



POLITECNICO

MILANO 1863

School of Architecture Urban Planning Construction Engineering

MSc Building and Architectural Engineering

APPLICATION OF DRONE-MOUNTED VIDEO AND THERMAL CAMERAS FOR BUILDING DIGITALIZATION AND DIAGNOSTICS

Candidato:

Kevin Autelitano

Matricola:

920780

Relatore:

Prof. Ing. Enrico De Angelis

Correlatori:

Prof. Ing. Lorenzo Mario Fagiano

Prof. Ing. Marco Scaioni

Dott. Michele Bognini

Accademic Year: 2020/2021

Abstract

Compared to the past, in most parts of the world, new constructions are designed and realized with very high energy performances; moreover, plans and incentives for a deep energy retrofit of the built assets have been set, to enable local and national communities reaching the energy efficiency and GHG emission reduction targets and, as additional benefits, to enhance their health (pollution reduction) and their energy security, and to minimize life cycle costs and environmental impact in general.

In both cases, despite we live in a world where it is possible to find a great multitude of data, there is a huge need for low-cost and reliable control methodologies and tools: to understand retrofit needs, to validate design choices and to test construction results, for an effective basic and condition-based maintenance and for a continuous commissioning of operating buildings. With these aims, in the last years, the use of Unmanned Aerial Vehicles (UAV) equipped with infrared cameras has been implemented, but the use of these tools among practitioners is still at a very qualitative level.

Starting from a literature review about this topic, the thesis presents a methodology to support the following steps: (1) the flight inspection planning; (2) the RGB and infrared pictures data gathering and post processing; (3) the analysis of the data to estimate the impact of thermal bridges and the heat flux correction; (4) the analysis of the data to estimate the conduction heat loss coefficient of a building and its uncertainty; (5) the identification of maintenance needs and building pathologies.

This thesis is based on a case study developed on a building in Italy (Spino d'Adda), hosting offices, a little data centre and telecommunication devices with no-fly zones. A commercial quadcopter equipped with a standard RGB camera has been used, to collect a first set of trial pictures.

Their results have been used to set a flight planning methodology, based on Ground Sampling Distance (GSD), topographical limitations, technical characteristics of the cameras and the detailing needs of the building.

Then, the model of the building is reconstructed through Structure from Motion technique using Agisoft Metashape. A sparse point cloud representing the building is realized calibrating all the images previously collected and, with an operation of feature detection using Scale-Invariant Feature Transform algorithm (SIFT), the interesting points are found and matched solving triangulation of homologous points problem. A 3D dense reconstruction of the building is then performed.

Colour information from thermal images, which are representing the temperature captured by the thermal camera implemented on the drone, are integrated on the model by the using known points identified as baseline.

A set of points with a known position is manually identified both in a set of RGB images and in thermal ones. The distance between them is measured so that the thermal images are scaled precisely and projected on the points composing the building.

The result is a 3D point cloud model composed by a matrix that identifies, for each point, an information about its position in the space and a colour under RGB parameters coming from thermal camera.

In the end, the matrix of colour is converted into matrix of temperature and compared to a model of ideal temperature which is realized by making considerations about the package for opaque and transparent components of the envelope, air temperature registered during the phase of picture collection and internal temperature set inside the building.

This step provides a starting point for implementing an automated procedure that is able to identify from thermal anomalies a possible failure in a building envelope component and making some considerations about its possible repairing or substitution.

Sinossi

Rispetto al passato, nella maggior parte del mondo, le nuove costruzioni sono progettate e realizzate con prestazioni energetiche molto elevate. Per un profondo risanamento energetico del patrimonio costruito, nonché permettere alle comunità locali e nazionali di raggiungere gli obiettivi di efficienza energetica, sono stati stabiliti piani di riduzione delle emissioni di gas serra al fine di migliorare la salute dei cittadini (riduzione dell'inquinamento) e la loro sicurezza energetica, nonché per minimizzare i costi del ciclo di vita e l'impatto ambientale in generale.

In entrambi i casi, c'è un enorme bisogno di metodologie e strumenti di controllo affidabili e a basso costo per capire le esigenze di retrofit, convalidare le scelte progettuali e testare i risultati di costruzione, per un'efficace manutenzione di base e basata sulle condizioni e per un continuo commissioning degli edifici operativi. Con questi obiettivi, negli ultimi anni, è stato implementato l'uso di Unmanned Aerial Vehicles (UAV) dotati di telecamere a infrarossi, ma l'uso di questi strumenti tra i professionisti è ancora a un livello molto qualitativo.

Partendo da una revisione della letteratura su questo argomento, la tesi presenta una metodologia per supportare le seguenti fasi: (1) la pianificazione dell'ispezione del volo; (2) la raccolta dei dati delle immagini RGB e infrarosse e il post processing; (3) l'analisi dei dati per stimare l'impatto dei ponti termici e la correzione del flusso di calore; (4) l'analisi dei dati per stimare il coefficiente di perdita di calore per conduzione di un edificio; (5) l'identificazione delle necessità di manutenzione e delle patologie dell'edificio.

Questa tesi si basa su un caso di studio sviluppato su un edificio in Italia (Spino d'Adda), che ospita uffici, un piccolo data center e dispositivi di telecomunicazione e no-fly zones. È stato utilizzato un drone commerciale dotato di una telecamera RGB standard, per raccogliere una prima serie di immagini di prova.

I loro risultati sono stati utilizzati per impostare una metodologia di pianificazione del volo, basata sulla distanza di campionamento a terra (GSD), i limiti topografici, le caratteristiche tecniche delle telecamere e le esigenze di dettaglio dell'edificio.

Poi, il modello dell'edificio viene ricostruito attraverso la tecnica Structure from Motion utilizzando Agisoft Metashape. Una nuvola di punti sparsa che rappresenta l'edificio viene realizzata calibrando tutte le immagini precedentemente raccolte e, con un'operazione di feature detection per mezzo dell'algoritmo Scale-Invariant Feature Transform (SIFT), i punti interessanti vengono trovati e abbinati risolvendo il problema della triangolazione dei punti omologhi. Viene quindi eseguita una ricostruzione 3D densa dell'edificio.

Le informazioni di colore dalle immagini termiche, che rappresentano la temperatura catturata dalla telecamera termica implementata sul drone, sono integrate sul modello utilizzando punti noti identificati come linea di base.

Un insieme di punti con una posizione nota è identificato manualmente sia in un insieme di immagini RGB che in quelle termiche. La distanza tra loro viene misurata in modo che le immagini termiche siano scalate con precisione e proiettate sui punti che compongono l'edificio.

Il risultato è un modello di nuvola di punti 3D composto da una matrice che identifica, per ogni punto, un'informazione sulla sua posizione nello spazio e un colore sotto i parametri RGB provenienti dalla telecamera termica.

Alla fine, la matrice di colore viene convertita in matrice di temperatura e confrontata con un modello di temperatura ideale che viene realizzato facendo considerazioni sul pacchetto per i componenti opachi e trasparenti dell'involucro, la temperatura dell'aria registrata durante la fase di raccolta delle immagini e la temperatura interna impostata all'interno dell'edificio.

Questa fase fornisce un punto di partenza per l'implementazione di una procedura automatizzata che sia in grado di identificare dalle anomalie termiche un possibile guasto in un componente dell'involucro dell'edificio e fare alcune considerazioni sulla sua eventuale riparazione o sostituzione.

Index

1 - Introduction	1
2- Work and case study presentation	6
2.1 - <i>Main topic: PhD Thesis and MSc thesis contribution</i>	6
2.2 - <i>Case study: complex in Spino d'Adda (Italy)</i>	9
3- Review about the drones used	11
3.1 - <i>Review about drones: characteristics present on the market</i>	11
3.2 - <i>Case study: drone used for the test</i>	15
4- Stage 1: flight path planning	17
4.1 - <i>Literature review on flight path</i>	17
4.2 - <i>Case study: flight path planning description</i>	20
5 - Stage 2: 3D model reconstruction	24
5.1 - <i>State of the art</i>	24
5.2 - <i>Structure from Motion: how it works and examples</i>	25
5.3 - <i>SIFT</i>	38
5.4 - <i>Thermography and its integration into a 3D environment</i>	44
5.5 - <i>Case study: Structure from Motion application using Agisoft Metashape</i>	49
6- Stage 3: Façade modelling	60
6.1 - <i>Literature review: k-means</i>	60
6.2 - <i>K-Means: application and discussion</i>	64
6.3 - <i>Literature review: MSAC</i>	72
6.4 - <i>Case study: MSAC and its application on Matlab</i>	74
7 - Stage 4: data analysis	80
7.1 - <i>Goal and strategy</i>	80
7.2 - <i>Reconstruction of the external temperature from analysis</i>	81
7.3 - <i>Reconstruction of the external surface temperature of control</i>	84
7.4 - <i>Analysis of the result</i>	94
8-Future update: real-time model and process automation	106
8.1 - <i>Implementation of SLAM for 3D reconstruction</i>	106
8.2 - <i>Real-time model generation for time scheduling reduction</i>	112
Bibliography	115
Acknowledgement	118

1 - Introduction

For satisfying its needs, mankind spontaneously tends to grow up by looking at many aspects: from a theoretical point of view to a technological one. Since its knowledge and its experience is increasing, the world around becomes more and more complex, while its requirements are becoming more ambitious and specific.

This natural tendency took the history of the industrial field to be divided into three periods.

The First Industrial Revolution, around the '700, introduced the concept of mechanical industry, while discoveries and inventions occurred during the Second Industrial Revolution (1870) showed in the period of mass consumes and modern factories.

In the second half of the last century, the Cold War, with the competition between USA and URSS, was one of the causes of the Third Industrial Revolution. In this period, computer and automation industry were developed.

Nowadays we are looking at the development of what it was named "Industry 4.0" or "Fourth Industrial Revolution" in order to identify an era mainly trained by two concepts: Data Management, sharing and interconnection of information, and Cybernetic Industry, that is developing the previous concept of automation in order to produce everyday more modern, artificially intelligent and complex cybernetic devices.

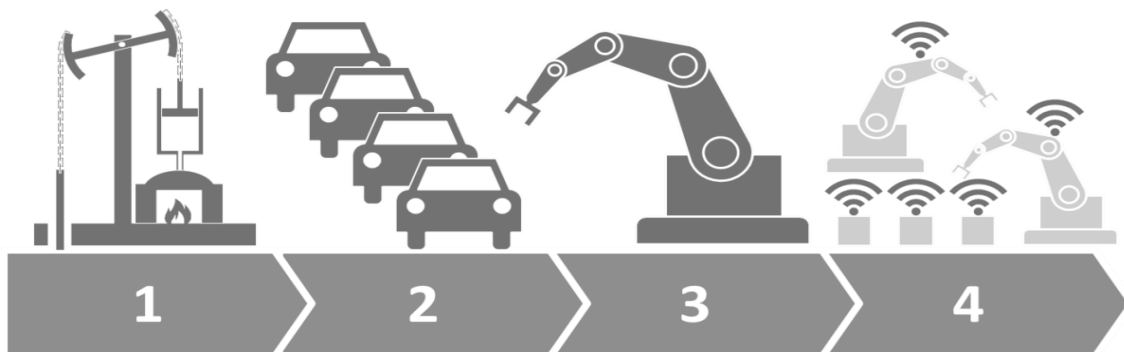


Figure 1- Industrial Revolutions in the history (Source: <https://www.allaboutlean.com/industry-4-0>)

Main topic of industry 4.0	
Interconnection	Collaborative, interconnected and programmable rovers
Addictive manufacture	3D printers connected with updating software
Augmented reality	Support to the production
Simulation	Simulation between interconnected machines
Internet of the things	Connection between products and processes
Cloud computing	Management information on opened systems
Cyber Security	Analysis and protection during the phases
Big data and analysis	Data analysis for product optimization

Table 1- Main topic of Industry 4.0 (Source: <http://www.ingfrancescopavan.it/industria-4.0.html>)

Consequently, robotics and artificial intelligence are used to substitute man's effort working in more precise and safer conditions.

This methodology is affecting several industrial and commercial fields, with military and space exploring fields as pioneers. From an historical point of view, one of the first UAV (Unmanned Aerial Vehicle) consisted of a group of air balloon full of explosive and without pilot used by Austrians to defeat Venice during the battle of 22nd august 1849.

Moreover, there are some reported cases in which the use of prototyped drones for inspection purposes were used in France at the end of the '700. (Robison, 2017).

Since this date, moved by historical reasons, the use of robotics in military industry has been growing and today national armies have many complex technologies available. Some examples could be tEODor, a remote-controlled rover used for defusing potential explosive devices providing also preliminary analysis of that, and MQ-9A Predator B, a UAV used for reconnaissance, surveillance and target acquisition operations.

Automation and cybernetic programs have also monopolized the space exploration field. Since the second half of last century, most of the missions were realized only by autonomous devices which were able to transfer data to a remote central directly placed on Earth or perform simple maintenance operations on space stations.

At the end of the century, researchers focused their efforts in implementing automation and its benefits also to civil field, creating a new business.

Different sectors have implemented the use of autonomous devices in their processes, like the automotive industry, and improved their business in terms of economic and technological growth.

Since many factors act in a work process influenced by a general concept of digitalization, the automation of different sectors is influenced by this phenomenon.

The gradual development and introduction of autonomous devices for civil operations has been studied also considering economic aspects, looking at their potential benefits for creating a new business. As a support, it is provided the chart below that shows the productivity growth trend of the most important economic areas related by the percentage of digitalized and automated work processes.

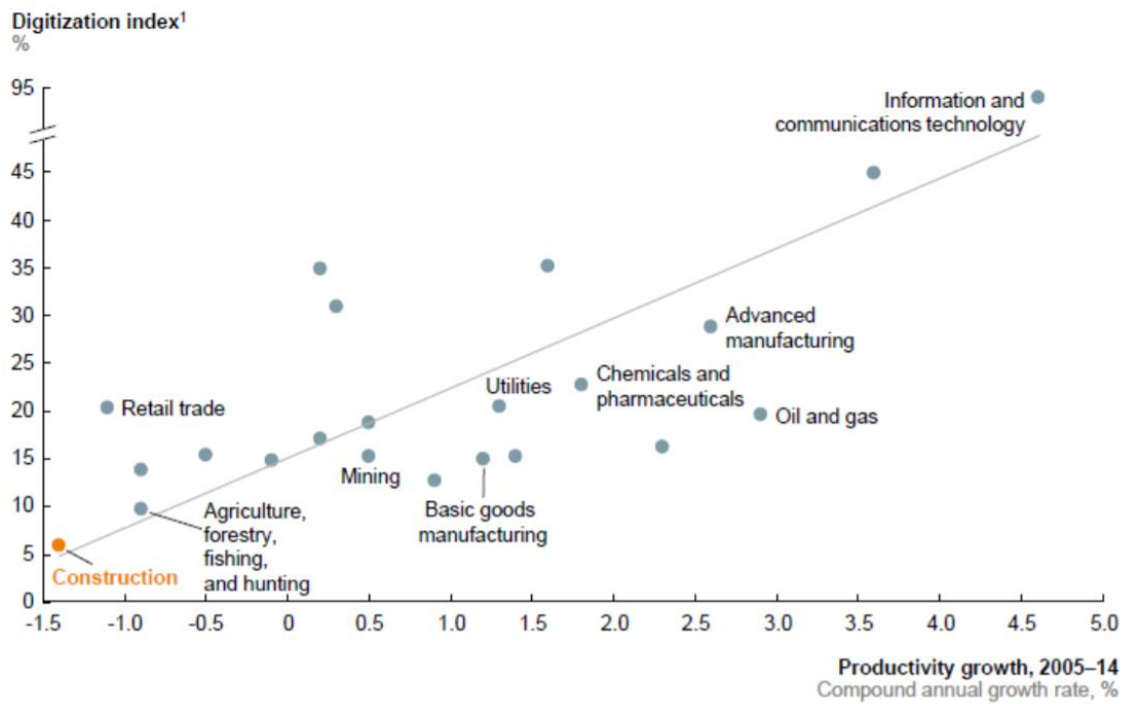


Figure 2 - Chart about productivity growth and digitalization index (Source: Building Process and Information Management by Prof. Francesco Pavan)

It is possible to notice that areas with higher percentage of digitalization index provided, during the period between 2005-2014, an increasing in their business. On the contrary, sectors like construction, that are represented by 6% of digitalization index only, have been invested by a reduction in terms of productivity growth. However, the economic crisis of 2008 did not help the situation.

If the situation looks dramatic, it is understandable how investing in digitalization, and so construction automation, could be interesting from a technological and economic point of view because of the unsaturated market.

Recently, the use of UAV (Unmanned Autonomous Vehicle) is becoming more and more frequent for operations that require the inspection of a civil construction or an infrastructure for the advantages that this technology can give.

A vehicle with a camera, RGB or infrared, and remotely-controlled by a licenced operator, can monitor the state of a civil construction, focusing the attention on potential defects without risks and using only non-destructive procedures.

Those kinds of operations are normally executed manually by a technician and all the considerations are obtained based on visual inspections, that considers the presence of a potential problem in a building after the failure has already occurred. A slow and risky operation could be the consequence of an increasing in maintenance costs.

In the following info-graphic it is shown how monitoring operations using drones are performed.

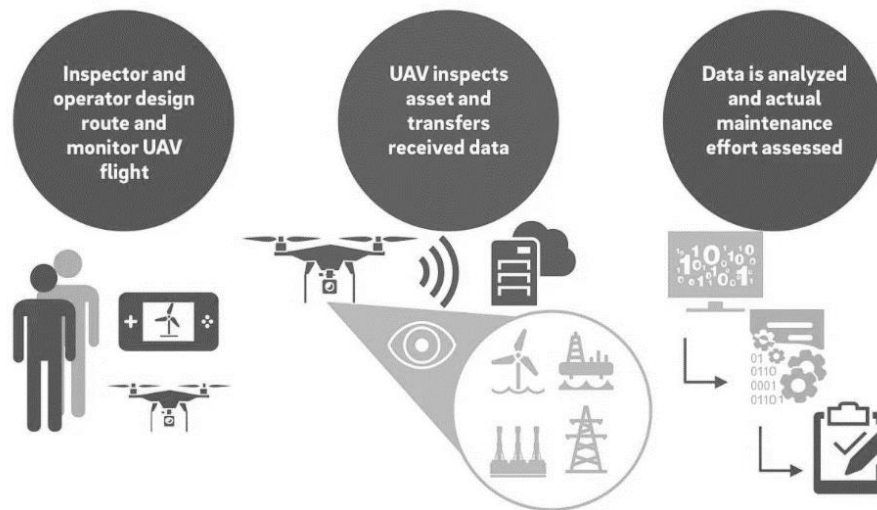


Figure 3 - Main steps for drone-based monitoring (Source: Weichenhain, Uwe in <https://www.rolandberger.com/it/Insights/Publications/Drones-The-future-of-asset-inspection.html>)

In a paper published by Roland Berger, Weichenhain presents the drone technology as the potential new standard in asset inspections, providing the results obtained in the oil and gas platform industry, where times required for inspection operations were reduced from 8 weeks to 5 days in average (Weichenhain, 2019).

In addition, since the use of drones was introduced recently, it is declared that it is not possible to estimate precisely how the decreasing in maintenance costs is, but the first projections are optimistic.

In fact, an estimated cost saving using drone-based assets was predicted in a range between 50% and 90% and the monitoring is going to be extended to other sectors, as it is shown in the chart below.

As it was said before, the growing trend is confirmed without the same velocity of other sectors, also for the construction sector (Weichenhain, 2019).

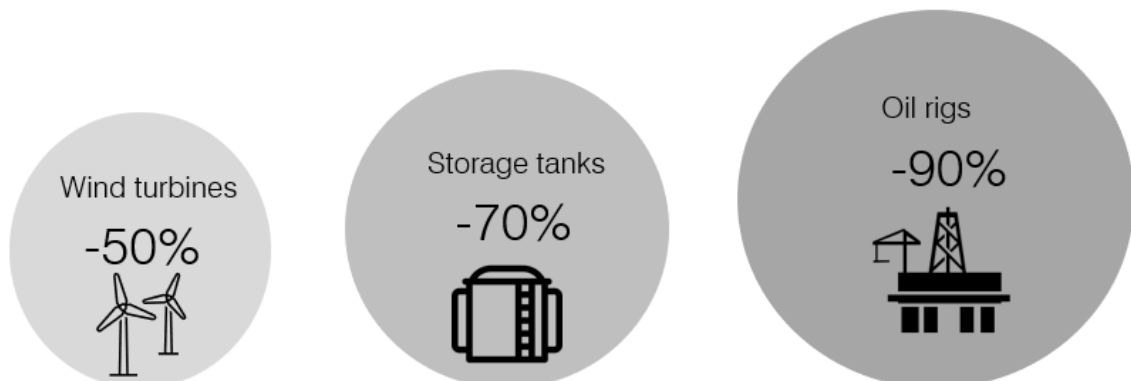


Figure 4- Economic savings in maintenance using drone-based monitoring

In conclusion, if only the construction sector is considered, it is important to notice where efforts to introduce automation inside construction places has been done.

Starting from the second half of 80s, especially in Japan, due to demographic factors, and the USA, where it was increasingly necessary to provide safer conditions to workers, robot was introduced. The development continued along the following decades.

According to a report provided by CNR in 2016, in Italy construction automation showed poor growth due to the setting of the economy mainly composed by Small to Medium-sized Enterprises with a low capitalization, low capability of investment and a natural tendency of surviving without risks.

The presence of a huge historical building heritage and the starting costs required by robotics are another cause of the contained growth of automation construction, mainly limited to supporting and monitoring operations (Canapini, Malaguti, & Ricchetti, 2010).

2- Work and case study presentation

2.1 - Main topic: PhD Thesis and MSc thesis contribution

The Thesis that is going to be presented is part of a main PhD research which is under development by DEIB and DABC of Politecnico di Milano and it is focused on the application of autonomous drone technology for building inspection.

As mentioned before, since a human inspection is considered risky, time consuming and expensive, the research aims to lay the groundwork for designing in future a fully autonomous hybrid multi-copter network which is able to map an area and planning periodic inspection of a building by acquiring and reading multi-spectral data from a variety of sensors such as Lidar and thermal sensors.

In order to reach the goal, the challenge to involve is realizing a robust system which is able to control simultaneously different actors taking into account a complex planning of flight path due to physical constraints.

The PhD Thesis is involved on exploring the field of tethered drones, which is actually, from a research point of view, a topic under open discussion.

The use of cable-connected drones would solve the problem of the autonomy of the system, performing even time-consuming inspections. On the other hand, automatically guiding the system through the space to a specific position avoiding obstacles is a challenging task.



Figure 5 - Cable connected drones and their movement (Source: Bolognini, Michele - PhD Thesis)

A further development of the topic is related to the optimization of the process in terms of time.

Even if working with cable-connected drones avoids the limitation caused by the autonomy of the battery, the less is the time required by the mission the low is the possibility of failure and reach out problems.

Finding the best path in the least time possible is the problem to solve.

For this, working with the approach of the Travelling Salesman Problem on a sample-space composed by 6 potential obstacles zones, a test has been performed. The solution shows that the approach adopted is independent of the number of zones.

In addition, a test about 9000 points comparing the application of a simple TSP approach with a hierarchical one was tested. Considering an optimal flight time of 15'06'' for the first and 15'36'' for the second approach, the application of a hierarchical TSP for the guide of the drones guarantees a total coverage of the selected points in 6.29 minutes, resulting a rising and interesting solution for movement topics.

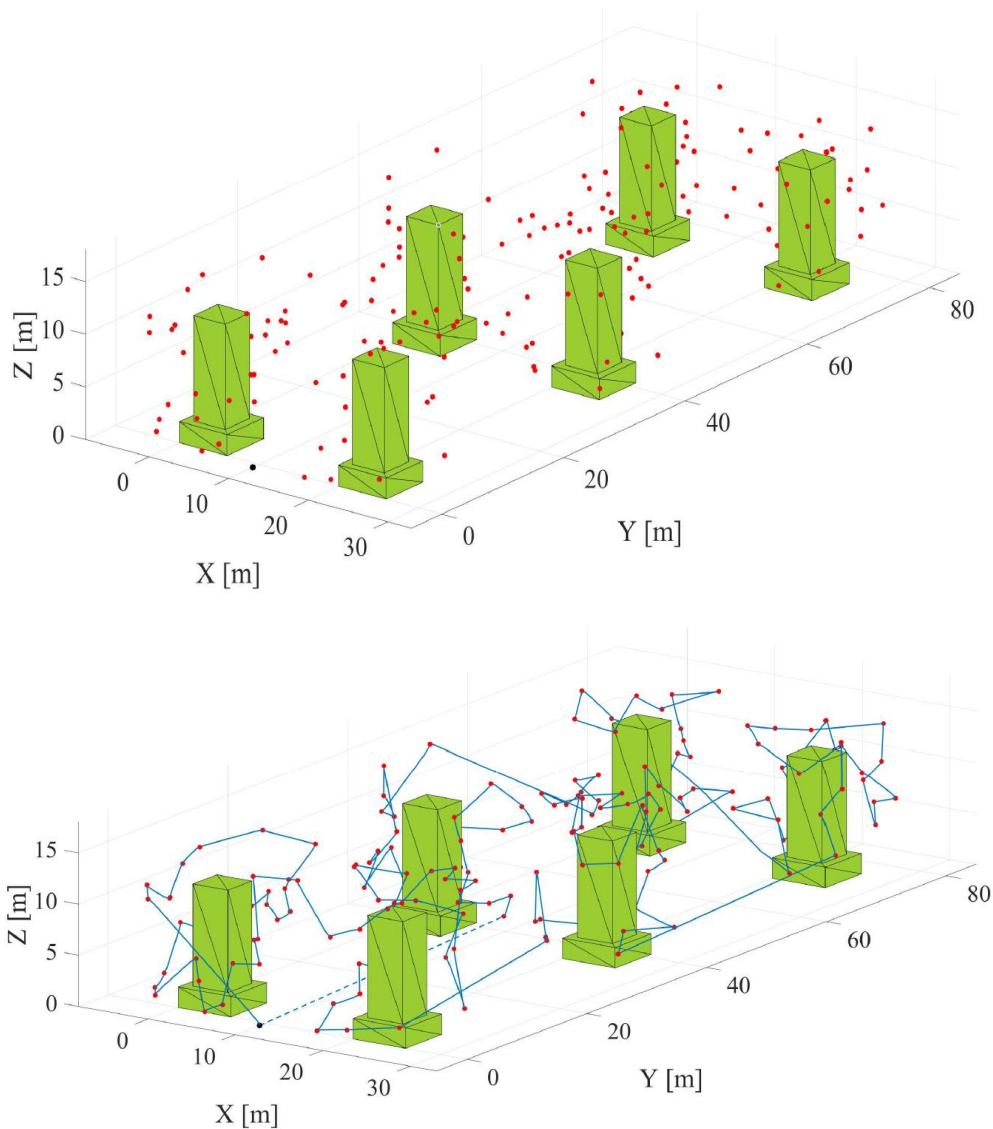


Figure 6 - Optimized movement through TSP approach (Source: Bolognini, Michele – PhD Thesis)

The second part of the PhD thesis is focused on the application on civil constructions in terms of surface defects detection.

To tackle the issue, a proper fusion of RGB and thermal sensors data is studied and developed.

This MSc Thesis aims to be a support on this last aspect, developing the conditions for the part concerning the reconstruction of the building in an IT environment and let the drone be able to read autonomously the results obtained.

The entire work has been divided into 4 main stages. Within the Stage 1, an analysis of flight paths optimization for collecting best data is performed considering the intrinsic characteristics of the UAV and the camera installed.

In Stage 2 the basis for the 3D reconstruction is studied analysing two of the main image-based computer vision techniques and applying the one that was considered the most adaptive for the purpose (structure from motion).

Some additional considerations about all the threads involved in the algorithm are also presented.

The stage is concluded applying an integrative technique able to adapt thermal sensors information in the proper part of the building by scaling them using some known points as baseline.

The result is a point cloud composed by those groups of information:

- Location (3 components that identifies the position of the point in the space);
- RGB-colour (3 components that represent the RGB parameters of the measured colour);
- IR-colour (3 components in RGB parameters that represent the colour from IR-sensors properly scaled).

Within stage 3, the point cloud, through M-Estimator Sample Consensus, is divided for modelling the facades composing the building. The identification of vertical and horizontal façade, in particular between opaque and transparent parts, is necessary for writing a proper thermal balance based on the characteristics adapted for that part.

This thread is propaedeutic to the last one: in fact, IR-colour information is collected in 10x10 cm areas and converted into a value of temperature.

At the end, localized variations of different areas are studied through proper algorithms.

Values measured are compared with the ones calculated considering a heat balance problem solved at the same climatic and geometrical conditions, so as the variation between the two samples can be only caused by failures into the stratigraphy of the envelope.

In the following scheme, all the passages are summarized.

It is important to mention that the main goal of this Thesis is to look at a multidisciplinary approach and try to implement concepts and applications of the energy modelling field into a different sector.

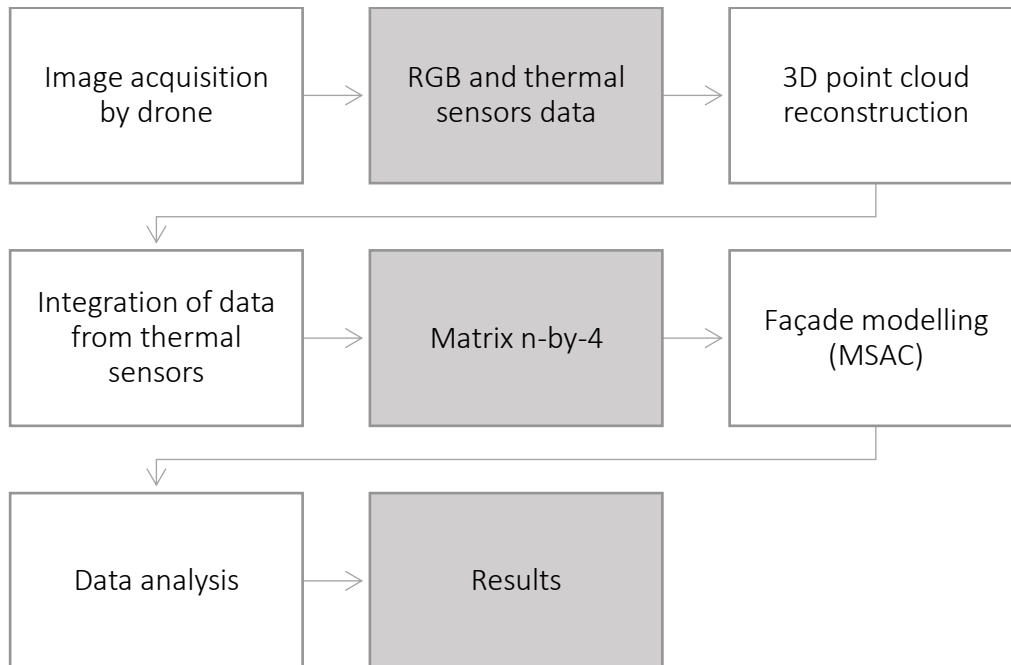


Figure 7- MSc contributions

2.2 - Case study: complex in Spino d'Adda (Italy)

In order to apply and verify all the threads involved, a complex based in Spino d'Adda was selected as case study.

The area, located outside the town, on the main street that connects Spino d'Adda and Rivolta d'Adda, is about 3000 mq and it is used mainly as office by the telecommunication department of Politecnico di Milano for research purposes.

The entire complex is composed by 4 buildings, which are used by the users in a non-continuous way.

The main one, called Block A, characterized by a squared plant of 8.5 m and three floors, is connected to Block B by a covered passage, characterized by regular plant and only one floor.

Both parts are composed by a system of reinforced concrete columns and beams, while slabs, included the roofs, are realized with traditional techniques.

Considering Block A, a system of 100x140 openable aluminium windows on every façade is installed.

In the area two more buildings are present: Block C and Block D. Due to their low dimensions compared to Block A and their destination of use, it was decided not to consider them in the tests.

So, only Block A and B are considered in the reconstruction of the model as interesting parts, while the others are just mentioned and taken into account as obstruction for the flight planning phase.

It is also important to take into account the presence of 6 antennas, 5 on ground and 1 on the top of Block A actually turned off, which are property of ASI (Agenzia Spaziale Italiana) that are constantly acquiring and registering data for research purposes.

The presence of these devices influenced fly path decisions: in fact, in order not to disturb their work, they had to be treated not only as simple obstacles during flight path planning. For this aspect, a no-fly zone had been arranged in advance.

In general, the site has been chosen between all the buildings assigned to Politecnico di Milano because there is a higher possibility of flying respecting the restrictions of the national code delivered by ENAC and the absence of frequented public spaces that could have been probable obstacles.

Despite of the obstacles given by adjacent buildings and the presence of antennas, Block A and the site was particularly suitable for the tests.

In the following pictures, it is shown the site and its integration with the context.



Figure 8 - Case study in Spino d'Adda

3- Review about the drones used

3.1 - Review about drones: characteristics present on the market

Before focusing the attention on the main topic of the research, it is important to illustrate a brief summary about the different type of drones that are present on the market.

UAV characteristics are suitable for the different needs required that make them optimal for specific operations. Thanks to that, it is possible to make some distinctions between drones according to their mechanical characteristics.

First of all, drones can be classified looking the way how the wings move: fixed or rotative.

The first one is composed by a rigid wing and equipped with an electric or combustion engine that is able to let the system having the proper thrust to fly.

The flight control is possible thanks to the presence of:

- Ailerons: on the bottom and on the wings, avoiding rotations caused by rolling;
- Elevator, that is controlling the pitch, the rotation of the front respect of the axis formed by the wings;
- rudder, that is controlling the yaw, specifically the rotation around the vertical axis.

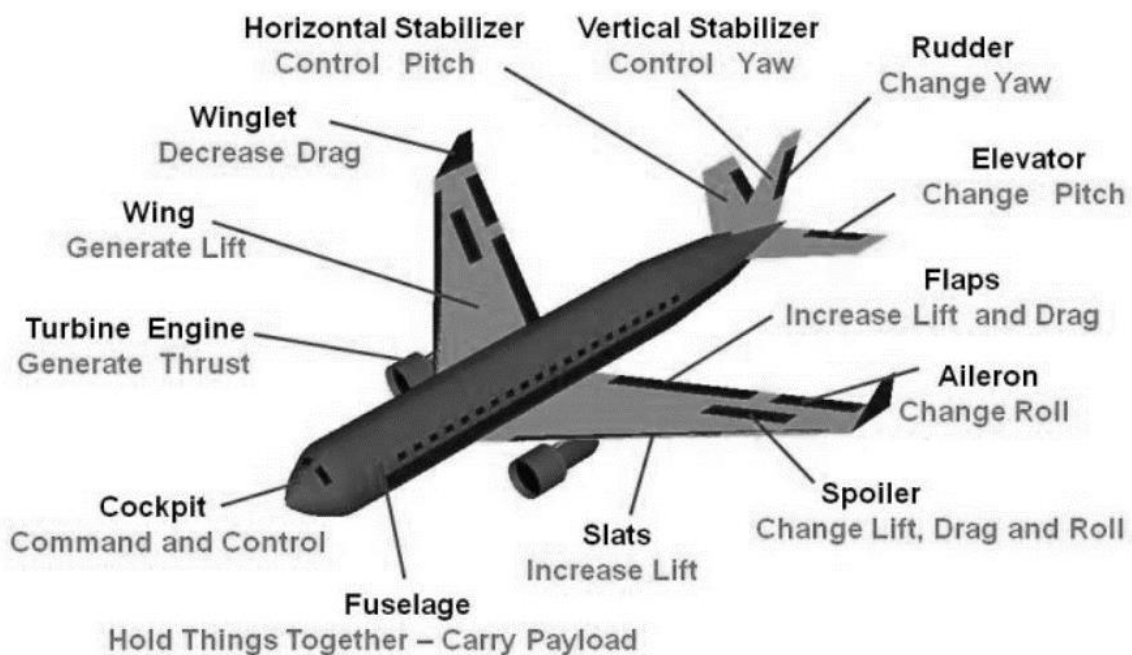


Figure 9 – Airplane parts and functions (Source: NASA - <https://www.grc.nasa.gov/WWW/k-12/airplane/airplane.html>)

Moreover, the aerodynamic of the system is able to stand high velocity wind up to 45 km/h. When a rotor gets damaged, due to its conformation, a fixed wing UAV can land safely reducing the impact on the soil.

3- REVIEW ABOUT THE DRONES USED

UAS at fixed wings are characterized by a high flight autonomy, in average around 45 minutes, and they can cover high areas, around 10 kmq, with one operation, collecting reliable pictures of the site.

On the contrary, fixed wing UAVs are characterized by higher costs and lower applications in case of building inspections due to their capacity of movement.

In fact, this system needs to move air for its displacement. Due to its conformation, the hovering is impossible.

Another issue is related to the place required for taking on and landing operations, in which a specific track is always necessary.

Actually, it is under development a fixed wing UAV model capable of working in VTOL (Vertical Take Off and Landing) and STOL (Short Take Off and Landing) but complexity and cost are still prohibitive.



Figure 10 - Examples of fixed wings drone (Source: Parrot – www.datamanager.it)

For rotative wing drones, stability and movement are obtained with couples of coaxial contra-rotating propellers, instead of two different rotors, that guarantee a relatively simple driving of the system also where the meteorological conditions are dangerous.

For this reason, the possibility of hovering is the main advantage of a rotative wing drone and this condition can help the users to perform analysis also where a stationary position of the camera is requested, for example for buildings or punctual infrastructures.

Rotative-wing drones can be differentiated within sub-groups depending on the number of propellers. The most diffused are the quadcopters, the exa-copters and the octocopters, with 4, 6 or 8 propellers each.

The differences between the three systems are in movements and autonomy. At the same conditions, a quadcopter is slower than an octocopter due to the lower number of propellers, but it can have higher autonomy because the system requires less power for the propellers.

In addition, in case of the failure of one of the propellers during flying operations the lack of propulsion is counterbalanced by the other rotors.

3- REVIEW ABOUT THE DRONES USED

For this reason, in an octocopter the incidence of repartition due to the failure of a propeller is about 12.5% so as the supplier redundancy requested to all the others increases up to 1.79%.

In a quadcopter, the same failure affects the flight about 25%, with an increasing in the supplier redundancy up to 6.25%.

