SOUTHERN JOURNAL OF SCIENCES

ESTABLISHED IN 1993

Original research paper

INTERACTIVE 3D RECONSTRUCTION AND DLT CAMERA CALIBRATION: A MANUAL REGISTRATION APPROACH

LIMA DA SILVA, Isaac Newton[1*]; FERNANDES, Éder[2]; GONZAGA, Henrique[3];

[1] Seção de Engenharia Legal, Instituto Geral de Perícias do Rio Grande do Sul; Escola Politécnica da Pontifícia Universidade Católica do Rio Grande do Sul. Brasil.

[2] Universidade do Vale do Rio dos Sinos. Brasil.

[3] Serviço Nacional de Aprendizagem Industrial - SENAI. Brasil.

** Autor correspondente*
e-mail: inls.perito@gmail.com

Received 29 August 2024; received in revised form 10 December 2024; accepted 19 December 2024

ABSTRACT

Background: This paper presents a straightforward and intuitive method for interactive 3D reconstruction and Direct Linear Transformation (DLT) camera calibration using a single image of a structured scene with known object dimensions. The method relies on manual registration of pairs of points on both the image and the terrain, allowing for precise alignment and calibration. **Aim:** By utilizing this method, users can easily reconstruct 3D scenes and calibrate cameras without the need for complex algorithms or extensive computational resources. Our approach offers a user-friendly solution for 3D reconstruction and camera calibration, making it accessible to a wider audience and applicable in a range of fields such as computer vision, augmented reality, and virtual reality. **Methods:** This work primarily focuses on the determination of the projection matrix, which plays a crucial role in mapping 3D points onto a 2D image plane. The projection matrix encapsulates both the intrinsic parameters of the camera (such as focal length and optical center) and the extrinsic parameters (such as camera position and orientation in the world coordinate system). By accurately determining the projection matrix, we can effectively project 3D points onto the 2D image plane, enabling tasks like 3D reconstruction, camera localization, and augmented reality applications. **Results:** We present experimental results obtained from testing the method on an image of a known object, demonstrating its effectiveness and accuracy in producing realistic 3D reconstructions. **Discussion:** The method's reliance on manual registration of point pairs allows for precise alignment and calibration without the need for complex algorithms or extensive computational resources. This user-friendly approach makes 3D reconstruction and camera calibration accessible to a wider audience and applicable in various fields. **Conclusions:** Overall, our approach offers a practical and accessible solution for 3D reconstruction and camera calibration, expanding the potential applications in computer vision, augmented reality, and virtual reality.

Keywords: *3D reconstruction; Camera calibration; Direct Linear Transformation (DLT); Single image.*

1. INTRODUCTION

Camera calibration is a fundamental process in computer vision and photogrammetry that aims to estimate the intrinsic and extrinsic parameters of a camera system. Among the various calibration techniques, Direct Linear Transformation (DLT) stands out as a robust and widely used method for accurately calibrating cameras. In this article, we delve into the intricacies of DLT camera calibration, exploring its principles, applications, advantages, and limitations.

In recent years, the field of computer vision and image processing has seen significant advancements, particularly in techniques related to 3D reconstruction and camera calibration (Hartley & Zisserman, 2003). Recent work by Silva et al. (2024) has demonstrated practical

DOI: 10.48141/2764-5959.01.v32.n38.2024_ISAAC_pgs_1_10.pdf

applications of these techniques in accident reconstruction and uncertainty analysis. These advancements have enabled researchers and practitioners to create realistic 3D models from 2D images and accurately calibrate cameras for various applications such as virtual reality, augmented reality, robotics, and more.

One of the key challenges in 3D reconstruction and camera calibration is the need for accurate and efficient methods that can handle complex scenes while maintaining simplicity and user-friendliness. Traditional approaches (Tsai, 1986; Tsai, 1987; Horn, 2004; Zhang, 2000) often require complex algorithms, extensive computational resources, and specialized equipment, making them less accessible to non-experts and limiting their practicality in certain scenarios.

