Journal of Information Systems Engineering and Management
2025, 10(16s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Research of Methods for Determining Key Points for 3D Modeling of Amber Samples

Yurii Podchashynskyi[1], Andrii Ryzhuk[2], Oksana Luhovykh[3], Chepiuk Larina[4]
[1]*Department of Information and Measurement Technologies, Zhytomyr State Technological University, Ukraine*
[2]*Department of Information and Measurement Technologies, Zhytomyr State Technological University, Ukraine*
[3]*Department of Information and Measurement Technologies, Zhytomyr State Technological University, Ukraine*
[4]*Department of Information and Measurement Technologies Zhytomyr State Technological University, Ukraine*
[1]*jup@ztu.edu.ua;* [2]*kmivt_rav@ztu.edu.ua;* [3]*auts_ksy@ztu.edu.ua;* [4]*chepyk@ztu.edu.ua*

ARTICLE INFO

Received: 01 Dec 2024

Revised: 25 Jan 2025

Accepted: 08 Feb 2025

ABSTRACT

3D visualization and modeling of amber samples is important for assessing shape, intersperses, color, and volume. The obtained results allow for amber evaluation and calculation of optimal stone processing. To create a machine vision system, the following steps must be performed: camera calibration, amber image acquisition, key point extraction, point triangulation, and polygon mesh creation.

The Nikon D3100 kit camera was calibrated using a checkerboard image and the cv2.calibrateCamera function (OpenCV library). The reprojection error is 0.3153 pixels (px), which is a good value.

To detect the control points of the image of the amber sample, two methods were researched: SIFT and ORB. 16 images of the sample were researched and the number of key points on the surface of the object, the calculation time for finding the key points and the number of correct key points were determined. The SIFT method was chosen, which is resistant to the influence of adverse factors and has a larger number of correct points.

To determine the spatial position of the camera, the Perspective-n-Point method was used. The errors for the key points of the camera position and orientation are in the range from 0.43 to 50 px. This indicates certain inaccuracies in determining the triangulation parameters, due to objective factors.

After completing all the steps, we have a set of triangulated points for the amber sample. Based on this set, we can calculate the steps for further processing the amber sample.

Keywords: computer vision, camera calibration, 3D visualization, point triangulation, object reconstruction.

I. INTRODUCTION

Amber is a fossilized resin of coniferous trees that formed 40-50 million years ago [1]. In general, amber is a resin that has become fossilized and contains succinite, which in Latin means pine.

Amber is precious because it took millions of years to form, it is a rather rare and limited phenomenon. Amber is found on the Baltic Sea coast, in Ukraine, Romania, Belarus, Mexico, Myanmar and the USA (Alaska).

Extracting amber from the ground is not easy, as it occurs deep in the ground in small quantities and requires specialized equipment. The high price is also influenced by the color of the specimen and the content of various ancient remains of insects and small animals (lizards, frogs).

Amber is widely used in jewelry, folk medicine, cosmetics, perfumery, electronics, and paleontology. For its application, amber is processed and given the desired shape. During processing, samples lose most of their mass, so the development of technology for evaluating the shape and pattern of a sample is very important. Estimation technologies will allow for the efficient use of raw materials, reduce the amount of amber extraction, and thereby reduce the burden on the environment [2,3,4,5].

Modern methods for evaluating the pattern of natural stones include:

– Optical microscopy. This allows the specialist to examine the surface of the stone and the smallest inclusions of the structure, it is used to assess transparency, color and internal defects. The disadvantage is the limited depth of view;
– Infrared spectroscopy. This allows the specialist to detect not only the structure of the stone, but also its chemical composition. The disadvantages of the method are the complexity of interpreting the spectra, which requires a highly qualified specialist;
– Computed tomography. This allows the specialist to create 3D models of the studied sample and allows to detect the internal structure of the stone. The disadvantage is the extremely high cost of equipment and special software;
– Stereovision. This allows a specialist to create a 3D model using a stereo camera and work with it. The disadvantage is the high cost of 3D scanning equipment.

Therefore, based on existing methods, the task is set to create an effective automatic solution for image processing of amber samples. Nowadays, technical visions are quite widely developed, which allows for non-contact evaluation and any type of geometric and colorimetric measurements of the object under investigation. It is also a fairly cheap method, given the achievements of microelectronics.

The purpose of the research is to evaluate the accuracy and time characteristics of methods for determining key points on the surface of amber samples based on their images and to provide suggestions for using these methods in the tasks of 3D visualization and modeling of amber samples in industrial use.

To achieve the goal, the following tasks are performed:

1. Geometric calibration of a digital camera based on standardized chessboard images;
2. Calculation of the reprojection error for the camera and estimation of values in the matrix of internal camera parameters;
3. Experimental research of two methods SIFT and ORB for defining key points, defining the best method for searching key points by parameters accuracy and speed;
4. Extracting control points from an image and comparing them for a set of generated images;
5. According to experimental research data, minimization of errors inherent in projection transformation when forming images of samples. Assessment of the accuracy of the triangulation process based on errors of reprojection, projection matrices, and distortion of the digital camera lens.
6. Obtaining graphic/schematic information about the spatial position of the camera.
7. Triangulation of key points to obtain their coordinates in space.

II. THEORY AND RESEARCH METHODOLOGY

The development of a machine vision system is based on the creation of mathematical models of images, research objects and errors in determining their geometric parameters [6,7,8].

The mathematical model of the technical means of forming images of samples is the model of the camera obscura. The transformation of the coordinates of a point on the surface of the sample when projected into the image plane can be described by the following equation [9]:

$$p = \frac{1}{z}(K\ 0) \times (R\ t) \times P_W, \quad (1)$$

where $P_W = (X_W\ Y_W\ Z_W\ 1)^T$ – homogeneous coordinates in three-dimensional space for a point belonging to the sample surface, $p = (u\ v\ 1)^T$ – homogeneous coordinates of the projection of this point, z – the distance from the optical center of the camera to the normalized image plane, K – matrix of internal parameters of the camera, R – matrix of rotation parameters (external parameters of the camera), t – translation vector.

If you open all the components of the equation, it will have the following form:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \frac{1}{z}\begin{bmatrix} kf & 0 & c_x & 0 \\ 0 & lf & c_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \times \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} X_W \\ Y_W \\ Z_W \\ 1 \end{bmatrix}, \quad (2)$$

where u and v – the coordinates of the point in the image plane with the origin of the two-dimensional coordinate system in the lower left corner, f – focal length of the camera, k and l – scale parameters reflecting the dimensions of the pixel area on the light-sensitive surface, c_x and c_y – coordinates of the optical center of the camera in the image plane, r and t – rotation and translation parameters (camera external orientation parameters).

Camera calibration is an important initial stage in setting up technical vision systems, which allows for increasing the accuracy of image processing and solving object reconstruction tasks. The calibration process includes determining the internal and external parameters of the camera and detecting various types of distortion, such as radial and tangential, caused by the imperfection of the optical system of the camera [9,10].

For calibration, standardized images are used - a template in which objects of the same size alternate [9,10]. The most popular template for calibration was taken as a basis - a chess grid. It consists of black and white squares of the same size that form a chessboard. An important aspect is that this template, like other templates for calibration, should not have a 180-degree unambiguity. Also, when forming images for calibration, several rules must be followed in order to avoid problematic calibration results:

1. Carry out fixation with constant focus;
2. Remember that the focal length will also change when zooming in;
3. Place the template at approximately the same distance at which the object under study will be fixed;
4. Check whether the template was not distorted during printing on paper;
5. There should be a white frame around the outline of the template, approximately the size of 1 grid element. Failure to comply with these conditions may lead to false determination of the grid under insufficient lighting conditions;
6. The rotation of the template to the camera should not exceed 45°;
7. To increase the accuracy of calibration, it is recommended to use 10-20 images, changing their position relative to the camera;
8. The calibration template should occupy at least 20% of the image size;
9. Do not resize or crop the edges of the images.