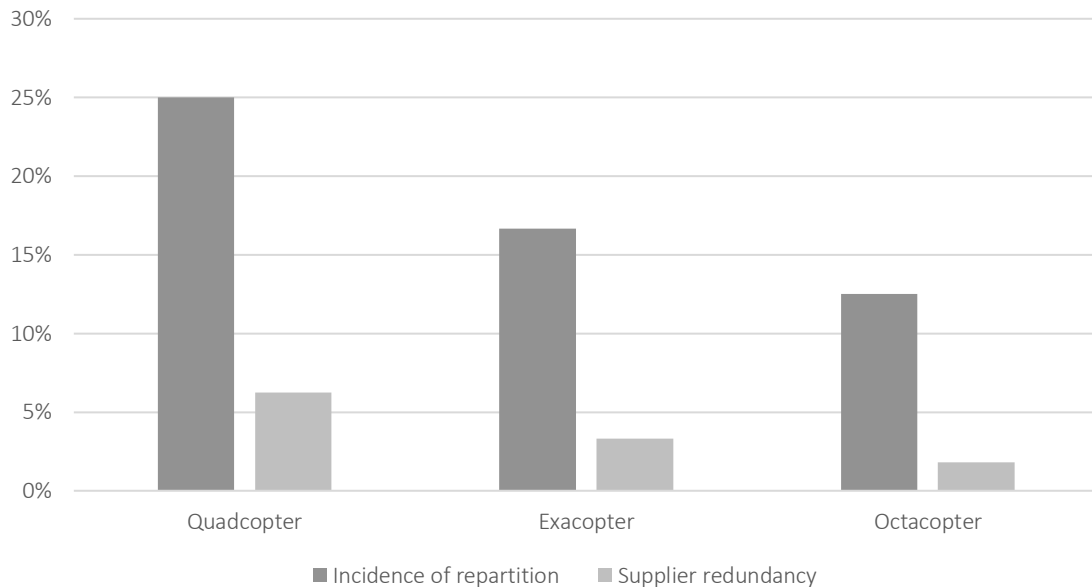


Chart 1 - Comparison between different drones in relation to the number of the blades

For a rotary-wing UAV, due to their variability, making comparisons about speed, maximum high reachable by a drone and autonomy of the battery requires to consider several aspects such as the weight of the system and of the instruments carried on the drone.

The maximum load capacity of a drone is affected by several aspects like the number of rotors or the length of a singular propeller. The power and the type of the engine installed plays an important role in this aspect.

Another important key is the condition of the site where drones are tested by the producers, the standard conditions applied and declared. For this reason, on the market they are present quadcopters which are able to sustain higher loads than an octocopter.

The same behaviour is encountered also for speed and autonomy.

In order to give an analytical and reliable comparison between different rotative-wing drones, a market research is obtained in order to find testers with the same boundary conditions that could have figured out reliable results for a comment.

With the support of an online market research instrument Aeroexpo, it was possible to focus the research on professional rotative-wing drones with 4, 6 or 8 rotors and an electric engine.

3- REVIEW ABOUT THE DRONES USED

Then, in order to narrow down the list obtained, a model realized by one producer and so tested in the same standard conditions is found. The purpose was to compare three different models with the same propulsion (it only changes the number of propulsors installed) and type of battery in terms of autonomy.

Finally, the producer was found, Challengers Aerospace systems and the models come from Hercules Series, and the results were compared (ChallengeAerospace, 2020).

Data input	H4 (quadcopter)	H6 (Exacopter)	H8 (Octocopter)
Empty weight	10.12 Kg	12.15 Kg	16.15 Kg
Max Take-off weight (MTOW)	21.46 Kg	30.29 Kg	52.43 Kg
Propulsion	4xCA-KU 200	6xCA-KU 200	8xCA-KU 200
Battery	2x2200 mAh	2x2200 mAh	4x2200 mAh

Performance	H4 (quadcopter)	H6 (Exacopter)	H8 (Octocopter)
Flight autonomy	60 min	75 min	120 min
Cruise speed	64.8 Km/h	72 Km/h	86.4 Km/h
Payload	11.34 Kg	18.14 Kg	36.3 Kg

Table 2 - Technical characteristics of the analysed drones

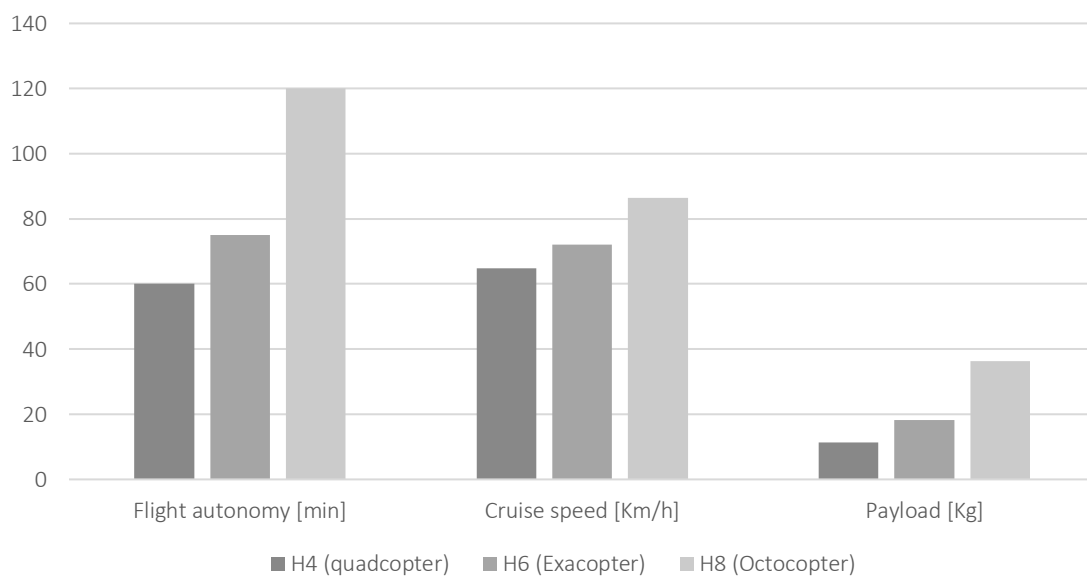


Chart 2 - Comparison between drones H4 H6 H8

In terms of performances, an octocopter is able to provide better results in speed and payload. Also, the autonomy is higher than exa-rotors, but in this case a most performed battery is installed.

The main disadvantage is the cost of the system, higher than the other two.

In conclusion, it is possible to affirm that each type of drone is suitable for particular operations. However, another main distinction between customized and commercial drones could be considered.

The first category mentioned tends to include drones which are realized in lab by assembling different parts, so as costs and stability can be variable for the driver.

The second category is the most accessible to the stakeholders, both amateurs and professionals, and the costs are known and fixed. In addition, since the product can be sold to everyone, stability in flight is tested by the engineers during previous phases in order to avoid possible accidents.

3.2 - Case study: drone used for the test

For linking the PhD and MSc Thesis, a commercial quadcopter by DJI and named DJI Mavic 2 Enterprise Dual is selected for the phase of data acquisition.

As mentioned before, Mavic 2 Enterprise Dual is a rotational wings quadcopter that is particularly suitable for many operations due to the possibility of installing several additional components.

The drone is characterized by a 743 g of payload, included propellers and gimbal, that is the load capability that an aircraft can lift without compromising stability.

Since it is a rotational wings drone, Mavic 2 can perform all its movement in all the directions and remain in hovering, particularly suitable for taking pictures. Under absence of wind, the drone can reach a maximum velocity of 65 km/h while the autonomy, under favourable conditions, is around 25 minutes.

This characteristic is extremely important because it requires that all the operations have to be planned considering the battery capability both of the drone and of the radio control. So, for long operations it is important to have more than one battery and manage their recharge.

The camera is installed on a 3-axis gimbal, so as the movements of roll, pitch and yaw are compensated and do not affect the precision of the picture. The device can control and stabilize all the inclinations between -90° and 30° , included horizontal and vertical roll between 0° and 90° .

Mavic Pro is equipped with an RGB camera M2E of 1/2.3" CMOS (complementary metal-oxide semiconductor), characterized by 12.35 Megapixels as sensor and a focal lens of 26 mm with a distortion lower than 1.5%.

3- REVIEW ABOUT THE DRONES USED

Opening of the focal lens is assumed to be between 24 mm and 48 mm, while pixel format of the images in 4:3 option is equal to 4056x3040 pixels.

The thermal camera of Mavic 2 Enterprise used is part of the optional included in the drone and it is interchangeable.

For obtaining a good result and avoid the influence of the external temperature or wind velocity, an uncooled VOx microbolometer is installed.

Due to its characteristics, the thermal camera can collect false colour pictures working between 800 nm and 1400 nm spectral band, catching the heat coming from the object. The dimension of the images is 4:3 is 640x480 pixels.

The drone is also implemented with proper sensors for obstacles detection in all the four sides. Characteristics are variables, passing from detection capabilities of the front and rear sensors which are able to detect an obstacle until 40 m in a fast way, while lateral ones reach 22 m detection.



Figure 11 - Photo of Mavic 2 Enterprise Dual (Source: DJI - <https://www.dji.com/it/mavic-2-enterprise>)

4- Stage 1: flight path planning

4.1 - Literature review on flight path

In this chapter, a general presentation of the basic and general rules for planning a flight path is presented. Since this Thesis work is focused on the fusion and analysis of data collected from pictures, the topic about the autonomous flight of a drone is secondary, even if it is studied by the main PhD research.

For this reason, the flight path studied and planned actually considers the presence of a user that pilot the drone according to national standards and all the possible obstacles in the aerial space.

This is an important operation that must regard several factors, from the geometric elements of the building analysed to the national, regional and local regulations.

An approach consists in realizing the flight path considering the intrinsic characteristics of the drone used, in particular taking into account the resolution of the camera installed and other limits of the device such as the resolution of the image, focal length depth and the objective of the camera.

For a complete data acquisition, a flight path able to cover all the building is required. Despite the research, where it is possible to find papers presenting special algorithms as universal solutions for finding the best flight planning, it is not possible to apply them completely to all the case studies because of the presence of many variables.

On the market many app launched with the scope of calculating the best trajectory are only limited at cases where the object to be focused is completely in plane.

However, when a 3D object like a building has to be reconstructed, the use of the devices described before is not so effective.

In any case, some good rules can be extracted in order to perform a good flight and obtain a good result.

The most diffused and efficient when a study of a single plane part is requested is the strip-method, where the drone is conducted by the pilot in a zig-zag movement, either vertical or horizontal, in relation of the façade (Rakha & Gorodetsky, Review of Unmanned Aerial System (UAS) applications in the built environment: Towards automated building inspection procedures using drones, 2018).

In particular, the first one is indicated for collecting information by buildings where high is much more developed than width and vice versa for an horizontal flight.

It is also demonstrated that a vertical flight tends to reduce the clarity and the quality of the images due to the unfavourable movement of the lens, while the strip horizontal method is most efficient when the speed of the drone is low.

4- STAGE 1: FLIGHT PATH PLANNING

Another type of flight path is the spiral movement, that forecasts spiral movements around the building. It is particularly indicated for preliminary recognitions and 3D model reconstruction.

Both the models consider all the device limitation and geometrical obstructions given by the external context.

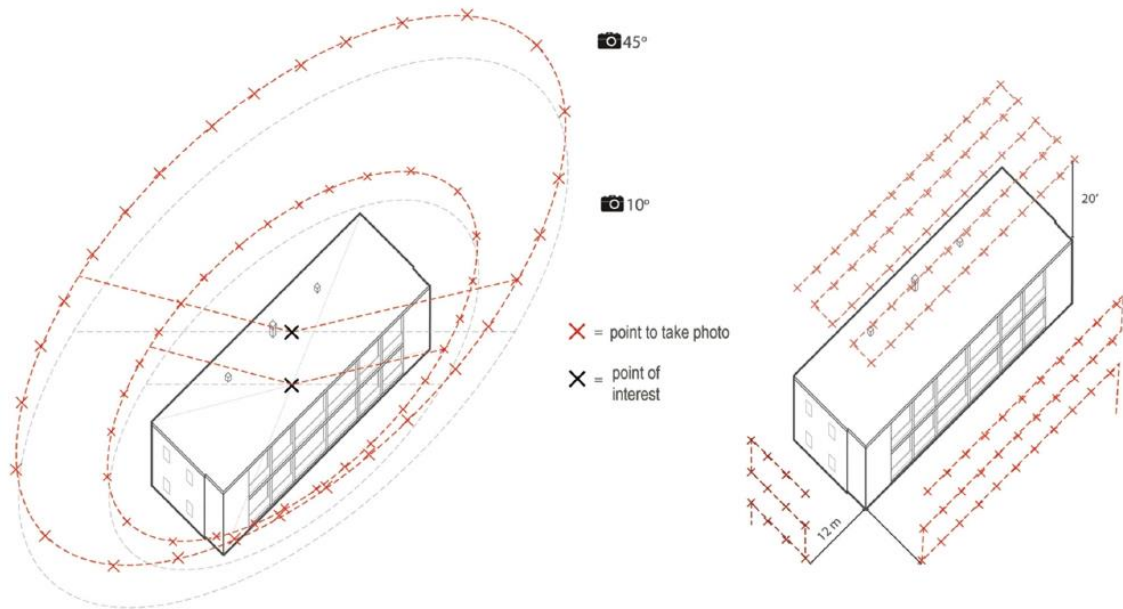


Figure 12 - Comparison between the two main flight procedures (Source: Rakha et Al. (2019) - Review of Unmanned Aerial System (UAS) applications in the built environment: Towards automated building inspection procedures using drones)

The inclination of the thermal camera during flight operations is another important aspect to be taken into account because, since all the system works analysing irradiation emissions of the objects, this output varies by changing the position of the drone (Rakha, Gorodetsky, & Kakillioglu, Campus as a Lab for Computer Vision-based Heat Mapping Drones: A Case Study for Multiple Building Envelope Inspection using Unmanned Aerial Systems (UAS), 2018).

For the acquisition of thermal images, it is important to regard at the position of the drone in terms of angle formed with the object. In fact, by an analysis of a photovoltaic field performed by positioning the drone in different places, different results up to 30% between them are obtained (Buerop-Lutz & H., 2015).

4- STAGE 1: FLIGHT PATH PLANNING

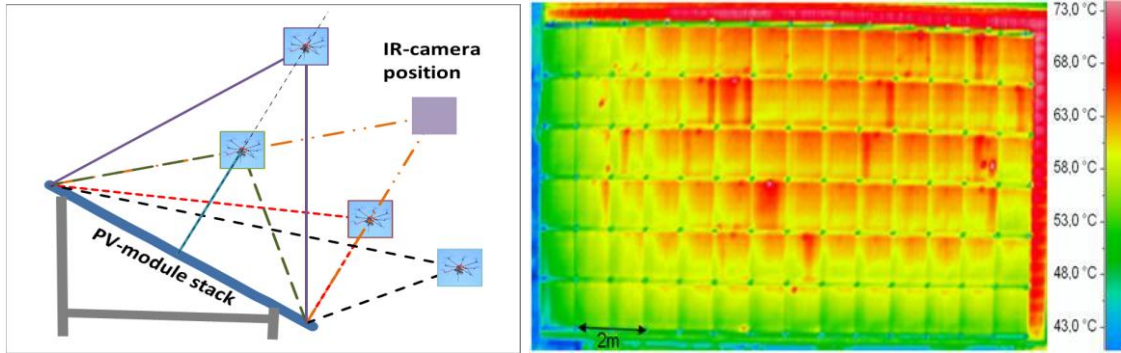


Figure 13 - Test on thermal results in relation of drone positions (Source: Buerop-Lutz H. - inspecting pv-plants using aerial, drone-mounted infrared thermography system)

Finally, forecast weather and adapting the operations in relation to them is important for conducting everything in safe conditions.

Wind speed and optimal visual has to be considered for conducting all the operations in a safe way.

If the first aspect can affect the stability of the drone and its driving, the second aspect can affect in some cases the visual requested by the pilot to avoid issues or dangerous situations.

For this last aspect, the Italian national standard prescribes that it is mandatory the constant monitoring and responsibility of the drone by a recognized and licenced pilot and the operations can be summarized into three categories:

- VLOS (Visual-Line Of Sight): operations conducted inside a proper distance, both vertical and horizontal, in which the remote pilot is able to maintain constantly the visual contact with the drone, driving all of its movements and avoiding collisions;
- EVLOS (Extended Visual-Line Of Sight): operations conducted over the limits required by VLOS in which the visual contact can be maintained thanks to alternative instruments;
- BVLOS (Beyond-Visual-Line Of Sight): Operations conducted over the maximum distance in which the operator is not able to maintain a constant direct contact with the UAV. It is not possible to manage its movement and avoid collisions (ENAC, 2020).

In addition, if the flight path is realized with the purpose of performing a 3D image-based reconstruction, the trajectory to be considered is strongly affected by the triangulation matching between the points of each single picture collected (Boller, Kuo, & Echsmann, 2012).

4.2 - Case study: flight path planning description

The following subchapter introduces how the drone flight has been planned for our data acquisition.

As mentioned before, a complex of four buildings located in the suburbs of Spino d'Adda, that is used as office and data monitoring station by Politecnico di Milano and Agenzia Spaziale Italiana, has been selected as case study.

This phase has been conducted thinking about two missions: collecting proper pictures for the 3D model reconstruction and provide a preliminary feedback for a future automation of the entire procedure.

As preliminary step, a map identifying the possible places to perform all the operations without disturbing the normal activities has been set.

As it was mentioned before, there are two institutions working in the site monitoring the data collected by a set of antennas that are constantly active. Due to the possible interferences caused by the drone with the installations, it was strictly forbidden to fly around them. For this reason, a no-fly zone in the south of the area is set in order to perform all the operations without asking an interruption of the other activities.



Figure 14 - Panoramic view of the site and no-fly zone identified

The flight planning is made for collecting proper pictures to reconstruct a reliable 3D model using photogrammetry knowledge. For this reason, considerations about the distance between camera and object, position and distance of the camera points has to be known by the analysis of GSD (Ground Sampling Distance, that is the distance between two consecutive pixels centres), calculated in the following way:

$$GSD = \frac{S_W * H * 100}{F_R * imW}$$

For qualitative reasons, in order to obtain a proper result in the 3D modelling, a maximum GSD equal to 1 cm/pixel has been set, so that a maximum distance between camera and object has to be taken into account.

In this case, due to the proximity of different buildings to Block A, object of studies, a maximum distance object-camera equal to 5 m during the movements parallel to the façade has been considered, while H equal to 6 m has been regarded for collecting pictures at the angles.

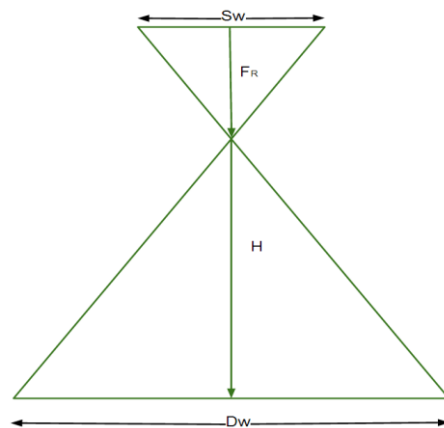
The other two important parameters are the following ones:

- 1- D_w (width of the single image footprint on the ground):

$$D_w = \frac{GSD * imW}{100}$$

- 2- D_h (high of the single image footprint on the ground):

$$D_h = \frac{GSD * imH}{100}$$



In the next table, the technical data of the camera used for the planning:

Parameter	Definition	Value
S_w	Sensor width of the camera	28 mm
H	Distance object-camera	5 m
F_r	Camera focal distance	24 mm
imW (RGB)	RGB image width	4056 pixels
imW (Thermal)	Thermal image width	640 pixels
imH (RGB)	RGB image high	3024 pixels
imH (Thermal)	Thermal image high	480 pixels

Table 3 - Parameters used for the planning

Camera	GSD [cm/pixel]	D_w [m]	H_w [m]	L_f [m]	H_f [m]
RGB	0.14	6.00	4.00	1.17	0.87
Thermo	0.91	6.00	4.00	1.17	0.88

Table 4 - Results obtained for flight path planning

The 3D model reconstruction of the building is based on an image-based technique called Structure from Motion, where every single point of the object is identified by analysing

its position in the space from different pictures. The more the point is represented in different pictures, the more its identification through triangulation is reliable.

For this reason, the overlapping of the pictures is fundamental and a proper parameter requires to be regarded in the preliminary stage before the flight.

For Structure from Motion, an optimal overlapping can vary from 66% to 85%; lower values can help at having a fast flight phase, but the resolution of the model can be problematic and the presence of holes caused missing triangulations can be frequent.

On the contrary, high values of overlapping ensure great redundancy and a reliable 3D reconstruction but the flight phase will be slower and problematic in terms of battery autonomy.

For this data acquisition, an overlapping equal to 80% is considered, so that images are taken every 1.20 m in horizontal and 1 m in vertical both for RGB and for thermal acquisition.

Another important aspect for the image-based 3D model reconstruction is the connection between the points of two non-planar objects such as two consecutive façades.

Nowadays, there are no algorithms that are able to automatically calculate a flight path considering this aspect, so it is important to connect the two consecutive sides of the building taking some angular pictures showing the two objects.

In this case, considering a 90° angle between the two facades, 4 pictures every 22.5° at 6m from the object are collected. This passage is repeated also when the drone is sent to take pictures of the roof.

Another important aspect to study is the starting point of the flight path. Since it is important to have a proper coincidence between the first and the last photo collected at the same level, the point 0 has to be set in a way that the photo taken can include part of the landscape.

The following images show the flight planning preliminary considered for Block A. A “chain movement” is repeated every 1 m for taking the building in a best way and avoid lack of information caused by the possible reflectance of metallic elements.

In yellow it is represented the position 0, while in blue we have angular positions.

During the acquisition of the pictures, it has been noticed that on the south-west side of the building, the 6-meters-distance preliminary set as optimal distance between the building and the camera for sides connection was not possible to be satisfied.

For this reason, just for that side, the acquisition of the photos was made at a distance of 5 m.

4- STAGE 1: FLIGHT PATH PLANNING

However, the correction performed on site did not affected the result reported in the following chapters.

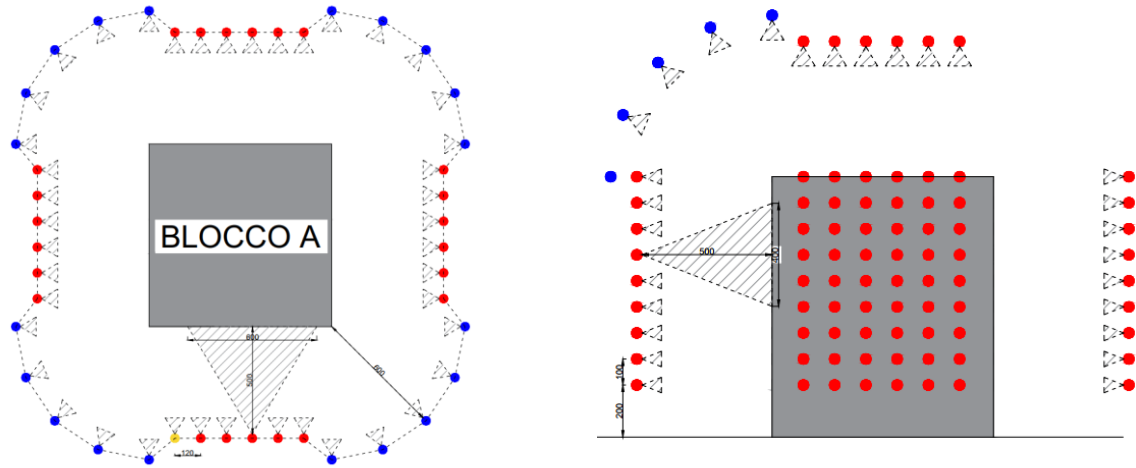


Figure 15 - Flight path planning for the building

5 - Stage 2: 3D model reconstruction

5.1 - State of the art

The discussion presented in this chapter is a propaedeutic introduction about the main computer-based techniques used for the 3D model reconstruction, starting from a 2D image information.

Computer vision 3D reconstruction is mainly diffused in medical fields for the analysis of human body, but also in the archaeological field, for the localization and mapping of the sites and cultural artifacts both on ground and undersea.

The application of 3D reconstruction techniques apport benefits in different aspects, such as limiting the destructive nature of excavating to the possibility of enriching the archaeological research (Ham, Wesley, & M., 2019).

From a research point of view, the first paper published was released in 1981 with an introduction of the possible problems to solve for meeting the target.

During the years, several approaches were presented and nowadays it is possible to declare that the result obtained is acceptable and satisfying.

By analysing the papers published in recent years, it is possible to notice an “horizontal” development of the topic: since all the main issues have now a solution, all the publications recently released present algorithm with alternative passages as hypothetical optimization.

However different sources such as the possibility of using electromagnetic waves scanners, the realization of the 3D model is possible using two main different ways: by laser scanning or using an image-based procedure. Comparing the two options, the difference between them is linked to efficiency and economic issues: while the first guarantees good results in terms of precision and efficiency, the second one is more economic and well-integrable to commercial drones (Inzerillo, Di Mino, & Roberts, 2018).

In any case, the output provided by each algorithm is a 3D point cloud representing the building as object of the reconstruction. Nowadays, on the market IT companies released specific products that are able to measure and rebuild the point clouds into a virtual environment like Agisoft Metashape or Pix4D.

The research is performed considering only image-based reconstructions mainly for its benefits.

5.2 - Structure from Motion: how it works and examples

Structure from Motion is a reconstruction technique based on computer vision algorithms that allows to build an object into a virtual 3D environment starting from the definition and automatic collimation of a series of fundamental points that are measured from a group of 2D pictures, usually taken by a camera or a drone.

The procedure takes a leaf from biology and human anatomy: in fact, human binocular vision catch the image in 2D that is elaborated by the brain into a 3D perception.

There are many methods for obtaining the result: from stereo methods to considerations about the motion.

Applications in Structure from Motion are various and different. In fact, it can be used for object recognition and in computer graphic also.

In addition, concept related to Structure from Motion are the basis for robotics application and automation motion. As mention, a huge number of SLAM (Simultaneous and Localization Mapping) development integrates SfM for their targets.

However, Structure from Motion needs complex geometric theories and algorithms to support it, and it needs to be improved in terms of speed.

Pinhole camera model

For understanding the logic behind Structure from Motion, it is necessary to start from the pinhole camera model, that is a simplified scheme where the light travels in a straight line. The light reflected by the object passes through the pinhole to form an inverted image on the imaging surface.

In general, theory considers that the light striking an object is reflected back in random directions in the scene but only a bundle of these rays traveling along the same direction enters the camera and strike the film in one single point.

However, the ideal model already presented is founded on the following rule: to every point in the scene corresponds a single point on the film.

The distance between the pinhole and the imaging surface is called focal length. Making a brief introduction: the smaller the pinhole, the smaller is the cone and the clearer the image. However, if the pinhole is too small, the image can be affected by blurring.

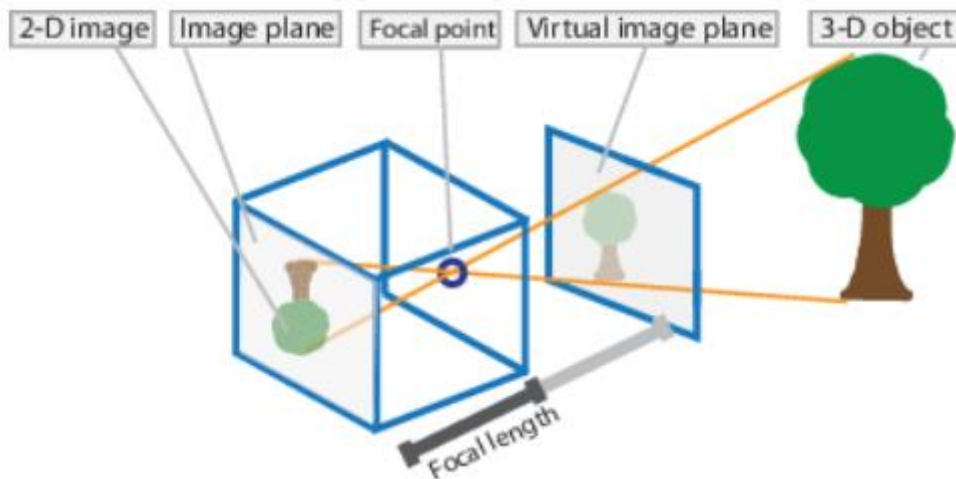


Figure 16 - Pinhole camera model (Source: Matlab - <https://it.mathworks.com/help/vision/ug/camera-calibration.html>)

In addition, a too small pinhole requires a longer exposure time because the amount of light passing through the hole and striking the surface of the film decreases.

As a general rule, the shorter the exposure time the better. There is a limit though to the size of the pinhole: when the size of the hole is about the same as the light's wavelength, light rays are diffracted.

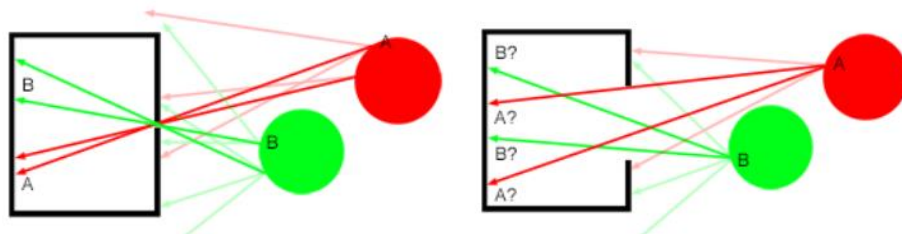


Figure 17 - Refraction based on exposure (Source: Madali – www.towardsdatascience.com)

Camera calibration

For studying Structure from Motion mathematically, it is important to introduce the concept of camera calibration in order to find a value for camera parameters and, consequently, a unique coordinate system. There are mainly two types of coordinate systems in SFM:

- the camera coordinate system: where the optical centre of the camera is taken as the origin. The x-axis is the horizontal direction, the y-axis is the vertical direction and the z-axis points to the direction viewed by the camera.
- the world coordinate system: where the origin can be selected arbitrarily, regardless of the specific location of the camera.

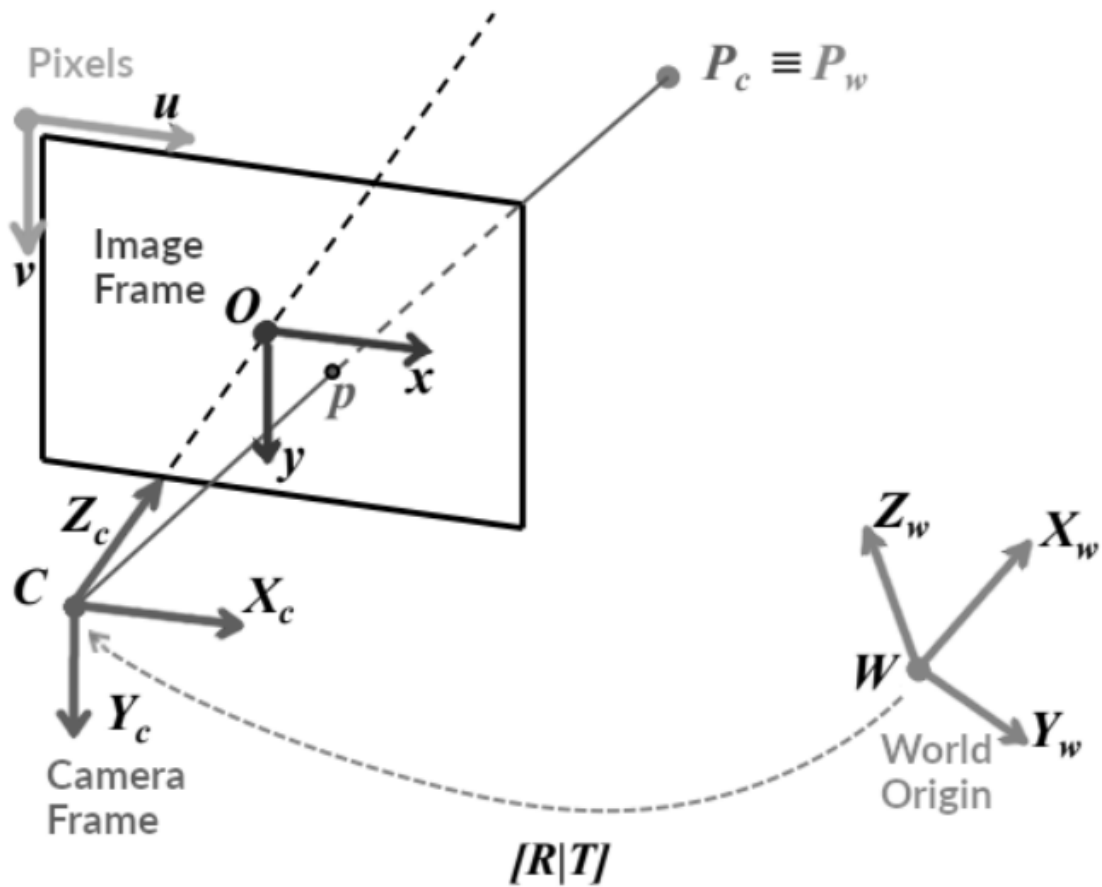


Figure 18 - Camera coordinates (Source: Madali – www.towardsdatascience.com)

In addition, two more coordinate systems are often involved in camera models:

- image pixel coordinate system: where the origin is at the upper left of the image, the u axis is right, and the v axis is down. The unit of the pixel coordinate system is the pixel, which is also the unit of measure for the resolution, as it is shown in the following image;
- imaging plane coordinate system, that are coplanar with the image pixel coordinate system. The origin is the intersection of the camera's optical axis and the imaging plane, which is usually the midpoint or principal point of the imaging plane.

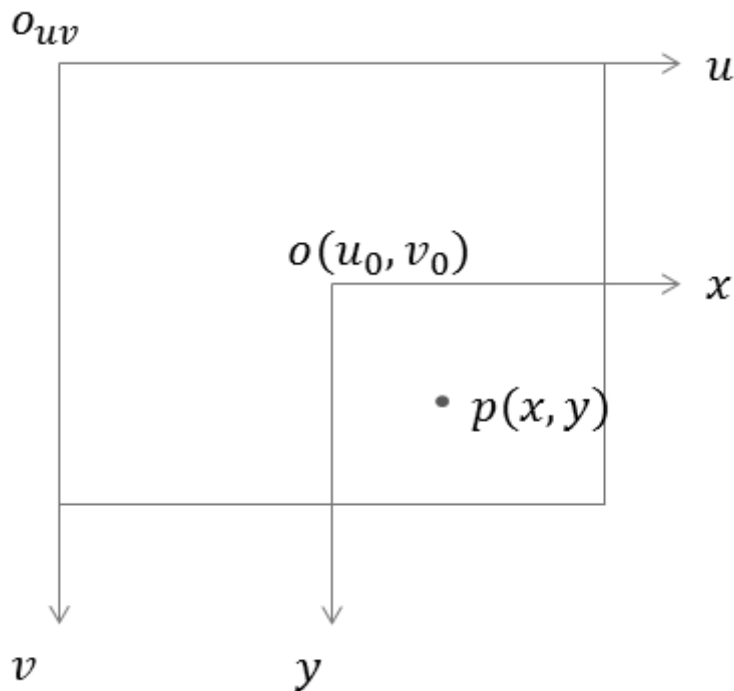


Figure 19 - Image coordinates (Source: Madali – www.towardsdatascience.com)

The unit is a physical unit, such as millimetres. Therefore, the imaging plane coordinate system and the pixel coordinate system are only different in the origin and measurement unit and the two coordinate systems differ by a zoom ratio and a translation of the origin.

Given $P(x, y)$ as the coordinates of the imaging plane, the following equation shows the conversion between the pixel coordinate system and the imaging plane coordinate system:

$$\begin{cases} u = \frac{x}{dx} + u_0 \\ v = \frac{y}{dy} + v_0 \end{cases} \rightarrow \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{dx} & 0 & u_0 \\ 0 & \frac{1}{dy} & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

Where:

- dx and dy represent the physical size of each pixel in the image in the imaging plane, specifically the zoom ratio previously mentioned;
- u_0 and v_0 are the coordinates of the origin of the imaging plane in the pixel coordinate system;

Camera coordinate system

The camera coordinate system is shown in the following scheme. It is possible to see from the figure that there is a perspective projection relationship between the points on the camera coordinate system and the points on the imaging plane coordinate system.

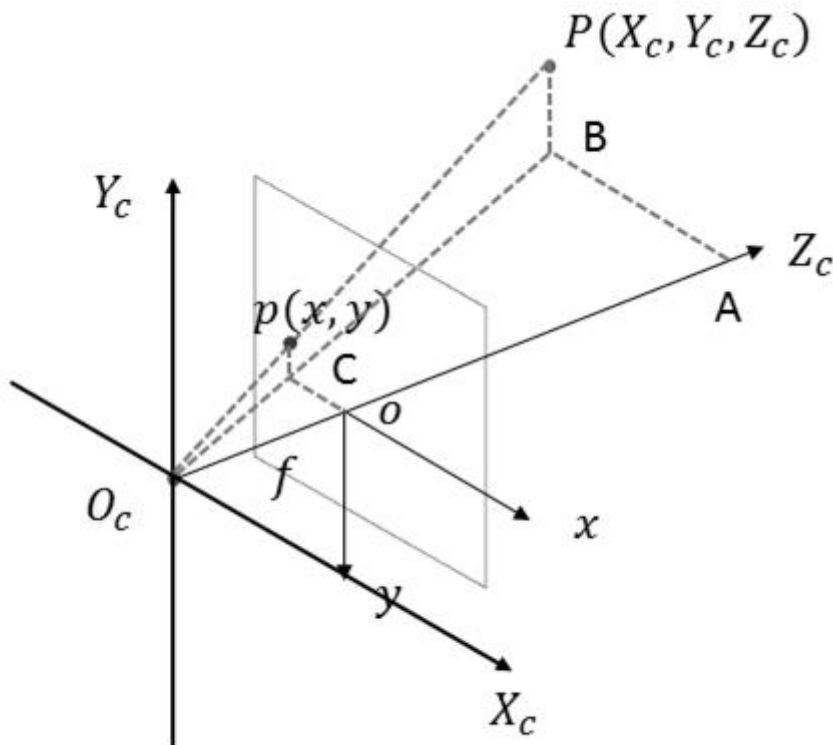


Figure 20 - Camera coordinates system (Source: Madali – www.towardsdatascience.com)

Assuming that the coordinate of the point P in the camera coordinate system corresponding to P is (X_c, Y_c, Z_c) , the following equation shows the conversion relationship between the imaging plane coordinate system and the camera coordinate system.

$$\begin{cases} x = f \frac{X_c}{Z_c} \\ y = f \frac{Y_c}{Z_c} \end{cases} \Rightarrow Z_c \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix}$$

Where:

- O_c is the origin of the camera coordinate system, assumed in the optical centre;
- X_c -axis, Y_c -axis and Z_c -axis are the optical axis of the camera, that are assumed to be parallel with the u -axis and the v -axis of the pixel coordinate system;
- f is the focal length

World coordinate system

It is now introduced the passage that shows the relationship between the camera coordinate system and the world coordinate system, that it can be described by the rotation matrix R and the translation vector t of the camera motion. A reference coordinate system is chosen in the environment to describe the position of the camera and objects.

