LightCam: ENLIGHTENING THE CAMERA OBSCURA – WHERE PHOTOGRAMMETRY, COMPUTER AND ROBOTIC VISION MEET

E. Nocerino [1]*, F. Menna [2], R. Hänsch [3]

[1] LIS UMR 7020, Aix-Marseille Université, CNRS, ENSAM, Université De Toulon, Marseille, France - erica.nocerino@univ-amu.fr
[2] 3D Optical Metrology (3DOM) unit, Bruno Kessler Foundation (FBK), Trento, Italy - fmenna@fbk.eu
[3] German Aerospace Center (DLR), Microwaves and Radar Institute, SAR Technology, Germany - ronny.haensch@dlr.de

Commission V, V/4

KEYWORDS: ISPRS ECB initiative, educational web-based resource, photogrammetry, computer vision, robotic vision

ABSTRACT:

In the last years, vision-based systems have flourished at an unprecedented pace, fuelled by developments in hardware components (higher resolution and higher sensitivity imaging sensors, smaller and smarter micro controllers, just to name a few), as well as in software or processing techniques, with AI (Artificial Intelligence) leading to a landmark revolution. Several disciplines have fostered and benefited from these advances, but, unfortunately, not always in a coordinated and cooperative way.

When it comes to image-based sensing techniques, photogrammetry, computer vision and robotic vision have many contact points and overlapping areas. Yet, as for people of different cultures and languages, communicating among the three different communities can be very harsh and disorienting - especially for beginners and non-specialists.

Driven by a strong educational and inclusive ambition, the LightCam project is funded by the ISPRS Education and Capacity Building Initiatives 2020 (ECB). The project's ambition is to act as an interpreter and ease the dialog among the three actors, i.e. photogrammetry, computer vision and robotics. Two intermediation tools will be developed to serve this aim: (i) a dictionary of concepts, terminology and algorithms, in the form of a knowledge base website, and (ii) a code repository, where pieces of code for the conversion between different formulations implemented in available software solutions will be shared.

1. INTRODUCTION

1.1 Background and motivations

We perceive the world through our senses, where vision is the one we rely on the most. Since time immemorial, vision has represented our primary resource to observe, investigate and understand the nature and essence of things. Not surprisingly, we have mimicked this natural characteristic, making vision the core sensing approach in visual arts, science and technology. Drawings, paintings and images were the preferred form of representation, documentation and communication before the invention of photography, which marked a milestone not only in art but also in engineering. In fact, it was not long after its invention that photography was introduced to map cities (Jensen, 2007) or survey buildings (Albertz, 2002), applications that represented the birth of photogrammetry.

Since then, imaging devices have significantly developed. The transition to digital and miniaturised sensors, coupled with algorithmic advances and increasing computing capabilities have opened new research and technological frontiers.

Vision, with its underlying mathematical and geometric principles, is a common ground for several scientific and technological domains, especially for photogrammetry, computer vision and robotics.

Photogrammetry is the art or science of obtaining reliable measurements by means of photography (ASPRS, 1952). Its main applications are mapping, surveying and 3D modeling, with great emphasis on accuracy, reliability and completeness of the results (Förstner & Wrobel, 2016).

Computer vision designs computational models able to reproduce the human visual system to support autonomous systems (Huang, 1996). It investigates theory and methods for the automatic extraction, analysis and understanding of information from single images or sequences.

Robotics is a vast engineering discipline which deals with robots, i.e. programmable machines designed to perform a series of tasks autonomously, or semi-autonomously. Robotic vision is one of the technologies that enables the robots with sensing capabilities and, strongly related to computer vision, is the focus of this project.

Whereas computer and robotic vision can be considered closely related, often sharing a common basis of terminology and theoretical formulations, in the past, these disciplines and photogrammetry have rarely crossed, growing up with different motivations in different backgrounds and developing their own nomenclatures and conventions.

At the beginning of the 90's, photogrammetric methodology was still limitedly applied in robot vision because, at that time, the severe constraints imposed to assure high precision and reliability contradicted the needs of high-speed performance in robotics (Gruen, 1992).

With the advent of the internet and the easier possibility to access information, attempts have been made to create a bridge between photogrammetry and computer/robotic vision.

