CONCEPT OF BINOCULAR STEREO VISION SYSTEM
FOR INSPECTION OF SURFACES

Key words
Fatigue tests, surface inspection, optical inspection, stereo vision.

Summary

The article presents the idea of a binocular stereo vision system for surface inspection and discusses the basic characteristics of epipolar geometry and the advantages and limitations of the methods of representation that use it. The structure of the vision system, basic optical parameters and the configuration of the equipment are presented. The assumed practical implementation possibilities are determined.

1. Review of methods of monitoring the materials destruction processes

One of the basic issues in materials and structure testing is the fatigue strength and cracking susceptibility when influenced by loads of different character and intensity. The problem of fatigue strength has a multi-scale character. Most of all, it concerns the operation and maintenance of technical objects and, with reference to supporting structures, machines and devices, is hence analysed on a macro scale. Together with the development of the advanced material technologies and mechatronic technologies, the problem of testing the fatigue strength on a micro scale is growing in importance. The multi-scale of the material destruction processes, most importantly fatigue cracking, generally results
from the very nature of these phenomena. The initiation of fatigue cracking has a local character and is discussed and tested on a micro scale. On the other hand, when the crack expands, it is then analysed on a macro scale.

As emphasised in literature, the science of the fatigue of materials is mainly experimental. One of the main sources of knowledge of materials is the fatigue test realised with the use of various devices and measurement apparatus. Many methods and specialised test devices and apparatus have been developed to detect and measure the fatigue cracks. Among the contact measurement methods [12], the most frequently applied techniques include ultrasonic and the resistively testing. The fibre optic sensors, when densely mounted, allow the tension of large structures in real time, are currently receiving more attention [5]. The principle of operation of these sensors is based on the change in the state of polarisation or spectral characteristics of the propagation of light. The second group consists of the non-contact methods, among the most often applied are the vision methods based on the computer analysis of the image, the acoustic methods, and the eddy current method. The tests carried out with the use of acoustic emission consist in the analysis of the elastic waves generated in solid state [10]. Their sources are, among others, the processes of the micro crack progress. An early signalling of the slowly growing structural defects is the advantage of this method. In the case of conductive materials, the technique of eddy current is applied, which enables the detection of the surface and subsurface defects [13]. This technique is characterised by the high sensitiveness of the detection of the discontinuity of the material and cracks.

Together with the development of optoelectronic technologies, the visual methods for the observation of destruction processes occurring on the surface of materials are more widely applied in laboratory tests and the inspection of the technical condition of structures at the time of their operation and maintenance [18]. They are quick and can be applied on a micro and macro scale and do not require the surface to be prepared in any special way. Advanced numerical methods including the digital image correlation (DIC) methods and AI methods are used for the analysis of images, the monitoring of the fatigue crack, and identifying the properties of materials [16]. One of the basic pros of the visual method is the possibility to automate the process of the inspection of the surface during tests [6]. The vast majority of currently applied solutions use the single camera optical systems for the 2D analysis of surfaces. Such monocular observation has various limitations. These include the following:

− The lack of information about the depth of the space that is important in the case of great deformations on the surface of the sample, the inspection of holes and the shift or rotation of the sample in relation to the optical system [17];
− The effect of the “covering” of the part of the area of the inspected surface with protruding bits of the material, which can take place in cracking processes.
In the past few years, the dynamic development of CCD cameras and the growing computer ability at a blistering pace allowed the design and practical application of vision systems for the 3D analysis of surfaces. The methods of the acquisition of images that are used in these systems can be divided into those that use [9]:

- depth from focus; their advantage, after the full scan is performed, is the possibility to create the image with maximum sharpness for each and every area observed;
- structural lighting, whose optical axis differs from the visual track; the disadvantage of these systems is the difficulty in choosing a proper lighting structure;
- chromatic aberration; these systems, due to their specific character are mainly used to measure single points; and,
- stereovision, which consists in the comparative analysis of two pictures of the same scene, but from a different reference point.

2. Representation method with the use of the binocular stereovision system

Several possible stereovision systems can be applied for surface analysis. In the systems that use two cameras, the canonical or epipolar camera systems are applied. In the canonical system, the cameras’ optical axes are parallel and the scanning lines of planes are on the same straight line (Fig. 1). Thanks to such solutions, the adjustment of individual pixels in the pair of stereoscopic images is much easier.

