POLITECNICO DI TORINO

Facoltà di Ingegneria
Corso di Laurea in Ingegneria Biomedica

Tesi di Laurea

3D reconstruction of chronic wounds from 2D images for in-depth clinical analysis

Relatori:
Prof. Fernando Corinto, PhD
Dott. Francesco Marrone, MS
Dott. Gianluca Zoppo, MS

Candidato:
Ada Palamà

Anno accademico 2019-2020

A mio padre,
alla mia famiglia
e ai miei amici.

"Il fiore che sboccia nelle avversità
è il più raro e il più bello di tutti"
- Mulan

Acknowledgments

At the end of this thesis, many are to thank for their precious work and help.

I would like to express my appreciation to my company supervisor Jacopo Secco for his help and encouragement, for have always believed in me and given me the necessary support in the bad moments.

I would like to acknowledge my supervisor Professor Fernando Corinto for his mentorship and his personal and technical support throughout the project.

I would like to thank to my supervisors Gianluca Zoppo and Francesco Marrone, my partners in crime, for all the teachings, the good times and comfort they've given me over the months.

I would like to thank to Dr. Elia Ricci and Monica Pittarello for the professionality and disponibility shown during the work in the clinic.

I would like to thank all the patients for their willingness and cooperation.

Thank you all.

Summary

In Italy, chronic wounds affect 1.5-2 million people and have an incidence of 4 million people. The study of wounds, the way they are cured and the employed resources, play an important role for the health of patients. Novel techniques and the optimization of wound's treatment can significantly impact the costs sustained by the National Health Service. It is estimated that the number of patients with skin ulcers will increase due to inadequate lifestyles, advancing age and increasing obesity. In this regard, current studies focus on the optimization and design of solutions that will allow the doctor's work to be accelerated, improve the ability to heal the wound and reduce the economic resources employed. Among the present tools that support the doctor in making treatment decisions, Wound Viewer is a novel, portable, fully automated device that is capable of performing an accurate and objective skin ulcer evaluation. Due to the recent progress made in the 3D image processing domain, powerful new tools for medical applications are now able to provide the geometry of the reconstructed object, 3D volume calculation and other important measurements. For these reasons, this work will focus on the 3D reconstruction of skin lesions that can be then used for additional medical evaluations.

The first chapter of this thesis opens with the description of skin chronic wounds and an analysis of causes, classification by Wound Bed Preparation Score, costs and treatments. There are numerous devices designed to support the specialist during the visit and to allow the data sharing and storing with time and costs savings. An innovative device in the vulnological field is the Wound Viewer.

In the second chapter is proposed an algorithm that is able to perform the 3D reconstruction using two images of the skin ulcer taken from the same camera from two different points of view. In this chapter are analyzed in detail all the algorithms used in the individual steps necessary for the construction of the final three-dimensional model. The final result is a dense point cloud positioned in 3D space.

The third chapter illustrates the results obtained, step by step, applying the algorithm to several initial Picture sets. For each Picture set is shown the three-dimensional point cloud obtained.

The fourth chapter provides an illustration of possible improvements starting from denoising of the point cloud to the surface reconstruction. Then, there are some examples of applications of the algorithm for in-depth clinical analysis.

Table of contents

List of tables

List of figures

Chapter 1

Chronic Wounds

In medicine, an ulcer is defined as a continuous solution of the skin that can involve the epidermis, dermis, hypodermis, tendons, muscle fascia, muscle tissue and underlying ligament, bone and/or cartilage structures. Skin ulcers are wounds or losses of skin tissue that have failed to activate a re-epithelialization mechanism.

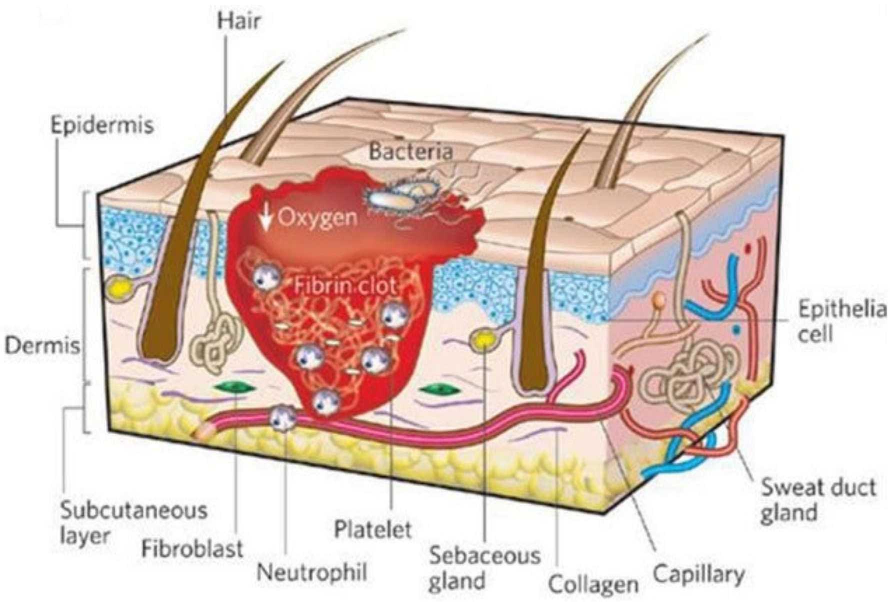


Figure 1.1: Skin tissue with ulcer from [4.1]

Reepithelialization consists of an orderly and timely reparative process to restore the functional and anatomical integrity of the affected area. A wound is defined as chronic if the time reparative process failed before a period of three months and it does not show a tendency of spontaneous healing.[4.2] An ulcer is always a consequence of a primitive lesion called a pre-ulcerative lesion. The study of the transition from this primitive state to the presence of a chronic lesion is considered as a fundamental phase for defining an appropriate therapeutic plan. However, sometimes this period is so short that the two different states have to be considered as an unique event. The analysis of the perilesional skin, *i.e.* the portion of skin beyond the margin of the wound, is fundamental in order to obtain a complete diagnosis and a correct therapeutic approach. However, one of the most important steps for the definition of a successful wound treatment is the identification of the correct etiology of the chronic wound. The risk of a wound becoming chronic depends on different factors such as local tissue hypoxia with repetitive ischemia-reperfusion injury. The recognition of a chronic wound requires an assessment of the appearance of the edge and the composition of the local wound environment. A chronic wound is characterized by a raised and hyperproliferative wound margin while the local wound environment, rich in anti-inflammatory products and pro-inflammatory cytokines, is characterized by an unbalanced enzymatic environment. Every wound could become chronic, especially when inadequately treated. As conseguence, it is necessary to correctly understand the wound's aetiology as the triggering causes may be multiple. In the following paragraph the main causes of ulceration are analyzed and the corresponding possible treatments are presented.

1.1 Wound causes

Haeger said that leg wounds are never a disease in itself, but are a symptom of an underlying disease.[4.3] This statement highlights that chronic wounds can be consequence of other underlying disorders. For instance, in the Western world, ulcers are mainly the result of:

- Venous insufficiency;
- Arterial Perfusion;
- Diabetes;
- Nutritional Status;
- Infection;
- Immunosuppression.

The lack of appropriate assessment for wounds' treatment could make time for healing longer. This results in increased patients pain and increased costs for the treatment. The first step in wound care is to make a proper assessment of wound condition and classification. Chronic wounds are characterized by the presence of a raised, hyperproliferative, nonadvancing wound margin.[4.2] Depending on the different mechanisms by which the lesion is caused, skin ulcers are divided into: vascular, arterial, pressure and diabetic ulcers.

- Vascular ulcers affect 5 million people in the United States and the cost of treatment accounts for 1-2 % of the health care budget. Vascular ulcers are the result of increased localized pressure in the veins and capillaries. Vascular wounds can be divided into arterial and venous and they could be discerned by accurate clinical assessment. Vascular skin ulcers are often the result of trauma, epithelial tumors, vascular or autoimmune diseases that trigger tissue loss. Other risk factors are obesity, sedentariness, heavy work and age. Most vascular ulcers are located in the lower limb between the ankle and calf. In the case of arterial ulcers, the intervention aims to restore blood flow by revascularization while venous ulcers are treated with compression therapy. Compression therapy improves venous ulcer healing rates and helps the patient by decreasing ulcer recurrences.[4.2]

- Arterial ulcers are more difficult to diagnose and therefore often increase the time needed for treatment. Blood flow in patients with arterial lesions is unable to meet oxygen demand due to low vessel dilation and obstruction. For this reason, this kind of ulcers has to be treated by reconstructive surgery or angioplasty.

- Pressure ulcers are also called decubitus ulcers because they are the consequences of a state of immobilisation for a prolonged period of time. They mainly affect the elderly. This kind of ulcers appears when a high-pressure area decreases the blood flow with a consequent accumulation of toxic products that can possibly lead to tissue hypoxia.

- Diabet wounds become chronic due to the presence of either small and larger vessels disease [4.2]. Diabetes ulcers are mainly located at the feet and have a neuropathic origin. They are mostly treated by diabetes control, medical therapies for neuropathy and surgical decompression which helps to improve vascular flow. A re-education to movement has shown important results assisted by the use of walkers, crutches, wheelchairs and stabilizing footwear.

As it will be discussed in the next paragraph, the cost of treating chronic wounds represents a major medical expense for society. Therefore, making a fast and valid assessment plays an important role both for the patient's health and for the reduction of the economic resources employed.

1.2 Socio-Economics Overview

A worrying fact about ulcers is the increasing number of people who will develop a chronic ulcer in their lifetime. In developed countries, it is estimated that 1-2 % of the population will develop an ulcer and this figure will increase due to incorrect lifestyles, increased obesity and diabetes.[4.4]
According to studies conducted in Great Britain and Denmark, there are about 4 people per 1000 inhabitants with one or more injuries. After one year, 15 % of the wounds remain unresolved and thus become chronic. A study conducted in the United States reported that the number of patients with skin lesions is 7.5 million and the total annual cost of treatment is around $ 22 billion.[4.5] The annual treatment cost for Americans is larger than in Europe since in the United States the average hospitalization of a patient with a skin lesion is about 3 times that of a patient treated in Europe. The average cost of an individual with a skin lesion is estimated at 14,426 $ while in Europe it is around € 6,650.
Also in Italy the vulnological sector uses important economic resources. The Italian patients with chronic skin ulcers are over two million and 30,000 of them are children. Ulcers need treatment that should be paid by the National Health Service, and according to the Association of Skin Ulcers (AIUC) during "UlcerDays" in 2009, 75 % of patients cannot afford the treatment because it is too expensive. As in the rest of the world, lower leg wounds interest, in particular, the over-65s. They have a wound prevalence of 3.6 % while in the general population the prevalence is of1 %.
The numbers change a lot for the ultra-octane class patient with skin ulcers. According to ISTAT data for 2011, in Italy there are 2.9 million members of the ultra-octane class. If these numbers look very high, even more worrying are the predictions about how many more patients will develop a skin ulcer. The number of patients is destined to reach 7.7 million in 2030, reaching worrying costs for the National Health Service. The cause of these numbers is in the lifestyles. These lifestyles play a great role and their impact is getting worse because of incorrect nutrition, stressful routines, increased obesity and diabetes.
The main cost of managing a patient with a skin ulcer is not the medication itself but the treatment time and hospital costs. The hospitalization represents 80-85% of the total cost of the treatment of patients with ulcers. These costs increase considerably in the case of complications such as infections or amputations.
The distribution of the costs of wound management is represented by the Figure 1.2. As reported, only 15 % of the total sum is spent on materials used for medication. 30-35 % of the resources correspond to the time spent by the doctor or nurse on the dressing and 50 % of the resources are used for hospitalization. The Figure 1.2 shows important numbers thus it is crucial to develop an operational plan that leads to greater efficiency in wound management and consequently lower treatment costs.

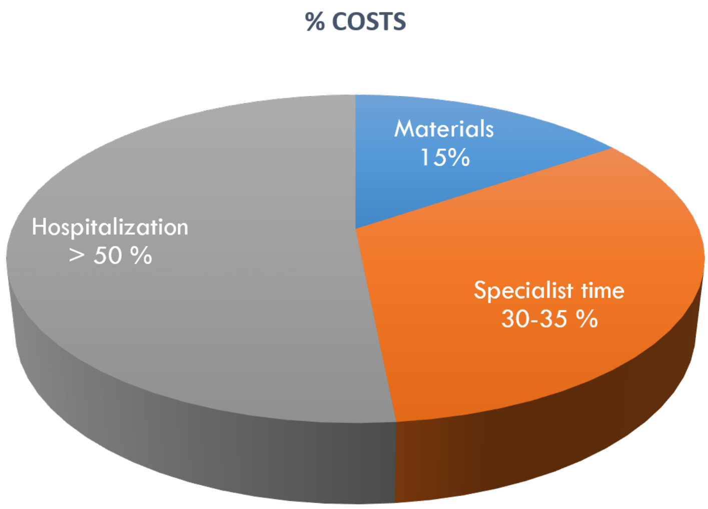


Figure 1.2: Cost distribution

The decisive factors in establishing a cost effective operational plan for both patients and costs are the following three: nurse's time, hospitalizations and dressings. Depending on the type of injury, the time and frequency of dressings are well enough determined.

The analysis of data has shown that it is important to focus on the incidence of the wound development and its complications. Each wound has its own evolution depending on the patient, the treatment and many other factors. It is therefore necessary to pay attention to the evolution of the wound.
Monitoring the wound and its temporal changes are therefore of fundamental importance in the treatment of skin ulcers. The current treatment consists of examining the granulation and colour of the tissue on a weekly basis, calculating the area, perimeter and depth of the lesion. These measurements are made by the doctor manually leading to high measurement inaccuracy. This leads to inappropriate monitoring, especially in detecting small variations in parameters and introducing an error (up to 40 % absolute variation) due to the experience of the physician and the human eye. These measurements are carried out manually with the help of instruments such as a ruler and gauge in order to measure the two longitudinal and transverse axes of the lesion. Then the two parameters are used to calculate the area and volume of the lesion. Frequently, the lesion does not have a perfect geometrical shape and therefore each value is calculated by approximating the lesion to be either circular or rectangular. A further method uses a thin film placed above the lesion to trace the perimeter and compute perimeter and area afterwards. The film is placed

on a millimeter grid for more precise measurement. However, wound's edges are often very difficult to be identified and therefore the evaluation strongly depends on the experience of the specialist. For this reason the measurement is subjective and characterised by low accuracy. An important parameter that nurses and specialists must derive from subjective visual evaluations is the classification of the lesion. The assessment is subjective and this is a problem when compared to an assessment as there are no common standards for doctors and health professionals to indicate whether the wound development is positive or negative. Another important factor is the time needed for the practitioner to derive all the parameters, therefore considering a life expectancy of at least 20 minutes. In order to improve the quality of the measurements and make them faster and as objective as possible, devices running algorithms capable of performing measurements have been developed over the years. A device capable of doing this is the Wound Viewer which will be described in the Section 1.4. The correct classification of a wound and its evolution would lead the doctors to prescribe the best treatment and a complete healing rate of 3-4 % faster for about 60 % of cases.

1.3 Wound classification

In wound's classification and definition of an appropriate treatment plan it is important to analyze all prognostic indicators. The most relevant parameters in assessing the wound's conditions are: change in wound's area or volume, tissue type and exudate control. The first two quantities are estimated manually whereas the evaluation of both the exudate and the tissue's granularity is classified by means of the Wound Bed Preparation (WBP) Score. The aim is to collect all the information about the wound in order to develop the best and fastest treatment strategy for identifying the most effective therapy for each wound. WBP uses the acronym TIME (Tissue, Infection or Infiammation, Moisture imbalance, Epidermal margin) to identify the parameters that the specialist has to evaluate.

- Tissue: the specialist is asked to analyze the wound's tissue by identifying the non-vital cells (necrotic or devitalized tissue) that need to be removed in order to recreate an appropriate environment for the growth of healthy tissue.
- Infection or Inflammation: prolonged inflammation and bacteria on the wound cause an increase in cytokines and protease activity. Cytokines are proteins able to bond to the cell membrane by receptors and give specific indications, such as cell differentiation or cell death. Proteases are proteins capable of catalyzing the disruption of the peptide link between the amino group (-NH_2) and the carboxylic group (-$COOH$) of proteins. The accumulation of these

molecules tends to slow down the activity of the growth factors and for this reason it is necessary to reduce bacterial infection and inflammation by protease inhibitors, anti-inflammatories and antimicrobials.

- Moisture imbalance: lesions are characterised by slower skin dehydration which must be counteracted by the removal of fluids. This limits possible aggravations due to an accumulation of fluid in the affected area such as maceration of the ulcer margins.

- Epidermal margin: keratinocytes are the most abundant cells in the epidermis, with protective functions against the aggression of pathogenic organisms. The wound healing process consists of the migration of these particles from the surrounding tissues towards the wound. For this reason, skin transplants, biological products and debridement are used in this case.

For the classification of the lesion, the specialist refers to the two tables 1.1 and 1.2 to assess the appearance of the wound and the amount of exudate present.

Score	Granulation	Fibrin	Necrosis
A	100%	-	-
B	50 -100%	+	-
C	< 50%	+	-
D	Quantity-independent	+	+

Table 1.1: Ulcer appearance

Depending on the extent of granulation, the wound is associated with the different bands (A, B or C). The lesions within which necrosis has formed belong to the D bandage, a condition in which the state of granulation is not assessed.
Another important factor in assessing the state of the skin lesion is exudate. It is the result of an inflammatory process, a fluid produced by the blood vessels, evidence that there is intense cellular activity in the area.

Once the anatomic-functional analysis has been carried out and the wound score has been assigned, the WBP classifier explains how to proceed. With these quantitative prerequisites established, the engineering of a support tool in this medical field was made conceptually possible.
To serve this purpose a device was developed: the Wound Viewer. This novel piece of medical equipment is able to evaluate the aforedescribed parameters through the use of artificial intelligence applied to an image of the lesion.

Score	State	
1	Under control	Little or no exudate. No demand for absorbent bandage; if clinically admitted the medication is changed on weekly basis.
2	Partially controlled	Moderate amount of exudate. It is necessary to change the bandage 2-3 times a week.
3	Out of control	Wound that produces a lot of exudate. It is necessary to apply absorbent medication once or several times a day.

Table 1.2: Wound exudate

1.4 Wound Viewer

Wound Viewer is a device for the treatment of skin ulcers developed by Omnidermal Biomedics, an innovative start-up founded in 2017 by three engineers. The project was followed by a doctor specialized in skin ulcers, Dr. Elia Ricci, and had an acceleration in 2017 thanks to the 'BioUpper Startup Accelerator'. This company is an Italian society that supports entrepreneurial projects helping them to enter more quickly and concretely into the market. Omnidermal received in 2018 in the presence of the President of the Republic Sergio Mattarella an important award, the "Leonardo Prize 2018". In this way the company has received international recognition, received a substantial number of pre-orders and caught the attention of an investor expert in the biomedical sector who has helped the company to become a reality thanks to an investment quantified in € 500k.

Figure 1.3: Wound Viewer

Wound Viewer is a class I, portable, fully automated device capable of performing an accurate and objective skin ulcer evaluation, providing the user with parameter values regardless of the operator using it.
The use of the device provides solutions to various problems encountered by professionals and patients in the vulnological field. In particular, the attention is focused on:

- Wound parameters;
- Data acquisition and storage;
- Data sharing;
- Time and costs;
- Granulation and tissue recognition;
- Measurement accuracy;
- Wound temporal evolution.