A prominent example is the book edited by Gruen & Huang (2001) entitled 'Calibration and orientation of cameras in computer vision'. It was the result of a workshop on the same topic held in 1992 and covering topical aspect of machine (or computer) vision. The articles collected in the book showed that these themes were core subjects to both photogrammetry and computer vision, at the same time showing how they were handled from different perspectives.
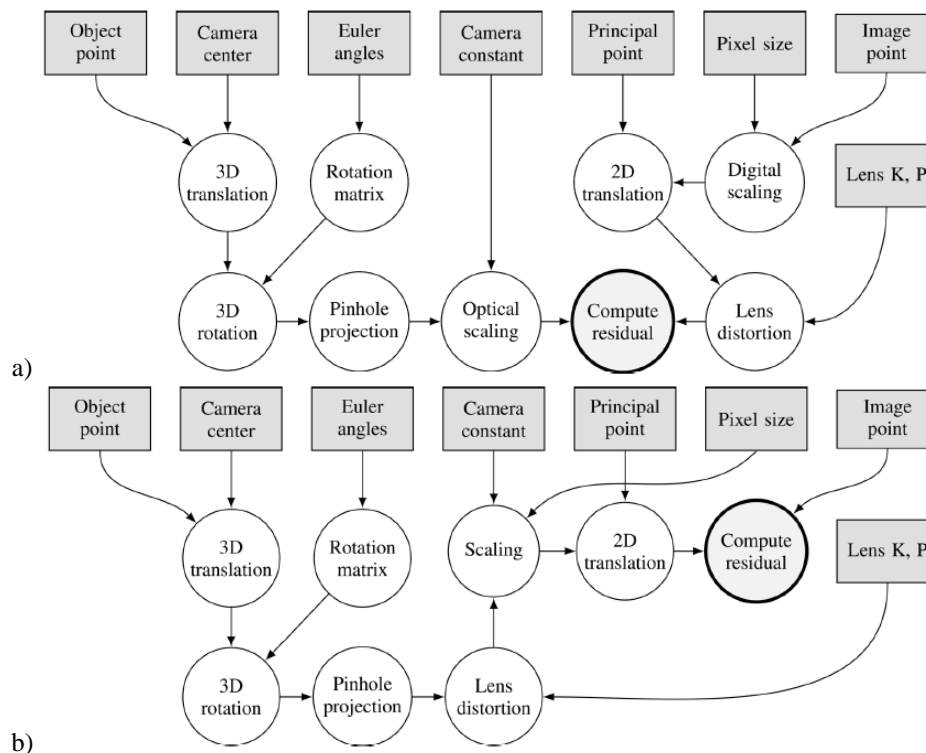
Figure 1: Computational chain according to the photogrammetric (a) and computer vision (b) formulation (Borlin et al., 2019):
In (a), an optical scaling in the camera coordinate system is applied, resulting in image coordinates expressed in mm. The image coordinates are scaled from pixels to mm before lens distortion parameters are used to "correct" the measured image coordinates. The difference between the projected ideal and corrected points provides the residual, highlighted in bold.
In (b), the lens distortion coefficients are used to 'distort' the projected ideal point in normalized units; the image coordinates are then scaled to pixel units. The residual (bold circle) is computed between the measured point and the "distorted" projected point.

In the computer vision community, the seminal work by Triggs et al. (1999) presented a survey of the theory and methods of photogrammetric bundle adjustment, addressed to potential developers in their community.

Davison (2015) recognised that, before the beginning of 2000, when his pivotal paper on SLAM with a single camera was published (Davison, 2003), the mobile robotics community had almost completely lost interest in pure vision-based navigation approaches and, at the same time, the computer vision community had almost abandoned researching on real-time and robotics applications.

In the last ten years, photogrammetry has tried to interact with computer vision; as a result, many tools developed in computer vision have been adopted in photogrammetry (Granshaw & Fraser, 2015), and many terms derived from computer vision have been included in the third edition of the photogrammetric terminology (Granshaw, 2016). The book by (Förstner & Wrobel, 2016), entitled "Photogrammetric Computer Vision", addresses concepts and methods developed in the computer vision community with a focus on the main application areas of photogrammetry, specifically mapping and image-based metrology.

Still, software solutions and algorithms coming from different communities, although sharing a common theoretical basis, may entail different formulations and implementations (Fig. 1; Drap & Lefèvre, 2016; Borlin et al., 2019).

1.2 The ISPRS ECB Initiatives

The International Society for Photogrammetry and Remote Sensing (ISPRS) has introduced the Education and Capacity Buildings initiatives in 2017. The aim of the funding program is to consolidate the Society's international position in education training and networking in the domains relevant to the Society, i.e. of photogrammetry, remote sensing and spatial information sciences. The funded projects, whose expected duration is 132 months, are expected to focus on topics such as: education and learning programs, teaching and training materials and resources, e-learning techniques and platforms and online databases (GEOInformatics, 2019). As the ultimate outcome, the supported initiatives should enhance the visibility of the ISPRS, broadly promote its mission, widen its public and network of collaborations.