Fig. 1. Geometry of the canonical camera system

In reality, however, these conditions cannot be fully executed because of the inaccuracy of the assembly of the structure of the vision system and the shortcomings of optics. Even the slightest angular aberration of optical axes may result in erroneous measurement. Thus, the camera system in such a configuration of cameras should also be considered as epipolar (Fig. 2) and
subject to the process of rectification, which mathematically guarantees the above mentioned conditions to be met. In order to ensure the best representation of the analysed surface, it is necessary to do the following:

− calibration process of the system, and
− the application of a fast and accurate algorithm for the matching of the pairs of stereo-point.

![Fig. 2. Geometry of the epipolar camera system](image)

There are two means of calibration applied: the separate calibration of cameras or the calibration of the entire system. Once the calibration finishes, in order to get the information about 3D coordinates of a given point of the analysed surface, it is necessary to identify the point in the images from both cameras. To limit the area of the search of the corresponding points, epipolar geometry is used [15]. Its theoretical basis is presented in Fig. 3.

![Fig. 3. Epipolar geometry](image)

We assume that pixel coordinates of the image of the $S$ point on the $P_L$ plane (marked with $S_{L'}$) and the point we are searching for is the $S_{P'}$ point. The $e_r$
line is marked on the basis of the statement that each real point in the space \((S_1, S_2, S_3)\) that corresponds to the \(S_L\) image has to be located on the optical axis defined by \(S_L\) and \(O_L\) points. The projection of that axis on the \(P_r\) plane marks the epipolar \(e_r\) line, which is the representation of all possible locations of the observed \(S\) point \((S_{P_1}, S_{P_2}, S_{P_3})\). To be more precise, the \(e_r\) line is formed by the intersection of the \(P_r\) projection and the plane marked by the points \(O_L\), \(O_P\) and \(S_L\). For the non-canonical system, the common point for all the epipolar lines on a given plane is the image of the central point of the second projection. This point is also known as the epipolar point. Epipolar geometry is determined on the basis of the information about the configuration of both cameras: the location of the central \(O_L\) and \(O_P\) points and the interrelation between the left and the right image plane. To properly record it, the so-called fundamental matrix \(F\) is applied. Its dimensions are 3x3, and it is defined to the constant value of the multiplier. This matrix connects the \(S\) point on one of the planes with the corresponding epipolar line on the other projection plane. For this, the following relationship is applied:

\[
F \cdot S = e
\]  

The coordinates of the \(S\) point are determined as homogenous for \(S=[u \ v \ w]\) for the element with \((x,y)\) coordinates\(=(u/w, v/w)\) and the epipolar line \(e\) with the equation \(ax+bx+c=0\) as \(e=[a \ b \ c]\).

The determination of the fundamental matrix is possible through the identification of a suitable amount of assigned images of the point from the cameras [7]. In the case when the camera system is canonical, the fundamental matrix is as follows:

\[
F = \begin{bmatrix}
0 & 0 & 0 \\
0 & 0 & 1 \\
0 & -1 & 0 \\
\end{bmatrix}
\]  

In the 3D reconstruction, the main target is to determine the coordinates of the points in images from both cameras that correspond to the analysed real point. When the canonical system is applied, the “y” coordinates are equal for both images of the points [3]. Then, the problem is reduced to determining the distance from the analysed point \(P\) to the camera system (the depth) on the basis of the analysis of the “x” coordinates of points \(P_1\) and \(P_2\) that are the representation of point \(P\) in camera 1 and 2 images (Fig. 4).
The equation describing the \( d_{1,2} \) divergence of points \( P_1 \) and \( P_2 \) is used:

\[
d_{1,2} = x_1 - x_2 = \frac{B \cdot f}{Z}
\]  

(3)

where:
- \( x_1, x_2 \) – coordinates of the images (in pixels),
- \( B \) – distance between the optical axes (the system basis),
- \( f \) – camera focal length (in pixels),
- \( Z \) – searched distance.

