

# Points Cloud Lidar measurements and forest tomography

Marcel Bariou [marcel.bariou@brasnah.fr](mailto:marcel.bariou@brasnah.fr)  
Brasnah sarl Mas Saragosse F-66170 Millas France

2020 February 6 ( english release)

## Summary

This note develops a complete approach for automatic processing, of LIDAR point clouds data collected in the forest undergrowth. These treatments are performed through a pipelined process. Two founder functions, are presented, the identification (Geo-localization) of each tree in the checked area, the evaluation of the diameter of their section, for each of these, at different heights. The goal is to prepare detailed forest inventory. Targeted performance are, in an area :

1. to identify (geolocate) all trees with a false alarm rate (first and second kind errors included) 5%, in the range diameter from 20 cm to 120cm in steps of 10 cm,
2. to provide the relative positions of the trees with +/- 50 cm of accuracy in an area forest,
3. to give the section sizes (diameters) at breast height with an accuracy of +/- 10cm,
4. to process, the equivalent of a measured area of 1 ha populated by 400 trees, in 30 minutes by providing for each tree the attributes pair : geolocation / diameter.

The associated algorithms are introduced, as well as the processing pipeline implemented. These treatments provide an output result dataset exportable to any geographic information system or geometric database.

This approach is the fruit of tests and learning started in 2018, the measures dating from 2016 were conducted at this period without specific protocol targeting a specific result. This increased the need of reflections for their analysis and led to some methodological recommendations concerning the measurement procedure and the calibration of the measurement chain (Instrument + software for creating \* **.las** files from data provided with the lidar inertial unit and measurement of the lidar pulse flight time) such as counting software and algorithms (i.e. the result of present works).

Results are issued in such a way to facilitate the production of future specifications for the development of dedicated industrial tools (software), based on the utilities and algorithms developed and introduced in this note.

The work target is to lead, on the first hand to a specialized autonomous application on standalone workstation, on the second hand to automated online processing on remote server. Some test have been done about remote processing, but these aspects are not covered in the current note.

“Where is the error, in the calibration or in the measurement process? Perhaps in both ?”  
**The unknown surveyor**  
Quote neither dated nor geolocated

# 1 Foreword

## 1.1 Purpose of the note and credits

The works presented in the article attempt to respond, in part to the expectations of the scientific team monitoring the National Nature Reserve of the Massane forest (R.N.N. la Massane). It is a beech forest of 330 ha, located in the Pyrenees Orientales on the Franco-Spanish border.

An area of 10 ha, fully protected, is the subject of scientific monitoring permanent for over 100 years. The most recent work related to measurement objectives of this note are presented in [\(1\)](#), [\(2\)](#), [\(3\)](#), [\(4\)](#). The following works were made possible by the provision of data from LIDAR measurement, by the Massane National Nature Reserve, carried out:

1. with aerial measurement by the company SINTEGRA (5) in 2016,
2. in the undergrowth by the company EXAMETRICS in 2016.

We will add a complement of 20 G.P.S points. performed by the author and the R.N.N. de la Massane in September 2019, on the measured spot. Although performed with modest and imprecise means, it nevertheless made it possible to assess the mathematical operator of transformation in a correct C.R.S (Coordinate Reference System) as we will see later.

A first series of works, [\(6\)](#) and [\(7\)](#), carried out as part of the company EXAMETRICS allowed to approach the subject, including an important bibliographic research (see [\(6\)](#)), and difficulties identification implement certain algorithms, in order to avoid pitfalls later, especially in the continuation that we gave them here. These works were conducted with the participation of students from the CHIPS Master of the University of Perpignan Via Domitia (U.P.V.D.) and the I.M.E.R.I.R Engineering school.

The software development parts of the the works, presented hereinafter, have been carried out exclusively with the use of free software under licenses GPL, LGPL or similar.

The workstation is a DELL Precision Tower 5810 station (64 GB of RAM) running under Ubuntu 18-04 equipped with a CUDA Quadro processor Pro K5000.

Among the main software libraries used we will name: PDAL, LIBLAS, EIGEN3, the C\_CUDA SDK, PyCuda. Prototypings were done in Python language and then rewritten in C++ or C\_CUDA.

The CloudCompare tool and viewer allows us to visualize LIDAR point clouds and to extract dimensional or distance attributes of objects in the points cloud.

Unless coincidence, as far as we know, the present work do not rely on an improvement of existing or documented techniques, they are the result of our learning of lidar technology and our experience in real-time 3D image synthesis geometry and the implementation of basic statistics and image filtering, all well known elsewhere.

They do not come under the complexity of algorithmic geometry. The current note is structured strictly in accordance with the chronology of the steps of the measurement and data processing, as shown in Figure 1.

We precede this presentation only with a brief reminder on the measurement techniques by Lidar and in particular the Terrestrial Mobile Lidar Scan (T.M.L.S) and its related technologies.

## **1.2 The forest an area of biodiversity and plant vigor**

The forest and the woods, in particular those which are left to the strict whims of nature or more exactly to the will of each tree to develop itself are characterized by an abundance of biodiversity.

There are large trees that have found their living space. They develop in cohabitation with their peers, with a sufficient surface on the ground, where many shrubs and bushes try to settle by desperately seeking the necessary space. In the absence of height growth, the bushes settle and grow, masking the soil. The shrubs, those that survive, stubbornly grow slowly. On the ground, a lot of debris, dry wood from normal growth, or from end-of-life fall, feeds a whole variety of fauna of auxiliaries by offering them habitats.

Whether trees, living or candles providing debris to the ground, these are all sources of diverse habitats for a whole fauna and insects contributing to the richness of biodiversity. At the level of the trees a fungal activity, accidents of life digging cavities other possible places of habitat. There are some elements of the richness of forest life. In terms of trees, species can predominate among a diversity, but the climatic zone favors some of them.

In the present zone of interest, it is essentially a beech forest. In any event, the diversity of objects of scientific interest is very wide. In this study we are only interested in trees, we still have to find them and get some attributes from.

## 1.3 Forest measurements, macro and micro-observation

Satellite, radar and aerial lidar measurements provide information on the canopy, or even more, including the ground, with hyperspectral, multispectral cameras, their average resolution (at best, meter or decimeter but in constant progress) classify them in what we can call macro-observation. On the other hand, a scientific or technical activity in the field, which endeavors to follow the development of this biotope by observing individual objects in the forest is a matter of detailed micro-observation.

In addition to direct observation, the collection of clues and / or samples for laboratory study, physical measures are gradually being implemented. The physical attributes of trees are taken by hand. The development of control plots, with a small surface area, receiving local species, allow by statistical induction to reasonably characterize larger forest areas. It is a common practice in forestry.

However, progressively data acquisition is automated, meteorological measurements, with all their palettes, have been the first arrivals, the strain gauges will make it possible to follow the evolution of the diameters of the growing trunks of trees, to name only the most common. The Internet of Things (I.O.T) will pave the way for a host of measures (physical, biochemical, etc.) that will have to be linked to a specific individual tree.

So, whether it is the forester or the scientist who travels through forest areas, their first questions are frequently:

- where is the tree?
- what is its state of growth?

These are the first two questions that the current works attempt to answer. But they will be the prelude to exploiting the possibilities to come from the proliferation of measures that the Internet of Things will bring.

## 1.4 Automate and generalize dendrometric measurement

In this note we introduce the methods implemented and the developments carried out to extract information from data collections carried out in the undergrowth with a terrestrial mobile lidar. Our objective ? Obtain reliable information (dimensional tree attributes data and their geolocation).

As a first approach, the targeted performances are as follows:

- Data collection (with full compliance with a strict measurement protocol see paragraph 4) in a forest area accessible in the undergrowth, by a walker, with a shrub and brushwood density still allowing walking lines spaced 10 m apart. The objective is to achieve the generation of \* .las file, the whole, in compliance with the protocol, for an average time of 2 hours per hectare (but to be appreciated more accurately due to methodological constraints).

- Analysis and processing of data, identification of trees (by their geolocation), estimation of their diameter at heights of your choice:
  1. generation of a digital terrain model (D.T.M.) of the area, placing tree bases at zero relief height,
  2. tree presence identification with a false alarm rate of less than 5% (first and second kind errors included),
  3. individual geolocation of each tree in the area with an absolute accuracy of +/- 10 m and a relative accuracy of the order of a meter or less.
  4. estimation of the diameter, with an accuracy of +/- 0.1 m, for each tree, a tree with a diameter less than 20 cm is detectable but without guarantee on the value of its diameter.
- exhibition and export of results for G.I.S or database populating.

As we will see (cf. §10) , the duration of the treatment process is strongly linked to the execution of the detection of each diameter by filtering, duration which increases in

$O(n^2)$  (N being representative of the ratio of the surface to be treated to the size of the detection filter), in the absence of precautionary treatment acceleration.

As standard we carry out 10 filtering operations with a diameter of 0.1 m to 1.20 meters, the objective is to process 1 ha per hour for an average population of 400 trees (1 tree every 5 m), all included, data collection and processing, approximately 30 seconds per tree, or approximately 3 hours per ha for the overall operations (Data collection + analysis).

Besides, beyond measurement, it is a question of driving these in an attractive economic cost conditions by allowing the individual traceability of each tree, to plan to consider introducing it commonly in current forest management.

## 1.5 Global data collection and non-intrusive measurements

The forest is a living body, we need to go into it in detail to get to know it better, the use of lidar allows this non-intrusive discovery in the undergrowth.

The associated measures, in first step, give access to certain fundamentals of each tree (dimensional attributes, geolocation). These measurements could be enriched on the one hand and on the other hand they can be combined with other types of measurements (provided by other sensors) or observations in order to increase our knowledge of the forest world.

Lidar measurements, especially in the undergrowth, offer the great advantage of being able to give an identity to each tree (its position is unique on earth) and to this identity we can associate all the attributes we want with it.

The work presented here constitutes a modest contribution to the forestry measure. Particularly dedicated, in the immediate future, to the automation of dendrometric measurement operations, they have five objectives:

1. Identify and develop LIDAR measurement protocols in the undergrowth, in order to achieve high productivity of reliable measurements.
2. Extract, from the cloud of these measured points, basic attribute data (diameter at breast height, geolocation), for each tree in the measured area, within the performance limits indicated.
3. Extend the automation of the point cloud analysis and analysis process as far as possible, in order to be able to offer a processing pipeline that is simple to implement and accessible to a wide variety of users.
4. Lay the groundwork for an online service offering.
5. Pave the way for environmental monitoring, as in logging, for an individual monitoring capacity for any tree, once identified by its geolocation, over its entire life cycle, by offering the possibility of gradually expanding the range of attributes that can be associated with it.

This work is part of technological research and development by implementing algorithms already proven elsewhere but by specifically adapting them to the processing of data from a LIDAR points cloud and by adding original elements.

## 1.6 Project partners

As indicated in paragraph 1.1, the project started on the initiative of the National Natural Reserve of the Massane forest in the Pyrénées Orientales (R.N.N. la Massane) and the company EXAMETRICS (St Estève, Pyrénées Orientales).

They made the decision on an exploratory study so that the R.N.N. La Massane could consider automating some of the surveys of study objects (trees, ground debris, elements of wildlife habitat ...) and in particular, firstly, the geolocation of each tree. To this end, the company EXAMETRICS carried out, in the undergrowth, during the second half of 2016 a LIDAR survey in the protected area (managed by R.N.N. la Massane) of 10 ha. In addition and in parallel, an overflight with LIDAR measurement was also carried out over the same area by the SINTEGRA Company, at the request of the R.N.N. la Massane. Thus at the end of 2016, the R.N.N. la Massane had two LIDAR datasets on its 10 ha study area:

1. a statement in the undergrowth,
2. an aerial survey.

## 1.7 Publication organization

The note is structured strictly in accordance with the chronology of the step of the treatment process, as presented in the figure below. We only precede it with a brief reminder of the measurement techniques by Lidar and in particular the terrestrial mobile lidar scan and the technologies associated with it.