To address these challenges, this paper introduces a streamlined and intuitive approach for interactive 3D reconstruction and camera calibration. Unlike traditional methods, our approach is based on the manual registration of pairs of points on both the image and the terrain. This manual registration allows users to map the image with the real-world scene accurately, providing a foundation for precise 3D reconstruction and camera calibration.

The core concept of our method revolves around leveraging known object dimensions within the scene. By incorporating object dimensions, users can input accurate scale information, further enhancing the accuracy of the 3D reconstruction and camera calibration process. This aspect is particularly beneficial when dealing with structured scenes where known object dimensions are available.

Furthermore, our approach eliminates the need for complex algorithms and extensive computational resources, making it accessible to a wider audience, including non-experts and researchers in various fields. The simplicity of manual registration combined with the use of object dimensions results in an intuitive and efficient workflow for generating realistic 3D models and calibrating cameras from a single image.

In this paper, we present the details of the DLT method, including the manual registration process, utilization of object dimensions, and the experimental results obtained from testing the method on a variety of test images. We demonstrate the effectiveness and accuracy of our approach in producing high-quality 3D reconstructions and accurately calibrating cameras, highlighting its potential for practical applications in computer vision and related domains.

Our approach for interactive 3D reconstruction and camera calibration through manual registration has wide-ranging applications, including but not limited to forensics, sports, and healthcare domains.

In forensics, the accurate estimation of the height and speed of objects or individuals captured in images is crucial for crime scene reconstruction and analysis. Our method enables forensic experts to create precise 3D models of crime scenes and accurately calibrate cameras to determine the height and speed of objects or persons involved in the incident. This information can be instrumental in reconstructing the sequence of events and aiding in forensic investigations.

In sports and health, our approach finds applications in postural estimation and analysis. By reconstructing 3D models from single images and calibrating cameras, our method can accurately estimate the posture and movements of athletes or patients during various activities. This information is valuable in sports performance analysis, injury prevention, rehabilitation, and ergonomics assessment in healthcare settings.

Furthermore, our method's simplicity and efficiency make it accessible to a wide range of users, including law enforcement agencies, sports coaches, healthcare professionals, and researchers. The intuitive manual registration process combined with the utilization of object dimensions provides a user-friendly workflow for generating detailed 3D reconstructions and precise camera calibration, enhancing the capabilities of applications in forensics, sports, and healthcare domains.

2. MATERIALS AND METHODS

2.1 Materials

In this study, the DLT camera calibration was performed using a set of corresponding points in the image and the scene. These points were used to solve the system of equations associated with the DLT method. The parameters of the camera were determined using the least squares optimization technique, which minimizes the residual errors between the observed image points and their reprojected counterparts in the calibration model.

The least squares method was implemented in custom software capable of solving the overdetermined linear system efficiently. This approach ensures accurate computation of the DLT coefficients, providing a robust calibration of the camera.

2.2 Methods

2.2.1 Direct Linear Transformation (DLT) Algorithm

DLT camera calibration is based on the Direct Linear Transformation algorithm, which is used to determine the transformation matrix between 3D object points and their corresponding 2D image coordinates captured by a camera (Abdel-Aziz & Karara, 1971) The primary goal of DLT calibration is to accurately map the relationship between the real-world coordinates of objects and their projections onto the camera's image plane.

2.2.2 Key Principles of DLT Calibration

The DLT calibration process involves several key principles:

1. **Camera Model**: DLT assumes a pinhole camera model, where light rays from the scene pass through a single point (the camera's optical centre) to form an image on the camera sensor.
2. **Calibration Object**: A known calibration object with precise 3D coordinates is essential for DLT calibration. Common calibration objects include checkerboard patterns, calibration grids, and specially designed targets with known marker positions.
3. **Parameter Estimation**: By analyzing the image correspondences and the known 3D coordinates of calibration points, DLT calculates the camera's intrinsic parameters (focal length, principal point, and, for some purposes, distortion coefficients) and extrinsic parameters (rotation and translation of the camera relative to the world coordinate system).

