



IMAGE-BASED METHODS FOR METRIC SURVEYS OF BUILDINGS USING MODERN OPTICAL SENSORS AND TOOLS: FROM 2D APPROACH TO 3D AND VICE VERSA

Pepe Massimiliano

University of Naples “Parthenope”,
Department of Sciences and Technologies, 80143 Naples, Italy

ABSTRACT

Nowadays, the advent of digital images and the diffusion of new tools for the transformation and the creation of 3D metric models has led to a higher quality and accuracy in the architectural survey of the buildings. Furthermore, the recent development of new sensors, such as smartphone or commercial Digital Single-Lens Reflex (DSLR) cameras or cameras incorporated in drones, the design of new algorithm and tools based on Image-Based Methods (IBM) has allowed an increasingly faithful representation of buildings. Of particular interest in the Close Range Photogrammetry (CRP), it was the introduction of the Structure from Motion (SfM) approach, which allowed to survey even buildings with complex architecture. However, the processing of construction of the 3D model building using this latter method requires High Performance Computation (HPC) and, sometimes, long processing time in order to build 3D models. Therefore, this approach is not always the better method in the construction of 3D models. In this paper, it was shown by diverse interesting case studies and wherever possible, if the building has a plane surface or it can be simplified as a plane surface, to create 3D models starting from 2D approach. In addition, a case study concerning the survey of a building with many glazed surfaces, such as skyscrapers, is discussed. This is due to the high reflectivity of glazed surfaces that do not allow (or make very complicated) an automatic alignment of multiple images using traditional software for the realization of both 2D and 3D models. Consequently, depending on the position of the operator and the position of the sensor, the image reflected in the glass changes. Therefore, a specific and suitable method should be used depending on the type of building. However, if the architecture of the building is complex, it is necessary to use a 3D approach, such as Structure from Motion. In this way, it was possible to build the 3D model and, subsequently, to export to single façade (2D) according to a traditional approach. Indeed, the use of the 2D orthophoto (or 2D CAD representation), thanks to its ease of management, is still a favourite method for many users, such as architects, engineers, restorers or maintenance technicians in many building activities.

Key words: Photogrammetry, IBM, smartphone, 3D model, drone, Projective, SfM

Cite this Article: Pepe Massimiliano, Image-Based Methods for Metric Surveys of Buildings using Modern Optical Sensors and Tools: From 2D Approach to 3D and Vice Versa. *International Journal of Civil Engineering and Technology*, 9(9), 2018, pp. 729-745.

<http://www.iaeme.com/IJCIET/issues.asp?JType=IJCIET&VType=9&IType=9>

1. INTRODUCTION

The use of images for the representation of the building represented, and still continues to play a key role in the design, management and restoration of the building architecture. In fact, the relationship between photography and architecture dates back to the invention of photography: as known, most likely the first photographic image was that of an architectural space, the courtyard of the Maison du Gras in a period between 1824 and 1826 [1]. This close link between photography and architecture is due to objectivity of this representation. Indeed, by the photo, it is possible to report (faithfully) the physical characteristics of the architectural subject taken into consideration. The history and the chronological development of the photographic method for buildings survey can be found in the diverse books of the photogrammetry, architecture and geomatics literature [2,3,4].

In the last years, the advent of digital images has increased the wide diffusion of photogrammetric techniques in the metric survey of the buildings of historical, cultural and social interest [5,6,7]. In this context, diverse approaches are become widespread, such as the transformation of the single digital image for the representation of the 2D object by projective transformations or Image-Based Methods (IBM) in order to realize 3D model [8, 9]. Especially for the construction of 2.5 and 3D models, diverse methods were implemented based on stereoscopic approach. The first experiences in this environment, taking into account a single camera (resting on a tripod) and performing stereo-photos [10]. Subsequently, some systems based on a digital camera able to move in a special cart placed on a calibrated steel bar have become spread [11]. In both cases, the camera must be calibrated or the typical geometrical parameters must be known: optical distortions, principal point, focal length, etc. Actually, the Structure from Motion (SfM) approach has become quite popular in CRP thanks to ability to determine the parameters of external orientation without any a priori knowledge of the approximate positions for cameras and 3D points [12]. SfM technique requires, in order to realize 3D models, a block of images with a high degree of overlap that capture the complete 3D structure of the scene viewed from several positions. In fact, even in this approach, calibrated camera is requested [13]. However, the development of the so-called *self-calibration* [14] allowed the application of this method with high precision also with non-professionals and calibrated cameras. In this approach, all the parameters of internal, external orientation and object coordinates of the points are unknown. This procedure considers the systematic errors due to the acquisition process of the frames and are calculated using the so-called “*Additional Parameters*” (APs). In this environment, the Brown (1971) model is most popular: it consists of 10 APs related to internal camera orientation (Δx_p , Δy_p , Δc), uncertainty about the shape of the pixel (skew factor S_x), no orthogonality of the reference system (shearfactor Λ), radial symmetric distortion of the lenses (k_1 , k_2 , k_3) and tangential distortion of the lenses (p_1 , p_2). Since all the terms of these equations can be attributed to errors due to real physical causes, the system is called “*physical model*” [15]. Therefore, solving a self-calibration with Bundle Adjustment (BA) means estimating the additional parameters in the collinearity equations and, at the same time, to determinate the parameters of external orientation of each image. Practically, the several steps to build 3D models are [16]: alignment of the images, building dense point clouds, mesh and texture. Nowadays, several free and commercial software allow to obtain 3D model in easy way. However, this

approach requires PC high computation performance: for example, Agisoft Photoscan software, that is one of the most widespread software in this environment, suggest the minimum technical features of the PC (indicated in section 3.2) in order to produce detailed three-dimensional models [17]. For this reason, the processing time using SfM approach on a large project can be long or even impossible. Therefore, depending on the type of building, the choice of a survey method must be identified beforehand. In addition, buildings with reflecting surfaces, such as high-rise buildings, could create many problems in the alignment phase and make the SfM approach inapplicable. This is due to the impossibility of the software to recognize homologous pixels between successive images. Therefore, in order to survey this special buildings, it was necessary to identify an alternative approach to build 3D model.