1 - data LIDAR acquisition on spot, landmarks localization (batch measure 15 to 20 mn). [Lidar Geoslam +GPS]	7 – PNG file denoising, closing of section tree convex envelop non-automated manual process, algorithm to develop and deepen . [GIMP]
2 - *.las file building and checking, size : 1 to 3 Go according to measurement time. (Geoslam + CloudCompare software)	8 -morphological erosion for increase density of points, on the *.png image. [GIMP G'MIC]
3 – D.M.T extraction (50 cm resolution). It will be also the ground footprint for GIS. [CreateMNT KNN Optimal CM]	9 – Activation of decreasing filter banks in size (120 cm to 20 cm), detection of accumulation pivots, erasing of detected sections.[FILT_CIRCLE_STANDARD]
4 – Setup to relief level zero (orthoprojection by DMT substraction from initial *.las file ). [Substract MNT_CM]	10 - Elimination of duplicates (accumulation pivots), with arbitration of selected diameter, Z tree extraction from DEM database or file . [FILT_CIRCLE_STANDARD]
5 – extraction of parallelepiped box of points, parallel to XOY, thick 10 cm, height 1.30 m from zero level. [Prj_Extract_Tile]	11 - Calculation of the transformation operator, from the landmarks selected on each batch. [WG84_2_RFG93]
6 - From parallelepiped box of points creation of image file (*.png) orthoprojection on XOY. Resolution 1cm ~ 1 px. [Create_PNG-X-Y-Zmoy]	12 - C.R.S Transformation (WGS84 => RFG93), of tree sections detected , by the previous operator for each zone and preparation of attribute data for transfer to G.I.S [Forest_Data_For_SIG]

Figure 1 : the global process

The bold characters denote, for the record, the utilities developed or the tools that may be used:

1. measurement tool: Lidar Geoslam, G.P.S Etrex Garmin,
2. software: Geoslam, CloudCompare, Gimp
3. utilities developed: Create\_MNT\_OTIMAL\_KNN\_CM, Substract\_MNT\_CM, Prj\_extract\_Tile, Create\_PNG\_X-Y-Z\_MOY, Filt\_Circle\_Standard, WG84\_2\_RFG93, COMPUTE\_Rot\_Trans, Forest\_Data\_For\_SIG

## 2 Quick reminder of LIDAR measure and processing

Commonly a lidar relies on an emission of monochromatic radiation emitted by a laser, from which it receives the return echo (s). The Euler angles of the transmitted pulse are well known, as is the time difference between transmission and reception, therefore the measuring instrument is able to position any echo in an identified 3D coordinate reference system.

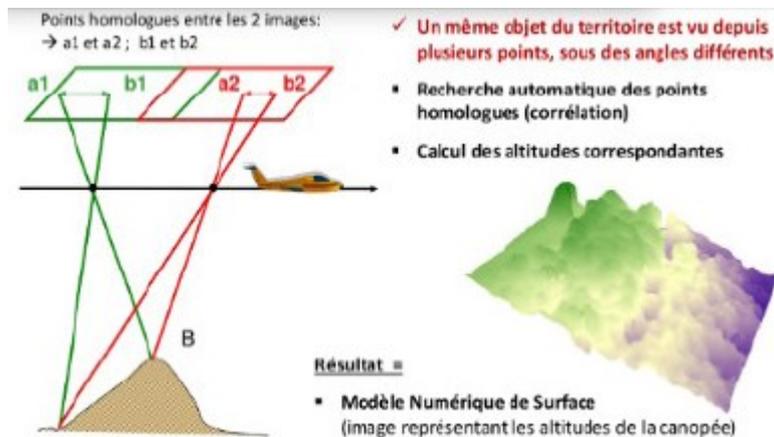


Figure 2 : Photogrammetry credit (IGN)

By providing this very precise information, it replaced the traditional photogrammetry, for point surveys, the photogrammetric evaluation is replaced by a real measurement on the spot.

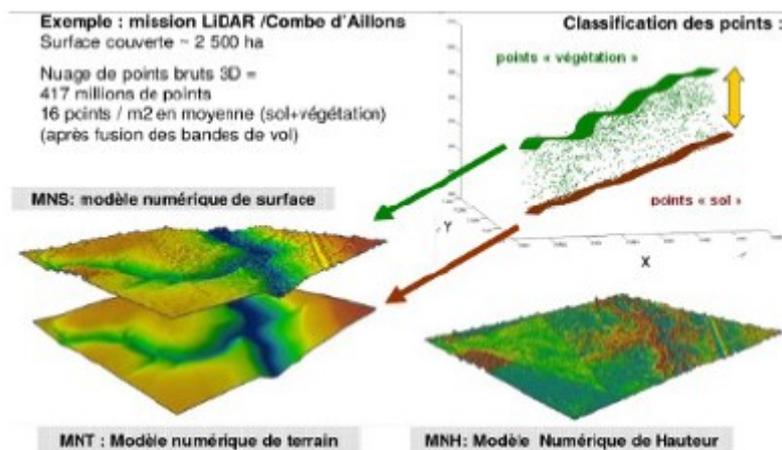


Figure 3 : Lidar data acquisition Credit (IGN)

It has a certain popularity for 3D renditions where the qualitative (the aesthetics) prevails over the quantitative (the precise measurement).

About our point of view, we are interested in measurements. They are made with a pulsed laser of which we collect only the first echo of all emissions and we must make measurements as precise as possible, on objects that we want, ultimately, locate in a point cloud. The latter is provided to us without any particular classification, it is up to

us to eliminate the outliers and conduct the appropriate grids, 2D or 3D in order to best calibrate ourselves with the objects to be analyzed. The processing of 3D objects is done particularly well by passing in 2D (by sections), which can be compared to tomographic processing. In fact as we will see later, this work is done in 2D, so on an image, which allows us to rely on very robust image processing algorithms.

### 3 Technological environment



Figure 4 : Terrestrial mobile lidar (T.M.L.) Credit(Geoslam)

The LIDAR GEOSLAM (figure: 4) used by the company EXAMETRICS provides the point cloud data that we analyze. These main characteristics are given below:

1.  $\lambda = 905 \text{ nm}$ , weight 1 Kg, displacement: 3 to 4 km / h.
2. Plan sweep area open to 270.
3. Fan rotation (2 Hz) volume (270 360).
4. Resolution 0: 625, ie 432 shooting angles.
5. The fan is scanned in 10ms (100 Hz). rate of acquisition 43,200 points per second.
6. Ambiguity distance 3500 m (erased by attenuation).
7. 1 single echo collected (the 1st per shot), therefore no classification.

Its laser scanning structure is a horizontal fan, the laser is carried by a head animated with a rotation of a frequency of 2 Hz. These laser shots, contained in a fraction of a sphere, return echoes in a completely disorderly manner, but with respect to the scanning.

We rely on the standard \* .las standard format de facto in the world of LIDAR [\(14\)](#). The construction of the \* .las file is carried out by GEOSLAM proprietary software linked to the measuring instrument and we have no control over this set.

## 4 Acquisition of data forest data : methodologies

The quality of the data obtained by LIDAR and the additional measurements is linked to the protocol observed during the measurement campaign. The measurement lidar is equipped with an inertial unit module (I.M.U. electronic component) which tends to drift over time, so it is recommended to avoid acquisitions lasting more than 20 minutes, between two readjustments of the inertial unit. In addition, 20 minutes of reading is 54. 10<sup>6</sup> points at the rate of 16 bytes per point, or approximately 900 MB of data.

Some recommendations:

1. Have an I.G.N. or equivalent and plan the physical route by gridding the zones, on any cartographic tool, identifying each zone identified after a readjustment of the inertial unit.
2. Bring a tape measure or decameter, laser rangefinder (an accuracy close to a few cm) and differential G.P.S. Ground measurements are necessary, using the range finder, this will remove certain ambiguities during the counting, the tape measure will make it possible to detect circumferences indicating the section of well identified trees, at breast height, for the reasons mentioned. A differential G.P.S., will ensure mandatory triangulation locations, with an accuracy close to a few cm.
3. On each reading, reset the I.M.U. (Inertial Module Unit), by placing the lidar on a tripod (itself georeferenced).
4. Ensure the horizontality of LIDAR during registration.
5. Orient the LIDAR North / South axis pointing north.
6. Read the position G.P.S. initial (Tripod position) (if necessary find the location of satellite reception, a G.P.S. type Etrex Garmin can do the trick at first, the luxury of the differential G.P.S. is probably a necessity),
7. Identify any landmark (Natural, by photo, LIDAR reflector, placed on the ground with compulsory GPS localization), necessarily detectable in the cloud to come, note at least 5 landmark points (which can also be singular trees easily identifiable in a point cloud) covering the measured area as well as possible and with the best possible accuracy. These five GPS reading points, are taken so as to build vectors as large as possible, non-collinear and orthogonal if possible for 2 of them (cf. [§11.1](#)).
8. Any statement G.P.S. of trees should be done at the foot of it and south of the tree (for example) to preserve the consistency of measurements.

9. If future aggregation (to be avoided if possible) envisaged, overlap 10 to 20% of the previous survey with homologous points (and landmark) common between surveys N and N + 1.
10. Identify any landmark (Natural, by photo, LIDAR reflector, placed on the ground with compulsory GPS localization), necessarily detectable in the cloud to come, note at least 5 landmark points (which can also be singular trees easily identifiable in a point cloud) covering the measured area as well as possible and with the best possible accuracy. These five GPS reading points, are taken so as to build vectors as large as possible, non-collinear and orthogonal if possible for 2 of them (cf. 11.1).
11. As far as possible identify G.P.S. start and ending point for each batch of measurement.
12. Refer to the statement and its associated data (all information relevant to geopositioning).
13. Create a cloud reassembly grid on the initial plan by perfectly georeferencing the zones.

These remarks are provided ex-post, because the current work revealed these prerequisites to be respected, in order to facilitate the extraction of data.

## 5 Extraction of the digital terrain model

The GEOSLAM terrestrial mobile lidar, due to its rotating movement, axis parallel to the ground, offers the advantage of providing many points on the ground. A so rich population of points allows reliable statistical extraction of terrain specific data. Knowledge of the digital terrain model (D.M.T), its relief, is fundamental in the context of the measures we perform, because on it depends the accuracy of the vertical measurements of trees that we calculate.

Indeed, in the particular context of the Massane forest, the relief is very rugged. So the global thresholding in Z (altitude) has no physical sense, about trees, on the initial surveys, hence the use of detailed knowledge of the terrain to bring the foot of trees to level zero.

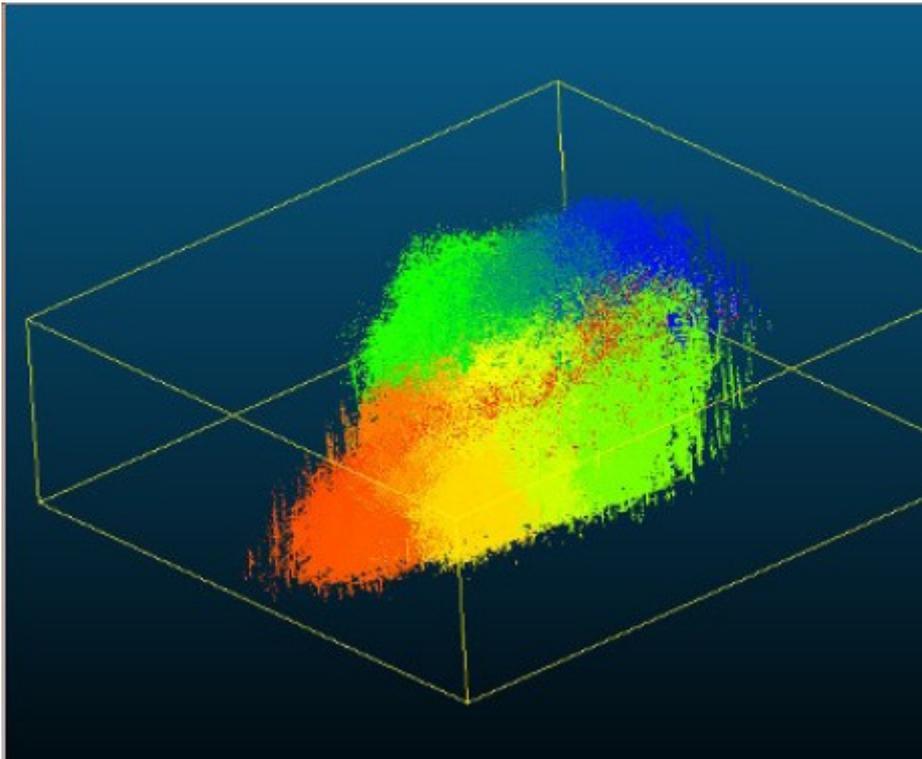


Figure 5 - Isometric view of the reference cloud in its parallelepiped container

in a very general context, this is also necessary for the same Reasons. Furthermore, in the measures at our disposal, noted Sintegra achievement, the deduced DEMs suffer from insufficient resolution, several possible reasons. We will cite two likely ones. the resolution of measures and the rules of adopted interpolation in connection with the choice of resolution, are insufficient. In consequence, we had to rebuild this DTM with a much better resolution, we made the choice to build 50cm  $\times$  50cm cells.