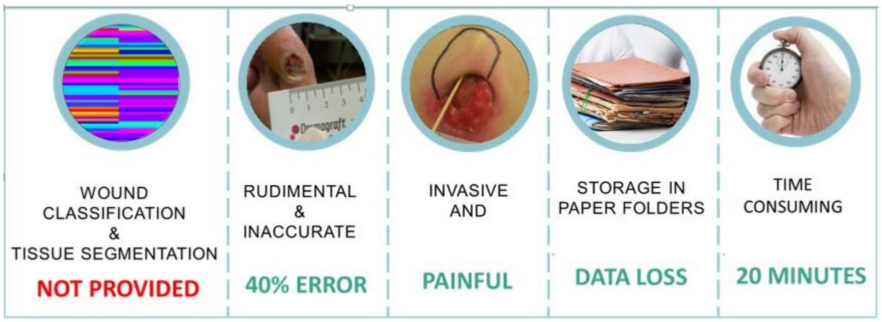


Figure 1.4: Manual measurement limits

Before the introduction of this device, the wound parameters were measured and calculated manually using invasive and painful methods for the patient. The longitudinal and transversal length of the wound were measured with a ruler and the area was calculated by approximating the wound to a known shape such as a rectangle or circle. This approach has shown an uncertainty up to 40 % on the parameters. In contrast, Wound Viewer uses an algorithm that is able to make measurements of the skin ulcer after the acquisition of wound's image. A 5MP color CMOS sensor is used for the acquisition phase. The information obtained from the device's high-precision IR distance sensors and the camera parameters are used within an algorithm that calculates the distance and therefore the depth of the lesion [4.6]. For a good analysis it is necessary a quality Picture set, made by well focused and

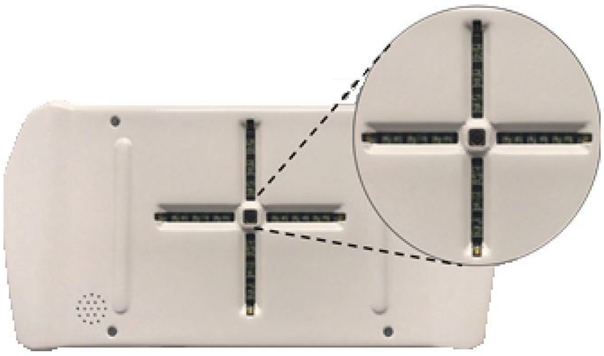


Figure 1.5: Wound Viewer sensors

correctly illuminated images. In order to have a good illumination of the wound and to avoid possible shadow cones that could change the image colors, the device uses a system consisting of 4 LED flashes. In the acquisition step all the leds are turned on to take well-lit and defined photos that will be used by the algorithm in the next step, the segmentation. In this phase the image is divided into different parts in order to separate the lesion from the background [4.6]. Once the lesion has been identified, the algorithm proceeds with the calculation of all the parameters of the wound in a more accurate way, overcoming the human error committed with manual measurements.
Data acquisition and storage has always been a major problem within the health care area. Today, more and more companies are providing a platform capable of storing, managing and storing data. All patient data has so far been maintained in paper medical file and made available only to the professional who treats the patient. Digitizing not only is mission-critical for the industry but is also a necessary progression for having data that are sharable, privacy-protected, unambiguous, objective, organized, quickly accessible and characterized by standardized processes. The software used by Wound Viewer makes possible to have an electronic medical history of each patient which stores all the data related to the patient, lesions and visits. The data is stored securely and privacy-protected in the cloud. The cloud is an online storage space, remotely accessible with different privileges for doctors and specialists. Each data stored on the device is synchronized and made available in real-time to authorized users. This resolves all the problems related to unreadable or missing documentation that commonly causes delays in the definition of a care plan and an increase in the possibility of having a wound becoming chronic in the meanwhile. Having an online data warehouse allows to have a clear idea of the time evolution of the wound because after each visit the acquired image of the wound is stored and as a consequence, the professional can make an evaluation of the prescribed therapy and confirm or modify it.

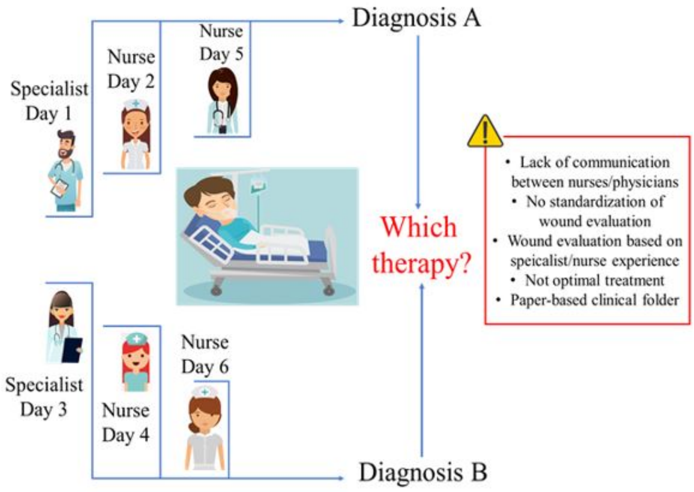


Figure 1.6: Communication problem in definition of treatment

The device has been designed for both hospital and home use and this is fundamental in relation to the cost analysis seen previously. Being able to carry out the home visit is a big step forward for the costs of hospitalisation and of the specialist's work hours. The examination can be made by a nurse directly at the patient's home, the data is then transmitted in real time to the doctor. This makes possible the exchange at distance of information and feedbacks thus optimizing costs and the amount of time spent providing assistance. The cost of a home visit by a nurse is around € 25 as opposed to the cost of a specialistic visit by a doctor, which is around € 150.
The application of the Wound Viewer for the domiciliary care allows constant communication between nurses and doctors (see Figure 1.7). Time optimization during a visit is crucial because time corresponds to important resources invested by the National Health System. The average time taken for a visit is 20 minutes. Measuring and calculating require time that the specialist could invest in other important activities. The optimization of time is fundamental and Wound Viewer is able to provide all the parameters related to the wound in 2 minutes only, reducing the time of 90 %. This time saving translates into considerable economic savings.
The introduction of an electronic device is useful to keep the information on wound evolution with a better quantification of the parameters such as the percentage of necrotic tissue or a more precise calculation of the area. The monitoring of the wound and all parameters over time is a fundamental step for the vulnological field to identify more precisely the type of tissue and consequently how the wound changes over time and how it should be classified according to standardized protocols. Wound Viewer has an additional feature: it provides the professional with an alert in case it

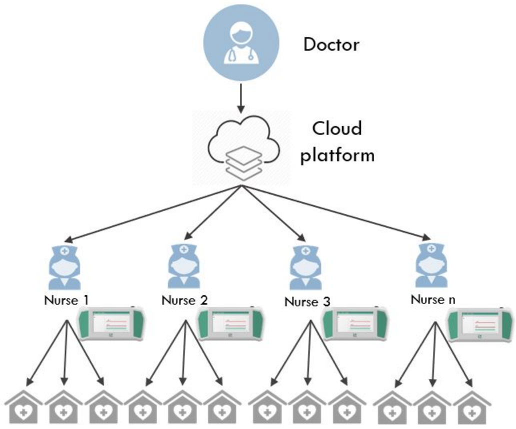


Figure 1.7: Wound Viewer domiciliary care

detects, after calculating the parameters, a dangerous lesion status. The alert helps the doctor to make a more accurate assessment of the ulcer and to decrease the risk that a dangerous parameter for the lesion status is erroneously neglected.
The device does not only calculate and keep track of the wound parameters but it also proposes the integration of an algorithm based on artificial intelligence able to use all the information collected during the training period to improve all its functionalities (calculation of granular tissue, area, volume) in order to obtain a trained classifier of the wound. Training is important because it allows the algorithm to acquire the necessary 'experience' and refine its ability to correctly recognize wounds based on all the information it has processed in the training period. The training of the algorithm is based on a multitude of data collected on ulcers and wound classifications performed by professionals. For this reason Omnidermal has started a program of clinical trials by providing the device to several public health care facilities (ASL Asti (Asti, Italy), ASL3 (Turin, Italy), Grenoble (France) and Madrid (Spain)) and collecting data on wound classification from a significant number of health care professionals. In this way, the device helps to standardize the classification, infact all the data collected represent the input to an Artificial Intelligence (AI) algorithm able to analyze them.

1.4.1 Wound Viewer algorithm

The Wound Viewer algorithm applies a Discrete Time Cellular Nonlinear Network (DT-CNN) calculation architecture. The algorithm is similar to a neural network, i.e. a model composed of artificial "neurons" which make up a data processing structure whose behavior is similar to biological nervous systems. Neurons are implemented as arbitrary and independent computing units. DT-CNN is a parallel

computing paradigm introduced by Chua and Yang.[4.7]
Each neural network goes through two main steps: the learning and the testing steps. In general, during the learning process the network is modulated in order to find the right conformation and the right connections for the type of input data while the test phase allows an evaluation of the algorithm performance.
The learning phase requires the use of an initial set of images to train the algorithm for wound recognition. The initial set must be composed of a large number of images. The algorithm is trained to perform automatic wound segmentation, i.e. to recognize the area of the ulcer within the two-dimensional color images.
This is performed by computing the number of occurrences g of each single RGB color in the wound area of the training set. The extracted statistical information about each color and its number of occurrences form the chromatic "knowledge" about the wounds.
The output of each neuron is computed by testing the image pixels in its neighborhood against the previously collected knowledge. The total number of pixels verified to be characteristic of a wound area are then compared to two thresholds. If the weighted counted number of pixels are higher than the two thresholds, then the pixel is reported to be part of a wound area (set to binary true) else it is rejected (set to binary false).[4.8] [4.9] The algorithm follows the equation:

$$O_{i,j} = h([\Sigma_{i-\frac{N}{2}}^{i+\frac{N}{2}} \Sigma_{j-\frac{N}{2}}^{j+\frac{N}{2}} h(g(I_{i,j}) - \theta)] - \rho) \tag{1.1}$$

$$\begin{cases} 0 \leq i \leq W-1 \\ 0 \leq j \leq H-1 \end{cases}$$

Where:

- I: original RGB image;
- O: black and white mask underlying the wound area;
- h: Heaviside function;
- N: integer number that defines the size of input's neighborhood;
- (i,j): coordinates of a single pixel;
- $I_{i,j}$: pixel RGB triplet code;
- $O_{i,j}$: binary output;
- θ: the cells threshold level;

- ρ: the automata threshold level;

The algorithm input is the pixel of original image I centered in (i,j). Considering a ROI of size $N \times N$ centered in (i,j), the algorithm counts how many pixels inside the ROI assume RGB values present in the training set's ulcer color database. Then, it compares the number of detected pixels with a θ threshold. If the number of counted pixels is higher than the θ threshold, the application of the Heaviside function returns 1, otherwise it returns 0. The number of pixels is compared with the second threshold ρ. If the numer of pixels is greater, the application of the Heaviside function returns $O_{i,j} = 1$, i.e. recognition of the pixel as part of wound, otherwise it will return $O_{i,j} = 0$. The result of this step is an image of the same size as the initial one representing the binary mask of the wound. Once the binary

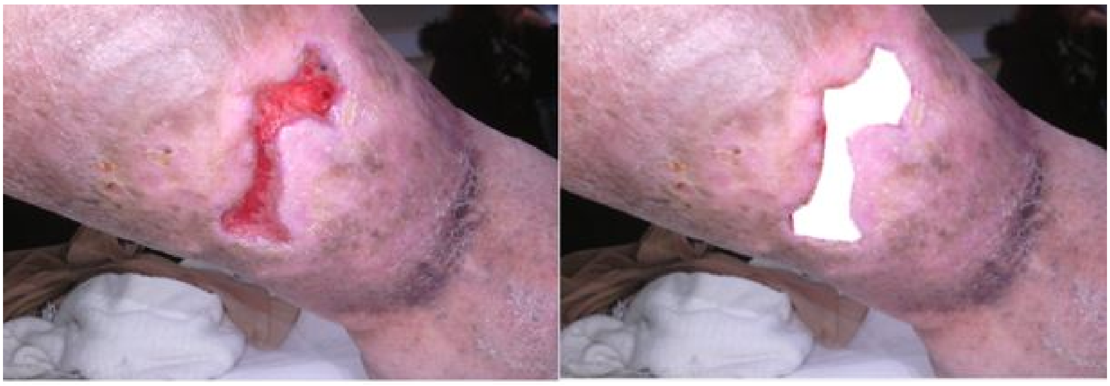


Figure 1.8: Result of Wound Viewer wound's area recognition

image is computed, the area of the recognized wound is computed by summing up the number of non-zeros elements of the mask and the wound's mask can be superimposed on the original image as shown in the image Fig.1.8.
The recognition of the tissue of the wound by the algorithm is used for the segmentation of the wound. Once the edge of the ulcer has been correctly recognized, it is possible to proceed with the calculation of parameters related to the lesion such as area in cm^2, depth in mm and granulation by WPB score. The area measurement is obtained from the reading of the external sensors while the wound depth is obtained by calculating the difference between these readings and the readings of the central sensors.
The area measurement is computed as

$$A = k * N_{pixels}$$

where N_{pixels} is the number of pixels identified as part of the wound and k is a distance dependant area density per pixel.

Figure 1.9: Result of Wound Viewer wound's edge recognition

The granularity of the wound is obtained as a result of the segmentation of the wound by evaluating each pixel recognized as part of the wound. Each pixel can be assigned to one of four macro groups: red, white, black and yellow. Given the granularity of the wound, the ulcer is then classified in terms of the WBP score. The described algorithm has been tested by comparing it with other methods in order to verify the accuracy with which it was able to detect the wound. The results show an accuracy of 97 % in Wound Bed Preparation with standardized and highly accurate morphological measurements of the lesions. The algorithm is also precise at identifying the lesion within a photo taken with flash, which has altered the values of the pixels making them brighter and therefore giving them higher values (remember that 1 corresponds to white and 0 to black). The system proposed by Omnidermal Biomedics is divided into two phases: initially there is the phase of data acquisition and storage, then the phase in which the algorithm just described is applied to identify the wound and its parameters. In the first phase are used: the camera, 12 white LEDs, a IR distance sensor and a temperature and humidity sensor. The algorithm identifies the pattern of each wound through the acquisition and analysis of several initial images. In the next phase, this pattern is used to test new images. In the acquisition of the first image of each wound all the LEDs are turned on and the stored photo is the main reference point because it is the one that will not present shadow nuances. Intuitively, the edge of each wound has a characteristic: it has higher vertical and horizontal gradient values in spite of the more homogeneous portion of tissue. The test phase showed that the higher gradient values are inside Blue and Green. After the analysis of a set of images, it has been demonstrated that the lesion is characterized by high values of the vertical and horizontal gradient

while the skin, characterized by a more homogeneous pigmentation, has lower values. Comparing the value of each pixel with a threshold (0.75) has first obtained a binary mask of the image. The values higher than the threshold 0.75 have been set to 1 and will constitute the seeds, while the lower ones have been set to 0 (black). For each element considered, it has been examined the neighborhood 5x5 going to attribute weights to the pixels of the neighborhood gradually decreasing as the distance from the pixel under examination increases.
Element C with weights attributed to the neighborhood:

$$C = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 0.5 & 0.5 & 0.5 & 1 \\ 1 & 0.5 & 0 & 0.5 & 1 \\ 1 & 0.5 & 0.5 & 0.5 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix} \tag{1.2}$$

Once the initial mask has been obtained, the algorithm averages of the surrounding area in order to include any false positives and obtain the binary mask of the final lesion. The sum of the 'positive' pixels of the mask permits to obtain the area of the lesion. For volume calculation, it uses other images acquired by illuminating the wound with some of the LEDs present on the device. The algorithm refers to four main colors for the analysis of the granulation: black corresponding to necrosis, red corresponding to granulation, yellow corresponding to fibrin and white corresponding to infection. Each colour has a degree of tolerance so as to include the pixels with the most similar colours and classify all the pixels of the lesion within the four categories described above. The algorithm test was performed on 200 patients, 36% of those were men and 64 % women, for a total of 213 ulcers. The ethnic groups of the patients were divided as follows: 95.5 % European Caucasian, 3 % African, 1.5 % Asian. The data obtained regarding the measurement of the area of the lesion carried out with this device showed an accuracy of over 85 % for more than 70 % of the wounds analyzed, demonstrating how reliable and useful the device can be for doctors, as well as not being invasive for the patient. The system developed by Omnidermal Biomedics is the result of a study carried out on memory systems which made possible to develop a skin ulcer detection system able to help the work of specialists proposing itself as a highly efficient tool in the treatment of ulcers with advanced data management.
The many advances have led to the question of what will be the next step in the field of vulnology, asking what will be another possible application of engineering to medicine. The answer might be in the study conducted in this thesis work with the proposal of an algorithm able to perform the 3D reconstruction of the wound from 2D images. The 3D model can provide a model that can perform morphological analysis of the ulcer.

Chapter 2

3D Reconstruction

2.1 3D Reconstruction introduction

In the last decade, the applications of engineering techniques to the medical sector have gained a lot of interest, pushing forward progress that would otherwise be impossible. Among the most impressive examples of integration between medicine and biomedical engineering one can cite 3D printing used in surgical planning such as the Hospital AM Solution system, or printing products to be implanted such as the bionic ear printed by Princeton University in 2013 [4.10] or the first skull transplant printed by University Medical Center of Utrecht in 2014 [4.11], the printing of custom limb prostheses or demonstration models. Other uses of 3D reconstruction are in virtual and augmented reality, providing specialists with a tool to practice with the use of surgical techniques.
3D reconstruction of objects has a large number of applications in diagnosis and monitoring in the medical field for the collection of anthropometric data. Having a tool capable of reproducing the object under examination in three dimensions makes the calculation of parameters (areas, volumes, chromatic distribution etc) objective and accurate. Reproducing human vision is a challenging process that requires the integration of optical, electronic and mechanical components in order to acquire images, process them and create a model. The most challenging part is the extraction of meaningful information from the acquired images. This challenge has been tackled by mean of physics-based techniques and probability theory so far. In recent years, research in this sector has developed a number of high quality algorithms that have enabled a much higher level of details. There are numerous techniques for the three-dimensional reconstruction of objects, from traditional photogrammetry to the new techniques derived from it such as *Computer Vision* (CV). Traditional photogrammetry reconstruction technique are widely used when the final goal is the metric accuracy of the reconstructed object. They are based on rigorous calculation

models allowing to estimate with quality and precision the necessary parameters. Computer Vision is an *image-based* branch of computer science, created to make the reconstruction automated and thus improving the computational time of operations. The procedure makes possible to automatically orient a large number of images of the object with geometry and dimensions not known a priori. Image-based techniques are becoming the main tools for the formulation of new algorithms aimed at improving performance both in terms of accuracy of the final result and computational time. *Structure from Motion* are methodologies derived from Computer Vision that have been developed for this purpose. They can obtain the final representation of the object as a point cloud distributed in a three-dimensional space. By using algorithms of *dense matching* it is possible to obtain point cloud with more and more points in order to represent the object in detail and allowing to make better measurements about wound's parameters. The goal of this thesis work research is to investigate the best *image-based* techniques in order to obtain an algorithm with good accuracy, calculation speed and performance for the reconstruction of skin ulcers.

The engineering approach is based on a detailed description of the problem and analysis of the characteristics of the object under investigation and of the known techniques in order to implement and evaluate the performance. Knowledge of the object to reconstruct from a physiological point of view is required together with a knowledge of reconstruction techniques already used in the past. After acquiring all the images of the object, information about the object will be obtained using the photographic parameters of the used camera such as focal length and depth of the sensor. If these parameters are not available a priori, it is anyway possible to obtain them by some algorithms during a calibration phase of the camera, as will be discussed in detail in the next chapter. Once the images are acquired and camera parameters are known, the next step is the Image Processing step. This is a fundamental step as it provides an a priori knowledge of the object inside the images. For example, in the case of 3D reconstruction in the architecture field, a search is carried out to find a certain shape in order to identify the building's facade from the background. In the case of this study, the knowledge concerns the shape, colour and depth of skin ulcers. The ulcers are distinguished from the rest of the skin by their shape and especially by their chromatic appearance. It is expected to find a graininess within the lesion and a shape designated by a rather evident edge. The knowledge of these characteristics is fundamental in the choice of the algorithm to be used. There are two main image-based reconstruction techniques of Computer Vision: Multiview and Photometric Stereo.