Assuming that the coordinates of P in the world coordinate system are (X_w, Y_w, Z_w) , it is possible to represent the conversion relationship between the camera coordinate system and the world coordinate system with the following equation:

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} = \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}$$

Coordinate system conversion

The conversion between the world coordinate system and the pixel coordinate system can be realized through the above four coordinate systems, as shown below:

$$Z_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{dx} & 0 & u_0 \\ 0 & \frac{1}{dy} & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & u_0 & 0 \\ 0 & f_y & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}$$

Where:

- Z_c can be considered a scale factor of the system;
- $\begin{bmatrix} u \\ v \\ 1 \end{bmatrix}$ represent the matrix including all the image points;
- $\begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}$ is the matrix representing the world points;

- $\begin{bmatrix} f_x & 0 & u_0 & 0 \\ 0 & f_y & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$ is the matrix including the intrinsic parameters of the camera (focal length and pixels dimension);
- $\begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix}$ is the matrix that includes the extrinsic parameters (translation and rotation);

$\begin{bmatrix} f_x & 0 & u_0 & 0 \\ 0 & f_y & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix}$ is also defined as *camera or fundamental matrix*

The internal parameter matrix of the camera is often known and fixed, while the external parameter matrix often needs to be figured out.

Calculation of the Epipolar constraint and essential matrix

Another important passage is to apply the equation that, starting from the Epipolar Constraint, calculate the matrix E, that is called Essential matrix.

The following scheme shows the relationship between two images collected by two different stages of the same point. The purpose of the essential matrix is to find a way to link the information for the 3D reconstruction.

It is possible to declare that space coordinates of the point and the distance from the point to the camera are not essential for the reconstruction. However, it can be seen from the following analysis that the relative relationship between the two cameras is required and the corresponding points evaluation in the two images also. So, as preliminary step, a problem of feature point extraction and matching has to be solved.

Given a P_w point in the world coordinate system. Considering the coordinate is X_w , its image in camera 1 is x_1 and its image in camera 2 is x_2 , as it follows:

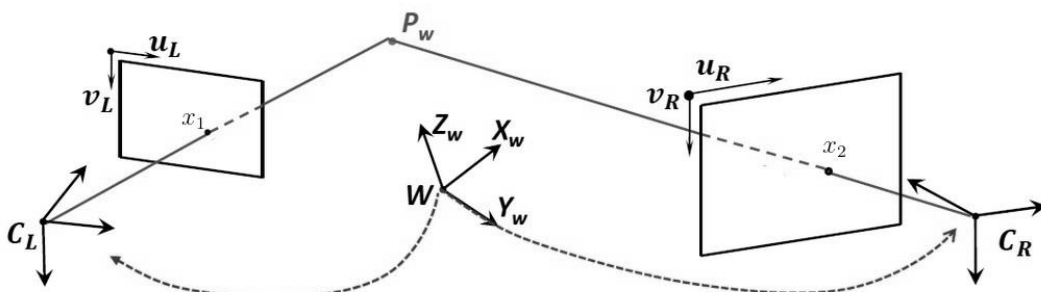


Figure 21 - Epipolar constraint scheme and reference adopted (Source: Madali – www.towardsdatascience.com)

It is also supposed that:

- the vertical distances from X to the image planes of the two cameras are s1 and s2 respectively;
- the two cameras have the same internal parameter matrix K (which is admissible, because it is supposed that the images are collected by the same camera),
- the transformation relationship with the world coordinate system is [R1, T1] and [R2, T2];

It is possible to get the following two equations, Since K is an invertible matrix, it is calculated the K inverse matrix to obtain:

$$\begin{aligned} s_1 x_1 &= K(R_1 X + T_1) & \Rightarrow & \quad s_1 K^{-1} x_1 = R_1 X + T_1 \\ s_2 x_2 &= K(R_2 X + T_2) & \Rightarrow & \quad s_2 K^{-1} x_2 = R_2 X + T_2 \end{aligned}$$

Then, it is possible to assume:

$$\begin{aligned} x'_1 &= K^{-1} x_1 & \Rightarrow & \quad s_1 x'_1 = R_1 X + T_1 \\ x'_2 &= K^{-1} x_2 & \Rightarrow & \quad s_2 x'_2 = R_2 X + T_2 \end{aligned}$$

Where:

- x'_1 and x'_2 are the normalized image coordinates.

Since the world coordinate system can be selected arbitrarily, it is chosen the world coordinate system coincident to the camera coordinate system of the first camera in order to write the second component in function of the first one. The following equation becomes:

$$\begin{aligned} s_1 x'_1 &= X \\ s_2 x'_2 &= R_2 X + T_2 & \Rightarrow & \quad s_2 x'_2 = s_1 R_2 x'_1 + T_2 \end{aligned}$$

In order to introduce the Essential matrix, it is important to illustrate the three-dimensional vector of the T matrix. By chance, it is started considering the perpendicular relationship between \widehat{T}_2 and x'_2 :

$$\widehat{T}_2 x'_2 = \begin{bmatrix} 0 & -t_3 & t_2 \\ t_3 & 0 & -t_1 \\ -t_2 & t_1 & 0 \end{bmatrix} x'_2$$

Then the inner product of both sides of the equation is calculated to obtain:

$$0 = s_1 (\widehat{T}_2 x'_2)^T R_2 x'_1 \Rightarrow x'_2 \widehat{T}_2 R_2 x'_1 = 0$$

If the Essential matrix is considered as the product of T2 and R2, so:

$$E = \widehat{T}_2 R_2 \Rightarrow x'_2 E x'_1 = 0$$

This equation is called Epipolar Constraint and the matrix E is called Essential matrix.

Derive essential matrix

It is important to understand how many point-correspondences are necessary for calculating the essential matrix and the research has looked into this problem for many years.

The first popular solution method uses 8 points to get a unique solution. It is popular due to its relative easy implementation. This algorithm is called 8-point algorithm or Longuet-Higgins algorithm.

There are many algorithm to solve this problem, both iterative and non-iterative, that were implemented in 2004, but the last mentioned is still the most popular.

Considering two points and applying the epipolar constraint to a pair of correspondence, the following equation based on the elements of matrix E is obtained:

$$E = \begin{bmatrix} e_{11} & e_{12} & e_{13} \\ e_{21} & e_{22} & e_{23} \\ e_{31} & e_{32} & e_{33} \end{bmatrix}$$

$$u_2 u_1 e_{11} + u_2 v_1 e_{12} + u_2 e_{13} + v_2 u_1 e_{21} + v_2 v_1 e_{22} + v_2 e_{23} + u_1 e_{31} + v_1 e_{32} + e_{33} = 0$$

The equation can be repeated for many points and the following structure is extracted:

$$\underbrace{\begin{bmatrix} \bar{u}_2^1 \bar{u}_1^1 & \bar{u}_2^1 \bar{v}_1^1 & \bar{u}_2^1 & \bar{v}_2^1 \bar{u}_1^1 & \bar{v}_2^1 \bar{v}_1^1 & \bar{v}_2^1 & \bar{u}_1^1 & \bar{v}_1^1 & 1 \\ \bar{u}_2^2 \bar{u}_1^2 & \bar{u}_2^2 \bar{v}_1^2 & \bar{u}_2^2 & \bar{v}_2^2 \bar{u}_1^2 & \bar{v}_2^2 \bar{v}_1^2 & \bar{v}_2^2 & \bar{u}_1^2 & \bar{v}_1^2 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \bar{u}_2^n \bar{u}_1^n & \bar{u}_2^n \bar{v}_1^n & \bar{u}_2^n & \bar{v}_2^n \bar{u}_1^n & \bar{v}_2^n \bar{v}_1^n & \bar{v}_2^n & \bar{u}_1^n & \bar{v}_1^n & 1 \end{bmatrix}}_{Q \text{ (this matrix is known)}} \begin{bmatrix} e_{11} \\ e_{12} \\ e_{13} \\ e_{21} \\ e_{22} \\ e_{23} \\ e_{31} \\ e_{32} \\ e_{33} \end{bmatrix} = 0$$

Once the essential matrix E is known, it is possible to decompose it into the rotation R and translation T for this we can use the SVD.

$$E = USV^T$$

Where:

- U and V are matrixes linked with the image;
- $S = \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ is a 3x3 diagonal matrix with the eigenvalues as diagonal elements;

On the other hand, it is possible to calculate T and R in function of the known matrix S:

$$\begin{aligned}
 - \hat{T} &= U \begin{bmatrix} 0 & \pm 1 & 0 \\ \pm 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} S V^T \\
 - \hat{R} &= U \begin{bmatrix} 0 & \pm 1 & 0 \\ \pm 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} V^T
 \end{aligned}$$

It is possible to calculate 4 possible solution each one, but only one of them has the points in front of both cameras.

3D Reconstruction

Once the transformation matrix between the two cameras is known, as well as the coordinates of each pair of matching points, a passage of triangulation of the matched points is started by restoring the coordinates of the matching points in space through this known information, and a 3D reconstruction of the captured space is generated. In the previous derivation, it is possible to obtain:

$$s_2 x_2 = K(R_2 X + T_2) \Rightarrow \widehat{x}_2 K(R_2 X + T_2) = 0$$

Then, it is possible to write linearly the equation. The following equation cannot be solved directly by inverse, so it is transformed into a sub-equation and solved using SVD.

The geometric meaning is equivalent to drawing the extension lines of x1 and x2 from the optical centres of the two cameras.

$$\widehat{x}_2 K R_2 X = -\widehat{x}_2 K T_2 \Rightarrow \widehat{x}_2 K (R_2 \quad T) \begin{pmatrix} X \\ 1 \end{pmatrix} = 0$$



Figure 22 - Example of 3D reconstruction by Matlab (Source: Matlab – Structure from Motion from multiple views - <https://it.mathworks.com/help/vision/ug/structure-from-motion-from-two-views.html>)

Multi-View SFM

It is now introduced the 3D reconstruction integration operated in Structure from Motion with more than 2 images. As the number of images increases, the difference between the newly added image and the first image may increase and the extraction of feature points becomes extremely difficult.

After the binocular reconstruction method to reconstruct the first two images, so that some points in the space are obtained and consequently the point cloud, it is calculated the pose for every subsequent frame using the P3P or DLT algorithm considering the point correspondences of the other keyframes.

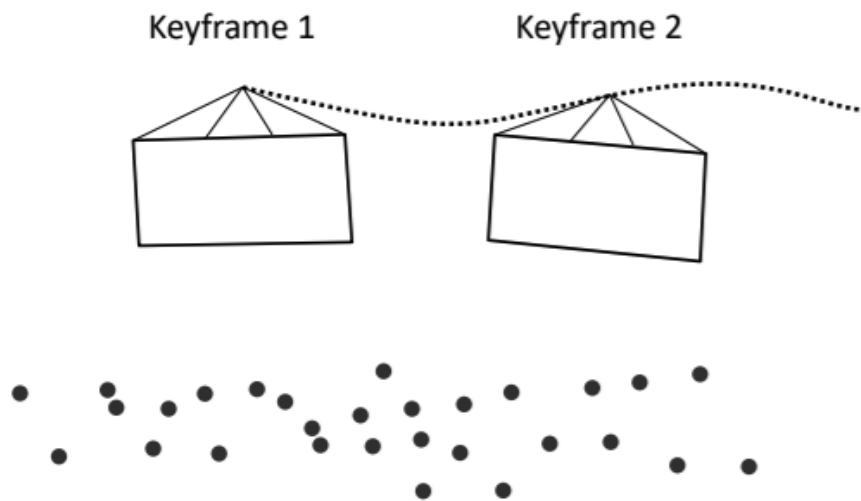


Figure 23 - Keyframe positioning (Source: Madali – www.towardsdatascience.com)

When there is not enough correspondence anymore, a new keyframe is set and it is matched with the previous keyframe. Among these matching points, the spatial coordinates of some of these matching points are known. In this way, new points are added at the point cloud by triangulation.

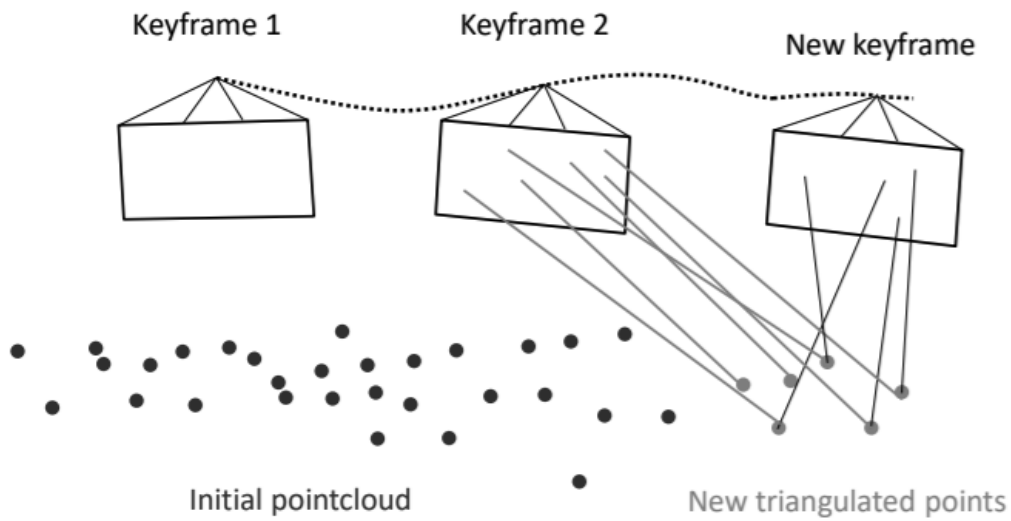


Figure 24 - Keyframe matching (Source: Madali – www.towardsdatascience.com)

Once the motion estimation is known, the structure can be refined optimizing the calculated trajectory. For doing that, a sliding window-method is used and a bundle adjustment is applied for it.

Bundle Adjustment (BA) is essentially a nonlinear optimization algorithm. It is usually used after the estimation of the transformation (R and T). BA refines motion and points at the same time the by minimizing the Sum of Squared Reprojection Errors:

$$(p^i, C_2) = \underset{p_i, C_i}{\operatorname{argmin}} \sum_{k=1}^K \sum_{i=1}^N \|p_k^i - \pi_k(p_i, C_k)\|^2$$

Where:

- C_i is the position of the camera in the world frame. For the actual minimization we can use the Levenberg-Marquadt algorithm.

Sometimes large reprojection errors have a bad influence for the optimization algorithm therefore the squared error is not advisable since the error increase.

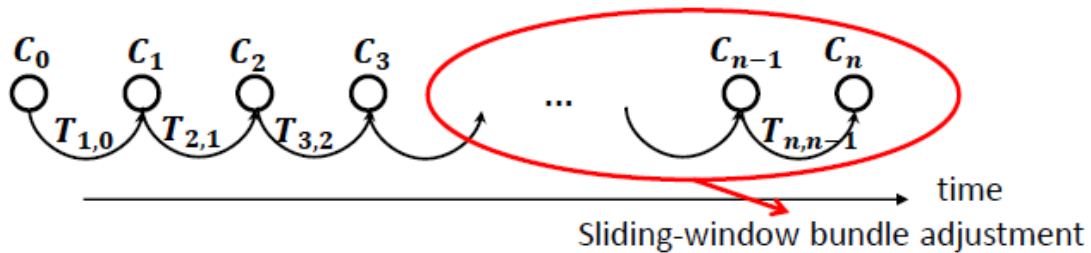


Figure 25 - Bundle adjustment correction after matching operations in Structure from Motion (Source: Madali – www.towardsdatascience.com)

SfM application: Pavement management

It is reported a test conducted by the University of Palermo with the purpose to build a 3D virtual model of cracks present in a street through PhotoScan for preventive maintenance considerations. The test is integrated in a more complex and general concept called Pavement Management, related to the possibility of operating in preventive maintenance, guaranteeing much more efficiency and money saving in a long period of the ways (Inzerillo, Di Mino, & Roberts, 2018).

The test was conducted by collecting photos in a street both with a drone equipped with a 12 megapixel photocamera and a Nikon equipped with an objective 34 megapixel. Then, after the creation of the 3D model through PhotoScan, its accuracy was tested first on MeshLab and then on Rhino considering the Haustroff histogram. The images show the results. Where there is red, it means that the alignment is perfect.

In this case, this condition is obtained more with the Nikon-Realized model, but this happens because of the dimensions of the object. It is important to remark that the targets were the potential cracks, but for an object of greater dimensions the result could have been different.

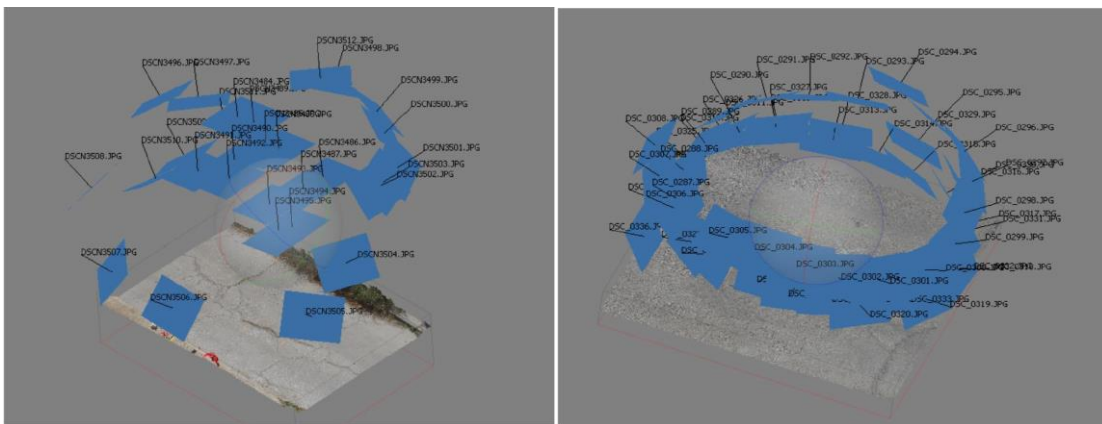


Figure 26 - Example of positioning for Structure from Motion reconstruction (Source: Inzerillo et Al. (2015) - Image-based 3D reconstruction using traditional and UAV datasets for analysis of road pavement distress)

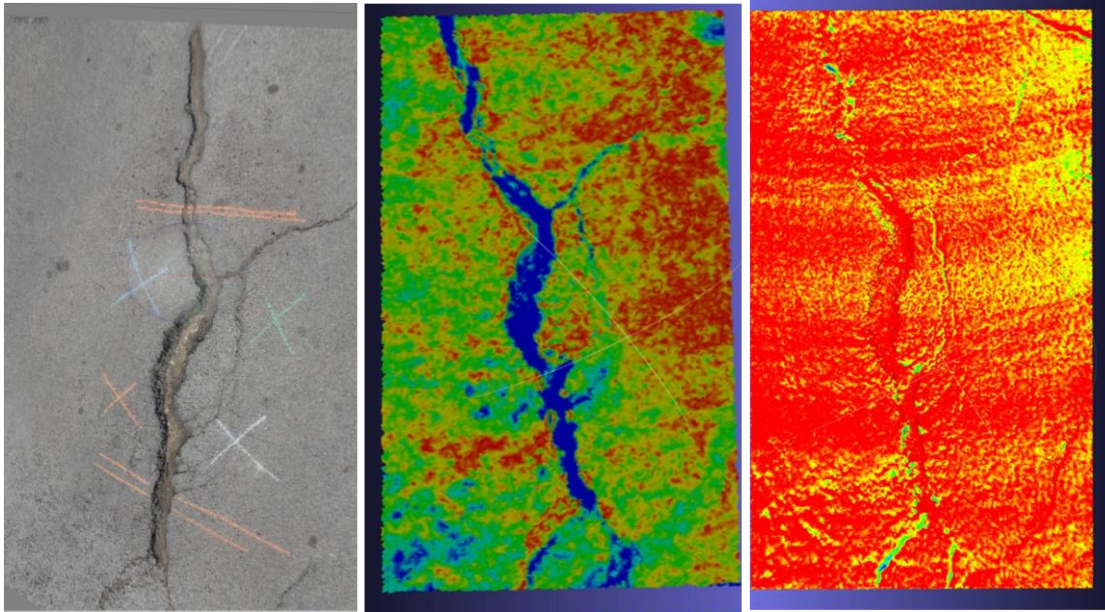


Figure 27 - Haulstroff diagram: Results from Drone-Model and Nikon-model (Source: Inzerillo et Al. (2015) - Image-based 3D reconstruction using traditional and UAV datasets for analysis of road pavement distress)

5.3 - SIFT

SIFT stands for Scale-Invariant Feature Transform and was first presented in 2004, by D.Lowe. SIFT is invariance to image scale and rotation and it is an algorithm useful for collecting the interesting points that are matched in the following 3D reconstruction.

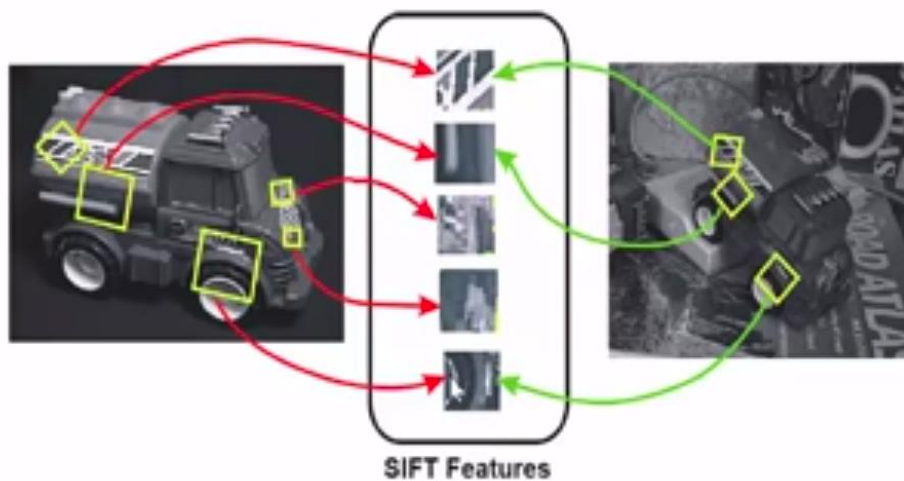


Figure 28 - Example of feature collection (Source: Tyagi - Introduction to SIFT(Scale Invariant Feature Transform) – www.towardsdatascience.com)

SIFT counts many advantages. For instances, since features are local, it is particularly resistant to clutter and occlusions.

Another important advantage is the way how it can distinguish and match individual features in dataset of large objects, without limitation in quantity and real-time performances.

For that reason, SIFT can be extended to a wide range of different features type, adding robustness.

SIFT can be synthesized into five threads:

- Scale-space peak selection: into this thread, the potential location of each single feature is found inside a common space for all of them;
- Keypoint Localization: The feature keypoints are accurately located in the space;
- Orientation Assignment: An orientation is assigned to every keypoint;
- Keypoint descriptor: After this step, it is possible to describe all the keypoints as an high dimensional vector;
- Keypoint Matching

Step 1: Scale-Space Peak Selection

In this step, a first filtering of potential feature is operated from convolution of the image.

Considering the image as a function $I(x,y)$, the scale-space is produced from the convolution of a Gaussian kernel considering different scales with the input image until blurring occurs. It is separated into octaves and the number of octaves, depending on the size of the original image. By default, several octaves of the original image, in which every image size is half of the previous one, is generated.

Within an octave, images are progressively blurred using the Gaussian Blur operator. Mathematically, "blurring" is the convolution of the Gaussian operator and the image. Gaussian blur has a particular operator that is applied to each pixel.

The scale space, can be represented as the convolution of the image into a Gaussian blurring, as it follows:

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y)$$

Where:

- $G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2+y^2}{2\sigma^2}\right)$ is the Gaussian Blur operator that governs the scale-space L ;
- $I(x, y)$ is the image, considered as a 2D set of points at x and y coordinates;
- σ is the scale parameter, set by default equal to 2;

By using a Difference of Gaussian operator, another set of images is set and considered for finding the interest keypoints.

The difference of Gaussian is obtained as the difference of Gaussian blurring of an image with two different σ , let it be σ and $k\sigma$. This process is done for different octaves of the image in the Gaussian Pyramid. It is represented in the following image:

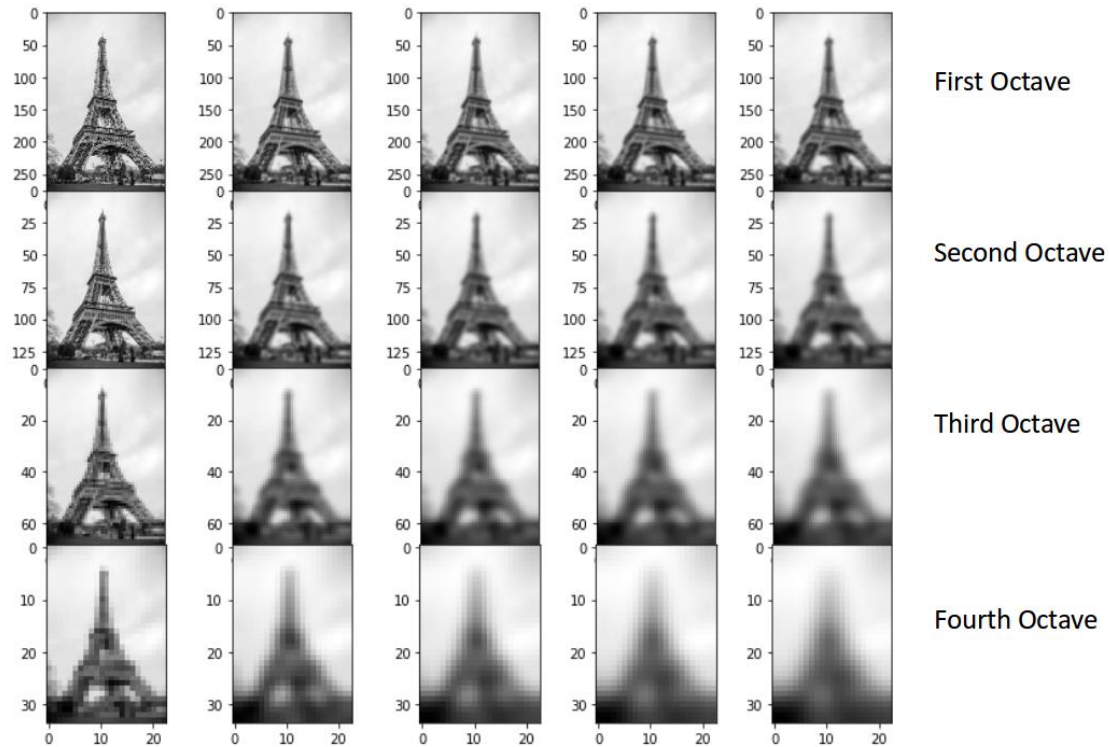


Figure 29 - Example of scale-space peak selection (Source: Tyagi - Introduction to SIFT(Scale Invariant Feature Transform) – www.towardsdatascience.com)

Step 2: Keypoints localization

Considering the previously generated Scale-space, the Laplacian of Gaussian approximation is calculated in order to find scale-invariant localized keypoints.

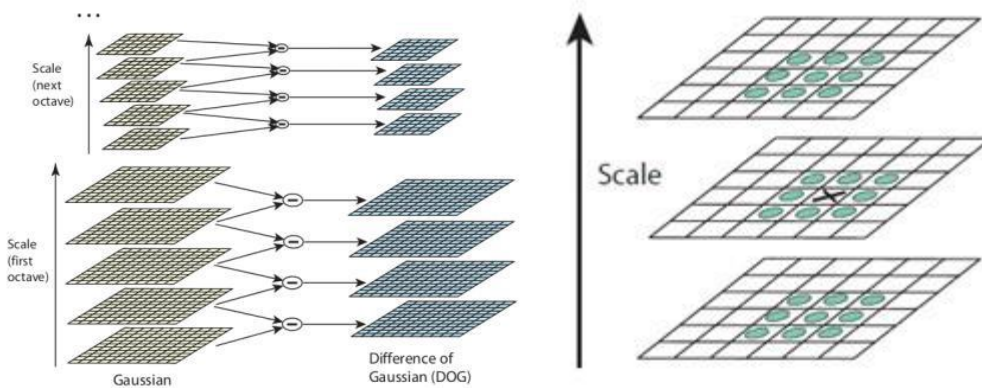


Figure 30 - Example of keypoints localization (Source: Tyagi - Introduction to SIFT(Scale Invariant Feature Transform) – www.towardsdatascience.com)

For all the images, each pixel is compared with its 8 neighbours as well as the 9 projected at the same point in the two adjacent scale. In total, for every pixel, 26 checks are computed. Pixels with local extremum, where the minimum and the maximum corresponds, are considered as potential keypoints because they are well represented in the scale.

Another additional sub-step, implemented by Loeb in 2004, passes through a Taylor series expansion of scale space to get a more accurate location of extrema in order to eliminate the ones without enough contrast.

If the intensity, which is linked to the keypoint, at this extremum is less than a threshold value (usually 0.03), the keypoint is rejected.

Difference of Gaussian has a higher response for edges that has to be removed if they are not accurate. A 2x2 Hessian matrix (H) to compute the principal curvature is used.

- Rejected flat: $|D(\hat{x})| < 0.03$
- Rejected edges:

$$H = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{bmatrix}$$

$$Tr(H) = D_{xx} + D_{yy} = \alpha + \beta$$

$$Det(H) = D_{xx}D_{yy} - (D_{xy})^2 = \alpha\beta$$

Given:

$$\left\{ \begin{array}{l} \frac{Tr(H)^2}{Det(H)} = \frac{(\alpha+\beta)^2}{\alpha\beta} = \frac{(r\beta+\beta)^2}{r\beta^2} = \frac{(r+1)^2}{r} \\ r = \frac{\alpha}{\beta} \end{array} \right.$$

Where:

- α is the eigenvalue of the Hessian matrix with larger magnitude;
- β is the eigenvalue of the Hessian matrix with the smaller magnitude;

It is possible to notice that the factor in r is at a minimum when the two eigenvalues are equal.

Step 3: Orientation Assignment

Once legitimate keypoints are calculated, their stability is tested. Known the scale at which the keypoint was detected, even scale invariance is known, it is then assigned an orientation to compute the rotation invariance.

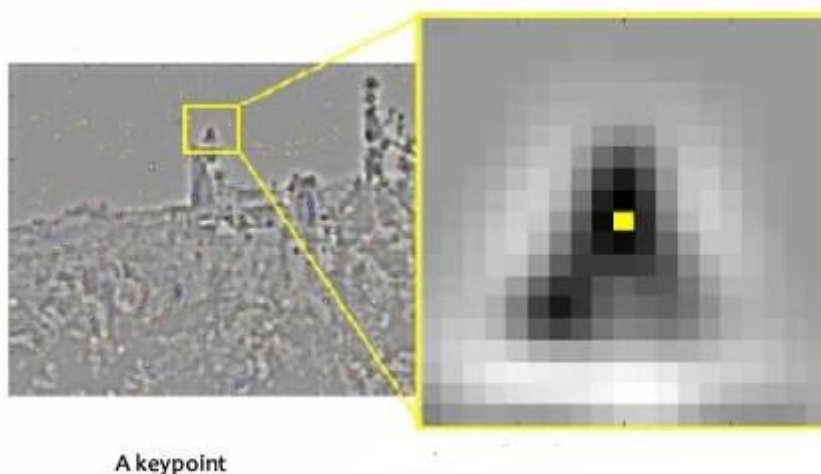


Figure 31 - Example of orientation assignment (Source: Tyagi - Introduction to SIFT(Scale Invariant Feature Transform) – www.towardsdatascience.com)

A neighbourhood is taken around the keypoint location depending on the scale, so as the gradient magnitude and direction are calculated in that region. After this stage, an orientation histogram with 36 columns, representing the range from 0 to 360 degrees, is figured out. Every gradient direction at a certain point is assigned in the previously created ranges and the “amount” that is added to the bin is proportional to the magnitude of the gradient at that point.

Once this step is made for all the pixels around the keypoint, the histogram will have a peak. The highest peak in the histogram is taken, also regarding any peak above 80%.

It is created keypoints with same location and scale, but different directions. It contributes to the stability of matching.

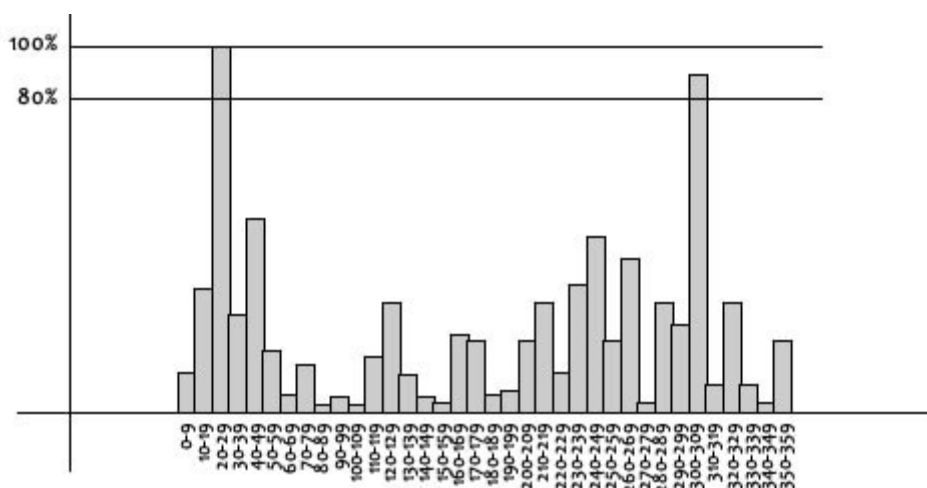


Figure 32 - Histogram for evaluating the stability of the features (Source: Tyagi - Introduction to SIFT(Scale Invariant Feature Transform) – www.towardsdatascience.com)

Step 4: Keypoint descriptor

Once each keypoint has its location, scale and orientation assigned, it is necessary to compute a descriptor for the local image region about each keypoint that should be extremely distinctive and invariant as much as possible to variations such as changes in viewpoint and illumination.

So that, a 16x16 window around the keypoint is taken and it is divided into 16 sub-blocks of 4x4 size.

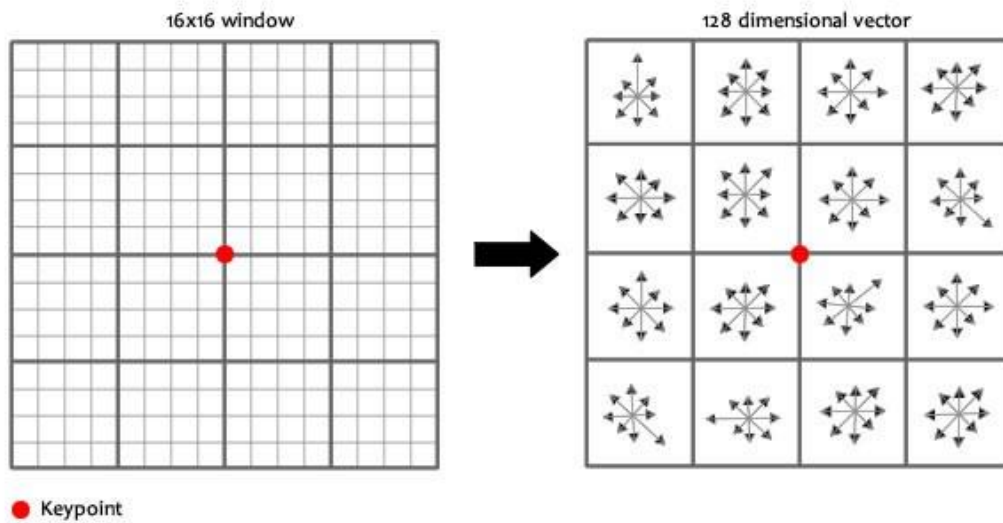


Figure 33 - Example of keypoints descriptor assignment (Source: Tyagi - Introduction to SIFT(Scale Invariant Feature Transform) – www.towardsdatascience.com)

For each sub-block, 8 bin orientation histogram is created.

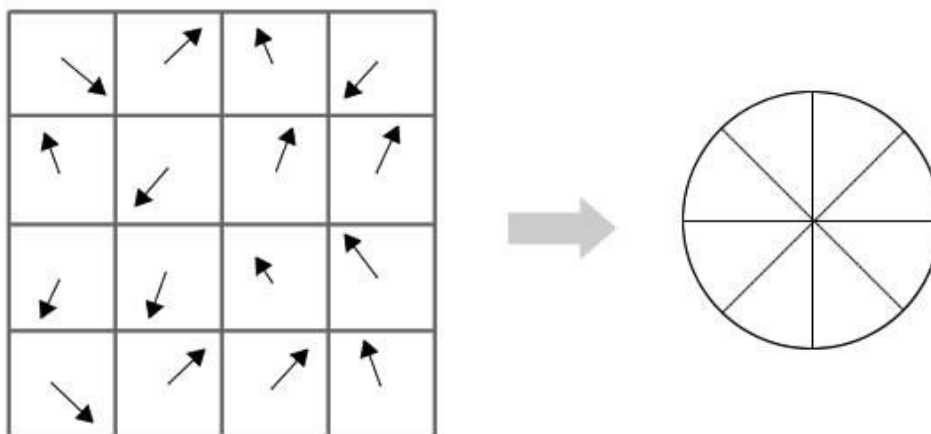


Figure 34 - orientation histogram generated for each bin orientation (Source: Tyagi - Introduction to SIFT(Scale Invariant Feature Transform) – www.towardsdatascience.com)

So, 4 X 4 descriptors over 16 X 16 sample array were used in practice. 4 X 4 X 8 directions give 128 bin values, which is represented as a feature vector to form keypoint descriptor. This feature vector introduces a few complications:

- 1- Rotation dependence: The feature vector uses gradient orientations. So, once the image is rotated, the value can be modified. For achieving rotation independence, the keypoints rotation is subtracted from each orientation so that, each gradient orientation is relative to the keypoints orientation.
- 2- Illumination dependence: if they are considered threshold numbers too high, it is necessary to achieve illumination independence. So, any number of the 128 values greater than 0.2 is changed to 0.2 in order to normalize this gap.

Step 5: Keypoint Matching

Keypoints between two images are matched by identifying their nearest neighbours but in some cases, the second closest match may be very near to the first. In that case, the ratio of closest-distance to second-closest distance is taken and, if it is greater than 0.8, they are rejected. By working in this way, around 90% of false matches are eliminated saving only 5% correct matches.

5.4 – Thermography and its integration into a 3D environment

Within this subchapter the use of thermography for energy detection purposes is introduced and it is also explained how it is integrated in a 3D virtual environment.

