3411

Elapsed time is 8380.654794 seconds.

fx >>
```

Function Name	Calls	<u>Total Time</u>	Self Time*	Total Time Plot (dark band = self time)
Complex DIIVINE feature	1988	7239.348 s	70.883 s	
Complex DIIVINE feature>RelativePhase HV	1988	206.274 s	59.781 s	I .
Complex DIIVINE feature>Scale CW SSIM	1988	1057.582 s	4.404 s	•
DIIVINE feature>WrapCauchyEstimate	47712	1099.457 s	1099.457 s	
DoGFeat	1988	204.570 s	0.282 s	1
MGGD ParaEstimate	15904	2511.643 s	525.022 s	-
MGGD ParaEstimate>f	64800	1973.996 s	1973.996 s	
addpath	2487	109.308 s	1.600 s	I.
angle	23856	70.929 s	70.929 s	Ĭ.
applycform	2485	42.258 s	1.927 s	
ty(stringValue)  ~ischar(stringValue)	7455	0.007 s	0.007 s	
applycform>check cform	2485	0.153 s	0.145 s	
applycform>check input image dimensions	2485	0.025 s	0.025 s	
axescheck	994	0.022 s	0.022 s	
binomialFilter	1988	0.064 s	0.019 s	
blanks	2984	0.735 s	0.735 s	
buildLpyr	7952	30.124 s	4.445 s	
buildSCFpyr	1988	350.974 s	62.873 s	1

## Relevant optimizations:

## **Function: Complex DIVINE feature**

x = 54.6%. It's optimization was 45.4%

### **Function: MGGD ParaEstimate**

5,840.662 ------100% 2,511.643 -----x

#### **Function: MGGD ParaEstimate>f**

```
5,426.849 ------- 100% 1,973.996 -----x
```

x = 36.37% .It's optimization was 63.63%

- The algorithm time was analyzed after using the embarrassing parallel programming technique, and the results obtained were:

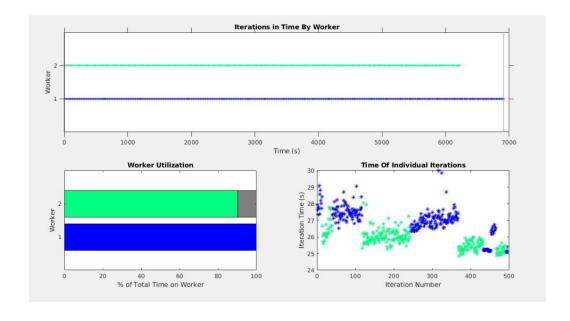
```
Command Window

>> example
La media de los scores es:
71.0440

El score final del video está entre:
66.7469 75.3411

Elapsed time is 6910.927229 seconds.

fx >>>
```



#### Conclusion

Throughout the analysis of the algorithm, modifications and optimizations were made in different ways, the algorithm was first executed without any modification and an execution time of 2.23 hours was achieved, then sequential optimizations were made and an execution time of 1.22 hours was achieved with them. and finally the embarrassing parallel programming technique was implemented, achieving an execution time of 1.05 hours, where it was observed that each worker, took a different frame for a certain amount of time, and thanks to the graph we could detail that a worker finished first than another, but both worked together simultaneously most of the time, this allowed to significantly improve the execution time of the algorithm, so we conclude that using this technique correctly will allow the tasks to be divided and processed between the available cores and therefore maximum use of computer resources can be made.

By performing the above procedure with a new video, whose duration is 15 seconds and which was divided into 497 frames, it was obtained that sequential modifications improved performance much more than the algorithm that uses the embarrassing parallel programming technique, as the percentage of improvement with sequential modifications is 42% which is significantly greater than that achieved by the parallel algorithm 52%. Therefore, we conclude that it is important to analyze the algorithm sequentially first, and make all the optimizations in this way, and then if it is possible to use parallel programming to improve it further, the combination of both optimization techniques allowed to obtain a good result. compared to the one that initially was.

By executing the sequential code over the first video without any modification and the parallel code the **speedup** calculated was:

(8051.25127)/(3783.603234) =**2.12793223** 

### Second part

In the second part of this workshop, three high complexity algorithms were implemented, which were initially run sequentially and then concurrently.

The first algorithm is responsible for finding all possible solutions of a linear equation of two variables until a number n (This equation is 3x + 5y = 100)

The second algorithm is responsible for finding all possible solutions of a linear equation of three variables until a number n (This equation is 3x + 5y + 6z = 100)

The third algorithm is responsible for finding all possible solutions of a linear equation of four variables until a number n (This equation is 3x + 5y + 6z + 7w = 120)

These algorithms have time complexity of:

