CAAD for Demolished Buildings Modeling with Reverse Engineering Functionality

Athanasios D. Styliadis, Department of Information Technology, The Alexander Institute of Technology & Education (ATEI), Thessaloniki, Greece, styl@it.teithe.gr

Abstract: *This paper is about single three-point perspective historic photography-based CAAD modeling (amateur camera calibration, pose and 3D reconstruction) of man-made environments, buildings and monuments, rich in geometrical regularity. The proposed method, takes profit from the presence in the image (historic photography) of three vanishing directions and two orthogonal object edges with known length ration, and then focuses on the graphical estimation of the skew intrinsic parameter of the uncalibrated camera (i.e. the angle of dot's x and y optical axes, in photography plane), dealing in this way even with the skew presence case (non-rectangle dot). The presence of skew is not a negligible factor in historic photography of early 20th century years, due to dot optical axes failure (carelessness manufacturing) or collapse, as well as the twist effect (distortion) from the undocumented film development process in these years. The graphical recovery of the skew factor is the main contribution of the paper to the pose and CAAD literature. It is shown that a single three-point perspective amateur photography, even with the presence of skew, is adequate for calibration, pose and planar structure (building façades) recovery, if the usual in building's architecture geometric clues are present (i.e. planarity, orthogonality and parallelism) and some metric data (e.g. length and width of demolished building's dimensions) are available. The proposed method was validated on a simulated cuboid and demonstrated on a number of demolished historical buildings for which only one uncalibrated (and skewed) photography was available. The accuracy evaluation shows that the method is suitable for CAAD modeling applications regarding demolished buildings and monuments of the early 20th century years (2% relative accuracy, i.e. 40cm for a 20m façade, included the metric data inaccuracy). The method is of interest for architecture, archaeology, reverse engineering and virtual reality.*

Keywords: Historic photography architecture; Camera calibration; Skew presence; Image-based CAAD; Demolished buildings

1. Introduction

There are many reasons and motives for 3D modeling of real-world objects, buildings and scenes, including: virtual reconstruction of historical buildings and monuments that no longer or only partially exist [1]; digital documentation of historical buildings and monuments for restoration purposes in case of fire, flood, war, earthquake, etc.; ability for virtual interaction without the risk of damage; production of e-learning data for educational resources; virtual tourism; virtual museum exhibits; and interactive on demand 3D visualization of the object, building or scene.

3D model based applications include: digital documentation of buildings, monuments and sites [2,3]; robot navigation and obstacle recognition [4]; augmented and virtual reality; architectural surveying; computer games; virtual tourism; forensics & inverse camera sciences [5]; etc.

In general, most of these applications specify ten requirements: high geometric accuracy; user-friendly interaction environment; all details capturing; quality photorealism; virtual video on demand; high automation and low user interaction; portability; low cost; model size efficiency and application flexibility. The order of importance of these requirements depends on the application's

objective. So, for digital documentation applications the *geometric accuracy* and the *all details capturing* are at the top of these requirements, whilst for virtual tourism the *virtual video on demand* and the *quality photorealism* must take special care.

A single system that can satisfy all these ten requirements is still in the future. In particular, accurately capturing all details with a fully automated system for a wide range of objects remains elusive, as well as elusive remains a system for 3D virtual reconstruction of demolished buildings, when only one uncalibrated photography is available and the building capturing geometry is not rich enough. Refs. to Sabry F. El-Hakim et al. [6] and Paul E. Debevec et al. [7] for the geometry-, image- and range-based available 3D modeling techniques, as well as for some hybrid and multiple techniques.

The current article addresses (into the image-based modeling domain) the problem of single uncalibrated historic photography-based 3D modeling of buildings that display a number of properties of geometric clues (planarity, parallelism, orthogonality, symmetry and planar or space topology). Such as the man-made architectures, buildings and monuments rich in geometrical regularity. This historic photography usually is just a single analogue photograph or a post-card which, nearly always, was captured from an uncalibrated camera for which the skew intrinsic parameter (i.e. the angle of the x and y optical axes of the *dot* in photography plane) is unknown. The knowledge of skew factor's value is essential for the evaluation of the *camera projection matrix* and its graphical recovery is the main contribution of the paper.

The presence of skew is not a negligible factor in historic photography of early 20th century years, due to *dot* optical axes failure (carelessness manufacturing) or collapse, as well as the twist effect (distortion) from the undocumented film development process in these years.

The graphical estimation of the skew intrinsic constraint is based on modern and effective CAD tools (i.e. a virtual camera embedded in a CAD platform and an open software design environment) for a controlled simulation of the photography projection procedure. These tools (virtual camera and open software functionality) are available only on todays modern CAD software controlled environments.

Modern and more effective CAD tools could assist architects and engineers, and the affordances they provide change the practice of modeling itself suggesting fundamentally new ways of thinking about the domain (design and engineering) [8,9,10].

1.1. Current approaches

The classical problem in the image-based techniques is to reconstruct the metric structure of the scene from two or more images by stereovision techniques [11,12]. However, this is a hard task due to the problem of seeking correspondences between different views (time-consuming and costly). Also, these techniques are not applicable in cases where only one, calibrated or not, image is available.

Methodology for 3D modeling from multi or single images, when specific restrictions are applied, is by default *constraint-based* and it tries to exploit geometric knowledge (properties) of the scene clues, like orthogonality, planarity, etc. Geometric properties are either detected automatically [13,14,15], or they are user-defined [16]. Some forms of symmetry have been exploid [17,44] and, in theory, general polynomial constraints on the 3D points could be used. Most often, however, only planarity, parallelism, alignment and angles topology are used [18].

Amongst these constraint-based methods, computation differs between *multi-view* and *single-view* methods. The former being usually modification of traditional 3D modeling methodology [19], while the latter rely on the possibility of expressing geometric properties as linear constraints on the estimated quantities. For this reason, the 3D directions, orthogonal or parallel to planes and edges of interest, should be estimated before the modeling (metric virtual reconstruction) itself [18,20,21]. These 3D directions are called "dominant directions" [22] and

form the *camera calibration matrix*. .

In recent years, a remarkable attention is focused on 3D reconstruction directly from a *single uncalibrated image*. It is well known that only one image cannot provide enough information for a complete and accurate 3D modeling. However, some metric quantities can be inferred directly from a single image with the prior knowledge of geometrical scene constraints. Such constraints may be expressed in terms of vanishing points or vanishing lines, co-planarity, special interrelationship of features and camera constraints [23,24].

There are many studies on the problem of *single view* based calibration, pose and modeling in the literature [25,26,27]. So, there are approaches on constraints-based *camera projection matrix recovery* (using geometric, space topology, photo-realistic and texturing cues) [5,18,20,28,29], as well as particular approaches on applying single view techniques for demolishing buildings modeling from a single historic photography [7,24,30,31,32]. Most of the studies are usually under the assumption of square (i.e. zero-skew and unit aspect ratio) or even rectangle (i.e. zero-skew and known aspect ratio) photography *dots*. However, these assumption are not valid for historic photography with a disorted film digitization and development, and may not be applicable to old fashion camera equipments or even to some off-the-shelf modern digital cameras with skew presence.

In particular: for *geometric cues* based modeling Wilczkowiak et al. [33] and Chen et al. [34] expand the idea to general parallelepiped structures, and use the constraints of parallelepipeds for camera calibration. Wilczkowiak et al. [21] also present a complete duality that exists between the intrinsic metric characteristics of a parallelepiped and the intrinsic parameters of a camera. Also, for a wide variety of man-made environment (architecture, façades, etc.), a *cuboid* is a reasonable model. Caprile and Torre [35] propose a method for camera calibration from vanishing points computed from the projected edges of the cuboids. These vanishing points correspond to three mutually orthogonal directions in space, which can provide three independent constraints to the intrinsic parameters of a camera.

For *line photogrammetry* based modeling several approaches that make use of vanishing points and vanishing lines have been proposed for either cameras calibration, or scene reconstruction. In this domain, the line features in the images are the observations. However, there is an explicit parameterisation for the edges in object space and this increase the complexity of the constraints formulation as it stated by Patias et al. [36], while Criminisi et al. [16] study the problem by computing 3D affine measurement from a single perspective image. The above approaches are based on the vanishing line of a reference plane and the vanishing point in a vertical direction.

Some of the traditional approaches for solving the camera projection matrix recovery problem utilize particular *photo-realistic cues*, such as lighting, shading, texturing and defocusing [28,29]. But, this kind of cues are usually not present in historic photography, and even more these methods make strong assumptions on shape, reflectance or exposure and tend to require a controlled environment, which is often not available and an extra software vertical application is needed.

Zhang et al. [25] propose a method, which combines a sparse set of user-specified *space topology cues*, such as surface position, normals, silhouettes and creases, to generate a well-behaved 3D surface satisfying these constraints.

Several approaches have been reported recently for the exploitation of the basic image geometry for damaged or destroyed buildings modeling using single image techniques [7,24,30,31,32]. All these approaches are based on a square dot assumption (zero-skew). In particular, L. Grammatikopoulos et al. refers to three camera calibration approaches using single images of man-made objects [30]. These three techniques are based on vanishing points, image line parameters and image point observations, all of them with the assumption of square *dots* in photography (i.e. zero-skew single photography). Frank van den Heuvel [24] reports for an accurate virtual reconstruction of the demolished building called "*Kommandantur*", located at the historic center of Berlin. In this paper a single uncalibrated image from the Albrecht Meydenbauer archives

captured in 1911 was used (http://www.geo.tudelft.nl/frs/architec/Meydenbauer/). Also, in this study skew absence was assumed as well.

1.2. The proposed approach

Inspired by the ideas of, new image-based modeling approaches [6,7], introduction of modern computing tools in CAAD [9,10], single-view based camera calibration and pose recovery [23,24], reverse image-to-model CAD-based projection for sensor atitude estimation [15,37,45], building modeling [38,39], and Refs. [40,44] for buildings' *a priori* geometric constraints, the proposed method is targeted on man-made structures, such as architectures, which typically contain three orthogonal principal directions (i.e. three-point perspective photography), and the corresponding vanishing points can be retrieved from the image of straight lines using maximum likelihood estimator [41,42]. the method aims at making full use of scene constraints to obtain a more accurate and photorealistic model of a 3D object.

It is assumed that there are one or several pairs of mutually orthogonal line segments, which lie in the pencil of planes defined by two of the vanishing directions in the scene and the pair of segments are of equal length or with known length ratio. This is not rare for most of man-made objects.

The proposed method mainly focuses on the problem of camera calibration (intrinsic parameters), pose estimation (extrinsic parameters) and façade structure recovery (3D virtual reconstruction) from an amateur (uncalibrated) single photography of rich in structure and geometry demolished buildings. It is based on: (a) the priori building spatial knowledge (vanishing points, topology-based constraints), (b) the automatic and interactive control of a virtual camera embedded in the design session for the recovery of the camera's skew factor, and (c) the use of parametric 3D modeling routines for the boundary representation of the non visible parts of the demolished building (assuming a symmetry cue).

It is well known that three (out of the total five needed for calibration) constraints on the intrinsic parameters of an uncalibrated camera with skew presence (i.e. an amateur photography of the early 20th century years) can be obtained from the vanishing points if three mutually orthogonal directions are defined in an image (three-point perspective photography). However, in geometry rich architecture, there usually exist one or many pairs of building segments, which are mutually in a known topological order (e.g. symmetry, repetition, array) and lie in the planes defined by any two of the three vanishing directions in the structured scene.

It is known [23] that a fourth, linear or quadric, independent constraint to the image of the absolute conic, can be obtained if one pair of the building edges is at a known topology defined order in space. For instance, one pair of equal-length or known length ratio parallel building edges can provide an additional, linear or quadric respectively, independent constraint to the image of the absolute conic. Hence, four independent constraints on an uncalibrated camera could be obtained from one historic photography rich in geometric clues.

The fifth dependent constraint to the image of the absolute conic, the skew factor (s), can be estimated iteratively by applying a software routine (embedded virtual camera into a software controlled CAD environment), which: (i) simulates the historic photography acquisition procedure, and (ii) analytically computes discrepancy vectors between a building's (façade) wireframe model and the image of the historic photography, and examines possible solutions for a penalty function minimization.

So, the camera can be calibrated without the assumption of zero-skew (i.e. a presence of skew-based calibration). Then, on estimating the positions and poses of a number of space planar surfaces (façades) from the recovered camera projection matrix and some scene topology constraints, the demolished building and the scene structure can be virtual reconstructed by combining these planar façades.

The proposed method was validated on simulated data (benchmarking on a pre-defined cuboid) and demonstrated on a demolished, in a 1917 fire, building for which only one historic photography with unknown camera's interior orientation is available, from local authority's photo-archives. The building was choosen according to the multi-cultural history of Thessaloniki, Greece It was a Greek meta-byzantine church (1863), located in a Jewish textile merchant area, close to an Ottoman military campus.

The accuracy evaluation shows that the method is suitable for reverse engineering applications regarding buildings and monuments, if the required modeling accuracy is not less than 2% (i.e. $2x10^{-2}$ or ≈40cm for a usual bulding façade).