2. OBJECTIVES AND METHODOLOGY

Built upon the last years of efforts of interaction between the three communities, the LightCam project wants to offer two platforms to promote a common itinerary towards the convergence of a unified theoretical basis and a coherent vocabulary. The developed tools, envisaged as support to education and teaching, will be mainly addressed to students, beginners, and non-experts.
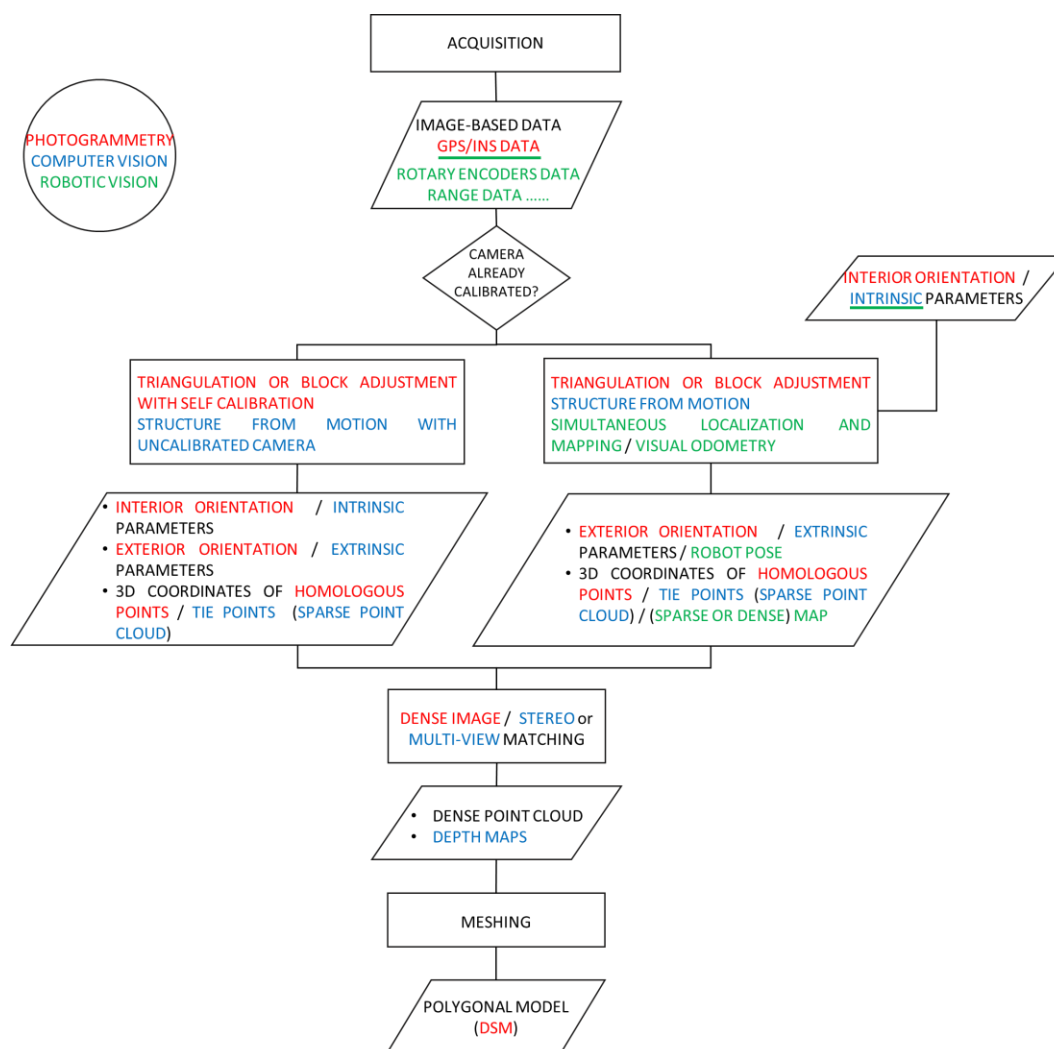
Figure 2: Typical vision-based 3D measurement and modelling pipeline.

PHOTOGRAMMETRY	COMPUTER / ROBOTIC VISION
Image matching	Stereo correspondence problem
Residuals of image coordinates	Re-projection error
Interior orientation parameters	Intrinsic parameters
Exterior orientation parameters	Extrinsic parameters / Robot pose
Image coordinates	Camera coordinates
Ground or object space coordinates	World coordinates
Dense image matching	Multi-view stereo / densification

Table 1: Examples of dictionary entries (adapted from Granshaw & Fraser, 2015).