According to the number of ground echoes ( $\geq 5\%$  of the points population), a statistical extraction of the terrain model is possible. This also allows consider to

substitute our statements to those of an air flight with lidar classifying multi-echo. It will be a significant savings in terms of costs!.

We worked on the whole point cloud (see figure: 5), from the zone, from which the totally outliers are removed. Firstly we have given ourselves the following rules:

- Tiling (grid) necessary with a 50cm × 50cm size.
- Respect the spacing of the trees widely.

The treatments we offer require two readings in file:

- a first reading extracts the footprint on the ground from the area measured (point cloud) and define the limits of the grid (mini, maxi X, Y),
- the second reading ensures the calculation of the index of each point and therefore its assignment to the right cell and its statistical association suitable.

The entire cell assignment procedure is described below. After some tests, we went from a resolution of 2m × 2m to a resolution 50cm × 50cm, using the centimeter as the unit of measurement and not the metre. We have followed the following processing steps:

- Establish the precise size of the area (MinMax X, Y) and the rules of grid building,
- Perform a grid (X, Y) with the attributes collected by cell:
  - min, max of the cell on (X, Y),
  - number of LIDAR points in the cell,
  - average value of altitude Z (meanZ),
  - median value of altitude Z (medZ), but not used,
  - standard deviation of altitude Z.
- For the allocation of points in the right cell, we adopted the following rules, to avoid any recursive operations and to carry out on the fly in one single file read pass:
  1. Establish the precise size of the area (MinMax X, Y) treated.
  2. Define, with the grid pitch known, the terminals Sup. and Inf. bycourse for coordinates [X, Y].
  3. Being able to do a 2D (cell) and 3D (voxel) grid according to the needs.
  4. Avoid recursive operations, the result must be directly accessible by random mode, the addressing is done by an index built by concatenating its coordinates (X, Y) [cell] and (X, Y, Z) [Voxel] modulo the step of grid along the coordinate axis.
- Define the assignment of the point, in the correct cell by the cell index deduced from the coordinates of the point as indicated below:
  1.  $X_{modulo} = [(int) (X_{real} / f_{pas})] * l_{pas}$ . (Integer I, floating point F).
  2. index = Xmodulo! Ymodulo! Zmodulo, a character string.

Thus in a direct triede (cf. fig. 6) one can collect the LIDAR points in a cell (2D) or a

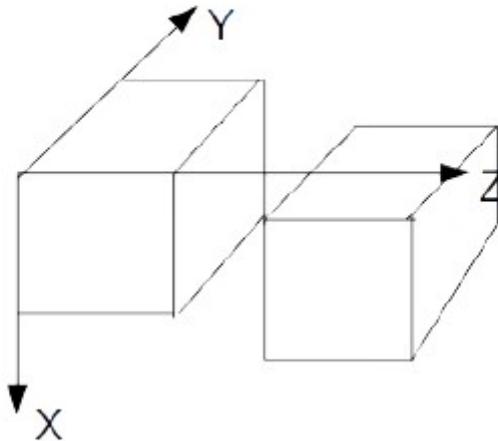


Figure 6 - 2D, 3D grid construction with point statistics per cell

voxel (3D) and have direct and fast access so random without recursive octree construction. This is greatly facilitated by C++ by using the MAP object <index, item> which fully dedicated to random access. The construction of the cell index (its identifier) turns out easy by a modulo operation the cell pitch, according to the different coordinates of the latter. Any point respecting this index is assigned to this cell.

At the end of this operation, we have a vision of our cloud initial organized in cells of 50cm × 50cm and no longer of elementary points. We have statistics at the level of each cell, number of points, Zmin, Zmax, Zmoy,  $\sigma_z$ , therefore:

1. Each point present in the cell contributes to statistics on the attribute Z (altitude).
2. All the points of the cloud have been classified in a cell.
3. We normally assume a normal distribution (To be deepened) of elevations.
4. We retain 5% of the population such as Zmax ( $Z_{mean} - 2 \times \sigma_z$ ) and we makes the assumption that it is the ground.
5. We impose a minimum density (number of points per cell) for making the calculation more reliable (criterion to be specified), to this end a vector of eligible cells is constructed.

At the end of this classification, a third treatment pass is carried out, with new file reading, respecting the previous criteria:

1. The point must be in an eligible cell by density,
2. Its Z coordinates is  $\leq (Z_{mean} + 2 \times \sigma_z)$ , which presupposes its appartnership on the ground.

We thus repopulate a new map (MAP C++ object) of cells containing only points presumed to belong to the ground, Statistical data Zmean and  $\sigma_z$

are calculated again for each cell. A filtering treatment allows to smooth neighborhood disparities, by adjusting the statistical mean .

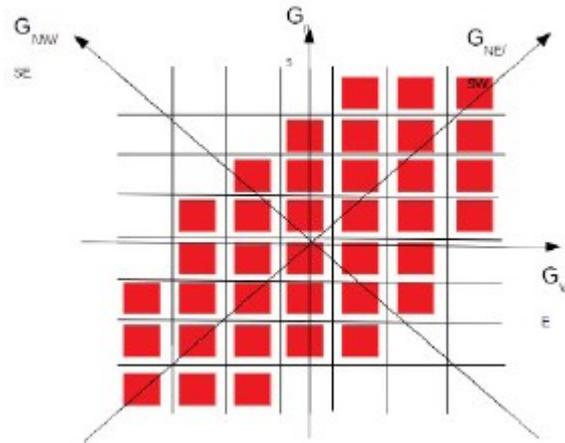


Figure 7 - Orientation of evaluation of the unlevel gradient coefficient

This treatment is provided by a standardized  $9 \times 9$  convolutional filter. Knowing that the coefficients weighting (see figure: 7) for calculating the central cell are deduced from gradients in overall elevation (over the whole area) in the directions:

- North / South,
- East / West,
- southwest / northeast,
- south-east / north-west.

The calculation of gradients is to be considered on a smaller area. This suite of operations gives the terrain model presented in the figure: 8.

On the other hand, the smoothing filter has no influence on the perimeter of the area, where we see outliers appear. As seen in Figure 9, outliers are eliminated by a KNN neighborhood filter:

1. 9 cell filter ( $3 \times 3$ ),
2. The central cell is elected if:
  - joint ownership report (xOy plan) with at least two other cells chosen in density,
  - difference in altitude, between this central cell and the closest cell concerned, less than or equal to the ground resolution (cell size rule, here 50 cm). This allows us to reach the next DEM (Digital Elevation Model) where the perimetric aberrations are removed.

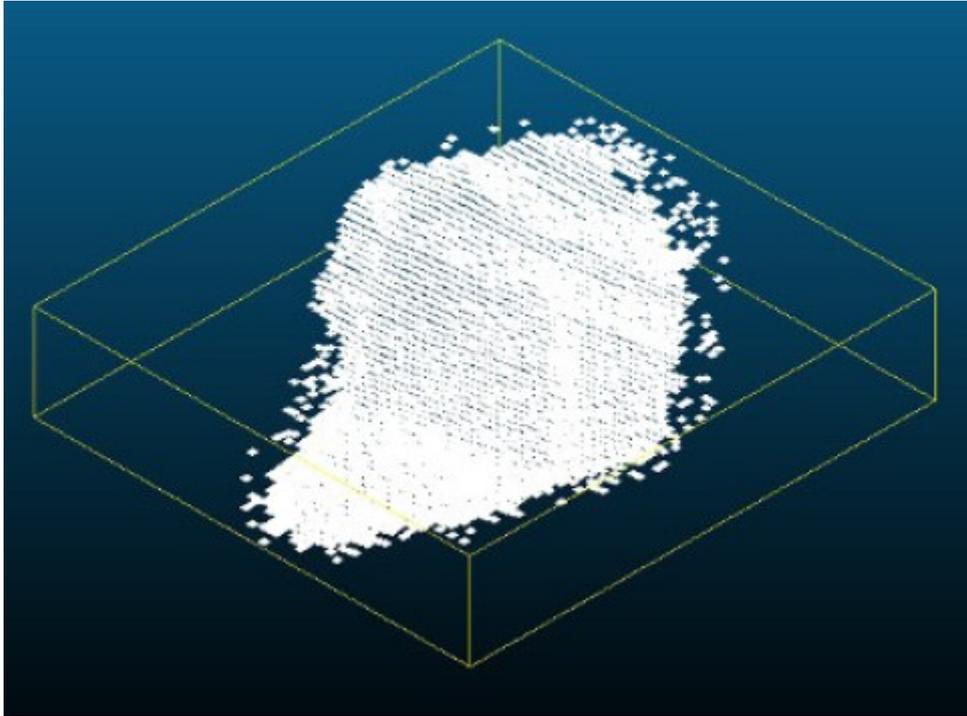


Figure 8 - Terrain model after original point cloud smoothing with the gradient filter

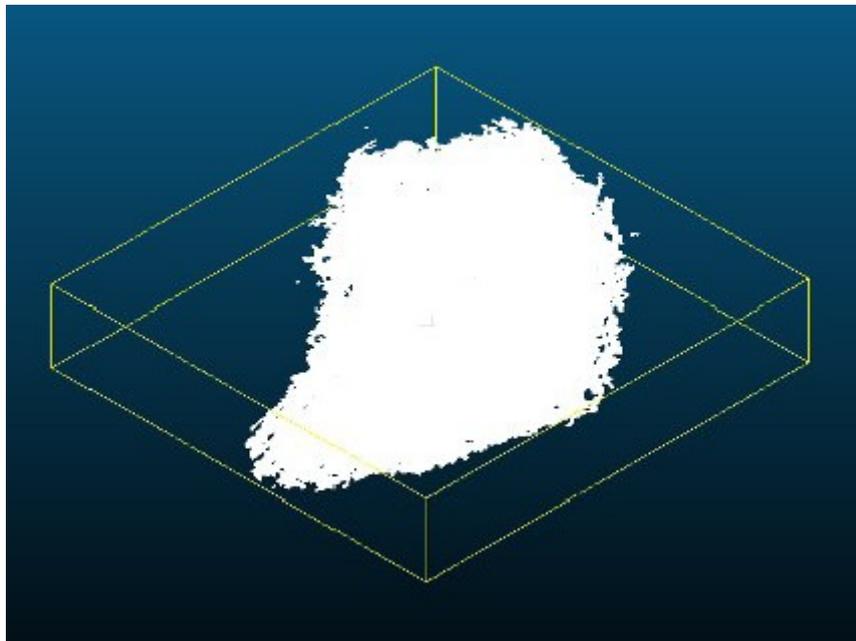


Figure 9 - Terrain model after smoothing with the KNN filter

## 6 Zero relief and tomographic sections

Below (see figure: 10), the digital terrain model covered by the initial point cloud.