DLT camera calibration is widely used across various domains due to its versatility and accuracy. In augmented reality, precise camera calibration is essential for seamlessly aligning virtual objects with the real-world scene, enhancing user experiences. In robotics, DLT calibration plays a vital role in enabling robots to perceive and navigate their environment accurately by establishing the relationship between camera images and the positions of 3D objects. Industrial metrology relies on DLT calibration to ensure precise measurements and inspections in manufacturing and quality control processes, leveraging vision-based systems for enhanced accuracy. Moreover, DLT calibration is extensively utilized in sports analysis, allowing analysts to track player movements, ball trajectories, and game tactics effectively from camera footage, leading to insightful performance evaluations and strategic improvements.

DLT camera calibration offers numerous advantages that make it a preferred choice in various applications. Firstly, its accuracy is commendable, especially when coupled with precise calibration data, allowing for high precision in estimating camera parameters and 3D object positions. Secondly, DLT calibration exhibits flexibility, as it can be applied to a wide range of cameras, from standard digital cameras to more specialized imaging systems, making it versatile across different setups. Additionally, DLT is robust against moderate distortions and noise in the calibration data, ensuring reliable performance in real-world scenarios. Lastly, DLT calibration is cost-effective compared to some complex methods, as it requires minimal specialized equipment and is relatively straightforward to implement, making it an accessible and practical solution for camera calibration needs.

DLT camera calibration, despite its numerous advantages, comes with certain limitations that should be considered. Firstly, it is sensitive to errors in calibration object placement and image correspondences, which can impact the accuracy of the calibration results. Secondly, while DLT can handle radial and tangential distortions, it may struggle with more complex lens distortions encountered in certain camera systems, potentially affecting the quality of the calibration (Devy et al., 1997; Weng, 1992; Barone et al., 2020). Lastly, DLT calibration primarily relies on 2D-3D point correspondences, which may limit its ability to accurately capture depth information, particularly in scenes with significant depth variations, leading to potential challenges in accurately reconstructing the scene's 3D structure.

2.2.3 Mathematical Foundation

The Direct Linear Transformation (DLT) algorithm, introduced by Abdel-Aziz and Karara (1971), stands out as the most widely used camera calibration method in motion capture applications. Notably, DLT demonstrates impressive precision and accuracy with a relatively simple mathematical model compared to other algorithms. Its solution method consistently yields satisfactory values, unlike many alternative calibration algorithms. This can be attributed to DLT's flexible approach, which imposes minimal restrictions on the scene and calibration object shape, allowing for adaptability to varying scene conditions. In contrast, other calibration models often encounter convergence issues with the numerical methods used to calculate calibration coefficients if these restrictions are not strictly adhered to.

The fundamental premise of the DLT model is based on the idea that a camera's image can be mathematically expressed as a geometric transformation between the world space and the image plane. This transformation is essentially the projection of a point from the world space onto the image plane. Figure 1 visually illustrates this projection, depicting two reference systems: the world space with coordinates X, Y, and Z and the image plane with coordinates (u, v). The camera's projection center is closely linked to its focal length and is also referred to as the "center of the camera." The optical system projects a point in world space onto the image plane, passing through the projection center. This collinear relationship between points and the camera center is known as the "collinearity condition," which serves as the foundational principle of the DLT method.

The two collinearity equations serve as the foundation of digital photogrammetry, as they define the connection between the exterior orientation parameters, the photographic coordinates of a point, and the three-dimensional coordinates of that point within the reference system of either the terrain or an object in space. The principle of collinearity asserts that, during the process of photograph capture, the object point P, the projection center O, and the image point p are aligned in a straight line.

$$u = u0 - \frac{c*r11(x-x0)+r21(y-y0)+r31(z-z0)}{r13(x-x0)+r23(y-y0)+r33(z-z0)}$$

(Eq. 1)

$$v = v0 - \frac{c*r12(x-x0)+r22(y-y0)+r32(z-z0)}{r13(x-x0)+r23(y-y0)+r33(z-z0)}$$

(Eq. 2)

“u” represents the x-coordinate of a point in the image;

“u0” is the x-coordinate of the camera perspective center in the image;

“v” denotes the y-coordinate of a point in the image;

“v0” is the y-coordinate of the camera perspective center in the image;

“c” represents a scalar value.