1.3. Paper contributions

The main contributions of this paper are: (a) the introduction of a novel graphical approach to skew factor estimation, as the fifth intrinsic parameter (a dependent constraint to the image of the absolute conic with respect to a given world coordinate system), (b) the introduction of the software controlled virtual camera concept (a new CAD utility controlled by a dialog box), and (c) the development of a number of parametric 3D CAAD modeling routines for the boundary representation of non visible, in historic photography, parts of early 20th century architectures (e.g. demolished building).

For these three contributions, the algorithmic basis and essential fragment code is presented in order to promote further research in this area.

So, the camera or/and historic photography skew factor is estimated by an automatic or interactive control of a virtual camera embedded in the design session and therefore the camera projection matrix is recovered. Also, the introduction of the visual camera as a new CAD software tool, utilizes the camera exterior orientation problem by presenting a simple approach for camera pose recovering with respect to a given world system.

Thus, on calculating the object-to-image homography, the demolished building's 3D model can be partly reconstructed by taking measurement on piecewise visible planar patches (façades) of it.

1.4. Paper organization

The remaining parts of the article organized as follows:

In Section 2, the notations, principles and the pinhole camera geometry (dominant directions, vanishing points, calibration and pose), which are similar to that of Refs. [17,18,23,24,40,53] are presented. Also, in this Section, some preliminaries on projection matrix and the absolute conic are reviewed similar to Refs. [19,22,23,24].

In Section 3, the historic photography-based camera recovery is discussed. So, the proposed calibration method, regarding the intrinsic camera parameters, is elaborated in detail with an algorithm and fragment coding.

In Section 4, the CAD-based pose recovery is discussed. Actually, a simple method to recover the pose (extrinsic parameters) and the camera projection matrix is given, together with a relative algorithm and fragment coding.

In Section 5, the historic photography-based demolished building's façade recovery, i.e. the measurement and the 3D reconstruction procedure (CAAD modeling), are presented. Also, in this Section, the development of a number of parametric 3D CAAD modeling routines for the boundary representation of non visible, in historic photography, parts of early 20th century architectures.with known geometry and symmetry is presented (algorithm and fragment coding).

In Section 6, the proposed method was benchmarked on synthetic data (cuboid) and an accuracy evaluation is discussed.

In Section 7, the functionality of the method is demonstrated and its effectiveness is shown on real-world data (a demolished building for which only one historic photography, with unknown interior orientation and skew presence, was available).

Finaly, conclusions and possible extensions are discussed in Section 8.

2. Pinhole camera geometry (definitions, notations)

In this Section the basic definitions and notations used in this paper, in connection with the pinhole camera geometry, are presented.

2.1. Definitions & Notations

Three-point perspective photography is a tilt photography (Fig. 1), where two or three façades of the building are visible; in this photography the 'horizontal' *image lines* (parallel horizontal *object edges*) if extended, in image space, meet at one of two vanishings points (v_x,v_y), as well as the 'vertical' *image lines* (parallel vertical *object edges*) if extended, in image space, meet at a third vanishing point (v_z).

A camera is said to be *calibrated* if the mapping between the image co-ordinates and the directions relative to the *camera optical center (photography vantage point)* are known. Camera's focal length is denoted by *f*, and the view angle of camera's perspective conic is denoted by FoV (Field of View). *Principal point* or *center of projection* is the incoming rays convergence, principal or nodal point (a perfect mathematical point in the pinhole model), and it is defined by the u_x, u_y co-ordinates. Old-fashion bellows cameras (with tilt-shift lenses) used to place the *center of projection* far from the *co-ordinate center* of the image. Film historic lens cameras deviate from the perfect mathematical pinhole camera model in several respects (e.g. lens distortion, film flop and locus of convergence).

Façade, is defined as any building face. So, by construction rules, a façade is a planar geometric structure and hence follows planar homography rules.

An observed at photography primitive pictural single element is denoted, either as 3D *point* (at object space or world plane), or as 2D *dot* (at photography or image plane), or as 2D *pixel* (at CAAD design session) (Fig. 1).

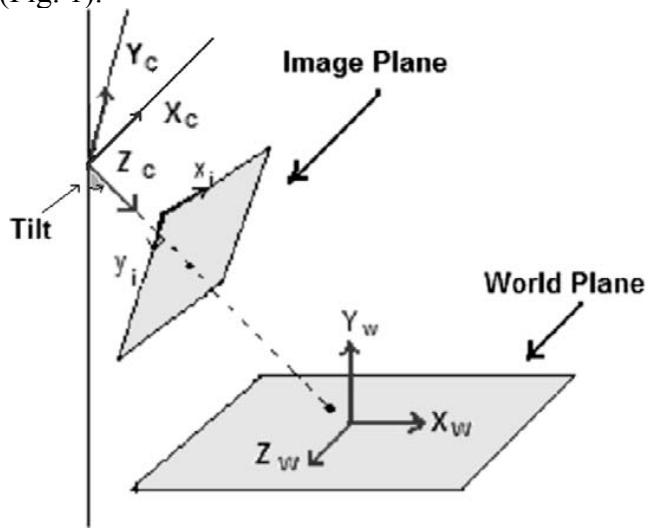


Fig. 1. Camera, image and world co-ordinate systems.

In the perspective camera projection model, a *point* is notated by the homogeneous vector $\mathbf{X} = [X,Y,Z,1]^T$, a *dot* is notated by the homogeneous vector $\mathbf{x} = [x,y,1]^T$, and a *pixel* is notated by the homogeneous vector $\boldsymbol{\chi} = [x,y,1]^T$.

A photography linear element is denoted, either as *object edge* (at object space or world plane), or as *image line* (at photography or image plane), or as *line segment* (at CAAD design session) (Fig. 2).

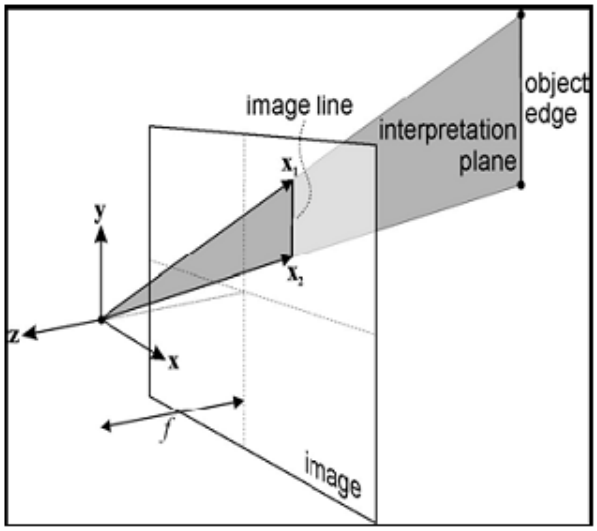


Fig. 2. The object edge - image line mapping.

Rectangular planar structure (patches) is an image of a real-world 3D *rectangle*. Historic photography of architectures is rich in rectangular planar structures (like building façades, doors and windows). A *rectangle* is defined by two object edges, whilst a *rectangular structure* is defined by four image lines which come from two different orthogonal line's groups.

A matrix is denoted by a bold italic upper case letter, e.g. ***K***, ***P***, ***Π, H*** (except the rigid transformation matrix, which is denoted by ***g***).

Projective geometry and *homogeneous co-ordinates* are used throughout this paper [12,40,41,42]. The world co-ordinate system is defined by the X_w, Y_w and Z_w axes, whilst the camera co-ordinate system is defined by the x_c, y_c and z_c axes (Fig. 1).

Also, **Ω**, is the absolute conic on the plane at infinity and **ω** is its image on the plane of photography. Metric entities are indicated with a subscript μ, λ is the unknown scale corresponding to the depth Z of the *point* **X**, and '≈' stands for equality up to scale (a non-zero scale factor).

2.2. Geometry for a single image of a planar building face (façade)

The geometry of a single image of a facade in a three-point perspective photography, is discussed in this sub-section.

2.2.1. The camera calibration matrix (K)

The upper triangle 3 x 3 ***camera calibration matrix*** (***K***) encodes the five intrinsic parameters of the camera, i.e. f_x, f_y, u_x, u_y, aspect ratio (r) and skew factor (s), in the form of:

$$\boldsymbol{K} = \begin{bmatrix} f_x & s & u_x \\ 0 & f_y = r f_x & u_y \\ 0 & 0 & 1 \end{bmatrix} \quad (1)$$

where: f_x, f_y represent the camera's focal length divided by the *dot's* width and height respectively (i.e. corresponding projections to the x_c- and y_c-axis of the camera co-ordinates system (Fig. 2)); u_x, u_y are the co-ordinates of camera's principal point; and s is the skew factor (where α is the included angle of the x_c and y_c *dot* axes); $r=f_y / f_x$ is termed as the *dot* aspect ratio.

The camera calibration matrix contains five unknowns (f_x, f_y, u_x, u_y, s). So, in the general case, five equations described by five linearly independent constraints obtained from the historic photography, are needed for its evaluation (i.e. the straightforward recovery of camera's parameters by Cholesky decomposition [11]).

But, just the first four of the five unknowns are independent and the fifth (s) is a dependent one ($s = f_y \text{ ctg } \alpha$). So, the five solution-equations could be described by four linearly independent constraints and another one dependent one, which could take values from the function ($s = f_y \text{ ctg } \alpha$) describing it in terms of (i) a constant estimated value for the f_y dependency, and (ii) a variable α taking values in a predefined domain.

In the proposed method, the four *linearly independent constraints* are obtained from the single historic photography of the (demolished) building, when the following four equations are satisfied (for more details please see next Section 3):

$$\begin{aligned} &v^T_x \boldsymbol{\omega}\, v_y = 0 \\ &v^T_y \boldsymbol{\omega}\, v_z = 0 \\ &v^T_z \boldsymbol{\omega}\, v_x = 0 \\ &v^T_1 \boldsymbol{\omega}\, v_2 = 0 \end{aligned} \qquad (2)$$

Also, the skew factor dependent constraint is described by the following two equations [43]:

$$\begin{aligned} &\text{(presence of skew case)} \\ &s = f_y \text{ ctg } \alpha \\ &\text{(absence of skew case)} \\ &[1\ 0\ 0]\, \boldsymbol{\omega}\, [0\ 1\ 0]^T = 0 \end{aligned} \qquad (3)$$

where: v_x, v_y, v_z are the three orthogonal vanishing points (a three-point perspective photography is assumed), v_1, v_2 are the two vanishing points resulting from the intersection of two equal-length object edges and a vanishing line in object space, and $\boldsymbol{\omega}$ is the image of the absolute conic (a symmetric matrix with five degrees of freedom) [11,27].

In the proposed method, the solution of these five constraint-equations -for the presence of skew case- is based on: a) a three-point perspective rich in geometry photography (for the first three independent constraint-equations), b) an observation in photography of two equal-length orthogonal object edges (for the fourth independent constraint-equation), and c) in the function $s = f_y \text{ ctg } \alpha$ describing the dependent s constraint in terms of (i) a constant estimated value for the f_y dependency, and (ii) a variable α taking values in the predefined domain 88°, 92°.

For most modern digital cameras literature assumes absence of skew (rectangle *dots*), i.e. $\alpha = 90^\circ$ or $s = 0$. Then, the camera becomes a simplified one with only four intrinsic parameters and the first four linearly independent constraints (Eq. 2) are adequate for *camera calibration matrix* evaluation. Even more, for images acquired by high quality cameras, literature assumes square *dots*, i.e. $s = 0$ and $r = 1$ ($f_x = f_y$), and the camera model is simplified to three parameters accordingly. In this case the first three linearly independent constraints are adequate for *camera calibration matrix* evaluation (Eq. 2).

However, these *dot* geometries are not applicable in historic photography of early 20th century years, because the presence of skew is not a negligible factor, due to *dot* optical axes failure (carelessness manufacturing) or collapse, as well as the twist effect (distortion) from the undocumented film development process in these years.

2.2.2. The camera pose matrix (g)

The ***camera pose matrix*** (***g***) encodes the six extrinsic parameters of the camera, i.e. the rotation angles (ω, φ, κ), and the camera optical center co-ordinates (X_0, Y_0, Z_0). It is a rigid transformation indicating the position and orientation of the camera in space (vantage point), and

factorizes as follows:

$$\boldsymbol{g} = (\boldsymbol{R}, \boldsymbol{t}) \mid (-\boldsymbol{R}, \boldsymbol{t}) \qquad (4)$$

where: $\boldsymbol{R}$ (3 x 3) is the rotation matrix which describes the orientation of the camera, and $\boldsymbol{t}$ (3 x 1) is the translation vector.

It is known [44] that from m images of a *rectangle* (in object space) 2^m possible solutions to the camera pose (relative to the object frame) could be obtained. So, for the single historic photography case, two alternative solutions (giving exactly the same image) for $\boldsymbol{g}$ are expected ($\boldsymbol{R}$, $\boldsymbol{t}$) and ($-\boldsymbol{R}$, $\boldsymbol{t}$).