Modelling the built environment is a complex challenge to achieve. Given the multiplicity of building parameters such as climate, envelope properties, equipment loads, and occupants behaviour, there is a potential for uncertainty to occur. Several studies have classified the causes of uncertainty in energy simulation models into four types as follows:

- Energy Modelling: uncertainty can result from either the simplification or assumptions in the building physics modelling processes or from the algorithm embedded in the simulation engine.
- Specification: inaccuracies in the specification of building systems such as geometry, envelope's thermal performance, occupancy schedules, material properties;
- Scenario Uncertainty: assumption of external conditions like outdoor climate or building level conditions like occupants behaviour;
- Numerical Uncertainty: numerical errors occur during the simulation process, such as approximation of material thermal properties.

In addition, it is important to determine the existing conditions building envelopes for the models to estimate retrofit savings reasonably using such methods. Building inspections measure and evaluate energy consumption data, examine energy use patterns, and detect energy losses.

Generally, there are three typical levels for building and site audit:

- Walkthrough: energy data are examined and quantified through a visual inspection of building energy systems to evaluate average energy use and establish benchmarks.
- Energy Audit: this includes a detailed analysis of building energy losses and the characteristics of energy systems using site measurements. Also, at this level, the efficiency of building systems and energy conservation measures are evaluated.
- Investment Grade: energy simulation models are developed and employed to predict annual energy use, considering weather conditions, building systems, and occupancy schedules.

The evaluation of the state of health of the envelope is important for the scheduling of the interventions for maintaining the performances designed at the beginning.

The process of envelope evaluation using thermography can be classified into three steps:

- 1- measurement step, that includes data collection and primary assessment of the envelope's existing condition;
- 2- quantification of data using numerical analysis to identify types of thermal abnormalities and their severity;
- 3- Documentation to summarize thermal anomalies detection, class, location, and overall assessment of the envelope.

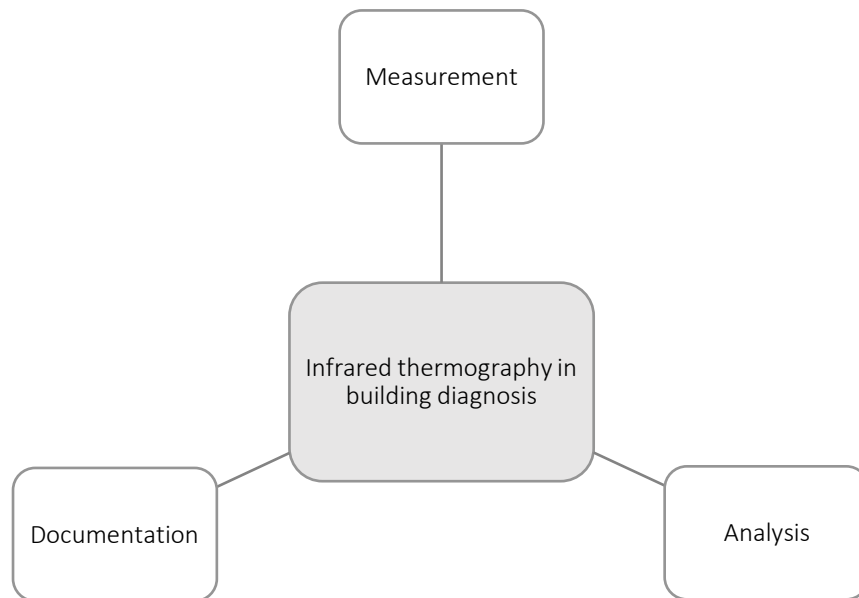


Figure 35 - A general introduction of the steps involved in a IR thermography building diagnosis

Generally, moisture, air leakage, and conduction heat loss are the main responsible of most of the thermal defects in the envelope.

Each one has different visual and thermal characteristics that are important to identify. In addition, external weather conditions such as wind, air temperature, and precipitation play an important role in the behaviour of them.

There are two main approaches to examine thermal anomalies: qualitative and quantitative. The difference between the two methods is the regulation of measurements' procedures, that need to be decided before launching the analysis. Qualitative IRT is utilized to determine the presence of thermal abnormalities using visual evaluation of temperature differences in the measured radiation spectrum.

It should be noted that qualitative measurements only distinguish potential thermal anomalies, and it is not to estimate the severity of defects.

Therefore, this approach is mainly used when thermal anomalies are easily detectable. Also, thermographers need to have a sufficient level of knowledge to decide whether thermal differences imply potential defects or not. Qualitative IRT is suitable for damage identification in historic buildings and envelopes deterioration.

Quantitative IRT is based on numerical analysis to quantify thermal anomalies. This method is used for envelope thermal diagnosis, such as the determination of thermal transmittance.

However, the main challenge in this method is that external weather conditions may influence the accuracy of the results.

In the building energy audit field, qualitative and quantitative methods are used at different process levels. At the walk-through level, qualitative IRT is used to collect information for the following parameters:

- 1- performance of buildings' envelope;
- 2- identification of air leakage and heat losses;
- 3- evaluation of thermal insulation;

Quantitative IRT is used for the development of simulation models and quantification of thermal anomalies through the evaluation of the following parameters:

- 1- thermal transmittance measurement;
- 2- severity and percentage of thermal defects;
- 3- examination of insulation quality;
- 4- moisture and air leakage determination

Thermography analysis techniques in literature are classified into two approaches: active and passive thermography. For thermal evaluation of the envelope, both methods are used to detect the existing condition of façade materials. Passive thermography focuses on the whole building envelope to identify energy losses problems by measuring temperature differences under normal solar radiation conditions.

Active thermography concentrates on particular building envelope areas, that are thoroughly examined to identify thermal defects. In this approach, an external energy source generates a temperature variation between the outdoor environment and the building's envelope.

After the review about the Infrared Thermography, the attention is focused on the integration of the images collected in the 3D model reconstructed by the drone in Structure from Motion.

The scope is to have a complete 3D preview of the site in order to let the drone performing different analysis, qualitative and quantitative, for the thermal energy monitoring evaluation of the surfaces involved.

In spite of what is done with RGB pictures, it is not possible to elaborate a 3D output by matching thermal 2D images due to their low resolution (commonly around 640x480 pixels against 4056x3040 pixels of a normal RGB photo) and other parameters.

Another solution that links the two types of information is necessary.

Two techniques are proposed and studied: the bi-camera method and a photo-texturing based on a scaling factor.

The first consists in process images that are collected simultaneously by two connected cameras, one acquiring RGB and the other one thermal, fixed in their positions. The calibration of the system is possible because their relative poses are known (Barazzetti L. et al. 2013).

By considering the rotation matrix R_{RGB} , the relative rotation matrix of the Infrared camera can be found out by considering a geometrical transformation based on the system stability condition.

The rotation matrix of the system can be computed as it follows:

$$R^* = R_{RGB}^T R_{IR}$$

Another constraint that is regarded is the centre of perspective of the camera. Since the system is composed by two cameras simultaneously acting, it is not possible to consider as centre the one of two of them.

However, by considering the fixed position, the translation vector t is assumed constant along all the transformation, so as:

$$t = R_{RGB}^T \Delta X$$

Where ΔX is equal to the difference between the perspective centre of the two cameras.

The system has been tested in Milan at Trifoglio Building of Politecnico. The geometric model was automatically generated by a software working with Terrestrial Laser Scanning and the bi-camera approach was used for the thermal model rebuilding.

At the end of the test, the result was considered not acceptable due to the lack of information that was not collected by the cameras.

In conclusion, since the strategy seems to be interesting, the manageability of the result results too tricky for its development.

In addition, it is important to remark that all the tests were performed considering an approach that was suitable for ground data gathering. All the considerations made before about the stability of the transformation matrix decade when the system is adopted for a drone acquisition. First, because two drones cannot fly close to each other for safety reasons, and second because of the instability of the flight.

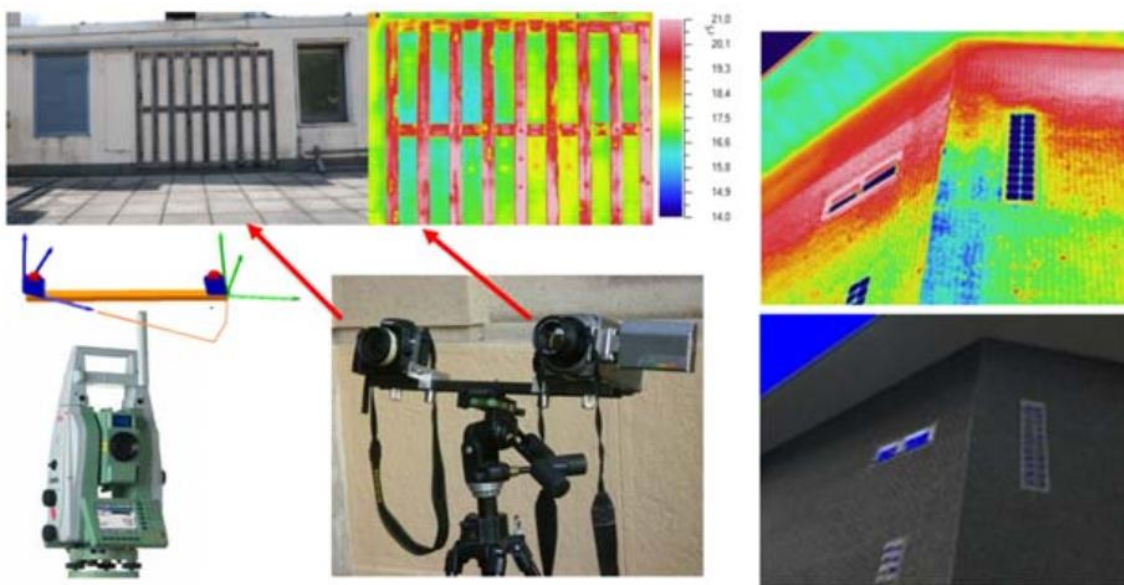


Figure 36 - building, bi-camera system with integrated RGB and thermal cameras, and textured 3D model (Source: Barazzetti et Al. (2015) - Mosaicking thermal images of buildings)

The second technique is based on the position of some fixing points, called baselines.

They are considered known and extremely visible points by which the position in the space is measurable.

Those points are measured in RGB images after the orientation and then they are reported also in the same position inside IR pictures for generating the same reference point even in the new reconstruction.

After the generation of the mesh, representing the geometric model, the IR photos are used for creating a texture model covering it and properly scaled considering the position of the baselines that are located in the same place.

By adopting this system, it is also avoided the problem related of the low resolution of IR images. On the contrary, in spite of the accuracy of the model, the position of the baselines still requires the user intervention.

In the following pictures an example of IR integration through the use of the baselines is presented. It represents an infrared analysis realized for the façade of Nave building in Politecnico di Milano.

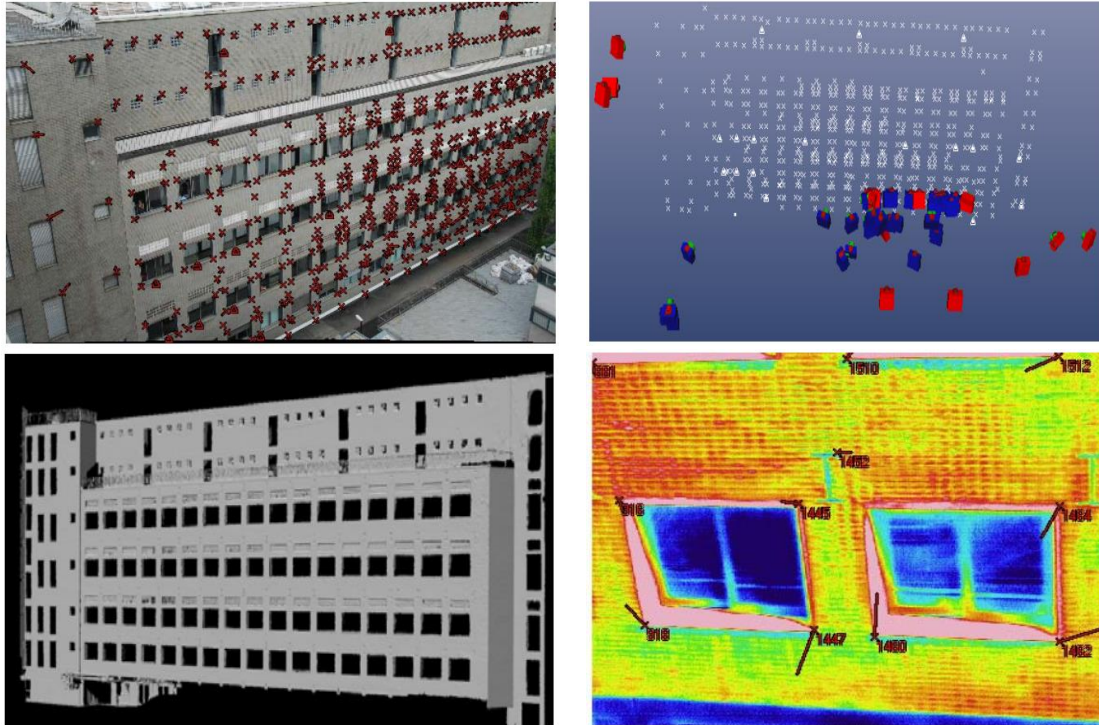


Figure 37 - Baseline approach for Edificio Nave (Source: Barazzetti et Al. (2015) - Mosaicking thermal images of buildings)

5.5 - Case study: Structure from Motion application using Agisoft Metashape

It is now presented the result obtained during the phase of reconstruction of the Spino d'Adda complex, object of study.

The work can be divided into two main sub-threads: the one related to the principal 3D model generation and a second part that involved the integration of the IR information.

In order to obtain the model, it was decided to use a licence software called Agisoft Metashape, that is able to perform all of the actions described in the previous chapters and give as output what we desire.

Two image datasets have been collected during one acquisition. In total, 656 photos have been collected, divided as it follows:

Type	Picture	Resolution (pixels)	Format
RGB	328	4056x3040	Jpeg
IR	328	640x280	Tiff

Table 5 - pictures used for the matching

It is now described the 3D model generation phase in which the first dataset has been involved.

As preliminary stage, the image quality has been evaluated in order to delete images resulted difficult to be read and that could have compromised the quality of the result.

For performing that, it has been used an internal evaluation algorithm that gives a mark to each single picture by scanning their own sharpness and resolution, representing a quantitative selection of the quality.

As it is shown in the following table, the lowest mark is 0 while the highest could be even higher than one.

Based on the resulted dataset, the results are collected and an analysis of the distribution has been performed.

RGB quality	Mark range	Distribution
Excellent	1.00+	24%
Acceptable	0.70-1.00	76%
Not acceptable	0.50-0.70	0%
Very bad	0.00-0.50	0%

Table 6 - Evaluation quality for RGB

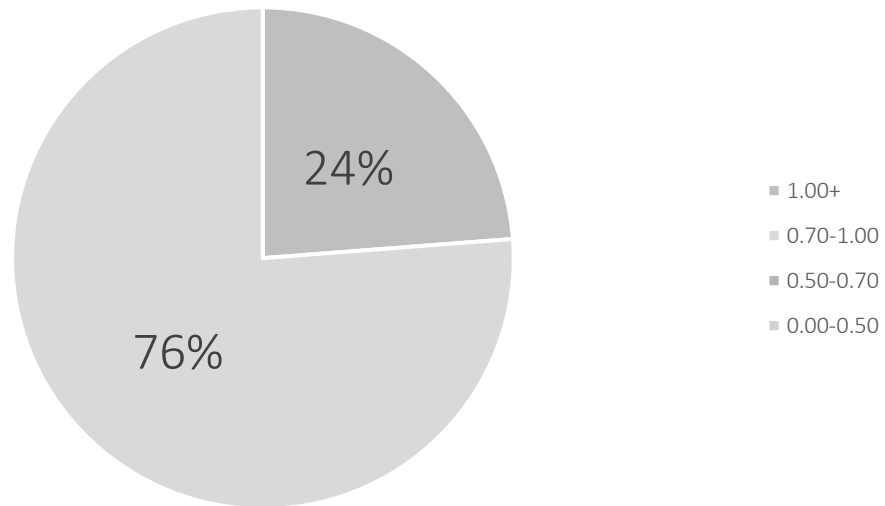


Chart 3 - Evaluation of the picture quality for the RGB images. The mark is calculated using an internal algorithm of Metashape

It is possible to notice that over 75% of the pictures are considered acceptable, while 24% of the overall number is even considered optimal by the algorithm.

It is important to mention that no pictures have been considered non-acceptable in terms of sharpness and resolution for the 3D reconstruction, but a qualitative analysis of every single photo has been performed in order to remove the ones that, accordingly to our sense of evaluation, could have created problems. Fortunately, all of them passed the selection.

Then, it has been possible to start the operations related to the orientation of the images, feature selections and matching step.

Using the EXIF file, which is intrinsic on every camera, it was possible to communicate to the software the type of camera used for the acquisition so as it could have extracted all the intrinsic parameters for the process. In the following table, the intrinsic data extracted by the software with their value.

Camera type		Frame
Pixel size (pixel)	640	480
Focal length	F	4.5
Perspective point on X	Cx	14.93
Perspective point on Y	Cy	14.27
Radial distortion coefficients	K1	-0.00255
	K2	0.0542
	K3	-0.1439
	K4	0.1026
Tangential distortion coefficients	P1	0.0003
	P2	0.0004
	P3	-13.73
	P4	-2.33

Table 7 - Camera parameters for SfM

For alignment and matching operations, the accuracy concerning how those steps are performed is important and it has to be considered. Generally, the more the detection is accurate, the higher the number of points.

On the contrary, the more the phase of feature detection is accurate, the higher is the time required by the software for performing the process.

The operation of feature detection is performed by Metashape applying SIFT, while matching operations are made with normal SfM techniques.

It is important to mention that in this phase we only know the pixel coordinates from where the camera points are extracted.

It has been conducted a test for understanding how the accuracy would have influenced the matching. Starting from the same picture's dataset, the alignment and matching procedure is repeated for four different accuracy levels:

- Highest: pixels are oversampled in order to research as many points as possible;

- High: pixels are processed in the same number;
- Medium: pixels are under-sampled about a factor equal to 2;
- Low: pixels are under-sampled about a factor equal to 4;

Since the number of points to be processed is high and the result is expected to be satisfactory, some boundary conditions are applied to the operations in order to make the algorithm quicker and reduce the processing time without losing the robustness of the system.

In particular, three main conditions were imposed at matching process, that includes a generic preselection of the images based on the georeferenced positions of the cameras in order not to make the process for distant images that have really low probability of matching points.

The filter has been refined also considering a reference preselection: by the georeferentiation, only closest images were compared between themselves.

As in our case, were the georeferenced system WGT84 was not precise, giving errors in the order of 10 m, a local coordination system based on the ground was selected by the software and all the coordinates were taken by it.

As last filter applied there is a limitation on the tie points measured. Because from a good matching there could be measure a huge number of points, making the system heavy and slow, it was asked to Metashape to calculate all the points and select the ones 40000 with the best precision and resolution.

The four scenarios mentioned were analysed and the results compared. In particular, it was studied the number of stable points found on an overall set of points and their ratio.

Some considerations about the time of processing, which is divided between time of alignment and time of matching, are also made and the results are shown in the following charts.

	Characteristics points	Total points	Ratio
Highest	318478	357116	89%
High	289789	339168	85%
Medium	232805	266548	87%
Low	48736	63834	76%

Table 8 - The four scenarios for each different accuracy

5 - STAGE 2: 3D MODEL RECONSTRUCTION

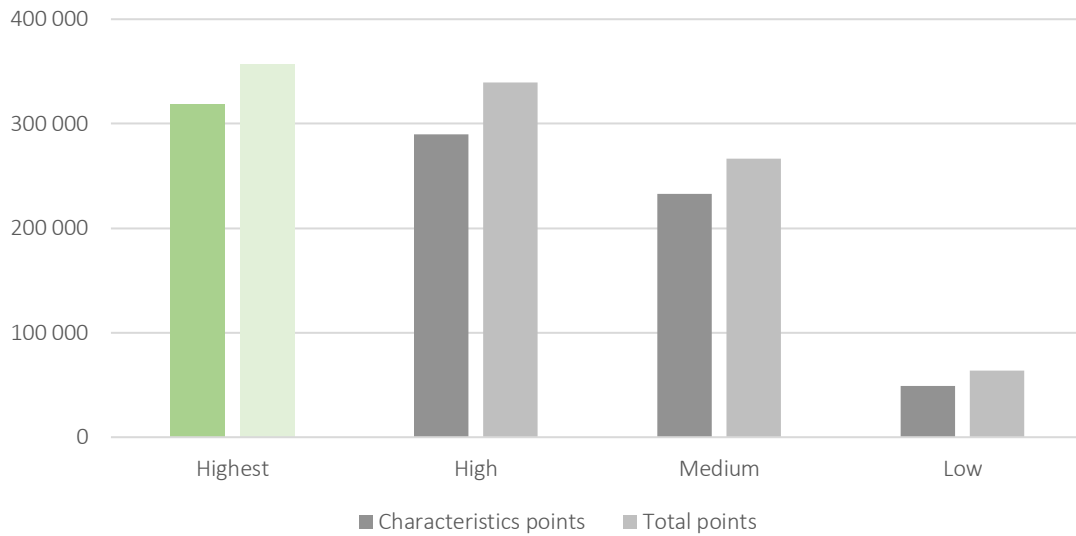


Chart 4 – Number of characteristics points measured on the overall total points found for each different accuracy in every scenario

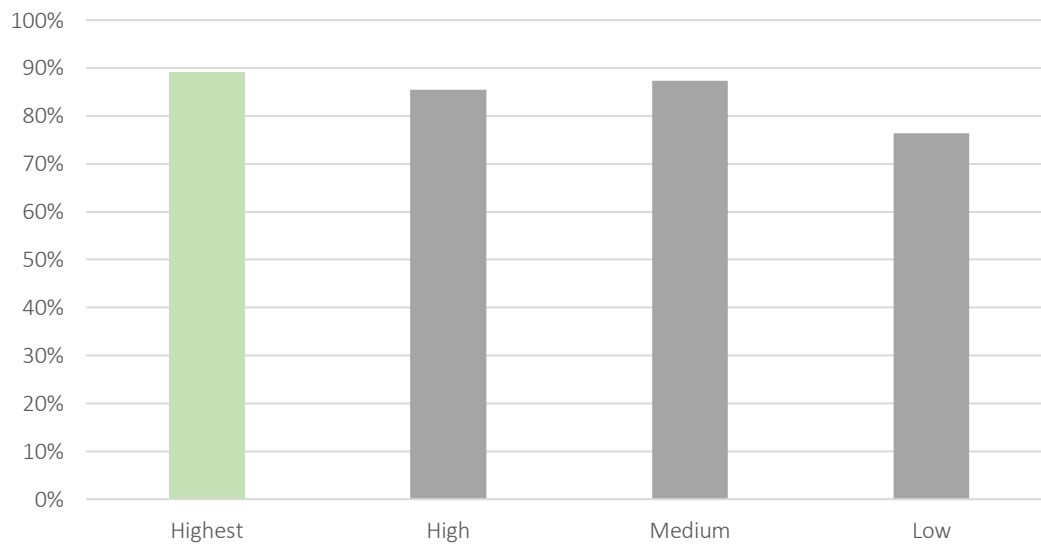


Chart 5 - Number of characteristics points measured on the overall total points found for each different accuracy in every scenario

From the comparison of the results obtained, the scenario with the accuracy set to highest results the best both in terms of number of points measured and also in terms of characteristics points confirmed.

Even the ratio between them is the highest measured, with 89% of the points found in a stable position in the space.

As predicted, the more the accuracy is reduced, less points are matched and found.

It is interesting to notice that setting the accuracy as medium results in a higher percentage of stabilized points compared to the scenario in which the accuracy was set

as high. It means that the process of feature detection in the mentioned scenario, even if it measured less points, it was able to stabilize them in a proper way.

As last, the scenario where it was asked a matching with low accuracy perform the worst results both in terms of measured points and stable points.

It is now presented the results obtained by analysing the time measured for performing the required operations by the algorithms.

	time of matching (min)	time of alignment (min)
Highest	12	54
High	12	17
Medium	2	11
Low	1	2

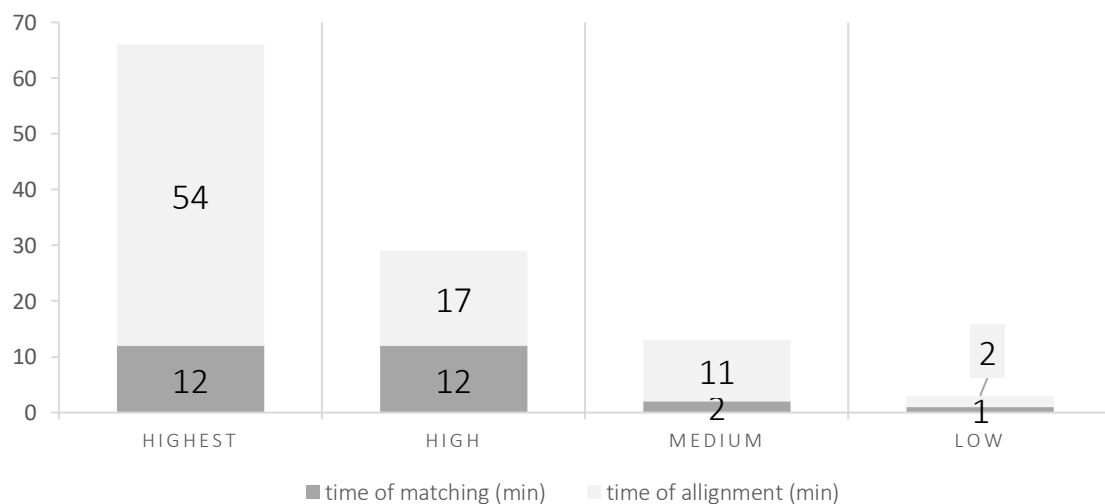


Chart 6 – Time of matching and alignment in minutes Figure 38 - Position of the markers measured in every image collected with the IR camera for each different accuracy in every scenario

As supposed, in the highest scenario the time measured is particularly high, with more than one hour required for the process. This result has to be regarded for future update, where the contemporaneity of some processes will be tested. A good compromise between precision and time of processing needs to be found.

Fast results are provided by the scenario in which it was asked a low accuracy, but it is important to mention that not all the images were correctly aligned.

In light of the last considerations, the scenario with high accuracy results the best compromise between precision and velocity. On the other hand, since this research needs to find a way for future updates, it was decided to go on with the first scenario.

5 - STAGE 2: 3D MODEL RECONSTRUCTION

In conclusion, a qualitative selection of the proper scenario is done looking at the output obtained, that they are shown in the following pictures.

This step was extremely useful for understanding which parts would have created noise in the creation of the mesh or dense point cloud.

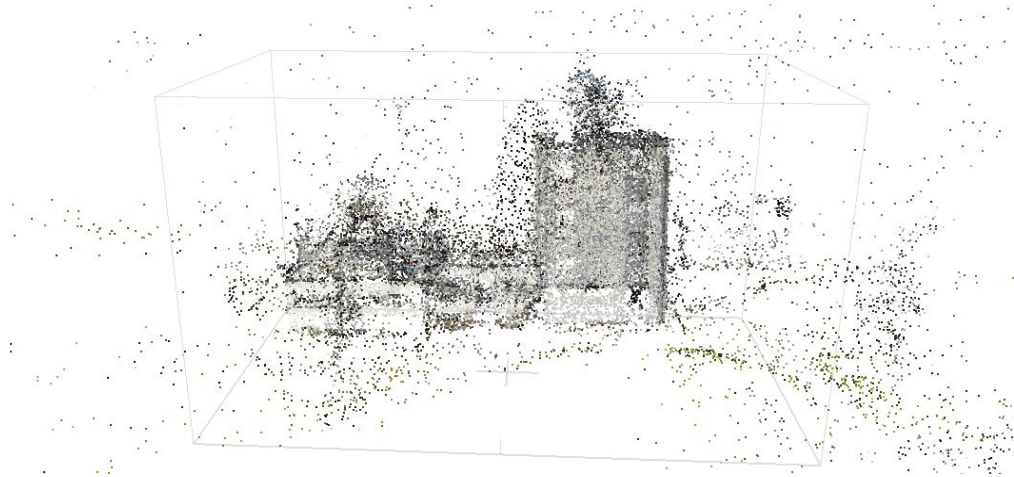


Figure 39 - Sparse point cloud for low accuracy



Figure 40 - Sparse point cloud for medium accuracy



Figure 41 - Sparse point cloud for high accuracy

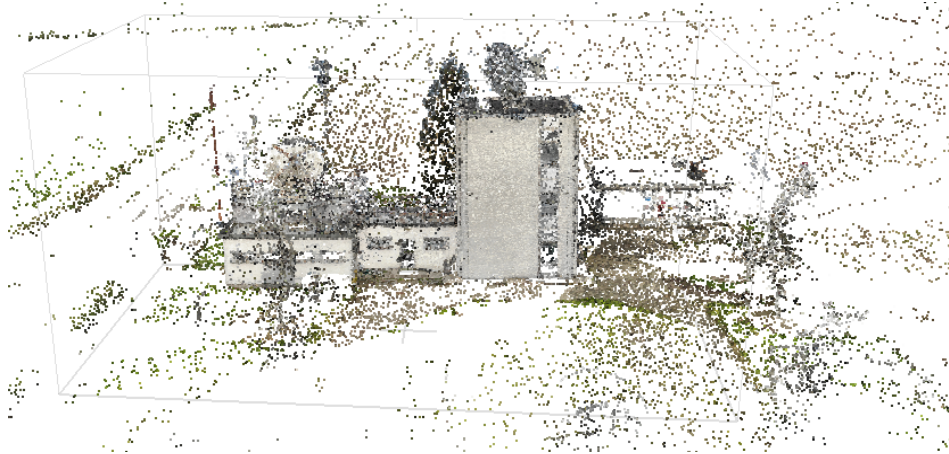


Figure 42 - Sparse point cloud for highest accuracy

In addition, it was performed a checking by analysis the projections of each measured point. In fact, for considering the reconstruction reliable, it is important that all the points are projected by at least 3 pictures. In this case, none of the pictures reported problems in this sense.

Once the matching operations are ended and the sparse point cloud is obtained, the goal of the mission is to check the accuracy of the 3D object and make it as similar as possible to the reality in terms of dimensions.

As mentioned before, it is not possible to trust completely in the georeferenced system because of its inaccuracy.

In addition, in our case it was decided to work with local coordinates internal to the software, without knowing precisely the error calculated.

In order to improve the robustness of the system, some points in field are measured and the distance between them is inserted as boundary condition.

The scope of this part of operation is to fix some markers and know distances and then ask the software to adapt all the reconstruction to them. In this way, errors of distances and positions in the 3D point cloud are strongly reduced.

In general, 8 markers were identified and placed into the model starting from the research of them inside every single picture.

Although this is a time-consuming operation, starting from the coordinate in every image, Metashape is able to filter all the pictures for working only with the ones in which the interested point has good probability to be fixed.

The operation is controlled looking at the error, metric and pixel, in order to contain it and have a good result. The positioning of the markers has been constantly checked

5 - STAGE 2: 3D MODEL RECONSTRUCTION

looking at the error in terms of pixels: in general, it was considered as acceptable an error lower than 1 pixel.

Another checking performed is on the projection. In fact, in order to identify a point in the space, at least 3 projections are required.

In the following table they are reported all the markers posed with the relative accuracy set and the error in pixels calculated and the checks performed.

It was decided to use 7 points from the corners in the top of the building, that were simple to identify in the pictures.

The distances between those points were measured, inserted as scalebar, and used for scaling the entire model.



Figure 43 - Baseline fixed in the site

Markers	Accuracy (m)	Projection	Error (pixels)	CHECK
point 1	0.005	58	0.41	OK
point 2	0.005	38	0.36	OK
point 3	0.005	42	0.40	OK
point 4	0.005	36	0.53	OK
Point 5	0.005	11	0.55	OK
Point 6	0.005	32	0.57	OK
Point 7	0.005	18	0.26	OK

Table 9 - Characteristics and checking of the baselines

Scale bar	Distance (m)	Accuracy (m)	Error (m)
point 1_point 2	8.00	0.05	-0.03
point 1_point 4	8.00	0.05	0.02
point 2_point 3	8.00	0.05	0.02
point 3_point 4	8.00	0.05	-0.04

Table 10 - Characteristics and checking of the scalebars

The error calculated by the scale bar is related to the expansion or compression of the distance before and after the optimization of the matching.

It is possible to notice that none of the errors considered is higher than 0.05 m, so due to the dimensions of the site, it can be considered a good result.

In the following tables, two other controls are inserted:

- 1- Control scale bar: an error that is calculated internally to the rigid 3D model;
- 2- Check scale bar: an error calculated in the scaling phase.

Total Error	Error (m)
Control scale bars	0.02
Check scale bars	0.03

Table 11 - Final check

All the baseline checked and calculated before are then imported in the dataset composed by the IR picture. The markers are placed in the pictures, that are scaled considering the baselines imported.

5 - STAGE 2: 3D MODEL RECONSTRUCTION

At the end, a texture covering the entire model is generated.

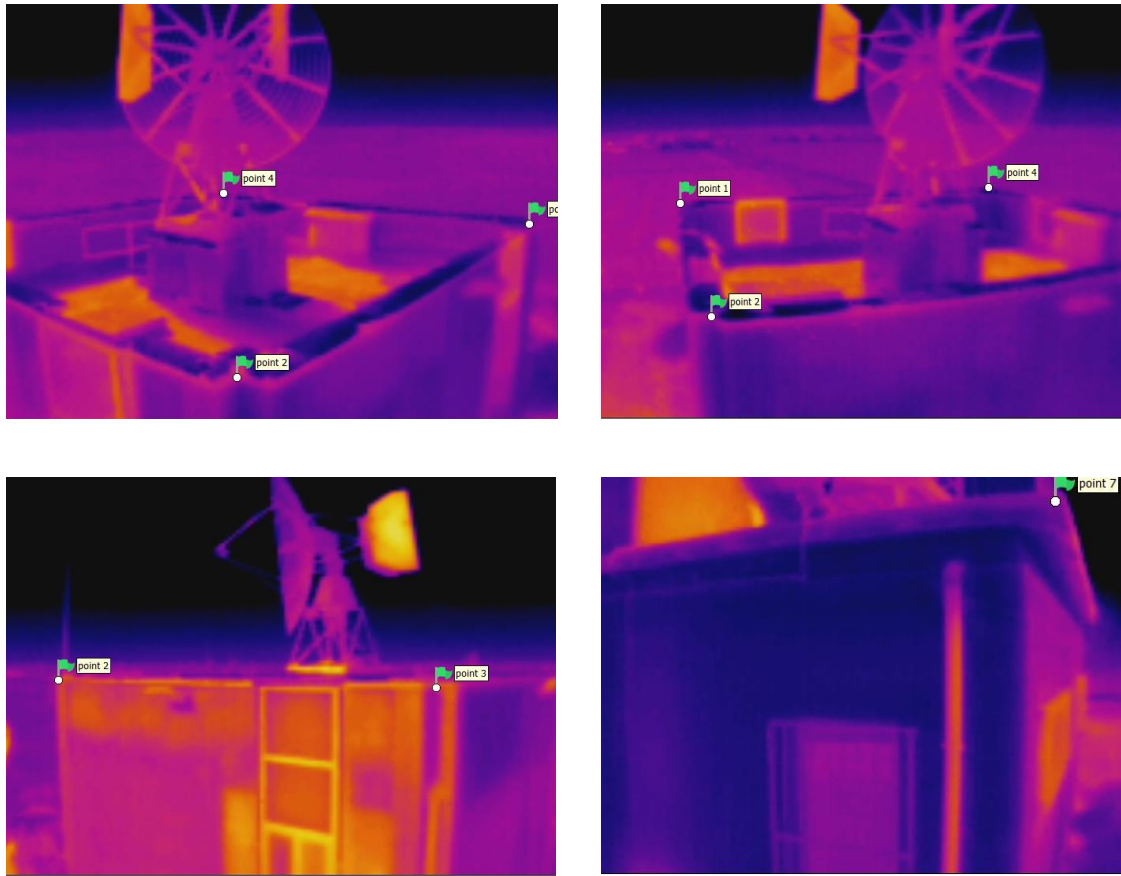


Figure 44 – Position of the markers identified in every thermal image collected with the IR camera

6- Stage 3: Façade modelling

After Stage 2 and the consequent punctual association of the thermal information punctually with the point cloud, the situation obtained is represented by a matrix n-by-9 that contains every point created during the reconstruction.

As briefing, each point is described with 4 values, in order:

- 1- Coordinates in the space (x, y, z);
- 2- Coordinates of colour, that represent the RGB parameters of the point;
- 3- Information of temperature obtained by the conversion from colour.

It is important, then, to analyse the information of temperature obtained in order to have some correspondence with the analytical model.

In order to do that, it is necessary to divide the point cloud into a set of façades, assigning them an orientation and proper information related to materials and dimensions, that are introduced in the next chapters.

In order to achieve the aim, starting from the characteristics of each single point, it is conducted a segmentation. Two main ways were tested:

- 4- Solving the problem using a clustering solution via k-means algorithm;
- 5- Solving the problem using a fitting solution. In this case, MSAC (M-Estimator Sample Consensus) is applied.

In the following chapter, the two algorithms and their relative results on the building are presented.

By comparing what is obtained, limited to a construction site, it results that MSAC is more reliable than k-means.

6.1 - Literature review: k-means

The first attempt was made trying to solve the problem performing clustering analysis, that is one of the most common exploratory data analysis technique used to get a solution about the structure of the data.

It can be defined as the task of identifying subgroups in a set of data, such as a point cloud. All the similar data are set into the same subgroup, which is called cluster. In other words, the goal is to find homogeneous subgroups within the data such that data points in each cluster are as similar as possible according to a similarity measure such as euclidean-based distance or correlation-based distance. The decision of which similarity measure to use is specific to every application.

Clustering is mainly used in image segmentation, where it is necessary to group similar regions together, or document clustering based on topics.

Unlike supervised learning, clustering is considered an unsupervised iterative learning method since there is not a starting measure to compare in order to set the output of the clustering algorithm to the true labels to evaluate its performance.

The structure of the data is just investigated by grouping the data points into distinct subgroups.

An example of clustering algorithm is K-means, that is considered as one of the most used clustering algorithms due to its simplicity.

K-means algorithm is an iterative algorithm that tries to partition the dataset into K pre-defined distinct non-overlapping subgroups (clusters) where each data point belongs to only one group.

It tries to make the intra-cluster data points as similar as possible while also keeping the clusters as different as possible. Every data point is assigned to a cluster such that the sum of the squared distance between the data points and the cluster's centroid (arithmetic mean of all the data points that belong to that cluster) is at the minimum. The less variation within clusters is recorded, the more homogeneous the data points are within the same cluster.

K-means algorithm works iterating three main steps as follows:

- 1- The number of clusters K is specified by the user;
- 2- Centroids are initialized by first shuffling the dataset and then randomly selecting K data points for the centroids without replacement;
- 3- Iteration is kept until there is no change to the centroids. In any case, assignment of data points to clusters does not change.