The *Photometric Stereo* is a technique capable of reconstructing 3D surfaces from frames acquired under different lighting conditions. The technique uses the analysis of the electromagnetic radiation flow of an object. It is computationally simple and

economically cheap because it requires only one camera. The acquisition is performed keeping the camera fixed and modifying for each frame the position of the light source. Once it acquired the picture set of the object, it reconstructs the map of the gradients and the map of the normals that are used to establish the orientation of each point of the object. This technique has limits related to the shape and the portion of the object to be reconstructed. The shape must be very regular in order to reflect light with equal intensity. Concerning the portion of the object to be reconstructed, it is intuitive to consider that it is impossible to make a 360° acquisition because the acquisition is performed with the fixed camera and only the portion of the object shot in the images can be reconstructed. This technique allows to acquire a large number of initial frames in order to have more information about the projection, at the expense of the computational load required. The Computer Vision technique able to overcome the limitations related to the partial reconstruction of the object is Multiview.
Multiview is a Computer Vision technique that reconstructs the 3D model of an object using multiple two-dimensional frames of the same object from different points of view. Depending on the modality adopted for the acquisition of images it is possible to obtain different levels of detail and it is possible to reconstruct totally or partially the object. The Multiview technique overcomes the constraint imposed by the position of the cameras and uses only one camera. The technique acquires multiple frames by changing the position of the camera for each shot.

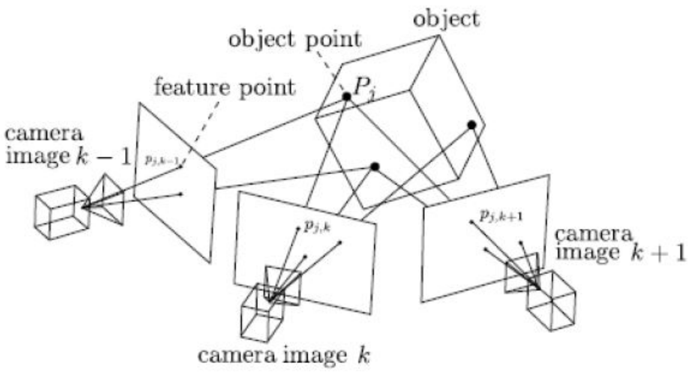


Figure 2.1: Multiview technique, from [4.12]

The use of this technique is economically affordable for everyone as the necessary instrumentation is a simple camera. The number of images needed for the reconstruction is not defined in advance but depends on the specifications required for the final result. Increasing the number of frames used increases the amount of information to be processed and consequently the computational time but there is a gain in terms of portion of object represented and detail accuracy of the final reconstructed object. By using only two photos, as in the stereo technique, it is possible to obtain the detailed 3D model of a portion of the object, losing the information at 360°. The camera position between one acquisition and the next is an important parameter for this technique to be evaluated during the acquisition of the initial frames. If the camera position between one shot and the next is changed significantly, as shown in the image below, there is a loss of the accuracy characteristic of this technique but it is possible to have a model that represents a larger portion of the object. If the two cameras are placed very close to each other, it results in a more accurate reconstruction of a smaller portion of the object, as shown in the figure below. The

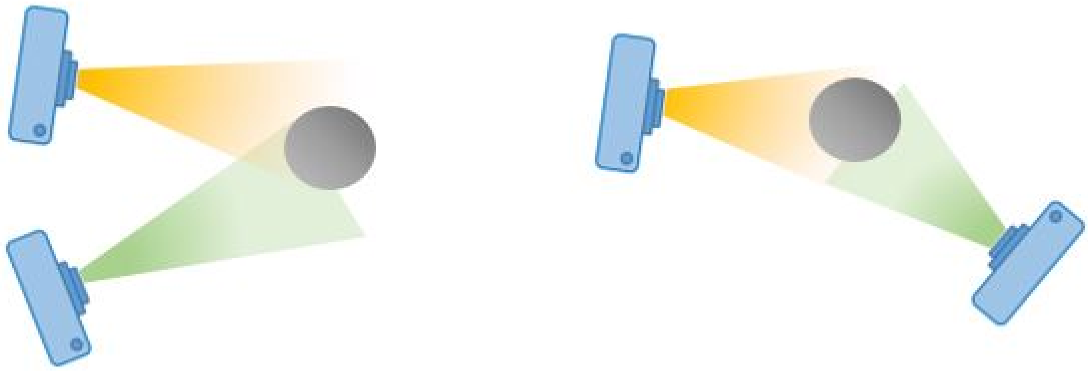


Figure 2.2: Partial reconstruction vs. lower accuracy

proposed algorithm is a Multiview image-based algorithm that uses only two input images of the object. The images are taken by moving the camera between the first and the second shot and keeping the object still. From these two images, it extracts bidimensional keypoints which are then used to predict the object's geometrical structure in the three dimensional space. In the triangulation step, the 2D points are placed in a 3D space to form a 3D point cloud. The camera used for the acquisition step is always the same and its parameters are obtained in the calibration step.
The proposed algorithm consists of the following main steps:

- Camera calibration;
- Images acquisition;
- Keypoints detection;

- Features extraction and matching;
- Essential matrix estimation;
- Camera Pose computation;
- Features densification;
- Triangulation;
- Surface reconstruction.

The algorithm has been implemented in Matlab.

2.2 Camera calibration

Calibration aims to find the correspondence between the 2D points of the image and the respective points in the 3D. During the camera calibration step an estimation is performed of the lens and camera sensor parameters. This estimation will be used to remove lens distortion from the images, derive the size of a captured object, determine the position of the camera in the scene and estimate the object's geometry. There are three groups of camera parameters, those are intrinsic parameters, extrinsic parameters and distortion coefficients. The extrinsic parameters represent the camera position in the three-dimensional space scene. They represent the transition from the world reference system to the camera reference system. Intrinsic parameters such as the optical center and focal length of the camera, represent the mapping between the geometric coordinates and the pixel coordinates of the image. Both extrinsic and intrinsic parameters are derived in the calibration step and they are used to switch from Camera coordinates to World coordinates and vice versa.

As shown in figure 2.3, the extrinsic parameters are used by the calibration algo-

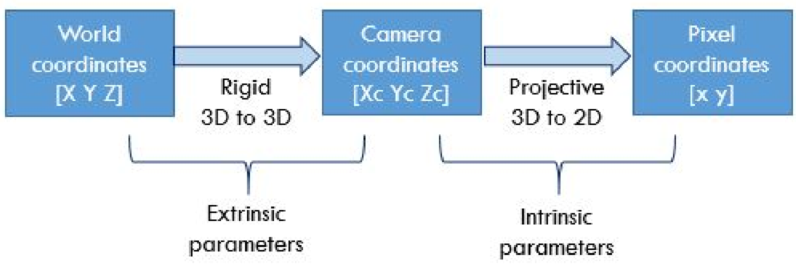


Figure 2.3: Coordinate switch using the camera parameters

rithm to switch from the real 3D coordinate system to the 3D coordinate system of the camera. Intrinsic parameters are used to switch from the 3D camera coordinates to the 2D image coordinates. With the camera calibration phase it is possible to obtain the extrinsic parameters that are the rotation matrix and the translation vector, while the focal length, the optical center and the inclination coefficients are obtained as intrinsic parameters. At last it is possible to obtain the information related to the radial and tangential distortion of the lens. The last two parameters are used to remove the distortion due to the lens from the image.
For the calibration step it has been used a Matlab toolbox called *Matlab Calibration Toolbox* [4.13]. Calibration is performed by recognizing known geometric patterns. In this phase, the camera was used to take several photos with different points of view of the pattern shown in the Figure 2.4. The toolbox works on different images

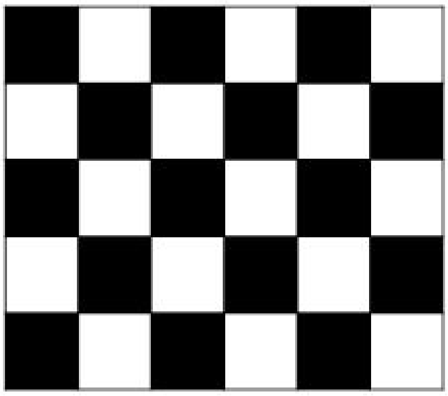


Figure 2.4: Pattern for calibration: the chessboard was printed on paper. A square side of 1 cm was chosen for each single square.

obtained by taking several pictures to a sheet with the printed pattern. The images are obtained by modifying camera position of an angle. In this case, the more pictures are taken, the more accurate the calibration output.
Equation 2.1 shows the relationship between the extrinsic and intrinsic parameters

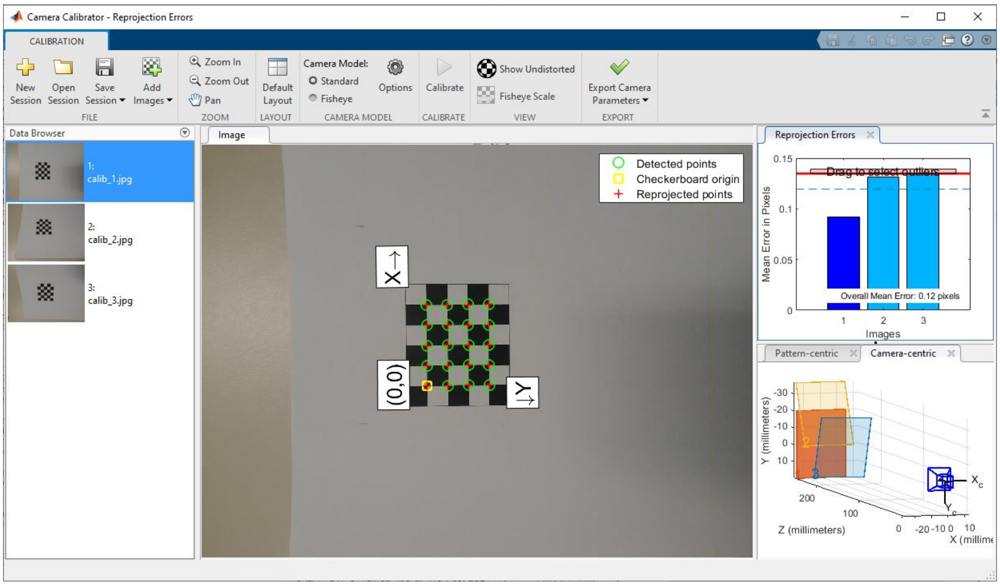


Figure 2.5: Matlab Camera Calibration Toolbox. On the left there are pictures of the pattern acquired with camera, on the right the retroprojection error is shown together with the relative position of camera and object in the three images.

of the camera used to obtain the Camera Matrix. The extrinsic parameters are represented by the Rotation matrix "R" and Translation "t" vector while the intrinsic parameters are expressed as the matrix K.

$$w \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = P \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \tag{2.1}$$

$$P = K \left[R|t \right] \tag{2.2}$$

$R \in \mathbb{R}^{3\times3}$, $t \in \mathbb{R}^3$, $K \in \mathbb{R}^{3\times3}$, $[R|t] \in \mathbb{R}^{3\times4}$.
where:

- ω: Scale factor
- $\begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$: vector of homogeneous coordinate of the image points
- $\begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$: vector of coordinates of world points

- P: Camera matrix
- $[R|t]$: Extrinsics matrix (Rotation and translation)
- K: Intrinsic matrix

2.2.1 The Intrinsic matrix parameters

The intrinsic matrix parameters include:

- Optical center C= (c_x,c_y): a point on the main axis of the lens where a light beam, passing through it, does not deviate.
- Focal length F: distance between lens optical center and sensor at which the image of an infinitely distant point is focused. It is typically expressed in millimeters.
- Coefficient of inclination α: indicates the angle between the x and y axis of the pixels that make up the sensor.
- Skew coefficient s, which is non-zero if the image axes are not perpendicular.

The K intrinsic matrix of the camera is defined as:

$$K = \begin{bmatrix} f_x & 0 & 0 \\ s & f_y & 0 \\ c_x & c_y & 1 \end{bmatrix} \tag{2.3}$$

where:

- p_x,p_y: size of the pixel in world units;
- f_x,f_y: focal length where:

$$\begin{cases} f_x & = \frac{F}{p_x} \\ f_y & = \frac{F}{p_y} \end{cases} \tag{2.4}$$

- s = $f_x \tan(\alpha)$, where α is the inclination coefficient shown in Fig.2.6.

During the calibration phase, the focal length and optical center are derived and used to estimate the K matrix as shown in (2.3).

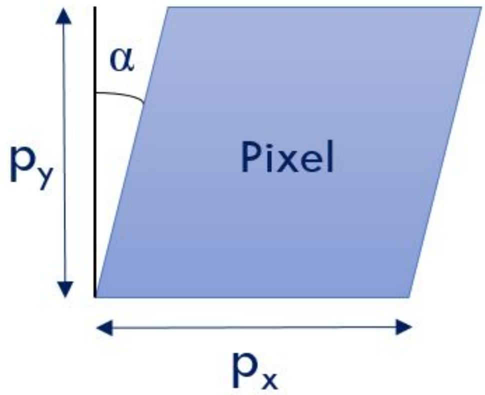


Figure 2.6: Pixel skew, with α inclination coefficient

2.3 Images acquisition

In this phase all the acquisitions are performed with the same camera that allows them to obtain digital colour data. A digital image is the numerical representation of a two-dimensional image. In the raster representation the image is a matrix consisting of elements called *pixel*. The number of basic elements of each image depends on the resolution of the camera used for the representation. Each digital image will consist of a Y number of rows and an X number of columns as shown in Fig.2.7. The resolution will be given as $X \times Y$ and each pixel will be characterized by the position (x,y) and a numeric coding depending on the type of image. The numeric encoding of pixel keeps the information about the color.

In the black and white image, each pixel has a binary representation, which can only assume two values: 0 corresponding to black and 1 corresponding to white. In this case the digital image is a matrix consisting only of 0s and 1s.
In grayscale representation, the image consists of a matrix of values coded on N bits per pixel, for a total of 2^N shades of gray.
In color representation, the digital image is the result of the overlapping of three matrices of the same size also known as channels. The three RGB matrices correspond to: R - red, G - green, B - blue. Each matrix consists of pixels encoded on N bits for a total of 2^N shades of color. The final color of each pixel is given by three values: each in the range [0,$2^{N-1} - 1$], as shown in Fig. 2.8. Today's cameras are capable to take high-resolution RGB images. For a good initial picture set it is necessary that the subjects of the photo are still, on the foreground, in focus and adequately illuminated, paying attention not to create shadows that would alter the colors of the image. This is crucial for a successful reconstruction of the

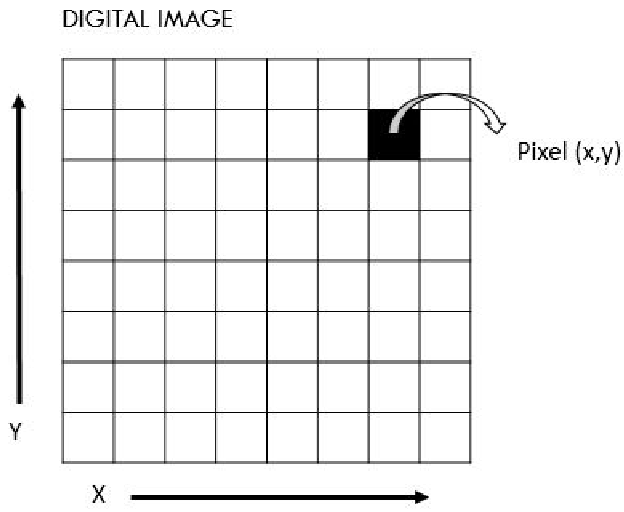


Figure 2.7: Digital Image: consists of $Y \times X$ pixels where Y represents the number of rows and X the number of columns. Each pixel is identified by the pair (x,y) of coordinates corresponding to the x-th row and y-th column.

object as an incorrect number of initial frames and inadequate image quality could lead to the failure of the reconstruction. In case of using an automatic camera, it is recommended if possible to take pictures in manual mode in order to avoid to change the parameters between one capture and the next. The camera position, as mentioned above, changes with little angle's adjusting between the first acquisition and the next one. After taking all the initial images, the algorithm goes to the next step and, if necessary, processes the image. During image processing it is possible to transform the image into shades of gray, change the brightness, contrast or any other

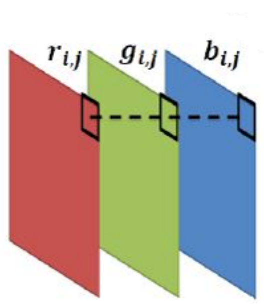


Figure 2.8: RGB image split into its three components

parameters until the object to be reconstructed is clearly visible and distinguishable from the background. The first step to obtain a good reconstruction is to have a quality picture set. Errors at this stage could lead to a failure of the reconstruction. On the one hand the small number of images necessary for the reconstruction has an extremely low computational cost but on the other hand this limits the possibility to obtain further important information for extracting the correct keypoints. The quality of the photo depends both on the camera characteristics and on the way the image is captured. A camera capable of providing high-resolution images will increase computation time and get more information from the images. The shooting technique used for the acquisition is also important to obtain images in which the object is in the foreground, well lightened and properly focused.
The images are acquired by holding the object of the reconstruction immobile and subsequently changing the angle of the camera between one acquisition and the next. In the previous chapter it has been illustrated the difference in case the camera angle between one shot and the next is very large or small. The way the initial picture set is captured depends on the subject of the pictures. In the case of skin ulcers, there is no need to reconstruct a 3D view of the entire affected body portion because the ulcers are located in a portion of the limb. Therefore, a rotation angle of 10-20 degrees provides already good overlap of the input images. The operator performing the acquisition must ensure that the patient keeps the limb affected by the lesion immobile during the shots. Another important parameter is the illumination. The lesion must be well illuminated and there must be no projected shadow cones that would change the colors of the image and cause the reconstruction to fail. Images are captured in manual mode with the flash activated so that the camera settings are not changed. Before capturing each image, make sure that the ulcer is in the foreground and in focus. A blurred image would lead to a failure of the algorithm in the feature matching phase. After the acquisition of the initial high-resolution color Picture set, the images are imported in .jpg format into the workspace using the *imread* function. RGB images are displayed within the workspace as three-dimensional matrices of values in which each cell of the matrices corresponds to the intensity levels of each pixel. For each pixel coordinate there is one intensity level in red, one in green and one in blue. The acquired images are processed to remove the distortion due to the lens that the camera introduces. The distortion can affect the accuracy of the final reconstruction. For this reason, the algorithm uses radial and tangential distortion parameters obtained during camera calibration to remove lens distortion. Removing the image distortion due to the lens means straightening the bent lines from the radial distortion of the lens. There is a Matlab function used in this step called *undistortImage* that can remove the distortion. The function takes in input the camera's parameters and a single picture to return an output image of the same size as the original one from which the distortion of the lens has been removed. This process needs to be performed for both the two images of the initial

Picture set. The algorithm calculates the position of each pixel of the input image within the new image with the following relationships:

- Radial distortion removal equations

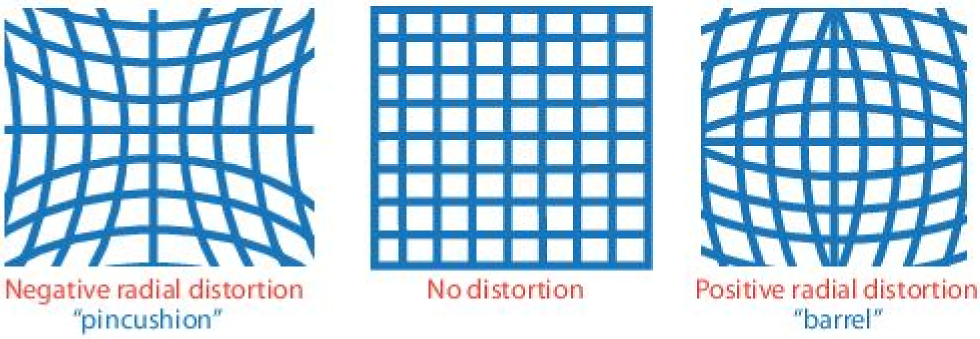


Figure 2.9: Radial distortion

$$\begin{cases} r^2 & = x^2 + y^2 \\ x_{\text{distorted}} & = x(1 + k_1 r^2 + k_2 r^4 + k_3 r^6) \\ y_{\text{distorted}} & = y(1 + k_1 r^2 + k_2 r^4 + k_3 r^6) \end{cases} \tag{2.5}$$

where x, y are the coordinates of the undistorted pixels. They are normalized non-dimensional coordinates calculated from pixel coordinates by translating to the optical center and dividing by the focal length. k_1, k_2, k_3 are the radial distortion coefficients of the lens. Typically, two coefficients are sufficient for calibration. For severe distortion, such as in wide-angle lenses, you can select three coefficients to include k_3.