“rij’ represents the elements of the rotation matrix R.

$$R = \begin{bmatrix} r11r12r13 \\ r21r22r23 \\ r31r32r33 \end{bmatrix} \qquad \text{(Eq. 3)}$$

The values of u, v, u0, and v0 used in equations 1 and 2 can be expressed in any unit. It's important to note that the two conversion factors may differ from each other, as the pixel may not be a square figure. By rearranging the equations, we derive the fundamental equations of the DLT method as illustrated in equations 4 and 5

$$u = \frac{L1*x+L2*y+L3*z+L4}{L9*x+L10*y+L11*z+1} \qquad \text{(Eq. 4)}$$

$$v = \frac{L5*x+L6*y+L7*z+L8}{L9*x+L10*y+L11*z+1} \qquad \text{(Eq. 5)}$$

The eleven constants (L1,...,L11) are commonly referred to in technical literature as "DLT coefficients." The primary goal of DLT calibration is to precisely ascertain the values of these eleven coefficients, which are subsequently utilized in the three-dimensional reconstruction process. These coefficients are directly associated with the intrinsic and extrinsic parameters of a camera, but that will be addressed in future work.

Equations 4 and 5 could be represented in matrix form as:

$$\begin{bmatrix} xyz10000-ux-uy-uz \\ 0000xyz1-vx-vy-vz \end{bmatrix} \begin{bmatrix} L1 \\ L2 \\ L3 \\ L4 \\ L5 \\ L6 \\ L7 \\ L8 \\ L9 \\ L10 \\ L11 \end{bmatrix} = \begin{bmatrix} u \\ v \end{bmatrix}$$

(Eq. 6)

To calibrate the camera, one simply needs to solve the linear system defined in equation 6 to determine the eleven calibration coefficients (L1,..., L11). However, the number of unknowns exceeds the number of equations because only one control point was considered, which is insufficient for calibration purposes. Equation 6 highlights that a single point (u,v) yields two equations. Therefore, a minimum of six calibration points is necessary to generate at least eleven equations, enabling the system to be solvable since there are eleven unknowns.

When six or more control points are used, the system defined in 6 becomes overdetermined, allowing for multiple possible solutions. While SVD typically yields satisfactory solutions, further optimization of the system solution can be achieved by minimizing residual errors through the least squares technique. In this study, the chosen approach for solving the system was the least squares method (Polidório et al., 1998).

3. RESULTS AND DISCUSSION

3.1. Results

As an example of camera calibration using DLT, a frame containing the image of a calibration object is presented in Figure 2. In technical literature, the frame or image used for camera calibration is referred to as a "keyframe." Table 1 displays the coordinates of seven selected control points extracted from Figure 2. The world coordinates of these points were directly measured on the calibration object, while the image coordinates were estimated by clicking on the center of each point in the image using a mouse, such as in (Pinheiro, 2008).

In this example, the data includes the coordinates of control points in both the world space (x, y, z) and the image plane (u, v). The data is provided in Table 1.

To calibrate the camera using this data, we will follow these steps:

1. Input the world coordinates (x, y, z) and image coordinates (u, v) into the calibration algorithm.
2. Utilize the least squares method to solve the overdetermined system of equations and determine the calibration coefficients.
3. Apply the calibration coefficients to accurately map points from the image plane to the world space.

Once the camera is calibrated, it can be used for tasks such as 3D reconstruction, object tracking, and augmented reality applications with improved accuracy and reliability. The Projection matrix, obtained from the DLT coefficients, is presented in Table 2.

The mathematical formulation required for the coordinate reconstruction process is derived from the equivalent equation represented in matrix form.

$$\begin{bmatrix} L1-uL9L2-uL10L3-uL11 \\ L5-vL9L6-vL10\ L7-vLu \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} u-L4 \\ v-L8 \end{bmatrix}$$

(Eq. 7)

The linear system presented in Equation 7 involves three unknowns (x, y, z) and two equations. However, it's widely recognized that at least three equations are necessary for solving such a system. A practical resolution to this issue involves providing information about some of the unknowns. In other words, if one of the coordinates is known, the remaining two can be determined accordingly.

