Multi-camera System for Stone Slab Scanning

João Henrique de Agrela Vital
joao.vital@tecnico.ulisboa.pt

Instituto Superior Técnico, Universidade de Lisboa, Lisboa, Portugal

May 2018

Abstract

At the dawn of Industry 4.0, some sectors found its methods completely revolutionized. However, the stone industry still relies on old, sub-optimal processes. For being traditionally a family business and dealing with non-standardised raw-materials, this industry is very resistant to change. Companies like Frontwave, S.A. are dedicated to rising the stone industry to higher standards and believe that the first step is to create an accurate description of the geometry and colour of the final products, in the form of an image. The data will allow to carefully plan the next operations, avoiding waste. These plans can then be sent over to a CNC machine or any other processing machine, for a clean and planned execution. Additionally, the image may be used for product classification, stock management, non-store retailing and post processing planning. This thesis proposes a new solution for the acquisition of a picture describing a stone slab. The development was driven by achieving the highest image resolution with minimal costs. The resulting system consists of an array of cameras and respective controllers. The controller modules serve as a primary processing stage, sending the outputs to a PC, which reconstructs the final image. State-of-the-art methods like coarse-to-fine matching, radial basis function warping and multi-resolution splining were adapted and implemented to achieve the best results with the least computational expense. Comparing with the current scanning machine developed by the company Frontwave, S.A., the solution proposed achieves ten times more resolution and saves 30% in imaging equipment costs. Additionally, the camera to slab distance was reduced by 80%, allowing for a much slimmer scanner.
Keywords: Industry4.0, Computer Vision, Ornamental Stone, Scanner

1. Introduction

During the last century, industrial activities have experienced significant development through standardisation, automation and production management. Although developments are available to any kind of industry, the nature of the business makes it easier or harder to implement the strategies. Up to the current date, the technological development has been split up in four revolutions, schematised in Figure 1.

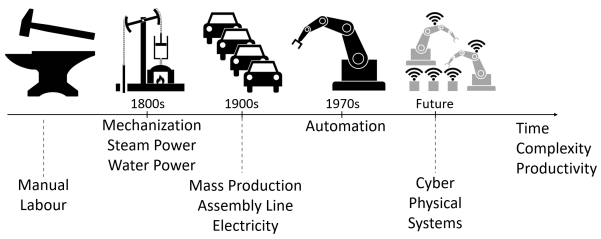


Figure 1: Industrial evolution time-line.

The fourth revolution is an on-going process and is closely related to the implementation of cyber physical systems to describe and track the processes and products inside a business.

1.1. Ornamental Stone Resources and Production in Portugal

Portugal is one of the world's leading producers of ornamental stones being placed 9^{th} in the worldwide production rank. It has internationally renowned products such as the white and pink marbles, and produces large quantities of light cream limestones, grey, yellow and pink granites, and dark grey slate. Figure 2 shows a SWOT analysis highlighting the value of the market and the importance of investing in renewed technologies for this sector.

1.2. The Company

Frontwave SA. [9] is a company dedicated to developing solutions which bring the stone industry to higher standards. Over the years of activity, the owners of stone processing factories showed the need to describe their final products digitally, in the form of an image, with the primary objective of accounting for stock and serving as marketing material. Additionally, it could be used for quality

Figure 2: SWOT analysis on the stone processing industrial sector.

control, material classification and as a reference for future processing of the slab. To meet the client's necessities, the company developed a machine to be placed at the end of the production line, with the purpose of acquiring an image of the final product. The images acquired may be used to showcase the factory's products on-line, making the information available worldwide and portraying an appealing and clean view of the product, leading to an easier communication with a potential client anywhere in the world.

1.3. Contributions

This project proposes a new machine which achieves images with higher resolution while sporting a structure designed for modularity and flexibility, allowing to fit conveyor belts of different sizes. Moreover, the structure volume was reduced and the price of the imaging equipment was reduced by 30%. This was done while aiming for a fast processing, required for acceptance and usability in the market.

2. Background

This section presents the theory supporting the development of this project.