The paper is organized as follows. The first section called “*Geometry background and accuracy*” describes the basic concept of the 2D and 3D approach and the formulas that are necessary to implement in order to calculate the spatial errors. The section called “*Methods and sensors for photogrammetry project*” describes the workflow that it is necessary to implement in order to obtain a metric representation 2D and 3D of the building under investigation and the features of the cameras used in these experimentations for photogrammetric purpose. The section called “*Experimental Results*” shows three case studies of application of survey methods on different buildings. The *Conclusions* are summarizing at the end of the paper.

2. GEOMETRY BACKGROUND AND ACCURACY

2.1. Concept Base and Perspective Projection

In general, the lens adopted in the photogrammetry are thins lenses. In this case, by simplifying the assumption of Gauss, the Huygens’s equation becomes [18]:

$$\frac{1}{c} = \frac{1}{d_0} + \frac{1}{d_i} \quad (1)$$

where

- c focal length;
- d_0 distance between the object and the centre of the lens;
- d_i distance between the image of the object and the centre of the lens.

In the case d_0 is very large compared to c and d_i (or realized with special lenses), the plane where the image is formed is practically coincident with the focal plane of the objective. Hence, the relationship between image and object space can be realized by *central projection*. The most important feature in the perspective centre is that the image point and the corresponding object point all lie on a straight line in space. The mathematical formulation can be described by the *collinearity equations*:

$$\begin{aligned} X &= X_0 + (Z - Z_0) \frac{r_{11}(x' - x'_0) + r_{12}(y' - y'_0) + r_{13} c}{r_{31}(x' - x'_0) + r_{32}(y' - y'_0) + r_{33} c} \\ Y &= Y_0 + (Z - Z_0) \frac{r_{21}(x' - x'_0) + r_{22}(y' - y'_0) + r_{23} c}{r_{31}(x' - x'_0) + r_{32}(y' - y'_0) + r_{33} c} \end{aligned} \quad (2)$$

where

- X, Y, Z 3D object coordinates;
- x', y' 2D coordinates of the initial image;
- X_0, Y_0, Z_0 position of the perspective center;
- c focal length;

x'_0, y'_0 coordinates of the principal point;

$r_{11}, r_{12}, \dots, r_{33}$ elements of the rotation matrix.

The focal length and the coordinates of the principal point represent the interior orientation parameters while position of the perspective center and the elements of the rotation matrix denote the *exterior orientation* parameters. In the case the object is similar to a plan, the transformation between the image space and the object space can be realized by projective transformation (or also called homographic). In this approach, the transformation between the image space and the object space (plane) the object coordinate (X, Y) can be calculated from the previous relation (2) if Z=0 for the XY plane and assuming that the internal orientation parameters of the camera are known. Therefore, re-arranging the terms of the equation (2), the central projection can be representing by the equations [19]:

$$\begin{aligned} X &= \frac{a_1 x' + a_2 y' + a_3}{a_7 x + a_8 y + 1} \\ Y &= \frac{a_4 x' + a_5 y' + a_6}{a_7 x' + a_8 y' + 1} \end{aligned} \tag{3}$$

From the previous report it emerges that through a single image it is possible to reconstruct an object and it is necessary to determine 8 independent parameters in order to obtain the perspective of a plane object. This equations are non-linear in the a_i coefficients. To solve a system in eight unknowns with two equations, it is necessary to use at least four control points. Of course, if in the project are known more of four points, as is desirable, it is possible to perform an approach to least squares method (LSM). The relation (2) can be obtain in the matrix form:

$$\begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} x' & y' & 1 & 0 & 0 & 0 & -x'X & -YX \\ 0 & 0 & 0 & x' & y' & 1 & -x'Y & -Yy' \end{bmatrix} [a_1 \ a_2 \ a_3 \ a_4 \ a_5 \ a_6 \ a_7 \ a_8]^t \tag{4}$$

In order to calculate the vector of residuals (v), it is possible to write the previous equation in the following way [20]:

$$A \cdot x - L = v \tag{5}$$

where:

$$A = \begin{bmatrix} x' & y' & 1 & 0 & 0 & 0 & -x'X & -YX \\ 0 & 0 & 0 & x' & y' & 1 & -x'Y & -Yy' \end{bmatrix};$$

$$L = \begin{bmatrix} X \\ Y \end{bmatrix};$$

$$x = [a_1 \ a_2 \ a_3 \ a_4 \ a_5 \ a_6 \ a_7 \ a_8]^t.$$

In the equation (5), it is present the vector of unknowns x that contains the corrections of the transformation parameters, the *design matrix* A , the observation corrections are stored in vector v and L is the observation vector. Generally, a weight matrix is added in order to assign a weight to the observations and unknown parameters during the estimation procedure. Using identical weights for all observations, the weight matrix becomes the identity matrix. Therefore, it is possible to derive the following equation by several and suitable algebraic steps and to obtain the best estimate [21]:

$$\hat{x} = (A^t A)^{-1} \cdot A^t \cdot L \tag{6}$$

Lastly, it is possible to obtain the residual *a posteriori* (v_{post}) and to check the quality of the transformation, as show in the following relation [22]:

$$v_{post} = A \cdot \hat{x} - L \tag{7}$$

The analysis of residuals allows to verify, point by point, that these values are lower than the one designed. In addition, the average value and the standard deviation of residues also play an important role in the control of the quality of the transformation.

Because lenses in photographic lenses differ from the ideal case by producing aberrations, it follows that the projected rays are no longer straight lines. In fact, one of the most important aberrations in the photogrammetric process is the *distortion*. The distortion, can be divided in two components: radial and tangential. In general, the tangential distortion is usually insignificant and is not included into computing of distortion correction. This phenomenon is very evident with the non-metric digital cameras or sensors with fish-eye lens [23]. Therefore, the images coordinates must be corrected for lens distortion before applying the projective transformation. In other words, a correction of this distortion should be the first step in image processing.

2.2. Stereoscopic Approach and Estimation of the Accuracy

The stereoscopic approach allows to determinate the third dimension of an object. Indeed, performing two or more shooting (according to a direction normal to the object) for each side of the object of the survey, it is possible to build a 3D model. In this case, the distance between object (d) and camera should be 3 or 4 times the base (B), i.e. 3 or 4 times the distance between the two camera stations [24]. In this case, known as “*normal case*”, the spatial accuracy can be calculated by following relations [25]:

$$\begin{aligned}\sigma_x &= \sqrt{\left(\frac{\xi_1}{c} m_b \frac{d}{B} \sigma_{P\xi}\right)^2 + (m_b \sigma_\xi)^2} \\ \sigma_y &= m_b \frac{y}{B} \sigma_{P\xi} \\ \sigma_z &= \sqrt{\left(\frac{\eta_1}{c} m_b \frac{d}{B} \sigma_{P\xi}\right)^2 + (m_b \cdot \sigma_\eta)^2}\end{aligned}\quad (8)$$

where σ_x is mean square error of the X-coordinate, σ_y mean square error of the Y-coordinate, σ_z mean square error of the Z-coordinate, c focal length, m_b photo-scale, ξ_1 and η_1 the image coordinates, $\sigma_{P\xi}$ accuracy of measured parallaxes, σ_ξ and σ_η are the standard deviation of the image coordinates.