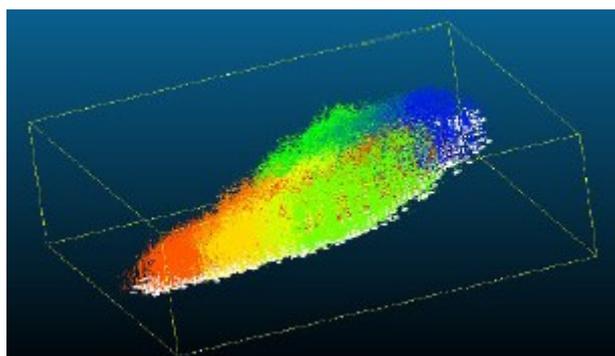


Figure 10 - Terrain model covered by the initial cloud

Once we have the digital terrain model the relief of the initial cloud is the result of a simple subtraction between the initial file cloud and the DTM file (see figure: 11). In the figures below, the colors, for the first image, are representative of the G.P.S. hours of data collecting on spot and for the second of the altitude

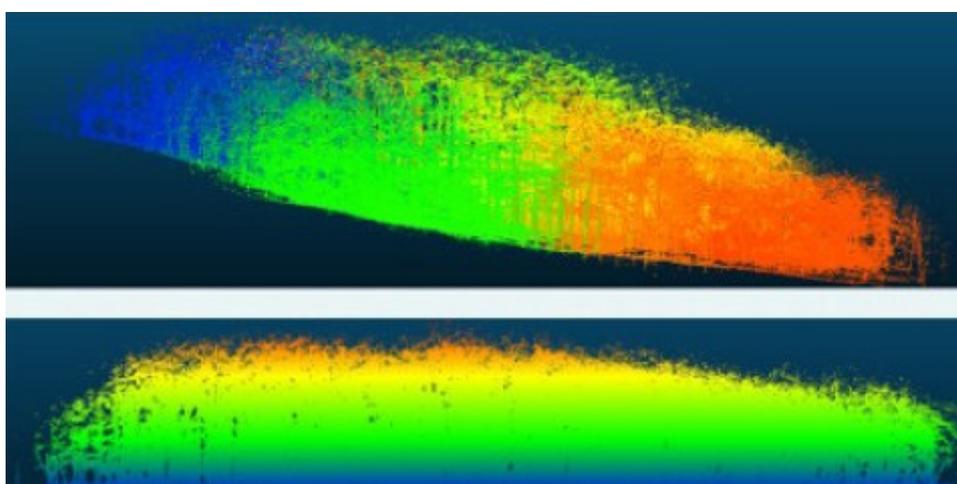


Figure 11 - Points cloud before and after zero relief operation

As seen in the image above, it is now possible to work uniformly and homogeneously on the height of the trees in relationship to the ground that carries them. To this end, we can perform extractions of thin section (some cm) , according to the geometry of our choice to identify trees or section of trees according to horizontal (see figure: 12) or vertical cutouts (see figure: 13). This type of normalized relief referenced to zero, gives us the opportunity to drive **tomographic** sections (sections), following the three ZOY section plans, ZOY, ZOY, ZOY by moving on their respective normal  $\vec{I}_x, \vec{J}_y, \vec{K}_z$  on the coordinate of our choice. Figure 12, taken from a CloudCompare screen, shows a horizontal section of a fraction [60m x 35 m] of the initial area and on which



Figure 12 - Example of horizontal section of cloud with zero relief

we can perfectly distinguish a tree lying on the ground and many convex envelopes (pseudo circles) which are, in fact, sections of trees 70 cm above the ground. These are our first objects of interest.

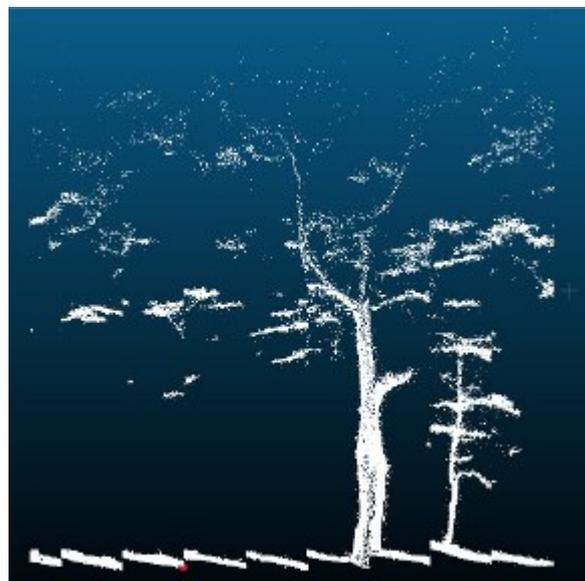


Figure 13 - Example of vertical section of cloud with zero relief

Figure 13, also taken from a cloud compare screen, shows a vertical section tree, located in an area of our choice when defining of the extraction of a vertical section. We will note at the foot of the tree the necessary improvement of resolution (in the example the grid size is 2m x2m). A need for resolution improvement of our grid will bring better result for the DTM gradient. This augurs well for all treatment options we can consider to provide.

These tomographic sections allow us to build images \*.png or others for the conduct of future classification algorithms and or shape recognition in horizontal or vertical section. The projections are driven on one of the three planes ZOY, ZOX, XOY, as required, the axis of projection being their respective normal  $\vec{I}_x, \vec{J}_y, \vec{K}_z$ . This allows us to transform our processing from a 3D point cloud into a 2D image processing.

## 7 Points cloud sectioning and image creation

So, in order to simplify the processing and to be able to rely on technology proven, as explained above, the necessary is done to work in image mode and in 2D, using orthogonal fitted projections. The tomographic sections performed allow us to lead to two complementary methods for tree attributes definition:

1. The vertical sections allow us to build histograms density along the Z axis and build morphological profiles individual for trees, allowing to quantify for example the volume of construction lumber like that of biomass. This exploitation is clearly part of the R&D tasks to come.
2. Horizontal sections highlight variable tree sections depending on the height, these cutting planes are transformed into image files (JPEG, PNG) on which we apply circular counting filters; the number of pixels which allows us to make a classification of the diameter of the trees at section heights, with precision +/- 10 cm. This approach is the result of an attempt to apply the circular Hough transform (cf. [\(8\)](#) which proved to be very imprecise, and subject to a large number of errors due to:
  - the low circular similarity of the sections,
  - frequent non-closing of the convex section envelopes.

Consequently, other approaches were necessary to allow the identification (location and measurements) of pseudo circles (convex envelopes) of tree sections. We adopted a deterministic correlation approach by sliding window of the size of the diameters sought on the section images horizontally. Furthermore, in order to increase the probability of closure convex envelopes, we have increased the vertical thickness of the horizontal section which has the effect of increasing the density of points projected on the ground, this increase the capability of tree detection at the expense of diameter accuracy. This will be corrected in the near future, some test are on progress about. As we know in the image, this image files being standardized, in terms of geolocation, we find easily the precise geographic location on the correlation peaks, using the accumulation pivot method explained below (Section 10.3), in order to refine this position with the best possible accuracy.

## 8 Construction of image files, standardization, noise reduction

The gross transformation, by projection onto the XOY plane of the parraleliped extracted from the point cloud, provides a noisy image. The figures 14, 15, 16 that follow give, except a light shift mirror error, and in order:

- A point cloud (CloudCompare screen) in top view of a thickness in Z of 30 cm, the objects of interest are the pseudo-circular sections which corresponds to horizontally sectioned tree trunks above the ground, lying trees, the rest being leaf echoes or branches or even measurement noise or artefact.
- The same point cloud projected in the XOY plane, extracted PNG image by calculation of projection on the ground of the point cloud for a section of 30 cm at a given height.
- the same denoised image, projection of the same point cloud cleaned and presenting the only objects of interest (convex envelope) from which we will extract automatically, the position (X, Y, Z) transformable into RFG 93, and the diameter depending on the filter size used.



Figure 14 - Example of horizontal section of cloud with zero relief

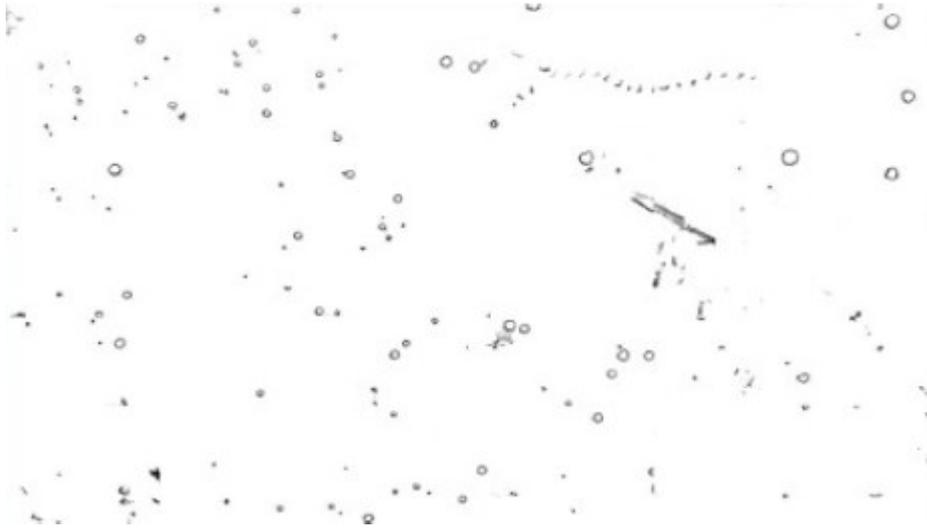


Figure 15 - Example of generated raw image

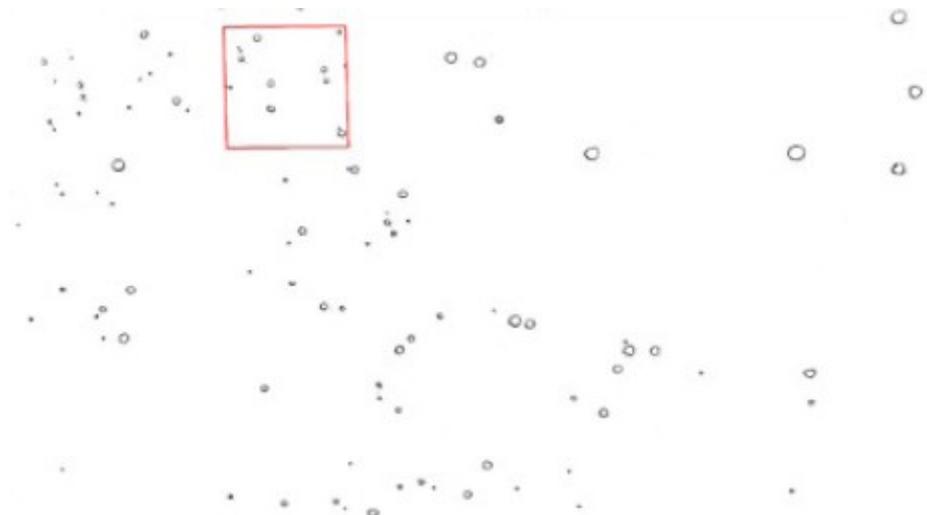


Figure 16 - Previous image denoised

## 9 Morphological erosion and tree sections library

Depending on the quality of the collected data on spot, the density of points on the convex envelop of a tree section may be insufficient. In the future, this could be partially corrected, by respecting a strict methodology in terms of field surveys. but immediately an erosion nucleus has allowed to artificially improve the density of the projected points, a square core erosion (1.0.1; 0.2.0; 1.0.1) was implemented before processing the image. The section of trees projected on the ground are pseudo-circular

sections whose collections of convex envelopes could ultimately build shape libraries for machine learning. But in the current state some of our work, we do not have enough forms, for us to embark on this path. This line of work which was envisaged in (7), could resume deep learning, drawing inspiration from the work of (9). This will be discussed in later R&D tasks. The figures below give an overview of the diversity of some convex envelopes, representative of tree sections. We let's give the appearance before and after erosion for a grid of examples.