- Tangential Distortion removal equations

$$\begin{cases} r^2 & = x^2 + y^2 \\ x_{\text{distorted}} & = x + (2p_1 xy + p_2(r^2 + 2x^2)) \\ y_{\text{distorted}} & = y + (p_1(r^2 + 2y^2) + 2p_2 xy) \end{cases} \tag{2.6}$$

where x,y are the undistorted pixel coordinates. They are normalized image coordinates calculated from pixel coordinates by translating to the optical center and dividing by the focal length. p_1 and p_2 are the tangential distortion coefficients of the lens.

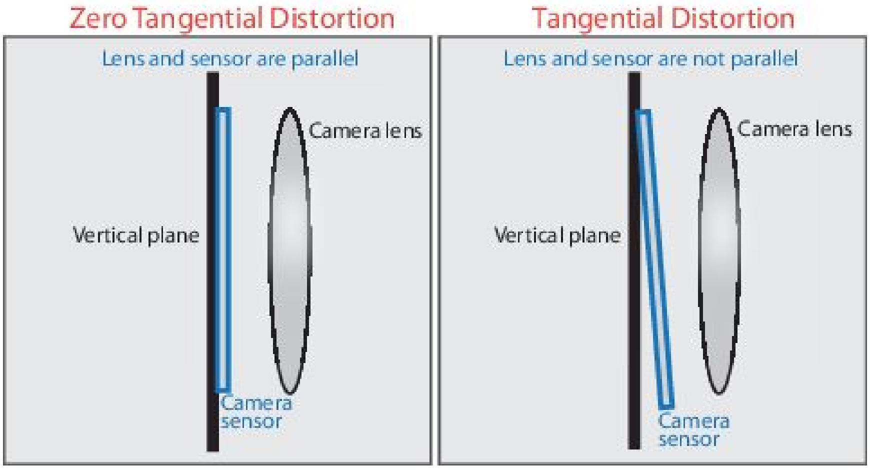


Figure 2.10: Tangential Distortion

2.4 Keypoint Detection

Once the entire initial picture set has been acquired, the images are analyzed individually in order to identify those points recognizable within the picture, the *keypoints*. Keypoints are those points of the image characterized by a high probability of being identified without ambiguity in each image. They will be used for the comparison between the different frames. There are many operators called *interest operators* able to perform this function with different mathematical models depending on the type of object one wants to recognize within the image. It is intuitive to think that the edges of a lesion, as well as of any object, are the most used points for recognition during this phase. The Multiview technique consists of a flow of initial images among which it is necessary to identify matches. It is necessary to assume that the points forming part of the object must be visible and distinguishable also from the other points of the same image.

2.4.1 Detect Corner Points inside the first image

The first step in features extraction is to identify those points within the image that are easily recognizable inside the other images. Searching for a pixel inside the second frame is not easy because of noise and because moving the camera leads to a different image intensity in the next frame. In order to be able to identify the point in the second image it is necessary that it has very evident/marked intensity characteristics. The search is not only local but uses a window of pixels around

the central pixel. Once identified the window in the first image I and the window in the second image J the algorithm performs a search for dissimilarity in order to minimize the residual error between the two. The algorithm used in this phase was developed by Shi and Tomasi in [4.14] and is defined by the "Minimum Eigenvalue Algorithm" as it uses eigenvalues and a threshold to mark a given pixel as corner (keypoint) or not.
The image I is transformed in grayscale with Matlab's function *rgb2gray* and then it is used for the keypoints search. The goal of this method is to find those regions called "windows" inside the image whose displacement generates a great variation in terms of itensity. Consider taking an image patch over the area $\mathbf{x} = (x{,}y)$ and shifting it by $\mathbf{z} = (u{,}v)$. The weighted sum of squared differences E between the intensity of these two patches is:

$$E(\mathbf{z}) = \sum_{\mathbf{x}} w(\mathbf{x})[I(\mathbf{x}+\mathbf{z}) - I(\mathbf{x})]^2 \tag{2.7}$$

where $w(\mathbf{x})$ is the window in position $\mathbf{x}$ that acts like a mask, ensuring that only the portion of the image in the window is used. The image gradient ∇I is defined as:

$$\nabla I = \left(\frac{\partial I}{\partial u}, \frac{\partial I}{\partial v}\right) \tag{2.8}$$

To maximize the difference function E, the method approximates by a Taylor expantion the function I as follow:

$$I(\mathbf{x}+\mathbf{z}) \approx I(\mathbf{x}) + \nabla I(\mathbf{x})^T \mathbf{z} \tag{2.9}$$

Therefore, equation (2.7) becomes:

$$E(\mathbf{z}) = \sum_{\mathbf{x}} w(\mathbf{x})[\nabla I(\mathbf{x})^T \mathbf{z}]^2 \tag{2.10}$$

or in a matrix form

$$E(\mathbf{z}) = \mathbf{z}^T A \mathbf{z} \tag{2.11}$$

where A

$$A = \sum_{\mathbf{x}} w(\mathbf{x}) \nabla I(\mathbf{x}) \nabla I(\mathbf{x})^T \tag{2.12}$$

To assign a pixel as a corner it is necessary to calculate the λ_1 and λ_2 eigenvalues of A. If both eigenvalues are small the region is quite uniform then there is no feature of interest. If eigenvalues are one large and one small then an edge is found. To identify the pixel as a corner both eigenvalues have to be large. Assuming both are large, the algorithm compares the minimum of the two with a λ threshold value. The threshold value is assigned as follows:

- Eigenvalues are calculated in a rather uniform region of the image in order to find the lower limit of λ.
- Eigenvalues are calculated in a high texture image region to identify the upper limit of λ.

The threshold must be chosen in order to identify more or less points.

$$\min(\lambda_1,\lambda_2) > \lambda \tag{2.13}$$

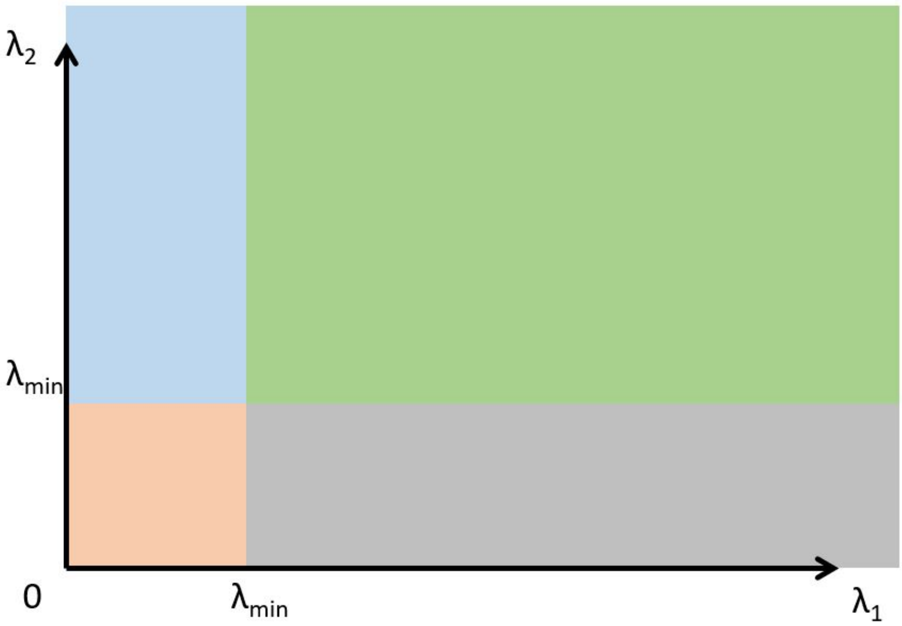


Figure 2.11: Corner allocation regions

- Green region: pixels accepted as corners;
- Blue and Grey regions: λ_1 or λ_2 is less than the minimum required.
- Red region: both λ_1 or λ_2 are below the minimum required.

2.5 Feature extraction and matching

After identifying the keypoints in the image, the process moves on to the features extraction phase. Each key point is described by *image descriptors* by using vectors. The vectors contain the highly descriptive local characteristics that, depending on the type of algorithm used, are more or less invariant to the alterations of scale or other characteristics of the image. Each descriptor must consider the variation in scale, orientation and illumination within the image. Depending on the mathematical model used to perform this function, there are several operators more or less sensitive to changes in scale, illumination, rotation or translation. The extraction and description of keypoints is based on the principle that each image is reproduced in different resolutions and worked with different filters in order to highlight details at different scales. This creates the image pyramid *space-scale*. Each layer is analyzed to extract the descriptor by analyzing the neighborhood of each pixel, calculating the magnitude of the gradient and the orientation of each keypoint. In this way it is possible to obtain the orientation histograms to describe the directions of the gradient in a neighborhood sub-region.
There are several algorithms that can be used at this phase to calculate different descriptor parameters that differentiate each other depending on the choice of the neighborhood and the sub-regions used.
After identifying and describing all the keypoints of each image, it is necessary to identify the matches between the images. Regardless of the number of images in the initial picture set, the match is made by considering one pair of images at a time. The similarity between the images, i.e the points matched between them, is evaluated using the descriptors. Each pixel uses a descriptor that is exploited to find the corresponding point in the second image. The correspondence can be found using the Euclidean distance between the descriptors. The exact match is unambiguous but it is possible that the search of the point to be associated will result in several matches and consequently a series of outliers will have to be removed later. From the computational point of view this process is very expensive and the time increases if increase the number of images used as input or the number of identified and described points increases for each image. A too strict search leads to have few matches and consequently an insufficient number of points for the reconstruction, on the contrary a not rigid search would lead to a high number of outliers.
In this phase it is important to limit the number of incorrect matches that would lead to the failure of subsequent reconstruction steps. For this reason it is performed a *bundle-adjustment* to minimize the overall error of repetition of the 3D points in the images by keeping only the strongest matches.
In this step it has been used the Kanade-Lucas-Tomasi (KLT) algorithm for feature extraction and matching.[4.16]
The displacement of the camera often leads to variations in terms of light on the

object. This may result in a variation of the pixel intensity among the corresponding areas of the two images and therefore the impossibility to recognize the same pixel in the second image, unless particular characteristics are evident. Moreover, it is also necessary to consider noise interference with the recognition of pixels. For this reason the search for matches cannot be performed on a single pixel but becomes a search for a window of pixels that have the required consistency. The recognition is done by comparing the window in the first image and the one in the second image. To recognize a pixel it is necessary that the two windows are similar, if this is not the case then the window is discarded. The size of the window depends on the number of parameters to estimate. A larger window allows you to estimate a higher number of parameters but increases the possibility that pixels will not be recognized for the aforementioned reasons. On the contrary, a smaller window reliably estimates only a few parameters. Let us now go deeper in details on the method.
The first image I, called reference image, contains a feature window w. The feature is tracked in the second image J, called current image. The coordinates of all pixel are $\mathbf{x} = (x,y)$. The algorithm calculates first the difference of the image intensity at the corresponding matched pixel position. Then it finds the parameters able to minimize the difference function.
The difference motion equation is:

$$E_{KLT}(d) = \sum_{\mathbf{x} \in w} (I(\mathbf{x}) - J(\mathbf{x} + d))^2 \tag{2.14}$$

However, the Kanade-Lucas-Tomasi algorithm iteratively updates the parameters of the model by considering a slightly different error function:

$$E_{KLT}(\Delta d) = \sum_{\mathbf{x} \in w} (I(\mathbf{x}) - J(\mathbf{x} + d + \Delta d))^2 \tag{2.15}$$

with $d,\Delta d \in R^2$.
Expanding $J(x + d + \Delta d)$ in $\mathbf{x} + d$ at the first order the equation become:

$$J(\mathbf{x} + d + \Delta d) = J(\mathbf{x} + d) + \nabla J(\mathbf{x} + d)^T \Delta d \tag{2.16}$$

Substituing (2.16) in (2.15) the error equation becomes:

$$E_{KLT}(\Delta d) \approx \sum_{\mathbf{x} \in w} (I(\mathbf{x}) - \nabla J(\mathbf{x} + d)^T \Delta d - J(\mathbf{x} + d))^2 \tag{2.17}$$

and in order to find its minimum the function is derived and set equal to zero. This results in:

$$\frac{\partial E_{KLT}}{\partial \Delta d} = 0 \tag{2.18}$$

$$\sum_{\mathbf{x}} \nabla J(\mathbf{x}+d)(I(\mathbf{x}) - \nabla J(\mathbf{x}+d)^T \Delta d - J(\mathbf{x}+d)) = 0 \tag{2.19}$$

and the solution is given by solving in terms of Δd.

$$\Delta d = H_{KLT}^{-1} \sum_{\mathbf{x} \in w} \nabla J(\mathbf{x}+d)(I(\mathbf{x}) - J(\mathbf{x}+d)) \tag{2.20}$$

where

$$H_{KLT}^{-1} = \sum_{\mathbf{x} \in w} \nabla J(\mathbf{x}+d) \nabla J(\mathbf{x}+d)^T \tag{2.21}$$

The value of the parameter d obtained is the one that minimizes the error function. The parameter d is used to obtain the position of the pixel in the second image $J(\mathbf{x}+\mathbf{d})$ corresponding to pixel $I(\mathbf{x})$ of the first image. This procedure is performed for all the keypoints identified in the first image in the previous step.

2.6 Essential Matrix Estimation

The Epipolar geometry described by the camera in two positions expresses the fundamental relationship between any pair of points in the planes of the two images and carries the main constraints related to the coordinates of these points. Epipolar geometry is independent of the structure of the scene and is defined by the relative position of the camera in the acquisitions. If we take two views, the Epipolar geometry is the intersection between the planes on which the three-dimensional points of the scene are projected with respect to the center of the camera and the bundles of planes having as a common line the line joining the centers C and C' shown in Fig. 2.12. where:

- X: point in 3D-space;
- x: point X in the first image;
- x': point X in the second image;
- C, C': camera centers in image 1 and 2;
- π: epipolar plane;
- l:*epipolar line*, intersection between the plan π and the first image;
- l':*epipolar line*, intersection between plane π and the second image.

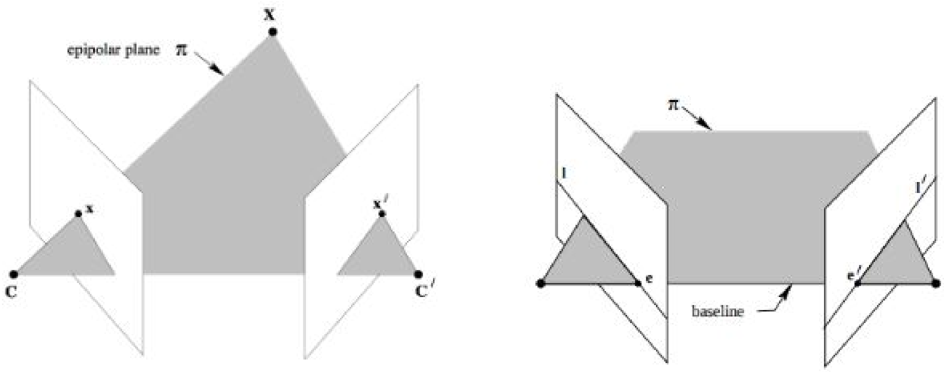


Figure 2.12: Epipolar Geometry.

x, x' and X are coplanar and satisfy the following equation:

$$\overrightarrow{Cx} \cdot (\overrightarrow{CC'} \times \overrightarrow{C'x'}) = 0 \tag{2.22}$$

which is also known as *Epipolar constraint*. The Epipolar constraint is helpful for matching corresponding points in a pair of images. However, it is interesting to highlight that it is possible to proceed even in the inverse direction: to infer Epipolar geometry from point correspondence. The Epipolar Geometry has an algebraic representation that is a 3×3 matrix called the *Fundamental matrix*.
Considering a point x on the first image, the aim is to look for the corresponding point x' on the other image. Given a pair of images, for any point x in one image there exists an *Epipolar line* l', intersection of the plane π and the second image where the matched point x' must lie. This is true also in the opposite direction. For this reason we can consider a map:

$$x \mapsto l' \tag{2.23}$$

from a point in one image to its corresponding Epipolar line in the other image. This map is a projective mapping from the points to lines which is represented by the Fundamental matrix F, a 3×3 rank-2 matrix.
If x and x' correspond, then x' lies on the Epipolar Line $l' = Fx$ corresponding to x. This means that

$$(x')^T l' = (x')^T Fx = 0 \tag{2.24}$$

This represents a necessary condition for points to correspond.
The equation (2.24) gives a way of characterizing the Fundamental matrix only in

terms of corresponding images, without mentioning the intrinsic K or extrinsics $(R|t)$ camera parameteres. Various methods are proposed for computation of the Fundamental matrix however the Epipolar equation can be rewritten as a *linear* and *homogeneous* equations using:

$$\begin{bmatrix} x'_i & y'_i & 1 \end{bmatrix} \begin{bmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{bmatrix} \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix} = 0 \tag{2.25}$$

$$x_i x'_i f_{11} + x_i y'_i f_{21} + x_i f_{31} + y_i x'_i f_{12} + y_i y'_i f_{22} + y_i f_{32} + x'_i f_{13} + y'_i f_{23} + f_{33} = 0 \tag{2.26}$$

for m correspondences the equation is:

$$\begin{bmatrix} x_1 x'_1 & x_1 y'_1 & x_1 & y_1 x'_1 & y_1 y'_1 & y_1 & x'_1 & y'_1 & 1 \\ .. & .. & .. & .. & .. & .. & .. & .. & .. \\ x_m x'_m & x_m y'_m & x_m & y_m x'_m & y_m y'_m & y_m & x'_m & y'_m & 1 \end{bmatrix} \begin{bmatrix} f_{11} \\ f_{21} \\ f_{31} \\ f_{12} \\ f_{22} \\ f_{32} \\ f_{13} \\ f_{23} \\ f_{33} \end{bmatrix} = 0 \tag{2.27}$$

The above equation can be solved with a number of matches between the two images by considering at least 8 matched points (i.e. m=8).
In matrix form (2.27) can be written in:

$$Af = 0 \tag{2.28}$$

and by ignoring the rank-2 constraint, this equation can be solved by considering a least-squares approach:

$$\min_f ||Af||^2 \tag{2.29}$$

where the vector f is defined up to a constant and it has to be different from the trivial solution $f = 0$. The solution for f that solves the optimization problem in (2.6) corresponds to the eigenvector of the smallest eigenvalue of AA^T. However f is singular and rank 2. Since f is computed based on discretized noisy data, these properties are not easily satisfied by the current solution. Therefore by performing a Singular Value Decomposition (SVT) of the estimated F:

$$F = U\Sigma V^T \tag{2.30}$$

$$\Sigma = \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & \sigma_3 \end{bmatrix} \tag{2.31}$$

it is possible to force the last diagonal eigenvalues of Σ to be zero and then recompute a new F as:

$$F = U\Sigma' V \tag{2.32}$$

where

$$\Sigma' = \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{2.33}$$

Suppose now that the camera is calibrated and the matrix K is know, than for every corresponding pair of points it holds that

$$(x')^T E x = 0 \tag{2.34}$$

$$E = K^T F K \tag{2.35}$$

where E is the essential matrix which can be decomposed into:

$$E = [t]_x R \tag{2.36}$$

where R is the rotation matrix and $[t]_x$ is a matrix that contains the translation components:

$$[t]_x = \begin{bmatrix} 0 & -t(3) & t(2) \\ t(3) & 0 & -t(1) \\ -t(2) & t(1) & 0 \end{bmatrix} \tag{2.37}$$

Here, again $det(E) = 0$ (i.e. rank 2), and in addition the two non-zero singular values are equal. For arbitrary scaling reasons it is possible to assume that they are both equal to 1. The essential matrix is used to define the two camera poses. It is possible that a pair of matched points does not respect the Epipolar Geometry and for this reason the two points are rejected. The matchings that respect epipolar geometry, i.e. the "inliers points", are kept for the next step.