In addition, three parallel threads are performed by the algorithm in order to be maintained as robust as possible:

- Computation of the sum of the squared distance between data points and all centroids;
- Assignment of each data point to the closest cluster (centroid);
- Computation of the centroids for the clusters by taking the average of all data points that belong to each cluster.

The approach K-means selected to solve the problem is called Expectation-Maximization. The first step assigns the data points to the closest cluster, while the second one computes the centroid of each cluster.

It is presented in the following part the mathematical solution of the problem.

Considering the function in object:

$$J = \sum_{i=1}^m \sum_{k=1}^K w_{ik} \|x^i - \mu_k\|^2$$

Where:

- $w_{ik}=1$ for data point x^i if it belongs to cluster k ; otherwise, $w_{ik}=0$;
- μ_k is the centroid of x^i 's cluster.

The E-step works performing a minimization of two parts.

First, it is minimized J by varying w_{ik} and treating μ_k fixed. Then, the same passage is repeated by varying μ_k and treating w_{ik} fixed. After the passage, the cluster assignment is performed.

The second step is called M-step and it is performed with the differentiation of J varying μ_k and recomputing the centroids after the cluster assignments from previous step. Therefore, E-step is:

$$\frac{dJ}{dw_{ik}} = \sum_{i=1}^m \sum_{k=1}^K \|x^i - \mu_k\|^2 \Rightarrow w_{ik} = \begin{cases} 1 & \text{if } k = \operatorname{argmin} \|x^i - \mu_j\|^2 \\ 0 & \end{cases}$$

In other words, assign the data point x^i to the closest cluster judged by its sum of squared distance from cluster's centroid.

And M-step is:

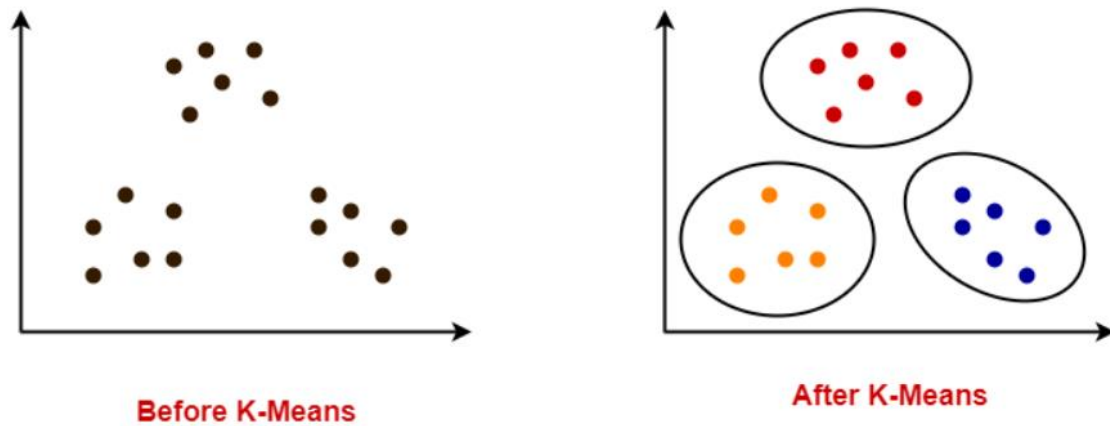
$$\frac{dJ}{d\mu_k} = 2 \sum_{i=1}^m w_{ik} (x^i - \mu_k) = 0 \Rightarrow \mu_k = \frac{\sum_{i=1}^m w_{ik} x^i}{\sum_{i=1}^m w_{ik}}$$

Which translates to recomputing the centroid of each cluster to reflect the new assignments.

It is important to add some additional information:

- Since clustering algorithms including k-means use distance-based measurements to determine the similarity between data points, it is recommended to standardize the data in order to have a mean of zero and a standard deviation of one since the features in any dataset would have different units of measurements.
- Given k-means iterative nature and the random initialization of centroids at the start of the algorithm, different initializations may lead to different clusters since k-means algorithm may stuck in a local optimum and may not converge to global

optimum. Therefore, it is recommended to run the algorithm using different initializations of centroids and pick the results of the run that that yielded the lower sum of squared distance (Dabbura, 2018).



Since the algorithm is iterative and needs a convergency, it is important to analyse how the clusters are identified.

Silhouette analysis can be used to determine the degree of separation between clusters. For each sample:

- Compute the average distance from all data points in the same cluster (a_i).
- Compute the average distance from all data points in the closest cluster (b_i).
- Compute the coefficient:

$$SI = \frac{b^i - a^i}{\max(a^i; b^i)}$$

Where:

- A_i is the average distance of the point from all data points;
- B_i is the average distance of the point from all the data points in the closer cluster;

SI gives back results included between -1 and 1. The following resumes the reliability degree of a clustering in relation of the index.

Range	Reliability
-1.00 < SI,average < 0.50	Not reliable
0.51 < SI,average < 0.80	Reliable
0.81 < SI,average < 1.00	Extremely reliable

Table 12 - Silhouette index evaluation

6.2 - K-Means: application and discussion

The following subchapter presents the results of the k-means application on the 3D point cloud are presented.

In order to do this, the datapoints were considered in matrix form n-by-6 looking both the information about the location and the one about the RGB colour.

In total three main tests were conducted using a Matlab script first considering the information collected without more importance to location or colour. It was predicted that k-means was able to select autonomously the location part and then differentiate every point into subclusters relating colours.

The test was conducted by setting a number of clusters equal to 7 and considering in a random way only 20k points on a total of 600k.

Despite the average silhouette index equal to 0.65, equal to a quite reliable result, the test was considered a failure due to different negative aspects.

First of all, due to the high quantity of points and their dimensions, the algorithm performance was too slowly and gave back results after 8 hours. The long time required by the code for processing data discouraged of repeating the test again.

Second, the low quality of the result was influenced by problems of comparability between spatial and colour dimensions.

Finally, the clustering was influenced mainly by colour information of every point: in fact, it is possible to notice in the following pictures that points are divided only taking the colour as discriminant.

There are some zones in which k-means tried to cluster also following the location, but in general it did not happen.

In the following pictures the results of the k-means clustering are shown, while the table illustrates the range where the information is included.

Information	Range
Location of points along X	0-40
Location of points along Y	0-40
Location of points along Z	0-20
Colour	0-255

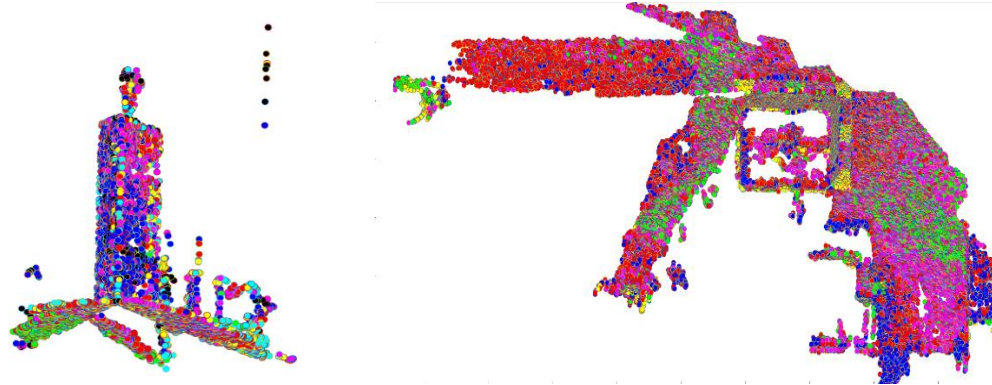


Figure 45 - Clustering division: first attempt

After the experience it was decided to introduce a preliminary step: all the values present in the three last columns, collecting RGB information, would have been reduced by 0.1 or 0.01 in order to drive the algorithm at performing the clustering first considering the location and then the colour of each point.

The test was performed for k included between 2, which is by logic the minimum number of clusters, and 15. Other tests with higher k were conducted but, since the results were not so different from the ones with k equal to 15, it was decided not to include them in the discussion.

For each clustering, the columns related to colour information were reduced once by 0.1 and then by 0.01. After that, the results were compared and the average silhouette index for each sample are summarized in the chart shown below.

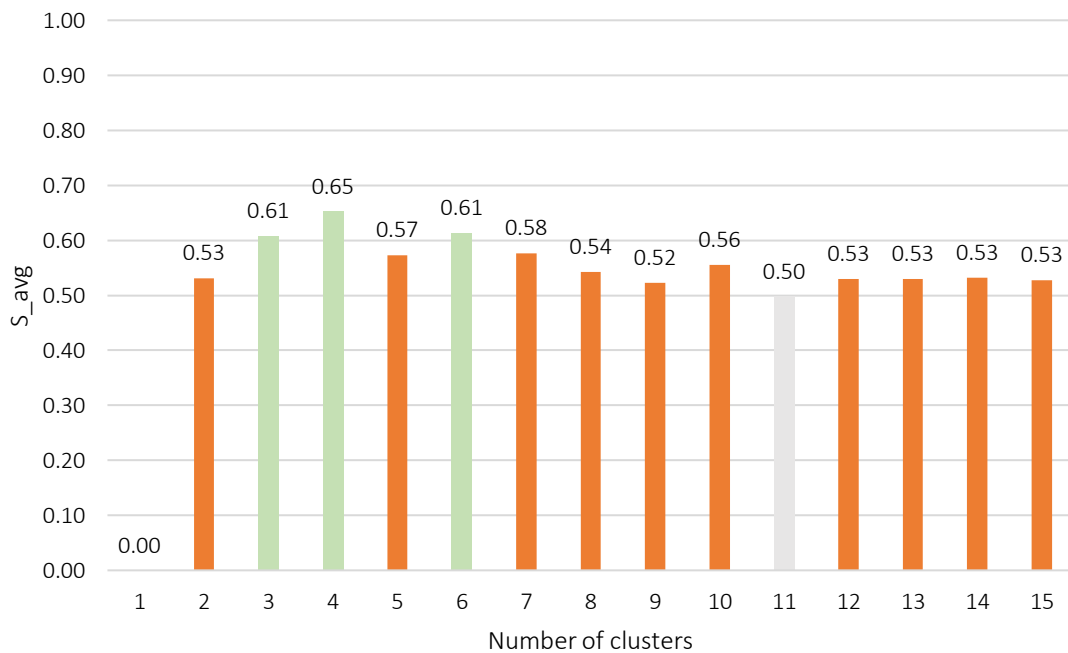


Chart 7 - Evaluation of the silhouette index in after the reduction of 0.01

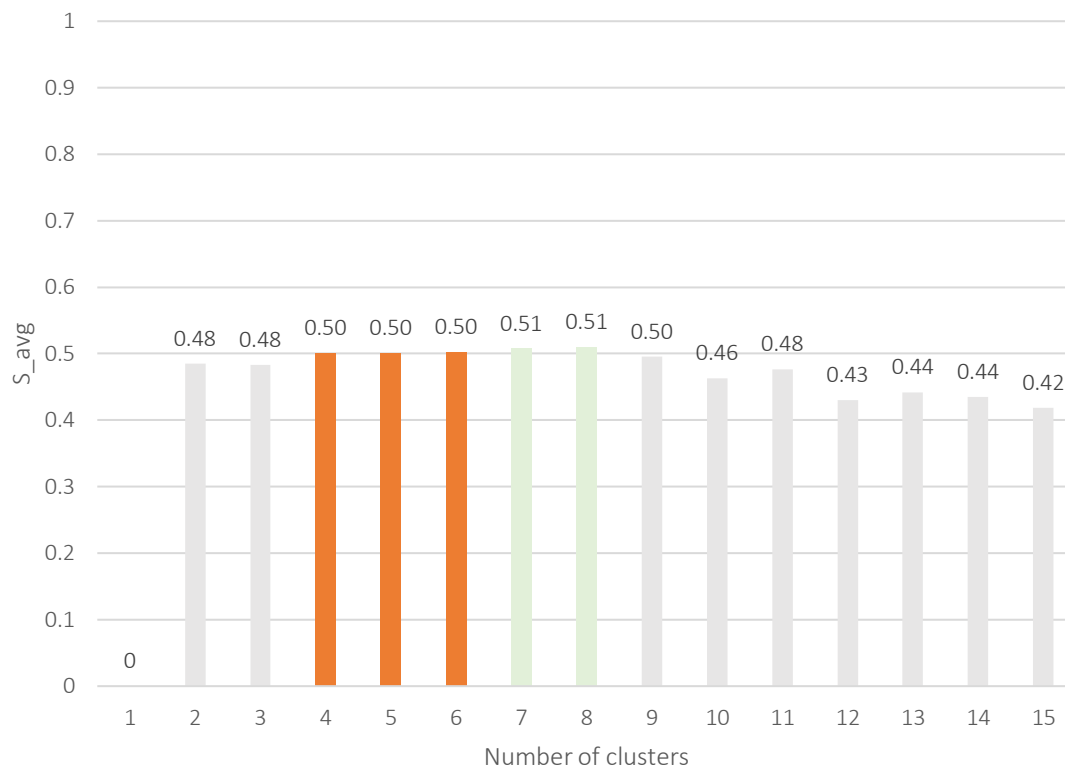


Chart 8 - Evaluation of the silhouette index after the reduction of 0.1

In general, by the analysis of the average silhouette indexes obtained, it was possible to notice a better trending when the RGB values were multiplied by 0.01. All the samples gave back silhouette indexes included between 0.50 (sample with k equal to 7) and 0.65 (sample with k equal to 4).

All the samples can be classified as reliable from a quantitative point of view, but none of them returned a silhouette index result higher than 0.81.

Then, it was analysed the average silhouette index of the tests conducted on the same number of samples with the same characteristics described before. In this case, the RGB columns of the matrix were multiplied by 0.1.

Looking at the results, it was immediately possible to notice, from a quantitative point of view, that 5 samples over 15 reached the threshold fixed at 0.50. In particular, it happened for samples with number of clusters set between 4 and 8.

However, a qualitative analysis of the result has to be performed because of the goal: since it was important to see how the datapoints are divided, an analysis of the distribution of points in the cluster is performed.

The best results of the tests are reported here.

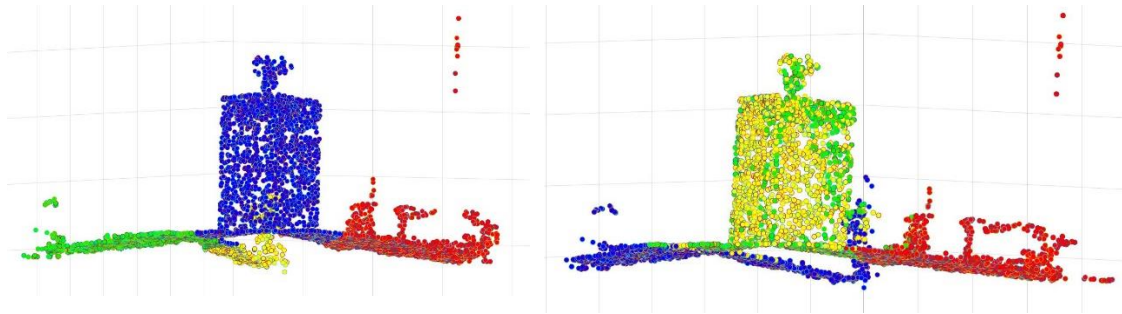


Chart 9 - K=4: clustering for 0.01 and 0.1 reduction

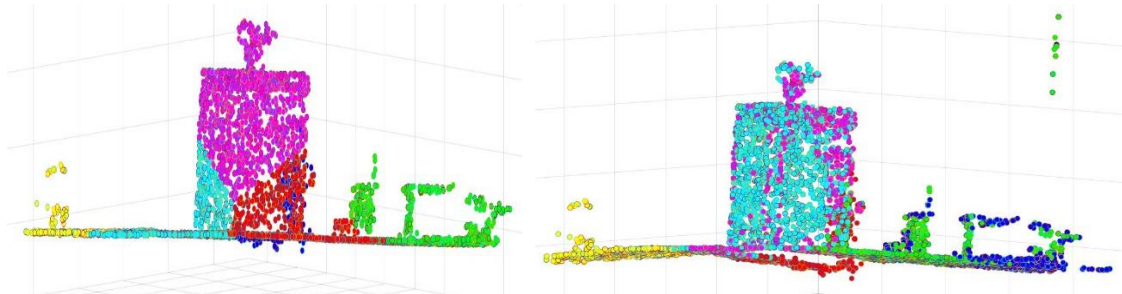


Chart 10 - K=6: clustering for 0.01 and 0.1 reduction

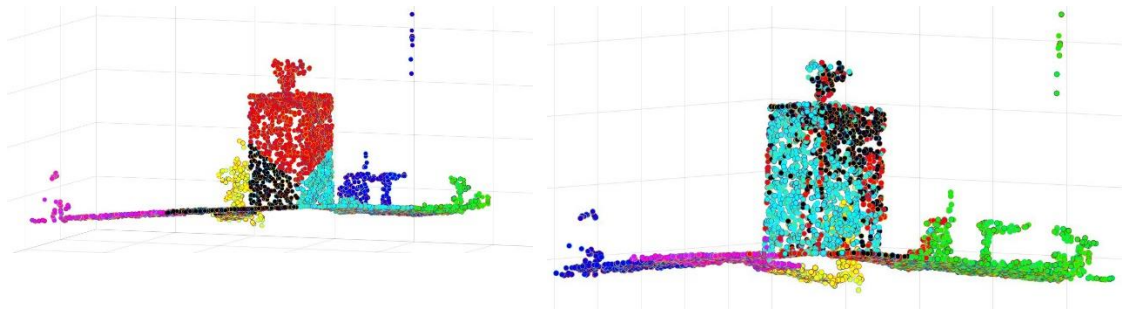


Chart 11 - K=7: clustering for 0.01 and 0.1 reduction

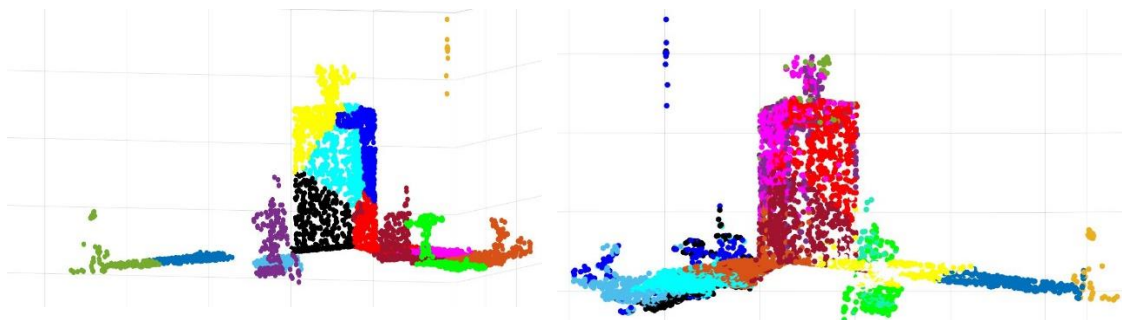


Chart 12 - K=15: clustering for 0.01 and 0.1 reduction

The qualitative analysis of the images tends to confirm the general trend supposed before: agreeing with the average silhouette index analysis, in the samples where the RGB values are multiplied by 0.01 the clusters are more localized than the second case and the results seem to be more reliable.

From a qualitative point of view, compared to the first test, it is possible to notice a general improvement in all the tests performed in the second rush.

However, the qualitative analysis shows that even clusters supported by high $SI_{average}$ are not suitable for façade modelling purposes. It is the case of the sample 4.0.01: all the points representing the building are collected in a unique cluster and differentiated by the ones of the ground, without a detailed differentiation. In addition, part of the sidewalk which turns around the building is included as part of the façade.

While adding number of clusters, the points are not divided uniformly by orientation but the Euclidean distance is calculated only considering the points in a specified area, giving back a “patchwork texture” of the envelope.

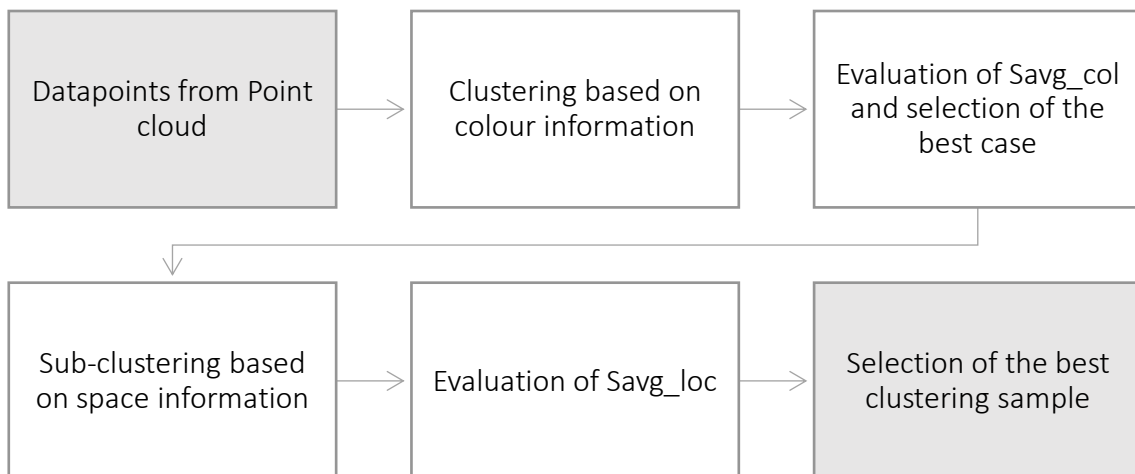
In general, despite the improvements, it is not possible to consider the result satisfactory.

Starting from this experience, it has been performed a third test with the purpose of unifying all the different areas in order to reach the modelling searched.

The application of an algorithm created ad hoc for this case formed by a double k-means working has been performed.

Points would have been distributed into a general cluster by using the reduced colour information of 0.01. Once the best result in terms of $S_{average}$ was found, k-means would have processed a second clustering on the groups that were still obtained.

In the following scheme it is shown the procedure adopted in order to reach the goal.



The best clustering is selected considering the best case resulting from what it was named Silhouette index matrix, which is made by k_{colour} columns and $k_{location}$ rows.

Every term in the matrix represents the final average silhouette index after the two clustering passages.

However, for theoretical reasons and in order to understand the effective efficiency of the algorithm, an analysis of each colour-clustering has been generated. For example, if the number of clusters related to colour is set to 4, the sub-clustering is performed also in the cases with k equal to 2 and 3. If the number of clusters is set to 7, then 7 different solutions representing the best sub-clustering are generated.

After a first analysis based on first colour-cluster it has been possible to have a first result about the trending of the algorithm and excluding some k -values from the off.

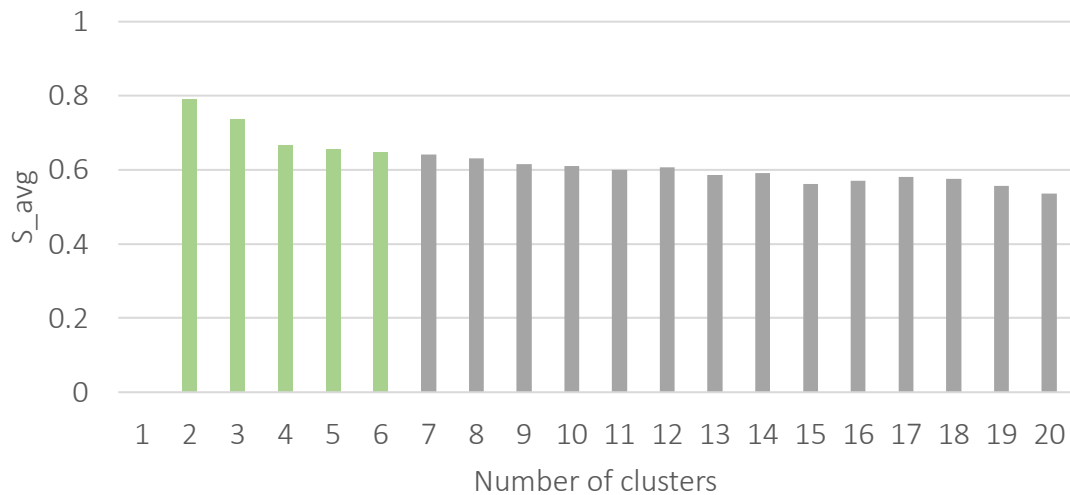


Chart 13 - Silhouette index evaluation for k -colour from 1 to 20

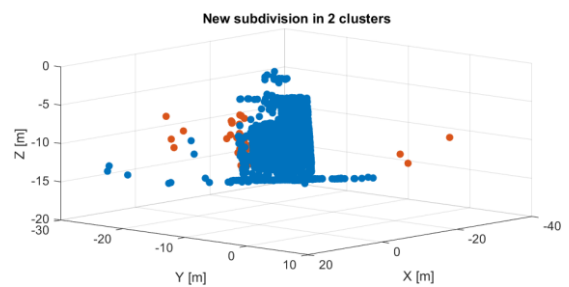
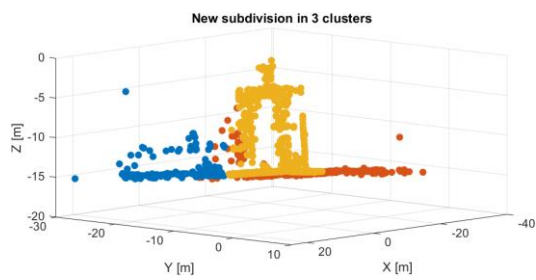
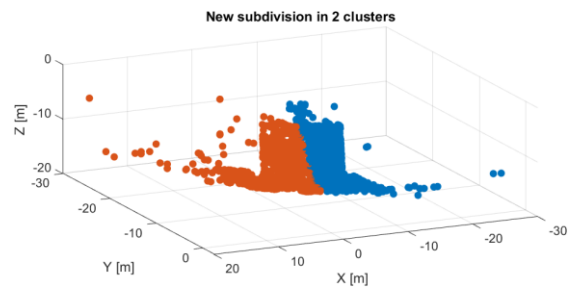
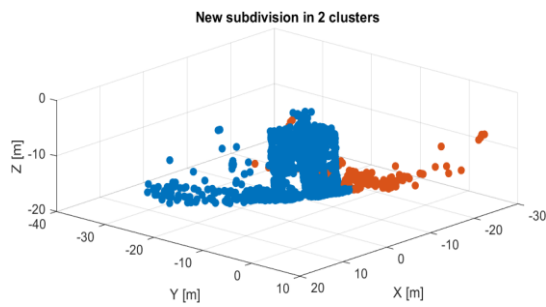
Looking at the chart, it is possible to notice that, in general, the trend of S_{avg_col} is reducing while the number of colour-clusters predetermined is increasing.

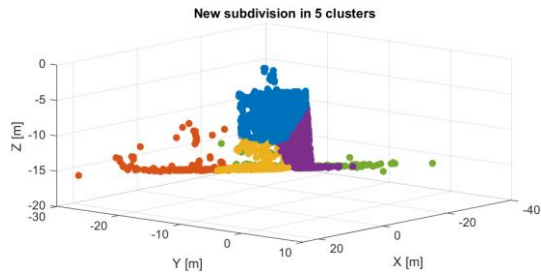
Average silhouette indexes values vary from 0.78, obtained considering only 2 colour-clusters and considered strongly unrealistic, to 0.52, with k -colour equal to 20.

For the later sub-clustering, only colour-clustering with k -values between 2 and 6 have been considered. In the following pictures, the best result is illustrated and all the considerations are presented.

6- STAGE 3: FAÇADE MODELLING

Color_clustering a				5	
Spatial_clustering b				6	
	1	2	3	4	5
1	0.00	0.00	0.00	0.00	0.00
2	0.72	0.49	0.67	0.93	0.51
3	0.51	0.44	0.67	0.35	0.56
4	0.61	0.47	0.49	0.52	0.55
5	0.65	0.46	0.60	0.48	0.58
6	0.67	0.46	0.64	0.52	0.50





By the analysis, it was possible to do some few considerations looking at the matrix.

Applying k-means twice reported many reliable silhouette index values, even higher than 0.60 in some cases. In one case (k_{colour} equal to 6 and k_{location} equal to 2, representing the possibility of dividing distant subgroups) 0.93 as S_{avg} has been reported, resulting the only extremely reliable value considered.

By the experience it was also possible to study how clustering with the same characteristics can be different. An example can be samples 12 and 22, where the optimal number of clusters recognized is 2.

While the first gives back S_{avg} equal to 0.93 and so the possibility of considering as reliable the process, the second one gave back S_{avg} equal to 0.49.

Sample 22 presents a points division that cut the point cloud into two similar parts, while sample 42 is able to separate the building from the rest of the space.

However, since the result obtained is far distant from what it was searched, the test cannot be considered a success.

At the basis of the failure there is the limit of the algorithm: since every point is studied with the closest performing 3D considerations, considering only one façade, points representing the third floor cannot be grouped together with the ones of the first floor without considering the other facades.

The problem would have been maybe solved applying a preliminary passage able to set the projections of the 3D space in a 2D environment, but in this case, k-means would have been useless for the scope proposed.

Another possible solution would have taken into account the possibility of scaling the building of an order of magnitude, but in this case, there was the risk of losing the principal position after the process and so perform inefficient thermal analysis, explained later.

In conclusion, since the dimensions of the point cloud, it would have been quite impossible to renounce at the random choose of only some illustrative points without losing in velocity of processing, which is necessary in all the project.

Since the façade modelling and its relative point cloud segmentation due to clustering process has been considered a failure, it is decided to move to other possible solutions such as fitting with MSAC.

6.3 - Literature review: MSAC

M-Estimator Sample Consensus (MSAC) is an upgrade of RANSAC (RANdom Sample Consensus), proposed by Fischler and Bolles.

It is a general parameter estimation approach designed to make a distinction of the data based on some fixed characteristics.

MSAC can be considered as an iterative resampling technique that generates candidate solutions by using the minimum number observations (data points) required to estimate the underlying model parameters.

Compared to conventional sampling techniques that use as much of the data as possible to obtain an initial solution and then proceed the research of the outliers, MSAC, like RANSAC, uses the smallest set possible and proceeds to enlarge this set with consistent data points.

At the end of the steps, it is possible to see how the datapoints are divided into two groups:

- Inliers: all the data grouped due to common characteristics;
- Outliers: what is presented as an anomaly;

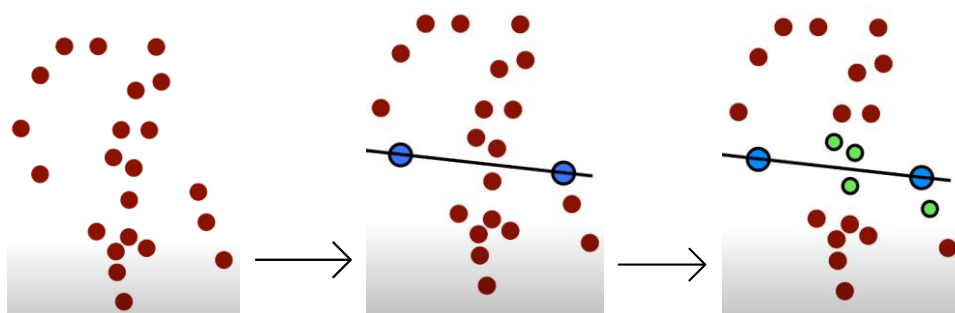
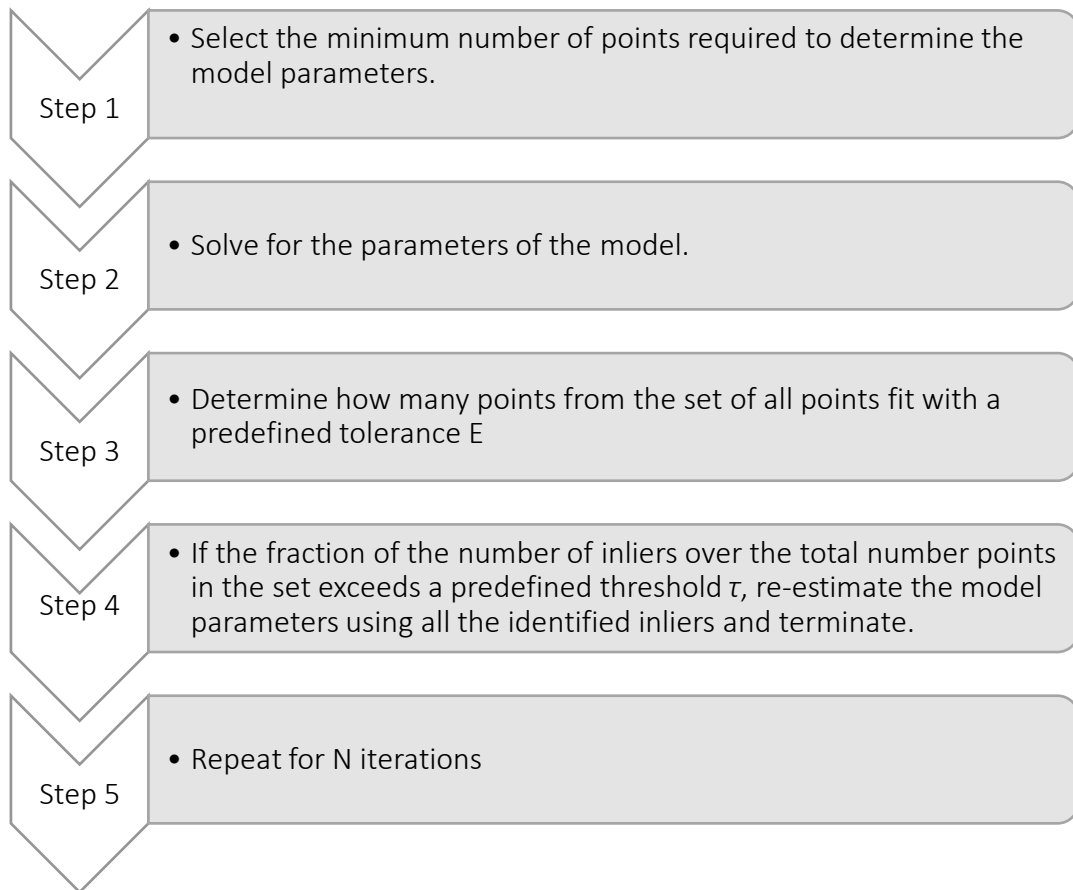


Figure 46 - An example of fitting analysis using RANSAC. In this case, it was asked to find the closest points to the line. In the third sequence, it is possible to see the division between the inliers and the outliers

It is possible to summarize the basic passages of the algorithm as it follows:



Where:

$N = \frac{\log(1-p)}{\log(1-(1-v)^m)}$ is the number of iterations. It is also set a maximum value in order not to exceed with the process. If the algorithm does not find the solution after N_{max} iterations, an error message is sent to the user who is invited to modify some parameters and try again;

- 3- p , which is usually set at 0.99, is the probability of success, identified as the possibility to find at least one sample with an outlier
- 4- u is the probability that any selected data is an inlier;
- 5- $v = 1 - u$ is the probability of observing an outlier;
- 6- m is the number of points;

In conclusion, comparing this parametrization with k-means, it is possible to notice that MSAC is extremely fast to process and uses even using huge number of samples.

The number of iterations required is extremely high and it is considered particularly manageable and modifiable.

On the other hand, all the samples have to be set in relation with a priori fixed characteristics, which are suitable for the project in object.

In the following chapter it is presented the approach adopted for modelling the point cloud and associate all the points at a particular case.

6.4 - Case study: MSAC and its application on Matlab

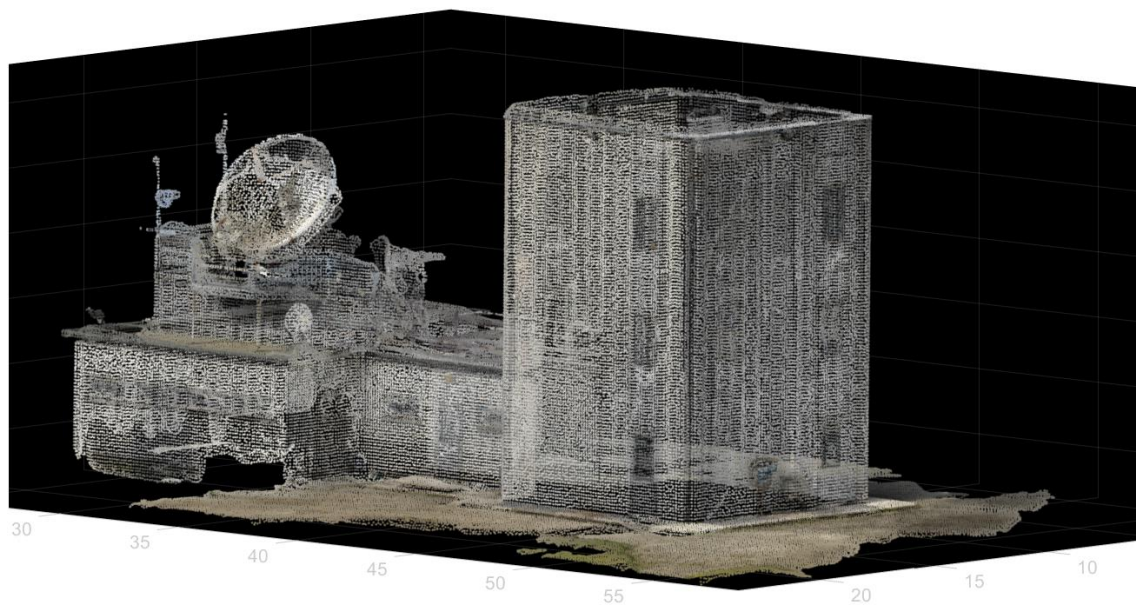
The topic of this subchapter is focused on the façade segmentation and object recognition of what is present in the point cloud.

As preliminary step, since working with punctual data is not interesting from an engineering point of view, it has been found a way for moving to values characteristics of a zone with selected dimensions which could describe the situation in a better way.

It is important to understand how to dimension the zone. In fact, areas with low dimensions are not interesting. On the contrary, areas with higher dimensions could compromise the reliability of the model averaging points that are parts of different parts and subjected to different behaviours.

For this test, it has been decided to divide the envelope into 10-by-10 cm² and, for each of them, make the average between all the colour values present inside it.

In this way, from 942660 points, 137594 zones are generated as it follows:



Step	Number of points
Before	942660 points
After	137594 points

Figure 47 - Point cloud after the reduction into zones

From a qualitative point of view, the point cloud has been divided starting from a spatial division to reach the identification of each single component.

Since we decided to perform all the analysis operations only on the main block, a first selective distinction needed to be performed.

For doing that, a fitting segmentation based on the parametrization of the Region of Interest is conducted, where ROI is obtained starting from the analysis of the spatial limits imported by Metashape and restricted in order to focus the research only on the interest part.