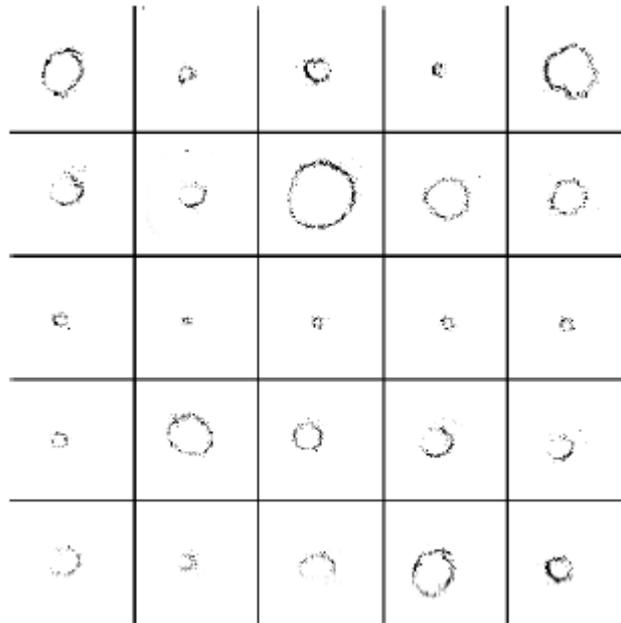


Figure 17 - Example of convex envelopes of tree sections

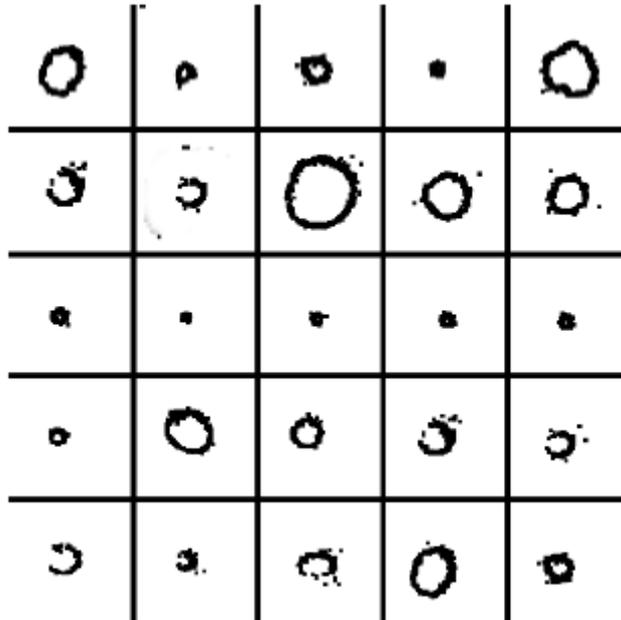


Figure 18 - Example of convex envelopes of tree sections after erosion

## 10 Location and evaluation of tree diameters

### 10.1 Technical filtering implemented method

We ensure, the mode of detection of trees, in an image, by a correlation circular filter. The main characteristics, to remember, of this filter are:

- It is a sliding filter, the pixel is equivalent to the centimeter.
- The filtering principle is based on counting common pixels between the filter and the image at a given pixel position in the image and not on the weighted contribution of the neighborhood, this is binary detection and counting resulting in a surface histogram, it can be presented as a raster image whose correlation peaks (Figure 20) emphasize a presumption of the presence of trees.
- An adjustment parameter related to the size of the filter (diameter of the tree) whether or not to elect a filter output as a presumption of the presence of a tree of a given diameter.
- The smaller the size of the desired tree diameter, the longer the processing grows as the square of the image size ratio over the filter size. As a result, acceleration mechanisms are put in place to quickly exit all areas with low pixel density or without pixels.
- At each chosen point, we retain its position in the image and the number of pixel intersection between the filter and the convex envelope.

- The filtering strategy recommends doing the detection treatments starting with the filters of larger diameters and progressing by decreasing size.
- After processing a given diameter, the convex envelopes of the image specific to this diameter, and detected, are deleted from the image being processing, speed is increased and ambiguities detection with the filters for smaller diameters reduced.

The figure below shows:

- image 1: an example of a convex envelope after erosion,
- image 2: a filter adapted to this convex envelope,
- image 3: intersection of the filter with the convex envelope, giving no correlation peak,
- image 4: position of the filter giving a maximum for the correlation peak, the coordinates of the correlation peak correspond to the position of the center of the filter in the processed image, that is to say the location of the tree.

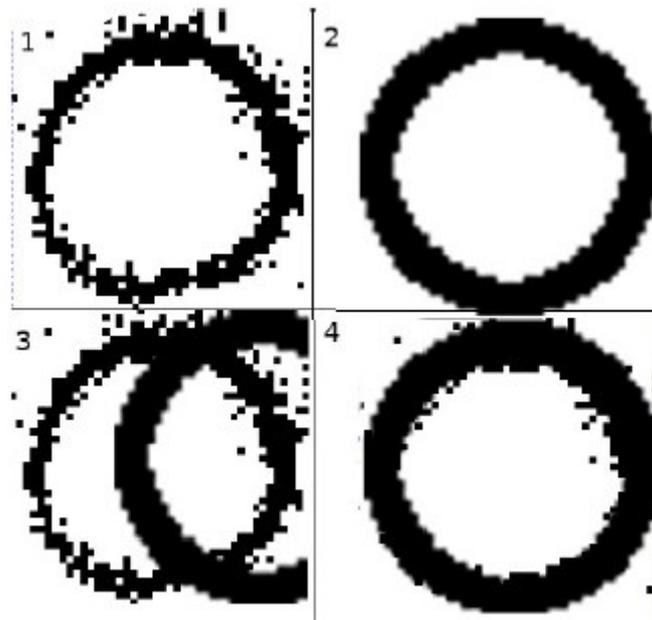


Figure 19 - Tree detection elements by diameter filtering

## 10.2 Construction of the filter and methods of utilization

The filter is a circular crown of 4 pixels, centered on the diameter of reference  $D$  that we are looking for in the image. The previous filter feels like a square image whose side measures  $(D \text{ diameter in cm} + 4 \text{ cm})$ . Indeed in our measurements, we took the centimeter as a reference and in our images 1px is equivalent to 1 cm. Consequently, the filter for detecting 80 cm diameter is a square image of 84 px side. Coded in level of gray (2 values 0, 255), it is in the form of a circular crown

centered, 4 px thick and black color, on a white background. As we have presented, a tree section is not perfectly circular. Our best fitting filter identifies the proximity of diameter and once the pivots of localized accumulations is completed, a complementary and fast control ensures the best classification between 3 possible diameters:  $D$ ,  $D + 10\text{cm}$ ,  $D - 10\text{cm}$  for the identified position, this also contributes to our tolerance of  $\pm 10\text{ cm}$ , on the diameter.

### 10.3 Search for accumulation pivots

The previous filtering, elects a certain number of points as correlation peaks. Figure 20, on the left, presents the raster image resulting from one filtering diameter of 30 cm, where the clear points correspond to the peaks of correlation.

On the right, an example of a correlation peak [position : 60, 110], where the accumulation is done in a neighborhood of 8 to 10 px around the peak (1 pixel is 1cm). This neighborhood feature is used for point elections. In fact, these are clusters of points surrounding the area where there is a strong presumption of presence of a tree the size of the filter template. This cluster of points is generally in a limited area of a few pixels in the correlation image.

The correlation raster is an archipelago of clusters of points, which we have to find for each one its center of gravity which will correspond to the effective position of a tree the diameter given by the filter attribute diameter. The principle of the accumulation pivot is very simple for the treatment that we will describe:

- a buffer A received the elementary points of all the clusters, neighborhood points to pivot are searched,
- a buffer B (originally empty) which works in flip-flop with A receives all points not related to the current pivot,
- a buffer C makes the weighted accumulation of the coordinates of the linked points at the pivot, the weighting is ensured by the number of points linked to the selection of the pivot,
- it is strictly a barycentric calculation,
- the population of buffer A being exhausted (A is empty), B contains all the points of untreated clusters, we perform the calculation of the barycenter of the data of C, we preserve the result, we empty C, we swap the functions of A and B (toggle or Flip flop) and we resume the operation with the new buffer of elementary points untreated clusters,
- this sequence of operations is continued until the piles are exhausted.

Clusters are groups of accumulation points separated by at most a distance of  $D$  pixels (8 to 10 px). During the examination of the cluster buffer at process, the first point encountered becomes the pivot and we retain all the points at a distance  $d$  ( $d \leq D$ ) and their accumulation will allow us to calculate the barycenter of the current cluster, i.e. the position of the tree. At the end of this treatment we have for each identified tree:

1. from its section diameter to the extraction height of the cloud about relief zero,
2. its location (X, Y) in pixels in the reference coordinate system of the image,
3. the possibility of deducing its coordinates (X, Y) in the point cloud source previously projected in the RFG93 standard,
4. the possibility of knowing its altitude, by mean of the terrain file D.M.T (cf. §5).

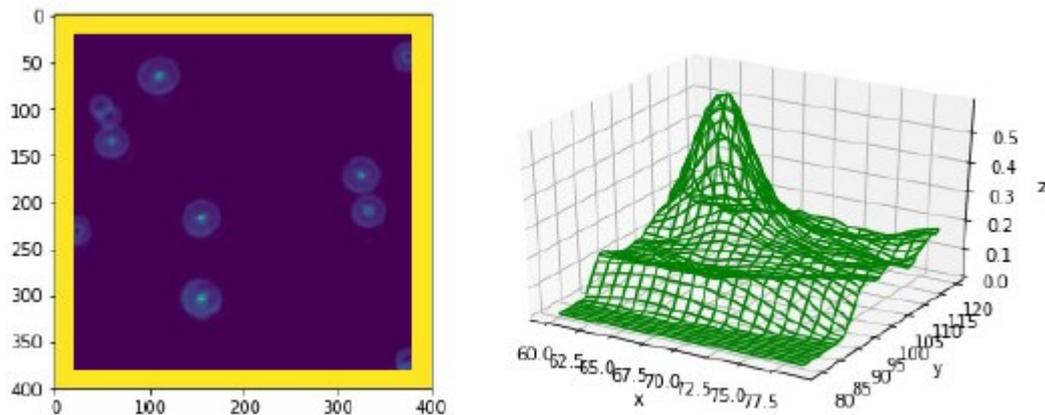


Figure 20 - Raster image resulting from correlation and a correlation peak

We have a geometrical transformation to ensure positioning trees identified in their own place using the transformation operator, whose evaluation mode will be given below.

## 10.4 Final elections and removal of duplicates

Our filter builds a cluster of tree presence likelihood points, to the diameter of the filter. Each point of the cluster (set on the central position of the filter sliding at a given position) is itself characterized by the number of the points of intersection it represents between the filter and the presumed tree section. It works like a likelihood function with complementary refinement.

The calculation of the barycenter of the points belonging to this cluster (in a neighborhood of  $7\text{px}$ ), will give us the center of the presumed tree section at best. At once we have this position, direct counting by precisely calibrating the filters ( $D+10\text{cm}$ ) and ( $D-10\text{cm}$ ) is immediate, this allows us to make an election between 3 diameters (best-fitting search). We retain the one represented by the filter which collects the more points by being wedged on the previous barycenter. Furthermore, the data trees (currently limited to only diameter) are stored in a table indexed by position, which excludes any duplication.

Finally, we practice a search by decreasing diameter, once a section of a tree classified to a given diameter is detected, we erase it from the image, this allows to speed up (the blank areas in points are skipped) the filtering operations very time consuming. A construction by linked lists of the only areas contained points by algorithms and their variant of type KD-Tree, R-tree should be considered in later work to speed up calculations without neglect a complete computing parallelization.

# 11 Transform operator between C.R.S

## 11.1 Ground checking of readjustment measures

Basically, the problem to be dealt with is the operation designated by "registration" which consists to readjust, aggregate two point clouds with common elements. A priori one of the clouds is considered reliable, representative of reality. It is the reference or the target, and it's supposed to be correctly defined, with the best possible accuracy in the correct C.R.S (coordinate reference system), in other words the right "datum" in terms of geography and/or geolocation. The other cloud, defined in the same C.R.S is positioned in a wrong way, it is a question of realigning it correctly by an operation associating :

1. one translation,
2. two rotations.

Considering that within this point cloud, the relative measures are correct, we operate a rigid transformation. To be precise, we will consider that the data collected by the measuring instrument, data to be stall or registred, constitute the source cloud. However, the GPS readings taken on the terrain will constitute the target cloud. Figure 21 illustrates this in a stylized way. **Ns** denotes the source cloud and **Nc** the target cloud.