Another way to determinate 3D models by stereo images is the so-called SfM approach [26]. Evolved from the machine vision community, specifically for tracking points across sequences of images captured at different positions, the SfM approach owes its existence to mathematical models, such as the collinearity [27, 28]. The equations of collinearity (see eq. 2) extended with the terms of correction ($\Delta x, \Delta y$), i.e. the additional parameters calculated by Brown model, expressed in relation to the image coordinates, can be obtained in the following form:

$$\begin{aligned}x' &= x'_0 - c \frac{r_{11}(X - X_0) + r_{12}(Y - Y_0) + r_{13}(Z - Z_0)}{r_{31}(X - X_0) + r_{32}(Y - Y_0) + r_{33}(Z - Z_0)} + \Delta x \\ y' &= y'_0 - c \frac{r_{11}(X - X_0) + r_{22}(Y - Y_0) + r_{32}(Z - Z_0)}{r_{13}(X - X_0) + r_{23}(Y - Y_0) + r_{33}(Z - Z_0)} + \Delta y\end{aligned}\quad (9)$$

where

- X, Y, Z 3D object coordinates;
- x', y' 2D coordinates of the initial image;
- X_0, Y_0, Z_0 position of the perspective center;

- c focal length;
- x'_0, y'_0 coordinates of the principal point;
- $r_{11}, r_{12}, \dots, r_{33}$ elements of the rotation matrix;
- $\Delta x, \Delta y$ terms of correction.

As regards the standard error on the Z object coordinate of a generic 3D point may be evaluated by the relation proposed in Fraser (1996) [29]:

$$\sigma_z = \frac{q \cdot Z}{c\sqrt{k}} \sigma_{p\xi} \tag{10}$$

where k is the number of images used to determine the same point and q is design factor expressing the strength of the camera network.

It is important to underline that the collinearity principle and Gauss-Markov model of the least squares are valid and developed on the images acquired by frame camera, i.e. sensors design with central perspective. In this paper, only frame cameras are taking into consideration. The standard error σ_X and σ_Y , respectively in X and Y coordinates, can be calculated by following formula [30]:

$$\sigma_X = \sigma_Y = \frac{d}{c} \sigma_{x'} \tag{11}$$

where $\sigma_{x'}$ is the image measurement precision.

3. METHODS AND SENSORS FOR PHOTOGRAMMETRY PROJECTS

3.1. Methods for 2D or 3D Representation

The choice of the Scale of Representation (SR) and, of consequence, the level of detail, is the first step in the survey phase and it is related to the type of survey to be carried out on the building. From traditional approach, considering the graphical error drawing, variable from 0.1-0.2 mm but in general established in 0.2mm, it is possible to calculate the Graphical Error (E_{GR}) using the following relation:

$$E_{GR} = SR \cdot 0.2 \tag{12}$$

In photogrammetry field, once established the tolerable error, the next step concerns the determination of the Ground Sample Distance (GSD) i.e. the pixel resolution measured on the structure. The value depends mainly on the camera sensor resolution (CCD pixel size), focal length (c) and distance object and sensor (d), according the formula:

$$GSD = \frac{d}{c} CCD \text{ pixel size} \tag{13}$$

Of course, the GSD value should be lower the Graphical Error. From the previous relation, it is possible to obtain the distance (d) at which to perform the photogrammetric survey. In general, in the acquisition step, it is used a value of about half of that GSD planned. Therefore, from the relation (13) it is possible to extract the maximum distance (d) between the camera and object:

$$d = \frac{c \cdot GSD}{CCD \text{ pixel size}} \tag{14}$$

Therefore, once established these parameters, it is possible to choose a specific approach for the building survey and, of consequence, the appropriate strategy in order to construct the individual façades of the building or the 3D model and vice versa (Figure 1).

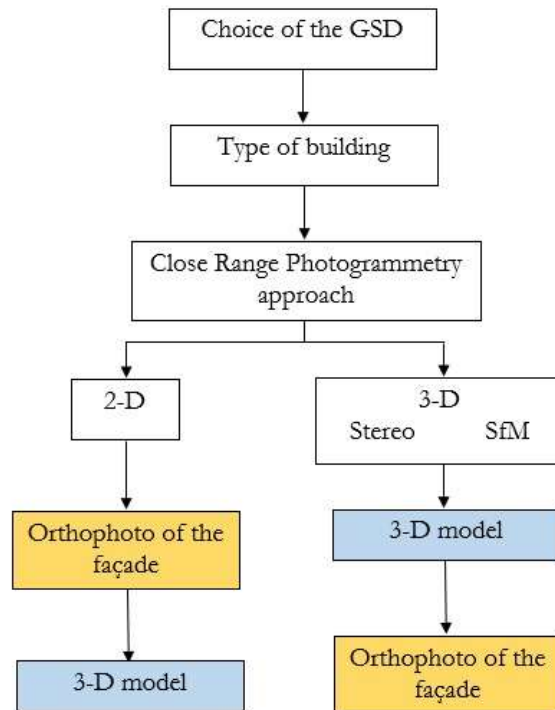


Figure 1 Workflow developed in order to obtain 3D models or orthophoto of the single façade.

In order to show the diverse approaches, three case studies were analysed. In particular, the following surveys were taken in consideration:





- study of a historical building using 2D approach by projective transformation (case study 1);
- study of a modern building using 2D approach by projective transformation (case study 2);
- study of a church belongs Cultural Heritage using SfM approach (case study 3).