2.1. Stone Scanning Machines

Currently there are four solutions on the market, the Bstone Scaner [2] by Bstone , Taglio Scanner [6] by Taglio, MapaScan [7] by MapaStone and Iris StoneScan by D2 Technologies [8]. The Bstone is a portable device, capable of scanning slabs up to 500x600 mm, manually, outside of the production line. The Taglio Scanner and MapaScan are very similar solutions, to be implemented on the production line and relying on a single high resolution camera to perform the acquisition. The StoneScan differs from the latter two by the usage of two high-end cameras instead of just one.

2.2. Image Pyramids

Image pyramid is the term used to describe an image which was subsequently filtered using a gaussian filter and sub-sampled. An image pyramid is representative of its scale space, developed from the necessity to detect features at different scales. The most common scale-space is the Gaussian scale-space, which is generated by consecutively filtering an image with increasingly strong filters in order to progressively average out the highest frequencies in the images. Figure 3 shows a Laplacian Pyramid, obtained using the difference of Gaussians method proposed by Burt and Adelson in [3].

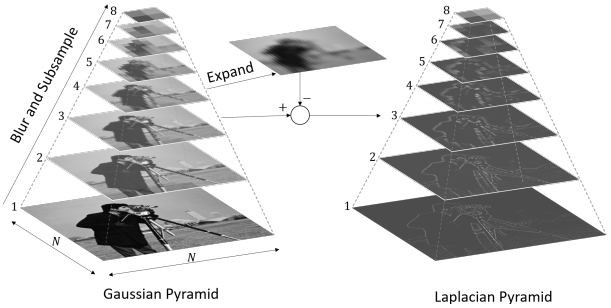


Figure 3: Example of Gaussian and Laplacian pyramids built using the Burt and Adelson's method.

2.3. Multi-resolution Blending

Multi-resolution blending is a method proposed by Burt and Adelson in [4] which adjusts the feathering window for image blending according to the frequency content of the images to blend. It is done by decomposing the images in frequency bands using the method explained in section 2.2. Figure 4 shows a qualitative representation of the feathering windows to use over the frequency bands. The steps to achieve multiresolution splining consist of:

1. Build laplacian pyramids for image A and B denoted by LA and LB.
2. Build a gaussian pyramid for the weighing function, GM. The weighing function can be converted to an image by attributing the function value to a pixel.
3. Combine the pyramid levels by doing $LS^{level} = LA^{level} * GM^{level} + LB^{level} * (1 - GM^{level})$.
4. Obtain the splined image by expanding and summing the levels of the pyramid.

2.4. Colour Modelling

When creating panographies, it is necessary to form seamless transitions between images. This is chiefly done by modelling the colour of the target overlapping region and transforming the colour of the source image using this model, in order to match the target. This is usually done using a linear model M, seen in Equation 1.

$$M^{linear-affine} = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \\ a_1 & e_1 & i_1 \end{bmatrix} \quad (1)$$

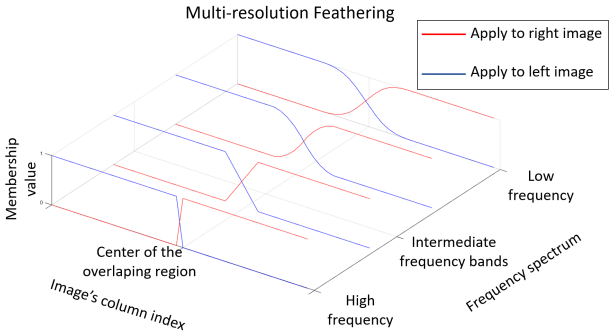


Figure 4: Graphical representation of the feathering windows used by the different methods for image blending.

Where the offset can be found by

$$\begin{bmatrix} a_1 \\ e_1 \\ i_1 \end{bmatrix}^T = \begin{bmatrix} \text{mean}(R_2) \\ \text{mean}(G_2) \\ \text{mean}(B_2) \end{bmatrix}^T - \begin{bmatrix} \text{mean}(R_2) \\ \text{mean}(G_2) \\ \text{mean}(B_2) \end{bmatrix}^T \cdot \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \quad (2)$$

Where R_i, G_i and B_i correspond to the red, green and blue pixel values of image i. Depending on the application, the model's parameters can be forced, for example, to generate a diagonal or strictly linear model.