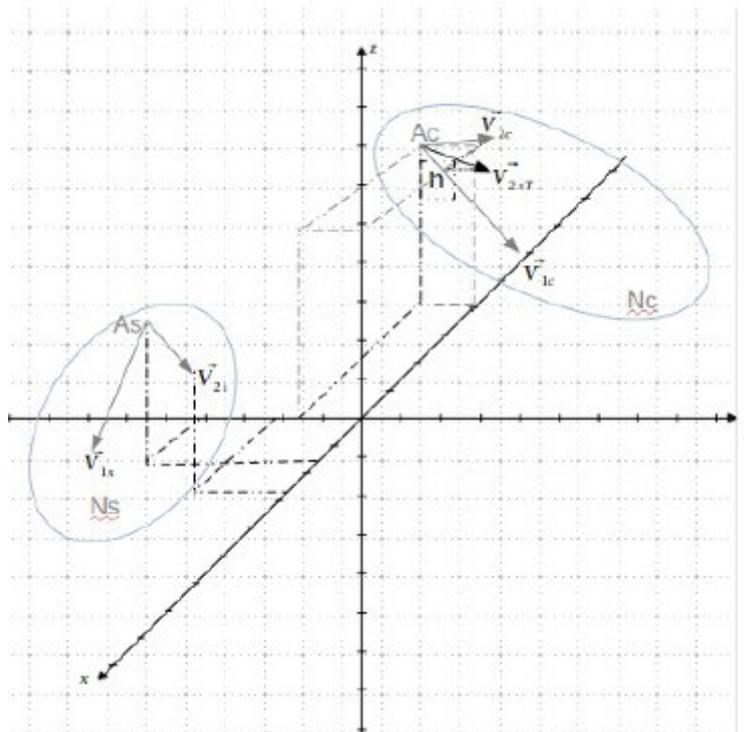


Figure 21 - Source cloud and target cloud

The data, collected by a terrestrial mobile Lidar scanning points with their coordinate [X, Y, Z] in 3D C.R.S , are geolocated on the basis of a construction exploiting

accelerometric and gyrometric data provided by the inertial unit equipping the LIDAR. However these geolocation data are given in the instrument euclidian reference system, with its origin calibrated 0, 0, 0 but in compliance with the reference of the coordinates specific to the device (WGS84 in this case either srid 4326) provided by the post-acquisition process. Indeed, this type of mobile instrument is dedicated, among other things, to measurements in areas not receiving permanent GPS of reliable quality reception. This is the case in a forest environment, but here the concealment of the signal GPS is not permanent. This allows us to make occasional geolocation (see paragraph 4 which details the rules facilitating future readjustments).

As indicated, in paragraph 4, on the spot measurements are combined to singular G.P.S points collecting, acting as landmarks of which we could, a priori, easily identify the counterparts (conjugate point) in the resulting point cloud of the LIDAR measurement. Indeed it is necessary to correctly readjust the cloud of measurement points in the geographic C.R.S with which we want to work. The LIDAR measuring instrument, with its inertial unit (I.M.U) , provides the source cloud. Conversely, the G.P.S collected data are part of the target cloud. However, on starting, two aspects are very important :

- for each point measured, by GPS, they are given in a format linked to a projection datum on a given ellipsoid which is not necessarily the same in the measuring instrument (WGS84, srid 4326) and / or and the geo format geographic location we want to use (RFG93 or Lambert93 in our case or srid 2154), this also calls for a transformation that apply both to LIDAR data and to data received from G.P.S surveys.
- concerning the C.R.S in which we work, LIDAR measures are conducted in the instrument's own 3D euclidian reference system of measurement relative to its own origin, the geolocated data of the terrain, provided by GPS are also provided in a 3D C.R.S, has the advantage of being well positioned and correctly oriented. Finally the two C.R.S must be the same, this is obtained by the transformation to carry out, we must to make coinciding by this transformation (one translation, associated to two rotation) the two C.R.S.

Listing 1 - C ++ code using listings, as a reminder:

```
char * srid2154 = "+ proj = lcc + lat_1 = 49 + lat_2 = 44 + lat_0 = 46.5 + lon_0 = 3"

srid2154 += " x_0 = 700000 + y_0 = 6600000 + ellps = GRS80 + towgs84 =
0,0,0,0,0,0 ";
srid2154 += " units = m + no_defs " ;
char * srid4326 = "+ proj = longlat + ellps = WGS84 + datum = WGS84 +no_defs" ;
projPJ target = pj_init_plus (srid2154);
projPJ source = pj_init_plus (srid4326);if (source == NULL || target == NULL)
    return false ;
```

Registration is done by identifying similarities in the source clouds (the homologous points) and in the target cloud (landmarks measured on the ground). Concerned the measured landmarks which will be used to readjust the cloud of measurements (cloud source) in the desired geographic repository, we work on the basis triangles (3 points: insurance of a planar surface). The principle is to readjust the measurement cloud, the source cloud, in the geographic reference system we want to exploit or target cloud. So the target cloud collects the readings GPS performed in the field (landmarks), while the source cloud is provided by the LIDAR measuring instrument and contains a priori the points homologous to previous landmark and easily identifiable. Finally, we end up with two pairs of vectors (3 points) one in the source cloud (the counterparts) and the other in the target cloud (the landmarks) that we must necessarily superimpose the two couple of vectors for the best fit.

## 11.2 Selection of satisfactory GPS measurement points

However, we carried out various ground surveys. However, some cares are needed, concerning GPS data collected. Peer to peer vector association must be carry out with some checking, angle and length conservation of conjugate vectors (In target and source cloud) must be preserved with the best possible accuracy. We are in the the field of measurement, so we have to discard the outliers of these couples (landmarks / conjugate point). We will do this in two steps. Beforehand remember some basic rules:

- a line requires knowledge of 2 points, with N measured points on spot, we are able to build  $(N-1)!$  geometric vectors,
- building a triangle requires two vectors with a common anchor point, starting from each point recorded by G.P.S on spot, it is then possible to build  $(N-1)$  triangles with the same anchor point, knowing that we have N points this offers the possibility of  $N * (N-1)$  triangles (couple vectors with the same anchor point)
- However, all these triangles does not have the same validity (length, angle disparity to check) for the calculation of a transformation matrix that is reliable, so we have to ensure a selection among these.

In terms of evaluation and calculation of the transformation operator, a single couple triangles is enough, the first from 3 landmarks in the target cloud, associated with the triangle built in the source cloud with the conjugate points of the three previous landmarks from the target cloud. More, the triangles of this couple must be as similar as possible, in order to respect the identity and conservation, so best fit of the lengths and angles produce a suitable transformation operator. The superimposing and overlapping of the two triangles require two rotations and a translation, operations of which we have to calculate the matrices  $(4 \times 4)$  in corresponding homogeneous coordinates. The election of several triangles would allow to control somewhat a priori, the validity about data of the resulting computed matrices, but the criteria of identity and similarity are paramount. Before calculating transformation matrices, for validation triangles, we will rely on two measurement criteria, on each triangle, in order to select the optimum couple:

- minimize the deviation of the module length calculation of each similar vector taken two by two (source and target),
- minimize the difference in the calculation of the angle at the anchor point between a triangle of the target and its counterpart in the source cloud. The accuracy of the calculations is also linked to the relative sizes of the modules of the vectors (retain a ratio  $> 0.7$ ) and that of the angle at the top (favor  $45^\circ$ ).

For the record, the minimization of the difference in the area of the surfaces of the triangles, taken two by two, is consistent with the two previous checks. As indicated in § 4, we take care to note several points on the ground, in order to have an optimum choice for the evaluation of the matrix specific to the transformation operator.

## 11.3 Possible sources error and its quantification

### 11.3.1 Sources of errors

As we have just presented, we note that the risks of errors are constantly present. In the previous paragraph, we highlight cause, measurement anomalies during GPS landmarks point readings on the field, which further confirms the need for maximum precautions to take in this phase. But this is not the only source of errors, we will add :

1. rear projection errors, when switching from one ellipsoid to another especially when the projection modes are different, which has an impact on the conservation of geometric attributes (length, angle, area of surface) the conservation rules change depending on whether the projection is conical, azimuthal or cylindrical, a priori this is integrated in the space reference data supplied by OSGEO. These points are to be deepened explored concerning tolerances of the measurement chain, which is not done in this note,
2. chain calibration problems: measuring instrument + software processing of inertial and distance data (flight time), because in the in this case, the data is provided by an inertial unit. On the other hand in this second case, it is entirely possible for us to calibrate it, as we let us indicate it below, even if these operations were not carried out initially, but they will need to be conducted in the future in order to qualify the measurement procedures.

The principle of calibration that we propose makes it possible to evaluate the linked quadratic error:

1. non-conservation of angles at the anchor point (common point of the vector couple),
2. non-conservation of lengths.

We will be interested in the first, because it is the essential problem that we meet with our landmarks statement, the second aspect is treated in a way similar we will not develop it here.

### 11.3.2 Proposal for post-measurement correction

Account of the absence of initial methodology in the measurement reports LIDAR, in the undergrowth, which we have, we conducted GPS surveys ex post ion spot. We reverse the method, that is, we have identified in the point cloud of the singular elements which we have noted, afterwards, the position on the ground. We note on the 20 points raised that there is:

1. reasonable preservation of the length, the gap between the point cloud and the ground readings are on average 0.7%,
2. on the other hand, concerning the angle of the anchor point, it can vary within a large proportion, from 0 to 75%.

Consequently in the present case, taking into account the history of the statements, we offer an ad hoc correction allowing a balanced data analysis. But in the future, any calibration economy should be ruled out, the correction method that we propose here is only the constraint of this which must be done for the preliminary calibration, the details of which will be given in the work to come, but that will be similar and reversed from what we present here. The calibration proposal and its opposite say:

- if the conjugate points of the point cloud are set on the landmark points found on spot then the instrument is calibrated, otherwise we correct the lidar instrument or its measurement data,
- if the conjugate points of the point cloud are not set on the landmark points readings on spot then the GPS reading is false and we correct it or its measurements, otherwise the calibrated measurement chain is estimated.

All proportion kept and the talent less, we would be in the syndrome possible of Michelangelo's error. Would he have made a mistake by sculpting the nose of the statue of David. What should he do? Correct the statue's nose at risk to alter it or that of the model? This joke to recall the fundamental dimension of measurement and calibration methodologies.

Given our history, we are in the second case proposals for calibrations we proceed as follows. Among our field readings, we will take the one or the angle error at anchor point is the weakest, we will denote it by  $\Theta_\epsilon$ . Either at the point anchor, B and C the other points whose GPS readings were taken on field. So we have the triptych:

- the vector  $\vec{AB}$ ,
- the vector  $\vec{AC}$ ,
- the angle  $\theta$ , at the anchor point A, on which we consider to have an error  $\theta_\epsilon$ .

Consequently, this correction is made around the normal  $\vec{N}$  at both previous vectors  $\vec{N} = \vec{AB} \otimes \vec{AC}$ , whose matrix we have to compute for rotation anchored in A around the unit vector  $\vec{N}$  by an angle  $\theta_\epsilon$ . A priori, we apply this rotation to the vector most affected by error according to our appreciation. Suppose that it is the end C which will pass to the position  $C_c$ ,

This will give the new triptych:

1. the vector  $\vec{AB}$  ,
2. the vector  $\vec{AC}_c$  ,
3. anchor point A.

which will then be used for the evaluation of the transformation operator, calculation that we see in the next paragraph.

## 11.4 Calculation of the transformation operator

As a reminder, two rotation matrices are to be calculated around two unit vector to choose. For both, as we will see, the anchor point is common, if the first rotation vector is normal to the first couple (source / target) similar, the second rotation vector becomes the superposition of these two vectors once normalized, but the plane and the angle of rotation are to be calculated.

The choice of vector pairs (landmark points and conjugate) was made in compliance with the previous criteria. Figure 21 shows the source clouds ( $N_s$ ) and target ( $N_c$ ) with pairs of similar vectors, respectively  $[\vec{V}_{1s}, \vec{V}_{2s}]$  anchored in  $A_s$  and  $[\vec{V}_{1c}, \vec{V}_{2c}]$  anchored in  $A_c$ . The calculation procedure, respects the following steps:

1. Translate the target anchor  $A_c$  on the source anchor  $A_s$   $\vec{T}_{as}$  and all on zero  $[0, 0, 0]$  of the frame of reference,  $\vec{T}_{s0}$  translation vector, which becomes the anchoring of rotations (origin of the source and target vectors), to evaluate the matrices of rotation.
2. the axis of the 1st rotation is the unit normal  $\vec{N} = \vec{V}_{1sg} \otimes \vec{V}_{1t}$  , the angle of rotation being the angle between these two vectors, this gives the matrix [M1] which ensures the superposition of  $\vec{V}_{1s}$  on  $\vec{V}_{1c}$  ,
3. the second axis of rotation becomes  $\vec{V}_{1c}$  ,
4. the transformation of  $\vec{V}_{2s}$  is calculated such that  $\vec{V}_{2sT} = [M1] * \vec{V}_{2s}$  ,
5. the second rotation matrix is obtained by superimposing  $\vec{V}_{2sT}$  with  $\vec{V}_{2c}$  , the angle of rotation being that separating their orthogonal projections on the axis of rotation  $\vec{V}_{1c}$  in h (see figure: 21), this allows us to obtain the matrix [M2].
6. the global transformation matrix  $[M] = [M2 * M1]$ ,
7. In the end, the transformation from source points to target points is defined with the following expression:  $[Point_{Target}] = [M] * [Point_{Source} - T_{s0}] - T_{as}$

## 12 Putting attribute data in the coordinate reference system workspace

In paragraph 10.3, we end up with the list of identified trees with their attributes (geolocation in RFG93 [PointSource] and the equivalent diameter of their section). Their position in the final C.R.S is given by the expression evaluated previously.

$$[\text{TargetPoint}] = [M] * [\text{PointSource} - \vec{T}_{s0}] - \vec{T}_{as}$$

This final result, given in the form of a CSV file makes it possible to transfer your data in the GIS of your choice.