3.2. Features of the Sensors Tested

Nowadays, the use of new passive sensor technologies for photogrammetric surveying is increasingly evolving. Considering the stability of the lens and the high performance of the sensors, the Digital Single-Lens Reflex (DSLR) cameras are widespread in this field. For these experiments, the photogrammetric surveys were carried out by Pentax K-x DSLR camera with 18-55mm f/3.5-5.6 lens [31] and by Canon EOS 100D DSLR camera with 18-55mm f/3.5-5.6 IS STM lens [32]. In addition, considering the high performance achievable from the camera contained in the smartphone (see Table 2), the Huawei *P20 Pro* smartphone was used. The Huawei *P20 Pro* smartphone optical sensor, co-engineered with Leica company, consists of three cameras which working in conjunction to improve the quality of the final image: the primary camera is a large, 40-megapixel colour camera with an f1.8 lens, which is joined by a 20-megapixel monochrome camera with an f1.6 lens and an 8 megapixel telephoto camera with an f2.4 lens. The fusion of the images produce in the final capture excellent detail, low noise and accurate colour [33]. Indeed, the monochrome camera adds extra light, detail and depth information. Lastly, an other camera taken into consideration is the one supplied with Xiaomi Mi Drone 4K UHD WiFi FPV Quadcopter drone which it was used for aerial survey of a building. The optical sensor incorporated with the drone is a 4K camera, which is able to acquire images in the dimension of 3840 x 2160 pixels [34]. Beyond to the EXIF (Exchangeable image file format) information associated to each image (from which it is possible to extract some technical characteristics of the sensor) it is very important,

for photogrammetric purposes, to also determine the internal orientation parameters and the radial distortion with high precision. For this reason, Agisoft Lens software [35] was used for camera calibration. Agisoft Lens is an automatic lens calibration software that supports estimation of the full camera calibration matrix, including non-linear distortion coefficients. The feature of each sensor tested and the results of the calibration step (generated by Agisoft lens report) can be summarizing in the following Table 1.

Table 1 Features of the sensors tested and results of the calibration step.

				
	Pentax k-x (f=18mm)	Huawei P20 Pro	Canon EOS 100D (f=18mm)	Xiaomi Mi camera
Applied in the case study	1	2	3	3
Sensor type	CMOS	CMOS	CMOS	CMOS
Pixel dimension (µm)	5.49	1.00	4.29	4.54
Effective pixels (megapixels)	12	40	18	8.2
Max resolution (pixel)	4288 x 2848	7296 x 5472	3872 x 2592	3840 x 2160
Focal length (mm)	18	3.95	18	3.5
Focal length (pixel)	2860	9700	2831	2532
Principal point – x- (pixel)	1944.79	3797.32	1938.57	2004.76
Principal point – y- (pixel)	1325.43	2783.57	1329.66	1482.53
Radial k1	-0.095007	0.139159	-0.092700	0.120684
Radial k2	0.119014	-0.722893	0.069329	-0.188559
Radial k3	-0.250438	-1.369630	-0.026744	0.009318
Tangential P1	0.000683	-0.000434	0.000487	-0.000698
Tangential P2	0.000335	0.001312	0.000146	0.001328

The curves representing the radial distortions of the camera tested are reported in Figure 2: of course, the radial distortion curve goes upward to indicate that the lens has a *barrel* distortion type. From the Figure 2.a, 2.b and 2.c, it is easy to note the distortion values obtained with the camera system in the *P20 Pro* smartphone are lower than those obtained by the Pentax K-m lens and comparable with those of the Canon EOS 100D lenses. Indeed, the maximum value of the distortion achievable *P20 Pro* smartphone is about 15 pixels. Instead, the camera of the Xiaomi Mi drone shows distortion values quite high, with a maximum value of 30 pixel around a radius of 1500 pixel.

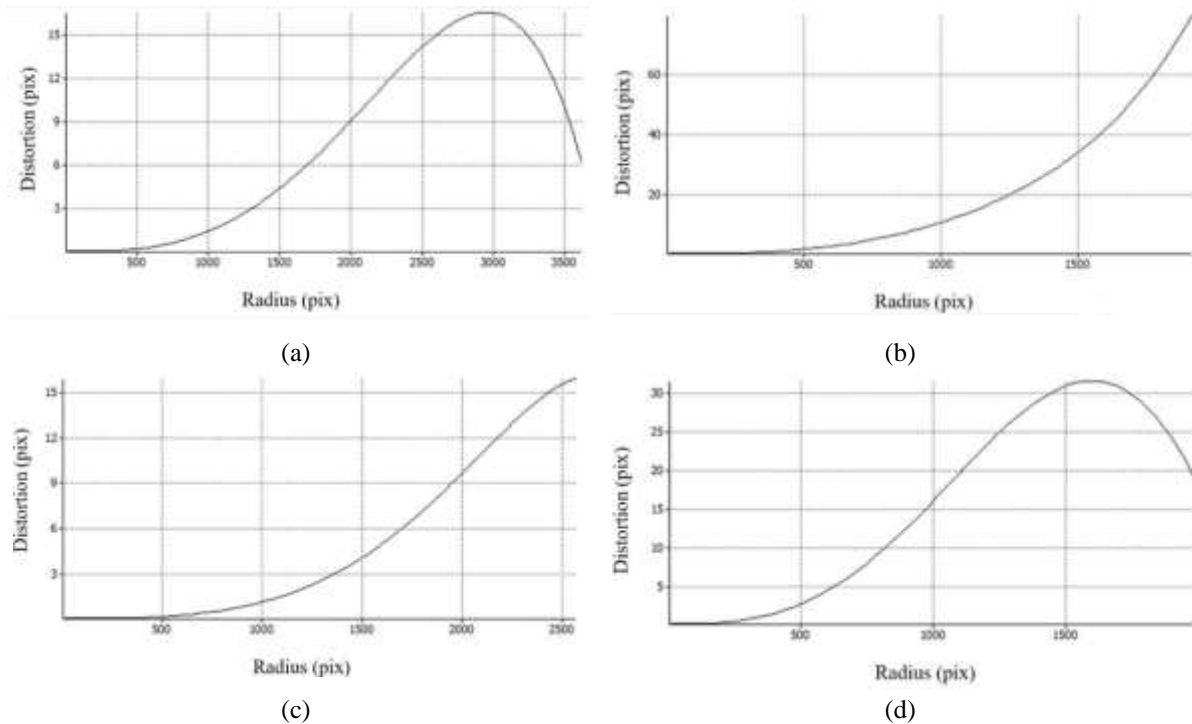


Figure 2 Radial distortion curves of the camera tested.