After the geometric calibration and obtaining the camera parameters based on the SFM (Structure from Motion) method, key points on the sample surface should be found. These points further serve as vertices of polygons approximating the surface of the sample in its 3D model.

First, the initial points of the surface are determined for each formed image, which can later be key points. By comparing the images in pairs, the coincidence of the initial points is determined, which are ultimately the key points of the surface of the sample.

The Kanade–Lucas–Tomasi method is suitable for solving this problem, but it works well only when the distance between the points of consecutive camera locations is small. In other cases, other functions for extracting and mapping control points on the image work well. The most famous of them are: SIFT, ORB [11,12,13].

Based on the task and the location of the camera relative to the sample in the process of forming a sequence of images, we will investigate the SIFT and ORB methods.

The SIFT method is designed to detect control points based on their descriptors. This method identifies key points even if rotation, scaling, or other affine transformations are applied to the images. The basis is the search for descriptors based on the search for local extrema. It searches for the most stable points in two images, that is, invariant to the influence of affine transformations. Gaussian noise is applied to the image and using the Difference of Gaussian or DoG method, these most persistent points are determined. After that, a direction histogram is calculated for each key point based on the magnitude and orientation formulas [11].

$$Magnitude = \sqrt{G_x^2 + G_y^2}, \quad (3)$$

$$Orientation = \arctan\left(\frac{G_y}{G_x}\right),$$

where *Gx, Gy* – partial derivatives of intensity functions *B(x,y)* by two spatial coordinates *x* and *y*.

At the same time, the orientation is the direction of the gradient of the image intensity function, and the magnitude is its value.

Orientation and magnitude characterize the relief of the video signal intensity function on the image.

After receiving this data, we create a descriptor for the point. As a rule, the descriptor in the SIFT method is a set of 16 histograms, which are aligned in a 4x4 grid and each element has 8 directions of orientation. Next, the descriptors of the most stable control points, which are invariant to rotation and scaling, are compared.

The ORB method includes a combination of two methods, namely the FAST (Features from Accelerated Segment Test) point detection method and the BRIEF (Binary Robust Independent Elementary Features) descriptor method [13]. The FAST method detects key points, namely corners, based on their intensity. A pixel is selected from the array. After that, a Bresenham circle[14] is built around this pixel, the radius of which is three. Now the pixels included in the circle are compared with the intensity of the central pixel and divided into three classes: darker, similar, lighter to the central pixel. And if eight pixels are lighter or darker than the central one, the method marks this point as key. But FAST lacks an orientation component and many large-scale functions. That is, when changing the orientation and scale of the image, the key points may be different. In the ORB modification, the FAST method has improvements. For one image, a pyramid of these images is built, where each element of the pyramid also contains the same image, but with a different resolution, and each subsequent element has reduced sampling parameters than the previous ones. The FAST method is used for imaging and key points are identified at each level of the pyramid. Due to this, the method is partially invariant to scale. After finding the points, their direction is determined depending on the change in the intensity level around the determined point. For this, the centroid of the intensity is used, thanks to which we can find out the direction of the displacement of the intensity from its center. Before determining the centroid, we determine the moment of order *(p,q)* for the image fragment around the point (patch) using the formula [14]:

$$m_{pq} = \sum_{x,y} x^p y^q B(x,y) \quad (4)$$

and based on this is the center of mass of the patch:

$$C = \left(\frac{m_{10}}{m_{00}}, \frac{m_{01}}{m_{00}}\right).$$

And the orientation vector itself, which comes from the center, is calculated as follows:

$$\theta = \arctan 2(m_{01} \mid m_{10})$$

After all the key points are found, they are further converted into a binary feature vector containing only 1 and 0 using the BRIEF method.

$$\tau(b;x,y) = \begin{cases} 1 & : b(x) < b(y) \\ 0 & : b(x) \ge b(y) \end{cases}, \quad (5)$$

where τ – binary feature vector; *b(x)* and *b(y)* – intensity in points.

Match point pairs are selected using one of 5 approaches (Uniform, Gaussian, Gaussian2, Coarse Polar Grid, Coarse Polar Grid2) using a binary feature vector [12,13,15].

After these operations, we get descriptors that can be used to match images, calculating the Hamming distance, which can be from 0 to 256 bits [12,13]. In general, the descriptor formula looks like this:

$$f_{n_d}(b) := \sum_{1 \le i \le n_d} 2^{i-1} \tau(b; x_i, x_y).$$

After finding the key points, the process of filtering them takes place, as false comparisons may occur due to noise. To avoid this, the ratio of the Hamming distance to the other closest distance for neighboring points is chosen. If the ratio exceeds 0.8, the key points are discarded. This eliminates about 90% of errors [15,16].

Then, based on this, the Essential matrix is calculated based on the RANSAC method. The Essential matrix describes the relative position and orientation of the cameras. The main idea of the RANSAC method is to find the best value of the Essential matrix E, in which the following condition is satisfied [literature]:

$$x_2^T E x_1 = 0,$$

where x_1, x_2 – coordinates of points found using the SIFT method on 2 images.

The method works in an iterative mode, in such a way as to ensure the best matching of key points. The maximum number of iterations, the threshold value of the error in calculating the projections of points and the number of points sufficient to determine a reliable point by its neighborhood are selected. There are variants of the method based on 2, 3, 5, 6, 7, 8 points. Also a subset of key points is also selected per iteration and the matrix E is found. Next, the error is calculated for each key point match. A pair of key points for which the error does not exceed a given threshold is considered to be an established match. E is calculated at each iteration and compared with the previous value. If an iteration contains a larger number of established key point correspondences, it is considered better than the previous one. Finally, the final matrix value is calculated based on all established correspondences. The obtained result of the matrix calculation is used to obtain the relative parameters of rotation and movement of the camera relative to the sample in the process of forming a sequence of images.

On the basis of the obtained essential matrix, it is now possible to restore two camera positions from which two images were formed. That is, we get the matrices of rotation and displacement. For this, we need the essential matrix calculated above, the key points on the two images, and the calibration matrix. Moreover, it is important to note that after calculating the rotation and displacement matrix, only the direction of the camera movement and the relative scale are given, so the values obtained are relative.

The method consists of several steps [15,16].

First, the main matrix is factorized, which was calculated at the previous stage and has the following form:

$$E = R[t]_x,$$

where R – rotation matrix 3x3, $[t]_x$ – asymmetric transfer matrix of the vector t.

Since the essential matrix has the 2nd rank, it can be decomposed into 2 rotation matrices and 2 translation matrices. For this, the singular expansion of this matrix is used, which has the following form:

$$E = U \sum V^T,$$

Where U,V – orthogonal matrices, and $\sum$ – diagonal matrix of singular values. And the very process of singular expansion allows to get estimates of the values of the rotation matrix R and the transfer vector t.