In the matrix above, the values used for the ROI searching are reported and the result obtained is shown. It is possible to notice that the site has been divided into two blocks.

$$ROI = \begin{bmatrix} X_{min} & Y_{min} & Z_{min} \\ X_{mas} & Y_{max} & Z_{max} \end{bmatrix} = \begin{bmatrix} 7 & 40 & 75 \\ 15.5 & 49 & 95 \end{bmatrix}$$

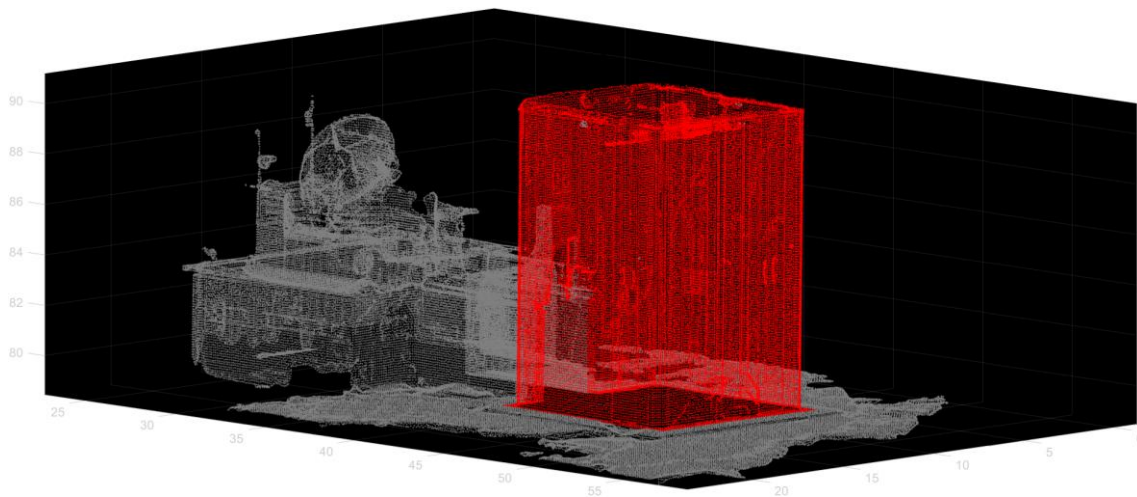


Figure 48 - Point cloud after the first division

The second step was related to the segmentation of the envelope into the single façade composing it. Even in this case, MSAC was used and, through the division into inliers and outliers based on some parameters, the result is obtained.

It has been tried to reach the goal by using two ways:

- working on a constantly reducing point cloud through the intermediate calculation of a new point cloud without the inliers;
- Working always with the same point cloud and, starting from it, proceeding with the plane fitting;

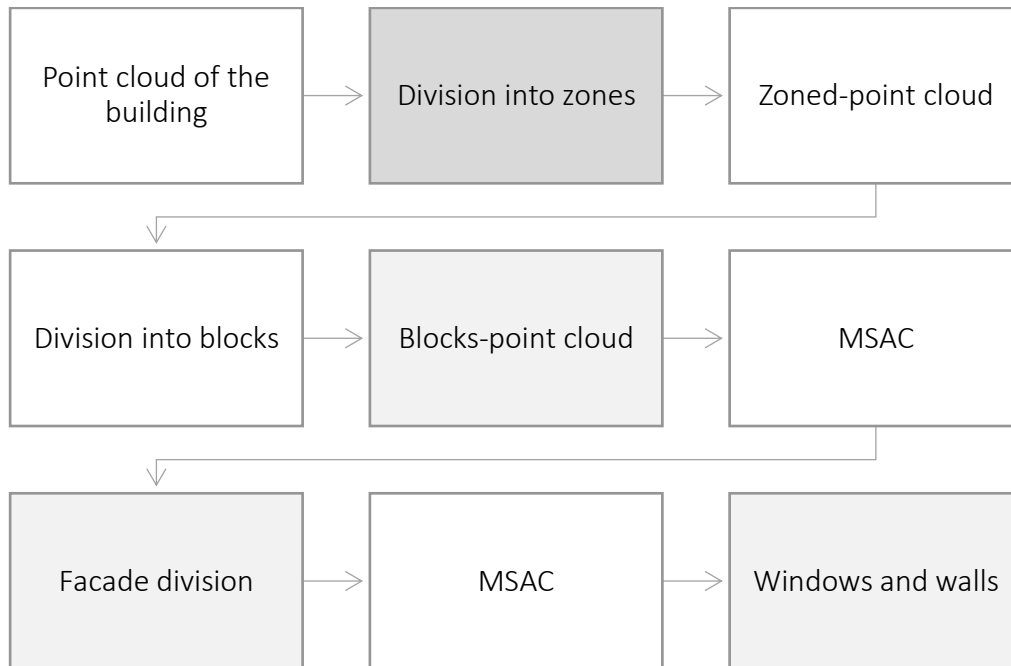
The second strategy resulted the best because of a series of simplifications that was possible to fix. First of all, despite the research of a precise parameter in terms of maximum distance from the centre, it has been possible to work with predetermined ROI that, in the second case, they would have changed due to the modification of the point cloud.

For the façade identification, the MSAC procedure considers for the inliers research:

- Maximum distance between points and plane fitting, set at 0.25 m;
- Maximum angle rotation between points and plane fitting, set at 1°;
- Reference vector;

- Region of Interest;
- Maximum number of iterations, set at 10000;
- Confidence factor, set at 0.99.

In the following tables the strategy adopted is illustrated and the parameters used for every plane fitting are showed.



South facade	
Maximum distance	0.25 m
Maximum angle	1°
Maximum number of iterations	10000
Confidence factor	0.99
Reference Vector	[1 0 0]
Region of Interest	$\begin{bmatrix} 7 & 40 & 75 \\ 7.5 & 49 & 95 \end{bmatrix}$
Number of inliers	11813

North facade	
Maximum distance	0.25 m
Maximum angle	1°
Maximum number of iterations	10000
Confidence factor	0.99
Reference Vector	[1 0 0]
Region of Interest	$\begin{bmatrix} 15 & 40 & 75 \\ 15.5 & 49 & 95 \end{bmatrix}$
Number of inliers	11259

West facade	
Maximum distance	0.25 m
Maximum angle	1°
Maximum number of iterations	10000
Confidence factor	0.99
Reference Vector	[0 1 0]
Region of Interest	$\begin{bmatrix} 7 & 40 & 75 \\ 15.5 & 40.5 & 95 \end{bmatrix}$
Number of inliers	10423

East facade	
Maximum distance	0.25 m
Maximum angle	1°
Maximum number of iterations	10000
Confidence factor	0.99
Reference Vector	[0 1 0]
Region of Interest	$\begin{bmatrix} 7 & 48.5 & 75 \\ 15.5 & 49 & 95 \end{bmatrix}$
Number of inliers	11987

Starting from the inliers already found out, using the ROI accurately selected, the identification of the windows in the façade is figured out.

This operation is important because it allows to select the proper thermal analysis tool, described in the following chapters.

Name	Window 0101
Facade	South
ROI	$\begin{bmatrix} 7 & 41.5 & 83 \\ 7.5 & 43 & 85 \end{bmatrix}$
Inliers	424

Name	Window0102
Facade	South
ROI	$\begin{bmatrix} 7 & 41.5 & 87 \\ 7.5 & 43 & 89 \end{bmatrix}$
Inliers	341

Name	Window0201
Facade	North
ROI	$\begin{bmatrix} 15 & 46 & 83 \\ 15.5 & 47.5 & 85 \end{bmatrix}$
Inliers	429

6- STAGE 3: FAÇADE MODELLING

Name	Window0202
Facade	North
ROI	$\begin{bmatrix} 15 & 46 & 87 \\ 15.5 & 47.5 & 89 \end{bmatrix}$
Inliers	307

Name	Window0301
Facade	West
ROI	$\begin{bmatrix} 12 & 40 & 83 \\ 13.5 & 40.5 & 85 \end{bmatrix}$
Inliers	339

Name	Window0401
Facade	East
ROI	$\begin{bmatrix} 7.5 & 48.5 & 83 \\ 9 & 49 & 85 \end{bmatrix}$
Inliers	316

Name	Window0402
Facade	East
ROI	$\begin{bmatrix} 7.5 & 48.5 & 87 \\ 9 & 49 & 89 \end{bmatrix}$
Inliers	335

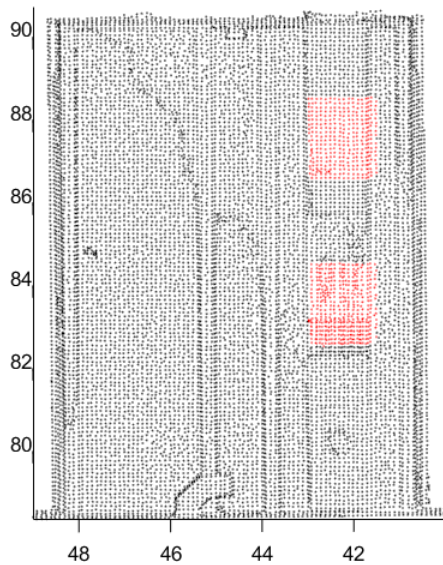


Figure 49- South façade

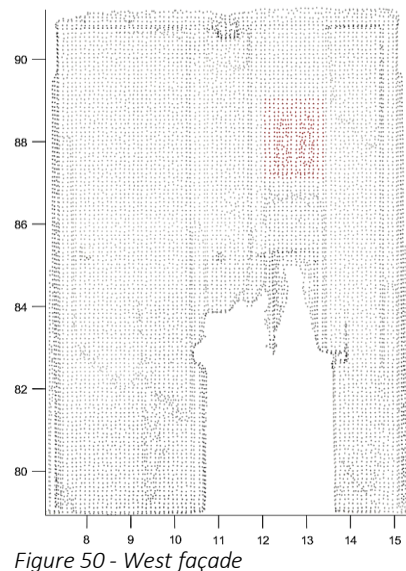


Figure 50 - West façade

6- STAGE 3: FAÇADE MODELLING

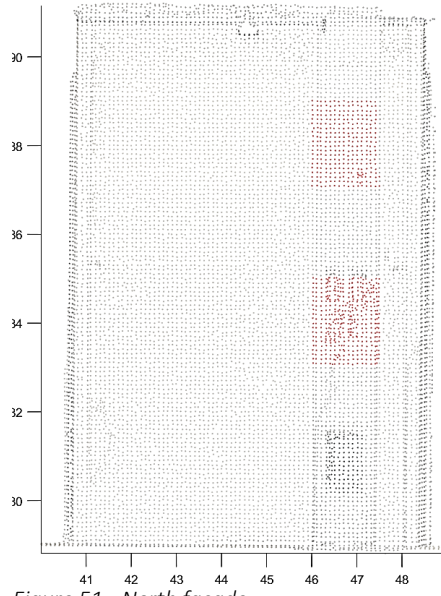


Figure 51 - North façade

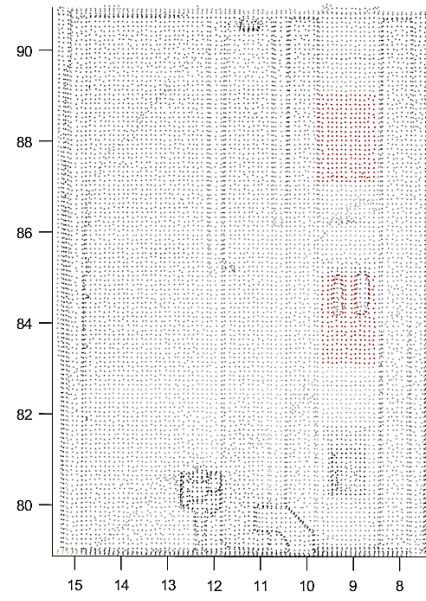


Figure 52 - East façade

7 - Stage 4: data analysis

7.1 - Goal and strategy

The last aspect to design and study is concerned on the analysis of the result acquired and rebuilt by the drone.

The colour-information coming from the thermography is read by a script and converted into temperature considering the colour scale used by the thermo-camera using an approach that, starting from the RGB code extracted from ironbow palette, that is the typical colour palette used for thermography, and assigning a fixed value based on the two extremes, the external temperature given by the thermo-camera is found.

This passage is necessary because it allows the script to translate what a thermo-camera is doing moving from the main matrix n -by- 6 to a n -by- 4 , where to each micro-zone is assigned a temperature value.

For the analysis, the procedure followed is a variation of the one proposed by Rakha et al. for the evaluation of the real U-value. In their experience, they considered the external surface temperature after a daily monitoring the envelope of their case study gathered by a drone survey and they compared it with a value coming from an analytical procedure for finding an uncertainty. By fixing value of T internal and external, a real U-value is estimated, so as energy saving potentials and suitable retrofitting strategies are performed.

In the study, internal energy loads were not considered as well as the contribution of sun radiation.

This could be a problem because, as mentioned by Melis et al, the contribution of the sun can affect the result of the radiative apparent temperature, and the external surface temperature as well, of $4K$, that is a significant difference.

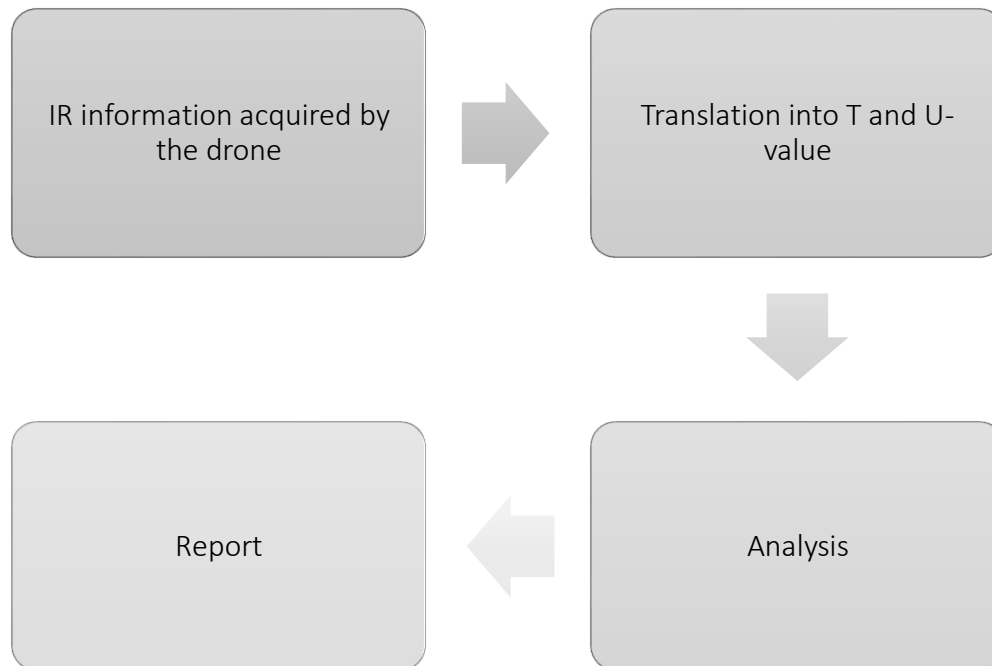
For each zone, the IR information is translated into a T value and compared with a model that calculates a temperature value at the same conditions.

In particular, this operation is finalized at helping the drone understand immediately where the problem is.

The analysis involved are, as it follows:

- Quantitative analysis of the external surface temperature measured;
- Quantitative analysis of the U-value estimated;
- Qualitative understanding of the type of thermal defect involved

For each of these studies, a Matlab script is generated in order to put the basis for an automatic analysis directly performed by a drone during a future upgrade.



7.2 - Reconstruction of the external temperature from analysis

As mentioned before, the colour, using the ironbow palette, is converted in temperature. This passage has been performed making a first grayscale conversion of the texture, where black is associated to the lowest temperature and white to the highest measured, and the T_{rad} value of each point is found out by linear interpolation between the two range extremes.

In this case, the range used for measuring the temperature is included between 5°C (black) and 41°C (white).

It is now studied how the IR camera converts the radiation into Trad.

The interpretation of thermal data is based on the comprehension of the physical processes enabling the energy transfer, such as radiation.

Since every object at a temperature higher than 0 K continuously emit electromagnetic radiation, the sensors designed for measuring the thermal radiation can record the TIR radiation emitted by surfaces and absorbed mainly in form of radiation coming from the Sun. This radiant energy results from the conversion of the internal kinetic energy at molecular level, that is measured by the radiant flux.

This last quantity is regulated by the Stefan–Boltzmann law, that states the radiant flux density emitted by an object proportional to the fourth power of the object's surface temperature, fixing that a hotter object radiates much more energy than a cooler one at every emission.

According to Wien's Displacement Law, the maximum peak wavelength at which a black body emits is inversely proportional to the kinetic temperature, as it follows:

$$\lambda_m = \frac{A}{T_{kin}}$$

Where:

- λ_m is the maximum wavelength at which the emission is present;
- A is the constant of Wien, which is equal to 2898.76 μmK ;
- T_{kin} is the absolute kinetic temperature;

In general, the higher is the emittance temperature, the higher is the radiation.

The spectral radiance can be calculated according to Planck's function, considering the spectral radiance emitted by a blackbody at a temperature T and wavelength λ can be calculated as:

$$L_\lambda(T) = \frac{C_1}{\lambda^5 \left(\exp\left(\frac{C_2}{\lambda T}\right) - 1 \right)}$$

Where:

- C_1 and C_2 are constants, equal to $3.7413 \cdot 10^8 \text{ W m}^{-2} \mu\text{m}^4$ and $1.4388 \cdot 10^4 \mu\text{m K}$, respectively;

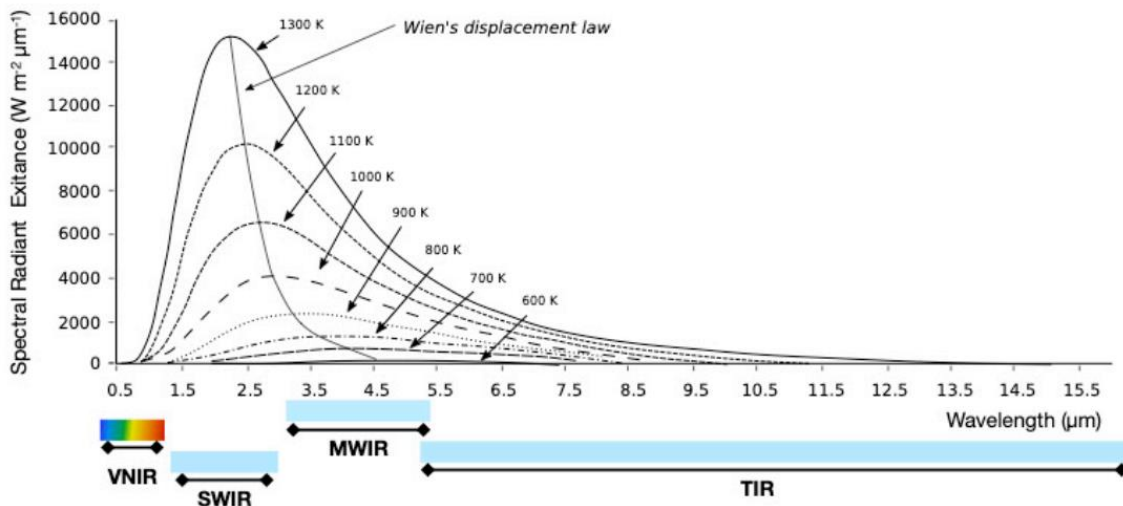


Figure 53 - Spectrum of the electromagnetic wave

All the equations introduced until now work with the idealization of the object into a blackbody, so it is considered that all the incident radiation is totally absorbed.

In the reality, part of the energy is reflected depending by the wavelength, the temperature and the type of surface. So, it is introduced the emissivity ϵ as the ratio between the radiative flux measured and the one of a black body at the same T:

$$\epsilon = \frac{Fr}{Fbb}$$

that is lower than one for the gray bodies and works under the following relation:

$$\alpha(\lambda) + \rho(\lambda) + \epsilon(\lambda) = 1$$

So, in conclusion, the measured radiative temperature has to be considered in the following relation for finding the kinetic temperature:

$$T_{rad} = \epsilon^{0.25} * T_{kin} * (\epsilon(\lambda) + \rho(\lambda))$$

Where:

- $\epsilon(\lambda)$ e $\rho(\lambda)$ are the emissivity and reflectivity coefficients related to the wavelength and dependent with the surface geometry and temperature.

Their values are measured experimentally on the field through a hypersensitive spectrometers.

Nowadays, all the thermal sensors implemented in the camera are equipped with a spectrometer, included the FLIR used in the drone, that are able to estimate the emissivity and the reflectivity and correct automatically the temperature measured.

If we are interest in the real reflective apparent temperature, before the data acquisition phase, it is necessary to set the emissivity equal to 1 and make all the considerations.

It is important to mention that the emissivity has to be measured on field using particular techniques, but, since this is an approach studied for a drone-based technique, tabular values are used.

However, many of the materials used for opaque traditional facades have higher values of emissivity, within the range of 0.8 and 0.95 and low dependency on temperature variation. So, it can be possible not to consider for building energy audit this last aspect, so tabular values of emissivity related to standard temperature conditions are sufficient and considered for the whole analysis (Riminesi, 2018).

It is important to mention that, by using this approach, we are considering an error but for construction monitoring and inspection aspects it can be neglected due to its low interference.

If the opaque façade would have been composed by metallic materials, the emissivity should have been monitored experimentally.

So, in conclusion the measured value of temperature has to be reduced by a constant factor before being analysed.

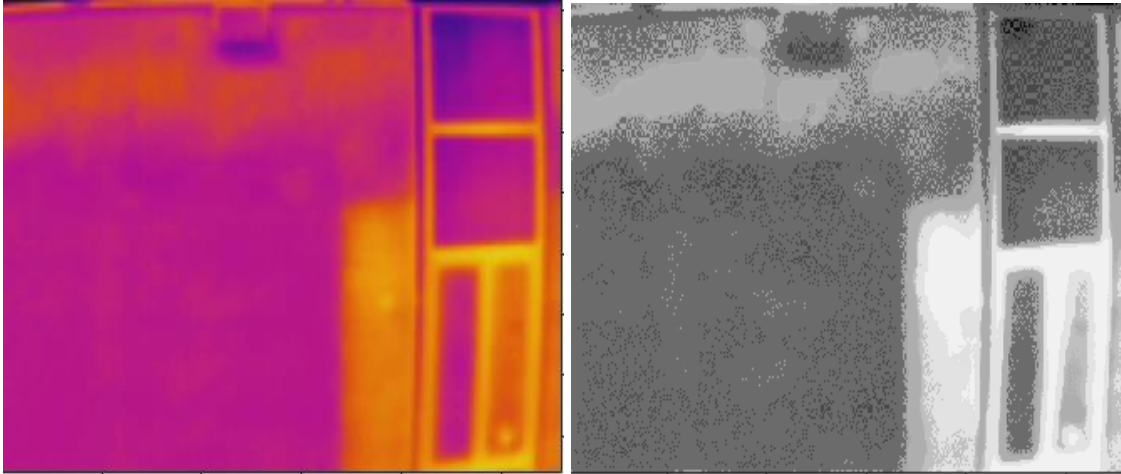


Figure 54 - Example of conversion from colour to temperature

7.3 - Reconstruction of the external surface temperature of control

It is now explained how it has been reconstructed a reliable tool that was able to let the drone reading the result and perform the analysis on it.

It is possible to notice that actually no tools with those capabilities are present in the market and, for this reason, a dedicate script had to be programmed for the mission.

Moreover, this research does not want to add new notions to the theoretical aspects, that are already well explained by two decades, but it has as main goal the generation of a reading instrument that is based on the already mentioned aspects and is the base for a future automation of the entire aspect.

In any case, even if the mission goes in the direction of the automation, part of the preliminary information about position and composition of the building has to be set manually by the user.

The strategy is involved in talking the same language of the software, so the control value is reconstructed like in a virtual environment, by analysing all the possible components that could have influence on the object.

For reconstructing the external temperature acting on a surface, a thermal balance between the building and the environment around is studied by summing the following contributions:

$$q_{a_{sw}} + q_{LWR} + q_{conv} + q_{cond} = 0$$

Where:

- $q_{a_{sw}}$ is the quantity of energy exchanged through shortwave radiation;

- q_{LWR} is the quantity of energy exchanged through longwave radiation. In this case, due to the morphology of the site, only air and ground were considered as exchanging bodies;
- q_{conv} is the quantity of energy exchanged by convection between the building and the air;
- q_{cond} is the quantity of energy exchanged through convection from the internal side of the building.

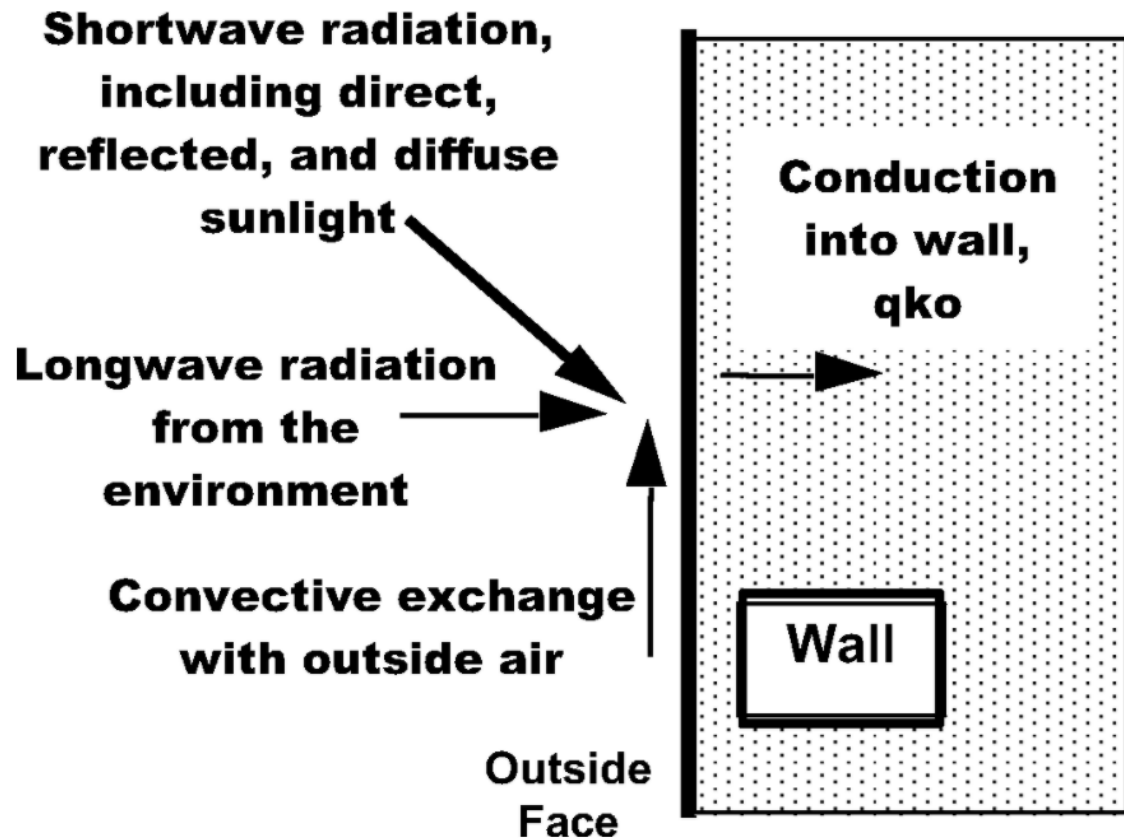


Figure 55 - Thermal balance for outdoor external surface temperature

Each component is calculated considering the data measured during the acquisition phase and the external temperature is found out. In detail, they are calculated as it follows.

7.3.1 - Longwave radiation

The longwave radiation is the part of heat exchanged by the building with the external through radiation.

The model used for the measurement is a standard analytical equation installed also in software like EnergyPlus. Due to the simple geometry of the building, some reliable assumptions can be set.

For the façade, the building exchanges through longwave radiation mainly with three actors: air, ground and sky; while for the roof the same exchange does not take into account the heat transfer through the ground.

In general:

$$\begin{aligned}
 q_{LWR} &= q_{Ground} + q_{Sky} + q_{air} \\
 &= \frac{\varepsilon * \sigma * F_{gnd} * (T_{se}^4 - T_{air}^4)}{T_{se} - T_{air}} * (T_{air} - T_{se}) \\
 &+ \frac{\varepsilon * \sigma * F_{sky} * B * (T_{se}^4 - T_{sky}^4)}{T_{se} - T_{sky}} * (T_{sky} - T_{se}) \\
 &+ \frac{\varepsilon * \sigma * F_{sky} * (1 - B) * (T_{se}^4 - T_{air}^4)}{T_{se} - T_{air}} * (T_{air} - T_{se})
 \end{aligned}$$

Where:

	Input	Equation	Value	Unit of measure
ε	Emissivity	-	Variable	-
σ	Boltzmann constant	-	$5.67 * 10^{-8}$	W/mq*K ⁴
Fgnd	Ground view factor	$0.5 * (1 - \cos\phi)$	0.50	-
Fsky	Sky view factor	$0.5 * (1 + \cos\phi)$	0.50	-
B	Ground-sky correlation factor	$\text{Sqrt}(0.5 * (1 + \cos\phi))$	0.707	-

Table 13 - Values used as input for longwave radiation

The temperature of the air is calculated and considered following the procedure presented by the US Standard (2007), that put in relation the temperature obtained from the meteorological station with the other factors such as the altitude and the air gradient.

Everything is calculated as in the following way:

$$T_{air} = T_b + L(H_z - H_b)$$

Where H_z is found out by considering: $H_z = \frac{E * z}{E + z}$

	Input	Value	Udm
Tb	T_air ground level	7	C
L	Air gradient	-0.00065	K/m
H _z	Geopotential altitude	-	m
H _b	Offset altitude	0	m
E	Earth radius	6356	Km
z	Altitude	84	m

Table 14 - Values used as input for air temperature calculation

A simplification has been set by considering the altitude equal for all the zones instead of measuring the real one for every square.

After the calculation, it has been noticed that the variation between the calculated value and the one collected for meteorological station differentiated about 0.0004 °C, due by the low-rise building, so it was possible to consider the second one for all the analysis.

7.3.2 - Shortwave radiation

The presence of the shortwave radiation is considered one of the most important influencing factors for a thermal balance. For this reason, since the external temperature of a surface under direct irradiation is subjected to important variations in the time, it is necessary to plan the acquisition phase in the part of the day when the sun radiation is minimal.

In addition, the radiation coming from the sun is influenced by many aspects such as the meteorological conditions, the location with the presence of some geometrical obstacles and the pollution level measured.

At a design level, values according to predetermined models can be used for obtaining an average daily value closed to the worst condition (clear sky at different orientations). In this case, it was not possible to adopt this solution because it was necessary to have the real value simulating the conditions measured during the acquisition phase.

The use of a pyranometer could be necessary for reaching the most reliable irradiation value, but it is not possible to bring the instrument in the site.

The atmosphere is then modelled considering the application of MODTRAN (Moderate resolution atmospheric transition), that is a software used for measurement through the atmosphere worldwide by researchers and agencies.

The code is designed for setting all the relevant parameters in order to model the atmosphere as a unique homogeneous layer and provide radiance and transmittance values useful for the scopes.

It has been tested that the influence of the radiation coming from the sun can create differences in the results even higher than 4 K.

Using MODTRAN, this gap can be reduced until 1K, that is an acceptable deviation for our purposes (Alba, Barazzetti M., & Rosina, 2011).

In this case, the part of shortwave radiation was calculated by applying the value coming from MODTRAN, accurately studied using other design model so as to find a reliable correspondence between them.

In the table showed below all the parameters inserted and used as input are illustrated as well as the result obtained.

It is important to mention that temperature, water column and ozone input required by the code were calculated considering the latitude where the site is in relation with the altitude.

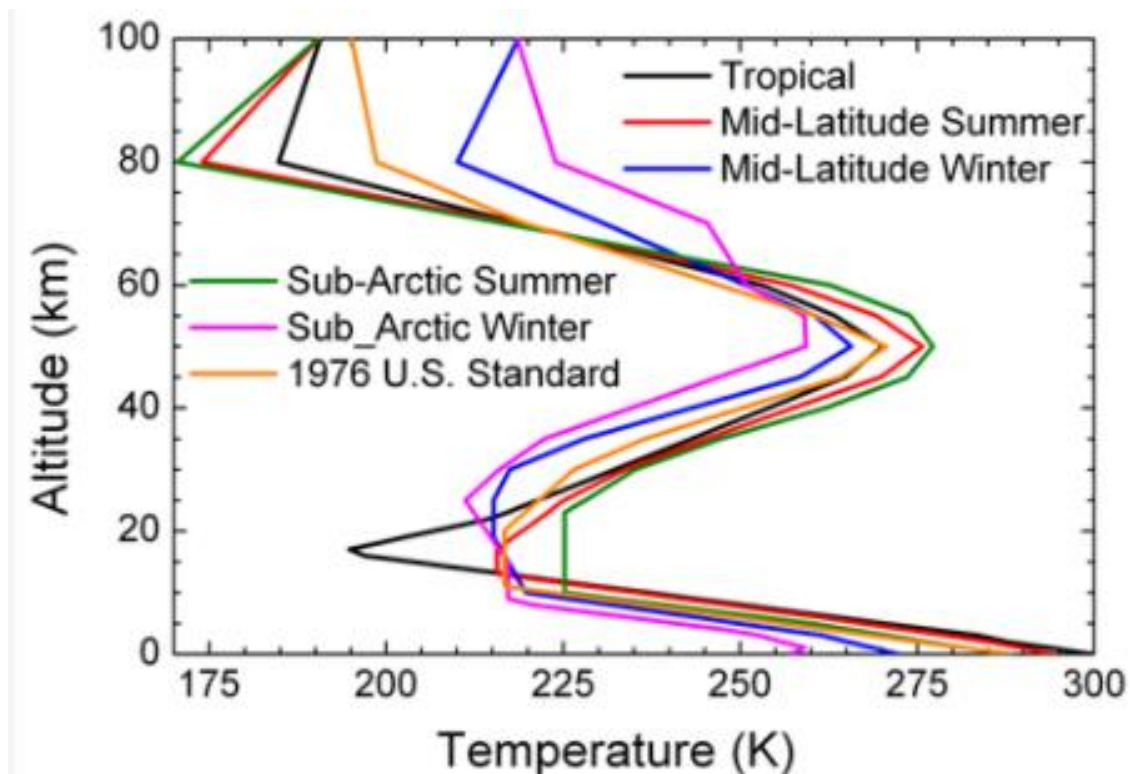


Chart 14 - MODTRAN: Temperature chart in relation with latitude and altitude

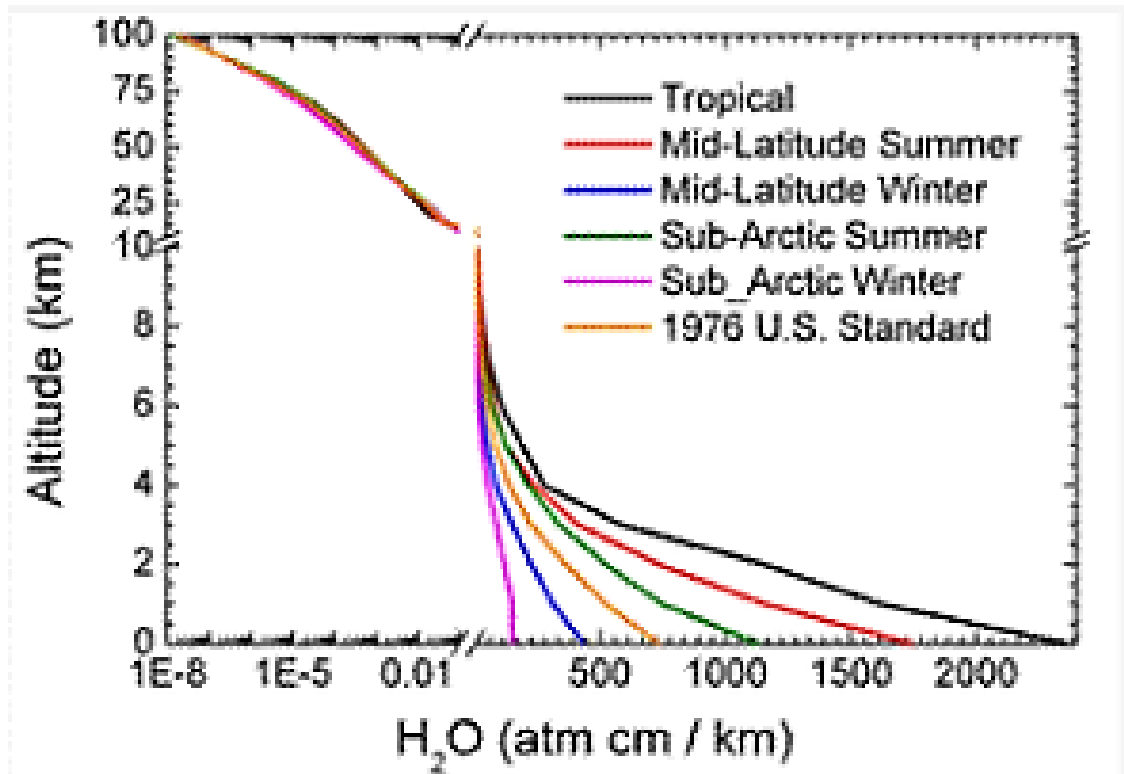


Chart 15 - MODTRAN: water pressure in relation with latitude and altitude

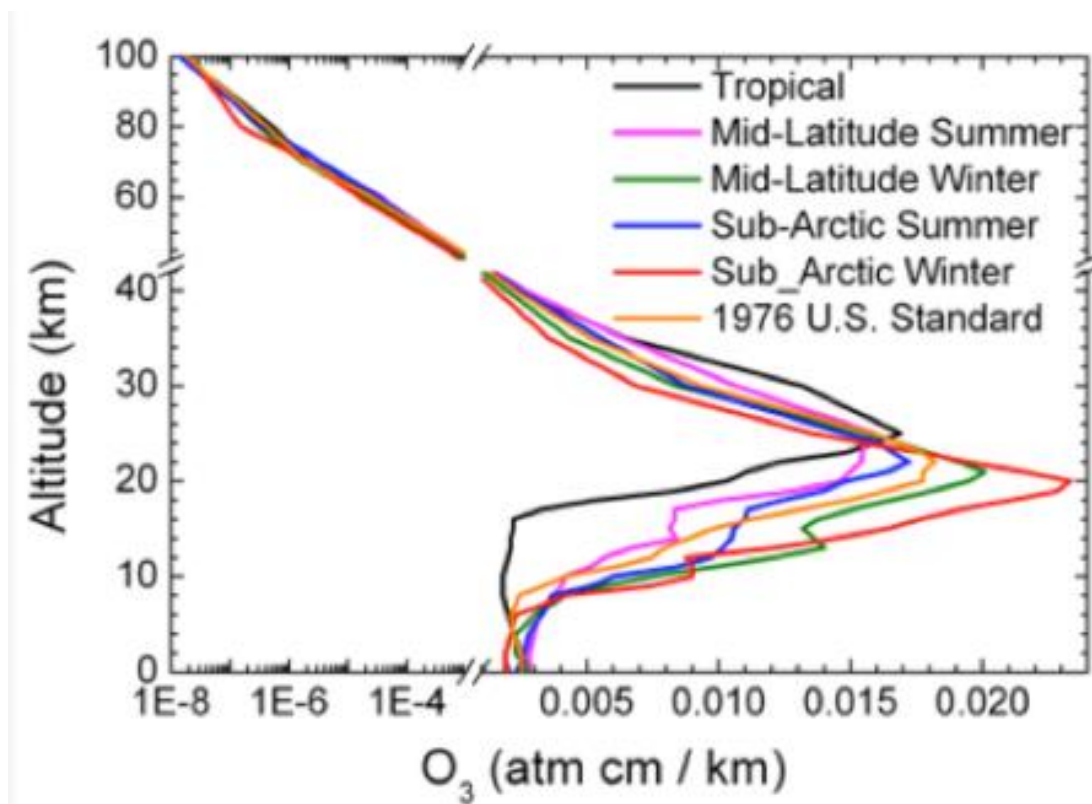


Chart 16 - MODTRAN: Ozone quantity in relation with latitude and altitude

Input	Value	Unit of measure
Atmospheric model	Mid-latitude winter	-
Water column	500	Atm/cm
O3 column	0.004	Atm/cm
CO2	400	ppm
CO	0.15	Ppm
CH4	1.8	ppm

Input	Value	Unit of measure
T_ground	10	C
Ground albedo	0.25	-
Aerosol model	Rural	-
High of the sensor	Range 50 km	-
Zenith sensor	Range 180°	-

Table 15 - Values used as input for the calculation

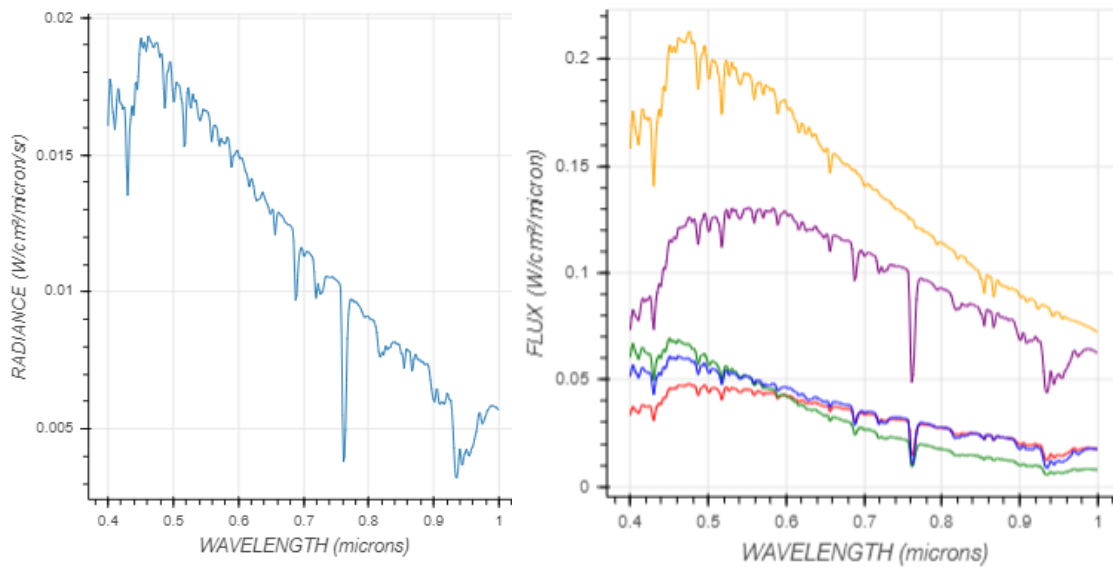


Chart 17 - Distribution chart of radiance and suggested flux per each wavelength considered

As suggested by the code, a flux incident to the wall equal to 140 W/mq was suitable for our case.