a) Huawei P20 Pro; b) Pentax K-m; c) Canon EOS 100D; d) Xiaomi Mi camera

In order to evaluate the accuracy and, at the same time, to verify the coherence of the interior parameter obtained with the 2D test field used in the previous procedure, a further calibration session by using a 3D test field was also performed. The 3D test field created for this calibration session was composed by 16 circular (12 bit) coded targets positioned at different heights from the ground. The coordinates of each target were measured by Pentax R-325(N) total station equipped with a laser sensor for non-prism readings and reported in a local reference system by apposite tool developed in Matlab environment. For each sensor taken into consideration, a block of the images with very high overlap and according diverse positions (high convergent network) was realized. The elaboration of the images was carried out by Agisoft Photoscan software and using a PC whose technical characteristics correspond to the minimum configuration suggesting by software manufacturer, i.e. CPU: Quad-core Intel Core i7 CPU, Socket LGA 1150 or 1155, Motherboard: Any LGA 1150 or 1155 model with 4 DDR3 slots and at least 1 PCI Express x16 slot; RAM: DDR3-1600, 4 x 4 GB (16 GB total) or 4 x 8 GB (32 GB total); GPU: Nvidia GeForce GTX 980. The average Root Mean Square (RMS) was calculated in two type conditions: using the internal values parameters obtained by 2D test field and in Self-Calibration mode (Table 2). In relation to the optimal shooting conditions, the high overlap and the optimal network configuration of the images, the accuracy values reached by the different sensors tested, and shown in Table 2, can be considered maximum. In addition, using this configuration, the difference between the accuracy achieved by the calibration methods is really small. This mean that on the field, in order to obtain an elevate accuracy and reliability, it is necessary to avoid a weak geometry of the image configuration. As regards the camera tested, from the Table 2, it is possible to note the high performance achievable from the smartphone Huawei *P20 Pro*. Indeed, already in the self-calibration mode, the precision was of the order of few millimetres.

Table 2 Accuracy on target. Average values of RMS achieved using diverse camera

Type of camera	N. Photos #	Calibration Method	
		Average RMS using Pre-calibrated parameters (mm)	Average RMS using Self-calibration (mm)
Pentax k-x (f=18mm)	77	3	2
Huawei P20 pro	74	2	2
Canon EOS 100D DSLR (f=18mm)	79	1	1
Xiaomi Mi camera	83	5	6

4. EXPERIMENTAL RESULTS

4.1. Case Study 1: Construction of the Photo-Plane and 3D Model using Projective Transformation

The first example discussed in the paper concerns a historical building situated in the old part of a small town (belong the medieval period, i.e. the period of X-XI century) near Salerno city (Italy). The photogrammetric survey was carried out by Pentax K-x camera (approximately 12.4 effective megapixels) with 18-55mm f/3.5-5.6 lens. The DSLR K-x camera incorporates a special developed CMOS image sensor to assure high-speed image data read-out at varying ISO sensitivity levels. Using the equation 14, it was determined the maximum distance between the camera and the object to comply with the project GSD. Because the final scale of the building representation was 1:50, the pixel the reference pixel was 12.5mm. For each side of the building, one shot was made with the exception of the main façade where a mosaic was realized.

In order to simplify the operations of recognizing the control points on the building, beyond to determinate the position of some elements of the building (edges of windows, edges of doors, etc.), special (adhesives) numbered markers was adopted as control points. These markers, equally distributed on the façades of the building, were measured by Pentax R-325(N) total station. The accuracy of this total station in the measure of the distance is 3mm while the angular precision is 5". Because each façade has a local reference coordinate system, the elaborations of the coordinates of the point of the controls were realized by apposite tool developed in Matlab environment.

The software used for the projective transformation of the images was Quantum GIS software (Open Source). This software, thanks the *Georeferencer Plugin*, allows more types of the transformations of the image [36] including also projective transformation. In this project, it was carefully checked not to exceed the tolerance value that, in the building surveys, was placed 2.5 times GSD. The result of the application of the method on this type of building is shown in the Figure 3a where it is possible to note as the upper part of this figure is represented the color orthophoto of the single façades. This type of environment allowed to draw, with accuracy and in a way quite fast, the different architectural elements (windows, doors, plaster, stones, downpipes, railings, etc), as shown in the lower part of the Figure 3a. Every architectural element was recorded on a different layer by a closed polyline (polygons) or polylines and on the same layer, it was reported a unique identification code. In this way, it was possible to connect the graphical information with other informations, i.e. notes, images or tables drafted by other specialists involved in the restoration. However, in order to connect the different informations, it was necessary to transform the vector CAD file in Esri shape files and associate to each polygon the right identifier. This operation was realized by apposite

tool implemented in ArcMap® software. Therefore, the single façade of the building was represented by raster image (orthophoto) and vector file (CAD and shape file).

In order to build a three-dimensional model, the various orthophoto of each façade were assembled. This task was carried out using *SketchUp* software, manufactured by *Trimble* company. This software allowed to create and visualize the 3D representations of the building in a relatively quick and easy way (Figure 3b). Indeed, starting from the metric survey of the building, it was possible to build the volumes of the building and, subsequently, add the texture on each façade.