After applying the singular expansion, we get possible rotation matrices according to the relations:

$$R_1 = UWV^T$$

$$R_2 = UW^T V^T,$$

where the matrix W has the form:

$$W = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

The transfer vector is defined as the 3rd column of the matrix U. After calculation, we have 4 possible variants of the combined matrix *(R1,t), (R1,-t), (R2,t), (R2,-t)*, which are checked to find the correct combination of their components, correct according to the criterion of a positive value of the scene depth. Correctness is determined using Positive Depth Constraint [17]. That is, the corresponding points are triangulated and the condition of positivity of depth is checked. The formula is applied to the first projection matrix (first camera position):

$$P_1 = [I|0],$$

where I – unit matrix.

Formula for the second camera position:

$$P_2 = [R|t].$$

If the depth of the scene is positive, it means that the points are in front of the camera and they coincide. If two more than one pair of points satisfy the condition of positivity of the scene depth, then the one with the smallest coincidence error is selected.

Further, after obtaining projection matrices of size 3x4 for each position of the camera, which contain information about the internal and external parameters of this camera (focal length, optical center, rotation, and displacement), it is possible to triangulate the full-fledged key points of the 3D model from the corresponding matches on the two-dimensional images [16,18]. Coincidences of points are represented in the form of uniform coordinates (*x,y,w*), where *w* is a coefficient that can be normalized to 1 and converts uniform coordinates into ordinary

$$x_1 = \begin{bmatrix} u_1 \\ v_1 \\ 1 \end{bmatrix} \text{ and } x_2 = \begin{bmatrix} u_2 \\ v_2 \\ 1 \end{bmatrix}$$

Where *u,v* – pixel coordinates of points in the image. According to formulas (1) and (2), we have to find X with known x_1, x_2, P_1, P_2:

$$x_1 = P_1 X \text{ and } x_2 = P_2 X .$$

For this, a system of linear equations is formed on the basis of projection coincidences for each point, which satisfy the following conditions:

$$x_1 P_1 X = 0 \text{ and } x_2 P_2 X = 0 .$$

The system of equations for each point can be written as:

$$\begin{cases} u_1 p_{31} X_1 + u_1 p_{32} X_2 + u_1 p_{33} X_3 + u_1 p_{34} - p_{11} X_1 - p_{12} X_2 - p_{13} X_3 - p_{14} = 0 \\ v_1 p_{31} X_1 + v_1 p_{32} X_2 + v_1 p_{33} X_3 + v_1 p_{34} - p_{21} X_1 - p_{22} X_2 - p_{23} X_3 - p_{14} = 0 \\ u_2 p_{31} X_1 + u_2 p_{32} X_2 + u_2 p_{33} X_3 + u_2 p_{34} - p_{11} X_1 - p_{12} X_2 - p_{13} X_3 - p_{14} = 0 \\ v_2 p_{31} X_1 + v_2 p_{32} X_2 + v_2 p_{33} X_3 + v_2 p_{34} - p_{21} X_1 - p_{22} X_2 - p_{23} X_3 - p_{14} = 0 \end{cases}$$

The given equations are combined into a common matrix and with the help of the SVD (singular decomposition) method, the desired value of the vector $X(X_1, X_2, X_3)$ with the smallest deviation from the coordinates of the points on both images is found. After finding the coordinates of the 3D model points, an important step is to calculate the reprojection error.

There are several approaches to solving this PnP problem [16,19]. Using the DLT (Direct Linear Transform) method, as a classic method for solving the PnP problem, allows finding a projective transformation according to formula (2), based on connected 3D points with their 2D projections on the image. A system of equations based on SVD is solved. DLT is quite sensitive to noise, as it is a linear method. Also depends on where these 3D points are located. If they are located close to each other, the solution data will be unstable.

Another method of solving the PnP problem contains an improved version of EPnP [15]. The main idea of the improvement is that all 3D points are represented as a linear combination of several base points. This in turn reduces the dimensionality of the problem and speeds up the search for the projection matrix. The displacement and rotation matrices R and t provide the transformation of a set of points P_i in into their corresponding 2D points in the image p_i.

$$p_i = K * (R * P_i + t) .$$

An improvement of this method is that key points of the images are used to approximate the coordinates of all 3D points. In order not to work with a large number of 3D points, this method represents the coordinates of each 3D point as a linear combination of four coordinates of key points:

$$P_i = \sum_{j=0}^{3} a_{ij} C_j ,$$

where a_{ij} – weighting coefficients that determine the influence of a key point C_j on the placement P_i.

These weight coefficients are calculated so that the coordinates of each 3D point can be found through the coordinates of the key points. Now the equation looks like this:

$$p_i = K * (R * (\sum_{j=0}^{3} a_{ij} C_j) + t) \cdot$$

Now the coordinates of all 3D points are expressed through a combination of the coordinates of the key points. After that, a system of linear equations is constructed and the least squares method is used to solve it. The disadvantage of this method is the placement of 3D points near each other and sometimes the need to apply additional nonlinear optimization to achieve better accuracy.

The next method for computing the PnP problem is RANSAC PnP [16,19]. It is based on repeatedly randomly selecting a subset of the data to evaluate the model and then determining the number of data that fit the three-dimensional model of the sample. A search is made for the subset of points that best matches the correct correspondence between the 3D points and their projections in the image. The subset of points that best satisfies the correct correspondence between the 3D points and their projections in the image is searched. The method begins by randomly selecting a minimum number of points to estimate this model. Next, based on these points, the projection matrix R and t are calculated. For their calculation, the PnP or EPnP mentioned above is used to pre-estimate the projection of the points. After the calculation, based on the found projection matrix, the projection points are searched and compared with the real 2D points. The error is also calculated. If this error is less than a threshold value, then these points are determined as key points. This is done for many points and if the number of control points exceeds a set threshold, the 3D model is acceptable. After all iterations are completed, the model with the largest number of key points is selected and an estimate is calculated using nonlinear optimization to more accurately find the camera parameters. The method is robust to interference and reliably processes noisy images. It can work in real time, as its computational efficiency is high for tasks with a relatively small number of 3D points.

Thus, the given theoretical models allow to describe the process of formation of key points of images. On the basis of key points, the surface of amber samples is approximated in 3D space.

III. CALCULATION

The subject of the computational experiment is a comparison of the accuracy and speed of SIFT and ORB methods. These methods find key points on the formed images and are an integral part of 3D visualization and modeling of amber samples. The characteristics of these methods ultimately determine the effectiveness of visualization and modeling of samples.

The first stage for the task of three-dimensional visualization of amber samples is the calibration of the camera.

To calibrate the Nikon D3100 kit camera, which includes the Nikon 18-55mm f/3.5-5.6G VR AF-S DX lens, a test image was used, namely a chessboard with a size of 9x7 squares.

A group of 16 images was formed (Fig. 1).

After loading the chessboard frames into the computer's memory, the following calculations are applied to them. First, the image is converted to grayscale to simplify processing. After that, blurring is applied to the image, most often a Gaussian filter is used to reduce noise. When the image is ready for processing, chessboard’s countours need to be found. In order to do this, the image goes through binary inversion so that 0 is black, and 255 is white and contours could be seen. It is possible to use the Canny method [9,10]. After that, the image will contain a white outline of the object we research. Next, the angles are determined, namely the intersection between the lines of squares.

The method works based on the search for local minima of the intensity difference. After the found angles between the contours on each image, 3D coordinates are constructed accordingly, in which z = 0, which means that the points lie on the same plane. Next, the method finds camera parameters that minimize the difference between the points found on the 2D image and the corresponding 3D points. To perform all these actions, the OpenCV library provides ready-made camera calibration functions cv2.calibrateCamera, which return the reprojection error value, the matrix of internal parameters, distortion coefficients, rotation and translation matrices. Two values are important for further calculations: the reprojection error and the matrix of internal parameters. After executing the cv2.calibrateCamera function, the reprojection error is 0.3153 px.