3.2. Discussions

After performing camera calibration using the provided object data, we obtained the calibration coefficients necessary to accurately map points from the image plane to the world space. The results obtained from the calibration process are crucial for ensuring accurate measurements and reliable imaging in various applications.

One key aspect to discuss is the accuracy of the calibration coefficients. A higher accuracy in the calibration coefficients implies a better alignment between the real-world scene and the captured images, leading to improved precision in subsequent tasks like 3D reconstruction and object tracking.

Another point of discussion is the residual error or discrepancy between the observed and calculated coordinates after applying the calibration coefficients. While calibration algorithms strive to minimize these errors, it's essential to evaluate the residual errors to assess the overall quality of the calibration process. Lower residual errors indicate a more effective calibration, whereas higher errors may suggest inaccuracies in the calibration model or data.

Calibrated cameras are vital in fields such as computer vision, robotics, and augmented reality, where accurate spatial mapping and measurements are critical. The calibration results enable precise object localization, motion tracking, and scene reconstruction, enhancing the capabilities and reliability of these applications.

It's important to note that while this work emphasizes the determination of the projection matrix, future work will delve into refining and estimating the intrinsic and extrinsic parameters of the camera system. The intrinsic parameters pertain to the internal characteristics of the camera, while the extrinsic parameters relate to its external positioning and orientation. By addressing these parameters, we aim to enhance the overall accuracy and performance of the camera system, enabling more robust and reliable applications in computer vision, robotics, and related fields.

4. CONCLUSIONS:

DLT camera calibration is a powerful technique for accurately estimating camera parameters and mapping 3D scenes to 2D images. Its principles, applications, advantages, and limitations highlight its importance in computer vision, robotics, metrology, and other domains where precise spatial understanding is essential. By understanding DLT calibration, researchers, engineers, and practitioners can leverage its capabilities to develop advanced imaging systems and applications that rely on accurate camera perception.

In conclusion, camera calibration plays a crucial role in ensuring the accuracy and reliability of imaging systems in various applications. Through the calibration process, we obtained calibration coefficients that enabled us to correct map points from the image plane to the world space. The results of our calibration process are essential for tasks such as 3D reconstruction, object tracking, and augmented reality applications, where precise spatial mapping and measurements are critical.

The discussion over the calibration results highlighted the importance of accuracy in the calibration coefficients and the evaluation of residual errors. A higher accuracy in the coefficients and lower residual errors indicate a more effective calibration process, leading to improved precision in subsequent tasks. Additionally, the practical implications of the calibration results were discussed, emphasizing the significance of calibrated cameras in fields such as computer vision, robotics, and augmented reality.

Overall, camera calibration is a fundamental step in ensuring the quality and reliability of imaging systems, allowing for enhanced capabilities and improved performance in various applications. The results obtained from the calibration process validate the effectiveness of the calibration method used and provide a solid foundation for achieving accurate spatial mapping and measurements in real-world scenarios.

5. DECLARATIONS

5.1. Study Limitations

No limitations were known at the time of the study.

5.2. Funding source

The authors funded this research.

5.3. Competing interests

The authors declare that there are no conflicts of interest regarding this research.