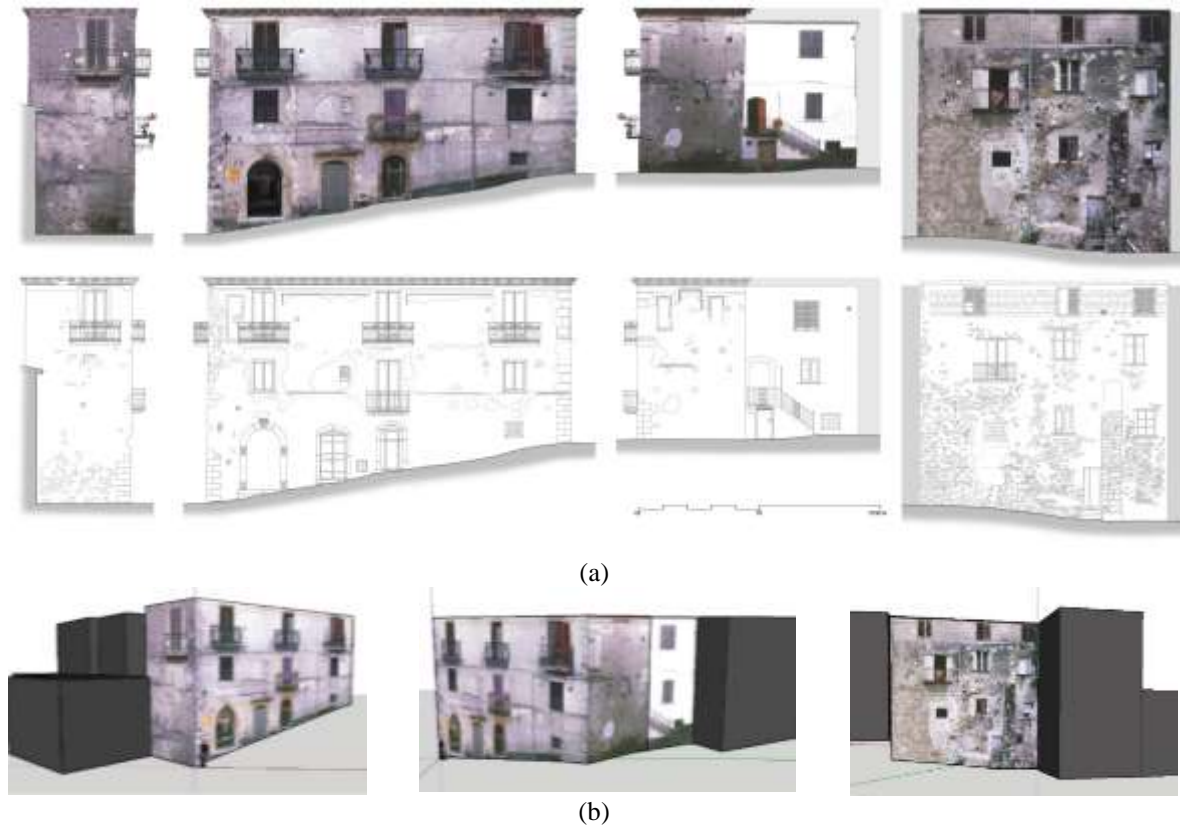


Figure 3 Metric survey of the building.

a) Orthophotos and CAD representation of the single façades surveyed; b) 3D model in different views (the gray volumes are external to the building being investigated)

4.2. Case Study 2: Construction of the Photo-Plane and 3D Model Using Projective Transformation

The structure under investigation concerns a building made with many glass surfaces. This type of structure is very interesting because the reflections of the surrounding environment in the glass surfaces implies the impossibility in merging the images. In other words, the recognition of homologous pixels between two subsequent photos (eg stereo approach) is very complicated. The structure under investigation occupies an area of about 300 square meters and has an elevation of about 12 mt. In this case study, the scale of representation of the project was 1:100 and, of consequence, the GSD adopted was 20 mm. For each side of the building, one shot was made. In particular, the images were generated by Huawei *P20 Pro* smartphone at a resolution of 40-megapixel. The usefulness to perform a survey through a smartphone are manifold: starting with the development of new apps for Google's Android Operating System (OS), such as *Bosch toolbox* (developed by Robert Bosch GmbH) or *ON*

3D-CameraMeasure tool or *ContextCapture* (developed by Bentley Systems Incorporated) which allow to report measurements and notes directly on the photo or obtain approximate measures of an object or create 3D model (for small project), until to the possibility immediate to share and storage the data acquired. In this case study, *Bosch toolbox* was used because it allowed to report the number and the position of the control points directly on the photo as well as to show the approximate measurements of the building (length, width and height) on the photo obtained by traditional instruments for the direct measurement of distances, such as metric roll and Leica DISTO Pro laser distance. In order to verify the GSD value in the acquisition step was smaller than the design one, it is necessary to consider the length of the single facades. The longest façade is that of the main facade which has a length of 27 m. Dividing this value of length (transformed into millimeters) by the pixel number of the sensor (7296 pixel), it is possible to obtain a GSD value lower than the design value, i.e. about 5mm. Of course, if this situation is verified for the longer façade, it follows that it is verified for the other facades. In addition, points easily recognizable in the photos were used as a control point in order to rectify the images of the facades. These points were measured by Pentax R-325(N). As in the previous case, each façade has its own reference coordinate system. On each image the control points were identified and the relative spatial coordinates derived from the topographic survey were inserted in the software. To realize the projective transformation of the images, RDF software was used. This software, developed by the photogrammetric laboratory of the IUAV university [37], allows the projective transformation by two approaches: analytically (calculation of parameters using the least squares method) or geometrically. In this case study, the approach analytically, which use Least Mean Square (LMS) method, was adopted. In this way, the images were transformed and was verified that the average error achieved on the single façade is lower than the tolerable one. The maximum error recorded on the different facades of the building was 20 mm. In the Figure 4, it is shown how it was possible to build the 3D models through assembly of the single orthophoto of the façade.

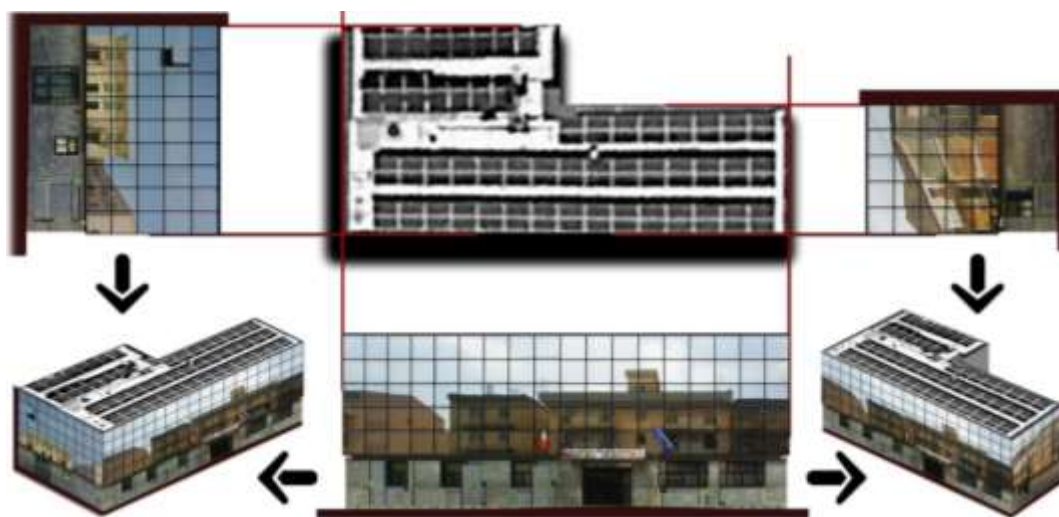


Figure 4 Survey of the building with many glazed surfaces in 2D and 3D views

4.3. Case Study 3: Construction of the 3D Model and Extraction of the Orthophoto of the Façades

The last survey taken into consideration concerns an historical building located in Serre (Italy) and called "*Madonna dell'Olivo*" which is used for Christian religious activities. This building (church) is mentioned in documents dating back to 1508, but the sacred construction is earlier. Therefore, the church is possible to place in the history between the *Swabian* and *Angevin*

period. The approximate dimensions of the structure are: 30m of length, 18m of width and 8 in elevation. Because of the geometric irregularity of the structure, it was not possible to apply the simple projective transformation and, of consequence, a SfM approach was used.