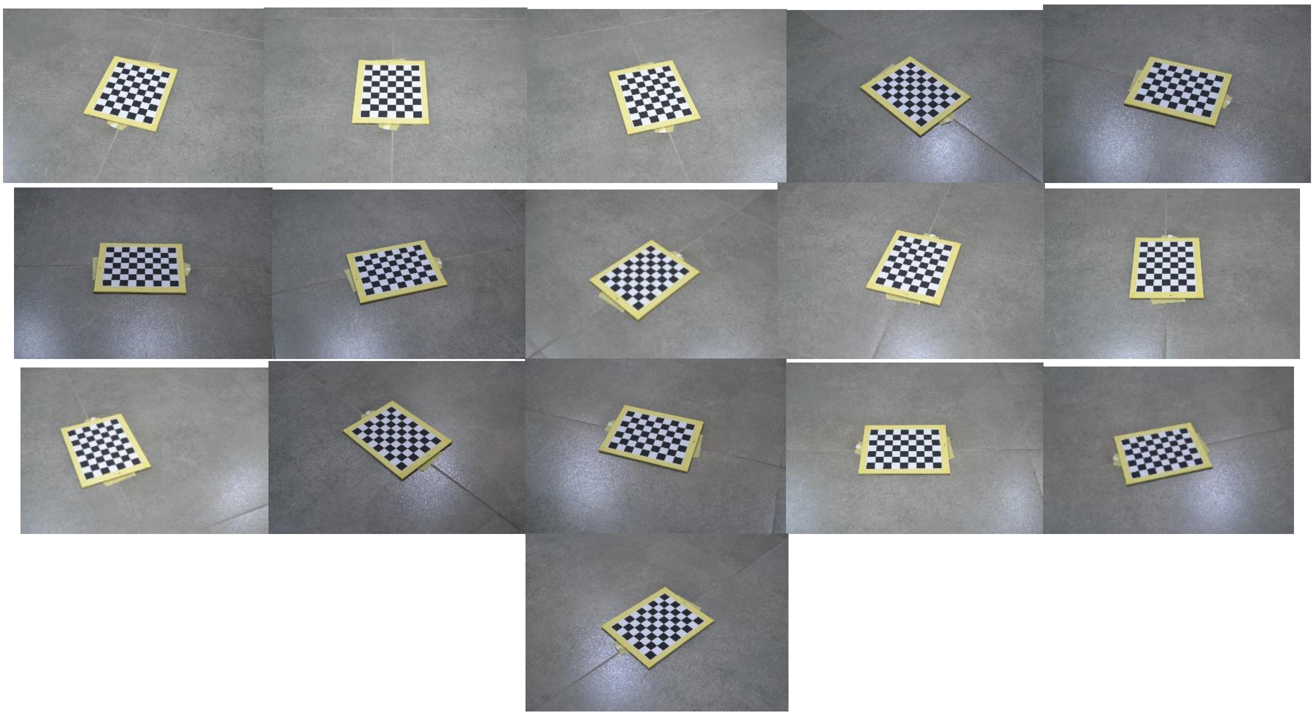


Fig.1 Test images of a chessboard from different perspective for camera calibration

This is a pretty good value, as the smaller this error, the more accurately the elements in the matrix of the camera's internal parameters are selected. The matrix has the following form:

$$p = \begin{bmatrix} 6.62490476e+03 & 0 & 2.30447768e+03 \\ 0 & 6.57219658e+03 & 1.53356815e+03 \\ 0 & 0 & 1 \end{bmatrix}$$

The obtained coefficients satisfactorily match the parameters contained in the image metadata.

After the camera is calibrated, it is necessary to find key points in the two working images. These are the elements that match between the two images. To research the separation and comparison of control points in the image, it is proposed to apply the SIFT and ORB methods [11, 15,16].

For the SIFT method, research was conducted on one sample of amber images. For the research, a comparison of two images of the same amber sample at different camera positions was carried out (16 images from different angles (Fig. 2).

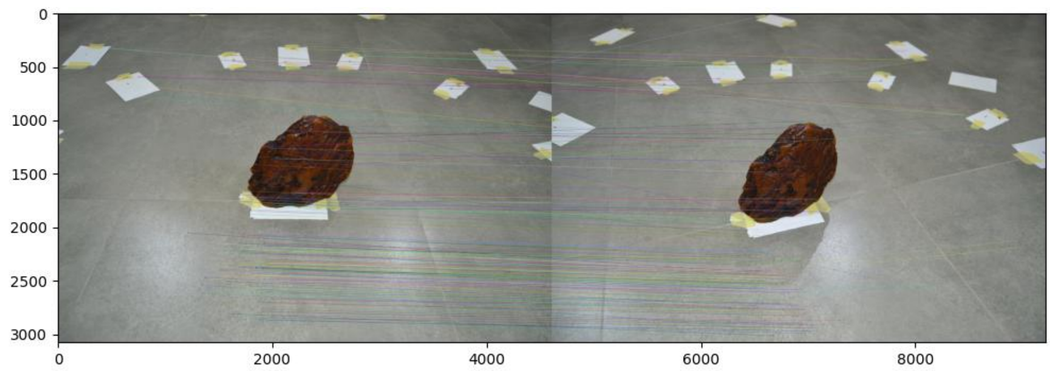


Fig.2 Matching two images of an amber sample using the SIFT method

Experimental research of the time characteristics of the methods were conducted on a computer platform with the following characteristics: Intel i5 3.3 GHz processor, 32 GB of RAM.

The program was written in the Python programming language using the OpenCV library.

Matching points in two images of the same object is a fundamental step in image reconstruction. Identifying similar points of the same object in two images provides information about the 3D position in space, since the projection on the photo is made to different places in the image. Analyzing these coincidences using epipolar geometry allows us to estimate the spatial location of the point.

The results obtained by the program are presented in Table 1: execution time when finding the key points of the first and second images, the number of key points of the first and second images, the number of correctly matched points based on the coincidence of the key points of the first and second images.

Table 1

Results of research on the SIFT key points search method

Image pair number for key points search (search iteration)	Time to search for key points in images, s		Number of key points found		Number of key point matches for two images
	First image of the couple	Second image of the couple	First image of the couple	Second image of the couple	
0	1,619040251	1,645991325	5965	7636	122
1	1,56699872	1,555999517	7636	10588	61
2	1,620998859	1,556033373	10588	3063	46
3	1,570999622	1,627004623	3063	4403	145
4	1,606998682	1,600999832	4403	5458	144
5	1,548999071	1,568998814	5458	5091	67
6	1,582999468	1,627999306	5091	19005	43
7	1,637000084	1,561990023	19005	9489	286
8	1,587002754	1,568997383	9489	3676	293
9	1,544998407	1,530999422	3676	2559	84
10	1,578999996	1,543001652	2559	2677	36
11	1,542999983	1,535998583	2677	3402	37
12	1,53599906	1,53000021	3402	1796	51
13	1,53499794	1,598999739	1796	3491	28
14	1,617999792	1,558999777	3491	1862	39
15	1,530999184	1,567026138	1862	5965	36
Total	25,22803187	25,17903972			1518

For the ORB point detection method, research was conducted on one sample of amber images. For the research, two images of the same amber sample were compared at different camera positions (Fig. 3).