Fig. 4. Geometry for marking the divergence in the canonical camera system

Relation (3) determines the \( Z \) distance of the \( P \) point and can be show as:

\[
Z = \frac{B \cdot f}{d_{1,2}}
\]

(4)

In order to enable the application of these relationships for the represented planes that do not agree with the assumptions of the canonical system, the rectification procedure is applied [8]. Once it is conducted, the epipolar point of a given plane is transferred to infinity and the epipolar lines become parallel. Let us assume that \( P_1 \) and \( P_2 \) are the corresponding points in the pair of the basic images, while \( R_1 \) and \( R_2 \) – after their rectification. The following equation is thus true:

\[
R_1^T \cdot F \cdot R_2 = 0
\]

(4)
Rectification is conducted for all basic images with the application of the suitable transformation matrix — $H$. It joins the points before and after the transformation of the plane of the image into its canonical representation, which is presented in the following:

$$ R_1 = H_1 \cdot P_1 $$ (5)

$$ R_2 = H_2 \cdot P_2 $$ (6)

Adding (5) and (6) to equation (4) we get:

$$ P_1^T \cdot H_1^T \cdot F \cdot H_2 \cdot P_2 = 0 $$ (7)

Knowing the coordinates of the sufficiently large set of assigned points, it is possible to determine the $H_1$ and $H_2$ matrixes on the basis of the equation (6). However, this may give erroneous results. Various algorithms that enhance the determination of suitable matrixes are applied. One of the most popular among them is the method that uses the epipolar geometry of the non-calibrated cameras, which was proposed by Harley and Bouget’s algorithm [1] using the shift and rotation parameters of the system after calibration. The advantage of the first method is its ability to transform the projection plane in real time on the basis of selected points of the currently observed scene. Its disadvantage, on the other hand, is the lack of information concerning the scale of the analysed space, which is the result of the non-determined internal parameters of the camera. The second algorithm assumes the minimisation of the changes in images undergoing transformation, and at the same time aims at increasing the common field of vision. Once the rectification process is properly conducted, it is necessary to implement an effective algorithm that would, for the coordinate data of the points on one projection plane, identify the coordinate of the collinear point on the second plane of the image and calculate its distance from the camera system.

The methods used for the creation of the maps of the depth of the analysed area can be divided into [4]:

1. Direct – the divergence is calculated on the basis of the intensity of the pixels. The result of such an action is a dense map of divergence. We can distinguish between the following groups of algorithms: block, dynamic, gradient, neuron, relaxation, probabilistic and diffusion.
2. Applying the characteristics of the object – the divergence is calculated on the basis of the characteristic properties of a given object, i.e. the corners. The result is the point map of divergence. The most common algorithms are the Marr-Poggio-Grimson method that uses the multistage LoG (Laplacian of
Gaussian) controller and the Shirai method that consists in the adjustment of
the edge points and the determination of their degree of probability.

In the case when we have detailed knowledge on the application of the
developed stereovision system, it is advisable to limit the range of the possible
values of divergence, since it will considerably increase the speed and reliability
of the implemented algorithm. It is especially advantageous in the observation
of the objects with particularly restricted dimensional ranges.

For the optimal configuration of the vision system, it is necessary to take
several relations and specific characteristics of stereovision systems into
consideration. The most crucial is the reduction of the image resolution $R$ at the
time of the depth measurement and simultaneously to the increase in distance $Z$.
The $R$ resolution simultaneously determines $Z$ distance measurement error.
When assumed requirements are taken into account, $Z$ distance can be defined
with the use of the following:

$$R = \frac{Z^2}{\frac{R_h \cdot B}{2 \tan \frac{\alpha}{2}} - Z}$$

where:

$R$ – resolution of the depth measurement,
$B$ – distance between optical axes of the cameras,
$R_h$ – horizontal resolution of the camera,
$\alpha$ – horizontal angle of camera’s view.

The proper setting of the orientation of the cameras in interrelation is yet
another problem. Increasing the angle between the optical axes creates the risk
of different light reflection from a given surface, which in turn may result in
significant aberration of the intensity of corresponding pixels. The camera
system similar to the canonical system enables, on the other hand, the
rectification process to be correctly conducted. However, to ensure correct
measurement at the time of the sample shift outside the plane, the angle between
optical axes of the cameras should be greater than $30^\circ$ [2]. It also allows better
representation of the edges of the holes in the tested sample and in turn, quicker
detection of the cracks that occur in this particularly damage prone area [11]. On
the other hand, in the case of samples with uneven edges, the greater the angle
between the axes, the greater possibility for the occlusion phenomenon to take
place. This makes the process of adjusting stereo points difficult and results in a
partial loss of data that are invisible in the image received from one camera.
3. Idea of the stereovision system for the inspection of surfaces in micro and macro scale

The general idea of the system consists in the application of the vision system composed of two high-resolution cameras. The inspection of the surface is conducted in real time on the basis of collected images, with the use of the computer software for the PC that controls the process. The general structure of the vision system is presented in Fig. 5. The computer will supervise the control over the monitoring process and the analysis of images. In the system, the possibility to communicate with the device through the Internet network has been assumed.