2.5. Radial Basis Functions

Radial basis functions are commonly used to interpolate highly non-linear data sets. These functions are defined by parameters and take the form seen in Equation 3.

$$s(x) = p(x) + \sum_{i=1}^{N} \lambda_i \Phi(||x - x_i||) \quad (3)$$

where $p(x)$ is a low degree polynomial and $\Phi(x)$ is a form function, chosen according to the problem. For fitting functions of three variables, the case of this problem, the bi-harmonic ($\Phi(r) = r$ and tri-harmonic ($\Phi(r) = r^3$) are the advised choices. The method for computing the parameters defining the function is described thoroughly in [5].

3. Implementation

This section provides a summary on the algorithm and procedures implemented and chained to obtain an image of a stone slab. Some algorithms like the one used for image matching, image fusion and colour correction were adapted from existing state-of-the-art methods.

3.1. Video Reconstruction

The frames were reconstructed using the measured conveyor belt's velocity, pixel to millimetre resolution and relative frame capture instant. The slab's displacement was calculated using Equation 4.

$$d^{pixels} = d^{mm} \cdot resolution = v \cdot \Delta t \cdot resolution \quad (4)$$

Where the resolution is the pixel to millimetre ratio and Δt is the time interval between each frame capture. The frames were interpolated linearly over a window of ± 3 pixel around the intersection, in order to create a smooth reconstruction. Figure 5 shows a scheme of the reconstruction and smoothing process.

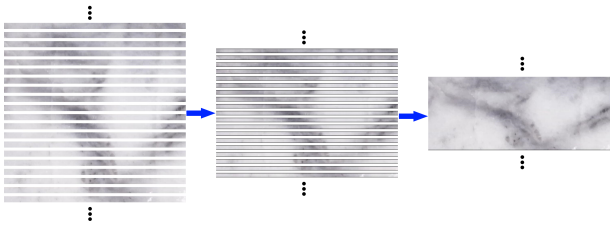


Figure 5: Example of the frame reconstruction process

The result may be distorted due to a camera rotation over the sensor plane. An example of this distortion can be seen in the results section of the video reconstruction, in Section 4.1. The next section explains how to correct this distortion.

3.2. Affine Transformation

A rotation in the sensor plane, parallel to the slab plane, introduces a geometric distortion due to the fact that the slab is scanned in movement and not in a single shot. To better understand this effect, Figure 6 shows it in a simple flow sequence.

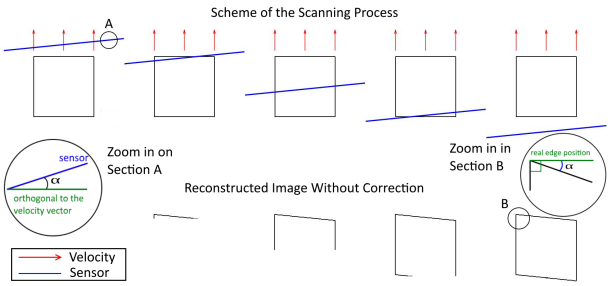


Figure 6: Scheme of the scanning process.

The distortion was shown to be proportional to the angle between the sensor and the direction orthogonal to the velocity, denoted from now on as α. For further columns of an image, the distortion's length is of $\tan(\alpha) \cdot j$, where j is the column index. Subtracting the distortion from the current pixel position leads to the real pixel location. This was done by multiplying an affine transformation matrix by the pixel coordinates of an image of size i x j, such that the real coordinates (i, j) can be found by:

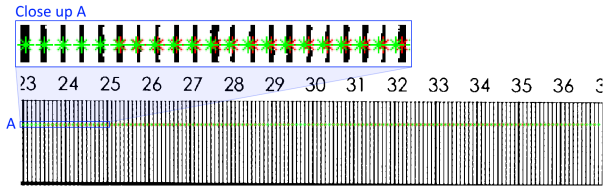
$$\begin{bmatrix} 1 & \tan(\alpha) \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} i' \\ j' \end{bmatrix} = \begin{bmatrix} i \\ j \end{bmatrix} \quad (5)$$