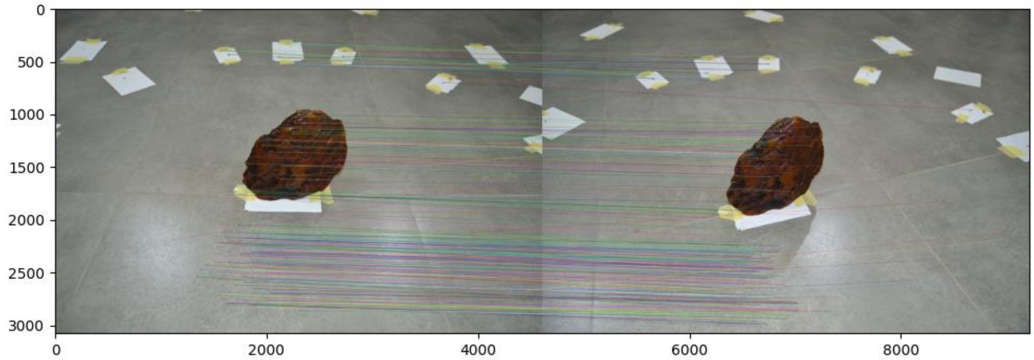


Fig. 3 Comparison of two images of an amber sample using the ORB method

The results obtained by the program are presented in Table 2: execution time for finding key points of the first and second images, number of key points of the first and second images, number of correctly matched points.

We compare the two methods SIFT and ORB for determining key points. For comparison, we will build comparison graphs: the graph of dependence for the time of calculating key points on the iteration of the first and second images is presented in Fig. 4, the graph of dependence on the number of key points on the iterations of the first and second images is presented in Fig. 5, the graph of dependence on the number of correctly matched points of two images is presented in Fig. 6 and 7.

Table 2

Results of research on the ORB key points search method

Image pair number for key points search (search iteration)	**Time to search for key points in images, s**		**Number of key points found**		**Number of key point matches for two images**
	First image of the couple	**Second image of the couple**	**First image of the couple**	**Second image of the couple**	
0	0,276781082	0,101001501	17384	19411	159
1	0,092030048	0,092998981	19411	19339	27
2	0,097999334	0,075999022	19339	9452	32
3	0,069000483	0,081983328	9452	13500	172
4	0,074025869	0,086037397	13500	16590	171
5	0,082998991	0,080023766	16590	14165	42
6	0,088000059	0,111000299	14165	19929	27
7	0,103029251	0,094996691	19929	18910	310
8	0,087000847	0,078025818	18910	10568	325
9	0,074026346	0,073998451	10568	6697	52
10	0,066026449	0,070000172	6697	6431	25
11	0,066026688	0,074973106	6431	8144	37
12	0,068027735	0,068000555	8144	4853	43
13	0,062025309	0,078000307	4853	10039	14
14	0,076025486	0,067998886	10039	4526	15

15	0,067026615	0,08699894	4526	17384	29
Total	1,450050592	1,32203722			1480

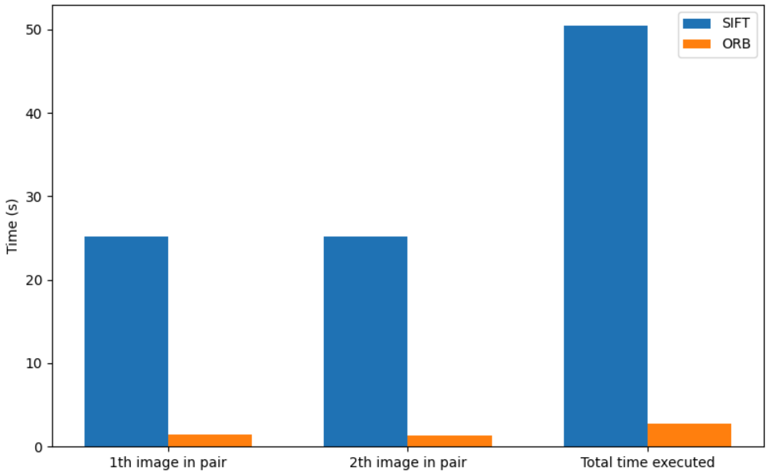
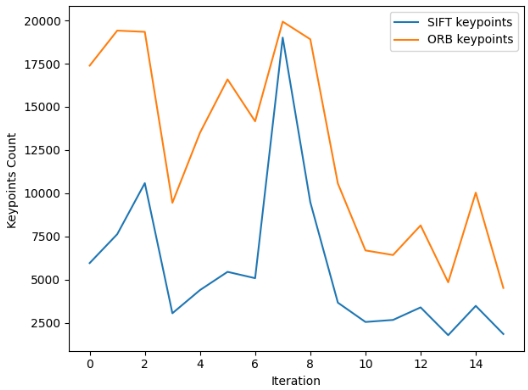
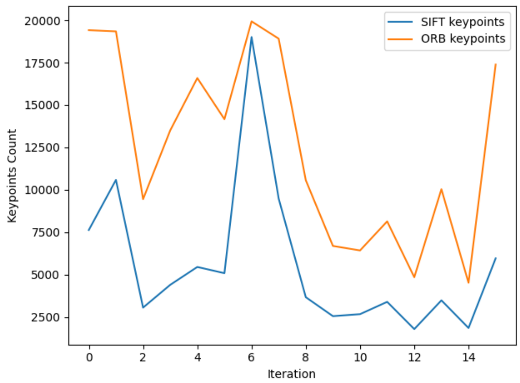


Fig.4 Comparison of time characteristics of key point search methods



a



б

Fig. 5 Dependence of the detected number of key points on the number of iterations of processing the formed set of images: a) the first image of the pair; b) the second image of the pair

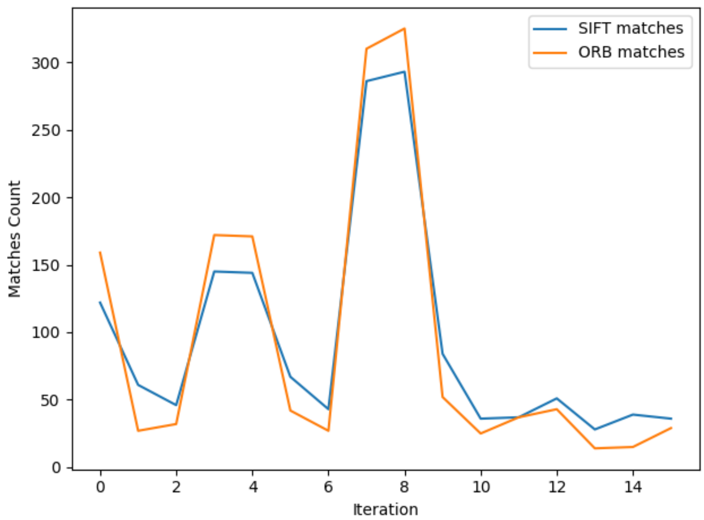


Fig. 6 Dependence of the number of detected matches of key points in a pair of images on the number of iterations for SIFT and ORB methods