5.4. Open Access

6. REFERENCES:

1. Hartley, R., Zisserman, A. Multiple View Geometry in Computer Vision. Cambridge University Press, 2003.
2. Tsai, Roger Y. *An Efficient and Accurate Camera Calibration Technique for 3D Machine Vision.* Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Miami Beach, FL, pp. 364-374, 1986.
3. Tsai, Roger Y. A Versatile Camera Calibration Technique for High Accuracy 3D Machine Vision Metrology Using Off-the-Shelf TV Cameras and Lenses. IEEE Journal of Robotics and Automation, Vol. RA–3, No. 4, August 1987, pp. 323–344.
4. Horn, B. K. P. Tsai's camera calibration method revisited. https://ocw.mit.edu/courses/6-801-machine-vision-fall-2004/5d74a1dd1b06a858ab5c2643684ef7ab_tsaiexplain.pdf.
5. Z. Zhang. *A flexible new technique for camera calibration.* IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 22, No. 11, pp. 1330-1334, 2000.
6. ABDEL-AZIZ, I. Y. & KARARA, H. M. Direct Linear transformation from comparator coordinates into object coordinates in close-range photogrammetry. Proceedings of the Sjoommisnot ox ChscRonv Photopmmartg, p. 1-18, 1971.
7. M. Devy, V. Garric, J.J. Orteu. Camera calibration from multiple views of a 2D object, using a global non-linear minimization method. Conference: Intelligent Robots and Systems, 1997. IROS '97., Proceedings of the 1997. DOI: 10.1109/IROS.1997.656569.
8. Weng, J. Camera Calibration with Distortion Models and Accuracy Evaluation. IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL. 14, NO. 10, OCTOBER 1992
9. Francesco Barone, Marco Marrazzo, and Claudio J. Oton. Camera Calibration withWeighted Direct Linear Transformation and Anisotropic Uncertainties of Image Control Points. Sensors 2020, 20, 1175; doi:10.3390/s20041175.
10. Airton Marco Polidório, Flavio Bortolozzi e Maurício Figueiredo. Calibração geométrica de câmeras, aplicando o método dos mínimos quadrados. Acta Scientiarum 20(4):495-503, 1998. ISSN 1415-6814.
11. Pinheiro, A. P. Desenvolvimento de um sistema de captura e análise de movimentos baseado em técnicas de visão computacional. 2008. 109 p. Thesis (Doctorate in Sciences) - 2010. 80 f. Thesis (Doctorate in Sciences) – Faculty of Electrical Engineering, Federal University of Uberlândia.
12. Gonzaga, H., Fernandes, É., & Silva, I. N. L. da. (2024). ANALYSIS OF UNCERTAINTY IN THE DETERMINATION OF THE CRITICAL YAW RADIUS USING MONTE CARLO SIMULATION. In PERIÓDICO TCHÊ QUÍMICA (Vol. 21, Issue 47, pp. 61–68). Dr. D. Scientific Consulting. https://doi.org/10.52571/ptq.v21.n47.2024_05_isaac_pgs_61_68.pdf
13. Silva, I. N. L. da, Fernandes, É., & Gonzaga, H. (2024). ESTIMATION OF VEHICLE YAW SPEED USING PHOTOGRAMMETRY OF TIRE TRACKS AT TRAFFIC ACCIDENT SCENES. In PERIÓDICO TCHÊ QUÍMICA (Vol. 21, Issue 48, pp. 32–41). Dr. D. Scientific Consulting. https://doi.org/10.52571/ptq.v21.n48.2024_01_silva_pgs_32_41.pdf