## 13 Additional geometric precautions

As we saw earlier extracting position data and of diameter is made from an image, therefore information is in 2D. This image is the fruit of the orthoprojection of a slice of cloud (parallelepiped of a certain thickness **dZ**) on the horizontal plane, following the flattening of the relief by the withdrawal of the D.M.T. For the construction of this image, we define, its limits [inf., sup.] :

1. on the X axis or longitude [Xmin, Xmax],
2. on the Y axis or latitude [Ymin, Ymax],

it should be present in mind that depending on the height **Z** of the sectioning **dZ**, of the fact that we collect points in a 3D cloud, the previous bounds may vary, but distances and angles are preserved when projected. Consequently, in order to be able to use the same previous transformation operator, if we work on images from sectioning carried out at altitudes Z different, corrective translational patches are required. But this is easily done by the following linear operation which transpose the coordinates of our image points (results of filtering and accumulation pivot) in the CloudCompare display C.R.S ( $X_{\text{cloud}}, Y_{\text{cloud}}$ ) which is our measurement instrument on the initial point cloud (source cloud). The linear transforms to be implemented respects the following equations :

$$(a) X_{\text{cloud}} = X_{\text{min}} + X_{\text{image}},$$

$$(b) Y_{\text{cloud}} = Y_{\text{max}} - Y_{\text{image}},$$

where in respect of the reference structures of the image and cloud coordinates of points:

1.  $X_{\text{min}}$  is the minimum abscissa of the source point cloud,
2.  $Y_{\text{max}}$  is the maximum ordinate of the source point cloud,
3.  $X_{\text{image}}, Y_{\text{image}}$  are the coordinates of a point of the image to be found in the cloud source,
4.  $X_{\text{cloud}}, Y_{\text{cloud}}$  are the coordinates obtained in the initial point cloud (source) and to which the transformation operator will be applied for place them in the final C.R.S. (target cloud).

## 14 Presentation of final results

### 14.1 Geolocation



Figure 22 - Top view of the cloud

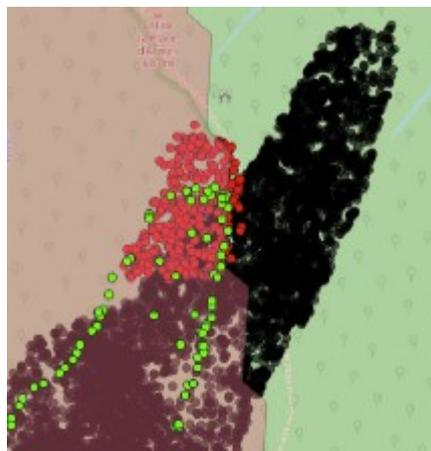


Figure 23 - Results in Qgis

Above, the result of the analysis of an area of about of 1.5 ha, gathered about 420 trees whose diameters vary from 20 cm to 120 cm. On the first image (figure: 22) we have a top view of a section (10 cm thick) of point cloud at 65 cm from the ground. It's a noisy image, the small circles are sections of trees. The three lines are profiles of the fence which delimits the protected north area of the Massane reserve.

The LIDAR statement incorporates information internal and external to the fence. On the second image above (figure 23) we find the same area included in the QGIS geographic information system. The perimeter of the fence (GPS survey) is indicated by the green dotted lines. Red dots (from our extraction operations) indicate the trees on the ground that are well in the expected area on each side of the fence, as shown in the previous figure.

This validates the principle of the pipelined transformation, allowing for each tree geolocation information and their diameter extraction automatically from a point cloud of trees in forest. The initial point cloud not having been the subject of specific collection rules, we nevertheless manage to restore with a certain precision the geolocation of trees as well as their diameters. The results of our calculations are given, for example, in the attribute table below, excerpt from a C.S.V file.

NumArbre	Longitude-X	Latitude-Y	Altitude-Z	Cloud-X	Cloud-Y	Cloud-Z	img-X	img-Y	Sid
237	702669,03...	6154747,4...	672,523	-53,41	-25,200001	-5,290004	5317	10251	-5350-2550
238	702647,04...	6154733,5...	671,544222	-78,230003	-25,780001	-9,37267	2835	10301	-7850-2600
231	702651,38...	6154746,0...	669,04496	-67,849998	-17,23	-9,49162	3873	9446	-6800-1750
232	702657,10...	6154749,1...	669,86700	-61,43	-17,73	-7,94258	4515	9496	-6150-1800
233	702642,19...	6154738,2...	669,202871	-79,699997	-19,02	-11,0155	2688	9625	-8000-1950
234	702653,28...	6154744,9...	669,775413	-66,870003	-19,219999	-8,89993	3971	9645	-6700-1950
227	702649,83...	6154762,0...	664,409086	-60,41	-2,31	-11,277	4617	7954	-6050-250
228	702647,83...	6154737,3...	665,2621	-64,640004	-5,38	-11,3227	4192	8261	-6500-550
229	702655,34...	6154739,7...	666,292159	-57,150002	-7,42	-9,65803	4943	8465	-5750-750
230	702643,00...	6154741,5...	668,798374	-77,199997	-16,639999	-16,0	2910	9387	-7750-1700
223	702698,48...	6154795,6...	667,74422	-2,27	-0,07	-0,509254	10431	7730	-2501-50
224	702699,45...	6154795,8...	667,942916	-1,33	-0,38	-0,234915	10525	7761	-1501-50
225	702654,97...	6154765,9...	664,693138	-54,099998	-1,8	-10,1514	5248	7903	-5450-200
226	702626,63...	6154747,7...	663,196762	-87,370003	-2,08	-15,7585	1921	7931	-8750-250
219	702649,17...	6154767,6...	663,33517	-57,990002	2,87	-11,3723	4859	7430	-5800-250
220	702626,78...	6154732,2...	662,015477	-84,769997	1,83	-16,1131	2181	7540	-8500-150

Figure 24 - Attribute data calculated by the processing application

In the attribute table above, we recognize for each tree:

- its identifier,
- its geolocation RFG93 on a forest site, resulting from transformation after extraction from point cloud,
- its position in the cloud of measurement points (Cloud-XYZ),
- its position (X, Y) in the ground projection image,
- the sid index to extract the altitude (Cloud-Z) from the database D.T.M. of the area, calculated by the application and then transformed into Altitude-Z.

The in-depth analysis of the results validates our initial objectives, presented in our introductory summary with the exception of the calculation times which remain the double those targeted due to lack of tasks, parallel calculations, threadings. Future tasks will address this problem.

## 15 Statistical assessment of the results

Based on the results obtained where we assume the possibility of error due to the inaccuracies of our GPS readings, it is still possible to conduct a statistical evaluation of the validity of our results. To this end we conduct the following calculation:

1. we take in order, two by two, the successive geolocated points provided by our algorithm,

2. we calculate the distance between these two points on the ground (RFG93 in meter) and on the projected image (1px = 1 cm), these are conjugate segments,
3. a priori these distances should be identical, we make the difference to have the absolute error,
4. to normalize the assessment, we do the relative calculation of the error (in thousandth).

The calculation is carried out on 452 segments, the table below provides the statistics of the relative errors that we evaluated on the comparative calculations: RFG93 distance versus projected image distance.

<b>STATISTIQUE D'ERREUR RELATIVE (Millième)</b>	
Moyenne	18,5854649490152
Erreur type	1,20288485374828
Médiane	15,4632424477124
Premier quartile	12,2218694914485
Troisième quartile	18,8934024456935
Variance	654,013251062417
Écart type	25,5736827825485
<u>Kurtosis</u>	227,578750683225
Asymétrie	13,3514336574673
Plage	474,871302876365
Minimum	0,000579121928813
Maximum	474,871881998294
Somme	8400,63015695489
Nombre de points	452

Figure 25 - Statistics on relative geolocation calculation errors

A quick analysis of the error statistics gives:

1. the average relative error is less than 2%,
2. half of the errors are less than 1.25%,
3. the third quartile begins with errors > 2%.

There are a few outliers, but overall the measurement accuracy relative respects our objectives. Of course, additional validation on the field will conclude this assessment.

## 16 Measurement, calibration results which method ?

### 16.1 Statement

Although not benefiting from calibration of the measurement chain, we considered that the data (position in the C.R.S of the measurement, and distances between points) presented in the CloudCompare viewer were correct (checking with new measure on spot). Regarding GPS ground surveys, even with resources restricted, we have taken sufficient care about measures (waiting for the GPS setting and comparison between 3 instruments), to record the 7 or 8 points correct landmark points including 3 to calculate the transformation operator. In addition, two post-measurement operations:

1. a validation sort (see § 11.2),
2. a correction on the angles of vector pairs (target) (cf. § 11.3.2),

allowed us to work on two triangles (source and target) strictly similar (length in meters and angle in radians identical within a tolerance of  $10^{-4}$ ). In total we had recorded around twenty landmark points, we have retained the most likely, in respect of our means of measurement.

As indicated in §11.2 we have analyzed the validity of the points raised by controlling the conservation of distances (in meters) and angles (in radians) with a tolerance of  $10^{-4}$ . This confirms a few points, those who were recorded with the greatest care. So with our 20 points raised, we built 190 segments. For 41 of them within +/- 2 m (maximum) we have the same lengths:

1. in the source cloud,
2. in the point cloud transformed by the calculated operator,
3. in ground measurements.

So for the 190 segments their lengths are evaluated in the source cloud and the cloud from the transformation operator. This also makes it possible to evaluate the validity of our calculated transformation operator. For the previous 41 (on spot measurements), 14 segments are in absolute tolerance  $10^{-2}$ m, the others at +/- 2 m at most (+/- 2% in relative error). For the rest, 139 segments, we consider them as associations with at least one point aberrant from our G.P.S. measurements on the ground with insufficient means.

At the count, we obtain without surprise, the positional extractions-tree, strictly respecting the point cloud. It clearly appears in our works, an imprecision on absolute geographic positioning, landmarks (insufficient measurements), our targeted objectives in term of :

- absolute geolocation are to be improved,
- distance measurement we respect those of the point cloud,

- value of angles between segments (relative position of the trees) we respect the point cloud.

If the point cloud is representative of reality, our approach is good. So the weakness that remains are linked to the strict account of our weak method of landmarks survey which can be improved. In any case, outliers landmarks are not used. We get results that reflect a profile of the reality on the ground, the approach adopted seems interesting and promising because it lets us see the possibility of fully achieving this that we're targeting.

But we face the pitfalls of method and calibration that we need having to prepare to get up right now, clearly laying down an obligation of measurement and calibration method:

1. first, the absence of a measurement and calibration protocol for the instruments strongly impairs the reliability of the data collected, but it is necessary to the credit of those who carried out these measures for having been able to provide data which to date support many points of reflection, objects of this note (No data: no verifiable reflections!),
2. secondly, the prospect of the economic attractiveness of such a measurement method, also led us to a certain neglected simplism (especially the ground readings made by the author) due to the low precision of the method used, we worked on economics, but this can easily be improved with more efficient instruments without really burdening the costs.

Regarding the problem of measurement and calibration methods we need know how to define, implement and qualify them. We give some way for improvements below. They will be taken up in subsequent works.