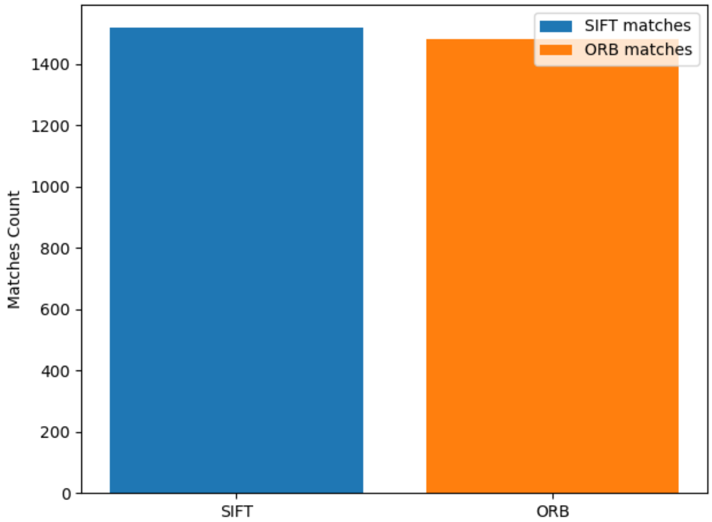


Fig. 7 Number of key point matches on image pairs using SIFT and ORB methods for the entire set of images

As a result, a set of images of two amber samples was processed and the research of two methods, SIFT and ORB, was conducted to detect key points. According to the results of calculations and comparison of the SIFT and ORB methods, it can be determined that:

1. The ORB method works faster in terms of time according to the dependence graphs for the time of calculation of key points from the iteration (Fig. 4);
2. The ORB method selects a larger number of key points according to the graph of the dependence on the number of key points from iterations (Fig. 5);
3. The ORB method has maxima according to the graph of the dependence on the number of correctly mapped points from iterations, but the SIFT method overlaps the curve of the ORB method in most cases (Fig. 6);
4. The SIFT method has a larger total number of correct matched points according to the graphs of the number of correct matches in general for the entire period of passing the methods (Fig. 7);
5. The SIFT method works better for images with high detail, texture and complex scenes.
6. The SIFT method works quite well with scaling and minor rotations (in the case of the research, images were formed with steps of 15, 30 degrees).
7. The SIFT method is less sensitive to sample illumination deficiencies due to the use of gradients when calculating descriptors. In the image (Fig. 2) there were minor changes in illumination, but the method performed well and found the correct key points.
8. The SIFT method uses a higher quality and more informative descriptor compared to ORB, which increases the accuracy of finding relevant points on the image.
9. The SIFT method is highly demanding on hardware resources.

Considering all the points of comparison between the SIFT and ORB methods, we choose the SIFT method to find the set of key points, since it finds a larger number of correctly matched points, provides work with complex textured sample surfaces, and is resistant to interference and image defects.

To sift out false key point results, filtering with a coefficient of 0.7, was used for better filtering of the results.

Then the reprojection error is calculated. This is a key indicator for evaluating the quality of the triangulation process. Error values are affected by noise and distortion in the camera lens. They add inaccuracies in the estimation of projection matrices. According to the obtained values of the reprojection error, graphs were constructed for the two methods (Fig. 8 and 9).

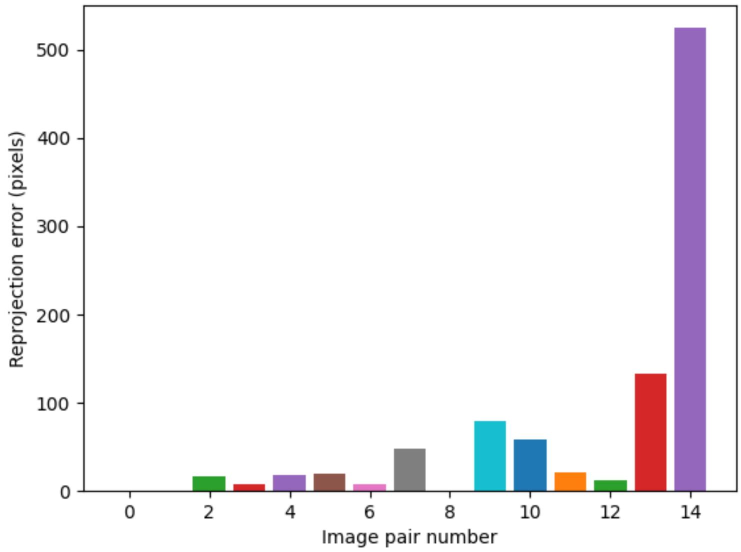


Fig. 8 Reprojection errors for key points found by the SIFT method

From Fig. 8, we can see that the errors vary in the range from 0.43 px to 500 px. In general, the ideal value of the reprojection error is less than 1 px. Low errors up to 10 px indicate that the triangulation is calculated quite accurately. From 10 px to 50 px, average accuracy values may indicate some inaccuracies in the triangulation parameters. From 50 px to 520 px, significant deviations may be caused by errors in finding the corresponding points on the image.

From the graph, we can conclude that most of the points are in the range from 0.43 px to 50px, which indicates some inaccuracies in the triangulation parameters.

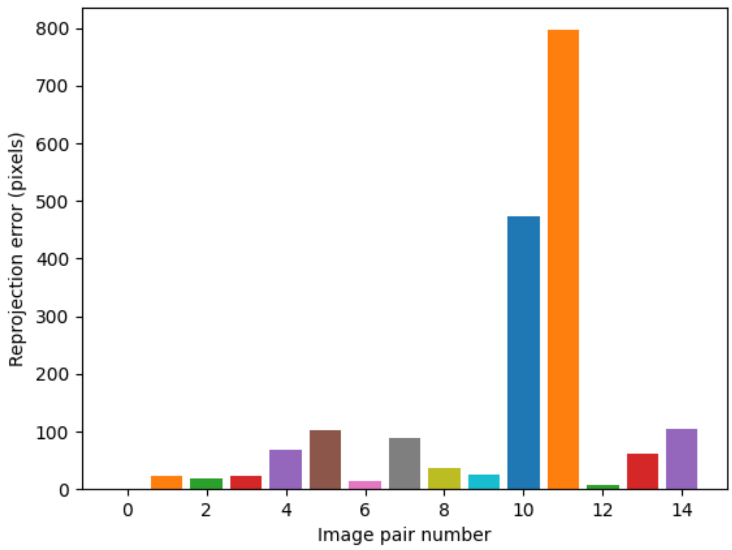


Fig. 9 Reprojection errors for key points found by the ORB method

It can be seen from Fig. 9 that the errors vary in the range from 0.63 px to 802 px.

It can be concluded from the graph that most of the points are in the range from 0.63 px to 100 px, and the maximum reprojection error reaches 800 px, which indicates certain inaccuracies in the triangulation parameters due to objective factors.

Also, camera calibration has a fairly strong positive effect on the error, which increases the accuracy of determining the motion and rotation matrix:

Rotation matrix

[-0.93299267 0.09844558 -0.34616924]

[-0.30827359 0.27775455 0.90984603]

[0.1857204 0.95559452 -0.22879478]

Vector of motion

[[1.87444477]

[-4.72681567]

[6.99074454]]

After calculating the errors, an important step is to refine the position and orientation of the camera based on the obtained coordinates of the points after triangulation. The Pnp (Perspective-n-Point) method provides a reliable way to solve this problem. This method allows not only to obtain data about the position of the camera, but also to improve the quality of the reconstruction by minimizing the errors that occur during the projection transformation. If we apply this to all images, we will get the coordinates of all points from which the images of the amber sample were formed. (Fig.10).