2.2.3. The camera projection matrix (P)

In the general case, the ***perspective (central) projection*** is described by the following equation:

$$\mathbf{x} \approx \boldsymbol{\Pi}\, \mathbf{X} \qquad (5)$$
$$(\lambda \mathbf{x} = \boldsymbol{R}\mathbf{X} + \boldsymbol{T})$$

where: $\boldsymbol{\Pi}$ ($\boldsymbol{\Pi}$ = R T) is the *perspective projection matrix*; $\mathbf{X} = [X, Y, Z, 1]^T$ and $\mathbf{x} = [x, y, 1]^T$ are vectors containing the homogeneous co-ordinates of the world *points* and image *dots* respectively; '≈' stands for equality up to scale (a non-zero scale factor); λ is a non-zero scalar; and $\boldsymbol{R}$ and $\boldsymbol{T}$ are the rotation and translation matrices respectively.

In the metric pinhole camera case, the ***historic photography*** or ***camera perspective projection matrix*** ($\boldsymbol{P}$), describing the perspective projection process from the Euclidean 3D space to an image, factorizes as follows:

$$\boldsymbol{P}_{\mu} = \boldsymbol{K}\, [\boldsymbol{R}\, \boldsymbol{t} \mid -\boldsymbol{R}\, \boldsymbol{t}] \quad (6)$$
$$(\boldsymbol{P} = \boldsymbol{K}\, \boldsymbol{g})$$

where: $\boldsymbol{P}$ (a 3 x 4 matrix in homogeneous co-ordinates) is the, 11-degree-of-freedom, *camera projection matrix* factored into the intrinsic (five) and extrinsic (six) camera parameters, and the subscript μ is for the metric entities case.

So, in the camera perspective projection model, the 3D *points* $\mathbf{X}$ (in object space) are related to their historic image projections 2D *dots* $\mathbf{x}$ (in photography plane) in the following way:

$$\lambda \mathbf{x} = \boldsymbol{K}\, \boldsymbol{g}\, \mathbf{X} \qquad (7)$$
$$(\lambda \mathbf{x} = \boldsymbol{P}\, \mathbf{X})$$

where: λ is a non-zero scalar related to the depth in space Z of the point $\mathbf{X}$, $\boldsymbol{K}$ is the camera calibration matrix, $\mathbf{g} = (\boldsymbol{R}, \boldsymbol{t})$ is a rigid body transformation represented by a 4 x 4 matrix using homogeneous co-ordinates, and $\boldsymbol{R}$ and $\boldsymbol{t}$ are the rotation matrix and the translation vector from the world co-odinate system to the camera co-ordinate system, known as the exterior orientation parameters or camera attitute (pose) parameters.

2.2.4. The planar homography projective transformation (H)

A **planar homography** is a, 8-degree-of-freedom, projective transformation (described by the 3 x 3 matrx $\boldsymbol{H}$) that operates on planar points (i.e. $\boldsymbol{H}$ is a special case of $\boldsymbol{\Pi}$). This transformation is of particular interest when planar surfaces (like building façades) are imaged. Because it is not possible to extract 3D projective quantities from a single image, if the scene is not planar, the homography plays an important role in the extraction of 3D geometric data from a historic photography image regarding building planar faces (façades).

The planar mapping which relates the co-ordinates of the planar building façade's *points* with the corresponding *dots* at image plane, is characterized by the planar homography:

$$\boldsymbol{H} = \boldsymbol{K}\, \boldsymbol{g} \qquad (8)$$

Homographies, which model the effects of projecting, in perspective, a set of coplanar points onto an arbitrary plane, preserve the straightness of lines.

Lemma 1: *The 3 x 3 homography matrix can be computed linearly by using a known*

rectangle on the object plane, together with the mapped rectangular structure in the image.

Proof: Without loss of generality it is assumed that the 3D *points* (world plane) are specified by homogeneous co-ordinates $\mathbf{X} = [X, Y, 1]^T$ and the 2D *dot*s co-ordinates in photography (image plane) are denoted by $\mathbf{x} = [x, y, 1]^T$. Then, using the homography the *points* and *dots* are related as follows:

$$\mathbf{x} \approx \boldsymbol{H}\,\mathbf{X} \qquad (9)$$
$$(\lambda\mathbf{x} = \boldsymbol{H}\,\mathbf{X})$$

where: '≈' denotes an equality up to scale.

Consider a co-ordinated *rectangular* frame (*rectangle* in object space with known dimensions *a* and *b*) associated with the building façade's plane and observed in the historic photography (Fig. 3).

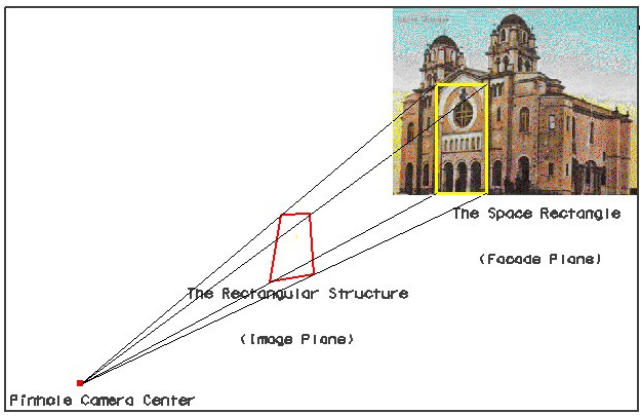


Fig. 3. The homography between the space *rectangle* and the image *rectangular* structure.

Then, in a local co-ordinate system, if one of the four rectangle points, say the upper left, being the origin and the two axes are aligned with the sides of the rectangle. The co-ordinates of the four *rectangle points* are expressed in the following 3 x 4 matrix (***S***):

$$S = \begin{bmatrix} 0 & a & 0 & a \\ 0 & 0 & b & b \\ 1 & 1 & 1 & 1 \end{bmatrix} \qquad (10)$$

So, if the four corresponding *dots* in image plane and the dimensions *a* and *b* in object space are known, the 8-degree-of-freedom matrix ***H*** can be recovered linearly from Eq. (9). However, the homography ***H*** will be recovered only up to scale. Also, ***H*** is a plane (i.e. particular façade)-dependent matrix and could not be used for all the 3D points recovering (because some of them are located in other façades).

The proposed in this paper method is based on the observation in the historic photography of a known dimensions rectangle (e.g. window or door) noticed on one of the building's planar façade s. Then according the Eqs. 9 and 8 the planar homography ***H*** can be linearly recoverd for this particular façade .

Then, due to the camera perspective projection model, the 3D *points* **X** (in object space) of the (demolished) building's particular planar façade are related to their historic image projections 2D *dots* **x** (in photography plane) on this particular planar homography ***H***, so they could recovered linearly by Eq. 9.

For the rest building façade s' points, first the camera pose **g** (***R***, ***t***) for the particular historic photograph is calculated by using this recovered ***H*** (please see Section 4), and then, according to Eq. 7, all the rest bulding façades' *points* (actually the points located in the planar building's surfaces) appeared in the historic photography could be uniquely defined in space (but only in an up-to-scale basis) from their correspondings *dots* in photography, if the ***K*** (camera calibration

matrix) is known.

2.2.5. The image of the absolute conic ω

Lemma 2: *The plane at infinite in space can be expressed as* $\boldsymbol{A} = [0,0,0,1]^T$*, and the mapping between* $\boldsymbol{A}$ *and the image of the infinite plane* $\boldsymbol{a}$ *is the planar homography* $\boldsymbol{H} = \boldsymbol{K}\,\boldsymbol{R}$.

Proof: Readers are refered to [11,27] for the proof of this *lemma*.

Lemma 3: *The image of the absolute conic* $\boldsymbol{\omega}$ *depends only on the camera calibration matrix* $\boldsymbol{K}$.

Proof: The absolute conic $\boldsymbol{\Omega}$, is a conic on the plane at infinity in space **A**, and $\boldsymbol{\omega}$ is its image on the plane of photography. The absolute conic $\boldsymbol{\Omega}$ satisfies the equation:

$$\boldsymbol{\Omega} = \{ x \mid x^T x = 0 \} \qquad (11)$$

where: x is an infinitive *dot* on $\boldsymbol{A}$.

The image $\boldsymbol{\omega}$ of the absolute conic $\boldsymbol{\Omega}$, under the $\boldsymbol{H}$ homography (lemma 1), is:

$$\boldsymbol{\omega} = \boldsymbol{H}^{-T}\,\boldsymbol{\Omega}\,\boldsymbol{H}^{-1} \qquad (12)$$
$$\boldsymbol{\omega} = (\boldsymbol{K}\,\boldsymbol{R})^{-T}\,\mathbf{I}\,(\boldsymbol{K}\,\boldsymbol{R})^{-1}$$
$$\boldsymbol{\omega} = \boldsymbol{K}^{T}\,\boldsymbol{R}^{-T}\,\boldsymbol{R}^{-1}\,\boldsymbol{K}^{-1}$$
$$\boldsymbol{\omega} = (\boldsymbol{K}\,\boldsymbol{K}^{T})^{-1}$$

Remark: It is clear from Eq. 12 that, the image of the absolute conic $\boldsymbol{\omega}$ depends only on the camera calibration matrix $\boldsymbol{K}$.

3. Historic photography: camera recovery (calibration)

In this Section the camera calibration procedure is discussed after the introductin of a number of *lemmas*, *proofs* and *remarks* regarding the five *linearly independent constraints*, which could be obtained from a single photography uncalibrated view of a (demolished) building. These constraints are also known as the five object-to-image internal camera's parameters constraints. Proofs of *lemmas* 4 and 7 could be found in [11,27] and for *lemma* 5 in [23].

3.1. Lemmas, proofs and remarks

Lemma 4: *The first three columns of projection matrix* $\boldsymbol{P}$*, are images of the vanishing points corresponding to the* Xw*,* Yw *and* Zw *axes of the world co-ordinate system respectively. The last column* $\boldsymbol{P}_4$*, is the image of the origin of the world co-ordinate system.*

Remark: In a two perspective (historic) photography of many man-made architectures, three mutually orthogonal pairs of parallel image lines could be obtained. Consequently, the three orthogonal vanishing points v_x, v_y, v_z, can be defined easily. Hence, three linearly independent constraints on the ω ($v^T_x\,\boldsymbol{\omega}\,v_y = 0$; $v^T_y\,\boldsymbol{\omega}\,v_z = 0$; $v^T_z\,\boldsymbol{\omega}\,v_x = 0$) could be obtained from only a single photography of the man-made architecture (e.g. historical building).

Lemma 5: *If two equal-length object edges are orthogonal, then* $v^T_1\,\boldsymbol{\omega}\,v_2 = 0$*, i.e. the vanishing points of two lines with orthogonal directions are conjugate with respect to the ω.*

Remark: Each pair of orthogonal vanishing points can provide one linearly independent constraint on the ω ($v^T_1\,\boldsymbol{\omega}\,v_2 = 0$).

Lemma 6: *The skew is present in historic photography of early 20th century years.*

Proof: The presence of skew is not a negligible factor in historic photography, due to *dot* optical axes failure (carelessness manufacturing) or collapse, as well as the twist effect (distortion) from the undocumented film development process in these years.

Remark: It is well known from camera equipment manufacturers' prospects that, even for the old-fashion camera equipments, the *dot* shape is a square or rectangle or slant (skewed *dot*) with the distorted skew angle (α) taking values between 88° and 92°.

So, in a software controled *virtual camera* (embebbed into a CAD) environment, all the discrete possible values for *dot* optical axes angle (α), could be regarded as input data in a

repetitive, with a user-defined step, calculation process. The goal of this process would be: (*i*) the simulation of the historic photography acquisition procedure, (*ii*) the analytical computation of the discrepancy vectors between a building's (façade) wireframe model and the image of the historic photography, and (iii) a graphical search for the best (calibration and pose recovery) solution or solutions according to a penalty function minimization procedure.

Lemma 7: *The image of the absolute conic (ω) is a symmetric matrix with five degrees of freedom (this is because the ω is defined up to a scale).*

Remark: If four linearly independent constraints can be obtained from an historic photograph and a fifth one (skew factor) could take values from the function ($s = f_y \text{ ctg } \alpha$) describing it in terms of (*i*) a constant estimated value for the f_y dependency, and (*ii*) a variable α taking values in a predefined domain, then the ω can be computed linearly and the intrinsic parameters of the camera can in turn be recovered straightforwardly from ω by Cholesky decomposition [11].

3.2. Calibration from three vanishing points & building a priori geometric constrains (three-point perspective photography)

Camera calibration is refered as the estimation (determination) of the intrinsic (interior orientation) parameters of the camera [24,53]. Following, three general techniques for camera calibration are presented in short, and then, the calibration technique used in paper's proposed method (a vanishing points & building *a priori* geometry-based camera calibration technique) is discussed.

3.2.1. Test field targeted points-based camera calibration

This *test field with targeted points* calibration leads to the most accurate results. Only approximate values for the co-ordinates need to be available for the targets due to the self calibration procedure which determines all parameters simultaneously.

Planar test fields with planar targets make the use of projective geometry advantageous, because the homogeneity of the projection equations leads to linear systems to be solved when determining transformation parameters. This holds true for both, target recognition for reconstruction and for determining approximate values.