3.3. Lens Distortion

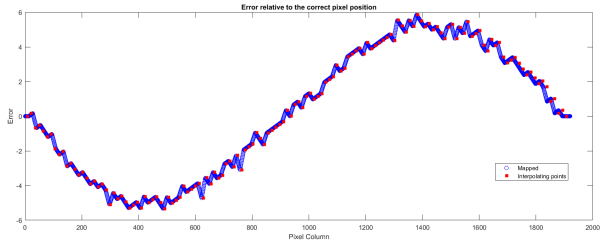
When an image is captured, there is a distortion due to lens curvature. This means that the resolution is not even along the image. In order to match all the image contributions, it is necessary to correct all geometric distortions. The lens distortion was measured with the help of a calibration ruler with equally spaced marks. An algorithm was used to extract the marks and measure the number of pixels in each interval. The images were corrected by placing the pixels in their real positions, evenly spaced. Equation 6 were used to calculate the mean spacing between markers and the error curve.

$$m_{res} = \frac{p_{max} - p_{min}}{N_{markers}}$$
$$E_i = p_i - (p_0 + m_{res} * i), \quad i = 0, ..., N \qquad (6)$$

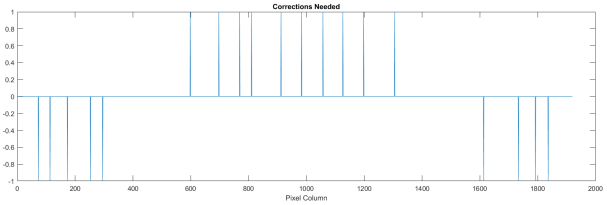
Where, m_{res} is the mean resolution, in pixels per marker, p_{max} and p_{min} correspond to the maximum and minimum horizontal coordinates and N is the total number of markers in the image. Additionally, E_i corresponds to the error of the i^{th} marker, counting from left to right. Figure 7 shows plots of the error curve and the graphical representation of the corrections to be made.



(a) Analysis of the ruler, where the green markers represent the correct position and the red markers present the measured position.



(b) Curve representing the error relative to the correct positions. Note that the curve is symmetric over the middle of the image, $x = 960$, and the axis $y = 0$.

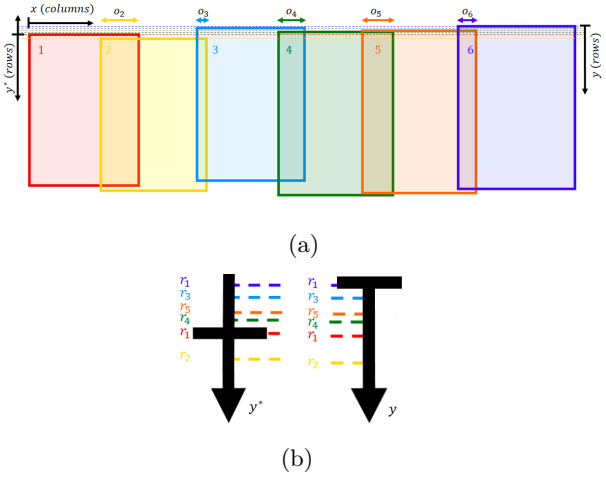


(c) Plot of the corrections to be made in each column.

Figure 7: Lens distortion analysis for a well aligned camera.

3.4. Image Matching

After acquisition and correction, the images were matched. This was achieved by making a preliminary search, inspired in the Anandan coarse to fine approach [1], by resizing the images by a factor of ten. Matching the images at a lower resolution provides an initial estimate with an uncertainty of ± 10 pixels. This search is significantly faster since this reduces the number of pixels processed. The method proposed matches images as a block, using an initial estimation, hence will be referred to as Block Matching with Initial Estimation(**BMIE**). Figure 8 show an example of the resulting match.



(a)

(b)

Figure 8: Scheme of the processing of the matching results.