## **16.2 The error source elements in the measurement and processing chain**

The basis for the reliability of measurement readings are based on the correct calibration of the measurement chain and associated treatments. In this case, this includes the following method that have been used to:

1. Collect the data (accelerometry, gyrometry, Euler angle of the shooting laser, flight time) => the LIDAR GEOSLAM ZEB-REVO, this is of the manufacturer concerns in the performances announced in its specifications,
2. condition the previous data => GEOSLAM software also supplied by the manufacturer, in order to prepare the data supply file in a standard \* **.las** format and faithful to the data collected, this is also the manufacturer concerns about measuring instrument, always respecting performances announced in its specifications,
3. take the GPS geographic coordinates (WGS 84) on the ground, in order to associate them with singular points from the previous point cloud => GPS Garmin Etrex (clearly we are here in a situation of inaccuracy), a differential reading is essential, like the use of telemetry laser,

4. visualize the point cloud, used in particular for distance measurements in a given C.R.S and point location records (x(longitude), y (latitude), Z (height)) and the distances between points in an Euclidean C.R.S. => CLOUDCOMPARE,
5. ensure the geometric transformations resulting from the algorithms of the present works.

### 16.3 Possible calibrations (annual frequency)

Let's take the previous points one by one, merging the first two to talk about GEOSLAM chain:

1. for the GEOSLAM chain we offer to drive two types of calibration:
  - a static calibration in fixed point, the lidar on tripod emits in direction of LIDAR reflectors (geolocated) arranged at distances and distinct heights, within a maximum distance of 35 m. One statement 20 minutes, knowing that each lidar shot is dated allows you to appreciate the possible drift of the inertial unit over the entire chain of measures,
  - a dynamic calibration, LIDAR reflectors (geolocated) are arranged on an area of 2500 m<sup>2</sup>(50m × 50m) or more (to validate), the lidar walks in the area and we check the validity of the measures distances between points as well as respect (initial calculation and measurements) of the angles at the top of the reconstituted triangles between the different points.
2. for CLOUDCOMPARE, we note the location of two points as their distance, the consistency of these two data is easily verified : two points give a line, the Euclidean distance calculation is immediate,
3. for the ground point recordings, we will use other equipments, laser rangefinders for distance between points and differential GPS to locate the landmark points, these two measures being able to qualify with reciprocity.
4. the landmark points being collected, we must ensure the consistency of distances between these landmarks and their counterparts (conjugate) in the measurement cloud and correct it if necessary, without forgetting to check the identity of the conjugate angles in constructions by the trio of similar points, we did it in the present tasks, but in the absence of a calibration element, the use is not very convincing.
5. concerning the transformation operators of this work, their validity has been tested by translation / rotation operations between triangles alike distributed in 3D space and their superimposition is done with a accuracy of 10<sup>-5</sup> on the meter (triangle of approximately 1000m<sup>2</sup> and placed at distances of 700 km in the C.R.S).

## 17 About the bibliography

The present works are a logical continuation of the learnings mentioned in 1.1, (7) and bibliographic research cited in (6), they have not been the subject additional bibliographic research to be conducted. The reflection earlier dictated their construction. On the other hand, in 2019/2020 a new state of the art is essential in particular for subsequent work and for a better promote of the current achievement, in order to clearly identify what should be our areas of progress, in our future work (Paragraph 17).

Additional bibliographic references, not referenced in the present article, are given for the record.

We will be particularly interested in the bibliography cited in (15) just as at the work of the University of Sherbrooke in Canada in the Province of Quebec in (16).

## 18 Outlook and future work

In terms of algorithmics and processing, all the steps have be done, from the starting cloud to the extraction of the geolocated tree with these attributes for G.I.S database. Previous work, especially for machine learning which is not discussed here, about the detection and classification of ground forms as cavities and parasites on tree (fungi), will be taken up and completed. All elementaries modules specialized are rewritten in C ++, even in C\_CUDA. The robustness of the overall size must be verified, with great care, with from measures spot feedback.

Global integration is to be done with PyQGIS & PyGTK which can easily call all these modules with a more user-friendly interface, which is still a lot of work, but a priori without big pitfalls. In this perspective, in addition to the PyQGIS hypothesis, specifications are to be considered for choices of industrial development:

- QGIS plugin with PyQgis,
- Dedicated software produced under Qt ( license difficulties),
- specialized software produced under GTK (Free license),
- online service under remote server, this last aspect has been the subject of much investigation, not addressed in the article, which highlighted the problems of sustainable compatibility of libraries, the implementation container (Docker), and local repatriation of 3D visualization.

About technological matters, integration and parallelization work (Cuda and OpenMp, theadings) about executables, remain very good subjects for work-study students or engineering or Master 2 trainees, without forgetting acceleration algorithms to improve.

## 19 Conclusions

The transformation operator being defined, we manage to minimize the quantity of calculations because we apply this transformation operator to only tree location points in the area (a few hundred) against all the points of the cloud (a few tens of millions per ha).

As indicated, the processed images are normalized (the position of the pixel is linearly linked to geolocation RFG93):

1. A position point finds its position [X, Y] in the point cloud(lidar C.R.S), to which the transformation operator is applied.
2. The position Z of the terrain, at position [X, Y], is found in the corresponding cell in the DTM database.
3. The operator transforms the whole data by rotations and translations the trees source file location points in the RFG93 C.R.S.
4. The transformation (rotations and translation) is carried out in a single operation, with sufficient precision.
5. The data of the tree, its attributes in the end, are given in a format allowing exporting to PostGis / Postgres or Spatialite databases, or any type of GIS.

Most of the applications are developed in C ++ as modules on line command. Some filtering is written in C\_CUDA and the possible parallelization of the tasks will be completed. The heaviest treatments are:

1. the generation of the DTM,
2. the filtering of the horizontal tomographic sections with the extraction of the positions of tree by diameter classes.

In the current state of development, the complete process leading to the extraction of tree positions for all diameter filters (11 in total) requires about 90 minutes per ha on average, without any parallelization. The process starts from the raw \*.las file from the stand data acquisition, until the generation of the QGIS attribute layer of trees with for each (identity, geoloc. RFG93, section diameter at several given height), and the SQL query for setting base geometric data of this information.

Industrialization, will integrate these processes and proposes a sequence of tasks in a more user-friendly way will rely on PyQGIS for the man / machine interface, in particular for a direct integration into QGIS. In addition, the assumption of specialized applications under Qt or GTK interface will offer openings to other GIS

# Bibliography

**[1]** Jean-André MAGDALOU, Joseph GARRIGUE (2002)

Works 55: Forest monitoring and cartography assisted by information system geographic

Association of Friends of La Massane, Arago Laboratory - 66650 -BANYULS-SUR-MER

**[2]** Jean-André MAGDALOU, Christophe HURSON and Joseph GARRIGUE (2002)

Works 62: Riparian monitoring and cartography assisted by information system geographic

Association of Friends of La Massane, Arago Laboratory - 66650 -BANYULS-SUR-MER

**[3]** Jean-André MAGDALOU, Christophe HURSON and Joseph GARRIGUE (2005)

Work 72: Study of the effects of the 2003 drought / heat wave conjunction on the Massane forest and cartography assisted by an information system geographic

Association of Friends of La Massane, Arago Laboratory - 66650 - BANYULS-SUR-MER

**[4]** Jean-André MAGDALOU \*, Christophe HURSON \*\* and Joseph GARRIGUE (2009)

Works 80: Dynamics of dead wood and impact of floods on some areas species bordering a small Mediterranean coastal river "La Massane" Association of Friends of La Massane, Arago Laboratory - 66650 -BANYULS-SUR-MER

\* National Natural Reserve of the Forêt de la Massane (massane.rn@wanadoo.fr)

\*\* Federation of Catalan Natural Reserves (Christophe.hurson@espaces-naturels.fr)

**[5]** Stéphane PIGNON (2016)

Forêt de la Massane, LIDAR survey, technical report.

Sintégra 11 Chemin des Prés F-38241 Meylan

**[6]** Mahdi SMIDA (2018) Aggregation, Deformation and irregularities Manipulation of point clouds LIDAR. UPVD, Master II thesis, 2018.

**[7]** Yasser BOUTALEB (2018) Point cloud treatments applied in forestry and arboriculture UPVD, Master Thesis, 2018.

**[8]** HK Yuen, J. Princen, J. Illingworth and J. Kittler (1989)

A comparative study of hough transform methods for circle finding

Department of Electronics and Electrical Engineering University of Surrey,

Guildford, GU2 5XH. UK

**[9]** Güray ERUS (2008)

Recognition of cartographic objects in satellite images to high resolution

PARIS DESCARTES University University Center of Saints-Pères UFROF MATHEMATICS AND COMPUTERS

**[10]** Monica Herrero-Huerta<sup>1</sup>, Roderik Lindenbergh, Pablo Rodriguez-Gonzalvez (2018) Automatic tree parameter extraction by a Mobile LiDAR System in an urban context

1. TIDOP Research Group, Higher Polytechnic School of Avila, University of Salamanca, Avila, Spain,

2. Department of Geoscience and Remote Sensing, Delft University of Technology, Delft, The Netherlands,
3. Department of Agronomy, Purdue University, West-Lafayette, Indiana, United States of America,
4. Department of Mining Technology, Topography and Structures, Universidad de León, Ponferrada, Spa

**[11]** Anne Bienert, Hans-Gerd Maas (2009)

#### METHODS FOR THE AUTOMATIC GEOMETRIC REGISTRATION OF TERRESTRIAL LASER SCANNER POINT CLOUDS IN FOREST STANDS

Dresden University of Technology, Institute of Photogrammetry and Remote Sensing, 01062 Dresden, Germany [anne.bienert@tu-dresden.de](mailto:anne.bienert@tu-dresden.de)

**[12]** Bisheng Yang 1,2, \*, Wenxia Dai 1, \*, Zhen Dong 1 and Yang Liu 1

(2016) Automatic Forest Mapping at Individual Tree Levels from Terrestrial Laser Scanning Point Clouds with a Hierarchical Minimum Cut Method

1 State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China; [dongz-henwhu@whu.edu.cn](mailto:dongz-henwhu@whu.edu.cn) (ZD); [lyang@whu.edu.cn](mailto:lyang@whu.edu.cn) (YL) 2 Collaborative Innovation Center of Geospatial Technology, Wuhan University, Wuhan 430079, China \* Correspondence: [bshyang@whu.edu.cn](mailto:bshyang@whu.edu.cn) (BY); [daiwen-xia@whu.edu.cn](mailto:daiwen-xia@whu.edu.cn) (WD); Phone. : + 86-139-8609-9632 (BY); + 86-187-7106-1723 (WD)

Remote sensing MDPI Academic Editors: Nicolas Baghdadi, Lars T. Waser, Josef Kellndorfer and Prasad S. Thenkabail Received: January 4, 2016; Accepted: 25 April 2016; Published: May 3 2016

**[13]** Anne Bienert 1, \*, Louis Georgi 2, Matthias Kunz 2 [OrCID], Hans-Gerd Maas 1 [OrCID] and Goddert Von Oheimb 2 [OrCID] (2018)

Comparison and Combination of Mobile and Terrestrial Laser Scanning for Natural Forest Inventories

Institute of Photogrammetry and Remote Sensing, Technische Universität Dresden, D-01062 Dresden, Germany

Institute of General Ecology and Environmental Protection, Technische Universität Dresden, D-01735 Tharandt, Germany

**[14]** OGC Community Standard (2018)

LAS Specification 1.4

Copyright notice Copyright c 2018 Open Geospatial Consortium

<http://www.opengeospatial.org/legal/>.

**[15]** 2-year Post-doctoral Offer Multi-layer reconstruction of 3D point clouds:

Application to forest 3D structure and biodiversity metrics extraction factory using airborne LiDAR and multispectral images Workplace

IGN - Lab LaSTIG - Team ACTE73 avenue de Paris 94165 Saint Mandé

<https://tinyurl.com/postdoc-lidar>

**[16]** Richard Fournier (2019)

See works and bibliographical references of the applied geomatics department

Department of Applied Geomatics Faculty of Letters and Human Sciences 2500,  
boulevard de l'Université Sherbrooke (Quebec) J1K 2R1

<https://tinyurl.com/uozy4ct>