2.1 The LightCam website

The first tool will be implemented as a knowledge base website, which is currently under construction (www.lightcam.fbk.eu), with the following envisaged sections:

- Birth and history of photogrammetry, computer and robotic vision, with an emphasis on their contact points
- Image formation with basics of optics and photography
- Typical vision-based 3D measurement and modelling pipeline (Fig. 2), with emphasis on contact points and differences between the three communities
- Dictionary for common concepts (Tab. 1)
- List of relevant software solutions, websites, available courses resources and seminal articles from the three communities.

2.2 The LightCam code repository

The repository will contain pieces of code to transform from one common mathematical representation to another (e.g., different angle definitions, from one coordinate system to another, etc.). It will also show examples for the conversion between different software applications adopting different implementation strategies (e.g., Bingo, Photomodeler, Australis, OpenCV, etc.) and format files (*.nvm, *.out, etc.).

The codes will be written in the most common programming language (e.g., Matlab, Python, C++, Java). An example is provided in Fig. 3.

```
% In the followings, how to convert from interior and exterior orientation
% parameters from
% photogrammetric notation to the computer vision projective matrix
% (to be used for example as input for CMVS/PMV2, https://www.di.ens.fr/pmvs/)
% The detailed explanation will be provided in the LigthCAM knowledge base website

% PREMISE:
% The images must be idealized or undistorted, i.e. the lens distortion
% should be removed. The principal point should not be moved in the image
% center.

% REFERENCE SYSTEM:
% The photogrammetric image reference system is centered in the center of the
% image.
% The x-axis points to the right and the y-axis points to the up.
% The PMVS2 image reference system is centered in the top left corner of an image
% (center of the pixel at the top-left corner of the image).
% The x-axis points to the right and the y-axis points to the bottom

% INPUT VARIABLES:
% c = principal distance in mm
% PPx = x coordinate of the principal point in mm in the photogrammetric image
% reference system
% PPy = y coordinate of the principal point in mm in the photogrammetric image
% reference system
% px_size_x = pixel size or pixel pitch in mm in the x axis
% px_size_y = pixel size or pixel pitch in mm in the y axis
% Sensor_W = camera sensor width in mm
% Sensor_H = camera sensor height in mm
% T = column vector containing the 3D coordinates of the image perspective
% center from photogrammetry
% R = R_3D2D image rotation matrix from photogrammetry (see definition below)

PPx_PMVS2 = ((PPx + (Sensor_W/2)) / px_size_x) + (px_size_x/2); % x coordinate of the principal point in
pixel in the PMVS2 image reference system
PPy_PMVS2 = (((Sensor_H/2) - PPy) / px_size_y) + (px_size_y/2); % y coordinate of the principal point in
pixel in the PMVS2 image reference system

cx = px_size_x * c; % principal distance in pixel in the x axis
cy = px_size_y * c; % principal distance in pixel in the y axis

% K = camera matrix
K = [cx, 0, PPx_PMVS2;...
     0,  cy, PPy_PMVS2;...
     0,  0,   1];

 % P = projection matrix
 P = K * ([1 0 0; 0 -1 0; 0 0 -1]*([R,-R*T]));

% Rotation matrix from object space (i.e., global frame) to image space (i.e., local frame, as defined in
% Mikhail et al. (2001)

R_3D2D = [ cos(phi)*cos(k), cos(omega)*sin(k)+sin(omega)*sin(phi)*cos(k), ...  sin(omega)*sin(k)-
cos(omega)*sin(phi)*cos(k);...
-cos(phi)*sin(k),cos(omega)*cos(k)-sin(omega)*sin(phi)*sin(k), ...
sin(omega)*cos(k)+cos(omega)*sin(phi)*sin(k);...
sin(phi), -sin(omega)*cos(phi), cos(omega)*cos(phi)];

% Please note, that in Luhmann et al. (2013) the matrix R denotes the rotation matrix from image to
% object space, corresponding to the inverse of R_3D2D

R_2D3D = R_3D2D';

% R_2D3D = [  cos(phi)*cos(k)        -cos(phi)*sin(k)      sin(phi);...
%           cos(omega)*sin(k)+sin(omega)*sin(phi)*cos(k)   cos(omega)*cos(k)-sin(omega)*sin(phi)*sin(k)    %
-sin(omega)*cos(phi);...
%           sin(omega)*sin(k)-cos(omega)*sin(phi)*cos(k)   sin(omega)*cos(k)+cos(omega)*sin(phi)*sin(k)    %
cos(omega)*cos(phi)];

% omega, phi and k are the image exterior orientation angles in radians. If the angles are provided in %
degree, the functions to use are cosd and sind
```

Figure 3: Example of exchange code, which will be included in the LightCam code repository.