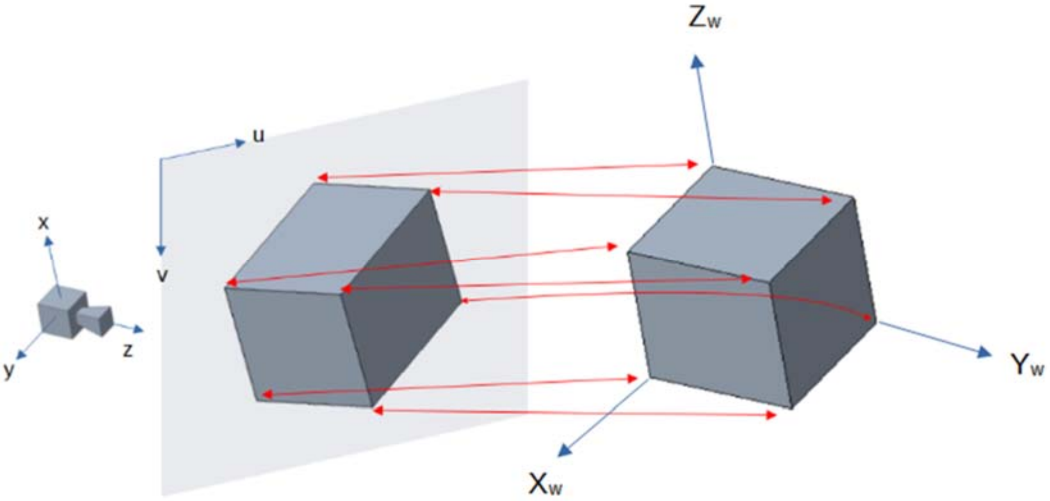
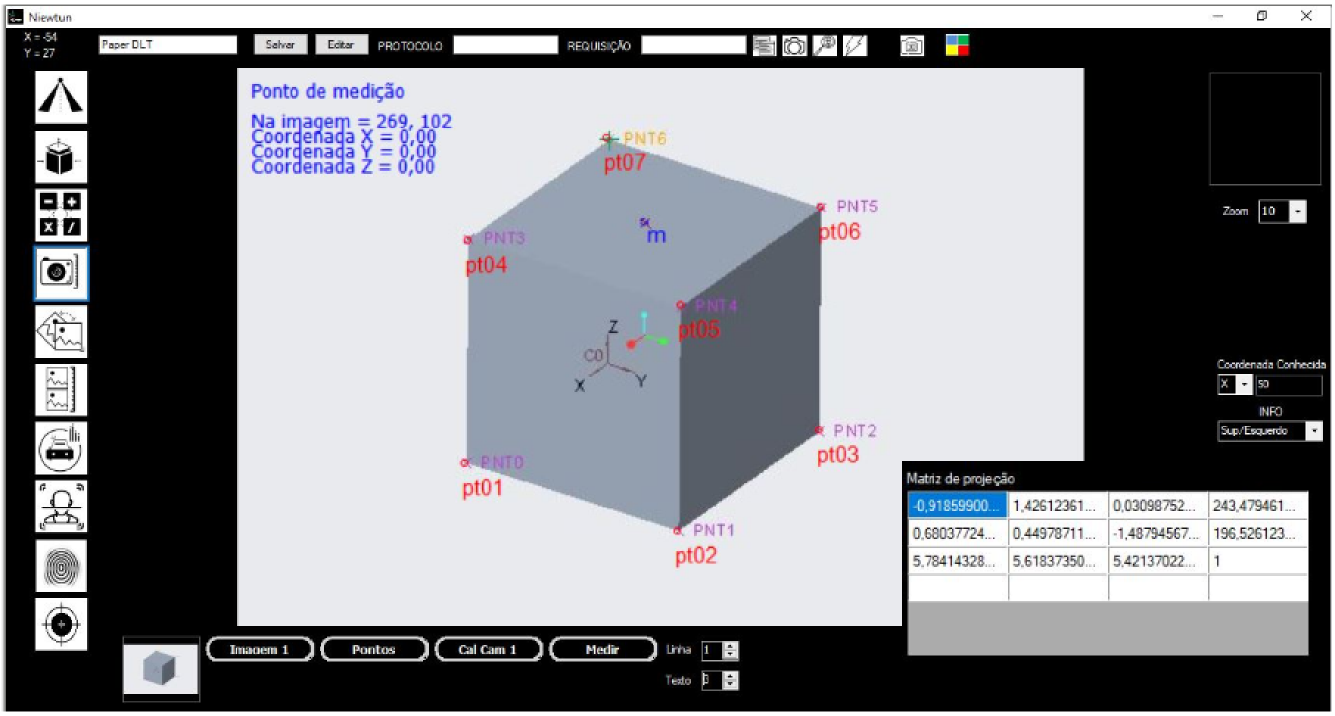


Figure 1. Perspective projection of points onto the image plane.