7.3.3 - Convection

The part of heat transfer through convection considers the relation between the building and the air.

For this part, the flux was calculated analytically, while the convective heat transfer coefficient is figured out through experimental equations.

$$q_{conv} = h_c(T_{ext} - T_{air})$$

After a research, it was figured out that 10 different ways for finding h were found and each of them considered different characteristics and were applied to some selected cases.

TARP algorithm (Thermal Analysis Research Program) proposed by Walton, linearized the convective coefficient as the sum of a natural and a forced component, where the second one was influenced by the characteristics of the surface.

Another important linearization was proposed by Yadzanian and Kleems, called MoWitt (Mobile Window Thermal Test), correlated the heat transfer convection with air velocity and difference of temperature between the body and the fluid.

However, since this correlation was considered extremely reliable, it was possible to test it only for low-rise and smooth buildings.

In general, other correlations were studied and tried to set several factors.

For our scope and purposes, it was decided to adopt the correlation suggested by ISO6946 called Simple Combined correlation.

In this case, the convective coefficient is figured out through the combination of the air velocity and the characteristics of the external layer, such as:

$$h = D + E * V_z + F * V_z^2$$

Where:

Surface type	D	E	F
External plaster	10.22	3.1	0.0
Impermeabilization	10.79	4.192	0.0

Table 16 - Value used as input for convective heat transfer coefficient

V_z , which is the velocity of the air is figured out analytically from the following equation.

All the values were calculated by using the one collected during the acquisition phase by the closest meteorological station.

$$V_z = V_{met} * \left(\left(\frac{d_{met}}{z_{met}} \right) \exp(a_{met}) \right) * \left(\frac{z}{d} \right) \exp(a)$$

	Input	Value	Udm
Vmet	W.S meteo station	2.46	m/s
dmet	Air gradient coefficient	270	-
amet	Meteo station profile coefficient	0.14	-
a	Site profile coefficient	0.14	-
d	Boundary velocity coefficient	270	-
z	Altitude	84	m

Table 17 - Values used as input for wind velocity calculation

7.3.4 - Conduction

The last part composing the thermal balance is about the heat transfer through conduction between the internal and the external side of the wall.

As mentioned before, following the same procedure adopted by Rakha, an internal fixed temperature has been set without considering any internal loads variation for simplification.

In this case, due to the characteristics of the heat generator, it was possible to set that the machine used to work until the internal temperature reaches 20°C and then the regulation system stops immediately the heat insertion measuring the internal temperature through a system of perception and regulation.

The stratigraphy of each component plays an important role. For this reason, the envelope has been divided into the technical elements and the analysis has been separated.

Since it was impossible to make some invasive analysis for discovering the layers composing the envelope, they were hypothesized starting from the year of construction of the building and then using as model the typical way of building.

Since the office centre of Spino d'Adda was realized around 1970s, the external wall considers the presence of 4 cm of air gap through two layers of bricks, while the roof is calculated considering the presence of 10 cm of thermal insulation and an impermeabilization layer, as it is possible to see from the picture.

In the following scheme the behaviour of the heat transfer is presented and the values used for the stratigraphy are summarized in the tables.

In general, the heat transfer through conduction is calculated as it follows:

$$q_{cond} = U(T_{int} - T_{ext})$$

Where:

- U, the thermal transmittance, is calculated as it follows:

$$U = \frac{1}{R} = \frac{1}{\sum \frac{s_i}{\lambda_i}}$$

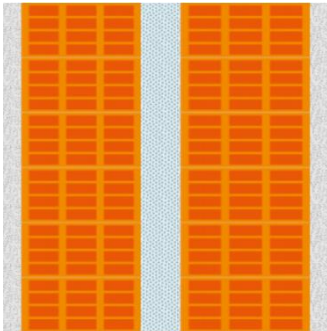
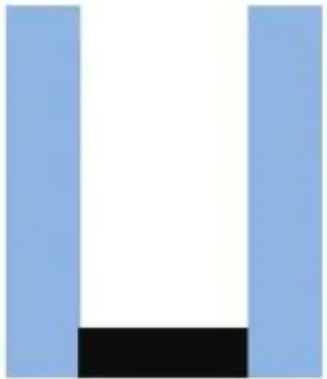
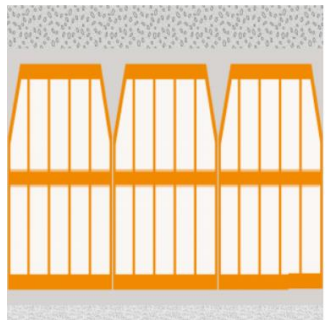
Type	External wall	
Composition	Double brick with air gap	
Thickness	32 cm	
U-value	1.1 ± 0.2 W/mqK	
Type	Window	
Composition	Double clear float with argon gap and opaque aluminium structure	
Thickness	2.4 cm	
U-value	2.5 ± 0.2 W/mqK	
Type	Roof	
Composition	Masonry slab with impermeabilization layer	
Thickness	38 cm	
U-value	0.53 ± 0.2 W/mqK	

Table 18 - Stratigrafies of the principal elements composing the envelope

The thermal balance already presented can be solved as an equation with only one unknown. The external surface temperature is found for all the cases presented and used as control value.

In the next chapter, T_{se} is compared with the measured value of external surface temperature and the considerations are done.

By the analysis of the delta generated, it is possible to understand the entity of the potential damage and establish the priority of intervention.

7.4 - Analysis of the result

In this subchapter, the results obtained during the tests are presented and some considerations are mentioned.

First, because of the necessity of testing the proper working of the analysing code, all the trainings have been performed manually.

7.4.1 – Qualitative analysis

By using this type of analysis, it was possible to search and check potential failures inside the envelope just looking at colour difference between the parts.

It was possible to notice different areas considered potentially interesting from an analytical point of view. The pictures show the entity of the potential defect.

In the east façade, it was possible to notice an important difference in terms of temperature in the areas closed to the windows. It was supposed that in that part, during some renovation works, the length of the wall was reduced in order to install the windows.

However, a continue violet line running along all the perimeter of each single window demonstrates that no air leakage is present.

The attention was even focused on the corners of the envelope, where the possibility of thermal bridges can occur due to the bi-axial heat transfer.

In general, this checking was made also in correspondence of the columns and slabs, where the difference between the construction materials would have created a thermal bridge.

None of them presented remarkable problems, except in the south-east corner, where the downpipe ran along the envelope.

In that area, the difference of colour, a bit clear than how it should be, suggested the presence of a thermal bridge. On the contrary, it was also supposed that the metallic downpipe influenced the analysis with heat transfer phenomena of conduction.

Looking at the top of the envelope, an important thermal defect is measured in correspondence with the roof. In fact, it was possible to notice that a general thermal

bridge involved in that part of the building, causing potential problems in the internal side of the offices.

Another important thermal defect was measured on the roof, in correspondence with the antenna.

In fact, it was possible to notice that part of the roof presented a huge difference of temperature compared to the rest of the element. The form and the gravity of the thermal anomaly propended the diagnosis for a general failure in the impermeabilization layer.

Due to water leakage, that stationed there after rains due to the inclination of the roof, part of the water has maybe gone into the stratigraphy, affecting the rest of the elements and reducing their performances in terms of thermal insulation.

Another thermal defect that could be seen in at the bottom of the envelope, between the sidewalk and the opaque east façade.

The presence of a short part where the temperature measured is much higher than the rest of the envelope suggested that there is a thermal bridge caused by a possible problem in the realization phase.

Looking at the windows, doors and transparent parts, it was possible to notice the colours closed to orange for all the struts. It suggested that the material adopted for realizing them is a metallic alloy such as aluminium. On the other hand, since those materials tend to reflect a bit the light and create false positive images, it is not possible to assert a problem in the insulation cameras inside the windows and the door.

In the following images, all the defects described before are mentioned and a summary table is created.

It is important to mention that all these preliminary hypotheses are made by analysing the photos, but it could be interesting to implement a sort of image recognition for thermal defects.

A first attempt was tested directly on the IR pictures, trying to ask to a Matlab code to recognize the typical defect performing the steps required based on a pre-uploaded comparative sample of potential defects.

The test did not reported the expected results but it could be interesting to perform a focusing into this argument and in general, on the machine learning applied to building construction field.

All the results performed in the qualitative analysis are then tested again in a quantitative analysis, with the purpose of extracting the external surface temperature measured and a variation in the transmittance of the façade.

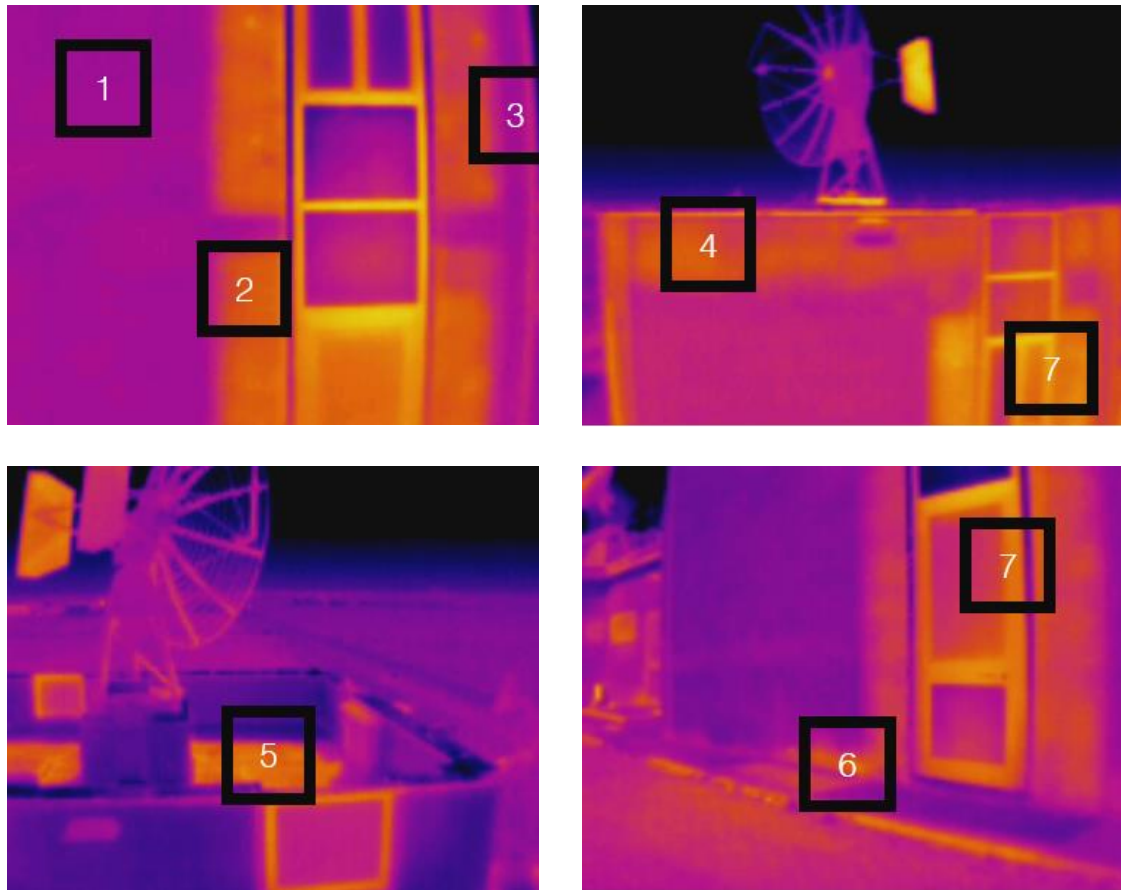


Figure 56 - Localized defect in the envelope

Defect	Estimated defect	Possible cause
1	General lower decay in the opaque part	Negligible time decay
2	Localized thermal bridge at opaque part	Defect of insulation in the wall
3	Thermal bridge	Non countable
4	Thermal bridge	Failure in the impermeabilization
5	Water leakage	Failure in the impermeabilization
6	Thermal bridge	Possible water rising
7	Possible high heat transfer	Non easily countable

Table 19 - Preliminary estimation of the cause of the failure

7.4.2 – Quantitative analysis

It was performed a quantitative analysis of the thermal defects inspected before with the purpose of evaluating level of damage and propose a proper maintenance intervention.

As mentioned before, using the control model properly realized for the case study, it is evaluated the entity of the damage through the external surface temperature.

Two values are extracted locally: the peak of temperature measured and the most diffused one. Then, a comparison with the value coming from the analysis is performed and the variation is calculated.

Finally, it was decided to substitute in the thermal balance the values obtained for figuring out the real value of transmittance operating, representing it in a range that includes both the most diffused cases and the worst one.

In the end, looking at the variations obtained, it was possible to evaluate the entity of the damage and scheduling a priority of intervention.

A comparison with the minimum values of transmittance required by the national code has also been presented.

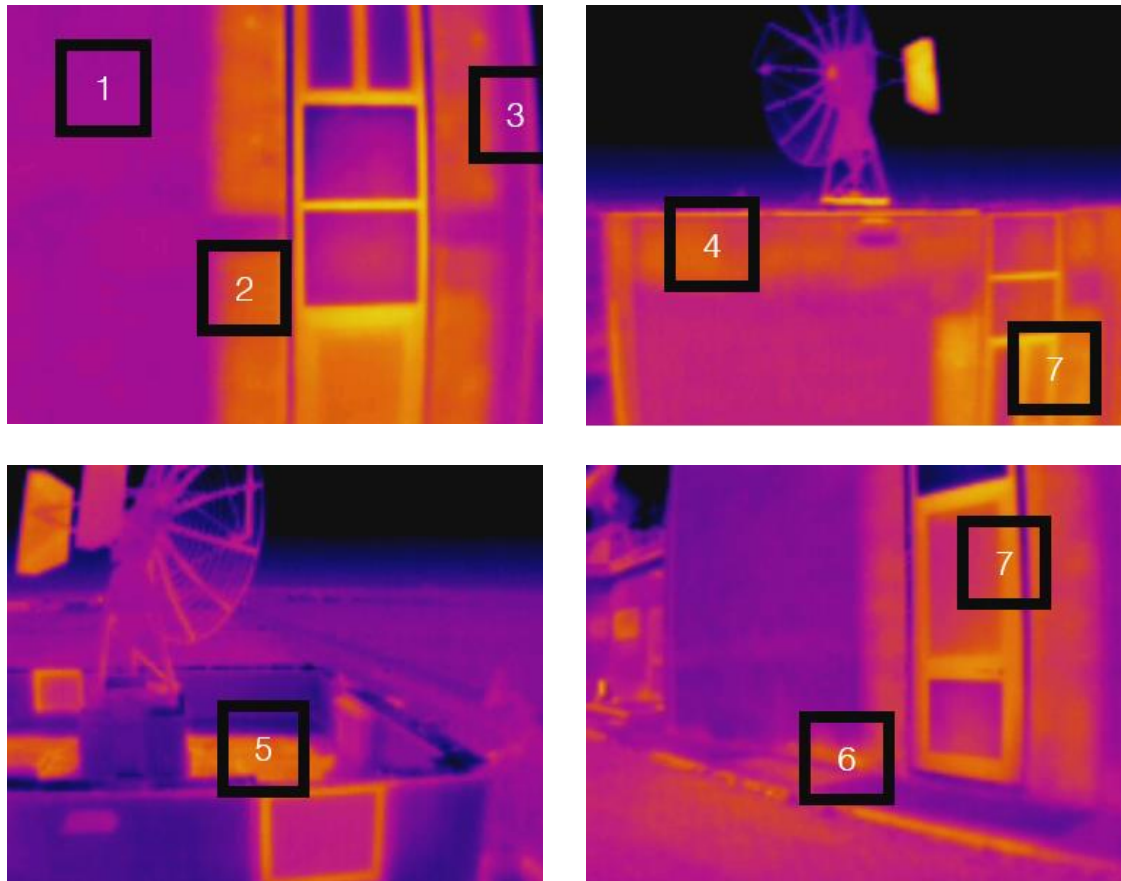


Figure 57 - Anomalies studied in the envelope

	T _{se_ctrl}	T _{se_diffused}	Variation (%)	T _{se_peak}	Variation (%)
Test 1	12.84	12.32	-4%	11.32	-4%
Test 2	12.84	15.26	16%	16.32	21%
Test 3	12.84	15.88	19%	15.88	19%
Test 4	12.84	15.20	16%	15.2	16%
Test 5	13.58	16.55	18%	17.24	21%
Test 6	12.84	14.77	13%	15.20	16%

Table 20 - Results obtained after the quantitative analysis for external surface temperature

7 - STAGE 4: DATA ANALYSIS

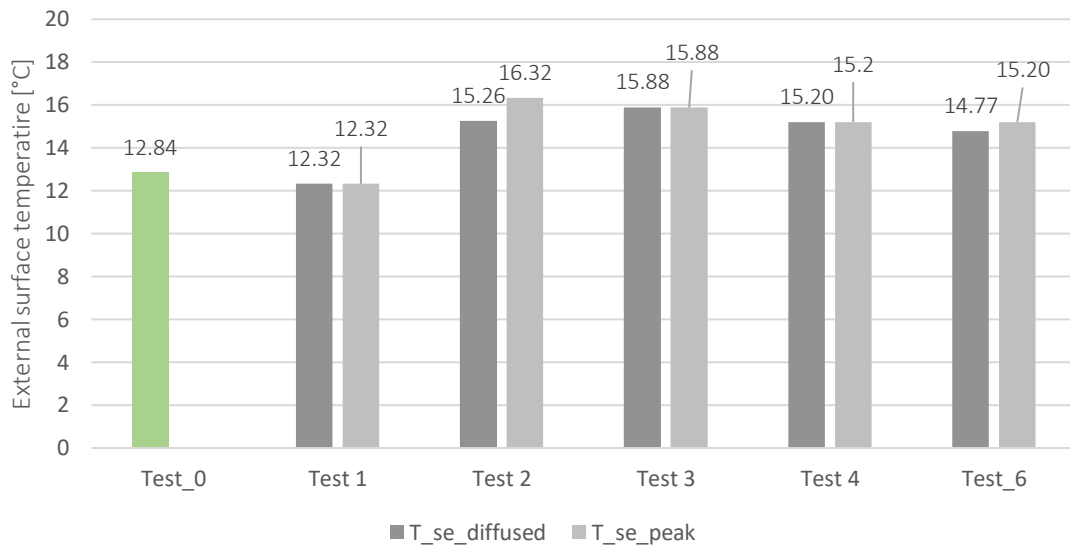


Chart 18 - Results obtained after the quantitative analysis for external surface temperature: general tests

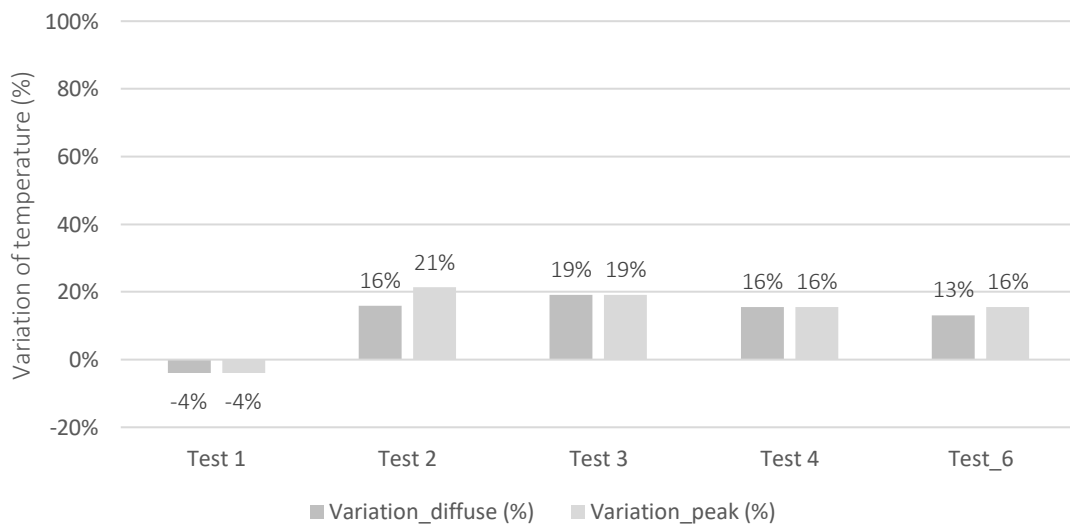


Chart 19 - Results obtained after the quantitative analysis for external surface temperature: general tests (variation in %)

7 - STAGE 4: DATA ANALYSIS

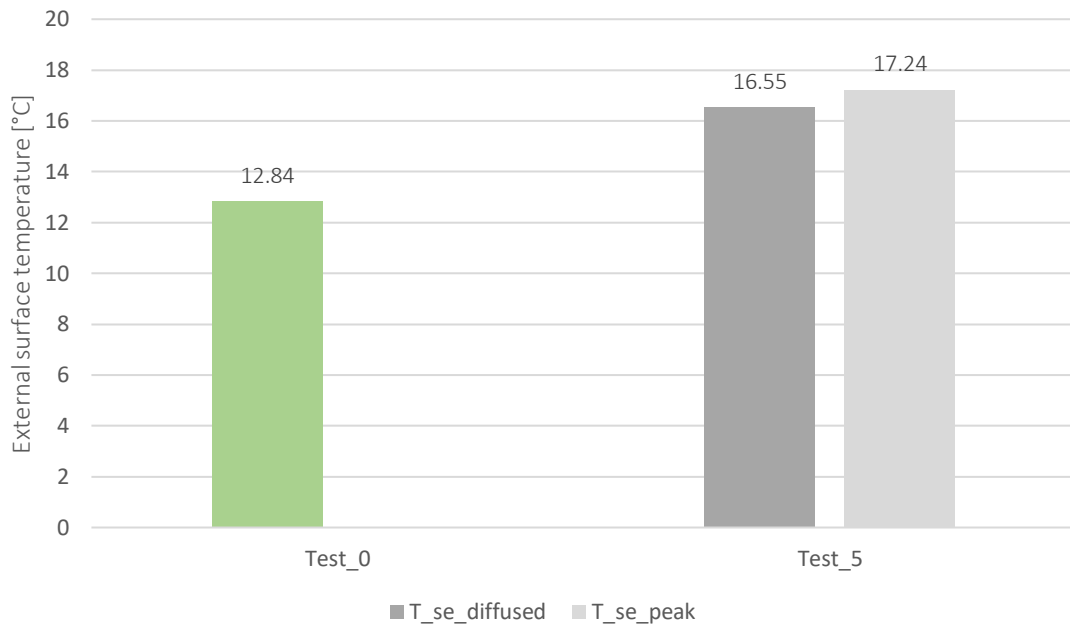


Chart 20 - Results obtained after the quantitative analysis for external surface temperature: Test 5

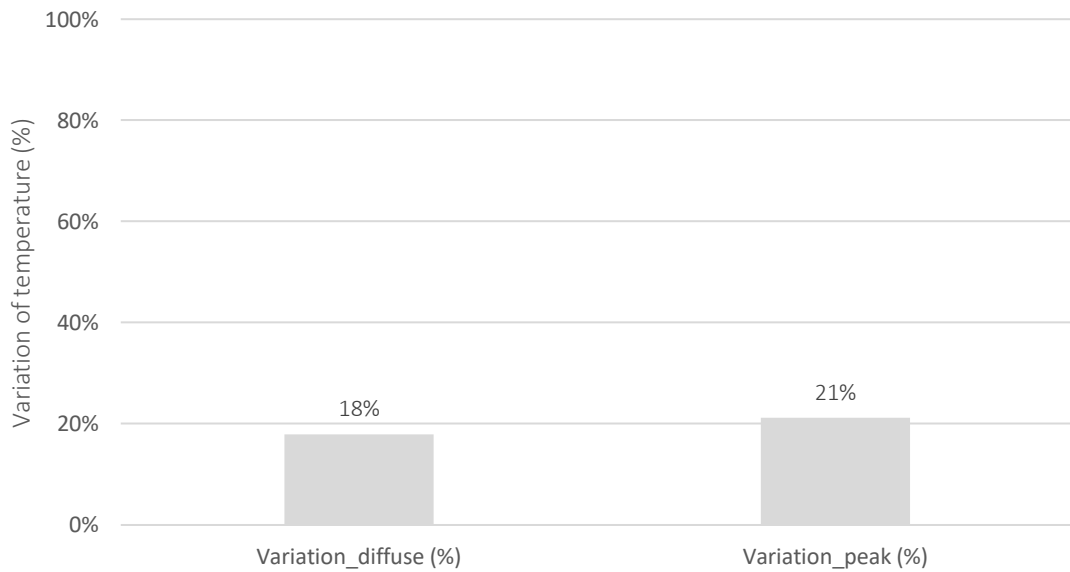


Chart 21 - Results obtained after the quantitative analysis for external surface temperature: Test 5 (variation in %)

	U_value	U_real_diffused	Variation (%)	U_real_peak	Variation (%)
Test 1	1.14	1.12	-2%	1.12	-2%
Test 2	1.14	2.04	44%	2.63	57%
Test 3	1.14	2.35	51%	2.35	51%
Test 4	1.14	2.02	44%	2.02	44%
Test 5	1.52	2.81	46%	3.51	57%
Test 6	1.14	1.85	38%	2.02	44%

Table 21 - Results obtained after the quantitative analysis for U-value: General tests

7 - STAGE 4: DATA ANALYSIS

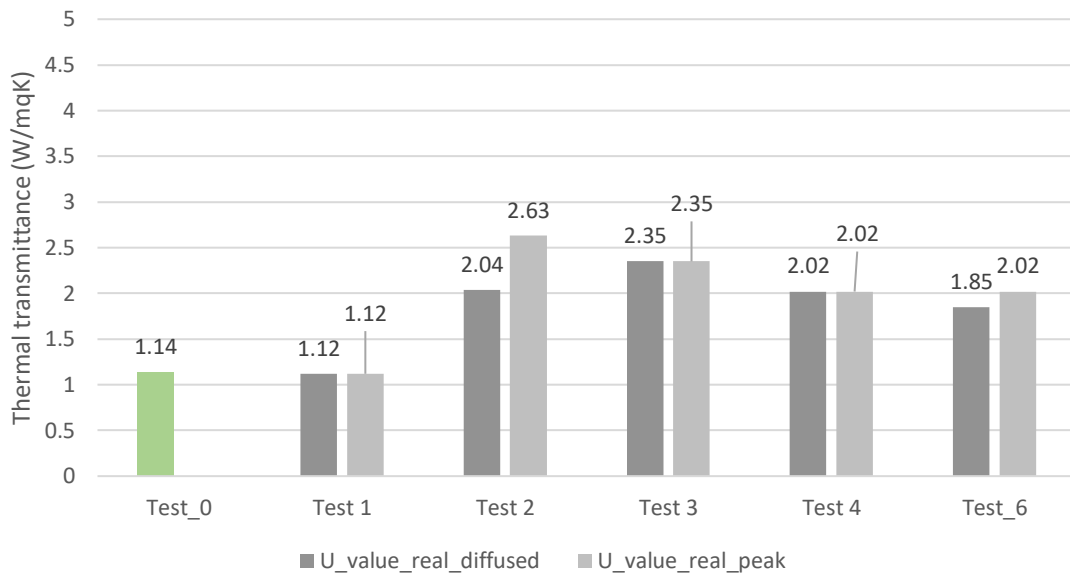


Chart 22 - Results obtained after the quantitative analysis for U-value: General tests

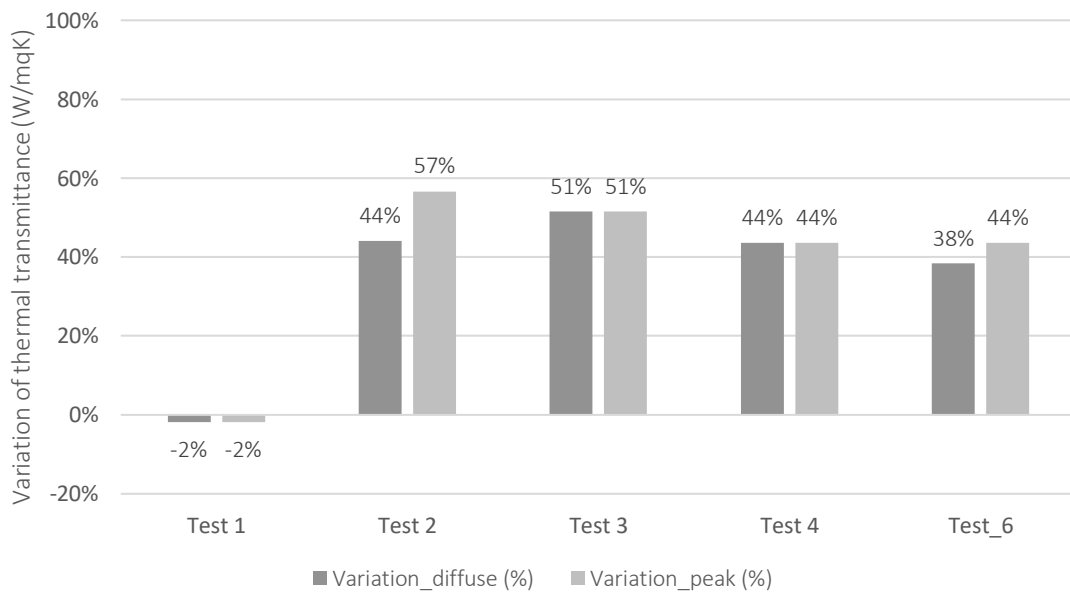


Chart 23 - Results obtained after the quantitative analysis for U-value: General tests (variation in %)

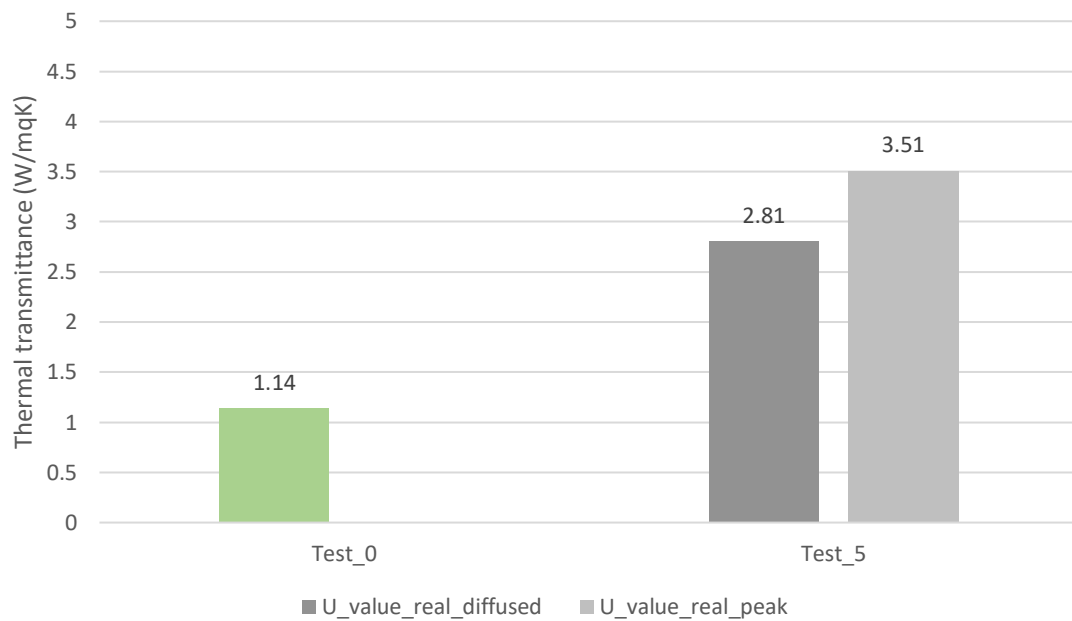


Chart 24 - Results obtained after the quantitative analysis for U-value: Test 5

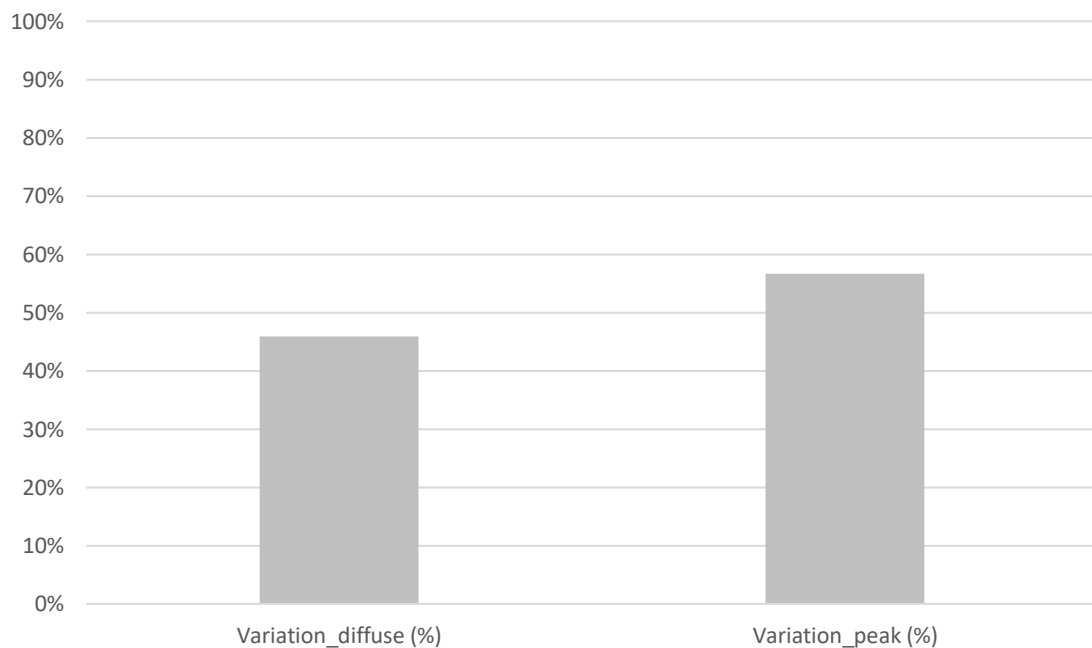


Chart 25 - Results obtained after the quantitative analysis for U-value: Test 5 (variation in %)

After the first analysis of the external surface temperatures, it was possible to notice local variations up to 20% more than the predicted both for the envelope and for the roof.

Test 1, that was also a general trial for testing the efficiency of the thermal balance, reported a low difference equal to -4% between the measured and the calculated values. It has been interpreted as the good disposition of the stratigraphy. However, it is known

that all the characteristics composing the wall change during time, but in this case the performance between the predicted and the estimated are similar.

In general, all the variations measured in local thermal anomalies are around 15% and 20%, even in the test performed on the roof, identifying the presence of an important anomaly to take into account.

The analysis of the transmittance values reported a different scenario, even more dangerous than the one analysed before.

Starting from Test 1, it was possible to estimate a reduction in the transmittance value under real conditions equal to 2%, so a negligible contraction that shows how the wall, in most part of the envelope is currently working at initial design conditions.

Focusing the attention on Test 2, that resumes the potential anomaly closed to the windows, returned U-values variation in a range between 44% and 57%. The presence of a potential thermal bridge windows and wall does not affect the flux in this way, so it is only possible to declare that part of this envelope presents a severe lack of performances that has to be repaired immediately.