3.2.2. Structured domain-based camera calibration

In this case the calibration is based on a number of well identifiable natural (non-targeted) points. The only requirement for the domain is that enough distinct points are detectable, which can be identified as homologous point in the images. Then, fully automatic procedures are able to perform the matching and to determine approximate values for a subsequent self-calibrating bundle adjustment. In this case, the matching procedure relies on projective geometry knowledge, in order to prune the number of the matching candidates (e.g. by exploiting the epipolar constraint between three views using the trilinear tensor). Obviously, the accuracy of this calibration method is less than the targeted points-based one. but it is suitable for scenes 3D reeconstruction image-based methods.

3.2.3. Vanishing points-based camera calibration

If the aspect ratio is assumed to be 1, on the absence of skew, the camera calibration matrix can be computed from the vanishing points corresponding to three perpendicular directions in the building's space [20,36,53].

The vanishing self-polar triangle

For a geometrical approach, regarding vanishing points-based camera calibration (aspect ratio=1 and absence of skew is assumed), simply needs to construct the triangle connecting the

three vanishing points (on the image plane), and then intersect the image with a corner of a cube, in such a way that, each side of this vanishing triangle is coincident with a different cube face. Then, the corner of the cube will be at the original camera center; the *center of projection* (camera's *principal point*) is obtained by dropping a perpendicular line from the camera center to the image plane. This point is actually the orthocenter (intersections of the heights) of the vanishing self-polar triangle. Then, using this principal point with two vanishing points, the *focal length* (*f*) is calculated. Actually, the value of the *focal length* is the length of the perpendicular line.

3.2.4. Vanishing points & building a priori geometric constraints-based camera calibration

In *three-point perspective photography* it is not a hard task to obtain three mutually orthogonal pairs of parallel lines from only one image of many man-made objects, such as architectures (buildings and monuments). Consequently, the three orthogonal vanishing points, say v_x, v_y, v_z, can be computed easily. Therefore, three linear constraints on the **ω** (image of the absolute conic) are obtained from only a single view of the scene.

Most researchers use these constraints to calibrate the camera under the assumption of square *dots* (absence of skew). Since under this assumption, two additional constraints can be obtained, thus, the image of the absolute conic **ω** can be computed linearly by the three linear constraints from the vanishing points. In this case, it is also known that the camera's principal point is the ortho-center of the triangle with the three orthogonal vanishing points as vertices [11]. Similarly, under the assumption of rectangle *dots*, just one additional constraint can be obtained for the absolute conic **ω**.

However, the square and rectangle *dots* assumption, is much less tenable and does not apply for the old-ages cameras, as well as for some off-the-shelf modern digital cameras. So, in these cases, the above assumptions may fail or give a poor solution.

Hence, an obvious question arises: "*can some additional constraints on the ω be extracted from a single uncalibrated view of a demolished building*"?

The answer to this question could be "yes!" in most cases regarding three-point perspective historic photography. This is because most of the man-made architectures (buildings) have rich geometric properties (e.g. symmetry [22,24,40,44] or contain some line segments with equal length or known length ratio [23]). Two equal-length or known length ratio orthogonal object edges (image lines) provide, for the image of the absolute conic, another linearly or quadric independent constraint respectively. More detailed description on this type of constraints can be found in [23,53]. In the proposed method, the two equal-length orthogonal object edges constraint is used as the 4th linear constraint on the image of the absolute conic (**ω**). Also, for the angle α, of the skew factor, a linear array of discrete values between 88° and 92° is used according to an algorithm presented at the end of this Section.

Hence, when a three-point perspective (totally uncalibrated) photograph of a (demolished) building is available and two equal-length orthogonal edges are observed, a 2D array (the intrinsic parameters array, ***IPA***) of possible linear solutions for the camera calibration matrix ***K*** could be obtained.

For the ***IPA*** array the following algorithm (pseudo-code) is used. As a result the array ***IPA*** will has, as far as, 400 intrinsic parameters 'solutions' but only one of them is the appropriate for an accurate camera recovering.

IPA [400,5] array of constraints
Procedure ***Intrinsic_Parameters_Array (IPA)***
begin
i:=1;
α := 88.00 to 92.00 step 0.01
calculation of f_x, f_y,, u_x,, u_y, according

to the equations (2) and the four linear
independent constraints

skew factor evaluation
as a function of f_x

$s := f_x$ ctg α;

***IPA**[i,5] := s;*
i := i+1;

end;

Finaly, some notes for this Section's discussions, regarding single view camera calibration.

- If more pairs of equal-length edges, parallel in space, are available in the building, the vanishing points v_1 and v_2 can be computed by maximum likelihood estimation [41], or other estimation methods, in such a way as to obtain a more faithful result for this linearly independent constraint on the ω.
- If the line segments in Remark for *Lemma* 5 are with known length ratio rather than of equal length, then the angle between the diagonals of its structure (rectangle) can be computed from the length ratio. In this case, v_1 and v_2 can also provide a constraint on the ω, but this constraint is a non-linear. Then, it can be converted into a linear independent constraint, on the **ω**, if the imaged circular points are recovered [23].
- Degeneration will occur if one or more of the four vanishing points v_x, v_y, v_1, v_2 is located at infinity or near infinity in the image. These are the cases when, the camera pose *tilt* (ω) (x-axis) or *pan* (φ) (y-axis) angles, take values apart from the 5° to 40° domain (Fig. 1). The camera's *swing* angle (κ) (z-axis) is related with the vz vanishing point and do not cause any degeneration. A. Zisserman et al. [43] gives a good discussion on the degeneration and ambiguities arising in camera calibration.

4. Historic photography: CAD-based pose recovery

Recovering a 3D planar structure (façade) from a single image is acomplished only when the camera has been calibrated, as well as its position in space (pose) is known.

Pose (***g***) recovery is the calculation of the ***R*** matrix and the ***t*** vector of the pinhole camera with respect to the (historic) photography.

From *lemma* 1, the homography ***H*** between a known rectangle (in façade plane) and the respective rectangular structure (in photography plane) can be computed. Then from Eq. (8):

$$\boldsymbol{H} = \boldsymbol{K}\,\boldsymbol{g}$$

so

$$\boldsymbol{H} = \boldsymbol{K}\,(\boldsymbol{R} \mid \boldsymbol{-R}, \boldsymbol{t})$$

and

$$(\boldsymbol{R}, \boldsymbol{t}) \mid (\boldsymbol{-R}, \boldsymbol{t}) = \boldsymbol{K}^{-1}\,\boldsymbol{H} \qquad (13)$$

Let,

$$[a1, a2, a3] = \text{Function}\,(\boldsymbol{K}, \boldsymbol{H}) \qquad (14)$$

where, a1, a2 and a3 are variables related to ***K*** and ***H*** array elements, and computed as far as the homography ***H*** and the intrinsic camera parameters (***K***) are known (please refer to [23] for more details). Then, the matrix product $\boldsymbol{K}^{-1}\,\boldsymbol{H}$ can be writen as:

$$\boldsymbol{K}^{-1}\,\boldsymbol{H} = [a1, a2, a3]$$

and, from Eq. (13)

$$(\boldsymbol{R}, \boldsymbol{t}) \mid (\boldsymbol{-R}, \boldsymbol{t}) = [a1, a2, a3] \qquad (15)$$

Hence, the two solutions for the pose $\boldsymbol{g}1 = (\boldsymbol{R}, \boldsymbol{t})$ and $\boldsymbol{g}2 = (\boldsymbol{-R}, \boldsymbol{t})$, can be computed from Eqs.

(15) as a function of a1, a2 and a3.

These two alternative solutions give exactly the same image of building's façade in photography, but only one 'puts' the façade in front of the camera and this is the correct one (the other *wrong* (***R***, ***t***) pair places the façade at the back of camera's optical center) [44].

4.1. The algorithm for interior & exterior camera orientation parameters

Lemma 8: *From m images of a rectangle (in object space)* 2^m *possible solutions to the camera pose (relative to the object frame) could be obtained*.

Proof: For the proof of this *lemma* the reader is referred to [44].

Remark: For the classical two images (stereoscopic) case regarding a reflectively symmetric rectangle structure , homography ***H*** is decomposed into four solutions (vantage points) from which exactly the same image is acquired. But, only two of them give *positive depths* of the normal vector under the perspective transformation.

Accordingly, and for the single (historic) photography case, the general solution of the calibration matrix gives two vantage points with *positive* and *negative* normal vector depths respectively. So, for the final *solution*, some knowledge about the topology of the demolished building or the surrounding scene or an estimation of camera's position is needed [19]. These two solutions are related to the two equivalent real camera's vantage points, from which exactly the same building's historic image could be acquired.

The algorithm for intrinsic and extrinsic camera parameters calculation has the following five steps:

Step A: Edge *points* (image *dots*) detection. For edge points detection the *Canny edge detector* could be used, and then the *Hough transformation* would be applied in order to fit the detected *points* (*dots*) into straight lines using the least squares technique.

Step B: Interpretation of the single three-point perspective (historic) photography. This step involves:

- Vanishing points detection (three linear independent constraints).
- two orthogonal façade edges with equa-length or with known length ratio detection (one linear independent constraint).

Then, from the above four linear independent constraints calculation of f_x, f_y, u_x, u_y, according to the Eq. (2) (regarding for the shake of simplicity, *absence of skew*).

Step C: Calculation of the planar homography matrix ***H***. This calculation is based on a known dimensions rectangle in façade's plane (please see *lemma* 1). ***H*** maps the building façade and the historic photography, and it is a façade-depemdent array.

Since the rectangle structure is reflectively symmetric, then there exists two equivalent vantage points from which exactly the same image is acquired (please see *lemma* 8).

Step D: Under the assumption that the skew angle (α) takes values between 88° and 92°, the procedure: ***Intrinsic_Extrinsic_ParamsRecovery (K,G1,G2)***, is called, for the 5th intrinsic (skew) and the six extrinsic camera parameters calculation, as well as the filling up of the relative arrays ***K*** and ***G1***, ***G2*** (Fig. 4).

The six extrinsic parameters are calculated with respect to a local co-ordinate system.

```
K (400,5) array of real (for the 5 camera interior orientation parameters: fx, fy, ux, ux, s)

G1, G2 (400,6) arrays of real (for the 6 camera exterior orientation parameters:ω,φ,κ, Xo,Yo,Zo)

Intrinsic_Extrinsic_ParamsRecovery(K,G1,G2)
begin
    i:=1;
    for α := 88.00 to 92.00 step 0.01
          begin
             K[i,1] := fx;  K[i,2] := fy;
             K[i,3] := ux; K[i,4] := uy;
(skew evaluation as a function of fx and then, filling up the array K at its 5th column)
             s := fx ctg α;
             K[i,5] := s;
(calculation of [a1, a2, a3] from Eq. (14))
             a1, a2, a3 := .....
(calculation of G1 and G2 from Eqs. (15))
             G1[i] := [a1, a2, a3];
             G2[i] := [a1, a2, a3];
              i := i+1;
          end;
end; (proc. IO_Camera_Orientation Recovery)
```

Fig. 4 The intrinsic and extrinsic parameters recovery procedure.

Step E: According to available metric data, setting up in a CAAD design session an initial simple 3D wireframe model of the building or any, visible in photography, part of it (façade , window, door, etc.).

Step F: Virtual camera - Discrepancy vectors. All the ***G1-*** and ***G2****-based* camera perspective building's (façade, window, door) wireframe models are displayed, in turn, on screen in the CAAD design session, and the discrepancy vectors between them and the image of the historic photography are calculated.

Then, from all these ***G1*** and ***G2*** possible solutions the best two (identical), according to a penalty function minimization procedure, are selected. These are the solutions for the two equivalent virtual camera's vantage points, from which exactly the same building's wireframe model image is acquired.

Step G: Final vantage point selection. In this step, the final camera pose selection between the two options from Step E is based, either to a user interaction, or to some approximated initial values for the exterior orientation paramaters.

Step I: Mean values calculations. All the steps between **Step A** and **Step G** are repeated N times (say, N=100) in order to obtain more statistically meaningful mean values, for all the interior and exterior camera orientation parameters, in perspect with the historic photography.

The whole procedure is controlled by a user-friendly dialog, incorporated into the CAAD HCI interface environment (Fig. 5).

Fig. 5 The dialog for virtual camera control (HCI dialog box).

At the heard of the algorithm, for interior & exterior camera orientation parameters calculation, is a software controled *virtual camera* embebbed into a CAAD environment [15,37]. The goal of this *virtual camera* is: (i) the simulation of the historic photography acquisition procedure, (ii) the analytical computation of the discrepancy vectors between a building's (façade) wireframe model and the image of the historic photography, and (iii) a graphical search for the best (calibration and pose recovery) solution or solutions according to a penalty function minimization procedure.

4.2. Pose vector automatic recovery & the role of building's a priori geometric knowledge

In general, the camera orientation estimation requires some building's *a priori* geometric knowledge. In most cases this knowledge is a number of control points, used for the relation of the camera co-ordinate system to the building's façade one. But, control points identification is usually a man-based procedure, which is not so easy to be automated, as control information has to be described in a general manner, which due to the different application areas is not possible. This contrasts with the determination of the relative position of the cameras which can be automated under certain conditions [24,36,37].