3.5. Colour Balancing

To homogenize the colours between adjacent images, a diagonal model was used, as described in Section 2.4. Since the panorama consists of more than two images, the gains were propagated over the whole chain of images. Equation 7 shows the matrix of gains referring to the next image, and the gain applied to each image is computed using Equation 8.

$$G = \begin{Bmatrix} | & | & | & | & | & | \\ g_1 & g_2 & g_3 & g_4 & g_5 & g_6 \\ | & | & | & | & | & | \end{Bmatrix}, \text{ where } g_i = \begin{Bmatrix} g_{ir} \\ g_{ig} \\ g_{ib} \end{Bmatrix} \qquad (7)$$

Being g_i the vector of gains to apply to an image's colour channels, red, green and blue, so that the area overlapping with the previous image has the same colour distribution. The gain g_1 is $\{1 \quad 1 \quad 1\}$ as this is the reference image. The gain applied to each image contribution is:

$$\Gamma_i = \prod_{1}^{i} g_i \qquad (8)$$

where the multiplication is made piecewise.

3.6. Image Fusion

Natural stone has very high variability, hence choosing a fixed window for blending would not be the best option. To deal with this characteristic, the images were blended using multi-resolution splinning, according to the method described in Section 2.3. Figure 9 shows sample images decomposed in bandwidths and with the membership values applied. The Pyramids were built using Burt and Adelson's [3] method for compact Pyramid generation and reconstruction which aims for a fast and computationally cheap implementation. Figure 10 shows the reconstruction of the blended levels of the compound pyramid.

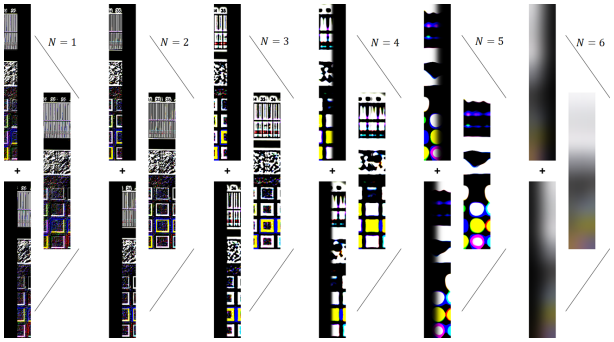


Figure 9: Representation of the merging of each pyramid level.

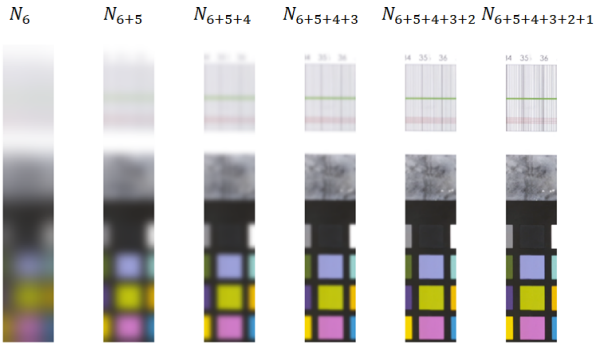


Figure 10: Reconstruction of the pyramid into the blended image.

3.7. Colour Correction

The colour values of the images acquired may not match the real RGB values of said colours, representing a colour distortion. This colour distortion is corrected by measuring the RGB values of a captured colour checker and mapping the colours to their real values. In this project, the colour checker used was the *Digital SG* from X-Rite, as seen in Figure 11.

Figure 11: Model of the colour checker used in the project.

Figure 12 portraits the error analysis over the colour space, showing a highly non-linear behaviour.

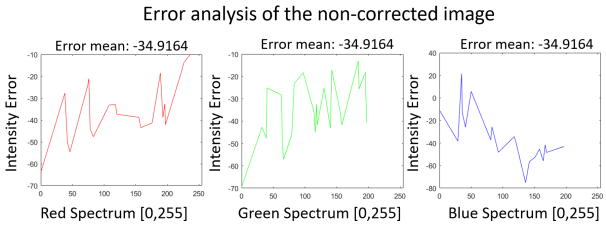


Figure 12: Mean error for each colour channel.