3. EXPECTED OUTCOMES

LightCam specifically addresses the aim of 'generation of computer aided teaching and learning material (curriculum development, text books, etc.) in the photogrammetry, remote sensing and spatial information sciences' by providing cross community educational resources aiming at a stronger interconnection among photogrammetry, computer vision and robotic vision. LightCam will set an easy and lightweight internet platform covering from nomenclature to principles of vision-based techniques, as well as main algorithms from most popular open source framework or commercial software packages developed in the three different communities.

The web resources will be targeted to students, young researchers, new users and practitioners, fostering an easier

dialogue and cooperation among the different communities. The benefits of the developed tools are expected to be even broader: Users coming from other fields (e.g. heritage, energy, water, healthcare, autonomous systems, transportation, traditional/robotic manufacturing, etc.) who employ vision-based methods will profit from the LightCam outcomes. This will further promote the ISPRS itself and its activities, pushing the liaisons with other communities beyond the current status quo.

In line with the objectives of ISPRS, the LightCam ambition is to provide and disseminate photogrammetric knowledge to a more and more vast audience, highlighting the common interests with computer vision and robotics, but at the same time affirming the specificities of the three communities.

The implementation simplicity of the LightCam web resources will facilitate its diffusion to a wide public, but, at the same time, the rigorousness of the shared information will be assured to guarantee high quality standards of the learning material.

REFERENCES

Albertz, J., 2002. Albrecht Meydenbauer-Pioneer of photogrammetric documentation of the cultural heritage. *International Archives of Photogrammetry Remote Sensing and Spatial Information Sciences*, 34(5/C7), pp.19-25.

APRS, American Society for Photogrammetry and Remote Sensing, 1952. *Manual of Photogrammetry*, Bethesda: ASP&RS.

Börlin, N., Murtiyoso, A., Grussenmeyer, P., Menna, F. and Nocerino, E., 2019. Flexible Photogrammetric Computations Using Modular Bundle Adjustment: The Chain Rule and the Collinearity Equations. *Photogrammetric Engineering & Remote Sensing*, 85(5), pp.361-368.

Davison, A. J, 2003. Real-Time Simultaneous Localisation and Mapping with a Single Camera. *Proc. Ninth Int. Conf. on Computer Vision*.

Davison, A. J, 2015. *15 Years of Visual SLAM. Presentation at The Future of Real-Time SLAM*: 18th December 2015 (ICCV Workshop), http://wp.doc.ic.ac.uk/thefutureofslam/wp-content/uploads/sites/93/2015/12/slides_ajd.pdf

Drap, P. and Lefèvre, J., 2016. An exact formula for calculating inverse radial lens distortions. *Sensors*, 16(6), p.807.

Förstner, W. and Wrobel, B.P., 2016. *Photogrammetric computer vision*. Springer International Publishing Switzerland.

Granshaw, S.I. and Fraser, C.S., 2015. Computer vision and photogrammetry: Interaction or introspection?. *The Photogrammetric Record*, 30(149), pp.3-7.

Granshaw, S.I., 2016. Photogrammetric terminology. *The Photogrammetric Record*, 31(154), pp.210-252.

Jensen, J. R. 2007. *Remote Sensing of the Environment: An Earth Resource Perspective*. Pearson Prentice Hall.

GEOInformatics, 2019. ISPRS Educational and Capacity Building Initiatives. https://geoinformatics.com/isprs-educational-and-capacity-building-initiatives

Gruen, A., 1992. Recent advances of photogrammetry in robot vision. *ISPRS journal of photogrammetry and remote sensing*, 47(4), pp.307-323.

Gruen, A. and Huang, T.S. eds., 2001. *Calibration and orientation of cameras in computer vision* (Vol. 34). Springer Science & Business Media.

Huang, T., 1996. *Computer Vision: Evolution And Promise*. 19th CERN School of Computing. Geneva: CERN. Vandoni, Carlo, E, ed.

Luhmann, T., Robson, S., Kyle, S. and Boehm, J., 2013. *Close-range photogrammetry and 3D imaging*. Walter de Gruyter.

Mikhail, E.M., Bethel, J.S. and McGlone, J.C., 2001. *Introduction to modern photogrammetry*. New York.

Triggs, B., McLauchlan, P.F., Hartley, R.I. and Fitzgibbon, A.W., 1999, September. Bundle adjustment—a modern synthesis. *In International workshop on vision algorithms* (pp. 298-372). Springer, Berlin, Heidelberg.