As far as the rotation is consernt, it can be derived automatically from images, in case the scene reveals regularities, specifically if parts of it are of a *Legoland* nature, i.e. they show three dominant mutually orthogonal directions. Then, with a vanishing point analysis, the rotation matrix can be derived.

The idea of deriving camera parameters is older than photography, going back to Lambert in the 18th century (Lambert, 1759), and has attracted many researchers especially in computer vision and photogrammetry [4,11,12,14].

The idea of deriving the rotation matrix from vanishing points, based on the fact that the direction of sets of lines in 3D is parallel to the direction from the projection center to the corresponding vanishing point and that the three orthogonal vectors to the three vanishing points,

form the columns of the rotation matrix, assuming the intrinsic parameters of the camera are known.

In the case where no some metric data of the demolished building (e.g. length and width building dimensions) are available, only an interactive solution seems to be feasible [37]. In this case a CAD model could be projected into the image (historic photography) in a reverse projection procedure [37]. In this way by fitting the CAD model to image, its dimensions, position and orientation are estimated [7,37,45].

However, it requires only the precise measurement of two points for specifying the rectangle in the rectified image and the indication on which surface element of the CAD-model it has to be pasted.

4.3. Notes on pose vector estimation

Following, some interesting notes on estimating the orientation of a camera (pose vector) are presented:

- If the intrinsic parameters are not known, the three coordinates of the projection center in the image coordinate system can be determined, together with the rotation matrix, by using a spatial resection [20,26,30,37].
- If the principal point is known, the principal distance and the rotation matrix can be determined from two vanishing points [31,32].
- If the intrinsic parameters of the camera are known, the rotation matrix can be determined from two straight image line pointing to one vanishing point, and one straight image line pointing to one of the other two vanishing points (assuming perpendicular 3D orientation) [36].

5. Historic photography: Demolished building's façade recovery

In this section, the methodology for the recovery and the 3D virtual reconstruction, for all visible (into the historic photography) façades, will be presented. It is assumed that the World Co-ordinate System (WCS) is defined on the particular façade, on which the rectangular structure used for the planar homography was observed (Section 4). This façade is regarded as the *main* one, whilst the historic photography (camera) projection matrix (***P***) is calculated from this main façade with respect to the WCS.

5.1. Building's geometry: hypotheses & constraints

Under the assumption of a number of hypotheses, a single photograph contains valuable information for building's 3D analysis and documentation [20,24], even when it is taken with an uncalibrated camera [30,31,32,53].

5.1.1. Hypotheses (façades geometry)

The proposed method is based on the following hypotheses, as far as the building's geometry is concernt:

Hypothesis 1. *Façades are planar geometric structures*. This hypothesis implies that a polyhedral boundary representation or a B-rep is a suitable type of representation for the (demolished) building.

Hypothesis 2. *Lens distortion is absent*.

Hypothesis 3. *Façade edges are straight*. The intersection of two *façades* leads to a straight edge and, therefore, straight building edges result in straight image lines (photography).

According to a well-known architectural photogrammetry rule, in recovery procedures, *line features* show advantages over *point features* [36]. Therefore, only line features were used as observations in the method.

5.1.2. Constraints (façades geometry)

The proposed method is based on a number of constraints, as far as the building's geometry is concernt. For the single image case, these constraints are::

- Parallelism of building's edges and façades.
- Perpendicularity of building's edges and façades.
- Coplanarity of façade's points.
- Symmetry of building's edges, façades and other featured structures.
- Grouding (landing) of building's featured structures (i.e. a point-in-plane constraint).

Also, for the multiple images [11,13,32], or the image (video) sequence [4,6,19] cases, the so called *shape constraints* could be used.

5.2. Main building's façade recovery

From the recovered historic photography (camera) projection matrix (***P***) and Eq. (7), the single view metrology can be extended to the Euclidean space (façade plane), in order to estimate the position in space of the main façade (main façade's pose), and to calculate the 3D co-ordinates of a number of critical or control points observed at this main façade's image (main façade's recovery).

Nevertheless, this recovery is a relative one (relative façade recovery), as well as it is an only up to scale factor λ recovery. So, a priori geometry knowledge of the main façade is needed for a full recovery.

Lemma 9: *If some prior geometry knowledge in an arbitrary façade plane, such as: a couple of orthogonal lines or the co-ordinates of a point (say: origin or a control point), can be retrieved from an image taken with a calibrated and known-pose camera (i.e. the matrix **P** is known), then the scalar λ, as well as the co-ordinates of many other façade's control points, can be computed*.

Proof: Suppose the co-ordinates of a façade, as a planar structure, are:

$$\boldsymbol{F}_i, \text{ i.e. } \boldsymbol{F}^{\mathrm{T}}_{\mathrm{i}}\,\mathbf{X} = 0$$

with

$$\mathbf{X} = [\mathrm{X}, \mathrm{Y}, \mathrm{Z}, 1]^{\mathrm{T}}$$

then, from an image point $\mathbf{x}_j$ on the photography, its corresponding façade point $\mathbf{X}_{\mathrm{j}}$ can be easily computed by the intersection of the *back-projected line* and the plane of the façade, as:

$$\lambda\mathbf{x}_{\mathrm{j}} = \boldsymbol{P}\,\mathbf{X}_{\mathrm{j}} \qquad \text{(from Eq. 7)}$$
$$\boldsymbol{F}^{\mathrm{T}}_{\mathrm{i}}\,\mathbf{X}_{\mathrm{j}} = 0$$

where: λ is a non-zero scalar related to the depth in space Z of the point **X**j, and ***P*** is the photography (perspective projection) recovered matrix.

Lemma 10: *Some façade entities, such as: the distance between two lines, distance from a point to a plane, angle formed by two lines or two planes, angle formed by a line and a plane, etc., can also be recovered by combining the available geometric building constraints*.

Proof: Similarly to the proof for *lemma* 9, for a line ***l*** in the image (photography), its *back-projection* is a plane in space $\boldsymbol{A} = \boldsymbol{P}^{\mathrm{T}}\boldsymbol{l}$; for a conic ***C*** in the image (photography), its *back-projection* is a cone in space $\boldsymbol{Q} = \boldsymbol{P}^{\mathrm{T}}\boldsymbol{C}\boldsymbol{P}$. Also, like *lemma* 9, the corresponding, to ***l*** and ***C*** image structures, façade co-ordinates, can also be computed linearly by the intersection of their *back-projections* and the façade's plane.

5.3. Other building's façades recovery

The buildings are usually composed of planar surfaces (façades, windows, doors, etc.). So, according to *lemma* 9, if the WCS co-ordinates of each surface can be obtained, then all the 3D information on the particular façade can be recovered. Thus, if and the rest building's facades are recovered accordingly, then the whole building would be assembled by merging the recovered façades into a 3D structure [16,27].

Lemma 11: *If two façades are perpendicular and one of them has known co-ordinates (i.e. façade's pose has been recovered) then, and the other façade can also be recovered.*

Proof: Suppose the façades $\boldsymbol{F}_1$ and $\boldsymbol{F}_2$ are perpendicular, the façade $\boldsymbol{F}_1$ has already be recovered, and the façade $\boldsymbol{F}_2$ passes through line $\boldsymbol{L}$ perpendicular to $\boldsymbol{F}_1$. Say,

$$\boldsymbol{F}_1{}^{\mathrm{T}}{}_{\mathrm{i}}\,\mathbf{X} = 0 \qquad (\mathbf{X} = [\mathrm{X}, \mathrm{Y}, \mathrm{Z}, 1]^{\mathrm{T}})$$

and

$$\boldsymbol{L} = [\mathrm{c}, 0, \mathrm{Z}, 1]^{\mathrm{T}}$$

with respect to $\boldsymbol{F}_1$ co-ordinate system.

Then, $\boldsymbol{F}_2$ must have the same X-co-ordinates as $\boldsymbol{L}$, i.e. the co-ordinates (pose) of the second façade $\boldsymbol{F}_2$ will be:

$$\boldsymbol{F}_2{}^{\mathrm{T}}{}_{\mathrm{i}}\,\mathbf{X} = 0, \text{ with } \mathbf{X} = [\mathrm{c}, \mathrm{Y}, \mathrm{Z}, 1]^{\mathrm{T}}$$

with respect to $\boldsymbol{F}_1$ co-ordinate system.

Then, from an image point $\mathbf{x}_j$ (at the photography), its corresponding at $\boldsymbol{F}_2$ façade point $\mathbf{X}_i$ can be easily computed by the intersection of the *back-projected line* and the plane of the $\boldsymbol{F}_2$ façade, as:

$$\lambda \mathbf{x}_i = \boldsymbol{P}\,\mathbf{X}_i \qquad \text{(from Eq. 7)}$$

$$\boldsymbol{F}^{\mathrm{T}}{}_2\,\mathbf{X}_i = 0, \mathbf{X} = [\mathrm{c}, \mathrm{Y}, \mathrm{Z}, 1]^{\mathrm{T}}$$

where: c is the distance of $\boldsymbol{F}_2$ from the origin of WCS (defined in $\boldsymbol{F}_1$), λ is a non-zero scalar related to the depth in space Z of point $\mathbf{X}_i$, and $\boldsymbol{P}$ is the photography (perspective projection) recovered matrix for façade $\boldsymbol{F}_1$.

Remark: From *lemma* 11, if the WCS co-ordinates of one façade can be obtained using the recovered $\boldsymbol{P}$ matrix, then all the 3D information on other, perpendicular to it, façades can also be recovered. Thus, for cuboid-like building geometries, the whole building 3D model is assembled by merging the recovered façades.

5.4. Parametric CAAD routines for demolished building's non-visible parts modeling

For the non-visible façades a featured library [15] with parametric CAAD models, could be used (for symmetric building structures).

As an example, the code and the 3D model of a parametric cruciform dome as a CAAD element are presented (Figs. 6 and 7).

A) Place a LINE SEGMENT CAAD element & copy it 20 m away on X-axis.
co=0;lc=0;wt=1
place line;xy=0,0,0;dl=10
choose element;xy=5,0,0
copy;xy=5,0,0;dl=20

B) Place an ARC & Create an (open) COMPLEX CHAIN (CAAD elements).
place arc;xy=10,0,0;xy=15,0,5;xy=20,0,0
choose element;xy=5,0,0
create chain xy=5,0,0;xy=15,0,5;xy=25,0,0
xy=100,100,0

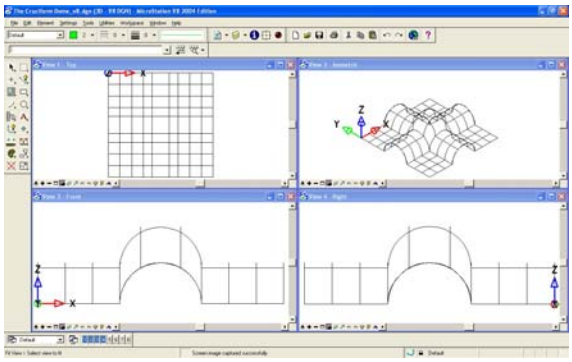
C) Copy the (open) COMPLEX CHAIN CAAD element another three times.
aa=90;choose element;xy=5,0,0
rotate copy
xy=30,0,0;xy=30,-30,0;xy=0,-30,0;reset

D) Display & Construct a B-Spline
choose all; construct surface edge
xy=100,0,0;xy=100,0,0

E) Render the Crociform Structure.
set view phong antialias | render all

Fig. 6. The CAAD command-based code for parametric cruciform dome modeling

A



B

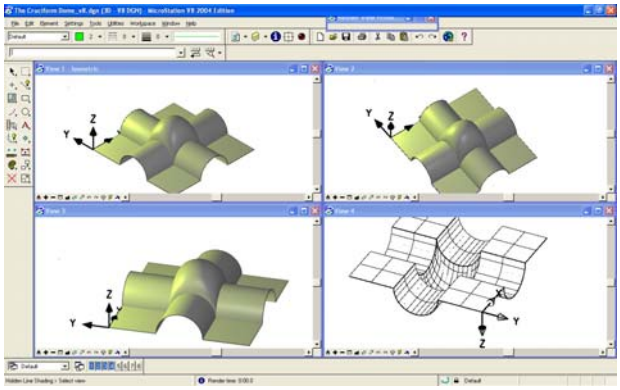


Fig. 7. A. The cruciform dome in wireframe.

B. The cruciform dome in rendered perspectives.

6. Photography-based simulated experiment (method benchmarking)

The proposed method was validated on simulated (synthetic) data. For this case, a CAD cuboid with many parallel edges, and known size (10.0m x 14.0m x 6.0m) and pose was used (Fig. 8). For edge points detection the Canny edge detector was used, and then the Hough transformation is applied in order to fit the detected points into straight lines. The three vanishing points are computed as the intersection of each set of these 'parallel' lines (CAD line segments).