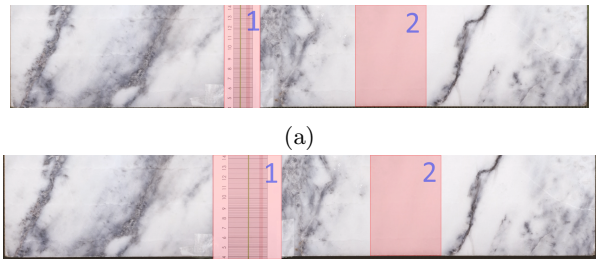
To cope with the non-linear fashion of the error, the colours were mapped using radial basis functions for each colour channel, using a bi-harmonic kernel.

4. Results

This section presents examples of images resulting from each step of the overall procedure. When relevant, the results are accompanied by comparisons against state-of-the-art solutions with the purpose of validation of the used method.

4.1. Video Reconstruction

The videos from each camera were reconstructed using the parameters from the system to calculate the slab displacement, as explained in Section 3.1. In order to prove that this is the method which yields the best results and performance, a method using image matching was implemented, for comparison. Figure 13 shows the results of both methods. Due to the aperture problem, the image matching method fails in areas where the texture is insufficient for unambiguous displacement estimation.



(a)

(b)

Figure 13: Image reconstruction using different approaches. (a) Stitching by search. (b) Stitching using the system's parameters. The images were reconstructed from left to right.

In the examples of the full reconstruction, the regions highlighted in red represent zones where the automatic search is prone to fail due to insufficient features. Area 1 is visibly distorted in Figure 13(a) and area 2, although harder to notice, is expanded by approximately 100 pixels, which is translated to roughly 10 mm. The error resulting from the search method would break the process since the reconstructed images would miss information or have information that would not match with the other cameras. Figure 14 shows a full reconstruction of an image, where it is possible to see the distortion caused by camera rotation, presented in Section 3.2.

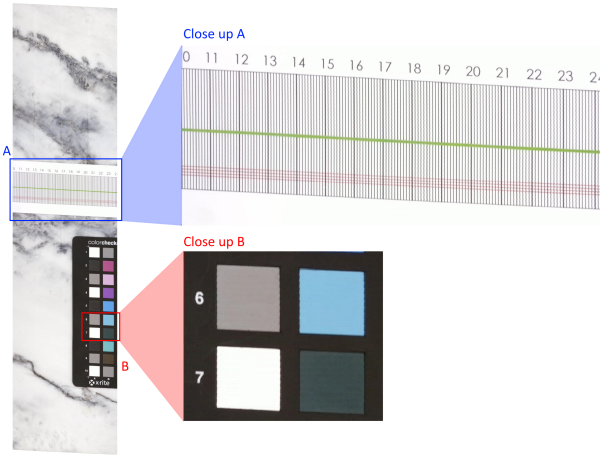


Figure 14: Reconstructed image, without any corrections. Close-up A: a closer look reveals the distortion present, where vertical lines remained vertical whereas horizontal lines appear with an angle. Close-up B: Notice that the squares are distorted into a rhombus.

4.2. Affine Transformation

The method described in Section 3.2 was implemented to correct the affine distortion. Figure 15 shows an example of a distorted image, on top, and its corrected version, at the bottom.

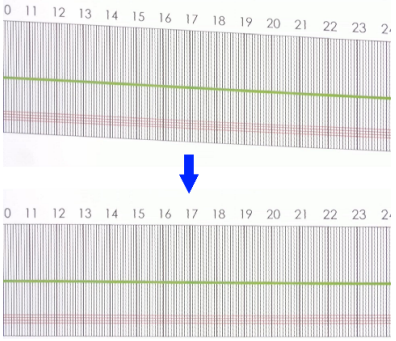


Figure 15: On top, an image of the distorted ruler, at the bottom, the corrected one.

This correction was applied to the reconstructions from each camera, using the rotation angle, measured with the help of a calibration ruler with a horizontal line. This puts all the images under the same reference and makes it possible to proceed with the image matching.

4.3. Image Matching

The corrected images were matched using the **BMIE** method explained in Section 3.4. Figure 16 shows the resulting reconstruction, before any colour corrections. Note that the colour balance in the leftmost part of the image is clearly different from the colour in the right.

Figure 16: Figure showing the result from merging the images without colour correction.