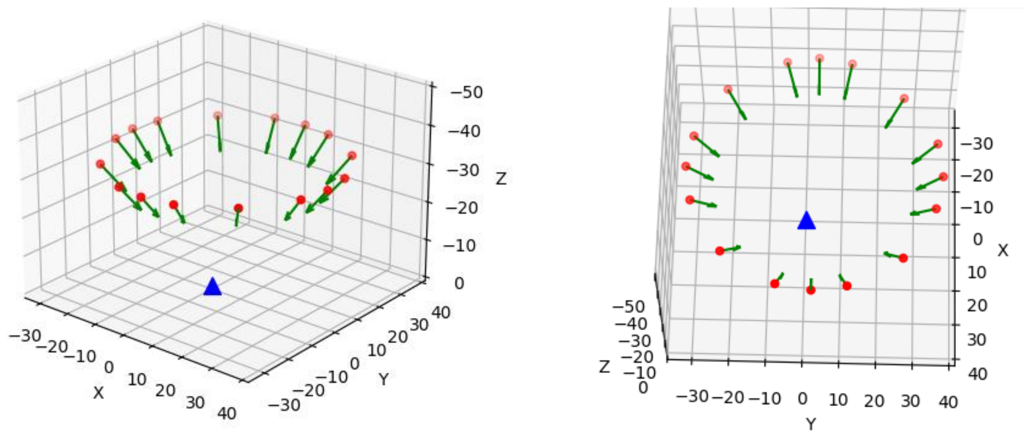


Fig.10 Position of the digital camera relative to the amber sample during the process of forming a set of its images (x,y,z – three coordinates of 3D space, Δ – center of mass of the sample, ° – center of the camera optical system, → – direction of the camera's optical axis)

After completing all the above steps, we get a set of triangulated 3D points for the surface of the amber sample (Fig.11).

a) b) c)

Fig..11 The set of points belonging to the surface of the amber sample: a) 1 perspective, b) 2 perspective, c) 3 perspective

Fig. 11 a,b,c shows the set of points obtained using the SFM algorithm. It represents points in 3D space for the object under study based on processed images. Each point in the set is a key point in two adjacent images, selected using the SIFT method. For each point, its spatial position is determined, namely the x, y, z coordinates, and color information.

The images represent points in 3D space that were identified as key points in 2 adjacent images. Not all of them belong to the studied sample. Some points correspond to the paper on which the sample is located, as they contain persistent gradation properties that are quite stable. To get the surface of the sample and get rid of the points that do not belong to it, you need to use filtering by distance. A sphere of a given size is constructed, which will completely reject those points that do not belong to the sample. It is also necessary to set filtering of points by color, as the color spectrum of amber is quite narrow.

As a result, we get a set of points in 3D space that correspond to the surface of the amber sample. Further, based on these points, we can construct approximations of the surface using polygonal elements, faces, and vertices. This approximation will be the subject of further research by the authors of the article and will allow solving practical problems related to deposits and amber processing processes.

IV. RESULT AND DISCUSSION

The obtained research results indicate a successful solution to the problem of developing a method (program) for three-dimensional visualization of amber samples.

1. The calibration of the Nikon D3100 kit digital camera, with the Nikon 18-55mm f/3.5-5.6G VR AF-S DX lens, was carried out. A method based on the processing of a test image was used, namely a chessboard with a size of 9x7 squares. 16 test images were used, which ensures precise calibration accuracy.

Based on the OpenCV library, ready-made solutions for camera calibration were applied using the cv2.calibrateCamera function. Reprojection errors, the matrix of internal parameters, distortion coefficients, rotation and translation matrices were calculated. The reprojection error is 0.3153 px, which is a fairly accurate value. The matrix of internal parameters of the digital camera is obtained. The obtained coefficients approximately correspond to those parameters that satisfactorily correspond to the metadata of the generated images.

2. Sixteen images formed for each amber sample with a step of 15°, 30°. On this basis, key points are found for neighboring images, which will be used in the triangulation algorithm. At this step, incorrect or inaccurate detection of these points may occur, which may cause a shift in the reconstruction of the sample surface.

3. Two methods SIFT and ORB for determining key points were experimentally investigated. Two neighboring images were selected for the research, the position of which differs by 15°, 30°. And the comparison was carried out in a circle, clockwise. Tables were constructed for the amber sample of two matched images. The tables display data from sixteen iterations, execution time for finding key points of the first and second images, number of key points of the first and second images, number of correctly matched points.

4. During the research of the SIFT and ORB methods. Their time and accuracy characteristics were obtained and compared. In terms of time, the ORB method works better for determining key points. The ORB method also works better for determining the number of key points. According to the graph of the number of correctly mapped points, it can be seen that the ORB method has a peak, but the SIFT method has an advantage in larger sections of the graph. To clarify the situation, bar charts have been constructed. It is clearly seen from the bar charts that the SIFT method has a higher number of correct key points that matched in the images.

The SIFT method was analyzed. According, the following advantages were taken into account:

- The SIFT method has a qualitative and more informative descriptor of points;
- SIFT method works better with high detail, texture and complex scenes in amber images;
- The SIFT method works well for image scaling and slight camera rotations;
- SIFT method is less sensitive to lighting imperfections in amber samples.

The SIFT method was chosen to determine key points, taking into account the larger number of correct key points found and the listed advantages.

5. Key points in images are searched and matched based on the SIFT method. This step includes detecting the most stable features of points that are insensitive to rotation and illumination, determining histograms for the features of the found points, and comparing the histograms of the found features for the two images.

6. Data on the spatial positions of the camera were obtained and the accuracy and quality of the reconstruction of the 3D surface of the sample were improved. The reconstruction is based on 16 formed images. The spatial coordinates of all camera positions from which images were formed were obtained. All spatial camera positions are obtained using the PnP (Perspective-n-Point) method solution to determine the relative positions of the camera between a pair of images and represent data on the movement and rotation of the cameras. After displaying these parameters on the graph, you can visually assess the correctness of the solution of the PnP problem for the camera and the two images.

7. Triangulation was applied to the key points identified using the SIFT method. This allowed the coordinates of the key points in 3D space to be determined.

8. The received triangulation data were checked by re-projecting them onto the 2D plane and the error of the received data was determined. The accuracy of the triangulation process was evaluated based on the reprojection errors. It can be concluded that for most points the error is in the range from 0.43 px to 40 px, which indicates some inaccuracies in the triangulation parameters. The magnitude of the error can be affected by errors in determining key points as well as inaccuracies in camera calibration. The accuracy of projection matrices and lens distortions was assessed based on the definition of the motion and rotation matrix.

9. The main result is the formation of a set of points belonging to the surface of the amber sample. The set gives an idea of the shape of the amber and the ability to programmatically interact with the parameters of this shape. It is also possible to obtain an estimate of the volume of the sample and plan the steps for its processing to the state of an industrial product.