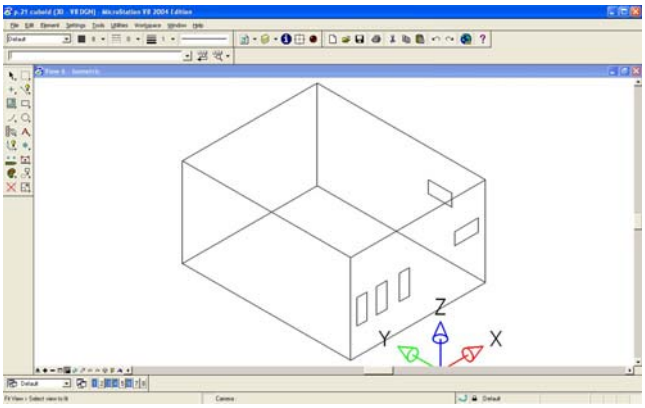
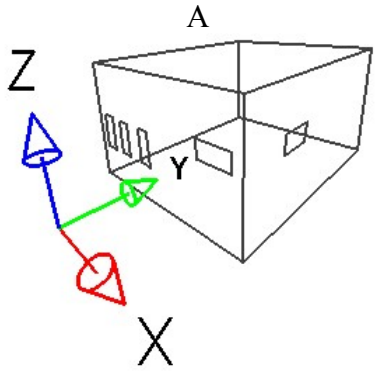
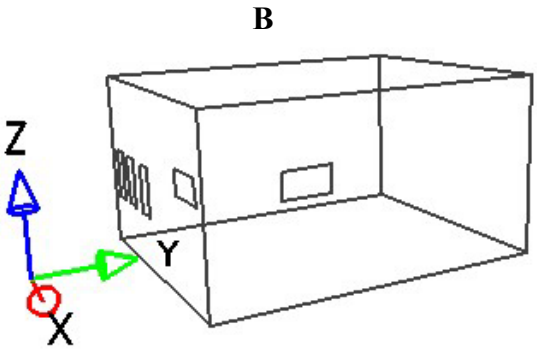


Fig. 8. The virtual cuboid in a CAD controlled environment.

In order to test the method in different imaging (historic photography) conditions, three different images of the cuboid were acquired (3-point projection mode) using the virtual camera CAD tool of MicroStation v8 software (www.Bentley.com). For this purpose an MDL program (MicroStation Development Language, C-like coding) reset the intrinsic parameters of the virtual camera for skew presence uncalibrated images of the cuboid (Figs. 9A and 9B). Notice that the perspective distortion is most visible in the first image (short focal length: 17.10mm / wide angle: 102.20°). For the third cuboid image the default (calibrated, absence of skew) virtual camera was used (Fig. 9C).





C

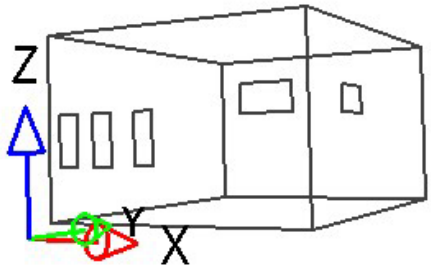


Fig. 9. A. The 1st cuboid image (uncalibrated virtual camera; presence of skew)..B. The 2nd cuboid image (uncalibrated virtual camera; presence of skew). C. The 3rd cuboid image (calibrated virtual camera; absence of skew).

For a more accurate evaluation procedure, the proposed method was comparatively tested -for all these three cuboid images- against the ground truth parameters of the CAD virtual camera, as well as against the DLT (the Direct Linear Transformation, proposed by Abdel Aziz and Karara in 1971) classic camera calibration method.

Also, in order to obtain more statistically meaningful results, one hundred independent tests were performed. Tables 1, 2 and 3 present the means and standard deviations of the relative errors of the five intrinsic and six extrinsic virtual camera parametres.

Table 1

The 1st uncalibrated camera's image (presence of skew): Accuracy evaluation of intrinsic and extrinsic camera parameters recovery (virtual camera as ground truth vs. DLT and proposed method). Mean values (100 tests) ±std. dev.

f=17.10 mm, FoV=102.2°	*f* (mm)	r	u_x (mm)	u_y (mm)	skew (°)	ω (°)	φ (°)	κ (°)	X_0 (m)	Y_0 (m)	Z_0 (m)
virtual camera	17.10	1.00	0.00	0.00	89.00°	-20.00°	15.00°	0.00°	12.00	-23.00	7.00
DLT method	16.92	1.06	-0.33	0.18	90.00°	-20.97°	13.49°	0.45°	12.93	-21.77	7.28
proposed method	16.94	1.05	-0.28	0.15	89.15°	-20.65°	14.52°	0.04°	12.62	-21.90	7.23
±std. dev.	*±0.04*	*±0.02*	*±0.02*	*±0.02*	*±0.12*	*±0.07*	*±0.08*	*±0.01*	*±0.08*	*±0.12*	*±0.06*

Note: The virtual camera's intrinsic and extrinsic parameters are regarded as *ground truth* values.

Remark: In this, like real-world, case (i.e. similar to a historic photography acquired with an uncalibrated and with presence of skew camera) the proposed method behaves well (relative errors for pose in the worst case: 5% and 0.75°) and better than the DLT one; which can only be applied to the absence of skew cases and where more space points (control points) are precisely located.

Table 2

The 2nd uncalibrated camera's image (presence of skew): Accuracy evaluation of intrinsic and extrinsic camera parameters recovery (virtual camera as ground truth vs. DLT and proposed method). Mean values (100 tests) ±std. dev.

f=35.00 mm FoV=62.4°	*f* (mm	r	u_x (mm)	u_y (mm)	skew (°)	ω (°)	φ (°)	κ (°)	X_0 (m)	Y_0 (m)	Z_0 (m)
virtual camera	35.00	1.00	0.00	0.00	89.00°	-10.00°	10.00°	0.00°	20.00	-15.00	6.50
DLT metho	36.02	1.07	-0.39	0.16	90.00°	-10.54°	10.49°	0.15°	21.26	-14.70	6.68
proposed	35.88	1.04	-0.24	0.14	89.21°	-10.21°	10.34°	0.02°	20.92	-15.22	6.62

method											
±*std. dev.*	±*0.04*	±*0.01*	±*0.02*	±*0.02*	±*0.17*	±*0.08*	±*0.06*	±*0.01*	±*0.09*	±*0.08*	±*0.05*

Note: The virtual camera's intrinsic and extrinsic parameters are regarded as *ground truth* values.

Remark: Similarly, like the 1st cuboid image, the proposed method dehaves better than the DLT, as a result of the presence of skew (relative errors for pose in the worst case: 4% and 0.40°).

Table 3

The 3rd calibrated camera's image (absence of skew): Accuracy evaluation of intrinsic and extrinsic camera parameters recovery (virtual camera as ground truth vs. DLT and proposed method). Mean values (100 tests) ±std. dev.

f=35.00 mm FoV=62.4°	*f* (mm)	r	u_x (mm)	u_y (mm	skew (°)	ω (°)	φ (°)	κ (°)	X_0 (m)	Y_0 (m)	Z_0 (m)
virtual camera	35.00	1.00	0.00	0.00	90.00°	20.00°	30.00°	0.00°	10.00	-30.00	2.00
DLT metho	34.82	1.02	-0.02	-0.08	90.00°	20.14°	30.12°	0.02°	10.21	-29.90	2.08
proposed method	35.74	1.03	-0.18	0.12	89.29°	20.20°	29.58°	0.02°	9.52	-30.13	2.16
±*std. dev.*	±*0.03*	±*0.01*	±*0.02*	±*0.01*	±*0.12*	±*0.08*	±*0.06*	±*0.01*	±*0.06*	±*0.06*	±*0.04*

Note: The virtual camera's intrinsic and extrinsic parameters are regarded as *ground truth* values.

Remark: In this case (calibrated camera and *absence of skew*) DLT method behaves better than the proposed one (relative error in the worst case: 0.5% and 0.14°), but the accuracy of the introduced method is still acceptable (relative error in the worst case: 1% and 0.30°). This is due to vanishing points detection errors and to the robustness of DLT in calibrated cameras and absence of skew cases. In multi-images calibration procedures the DLT method gives up to a 0.25% relative accuracy (absence of skew cases).

The above results demonstrate the metric quality of the proposed method in general, and its relatively better performance for the uncalibrated *presence of skew* cases (historic photography). Mainly, this good performance is due to the CAD-supported robustness of the method, as well as to the error-free simulation of the pinhole model of the built-in-CAD virtual camera. Even more, a better accuracy is expected when the target object is a rich in geometric constraints architecture (building).

7. Historic photography-based CAAD Modeling Applications

In this Section a number of figures will demonstrate the potentiality of the proposed method. For this purpose the Aghios Nikolaos Tranos (Thessaloniki, Greece), a non-existing today meta-Byzantine church (1863), was choosen [47,48,49,50]. The church was demolished in the great Thessaloniki fire (August 18th (NS) or August 5th (OS)., 1917) (Fig. 10). The fire victims (homeless) of the three main communities living in Thessaloniki were 50.000 Jews, 12.500 Greek Orthodox and 10.000 Ottoman Muslims, of a total population of 270.000 people (please see: http://en.wikipedia.org/wiki/Thessaloniki). Also, the main consequence of this great fire was the fact that nearly half of the city's Jewish homes and livelihoods, as well as many Ottoman and Byzantine architectures were destroyed, leading to a massive emigration abroad. So, some of the victims (Greeks, Jews, Ottomans) stepped onto the Orient Express to Paris, whilst some other found their way to America (please see:

http://en.wikipedia.org/wiki/Great_Thessaloniki_Fire_of_1917).

A

B

C.

Fig. 10. A. Aghios Nikolaos Tranos on fire! B. Photography from the great Thessaloniki fire (1917). C. A map showing the huge Thessaloniki area destroyed by the fire (1917). In the middle of this area is the location of Aghios Nikolaos Tranos church.

The only available historic photography of Aghios Nikolaos Tranos demolished church (Fig. 11A) can be found in [48]. The digitized, with resolution 484 x 319 pixels, image of this photography has been published to the Internet by the Municipality of Thessaloniki. For the current application, this image has been downloaded from the Web site: http://www.thessalonikicity.gr/eikones/. Also, from a recent excavation, the metric data (dimensions) of this church were found to be 35.50m x 15.50m [49,50].

According to the proposed method results (mean values of 100 trial and error processes), the historic photography was acquired by a camera with: focal length f=33.33mm, aspect ratio r=1.03, field of view FoV=43.65°, principal point co-ordinates u_x=20.25mm, u_y=13.35mm, skew angle α = 89.07° and film format 40.5mm (width) x 26.7mm (height).

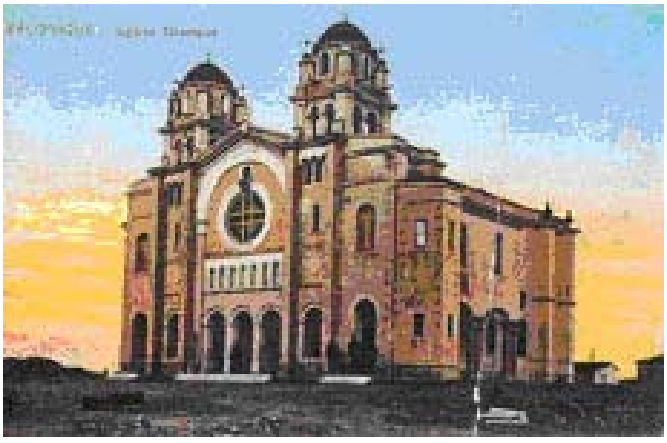
Also, the recovered camera pose (historic photography vantage point) has the following values for its six orientation parameters: X_0=14.79m, Y_0=-27.73m, Z_0=-1.78m, ω=15.96°, φ=35.28° and κ=1.52°.

The error (measured as in the previous Section in respect with the known building's dimensions) in the metric data for camera pose and façades 3D control points, is approximately: 0.32cm and 0.40° for historic photography vantage point pose recovery and 0.65cm for façades 3D control points recovery (i.e. a relative error in the worst case of 2 % for this 35.50m x 15.50m church).

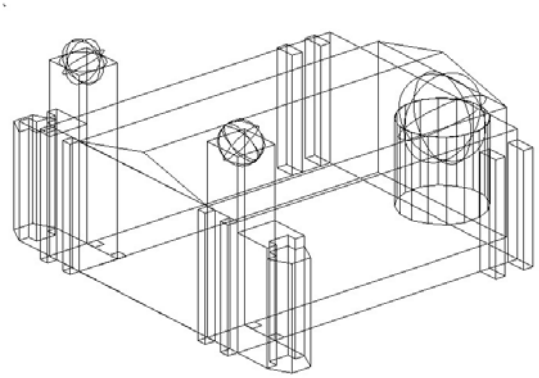
Remark: The performance of the proposed method, in this historic photography case, is better than the simulated cuboid (Section 6), because the structure of the church as a building is much more rich in geometric constraints, rather than the simulated simple cuboid.

Figs. 11, 12 and 13 present the virtual reconstruction result of Fig. 11A by the proposed method. The 3D model is displayed under different viewpoints.