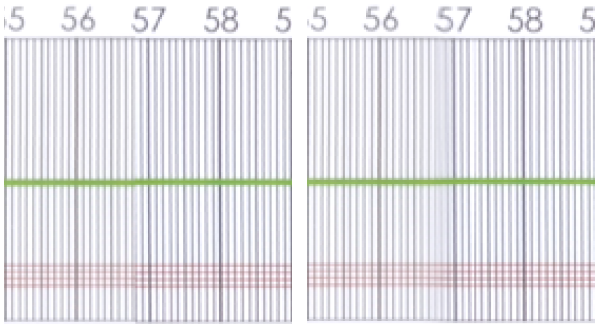
The method was tested against state-of the art approaches to motivate its usage in the application. Table 1 shows the results of the tests. The method proposed achieved errors similar to the SIFT implementation, outperforming it when it comes to the run time.

Method	E_v	σ	E_h	σ	Δt
BMIE	0.5	0.54	2.17	2.32	1.55
SIFT	0.6	1.03	2.17	2.14	5.7
SURF	1.7	1.21	76.17	59.35	1.4
MSER	0.6	0.52	108.50	90.61	3
Harris	1.5	1.37	77.67	60.60	1.4

Table 1: Matching methods comparison. Relative to the reference match, made by hand. Where E_i stands for the vertical and horizontal error, in pixels, and Δt to the run time, in seconds.

4.4. Image Fusion

The images were blended using the multi-resolution splinning method presented in Section 3.6. This method was used since it avoids both ghosts and seams in the merged image. Using a fixed window would eventually produce one of these artefacts given the unpredictability of the patterns present in natural stone. Figures 18 and 20 show examples of images with high and low frequency content, respectively, resulting from multi-resolution splinning. Figures 17 and 19 provide a comparison using a sharp transition and a fixed window cumulative distribution function blending.



(a) No feathering. (b) Fixed window feathering.

Figure 17: Images containing high frequency information blended using the tested methods.

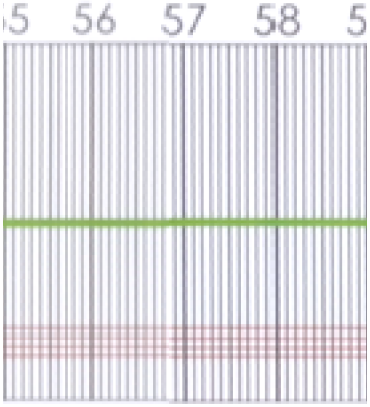
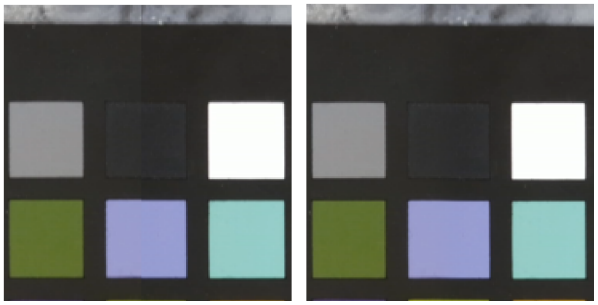


Figure 18: Multi-resolution feathering of a high frequency region.



(a) No feathering. (b) Fixed window feathering.

Figure 19: Images containing low frequency information blended using the tested methods.

Figure 20: Multi-resolution feathering of a low frequency region.

4.5. Colour Corrections

The colour corrections consist on homogenizing the colour distribution, following the method presented in Section 3.5, which is done pre-image fusion, and

the colour correction explained in Section 3.7. Figure 21 shows the image resulting from the reconstruction after going through the colour correction stage.