The survey of the church was realized by terrestrial and aerial cameras taking into account that the scale of the representation of the single façades was 1:50. Therefore, the maximum GSD value of the project was 5mm. The terrestrial survey was carried out by Canon EOS 100D DSLR camera with 18-55mm f/3.5-5.6 IS STM lens whose sensor is able to produce 18 megapixel images with a fully-featured (APS-C Hybrid AF II CMOS). In order to cover the whole building, it was necessary to carry out 226 photos. As regards the survey of the upper part of the building, it was realized using Xiaomi Mi Drone 4K UHD WiFi FPV Quadcopter. This Unmanned Aerial System (UAS) is equipped with a 4K camera able to acquire image in the dimension of 3840 x 2160 pixels. In order to obtain clear images, the camera is attached to a 3-axis gimbal, which it stabilizes up to 2000 vibrations per second. The flight step was carefully planned [12] and carried out by means of more strips (or flight lines) varying the position of the camera in order to cover the area of interest by tilt (45°) and nadir photos using an overlap of the 80% and a sidelap of the 60%. In this way, a total amount of 96 aerial photos were carried out.

A topographic network was realized in order to calculate the coordinates of some points of the structures and, of consequence, to obtain an accurate measure of the markers or points of control or points of detail of the structure. In particular, the network was realized using the Pentax total station used even in the cases described above. The topographic network, constituted by 4 vertexes, was materialized through steel topographic nails. The adjustment of the network was realized by *Leica Geo Office* software. The root mean squares of the measures elaborations were resulted lower than 3 mm for xy and z coordinates. On the building, 16 points easily recognizable in the photos were chosen and the relative coordinates were calculated in a local reference system.

The elaboration of the terrestrial and aerial images was carried out by Agisoft Photoscan software which it allows to build 3D model using several straightforward processing steps. In particular, in the first step, Agisoft PhotoScan detects points in the source photos which are stable under viewpoint and lighting variations and generates a descriptor for each point based on its local neighbourhood. This approach is similar to SIFT (Scale-Invariant Feature Transform) algorithm [38] but uses different algorithms for a little bit higher alignment quality. In order to increase the accuracy of the position and orientation of the cameras and to select the effective object of the survey, apposite masks were created on the part of images that are uncorrelated to area under investigation. In self-calibration mode, all images (terrestrial and areas) were aligned with elevated accuracy. In particular, the average RMS achieved in the align step was 22mm. Once the images were aligned, Agisoft PhotoScan built a sparse point clouds which is a 3D approximation of the scene in the images (about 2 million of points). Subsequently, using a greedy algorithm to find approximate image locations and refines those later using a bundle-adjustment algorithm, Agisoft PhotoScan built a dense point clouds. In this step, the software offers different options to generate dense points clouds (Low, Medium, High, Very High). The choice of an option is related to the morphological complexity of the object to be surveyed and the type of PC used for data processing. In this case study, the option “*High*” was used and, at the end of the processing, about 53 millions of points were generated. In order to efficiently reduce computation time of the construction of the mesh, starting from the dense point clouds obtained in the project, it was exploited the network processing available in Agisoft Photoscan. Afterwards, it was possible to build the textured model. In particular, the “*Generic*” mapping mode was chosen. In this way, the software allows to parametrize texture atlas for arbitrary geometry. The software allowed the

export of the single orthophoto of the façades according to more points of view. Subsequently, in CAD environment (using the orthophoto of the façade as a background), the information containing the different architectural elements were represented. The result of the 3D elaboration of the building and the CAD representation of the several façades are shown in the Figure 5.



Figure 5 Metric survey of the church: from 3D to 2D representation

The 3D model obtained by SfM approach requires rather important processing times. Indeed, the only processing times is equal to about 210000s, i.e. about three days. In addition, the dimension of the file of 3D model is very important. For example, the file dimension of the point clouds in the *.LAS format [39] is of 1.54GB. Consequently, the management of these files is rather complex. Therefore, the transformation of these files into 2D files improves the management and sharing of information among the various stockholders involved in the project.

5. CONCLUSIONS

In this paper, in order to build 3D models and/or single orthophoto of the façades, several image-base methods were investigated. In addition, in order to optimize the time of the elaboration and to identify the correct geometry approach, it was necessary to use a suitable method as discussed and suggested in the different case studies. In particular, the orthophoto of the façade can be obtain using several approach depending the type of the building. Indeed, it was possible to obtain the orthophoto of the single façade from 3D model or directly using methods for the transformation of the images. Vice versa, it was possible to build 3D model starting from the single orthophoto of the façade.

Lastly, as shown in the third case study, the potential offered by drones in the survey of the buildings is enormous and represent a valid tool every time that it is necessary to acquire information relative at the high parts of a building which cannot be surveyed through a terrestrial survey.

ACKNOWLEDGMENTS

I would like to thanks G. D'Angelo and M. Corso for the cooperation in the survey of the *Madonna dell'Olivo* building.