2.7 Camera Pose computation

Once the estimation of the essential matrix is finally performed, the motion can be recovered with an SVD Decomposition. The matrix E can be decomposed as:

$$E = Z \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} Q^T \tag{2.38}$$

when Z, Q an orthonormal matrices.
Since $E = [t]_x R$, the next step is to factor E into the product of a skew symmetric matrix (i.e. $[t]_x$) and a rotation matrix (i.e. R).
After the definition of the matrix:

$$D = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{2.39}$$

The extrinsic parameters location and orientation of the camera in the second frame relative to the first one are obtained as:

$$R_2 = ZDQ^T \tag{2.40}$$

or

$$R_2 = ZD^TQ^T \tag{2.41}$$

$$t_2 = \begin{pmatrix} Z_{13} & Z_{23} & Z_{33} \end{pmatrix}^T \tag{2.42}$$

or

$$t_2 = -\begin{pmatrix} Z_{13} & Z_{23} & Z_{33} \end{pmatrix}^T \tag{2.43}$$

All these four combinations of t_2 and R_2 satisfy the epipolar costraints but the geometrically correct solution is the one that produces points in front of the camera of the two frames. Once obtained the the extrinsic parameters, it is possible to compute camera matrix. The camera in the first image is at the origin looking along the Z-axis. Thus, its rotation matrix is identity and its translation vector is null.

$$R = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{2.44}$$

and

$$t = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix} \tag{2.45}$$

The camera matrix is therefore defined as

$$P = K(R|t) \tag{2.46}$$

where R and t correspond to (2.44) and (2.45) for the camera in the first image and to R_2 and t_2 for the camera in the second image. The two camera matrices will be used in the triangulation phase to obtain the 3D position of the image points.

2.8 Features densification

The first part of the algorithm allowed to obtain the location and orientation of the camera in the second image relative to the first image. To obtain a dense point cloud in three-dimensional space, it is necessary to perform a "Features densification" and then proceed to the triangulation.
In the *Structure From Motion* there are two categories of triangulation algorithms: Incremental (or sequential) algorithms and Global algorithms.

- Incremental algorithms reconstruct the point cloud from an initial pair of images and iteratively add one image at a time, independently of the number of images in the picture set. All points identified as keypoints are placed in the three-dimensional space to estimate the position and orientation of the camera. Each time a new image is introduced, the steps listed above are repeated and a *bundle adjustment* is performed to minimize the error. For a good final result it is good to choose as initial image pair the one with the highest number of matches found in the previous steps.

- The global algorithms perform all the steps in a single one by considering all the images with the matched features and performing a single final bundle-adjustment. The positions and rotations of the cameras are also calculated in a single step, which makes the algorithm very expensive from a computational point of view.

The proposed algorithm makes use of only two images of the object so there is no necessity to make a choice of the pair of images to be considered. The three-dimensional point cloud must have a very high number of points. This is obtained by applying a point cloud *densification* algorithm.
The first step is to perform a new keypoints search inside the first image. The search is done as described in section 2.4 using the "Minimum Eigenvalue Algorithm" algorithm with lower λ eigenvalue threshold in order to select more points as corners. This operation will require more computational times and more computational power in the next steps. The algorithm creates a ROI that excludes from the search only the external pixel frame of the image in order to ignore any keypoints on the edge. The new points are used to perform "Features extraction and matching" with the same method described in 2.5, using the KLT algorithm.

2.9 Triangulation

For the points Triangulation the algorithm uses the matched points and the two projection matrices of the camera in the first and the second image.
The projection matrix of the camera is used to project the points of the 3D space in homogeneous coordinates into the corresponding point in the camera's image. The camera matrix describes the location and orientation of camera in world coordinate system. Camera matrix is calculated using the camera parameters, rotation matrix and translation vector found previously and returns a nonsparse numeric 4×3 matrix.
The first camera pose is assumed at the origin along the z-axis so it is characterized

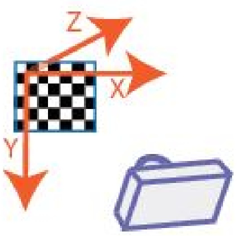


Figure 2.13: Camera matrix estimation, from [4.15]

by a rotation matrix which is an identity matrix and a null translation vector. Calling P the camera matrix, in the stereo triangulation step the equation (2.1) is developed in this way:

$$w \times \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix} = \begin{bmatrix} P_1 \\ P_2 \\ P_3 \end{bmatrix} \times X_i \tag{2.47}$$

with X_i world point cordinates and P_1, P_2, P_3 rows of the matrix P

$$\begin{cases} wx_i &= P_1 X_i \\ wy_i &= P_2 X_i \\ w &= P_3 X_i \end{cases} \tag{2.48}$$

replacing the third equation of (2.48) in the first two become:

$$\begin{cases} P_3 X_i x_i &= P_1 X_i \\ P_3 X_i y_i &= P_2 X_i \end{cases} \tag{2.49}$$

the two equations become:

$$\begin{cases} P_3 X_i x_i - P_1 X_i &= 0 \\ P_3 X_i y_i - P_2 X_i &= 0 \end{cases} \tag{2.50}$$

Using the constrints from both images, it is possible to riformulate (2.50) as the following linear equation

$$\begin{cases} [P_3 x_i - P_1] X_i &= 0 \\ [P_3 y_i - P_2] X_i &= 0 \\ [P'_3 x'_i - P'_1] X_i &= 0 \\ [P'_3 y'_i - P'_2] X_i &= 0 \end{cases} \tag{2.51}$$

for $i = 1,...m$ for the m correspondence.

where

- $P = (I,0)$ camera matrix in the first image;
- P_1, P_2, P_3 are the rows of matrix P.
- $P' = (R,t)$ camera matrix in the second image;
- P'_1, P'_2, P'_3 are the rows of matrix P'.
- $x = (x_i,y_i)$ point in the first image;
- $x' = (x'_i,y'_i)$ point in the second image;
- X: the corresponding 3D point in the "World space".

This system can be solved using again the least squared error method.

The algorithm proposed resulted in dense point clouds in three-dimensional space.

2.10 Surface Reconstruction

The next step is process the point cloud in order to remove any remaining defects created during the reconstruction of the 3D model.
The errors that have the greatest impact on the surface reconstruction are shown in the Fig.2.14.

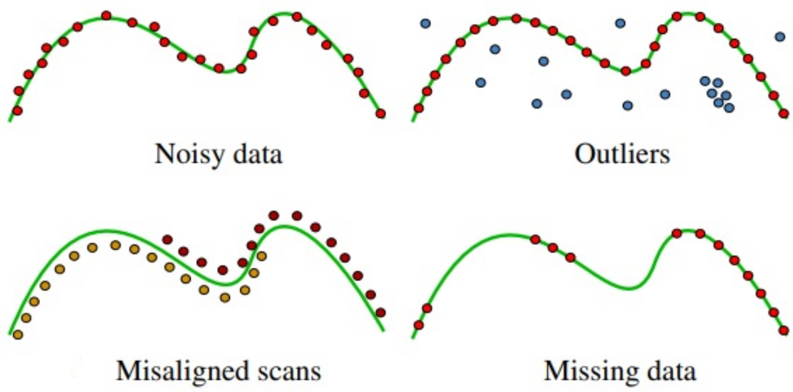


Figure 2.14: Point cloud artifacts on 2D line, from [4.17]

- Noisy data: those points that are not really part of the object that are randomly distributed on the surface. These points can be removed by a quite rigid surface model.
- Outliers: those points belonging to the true surface that have been incorrectly positioned in the triangulation step. Their displacement from the correct position is sometimes very high. Therefore, these points should not be used for surface reconstruction. This is performed by choosing a more rigid surface model.
- Misalignment: points that belong to the true surface but have reported an error in triangulation step that resulted in a small final positioning error.
- Missing data: having missing pixels can be traced back to different phases of the algorithm. It is possible that the pixel was not selected as keypoint in the first image, it did not find a match in the second image or it may have been eliminated by inlier points in the essential matrix estimation phase.

To reduce the listed artifacts it is necessary to process the point cloud in order to remove any duplicate points (points with the same coordinates (x,y,z)) that do not

give additional information but make the computational load heavier. Any errors in the reconstructed surface and any noise could be removed by filters.
After the artifacts removal, the point cloud can be processed to obtain the surface reconstruction by creating a surface mesh. There are several open-source software that can be used for this purpose, let us see some examples:

- Greedy Projection Triangulation: the surface is created considering three close points. They are used to create the initial triangle. Starting from the first triangle, the algorithm connects all the adjacent triangles until the surface is completed. The algorithm uses a threshold distance to assign the point to the nearest surface. The threshold is editable and can be set according to the density of the region. If the point cannot be assigned as part of the surface, then the algorithm generates a new triangle and starts to expand it again.

- Ball Pivoting: the surface is reconstructed from three connected points to form a triangle. The algorithm performs a control by a sphere of radius ρ that touches the three points. The sphere is rotated around one side until it meets another point. Once found, the point is connected to the other two points to form a new triangle and so on until the entire surface is reconstructed.

The quality of the mesh reconstruction depends on the choice of the algorithm adopted and on the quality of the pre-processing performed on the point cloud. For this reason it is necessary to perform tests with different algorithms in order to identify the best one for each step. Once the surface has been reconstructed, it is possible to proceed with the calculation of the wound parameters, as shown in the last chapter.

Chapter 3

Application of the algorithm to skin ulcer images

3.1 Camera calibration

The algorithms described in Chapter 2 have been applied to the acquired images of skin ulcers. The images were acquired on patients' wounds through the camera of a smartphone. The model of the mobile phone is LG G6 and all the images were acquired with the same phone. For this reason the calibration of the camera has been performed only once and all the parameters related to the camera have been used in the algorithm also for the subsequent acquisitions. The photos in Fig.3.1 represents a chessboard used for the calibration phase. The side of each square is 1 cm wide and its precise measure is required as input to the Camera Calibrator Toolbox.

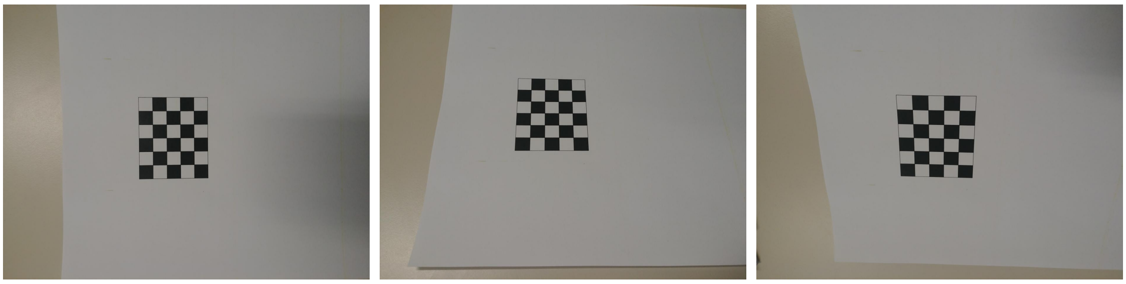


Figure 3.1: Picture set for calibration: the chessboard was printed on a paper sheet and several images were captured at different camera rotation angles.

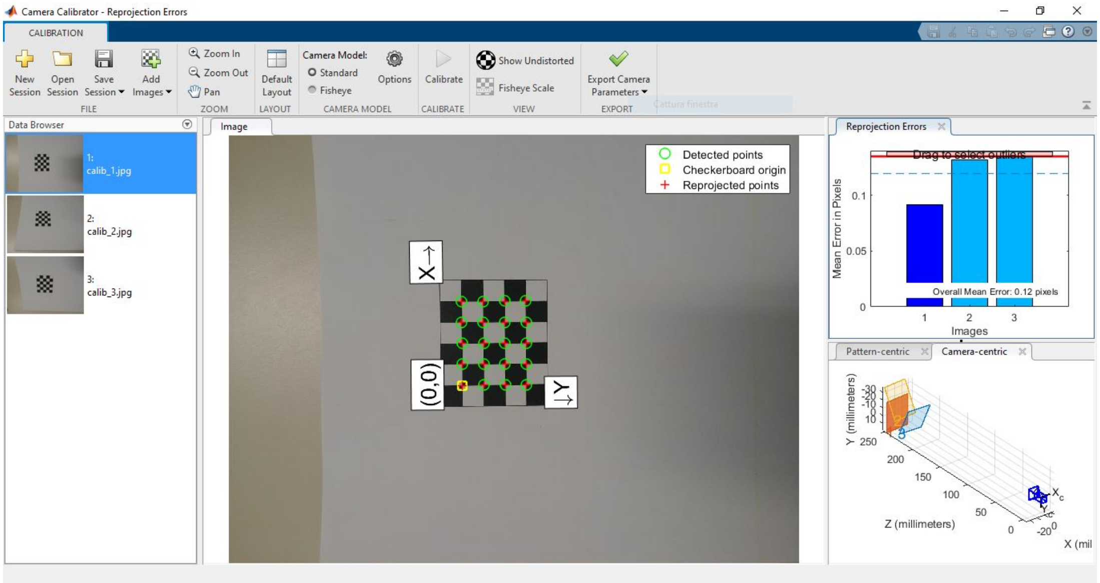


Figure 3.2: Calibration toolbox screenshot

Camera Parameters:

- Radial distortion coefficients: $\begin{bmatrix} 0.4475 & 1.8286 \end{bmatrix}$
- Tangential distortion coefficients: $\begin{bmatrix} 0 & 0 \end{bmatrix}$
- Focal lenght: $(f_x, f_y) = (3.8657,\ 3.7983)$
- Image size: 3120×4160.
- Optical center (or "Principal Point"): $(c_x, c_y) = (2.1220,\ 1.5239)$
- Skew coefficient s: 0
- Intrinsic Matrix K:

$$K = \begin{bmatrix} f_x & 0 & 0 \\ s & f_y & 0 \\ c_x & c_y & 1 \end{bmatrix} \tag{3.1}$$

$$K = \begin{bmatrix} 3.8657 & 0 & 0 \\ 0 & 3.7983 & 0 \\ 2.1220 & 1.5239 & 0.0010 \end{bmatrix} \tag{3.2}$$

- Rotation Matrices

$$R_1 = \begin{bmatrix} -0.0084 & -0.9999 & -0.0094 \\ 0.0094 & -0.0087 & 0.0331 \\ -0.0332 & -0.0091 & 0.9994 \end{bmatrix} \tag{3.3}$$

$$R_2 = \begin{bmatrix} 0.0040 & -0.9109 & -0.4127 \\ 0.9990 & 0.0221 & -0.0392 \\ 0.0448 & -0.4121 & 0.9100 \end{bmatrix} \tag{3.4}$$

$$R_3 = \begin{bmatrix} -0.0196 & -0.8923 & 0.4510 \\ 0.9991 & -0.0006 & 0.0422 \\ -0.0374 & 0.4514 & 0.8915 \end{bmatrix} \tag{3.5}$$

- Translation Vectors:

$$t_1 = \begin{bmatrix} -26.7223 & 19.6235 & 249.3022 \end{bmatrix} \tag{3.6}$$

$$t_2 = \begin{bmatrix} -19.6430 & 17.6781 & 235.6917 \end{bmatrix} \tag{3.7}$$

$$t_3 = \begin{bmatrix} -25.1630 & -0.7340 & 236.5726 \end{bmatrix} \tag{3.8}$$

3.2 Picture set 1

The first example shows the ideal ulcer for reconstructive purposes as it has a slightly depth and very marked wound's edges. The acquisition was performed by modifying the angle between the first and second shot. The two images are well focused and captured at a close distance from the patient.

3.2.1 Read the image

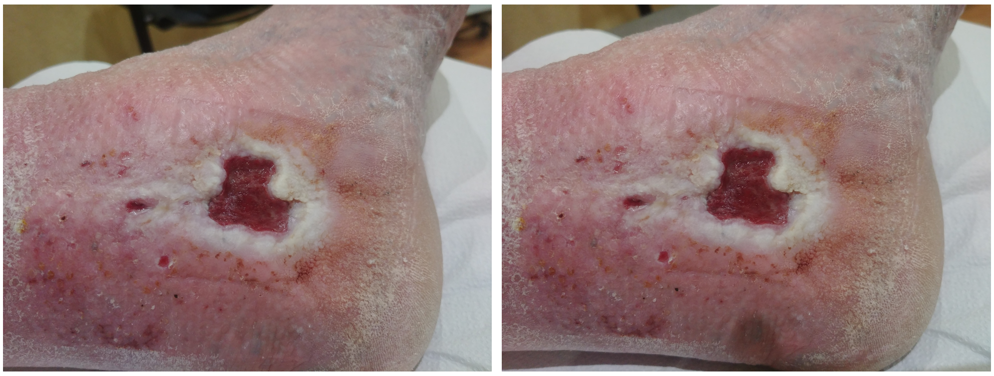


Figure 3.3: Images of Picture set 1

3.2.2 Lens distortion removal

The equations listed in (2.5) and (2.6) are used to calculate the new pixel position in the undistorted images.