![Fig. 5. General structure of the visual inspection system with the binocular vision system](image)

The modelling of the structure of the vision system was conducted simultaneously to the analysis of the possibility of the application of digital cameras and lenses in configurations that would allow one to reach the targets of the task. The mathematical simulation of the visual tracks for the assumed conditions for the observation of the surface of the samples was conducted. The determined basic parameters of optical systems and digital cameras ensure the achievement of the planned and desired measurement resolution and the operational scope of the vision system. High-resolution digital cameras with GigaEthernet interface, that allow the analysis of the surface in the micrometer scale and the measurements with the assumed resolution, were selected. Their additional advantage is the ability to synchronise the image acquisition both by the programme and the equipment, which is extremely crucial for the application in stereovision systems. The selected objective – Zoom 6000 [14] has a modular structure with a reconfiguration option and is equipped with a mechatronic system for the regulation of the zoom and the focus. The electronic controllers with the built-in RS interface are used to control the settings of the apparatus. The basic parameters of the vision system are presented in Tab. 1.
Table 1. List of basic parameters of the vision system

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Maximum dimension of the vision filed</td>
<td>50 x 50 mm</td>
</tr>
<tr>
<td>2 Minimum dimension of the vision filed</td>
<td>1 x 1 mm</td>
</tr>
<tr>
<td>3 Maximum optical resolution</td>
<td>max. 1 µm</td>
</tr>
<tr>
<td>4 Image record frequency</td>
<td>max. 17 frames/s</td>
</tr>
</tbody>
</table>

The developed stereovision system in its experimental version is presented in Fig. 6.

![Test stand: a) general view, b) visual head mounted onto the positioning module](image)

Each of the cameras has three degrees of freedom, which allows them to be optimally arranged in relation to the tested surface, and it gives them the possibility to move linearly along the optical axis, regulate the aberration angle and change the angle between the modules for the positioning of the head in the horizontal plane. The LED ring illuminators and the directional illuminators have been applied for the illumination of the tested surface. The examples of the images of the surface recorded by the developed system are presented in Fig. 7.

The GeForce GTX 480 graphic card based on the NVIDIA chipset has been selected for the processing and the analysis of the stereovision images. The application of the CUDA technology incorporated within the card and the direct transfer of calculation onto its graphic processor makes it possible to simultaneously conduct much more operations than in the case of classic processors. This allows the application of the most advanced algorithms operating on huge size files with the minimum time for their processing.
Fig. 7. The example of image recording from two cameras (sample with the crack on the edge of the hole)

Conclusions

The developed structure of the surface inspection system consists in the use of a binocular vision system and the generation of the stereovision image. It allows the observation of the process of the creation of surface defects (cracks, loss) and the deformation of the tested sample. The proposed system is insensitive to rotation and shifts of the sample outside its plane. This widens the functionality of the system and increases the measurement accuracy in comparison to monocular vision systems. To reconstruct the entire 3D structure of the tested object, it is necessary to make sure that the texture on its surface has suitable density and contrast. Otherwise, the measurement of the deformations will reduce the characteristic points occurring on the sample and the possible interpolation of the deformation between the areas of their occurrence. The proposed system is mainly meant for the testing of samples subject to fatigue loads under laboratory conditions. In a modified version, it could be used for the detection and tracking of destruction processes occurring in structures in real conditions.

Scientific work executed within the Strategic Programme “Innovative Systems of Technical Support for Sustainable Development of Economy” within Innovative Economy Operational Programme.
References


Recenzent:
Stanisław MROZIŃSKI

Koncepcja dwukamerowego stereowizyjnego systemu do inspekcji powierzchni

Słowa kluczowe
Badania zmęczeniowe, inspekcja powierzchni, optyczna inspekcja, stereowizja.

Streszczenie
W artykule przedstawiona została koncepcja dwutorowego systemu stereowizyjnej inspekcji powierzchni. Omówione zostały podstawowe cechy geometrii epipolarnej oraz zalety i ograniczenia metod obrazowania z jej wykorzystaniem. Zaprezentowana została struktura układu wizyjnego, podstawowe parametry optyczne oraz konfiguracja sprzętowa. Określane zostały przewidywane możliwości zastosowań praktycznych.