REFERENCES

- [1] L. Albertini; P. Sandrini, L'uso della fotografia. Tecniche, autori e applicazioni. 2012.
- [2] G. Konecny, Geoinformation: remote sensing, photogrammetry and geographical information systems. Taylor & Francis, London. 2002.
- [3] M. Gomarasca, Elementi di geomatica. Associazione Italiana di Telerilevamento (AIT) Edizioni, Milano, p. 618, 2004.
- [4] G. Fangi, The book of spherical photogrammetry: Theory and experiences. Edizioni Accademiche Italiane. 2017.
- [5] A. Wiedemann, Digital orthoimages in architectural photogrammetry using digital surface models. International Archives of Photogrammetry and Remote Sensing, 31, pp. 605-609, 1996.
- [6] N. Haala, M. Kada, An update on automatic 3D building reconstruction. ISPRS Journal of Photogrammetry and Remote Sensing, 65(6), 570-580, 2010.
- [7] M. Pepe, S. Ackermann, L. Fregonese, C. Achille, New perspectives of Point Clouds color management–The development of tool in Matlab for applications in Cultural Heritage. The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 42, 567, 2017.
- [8] P. Arias, H. Lorenzo, C. Ordoñez, Simple methods for close range photogrammetry surveying of rural industrial constructions. In XXth International Society for Photogrammetry and Remote Sensing (ISPRS) Congress: Geo-Imagery Bridging Continents, Vol. 4, 2004.
- [9] P. Rodríguez-Navarro, T. Gil Piqueras, Preservation Strategies for Southern Morocco's At-Risk Built Heritage. Buildings 8(2), 16, 2018.

- [10] M. Ishii, Application for Close-Range Photogrammetry Using a Camera System Attached on Transit and Stereo Image System (TOPCON PS-1000/PI-1000). *International Archives of the Photogrammetry, Remote Sensing*, 29, 30-30, 1993.
- [11] A. Conti, L. Fiorini, A. Nobile, L. Menci, F. Ceccaroni, D. Bianchini, M. Ghezzi, Esperienze di rilievo e modellazione 3D per un allestimento interattivo e accessibile da web, *Archeomatica*, Vol.3, 2010.
- [12] M. Pepe, L. Fregonese, M. Scaioni, Planning airborne photogrammetry and remote-sensing missions with modern platforms and sensors, *European Journal of Remote Sensing*, 51(1), pp. 412-435, 2018.
- [13] T. Luhmann, S. Robson, S. Kyle, I. Harley, *Close range photogrammetry*, Wiley, 2007.
- [14] A. Grün, H. A. Beyer, System calibration through self-calibration. *Calibration and Orientation of Cameras in Computer Vision*, Springer, Berlin Heidelberg, pp. 163-193, 2001.
- [15] E. Tufarolo, Auto-calibrazione di fotocamere digitali amatoriali applicata a tecniche di acquisizione multi-scala in fotogrammetria dei vicini. EUT Edizioni Università di Trieste, 2014.
- [16] S. Gindraux, R. Boesch, D. Farinotti, Accuracy assessment of digital surface models from unmanned aerial vehicles' imagery on glaciers. *Remote Sensing*, 9(2), 186, 2017.
- [17] Agisoft, Agisoft PhotoScan user manual: professional edition, 2018.
- [18] K. Kraus, *Photogrammetry: Geometry from Images and Laser Scans* (2nd ed., p. 459). Germany: Walter de Gruyter, 2007.
- [19] D. Liebowitz, A. Criminisi, A. Zisserman, Creating architectural models from images. In *Computer Graphics Forum*, Blackwell Publishers Ltd, Vol. 18, No. 3, pp. 39-50, 1999.
- [20] E. M. Mikhail, J. S. Bethel, J. C. McGlone, *Introduction to modern photogrammetry*. New York, 2001.
- [21] F. Bethmann, T. Luhmann, Least-squares matching with advanced geometric transformation models. *Photogrammetrie-Fernerkundung-Geoinformation*, (2), 57-69, 2011.
- [22] H. B. Papo, Free net analysis in close-range photogrammetry. *Photogrammetric engineering and remote sensing*, 48(4), 571-576, 1982.
- [23] M. Pepe, S. Ackermann, L. Fregonese, F. Fassi, A. Adami, Applications of action CAM sensors in the archaeological Yard. *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*, 42(2), 861-867, 2018.
- [24] V. A. Girelli, *Tecniche digitali per il rilievo, la modellazione tridimensionale e la rappresentazione nel campo dei beni culturali* (Doctoral dissertation, alma), 2007.
- [25] K. Kraus, *Photogrammetry: geometry from images and laser scans*. Walter de Gruyter. 2011.
- [26] D. Fritsch, M. Klein, 3D preservation of buildings—reconstructing the past. *Multimedia Tools and Applications*, 77(7), 9153-9170, 2018.
- [27] J. H. Chandler, S. Buckley, *Structure from motion (SfM) photogrammetry vs terrestrial laser scanning*, 2016.
- [28] K. Al-Ismaeil, *Structure from Motion & Camera Self-Calibration* (Doctoral dissertation, Université de Bourgogne), 2011.
- [29] C.S. Fraser, *Network design. Close range photogrammetry and machine vision*, Caithness: Whittles Publishing. (Ed. K.B. Atkinson), pp. 256–281, 1996.

- [30] M. L. Brutto, G. Dardanelli, Vision metrology and Structure from Motion for archaeological heritage 3D reconstruction: a Case Study of various Roman mosaics. ACTA IMEKO, 6(3), 35-44, 2017.
- [31] Pentax, 2018. <https://mypentax.rioh-imaging.eu/> (Last accessed, 02/08/2018)
- [32] Canon, 2018. https://www.canon.ie/for_home/product_finder/cameras/digital_slr/eos_100d/specification.aspx (Last accessed 14/08/2018)
- [33] Huawei, <https://consumer.huawei.com/en/phones/> (Last accessed 14/08/2018), 2018.
- [34] Xiaomi, <https://www.mi.com/en/drone/> (Last accessed 14/08/2018), 2018.
- [35] Agisoft lens, <http://downloads.agisoft.ru/lens/doc/en/lens.pdf> (Last accessed 10/12/2017), 2017.
- [36] Quantum, G. I. S. Development Team. Quantum GIS geographic information system. Open Source Geospatial Foundation Project, 2013.
- [37] IUAV, <http://www.iuav.it/SISTEMA-E/Laboratori2/cosaoffri/software/index.htm>, (Last accessed on 07/07/2018), 2018.
- [38] Lowe, David G. Object recognition from local scale-invariant features. Proceedings of the International Conference on Computer Vision. 2., pp. 1150–1157, 1999.
- [39] M. Pepe, A survey by Airborne Laser Scanner of open large structure: A case study of Pompeii Amphitheatre. ARPN Journal of Engineering and Applied Sciences, 12(21), 1-11, 2017.