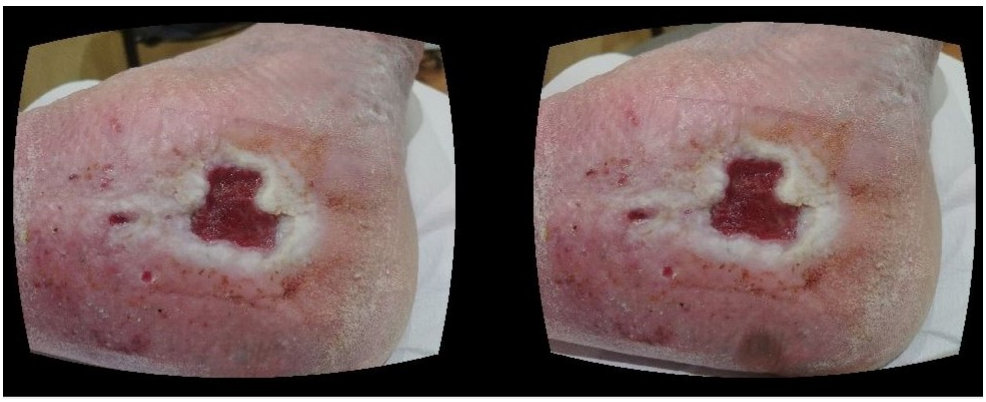


Figure 3.4: Lens distortion removal of Picture set 1

3.2.3 Features extraction and matching

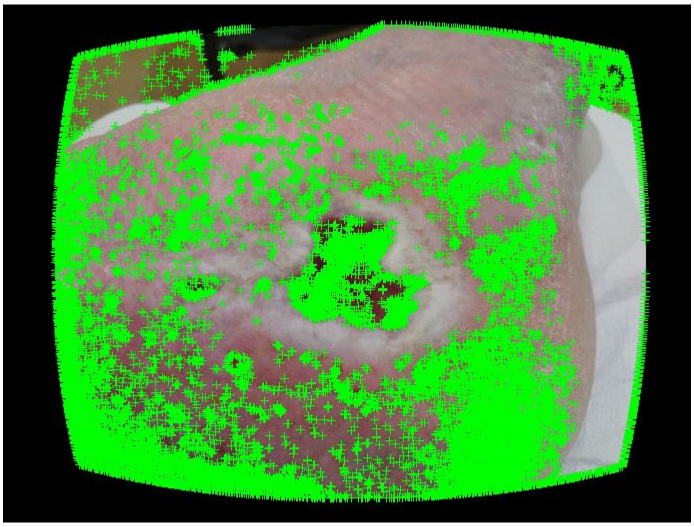


Figure 3.5: Features extraction of Picture set 1

The *Minimum Eigenvalue algorithm* has identified the 18536 keypoints. The algorithm identified as keypoints the tissue within the main wound, some small wounds around it and the lower edge of the foot.

The KLT algorithm was used to find the matches between the first and the second image. It returns 10640 matched points that amount for 57% of the initial points.

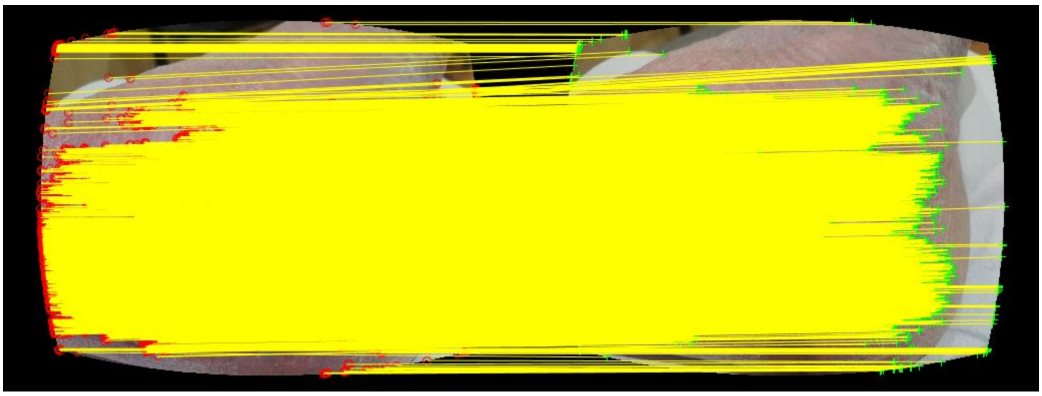


Figure 3.6: Features matching of Picture set 1

3.2.4 Essential Matrix Estimation

The matched points must respect the epipolar geometry. They are used to estimate the fundamental matrix F, which in turn is used to derive the Essential matrix E.

$$E = \begin{bmatrix} -0.0233 & 0.2256 & 0.9316 \\ -0.1464 & 0.0242 & 0.2798 \\ -0.9454 & -0.2945 & 0.0050 \end{bmatrix} \tag{3.9}$$

Points that do not respect the epipolar geometry are rejected. As consequence, in this phase remained only 3601 points called "Epipolar inliers". Considering the initial 10640 points, approximately 67% of the matched points were rejected.

A comparison of the Figures 3.6 and 3.7 shows that most of the rejected points

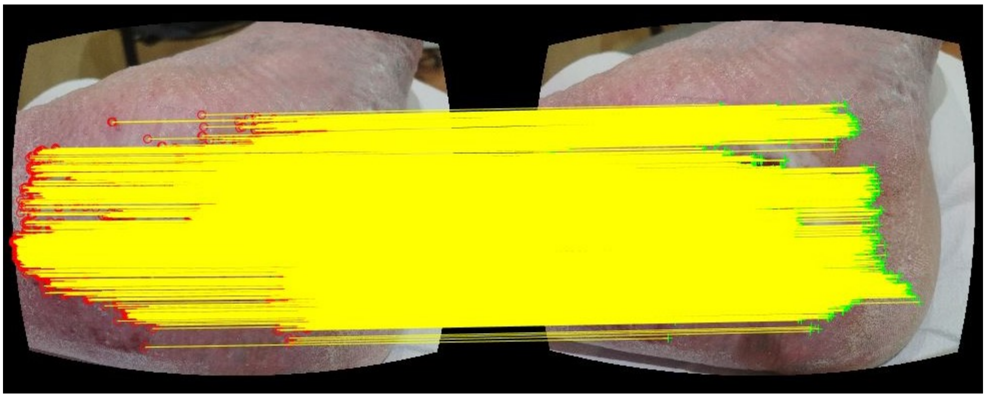


Figure 3.7: Epipolar inliers of picture set 1

were not part of the region affected by the ulcer.

3.2.5 Camera Pose

The Essential matrix (3.9) is used with the Epipolar inliers and the camera parameters to compute che location and orientation of the camera in the second image relative to the first one. The location and orientation are expressed by the rotation matrix R_2 and the translation vector t_2.

$$R_2 = \begin{bmatrix} 0.9997 & 0.0121 & 0.0220 \\ -0.0101 & 0.9958 & -0.0912 \\ -0.0230 & 0.0910 & 0.9956 \end{bmatrix} \tag{3.10}$$

$$t_2 = \begin{bmatrix} 0.2904 & -0.9283 & 0.2321 \end{bmatrix} \tag{3.11}$$

3.2.6 Features densification

To densify the found features, the algorithm applied a new keypoints extraction with a less restrictive λ threshold. This process identified 212699 points used for the matching phase in which a total of 11074 matches were identified. In the images 3.8 and 3.9 are shown the first feature matching and the second one performed for the densification, respectively.

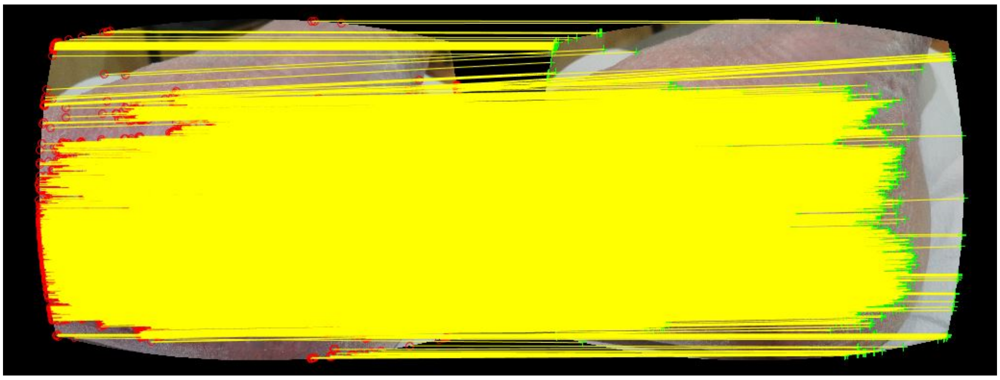


Figure 3.8: First feature matching

Figure 3.9: Second dense feature matching

3.2.7 Camera matrix estimation

Once identified the correspondences between the two images, it is possible to estimate the camera matrix P of the first image. The camera is assumed to be at the origin looking along the z-axis with null translation vector.

$$P = \begin{bmatrix} 3865.67 & 0 & 0 \\ 0 & 3798.23 & 0 \\ 2122.05 & 1523.91 & 1 \\ 0 & 0 & 0 \end{bmatrix} \tag{3.12}$$

3.2.8 Compute extrinsics of the second camera

The extrinsics of the second camera are calculated using (3.10) and (3.11). The following R and t are used to obtain the matrix of camera in the second image.

$$R_2 = \begin{bmatrix} 0.9997 & -0.0101 & -0.0230 \\ 0.0121 & 0.9958 & 0.0910 \\ 0.0220 & -0.0912 & 0.9956 \end{bmatrix} \tag{3.13}$$

$$t_2 = \begin{bmatrix} -0.2841 & 0.9485 & -0.1399 \end{bmatrix} \tag{3.14}$$

The rotation matrix and the translation vector are derived from the orientation and position of the camera calculated in the section 3.2.5. R and t are used to derive the camera matrix P' in the second image as follows:

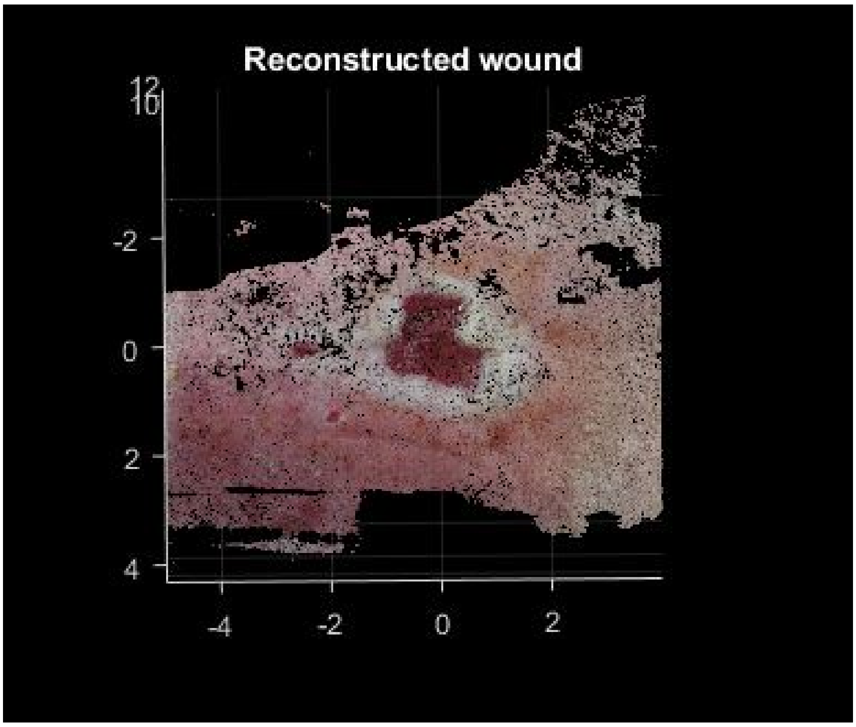
$$P' = \begin{bmatrix} R_2 \\ t_2 \end{bmatrix} K \tag{3.15}$$

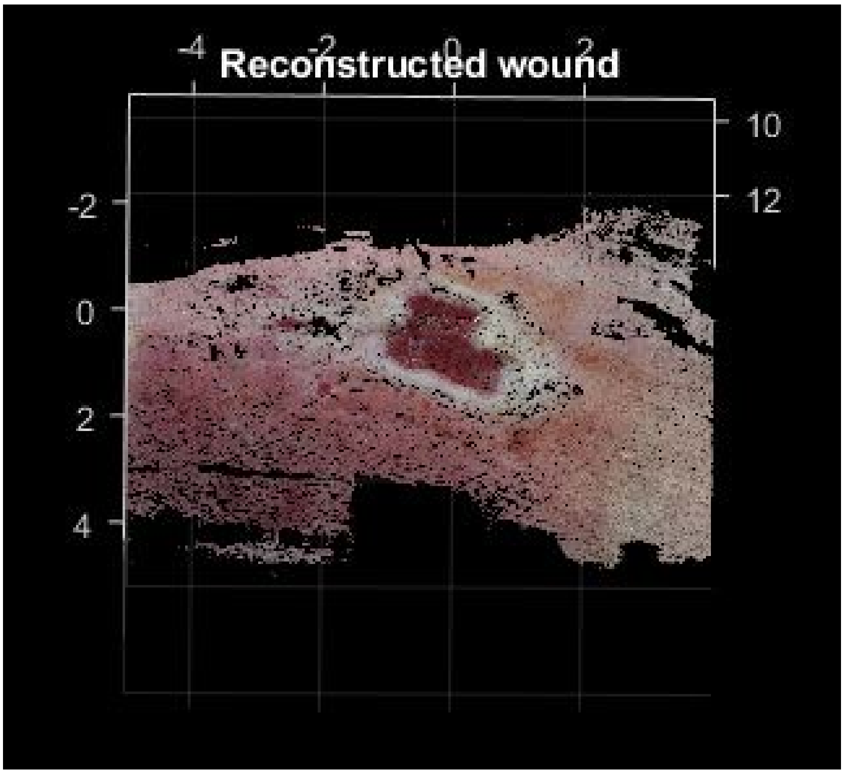
with K intrinsic matrix.

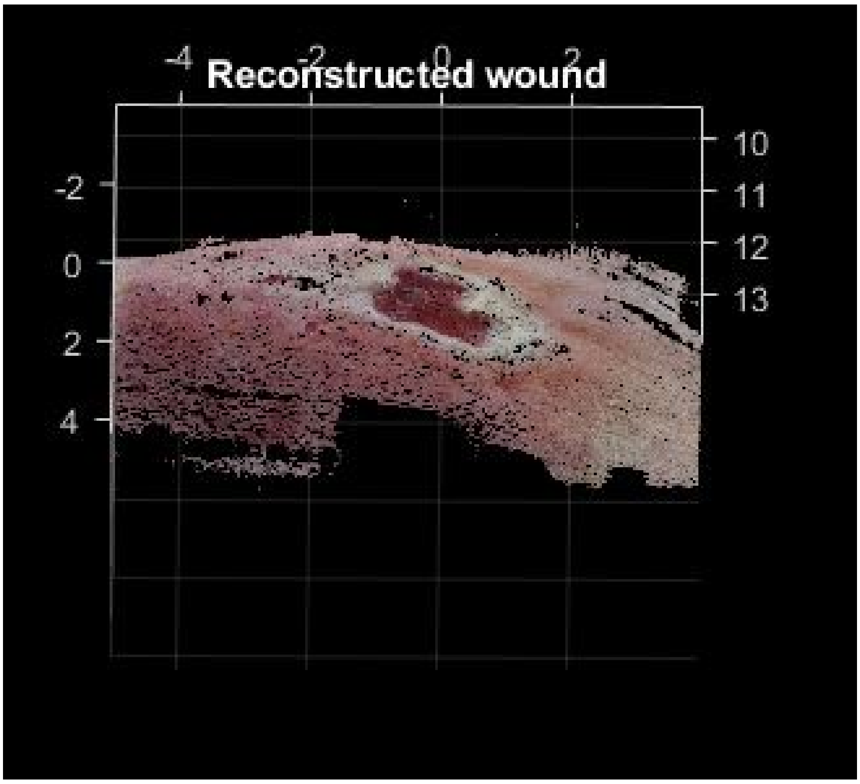
$$P' = \begin{bmatrix} 3819.3 & -76.5466 & -0.0213 \\ 252.1979 & 3924.5 & 0.0941 \\ 2189.9 & 1158.2 & 0.9953 \\ -1273.6 & 3421.7 & -0.1409 \end{bmatrix} \tag{3.16}$$

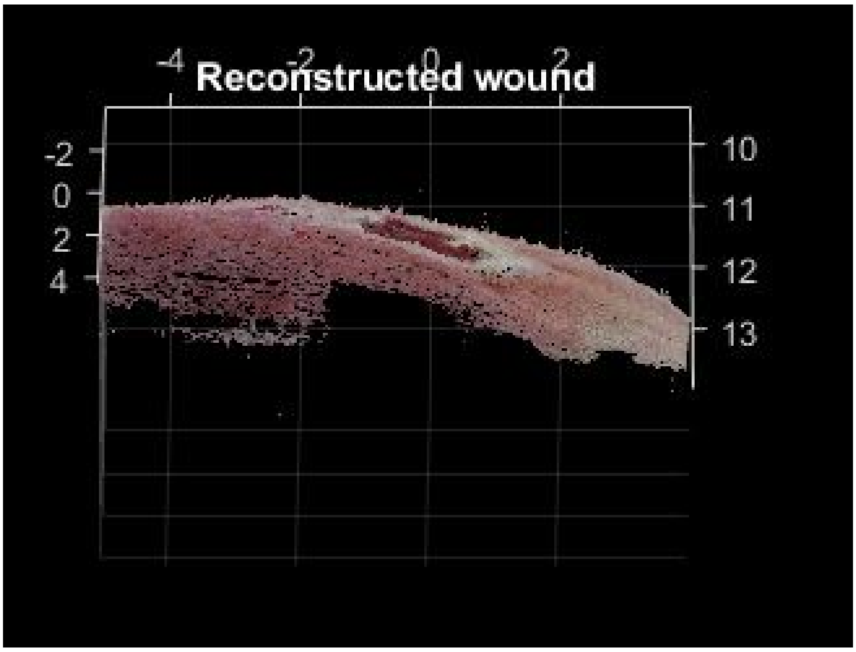
3.2.9 Triangulation

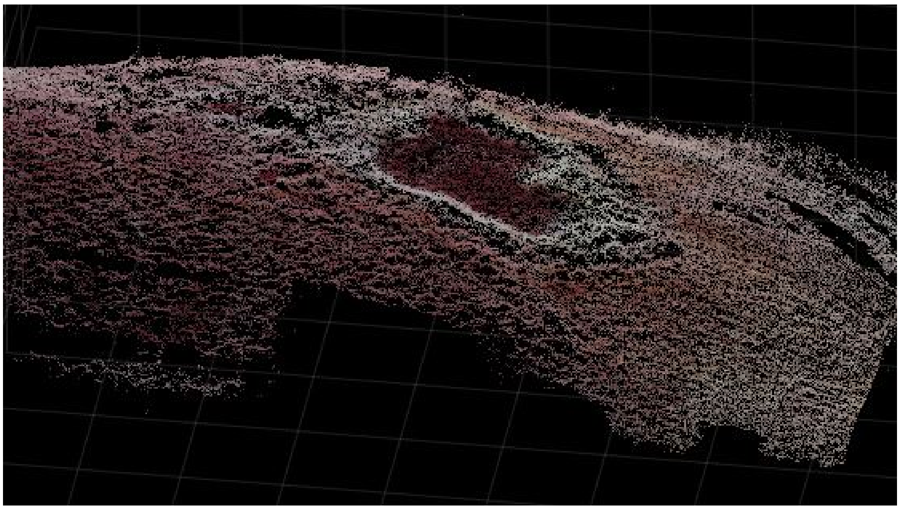
In the triangulation phase the camera matrix and the matched points of the first and second image are used to obtain the point cloud in the three-dimensional space.


Reconstructed wound
12
10
-2
0
2
4
-4
-2
0
2


Reconstructed wound
-4
-2
0
2
10
12
-2
0
2
4


Reconstructed wound
-4
-2
0
2
-2
0
2
4
10
11
12
13


Reconstructed wound
-4
-2
0
2
-2
0
2
4
10
11
12
13



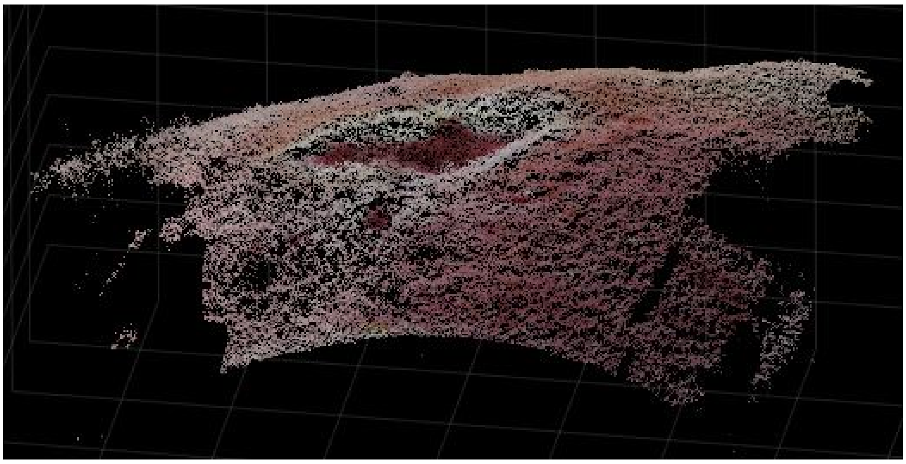


Figure 3.10: 3D Reconstruction of Picture set 1

3.3 Picture set 2

3.3.1 Read the image

The wound of the second Picture set has a less defined edge and less depth than the wound of Picture set 1. The tissue inside the wound is not homogeneous and some regions of the wound have pigmentation similar to the outer skin region. This has resulted in the error of the algorithm because it did not consider all the points of the ulcer as keypoints.