V. CONCLUSIONS AND FUTURE SCOPE

Based on the results of the research of methods for determining key points in the task of three-dimensional visualization and modeling of amber samples, the following conclusions can be drawn:

1. To prepare technical means for work as part of a technical vision system, it is necessary to calibrate a digital camera. The possibility of calibration based on a set of images of a test object (chessboard) and the OpenCV library software has been proven. The results of the calibration procedure are the matrix of the camera's internal parameters, the distortion coefficients of the optical system, and the matrix of the camera's rotation and movement relative to the object in the process of forming a set of images. In further calculations, this data allows determining the three-dimensional coordinates of key points of the amber sample based on its two-dimensional images.
2. The initial data for creating a three-dimensional model is a set of images of an amber sample formed by a digital camera. Numerical modeling and calculations confirmed the possibility of solving the specified problem on the basis of 16 images obtained when the camera was moved step by step around the sample in one plane.
3. Through a preliminary analysis of existing methods for searching and matching key points, two methods (SIFT and ORB) were selected for detailed studies. The two methods were compared based on the accuracy of coordinate determination, calculation time, and resistance to adverse factors (noise in the images, uneven lighting of the samples, etc.). The SIFT method was preferred on the basis of: detecting a larger number of key points and acceptable accuracy; resistance to adverse factors; better compliance with the conditions of studying amber samples. Although it should be noted that this method is computationally complex, and the possibility of its implementation in real time depends on the degree of detail of the images.
4. The results of using the SIFT method make it possible to obtain the spatial position of the camera at all points and improve the quality of the reconstruction of the three-dimensional model. In this practical task, the reconstruction consists of calculating the coordinates of the vertices for a three-dimensional approximation of the shape of the amber sample. This calculation is carried out using the two-dimensional coordinates of the key points and data on the spatial position of the camera.
5. Triangulation was performed and the coordinates of the vertices of the three-dimensional spatial model of amber samples were determined. To assess the accuracy of determining the coordinates of the vertices of the spatial model, it was proposed to perform a repeated procedure of their projection transformation into image projections. In contrast to the process of forming initial images by the optical system of the camera, this conversion is carried

out numerically in the computing environment of the technical vision system. For most keypoints, the error of such reprojection is in the range from 0.43 px to 40 px. This indicates certain inaccuracies in determining the parameters of the triangulation procedure. It is possible to eliminate this shortcoming by improving image forming procedures.

6. The final result of the study is to obtain a set of points belonging to the surface of the sample and corresponding to the vertices of its three-dimensional model. On the basis of the three-dimensional model, it is possible to analyze the shape of the geometric dimensions and volume of the samples, visualization of the samples, as well as the planning of the processes of their processing to obtain industrial amber products.

Further research plans on the topic of the article

Several key tasks will be solved in further studies of the reconstructed amber sample.

First, the color of each fragment of the sample will be determined, which will allow us to investigate the chemical composition and possible natural features of the amber, such as the presence of impregnation.

Secondly, a detailed analysis of the geometric shape of the samples will be carried out, using both standard geometric approaches and fractal dimension measurements. Based on these calculations, areas with impregnation and defects will be determined, which will allow identifying possible areas for further processing or restoration of the amber.

Thirdly, it is necessary to investigate the methods of fitting the geometric shapes of future products into the three-dimensional model of the samples, which will allow to accurately calculate the further processing of the samples, with the aim of minimizing the loss of precious material.

REFERENCES

[1] Laboratory "InConsulting". Diagnosis and examination of amber. 2024. https://inconsulting.com.ua/uk/gemologchna-ekspertiza/diahnostyka-ta-ekspertyza-burshtynu.html.

[2] Remezova O.O., Naumenko U.Z. The problem of choosing methods for developing amber deposits in Ukraine. X International Scientific and Practical Conference "Technologies and Processes in Mining and Construction", Lutsk, 2022. 82-90. https://donntu.edu.ua/wp-content/uploads/2023/07/zbirkatez_2022_compressed.pdf#page=82.

[3] Naumenko U.Z., Matsui V.M., 2020. Stages of geological development and genetic types of the root primary source of amber-succinite placers of Ukraine. Geological Journal (Ukraine), No. 4 (373), pp. 76—86. https://doi.org/10.30836/igs.1025-6814. 2020.4.209022.

[4] Environmental consequences of the adverse impact of unauthorized amber mining on forest areas of Zhytomyr region [Електронний ресурс] / S.Kovalevskii, Y. Marchuk, K. Maevskyi, A. Churilov. – 2021 https://doi.org/10.31548/forest2021.01.007.

[5] Levytskyi, V., Sobolevskyi, R., Korobiichuk, V. (2018). The optimization of technological mining parameters in a quarry for dimension stone blocks quality improvement based on photogrammetric techniques of measurement. Rudarsko Geolosko Naftni Zbornik, 33 (2), 83–89. doi: https://doi.org/10.17794/rgn.2018.2.8.

[6] Sheikh, H. R., Sabir, M. F., Bovik, A. C. (2006). A Statistical Evaluation of Recent Full Reference Image Quality Assessment Algorithms. IEEE Transactions on Image Processing, 15 (11), 3440–3451. doi: https://doi.org/10.1109/tip.2006.881959.

[7] Korobiichuk, I., Podchashinskiy, Y., Shapovalova, O., Shadura, V., Nowicki, M., Szewczyk, R. (2015). Precision increase in automated digital image measurement systems of geometric values. Advances in Intelligent Systems and Computing, 335–340. doi: https://doi.org/10.1007/978-3-319-23923-1_51.

[8] Korobiichuk, I., Podchashinskiy, Y., Luhovykh, O., Levkivskyi, V., Rzeplińska-Rykała, K. (2020). Theoretical Estimates of the Accuracy of Determination of Geometric Parameters of Objects on Digital Images. Automation 2020: Towards Industry of the Future, 289–299. doi: https://doi.org/10.1007/978-3-030-40971-5_27.

[9] David Forsyth, Jean Ponce Computer Vision: A Modern Approach 2nd Edition, Pearson Education, 2011, 800 p.

[10] Gonzalez and. Richard E. Woods' Digital Image Processing, Fourth Edition, Pearson Education Limited 2018, - 1022 p.

[11] Singh A. SIFT Algorithm: How to Use SIFT for Image Matching in Python. Aishwarya Singh. 2019. https://www.analyticsvidhya.com/blog/2019/10/detailed-guide-powerful-sift-technique-image-matching-python.

[12] Detecting and Tracking Objects with ORB Algorithm using OpenCV. 2024. https://medium.com/thedeephub/detecting-and-tracking-objects-with-orb-using-opencv-d228f4c9054e.

[13] Tyagi D. Introduction to SIFT(Scale Invariant Feature Transform). Deepanshu Tyagi. 2019. https://medium.com/@deepanshut041/introduction-to-sift-scale-invariant-feature-transform-65d7f3a72d40.

[14] Bresenham's circle drawing algorithm [Електронний ресурс]. 2024. https://www.geeksforgeeks.org/bresenhams-circle-drawing-algorithm.

[15] Caldelli R, Amerini I, Costanzo A.SIFT match removal and keypoint preservation through dominant orientation shift. 2015 23rd. European Signal Processing Conference (EUSIPCO) 2015: 2062-2066. DOI: 10.1109/EUSIPCO.2015.7362747.

[16] Andrea Costanzo, Amerini Irene, Caldelli Roberto, Barni Mauro. Forensic Analysis of SIFT Keypoint Removal and Injection. IEEE Transactions on information forensics and security 2014: 1450-1464. DOI: 10.1109/TIFS.2014.2337654.

[17] McInroy J. Robust pose estimation which guarantees positive depths. J. McInroy, C. Li. 2023. https://www.nature.com/articles/s41598-023-49553-9.

[18] Lobo T. Understanding Structure From Motion Algorithms .Teresa Lobo. 2023. https://medium.com/@loboateresa/understanding-structure-from-motion-algorithms-fc034875fd0c.

[19] Lepetit V. EPnP: An accurate O(n) solution to the PnP problem. V. Lepetit, P. Fua, F. Moreno-Noguer. 2009.https://www.researchgate.net/publication/49458221_EPnP_An_accurate_On_solution_to_the_PnP_problem.