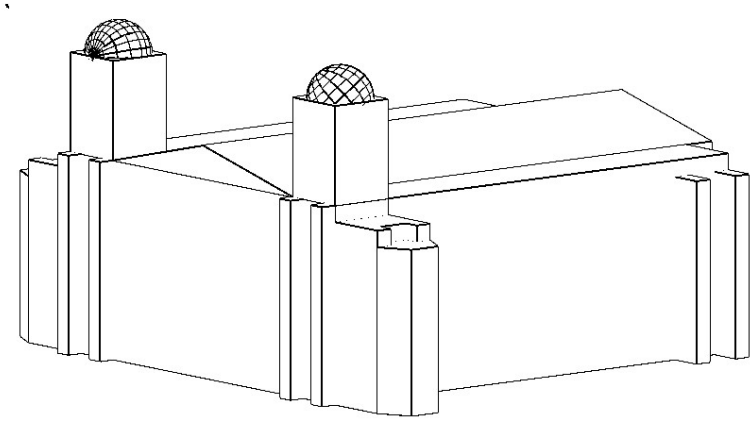
A



B



C



D

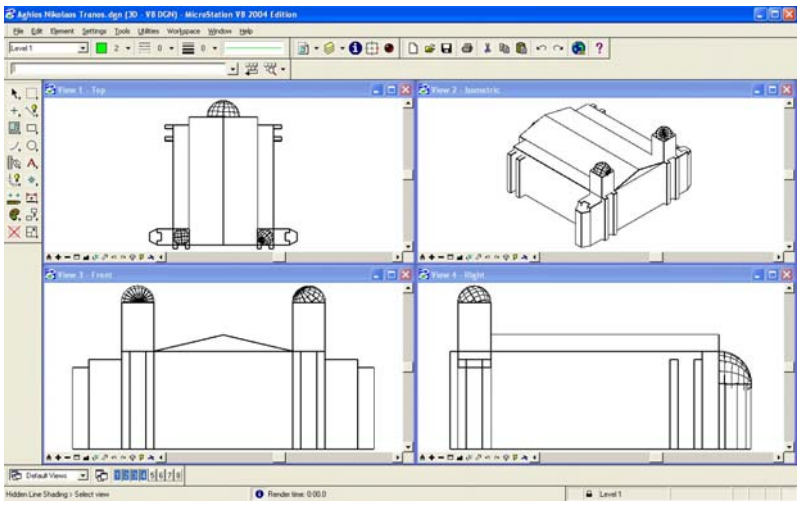
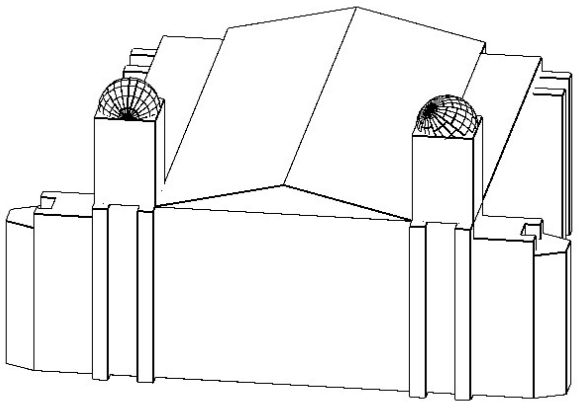
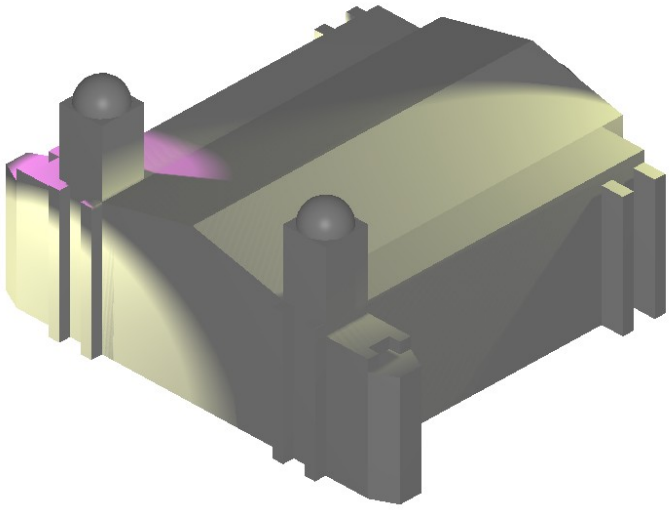


Fig. 11. A. The historic photography of the Aghios Nikolaos Tranos church (two façades are visible). B. The wireframe 3D model. C. The hidden-lines removal 3D model. D. Four architectural plans (views) of the 3D model.

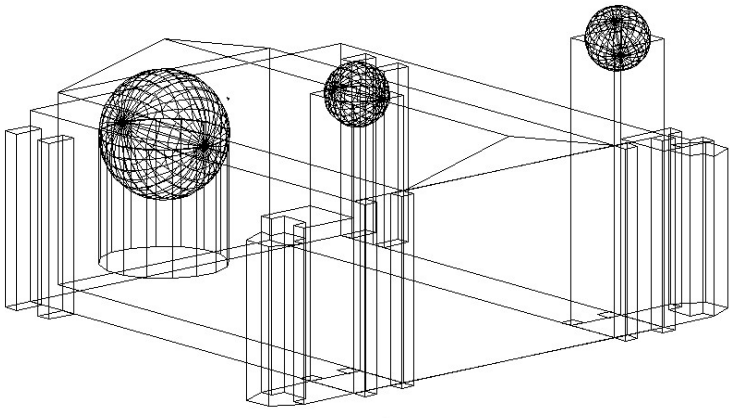
A



B



C



D

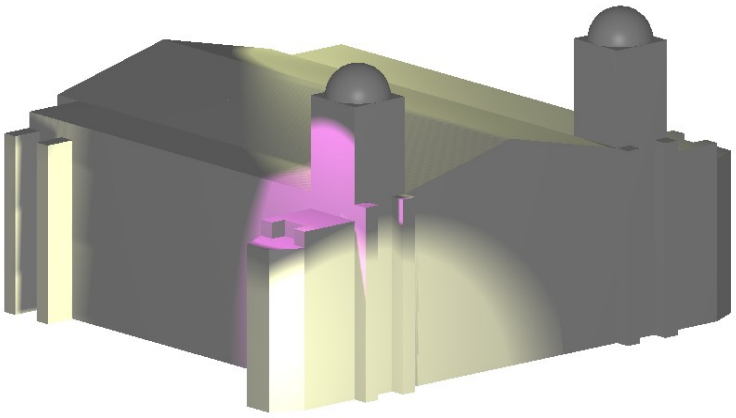
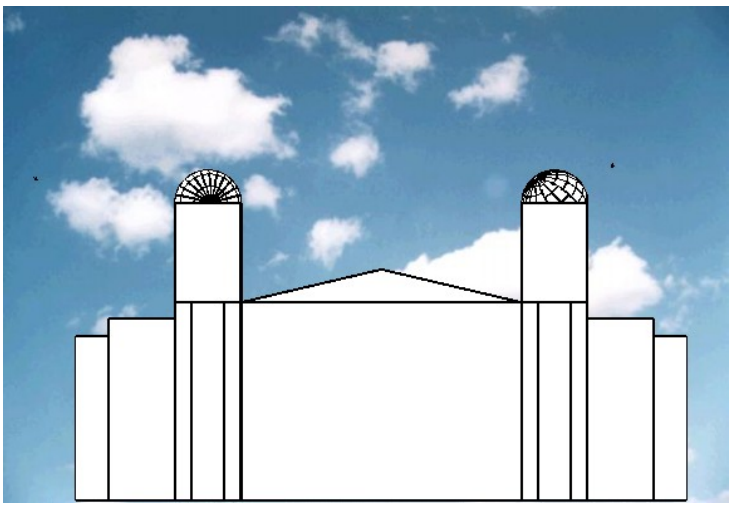


Fig. 12. A. A front - right isometric hidden-lines removal model. B. A right isometric phong rendered view (without materials and texturing). C. A left isometric wireframe model. D. A left isometric rendered view (without materials and texturing).

A



B

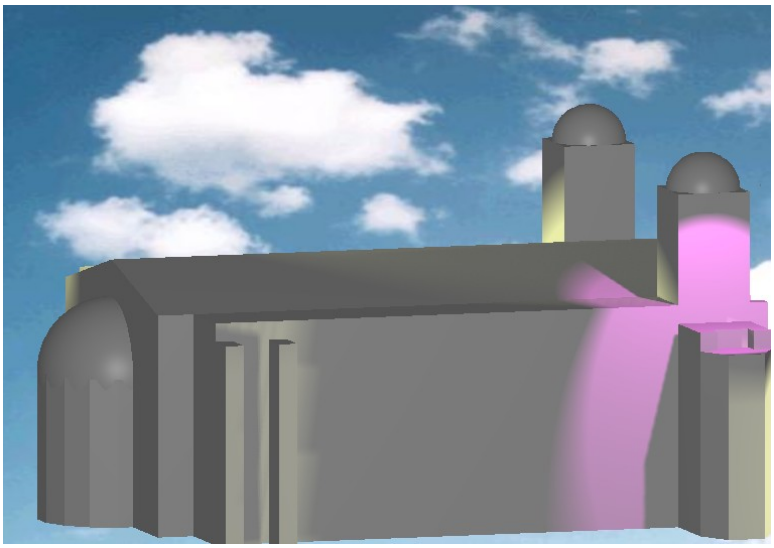


Fig. 13. A. The front façade of the 3D model with a sky backdrop. B. The Aghios Nikolaos Tranos 3D model phong rendered with a sky backdrop.

8. Conclusions

In this paper, after an overview of the current status and prospects of the camera calibration methods, the problem of calibration with the presence of skew is studied in connection with the 3D virtual reconstruction, from a single three-point perspective uncalibrated historic photography, of a structured (rich enough in geometric properties) demolished building.

The presence of skew is not a negligible factor in historic photography of early 20th century years, due to *dot* optical axes failure (carelessness manufacturing) or collapse, as well as the twist effect (distortion) from the undocumented film development process in these years.

The proposed method is based on the use of three vanishing points as linearly independent constraints, as well as on an additional linearly or quadric independent constraint resulting from the observation, in photography, of two orthogonal object edges with equal length or known length ratio respectively. All these four constraints are independent constraints on the image of the absolute conic.

The skew factor (s) (i.e. the 5th linear dependent constraint on the image of the absolute conic) is a linear function of the focal length corresponding to *dot* y-axis (f_y) and the *dot* optical axes angle (α). The f_y is estimated first using the three vanishing points constraints, whilst the α is regarded as a variable taking values from an array of 400 discrete values between two limits (say 88°, 92° with a step of 0.01°). Hence, all the possible values for the skew factor (s) are computed linearly from the estimated f_y value and these discrete α values, and then, these possible s values are regarded as input data for the 5th intrinsic parameter in a repetitive virtual camera-based calibration process.

The camera calibration matrix calculation is based on the above five constraints and then by applying a simple method for the recovery of camera extrinsic parameters (i.e. the camera projection matrix, pose), with respect to a given world co-ordinate system, the homography matrix between the building façade and the historic photography is evaluated.

This single view metrology is then extended to the Euclidean space (façade recovery), in order to estimate the position (pose) and the 3D co-ordinates of a number of critical points of an observed in the photography building's façade (i.e. a space planar surface or patch) from the recovered projection matrix and some building geometry-based constraints. This procedure applied, as well as to any other façade captured into the photography and finaly, the building can be virtual reconstructed by combining these façade 3D models.

At the hard of the proposed method is a software controled *virtual camera* embebbed into a CAAD environment. The goal of this virtual camera is: (i) the simulation of the historic photography acquisition procedure, (ii) the analytical computation of the discrepancy vectors between a building's (façade) wireframe model and the image of the historic photography, and (iii) a graphical search for the best (calibration and pose recovery) solution or solutions according to a penalty function minimization procedure.

The fifth dependent constraint to the image of the absolute conic, the skew factor (s), is estimated as a result of the above routine, and hence, the building's façade 3D model in connection with the historic photography is constructed.

This skew presence consideration is, actually, a contribution to the popular calibration method based on three orthogonal vanishing points and assuming a zero skew factor (skew absence methodology).

Façade's recovery is robust since it is based on a high-speed today's computing iterative process, and very effective as fine details like the depths of windows in a building can also be found, since additional building constraints are utilized during the reconstruction [53]. Also, the method is nearly full automatic, since only a simple user interaction is needed for the final selection between two identical camera pose solutions (with positive and negative normal vectors).

An extensive comparative experimental study with another clasical method is also performed on simulated data [37]. This validation procedure shows that, in case of a monocular image (photography) with presence of skew, the proposed approach performs better than the typical three vanishing points based methodology, for historic photography-based modeling.

Also, for demonstration purposes, a historic photography of a demolished building was used and the created 3D model is presented in Section 7.

The proposed approach will have wide applications in 3D modeling, since a great number of historic amateur photographs could be found in photo-archives regarding demolished man-made objects (buildings and cultural monuments). These architectures tends to have rich geometry and this property is fully exploited by the method

It is a simple and convenient method using a single amateur image and the difficult and costly matching problem of the traditional methods has been avoided at the expense of a minimal human interaction for the observation of geometric constraints.