Even in Test 3 U-values under real conditions overpass 50% compared to the expected one, but, as mentioned before, the presence of a potential thermal bridge in the picture could have been affected by the copper downpipe. A proper increasing in the insulation could help in reducing the possible heat transfer through that part of the envelope.

Test 4, performed in correspondence of the parapet of the east façade, affected by a water leakage due to a problem in the impermeabilization layer, returned in increasing equal to 44% in terms of thermal transmittance.

Even if it is not a part that divides the internal of the building with the external, it has to be considered strongly damaged and it needs to be repaired immediately, otherwise the failure could damage other parts of the envelope and reduce the internal performances.

The damage can be repaired through the inspection and the substitution of the impermeabilization layer. In addition, the maintenance of the aluminium cover along the perimeter of the parapet could guarantee a higher protection of the mentioned part.

In conclusion, Test 6, where a probable water leakage between the sidewalk and the envelope was noticed, returned an increasing in transmittance values measured in a range between 38% and 44%, representing a severe damage.

Neglecting the problem could generate the presence of mould by a probable humidity rising.

At the end, Test 5, conducted on the roof, where the presence of a localized thermal anomaly was underlined, returned an increasing of the U-value measured in a range between 44% and 57%, representing a potentially extremely dangerous defect.

As mentioned before, the damage has been caused by a failure in the impermeabilization layer, maybe caused by its laceration during successive phases.

The immediate repairing of the damaged layer could decrease the loss of performances.

In the following table they are collected all the possible damage, classified by an index of gravity in relation of the transmittance variation measured.

The range of the index goes from 0 to 1 and it is divided in this way:

Range	Entity of the damage
0% - 15%	Negligible
16% - 25%	Mild
26% - 50%	Serious
50% - 100%	Severe (immediate repairing)

	Ureal_peak	Var (%)	Entity	Possible intervention
Test 1	1.12	-2%	Negligible	-
Test 2	2.63	57%	Severe	Insulation addition
Test 3	2.35	51%	Severe	Insulation addition
Test 4	2.02	44%	Serious	Impermeabilization repairing
Test 5	3.51	57%	Severe	Impermeabilization repairing
Test 6	2.02	44%	Serious	Impermeabilization repairing

Table 22 - Summary of the results

7.4.3 – Energy saving

Looking at the results obtained in terms of measured and calculated thermal transmittance, some few considerations can be carried out.

The results obtained analysing all the parts of the envelope not affected by local damages, demonstrated that the behaviour calculated and predicted with the thermal balance corresponds with the measured one without regarding at some uncertainties.

However, the performance analysed and found out can carry out some considerations about a general energy behaviour of the entire building.

Even if the site is not frequented daily by the users, except for some guardians, it is possible to notice that it is a high energy-consuming construction and it is well demonstrated by the values of transmittance measured during this research.

By studying the transmittance value obtained for each single zone, considering a proper difference of temperatures, it is possible to calculate a flux in kW/mq, and, after an acceptable simplification that considers the external temperature equal hourly, we can reach a value in kWh/mq.

On the other hand, looking at energetic indexes, it is necessary to study the phenomena dynamically along the time.

Actually, some computational programmes, such as Edilclima, have been designed for that.

In the software, that works with designed values provided by different national standards according to the location of the building, after having inserted some input information such as the stratigraphy and the building services installed, it is possible to evaluate the state of the art of the construction and figure out an energy certification.

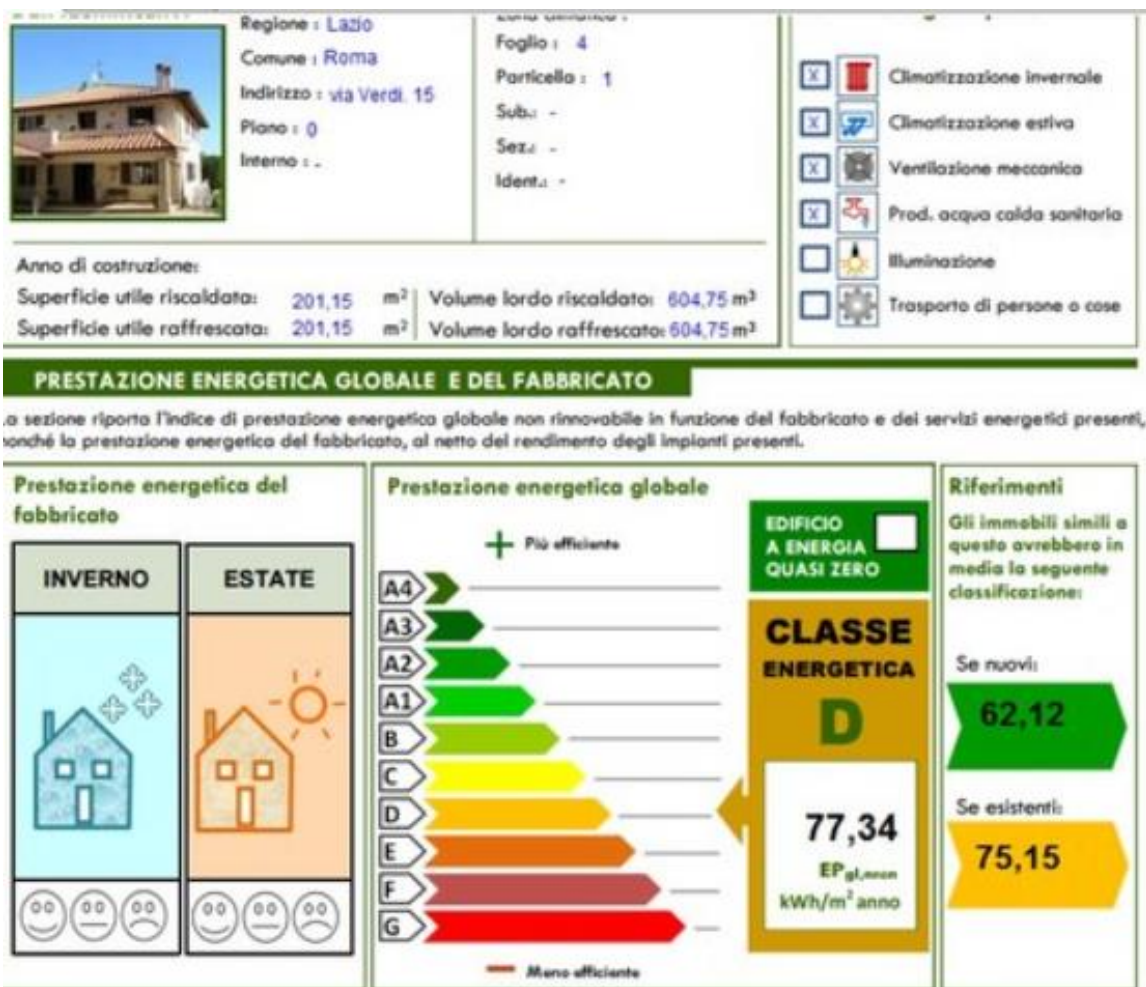


Figure 58 - Example of energy certification in Edilclima (Source: <https://www.archiproducts.com>)

For this reason, actually, an analysis like the one already performed can only be a support and not a complete substitution of a competitive system like the one produced by Edilclima. In addition, the evaluation of the heat transfer for certification purposes with the use of infrared camera needs to be coupled with another technique called "blower door".

A big ventilator is installed on the door, while all the other openings are closed in order to generate a difference of pressure and evaluate the also the possible infiltrations.

In case of possible renovation of the site, the national government in 2015 with “DM 25 giugno 2015” established minimum range of transmittance to be adopted by the designers for the renovation of buildings starting from 1st January 2021.

In the following table the minimum values of thermal transmittance are reported for our case, compared with the ones from the state of the art.

Geographic zone: E			Year: 2021
Part of the envelope	U (DM 25/06/2015)	U_measured	Variation (%)
Vertical	0.28 W/mqK	1.14 W/mqK	+ 75%
Horizontal	0.24 W/mqK	1.52 W/mqK	+ 84%

Table 23 - Comparison between U-value measured and minimum U-value required by the national standard

8-Future update: real-time model and process automation

After the tests performed and the results obtained, some considerations can be presented as a sort of starting point for a future update of the entire work.

The possible solutions require general improvements on the flight path planning, matching and feature detection phase and also in the analysis part.

8.1-Implementation of SLAM for 3D reconstruction

Looking at the flight path planification, a possible update can regard the generation of a hybrid algorithm that is based on the intrinsic characteristics of the camera adopted and the geometric obstacles that are present on the site.

SLAM (Simultaneous Localization and Mapping) can help in finding that last information through an algorithm that is able to generate an approximate map of the site useful for the movement and the calculation of the proper distance of the camera for the following phases.

The acronym SLAM, as explained before, means “Simultaneous Localization and Mapping”, identifying a series of algorithms developed to solve two problems:

- Mapping: realization of an outdoor environment map. It is usually performed by the analysis of the data coming from the internal sensors of the drone, that can return useful information of the visited space;
- Localization: positioning and device orientation, known the space visited. It is necessary a map of the place.

It is possible to notice that the two threads are strongly dependent each other and, as the “Chicken-or-egg” philosophical problem, the first requires to be preceded by the second one and vice versa.

SLAM can solve this problem performing a simultaneous solution of the two problems through the identification and visualization during the time of static points called landmarks or keyframes, realizing a map without any pre-defined knowledge meanwhile it is localized in the space.

The output obtained can be either a 2D scheme or a 3D reconstruction. In any case, it is important to regard possible mistakes in the data acquisition, mainly caused by time-loop data transmission.

Finally, to obtain a reliable result, as for Structure from Motion, it is necessary that the drone path is closed to have a coincident starting and final point.

Sensors used for data gathering are mainly rangefinder lasers and sonars. Recently, for economic and ergonomic point of view, camera started to be used (we talk about visual SLAM or monocular SLAM if there is only one camera). Odometers are used for complementary information such as drone velocity.

A sonar is useful for collecting information about point distance, but it is not so precise in understanding the direction. On the contrary, a camera gives efficient information about direction but imprecise distance information.

The probability of statistical error must be considered. Since the system requires the maximum precision in the mapping, two techniques can be adopted for the purpose: the EKF (Extended Kalman Filter), object of the focus, and the particellar filter.

The EKF is an evolution of the Kalman Filter that permits to extend this powerful mathematical model also to non-linear systems and it allows to estimate the state of the dynamic system under conditions of the equation used.

The input of the system are control signals and a “noise”, that expresses the uncertainty applied on the model of the adopted system. Sensors used for the indirect measure of the systems also introduce uncertainty and so the possibility of statistical error.

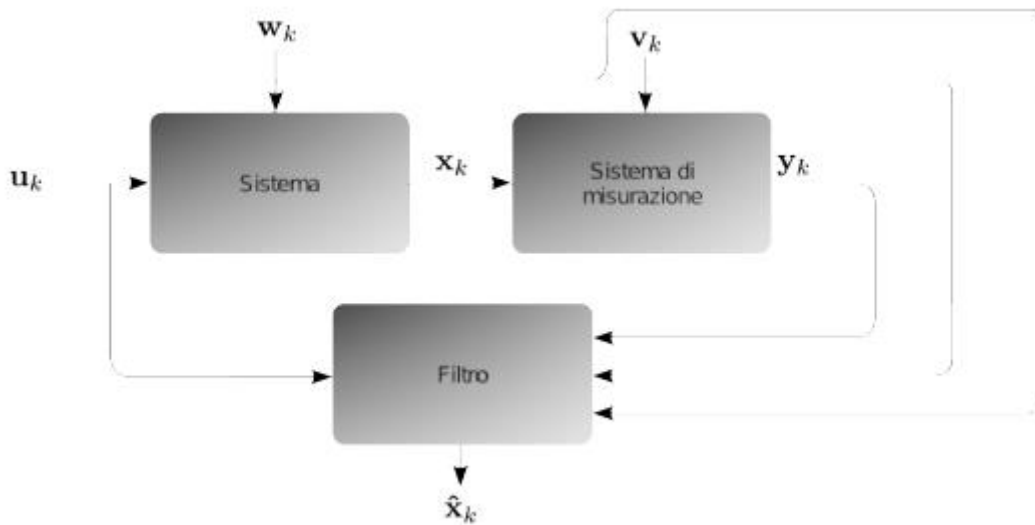
The generic filter estimates the state of the system starting from the uncertainty of the measure and the model and the input signal.

For presenting in a better way how the filter works, it is provided the equation and an explicative scheme.

$$\begin{cases} x_k = f(x_{k-1}; u_{k-1}; w_{k-1}) \\ y_k = h(x_k; v_k) \end{cases}$$

Where:

- x_k : state of the system;
- u_k : input signal of the control system;
- w_k : input noise which is responsible of the uncertainty of the system;
- f is the dynamic system;
- v_k : statistical error given by the measurement tool;
- h is the function of the measurement system;
- y_k is the measure



In this case, EKF can be adopted because it is applied to a time-varying system which is impossible to measure directly with the instrumentation.

Nowadays, since from a theoretic point of view the research is considered closed, SLAM seems more a concept than a problem solving.

For SLAM, it is necessary to have two types of devices:

- Measurement devices (sensors): they collect fundamental data for the 3D reconstruction;
- Calculation devices: devices with an efficient computational capacity for the execution of the algorithm. It is usually a rover or an autonomous vehicle that seems like a CPU desktop.

It is also possible to use smartphone or tablet if they have enough computational capacity.

Finally, all the devices are dynamic, but it is enough that only the sensor is in movement.

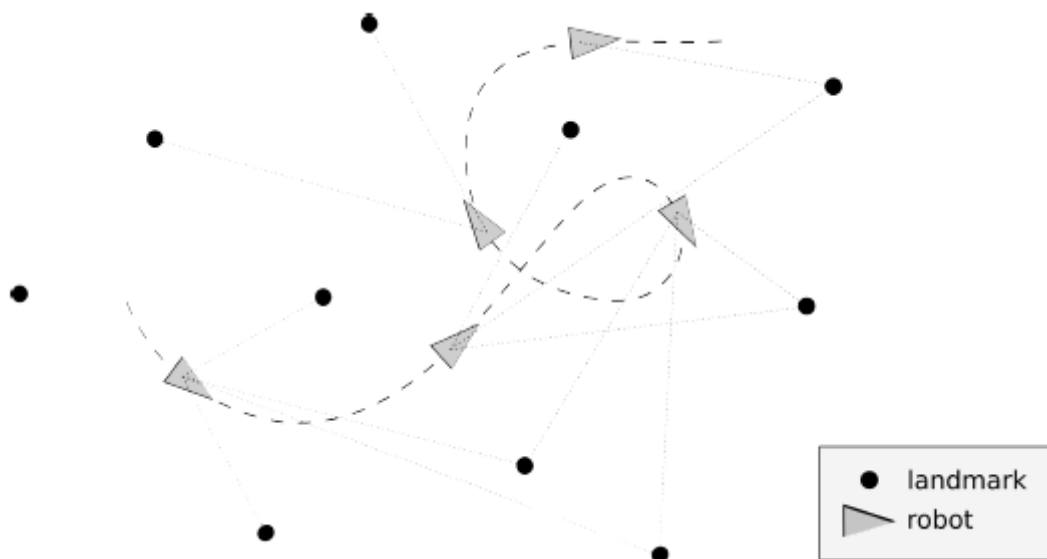


Figure 59 - Path of the rover in relation with landmarks using SLAM

Since SLAM is a process that requires movements of the sensors for realizing a 3D map, the path of the drone must be carefully considered to cover all the interested zones in a reasonable time requested by the autonomy of the battery. So, it is important to find the best optimized path.

Engel et al. presents two ways for optimizing the drone path dividing its action in local and global mission.

The purpose of the work was to estimate the best map trajectory while creating a semi-dense 3D reconstruction of the room in the lowest time possible, covering as much voxels as possible, using as algorithm the LSD-SLAM (Large-Scale Direct – Simultaneous Localization and Mapping).

This method is considered as an evolution of the traditional monocular SLAM. Although the purpose is the same, LSD-SLAM generates directly a consistent global map, using direct image alignment and probabilistic, semi-dense depth maps instead of keypoints. The trajectory is closed making the starting and final keyframes overlapping and in this case, the data obtained both by SLAM and by the internal devices of the drone, are filtered through EKF approach.

Looking at the movement, a local approach in the room was conducted performing a simple lateral movement for covering the maximum amount of occupancy map's voxels, as shown in figure.

The lateral motion helps to reduce the formation of indentation at the edges (in particular, closed to textureless elements such as walls) in the semi-dense image of the occupancy map, that is a phenomenon caused by the probabilistic approach used for supporting the hypothesis. In fact, the program considers a voxel as occupied only if there is at least 86% of probability that it is effectively occupied. This high conservative

threshold and the effectiveness of the lateral motion are based on Monte-Carlo simulation approach.

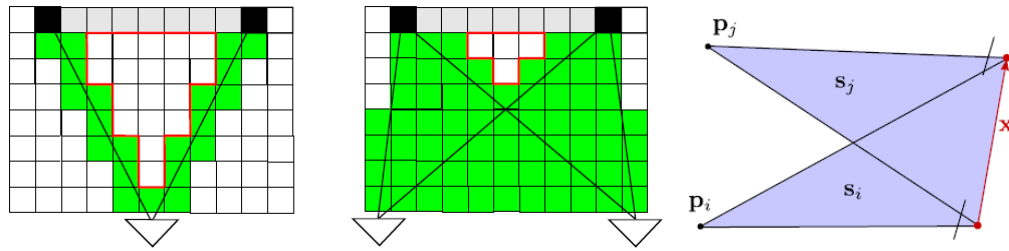


Figure 60 - Start discovery movement tested using Monte-Carlo simulation (Source: Von Stumberg et Al - From Monocular SLAM to Autonomous Drone Exploration)

The test is then repeated moving from a lateral motion to a 3D motion, called Star Discovery. The drone is flown inside the room following a motion star-shape trajectory and illustrated in the picture below.

At a global point of view, an algorithm permitted to the drone of operating and finding the occupied voxels and, at the same moment, looking for interested and unexplored zones of the room.

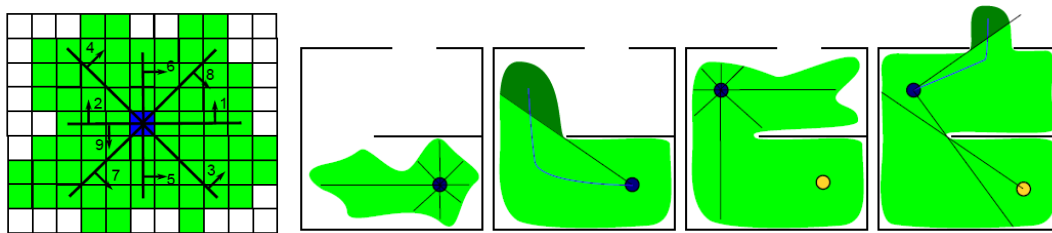


Figure 61 - Star discovery movement (Source: Von Stumberg et Al - From Monocular SLAM to Autonomous Drone Exploration)

It results that star discovery can guarantee a higher coverage of the room and so a more efficient and precise 3D LSD-SLAM occupancy map. The results are shown in the pictures below and it is possible to notice the comparison of analyzed voxels found after star discovery motion.

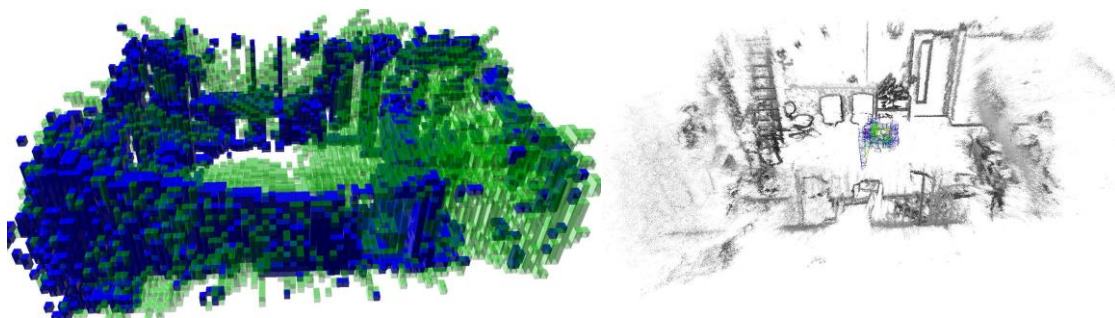


Figure 62 - Comparison between the voxels mapped with the two movements tested (Source: Von Stumberg et Al - From Monocular SLAM to Autonomous Drone Exploration)

SLAM technology is used also for the localization of outdoor places with specific requirements through the creation of a 2D grid map. For example, several studies are developing the possibility that a drone could select an area and land autonomously under emergencies conditions.

This possibility becomes fundamental for drones designed for operating in war contexts because the GPS technology can be easily intercepted and crypted by disturbing tools such as jammers. However, this problem could be present also in civil contexts.

In the paper analyzed, it was decided to work with a M100 UAV as drone accessorized with a camera and a barometer for evaluating the altitude of UAV from the terrain.

For the test it was decided to work with SVO algorithm, that uses two parallel threads, one for estimating the camera motion, and a second one for mapping as the environment is being explored. This separation allows fast and constant-time tracking in one thread, while the second thread extends the map, decoupled from hard real-time constraints.

It was possible to divide the test into three main steps: starting from the 3D points cloud, the grid map was extracted and then, due to the probabilistic estimation of each point caught, the suitable area for emergency landing was selected.

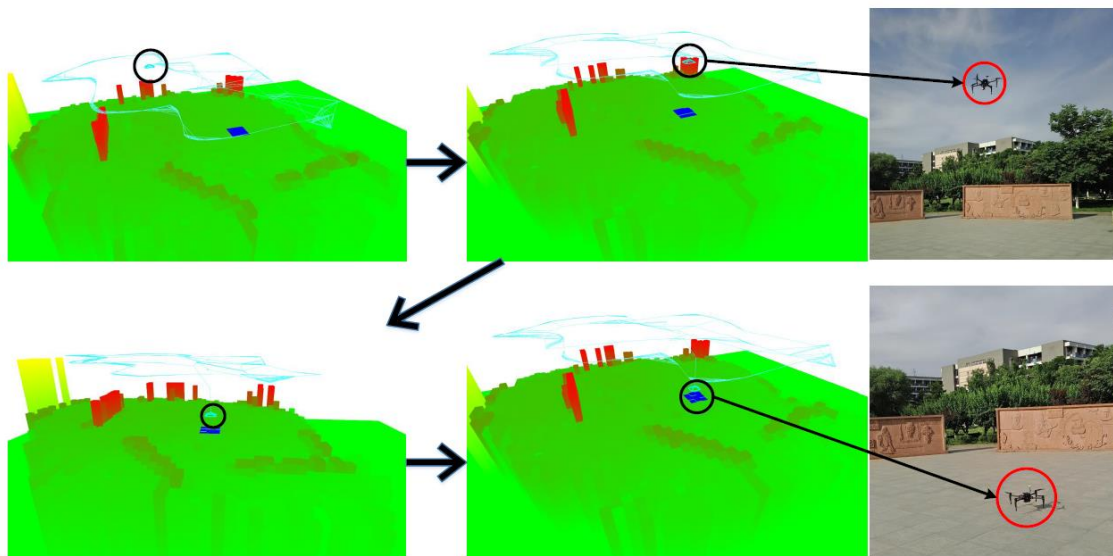


Figure 63 - Test on autonomous landing of the drone (Source: Yang et al - Monocular Vision SLAM-Based UAV Autonomous Landing in Emergencies and Unknown Environments)

As previously discussed about, the implementation of SLAM algorithm can be an interesting updated on the research. Since the literature review has not presented a possible universal application in the issue of the optimal movement of the drone, the integration of the two image-based techniques can be studied.

By the results obtained during this experience, it is possible to think about the combination of SLAM and SfM in order to launch an algorithm which is able to plan an

autonomous information acquisition phase without the presence of the pilot and on the basis of the technical characteristics of the cameras and the geometrical limitations present in the site.

The process can be hypothesised by considering two consequent thread. During the first step, the UAV scans the entire area from the top of it making a summary reconstruction of the obstacles.

The result is the generation of a computer vision object which is able to communicate to the drone which are the distances between the obstacles present in the site studied. An additional identification of the safe areas where to land autonomously is also considered by analysing the information gathered during the preliminary flight.

In order to perform this and reduce at the minimum the possible collision with the elements present in the area, an elliptic flight at the top of the building can be planned.

All the data acquired are used as boundary conditions for the solution of the second part of the process, that is the one related to model reconstruction, both thermal and RGB.

In the next stage, a flight path for image acquisition is automatically processed computing the optimal GSD and footprint value on the basis of the geometrical data collected before.

8.2 - Real-time model generation for time scheduling reduction

After all the processes involved and presented in the research, the drone is able to produce, read and analyze a 3D object with all the information required.

By the way, all the steps in sequence ask for time and this may take disadvantageous using this process instead of a conventional visual analysis of pictures or an energy analysis using software like Edilclima or IES-VE.

By the results obtained during this experience, it is possible to compute the time of the process in 9-10 hours, without considering the one required by image acquisition phase, distributed in this way:

- Image acquisition: variable (2 hours, in this case);
- Image uploading: 1 hour;
- Orientation of the images: 40 minutes;
- Image matching: 15 minutes;
- Markers and scale bar positioning: 30 minutes;
- Dense point cloud reconstruction: 2 hours;
- Mesh reconstruction: 1 hour;
- Thermal texture integration: 2.5 hours;
- Façade modelling: 3.5 minutes;
- Analysis: 5 minutes

Even in a small site like the office building in Spino d'Adda the time asked for the process can be reduced making some hypothesis on possible updates and improvements of the entire code.

First, it is possible to declare that working all the steps into a unique software instead of the two required now can help in reducing time asked by their communication.

In order to do that, writing everything in a unique code in Matlab or other IT environments, that are able to perform both the part related to the 3D reconstruction and the analysis of the point cloud, could be planned.

Since the Structure from Motion part requires a high level of complexity, it will be hard to obtain the same precision than realizing a model in Metashape.

However, since the scope of the software is to perform analysis and not provide precise and aesthetically-attractive 3D objects, a part of precision can be sacrificed for time reduction and simplification updates.

In addition, the experience demonstrated that the dense point cloud reconstruction has not given back useful and solid results, so as, for our purpose, we can think to remove it and gain time passing directly on mesh generation from the sparse point cloud.

In the project of an updated script, the phase of image acquisition could start in parallel with the first Structure from Motion steps: once an image is collected, the process of image quality evaluation, feature extraction and evaluation could be started.

By the way, it is important to notice that a pre-calibration of the cameras is required and easily possible to do because all the intrinsic parameters are known values.

Matching can be processed after the acquisition of at least four sequential images and continued by the time the data gathering phase is performed. In this way, part of the process can be anticipated and the time reduced.

However, mesh and texture generation have to be performed once all the points are matched and well-identified in the space, so as the sparse point cloud is generated.

In general, working by considering the contemporaneity of some steps can be useful for reducing time on the process, estimated in 60% of the overall time scheduled now.

In addition, even the part of the algorithm related to the analysis has good opportunities of technical improvements.

First of all, the façade segmentation can be generalized and extended by improving the sensibility of the MSAC part. A great part of the effort has to be focused on the integration of the algorithm in the segmentation of complex facades.

Then, even the part related to analysis could be generalized. Actually, all the analysis are performed considering a predicted value of external surface temperature based on the thermal characteristics of the envelope and meteorological conditions measured on site.

Since the first group of variables are inserted before the analysis, the second is too dependent on the availability of a reliable station.

The automatic implementation of a service that extracts the required climatic data from the closest station only considering the location has to be taken into account.

In conclusion, since the morphology of the site and the absence of other relevant obstacles, an additional study of the shading has not been performed.

Bibliography

- Ernest Orlando Lawrence Berkeley National Laboratory. (2016, Febbraio 5). *BigLadderSoftware*. Retrieved from Outside Surface Heat Balance: <https://bigladdersoftware.com/epx/docs/8-4/engineering-reference/outside-surface-heat-balance.html>
- Alba, M., Barazzetti, L., M., S., & Rosina, E. (2011). Mapping Infrared Data on Terrestrial Laser Scanning 3D Models of Buildings. *Remote Sensing*, 1848-1870.
- Barazzetti, L., Previtali, M., Brumana, R., Cuca, B., Oreni, D., Roncoroni, F., & Scaioni, M. (2013). AUTOMATIC FAÇADE SEGMENTATION FOR THERMAL RETROFIT. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, (pp. 197-204). Trento: 3D-ARCH 2013.
- Barazzetti, P., Previtali, M., Rosina, E., & Scaioni, M. (2013). Mosaicking thermal images of buildings. *Proceedings of SPIE - The International Society for Optical Engineering*.
- Bayomi, N., Nagpal, S., T., R., & J.E., F. (2021). Building envelope modeling calibration using aerial thermography . *Energy e buildings*, 2-12.
- Boller, C., Kuo, U., & Echsmann, M. (2012). Unmanned Aircraft Systems for Remote Building Inspection and Monitoring. *6th European Workshop on* (pp. 58-64). Dresden: German society for non-destructive testing.
- Buerop-Lutz, C., & H., S. (2015). Inspecting PV-Plants using aerial drone-mounted infrared thermography system. *Third Southern Africa Solar Energy Conference*. Krueger National Park, South Africa: SASEC2015.
- Canapini, D., Malaguti, F., & Ricchetti, R. (2010). *Automazione e Robotica nelle Costruzioni: Presente e Futuro*. Retrieved from Piattaforma Tecnologica italiana delle costruzioni: www.ptic.it
- Casana, J., Wiewel, A., Cool, A., Chad Hill, A., & Fisher, K. (2017). Archaeological Aerial Thermography in Theory and Practise. *Advances in Archaeological Practice* 5, 310-327.
- ChallengeAerospace. (2020, september 20). *Products*. Retrieved from ChallengeAerospace.com: <http://challengeraerospace.com/innovations/multi-rotor-aircraft-systems-2/>
- Dabbura, I. (2018, settembre 17). *K-means Clustering: Algorithm, Applications, Evaluation Methods, and Drawbacks*. Retrieved from Towards data science:

<https://towardsdatascience.com/k-means-clustering-algorithm-applications-evaluation-methods-and-drawbacks-aa03e644b48a>

- De filippo, M., Harris, S., Asadiabadi, S., & Nigel, K. (2019). Concept of Computer Vision Based Algorithm for Detecting Thermal Anomalies in Reinforced Concrete Structures. *Proceedings*, 2-10.
- ENAC. (2020, luglio 9). Regolamento Enac - Mezzi aerei a pilotaggio remoto. Roma, Italia.
- Entrop, A., & A., V. (2017). Infrared drones in the construction industry: designing a protocol for building thermography procedures. *11th Nordic Symposium on Building Physics* (pp. 64-68). Trondheim: ScienceDirect.
- Frattoilillo, E. (2018). La radiazione solare: utilizzo, selettività e schermi. Cagliari, CA, Italia.
- Frodella, W., Elashvili, M., Casagli, N., & AAVV. (2020). Combining InfraRed Thermography and UAV Digital Photogrammetry for the Protection and Conservation of Rupestrian Cultural Heritage Sites in Georgia: A Methodological Application. *Remote sensing*, 12-37.
- Garbade, M. J. (2021, gennaio 8). *Understanding k-means clustering in machine learning*. Retrieved from Towards data science: <https://towardsdatascience.com/understanding-k-means-clustering-in-machine-learning>
- Ham, H., Wesley, J., & M., H. (2019). Computer vision based 3D reconstruction : A review. *International Journal of Electrical and Computer Engineering (IJECE)*, 2394-2402.
- Inc., S. s. (2021, gennaio 30). *Spectral science*. Retrieved from MODTRAN: http://modtran.spectral.com/modtran_index
- Inzerillo, L., Di Mino, G., & Roberts, R. (2018). Image-based 3D reconstruction using traditional and UAV datasets for analysis of road pavement distress. *Automation in construction*, 457-469.
- Kostantinos, D. G. (2010). *Overview of the RANSAC Algorithm* .
- Lindenberg, T. (2012). Scale-Invariant Feature Transform. *Scholapedia*, 1-18.
- Melis, M. T., Da Pelo, S., Ivan Erbi, I., Loche, M., Deiana, G., Demurtas, V., . . . Scaringi, G. (2020). Thermal Remote Sensing from UAVs: A Review on Methods in Coastal Cliffs Prone to Landslides. *MDPI - Remote sensing*, 2-29.
- Otiz-Sanz, J., Gil-Docampo, M., Arza-Garcia, M., & Canas-Guerrero, I. (2019). IR Thermography from UAVs to Monitor Thermal Anomalies in the Envelopes of Traditional Wine Cellars: Field Test. *Remote sensing*, 2-18.

- Papakonstantinou, A., Kavroudakis, D., Kourtzellis, Y., Chtenellis, M., Kopsachilis, V., Topouzelis, K., & Vaitis, M. (2019). Mapping Cultural Heritage in Coastal Areas with UAS: The Case Study of Lesvos Island. *MDPI*, 19.
- Previtali, M. (2013). *Automatic segmentation, classification and extraction of repeated patterns for building façades modelling*. Milano : - .
- Previtali, M., Barazzetti, L., Brumana, R., Cuca, B., Oreni, D., Roncoroni, F., & Scaioni, M. (2013). Automatic façade modelling using point cloud data for energy-efficient retrofitting. *Appl Geomat*, 95-113.
- Previtali, M., Barazzetti, L., Redaelli, V., Scaioni, M., & Rosina, E. (2013). Rigorous procedure for mapping thermal infrared images on three dimensional models of building façades. *Journal of applied remote sensing*.
- Rakha, T., & Gorodetsky, A. (2018). Review of Unmanned Aerial System (UAS) applications in the built environment: Towards automated building inspection procedures using drones. *Automation in construction (93)*, 252-264.
- Rakha, T., Gorodetsky, A., & Kakillioglu, B. (2018). Campus as a Lab for Computer Vision-based Heat Mapping Drones: A Case Study for Multiple Building Envelope Inspection using Unmanned Aerial Systems (UAS). *7th International Building Physics Conference* (pp. 517-522). Syracuse: IBPC2018.
- Riminesi, C. (2018). *Tecniche e metodologie diagnostiche non-invasive: Termografia IRT*. Firenze: Consiglio Nazionale delle Ricerche.
- Robison, R. (2017). *Droni nel cielo - Storia degli aeromobili a pilotaggio remoto*. Tricase: - .
- Von Stumberg, L., Usenko, V., Engel, J., Struckler, J., & Cremers, D. (2018). From Monocular SLAM to Autonomous Drone Exploration. *Arxiv*.
- Weichenhain, U. (2019, gennaio 29). *Cutting costs in infrastructure maintenance with UAVs*. Retrieved from RolandBerger.com: <https://www.rolandberger.com/en/Insights/Publications/Drones-The-future-of-asset-inspection.html>
- Yang, Li, Zang, LI, & Li. (2018). Monocular Vision SLAM-Based UAV Autonomous Landing in Emergencies and Unknown Environments. *MDPI*, 1-18.
- Yıldırım, S. (2021, gennaio 8). *k-means clustering*. Retrieved from Towards data science: <https://towardsdatascience.com/k-means-clustering-explained>
- Zerol. (2020, ottobre 4). *A Simple Explanation of K-Means Clustering*. Retrieved from Analytics Vidhya: <https://www.analyticsvidhya.com/blog/2020/10/a-simple-explanation-of-k-means-clustering/>

Acknowledgement

Ho provato a scrivere una memoria strutturata per quest'ultimo "capitolo", ma poi mi sono reso conto che suonava troppo formale, fredda e distante. In altre parole: troppo inespressiva per poter raccontare con poche righe cosa è stata per me questa esperienza.

Credo che proseguirò a braccio e non me ne pentirò.

Ricordo ancora il primo giorno che sono arrivato al Politecnico. Pioveva, faceva freddo e il treno arrivò a Milano in ritardo.

Mi avvicinavo a quel nuovo mondo con piglio curioso, molto ingenuo, e forse un po' timoroso di ciò che di lì a poco sarebbe accaduto. Posso dire che mi presentavo, dal punto di vista delle esperienze, dopo cinque anni di liceo, come un foglio bianco.

Oggi, invece, c'è il sole, non posso più prendere quel treno e il vento sta cercando di portarmi via i fogli sul quale, tra poco, dovrò rimettermi a lavorare.

Nel frattempo, il foglio bianco è stato riempito di nozioni e teoremi, gioie e qualche delusione. Se immaginassi questa parte di vita come una grande stazione, direi che tanta gente è scesa e qualcuno è partito. Tante cose sono cambiate e mi hanno cambiato.

Alla fine di questo percorso, sento di non essere più un ragazzo ingenuo, ma una persona in grado di affrontare le sfide che il mondo della Post-Pandemia ci presenterà.

Per questo dovrò ringraziare il Politecnico, non semplice università ma vera e propria scuola di Vita.

Non illudiamoci che tutto tornerà alla normalità. Nulla, dopo questa Pandemia, sarà veramente come prima.

Continueremo a guardare le cicatrici di questa lotta da cui, per forza di cose, avremo tutti qualcosa da imparare, e ne usciremo migliori solo se, parafrasando un artista che io stimo parecchio, su di esse "ci attaccheremo le ali".

Io sono fiducioso. Del resto, la base per i nostri computer e la Nutella sono state inventate per necessità imposte da un momento di crisi.

Mi sono dilungato troppo. Credo che ora sia il caso di iniziare con i molti grazie che devo dire.

Un ringraziamento speciale va ai Proff. De Angelis, Fagiano, Scaioni, e ovviamente a Michele, che mi hanno seguito passo dopo passo per scrivere questa tesi, dispensandomi con pazienza consigli utili e mostrandomi la giusta via.

ACKNOWLEDGEMENT

Ci tengo anche a ringraziare l'Arch. Emanuele Monti di Arona e lo studio Beta Progetti di Firenze, per aver scommesso e creduto nelle mie capacità, permettendomi di lavorare con loro.

Un sentitissimo ringraziamento va anche all'Ing. Marchisio e al Comune di Arona, per avermi dato l'opportunità di svolgere con loro il tirocinio curriculare, permettendomi di fare un'esperienza meravigliosa e altamente formativa.

E poi grazie a tutti i famigliari, gli amici e i colleghi, per il loro supporto e i loro straordinari consigli nei momenti in cui tutto sembrava crollare.

L'ultimo ringraziamento, che poi è una dedica, va ai miei genitori, per tutti i sacrifici, il supporto, la pazienza e il tempo che hanno saputo regalarmi e dedicarmi.

A volte lo si dà troppo per scontato, ma avere una famiglia che ti appoggia e ti aiuta a crescere è un raro privilegio.

Mamma e Papà, noi siamo quello che impariamo.

E io sono felice di essere ciò che sono grazie a quello che mi avete trasmesso.

Kevin