```
O(n^2)
```

 $O(n^3)$  and

O (n ^ 4)

respectively

## Sequentially

```
File Edit View Search Terminal Help

(base) juan2203@JuanJose:-/Documents$ g++ algorithms.cpp -o secuencial
(base) juan2203@JuanJose:-/Documents$ ./secuencial

Ejecutar los tres algoritmos de manera secuencial toma un tiempo de: 66.15 segundos
(base) juan2203@JuanJose:-/Documents$ [
```

Running the algorithms sequentially takes a runtime of 66.15 seconds

They were executed in a main function one by one:

```
int main(){
    auto start = std::chrono::system_clock::now();
    ecuacion lineal_2var(100000);
    ecuacion_lineal_3var(2000);
    ecuacion_lineal_4var(300);

auto end = std::chrono::system_clock::now();
    chrono::duration=duble> elapsed = end - start;
    printf("Ejecutar los tres algoritmos de manera secuencial toma un tiempo de: %.2lf segundos\n", elapsed.count());
    return 0;
```

The processor looked like this when we ran the algorithms sequentially:

As we can see, core 1 is being 100% used, that is, all this resource is being used to execute the algorithms sequentially, and it is evident that the other core is being wasted.

### **Concurrently Execution**

```
File Edit View Search Terminal Help
(base) juan2203@JuanJose:-/Documents$ g++ algorithms_thread.cpp -o paralelo -lpthread
(base) juan2203@JuanJose:-/Documents$ ./paralelo
Ejecutar los tres algoritmos de manera paralela toma un tiempo de: 26.95 segundos
(base) juan2203@JuanJose:-/Documents$ [
```

Running the algorithms concurrently takes a runtime of 26.95 seconds

They were all executed in a different thread each one:

```
int main(){
    auto start = std::chrono::system_clock::now();
    pthread t hilo1;
    pthread_t hilo2;
    pthread_t hilo3;

int nl = 100000,n2 = 2000,n3 = 300;
    pthread_attr_t attr;
    pthread_attr_init(Sattr);

    pthread_attr_init(Sattr);

    pthread_create(Shilo1,NULL,ecuacion_lineal_2var,(void*)&n1);
    pthread_create(Shilo2,NULL,ecuacion_lineal_3var,(void*)&n2);
    pthread_create(Shilo3,NULL,ecuacion_lineal_4var,(void*)&n3);

    pthread_join(hilo1,NULL);
    pthread_join(hilo2,NULL);
    pthread_join(hilo3,NULL);
    p
```

The processor looked like this when we ran the algorithms concurrently:

As we can see, there are 3 cores that are being 100% used, this is how the great execution of the 3 algorithms is achieved concurrently, reducing sequential time in a drastic way

### How fast does concurrent vs. sequential programming work in this case?

$$x = ((26.95)*100)/66.15 = 40.74\%$$

This means that the concurrent algorithm only takes 40.74% of the time the sequential algorithm takes, therefore, we can say that the concurrent algorithm is 59.26% faster than the sequential algorithm.

#### Conclusion

We can then conclude that the use of concurrent programming in the execution of independent programs using threads is a very good solution for the problem of the large amount of time that these algorithms require when executing them sequentially, in addition, we can take the maximum advantage of computational resources that we have at home, making optimal use of them, thus increasing the performance of our programs.