The method described here assumes perfect data and that façade's image formation is an exact homographic transformation even with the presence of skew. Extensions would be provided

in order to integrate imprecise data (noisy façade data and lens distortion), or to use redundant ones. So, it is clear that the proposed methods are based on some known specific geometrical information about the building, and the accuracy of the approach depends greatly on the image preprocessing, such as image lines (building's edges) selection, line fitting and vanishing point detection. Hence, in order to improve the accuracy of the calibration and reconstruction procedures, the selection of a robust algorithm for lines and vanishing points detection is crucial [18,19,26,35,36,38].

Also, it should be noted that, *lens distortion* has not been considered in the proposed method, since building's photography seems to has a non negligible but predictable lens distortion, even in historic photography, due to predefined rules in the building industry (façades and walls with straight up geometry, etc.). However, the image (photography) should be rectified first (radial lens distortion parameters estimation using image lines with known geometric properties) for lens distortion removal in high-accuracy modeling projects [36,24].

Also, a simple and feasable extension of the proposed method, is the case where more than one historic photographs or post-cards are available. In this case, the method could be profited from this redundancy by applying a least squares adjustment procedure for noise reduction [24] only for façade (planar structure) recovery; since, the calibration and pose recovery are image (photography) dependent procedures.

For future research, a further assessment of the camera pose recovery algorithm using multiple (historic) images [46], as well as the relations between the calibration methods, the collaborative Web-based engineering modeling, the collaborative groupware e-learning functionality and potentiality [51,52], and the building, material and meta-documentation reverse engineering functionalities [53], of any (probably destroyed) historical building captured in an historic photography, must be examined, formulated and documented under the skew presence constraint.

Finaly, some words for CIPA's target group 2. CIPA, an ICOMOS committee for Architectural Photogrammetry, established recently a task group named "*Single Images in Conservation*" for virtual reconstructions of monuments and sites from single historic photography. The website of this task group contains information on various approaches (http://cipa.icomos.org/).

9. Acknowledgements

The current paper is supported by the EPEAEK II - Archimedes research project (Action 2-2-17), "*Personalized learning in a reusable way*", of the Alexander Institute of Technology & Education (ATEI), Department of Information Technology, Thessaloniki, Greece. The EPEAEK II project is co-funded by European social funds (75%) and Greek national resources (25%) (the Greek ministry for education and religious affairs).

10. References

[1] S. Potier, J.L. Maltret, J. Zoller, Computer graphics: assistance for archaelogical hypotheses, Automation in Construction 9 (2000) 117-128.

[1] N.M. Alves, P.J. Bartolo, Integrated computational tools for virtual and physical automatic construction, Automation in Construction 15 (2006) 257-271.

[3] A.D. Styliadis, Digital documentation of historical buildings with 3-d modeling functionality, Automation in Construction 16 (2007).

[4] R. Laganiere, H. Hajjdiab, A. Mitiche, Visual reconstruction of ground plane obstacles in a sparse view robot environment, Graphical Models 68 (2006) 282-293.

[5] Software for photography-based modeling: (a) Façade - Modeling and rendering from photographs, Paul E. Debevec et al., California (http://www.debevec.org), (b) PhotoModeler Pro - Forensics & Inverse Camera, Eos Systems Inc., Vancouver, Canada (http://www.photomodeler. com).

[6] S.F. El-Hakim, J.-A. Beraldin, M. Picard, G. Godin, Detailed 3d reconstruction of large-scale heritage

sites with integrated techniques, IEEE Computer Graphics and Applications (2004) 21-29.

[7] P. Debevec, C.J. Taylor, J. Malik, Modeling and rendering architecture from photographs: a hybrid geometry-and imaged-based approach, in: Proceedings of SIGGRAPH, 1996, pp. 11–21.

[8] G. Carrara, Y.E. Kalay, G. Novembri, Multi-modal representation of design knowledge, Automation in Construction 1 (1992) 111-121.

[9] T. Chastain, Y.E. Kalay, C. Peri, Square peg in a round hole or horseless carriage? Reflections on the use of computing in architecture, Automation in Construction 11 (2002) 237-248.

[10] Y.E. Kalay, The impact of information technology on design methods, products and practices, Design Studies 27 (2006) 357-380.

[11] R. Hartley, A. Zisserman, Multiple View Geometry in Computer Vision, Cambridge University Press, Cambridge, MA, 2000.

[12] O.D. Faugeras, Three-dimensional computer vision: a geometric viewpoint, MIT Press, Cambridge, MA, 1993.

[13] M. Zucchelli, J. Santos-Victor, H. Christensen, Multiple plane segmentation using optical flow, in: British Machine Vision Conference, 2002, pp. 313-322

[14] A. Dick, P. Torr, R. Cipolla, Automatic 3D modelling of architecture, in: British Machine Vision Conference, 2000, pp. 372-381

[15] H. Li, W. Wu, A new approach to image-based realistic architecture modeling with featured solid library, Automation in Construction 13 (2004) 555-564.

[16] A. Criminisi, I. Reid, A. Zisserman, Single view metrology, Intl. Journal of Computer Vision 40 (2000) 123–148.

[17] W. Hong, A.Y. Yang, K. Huang, Y. Ma, On symmetry and multiple view geometry: structure, pose and calibration from a single image, Intl. Journal of Computer Vision 60 (2004) 241-265.

[18] A. Criminisi, Accurate visual metrology from single and multiple uncalibrated image, Ph.D. thesis, University of Oxford, 1999.

[19] M. Pollefeys, R. Koch, L.V. Gool, Self-calibration and metric reconstruction inspite of varying and unknown intrinsic camera parameters, Intl. Journal of Computer Vision 32 (1999) 7-25.

[20] F.A. van den Heuvel, 3D Reconstruction from a single image using geometric constraints, ISPRS Journal of Photogrammetry and Remote Sensing 53 (1998) 354–368.

[21] M. Wilczkowiak, E. Boyer, P. Sturm, Camera calibration and 3D reconstruction from single images using parallelepipeds, in: Proceedings of Intl. Conference on Computer Vision, Vancouver, Canada, 2001, pp. 142–148.

[22] E. Grossmann, J.S. Victor, Least-squares 3D reconstruction from one or more views and geometric clues, Computer Vision and Image Understanding 99 (2005) 151-174.

[23] G. Wang, H.-T. Tsui, Z. Hu, F. Wu, Camera calibration and 3D reconstruction from a single view based on scene constraints, Image and Vision Computing 23 (2005) 311-323.

[24] F.A. van den Heuvel, Automation in architectural photogrammetry (line-photogrammetry for the reconstruction from single and multiple images), Ph.D. thesis, KNAW / Netherlands Geodetic Commission - Publications on Geodesy 54, 2003.

[25] L. Zhang, G. Dugas-Phocion, J.S. Samson, S.M. Seitz, Single view modeling of free-form scenes, in: Proceedings of Computer Vision and Pattern Recognition, Kauai, Hawaii, 2001, pp. 990–997.

[26] R. Cipolla, E. Boyer, 3D model acquisition from uncalibrated images, in: Proceedings of IAPR Workshop on Machine Vision Applications, Chiba, Japan, 1998, pp. 559–568.

[27] G.H. Wang, Z.Y. Hu, F.C. Wu, Single view based measurement on space planes, Journal of Computer Science and Technology 19 (3) (2004).

[28] S.K. Nayar, Y. Nakagawa, Shape from focus, IEEE Transactions on Pattern Analysis and Machine Intelligence 16 (1994) 824–831.

[29] B.J. Super, A.C. Bovik, Shape from texture using local spectral moments, IEEE Transactions on Pattern Analysis and Machine Intelligence 17 (1995) 333–343.

[30] L. Grammatikopoulos, G.E. Karras, E. Petsa, Camera calibration approaches using single images of man-made objects, in: CIPA Proceedings (http://cipa.icomos.org), Antalya, Turkey, 2003.

[31] G.E. Karras, E. Petsa, Metric information from uncalibrated images, in: CIPA Proceedings (http://cipa.icomos.org), Olinda, Brasil, 1999.

[32] I. Kalisperakis, M. Rova, E. Petsa, G.E. Karras, On multi-image reconstruction from historic photography, in: CIPA Proceedings (http://cipa.icomos.org), Antalya, Turkey, 2003.

[33] M. Wilczkowiak, E. Boyer, P. Sturm, 3D modeling using geometric constraints: a parallelepiped based approach, in: Proceedings of European Conference on Computer Vision Vol. IV, Copenhagen, Denmark, 2002, pp. 221–237.

[34] C.S. Chen, C.K. Yu, Y.P. Hung, New calibration-free approach for augmented reality based on parameterized cuboid structure, in: Proceedings of Intl. Conference on Computer Vision, Corfu, Greece, 1999, pp. 30–37.

[35] B. Caprile, V. Torre, Using vanishing points for camera calibration, Intl. Journal of Computer Vision 4 (1990) 127–140.

[36] P.G. Patias, E. Petsa, A. Streilein, Digital line photogrammetry, IGP Bericht (Report) 252, ETH Zurich, 1995.

[37] A.D. Styliadis, P.G. Patias, I.G. Paraschakis, CAD-supported determination of sensor attitude in terrestrial photogrammetric applications, in: Intl. Archives of Photogrammetry and Remote Sensing XXXI (B5), Vienna, Austria, 1996, pp. 451-456.

[38] G. Suter, K. Brunner, A. Mahdavi, Building model reconstruction based on sensed object location information, Automation in Construction 16 (2007).

[39] L.D. Luca, P. Veron, M. Florenzano, Reverse engineering of architectural buildings based on a hybrid modeling approach, Computers & Graphics 30 (2006) 160-176.

[40] J. Košecká, W. Zhang, Extraction, matching, and pose recovery based on dominant rectangular structures, Computer Vision and Image Understanding 100 (2005) 274-293.

[41] D. Breaz, Integral opeators on the TUCD(α)-class, ACTA Universitatis Apulensis (Mathematics - Informatics) 8 (2004) 40-49.

[42] M. Risteiu, A. Iosif, Web-based modeling of coal mine-CFTPP virtual consortium: knowledge management for a best representation of an enterprise, ACTA Universitatis Apulensis (Mathematics - Informatics) 7 (2004) 379-388.

[43] A. Zisserman, D. Liebowitz, M. Armstrong, Resolving ambiguities in auto-calibration, Philosophical Transactions of the Royal Society of London, Series A 356 (1740) (1998) 1193-1211.

[44] A.Y. Yang, K.H. Huang, S. Rao, W. Hong, Y. Ma, Symmetry-based 3-D reconstruction from perspective images, Computer Vision and Image Understanding 99 (2005) 210-240.

[45] D.G. Lowe, Fitting parameterized three-dimensional models to images, IEEE Transactions on Pattern Recognition and Machine Intelligence 14 (1991) 441-450.

[46] D. Cooper, T.P. Pridmore, N. Taylor, Assessment of a camera pose algorithm using image of brick sewers, Automation in Construction 10 (2001) 527-540.

[47] T.S. Mantopoulou-Panagiotopoulou, The monastery of Aghios Menas in Thessaloniki, Dumbarton Oaks (Trustees for Harvard University) papers, 50 (1996) 239-262.

[48] G. Moutsopoulos, Thessaloniki 1900-1917, M. Molho Editions [in Greek].

[49] E. Marki-Aggelkou, Aghios Nikolaos Tranos excavation, Macedonika 19 (1979) 271-298.

[50] T.S. Mantopoulou-Panagiotopoulou, New data for the Aghios Nikolaos Tranos church in Thessaloniki (1863) - A meta-byzantine typology approach, Makedonika 29 (1993-1994) 132-174.

[51] N. Doukas, A. Andreatos, Advancing electronic assessment, IJCCC Vol. II (1) (2007) 53-62.

[52] I. Kitagaki, A. Hikita, M. Takeya, Y. Fujihara, Development of an algorithm for groupware modeling for a collaborative learning, IJCCC Vol. II (1) (2007) 63-69.

[53] P.E. Debevec, Modeling and rendering architecture from photographs, Ph.D. thesis, University of California at Berkeley, 1996 (www.debevec.org).

The Author:

Athanasios D. STYLIADIS *is an associate professor of computer graphics, at the Department of Information Technology at the Alexander Institute of Technology & Education (ATEI), Thessaloniki, Greece. He was born in 1956 in Florina, Greece and he received a Diploma in rural and surveying engineering (Aristotle University of Thessaloniki, Greece, 1980), an M.Sc. in computer science (Dundee University, Scotland, 1987), and a Ph.D. in engineering for his research in CAAD, GIS, and computer modeling (Aristotle University of Thessaloniki, Greece, 1997). Prof. Styliadis was a fellow research scholar at the Department of Geomatics, University of Melbourne, and at the Center for GIS and Modeling (CGISM), Australia. Also, he has been worked at the Hellenic Army Geographical Service (HAGS, Athens) for three years as CAD and GIS system analyst and programmer. He has over 60 journal and conference proceedings publications and he is the author of three books in: (i) Computer Graphics (424 p., 1999, 2002, 2006), (ii) Human-Computer Interaction - HCI Programming (488 p., 2002), and (iii) GIS - Spatial Reasoning & Geomatics Engineering (416 p., 2003). Prof. Styliadis' current research interests are mainly on design computing and digital architecture (computer modeling, CAAD, hybrid CAD/GIS systems, digital documentation of historical living systems, monuments and sites).*