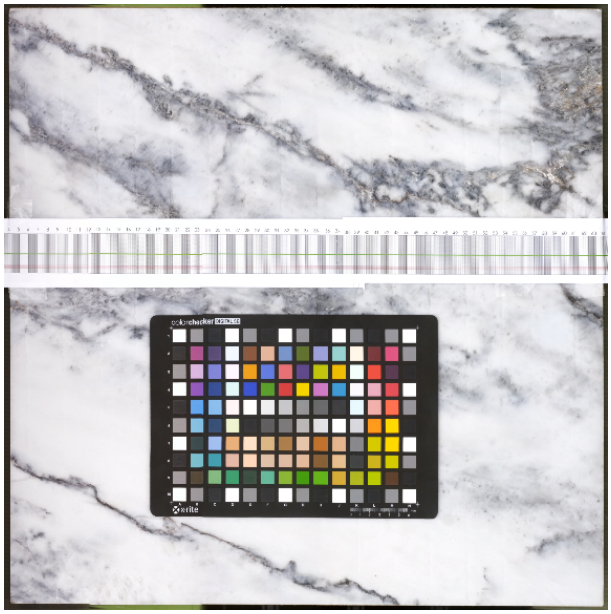


Figure 21: Resulting panorama after colour corrections.

5. Conclusions

This document proposes an innovative system for scanning stone slabs. The system consists of an array of cameras and respective controllers, and a PC. This was achieved by creating a network to enable device communication and capture synchronization. The system allows to scan slabs up to 5 meters in length and since it was built for modularity, can be adapted to fit any conveyor belt. It was inspired by the existing stone scanner Iris StoneScan, by D2 Technology in partnership with Frontwave, and its development and design were driven towards an increase in resolution and decrease in price and volume. Indeed, the resolution achieved was of 14.9 $\frac{pixel}{mm}$, showing a nearly ten times increase compared to the 1.55 $\frac{pixel}{mm}$, achieved by the current version. The price is around 880 $\frac{€}{m}$, hence the cost of a 2.4 m array would be 2112 €, representing 30% savings in the imaging equipment employed in the Iris StoneScan. Regarding the size, the cameras stand at approximately 20 cm to the slab, while the Iris StoneScan has its cameras placed at about 90 cm from the slab, representing a reduction of about 80%. The processing time from the end of acquisition to the final output is around 10 seconds, 5 of which are in the acquisition modules, hence there is availability for a new slab to be scanned every five seconds. This is sufficient for a regular stone processing factory. These achievements were made at the cost of the assumption that the slab's velocity is constant. This assumption would be more easily satisfied if the cameras translated over the immobile slab, and not the other way around. This is because the reduced weight and inertia allow for an easier control over the motion of the cameras. One of the versions of the Iris StoneScan works in the vertical position, where the slab is placed in a support while the cameras move over it at a constant velocity, hence the solution proposed would have its best results when employed in this kind of system.

Acknowledgements

Firstly, I would like to thank my supervisors for the guidance and support. A sincere acknowledgement to the Frontwave, S.A. team, for welcoming and tutoring me. A warm thank you to my family and friends for being supportive and cheering me on throughout the whole project.

References

[1] P. Anandan. A computational framework and an algorithm for the measurement of visual motion. *International Journal of Computer Vision*, 2(3):283–310, Jan 1989.

[2] BStone. Bstone scanner - the world's first handy stone scanner. `http://www.bstone.com/`. Accessed: 2017-03-20.

[3] P. J. Burt and E. H. Adelson. The laplacian pyramid as a compact image code. *IEEE TRANSACTIONS ON COMMUNICATIONS*, 31:532–540, 1983.

[4] P. J. Burt and E. H. Adelson. A multiresolution spline with application to image mosaics. *ACM Trans. Graph.*, 2(4):217–236, Oct. 1983.

[5] J. C. Carr, R. K. Beatson, J. B. Cherrie, T. J. Mitchell, W. R. Fright, B. C. McCallum, and T. R. Evans. Reconstruction and representation of 3d objects with radial basis functions. In *Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH '01, pages 67–76, New York, NY, USA, 2001. ACM.

[6] T. S. House. Scanner - marble and stone scanning. `http://www.taglio.it/en/stone/scanner-2/`. Accessed: 2017-03-20.

[7] M. Scan. Mapascan, the 1st scanner in the world of stone. `http://www.mapastone.com/`. Accessed: 2017-03-20.

[8] D. Technology. Stonescan iris. `http://www.d2technology.com/visao.html`. Accessed: 2017-03-20.

[9] F. Technology. Stonescan. http://frontwave.pt/technology/project/stonescan/. Accessed: 2017-03-20.