X = -54
Y = 27

Paper DLT

Ponto de medição

Na imagem = 269, 102
Coordenada X = 0,00
Coordenada Y = 0,00
Coordenada Z = 0,00

Zoom 10

Coordenada Conhecida
X 50

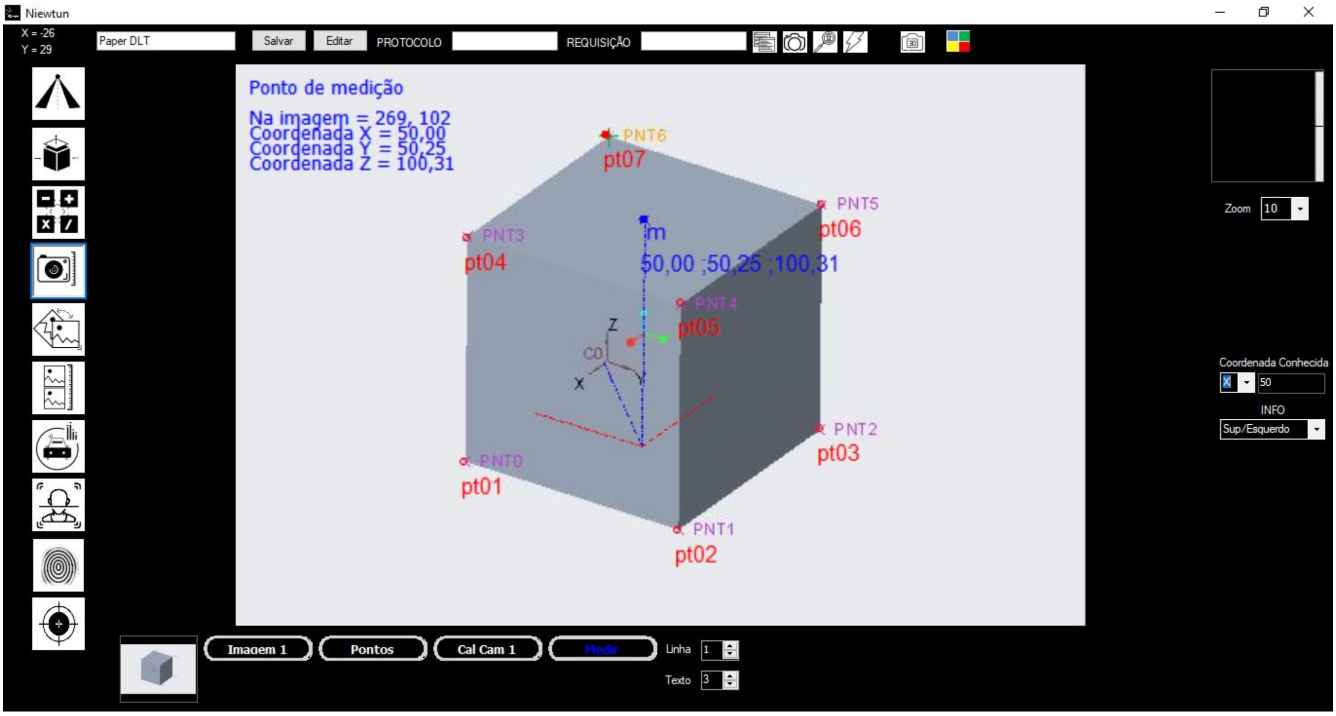
INFO
Sup/Esquerdo

Matriz de projeção

-0,91859900...	1,42612361...	0,03098752...	243,479461...
0,68037724...	0,44978711...	-1,48794567...	196,526123...
5,78414328...	5,61837350...	5,42137022...	1

Linha 1
Texto 8

Figure 2. The keyframe for the camera calibration process.



Paper DLT

Ponto de medição

Na imagem = 269, 102
Coordenada X = 50,00
Coordenada Y = 50,25
Coordenada Z = 100,31

m
50,00 ;50,25 ;100,31

Zoom 10

Coordenada Conhecida X 50

INFO Sup/Esquerdo

Linha 1

Texto 3

Figure 3. A measuring point showing the result of the centre point of the top face of the object at the calculated coordinates of (50, 50, 100).

Table 1. The control points.

Point	x (world)	y (world)	z (world)	u (image)	v (image)
PT01	100	0	0	151	263
PT02	100	100	0	292	308
PT03	0	100	0	386	241
PT04	100	0	100	153	115
PT05	100	100	100	294	158
PT06	0	100	100	387	93
PT07	0	0	100	245	47

Table 2. The calculated Projection Matrix.

Projection Matrix			
-0.91859901	1.42612362	0.03098753	243.47946167
0.68037724	0.44978711	-1.48794568	196.52612305
0.00005784	0.00000562	0.00005421	1.00000000