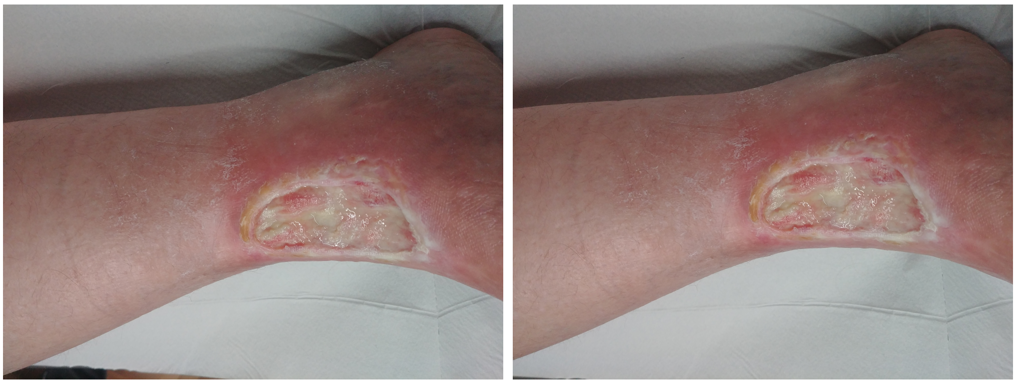


Figure 3.11: Images of Picture set 2

3.3.2 Lens distortion removal

The equations listed in (2.5) and (2.6) are used to calculate the new pixel position in the undistorted images.

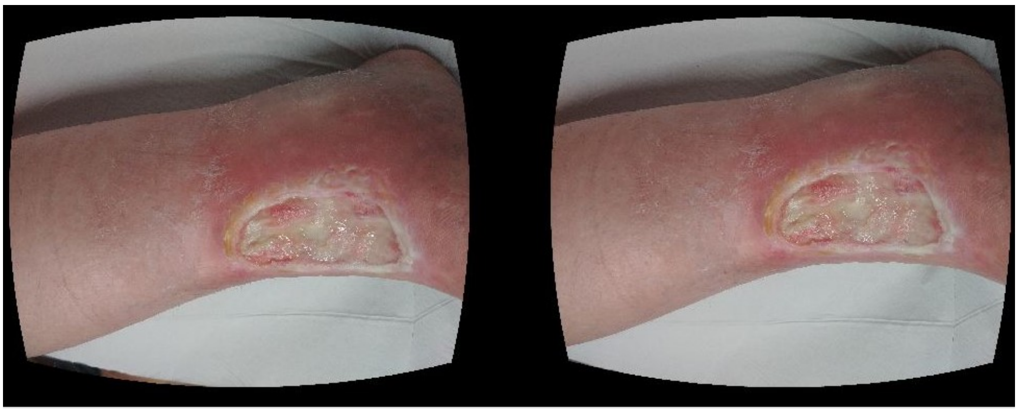


Figure 3.12: Lens distortion removal of Picture set 2

3.3.3 Features extraction and matching

The keypoint detection phase results in 15371 points identified. These are shown in the Fig.3.13.

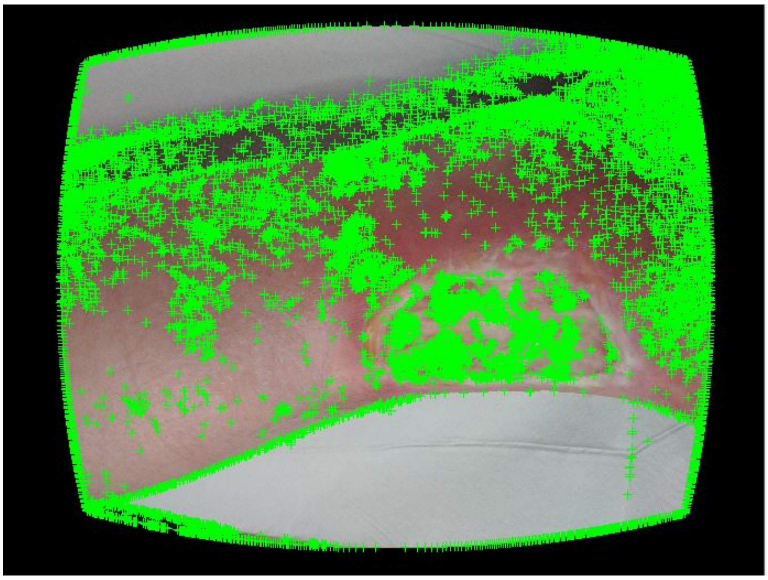


Figure 3.13: Features extraction of Picture set 2

The search of correspondence for the 15371 points of the first image results in 6996 matches. In this step the KLT algorithm has been able to match lots of points near wound area but also lots of points near the image's edge.

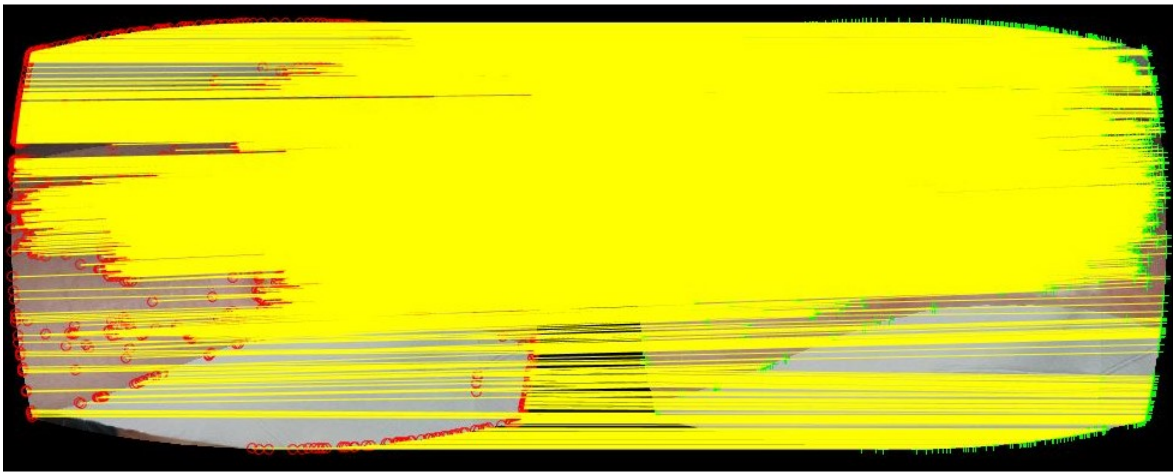


Figure 3.14: Features matching of picture set 2

3.3.4 Essential Matrix Estimation

The matched points shown in 3.14 are used to obtain the Essential matrix and also to reject the pairs of points that did not respect the Epipolar Geometry.

$$E = \begin{bmatrix} 0.0037 & -0.0016 & 0.9997 \\ -0.1155 & -2.8472 & 0.0283 \\ -0.9933 & -0.0020 & 4.4053 \end{bmatrix} \tag{3.17}$$

Starting from 6996 intial correspondences, only about 39.6 % of them respect the

Figure 3.15: Inlier points of Picture set 2

Epipolar geometry, i.e. 2773 points. The figure 3.15 shows the Inlier points matched. Lots of points internal to the wounds are rejected, this affects the final result. A situation like this can be solved in the post-reconstruction processing stage of surface reconstruction.

3.3.5 Camera Pose

The essential matrix (3.17) is used with the Epipolar inliers and the camera parameter to compute che location and orientation of the camera in the second image relative to the first one. The location and orientation are expressed by the rotation matrix R_2 and the translation vector t_2.

$$R_2 = \begin{bmatrix} 0.99965 & -0.02603 & -0.00376 \\ 0.02629 & 0.99277 & 0.11711 \\ 6.82647 & -0.11717 & 0.99311 \end{bmatrix} \tag{3.18}$$

$$t_2 = \begin{bmatrix} -0.00206 & 1 & 0.00162 \end{bmatrix} \tag{3.19}$$

3.3.6 Features densification

To obtain a dense final point cloud, the keypoints extraction algorithm is applied with a less restrictive threshold λ. The process has identified 232694 points unlike the 212699 previously identified. The new keypoints are used for the Feature matching that results in 86875, a value more than twelve times greater than the previous matching (6996 pairs of points). In the image 3.16 is shown the first features matching, while in Fig. 3.17 is shown the second one performed after densification.

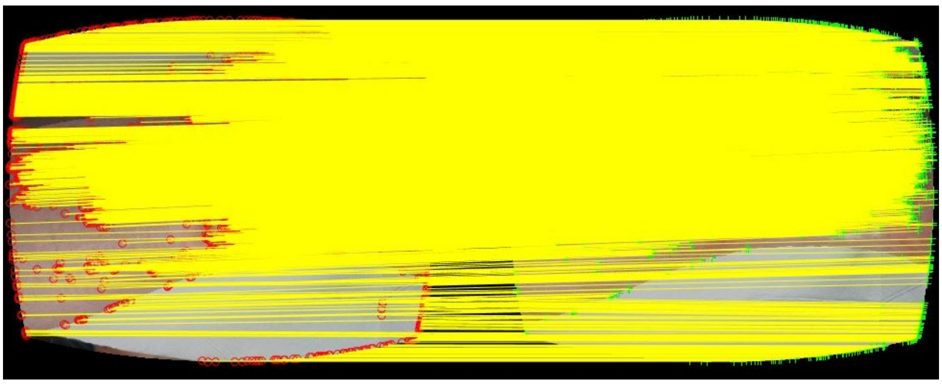


Figure 3.16: First features matching

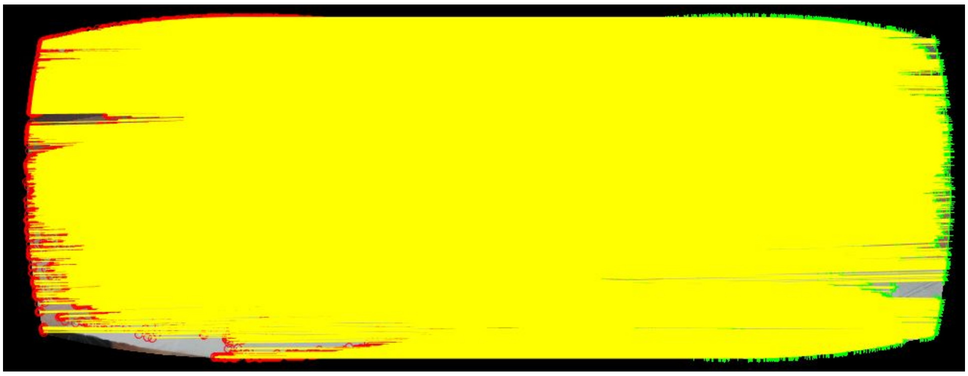


Figure 3.17: Second dense features matching

3.3.7 Camera matrix estimation

Once identified the correspondences between the two images, it is possible to estimate the camera matrix P of the first image. The camera is assumed to be at the origin looking along the z-axis with null translation vector.

$$P = \begin{bmatrix} 3865.66 & 0 & 0 \\ 0 & 3798.25 & 0 \\ 2122.04626 & 1523.90592 & 1 \\ 0 & 0 & 0 \end{bmatrix} \tag{3.20}$$

3.3.8 Compute extrinsics of the second camera

The extrinsics of the second camera are calculated using (3.10) and (3.11). The following R_2 and t_2 are used to obtain the matrix of camera in the second image.

$$R_2 = \begin{bmatrix} 0.99968 & 0.0263 & 0.00068 \\ -0.02603 & 0.99277 & -0.11716 \\ -0.00375 & 0.11711 & 0.99311 \end{bmatrix} \tag{3.21}$$

$$t_2 = \begin{bmatrix} 0.0281 & -0.9929 & 0.1156 \end{bmatrix} \tag{3.22}$$

The rotation matrix R_2 and the translation vector t_2 are used to derive the camera matrix P' in the second image as follows:

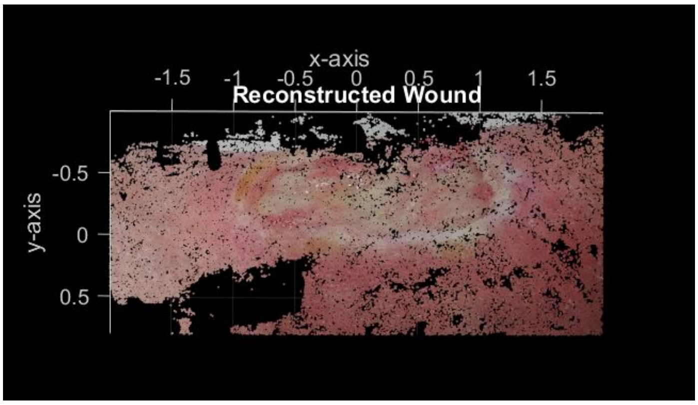
$$P' = \begin{bmatrix} R \\ t \end{bmatrix} K \tag{3.23}$$

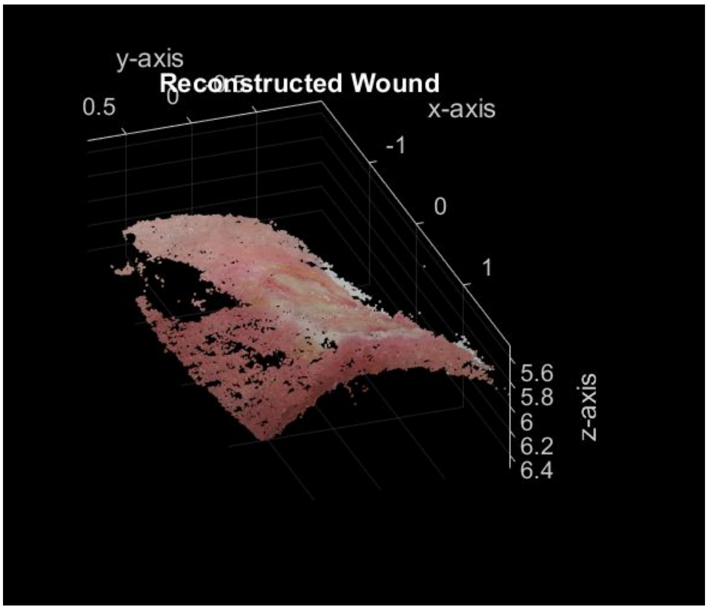
with K intrinsic matrix.

$$P' = \begin{bmatrix} 3867.7 & 102.6 & 0.0016 \\ -353.63 & 3588 & -0.1193 \\ 2088.6 & 1965.8 & 0.9929 \\ 378.6 & -3591.4 & 0.1172 \end{bmatrix} \tag{3.24}$$

3.3.9 Triangulation

The algorithm was able to reconstruct the wound almost entirely.





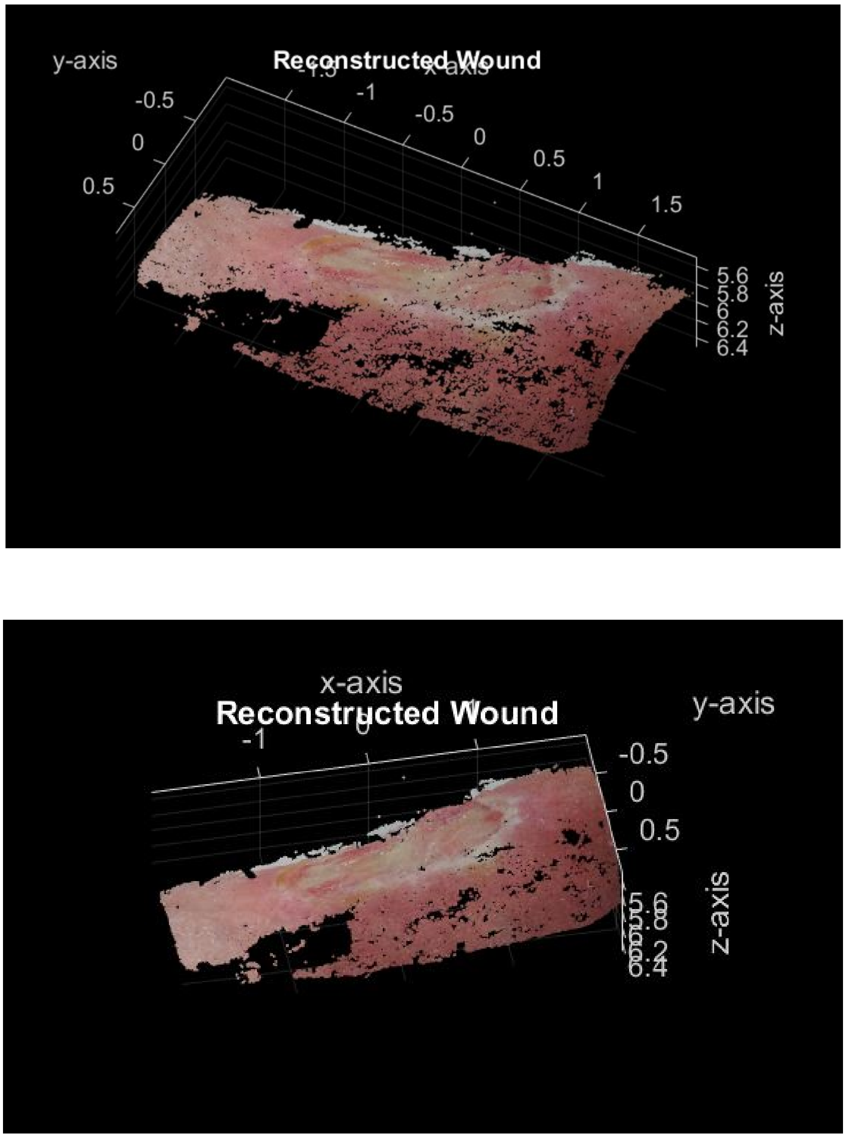


Figure 3.18: 3D reconstruction of Picture set 2

3.4 Picture set 3

The two images used for the 3D reconstruction show a cutaneous wounds with a rather homogeneous tissue and well-defined edges.

3.4.1 Read the image

Figure 3.19: Images of Picture set 3

3.4.2 Lens distortion removal

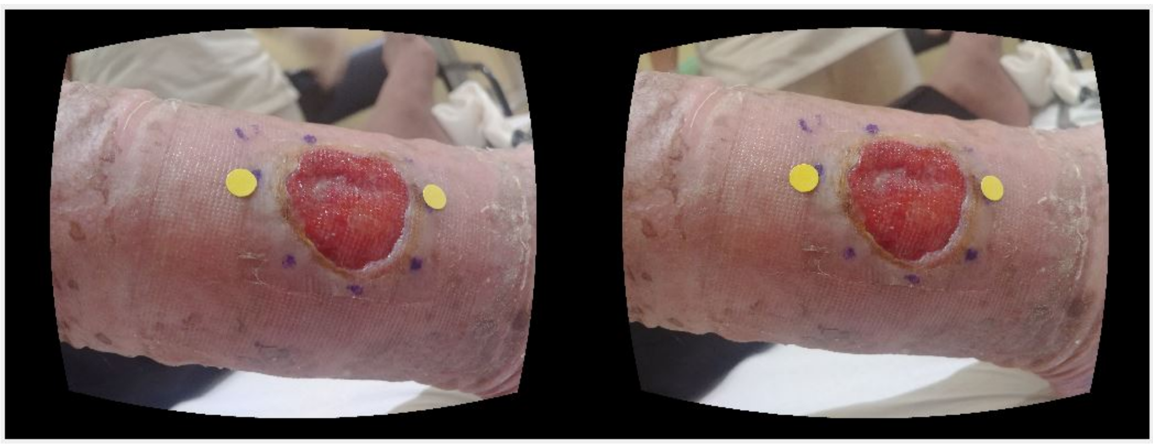


Figure 3.20: Lens distortion removal of Picture set 3

3.4.3 Features extraction and matching

The *Minimum Eigenvalue algorithm* has identified the 9985 keypoints that are shown in the Fig. 3.21. The extraction of the keypoints in the first image identified many

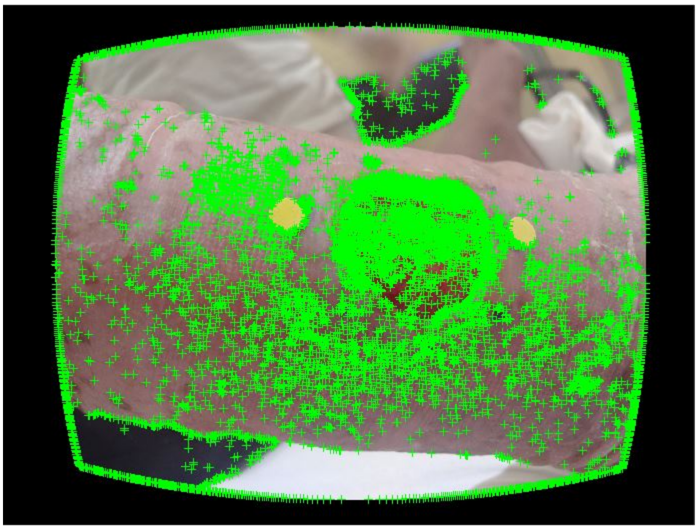


Figure 3.21: Features extraction of Picture set 3

points inside the wound as the tissue has different intensity values compared to the rest of the leg.
The KTL algorithm allow to identify 5062 matched points that are shown in Fig. 3.24

Figure 3.22: Features matching of Picture set 3

3.4.4 Essential Matrix Estimation

The matched points are used to estimate the fundamental matrix F, which in turn is used to derive the essential matrix E. Points that do not respect the Epipolar geometry are rejected and the remaining matched points are called "epipolar inliers".

$$E = \begin{bmatrix} -0.00357 & -0.85102 & 0.5222 \\ 0.8703 & -0.0128 & 0.0371 \\ -0.4917 & -0.0469 & 0.0122 \end{bmatrix} \tag{3.25}$$

The Figure 3.23 shows the 9985 Epipolar Inlier points, i.e. the points that respect the Epipolar Geometry. They are also the points used to obtain the Essential Matrix. The matched points are centred in the wound area except for a small area that will not be reconstructed.

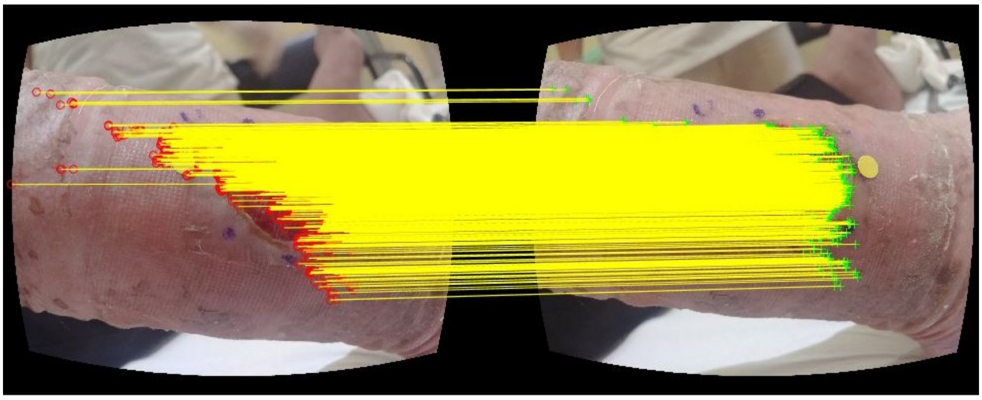


Figure 3.23: Inlier points of Picture set 3

3.4.5 Camera Pose

The essential matrix (3.9) is used with the Epipolar inliers and the camera parameters to compute the location and orientation of the camera in the second image relative to the first one. The location and orientation are expressed by the rotation matrix R_2 and the translation vector t_2.

$$R_2 = \begin{bmatrix} 0.9996 & -0.0170 & -0.0208 \\ 0.0162 & 0.9992 & -0.0364 \\ 0.0214 & 0.0360 & 0.9991 \end{bmatrix} \tag{3.26}$$

$$t_2 = \begin{bmatrix} 0.0286 & -0.5229 & -0.8519 \end{bmatrix} \tag{3.27}$$

3.4.6 Features densification

Applying again the keypoints extraction algorithm with a less restrictive threshold on the initial Picture set 3 allow to obtain an higher number of keypoints. They are the starting point for the densification of the final point cloud. This process identified 136360 points used for the matching phase in which a total of 49488 matches were identified. In the Fig. 3.24 is shown the first features matching while in Fig. 3.25 is shown the second features matching performed after the densification.

Figure 3.24: First feature matching

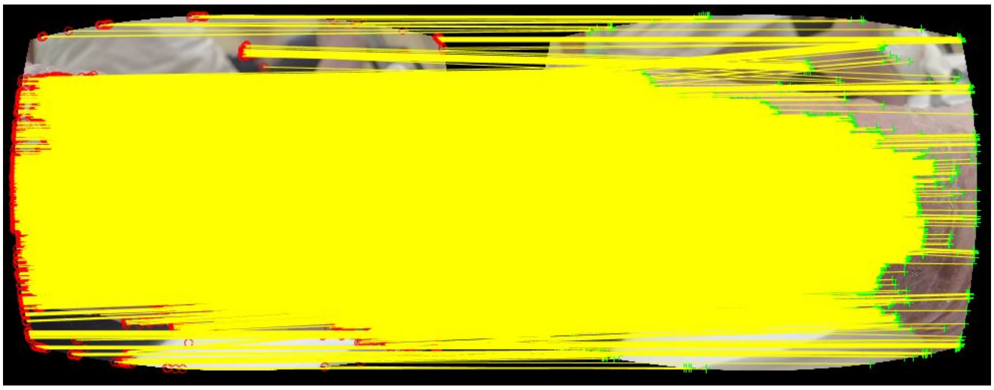


Figure 3.25: Second dense feature matching

3.4.7 Camera matrix estimation

Once identified the correspondences between the two images, it is possible to estimate the camera matrix P of the first image. The camera is assumed to be at the origin looking along the z-axis with null translation vector.

$$P = \begin{bmatrix} 3865.63 & 0 & 0 \\ 0 & 3798.25 & 0 \\ 2122.04626 & 1523.90592 & 1 \\ 0 & 0 & 0 \end{bmatrix} \tag{3.28}$$

3.4.8 Compute extrinsics of the second camera

The extrinsics of the second camera are calculated using (3.10) and (3.11). The following R_2 and t_2 are used to obtain the camera matrix of camera in the second image.

$$R_2 = \begin{bmatrix} 0.9996 & 0.0162 & 0.0214 \\ -0.0170 & 0.9992 & 0.0360 \\ -0.0208 & -0.0364 & 0.9991 \end{bmatrix} \tag{3.29}$$

$$t_2 = \begin{bmatrix} -0.0553 & 0.4910 & 0.8694 \end{bmatrix} \tag{3.30}$$

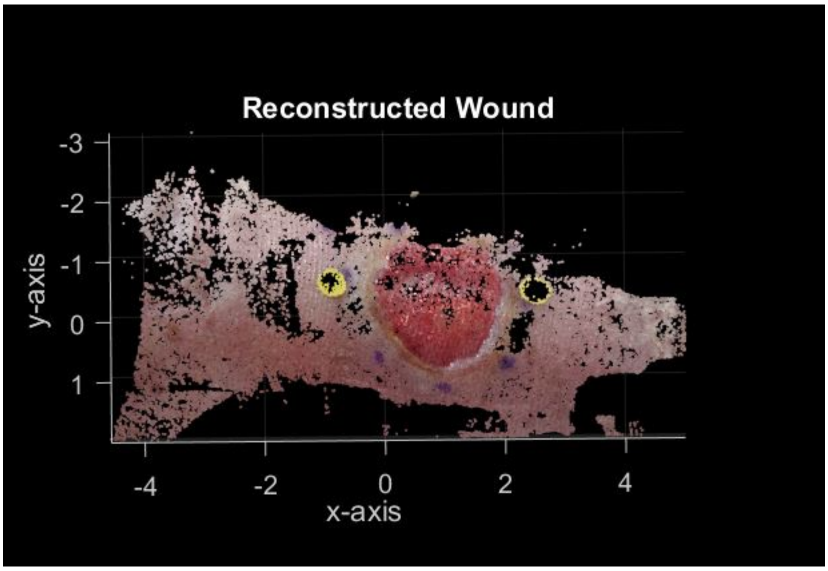
The rotation matrix and the translation vector are derived from the orientation and position of the camera calculated in the section 3.2.5. R_2 and t_2 are used to derive the camera matrix P' in the second image as follows:

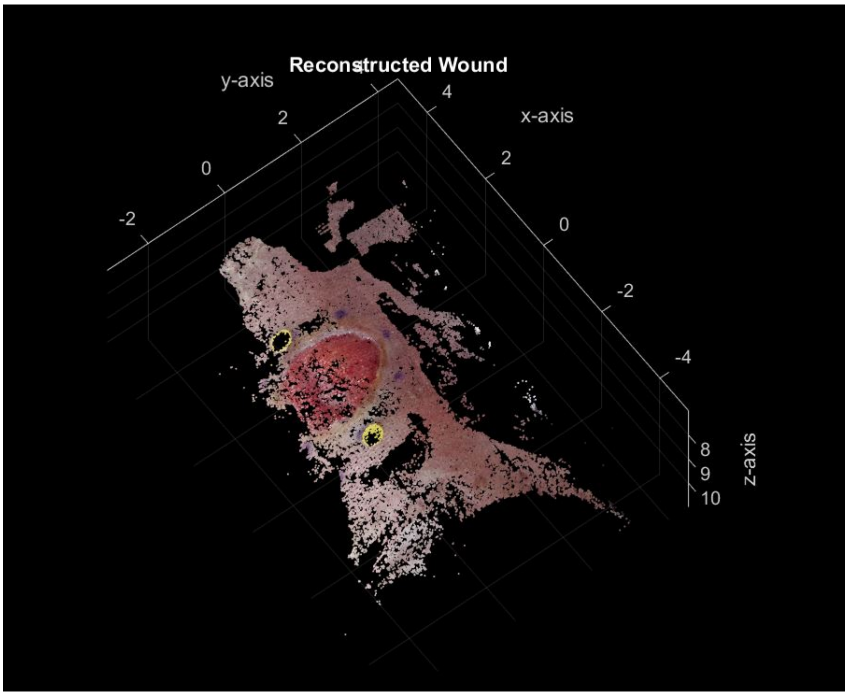
$$P' = \begin{bmatrix} R_2 \\ t_2 \end{bmatrix} K \tag{3.31}$$

with K intrinsic matrix.

$$P' = \begin{bmatrix} 3911 & 91.05 & 0.022 \\ 14.4431 & 3849.9 & 0.0358 \\ 2037.3 & 1385.2 & 0.9991 \\ 1674.1 & 3229 & 0.8628 \end{bmatrix} \tag{3.32}$$

3.4.9 Triangulation





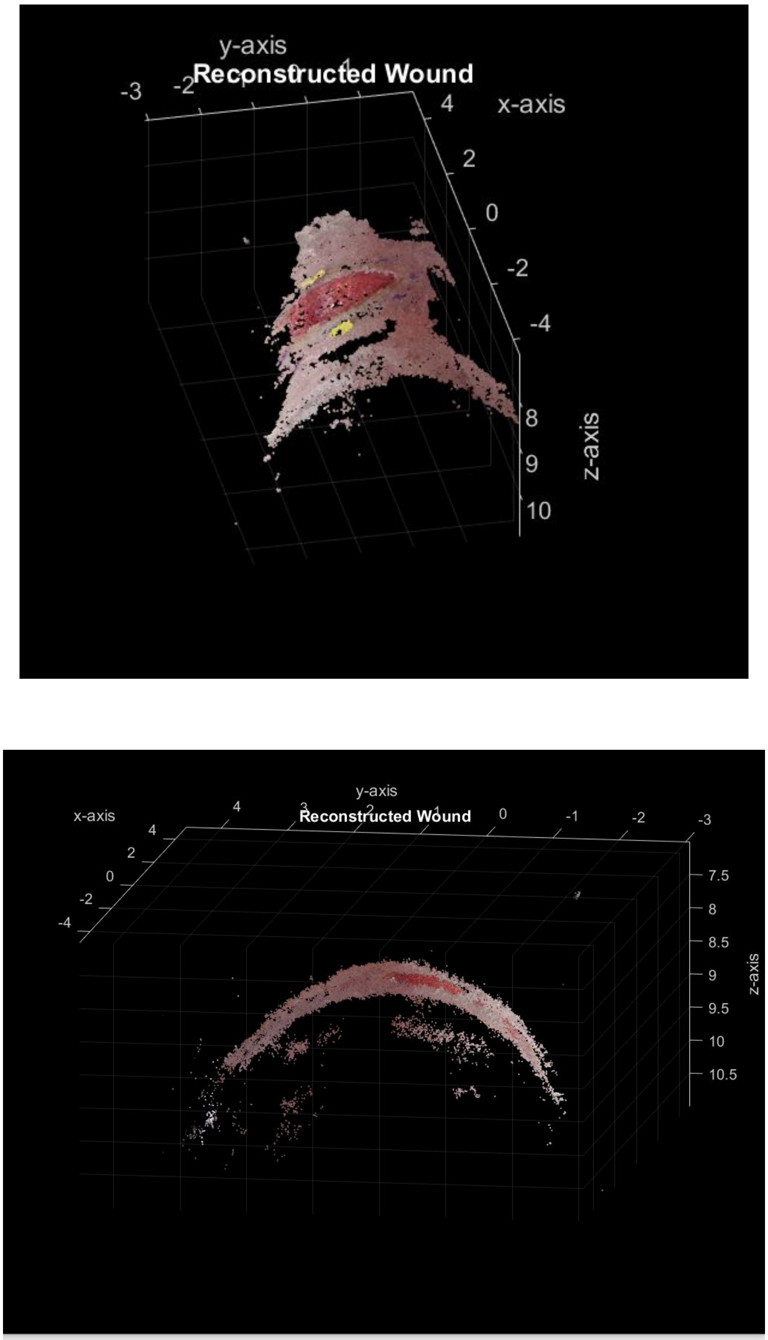


Figure 3.26: 3D reconstruction of Picture set 3

From the 3D reconstruction's images it is possible to see how the algorithm is able to represent the tendentially cylindrical geometry of the leg. The wound has been reconstructed almost entirely with the exception of the lower portion.

Chapter 4

Conclusions and future works

In this thesis work, it has been proposed an algorithm that can reconstruct a three-dimensional model of the ulcer from two single images. As shown in the previous chapter, the algorithm obtains dense point clouds. The point clouds surface mesh are the basis for a morphological analysis in order to identify the parameters related to the ulcer.
The idea is to perform a segmentation of the ulcer in order to recognize it within the three-dimensional model and work on the next steps only on the wound's model. With the coordinates of the model and the edge obtained from the segmentation it is possible to calculate parameters such as area (counting the number of superficial pixels inside the edge), depth measurement (i.e. the maximum distance between two pixels along the normal to the upper surface) and using these two to estimate the volume of the ulcer.
The algorithm proposed in this thesis is written using the software Matlab. However, once the entire routines will be completed together with the calculation of the wound parameters, it will be necessary to perform a translation of the entire algorithm in the C language. This in order to begin the test phase of the algorithm on the Wound Viewer device. A successful reconstruction with the device would represent an important step forward in the vulnological sector, giving the specialist the possibility to obtain objective wound parameters in a faster, non-invasive and painless way. This together with an online data warehouse will allow us to have a clearer idea and follow-up of the time evolution of the patient's wound.

Bibliography

[4.1] https://www.dermaclub.it/malattie/ulcere-e-decubiti

[4.2] F. Werdin, M. Tennenhaus, H.E. Schaller and H.O. Rennekampff, *Evidence-based Management Strategies for Treatment of ChronicWounds*, Eplasty, 2009, Vol.9, pp.e19 Venetia, 1612.

[4.3] K. Haeger, *Leg ulcers* in *The treatment of Venous Disorders* (ed. J.T.Hobbs), MTP, Lancaster,1977.

[4.4] F. Gottrup, *Optimizing wound treatment through health care structuring and professional education*, Wound Repair and Regeneration, March 2004, Vol.12(2), pp.129-133.

[4.5] J. Cornwall, C.J. Doré and J.D. Lewis, *Leg ulcers: epidemiology and aetiology*, British Journal of Surgery, 73(9):693–696, 1986.

[4.6] M. Farina and J. Secco, *Live demonstration: 3D wound detection & tracking system based on artificial intelligence algorithm*, 2017 IEEE Biomedical Circuits and Systems Conference (BioCAS), Turin, 2017, pp. 1-1.

[4.7] L.Chua, L. Yang *Cellular neural networks: Applications*, IEEE Transactions on circuits and systems, 35(10): 1273-1290, 1988.

[4.8] , J. Secco, M. Farina, D. Demarchi, F. Corinto and M. Gilli, *Memristor cellular automata for image pattern recognition and clinical applications*, 2016 IEEE International Symposium on Circuits and Systems (ISCAS), Montreal, QC, 2016, pp. 1378-1381.

[4.9] , J. Secco, M. Farina, D. Demarchi and F. Corinto, *Memristor cellular automata through belief propagation inspired algorithm*, 2015 International SoC Design Conference (ISOCC), Gyungju, 2015, pp. 211-212.

[4.10] Manu S. Mannoor, Ziwen Jiang, Teena James, Yong Lin Kong, Karen A. Malatesta, Winston O. Soboyejo, Naveen Verma, David H. Gracias, Michael C. McAlpine, *3D printed bionic ears*, 2013, Nano Letters, 13(6), 2634-2639.

[4.11] https://www.umcutrecht.nl/en/about-us/news/research/3d-printed-skull-implanted-in-patient

[4.12] P. Moulon, P. Monasse and R. Marlet *Adaptive Structure from Motion with a contrario model estimation*, ACCV 2012: Computer Vision – ACCV 2012 pp 257-270.

[4.13] Bouguet, J.Y. *Camera Calibration Toolbox for Matlab*, Computational Vision at the California Institute of Technology.

[4.14] J.Shi, C. Tomasi, *Good Features to Track*, IEEE Conference on Computer Vision and Pattern Recognition, June 1994, pp. 593–600.

[4.15] https://it.mathworks.com/help/vision/ug/camera-calibration.html

[4.16] T. Kanade, C. Tomasi, *Detection and tracking of Point Features*, Technical Report CMU-CS-91-132, April 1991.

[4.17] M. Berger, A.Tagliasacchi, L.Seversky, J. Levine, A. Sharf, C. Silva *State of the Art in Surface Reconstruction from Point Clouds*, Eurographics 2014-State of the Art Reports.